

**FUNCTIONAL AND STRUCTURAL STUDIES OF PLANT
 β -GLYCOSIDASE SUBSTRATE SPECIFICITY**



Sompong Sansenya

**A Thesis Submitted in Partial Fulfillment of the Requirements for
The Degree of Doctor of Philosophy in Biochemistry
Suranaree University of Technology
Academic Year 2013**

การศึกษาหน้าที่และโครงสร้างของเอนไซม์เบตาไกลโคซิเดสจากพืชต่อ

ความจำเพาะกับสับสเตรท



นายสมพงษ์ แสนเสนา

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาชีวเคมี

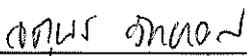
มหาวิทยาลัยเทคโนโลยีสุรนารี

ปีการศึกษา 2556

**FUNCTIONAL AND STRUCTURAL STUDIES OF PLANT β -
GLYCOSIDAS SUBSTATE SPECIFICITY**

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

Thesis Examining Committee



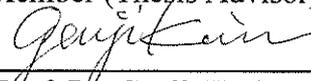
(Assoc. Prof. Dr. Jatuporn Wittayakun)

Chairperson



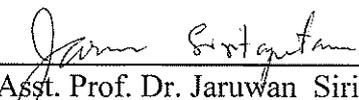
(Prof. Dr. James R. Ketudat-Cairns)

Member (Thesis Advisor)



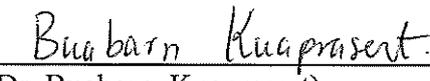
(Prof. Dr. Genji Kurisu)

Member



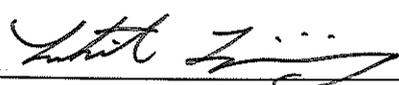
(Asst. Prof. Dr. Jaruwat Siritapetawee)

Member



(Dr. Buabarn Kuaprasert)

Member



(Prof. Dr. Sukit Limpijumnong)

Vice Rector for Academic Affairs



(Assoc. Prof. Dr. Prapun Manyum)

Dean of Institute of Science

สมพงษ์ แสนเสนา

:การศึกษาหน้าที่และโครงสร้างของเอนไซม์เบตาไกลโคซิเดสจากพืชต่อความจำเพาะกับสับสเตรท (FUNCTIONAL AND STRUCTURAL STUDIES OF PLANT β -GLYCOSIDASE SUBSTRATE SPECIFICITY)

อาจารย์ที่ปรึกษา : ศาสตราจารย์ ดร.เจมส์ เกตุทัต-คาร์นส์, 137 หน้า

ผลึกของเอนไซม์ Os4BGlu12 อิสระ และเอนไซม์ที่รวมอยู่กับตัวยับยั้ง 2,4-dinitrophenyl-2-deoxy-2-fluoroglucoside (DNP2FG) และ 2-deoxy-2-fluoroglucose (G2F) และเอนไซม์กลายพันธุ์ Os4BGlu12 E179Q ที่รวมอยู่กับ TAG ได้ถูกผลิตขึ้น ผลึกทั้ง 4 ชนิดสามารถหักเหรังสีเอกซเรย์ได้ความละเอียดถึง 2.50 2.45 2.40 และ 3.2 อังสตรอม ตามลำดับจากการเปรียบเทียบโครงสร้างสามมิติกับเอนไซม์ชนิดอื่นที่อยู่ในตระกูลเดียวกันพบว่าโครงสร้างโดยรวมมีลักษณะเหมือนกัน แต่ตรงบริเวณ loop B ของเอนไซม์ Os4BGlu12 จะมี disulfide bridge เพิ่มขึ้นมา บริเวณเร่งปฏิกิริยาของเอนไซม์มีลักษณะเป็นช่องและมีความลึกประมาณ 20 อังสตรอม ตรงบริเวณที่ลึกสุดของตำแหน่งเร่งปฏิกิริยาของเอนไซม์กับตัวยับยั้ง G2F จะพบโครงสร้างแบบ 4C_1 chair ซึ่งเกิดพันธะโควาเลนต์กับกรดอะมิโนที่ทำหน้าที่เป็น nucleophile แต่ในขณะที่ตรงบริเวณเร่งปฏิกิริยาของเอนไซม์กับตัวยับยั้ง DNP2FG จะพบโครงสร้างแบบ 1S_3 skew boat ซึ่งสอดคล้องกับโครงสร้างแบบ 4H_3 half-chair ในสถานะทรานซิชันของปฏิกิริยาการย่อยสลาย เมื่อบริเวณเร่งปฏิกิริยาของกรดอะมิโนที่ทำหน้าที่เป็น nucleophile ในโครงสร้างของเอนไซม์ Os4BGlu12 ที่จับอยู่กับตัวยับยั้ง G2F มีความคล้ายคลึงกับที่พบในเอนไซม์ *Sinapsis alba* myrosinase มากกว่าเอนไซม์ในกลุ่ม O-glycosidase เช่น เอนไซม์ Os3BGlu6 หรือเอนไซม์ Os3BGlu7 ซึ่งสอดคล้องกับการที่เอนไซม์ Os4BGlu12 สามารถย่อยสับสเตรท S-glycoside ได้โดยไม่พบการย่อยสับสเตรท S-glycoside ในกลุ่ม O-glycosidase ตัวอื่น จากการเปรียบเทียบตำแหน่ง aglycon binding site ของเอนไซม์ Os4BGlu12 กับเอนไซม์ตัวอื่นที่อยู่ในตระกูลเดียวกันพบว่าบริเวณ aglycon ของตัวยับยั้ง DNP2FG และสับสเตรท TAG ในโครงสร้างของเอนไซม์ Os4BGlu12 ถูกล้อมรอบด้วยกรดอะมิโนชนิดไม่ชอบน้ำ และกรดอะมิโนชนิดมีขั้วซึ่งกรดอะมิโนเหล่านี้มีความแตกต่างกับกรดอะมิโนที่พบใน aglycon binding site ของเอนไซม์ตัวอื่นที่อยู่ในตระกูลเดียวกัน

จากการเปรียบเทียบโครงสร้างของ เอนไซม์ Os3BGlu6 และ เอนไซม์ Os4BGlu12 กับ เอนไซม์ Os3BGlu7 ที่จับกับโอลิโกแซคคาไลด์ แสดงให้เห็นว่ากรดอะมิโน Met251 ในเอนไซม์ Os3BGlu6 กีดขวางการเข้าจับของเซลล์โอลิโกแซคคาไลด์ที่บริเวณ subsite +2 ขณะที่กรดอะมิโน His252 ในเอนไซม์ Os4BGlu12 ในบริเวณสอดคล้องกันนี้สามารถสร้างพันธะไฮโดรเจนกับโอลิโกแซคคาไลด์ได้เหมือนกับเอนไซม์ Os3BGlu7 การเปลี่ยนกรดอะมิโน Met251 เป็น Asn ของ เอนไซม์ กลายพันธ์ Os3BGlu6 ทำให้ค่า k_{cat}/K_m ในการย่อย laminaribiose เพิ่มขึ้น 15 เท่าเมื่อเทียบกับเอนไซม์ Os3BGlu6 ดั้งเดิม และค่า k_{cat}/K_m เพิ่มขึ้น 9 ถึง 24 เท่าในการย่อย เซลล์โอลิโกแซคคาไลด์ที่มีหน่วยความยาว 2 ถึง 5 หน่วย ในทางกลับกันการเปลี่ยนกรดอะมิโน Ans245 เป็น Met ในเอนไซม์ Os3BGlu7 ทำให้ค่า k_{cat}/K_m ในการย่อย laminaribiose ลดลง 6.5 เท่า และ 17 ถึง 30 เท่าสำหรับเซลล์โอลิโกแซคคาไลด์ที่มีหน่วยความยาวสายมากกว่า 2 หน่วย ในขณะที่การเปลี่ยนกรดอะมิโน His252 เป็น Met ในเอนไซม์ Os4BGlu12 ค่า k_{cat}/K_m ลดลง 2 ถึง 6 เท่า

สาขาวิชาชีวเคมี
ปีการศึกษา 2556

ลายมือชื่อนักศึกษา _____

ลายมือชื่ออาจารย์ที่ปรึกษา _____



SOMPONG SANSENYA : FUNCTIONAL AND STRUCTURAL STUDIES
OF PLANT β -GLUCOSIDASE SUBSTRATE SPECIFICITY.

THESIS ADVISOR: PROF. JAMES R. KETUDAT-CAIRNS, Ph.D. 135 PP.

GLYCOSIDE HYDROLASE FAMILY 1/ β -GLUCOSIDASE/RICE/PROTEIN
CRYSTALLIZATION/THREE DIMENSIONAL STRUCTURE/
MUTATION/KINETIC STUDY

The crystal structures of apo wild type Os4BGlu12, and its complexes with 2,4-dinitrophenyl-2-deoxyl-2-fluoroglucoside (DNP2FG) and 2-deoxy-2-fluoroglucose (G2F) and the acid/base mutant Os4BGlu12 E179Q complex with TAG were solved at 2.50, 2.45, 2.40 and 3.2 Å resolution, respectively. The overall structure of rice Os4BGlu12 is typical of GH1 enzymes, but it contains an extra disulfide bridge in the loop B region. The active site is located at the bottom of an approximately 20 Å deep slot-like pocket surrounded by a large surface loop. In the innermost part of the active site (the -1 subsite), the glucose ring of the G2F in the covalent intermediate was found in a 4C_1 chair conformation, while that of the noncovalently bound DNP2FG had a 1S_3 skew boat, consistent with hydrolysis via a 4H_3 half-chair transition state. The 1S_3 skew boat conformation of the glucose ring of TAG also fit to the electron density best with the lowest B-factor value when compared to the 4C_1 chair conformation. The position of the catalytic nucleophile (Glu393) in the G2F structure was more similar to that of the *Sinapsis alba* myrosinase G2F complex than to that in covalent intermediates of other O-glucosidases, such as rice Os3BGlu6 and Os3BGlu7 β -glucosidases. This correlated with a significant thioglucosidase activity for Os4BGlu12, although with 200-

to 1200-fold lower k_{cat}/K_m values for *S*-glucosides than the comparable *O*-glucosides, while hydrolysis of *S*-glucosides was undetectable for Os3BGlu6 and Os3BGlu7. The aglycones of DNP2FG and TAG were in contact with both hydrophobic and polar residues, which had little similarity to those aglycone-binding residues in other known GH1 β -glucosidase structures.

Superimposition of the structures of Os3BGlu6 and Os4BGlu12 with those of Os3BGlu7 bound to oligosaccharides showed that the corresponding Os3BGlu6 residue, Met251, appears to block the binding of cellooligosaccharides at the +2 subsite, whereas His252 in this position in Os4BGlu12 could hydrogen bond to oligosaccharides. Mutation of Os3BGlu6 Met251 to Asn resulted in a 15-fold increased k_{cat}/K_m value for hydrolysis of laminaribiose compared to wild type Os3BGlu6 and 9 to 24-fold increases in the k_{cat}/K_m values for cellooligosaccharides with degrees of polymerization (DP) of 2-5. On the other hand, mutation of Os3BGlu7 Asn245 to Met decreased the k_{cat}/K_m of hydrolysis by 6.5-fold for laminaribiose and 17 to 30-fold for cellooligosaccharides with DP > 2, while mutation of Os4BGlu12 His252 to Met decreased the corresponding k_{cat}/K_m values 2 to 6-fold.

Together, these studies have clarified the binding and hydrolysis of various substrates in plant GH1 β -glucosidases.

School of Biochemistry

Academic Year 2013

Student's Signature _____

Advisor's Signature _____

CONTENTS

	Page
ABSTRACT IN THAI.....	I
ABSTRACT I ENGLISH.....	III
ACKNOWLEDGEMENT.....	V
CONTENTS.....	VII
LIST OF TABLES.....	XII
LIST OF FIGURES.....	XIII
LIST OF ABBREVIATIONS.....	XVI
CHAPTER	
I INTRODUCTION.....	1
1.1 Glycosyl hydrolase.....	1
1.2 Glycoside hydrolase classification and structure.....	2
1.3 Morphology of active sites in glycoside hydrolases.....	5
1.4 Glycoside hydrolase mechanisms.....	8
1.4.1 Inverting mechanism.....	8
1.4.2 Retaining mechanism.....	8
1.4.3 Acid/base and nucleophile mutants with rescued activity.....	9
1.5 GH 1 substrate specificity and its structural basis.....	11
1.6 Rice β -glycosidases.....	18
1.7 Research objectives.....	25

CONTENTS (Continued)

	Page
II MATERIALS AND METHODS	25
2.1 Materials.....	25
2.1.1 Recombinant plasmids and bacterial strains.....	25
2.1.2 Site-directed mutagenesis and oligonucleotides primers.....	26
2.1.3 Chemicals and reagents.....	26
2.2 General methods.....	27
2.2.1 Transformation of expression host cells with recombinant expression plasmid.....	27
2.2.2 Recombinant expression.....	28
2.2.3 Protein extraction from <i>E. coli</i>	29
2.2.4 Purification of recombinant thioredoxin- β -glucosidase fusion Proteins.....	37
2.2.5 Purification of tag-free β -glucosidases.....	30
2.2.6 Protein analysis.....	31
2.2.6.1 Enzyme activity.....	31
2.2.6.2 SDS-PAGE electrophoresis.....	32
2.2.6.3 Protein concentration determination.....	32
2.3 Temperature stability assay.....	33
2.4 Kinetic analysis.....	33

CONTENTS (Continued)

	Page
2.5 Tris inhibition.....	35
2.6 Protein crystallization.....	35
2.6.1 Preparation of protein for crystallization.....	35
2.6.2 Apo-Os4BGlu12 crystallization.....	35
2.6.3 Crystallization of wild type Os4BGlu12 and Os4BGlu12 E179Q Complexes.....	36
2.6.4 Crystallization of Os3BGlu6 E178Q and E394D complexes with gibberellin GA ₄ β-D-glucosyl ester.....	37
2.6.5 Microseeding.....	38
2.6.6 Soaking of crystal with ligand.....	38
2.7 Data collection and processing.....	39
2.7.1 Flash vitrification.....	39
2.7.2 Synchrotron X-ray diffraction.....	40
2.8 Structure solution, refinement and validation.....	41
2.9 Docking studies.....	43
III RESULTS	44
3.1 Protein expression and purification.....	44
3.1.1 Purification of recombinant Os4BGlu12.....	44
3.1.2 Purification of Os4BGlu12 E179Q and H252M.....	47
3.1.3 Purification of Os3BGlu6 E178Q and M251N mutant proteins.....	48

CONTENTS (Continued)

	Page
3.2 Substrate specificity of Os4BGlu12.....	50
3.3 Stability study comparing two rice β -glucosidase enzymes.....	52
3.4 Crystals of apo-Os4BGlu12.....	54
3.5 Crystals of Complexes of Os4BGlu12 with DNP2FG and G2F.....	54
3.6 Crystals of Os4BGlu12 E179Q with TAG.....	55
3.7 Crystals of apo Os3BGlu6 E178Q.....	55
3.8 Os4BGlu12 structure determination.....	62
3.8.1 X-ray analysis.....	62
3.8.1 Model quality and overall structure.....	65
3.9 Os4BGlu12 interactions with ligands.....	72
3.9.1 Apo-Os4BGlu12 binding of Tris.....	72
3.9.2 Os4BGlu12 complex with DNP2FG.....	74
3.9.3 Os4BGlu12 complex with G2F.....	77
3.9.4 Os4BGlu12 E179Q TAG complex.....	79
3.10 Determination of Os3BGlu6 E178Q complex structures.....	82
3.10.1 X-ray analysis and structure refinement.....	82
3.10.2 Structure of Os3BGlu6 E178Q soaked with <i>p</i> NPGlc.....	85
3.10.3 The structure of Os3BGlu6 E178Q with GA ₄ GE.....	86
3.11 Kinetic comparisons of rice wild type β -glucosidases and their cello- oligosaccharide-binding subsite +2 mutants.....	91

CONTENTS (Continued)

	Page
IV DISCUSSION	95
4.1 The crystal structure of Os4BGlu12.....	95
4.1.1 Overall structure.....	95
4.1.2 The binding of ligand in the active site of Os4BGlu12.....	98
4.1.2.1 Michaelis complex of Os4BGlu12 with DNP2FG.....	99
4.1.2.2 The covalent intermediate of Os4BGlu12 with G2F.....	102
4.1.2.3 Thioglucosidase activity.....	103
4.1.2.4 The hydrolysis of natural glucoside substrate by rice Os4BGlu12.....	104
4.2 Structures of Os3BGlu6 covalent glucosyl-intermediate.....	106
4.3 Comparison of oligosaccharide binding by rice β -glucosidases.....	110
V CONCLUSION	116
REFERENCES.....	122
CURRICULUM VITAE.....	135

LIST OF TABLES

Table	Page
2.1 Mutagenic oligonucleotide primer.....	26
3.1 Substrate specificity of Os4BGlu12 for the hydrolysis of various substrates.....	51
3.2 The Os4BGlu12 crystallization conditions and X-ray diffraction.....	60
3.3 Os3BGlu6 crystallization conditions and X-ray diffraction experiments.....	61
3.4 Data collection statistics for Os4BGlu12 data sets.....	64
3.5 Refinement statistics for the structures of Os4BGlu12 apo enzyme and its complexes with G2F and DNP2FG and Os4BGlu12 E179Q with TAG	68
3.6 Hydrolysis of <i>O</i> -glucosides and <i>S</i> -glucosides by Os4BGlu12.....	79
3.7 Data-collection and structure-refinement statistics.....	84
3.8 Hydrogen bonds occurring between active site amino acid residues and glucose in the structure of Os3BGlu6 E178Q soaked with <i>p</i> NPGlc.....	86
3.9 Hydrogen bonds occurring between glucose or glycerol with amino acid binding residues in the structure of Os3BGlu6E178Q soaked with GA ₄ -GE.....	88
3.10 Kinetic constants of Os3BGlu6, Os3BGlu6 M251N, Os3BGlu7, Os3BGlu7 N245M, Os4BGlu12, and Os4BGlu12 H252M for hydrolysis of oligosaccharides.....	94

LIST OF FIGURES

Figure	Page
1.1 Generally catalytic mechanism of a glycoside hydrolase.....	2
1.2 The overall topology of the white clover cyanogenic β -glucosidase (PDB file 1CBG).....	5
1.3 The three types of active site shapes found in glycoside hydrolases.....	7
1.4 General pathways of the two major mechanisms of enzymatic glycosidic bond hydrolysis.....	10
1.5 Mechanisms of azide rescue of a nucleophile mutant: (A) and an acid/base mutant (B).....	11
3.1 SDS-PAGE of fractions from purification of recombinant thioredoxin-Os4BGlu12 protein by Co^{2+} -IMAC.....	45
3.2 SDS-PAGE of tag-free Os4BGlu12 protein fractions purified via a second IMAC step.....	46
3.3 SDS-PAGE analysis of Os4BGlu12 protein from two steps of purification.....	47
3.4 The purification profile of the two recombinant Os4BGlu12 mutant proteins E179Q(A) and H252M (B).....	48
3.5 SDS-PAGE profiles of the Os3BGlu6 E179Q and M251N mutants after the first and second IMAC steps.....	49
3.6 Comparison of the thermostabilities of Os4BGlu12 and Os3BGlu7.....	53
3.7 Photographs of apo-Os4BGlu12 crystals produced in various NaCl concentrations.....	56

LIST OF FIGURES (Continued)

Figure	Page
3.8 Photographs of Os4BGlu12 crystals produced in co-crystallization with 2 mM DNP2FG in various NaCl concentrations.....	57
3.9 Photographs of crystals obtained from co-crystallization of Os4BGlu12 E179Q with 2 mM TAG in various NaCl concentrations.....	58
3.10 Apo Os3BGlu6 E178Q crystals developed in precipitants with different pH values.....	59
3.11 Dynamic light scattering (DLS) spectrum of tag-free Os4BGlu12 protein.....	67
3.12 Ribbon diagram representation of the the Os4BGlu12 apo protein structure in the asymmetric unit.....	69
3.13 Superimposition of the 3D structures of different plant GH1 β -glucosidases.....	70
3.14 Binding of a Zn ²⁺ ion between the two molecules in the asymmetric unit	71
3.15 The active site region of native Os4BGlu12 with Tris bound.....	73
3.16 The Lineweaver Burk plots for hydrolysis of <i>p</i> NPGlc in different concentrations of Tris.....	74
3.17 Protein-ligand interactions in the active site of the Os4BGlu12 DNP2FG Complex.....	76
3.18 The active site region of the covalent complex of Os4BGlu12 with G2F.....	78
3.19 Active site of the crystal structure of Os4BGlu12 E179Q with TAG.....	81

LIST OF FIGURES (Continued)

Figure	Page
3.20 The electron density maps of the glycosyl-enzyme intermediate of Os3BGlu6 E178Q with <i>p</i> NPGlc (A) and GA ₄ -GE (B, C).....	89
3.21 Binding of glucose in the active site of Os3BGlu6 E178.....	90
3.22 Comparison of the structures of three rice β -glucosidases in complexes with substrates or inhibitors.....	93
4.1 Mechanism of retaining β -glucosidases, as seen in rice Os4BGlu12.....	99
4.2 Comparison of ligand binding by three rice β -glucosidases.....	101
4.3 Active site cleft of the crystal structure of Os4BGlu12 E179Q with TAG and the homology model of Os4BGlu13 with TAG.....	106
4.4 Comparison of the active site region of Os3BGlu6 bound with G2F (A) and those of the structures of Os3BGlu6 E178Q soaked with <i>p</i> NPGlc (B) and Os3BGlu E178Q soaked with GA ₄ GE (C).....	109
4.5 Comparison of cellooligosaccharide binding in the active sites of Os3BGlu7, Os3BGlu6 and Os4BGlu12.....	115

LIST OF ABBREVIATIONS

BSA	Bovine Serum Albumin
°C	Degree(s) Celsius
DNase	Deoxyribonuclease
EDTA	Ethylene diamine tetraacetic acid
GA ₄ GE	Gibberellin A4 β-D-glucosyl ester
h	Hour
kDa	Kilo Dalton
LB	Lysogeny broth (Luria-Bertani broth)
min	Minute
(m, μ)g	(milli, micro) Gram
(m, μ)L	(milli, micro) Liter
(m, μ)M	(milli, micro) Molar
(μ)mol	(micro) Mole
MW	Molecular Weight
nm	nanometer(s)
IPTG	Isopropyl-β-D-thiogalactopyranoside
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PEG	Polyethyleneglycol

LIST OF ABBREVIATIONS (Continued)

<i>p</i> NP	<i>p</i> -Nitrophenolate
<i>p</i> NPGlc	<i>p</i> -Nitrophenyl- β -D-glucopyranoside
S200	Superdex 200
SDS	Sodium dodecyl sulfate
s	second(s)
TEMED	Tetramethylenediamine
Tris	Tris-(hydroxymethyl)-aminoethane
V_m	Matthews' coefficient
v/v	Volume/volume
w/v	Weight/volume
w/w	Weight/weight
Å	Ångströms

CHAPTER I

INTRODUCTION

1.1 Glycoside hydrolases

Glycoside hydrolases (GH; EC 3.2.1-3.2.3) are enzymes that catalyze the hydrolysis of glycosidic linkages between sugars or between sugar and nonsugar aglycone moieties of glycosides, leading to the formation of a sugar hemiacetal or hemiketal and the corresponding free aglycone (Figure 1) (Esen, 1993). Glycoside hydrolases are also referred to as glycosyl hydrolases, glycosidases, and carbohydrases. Glycoside hydrolases can catalyze the hydrolysis of O-, N- and S-linked glycosides. The existence of glycosidases has been known for more than 100 years and they were the very first biological catalyst investigated. Glycosidases are truly remarkable enzymes. They are able to catalyze glycoside hydrolysis at rates of up to 1000 s^{-1} , which are rate enhancements of 10^{17} fold (Wolfenden *et al.*, 1998).

Different glycosidases form different products and display completely different ranges of bond specificity. For example, α -amylase is an enzyme that hydrolyses endo- α -(1-4) bonds in polysaccharide chains and produces α -(1-4)-linked oligosaccharides with α -D-glucosyl residues on the reducing end of the oligosaccharides released from the nonreducing end of the original polysaccharide molecules (it is a so-called retaining enzyme) (Boel *et al.*, 1990). β -amylase also hydrolyzes α -(1-4) bonds of amylose, but it cuts off only maltose from the nonreducing ends (exo) and forms a different anomeric configuration (β -maltose;

it is also called inverting enzyme) (Mikami *et al.*, 1992)

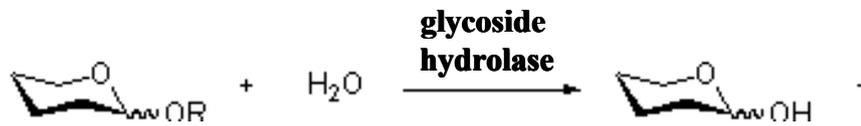


Figure 1.1 General catalytic mechanism of a glycoside hydrolase.

1.2 Glycoside hydrolase classification and structure

Glycoside hydrolases have been classified based on two independent criteria. The first criterion is catalytic specificity, as designated by their EC (Enzyme Commission) numbers, while the second criterion is sequence homology (Rojan *et al.*, 2004). On the basis of sequence similarity, GH have been grouped into 132 families by the Carbohydrate-Active Enzymes (CAZY) classification system (Henrissat, 1991; Henrissat and Bairoch, 1993, 1996; Henrissat and Davies, 1997). The up-to-date information is available on Carbohydrate-Active Enzymes database (CAZY at <http://www.cazy.org/CAZY/index.html>, Cantarel *et al.*, 2009). In the case of multiple-domain proteins, each catalytic domain is considered separately. Since three-dimensional structures are more strongly conserved than the sequence, GH families can be further grouped into clans (or superfamilies) of enzymes with related structures and catalytic mechanisms (Henrissat and Davies, 1997).

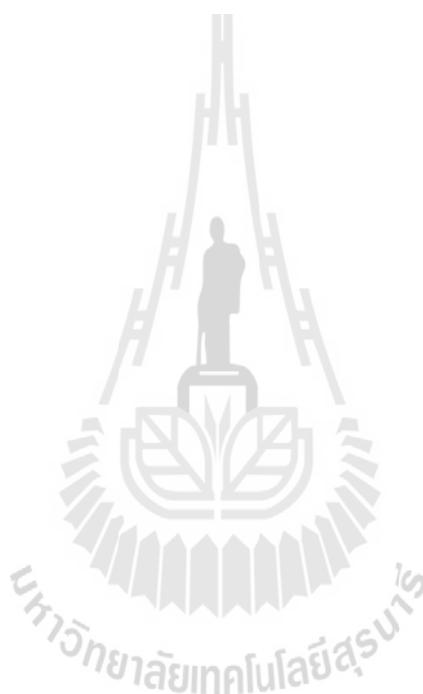
A family was initially defined as a group of at least two amino acid sequences displaying significant similarity and with no significant similarity with other families (Henrissat, 1991). Later, some related families of glycosidases have been combined.

into clans (Henrissat and Bairoch, 1996; Henrissat and Davies, 1997). According to its definition, a clan is a group of families that are thought to have a common ancestry, which is recognized by significant similarities in tertiary structure together with conservation of the catalytic residues and catalytic mechanism. For example, glycosidases catalyze hydrolysis of glycosidic bonds of their substrates via two general mechanisms, leading to either overall inversion or retention of the anomeric configuration at the cleavage point, and members of one clan will generally all be either inverting or retaining (McCarter and Withers 1994; Davies and Henrissat 1995; Henrissat and Davies, 1997).

Families of four clans (GH-A, GH-D, GH-H, and GH-K), as well as several other families that have not been assigned to any clan, contain proteins with a similar $(\beta/\alpha)_8$ barrel folds of their catalytic domains. The $(\beta/\alpha)_8$ barrel is also called the TIM barrel, derived from the first enzyme in which it was identified, triose phosphate isomerase (Banner *et al.*, 1975). This typical alpha/beta barrel comprises 8 α -helices and 8 β -strands creating a barrel-like structure, the core of which is formed by a barrel-shaped parallel β -sheet, which is surrounded by the α -helices (Figure 2).

GH Clan A is the largest clan and its members share a similar core $(\beta/\alpha)_8$ barrel structure. The active sites of these enzymes are located at the C-terminal portion of the β -barrels and are surrounded by loops connecting the α -helices to the β -strands. Two catalytic amino acid residues are located at the C-terminal ends of β -strands 4 and 7 of the barrel, and the enzymes hydrolyze the glycosidic bond with retention of the anomeric configuration (Jenkins *et al.*, 1995; Henrissat *et al.*, 1995; Coutinho and Henrissat, 1999). These residues are found near the bottom of a cleft or crater-shaped active site at the carboxy-terminal end of the β -strand of the central

barrel (Cantarel *et al.*, 2009, Ketudat Cairns and Esen, 2010). This group comprises more than 1000 enzymes with various substrate specificities from 19 different families, including GH families 1, 2, 5, 10, 17, 26, 30, 35, 39, 42, 50, 51, 53, 59, 72, 79, 86, 113 and 128 (Henrissat *et al.*, 1995; Jenkins *et al.*, 1995; Bolam *et al.*, 1996; Henrissat and Bairoch, 1996; Cantarel *et al.*, 2009).



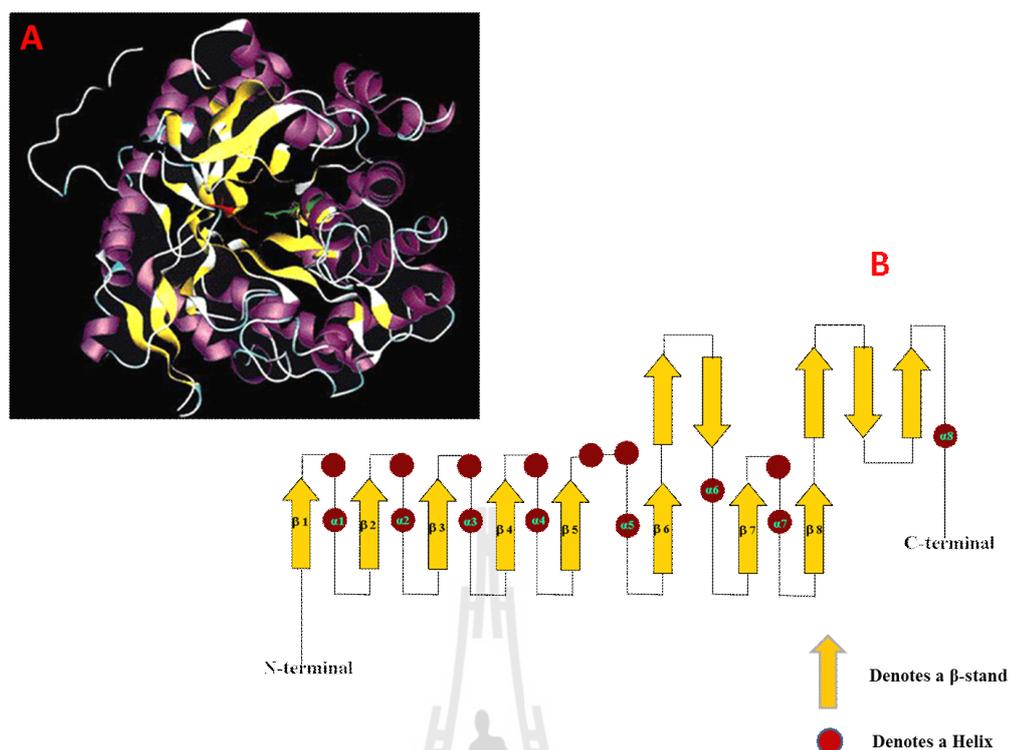


Figure 1.2 The overall topology of the white clover cyanogenic β -glucosidase (PDB file 1CBG).

A. Ribbon diagram showing the overall fold of the molecule. The catalytic residues Glu183 in green and Glu397 in red. **B.** Secondary structure diagram of white clover cyanogenic β -glucosidase showing the eight parallel β -strands that form the barrel core. The β -strands are represented by arrows and the α -helices by circles. The eight β -strands forming the barrel are labeled $\beta 1$ – $\beta 8$ and the eight peripheral helices $\alpha 1$ – $\alpha 8$ (Barrett *et al.*, 1995).

1.3 Morphology of active sites in glycoside hydrolases

Davies and Henrissat (1995) described the different active-site structures among GH. First, a pocket or crater-like active site is seen in exo-acting hydrolases, such as β -amylase, glucoamylase, and monosaccharidases, like β -glucosidase, β -

galactosidase and neuraminidase. This type of active site is capable of recognizing polysaccharide substrates; however, the specificity is determined by the depth and shape of the pocket, which in turn reflects the number of subsites that contribute to binding and to the length of the leaving group. Second, the cleft-like active site allows random binding of several sugar units in polymeric substrates and is commonly found in endo-acting polysaccharides, which include β -1,3-1,4-glucanases and β -1,3-glucanases. Third, the tunnel-like active site is created by the covering of an open groove by a long loop of the protein. This morphology is seen in the cellobiohydrolases (Rouvinen *et al.*, 1990) and α -carrageenases. In cellobiohydrolases, the tunnel enables the polysaccharide chain to be threaded through it, releasing the product while the enzyme is still attached to the polysaccharide substrate, thereby allowing for processivity. Hydrolysis will cease when the movement of the enzyme is blocked by steric factors or when the loops closing the active site move and free the polysaccharide chain.



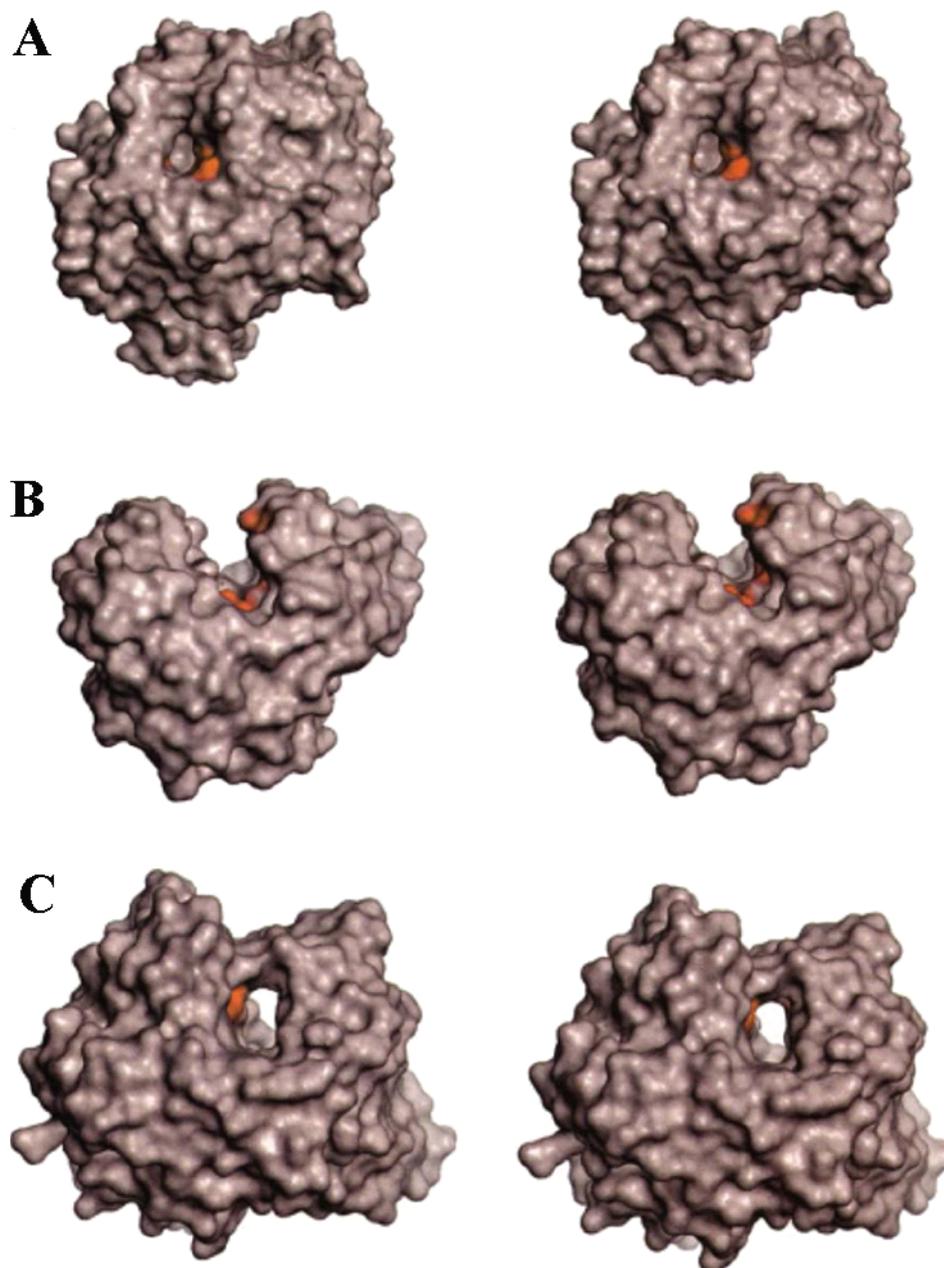


Figure 1.3 The three types of active site shapes found in glycoside hydrolases.

(A) The pocket (glucoamylase from *A. awamori*). (B) The cleft (endoglucanase E2 from *T. fusca*). (C) The tunnel (cellobiohydrolase II from *T. reesei*). Stereo views of the enzyme surfaces are shown in each case. The figure is from Davies and Henrissat (1995).

1.4 Glycoside hydrolase mechanisms

As mentioned above, two mechanisms of GH, inverting and retaining mechanisms, have been identified (McCarter and Withers, 1994). Each mechanism is classified by the anomeric configuration of the products generated from the reaction which is the opposite or the same as the substrate (Figure 4). Two active site carboxylic acids residues are involved in both mechanisms, but with different roles.

1.4.1 Inverting mechanism

Inverting glycosidases catalyze hydrolysis via a direct displacement of the aglycone (leaving group) by water (which acts as a nucleophile). During catalysis, one carboxylated residue acts as a general base to activate the water molecule as a nucleophile (Figure 4B). The activated water then attacks the anomeric carbon of the sugar, while another carboxylated residue acts as general acid to assist the departure of the aglycone part by protonating the oxygen atom at the glycosidic bond. The distance between these two catalytic residues is $\sim 10 \text{ \AA}$, which allows the water and substrate to bind simultaneously (Zechel and Wither, 2000).

1.4.2 Retaining mechanism

In retaining glycosidases, including GH1, the cleavage of the glycosidic bond is performed by a double displacement mechanism (Figure 4A). The distance between the two catalytic residues is $\sim 5\text{-}5.5 \text{ \AA}$. The catalysis is performed in two separate steps. In the first step, one carboxylated residue acts as an acid catalyst, protonating the glycosidic oxygen, while the other carboxylated residue acts as a nucleophile to attack at the sugar anomeric center to form a covalent glycosyl-enzyme intermediate.

In the second step, the resulting glycosyl-enzyme intermediate is hydrolyzed by a water molecule and this second nucleophilic substitution at the anomeric carbon generates a product with the same stereochemistry as the substrate (Davies and Henrissat, 1995; Zechel and Wither, 2000).

1.4.3 Acid/base and nucleophile mutants with rescued activity

The catalytic residues of retaining glycosidases can be identified by mutation and rescue of the activity by small nucleophiles (Wang *et al.*, 1994, 1995). In these studies, either of the two catalytic residues (glutamates in GH1) are generally mutated to alanine or glycine, and these mutations of either the catalytic nucleophile or acid/base show extremely low activity. The activity of the nucleophile to Ala or Gly mutants of *Agrobacterium* sp. β -glucosidase (Abg) was reactivated for cleavage of 2,4-dinitrophenyl β -D-glucoside by azide to produce β -D-glucosyl azide, a product with the inverted anomeric configuration (Figure 5A) (Wang *et al.*, 1994). In the case of acid/base mutations, a substrate with a good leaving group, which has a low pK_a and therefore does not require protonation, eliminates the requirement for acid catalysis in the glycosylation step, while the small nucleophile (azide) does not require basic assistance in the deglycosylation step (Wang *et al.*, 1995). Thus, the covalent intermediate can form normally in the glycosylation step and the product with the same anomeric configuration as the substrate (β -D-glucosyl azide in the case of Abg acid/base mutants) is produced by transglycosylation in the deglycosylation step. These results have been observed with mutant species of other retaining β -glycosidases, such as *Cellulomonas fimi* β -1,4-glycanase, *Geobacillus*

stearothermophilus T6 β -xylosidase and *Streptomyces* sp. β -glucosidase (MacLeod *et al.*, 1996; Bravman *et al.*, 2003; Vallmitjana *et al.*, 2001).

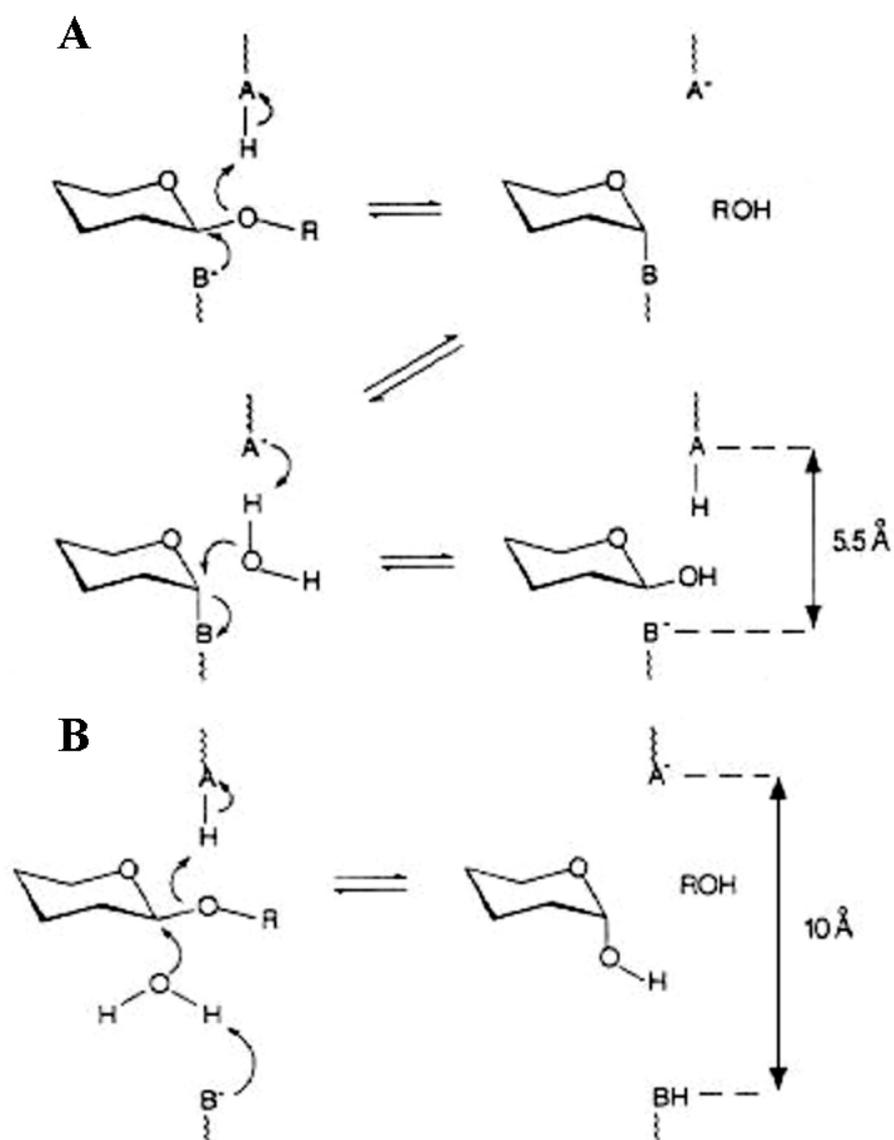


Figure 1.4 General pathways of the two major mechanisms of enzymatic glycosidic bond hydrolysis:

(A) retaining and (B) inverting mechanisms pathway (Davies and Henrissat, 1995).

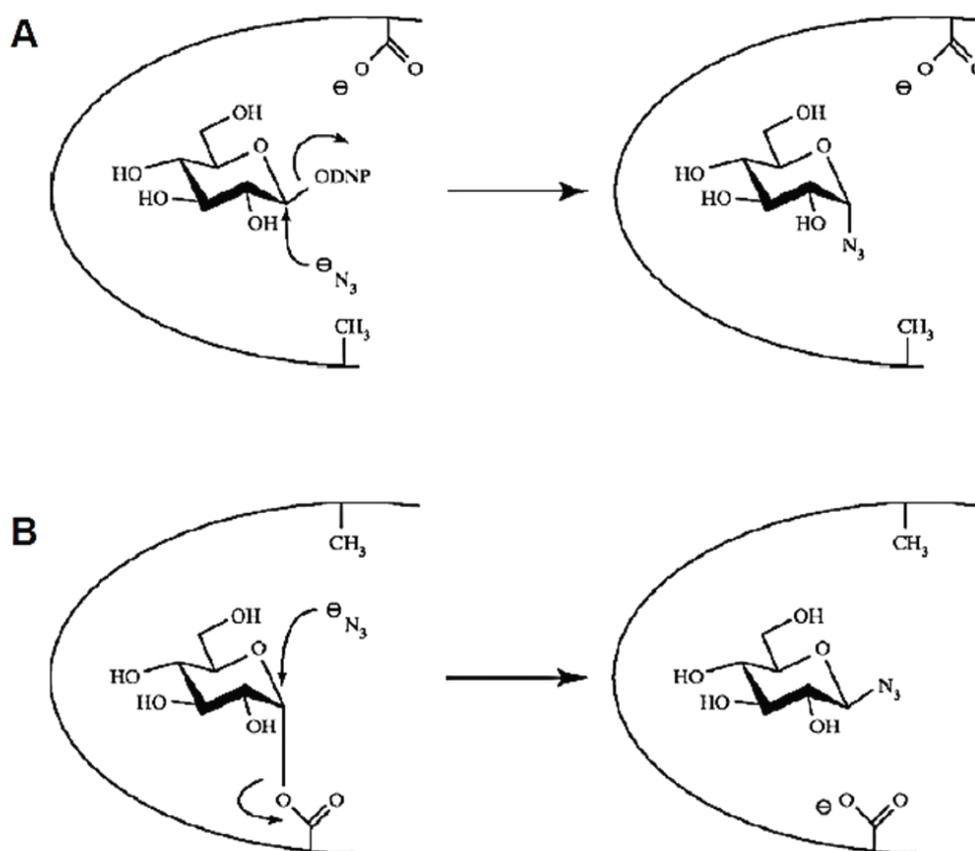


Figure 1.5 Mechanisms of azide rescue of a nucleophile mutant:

(A) and an acid/base mutant (B) (Ly and Withers, 1999).

1.5 GH 1 substrate specificity and its structural basis

Over the past 20 years, there has been a great increase in the number of 3D structures of glycoside hydrolases. GH1 includes retaining glycosidases from archaea, eubacteria and eukaryotes, with a wide range of substrate and reaction specificities, as well as proteins with other biological and enzymatic functions (Ketudat Cairns *et al.*, 2012). These include β -D-glucosidases, β -D-mannosidases, β -D-galactosidases,

phospho- β -D-glucosidases, disaccharidases like primeverosidase (Mizutani *et al.*, 2002) and furcatin hydrolase (Ahn *et al.*, 2004), myrosinases (thio- β -glucosidases) hydrolyzing the *S*-glycosidic bonds of plant 1-thio- β -D-glucosides (glucosinolates) (Burmeister *et al.*, 1997), and hydroxyisourate hydrolase, which hydrolyzes an internal bond in a purine ring, rather than a glycosidic bond (Raychaudhuri and Tipton, 2002).

The structures of around thirty-nine GH1 enzymes have been determined, and their catalytic domains all consist of a $(\beta/\alpha)_8$ barrel with a cleft or crater-shaped active site at the carboxy-terminal end of the β -strands of the central barrel (Cantarel *et al.*, 2009; Ketudat Cairns and Esen, 2010). The first structure from this family was that of the white clover (*Trifolium repens*) cyanogenic β -glucosidase, which is thought to perform a defensive function by hydrolyzing linamarin, leading to the release of hydrogen cyanide (Barrett *et al.*, 1995). Based on the three-dimensional structures, the active site of GH1 enzymes range from narrow slots to wide pockets and can be divided into two parts, the glycone and aglycone binding subsites. The glycone subsite (also called subsite -1) binds the nonreducing end monosaccharide of the substrate, whereas the aglycone binding site binds with the remaining part of the substrate. The aglycone binding region may be formed by several subsites (+1, +2, +3 and so on, where each subsite binds one monosaccharide residue in the case of oligosaccharide substrates). The substrate cleavage point is between subsites -1 and +1 (Marana, 2006). The glycones recognized by GH1 β -glycosidases include β -D-glucose, β -D-galactose, β -D-fucose, β -D-mannose, β -D-xylose, 6-phospho- β -D-glucose and 6-phospho- β -D-galactose (Ketudat Cairns and Esen, 2010). The diversity of aglycones is higher, including monosaccharides, oligosaccharides, and aryl or alkyl

groups (Coutinho and Henrissat, 1999). The enzymes may hydrolyze a broad range of substrates with different glycones and aglycones with different specificity, but some enzymes may be specific for only one type of glycone or aglycone. The fundamental substrate specificity of these enzymes depends on their overall binding shape, dimensions and geometry of the binding site, and conformation of the active site amino acids that are important for the substrate recognition and binding, as well as the structure of the aglycone and glycone moieties of the substrate (Czjzek *et al.*, 2000; Verdoucq *et al.*, 2004).

Since several crystal structures of enzyme-ligand complexes have been determined, the residues forming the glycone binding site (subsite -1) and the non-covalent interaction with the glycone are well identified (Burmeister *et al.*, 1997; Czjzek *et al.*, 2000; 2001; Zechel *et al.*, 2003; Verdoucq *et al.*, 2004; Isorna *et al.*, 2007; Chuenchor *et al.*, 2008; 2011 and Seshadri *et al.*, 2009). The interactions between the residues at subsite -1 and the monosaccharide moiety at the substrate nonreducing end involves a hydrogen bond network of at least 5 active site amino acids and 4 substrate hydroxyls (Marana, 2006). All hydrogen bonds and residues forming these interactions are conserved among GH1 enzymes. In *Z. mays* Glu1, amino acid residues critical for glucose binding, as identified by Czjzek *et al.* (2001), include Asn38, Trp465 and His142, which interact with OH3, Gln38, Trp457 and Glu464, which interact with OH4, while Glu464 also interacts with OH6, and the basal platform residue W457. This basal platform establishes stacking interactions with the glycone ring in the form of soft hydrogen bonds between the pi electrons of the indole ring and the slightly positively charged C-H protons of the glucose ring. In *T. maritima*, *Bacillus polymyxa*, *Sinapis alba*, *Lactococcus lactis*, *Z. mays* and *S.*

bicolor β -glycosidases, the OH2 forms a bidentate hydrogen bond with an asparagine and a glutamate residue (eg. N188 and E404 in *S. bicolor*). This glutamate also acts as the catalytic nucleophile of the β -glycosidases. A histidine residue also forms a hydrogen bond with OH2 in *B. polymyxa* (Isorna *et al.*, 2007) and *S. bicolor* (Verdoucg *et al.*, 2004). Although the amino acid residues forming the hydrogen bond network at subsite -1 are conserved, differences of glycone specificity of the GH1 enzymes have been reported (Czjzek *et al.*, 2000, Opassiri *et al.*, 2004; 2006).

Some cases of structural differences between the crystallographic structures of the apoenzyme and its complexes occur at the positions of the nucleophilic residues. In the covalent intermediate complex structures of the *P. polymyxa*, *S. solfataricus*, *T. maritima*, rice Os3BGlu6, and rice Os3BGlu7 with 2-deoxy-2-fluoroglucoside (G2F) and human cytosolic β -glucosidase's acid/base mutant with native D-glucoside, the side-chain position of the nucleophilic residue has moved to link with the anomeric carbon of the G2F or glucosyl residue. For example, a displacement of C δ by 1.5–1.6 Å is observed when the rice Os3BGlu7 covalent complex (2RGM) is compared to the apoenzyme structure, 2RGL (Chuenchor *et al.*, 2008). In these structures, the angles between the sugar anomeric carbon, O ϵ 1 and C δ of the nucleophilic residue, fall in the range of 113–117°. In contrast, in *S. alba* myrosinase, an S-glycosidase, the displacement of C δ is only 0.2 Å in the covalent G2F complex (PDB accession 1E73) compared to the apo protein (1MYR) and the angle between the anomeric carbon, O ϵ 1 and C δ of the nucleophilic residue is wider (138°) (Burmeister *et al.*, 2000).

The high diversity of aglycones recognized by β -glycosidases is correlated with the structural variability of the aglycone binding subsite (Marana, 2006). The study of the active site structure and the amino acids which are responsible for

substrate binding and recognition should reveal the substrate specificity of these enzymes. Recently, several β -glycosidases crystallized in complex with a ligand occupying the aglycone-binding region have been published (Czjzek *et al.*, 2000; Verdoucq *et al.*, 2004; Sue *et al.*, 2006; Isorna *et al.*, 2007; Barleben *et al.*, 2007; Chuenchor *et al.*, 2008; 2011; Seshadri *et al.*, 2009 and Sue *et al.*, 2011; Barleben *et al.*, 2007 and Xia *et al.*, 2012).

Maize β -glucosidase (ZmGlu1) and sorghum dhurrinase (SbDhr1) share 72% sequence identity, but ZmGlu1 shows broad substrate specificity, which includes its natural substrate, DIMBOA-Glc, but not dhurrin. SbDhr1, on the other hand, is specific to dhurrin. Swapping of the C-terminal domain between ZmGlu1 and SbDhr1 gave rise to a chimeric ZmGlu1 with dhurrinase activity. Amino acid sequence comparison and homology modeling of ZmGlu1 and SbDhr1 revealed three amino acids, Ser462, Ser463, and Phe469, at the C-terminus of SbDhr1 that are involved in dhurrinase specificity (Czjzek *et al.*, 2001). Later, comparison of the crystal structures of ZmGlu1 and SbDhr1 showed that these two β -glucosidases are different in their active site (Verdoucq *et al.*, 2004). The active site of ZmGlu1 appears to be a flattened crater or slot, while the active site of SbDhr1 is wider and smaller. In ZmGlu1, the aromatic side chains of Trp378, Phe198, Phe205, and Phe466 form the aglycone binding site and are responsible for binding to DIMBOA or other aromatic aglycones by aromatic stacking π -interactions (Czjzek *et al.*, 2000). This is different from SbDhr1, in which Asn259, Phe261, and Ser462 are crucial for aglycone recognition and binding via hydrogen bonding, as well as hydrophobic interactions and π -interactions (Verdoucq *et al.*, 2004).

One case of importance to plant secondary metabolism is the enzymes

involved in indole alkaloid metabolism, strictosidine- β -glucosidase (SG) and raucaffricine glucosidase (RG) (Barleben *et al.*, 2007 and Xia *et al.*, 2012). Strictosidine- β -glucosidase (SG) is known to hydrolyze its natural substrate strictosidine, while raucaffricine glucosidase (RG) is also involved in indole alkaloid metabolism and uses the downstream metabolite raucaffricine glucoside as a substrate. In terms of specificity, RG converts strictosidine, the substrate of SG, but SG does not accept raucaffricine, the substrate of RG (Barleben *et al.*, 2007 and Xia *et al.*, 2012). The inactive mutant crystal structures of both enzymes were solved, SG E207Q complexed with strictosidine glucoside (PDB; 2JF7) and RG E186Q with raucaffricine glucoside (4A3Y). Comparison of the two structures showed that the aglycone binding site of RG is a little similar to that of SG, but more similar to other β -glucosidases. The active site morphology of RG has a wider gate, which allows strictosidine to enter the catalytic site, whereas the slot-like entrance of the active site of SG prevents access by raucaffricine. The critical residues to make a difference of active site morphology of the two enzymes are Trp392 in RG and Trp388 in SG. The conserved Trp residues are positioned in opposite directions and the Trp392 direction makes a wider gate that allows both raucaffricine and strictosidine substrates to bind in the catalytic site of RG enzyme, while the Trp388 direction facilitates stronger hydrophobic, sandwich-like interactions with the strictosidine substrate. When this conserved Trp is compared to the corresponding Trp in other GH1 β -glucosidase structures, such as cyanogenic β -glucosidase (1CBG), Maize (ZmGlu1) β -glucosidase (1E4L), rice Os3BGlu7 β -glucosidase (2RGL), the position of Trp392 in RG is similar to other GH1 β -glucosidase, but has an opposite side chain chi angle when compared to the SG Trp388 (Xia *et al.*, 2012).

In the structure of the strictosidine β -glucosidase (mutant SG-Glu207Gln) complexed with strictosidine, other residues making interactions in the binding site have been described as follows (Barleben *et al.*, 2007). The aglycone part of the strictosidine substrate is surrounded by Phe221, Trp388, Gly386, Met275, Thr210, and Met297, most of which are hydrophobic (Barleben *et al.*, 2007). The critical residues for recognition of the aglycone are Gly386 and Trp388. Gly386 is in very close proximity to the indole system of strictosidine, hence replacing it with a larger residues like serine decreases the enzyme activity, as demonstrated for the G386S mutant. The glucose moiety interacts with the conserved residues found in other GH1 structures (Asn206, Gln207, Asn343, Tyr345, Glu416, Trp473, His161, Gln57, Trp465, Glu472, and Tyr481). Glu207 (or its substitute Gln207, in the E2207Q mutant) and Glu416 of SG are located within the pocket near the sugar moiety, with a distance of 5.2 Å between their carboxyl carbons, which is appropriate for the substrate to enter and to place the glucosidic bond in an optimal position for hydrolysis. His161, which is located 5.8 Å away from the anomeric carbon atom of the glucosyl moiety, is not a catalytic residue, but forms a hydrogen bond to the O3 of the sugar moiety of strictosidine, thereby helping the substrate to be in correct orientation for deglycosylation. Tyr481 corresponds to Tyr473 in ZmGlu1, which was suggested to be important for aglycone recognition by the maize enzyme (Czjzek *et al.*, 2000), and is hydrogen bonded with the amide group of Trp388 (Trp378 in ZmGlu1) (Verdoucq *et al.*, 2003).

1.6 Rice β -glycosidases

Database analysis has indicated the presence of 34 GH1 genes in the rice genome that are likely to encode active glycosidases (Opassiri *et al.*, 2006). Recently, three rice GH1 enzymes have been characterized at the molecular and structural level, Os3BGlu6, Os3BGlu7, and Os4BGlu12, which share 51–53% sequence identity, come from three distinct phylogenetic branches of plant GH1 enzymes, yet all hydrolyze oligosaccharides and glycosides, although with different preferences.

Rice Os3BGlu7, originally known as BGlu1, is encoded by a gene that was cloned from rice seedling and characterized by Opassiri *et al.*, 2003. The *Os3BGlu7* mRNA is highly expressed in shoot during germinating and flower during expansion (Opassiri *et al.*, 2003). The Os3BGlu7 fusion protein hydrolyzed β -(1,4)-linked gluco-oligosaccharides with increasing catalytic efficiency (k_{cat}/K_m) values as the DP increased from 2 to 6, while the hydrolysis efficiency of β -(1,3)-linked oligosaccharides decreased with the DP from 2 to 3 and was not detectable for longer laminarioligosaccharides (Opassiri *et al.*, 2003 and 2004). The purified Os3BGlu7 fusion protein also hydrolyzed several *p*NP glycosides with varying efficiency. The Os3BGlu7 hydrolysis efficiency was most similar to barley BGQ60 β -D-glucosidase/ β -D-mannosidase, with which it shares over 66% amino acid sequence identity, and both hydrolyze short β -(1,3)-and longer β -(1,4)-linked oligosaccharides. However, barley BGQ60 is a better β -mannosidase than β -glucosidase, whereas rice Os3BGlu7 is a better β -glucosidase than β -mannosidase (Leah *et al.*, 1995; Hrmova *et al.*, 1998; 2006 Opassiri *et al.*, 2003 and 2004). Os3BGlu7 hydrolyzed *p*NP- β -D-mannoside (*p*NPMan) with 10% relative activity when compared to *p*NP- β -D-

glucoside (*p*NPGLc). Other rice isoenzymes in the same phylogenetic cluster with Os3BGlu7 (Os3BGlu8 and Os7BGlu26) also hydrolyzed both *p*NPGLc and *p*NPMAN, but with different relative activities. Os3BGlu8 hydrolyzed *p*NPGLc much faster than *p*NPMAN, while Os7BGlu26 showed higher hydrolysis efficiency for *p*NPMAN than *p*NPGLc (Opassiri *et al.*, 2004; Kuntothom *et al.*, 2009).

The crystal structure of apo Os3BGlu7 and its covalent intermediate with 2-deoxy-2-fluoroglucoside was solved by Chuenchor *et al.* (2008). The overall structure of Os3BGlu7 was similar to the known structures of other GH1 β -glucosidase, but showed several differences in the loop around the active site (loop C). The amino acid residues surrounding the -1 subsite (glycone binding site) are conserved residues found in other GH1 enzyme, while docking studies with celotriose indicated that the residues Tyr131, Arg178, Leu183, Asp243 His267 and Trp337, all of which are conserved in barley BGQ60, might be crucial for binding of oligosaccharide substrates (Chuenchor *et al.*, 2008). Recently, the crystal structures of an Os3BGlu7 catalytic acid/base mutant (Glu176 mutated to Gln: E176Q) with oligosaccharides (laminaribiose, celotetraose and celopentaose) were solved (Chuenchor *et al.*, 2011). The amino acid residues crucial for binding of oligosaccharide substrates beyond the -1 subsite included Trp358 (the conserved Trp residue discussed for other plant GH1 enzyme in Section 1.5), Leu183, Asn190, Asn245 and Leu442 (at the +1 and +2 subsites), of which Asn245 showed two strong hydrogen bonds to Glc residues 3. The nonreducing end glucosyl residues (at subsite +3 and +4) interact with Tyr341 and Ser334 (Chuenchor *et al.*, 2011).

The mutation of the Os3BGlu7 catalytic nucleophile (Glu386) destroys the hydrolysis catalytic activity, but creates a glycosynthase that can synthesize long

oligosaccharides by transglucosylation, when α -glucosyl fluoride is used as a donor and *p*NP-cellobioside as acceptor (Hommalai *et al.*, 2007 and Pengthaisong *et al.*, 2012). Transglucosylation activity has also been observed in the native Os3BGlu7, which produced pyridoxine glucoside with *p*NPGlc as a glucosyl donor (Opassiri *et al.*, 2004). The crystal structures of the E386G glycosynthase with and without α -D-glucosyl fluoride and celooligosaccharides (cellotetraose and cellopentaose) were solved by Pengthaisong *et al.* (2012). In the complex of E386G glycosynthase with α -glucosyl fluoride, an ordered water molecule was found near the 2-OH (replaced by fluoride), and this water was calculated to be important to glycosynthase activity (Pengthaisong *et al.*, 2012; Wang *et al.*, 2013). In the complex structure of the Os3BGlu7 catalytic acid/base and nucleophile mutants with cellotetraose and cellopentaose, the aromatic ring of Tyr341 had interactions with glucosyl residues at the +3 and +4 subsites, but the mutation of Tyr341 to alanine and leucine showed only small effects on the hydrolysis of oligosaccharides (Chuenchor *et al.*, 2011; Pengthaisong *et al.*, 2012). However, Pengthaisong *et al.* (2012) observed that mutation of Os3BGlu7 Tyr341 (Y341A) resulted in cellotetraose binding in an alternative productive position, which could explain why the mutant enzymes hydrolyzed cellotetraose and cellopentaose with k_{cat}/K_m values similar to those of the wild type enzyme. However, the Y341A and Y341L mutations of the Os3BGlu7 E386G glycosynthase synthesized shorter *p*NP-celooligosaccharides than does the E386G glycosynthase, suggesting that Tyr341 plays an important role in the synthesis of long oligosaccharides by the glycosynthase. These studies suggest that aglycone-binding is complex and the active site shape may play at least as critical a role as the surface residues.

Rice Os3BGlu6 is a member of cluster At/Os1 in the phylogenetic analysis based on amino acid sequence similarity (Opassiri *et al.*, 2006). The first expression and characterization of Os3BGlu6 was done by Seshadri *et al.* (2009). Os3BGlu6 only hydrolyzes β -1-2- and β -(1-3)-linked disaccharides (sophorose and laminaribiose), while β -(1-4)-linked cellooligosaccharides were hydrolyzed extremely poorly. Among *p*NP glycosides, Os3BGlu6 had highest activity on *p*NP- β -D-fucopyranoside, followed by *p*NP_{Glc} and *p*NP- β -D-galactoside, respectively. Os3BGlu6 also hydrolyzes certain natural glycosides, including apigenin-7-*O*-glucoside, glycitin, diadzin, genistin, esculin, arbutin, coumaryl alcohol β -D-glucoside, coniferin and salicin, although at low rates. Os3BGlu6 had highest hydrolysis activity with *n*-octyl- β -D-glucoside, followed by *n*-heptyl- β -D-glucoside, respectively (Seshadri *et al.*, 2009). Recently, Os3BGlu6 was shown to hydrolyze gibberellin A4 glucosyl ester (GA₄-GE) to free GA₄ and glucose with the highest activity of five rice GH1 enzymes, which had been recombinantly expressed in our laboratory (Os3BGlu6, Os3BGlu7, Os4BGlu12, Os4BGlu18 and Os4BGlu31) (Hua *et al.*, 2013). The transesterification activity was seen in the acid/base mutants of Os3BGlu6 (E178Q and E178A) when GA₄-GE was used as a donor and sodium azide as an acceptor, despite the observation that wild type Os3BGlu6 had no transesterification activity (Hua *et al.*, 2013). The crystal structures of apo Os3BGlu6 and its complexes with inhibitors (2-deoxy-2-fluoro-glucoside and *n*-octyl- β -D-thioglucoside) were solved by Seshadri *et al.* (2009). The hydrolytic specificity of rice Os3BGlu6 was different from rice Os3BGlu7 in that Os3BGlu6 had little activity toward cellooligosaccharides, while Os3BGlu7 hydrolyzed β -(1-4)-linked oligosaccharides with increasing efficiency as the DP

increased up to 6 glucosyl residues (Opassiri *et al.*, 2004 and Seshadri *et al.*, 2009). A comparison of the structures of the covalent intermediates of rice Os3BGlu6 and Os3BGlu7 found that the amino acid residues at the -1 subsite are the same, except that Trp133 in the structure of Os3BGlu6 was replaced by Tyr131 in the structure of Os3BGlu7. In the aglycone binding site of the two enzymes, Os3BGlu7 Asn245, which acts at subsite + 2 in binding of cellooligosaccharides, is replaced with Met251 in Os3BGlu6 (Seshadri *et al.*, 2009). This Met251 appears to block the third glucosyl residue of cellooligosaccharides from binding to Os3BGlu6 in the same position as in Os3BGlu7, and the Os3BGlu7 N245V mutation gave a k_{cat}/K_m value for hydrolysis of cellotriose 15 times lower than wild type Os3BGlu7, due to a 42- fold increase in K_m (Chuenchor *et al.*, 2008). These effects show the importance of Asn245 in Os3BGlu7 binding of the third glucosyl residue in cellooligosaccharides (Chuenchor *et al.*, 2011).

Os4BGlu12 is a rice β -glucosidase that has been characterized and belongs to the plant GH1 phylogenetic tree cluster At/Os7 (Opassiri *et al.*, 2006). Os4BGlu12 showed the highest similarity to the cell wall associated β -glucosidase that had been previously characterized and shown to hydrolyze β -(1,3)- and β -(1,4)-linked gluco-oligosaccharides (Akiyama *et al.*, 1998). Os4BGlu12 that was expressed in recombinant *Escherichia coli* was shown to similarly hydrolyze gluco-oligosaccharides and also hydrolyzes certain glucosides with large, rather apolar aglycones, such as apigenin 7-*O*- β -D-glucoside and deoxycorticosterone 21-*O*- β -D-glucoside (Opassiri *et al.*, 2006, 2010). Northern blot analysis revealed the *Os4BGlu12* gene is highly expressed in rice seedling shoot and in the leaf sheath and stem of rice at flowering stage (Opassiri *et al.*, 2010). It was also shown to be wounding induced and to hydrolyze the β -(1,3),(1,4)-mixed linkage oligosaccharides

released by a wounding induced β -glucanase, suggesting a role for Os4BGlu12 in wound recovery (Opassiri *et al.*, 2010). Os4BGlu12 shares 51-53% amino acid sequence identity with rice Os3BGlu7 and its close relatives Os3BGlu8 and Os7BGlu26. However, unlike these enzymes, for which the catalytic efficiencies (k_{cat}/K_m) increase as the lengths of β -1,4-linked gluco-oligosaccharides increase from DP3 to DP6 (Opassiri *et al.*, 2004; 2006 and Kuntothom *et al.*, 2009), Os4BGlu12 shows only a slight increase in going from celotriose to cellopentaose, and drops at cellohexaose. Recently, Os4BGlu13, the rice isoenzyme that is most closely related to Os4BGlu12, was identified as a tuberonic acid glucoside (TAG) β -glucosidase (TAGG) (Wakuta *et al.*, 2010). Os4BGlu12 was later identified as another isoenzyme that hydrolyzed tuberonic acid glucoside and designated TAGG2 (Wakuta *et al.*, 2011). As tuberonic acid, a derivative of jasmonic acid, is glycosylated on the end of an unsaturated 5-carbon chain, it is possible the hydrolysis of apolar glycosides reflects a role in hydrolysis of tuberonic acid or similar compounds in the plant.

In order to understand the enzymes substrate binding mechanism leading to the substrate preference of Os4BGlu12 and the basis for the different substrate specificity among rice GH1 enzymes, the three-dimensional structure of Os4BGlu12 was determined in the current work. Moreover, the critical residues that appear to differentiate substrate specificity for celooligosaccharides were mutated between the rice enzymes for which structures are available. To investigate substrate-enzyme interactions further, the structures of catalytically compromised mutants of Os3BGlu6 and Os4BGlu12 were determined with putative natural substrates. This study will help to further clarify the enzyme-substrate binding mechanism by demonstrating the relationship of the active site architecture of the GH1 enzymes to substrate binding

and hydrolysis.

1.7 Research Objectives

The objectives of this thesis project are:

1. To determine the structure of Os4BGlu12 alone and in complex with a mechanism-based inhibitor in order to learn about substrate binding of this enzyme.
2. To clarify the basis for differential oligosaccharide hydrolysis by Os3BGlu6, Os3BGlu7 and Os4BGlu12 by site-directed mutagenesis to exchange residues apparently critical for differential oligosaccharide binding.
3. To identify the determinants of substrate-specificity of Os3BGlu6 and Os4BGlu12 by determination of the X-ray crystal structures of inactive mutants of these enzymes in complexes with synthetic and natural substrates.
4. To investigate the substrate specificity of Os4BGlu12 further by characterizing the activity of the tag-free form of the enzyme used for crystallization against natural substrates, including tuberonic acid β -D-glucoside and salicylic acid β -D-glucosides.

CHAPTER II

MATERIALS AND METHODS

2.1 Materials

2.1.1 Recombinant plasmids and bacterial strains

The recombinant plasmids for this work, pET32a+/DEST-*Os4bglu12* and pET32a+/DEST-*Os3bglu6* were previously made by Tassanee Onkoksoong (Opassiri *et al.*, 2006) and Supriya Seshadri (Seshadri *et al.*, 2009), respectively. These plasmids were used to produce recombinant Os4BGlu12 and Os3BGlu6 as N-terminal thioredoxin/His-tag fusion proteins. The mutations for this study were constructed by mutagenesis of these wild type recombinant plasmids.

The *Escherichia coli* strains Origami B(DE3) and Origami(DE3) were used to produce the recombinant thioredoxin β -glucosidase fusion proteins Os4BGlu12 and Os3BGlu6, respectively. The *E. coli* strain DH5 α was used for cloning plasmids.

2.1.2 Site-directed mutagenesis and oligonucleotides primers

The mutagenesis to produce cDNA encoding the six mutants including; Os3BGlu6 E178Q, Os3BGlu6 E394D, Os3BGlu6 M251N, Os3BGlu7 N254M, and Os4BGlu12 E179Q, Os4BGlu12 H252M was performed with the QuikChange® Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA) and the oligonucleotides primers shown in Table 2.1. These primer were ordered from Pacific Science Co. Ltd.

Table 2.1 Mutagenic oligonucleotide primers

Point mutation	Oligonucleotide sequences
Met251 → Gln	“F” 5’ GCTTGGGATAGCGTTCGACGTGAATTGGTTCGAGCCG 3’ “R” 5’ CGGCTCGAACCTTCACGTCTGAACGCTATCCCAAGA 3’
Glu178 → Gln	“F” 5’ GGATCACGCTCAACCAACCGCACACGGTGGC 3’ “R” 5’ GCCACCGTGCGGTTGGTTGAGCGTGATCC 3’
His252 → Met	“F” 5’ GATTGGAATAACTCTGGTCTCGATGTGGTTTGTTCCTTCTCC 3’ “R” 5’ GGAGAAGGGAACAAACCACATCGAGACCAGAGTTATTCCAATC 3’
Glu179 → Gln	“F” 5’ AGAGTGAAAAATTGGATCACCTTCAATCAGCCTTGGACTTTC 3’ “R” 5’ GAAAGTCCAAGGGTGATTGAAGGTGATCCAATTTTGGACTCT 3’
Gln245 → Met	“F” 5’ GAGTTGGAAAGAGCTTCATACCACATGAAAGTCCAGAACTATTCCAAC 3’ “R” 5’ GTTGAATAGTTCGGACTTCATGTGGTATGAAGCTCTTCCAAC 3’

2.1.3 Chemicals and reagents

Chemicals and reagents used for this work came from a variety of suppliers. Reagents used included: acrylamide, N,N’-methylene-bis-acrylamide, ammonium persulfate, ampicillin, acetonitrile, acetic acid, bacto agar, bacto peptone, calcium chloride, cellooligosaccharides, Coomassie brilliant blue R250, DNase I, enterokinase, glycine, glucose, HPLC-grade water, hydrochloric acid, imidazole, isopropyl thio-β-D-galactoside (IPTG), kanamycin, 2-mercaptoethanol, molecular

weight standards for SDS-PAGE, phenylmethylsulfonylfluoride (PMSF), polyethylene glycol 4000, sodium dodecyl sulfate, TEMED, tetracycline, Triton X-100, and yeast extract. The oligosaccharides and synthetic glycosides used in this thesis, including *p*-nitrophenyl (*p*NP) β -D-glucoside, *p*NP- β -D-fucoside, *p*NP- β -D-galactoside, *p*NP- β -D-xyloside, *p*NP- β -D-mannoside, sophorose, gentiobiose and cellobiose, were purchased from Sigma (St. Louis, Mo, USA). Cellooligosaccharides of degree of polymerization (DP) 3-6 and laminaribiose were purchased from Seikagaku Kogyo Co. (Tokyo, Japan) and salicylic acid β -glucoside (SAG), *meta*-salicylic acid glucoside (*m*-SAG), *para*-salicylic acid glucoside (*p*-SAG), and tuberonic acid glucoside (TAG) were provided by Dr. Shinji Wakuta, Dr. Wataru Saburi, and Prof. Hirokazu Matsui, Research Faculty of Agriculture, Hokkaido University. The crystallization screening kits used include the Sigma Crystallization Basic Kit, JBScreen HTS I and II (Jena Bioscience, Jena, Germany), Crystals Screen High Throughput HR2-130 and HR2-134 (Hampton Research, Aliso Viejo, CA, USA), JY Screen (Mahidol University, Thailand) and Emerald Biosystems Wizard I and II (Emerald BioSystems, Bainbridge Island, WA, USA).

2.2 General methods

2.2.1 Transformation of expression host cells with recombinant expression plasmid

The recombinant plasmids for both wild type Os4BGlu12 and its mutants (pET32a+/DEST-*Os4bglu12* and its derivatives) were transformed into Origami B(DE3) competent cells, while wild type Os3BGlu6 and its mutants (pET32a+/DEST-*Os3bglu6*) were transferred into Origami(DE3) competent cells.

One microliter (100 ng) of plasmid was added to 50 μ L thawed competent cells, mixed gently and the tube stored on ice for 30 min. The mixed cells were heat shocked at 42 $^{\circ}$ C for 60 second and the tube rapidly transferred to ice for 2 min. Five hundred microliters of LB broth was added to the transformed cells and they were incubated at 37 $^{\circ}$ C for 1 h. Then, 200 μ L of the transformation medium was spread directly onto an LB agar plate containing 100 μ g/mL of ampicillin, 15 μ g/mL kanamycin and 12.5 μ g/mL tetracycline and incubated overnight at 37 $^{\circ}$ C.

2.2.2 Recombinant expression

Selected Origami(DE3) colonies containing pET32a+/DEST-*Os3bglu6* and its mutants were grown overnight in 50 mL LB medium containing 50 μ g/mL ampicillin, 15 μ g/mL kanamycin and 12.5 μ g/mL tetracycline and cultured at 37 $^{\circ}$ C in an incubator shaker. The starter culture was then transferred into four 2 liter Erlenmeyer flasks containing 800 mL LB medium with the same antibiotics as above to give 1% culture and incubated at 37 $^{\circ}$ C until the optical density (OD) at 600 nm reached 0.5–0.6. Then, to induce the expression of the target proteins, the culture temperature was changed to 20 $^{\circ}$ C for 16 h.

The positive colonies of Origami B(DE3) containing pET32a+/DEST-*Os4bglu12* and its mutants were used to express Os4BGlu12 by the same methods as used for Os3BGlu6, except that when the OD at 600 nm reached 0.5–0.6, IPTG was added to a final concentration of 0.4 mM to induce the expression of the target proteins and the culture was grown for 16 h at 20 $^{\circ}$ C. Then the induced cultures were transferred to 50 mL precooled centrifuge tubes and chilled on ice for 10 min, then centrifuged at 3500 rpm for 15 min at 4 $^{\circ}$ C. The weights of the cell pellets were

recorded and the pellets stored at -80 °C.

2.2.3 Protein extraction from *E.coli*

Frozen bacterial cells pellets from expression of either Os3BGlu6 or Os4BGlu12 were thawed, and then resuspended in 5 mL per gram cells freshly prepared extraction buffer containing 20 mM Tris-HCl, pH 8.0, 200 µg/mL lysozyme, 1% Triton-X 100, 1 mM phenylmethylsulfonylfluoride (PMSF) and 0.25 µg/mL DNase1. The cell suspension was then incubated at room temperature for 30 min. The supernatant containing soluble protein was separated from the cell debris by centrifugation at 12000 rpm for 15 min at 4 °C, and the protein solution was kept at 4 °C.

2.2.4 Purification of recombinant thioredoxin-β-glucosidase fusion proteins

The thioredoxin-Os3BGlu6 and thioredoxin-Os4BGlu12 fusion proteins were purified by immobilized metal affinity chromatography (IMAC) on Co²⁺ (Talon resin, Clontech, Palo Alto, CA, USA) at 4 °C. In each round of protein purification, fifteen-milliliters of soluble protein extract was loaded onto a 3 mL Co²⁺ IMAC column, which was pre-equilibrated with equilibration buffer (20 mM Tris-HCl, pH 8.0, and 150 mM NaCl). Next, the column was washed with 5 column volumes of equilibration buffer and 5 column volumes each of 5 and 10 mM imidazole in equilibration buffer. The bound protein was eluted with 5 column volumes of 250 mM imidazole in equilibration buffer. All fractions from the purification steps were assayed for β-glucosidase activity with 1 mM *p*NPGlc as described by Opassiri *et al.*, (2003). Then the pure protein was pooled and concentrated in a 30 kDa molecular

weight cut-off (MWCO) Centricon centrifugal filter (Millipore, Bedford, MA, USA) and the buffer was changed to 150 mM NaCl in 20 mM Tris/HCl, pH 8.0.

2.2.5 Purification of tag-free β -glucosidases

For the crystallization and kinetic studies the thioredoxin and histidine fusion tags were cleaved off and the tag-free β -glucosidases were used. To remove the N-terminal fusion tag, Os4BGlu12 fusion protein was cleaved with 1:1000 (w/w) of enterokinase (New England Biolabs, Ipswich, MA, USA) and proteins in 150 mM NaCl, 20 mM Tris-HCl, pH 8.0, at 23 °C for 16 h. Os3BGlu6 fusion protein was instead cleaved with tobacco etch virus (TEV) protease at the ratio of 1:50 (w/w) TEV protease to the fusion protein in 20 mM Tris-HCl, pH 8.0, 150 mM NaCl for 16 h at 4 °C. The tag-free protein was separated from the fusion tag and uncleaved fusion protein by loading the digest onto a Co²⁺ IMAC column. The unbound protein fraction was collected and the resin was washed with 5 and 10 mM imidazole in equilibration buffer (150 mM NaCl, 20 mM Tris-HCl, pH 8.0). The unbound and wash fractions containing the tag-free protein were combined. The protein sample was concentrated and the buffer was changed to 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, in a 10 kDa MWCO Centricon filter by centrifugation at 1062 *g* at 4 °C. Finally, the tag-free protein was further purified by Sephacryl S200 chromatography on an ÄKTA Protein Purifier system (GE Healthcare, Sweden). The protein was equilibrated and eluted with 150 mM NaCl, 20 mM Tris-HCl, pH 8.0, at a flow rate of 0.25 mL/min. Fractions containing β -glucosidase activity were pooled.

2.2.6 Protein analysis

2.2.6.1 Enzyme activity

The β -glucosidase activity for all enzymes in this study were determined according to the method of Opassiri *et al.*, (2003). Activity toward *p*NPGlc was assayed in a reaction mixture containing 90 μ L of 50 mM sodium acetate, pH 5.0, 5 μ L of 10 mM *p*NPGlc (0.5 mM final concentration) and 5 μ L of crude enzyme. The reaction was incubated at 37 °C for 10 min, and then 50 μ L of 0.4 M sodium carbonate (Na_2CO_3) was added and to alkalize and stop the reaction. The free *p*-nitrophenolate was measured with an iEMS reader MF microplate photometer (Labsystem iMES MF, Finland) at 405 nm. A reaction without enzyme was used as a blank.

Glucose released from oligosaccharide substrates was measured with a coupled peroxidase-glucose oxidase assay. 25 μ L each of substrate (5 mM final concentration) and protein (0.1-1 mg/mL final concentration) in 50 μ L buffer (50 mM sodium acetate, pH 5.0) was mixed well in 1.5 mL tubes and incubated at 30 °C for 30 min. The reactions were stopped by heating at 100 °C for 5 min, and then transferred to a microtiter plate for measurement of the glucose. 100 μ L of peroxidase-glucose-oxidase enzyme and 50 μ L of ABTS were added to each microtiter plate well and the plate was incubated at 37 °C for 30 min. The release of glucose was determined by measuring the absorbance of the reaction at 405 nm.

The comparison of the hydrolytic activities of Os4BGlu12 toward various substrates (*p*NP glycoside, cellooligosaccharides and natural glycosides) at 2 mM concentration (Table 3.1) was done by the same assay methods described above, with 0.1-1 μ g of enzymes incubated at 30 °C for 5-60 min in 50 mM sodium acetate, pH

5.0, and terminated as described above.

2.2.6.2 SDS-PAGE electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to Laemmli (1970). The 12% acrylamide separating gel was prepared in 0.376 M Tris-HCl buffer, pH 8.8, while the 4% stacking gel contained was prepared in 0.125 M Tris-HCl buffer, pH 6.8. The protein sample was mixed with 1X SDS-PAGE loading buffer (50 mM Tris-HCl, pH 6.8, 10% (v/v) glycerol, 2% (w/v) SDS, 14.4 mM 2-mercaptoethanol and 0.1% bromophenol blue) and then boiled at 100 °C for 5 min. A 10-15 µL aliquot was loaded onto the gel, and then electrophoresed at 150 V. Protein bands were then visualized by staining of this gel in Coomassie brilliant blue R-250 staining solution (0.1% (w/v) Coomassie brilliant blue R-250, 45% (v/v) methanol, 45% (v/v) distilled water and 10% (v/v) glacial acetic acid) for 1 h at room temperature, and then destainin with destaining solution (45% (v/v) methanol, 45% (v/v) distilled water and 10% (v/v) glacial acetic acid) until excess of dye was removed. The size of protein was estimated by comparing its migration with those of the protein standard markers, which included bovine α -lactalbumin (14 kDa), trypsin inhibitor (21.5 kDa), bovine carbonic anhydrase (31 kDa), ovalbumin (45 kDa), bovine serum albumin (66 kDa), and phosphoryase (97.4 kDa).

2.2.6.3 Protein concentration determination

Protein concentrations in all the purification steps were estimated by the method of Bradford (Bradford *et al.*, 1976) with a Bio-Rad kit (Bio-Rad Corp., Hercules, CA, USA) using bovine serum albumin (BSA) as a standard (0-5µg). The

assay solution contained suitably diluted enzyme in distilled water in a total volume of 800 μL . The solution was mixed and incubated at room temperature for 10 minutes, and then the absorbance was measured at 595 nm with a Genesys 10UV spectrophotometer (Spectronic Instruments, Rochester, NY, USA).

For enzymological studies, purified tag-free Os4BGlu12 concentration was measured by absorbance at 280 nm. The extinction coefficient (ϵ)₂₈₀ of 107160 $\text{M}^{-1}\text{cm}^{-1}$ used in the protein concentration calculation was calculated by the equation of Gill and von Hippel (1989).

2.3 Temperature stability assay

For comparison of thermo-stability Os4BGlu12 with Os3BGlu7, tag-free Os3BGlu7 enzyme (which was produced as described by Chuenchor *et al.*, 2006) was compared with tag-free Os4BGlu12. In 50 μL reactions, 0.5 μg of Os4BGlu12 or Os3BGlu7 enzyme was incubated with or without 1 mM dithiothreitol (DTT) in 50 mM sodium acetate, pH 5.0, for 60 min at temperatures ranging from 0 to 60 $^{\circ}\text{C}$. The enzyme was then cooled on ice and then assayed at 30 $^{\circ}\text{C}$ with 0.3 mM *p*NPGlc as substrate for 10 min, and the reactions stopped and measured as described in section 2.2.6.1.

2.4 Kinetic analysis

All kinetic parameters were calculated from triplicate assays done at 30 $^{\circ}\text{C}$, in 50 mM sodium acetate, pH 5.0, as described in Section 2.2.6.1. For comparison of the hydrolysis efficiency of *O*- and *S*-glycosides by Os4BGlu12, the initial velocity was determined by measuring the enzyme activity at 3-6 different time points for 5 to 120

min, depending on the individual substrates used. Enzyme amounts in the range of 0.1 to 1 μg were used in the initial assay. The time point and amount of the enzyme where the reaction rate was linear and the absorbance value was in the range of 0.1-1.0 were then used to determine the kinetic constants. Substrate concentrations over a range of approximately one-third to three times the apparent K_m were included in calculations of kinetic parameters.

The activities toward a various substrates (β -1,3-linked (laminaribiose) and β -1,4-linked (cellooligosaccharides DP 2 to 6)) of Os3BGlu6, Os3BGlu6 M251N, Os3BGlu7, Os3BGlu7 N245M, Os4BGlu12 and Os4BGlu12 H252M were determined by measuring the activity at 3–6 different time points to determine the time giving the initial velocities (v_0). The enzyme concentrations used were 2 μg of wild type Os3BGlu6 and 0.3–0.5 μg of Os3BGlu6 M251N, Os3BGlu7, Os3BGlu7 N245M, Os4BGlu12 and Os4BGlu12 H252M. Substrate concentrations over a range of approximately 0.3 to 3 times the apparent K_m were used to determine the kinetic parameters. The apparent kinetic parameters including k_{cat} , K_m , k_{cat}/K_m were calculated by nonlinear regression of Michaelis–Menten plots with the Grafit 5.0 computer program (Erithacus Software, Horley, UK).

Changes in transition state binding energies in the mutations ($\Delta\Delta G_{S^*mut}$) were calculated as described by Fersht *et al.* (1987): $\Delta\Delta G_{S^*mut} = -RT [\ln(k_{cat}/K_m)_{mutant} - \ln(k_{cat}/K_m)_{wildtype}]$.

2.5 Tris inhibition

Competitive K_i values were determined by incubating of 0.5 μg enzyme with ten different concentrations of the inhibitor (20 to 200 mM) in presence of 0.1, 0.3, 0.5, 1.0, 3.0 and 5.0 mM *p*NPGlc as the substrate under the reaction conditions described in Section 2.2.6.1. A series of Lineweaver–Burk plots were drawn to verify the inhibition was competitive, and then Dixon plots were used to calculate the inhibition constants (Dixon, 1953).

2.6 Protein crystallization

2.6.1 Preparation of protein for crystallization

For crystallization screening, the purified Os4BGlu12 and Os3BGlu6 wild type and mutant proteins were adjusted to 3-10 mg/mL, in 20 mM Tris/HCl, pH 8.0, in a 10 kDa MWCO Centricon by centrifugation at 3000 rpm, 4 °C and the protein was filtered through an Ultrafree-MC 0.22 μm filter (Millipore) by centrifugation at 2500 rpm for 5 minute. The precipitant solutions were prepared with HPLC grade water, filtered with 0.45 μm filter and stored at 15 °C.

2.6.2 Apo-Os4BGlu12 crystallization

Conditions for the crystallization of apo-Os4BGlu12 were previously screened by the microbatch method with a systematic screen of monovalent and divalent salts (0.2 M) and pH (in the range 4.0–8.5 at 0.1 M buffer concentration) in the presence of 25% (w/v) polyethylene glycol (PEG) 4000 (Sansenya *et al.*, 2010). A promising condition (25% (w/v) PEG 4000 in 0.1 M Tris–HCl, pH 8.5, 0.2 M NaCl)

was optimized by hanging-drop vapor diffusion by varying the molecular weight and percentage of PEG and the concentrations of NaCl and protein, as described by Sansenya *et al.*, 2010. The crystallization trials were performed in 24 well plates (Greiner Bio-One, Germany). Glass cover slips (Menzel-Glaser, Germany) were siliconized to reduce the contact area of the drop protein with the glass surface of the cover slip. High vacuum grease (Dow Corning, USA) was applied to the top edge of each well. A hanging drop of 2 μ L of 2.5-5 mg/mL pure Os4BGlu12 was mixed with 1 μ L of precipitant solution on the center of a siliconized cover slip. The cover slip with the drop was inverted and placed on to the top of the well containing 0.5 mL of precipitant solution, and the grease seal on the top was twisted a few degrees to complete sealing. The drop was equilibrated against a reservoir of precipitant solution at 15 °C. Microseeding (Bergfors, 2003) was performed to produce larger crystals with pre-incubation of the drop and reservoir for 1 h before streaking in the crystal seeds as described in the Section 2.6.5.

2.6.3 Crystallization of wild type Os4BGlu12 and Os4BGlu12 E179Q complexes

To produce wild type Os4BGlu12 complexes with inhibitor or substrate, protein solution was mixed with 2-10 mM 2,4-dinitrophenyl-2-deoxy-2-fluoro- β -D-glucopyranoside (DNP2FG). In the case of Os4BGlu12 E179Q, the protein was crystallized with 2-10 mM tuberonic acid glucoside (TAG) solution, then the mixture was incubated at 4 °C for 10 min. Crystallization of the complexes was then performed as described for the native crystals.

2.6.4 Crystallization of Os3BGlu6 E178Q and E394D complexes with gibberellin GA₄ β-D-glucosyl ester

Purified Os3BGlu6 E178Q was concentrated to 3-5 mg/mL and hanging drop vapor diffusion crystallization trials were set up to optimize the initial crystallization condition (0.1 M Bis-Tris, pH 6.5, and PEG 5000MME; Seshadri *et al.*, 2009). For the crystal of complexes Os3BGlu6 E394D with gibberellin GA₄-β-D-glucosyl ester (GA₄GE), a 20 mM GA₄GE solution was prepared in the HPLC grade water, then was mixed to a final concentration of 2-5 mM of GA₄Glc with 3-5 mg/mL of Os3BGlu6 E394D (final concentration) protein solution and the mixture was incubated for 10 min at 4 °C. The hanging drops consisted of 2 μL of the protein substrate mixture solution and 1 μL of precipitant solution. Then, the drop was equilibrated with 0.5 mL of reservoir solution, and kept at 15 °C.

Hua *et al.* (2013) observed the release of free glucose from both GA₄GE and *p*NPGlc by the hydrolytic activity of Os3BGlu6 E178Q, but it had a slow rate when compared to the wild type Os3BGlu6 and Os3BGlu6 E178Q preferred transglycosylation of small nucleophiles with the glucosyl moiety over hydrolysis. This suggested a covalent intermediate might be built-up on the enzyme. The crystals of complexes of Os3BGlu6 E178Q with GA₄GE and *p*NPGlc were obtained from the apo Os3BGlu6 E178Q crystals, produced as described above with optimization conditions for crystallization trials that were the same as described for Os3BGlu6 E394D above.

2.6.5 Microseeding

The microseeding technique was performed to produce the large, single crystals. The small crystals or cluster crystals obtained from the first optimization conditions were used as a seed stock for microseeding. The crystals were removed from their initial drop, and then transferred to 10 μ L of reservoir solution. The crystal was crushed with a cat whisker, and then the suspension was added to a 1.5 mL microcentrifuge tube containing 100 μ L of reservoir solution. The tube was centrifuged for 5 min at 3000 rpm at room temperature, then the supernatant containing the microcrystals was serially diluted with mother liquor to 1/100, 1/500 and 1/1000 and the stock solution was stored at 4 °C. For crystal growth, the protein precipitant drop was pre-equilibrated with 0.5 mL of reservoir solution for 2 h at 15 °C. The equilibrated drops were streaked with the diluted microseed suspension on a cat whisker, and then incubated at 15 °C.

2.6.6 Soaking of crystal with ligand

To obtain complexes Os4BGlu12 E179Q with tuberonic acid glucoside (TAG) and salicylic acid glucoside (SAG), the crystals from co-crystallization with the same ligand were soaked in the precipitant containing 10 or 20 mM TAG or SAG, 10-30 min before flash vitrification. To obtain Os4BGlu12 crystals with 2-deoxy-2-fluoroglucofuranoside (G2F), the co-crystal with 2 mM DNP2FG was picked from the original drop with a nylon loop. The picked crystal was then soaked in the precipitant solution for 5-15 s without DNP2FG at room temperature on a glass cover slip. For the Os4BGlu12 complex with DNP2FG, Os4BGlu12 was crystallized in the presence of 2 mM DNP2FG, and the crystal was soaked in cryosolution containing 10

mM DNP2FG, and this protocol resulted in electron density for the noncovalent complex.

To try to obtain crystals of Os3BGlu6 E178Q with substrate, free Os3BGlu6 E178Q crystals were picked and then briefly (1-5 s) soaked in the precipitant solution containing either 10 mM *p*NPGlc or GA₄GE at room temperature. Os3BGlu6 E394D co-crystallized with GA₄GE was soaked in the cryosolution containing 10 mM GA₄GE 10-30 min at room temperature, before flash vitrification.

2.7 Data collection and processing

2.7.1 Flash vitrification

When the crystal was cooled in liquid nitrogen or during X-ray data collection, the temperature of the system is lower than the melting point of the precipitant solution. In this case the crystal might have been damaged by ice crystal formation. Prior to diffraction, the crystal was mounted in a nylon loop of the appropriate size (Hampton Research, 0.1-0.2 mm) and soaked briefly (5-15 s) in the fresh cryo-protectant, then flash cooled in liquid nitrogen (-180 °C) for X-ray diffraction. The cryo-protectant solution contained each of the crystallization precipitant components at a concentration increased from the original concentration. All the crystals of wild type Os4BGlu12 and Os4BGlu12 E179Q were soaked in the cryo-protectant including 21% increased of all precipitant components and 28% (v/v) glycerol, as described previously (Sansenya *et al.*, 2010). For the Os3BGlu6 mutant crystals, each crystallization precipitant component was increased 15% and 18% (v/v) glycerol was present in the cryo-solution, as described previously by (Seshadri *et al.*, 2009). In the case of crystals with ligands, the cryo-protectant contained the ligand at

the same concentrations in which the crystals had been soaked (Section 2.6.6).

2.7.2 Synchrotron X-ray diffraction

All crystals data sets reported in this thesis were collected with synchrotron radiation as the X-ray source and an ADSC Quantum 315 CCD (Charge-Coupled Device) detector at the BL13B1 beamline of the National Synchrotron Radiation Research Center (NSRRC), Hsinchu, Taiwan. The vitrified crystal in the cryo-loop was placed onto the goniometer head between the X-ray beam and the detector. During X-ray diffraction, the distance between the crystal and detector was 280-350 mm for Os4BGlu12 crystals and 250-300 mm for Os3BGlu6 crystals. The X-ray wavelength was set at 1.00 Å and crystals were kept at 100 K with a nitrogen stream from an Oxford Cryosystems Cryo-stream during X-ray diffraction. All crystal diffraction datasets were collected over 180° rotation or until the data were complete with 0.50° oscillations, and the exposure time per frame was 15-20 s for Os4BGlu12 and 5-10 s for Os3BGlu6. The maximum resolution limit of the diffraction data was judged based on the ratio of intensity measured to the background noise, referred to as I/σ and calculated as $I/\sigma = \Sigma(I - \sigma)/\Sigma\sigma$, where I is the intensity of a reflection and σ is the background noise surrounding it in the detector. When cutting of the resolution limit of the data, this ratio was kept (2 in the outer shell. The diffraction resolution was also limited by the overall quality of the data, which was assessed by the dataset completeness, average redundancy per shell and R_{merge} . In general, the overall linear R_{merge} was kept <12-14% and the R_{merge} in the outer shell below 35-50%, depending on the redundancy. All the diffraction images were indexed, integrated and scaled with the HKL-2000 program (Otwinowski and Minor, 1997).

2.8 Structure solution, refinement and validation

The crystal structure of the Os4BGlu12 apo enzyme was solved by molecular replacement with the *MOLREP* program (Vagin and Teplyakov, 1997) with the structure of cyanogenic β glucosidase from white clover (PDB code 1CBG, 63% amino acid sequence identity to Os4BGlu12) (Barrett *et al.*, 1995) as the search model. The crystals for the structures of Os4BGlu12 with DNP2FG, G2F and TAG were isomorphous with that for the apo Os4BGlu12 structure, allowing the structures to be solved by rigid body refinement with the native structure as a search model in the Refmac 5.0 program (Murdushov *et al.*, 1999; CCP4 suite). The three structures were also refined with Refmac 5.0. The analysis of the electron density and model building were done in the Coot program (Emsley and Cowtan, 2004). In the refinement and model building step, the structure factors, including the amplitudes from the reflection intensities and the phase estimates, were converted to electron density by Fourier transform (i.e. the FFT program of CCP4) or the automatic FFT calculation in Coot.

Subsequently, coot was used to manually rebuild the structural model. The $2m|F_{\text{obs}}|-d|F_{\text{calc}}|$ and $m|F_{\text{obs}}|-d|F_{\text{calc}}|$ Fourier maps were calculated from the observed structure factor amplitudes and the calculated structure factor amplitudes and phases. The model rebuilding was carried out based on these 2 maps, which were recalculated for each cycle of the rebuilding and refinement process. The atomic positions and B-factors of all atoms were refined to fit the observed diffraction data. The water molecules were added with the ARP/wARP program in the CCP4 suite. The water molecules were added or deleted based on the electron density ($2mF_o-dF_c$ map contoured at 1.5-2.0 σ), B-factor and hydrogen bonding distances between the

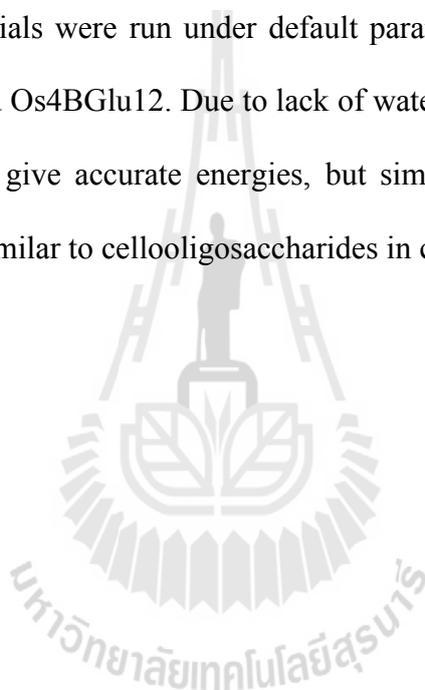
water molecules and neighboring amino acid residues. The occupancies of the DNP2FG molecules binding in each of the active sites of the Os4BGlu12/DNP2FG structure were refined to 0.8, by setting the occupancy at different values and refining the structures to determine which value gave the minimum values of the temperature factors (B-factors) of the ligand and R_{free} , and the glucosyl residue was refined to a 1S_3 conformation (the initial conformation was similar to 4C_1 , taken from *Bacillus polymyxa* (PDB: 1E4I) in which the whole DNP2FG was bound at the surface of this structure, not in the -1 subsite), which fit into the density at 0.8 occupancy. A G2F ring in the 4C_1 conformation was fit to the occupancy of 1 in the electron density map of the Os4BGlu12G2F complex structure. And the glucosyl residue of Os4BGlu12E178Q with TAG was refined as 1S_3 to the density, based on this conformation having the lowest of temperature factor (B-factor), when the 4C_1 and 1S_3 conformations were tried.

The structure of Os3BGlu6 E178Q with *p*NPG and Os3BGlu6 E178Q with GA₄GE were solved via rigid body refinement with the wild type Os3BGlu6 structure (PDB_3GNO; Seshadri *et al.*, 2009) from which the waters and heteroatoms were removed as a template model. Manual rebuilding of the model was performed in Coot and refinement was done with Refmac 5.0. The glucosyl residues at the -1 subsite of the structures of Os3BGlu6 E178Q with *p*NPGlc and Os3BGlu6 E178Q with GA₄GE were refined as 4C_1 , based on its fit into the electron density for both structures. The overall quality of each model was evaluated with the PROCHECK and MolProbity programs (Laskowski *et al.*, 1993 and Davis *et al.*, 2007). The refinement and model statistics are given in Tables 3.4, 3.5 and 3.7. The structural superimposition and figures were produced with PyMol (Schrödinger LLC, Portland,

OR, USA).

2.9 Docking studies

Docking trials were run in Autodoc 4.0 using the Lamarckian genetic algorithm (Morris *et al.*, 1998) with cellotriose with the nonreducing sugar in the 1S_3 skew boat configuration, as previously described for Os3BGlu7 (Chuenchor *et al.*, 2008). Ten docking trials were run under default parameters from Autodock 4.0 for each of Os3BGlu6 and Os4BGlu12. Due to lack of water in the simulation, these trials were not expected to give accurate energies, but simply to show the feasibility of binding in positions similar to celooligosaccharides in crystal structures of Os3BGlu7.



CHAPTER III

RESULTS

3.1 Protein expression and purification

3.1.1 Purification of recombinant Os4BGlu12

The thioredoxin/His₆-tagged Os4BGlu12 fusion protein was expressed in the soluble fraction in *E. coli* strain Origami B(DE3) cells from the pET32a+/DEST-*Os4BGlu12* expression vector under the optimized conditions (Opassiri *et al.*, 2006). The thioredoxin-Os4BGlu12 fusion protein contains a His₆-tag and a cleavage site for enterokinase following the thioredoxin at the N-terminus of the mature Os4BGlu12 protein. When the recombinant fusion protein was purified by a single step of IMAC on a Co²⁺ resin column, a major band of fusion protein at 69 kDa and many nonspecifically binding and contaminating proteins were seen on the SDS-PAGE (Figure 3.1). The yield of protein obtained from the elution fractions that contained β -glucosidase activity was approximately 5 mg protein per liter of *E. coli* culture. To remove the fusion tag at the N-terminus, which might interfere with crystallization, the protein obtained from the first IMAC was cleaved with enterokinase to generate the tag-free Os4BGlu12 protein. The thioredoxin fusion tag was removed from the protein solution by adsorption to Co²⁺ bound resin in the second IMAC step. The purification profile of tag-free Os4BGlu12 protein obtained from the second IMAC is shown in Figure 3.2. The fractions contain tag-free Os4BGlu12 protein, which was

detected at approximately 55 kDa on SDS-PAGE, along with different amounts of minor contaminating protein bands. These fractions were pooled and concentrated, and the buffer was changed to 20 mM Tris-HCl, pH 7.2. After the second IMAC step, the Os4BGlu12 protein was obtained with >95% purity, with a yield of about 2.5 mg per liter of culture. This protein is compared to that from the first IMAC in Figure 3.3. The protein from this second IMAC was used for crystallization screening and the production of sufficient crystals in the optimization experiment.

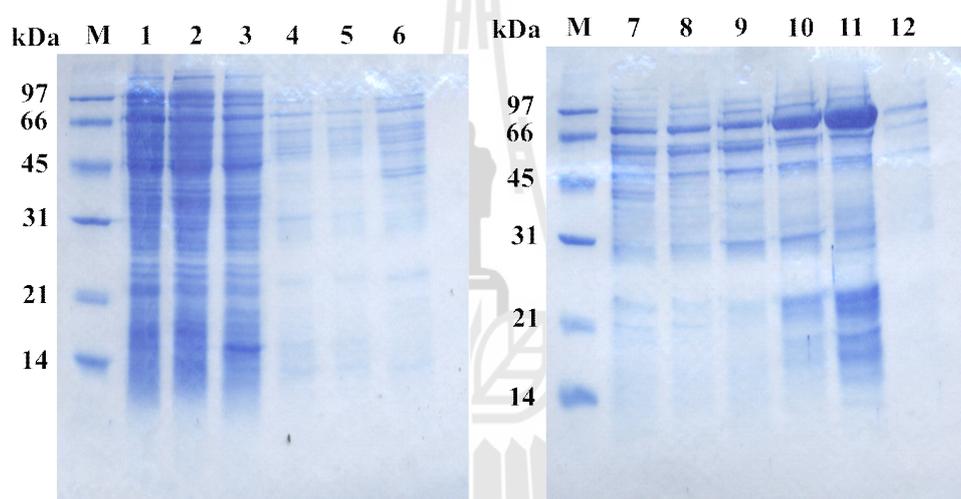


Figure 3.1 SDS-PAGE of fractions from purification of recombinant thioredoxin-Os4BGlu12 protein by Co^{2+} -IMAC.

Lane M, molecular weight protein marker. Lane 1, soluble fraction of induced *E. coli* cells containing pET32a+/DEST-*Os4bglu12*. Lane 2, flow-through of IMAC column. Lane 3, the protein eluted in the wash with equilibration buffer. Lanes 4-6, the protein fractions eluted in the wash with 5 mM imidazole. Lanes 7-9, the protein fractions eluted in the wash with 10 mM imidazole. Lanes 10-12, the elution fractions of thioredoxin-Os4BGlu12 protein eluted with 250 mM imidazole.

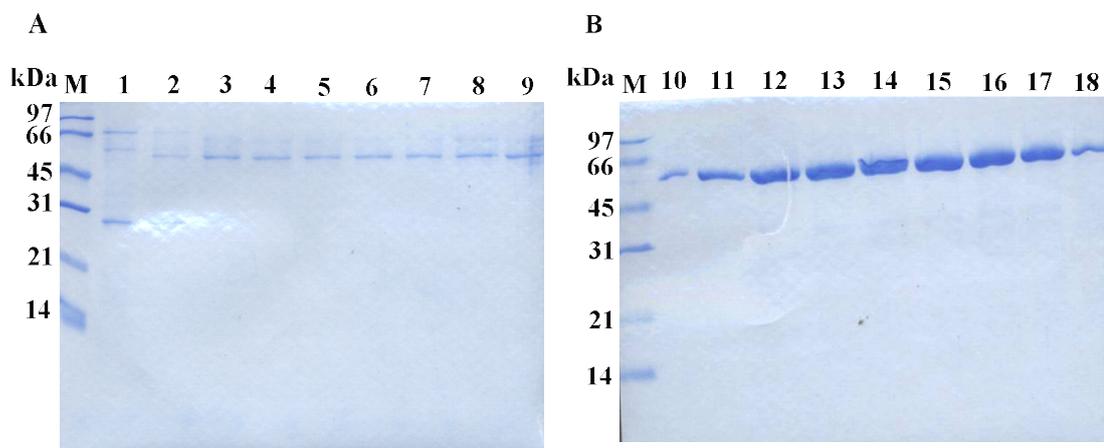


Figure 3.2 SDS-PAGE of tag-free Os4BGlu12 protein fractions purified via a second IMAC step.

The purified fusion protein was cleaved with enterokinase and purified by Co^{2+} -IMAC to remove the fusion tag. Lane M, molecular weight protein marker. A. Lane 1, flow-through. Lane 2, fraction washed from the column with equilibration buffer. Lanes 3-9, the protein fractions eluted with 5 mM imidazole. B. Lanes 10-18, the protein fractions eluted with 10 mM imidazole.

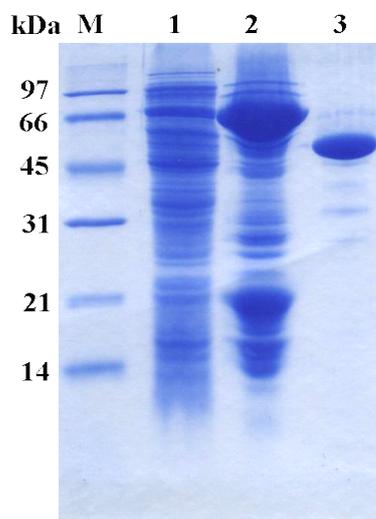


Figure 3.3 SDS-PAGE analysis of Os4BGlu12 protein from two steps of purification. Lane M, molecular weight protein maker. Lane 1, soluble protein extract of *E. coli* cells expressing Os4BGlu12. Lane 2, thioredoxin-Os4BGlu12 fusion protein after initial IMAC. Lane 3, tag- free Os4BGlu12 after second IMAC.

3.1.2 Purification of Os4BGlu12 E179Q and H252M

Purification of the Os4BGlu12 mutants Os4BGlu12 E179Q and Os4BGlu12 H252M was also carried out using the same two-step of immobilized metal affinity chromatography (IMAC) procedure, with similar yields of approximately 3-5 mg of highly purified protein per liter of bacterial culture. The purification profiles of the two mutants are show in Figure 3.4.

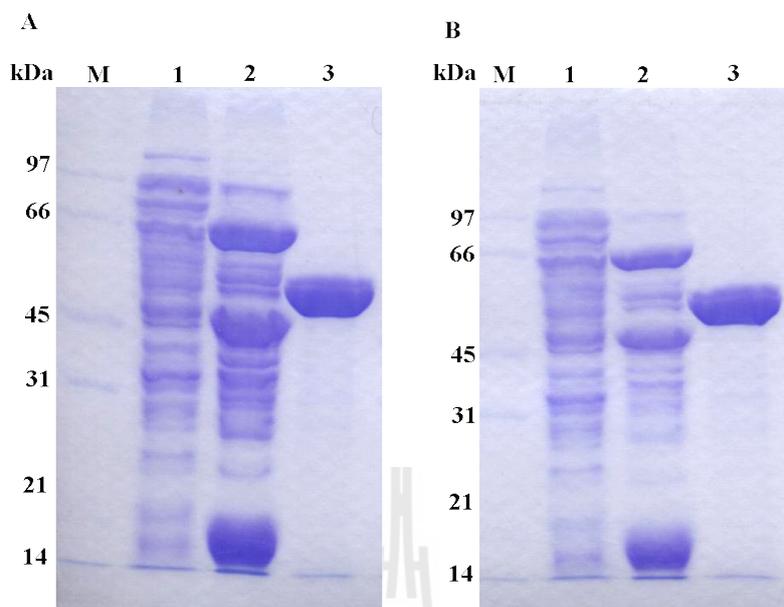


Figure 3.4 The purification profile of the two recombinant Os4BGlu12 mutant proteins E179Q (A) and H252M (B).

Purification fraction pools with the two steps of IMAC on Co^{2+} IMAC resin are shown.

Lane M, molecular weight protein maker. Lane 1, crude protein extract of Origami B(DE3) cells expressing the protein. Lane 2, the fusion protein after initial IMAC. Lane 3, protein after cleavage of tag and second IMAC.

3.1.3 Purification of Os3BGlu6 E178Q and M251N mutant proteins

The optimized expression and purification of wild type Os3BGlu6 recombinant protein was previously described by Seshadri *et al.*, 2009. The initial purification of two Os3BGlu6 mutants, E178Q and M251N, was also done by IMAC on a Co^{2+} -IMAC column, followed by TEV protease digestion to cleave between the thioredoxin fusion-tag and the Os3BGlu6 mutant proteins. The fusion-tag and TEV

protease were then removed by binding them to the IMAC column. The results obtained from large scale purification with these steps are shown in Figure 3.5. The overall yield obtained from 3.2 liters of bacterial culture was approx. 15-20 mg of the approximately 95 % pure Os3BGlu6 protein.

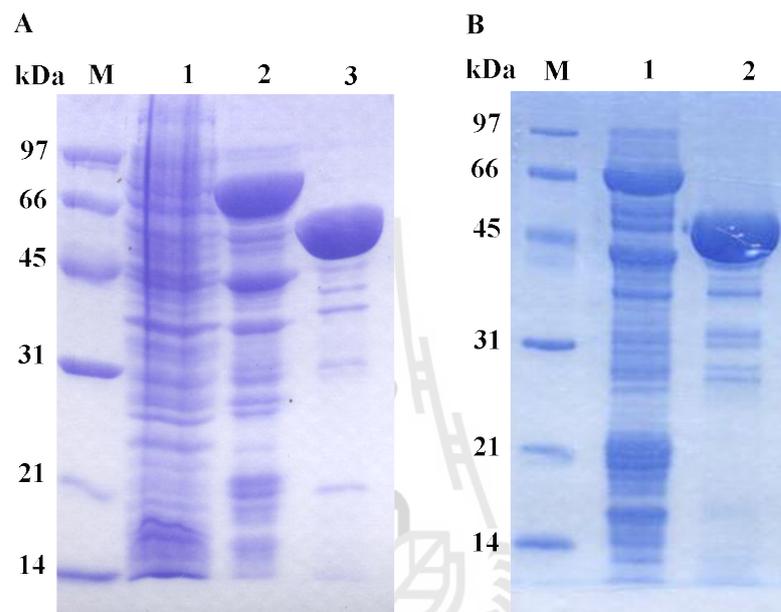


Figure 3.5 SDS page profiles of the Os3BGlu6 E179Q and M251N mutants after the first and second IMAC steps.

Panel A, Os3BGlu6 E178Q purification fractions, Lane 1, crude protein extract of induced Origami(DE3) cells. Lane 2, the recombinant thioredoxin fusion-tag protein after 1st IMAC column. Lane 3, non-fusion protein after it was cut with TEV and passed through the second IMAC column. Panel B, Os3BGlu6 M251N purification fractions. Lane 1, fusion protein after the 1st IMAC column. Lane 2, tag-free protein after it was cut with TEV and passed through a second IMAC column. Lane M is the molecular weight protein markers in both A and B.

3.2 Substrate specificity of Os4BGlu12

The of hydrolysis activity of tag-free Os4BGlu12 on *p*NP-glycosides and *o*NP- β -D-glucoside were tested, as shown in Table 3.1. Of the *p*NP derivatives, Os4BGlu12 showed highest activity for hydrolysis of *p*NP- β -D-fucoside (96.7 ± 9.9 $\mu\text{mol}/\text{min}/\text{mg}$ protein), while the hydrolytic efficiency for *o*NP- β -D-glucoside (23.0 ± 1.6 $\mu\text{mol}/\text{min}/\text{mg}$ protein) was approximate 3 times lower than that for *p*NP- β -D-glucoside (60.7 ± 1.1 $\mu\text{mol}/\text{min}/\text{mg}$ protein). Os4BGlu12 had the lowest hydrolysis activity for *p*NP- β -D-mannoside (0.37 ± 0.03 $\mu\text{mol}/\text{min}/\text{mg}$ protein). The hydrolysis of synthetic and natural glycosides and cello-oligosaccharides is also shown in Table 3.1. Among the β -(1,4)-linked oligosaccharides tested, Os4BGlu12 had the best activity toward cellotetraose (26.1 ± 1.2 $\mu\text{mol}/\text{min}/\text{mg}$ protein), followed by cellotriose (15.5 ± 0.49 $\mu\text{mol}/\text{min}/\text{mg}$ protein) and cellobiose (0.520 ± 0.003 $\mu\text{mol}/\text{min}/\text{mg}$ protein), respectively. In this study the hydrolysis of Os4BGlu12 toward β -(1,4)-linked oligosaccharides was only tested for DP 2 to 4, while Opassiri *et al.*, (2010) showed that Os4BGlu12 had the best hydrolytic efficiency (k_{cat}/K_m) on cellopentaose (4.9 ± 0.1 $\text{s}^{-1}\text{mM}^{-1}$), which was higher than cellotetraose (3.0 ± 0.2 $\text{s}^{-1}\text{mM}^{-1}$), and the catalytic efficiency was lower for cellohexaose (3.8 ± 0.2 $\text{s}^{-1}\text{mM}^{-1}$). Os4BGlu12 hydrolyzed the β -(1,3)-linked disaccharide laminaribiose (28.8 ± 0.24 $\mu\text{mol}/\text{min}/\text{mg}$ protein) best, with lower activities toward the β -(1,2)-linked disaccharide sophorose (4.28 ± 0.33 $\mu\text{mol}/\text{min}/\text{mg}$ protein) and the β -(1,6)-linked disaccharide gentiobiose (0.450 ± 0.006 $\mu\text{mol}/\text{min}/\text{mg}$ protein). Os4BGlu12 also hydrolyzed other glycosides, including helicin, *meta*-salicylic acid glucoside (*m*-SAG), *para*-salicylic acid glucoside (*p*-SAG), salicylic acid glucoside (*ortho*-salicylic

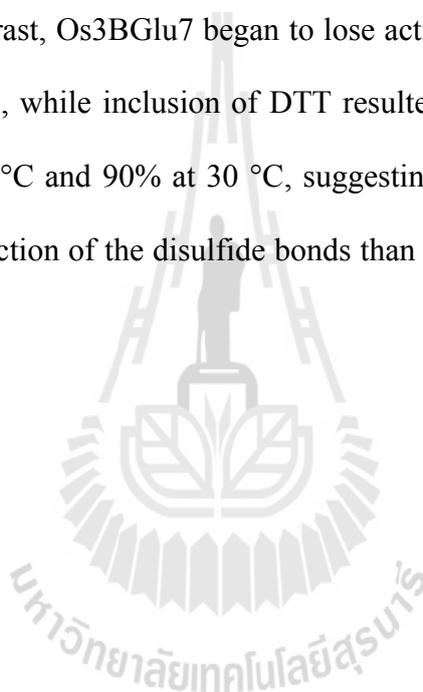
acid, SAG), tuberonic acid glucoside (TAG) and salicin. All the hydrolysis activities shown in Table 3.1 are higher than previously reported by Opassiri *et al.* (2010), but showed a similar trend for the hydrolysis efficiently toward various substrates. Opassiri *et al.* (2010) characterized hydrolysis by the fusion protein that had only one step of purification (1st IMAC), while in this study the tag-free protein was prepared with two steps of purification (1st and 2nd IMAC).

Table 3.1 Substrate specificity of Os4BGlu12 for the Hydrolysis of various substrates

Substrate	$\mu\text{mol min}^{-1} \text{mg}^{-1} \text{protein}$	Percent Relative activity
<i>p</i> NP β -D-Glucoside	60.7 \pm 1.1	100
<i>o</i> NP β -D-Glucoside	23.0 \pm 1.6	33
<i>p</i> NP β -D-Fucoside	96.7 \pm 9.9	159.3
<i>p</i> NP β -D-Galactoside	43.0 \pm 4.5	70.8
<i>p</i> NP β -D-Xyloside	13.4 \pm 0.32	22.1
<i>p</i> NP β -D-Mannoside	0.37 \pm 0.03	0.6
TAG	50.1 \pm 1.0	82.5
Helicin	186.1 \pm 3.0	309.6
Salicin	0.16 \pm 0.009	0.3
SAG	60.3 \pm 1.4	99.3
<i>m</i> -SAG	61.0 \pm 3.3	100.5
<i>p</i> -SAG	104.7 \pm 4.9	172.5
Cellobiose	0.52 \pm 0.003	0.9
Cellotriose	15.5 \pm 0.49	25.5
Cellotetraose	26.1 \pm 1.2	43
Sophorose	4.28 \pm 0.33	7
Laminaribiose	28.8 \pm 0.24	47.4
Gentiobiose	0.45 \pm 0.006	0.7

3.3 Stability study comparing two rice β -glucosidase enzymes

Figure 3.6 shows the activity levels for Os4BGlu12 (Figure 3.6 A) and Os3BGlu7 (Figure 3.6 B) after preincubation at temperatures of 0-90 °C for 60 min. The Os4BGlu12 activity remained stable after incubation at 20-50 °C for 60 min, whereas the activity dropped about 5% at 60 °C after 60 min. Inclusion of DTT in the enzyme solution resulted in decreases in the enzyme activity of 26% at 40 °C, and 80% at 60 °C. In contrast, Os3BGlu7 began to lose activity at 40 °C and was inactive at higher temperatures, while inclusion of DTT resulted in a decrease in the enzyme activity of 25% at 10 °C and 90% at 30 °C, suggesting that Os3BGlu7 stability was more sensitive to reduction of the disulfide bonds than that of Os4BGlu12 (Figure 3.6 A and B).



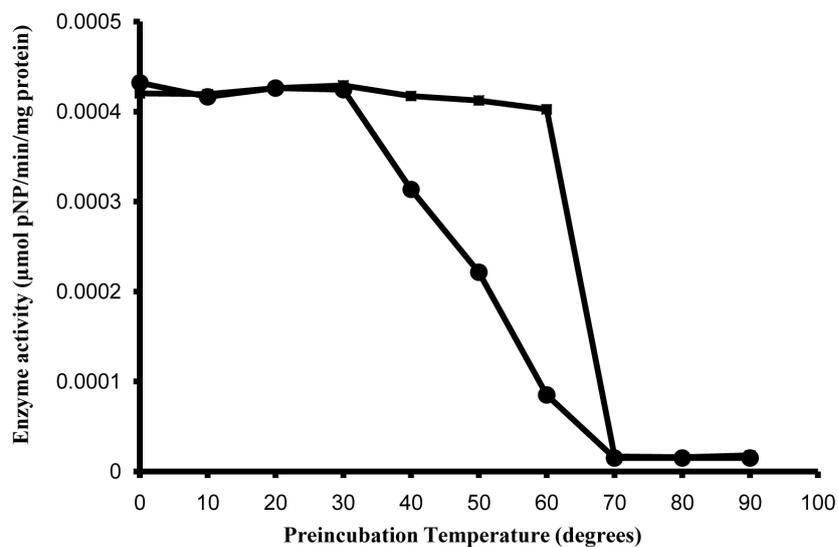
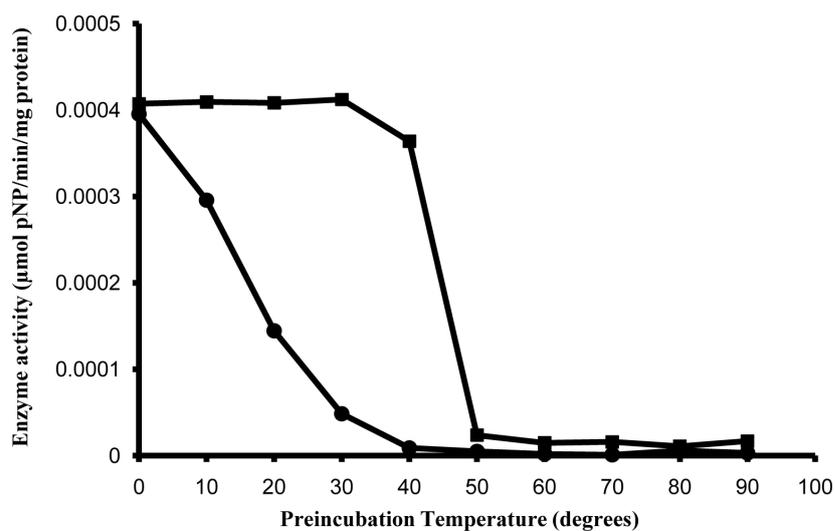
Os4BGlu12**Os3BGlu7**

Figure 3.6 Comparison of the thermostabilities of Os4BGlu12 and Os3BGlu7.

The enzyme activities were measured in the standard assay (0.3 mM *p*NPGlc in 50 mM sodium acetate, pH 5.0, for 10 min) after incubation of enzyme with (●) and without (■) 1 mM of DTT for 1 h at the designated temperatures in 50 mM sodium acetate, pH 5.0.

3.4 Crystals of apo-Os4BGlu12

The initial condition for crystallization of wild type apo Os4BGlu12 was previously identified by Sansenya (2008), by screening via the microbatch under oil technique. Small crystals were observed with a precipitant consisting of 25% (w/v) PEG 4000 in 0.1 M Tris-HCl, pH 8.5, 0.2 M NaCl within 17 days (Sansenya *et al.*, 2010). Optimization for this condition was carried out by hanging-drop vapor diffusion with microseeding and 3-9 mg/mL protein concentrations (Sansenya, 2008). The suitable apo-Os4BGlu12 crystals were obtained in 19% (w/v) PEG 3350 in 0.1 M Tris-HCl, pH 8.5, 0.16-0.26 M NaCl with 3.5 mg/mL protein concentration (Figure 3.7).

3.5 Crystals of Complexes of Os4BGlu12 with DNP2FG and G2F

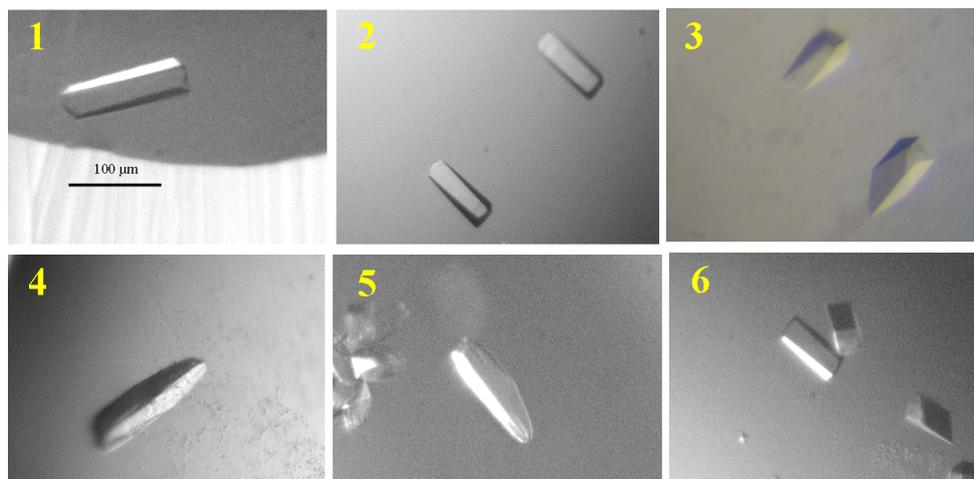
The crystals with DNP2FG and G2F were crystallized in the presence of 2 mM of DNP2FG. The suitable condition for both complexes was optimized starting from the condition used for the initial apo-Os4BGlu12 crystals with 3.5 mg/mL protein by hanging-drop vapor diffusion with microseeding. The maximum size crystals were obtained within 10-17 days in the optimized condition (Figure 3.8).

3.6 Crystals of Os4BGlu12E179Q with TAG

The inactive Os4BGlu12 E179Q variant was crystallized from a solution of 3.5 mg/mL protein with 2 mM of TAG in the condition optimized from that of the initial apo-Os4BGlu12 crystallization condition by hanging-drop vapor diffusion with microseeding. High quality crystals were obtained from several conditions within 10-20 days (Figure 3.9).

3.7 Crystals of apo Os3BGlu6 E178Q

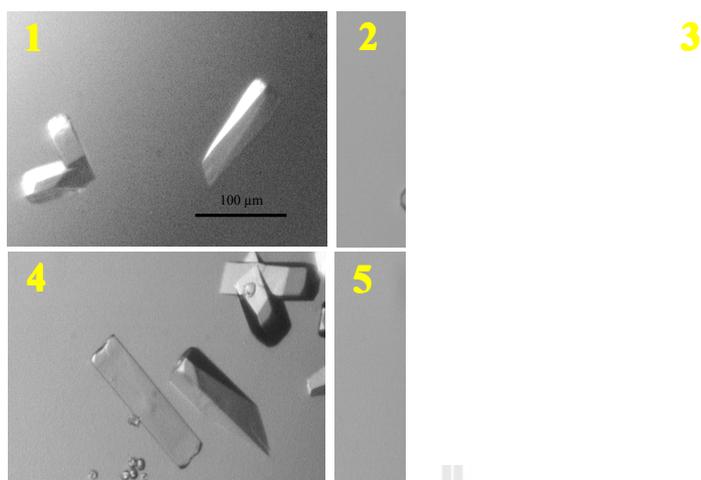
Seshadri *et al.* (2009) reported three crystal structures of apo Os3BGlu6 and its complexes with inhibitors. The crystallization conditions of Os3BGlu6 E178Q with *p*NPGlc and GA₄GE were optimized from the conditions in that previous report. Hanging drop vapor diffusion trials were run with 5 mg/mL of pure Os3BGlu6 E178Q and 13% PEG5000MME, 0.1 M Bis/Tris at different pH with microseeding. The crystals with maximum size grew within 2-5 days in the optimization conditions with different pH (Figure 3.10).



Condition number	Chemical composition	Maximum dimension (μm)
1	19% (w/v) PEG 3350 in 0.1 M Tris-HCl, pH 8.5, and 0.16 M NaCl	120 x 25 x 20
2	19% (w/v) PEG 3350 in 0.1 M Tris-HCl pH 8.5 and 0.18 M NaCl	80x 15 x 15
3	19% (w/v) PEG 3350 in 0.1 M Tris-HCl pH 8.5 and 0.20 M NaCl	80 x 30 x 30
4	19% (w/v) PEG 3350 in 0.1 M Tris-HCl pH 8.5 and 0.22 M NaCl	125x 30 x 25
5	19% (w/v) PEG 3350 in 0.1 M Tris-HCl pH 8.5 and 0.24 M NaCl	120x 25 x 20
6	19% (w/v) PEG 3350 in 0.1 M Tris-HCl pH 8.5 and 0.24 M NaCl	75x 20 x 20

Figure 3.7 Photographs of apo-Os4BGlu12 crystals produced in various NaCl concentrations.

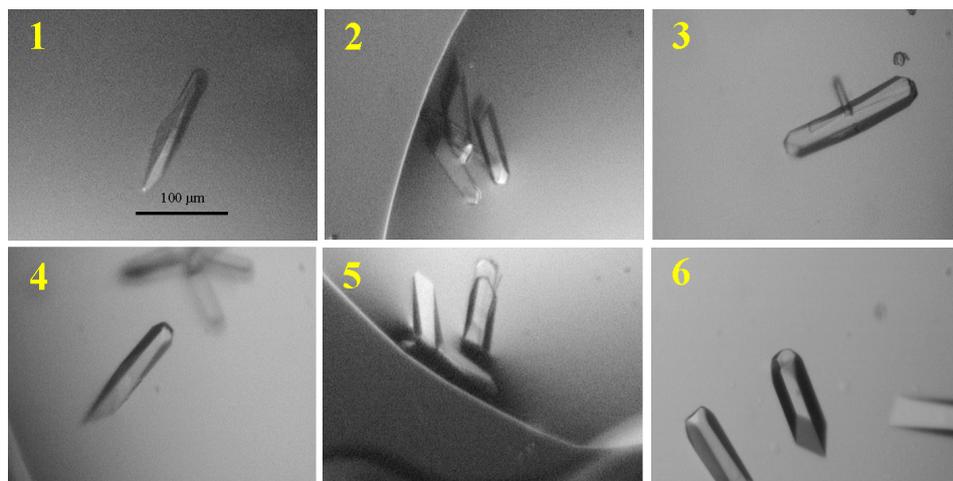
The table shows the composition of the precipitants in the conditions that produced the crystals. In all the conditions, small crystals were observed within 1 day and the maximum size crystals were observed within 10-15 days.



Condition number	Chemical composition	Maximum dimension (μm)
1	19% (w/v) PEG 2000 in 0.1 M Tris-HCl, pH 8.5, and 0.16 M NaCl	120 x 25 x 25
2	19% (w/v) PEG 2000 in 0.1 M Tris-HCl, pH 8.5, and 0.18 M NaCl	145 x 20 x 20
3	19% (w/v) PEG 2000 in 0.1 M Tris-HCl, pH 8.5, and 0.20 M NaCl	110 x 15 x 20
4	19% (w/v) PEG 2000 in 0.1 M Tris-HCl, pH 8.5, and 0.24 M NaCl	110 x 30 x 25
5	19% (w/v) PEG 2000 in 0.1 M Tris-HCl, pH 8.5, and 0.26 M NaCl	120 x 35 x 25

Figure 3.8 Photographs of Os4BGlu12 crystals produced in co-crystallization with 2 mM DNP2FG in various NaCl concentrations.

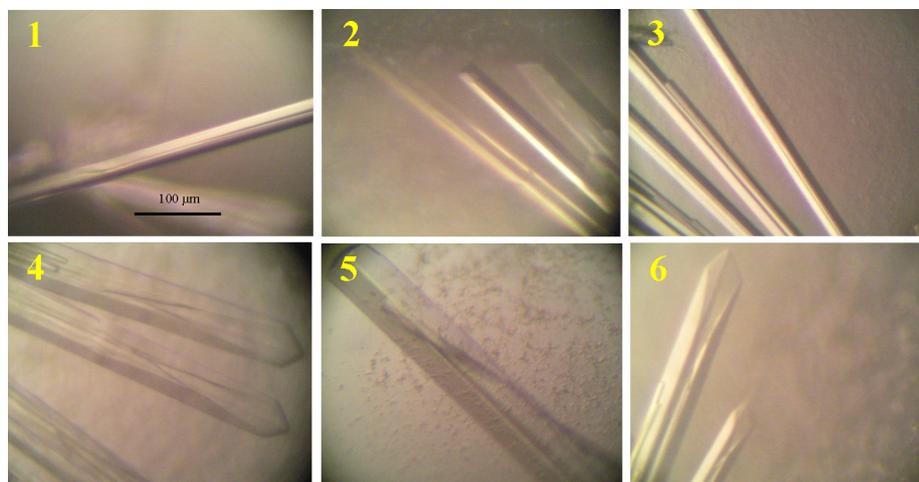
The table shows the composition of the precipitants in different conditions. In all the conditions, small crystals were observed within 1-2 day and the maximum size crystals were observed within 10-17 days.



Condition number	Chemical composition	Maximum dimension (μm)
1	19% (w/v) PEG 4000 in 0.1 M Tris-HCl, pH 8.5, and 0.16 M NaCl	120 x 20 x 20
2	19% (w/v) PEG 4000 in 0.1 M Tris-HCl, pH 8.5, and 0.18 M NaCl	95 x 20 x 20
3	19% (w/v) PEG 4000 in 0.1 M Tris-HCl, pH 8.5, and 0.20 M NaCl	110 x 15 x 20
4	19% (w/v) PEG 4000 in 0.1 M Tris-HCl, pH 8.5, and 0.22 M NaCl	110x 30 x 25
5	19% (w/v) PEG 4000 in 0.1 M Tris-HCl, pH 8.5, and 0.24 M NaCl	120x 35 x 25
6	19% (w/v) PEG 4000 in 0.1 M Tris-HCl, pH 8.5, and 0.26 M NaCl	120x 35 x 25

Figure 3.9 Photographs of crystals obtained from co-crystallization of Os4BGlu12 E179Q with 2 mM TAG in various NaCl concentrations.

The table shows the composition of precipitant in different conditions. In all the conditions, small crystals were observed within 1-2 day and the maximum size crystals were observed within 10-20 days.



Condition number	Chemical composition	Maximum dimension (μm)
1	13% PEG5000MME, 0.1 M Bis/Tris, pH 5.5	500 x 20 x 20
2	13% PEG5000MME, 0.1 M Bis/Tris, pH 6.2	200 x 35 x 35
3	13% PEG5000MME, 0.1 M Bis/Tris, pH 6.4	500 x 20 x 20
4	13% PEG5000MME, 0.1 M Bis/Tris, pH 6.5	500 X 55 X 40
5	13% PEG5000MME, 0.1 M Bis/Tris, pH 6.7	450 x 50 x 20
6	13% PEG5000MME, 0.1 M Bis/Tris, pH 6.9	350 x 45 x 20

Figure 3.10 Apo Os3BGlu6 E178Q crystals developed in precipitants with different pH values.

The table shows the composition of each precipitant. In all the conditions, small crystals were observed within 1 day and the crystals achieved maximum size within 2 days.

Crystals obtained from the above conditions diffracted X-rays from a synchrotron radiation source at the BL13B1 beamline of the National Synchrotron Radiation Research Center (NSRRC), Taiwan. The data sets with the resolutions high enough to produce quality electron density maps were used to solve the three dimensional structures, as summarized in Tables 3.2 and 3.3.

Table 3.2 The Os4BGlu12 crystallization conditions and X-ray diffraction.

Crystals	Precipitant composition	Cryo-solution	Soaking times (s)	Resolution (Å)
Apo-Os4BGlu12	19% (w/v) PEG 3350 in 0.1 M Tris-HCl, pH 8.5, and 0.16 M NaCl	23% (w/v) PEG 3350, 0.12 M Tris-HCl, pH 8.5, 0.19 M NaCl in 18% (v/v) Glycerol	5	2.50
Os4BGlu12 with G2F	19% (w/v) PEG 2000 in 0.1 M Tris-HCl, pH 8.5, and 0.18 M NaCl	23% (w/v) PEG 2000, 0.12 M Tris-HCl, pH 8.5, 0.19 M NaCl in 18% (v/v) glycerol	15	2.40
Os4BGlu12 with DNP2FG	19% (w/v) PEG 2000 in 0.1 M Tris-HCl, pH 8.5, and 0.16 M NaCl	10 mM DNP2FG in 23% (w/v) PEG 2000, 0.12 M Tris-HCl, pH 8.5, 0.19 M NaCl with 18% (v/v) glycerol	15	2.45
Os4BGlu12 E179Q with TAG	19% (w/v) PEG 4000 in 0.1 M Tris-HCl, pH 8.5, and 0.26 M NaCl	10 mM TAG in 23% (w/v) PEG 4000, 0.12 M Tris-HCl, pH 8.5, 0.19 M NaCl with 18% (v/v) glycerol	900	3.20

Table 3.3 Os3BGlu6 crystallization conditions and X-ray diffraction experiments.

Crystals	Precipitant composition	Cryo-solution	Soaking times (s)	Resolution (Å)
Os3BGlu6	13% PEG5000MME,	2 mM ρ NPGlc in	5	1.90
E178Q_	0.1 M Bis/Tris, pH	22% PEG5000MME,		
ρ NPGlc	6.5	0.2 M Bis/Tris, pH 6.5, with 18% (v/v) glycerol		
Os3BGlu6	13% PEG5000MME,	10 mM GA ₄ GE in	120	2.40
E178Q_	0.1 M Bis/Tris, pH	22% PEG5000MME,		
GA ₄ GE (2 min)	6.4	0.2 M Bis/Tris, pH 6.4, with 18% (v/v) glycerol		
Os3BGlu6	13% PEG5000MME,	10 mM GA ₄ GE in	180	1.97
E178Q_	0.1 M Bis/Tris, pH	22% PEG5000MME,		
GA ₄ GE (3 min)	6.9	0.2 M Bis/Tris, pH 6.9, with 18% (v/v) glycerol		
Os3BGlu6	13% PEG5000MME,	10 mM GA ₄ GE in	600	2.50
E178Q_	0.1 M Bis/Tris, pH	22% PEG5000MME,		
GA ₄ GE (10 min)	5.5	0.2 M Bis/Tris, pH 5.5, with 18% (v/v) glycerol		

3.8 Os4BGlu12 structure determination

3.8.1 X-ray analysis

The crystals of apo Os4BGlu12 and its complexes with DNP2FG and G2F and of Os4BGlu12 E179Q with TAG diffracted X-rays to give data sets with 2.50, 2.45, 240 and 3.20 Å resolutions, respectively. The statistics for all the X-ray data sets are summarized in Table 3.4. The diffraction patterns showed that the apo and complex crystals had tetragonal unit cells with similar unit-cell parameters of $a = b = 112.7$, $c = 182.8$ Å ($\alpha = \beta = \gamma = 90^\circ$) (apo Os4BGlu12), $a = b = 114.1$, $c = 184.5$ Å, ($\alpha = \beta = \gamma = 90^\circ$) for Os4BGlu12 with DNP2FG, $a = b = 114.0$, $c = 184.1$ Å, ($\alpha = \beta = \gamma = 90^\circ$) for Os4BGlu12 with G2F and $a = b = 113.2$, $c = 183.4$ Å, ($\alpha = \beta = \gamma = 90^\circ$) for Os4BGlu12 E179Q with TAG. The systematic absences were consistent with either space group $P4_12_12$ or $P4_32_12$. Two protein molecules were found in the asymmetric unit with a Matthews coefficient (V_m) of $2.46 \text{ \AA}^3 \text{ Da}^{-1}$ and solvent content of 50% for apo Os4BGlu12, while the crystals of Os4BGlu12 with DNP2FG and G2F and had solvent contents of 53–55%, due to their slightly larger unit cell volumes. The structure of Os4BGlu12 was solved by molecular replacement with the MOLREP program (Vagin & Teplyakov, 1997) with the cyanogenic β -glucosidase from white clover structure (PDB code 1CBG; Barrett *et al.*, 1995) as the search model to give an initial R_{factor} of 0.450, an R_{free} of 0.466 and a figure of merit (FOM) of 0.502. The space group of the two crystals was determined to be $P4_32_12$ because it gave better R values than the solution in the $P4_12_12$ space group. Further refinement with inclusion of water molecules and other heteroatoms found in the solvent gave the final residual values of $R = 0.203$ and $R_{\text{free}} = 0.248$ (Table 3.5).

The crystals of Os4BGlu12 complexed with DNP2FG and G2F and Os4BGlu12 E179Q with TAG were isomorphous with the apo Os4BGlu12 crystals. Rigid-body refinement was used as a simple molecular replacement protocol with the apo Os4BGlu12 structure from which the heteroatoms had been removed as a search model to give the initial complex structures. The Os4BGlu12/DNP2FG complex data yielded a solution with an initial R_{factor} of 0.388, an R_{free} of 0.390 and an FOM of 0.794. The solution of rigid-body refinement of the apo Os4BGlu12 structure into the Os4BGlu/G2F data yielded an initial R_{factor} of 0.380, an R_{free} of 0.390 and an FOM of 0.808, while that for the Os4BGlu E179Q/TAG structure factors gave an initial R_{factor} of 0.302, an R_{free} of 0.305 and an FOM of 0.763.

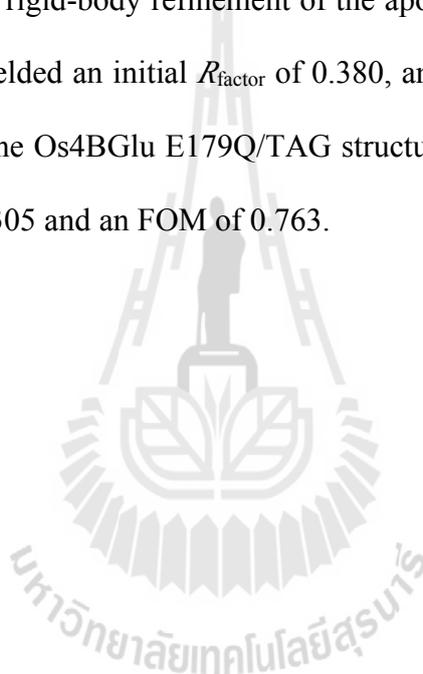


Table 3.4 Data collection statistics for Os4BGlu12 data sets.

Dataset	Os4BGlu12	Os4BGlu12_ DNP2FG	Os4BGlu12_ _G2F	Os4BGlu12_ E179Q_ TAG
Beamline	BL13B1	BL13B1	BL13B1	BL13B1
Wavelength (Å)	1.00	1.00	1.00	1.00
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Unit-cell parameters (Å)	$a = 112.6$ $b = 112.6$ $c = 182.7$	$a = 114.1$ $b = 114.1$ $c = 184.5$	$a = 114.0$ $b = 114.0$ $c = 184.1$	$a = 113.2$ $b = 113.2$ $c = 183.4$
Resolution range (Å)	30-2.50	30-2.45	30-2.40	30-3.20
Resolution outer shell (Å)	2.50-2.49	2.51-2.45	2.49-2.40	3.31-3.20
No. Unique reflections	39533	43131	48334	20392
No. Observed reflections	408534	381316	329557	181617
Completeness (%) ^a	100.0 (99.9)	100.0 (99.9)	99.6 (99.9)	100.0 (99.9)
Average redundancy ^a per shell	9.7 (10.0)	8.4 (8.6)	6.8 (6.2)	8.9 (9.2)
$I/\sigma(I)$ ^a	19.6 (6.4)	21.7 (4.4)	17.5 (3.4)	12.7 (3.8)
$R_{(merge)}$ (%) ^{ab}	10.6 (41.7)	9.4 (49.9)	10.4 (53.7)	15.1 (52.6)

^a Values in parentheses refer to the corresponding values of the highest resolution shell, from 10 shells of data. ^b $R_{merge} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \{I(hkl)\}|}{\sum_{hkl} \sum_i I_i(hkl)}$ where $I_i(hkl)$ is the intensity of the i th measurement of reflection hkl and $\{I(hkl)\}$ is the mean value of $I_i(hkl)$ for all i measurements.

3.8.2 Model quality and overall structure

The refinement and model quality statistics for all the Os4BGlu12 structures are shown in Table 3.5. In general, the electron density of almost all residues was seen clearly in the final models of all structures. The electron density for residues 9-486 were observed in the structures of the apo enzyme and its complexes with and DNP2FG and G2F and in that of the Os4BGlu12 E179Q acid/base mutant with TAG, but eight residues from Os4BGlu12, including Ala1, Tyr2, Asn3, Ser4, Ala5, Gly6, Glu7 and Pro8, and 19 amino acid residues remaining from the N-terminal fusion tag (Ala-Met-Ala-Asp-Ile-Thr-Ser-Leu-Tyr-Lys-Lys-Ala-Gly-Ser-Ala-Ala-Ala-Pro-Phe) were not visible. At the C-terminus, the side chains of Lys486 in all four structures had poorly defined electron densities. Residues Ser333, Asn334, Gly335, Leu336 and Asn337 in the loop C region had weak density, especially for molecule B, and the surface residues Arg76, Glu108, Glu228, Glu306 and Ile355 did not have defined electron densities for their side chains in all four structures. In fact, the poorly defined part of loop C sits at the entrance to the active site and was found to be cleaved between Asn334 and Gly335 in the protein purified from rice panicles, the same position as proteolysis of OsTAGG1 (Wakuta *et al.*, 2010, 2011), although no evidence of cleavage at this site was seen in the crystallized recombinant protein.

The overall structure of Os4BGlu12 portrayed in Figure 3.12 shows 478 residues for each monomer, starting at residue 9 after the predicted mature protein N-terminus and ending with residue 486 as the last well defined residue. The typical GH1 (β/α)₈-barrel structure, shown in Figure 3.12, placed the putative catalytic acid/base, Glu179, and nucleophile, Glu393, at the C-terminal ends of β -strands 4 and 7, respectively, on opposite sides of the glucose ring at the bottom of the active site,

as expected for a GH A enzyme (Jenkins *et al.*, 1995 and Henrissat *et al.*, 1995). The substrate binding pocket of Os4BGlu12 is shaped primarily by the four extended variable loops connecting strands and helices at the C-terminal side of the barrel (Sanz-Apparasio *et al.*, 1998) in Figure 3.12 and 3.13: loop A (Ser25-Asp66, between $\beta 1$ and $\alpha 1$), loop B (Glu179-Arg213, between $\beta 4$ and $\alpha 4$), loop C (Trp321-Pro370, between $\beta 6$ and $\alpha 6$) and loop D (Asn394-Asp412, between $\beta 7$ and $\alpha 7$). The structures of monomers A and B in the asymmetric unit are nearly identical, with an RMSD of 0.06 Å over all C alpha atoms.

In the asymmetric unit of Os4BGlu12 crystals, His69 of both molecules and Asp66 of molecule A and Asp36 of molecule B chelate a Zn ion asymmetrically between the two protein molecules, as seen in Figure 3.12, and shown in detail in Figure 3.14. The presence of zinc bound to the protein in the crystal packing was identified on the basis of its coordination (tetrahedral coordination through four residues, His69 and Asp36 from molecule A and Glu36 and His69 from molecule B) and small ligand-ion distances (1.90-2.08 Å) and the B factor of Zn ion (24.23 Å²), which is similar to that of the liganding amino acid residue atoms (20.67 Å² for O δ 2 of Asp66, 21.48 Å² for N δ 1 of His69 and 23.56 Å² for O ϵ 2 of Glu36 and 24.84 Å² for N ϵ 2 of His69). The Zn²⁺ stabilizing interaction is not expected to occur in solution. Os4BGlu12 is a monomeric protein in solution, as judged by gel filtration and dynamic light scattering (DLS) (Figure 3.11), and the maximum interface of 488 Å² between molecules A and B, 2.7% of the average monomer surface area of 18,307 Å², is predicted to not result in a dimer in solution by the PISA program (Krissinel and Henrick, 2007).

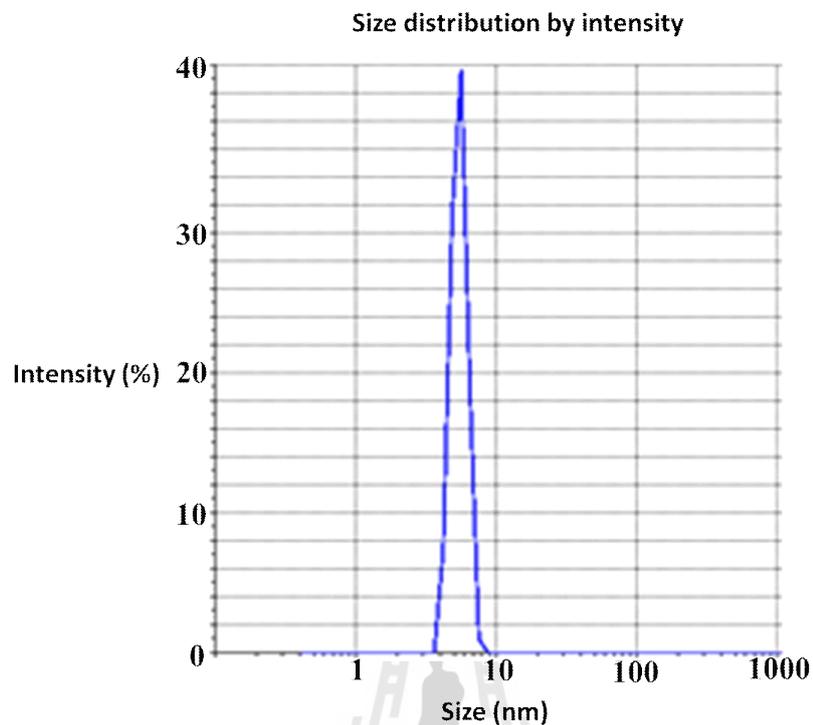


Figure 3.11 Dynamic light scattering (DLS) spectrum of tag-free Os4BGlu12 protein. DLS was used to determine the mean particle size and size distribution in a solution of 1 mg/mL protein in 20 mM Tris-HCl, pH 7.2, and was performed at 293K. The size distribution was plotted as intensity (%) versus size as the diameter of particle (d.nm). The protein gave a single peak (100% of scattering) at 5.454 nm diameter, which corresponds to the estimated molecular weight of tag-free Os4BGlu12 monomer (55 kDa). This result confirmed that Os4BGlu12 forms a functional monomeric enzyme in solution.

Table 3.5 Refinement statistics for the structures of Os4BGlu12 apo enzyme and its complexes with G2F and DNP2FG and Os4BGlu12 E179Q with TAG.

Dataset	Os4BGlu12	Os4BGlu12	Os4BGlu12	Os4BGlu12
		_DNP2FG	_G2F	E179Q_TAG
PDB code	3PTK	3PTQ	3PTM	Unpublished
R _{factor} (%)	20.3	20.9	19.6	23.4
R _{free} (%)	24.8	25.1	23.4	27.9
No. of residues in protein	486	486	486	486
	(2 molecules = 972)	(2 molecules = 972)	(2 molecules = 972)	(2 molecules = 972)
No. protein atoms	7720	7720	7720	7720
No. Ligand atoms	16 (Tris)	48 (DNP2FG)	22 (G2F)	34 (TAG)
No. Other hetero atoms	1 (Zn)	31 (GOL, Zn)	37 (GOL, Zn)	1 (Zn)
No. waters	228	280	511	64
<i>Mean B-factor</i>				
Protein	32.47	33.18	28.31	53.8
Ligand	52.70 (Tris)	62.69 (DNP2FG)	21.47 (G2F)	71.2 (TAG)
Other hetero atoms	24.23(Zn)	45.04 (GOL), 22.75 (Zn)	51.20 (GOL), 22.40 (Zn)	43.6 (Zn)
Waters	32.18	31.60	30.40	25.1
r.m.s. bond deviations (length)	0.009	0.010	0.006	0.006
r.m.s angle deviations (degrees)	1.19	1.26	1.02	0.892
Ramachandran plot	88.4	89.4	89.5	87.7
Residues in most favorable regions (%)				
Residues in additional allowed regions (%)	10.5	9.4	9.6	11.7
Residues in generously allowed regions (%)	1.1	1.2	0.9	0.9
Residues in disallowed regions (%)	0	0	0	0

Numbers in parentheses are statistics for the outer resolution shell. Ramachandran values were determined from PROCHECK (Emsley *et al.*, 2004). For the ligands, DNP2FG stands for dinitrophenyl 2-deoxy-2-fluoro- β -D-glucopyranoside, G2F for 2-fluoro- α -D-glucopyranoside and GOL for glycerol.

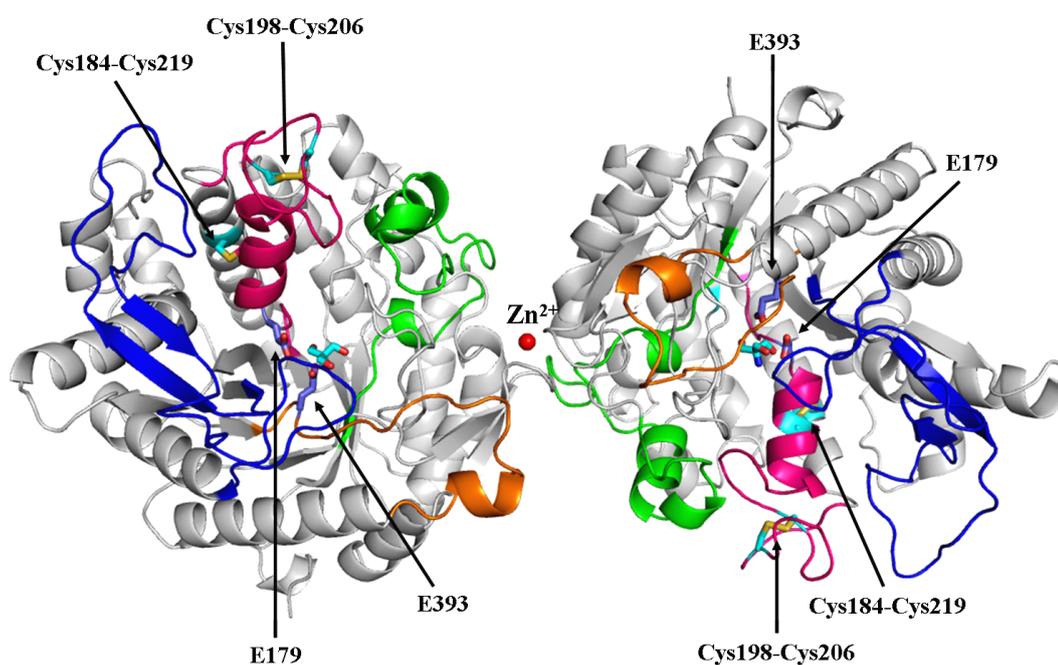


Figure 3.12 Ribbon diagram representation of the Os4BGlu12 apo protein structure in the asymmetric unit.

The catalytic acid/base, Glu179, and nucleophile, Glu393, are labeled and shown as sticks, colored in red for oxygen and purple for carbon. The four variable loops between the β -strands and α -helices of the $(\beta/\alpha)_8$ barrel are labeled and colored green for loop A, magenta for loop B, blue for loop C, and orange for loop D (including variable helices and strands that fall in these loop regions). The Zn^{2+} ion between the A and B protein molecules in the asymmetric unit is shown as a red sphere. Two molecules of Tris associated with the A and B molecules in the asymmetric unit of the Os4BGlu12 structure are represented by sticks with carbon in cyan, oxygen in red and nitrogen in blue. Two disulfide bridges located in the loop B region are represented by a thick bond (in gold with carbons in cyan). The Cys198-Cys206 disulfide bond is present in other structures of GH1 enzymes, while the Cys184-Cys219 bond is not observed in other GH1 enzymes but is located in the same loop region.

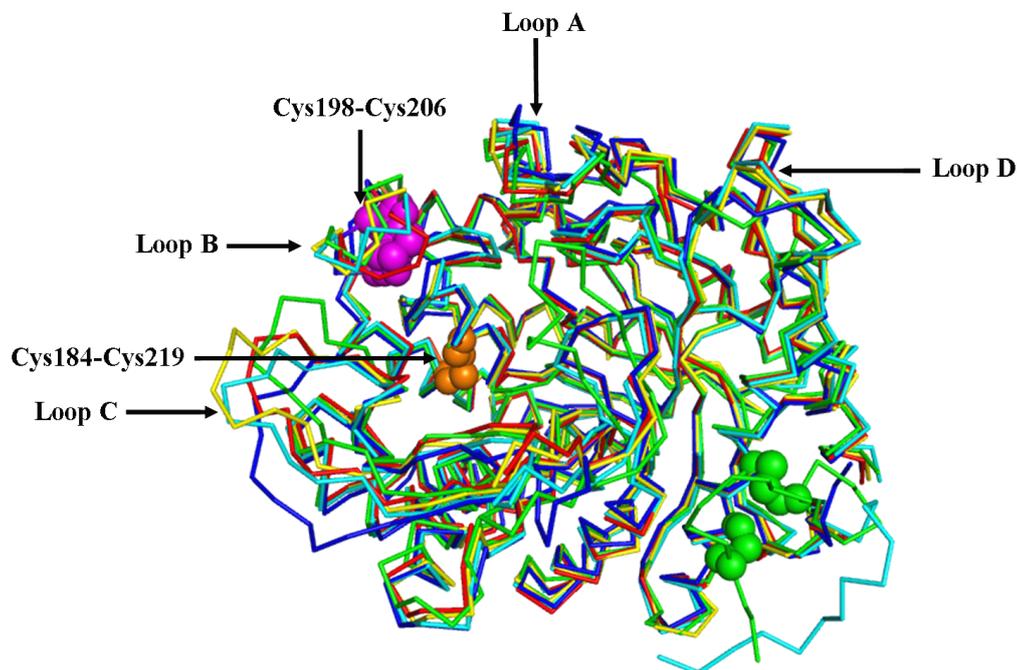


Figure 3.13 Superimposition of the 3D structures of different plant GH1 β -glucosidases.

The structures shown as alpha-carbon traces are: white clover (1CBG) in cyan, rice Os3BGlu7 (2RGM) in blue, rice Os3BGlu6 (3GNO) in yellow, *Sinapis alba* myrosinase (1MYR) in green, and Os4BGlu12 in red. Two disulfide bridges located at the loop B region are in pink space filling representation for the Cys198-Cys206 disulfide, which is present in other structures of plant GH1 enzymes, and in orange for the Cys184-Cys219 bond, which is not observed in other GH1 enzymes but is located in the same loop region. The green spheres represent another two disulfide bridges found in *S. alba* myrosinase between its N-terminal region and helix 7 of the $(\beta/\alpha)_8$ barrel.

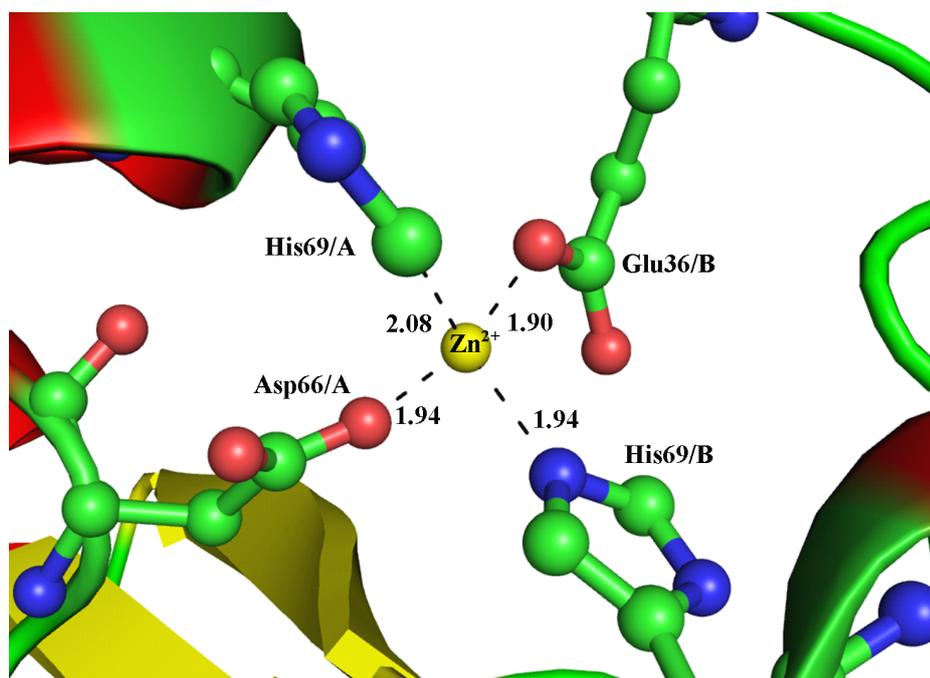


Figure 3.14 Binding of a Zn^{2+} ion between the two molecules in the asymmetric unit. Key residues interacting with the zinc ion are labeled and depicted as balls and sticks with carbons colored green, nitrogen blue, and oxygen red. The zinc ion is shown as yellow sphere, and the distances between interacting atoms and the Zn^{2+} ion are given in Å nströms.

3.9 Os4BGlu12 interactions with ligands

3.9.1 Apo-Os4BGlu12 binding of Tris

The active site of Os4BGlu12 is located at the bottom of an approximately 20 Å deep, slot-like pocket, with two walls approximately 8 Å apart at the narrowest point. As seen in Figure 3.12, a Tris molecule was observed at the bottom of the active site of the apo Os4BGlu12 structure in two different conformations in molecules A and B of the asymmetric unit (Figure 3.15). To support this assignment, Tris was found to be a competitive inhibitor (Figure 3.16), although with a high K_i of 144 ± 0.1 mM. Because Tris was present at 0.1 M in the crystallization condition, it could be found in the active site of the apo enzyme of Os4BGlu12, despite its low affinity, but was not likely to compete effectively with specific ligands. Tris is commonly a weak inhibitor of β -glucosidases, and it was previously reported that the active site of the apo enzyme of Os3BGlu6 (PDB accession 3GNO) contained Tris (Seshadri *et al.*, 2009), which was a component of the storage buffer of the protein and a competitive inhibitor with a K_i of 5.1 mM.

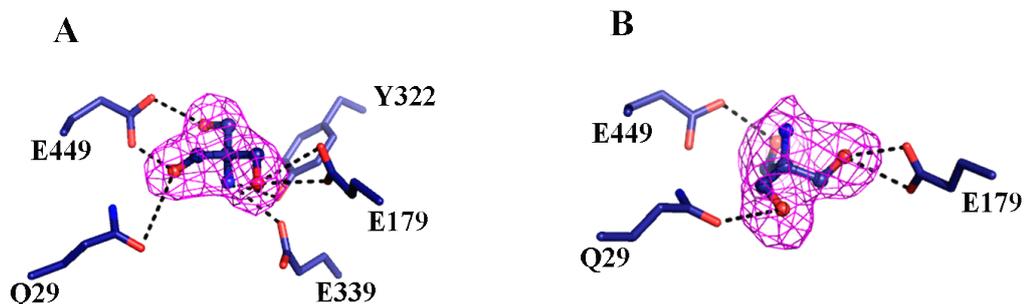


Figure 3.15 The active site region of native Os4BGlu12 with Tris bound.

Panels A and B show the $2mF_{\text{obs}} - dF_{\text{calc}}$ electron density maps of the Tris molecules bound in the active sites of monomers A and B are show. The Tris molecules are shown in balls and sticks with carbon in purple, oxygen in red and nitrogen in dark blue for both monomers. The maps are contoured at the 1σ level. The amino acids surrounding Tris are represented by sticks with carbon in purple, oxygen in red and nitrogen in dark purple. The dashed lines indicate interactions from amino acids within hydrogen-bonding distance of the ligands.

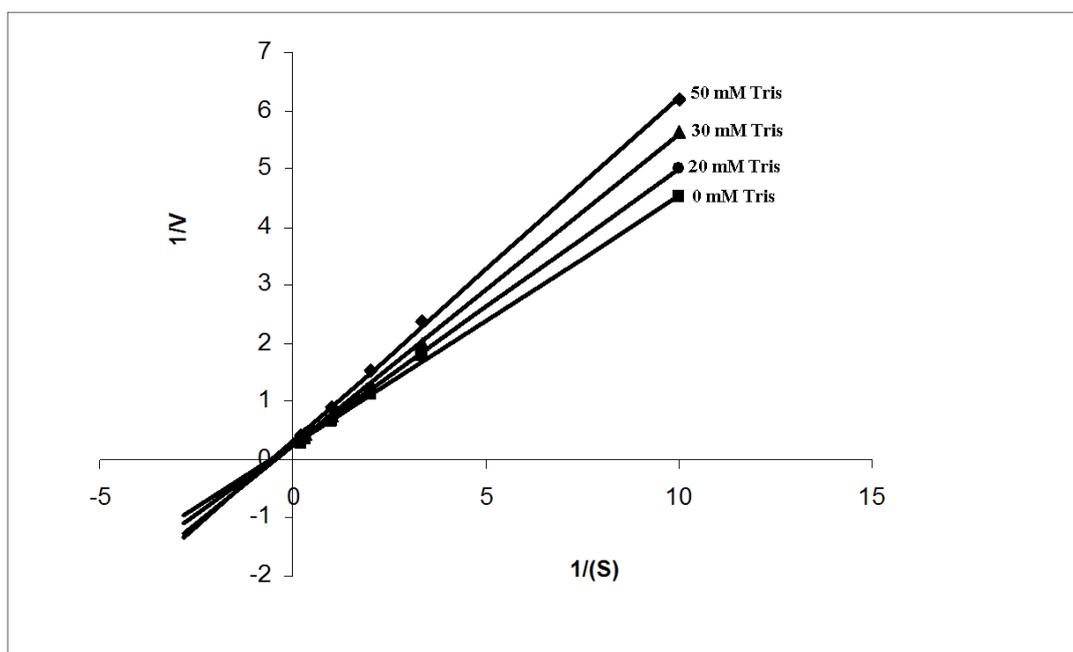


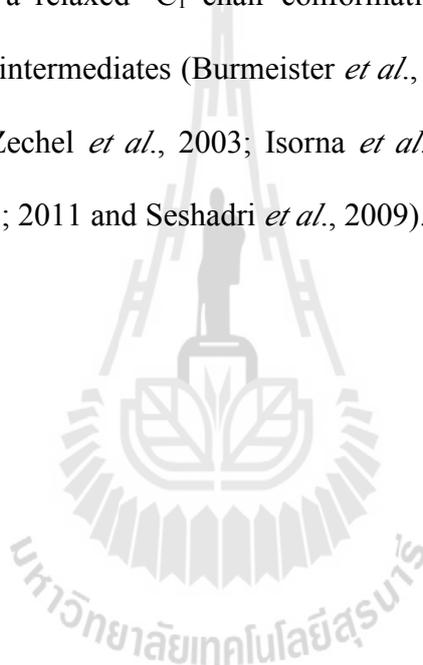
Figure 3.16 The Lineweaver Burk plots for hydrolysis of *p*NPGlc in different concentrations of Tris.

The intersection on the Y-axis demonstrates Tris acts as a competitive inhibitor of Os4BGlu12.

3.9.2 Os4BGlu12 complex with DNP2FG

When Os4BGlu12 was crystallized with precipitant solution containing 2 mM DNP2FG then soaked in cryoprotectant containing 10 mM DNP2FG before cooling and X-ray diffraction, the anticipated covalent intermediate of α -2-deoxy-2-fluoroglucoside was not observed in the electron density. Instead, the electron density map of Molecule A clearly revealed the presence of the 2-deoxy-2-fluoroglucose ring in a β -linkage to a weak density for the 2,4-dinitrophenol group, as shown in Figure 3.17 A and B. The whole ligand had a high average B-factor of 72 \AA^2 at full occupancy and refined to a partial occupancy of 0.8 with the B-factor of 63 \AA^2 , in line

with the transient nature of the Michaelis complex. The sugar ring was better defined than the nitrophenyl ring, and molecule B had poor density for the aglycone and higher B-factors, suggesting the active site was partially occupied with covalently linked or free 2-F-glucose, with different occupancies for the two molecules. However, the glucosyl moiety fit the electron density better in the 1S_3 skew-boat conformation previously observed for noncovalent β -glucosidase/substrate intermediates than in a relaxed 4C_1 chair conformation, which has generally been observed for covalent intermediates (Burmeister *et al.*, 1997; Malet and Planas, 1998; Czjzek *et al.*, 2001; Zechel *et al.*, 2003; Isorna *et al.*, 2007; Noguchi *et al.*, 2008; Chuenchor *et al.*, 2008; 2011 and Seshadri *et al.*, 2009).



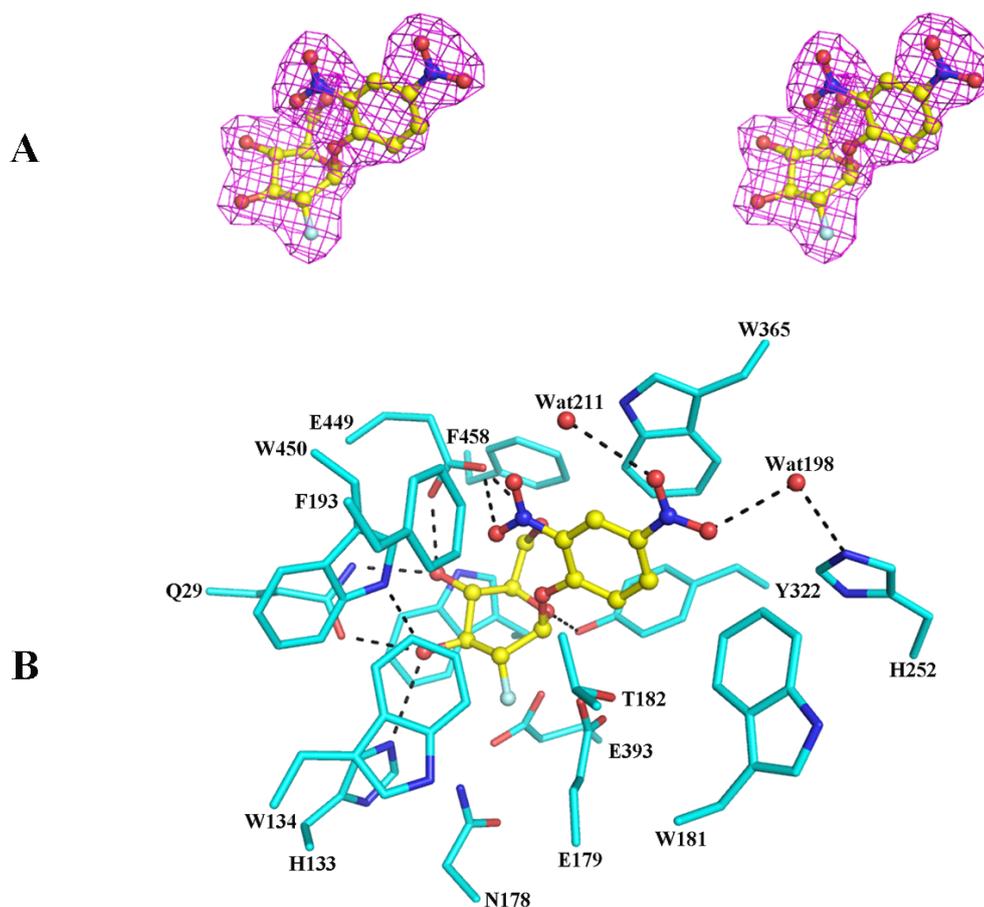


Figure 3.17 Protein-ligand interactions in the active site of the Os4BGlu12 DNP2FG complex.

(A) Stereo view of the electron density ($2mF_{\text{obs}} - dF_{\text{calc}}$ map contoured at the 1σ level) of DNP2FG, calculated with a model from which DNP2FG was omitted. (B) The amino acid residues surrounding DNP2FG are displayed in stick representation, with carbon, nitrogen and oxygen atoms colored in cyan, blue and red, respectively. DNP2FG is drawn in ball and stick representation in the same colors, except carbon atoms are in yellow and fluorine in green.

3.9.3 Os4BGlu12 complex with G2F

Because the crystal soaked in a high concentration of DNP2FG before freezing gave a noncovalent complex instead of the expected covalent intermediate, another crystal from cocrystallization with DNP2FG was soaked in the cryo-solution containing precipitant but without DNP2FG, and this crystal diffracted X-ray to yield electron density for the covalent intermediate. In the structure of Os4BGlu12 with a covalent bond between the anomeric carbon of G2F and the Glu393 nucleophile residue O ϵ 1, the 4C_1 chair conformation of the G2F ring was observed in subsite -1 (Figure 3.18 A and B). Aside from the covalent bonding of the nucleophile, the glucose was bound by hydrogen bonds similar to the Michaelis complex. Figure 3.18 B shows that the density was clear and Tyr322 was seen to have an alternative conformation, in which it can hydrogen bond to the catalytic nucleophile in this complex, in addition to a conformation hydrogen bonding to O5, similar to the noncovalent complex. The covalent complex of Os4BGlu12 is more similar to that of the covalent complex of *S. alba* myrosinase (1E73), the enzyme that hydrolyzed S-glucoside substrate, than to those of O-glucosidase enzymes, such as rice Os3BGlu7 (Chuenchor *et al.*, 2008). The similarity of the structures of the covalent complex of Os4BGlu12/G2F and the covalent complex of *S. alba* myrosinase is consistent with the fact that Os4BGlu12 had hydrolysis activity with S-glucoside substrates (Table 3.6).

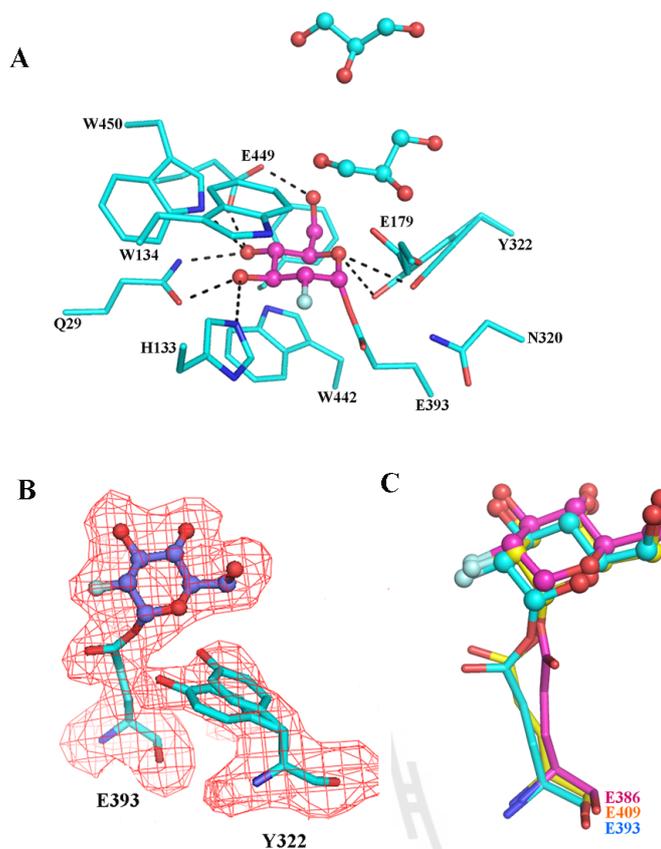


Figure 3.18 The active site region of the covalent complex of Os4BGlu12 with G2F. (A) Ligand binding of the monomer A. Amino acids surrounding the ligand are represented as sticks with carbons colored cyan. The ligands are represented in ball and stick and include G2F and glycerol with carbons in pink and yellow, respectively. (B) Electron density map (Omit $F_{obs} - F_{calc}$ contoured at 3σ) of the glycosyl-enzyme intermediate showing the two conformations of Tyr322 in the structure of Os4BGlu12 with G2F. (C) Superimposition of nucleophile and sugar residues of the glycosyl-enzyme intermediate of the Os4BGlu12/G2F complex with those of *S. alba* myrosinase and rice Os3BGlu7. Carbons are shown in cyan for Os4BGlu12/G2F (3PTM), in pink for rice Os3BGlu7 (2RGM) and in yellow for *S. alba* myrosinase (1E73).

Table 3.6 Hydrolysis of *O*-glucosides and *S*-glucosides by Os4BGlu12.

Substrate	V_{\max} ($\mu\text{mol min}^{-1}$ mg^{-1} protein)	K_m (mM)	k_{cat}^1 (s^{-1})	k_{cat}/K_m ($\text{mM}^{-1} \text{s}^{-1}$)
<i>p</i> NPG	20.1 \pm 0.6	0.56 \pm 0.03	18.5 \pm 0.15	33
<i>p</i> NPSG	0.29 \pm 0.007	1.64 \pm 0.10	0.27 \pm 0.015	0.17
<i>n</i> -heptyl- β -D-glucoside	13.2 \pm 0.13	0.91 \pm 0.07	12.1 \pm 0.05	13
<i>n</i> -octyl- β -D-glucoside	10.7 \pm 0.3	0.44 \pm 0.015	9.8 \pm 0.22	22
<i>n</i> -octyl- β -D-thioglucopyranoside	4.00 $\times 10^{-2}$ \pm 0.003	1.61 \pm 0.10	3.76 $\times 10^{-2}$ \pm 0.004	2.34 $\times 10^{-2}$

¹The k_{cat} values given are apparent values calculated by division of the V_{\max} by the molar amount of Os4BGlu12 protein (Sansenya *et al.*, 2011).

3.9.4 Os4BGlu12 E179Q TAG complex

The rice Os4BGlu12 β -glucosidase has recently been shown to hydrolyze the natural substrate tuberonic acid β -D-glucoside (TAG) and given the additional name TAG β -glucosidase 2 (TAGG2) (Wakuta *et al.*, 2011). The crystal structure of the inactive mutant Os4BGlu12 E179Q with TAG was determined to elucidate the aglycone binding determinants. The Os4BGlu12 E179Q was co-crystallized with TAG by hanging drop vapor diffusion, and its structure solved at 3.2 Å resolution. The TAG molecule was observed in the active site of Os4BGlu12 E179Q in this structure (Figure 3.19 A). The density of the glucose ring was best fit to a 1S_3 skew boat conformation with the average B-factor was 67.5 Å², while the density was fit to the TA aglycon with an average B-factor was 74.6 Å² (Figure 3.19 B), compared to the average B-factor of 53.8 Å² for the protein. The glycone residue was in a similar position to that of G2F in the G2F and DNP2FG complexes (Figure 3.19 C). At the -1 subsite, a group of conserved residues (Gln29, His133, Asn178,

Tyr322, Glu393, Trp450 and Glu449) form hydrogen bonds with O2, O3 and O4 of the TAG glucosyl moiety. The hydrogen bonding with glucosyl unit in the structures of Os4BGlu12 with DNP2FG and Os4BGlu12 E179Q with TAG was similar, except that Glu449 was displaced away from the glucosyl O6 of TAG in the structure of Os4BGlu12 E179Q with TAG (Figure 13.9 C). The aglycone binding subsite is defined by both hydrophobic and polar residues. The TA group was bound by hydrophobic and polar residues, including Trp181, Thr182, Asn186, His252, and Trp363. Thus, despite the relatively low resolution of the structure, residues active in natural substrate binding could be identified.



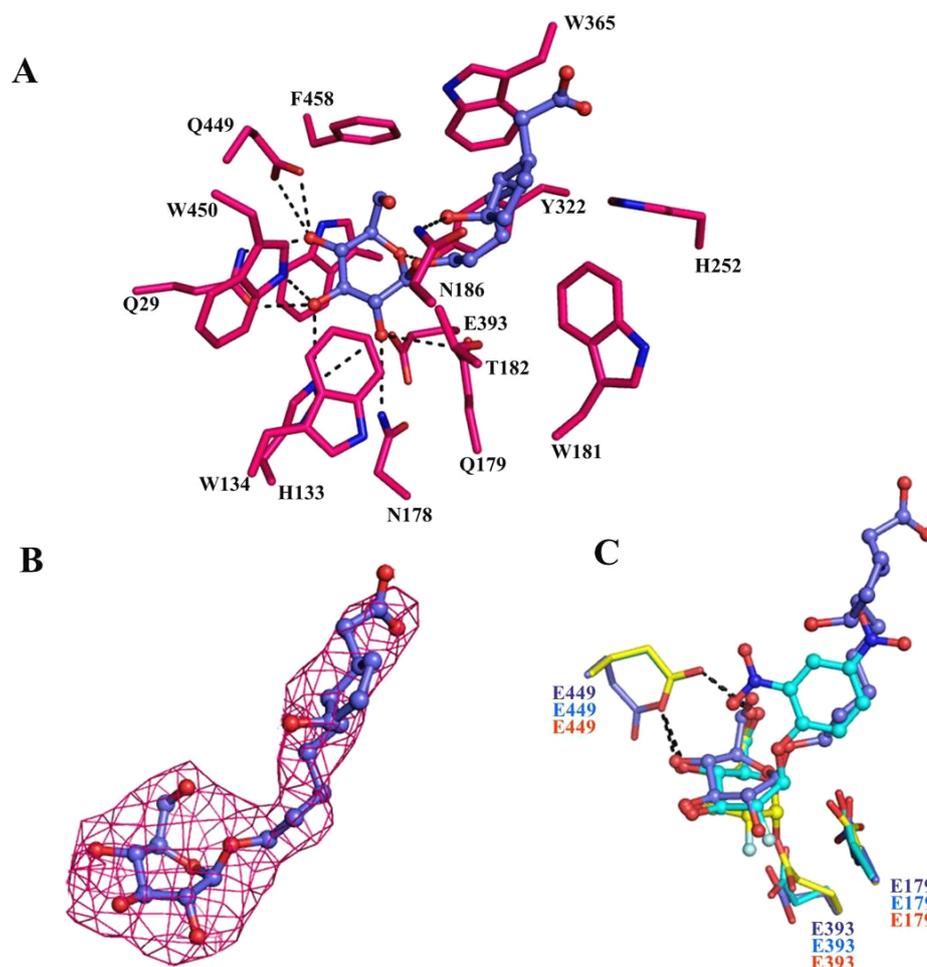


Figure 3.19 Active site of the crystal structure of Os4BGlu12 E179Q with TAG.

(A) The amino acid residues surrounding the TAG ligand in the active site. The amino acids are represented as sticks with carbon in hot pink, while TAG shown in ball and stick representation with carbon in purple,. (B) The $|F_{\text{obs}}| - |F_{\text{calc}}|$ electron density map around the TAG ligand is shown contoured at 3σ , as calculated before insertion of the ligand. (C) Superimposition of the complexes with the three ligands, DNP2FG (carbon in cyan), G2F (carbon in yellow) and TAG (carbon in purple). The catalytic acid/base (E179), nucleophile (E393) and E449 positions in each of the three structures are shown with the ligands. Oxygen is in red and nitrogen in dark blue in all pictures.

3.10 Determination of Os3BGlu6 E178Q complex structures

3.10.1 X-ray analysis and structure refinement

From previous study (Hua *et al.*, 2013), Os3BGlu6 had highest hydrolysis activity to GA₄GE among the enzymes expressed in recombinant systems in our lab. The purified Os3BGlu6 E178Q protein was crystallized with and soaked with GA₄GE and *p*NPGlc in an effort to obtain complexes with these substrates. The crystals of Os3BGlu6 E178Q with GA₄GE and *p*NPGlc diffracted X-rays to 1.97 Å and 1.90 Å resolutions, respectively, and the processing statistics are shown in Table 3.7. The crystals of Os3BGlu6 E178Q with GA₄GE and *p*NPGlc had *P*2₁2₁2₁ space group symmetry, as did that of the native structure (Seshadri *et al.*, 2009). The unit cell parameters were $a = 57.05$, $b = 91.22$, $c = 111.37$ Å ($\alpha = \beta = \gamma = 90^\circ$) for Os3BGlu6 E178Q with GA₄GE and $a = 56.99$, $b = 91.13$, $c = 111.22$ Å ($\alpha = \beta = \gamma = 90^\circ$) for Os3BGlu6 E178Q with *p*NPGlc, respectively. The crystals of Os3BGlu6 E178Q with GA₄GE and Os3BGlu6 E178Q with *p*NPGlc were isomorphous with that of the native Os3BGlu6 structure, allowing the structures to be solved by rigid body refinement with the native enzyme structure as a search model. Both crystals contained one molecule in the asymmetric unit, with a Matthews coefficient (V_m) of 2.59 Å³Da⁻¹ and solvent content of 52.49% for the crystal with GA₄GE and a V_m of 2.58 Å³Da⁻¹ and solvent content of 52.33% for Os3BGlu6 E178Q with *p*NPGlc. These crystal parameters are similar to those of the previous three Os3BGlu6 structures (Seshadri *et al.*, 2009).

The crystal structures of Os3BGlu6 E178Q with GA₄GE and *p*NPGlc were refined at 1.97 and 1.90 Å, respectively. The initial residual factors were: an R_{factor} of 25.90% and R_{free} of 27.11% for Os3BGlu6 E178Q with GA₄-GE and R_{factor} of

26.15% and R_{free} of 26.84% for Os3BGlu6 E178Q with *p*NPGlc. After the ligand, water and other hetero atoms were added, the refinement yielded a final R_{factor} of 14.4% and R_{free} of 18.1% for Os3BGlu6 E178Q with GA₄GE and R_{factor} of 14.7% and an R_{free} of 18.1% for Os3BGlu6 E178Q with *p*NPGlc. The protein structure was well defined from Gly11 to Thr488, except for the loop of residues 331 to 337, which was poorly defined by the electron density, as was the N-terminus (similar to the previously reported wild type structure). The final model gives the average B-factors for the protein atoms of 20.4 Å² for Os3BGlu6 E178Q with GA₄GE and 16.4 Å² for Os3BGlu6 E178Q with *p*NPGlc, as shown in the refinement statistics in Table 3.7.

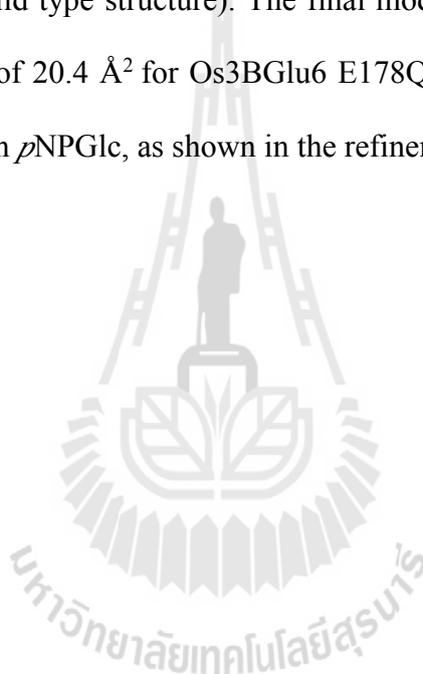


Table 3.7 Data-collection and structure-refinement statistics.

Dataset	Os3BGlu6E178Q_	Os3BGlu6E178Q_
	<i>p</i>NPG	GA₄-GE
PDB code	3WBA	3WBE
Beamline	BL13B1	BL13B1
Wavelength (Å)	1.00	1.00
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
	<i>a</i> = 56.99	<i>a</i> = 57.05
Unit-cell parameters (Å)	<i>b</i> = 91.13	<i>b</i> = 91.22
	<i>c</i> = 111.22	<i>c</i> = 111.37
Resolution range (Å)	30-1.90	30-1.97
Resolution outer shell(Å)	2.33-1.90	2.04-1.97
No. Unique reflections	46370	41777
No. Observed reflections	297256	297549
Completeness (%)	97 (99.9)	99.9 (100)
Average redundancy per shell	6.5 (6.3)	7.1 (7.2)
<i>I</i> / σ (<i>I</i>)	28.9 (4.5)	23.3 (6.6)
R _(merge) (%)	6.0 (41.0)	8.4 (27.6)
R _{factor} (%)	14.8	14.7
R _{free} (%)	17.9	18.0
No. of residues in protein	478	478
No. protein atoms	3917	3952
No. Ligand atoms	11 (Glc)	11 (Glc)
No. Other hetero atoms	36 (GOL, Tris)	36 (GOL)
No. waters	534	491
<i>Mean B-factor</i>		
Protein	16.6	20.6
Ligand	12.1	10.4
Other hetero atoms	35.3 (GOL)	44.2 (GOL)
Waters	34.3	36.8
r.m.s. bond deviations (length)	0.013	0.014
r.m.s angle deviations (degrees)	1.290	1.264
Ramachandran plot	88.9	88.6
Residues in most favorable regions (%)		
Residues in additional allowed regions (%)	10.7	10.9
Residues in generously allowed regions (%)	0.5	0.5
Residues in disallowed regions (%)	0	0

3.10.2 Structure of Os3BGlu6 E178Q soaked with *p*NPGlc

A glucosyl moiety was observed at the bottom of the active site of Os3BGlu6 E178Q in the structure derived from a crystal soaked with *p*NPGlc. The anomeric carbon of glucose is covalently bonded with O ϵ 1 of Glu394 at a distance of 1.50 Å (Figure 3.20 A). The hydrogen bonding network between the glucose moiety and the surrounding amino acid residues includes Gln31, His132, Asn177, Asn178, Tyr321, Glu451 and Trp452. The density clearly shows a 4C_1 chair conformation of the glucosyl ring in the -1 subsite. Its binding is mediated by nine hydrogen bonds (Figure 3.21 A and Table 3.8). O2 of glucose formed hydrogen bonds with His132 N ϵ 2 (3.30 Å), Asn177 OD1 (3.01 Å) and Gln178 N ϵ 2 (2.99 Å), while O3 of glucose is within hydrogen bonding distance of Gln31 O ϵ 1 (2.66 Å) and Trp452 N ϵ 1 (2.85 Å). The O5 atom of glucose formed a hydrogen bond with Tyr321 OH (2.70 Å) and O4 is hydrogen bonded to Gln31 N ϵ 2 (2.91 Å) and Glu451 O ϵ 1 (2.61 Å), while O6 forms a hydrogen bond with Glu451 O ϵ 2 (2.60 Å). Another hydrogen bonding interaction is mediated by a water molecule (water 52 in the structure file) that interacts with O2 of glucose (2.80 Å) and Glu394 O ϵ 1 (2.98 Å). A glycerol molecule was observed with two conformations in the +1 subsite. Following refinement with the two conformations of glycerol, the $|F_{\text{obs}}| - |F_{\text{calc}}|$ electron density difference map was relatively clean.

Table 3.8 Hydrogen bonds occurring between active site amino acid residues and glucose in the structure of Os3BGlu6 E178Q soaked with *p*NPGlc.

GLUCOSE		
Residues	Glucose	Distance (Å)
Gln31Oε1	O3	2.66
Nε2	O4	2.91
His132 Nε2	O2	3.30
Asn177OD1	O2	3.01
Gln178Nε2	O2	2.99
Tyr321OH	O5	2.70
Glu451Oε1	O4	2.61
Oε2	O6	2.60
Trp452Nε2	O3	2.85
Residues/Glucose	Wat52	Distance (Å)
Glu451Oε1	Wat52	2.98
Glucose	Wat52	2.80

3.10.3 The structure of Os3BGlu6 E178Q with GA₄GE

A glucose molecule was also observed in the active site of the Os3BGlu6 E178Q structure from the crystal soaked with GA₄GE. The glucose in a ⁴C₁ conformation was well refined to the density, with the anomeric carbon covalently bound with Oε1 (1.45 Å) of Glu394, the catalytic nucleophile residue (Figure 3.20 B and 3.21 B). The glucose was set to 0.5 occupancy, since glycerol appeared to occupy the active site in half of the molecules, and the average B-factor was 10.4 Å² for glucose and 22 Å² for glycerol (Figure 3.21 B and C). As seen in the Figure 3.20 B and 3.21 B, the catalytic nucleophile residue Glu394 refined to two different

conformations, one covalently bound with the anomeric carbon of the glucosyl moiety and another one hydrogen bonded with O2 of glycerol (2.48 Å). The distance between Oε1 of Glu394 in the noncovalently bound conformation and glucose C1 was 2.11 Å. The glucose residue formed hydrogen bonds with 7 amino acid residues around the -1 subsite (Gln31, His132, Asn177, Gln178, Trp321, Glu394, Glu451 and Trp452) (Table 3.8). The Oε1 and Nε2 atoms of Gln31 hydrogen bonded with O3 (2.73 Å) and O4 (2.90 Å) of the glucose ring, respectively. The O2 atom of glucose formed hydrogen bonds with Nε2 of His132 (3.19 Å), Oδ1 of Asn177 (3.04 Å), and Nε2 of Gln178 (3.22 Å), while O3 of glucose hydrogen bonded with Nε1 of Trp452 (2.76 Å). The O4, O5 and O6 atoms of glucose formed hydrogen bonds with Oε1 of Glu451 (2.50 Å), OH of Tyr321 (2.84 Å), and Oε2 of Glu451 (2.61 Å), respectively. These interactions are summarized in Table 3.9.

The glycerol molecule was present in the same positions as C2, C3 and C4 of the glucose ring, so that O1, O2 and O3 of glycerol can be compared with O2, O3 and O4 of glucose (Figure 3.21 A, B, C and D). Oxygen 1 and oxygen 2 of glycerol project at angles similar to those of O2 and O3 of glucose, but O3 of glycerol has a different position when compared to O4 of glucose. O3 of glycerol is in a position equivalent to the axial O4 of D-galactose, rather than the equatorial O4 of D-glucose. Hydrogen bonds to the glycerol molecule are formed by 6 amino acid residues (Table 3.9). O1 of glycerol is in the position to form hydrogen bonds with Oδ1 of Asn177 (2.78 Å) and Oε2 of Glu394 (3.03 Å). O2 of glycerol forms hydrogen bonds with Oε1 of Gln31 (2.49 Å) and Nε1 of Trp452 (3.04 Å), while O3 of glycerol hydrogen bonds with Oε2 of Glu451 (3.11 Å).

Table 3.9 Hydrogen bonds occurring between glucose or glycerol with amino acid binding residues in the structure of Os3BGlu6 E178Q soaked with GA₄GE.

GLUCOSE			GLYCEROL		
Residues	Glucose	Distance (Å)	Residues	Glycerol	Distance (Å)
Gln31Oε1	O3	2.73	Gln31Oε1	O2	2.49
Nε2	O4	2.90	Asn177OD1	O1	2.78
His132 Nε2	O2	3.19	Glu394Oε2	O1	3.03
Asn177OD1	O2	3.04	Glu451Oε2	O3	3.11
Gln178Nε2	O2	3.22	Trp452Nε1	O2	3.04
Tyr321OH	O5	2.84			
Glu451Oε1	O4	2.50			
Oε2	O6	2.61			
Trp452Nε2	O3	2.76			

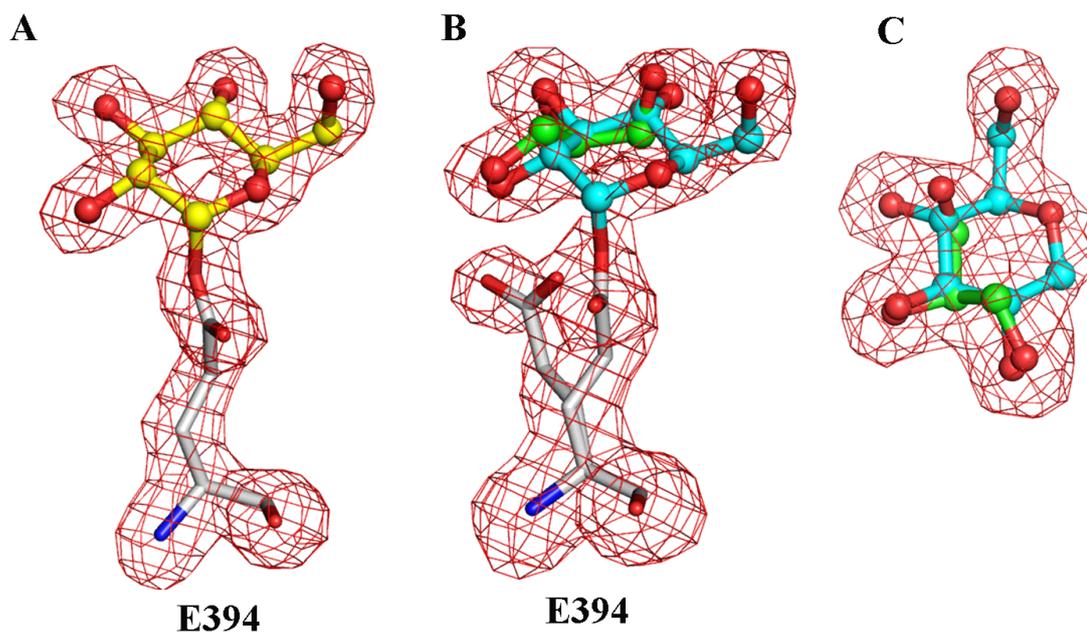


Figure 3.20 The electron density maps of the glycosyl-enzyme intermediate of Os3BGlu6 E178Q with *p*NPGlc (A) and GA₄GE (B, C).

The positive $F_{\text{obs}} - F_{\text{calc}}$ Omit maps are shown in red mesh contoured at 3σ . A. The glucose and two conformations of glycerol in the active site of Os3BGlu6 E178Q soaked with *p*NPGlc are represented by balls and sticks with carbon in yellow for glucose and in green for glycerol. The nucleophile residue (E394) is represented by sticks with carbon in gray, oxygen in red and nitrogen in dark blue. B. The mixed structures of apo Os3BGlu6 E178Q and Os3BGlu6 E178Q covalently bound to Glc from GA₄GE was modeled with two conformations of E394 and glucose and glycerol occupying the same position, by setting the occupancy at 0.5 for glucose and glycerol and each of the E394 conformations. The glucose and glycerol are shown in ball and stick representation with carbon in cyan for glucose and in green for glycerol. The nucleophile residue (E394) is represented as sticks with carbon in gray, oxygen in red and nitrogen in dark blue. C. Different view of the electron density of figure B.

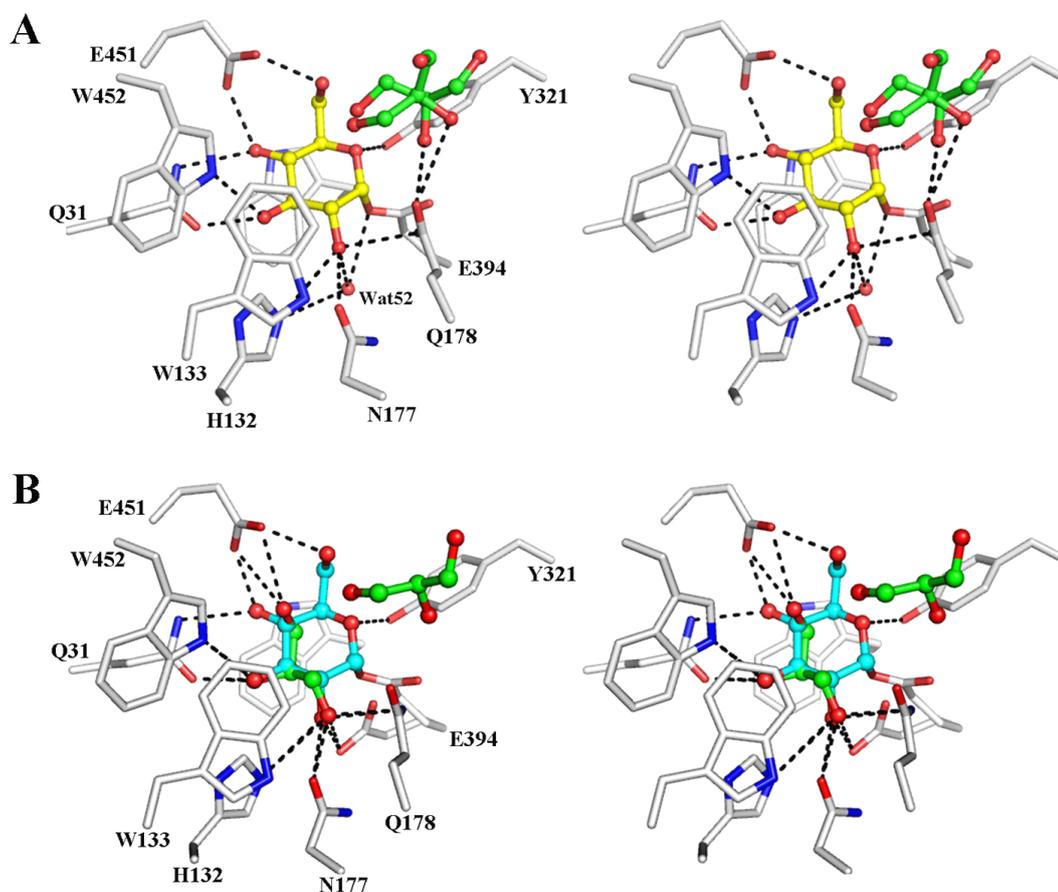


Figure 3.21 Binding of glucose in the active site of Os3BGlu6 E178.

A and B, Stereo view of protein-ligand interaction in the active site of the Os3BGlu6 E178Q soaked with *p*NPGlc and with GA₄GE, respectively. The amino acids surrounding the -1 subsite are represented as sticks with carbon in gray, nitrogen in dark blue and oxygen in red. The covalent glucosyl moiety is represented as balls and sticks with carbon in yellow for Os3BGlu6 E178Q soaked with *p*NPGlc and in cyan for Os3BGlu6 E178Q soaked with GA₄GE, and oxygen in red. The two glycerols in both complex structures are represented as balls and sticks with carbon in green and oxygen in red. Hydrogen bonding interactions of glucose and glycerol with amino acid residues are shown as black dashed lines

3.11 Kinetic comparisons of rice wild type β -glucosidases and their cello-oligosaccharide-binding subsite +2 mutants

Native and complex structures of rice Os3BGlu6, Os3BGlu7 and Os4BGlu12 have been solved (Chuenchor *et al.*, 2008; 2011; Seshadri *et al.*, 2009; this work). These include covalent intermediates of all three enzymes with 2-fluoroglucoside, oligosaccharide-bound complexes of an inactive Os3BGlu7 mutant, and substrate-like inhibitor complexes with Os3BGlu6 and Os4BGlu12. The amino acid residues surrounding the nonreducing end glucosyl residue in the -1 subsite of the three enzymes are all conserved, while their aglycone binding subsites are not. One important position for aglycone binding is that of Os3BGlu7 Asn245, which was seen to make hydrogen bonds to the glucosyl residue in subsite +2 in celooligosaccharides and to that in subsite +1 in laminaribiose (Chuenchor *et al.*, 2011). A comparison of the celooligosaccharide complex structures with Os3BGlu6 shows that the Met251 at the position corresponding to Os3BGlu7 Asn245 extends into the active site cleft in a position that would sterically block the binding of β -(1,4)-linked oligosaccharides in the same position in Os3BGlu6 (Figure 3.22) (Chuenchor *et al.*, 2011 and Seshadri *et al.*, 2009). In contrast, Os4BGlu12 has His252 in this position, which may be able to form hydrogen bonds to the celooligosaccharides.

To assess the effects of the residue corresponding to Os3BGlu6 Met251 and Os3BGlu7 Asn245, the kinetic constants for hydrolysis of the β -(1,3)-linked disaccharide laminaribiose and β -(1,4)-linked celooligosaccharides (DP 2-6) were determined for Os3BGlu6 M251N, Os3BGlu7 N245M and Os4BGlu12 H252M, and

compared to the wild type enzymes, as show in Table 3.10 Os3BGlu6 M251N showed a 10- to 24-fold increase in the k_{cat}/K_m of all the substrates compared with the wild type enzyme. The greatest change was seen for cellotriose, where the 24-fold increase in k_{cat}/K_m corresponds to an 8 kJ/mol improvement in the binding of the transition state ($\Delta\Delta G_{S^*mut} = -8$ kJ/mol). The k_{cat}/K_m of cellotetraose was within error of that of cellotriose, while that of cellopentaose was lower. In contrast, the Os3BGlu7 N245M mutant showed a 6.5-fold drop in k_{cat}/K_m for laminaribiose and 17- to 30-fold drop in k_{cat}/K_m for cellooligosaccharides with DP of 3-6 compared to wild type. Although Os3BGlu7 N245M did show an increase in k_{cat}/K_m with DP from 2 to 6, similar to the wild type enzyme, the largest decrease in k_{cat}/K_m was seen at cellohexaose. In Os4BGlu12, the corresponding H252M mutation decreased the k_{cat}/K_m 3 fold for laminaribiose and 2-6 times for cellooligosaccharide with DP of 2 to 5.

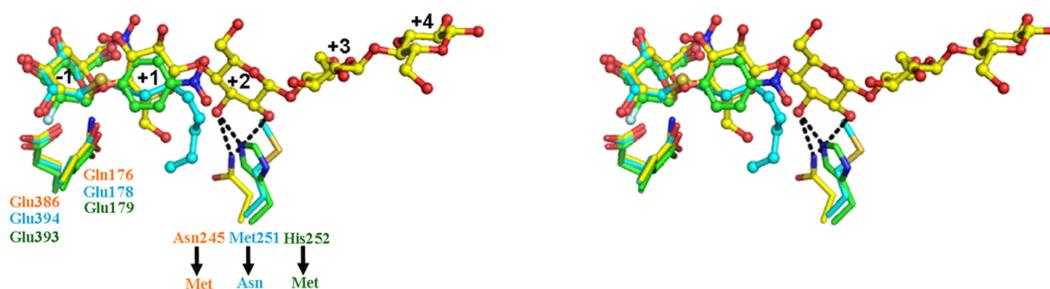


Figure 3.22 Comparison of the structures of three rice β -glucosidases in complexes with substrates or inhibitors.

The relaxed eye stereo view of the ligands in the active sites of the superimposed structures of the complexes of Os3BGlu7 E176Q with cellopentaose (PDB ID: 3F5K, yellow carbons; Chuenchor *et al.*, 2011); Os3BGlu6 with *n*-octyl- β -D-thioglucopyranoside (PDB ID:3GNP, cyan carbons; Seshadri *et al.*, 2009); and Os4BGlu12 with 2,4-dinitrophenyl 2-deoxy-2-fluoro- β -D-glucopyranoside. The acid/base and nucleophile residues are shown, as well as the residues at the positions mutated in this work, with the mutations designated by the arrows.

Table 3.10 Kinetic constants of Os3BGlu6, Os3BGlu6 M251N, Os3BGlu7, Os3BGlu7 N245M Os4BGlu12, and Os4BGlu12 H252M for hydrolysis of oligosaccharides.

Substate	Kinetic parameters	Os3BGlu6	Os3BGlu6 M251N	Os3BGlu7 ^b	Os3BGlu7N245M	Os4BGlu12 ^c	Os4BGlu12 H252M
Laminaribiose (β -1,3-linked Glc ₂)	K_m (mM)	8.7 \pm 0.07	3.5 \pm 0.2	2.05 \pm 0.10	4.5 \pm 0.3	5.1 \pm 0.2	7.2 \pm 0.4
	k_{cat} (s ⁻¹)	1.64 \pm 0.06	9.9 \pm 0.5	32 \pm 3	9.6 \pm 0.3	23 \pm 1	10.6 \pm 0.3
	k_{cat}/K_m (M ⁻¹ s ⁻¹)	180	2,830	15,700	2,410	4,500	1,460
Cellobiose (β -1,4-linked Glc ₂)	$\Delta\Delta G_S^{*mut}$ (kJ mol ⁻¹)		-6.9		+4.7		+1.3
	K_m (mM)	17.5 \pm 0.3	14.3 \pm 1.2	31.5 \pm 1.6	46 \pm 3	27 \pm 2	22.7 \pm 1.9
	k_{cat} (s ⁻¹)	0.27 \pm 0.004	1.97 \pm 0.09	1.52 \pm 0.13	1.18 \pm 0.09	5.9 \pm 0.4	2.42 \pm 0.11
Celotriose	$\Delta\Delta G_S^{*mut}$ (kJ mol ⁻¹)		-5.6		+1.3		+1.7
	K_m (mM)	25.1 \pm 0.8	3.2 \pm 0.2	0.72 \pm 0.002	1.51 \pm 0.10	5.0 \pm 0.5	11.2 \pm 0.4
	k_{cat} (s ⁻¹)	0.41 \pm 0.014	1.32 \pm 0.12	18.1 \pm 0.4	2.16 \pm 0.13	44.3 \pm 0.3	41.9 \pm 0.13
Cellobetraose	$\Delta\Delta G_S^{*mut}$ (kJ mol ⁻¹)		-8.0		+7.2		+2.3
	K_m (mM)	20.3 \pm 0.3	3.9 \pm 0.2	0.28 \pm 0.01	1.78 \pm 0.16	5.9 \pm 0.5	23.4 \pm 0.16
	k_{cat} (s ⁻¹)	0.54 \pm 0.004	1.62 \pm 0.10	17.3 \pm 0.6	5.5 \pm 0.2	89 \pm 3	87.8 \pm 0.2
Cellopentaose	$\Delta\Delta G_S^{*mut}$ (kJ mol ⁻¹)		-6.9		+7.5		+2.4
	K_m (mM)	27.0 \pm 0.5	4.6 \pm 0.4	0.24 \pm 0.01	1.32 \pm 0.11	5.5 \pm 0.4	20.25 \pm 0.11
	k_{cat} (s ⁻¹)	0.25 \pm 0.003	0.62 \pm 0.04	16.9 \pm 0.1	6.7 \pm 0.5	105 \pm 9	62.1 \pm 0.5
Cellohexaose	$\Delta\Delta G_S^{*mut}$ (kJ mol ⁻¹)		-6.9		+6.7		+4.6
	K_m (mM)	ND ^a	ND	0.11 \pm 0.01	1.52 \pm 0.08	ND	ND
	k_{cat} (s ⁻¹)	ND	ND	16.9 \pm 0.3	7.73 \pm 0.4	ND	ND
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	ND	ND	153,000	5,100	ND	ND
	$\Delta\Delta G_S^{*mut}$ (kJ mol ⁻¹)				+8.6		

^a ND means not determined. ^b Kinetics constants were taken from Opassiri *et al.*, 2004.. ^c Kinetic constants for wild type Os4BGlu12 were taken from Opassiri *et al.*, 2010. ^d The change in Gibbs free energy for transition state binding by the mutants compared to the wildtype enzyme was calculated as $\Delta\Delta G_S^{*mut} = -RT[\ln(k_{cat}/K_m)_{mutant} - \ln(k_{cat}/K_m)_{wildtype}]$ (Fersht *et al.*, 1987).

CHAPTER IV

DISCUSSION

4.1 The crystal structure of Os4BGlu12

4.1.1 Overall structure

To date, 39 structures of glycoside hydrolase family 1 enzymes are available in the structural database, including five structures from archaea, seventeen structures from bacteria and seventeen structures from eukaryotes (<http://www.cazy.org/GH1>). The overall folds of these known structures include a $(\beta/\alpha)_8$ TIM barrel fold for the catalytic domains, with the two catalytic residues located at the ends of β -strands 4 (acid/base) and 7 (nucleophile) (Henrissat *et al.*, 1995). Recently, three structures of rice enzymes were determined: rice Os3BGlu7, by Chuenchor *et al.*(2008); rice Os3BGlu6, by Seshadri *et al.*(2009); and Os4BGlu12, in the current work (Sansenya *et al.*, 2012). These enzymes also exhibited a $(\beta/\alpha)_8$ TIM barrel fold similar to known structures in GH1.

For structural comparison, cyanogenic β -glucosidase from *Trifolium repens* (1CBG) (Barrett *et al.*, 1995), rice Os3BGlu7 β -glucosidase (2RGM) (Chuenchor *et al.*, 2008) and Os3BGlu6 (3GNO) (Seshadri *et al.*, 2009) and myrosinase from *Sinapis alba* (1MYR) (Burmeister *et al.*, 1997) were superimposed on the Os4BGlu12 structure, as shown in Figure 3.13. The 1CBG, 2RGM and 3GNO structures represent plant *O*-glucosidases, and 1MYR is a plant *S*-glycosidase.

Os4BGlu12 showed the highest amino acid sequence identity with 1CBG (63%), followed by 3GNO (53%), 2RGM (51%) and 1MYR (43%). The C α atoms of molecule A from the 1CBG, 3GNO, 2RGL and 1MYR structures superimposed to rice Os4BGlu12 molecule A with RMSD of 0.61, 0.71, 0.77 and 0.78 Å for 444, 431, 428 and 413 C α atoms, respectively. The main-chain trace of loop C of Os4BGlu12 differs significantly from that of other GH1 structures. This agrees with previous reports that this loop is less conserved (Burmeister *et al.*, 1997). Loop C and its neighboring loops are mainly responsible for formation of the aglycone-binding site and thus contribute to the substrate specificity among GH1 enzymes (Burmeister *et al.*, 1997; Czjzek *et al.*, 2001; Husebye *et al.*, 2005; Sue *et al.*, 2006 and Chuenchor *et al.*, 2008).

As shown in Figure 3.12, the monomer structure of Os4BGlu12 contains two disulfide bridges between Cys198-Cys206, located in loop B, and Cys184-Cys219, connecting the α -helix at the beginning of loop B and α -helix 4 of the (β/α)₈ barrel, which follows loop B. The disulfide bond between Cys198 and Cys206 is conserved among the known structures of plant plastid and secretory pathway-targeted GH1 enzymes (Barrett *et al.*, 1995; Burmeister *et al.*, 1997; Cicek and Esen, 1999; Verdoucq *et al.*, 2004; Sue *et al.*, 2006; Chuenchor *et al.*, 2008 and Seshadri *et al.*, 2009). Most of these GH1 proteins have only this one disulfide bridge, but *S. alba* myrosinase (1MYR) has another two disulfide bridges between its N-terminal region and helix 7 of the (β/α)₈-barrel (shown in Figure 3.13), which probably confer additional stability to the enzyme (Burmeister *et al.*, 1997). The conserved disulfide bond was also reported to be necessary for proper folding of maize ZmGlu (Rotrekl *et al.*, 1999). It was hypothesized that the presence of two disulfide bridges in the loop B

region might help explain the fact that Os4BGlu12 has higher temperature stability than rice Os3BGlu7, which has only one disulfide bond (Opassiri *et al.*, 2003). However, the reducing agent dithiothreitol showed a larger effect on the stability of Os3BGlu7 than on that of Os4BGlu12. Thus, although the conserved and nonconserved disulfide bonds may contribute slightly to Os4BGlu12 stability, other factors appear to explain its higher stability than Os3BGlu7.

Aside from the disulfide bonds, the structure of Os4BGlu12 contains several salt bridges and hydrogen bonds that contribute to its stability. Of the 25 arginine residues, 18 residues form direct salt bridges with other residues within the same monomer, while 15 of 30 lysine residues and 3 of 10 histidine residues form such salt bridges. Fifteen salt bridges are obtained from 27 aspartate residues, while 15 are formed by the 20 glutamate residues. For comparison, the structure of rice Os3BGlu7 (PDB entry 2RGM) has only 8 of 21 arginine residues, 5 of 25 lysine residues, 3 of 11 histidine residues, 14 of 24 aspartate residues, and 4 of 12 glutamate residues involved in intramolecular salt bridges. Thermophilic proteins have been observed to have more salt bridges than their mesophilic relatives (Tanner *et al.*, 1996), so the greater number of salt bridges in Os4BGlu12 may contribute to its greater stability than Os3BGlu7, along with other factors.

In the asymmetric unit of Os4BGlu12 crystals, His69 of both molecules and Asp66 of molecule A and Asp36 of molecule B chelate a Zn^{2+} ion asymmetrically between the two protein molecules, as seen in Figure 3.12, and shown in detail in Figure 3.14, but this stabilizing interaction is not expected to occur in solution. Os4BGlu12 is a monomeric protein in solution, as judged by gel filtration and dynamic light scattering, and the maximum interface of 488 \AA^2 between molecules A

and B, 2.7% of the average monomer surface area of 18,307 Å², is predicted to not result in a dimer in solution by the PISA program (Krissinel and Henrick, 2007). The Zn²⁺ ion appears to have been a fortuitous contaminant in the purification solution or precipitant, which facilitated the contacts between the proteins in the crystal matrix, similar to our group's previous experience with Os3BGlu7 (Chuenchor *et al.*, 2008).

4.1.2 The binding of ligand in the active site of Os4BGlu12

DNP2FG has been used to show that the enzymatic mechanism of retaining glycosidases proceeds via a covalent glucosyl-enzyme intermediate, because the electron withdrawal by the 2-fluoro group slows both the glycosylation and deglycosylation steps (the first and second reactions of Figure 4.1, respectively), but the use of a good leaving group (2,4-dinitrophenolate) as the aglycone allows the covalent intermediate with the nucleophile residue to form faster than hydrolysis (Withers *et al.*, 1990). However, it has been shown that acceptors for transglycosylation may allow rescue of the enzyme by accepting the 2-fluoroglucosyl moiety (Malet and Planas, 1998; Okuyama *et al.*, 2002 and Hommalai *et al.*, 2007). In the case of the Os4BGlu12 complex structures, the inclusion of 10 mM DNP2FG in the cryoprotectant might have allowed rescue of the enzyme, and resulted in the unhydrolyzed DNP2FG being seen in the active site. When the co-crystals were flash cooled after a 15 s incubation in cryoprotectant without DNP2FG, their diffraction yielded an electron diffraction map that showed the covalent intermediate of α -2-fluoroglucoside (G2F) bound to the O ϵ 2 of the catalytic nucleophile, Glu393, in a relaxed ⁴C₁ chair conformation (Figure 3.18 A and B). Thus, the use of different cryoprotectant conditions allowed solution of the structures of both the Michaelis

complex of DNP2FG and the covalent intermediate with G2F, which supported the postulated conformational trajectory of the glucosyl residue through a 1S_3 skew boat to a 4H_3 half chair transition state to the 4C_1 chair covalent intermediate.

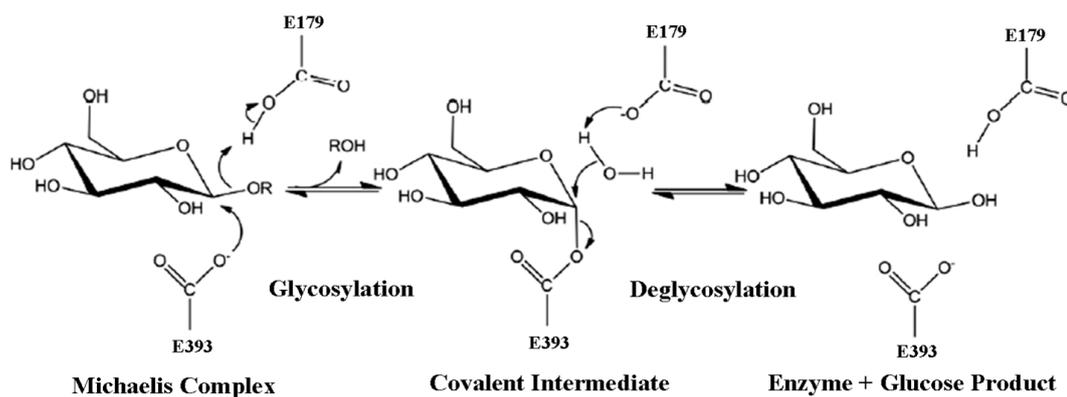


Figure 4.1 Mechanism of retaining β -glycosidases, as seen in rice Os4BGlu12.

Upon formation of the Michaelis complex between enzyme and substrate, the reaction proceeds via two steps, glycosylation to form the covalent intermediate, and deglycosylation to release the sugar from the protein. In Os4BGlu12, Glu179 serves as the catalytic acid/base, while Glu393 serves as the catalytic nucleophile. The figure is from Sansenya *et al.*, 2011.

4.1.2.1 Michaelis complex of Os4BGlu12 with DNP2FG

Figure 3.17 B, shows the Os4BGlu12 Michaelis complex with DNP2FG, in which the Glu393 O ϵ 1 was hydrogen bonded with O5 (3.09 Å) and within 3.2 Å of the anomeric carbon of the sugar ring. The O ϵ 2 of Glu393 was within hydrogen bonding distance of the F2 (2.62 Å), suggesting it would likely hydrogen bond to the O2 atom of a normal glucosyl moiety. The distance between O ϵ 2 of the catalytic

acid/base residue Glu179 and O1 of DNP2FG was 3.30 Å. Thus, the ¹S₃ conformation of the Michaelis complex placed the substrate in position for nucleophilic attack and protonation by the catalytic acid. The glycone residue formed hydrogen bonds with 6 residues around the -1 subsite (Gln29, His133, Tyr322, Glu393, Gln449 and Trp450). In addition to the interactions described above, O3 of the 2-F-glucose was hydrogen bonded with His133 Nε2 (2.88 Å), Glu29 Oε1 (2.75 Å) and Trp450 Nε1 (2.82 Å). O4 hydrogen bonds with Gln29 N ε 2 (2.71 Å) and Glu449 O ε 1 (2.54 Å), while O5 hydrogen bonds with the phenolic hydroxyl of Tyr322 (2.80 Å) and O6 with Glu449 Oε2 (2.63 Å). These hydrogen bonds between the glucose and the active site residues have been observed for other GH1 β-glucosidase/substrate complexes (Verdoucq *et al.*, 2004). Other conserved residues making up the glycone binding site include Trp134, Asn178, Asn320, Trp365, Trp442 and Phe458. The aglycone binding site is occupied by hydrophobic residues and a few polar residues. The 2,4-dinitrophenyl group was bound by hydrophobic and aromatic stacking interactions with Trp181 and Phe193 on one side and Trp365 on the other. The polar residues found in the aglycone binding subsite include Thr182, Asn186, Asn325, and Asn452. In GH1, Trp365 is a conserved residue, which corresponds to Trp358 in Os3BGlu7 and Trp366 in Os3BGlu6 (Figure 4.2 A and B).

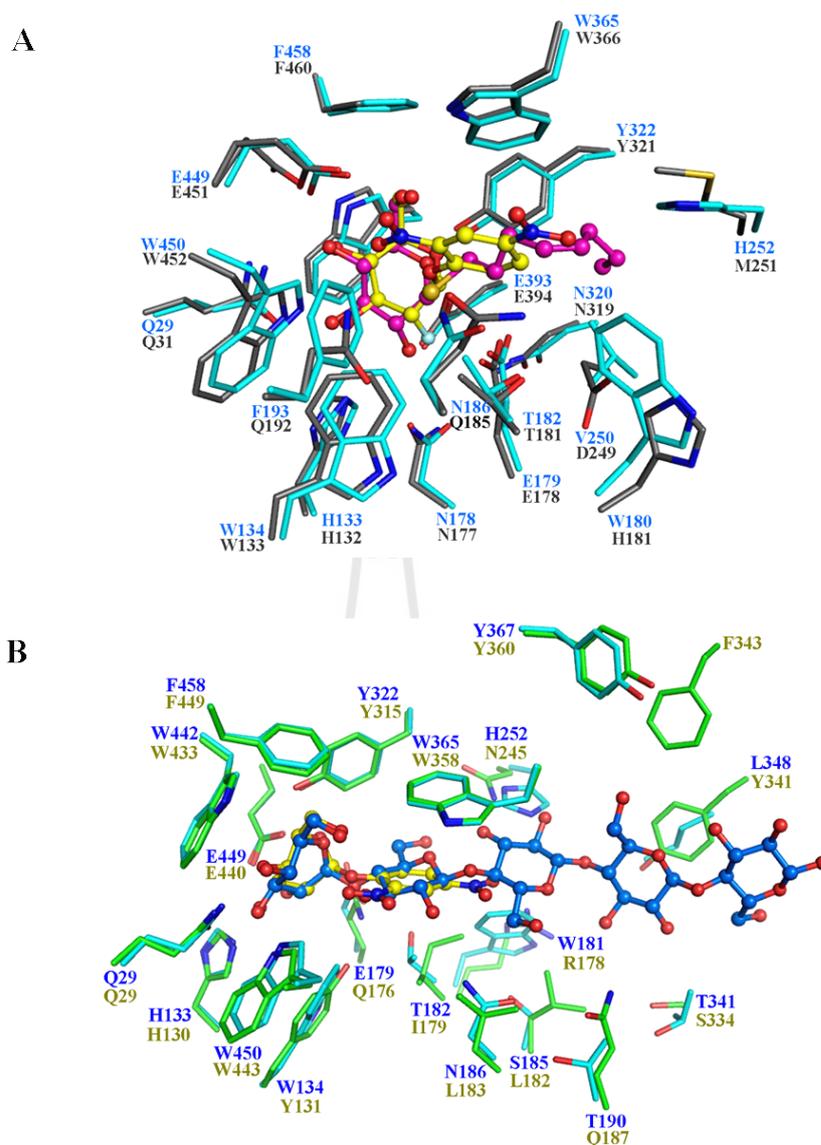


Figure 4.2 Comparison of ligand binding by three rice β -glucosidases.

(A) Superimposition of Os4BGlu12_DNP2FG (cyan carbons) with Os3BGlu6 with *n*-octyl- β -D-thioglucopyranoside (gray carbons). The ligands are shown in ball and stick representations, with pink carbons for DNP2FG and yellow carbons for *n*-octyl- β -D-thioglucopyranoside. (B) Superimposition of the binding of DNP2FG by Os4BGlu12 (cyan carbons) and cellopentaose by Os3BGlu7 (3F5K) (green carbons). The ligand carbons are drawn in yellow for DNP2FG and blue for cellopentaose.

4.1.2.2 The covalent intermediate of Os4BGlu12 with G2F

In the structure of Os4BGlu12 with a covalently bound G2F, the anomeric carbon formed a covalent bond with the Glu393 nucleophile residue O ϵ 1 and the G2F was observed in a 4C_1 chair conformation of the ring in subsite -1 (Figure 3.18 A and B). Aside from the covalent bonding to the nucleophile, the glucose is bound by hydrogen bonds similar to the Michaelis complex. Figure 3.18 B shows that the density was clear and Tyr322 was seen to have an alternative conformation, in which it can hydrogen bond to the catalytic nucleophile in this complex, in addition to a conformation hydrogen bonding to O5, similar to the noncovalent complex. The major structural differences between the apo enzyme and covalent intermediate complexes of many GH1 enzymes are the positions of the nucleophilic residues. In the complex structures of the *P. polymyxa*, *S. solfataricus*, *T. maritima*, rice Os3BGlu6, rice Os3BGlu7, and human cytosolic β -glucosidases, the side-chain position of the nucleophilic residue has moved to link with anomeric carbon of G2F. For example, a displacement of C δ by 1.5-1.6 Å is observed when the rice Os3BGlu7 covalent complex, 2RGM (pink structure in Figure 3.18 C), is compared to the apoenzyme structure, 2RGL (Chuenchor *et al.*, 2008). In these structures, the angles between the sugar anomeric carbon, O ϵ 1 and C δ of the nucleophilic residue, fall in the range of 113-117°. In Os4BGlu12, the C δ is only displaced 0.3 Å (molecule B) to 0.7 Å (molecule A), and the angle between the sugar anomeric carbon and Glu393 O ϵ 1 and C δ is strained (142°). This is similar to the *S. alba* myrosinase (shown in yellow in Figure 3.18 C), an *S*-glycosidase, where the displacement of C δ is only 0.2 Å in the covalent G2F complex (PDB accession 1E73) compared to the apo protein (1MYR) and the angle is similarly wide (138°) (Burmeister *et al.*, 2000).

4.1.2.3 Thioglucosidase activity

The similarity of the covalent intermediate of Os4BGlu12 to that of *S. alba* myrosinase suggested that Os4BGlu12 might show mechanistic similarities to the myrosinase, so its ability to hydrolyze thioglucosides was tested. Os4BGlu12 could hydrolyze the *S*-glycosides *p*NP- β -D-thioglucoside and *n*-octyl- β -D-thioglucoside, as shown in Table 3.6. Os4BGlu12 β -glucosidase catalyzed the hydrolysis of *n*-octyl- β -D-thioglucopyranoside with about 1000-fold lower efficiency than it hydrolyzed *n*-octyl- β -D-glucopyranoside and *p*NP- β -D-thioglucoside with about 200-fold less efficiency than it hydrolyzed *p*NP- β -D-glucoside. The K_m values of *S*-glycosides were higher than the corresponding *O*-glycosides by only about 3-fold, showing that most of the difference was due differences in the rates of the forward reactions.

Commercial sweet almond β -glucosidase has been shown to have thioglycohydrolase activities similar to Os4BGlu12, and the K_m values for the *S*- and *O*-glycosidase are also similar (0.9 mM for *p*NP- β -D-fucoside and 1.6 for *p*NP- β -D-thiofucoside), while the k_{cat}/K_m value was about 1000-fold lower for the *S*-glycoside (Shen *et al.*, 2007). In contrast, Os3BGlu6 and Os3BGlu7 exhibited no significant hydrolysis of *p*NP- β -D-thioglucoside under the conditions of the assay, although a slight hydrolysis of *p*NP- β -D-thioglucoside by Os3BGlu7 was recently reported at high concentrations of substrate and enzyme in NMR (Kuntothom *et al.*, 2010). This is consistent with the fact that their covalent intermediate structures were different from myrosinase and Os4BGlu12, but the reason for this correlation remains unclear.

It has been suggested in the myrosinase literature that classic myrosinases are derived from the defense-related β -glucosidases, such as cyanogenic β -glucosidases (Wang *et al.*, 2009). Indeed, the minimum evolution-based phylogenetic tree groups the myrosinases group with cluster At/Os7, which contains the cyanogenic β -glucosidases, although in our previous analyses (Opassiri *et al.*, 2006) and distance-based neighbor-joining and maximum likelihood trees of the same sequences failed to show a significance to this association. The structural similarity of the covalent intermediate to that of *S. alba* myrosinase and the higher thioglucosidase activity of Os4BGlu12 and almond β -glucosidase (Shen *et al.*, 2007) compared to the more distantly related Os3BGlu6 and Os3BGlu7 provide mechanistic support for this similarity.

4.1.2.4 The hydrolysis of natural glucoside substrate by rice Os4BGlu12

Recently a rice tuberonic acid glucoside (TAG) β -glucosidase has been identified and designated OsTAGG1 (Wakuta *et al.*, 2010). The OsTAGG1, which is designated Os4BGlu13 in the genome-base systematic nomenclature of Opassiri *et al.* (2006), has about 80% amino acid sequence identity with rice Os4BGlu12 β -glucosidase, and Os4BGlu12 was also identified as a tuberonic acid glucoside β -glucosidase (designated OsTAGG2) by Wakuta *et al.* (2011). The enzymic properties of the two enzymes are quite similar with pH optima of 4.5 for Os4BGlu12 and 4.0 for Os4BGlu13, and thermal stability at ≤ 55 °C for Os4BGlu12 and at ≤ 50 °C for Os4BGlu13, while the optimum temperature and pH stability of both enzymes are the same (Wakuta *et al.*, 2010 and 2011). The kinetic parameters for hydrolysis of TAG

substrate are different for the two enzymes. The K_m of Os4BGlu12 ($146 \pm 10 \mu\text{M}$) is four times higher compared to Os4BGlu13 ($31.7 \pm 1.4 \mu\text{M}$), while the V_{max} of Os4BGlu12 ($38.0 \pm 2.5 \mu\text{mol/mg/min}$) is 2.5 times higher than Os4BGlu13 ($14.7 \pm 0.4 \mu\text{mol/mg/min}$) (Wakuta *et al.*, 2011).

The amino acid residues at the -1 subsite in the crystal structure of Os4BGlu12 E179Q with TAG are conserved with those in the model structure of Os4BGlu13 with TAG (Figure 4.3). The glucosyl residues interacted with Gln29, His133, Trp134, Asn178, Glu179, Tyr322, Glu393, Trp442, Glu449, Trp450 and Phe458 in the crystal structure of Os4BGlu12 and with the corresponding Gln49, His153, Trp154, Asn198, Glu199, Tyr342, Glu413, Trp462, Glu469, Trp470 and Phe478 in the modeled structure of Os4BGlu13 with TAG. The hydrogen bonding patterns of the Os4BGlu12 crystal and modeled Os4BGlu13 structures are similar at the -1 subsite. The aglycone binding site of the crystal structure of Os4BGlu12 E179Q with TAG is occupied by both hydrophobic and polar residues, which are different from those in the modeled structure of Os4BGlu13 with TAG. The TA group was bound by Trp181, Thr182, Ser185, Asn186, His252 and Trp365 in the TAG complex with Os4BGlu12 E179Q, while these residues are replaced by Leu201, Ser202, Val205, Ala206, Asn272 and Trp385 in the modeled structure of Os4BGlu13 with TAG (Figure 4.3). The differences in the amino acid residues surrounding the TA group between the structure of Os4BGlu12 E179Q with TAG and the model of Os4BGlu13 with TAG may explain why the hydrolysis efficiency of TAG substrate by the two enzymes is different, although both can bind and hydrolyze this substrate.

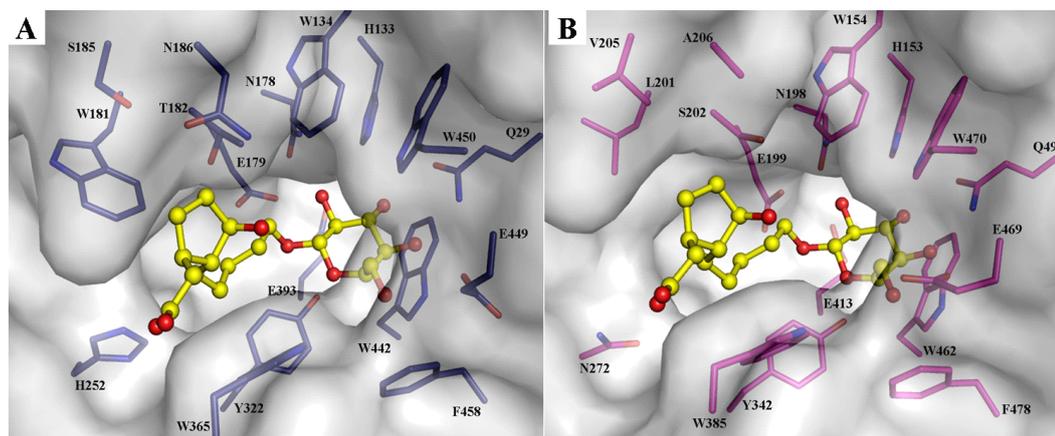


Figure 4.3 Active site cleft of the crystal structure of Os4BGlu12 E179Q with TAG and the homology model of Os4BGlu13 with TAG.

A, the active site surface cleft of Os4BGlu12 E179Q with TAG and B, Os4BGlu13 model with TAG. The amino acid residues are drawn as stick structures with carbon in purple for Os4BGlu12 E179Q and pink for Os4BGlu13, oxygen in red and nitrogen in blue. The TAG is represented in ball and stick with carbon in yellow and oxygen in red for both structures.

4.2 Structures of Os3BGlu6 covalent glucosyl-intermediate

Rice Os3BGlu6 shown higher hydrolytic activity on GA₄GE than other rice enzymes tested (Os3BGlu7, Os4BGlu12, Os4BGlu18 and Os9BGlu31), while its acid/base mutant Os3BGlu E178Q could produce β-D-glucosyl azide from transglycosylation reactions with GA₄GE as a glucosyl donor substrate and azide as an acceptor (Hua *et al.*, 2013). The crystals of the Os3BGlu6 mutants E178Q and E394D were produced to obtain the enzyme-GA₄GE complex. This noncovalent GA₄GE complex with an Os3BGlu6 mutant was not obtained, due to the high reactivity of this substrate, which resulted in poor density in the E394D nucleophile

mutant active site and the covalent glucosyl-enzyme intermediate in the E178Q acid/base mutant.

Figure 4.4 shows that the positions and interactions made by amino acid residues surrounding the native glucosyl moiety in the active site of Os3BGlu6 E178Q with glucose from GA₄GE (PDB: 3WBE) or *p*NPGlc (PDB: WBA) were similar to those interacting with the 2-deoxy-2-fluoroglucopyranoside residue (G2F) in the structure of wild type Os3BGlu6 with G2F (3GNR). In the Os3BGlu6/G2F complex (Seshadri *et al.*, 2009), the density of the 2-deoxy-2-fluoroglucosyl residue matched a relaxed ⁴C₁ chair conformation, as did that for the glucosyl moiety in the Os3BGlu6 E178Q covalent intermediate structure (Figure 4.4D). However, the distances between the anomeric carbon of glucose and the Glu394 nucleophile residue of Os3BGlu6 E178Q (1.5 Å) was significantly longer than in the G2F covalent complex (1.34 Å). A water molecule was observed between the 2-OH of glucose or fluorine atom of G2F and N177 and E394 in the Os3BGlu6 E178Q/Glc and Os3BGlu6/G2F structures. A water molecule in this position has been shown to be critical for the hydrolysis and glycosynthase reactions of Os3BGlu7 (Wang *et al.*, 2013). The hydrogen bonding pattern between the glucosyl moiety and residues at the -1 subsite of the Os3BGlu6 E178Q/Glc covalent complex structure is similar to that of the Os3BGlu6/G2F complex (Figure 4.4 A, B and C). The average distances between the 2-OH of the glucose and 2-fluorine of the G2F moieties and the polar atoms of the surrounding amino acid residues are similar (Figure 4.4 A-D). Many structures of GH1 glycosyl-enzyme intermediates with a 2-deoxy-2-fluoroglucosyl (G2F) moiety bound to the nucleophilic residue have been reported, including one for Os3BGlu6 (PDB code 3GNR) (Burmeister *et al.*, 1997, Zechel *et al.*, 2003, Gloster *et*

al., 2004, Isorna *et al.*, 2007, Chuenchor *et al.*, 2008, Seshadri *et al.*, 2009 and Sansenya *et al.*, 2011). These GH1 glycosyl enzyme intermediates used 2,4-dinitrophenyl 2-deoxy-2-fluoroglucoside, with which the substitution of an electronegative fluorine atom for a hydroxyl group at C-2 adjacent to the reaction center destabilizes the transition states and decreases the rates of both glycosylation and deglycosylation (Rempel and Withers, 2008). The human cytosolic β -glucosidase acid/base mutant glucosyl complex (2ZOX) is the only previously reported GH1 structure with the nucleophile covalently bound to native glucose (Noguchi *et al.*, 2008). The hydrogen bonding pattern between the G2F molecule and the enzyme at the -1 subsite is conserved in other reported GH1 covalent intermediates, but this pattern lacks the hydrogen bonds for which the 2-OH is the proton donor seen in human cytosolic β -glucosidase and Os3BGlu6 E178Q covalent glucoside complex structures. In general, the reports of G2F complexes do not report hydrogen bonds to the 2-F group, but the conservation of the water in proximity to the fluorine in these complexes supports the idea that the fluorine may accept a hydrogen bond from the water and other nearby hydrogen donors. On the other hand, the loss of possible hydrogen bonds from the 2-OH to O δ of N177 and O ϵ 2 of E394 in Os3BGlu6, which are seen in the native glucosyl-enzyme intermediates (Figure 4.4), may result in a further loss of stabilization of the transition state in the 2-F-glucoside reaction (Rempel and Withers, 2008).

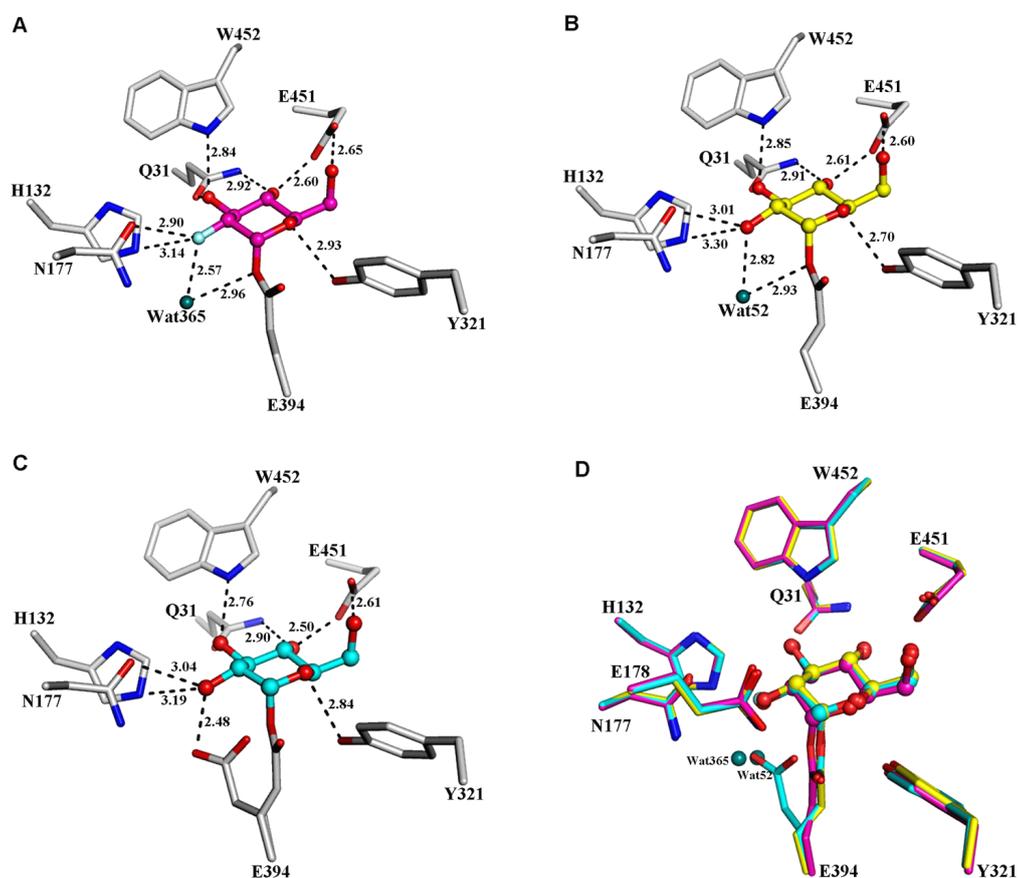


Figure 4.4. Comparison of the active site region of Os3BGluc6 bound with G2F (A) and those of the structures of Os3BGluc6 E178Q soaked with *p*NPGlc (B) and Os3BGluc6 E178Q soaked with GA₄GE (C) (Hua *et al.*, 2013).

The G2F and glucosyl moieties are represented as balls and sticks with carbon in pink for G2F, in yellow for Os3BGluc6 E178Q soaked with *p*NPGlc and cyan for Os3BGluc6 E178Q soaked with GA₄GE, oxygen in red and fluorine in pale cyan. The amino acid residues surrounding the ligands are represented by sticks with carbons in gray, nitrogens in dark blue and oxygens in red for all three structures. (D) Superposition of the three structures in A, B and C, with *p*NPGlc (yellow carbons) and GA₄GE (cyan carbons) and to the G2F moiety in Os3BGluc6 (3GNP, pink carbons).

4.3 Comparison of oligosaccharide binding by rice β -glucosidases

The three rice β -glucosidases for which structures have been determined, Os3BGlu6 (Figure 4.5 A), Os3BGlu7 (Figure 4.5 B) and Os4BGlu12 (Figure 4.5C), come from three distinct phylogenetic branches of plant GH1 enzymes, yet all hydrolyze oligosaccharides and glycosides, although with different preferences. Os3BGlu6 only hydrolyzes β -1,2- and β -(1,3)-linked disaccharides (sophorose and laminaribiose) efficiently, while Os3BGlu7 and Os4BGlu12 efficiently hydrolyze β -(1,4)-linked gluco-oligosaccharides (cellooligosaccharides) of 3-6 residues, as well as laminaribiose (Opassiri *et al.*, 2010). The superimposition of rice Os3BGlu6 and Os3BGlu7 found that Asn245 at subsite +2 for binding cellooligosaccharides in Os3BGlu7 (Figure 4.5 A) is replaced with Met251 in Os3BGlu6 (Figure 4.5 B) (Seshadri *et al.*, 2009). This Met251 in Os3BGlu6 is in the same position as Asn245 in Os3BGlu7, and the Os3BGlu7 N245V mutation gave a k_{cat}/K_m value for hydrolysis of cellotriose 15 times lower than wild type Os3BGlu7, due to a 42-fold increase in K_m (Chuenchor *et al.*, 2008). These effects show the importance of Asn245 in Os3BGlu7 binding of the third glucosyl residue in cellooligosaccharides (Chuenchor *et al.*, 2011). The corresponding residue in Os4BGlu12 is His252 (Figure 4.2 A and B), which also has the potential to hydrogen bond to the third glucosyl residue in glucooligosaccharides, thereby explaining the binding of Os4BGlu12 to cellotriose and longer oligosaccharides.

The superimposition also showed that the Met251, which extends into the Os3BGlu6 active site cleft, would interfere with binding of cellooligosaccharides in the same mode as in Os3BGlu7, since it comes within 0.9 Å of O2 of the third glucosyl residue of cellotetraose or cellopentaose, when the Os3BGlu6 structure is

superimposed on structures of those complexes, as illustrated in Figure 3.22. When celotriose was docked into the active site, the nonreducing glucosyl residue in the +2 subsite turned away from this position in Os3BGlu6, but the subsite -1 and +1 glucosyl residues were in similar positions to those in celooligosaccharides in the Os3BGlu7 active site (Figure 4.5 A and B). Replacement of Met251 in Os3BGlu6 with the hydrogen-bonding Asn of Os3BGlu7 afforded the opportunity to introduce a new glucosyl residue-binding subsite, while relieving the steric block. Consequently, Os3BGlu6 M251N showed a 10 to 24-fold increase in the $k_{\text{cat}}/K_{\text{m}}$ values for all the oligosaccharide substrates compared with the wild type enzyme. The Os3BGlu6 M251N mutation increased the efficiency of hydrolysis ($k_{\text{cat}}/K_{\text{m}}$) for laminaribiose and cellobiose by 15 to 16-fold. The greatest change was seen for celotriose, where the 24-fold increase in $k_{\text{cat}}/K_{\text{m}}$ corresponds to an 8 kJ/mol improvement in the binding of the transition state ($\Delta\Delta G_{\text{S}^*_{\text{mut}}} = -8$ kJ/mol). The $k_{\text{cat}}/K_{\text{m}}$ of celotriose was threefold that of cellobiose, but that of cellotetraose was within error of that of celotriose, while that of cellopentaose was lower. Thus, the Asn251 in the Os3BGlu6 mutant can carry out a similar role to Asn245 in Os3BGlu7 in stabilizing β -(1,3)- and β -(1,4)-linked gluco-disaccharide binding and produce a binding site for the third glucosyl residue in celooligosaccharides. The fact that cellotetraose and longer gluco-oligosaccharides were hydrolyzed with similar or lower $k_{\text{cat}}/K_{\text{m}}$ values suggests that only the +2 subsite was created by the mutation and no further binding sites are accessible for longer celooligosaccharides.

It was hypothesized that replacement of Os3BGlu7 Asn245 with Met might provide a steric block to celooligosaccharide binding, similar to the role of Met251 in Os3BGlu6. The Os3BGlu7 N245M mutant showed a 6.5-fold drop in the $k_{\text{cat}}/K_{\text{m}}$ value

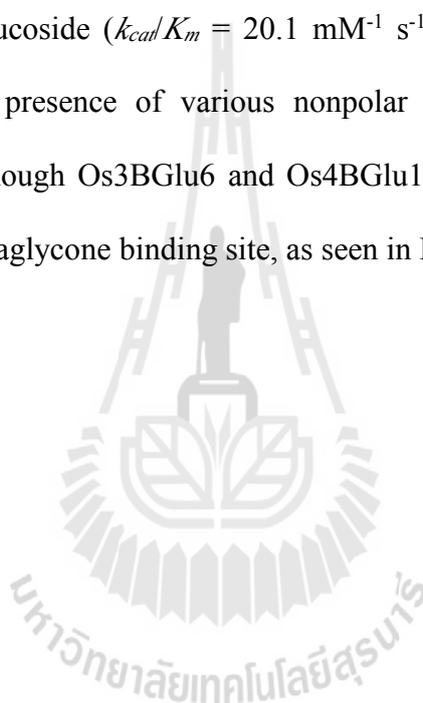
for laminaribiose and 17-fold to 30-fold drops in those for cellooligosaccharides with DP of 3–6 compared to wild type. Os3BGlu7 N245M did show an increase in k_{cat}/K_m values with increasing DP from 2 to 6, similar to the wild type enzyme, but the largest decrease in k_{cat}/K_m compared to wild type was seen at cellohexaose. The decreases in k_{cat}/K_m for cellobiose and cellotriose were similar to the 2-fold and 15-fold drops for these substrates in the Os3BGlu7 N245V mutant, which was not expected to introduce a steric block (Chuenchor *et al.*, 2008). This suggests that the introduced Met may not extend into the active site cleft in Os3BGlu7 as much as in Os3BGlu6, or that the substrate is able to adjust its position without further loss of energy. We recently observed that mutation of Os3BGlu7 Tyr341, which interacts with glucosyl residues in the +2 to +4 subsites (Chuenchor *et al.*, 2011) resulted in cellotetraose binding in an alternative productive position, which interacted less with Asn245 (Pengthaisong *et al.*, 2012). So, the steric influence of the Met could cause cellotriose and longer oligosaccharides to bind in this alternative position.

When the structure of Os4BGlu12 was superimposed on the Os3BGlu7 cellopentaose and laminaribiose complex structures, His252 of Os4BGlu12 appeared to be within range to form hydrogen bonds with the glucosyl residues in the +1 and +2 subsites. Docking of cellotriose into Os4BGlu12 placed the 3 glucosyl residues in nearly the same positions in the lowest energy conformation, including a possible hydrogen bond from O3 of the glucosyl residue in subsite +2 to His252 (Figure 4.5C). Mutation of Os4BGlu12 His252 to Met decreased the k_{cat}/K_m three-fold for laminaribiose and 2 to 6-fold for cellooligosaccharides with DP of 2–5. The relatively minor effect on cellotriose hydrolysis ($\Delta\Delta G_{\text{S}^*\text{mut}}$ of 2.3 kJ/ mol), compared to the corresponding mutation in Os3BGlu7 ($\Delta\Delta G_{\text{S}^*\text{mut}}$ of 7.2 kJ/ mol), shows weaker

interaction between His252 and the oligosaccharide. This suggests that the position of binding of cellooligosaccharides by Os4BGlu12 is similar to, but not the same as, that seen in Os3BGlu7, at least at the +2 subsite.

The efficiency of hydrolysis increases with the length of cellooligosaccharides from 3 to 6 for Os3BGlu7, but shows little change as the length increases in Os4BGlu12, which maintains a rather high K_m of approximately 5 mM for all these cellooligosaccharides (Opassiri *et al.*, 2010). The distance from the deepest part of the active site slot to the surface entrance of Os4BGlu12 is approximately 20 Å, but the lack of increase in affinity for longer cellooligosaccharides suggests the outer parts of this cleft have little interaction with these cellooligosaccharides. Figure 4.2 B shows that superimposition of the structures of Os4BGlu12 with DNP2FG and Os3BGlu7 E176Q with cellopentaose (3F5K) (Chuenchor *et al.*, 2011) indicated that, while all residues in the -1 subsite of these two enzymes are the same, except for the Trp134 in Os4BGlu12 being replaced by Tyr131 in Os3BGlu7, their aglycone-binding sites show little similarity. Although both contain the conserved Trp at the +1 and +2 subsites (Trp 365 in Os4BGlu12, Trp 358 in Os3BGlu7), the aglycone-binding subsite residues Trp181, Thr182, Ser185, Asn186, Thr190, Phe193, Phe254 and His252, in Os4BGlu12 are replaced by Arg178, Ile179, Leu182, Leu183, Gln187, Asn190, Asn245 and Tyr247, respectively, in Os3BGlu7. The Tyr341 residue, which stacks the residues in the +3 and +4 subsites in Os3BGlu7 (Chuenchor *et al.*, 2011) is replaced by Leu348 in Os4BGlu12, which may explain the relatively poor binding of cellotetraose and longer oligosaccharides in Os4BGlu12 compared to Os3BGlu7 (Opassiri *et al.*, 2010).

Os4BGlu12 had a 2-fold higher K_m for *n*-heptyl- β -D-glucopyranoside than for *n*-octyl- β -D-glucopyranoside, suggesting that Os4BGlu12 may prefer glycosides with more extensive apolar regions. This is similar to Os3BGlu6, which showed a similar preference for octyl over heptyl glucoside (Seshadri *et al.*, 2009), but Os4BGlu12 is over 10 times more efficient at hydrolyzing both substrates than Os3BGlu6. Os4BGlu12 was previously noted to efficiently hydrolyze the large, hydrophobic deoxycorticosteroid glucoside ($k_{cat}/K_m = 20.1 \text{ mM}^{-1} \text{ s}^{-1}$) (Opassiri *et al.*, 2010). This seems to reflect the presence of various nonpolar groups in the active site, as mentioned above, although Os3BGlu6 and Os4BGlu12 show little difference in the overall polarity of the aglycone binding site, as seen in Figure 4.2A.



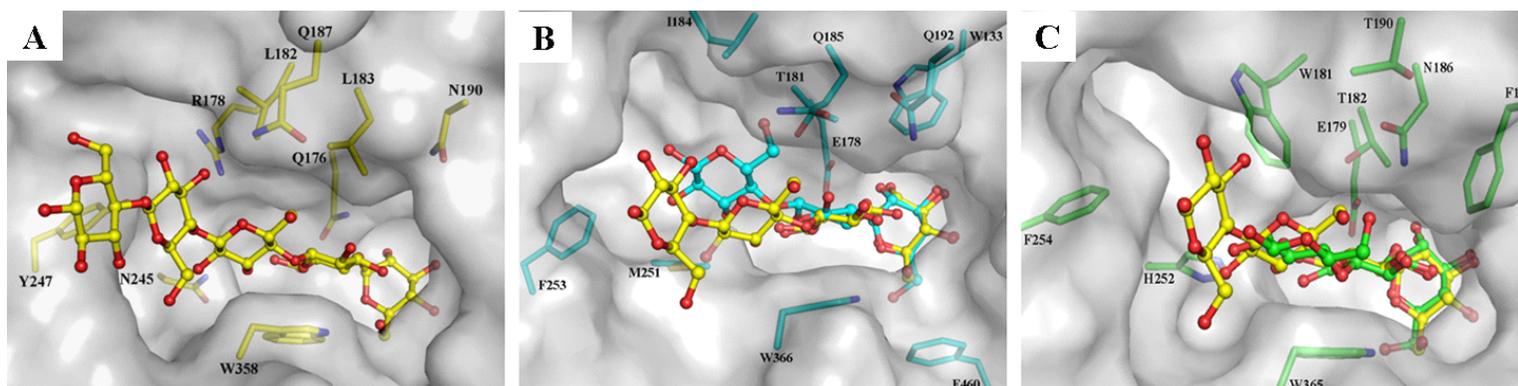


Figure 4.5 Comparison of cellobiosaccharide binding in the active sites of Os3BGlu7, Os3BGlu6 and Os4BGlu12.

(A) The active site of the Os3BGlu7 E176Q complex with cellopentaose showing the surface and underlying residues that interact with the ligand. (B) The active site of Os3BGlu6 with cellotriose docked into the active site and the cellotetraose from the superimposed complex of Os3BGlu7 (PDB ID: 3FKJ) shown for comparison. (C) The active site of Os4BGlu12 with cellotriose docked into it and cellotetraose from the superimposed complex of Os3BGlu7 with cellotetraose shown for comparison. In all panels, the amino acids are represented by sticks and ligands by balls and sticks with oxygen in red, sulfur in dark yellow and nitrogen in dark blue, carbon in yellow for Os3BGlu7 with cellopentaose or cellotetraose, cyan for Os3BGlu6 with *n*-octyl- β -D-thioglucopyranoside and green for Os4BGlu12 with 2,4-dinitrophenyl 2-deoxy-2 fluoro- β -D-glucopyranoside. Docking of cellotriose in B and C was done with Autodoc 4.0 (Morris *et al.*, 1998).

CHAPTER V

CONCLUSION

The thioredoxin/His₆-tagged Os4BGlu12 fusion protein was expressed in the soluble fraction in *E. coli* strain Origami B(DE3) cells and purified for crystallization. After purification, the Os4BGlu12 protein with the N-terminal tag removed was approximately 95% pure and had a yield of approximately 2.5 mg purified protein per liter of LB medium. Two mutations of Os4BGlu12 (E179Q and H252M) were generated and expressed in the same system with native enzyme and the yields and purities of both mutant proteins were similar to those of native Os4BGlu12 protein. The purified Os4BGlu12 from which the fusion tag had been removed was used to test the hydrolysis ability of various substrates, similar to the previous characterization of the full fusion protein (Opassiri *et al.*, 2010). Among *p*NP and *o*NP glycosides, Os4BGlu12 hydrolyzed *p*NP-β-D-fucoside, *p*NP-β-D-glucoside, *p*NP-β-D-galactoside, *o*NP-β-D-glucoside, *p*NP-β-D-xyloside and *p*NP-β-D-mannoside, in order of efficiency of hydrolysis. Among glucosyl disaccharides, laminaribiose (β-1→3-linked) was hydrolyzed most rapidly, followed by sophorose (β-1→2-linked) and gentiobiose (β-1→6-linked), while among β-1→4-linked, Os4BGlu12 hydrolyzed cellotetraose, cellotriose and cellobiose, in order of decreasing relative activity. Among natural glucosides, Os4BGlu12 had hydrolytic

p-SAG (172.5%), *m*-SAG (100.5%), SAG (99.3%), TAG (82.5%) and salicin (0.3%) respectively, but did not hydrolyze arbutin.

The initial native crystal of wild type Os4BGlu12 was obtained from a precipitant consisting of 25% (w/v) PEG 4000 in 0.1 M Tris-HCl, pH 8.5, 0.2 M NaCl within 17 days under with the microbatch under oil technique at 15 °C. Optimization of the native crystal conditions was carried out by hanging-drop vapor diffusion with microseeding to obtain crystals of sufficient size to allow X-ray diffraction. Native Os4BGlu12 crystals were obtained from the precipitant consisting of 19% (w/v) PEG 3350 in 0.1 M Tris-HCl, pH 8.5, 0.16 M NaCl with the maximum size of 120 x 25 x 25 µm. Both the Michaelis complex with DNP2FG and the subsequent covalent intermediate with G2F could be observed by including or omitting DNP2FG from the cryoprotectant, respectively. These two crystals were obtained from the co-crystallization condition with 2 mM DNP2FG and the precipitant consisting of 19% (w/v) PEG 2000 in 0.1 M Tris-HCl, pH 8.5, 0.16 M NaCl for the Os4BGlu12 crystal with DNP2FG and 19% (w/v) PEG 2000 in 0.1 M Tris-HCl, pH 8.5, 0.18 M NaCl for the Os4BGlu12 crystal with G2F, respectively. The crystal of Os4BGlu12E179Q with TAG was obtained from the crystallization condition including 19% (w/v) PEG 4000 in 0.1 M Tris-HCl, pH 8.5, 0.26 M NaCl and 2 mM TAG. Diffraction data sets for the apo Os4BGlu12 structure and inhibitor-bound (DNP2FG and G2F) and TAG complexes were collected to 2.50, 2.45, 2.40 and 3.2 Å resolution, respectively. The space group and the unit-cell parameters of the crystals indicated the presence of two molecules per asymmetric unit, with a solvent content of 50%. Os4BGlu12 crystals exhibited $P4_32_12$ space group symmetry. The apo Os4BGlu12 structure was solved by molecular replacement using the white

clover cyanogenic β -glucosidase structure (PDB code: 1CBG) as a search model. The structures of Os4BGlu12 in complexes with the ligands DNP2FG and G2F and the Os4BGlu12 mutant E179Q complex with TAG were solved by rigid body refinement with the apo-Os4BGlu12 structure as the template model. Several cycles of model rebuilding and refinement resulted in convergence to give the final R and free R (R_{free}) factors of 20.3% and 24.8% for apo Os4BGlu12, 20.9% and 25.1% for the Os4BGlu12 complex with DNP2FG, 19.6% and 23.4% for its complex with G2F, and 23.4% and 27.9% for Os4BGlu12 E179Q with TAG.

The overall structure of Os4BGlu12 exhibits the $(\beta/\alpha)_8$ barrel fold characteristic of GH1 enzymes. Os4BGlu12 is a monomer in solution, while the crystal contains a crystallographic dimer, which is stabilized by a Zn^{2+} ion bound between two molecules in the asymmetric unit, by four residues (Asp66 and His69 from molecule A and Glu36 and His69 from molecule B). The structure of Os4BGlu12 contains two disulfide bridges (non-conserved Cys184-Cys219 and conserved Cys198-Cys206), located in loop B, while most plant GH1 β -glucosidases only have the conserved disulfide bond. However, the extra disulfide bond did not appear to explain the higher stability of Os4BGlu12, compared to Os3BGlu7 produced in the same system. In the complex structures with DNP2FG, G2F and TAG, the glucosyl residues were located in a similar position at the -1 subsite and surrounding by conserved amino acid residues, to which the glucosyl group formed hydrogen bonds similar to those seen in other GH1 enzymes. The position of the catalytic nucleophile, Glu393, in the Os4BGlu12/G2F complex was similar to that of the nucleophile in the *S. alba* myrosinase, an *S*-glycosidase, consistent with Os4BGlu12 catalysis of the hydrolysis of *S*-glycosides. The aglycone binding subsite

in the structures of the complexes with DNP2FG and TAG was surrounded by both aromatic and polar amino acids, accounting for its ability to hydrolyze glycosides with a range of hydrophobic and polar aglycones.

The structure of Os4BGlu12 was compared to those of other rice GH1 enzyme with known structures (Os3BGlu6 and Os3BGlu7) to compare the residues affecting the substrate binding specificity. The hydrolysis pattern of oligosaccharides observed in the three rice enzyme was different and the amino acid residues located at the aglycone binding site were also different. At the +2 subsite the residue Asn245, which acts in binding oligosaccharides in the structure of Os3BGlu7, was replaced by His252 in Os4BGlu12, which can also hydrolyze long cellooligosaccharides, while the corresponding position was replaced by Met251 in Os3BGlu6, which hydrolyzes cellooligosaccharides very poorly. This Met251 appeared to block the third glucosyl residue of cellooligosaccharides from binding to Os3BGlu6, if they were bound in the same position as in Os3BGlu7 (Seshadri *et al.*, 2009; Chuenchor *et al.*, 2011; and Sansenya *et al.*, 2012). Conversion of the corresponding Met251 to Asn, created a new cellooligosaccharide binding subsite for Os3BGlu6 and allowed it to hydrolyze cellooligosaccharides with DP 2-5 at much higher rates, while the inverse replacements of Os3BGlu7 Asn245 and Os4BGlu12 His252 with Met decreased hydrolysis of cellooligosaccharides.

The aglycone-binding cleft of Os4BGlu12 appears to be less adapted for binding long cellooligosaccharides than Os3BGlu7 and the amino acid residues at their aglycone binding sites show little similarity. The aglycone binding subsite residues Trp181, Thr182, Ser185, Asn186, Thr190, Phe193, Phe254 and His252, in Os4BGlu12 are replaced by Arg178, Ile179, Leu182, Leu183, Gln187, Asn190,

Asn245 and Tyr247, respectively, in Os3BGlu7. The Tyr341 residue, which stacks the residues in the +3 and +4 subsites in Os3BGlu7 (Chuenchor *et al.*, 2011) is replaced by Leu348 in Os4BGlu12, which may explain the relatively poor binding of cellotetraose and longer oligosaccharides in Os4BGlu12 compared to Os3BGlu7 (Opassiri *et al.*, 2010).

The Os3BGlu6 mutants E178Q and M251N were expressed and purified with the same system in which wild type Os3BGlu6 was previously (Seshadri *et al.*, 2009). After purification, the overall yield obtained from 3.2 liters of bacterial culture was approx. 15-20 mg of protein with more than 95% purity. The crystals of Os3BGlu6 E178Q with *p*NPGlc and Os3BGlu6 E178Q with GA₄GE were obtained by optimizing around the conditions previously used for crystals of apo Os3BGlu6 and Os3BGlu6 complexes with G2F and *n*-octyl- β -D-glucoside (3GNO, 3GNP and 3GNR). The crystals of Os3BGlu6 E178Q with GA₄GE and *p*NPGlc diffracted X-rays to 1.97 Å and 1.90 Å resolution, respectively, and belong to the $P2_12_12_1$ space group, as did the native crystal structure. The overall structure gave visible electron density for residues from Gly11 to Thr488 except for the loop of residues 331 to 337, which was poorly defined by the electron density. When two crystals of Os3BGlu6 E178Q were soaked either with GA₄GE or *p*NPGlc the covalent intermediate was defined at the -1 subsite, between the anomeric carbon (C1) of glucose and O ϵ 1 of the catalytic nucleophile (E394) and the hydroxyls of the glucose ring, which took a ⁴C₁ chair conformation.

When the three Os3BGlu6 covalent intermediate structures Os3BGlu6 E178Q with glucose from GA₄GE (PDB: 3WBE) or *p*NPGlc (PDB: 3WBA) and wild type Os3BGlu6 with G2F (PDB: 3GNR), were superposed, the interaction between the

amino acid residues and glucosyl residues were similar. The catalytic nucleophile (E394) of Os3BGlu6 E178Q with glucose from GA₄GE (PDB: 3WBE) was refined to two conformations, one forming a covalent with anomeric carbon (C1) of the glucosyl moiety and the other one in the position of an apo enzyme catalytic nucleophile, noncovalently bound to a glycerol. In the later conformation, the E394 O ϵ 2 was located the same position as a water molecule (Wat52) in the Os3BGlu6 E178Q with *p*NPGlc structure, when the two structures were superimposed. The water molecule (Wat52) was also present in the structure of the Os3BGlu6/G2F complex and a similar water molecule has been shown to be critical for the hydrolysis and glycosynthase reactions of Os3BGlu7 (Wang *et al.*, 2013). So, the covalent intermediate of Os3BGlu6 E178Q with unmodified α -D-glucoside was observed with either *p*NPGlc or GA₄GE donor substrate, supporting the covalent double displacement mechanism for this enzyme with either glycoside or glycosyl ester substrates.

The structural information and enzyme kinetic evaluation of the function of rice Os4BGlu12 and its comparison to other rice GH1 enzymes (Os3BGlu6 and Os3BGlu7) from these studies provide a basis for understanding the substrate-enzyme interactions. At the glycone binding site, the amino acid residues were conserved when compare to other GH1 enzymes, while at the aglycone binding site rice Os4BGlu12 shows little similarity when compared to the other GH1 enzymes. From this observation, we can surmise that the differences in the aglycone binding subsites of GH1 enzymes lead to the varieties of substrate specificities and functions of these enzymes.

REFERENCES

- Ahn, Y.O., Mizutani, M., Saino, H., and Sakata, K. (2004). Furcatin hydrolase from *Viburnum furcatum* Blume is a novel disaccharide-specific acuminosidase in glycosyl hydrolase family 1. **J. Biol. Chem.** 279: 23405-23414.
- Akiyama, T., Kaku, H., and Shibuya, N. (1998). A cell-wall bound β -glucosidase from germinated rice: Purification and properties. **Phytochemistry.** 48: 49-54.
- Banner, D.W., Bloomer, A.C., Petsko, G.A., Phillips, D.C., Pogson, C.I., Wilson, I.A., Corran, P.H., Furth, A.J., Milman, J.D., Offord, R.E., Priddle, J.D., and Waley, S.G. (1975). Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 Å resolution using amino acid sequence data. **Nature.** 255: 609-614.
- Bergfors, T. (2003) Seed to crystals. **J. Struct. Biol.** 142: 66-76.
- Barleben, L., Panjikar, S., Ruppert, M., Koepke, J., and Stockigt, J. (2007). Molecular architecture of strictosidine glucosidase: the gateway to the biosynthesis of the monoterpenoid indole alkaloid family. **Plant Cell.** 19: 2886-2897.
- Barrett, T., Suresh, C.G., Tolley, S.P., Dodson, E.J., and Hughes, M.A. (1995). The crystal structure of a cyanogenic β -glucosidase from white clover, a family 1 glycosyl hydrolase. **Structure.** 3(9): 951-960.

- Boel, E., Brady, L., Brzozowski, A.M., Derewenda, Z., Dodson, G.G., Jensen, V.J., Peterson, S.B., Swift, H., Thim, L., and Woldike, H.F. (1990). Calcium binding in α -amylase: an X-ray diffraction study at 2.1 Å resolution of two enzymes from *Aspergillus*. **Biochemistry**. 9: 6244-6249.
- Bolam, D.N., Hughes, N., Viriden, R., Lakey, J.H., Hazlewood, G.P., Henrissat, B., Braithwaite, K.L., and Gilbert, H.J. (1996). Mannanase A from *Pseudomonas fluorescens* subsp. *cellulosa* is a glycosyl hydrolase in which E212 and E320 are the putative catalytic residues. **Biochemistry**. 35: 16195-16204.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Anal. Biochem.** 72: 248-254.
- Bravman, T., Belakhov, V., Solomon, D., Shoham, G., Henrissat, B., Baasov, T., and Shoham, Y. (2003). Identification of the catalytic residues in family 52 glycoside hydrolase, a β -xylosidase from *Geobacillus stearothermophilus* T-6. **J. Biol. Chem.** 278: 26742-26749.
- Brünger, T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., and Warren, G.L. (1998). Crystallographic and NMR system: a new software suite for macromolecular structure determination. **Acta Cryst.** D54: 905-921.
- Burmeister, W.P., Cottaz, S., Driguez, H., Palmieri, S., and Henrissat, B. (1997). The crystal structures of *Sinapis alba* myrosinase and of a covalent glycosyl-

enzyme intermediate provide insights into the substrate recognition and active-site machinery of an S-glycosidase. **Structure**. 5: 663-675.

Burmeister, W.P., Cottaz, S., Rollin, P., Vasella, A., and Henrissat, B. (2000). High resolution X-ray crystallography shows that ascorbate is a cofactor for myrosinase and substitutes for the function of the catalytic base. **J. Biol. Chem.** 275: 39385-39393.

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 glycogenomics. **Nucleic Acids Res.** 37: D233-D238.

Chuenchor, W., Pengthaisong, S., Yavaniyama, J., Opassiri, R., Svasti, J., and Ketudat Cairns J.R. (2006). Purification, crystallization and preliminary X-ray analysis of rice BGlu1 β -glucosidase with and without 2-deoxy-2-fluoro- β -D-glucoside. **Acta Cryst.** F62: 798-801.

Chuenchor, W., Pengthaisong, S., Robinson, R.C., Yuvaniyama, J., Oonanant, W., Bevan, D.R., Esen, A., Chen, C.J., Opassiri, R., Svasti, J., and Ketudat-Cairns J.R. (2008). Structural insights into rice BGlu1 β -glucosidase oligosaccharide hydrolysis and transglycosylation. **J. Mol. Biol.** 377: 1200-1215.

Chuenchor, W., Pengthaisong, S., Robinson, R.C., Yuvaniyama, J., Svasti, J., and Ketudat Cairns J.R. (2011) The structural basis of oligosaccharide binding in rice BGlu1 beta-glucosidase. **J. Struct. Biol.** 173, 169-176.

Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. **Acta Cryst.** D50: 760-763.

- Coutinho, P.M., and Henrissat, B. (1999). Carbohydrate active enzymes: an integrated database approach. In *Recent Advances in Carbohydrate Bioengineering* (Gilbert, H.J., Davies, G., Henrissat, B. and Svensson, B., eds.), pp. 3-12, The Royal Society of Chemistry, Cambridge.
- Czjzek, M., Cicek, M., Zamboni, V., Burmeister, W.P., Bevan, D.R., Henrissat, B., and Esen, A. (2000). The mechanism of substrate (aglycone) specificity in β -glucosidases is revealed by crystal structures of mutant maize β -glucosidase-DIMBOA, -DIMBOAGlc, and -dhurrin complexes. **Proc. Natl. Acad. Sci. USA.** 97: 13555-13560.
- Czjzek, M., Cicek, M., Zamboni, V., Burmeister, W.P., Bevan, D.R., Henrissat, B., and Esen, A. (2001). Crystal structure of a monocotyledon (maize ZMGlu1) β -glucosidase and a model of its complex with *p*-nitrophenyl β -D-thioglucoside. **Biochem. J.** 354: 37-46.
- Davies, G., and Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. **Structure.** 3: 853-859.
- Dixon, M. (1953) The determination of enzyme inhibitor constants. **Biochem. J.** 55: 170-171.
- Esen, A. (1993). **β -Glucosidases: Biochemistry and Molecular Biology, ACS Symposium Series 533.** ACS Symposium Series 533, American Chemical Society, Washington, D.C. pp 1-26.
- Emsley, P., and Cowtan, K. (2004). Coot-model-building tools for molecular graphics. **Acta. Cryst.** D60: 2126-2132.

- Fersht, A.R., Leatherbarrow, R.J., Wells, T.N.C. (1987) Structure-activity relationships in engineered proteins: analysis of use of binding energy by linear free energy relationships. **Biochemistry**. 19: 6030-8.
- Gill, S.C., and von Hippel, P.H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. **Anal. Biochem.** 182: 319-326.
- Gloster, T.M., Roberts, S., Ducros, V.M., Perugino, G., Rossi, M., Hoos, R., Moracci, M., Vasella, A., and Davies, G.J. (2004). Structural studies of the β -glycosidase from *Sulfolobus solfataricus* in complex with covalently and noncovalently bound inhibitors. **Biochemistry**. 43: 6101-6109.
- Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. **Biochem. J.** 280: 309-316.
- Henrissat, B., and Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. **Biochem. J.** 293: 781-788.
- Henrissat, B., and Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. **Biochem. J.** 316: 695-696.
- Henrissat, B., and Davies, G.T. (1997). Structural and sequence-based classification of glycoside hydrolases. **Curr. Op. Struc. Biol.** 7: 637-644.
- Hommalai, G., Withers, S.G., Chuenchor, W., Ketudat Cairns, J., and Svasti, J. (2007). Enzymatic synthesis of cello-oligosaccharides by rice BGlu1 β -glucosidase glycosynthase mutants. **Glycobiology**. 17: 744-753.
- Hrmova, M., and Fincher, G.B. (1998). Barley β -D-glucan exohydrolases. Substrate specificity and kinetic properties. **Carbohydr. Res.** 305: 209-221.
- Hrmova, M., Gori, R.D., Smith, B.J., Fairweather, J.K., Driguez, H., Varghese, J.N.,

- and Fincher, G.B. (2002). Structural basis for broad substrate specificity in higher plant β -D-glucan glucohydrolases. **Plant Cell**. 14: 1033-1052.
- Hua, Y., Sansenya, S., Saetang, C., Wakuta, S., and Ketudat Cairns, J.R. (2013). Structural and enzymatic characterization of Os3BGlu6 and its mutants hydrolyzing gibberellin glucosyl ester to release free gibberellin GA₄. **Arch. Biochem. Biophys.** 537(1): 39-48.
- Husebye, H., Arzt, S., Burmeister, W.P., Haertel, F.V., Brandt, A., Rossiter, J.T., and Bones, A.M. (2005). Crystal structure at 1.1 Å resolution of an insect myrosinase from *Brevicoryne brassicae* shows its close relationship to beta-glucosidases. **Insect. Biochem. Mol. Biol.** 35: 1311-1320.
- Isorna, P., Polaina, J., Latorre-García, L., Cañada, F.J., González, B., and Sanz-Aparicio, J. (2007). Crystal structures of *Paenibacillus polymyxa* β -glucosidase B complexes reveal the molecular basis of substrate specificity and give new insights into the catalytic machinery of family 1 glycosidases. **J. Mol. Biol.** 371: 1204-1218.
- Jenkins, J., Leggio, L.L., Harris, G., and Pickersgill, R. (1995). β -Glucosidase, β -galactosidase, family A cellulases, family F xylanases and two barley gluconases form a superfamily of enzymes with 8-fold β/α architecture and with two conserved glutamates near the carboxy-terminal ends of β strands four and seven. **FEBS Letters**. 362: 281-285.
- Ketudat Cairns, J.R., and Esen, A. (2010). β -glucosidases. **Cell. Mol. Life Sci.** 67: 3389-3405.
- Ketudat Cairns, J.R., Pengthaisong S., Luang, S., Sansenya, S., Tankrathok, A., and Svasti, J. (2012) Protein-carbohydrate interactions leading to hydrolysis and

- transglycosylation in plant glycoside hydrolase family 1 enzymes. **J. Appl. Glycosci.** 59: 51-62.
- Kuntothom, T., Luang, S., Harvey, A.J., Fincher, G.B., Opassiri, R., Hrmova, M., and Ketudat Cairns, J.R. (2009) Rice family GH1 glycoside hydrolases with β -D-glucosidase and β -D-mannosidase activity. **Arch. Biochem. Biophys.** 491: 85-95.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. **J. Appl. Cryst.** 26: 283-291.
- Ly, H.D., and Withers, S.G. (1999). Mutagenesis of glycosidases. **Annu. Rev. Biochem.** 68: 487-522.
- MacLeod, A.M., Tull, D., Rupitz, K., Warren, R.A.J., and Withers, S.G. (1996). Mechanistic consequences of mutation of active site carboxylates in a retaining β -1,4-glycanase from *Cellulomonas fimi*. **Biochemistry.** 35: 13165- 13172.
- Marana, S.R. (2006). Molecular basis of substrate specificity in family 1 glycoside hydrolases. **IUBMB Life.** 58: 63-73.
- McCarter, J., and Withers, S.G.X. (1994). Mechanisms of enzymatic glycoside hydrolysis. **Curr. Opin. Struct. Biol.** 4: 885-892.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. **J. Appl. Cryst.** 40: 658-674.
- Mikami, B., Sato, M., Shibata, T., Hirose, M., Aibara, S., Katsube, Y., and Marita, Y. (1992). Three dimensional structure of soybean α -amylase determined at 3.0

- Å resolution-preliminary chain tracing of the complex with α -cyclodextrin. **J. Biochem.** 112: 541-546.
- Mizutani, M., Nakanishi, H., Ema, J., Ma, S., Nogchi, E., Inohara-Ochiai, M., Fukuchi-Mizutani, M., Nakao, M., and Sakata, K. (2002). Cloning of β -primeverosidase from tea leaves, a key enzyme in tea aroma formation. **J. Plant Physiol.** 130: 2164-2176.
- Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K. and Olson, A.J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. **J. Comput Chem.** 19: 1639-1662.
- Murdushov, G.N., Vagin, A., Lebedev, A., Wilsona, K.S., and Dodson, E.J. (1999) Efficient anisotropic refinement of macromolecular structures using FFT. **Acta. Cryst.** D55: 247-255.
- Navaza, J. (1994). AMoRe: an automated package for molecular replacement. **Acta Crystallog. Sect. A.** 50: 157-163.
- Opassiri, R., Ketudat-Cairns, J.R., Akiyama, T., Wara-Aswapati, O., Svasti, J., and Esen, A. (2003). Characterization of a rice β -glucosidases highly expressed in flower and germinating shoot. **Plant Sci.** 165: 627-638.
- Opassiri, R., Hua, Y., Wara-Aswapati, O., Akiyama, T., Svasti, J., Esen, A., and Ketudat-Cairns, J.R. (2004). β -glucosidase, exo- β -glucanase and pyridoxine transglucosylase activities of rice BGlu1. **Biochem J.** 379: 125-131.
- Opassiri, R., Pomthong, B., Onkoksoong, T., Akiyama, T., Esen, A., and Ketudat Cairns, J.R. (2006). Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12 β -glucosidase, **BMC Plant Biol.** 6: 33.

- Opassiri, R., Maneesan, J., Akiyama, T., Pomthong, B., Jin, S., Kimura, A., and Ketudat Cairns, J.R. (2010) Rice Os4BGlu12 is a wound-induced β -glucosidase that hydrolyzes cell wall- β -glucan-derived oligosaccharides and glycosides. **Plant Sci.** 179: 273-280.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode, **Meth. Enzymol.** Academic Press, New York Volume 276: 307-326.
- Pengthaisong, S., Chen, C.F., Withers, S.G., Kuaprasert, B., and Ketudat Cairns, J.R. (2012) Rice BGlu1 glycosynthase and wild type transglycosylation activities distinguished by cyclophellitol inhibition. **Carbohydr Res.** 352: 51-59.
- Rempel, B.P., and Withers, S.G. (2008). Covalent inhibitors of glycosidases and their applications in biochemistry and biology. **Glycobiology.** 18: 570-586.
- Raychaudhuri, A., and Tipton, P.A. (2002). Cloning and expression of the gene for soybean hydroxyisourate hydrolase. Localization and implications for function and mechanism. **Plant Physiology.** 130: 2061-2068.
- Rhodes, G. (2000). **Crystallography made crystal clear. A guide for users of macromolecular models (2nd ed.)**. New York: Academic Press.
- Rojan, S.S., Yang, X., Collart, F., Yip, V.L., Withers, S.G., Varrot, A., Thompson, J., Davies, G.J., and Anderson, W.F. (2004). Novel catalytic mechanism of glycoside hydrolysis based on the structure of an $\text{NAD}^+/\text{Mn}^{2+}$ -dependent phospho-alpha-glucosidase from *Bacillus subtilis*. **Structure.** 12: 1619-1629.
- Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J.K.C., and Jones, T.A. (1990). Three dimensional structure of cellobiohydrolase II from *Trichoderma reesi*. **Science.** 249: 380-386.

- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). **Molecular Cloning: a Laboratory Manual, 2nd ed.** Cold Spring harbor Laboratory Press, New York.
- Sansanya, S. (2008) Protein crystallization and preliminary determination of structure of rice β -glucosidase. Master's degree Dissertation, Suranaree University of Technology, Nakhon Ratchasima, Thailand.
- Sansanya, S., Ketudat Cairns, J.R., and Opassiri, R. (2010). Expression, purification, crystallization and preliminary X-ray analysis of rice (*Oryza sativa* L.) Os4BGlu12 β -glucosidase. **Acta Cryst.** F66: 320-323.
- Sansanya, S., Maneesan, J, and Ketudat Cairns, J.R. (2012). Exchanging a single amino acid residue generates or weakens a +2 cellooligosaccharide binding subsite in rice β -glucosidases. **Carbohydr. Res.** 351: 130-133.
- Sansanya, S., Opassiri, R., Kuaprasert, B., Chen. C.J., and Ketudat Cairns, J.R. (2011). The crystal structure of rice (*Oryza sativa* L.) Os4BGlu12, an oligosaccharide and tuberonic acid glucoside-hydrolyzing β -glucosidase with significant thioglucohydrolase activity. **Arch. Biochem. Biophys.** 510: 62-72.
- Sanz-Aparicio, J., Hermoso, J.A., Martinez-Ripoll, M., Lequerica, J.L., and Polaina, J. (1998). Crystal structure of β -glucosidase A from *Bacillus polymyxa*: insights into the catalytic activity in family 1 glycosyl hydrolases. **J. Mol. Biol.** 275: 491-502.
- Seshadri, S., Akiyama, T., Opassiri, R., Kuaprasert, B., and Ketudat Cairns, J.R. (2009) Structural and enzymatic characterization of Os3BGlu6, a rice β -

- glucosidase hydrolyzing hydrophobic glycosides and (1-3)- and (1-2)-linked disaccharides. **Plant Physiol.** 151: 47-58.
- Shen, H., and Byers, L.D. (2007) Thioglycoside hydrolysis catalyzed by β -glucosidase. **Biochem. Biophys. Res. Commun.** 362: 717–720.
- Sue, M., Yamazaki, K., Yajima, S., Nomura, T., Matsukawa, T., Iwamura, H., and Miyamoto, T. (2006). Molecular and structural characterization of hexameric β -D-glucosidases in wheat and rye. **Plant Physiol.** 141: 1237-1247.
- Tanner, J.J., Lei, B., Tu, S.C., and Krause, K.L. (1996) Flavin reductase P: structure of dimeric enzyme that reduces flavin. **Biochemistry.** 42 :13531-9.
- Terwilliger, T.C., and Berendzen, J. (1999). Automated MAD and MIR structure solution. **Acta Cryst.** D55: 849-861.
- Terwilliger, T.C. (2000). Maximum likelihood density modification. **Acta Cryst.** D56: 965-972.
- Terwilliger, T.C. (2003). Automated main-chain model building by template matching and iterative fragment extension. **Acta Cryst.** D59: 38-44.
- Vagin, A., and Teplyakov. (1997). MOLREP: an automated program for molecular replacement. **J. Appl. Cryst.** D30: 1022-1025.
- Vallmitjana, M., Ferrer-Navarro, M., Planell, R., Abel, M., Ausín, C., Querol, E., Planas, A., and Pérez-Pons, J.A. (2001). Mechanism of the family 1 β -glucosidase from *Streptomyces* sp: Catalytic residues and kinetic studies. **Biochemistry.** 40: 5975-598.
- Verdoucq, L., Czjzek, M., Moriniere, J., Bevan, D.R., and Esen, A. (2003). Mutational and structural analysis of aglycone specificity in maize and sorghum β -glucosidases. **J. Biol. Chem.** 278: 25055- 25062.

- Verdoucq, L., Morinière, J., Bevan, D.R., Esen, A., Vasella, A., Henrissat, B., and Czjzek, M. (2004). Structural determinants of substrate specificity in family 1 β -glucosidases: Novel insights from the crystal structure of sorghum dhurrinase-1, a plant β -glucosidase with strict specificity, in complex with its natural substrate. **J. Biol. Chem.** 279: 31796-31803.
- Wakuta, S., Hamada, S., Ito, H., Matsuura, H., Nabeta, K., and Matsui, H. (2010) Identification of a β -glucosidase hydrolyzing tuberonic acid glucoside in rice (*Oryza sativa* L.). **Phytochemistry.** 71: 1280-1288.
- Wakuta, S., Hamada, S., Ito, H., Imai, R., Mori, H., Matsuura, H., Nabeta, K., and Matsui, H. (2011) Comparison of enzymatic properties and gene expression profiles of two tuberonic acid glucoside β -glucosidases from *Oryza sativa* L. **J. Appl. Glycosci.** 58: 68-70.
- Wang, M., Li, D., Sun, X., Zhu, Y.J., Nong, H., and Zhang. (2009) Characterization of a root-specific β -thioglucoside glucohydrolase gene in *Carica papaya* and its recombinant protein expressed in *Pichia pastoris*. **Plant Sci.** 177: 716-723.
- Wang, Q., Graham, R.W., Trimbur, D., Warren, R.A.J., and Withers, S.G. (1994). Changing enzymatic reaction mechanisms by mutagenesis: Conversion of a retaining glucosidase to an inverting enzyme. **J. Am. Chem. Soc.** 116: 11594- 11595.
- Wang, Q., Trimber, D., Grahnan, R., Warren, R.A.J., and Withers, S.G. (1995). Identification of the acid/base catalyst in *Agrobacterium faecalis* β -glucosidase by kinetic analysis of mutants. **Biochemistry.** 34: 14554-14562.

- Wang, J., Pengthaisong, S., Ketudat Cairns, J.R., and Liu, Y. (2013) X-ray crystallography and QM/MM investigation on the oligosaccharide synthesis mechanism of rice BGlu1 glycosynthases. **Biochim. Biophys. Acta.** 1834: 536-545.
- Wolfenden, R., Lu, X., and Young, G. (1998). Spontaneous hydrolysis of glycosides. **J. Am. Chem. Soc.** 120: 6914-6815.
- Xia, L., Ruppert, M., Wang, M., Panjekar, S., Lin, H., Rajendran, C., Barleben, L. and Stöckigt, L. (2012) Structures of alkaloid biosynthetic glucosidases decode substrate specificity. **ACS Chem. Biol.** 7: 226-234.
- Zechel, D.L., and Withers, S.G. (2000). Glycosidase mechanisms: anatomy of a finely tuned catalyst. **Acc. Chem. Res.** 33: 11-18.
- Zechel, D.L., Boraston, A.B., Gloster, T., Boraston, C.M., Macdonald, J.M., Tilbrook, D.M., Stick, R.V., and Davies, G.J. (2003). Iminosugar glycosidase inhibitors: structural and thermodynamic dissection of the binding of isofagomine and 1-deoxynojirimycin to β -glucosidases. **J. Am. Chem. Soc.** 125: 14313-14323.

CURRICULUM VITAE

NAME Mr. Sompong Sansenya

DATE OF BIRTH 21 July 1982

PLACE OF BIRTH Roi-Et, Thailand

EDUCATION Suranaree University of Technology,
Nakhon Ratchasima,
Thailand, 2006-2009,
Master of Science, Biochemistry.
Maejo University, Chiang Mai,
Thailand, 2001-2005,
Bachelor of Science, Chemistry.

AWARDS AND SCHOLARSHIPS

SUT-PhD. Fund grant SUT-PhD/06/2554
(National Research University Project of the
Committee on Higher Education to Suranaree
University of Technology).