

**FUNCTIONAL AND STRUCTURAL STUDIES OF RICE  
 $\beta$ -GLUCOSIDASE**

**Supriya Seshadri**

**A Thesis Submitted in Partial Fulfillment of the Requirements for the  
Degree of Doctor of Philosophy in Biochemistry  
Suranaree University of Technology  
Academic Year 2008**

การศึกษาหน้าที่และโครงสร้างของเอนไซม์เบตาไกลูโคซิเดสจากข้าว

นางสาวสุปรียา เขชาตรี

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาชีวเคมี

มหาวิทยาลัยเทคโนโลยีสุรนารี

ปีการศึกษา 2551

# FUNCTIONAL AND STRUCTURAL STUDIES OF RICE

## $\beta$ -GLUCOSIDASE

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. Malee Tangsathitkulchai)

Chairperson

---

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

Member (Thesis Advisor)

---

(Assoc. Prof. Dr. Palangpon Kongsaree)

Member

---

(Assoc. Prof. Dr. Wipa Suginta)

Member

---

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

Member

---

(Asst. Prof. Dr. Rodjana Opassiri)

Member

---

(Prof. Dr. Pairote Sattayatham)

Vice Rector for Academic Affairs

---

(Assoc. Prof. Dr. Prapan Manyum)

Dean of Institute of Science

สุปรียา เชชาตรี : การศึกษาหน้าที่และโครงสร้างของเอนไซม์เบต้ากลูโคซิเดสจากข้าว (FUNCTIONAL AND STRUCTURAL STUDIES OF RICE  $\beta$ -GLUCOSIDASE) อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร. เจมส์ เกตุทัต คาร์นส์, 144 หน้า.

เอนไซม์เบต้ากลูโคซิเดสในตระกูล Glycoside Hydrolase Family 1 (GH1) มีบทบาทสำคัญในกระบวนการต่างๆ ของพืช เช่น การสร้างสารต่อต้านแมลงศัตรูพืช การย่อยโอลิโกแซคคาไรด์ (oligosaccharide) ซึ่งเป็นส่วนประกอบของผนังเซลล์พืช การควบคุมไฟโตฮอร์โมน (phytohormone) และการสร้างลิกนิน (lignification) เอนไซม์ Os3BGlu6 เป็นหนึ่งในเอนไซม์ 34 ชนิดของเอนไซม์เบต้ากลูโคซิเดสใน GH1 ที่ตรวจพบในจีโนมของข้าว (*Oryza sativa* L.) และเป็นหนึ่งในแปดของ phylogenetic clusters ในจีนข้าว และ *Arabidopsis* เอนไซม์ Os3BGlu6 ถูกนำมาศึกษาคุณสมบัติทางชีวเคมีและวิเคราะห์หาโครงสร้างสามมิติ เอนไซม์ Os3BGlu6 ถูกผลิตใน *E. coli* Origami (DE3) ในรูปของโปรตีนที่มี thioredoxin และ His<sub>6</sub> tags อยู่ที่ปลายอะมิโนของโปรตีน เอนไซม์ Os3BGlu6 สามารถย่อยสับสเตรท *p*-nitrophenyl (*p*NP)  $\beta$ -D-fucoside ( $k_{cat}/K_m = 66.8 \text{ mM}^{-1}\text{s}^{-1}$ ) *p*NP- $\beta$ -D-glucoside ( $k_{cat}/K_m = 6.2 \text{ mM}^{-1}\text{s}^{-1}$ ) และ *p*NP- $\beta$ -D-galactoside ( $k_{cat}/K_m = 1.6 \text{ mM}^{-1}\text{s}^{-1}$ ) ได้ดี และสามารถย่อย alkyl glycosides *n*-octyl- $\beta$ -D-glucoside ( $k_{cat}/K_m = 2.7 \text{ mM}^{-1}\text{s}^{-1}$ ) ได้ดีกว่า *n*-heptyl  $\beta$ -D-glucoside ( $k_{cat}/K_m = 0.85 \text{ mM}^{-1}\text{s}^{-1}$ ) นอกจากนี้ยังสามารถย่อยไดแซคคาไรด์ที่มี  $\beta$ -(1 $\rightarrow$ 3)-linked ( $k_{cat}/K_m = 1.7 \text{ mM}^{-1}\text{s}^{-1}$ ) และ  $\beta$ -(1 $\rightarrow$ 2)-linked ( $k_{cat}/K_m = 0.96 \text{ mM}^{-1}\text{s}^{-1}$ ) ได้ดีกว่าไดแซคคาไรด์ชนิดอื่นๆ

ได้ศึกษาโครงสร้างสามมิติของเอนไซม์ Os3BGlu6 อีสระ เอนไซม์ที่รวมอยู่กับ 2-deoxy-2-fluoroglucoside และ *n*-octyl- $\beta$ -D-thioglucopyranoside ซึ่งเป็น non-hydrolyzable analogue ของ *n*-octyl- $\beta$ -D-glucoside โดยผลึกทั้ง 3 ชนิดสามารถหักหรั่งสี่เอกซเรย์ได้ความละเอียดถึง 1.83 1.81 และ 1.80 อังสตรอม ตามลำดับ พบว่าบริเวณเร่งของเอนไซม์ มีความยาวจากส่วนที่ลึกที่สุดของบริเวณเร่งถึงบริเวณทางเข้าของตำแหน่งเร่งปฏิกิริยาเท่ากับ 14 อังสตรอม ซึ่งสั้นกว่าความยาวของบริเวณเร่งปฏิกิริยาของเอนไซม์ Os3BGlu7 4 อังสตรอม ส่วนที่แคบที่สุดของบริเวณเร่งปฏิกิริยามีความกว้าง 5.3 อังสตรอม และมีส่วนที่กว้างที่สุดเท่ากับ 7.6 อังสตรอม พบ Tris 1 โมเลกุลซึ่งมีโครงสร้างที่แตกต่างกัน 2 แบบในโครงสร้างสามมิติของเอนไซม์อีสระ ซึ่งสอดคล้องกับการทำงานของโมเลกุลนี้ที่สามารถเป็นตัวยับยั้งแบบอ่อนของเอนไซม์ Os3BGlu6 โครงสร้างสามมิติของเอนไซม์กับ 2-deoxy-2-fluoroglucosyl พบว่า 2-deoxy-2-fluoroglucosyl อยู่ในโครงสร้างฟอนคลายของแก้อี้แบบ <sup>4</sup>C<sub>1</sub> ซึ่งเกิดพันธะโควาเลนต์กับกรดอะมิโนที่ทำหน้าที่เป็น nucleophile บริเวณตำแหน่งที่จับกับสับสเตรทด้านใน (glycone-binding subsite) คล้ายกับการเกิดพันธะโควาเลนต์เชิงซ้อนในโครงสร้างสามมิติของเอนไซม์ Os3BGlu7 จากข้าวกับ 2-deoxy-2-fluoroglucosyl ส่วนโครงสร้าง

สามมิติของเอนไซม์ Os3BGlu6 กับ *n*-octyl- $\beta$ -D-thioglucopyranoside พบว่ากรดอะมิโน M251 อยู่บริเวณทางเข้าของสัปสเตรทเป็นอุปสรรคในการจับโอลิโกแซคคาไรด์ที่มีพันธะ  $\beta$ -(1 $\rightarrow$ 4) สายยาวแต่กลับทำอันตรกิริยากับส่วนที่ไม่ชอบน้ำของ *n*-octyl- $\beta$ -D-thioglucopyranoside ได้ซึ่งสามารถบ่งบอกถึงความชอบของเอนไซม์ Os3BGlu6 กับโอลิโกแซคคาไรด์สายสั้นและไกลโคไซด์ที่มีส่วนไม่ชอบน้ำ ดังนั้นความแตกต่างของกรดอะมิโนในบริเวณเร่งปฏิกิริยาน่าจะมีส่วนทำให้เอนไซม์มีความจำเพาะต่อการย่อยสัปสเตรทแตกต่างกัน ซึ่งน่าจะช่วยอธิบายการทำงานของเอนไซม์เบต้ากลูโคซิเดสจากข้าวในเอนไซม์ตระกูล GH1 เช่น Os3BGlu6

SUPRIYA SESHADRI : FUNCTIONAL AND STRUCTURAL STUDIES  
OF RICE  $\beta$ -GLUCOSIDASE. THESIS ADVISOR : ASSOC. PROF. JAMES  
R. KETUDAT-CAIRNS, Ph.D. 144 PP.

GLYCOSYL HYDROLASE/RECOMBINANT EXPRESSION/ $\beta$ -GLUCOSIDASE/  
 $\beta$ -FUCOSIDASE/*n*-OCTYL- $\beta$ -D-GLUCOSIDE

In plants, glycoside hydrolase family 1 (GH1)  $\beta$ -glucosidases are believed to play important roles in many processes, such as chemical defense against herbivores, hydrolysis of cell wall-derived oligosaccharides, phytohormone regulation, and lignification. Os3BGlu6 is one of the 34 GH1  $\beta$ -glucosidases predicted from the rice (*Oryza sativa* L.) genome, and it represents one of the eight phylogenetic clusters including both rice and *Arabidopsis* genes. In this study, Os3BGlu6 was characterized for its biochemical properties and its 3D structure was crystallographically determined. Os3BGlu6 was expressed in *E. coli* strain Origami (DE3) as a recombinant fusion protein with N-terminal thioredoxin and His<sub>6</sub> tags. The purified Os3BGlu6 hydrolyzed *p*-nitrophenyl (*p*NP)  $\beta$ -D-fucoside ( $k_{\text{cat}}/K_{\text{m}}=66.8 \text{ mM}^{-1}\text{s}^{-1}$ ), *p*NP- $\beta$ -D-glucoside ( $k_{\text{cat}}/K_{\text{m}} = 6.2 \text{ mM}^{-1}\text{s}^{-1}$ ), and *p*NP- $\beta$ -D-galactoside ( $k_{\text{cat}}/K_{\text{m}}=1.6 \text{ mM}^{-1}\text{s}^{-1}$ ) efficiently. It also hydrolyzed the alkyl glycosides *n*-octyl- $\beta$ -D-glucoside with greater catalytic efficiency ( $k_{\text{cat}}/K_{\text{m}}=2.7 \text{ mM}^{-1}\text{s}^{-1}$ ) than *n*-heptyl  $\beta$ -D-glucoside ( $k_{\text{cat}}/K_{\text{m}}=0.85 \text{ mM}^{-1}\text{s}^{-1}$ ). Among the oligosaccharides,  $\beta$ -(1 $\rightarrow$ 3)-linked ( $k_{\text{cat}}/K_{\text{m}}=1.7 \text{ mM}^{-1}\text{s}^{-1}$ ) and  $\beta$ -(1 $\rightarrow$ 2)-linked ( $k_{\text{cat}}/K_{\text{m}}=0.96 \text{ mM}^{-1}\text{s}^{-1}$ ) disaccharides were hydrolyzed most efficiently.

The three-dimensional structures of native Os3BGlu6, its covalent intermediate with 2-deoxy-2-fluoroglucoside, and its complex with the non-

hydrolyzable analogue of *n*-octyl- $\beta$ -D-glucoside, *n*-octyl- $\beta$ -D-thioglucopyranoside, were determined at 1.83 Å, 1.81 Å and 1.80 Å resolution, respectively. The distance from the deepest part of the active site slot to the surface entrance is 14 Å, shorter by 4 Å from that of Os3BGlu7 (18 Å). The narrowest part of aglycone binding site, was about 5.3 Å across, while the broadest region is 7.6 Å. One molecule of Tris in two possible conformations could be seen in the active site of the native structure, consistent with its action as a weak inhibitor for Os3BGlu6. The density for the 2-deoxy-2-fluoroglucosyl residue in a relaxed  ${}^4C_1$  chair conformation covalently bound to the catalytic nucleophile was evident in the glycone-binding subsite and its position was similar to that in a 2-deoxy-2-fluoroglucosyl covalent complex of rice Os3BGlu7. The residue Met251, located at the entrance to the active site of Os3BGlu6, obstructed the binding of longer  $\beta$ -(1 $\rightarrow$ 4)-linked oligosaccharides, but interacted with the hydrophobic aglycone of *n*-octyl- $\beta$ -D-thioglucopyranoside, contributing to the preference of Os3BGlu6 for shorter oligosaccharides and hydrophobic glycosides. This observation suggests that small differences contributed by residues in the active site are likely to account for differences in substrate specificities, which in turn determine the functions of GH1  $\beta$ -glucosidases in rice, such as Os3BGlu6.

School of Biochemistry

Academic Year 2008

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

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

## **ACKNOWLEDGEMENTS**

I wish to express my heartfelt gratitude to my thesis advisor, Assoc. Prof. Dr. James R. Ketudat-Cairns for giving me an opportunity to work on the rice  $\beta$ -glucosidase research project, and in making this project a success. I am grateful for his timely and invaluable guidance, kind supervision, unstinted support and encouragement. I am also genuinely grateful to Asst. Prof. Dr. Rodjana Opassiri for her help and guidance throughout my project. I wish to express my sincere thanks Dr. Takashi Akiyama for his help with the cDNA library construction and for providing me with the cDNA clone for this project. I am grateful to Asst. Prof. Dr. Mariena Ketudat-Cairns for teaching me molecular biology, and constantly inspiring me and encouraging me throughout my project work. I am also grateful to Assoc. Prof. Wipa Suginta for teaching me one of the most challenging subjects, Enzymology.

I would like to specially thank Dr. Buabarn Kuaprasert at the Synchrotron Light Research Institute (SLRI), Thailand for providing me an opportunity to access the resources at SLRI, and for her valuable guidance, advice on x-ray crystallography and helping me on the structural analysis. I would like to extend my heartfelt thanks to Assoc. Prof. Dr. Chun-Jung Chen, Dr. Phimonphan Chuankhayan and the staff at National Synchrotron Radiation Research Center (NSRRC), Taiwan for their kind help and in assisting me crystal screening, and data collection.

I would like to thank the National Center for Genetic Engineering and Biotechnology and the National Synchrotron Research Center for providing the funding for the research work.

I would like to thank Assoc. Prof. Dr. Palangpon Kongsaree and Assoc. Prof. Dr. Malee Tangsathitkulchai for their valuable suggestions and advice.

Special thanks to the staff at the Center for International Affairs at the Suranaree University of Technology for their help in making arrangements for me to be here in Thailand for my research work and helping me with the visa process.

I wish to express my special thanks to all my friends in the School of Biochemistry, Suranaree University of Technology and for their constant support and helping me throughout the project and in making this journey of mine so easy.

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

Supriya Seshadri

# CONTENTS

	<b>Page</b>
ABSTRACT IN THAI.....	I
ABSTRACT IN ENGLISH.....	III
ACKNOWLEDGEMENTS.....	V
CONTENTS .....	VII
LIST OF TABLES.....	XII
LIST OF FIGURES .....	XIV
LIST OF ABBREVIATIONS .....	XVIII
<b>CHAPTER</b>	
<b>I INTRODUCTION .....</b>	<b>1</b>
1.1 Overview-Glycosyl Hydrolases.....	1
1.2 Glycosyl Hydrolase Classifications.....	2
1.3 Glycosyl hydrolases families and folds.....	3
1.4 Morphology of active sites in Glycosyl Hydrolases.....	5
1.5 Catalytic Mechanism and Catalytic Machinery.....	6
1.6 Overview of GH Family 1 (GH1) $\beta$ -glucosidases.....	10
1.6.1 GH1 .....	11
1.6.2 $\beta$ -Glucosidase functions .....	11
1.6.3 Substrate specificity and 3D structures .....	15
1.7 Rice GH Family 1 $\beta$ -glucosidases .....	21
1.8 Research Objectives .....	25

## CONTENTS (Continued)

	Page
<b>II MATERIALS AND METHODS</b> .....	26
2.1 Materials.....	26
2.1.1 Plasmids and bacterial strains.....	26
2.1.2 Reagents and Chemicals.....	26
2.2 General Methods .....	28
2.2.1 Preparation of <i>E. coli</i> strains DH5 $\alpha$ and Origami(DE3) for competent cells .....	28
2.2.2 Isolation of recombinant plasmid by alkaline lysis .....	29
2.2.3 Agarose gel electrophoresis.....	31
2.2.4 Purification of TEV protease.....	31
2.2.5 Os3BGlu6 activity assay .....	31
2.2.6 Bio-RAD Protein Assay .....	31
2.2.7 Determination of protein expression profiles by SDS-PAGE .....	33
2.3 Cloning and Expression of Os3BGlu6 .....	34
2.3.1 Cloning and PCR amplification of cDNA encoding mature Os3BGlu6.....	34
2.3.2 Expression of Os3BGlu6 in <i>E. coli</i> Origami(DE3).....	35
2.3.3 Os3BGlu6 recombinant protein extraction and purification .....	35
2.4 Os3BGlu6 pH and temperature optimum & stability studies.....	37
2.5 Os3BGlu6 substrate specificity and enzyme kinetics .....	38
2.6 Tris Inhibition.....	39

## CONTENTS (Continued)

	<b>Page</b>
2.7 Os3BGlu6 Crystallization .....	40
2.7.1 Preparation of Os3BGlu6 protein for crystallization .....	40
2.7.2 Preliminary screening by micro-batch technique .....	40
2.7.3 Optimization of crystals by hanging drop .....	41
2.7.4 Microseeding by the streak seeding method .....	45
2.7.5 Soaking of crystals with Inhibitors.....	45
2.7.6 Cryofreezing of crystals .....	45
2.7.7 Data collection, processing and structure refinement.....	46
2.7.8 Model building and structure validation .....	47
2.7.9 N terminal amino acids sequencing.....	49
<b>III RESULTS</b> .....	<b>50</b>
3.1 Cloning and expression of Os3BGlu6.....	50
3.2 Os3BGlu6 protein purification.....	55
3.3 Effect of pH and temperature on the activity and stability of Os3BGlu6.....	56
3.4 Substrate specificity and kinetic analysis of Os3BGlu6.....	59
3.5 Os3BGlu6 crystallization .....	64
3.5.1 Initial screening by the microbatch method .....	64
3.5.2 Optimization by hanging drop.....	65
3.5.3 Optimization by miroseeding .....	73
3.5.4 Soaking of Os3BGlu6 crystals with inhibitors.....	73
3.5.5 Diffraction, Data Collection, processing and refinement.....	74

## CONTENTS (Continued)

	<b>Page</b>
3.6 3D structures of Os3BGlu6 .....	78
3.6.1 Validation of structures .....	78
3.6.2 Overall structures of Os3BGlu6 and its complexes with G2F and <i>n</i> -octyl- $\beta$ -D-thioglucopyranoside .....	79
3.7 Tris inhibition by Os3BGlu6 .....	86
3.8 Active site analysis .....	88
3.8.1 Hydrogen bonding interactions of Tris .....	88
3.8.2 Os3BGlu6/G2F covalent intermediate .....	92
3.8.3 Os3BGlu6/ <i>n</i> -octyl- $\beta$ -D-thioglucopyranoside .....	95
<b>IV DISCUSSION</b> .....	<b>100</b>
4.1 Cloning, expression, and purification of recombinant rice Os3BGlu6 $\beta$ -glucosidase.....	100
4.2 Biochemical properties of Os3BGlu6 .....	103
4.3 Substrate specificity of Os3BGlu6 .....	105
4.4 Os3BGlu6 crystallization .....	108
4.5 Overall structure .....	109
4.5.1 Unit cell differences .....	109
4.6 Crystal structures of Os3BGlu6/G2F intermediate and Os3BGlu6/ <i>n</i> -octyl- $\beta$ -D-thioglucopyranoside .....	115
4.7 Comparison of substrate binding and recognition between Os3BGlu6 and other GH family 1 enzymes.....	116

**CONTENTS (Continued)**

	<b>Page</b>
4.7.1 Os3BGlu6 and Os3BGlu7 .....	113
4.7.2 Os3BGlu6 and Maize ZmGlu1 .....	116
4.7.3 Os3BGlu6 and Sorghum (SbDhr1) .....	118
4.7.4 Os3BGlu6 and Strictosidine- $\beta$ -glucosidase.....	119
<b>V CONCLUSION</b> .....	<b>126</b>
<b>REFERENCES</b> .....	<b>131</b>
<b>CURRICULUM VITAE</b> .....	<b>147</b>

## LIST OF TABLES

Table	Page
2.1	Oligonucleotide primers used for sequencing of recombinant pET32a(+)/DEST/Os3BGlu6 ..... 34
2.2	Screening kits for Microbatch method ..... 41
2.3	Optimization of Os3BGlu6 crystals in the condition 17 to 22% PEG 5000 MME and 0.1 M Bis Tris, pH 6.5 ..... 43
2.4	Optimization of Os3BGlu6 crystals with different concentrations of protein ..... 43
2.5	Optimization of Os3BGlu6 crystals with different types of PEG ..... 47
2.6	Optimization of Os3BGlu6 crystallization with different pH and PEG 5000 MME ..... 48
3.1	Kinetic parameters of Os3BGlu6 for the hydrolysis of pNP-glycosides ..... 60
3.2	Hydrolysis of glycosides and oligosaccharides by Os3BGlu6..... 63
3.3	Summary of positive conditions crystals of native Os3BGlu6 obtained from the commercially available screening kits ..... 64
3.4	Results of coarse screening with different concentrations of PEG 5000 MME obtained from condition D10 ..... 66
3.5	Results of coarse screening with different concentrations of PEG 5000 MME and 0.1 M Bis-Tris, pH 6.5, with five different concentrations of protein (1 to 10 mg/ml)..... 69

**LIST OF TABLES (Continued)**

<b>Table</b>	<b>Page</b>
3.6 Results of refined screening with different concentrations of PEG 5000 MME and 0.1 M Bis-Tris in pH range 5.5 to 6.9 .....	71
3.7 Data collection, processing and refinement parameters .....	77
3.8 Hydrogen bonds occurring between the Tris molecule, binding residues and waters in the enzyme's active site .....	91
3.9 Hydrogen bonds occurring between <i>n</i> -octyl-thioglucopyranoside and amino acid residues in the enzyme's active site .....	99
4.1 Unit cell dimensions for Os3BGlu6 crystals .....	114

## LIST OF FIGURES

Figure	Page
1.1 The overall topology of the white clover cyanogenic $\beta$ -glucosidase (PDB code: 1CBG) .....	4
1.2 The three different types of active sites found in glycosyl hydrolases .....	6
1.3 General mechanisms for glycosyl hydrolases .....	8
1.4 Structure of maize ZmGlu1 $\beta$ -glucosidase and its inactive Glu1E191Asp bound to DIMBOA-glc .....	18
1.5 Superimposition of the crystal structures of TaGlu1b (PDB code: 2DGA) and ZmGlu1-Glu191Asp (PDB code: 1E56) and the modeled structure of ScGlu. ....	19
1.6 Hydrogen bonding network between the glucosidic part of strictosidine and residues in the ligand structure of strictosidine $\beta$ -glucosidase (SG) inactive mutant Glu207Gln.....	20
1.7 Simplified phylogenetic tree of the amino acid sequences of eukaryotic GH1 proteins with known structures and those of rice and Arabidopsis GH1 gene products .....	23
3.1 The full-length cDNA sequence and deduced protein sequence of rice Os3BGlu6 with SapporoMstrtF and SapporoStopR primers .....	51
3.2 1% agarose gel electrophoresis of PCR product of Os3BGlu6 cDNA after PCR amplification.....	53

## LIST OF FIGURES (Continued)

Figure	Page
3.3 Expression cartridge of the pET32a(+)/Os3BGlu6 plasmid.....	53
3.4 Comparison of the $\beta$ -glucosidase activity of Os3BGlu6 in <i>E. coli</i> extracts after induction at 20°C at different concentrations of IPTG.....	54
3.5 The effect of temperature on the induction of Os3BGlu6.....	55
3.6 SDS PAGE analysis of Os3BGlu6 purification.....	56
3.7 pH dependence of Os3BGlu6 activity.....	57
3.8 pH stability of Os3BGlu6 activity.....	58
3.9 Temperature dependence of Os3BGlu6 activity.....	58
3.10 Thermostability of Os3BGlu6.....	59
3.11 TLC analyses of natural substrates hydrolyzed by Os3BGlu6.....	61
3.12 Os3BGlu6 crystals obtained from microbatch screening after 180 days.....	65
3.13 Clusters of Os3BGlu6 crystals obtained from optimization by varying the concentration of PEG 5000 MME in 0.1 M Bis-Tris, pH 6.5.....	67
3.14 Protein from Os3BGlu6 crystals run on SDS PAGE.....	68
3.15 Os3BGlu6 crystals obtained in different concentrations of PEG 5000 MME and protein in 0.1 M Bis-Tris, pH 6.5.....	70
3.16 Os3BGlu6 crystals obtained in different concentrations of PEG 5000 MME and 0.1 M Bis-Tris, pH 6.4, 6.5, and 6.7.....	72
3.17 A crystal of native Os3BGlu6 of dimensions 210 x 20 x 15 $\mu$ m obtained by microseeding in the condition 21% PEG 5000 MME, 0.1 M Bis-Tris, pH 6.5.....	73

## LIST OF FIGURES (Continued)

Figure	Page
3.18	Os3BGlu6 crystals soaked with inhibitors in 19% 5000 MME, 0.1 M Bis Tris, pH 6.5 and 8 mg/ml protein ..... 74
3.19	Diffraction images of Os3BGlu6 crystals ..... 75
3.20	Ramachandran plots obtained from PROCHECK validation of the refined Os3BGlu6 structure..... 79
3.21	Protein from Os3BGlu6 crystals electroblotted on PVDF (polyvinylidene fluoride) membrane for N-terminal sequencing..... 81
3.22	Stereo cartoon diagrams of the Os3BGlu6 structures..... 82
3.23	Superimposition of the overall structures of native Os3BGlu6, and the Os3BGlu6/G2F and Os3BGlu6/ <i>n</i> -octyl- $\beta$ -D- thioglucopyranoside complexes ..... 84
3.24	Superimposition of Loops A, B, C, and D from six known GH1 structures..... 86
3.25	Graphs indicating enzyme inhibition for Tris..... 87
3.26	The Fo-Fc omit map of Tris molecule in two conformations calculated from the final model of native structure contoured at 3.0 sigma..... 88
3.27	Active site structure of Os3BGlu6 with a Tris molecule with two different binding modes of Tris, A and B..... 90
3.28	The Fo-Fc omit map of 2-fluoroglucoside covalently linked to Glu394 contoured at 3 sigma..... 92
3.29	Binding of 2-fluoroglucoside in the Os3BGlu6 active site ..... 94

## LIST OF FIGURES (Continued)

Figure	Page
3.30	Fo-Fc omit map and 2Fo-Fc map of ligand <i>n</i> -octyl- $\beta$ -d-thioglucopyranoside contoured at 3 sigma and 1 sigma..... 96
3.31	Os3BGlu6 interactions with <i>n</i> -octyl- $\beta$ -D-thioglucopyranoside ..... 98
4.1	Chemical structures of some sugars, used in determining the substrate specificity of Os3BGlu6 in their glycosidic forms..... 105
4.2	Natural substrates hydrolyzed by plant GH1 enzymes..... 112
4.3	Comparison of Os3BGlu6 and Os3BGlu7 ligand interactions..... 115
4.4	Comparison of the aglycone binding pockets of Os3BGlu6/ <i>n</i> -octyl- $\beta$ -D-thioglucopyranoside and ZmGlu1/DIMBOA-Glc (PDB code: 1E56) in stereoview. .... 117
4.5	Comparison of the aglycone binding pockets of the Os3BGlu6/ <i>n</i> -octyl- $\beta$ -D-thioglucopyranoside complex and a SbDhr1 mutant bound to dhurrin (PDB code: 1V03)..... 119
4.6	Comparison of the aglycone binding pockets of Os3BGlu6/ <i>n</i> -octyl- $\beta$ -D-thioglucopyranoside complex and strictosidine- $\beta$ -glucosidase (SG) (PDB code: 2JF6) in stereoview ..... 121

## LIST OF ABBREVIATIONS

Abs	Absorbance
ABTS	2,2'-azinobis-3-ethylbenthaiazolinesulfonic acid
bp	Base pairs
BSA	Bovine serum albumin
°C	Degrees Celsius
cDNA	Complementary deoxynucleic acid
CV	Column Volume
DNA	deoxynucleic acid
dNTPs	Deoxynucleoside triphosphates
DNase	Deoxyribonuclease
DMSO	dimethyl sufoxide
DP	Degree of Polymerization
EDTA	Ethylenediamine tetraacetate
<i>g</i>	gravitational acceleration (relative centrifugal force)
GH	Glycosyl Hydrolase
GH1	Glycosyl Hydrolase Family 1
G2F	2-deoxy-2-fluoro- $\beta$ -D-glucopyranoside
HPLC	High Performance Liquid Chromatography
hr	Hour(s)
IMAC	Immobilized Metal Affinity Chromatography
kb	Kilo base pair(s)

**LIST OF ABBREVIATIONS (Continued)**

kDa	Kilo Dalton(s)
LB	Luria Petroni broth
min	Minute(s)
(m, $\mu$ )g	(milli, micro) Gram(s)
(m, $\mu$ ) l	(milli, micro) Liter(s)
(m, $\mu$ )M	(milli, micro) Molar(s)
MME	Monomethyl ether
( $\mu$ )mol	(micro) Mole(s)
MW	Molecular Weight
nm	nanometer(s)
IPTG	Isopropyl- $\beta$ -D-thiogalactopyranoside
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PEG	Polyethylene glycol
PIPES	Piperazine-1,4-bis-2-ethanesulfonic acid
PCR	Polymerase chain reaction
PMSF	Phenylmethylsulfonylfluoride
<i>p</i> NP	<i>p</i> -Nitrophenolate
<i>p</i> NPGlc	<i>p</i> -Nitrophenyl- $\beta$ -D-glucopyranoside
PVDF	Polyvinylidene Flouride
RNA	Ribonucleic acid
RNase	Ribonuclease

**LIST OF ABBREVIATIONS (Continued)**

rpm	rotations per minute
s	second(s)
SDS	Sodium dodecyl sulfate
SgDhr1	Sorghum $\beta$ -glucosidase (Dhurrinase)
TEMED	Tetramethylenediamine
TEV	Tobacco etch virus
TLC	Thin Layer Chromatography
Tris	Tris-(hydroxymethyl)-aminoethane
v/v	Volume/volume
w/v	Weight/volume
ZmGlu1	Maize $\beta$ -glucosidase

# CHAPTER I

## INTRODUCTION

### 1.1 Overview-Glycosyl Hydrolases

Carbohydrates are one of the most diverse groups of biological molecules, and exist mainly in the form of monosaccharides (e.g. glucose, ribose, galactose, etc.), disaccharides (e.g. sucrose, lactose, maltose, etc.), oligosaccharides, polysaccharides (e.g. cellulose, starch, glycogen, etc.), and glycoconjugates (glycosides, glycoproteins, and glycolipids). They are one of the most abundant classes of biomolecules existing in a number of stereochemical variations. As part of the biological function of many carbohydrates, selective hydrolysis of the glycosidic bond is important for processes like energy uptake, cell wall expansion and degradation and turn-over of signaling molecules (Davies and Henrissat, 1995), defense against pathogens, hydrolysis of structural polysaccharides during penetration of pathogens into cells, turn-over of cell surface carbohydrates, etc. (Henrissat *et al.*, 1995).

The glycosidic bond is one of the most stable kinds of bonds, with a half life of 5 million years (Wolfenden *et al.*, 1998) and the key enzymes that catalyze the hydrolysis of glycosidic bonds between two or more carbohydrates or between a carbohydrate and a noncarbohydrate residue are classified as glycosyl hydrolases, also referred to as glycoside hydrolases or glycosidases.

## 1.2 Glycosyl Hydrolase Classifications

The Enzyme Commission (EC) classification nomenclature assigns enzyme numbers based primarily on the reactions they catalyze and their substrate specificity. However, glycosyl hydrolases are nearly impossible to segregate by the above mentioned system, because of their multiple substrate specificities. For example, endoglucanases, which are known to be cellulases, are also active toward xylan, xyloglucan, beta-glucan and artificial substrates (Henrissat and Davies, 1997). Another example of how this system fails to show enzyme relationships is that of myrosinase (EC 3.2.3.1), which hydrolyzes S-glucosides but has a sequence, molecular mechanism and 3D structure similar to that of beta-glucosidases (EC 3.2.1.21) (Burmeister *et al.*, 1997). This substrate based classification fails to reflect evolutionary events and the structural features of glycosyl hydrolases. Hence, a new method to classify glycosyl hydrolases based on sequence similarities was developed by Henrissat (1991).

Based on the fact that a direct relationship exists between sequence and folding similarities and that useful structural and mechanistic information could be deduced from the amino acid sequence alone (Chothia and Lesk, 1986; Henrissat *et al.*, 1989; Henrissat, 1990; Svensson, 1988; MacGregor and Svensson, 1989; Raimbaud *et al.*, 1989). Henrissat (1991) started systematic comparison of the primary sequences of glycosyl hydrolases. In this system of classification, enzymes with different substrate specificities (polyspecific) are sometimes seen in the same family indicating evolutionary divergence [e.g. family 16 containing  $\beta$ -agarase and endo-(1,3)-glucanase] and enzymes with same substrate specificities are seen in different families indicating convergent evolution (e.g. cellulases are found in 11 different families,

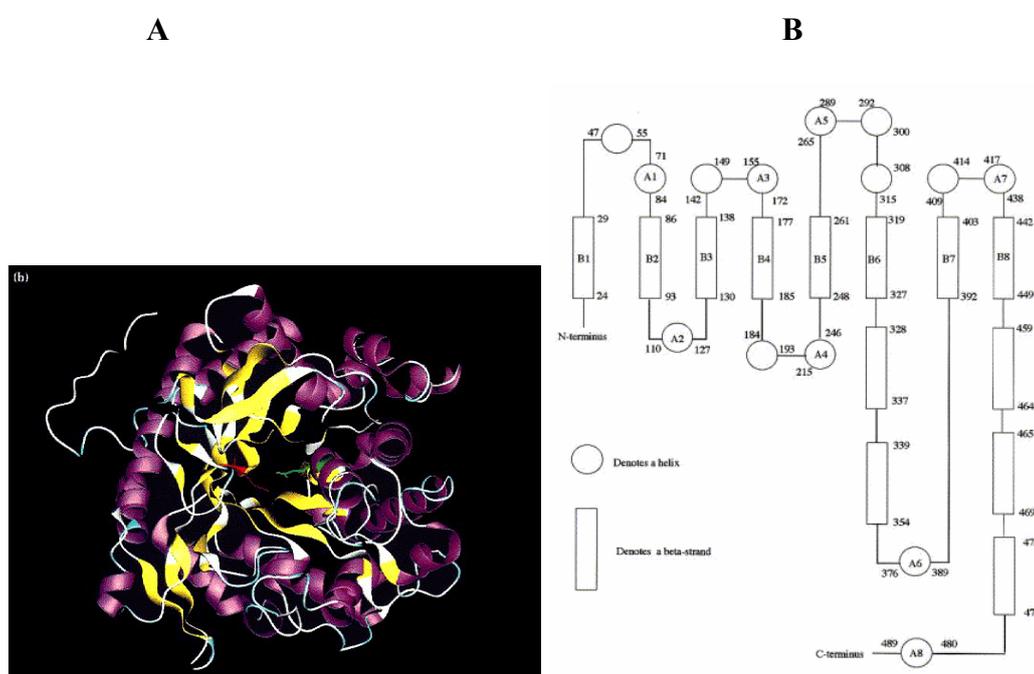
Davies and Henrissat, 1995). There has been a substantial increase in the number of known glycosyl hydrolase genes in a number of families recently and a classification reflecting the sequence and structural similarity is extremely useful.

The relationship between sequence and folding similarities also enhances the possibility of homology modeling, if the 3D structure of one enzyme in a family is known. A good example is that of family 20 chitobiase, the 3D structure of which is known. This family also contains human hexosaminidases, mutations of which are responsible for Tay Sachs disease. Hence from the known structure of chitobiase, Tews and coworkers (1996) were able to determine the residues responsible for this disease. Another example is that of glycosyl hydrolase family 3, where Harvey *et al.* (2000) used homology modeling to predict the structures of more than a hundred other enzymes based on the only known 3D structure of family 3 that was available, which was a barley  $\beta$ -glucan exohydrolase (Varghese *et al.*, 1999).

### **1.3 Glycosyl hydrolase families and folds**

Glycosyl hydrolases have been classified into a number of families by the system based on the similarities in the amino acid sequences (Henrissat and Bairoch, 1993). There has been a rapid increase in the number of 3D structures of a number of glycosyl hydrolases. Families whose structures are known may now be grouped into “Clans” with similar structures and mechanisms (Henrissat and Bairoch, 1996). So far, 14 clans have been assigned, comprising of 114 families ([http://www.cazy.org/fam/acc\\_GH.html](http://www.cazy.org/fam/acc_GH.html)). The largest of the glycosyl hydrolase clans is the GH- A clan, also called the 4/7 superfamily because the proton donor and the nucleophile are found on  $\beta$ -strands 4 and 7 of the  $(\beta/\alpha)_8$  barrel, respectively (Jenkin *et al.*, 1995). It is comprised of GH families 1, 2, 5, 10, 17, 26, 30, 35, 39, 42, 50, 51, 53,

59, 72, 79, 86, and 113, which possess different substrate specificities. The  $(\beta/\alpha)_8$  barrel is also called the TIM barrel, after the first enzyme in which it was identified, triose phosphate isomerase (Banner *et al.*, 1975). The typical alpha/beta barrel comprises of 8  $\alpha$ -helices and 8  $\beta$ -strands creating a barrel-like structure, the core of which is formed by a barrel-shaped parallel  $\beta$ -sheet, surrounded by the  $\alpha$ -helices (Figure 1.1).



**Figure 1.1** The overall topology of the white clover cyanogenic  $\beta$ -glucosidase (PDB code: 1CBG).

(A) Ribbon diagram showing the overall fold of the molecule. The catalytic residues are shown in stick representation with Glu183 in green and Glu397 in red; (B) Secondary structure diagram of white clover cyanogenic  $\beta$ -glucosidase showing the eight parallel  $\beta$ -strands that form the barrel core. The  $\beta$ -strands are represented by rectangles and the  $\alpha$ -helices by circles. The eight  $\beta$ -strands forming the barrel are labelled B1–B8 and the eight peripheral helices A1–A8. (Barrett *et al.*, 1995)

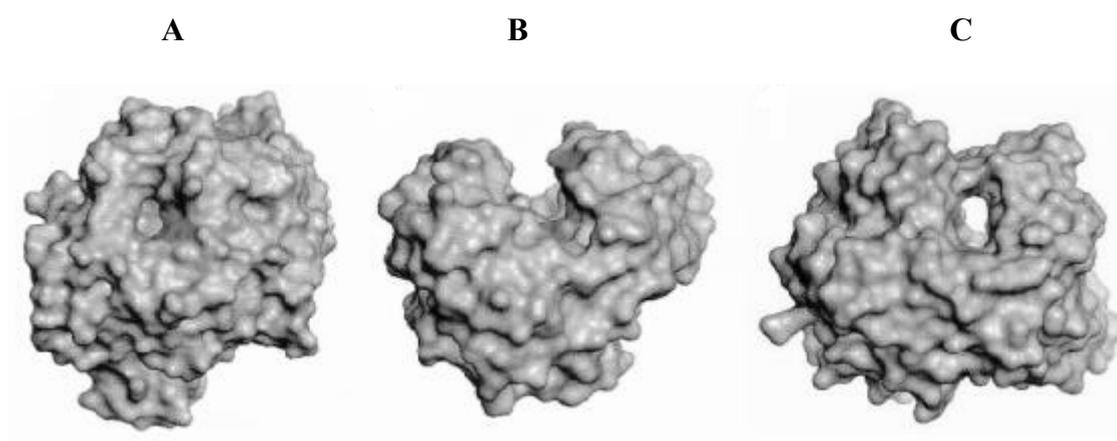
## 1.4 Morphology of active sites in Glycosyl Hydrolases

Despite the variety of protein folds seen in GH families, their active site structures, regardless of mechanism, fall into three general morphological classes: pocket crater, cleft or groove and tunnel. The pocket/crater (Figure 1.2 A) is seen in exo-acting hydrolases,  $\beta$ -amylase, glucoamylase (Aleshin *et al.*, 1992) and monosaccharidases, like  $\beta$ -glucosidase,  $\beta$ -galactosidase, and neuraminidase. It 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 (Davies and Henrissat, 1997).

A cleft or groove (Figure 1.2 B) is an open structure on the surface the enzyme, which allows for several sugar subunits of a polymeric substrate to randomly bind. This is typical for endo-acting hydrolases and includes lysozymes, endocellulases, chitinases,  $\alpha$ -amylases,  $\beta$ -(1 $\rightarrow$ 3,1 $\rightarrow$ 4)-glucanases and xylanases, among others. The endohydrolase usually has a substrate binding groove (cleft) that extends across its surface. Barley  $\beta$ -(1 $\rightarrow$ 3,1 $\rightarrow$ 4)-glucan endohydrolase has the oligosaccharide substrate bound into the open cleft extending across the surface of the enzyme. Because the catalytic amino acid residues are located in the substrate-binding cleft, the endohydrolase can bind anywhere along the polymeric substrate and subsequently hydrolyze the internal linkages.

A tunnel (Figure 1.2 C) 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  $\alpha$ -carrageenases (Michel *et al.*, 2001). 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 1.2** The three different types of active sites found in glycosyl hydrolases.

(A) The pocket (glucoamylase from *A. awamori*); (B) The cleft (endoglucanase E2 from *T. fusca*) and (C) The tunnel (cellobiohydrolase II from *T. reesei*). (Henrissat and Davies, 1995).

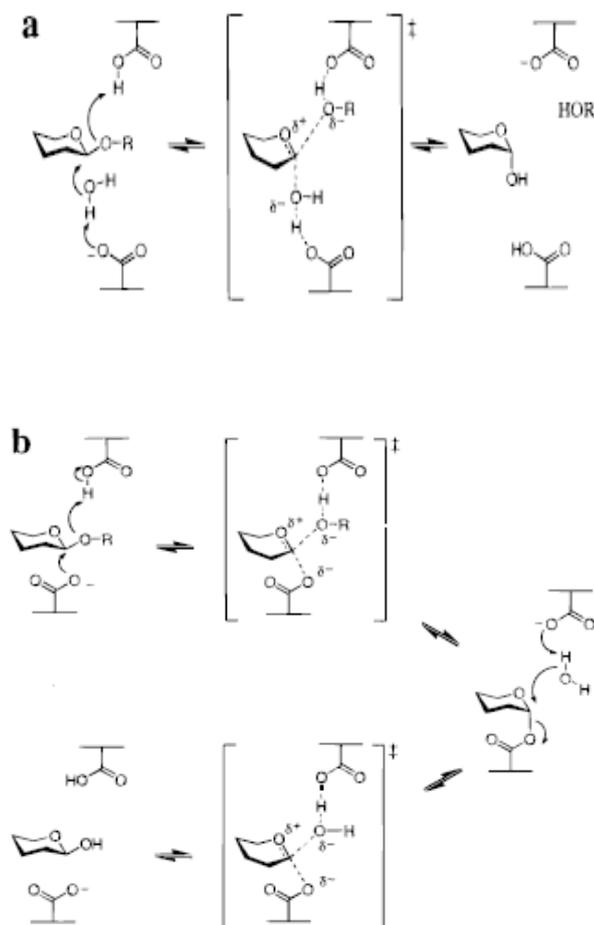
## 1.5 Catalytic Mechanism and Catalytic Machinery

The catalytic mechanism exhibited by GH family involves two general mechanisms with two possible outcomes, either the inversion or the retention of the stereochemistry around the anomeric C1 carbon. However, in both mechanisms, the position of the proton donor is similar, i.e., located within hydrogen bonding distance of the glycosidic oxygen.

Hydrolysis with inversion (Figure 1.3 a) is accomplished by a single displacement mechanism involving an oxocarbenium ion-like transition state, where

the general base deprotonates a water molecule, which then attacks the C1 carbon atom of the glycosidic bond (Zechel and Withers, 2000). The base is aided in cleaving the glycosidic bond by a general acid catalyst, which protonates the leaving group, the liberated aglycone. As a consequence of the inclusion of an activated water between the general base and the anomeric C1 atom in addition to the substrate, the two catalytic carboxyl groups in the protein are located on the opposite sides of the glycosidic bond, generally about 10.5 Å apart (McCarter and Withers, 1994; Wang *et al.*, 1994; White *et al.*, 1994).

The retaining mechanism (Figure 1.3 b) is a double displacement mechanism, which includes the formation of a covalently bound glycosyl enzyme intermediate that is further hydrolysed, with oxocarbenium-ion-like transition states at each step. Unlike the inversion mechanism, this is performed by the nucleophile directly attacking the C1 atom, rather than a water molecule. This requires a closer proximity of the two active carboxyl groups, generally around 5 to 5.5 Å. In the first step (glycosylation), the glycosidic oxygen is protonated by one of the carboxylate residues acting as the general acid catalyst, while the other carboxylate group is the nucleophile that attacks the anomeric carbon, resulting in the formation of a covalent intermediate. In the second step (deglycosylation), the general base deprotonates the incoming water molecule, which attacks at the anomeric center to cleave the covalent intermediate and displace the sugar (Zechel and Withers, 2000).



**Figure 1.3** General mechanisms for glycosyl hydrolases.

(a) The single displacement mechanism, which involves a single oxycarbenium ion transition state, is found in inverting glycosidases and (b) The double displacement mechanism, which involves two oxycarbenium ion transition states, is found in retaining glycosidases. O-R is the alkyl or aryl group, serving as the leaving group in the glycosylation step (Zechel and Withers, 2000).

Various techniques like covalent trapping intermediates, site directed mutagenesis and secondary methods like hydrophobic cluster analysis have been employed to identify the critical catalytic residues involved in the hydrolysis of the glycosidic bond and the regions surrounding the active site in retaining glycosyl

hydrolases. The 2-deoxy-2-fluoro glucoside inactivators (with reactive leaving groups like dinitrophenolate or fluoride) that function via the formation of stable glycosyl-enzyme intermediates have been used successfully in identification of the catalytic nucleophile of several retaining glucosidases (Withers and Street, 1988; MacLeod *et al.*, 1994). The mechanism involves replacing the 2-OH group with fluorine, which destabilizes the partial positive charge on C-1 and O5 in the transition states for glycosyl enzyme formation and hydrolysis slowing down both the steps of hydrolysis. This technique was first employed to identify the catalytic nucleophile in a GH family 1  $\beta$ -glucosidase from *Agrobacterium* (Withers *et al.*, 1987; Withers and Street, 1988), and later was used for family 3  $\beta$ -glucosidases from *Aspergillus niger* (Dan *et al.*, 2000) and *Flavobacterium meningosepticum* (Li *et al.*, 2002), among others (Withers and Aebersold, 1995).

Site-directed mutagenesis studies have been carried out to understand the functions of GH 1  $\beta$ -glucosidase's catalytic residues identified by sequence alignment, X-ray crystallography or homology modeling. From the known three-dimensional structure of family 1  $\beta$ -glucosidases, the two glutamates were located in the highly conserved LNEF and I/VTENG motifs positioned within the active site on opposite sides of the glycosidic bond (Davies and Henrissat, 1995, Henrissat and Bairoch, 1996). A highly substrate-specific strictosidine  $\beta$ -glucosidase playing a role in indole alkaloid metabolism has been studied by Barleben and coworkers (2007). They demonstrated the essential role of residues Glu-207, Glu-416, His-161, and Trp-388 in catalysis by site-directed mutagenesis. Site-directed mutagenesis has also been employed for other plant  $\beta$ -glucosidases, like prunasin hydrolase (Zhou *et al.*, 2002), maize  $\beta$ -glucosidase (Cicek *et al.*, 2000) and wheat  $\beta$ -glucosidase (Sue *et al.*, 2006).

Very recently, Noguchi and group (2008) have solved the structure of family 1 human cytosolic  $\beta$ -glucosidase KLRP acid/base mutant E165Q (1.9Å) with the glucose covalently bound to the nucleophile E373 by soaking it with a high concentration of *para*-nitrophenyl- $\beta$ -D-glucopyranoside substrate, further confirming the mechanism of double displacement in retaining  $\beta$ -glucosidases. Further, this structure also suggested the role of a water molecule in stabilizing the transition state 2-OH of glucose by the structural changes of the active site residues creating a space for the water to bind and extend the hydrogen bonding network.

In most of the mutagenic studies described above, the catalytic amino acids were first identified based on multiple sequence alignment. When the sequences are not similar enough to align directly, hydrophobic cluster analysis can be used. Hydrophobic cluster analysis has been employed to analyze the regions surrounding the nucleophile, for example, in many of the retaining glycosyl hydrolases (e.g. GH family 1). It has been revealed that the nucleophile is the glutamic acid residue located after a hydrophobic cluster (Henrissat *et al.*, 1995). The number of residues between the proton donor and the nucleophile varies from 75 to more than 200 amino acid residues. These variations could be due to loops of different lengths between the secondary structure elements of the  $(\alpha/\beta)_8$  barrel or to insertion of one or more extra domains.

## 1.6 Overview of GH Family 1 (GH1) $\beta$ -glucosidases

$\beta$ -glucosidases [ $\beta$ -D-glucoside glucohydrolase; (EC 3.2.1.21)] are glycosyl hydrolases that catalyze the hydrolysis of the glycosidic bond at the nonreducing terminal glycosyl residue of a glycoside or an oligosaccharide, releasing glucose and an aglycone or a shortened oligosaccharide.  $\beta$ -glucosidases are grouped into GH

families 1, 3, and 9 from the CAZY website (Henrissat, 1993; Coutinho and Henrissat, 1999; Opassiri *et al.*, 2007).

### 1.6.1 GH1

Most plant  $\beta$ -glucosidases that have been characterized fall in GH1. In addition to  $\beta$ -glucosidases, glycosyl hydrolase family 1 (GH1) also includes myrosinases (thio- $\beta$ -glucosidases) hydrolyzing the *S*-glycosidic bonds of plant 1-thio- $\beta$ -D-glucosides (glucosinolates), (Burmeister *et al.*, 1995),  $\beta$ -mannosidases,  $\beta$ -galactosidases,  $\beta$ -glucuronidases,  $\beta$ -fucosidases, 6-phospho- $\beta$ -galactosidases, diglycosidases like primeverosidase (Mizutani *et al.*, 2002), furcatin hydrolase (Ahn *et al.*, 2004) and isoflavone 7-*O*- $\beta$ -apiosyl- $\beta$ -1,6-glucosidase (Chuankhayan *et al.*, 2005), and hydroxyisourate hydrolase, which hydrolyzes an internal bond in a purine ring, rather than a glycosidic bond (Raychaudhuri and Tipton, 2002).

### 1.6.2 $\beta$ -Glucosidase functions

$\beta$ -Glucosidases are ubiquitous in the living world (having been found in archaea, bacteria, fungi, insects, plants and mammals) and play important role in biological processes (Esen, 1993). In bacteria and fungi, they are involved in cellulose and cellobiose metabolism and, thus, play a role in biomass conversion (Bèguin, 1990; Fowler, 1993). In plants, they have been found to be involved in numerous functions, including cell wall metabolism (Leah *et al.*, 1995), defense, lignification (Dharmawardhana *et al.*, 1995), aroma formation in tea leaves (Mizutani *et al.*, 2002), regulation of the biological activity of plant phytohormones, such as cytokinins, gibberellins and auxins (Brzobohaty *et al.*, 1993; Falk and Rask, 1995; Schliemann, 1984; Haberer and Kieber, 2002), control of biosynthesis of indole 3-acetic acid (Ljung *et al.*, 2001; Persans *et al.*, 2001), biosynthesis of alkaloids in indole alkaloid

metabolism (Barleben *et al.*, 2005; 2007) and others. In mammals, GH1 enzymes have been studied for their functions, which include those of human  $\beta$ -glucosidase (lactase phlorizin hydrolase) involved in flavonoid metabolism (Day *et al.*, 2000) and in the metabolism of dietary lactose (Naim *et al.*, 2001). In addition, deficiency of a lysosomal  $\beta$ -glucosidase (glucocerebrosidase), belonging to GH family 30, causes accumulation of glucocerebrosides in the lysosomes of some cells, which causes Gaucher's disease (Michelin *et al.*, 2004). Some of the plant functions are illustrated in detail below.

In many plants,  $\beta$ -glucosidases are known to play a significant role in defense mechanisms. In white clover (*Trifolium repens* L.) and cassava (*Manihot esculenta* Crantz), the cyanogenic glucosides, linamarin and lotaustralin occur in the cell walls. Upon tissue damage, these cyanogenic glucosides are cleaved by cyanogenic  $\beta$ -glucosidases, resulting in the production of the toxic compound hydrogen cyanide, which explains the process of cyanogenesis in response to herbivore or fungal attack. Similarly, hydroxamic acids released from their glucosides by  $\beta$ -glucosidases are the major defense compounds in plants like maize, wheat, rye, and barley (Niemeyer, 1988), and have been shown to inhibit bacterial and fungal growth, as well as insect development and reproduction (Argandona *et al.*, 1983). Other plant secondary metabolites like flavonoids also exist in plants as glucoconjugates, the aglyconic forms of which are released upon glycoside hydrolysis and play important roles in defense. In tea plants, a  $\beta$ -glucosidase ( $\beta$ -primeverosidase) hydrolyzes  $\beta$ -primeverosides to form aglycones like methyl salicylate, geraniol, linalool and 3-hexenol (Mizutani *et al.*, 2002). These aglycones play a role in defense and have anti-microbial and antifungal activities.

Isoflavone glycosides have also been suggested to act in defense, as well as other plant-microbe interactions, such as recruitment of rhizobium bacteria by legumes (Suzuki *et al.*, 2006). For example, a  $\beta$ -glucosidase with high specific activity toward isoflavone conjugates was purified from soybean. It hydrolyzed both the 7-O- $\beta$ -D-glucoside and 7'-O-(6''-O  $\beta$ -D-malonyl- $\beta$ -D-glucoside) forms of the isoflavones genistein and daidzein to release free aglycones, which may in turn play important roles in plant-microbe interactions and defensive mechanisms against pathogen infection (Hsieh and Graham, 2001; Suzuki *et al.*, 2006). A  $\beta$ -glucosidase highly specific for isoflavone 7-O- $\beta$ -D-glucosides have also been purified from garbanzo bean (*Cicer arietinum* L.) roots, leaves, and hypocotyls (Hösel and Barz, 1975). Other isoflavone specific enzymes include a glycosyl hydrolase family 1  $\beta$ -glycosidase purified from *D. nigrescens* Kurz, which was found to have isoflavonoid  $\beta$ -apiosylglucosidase activity and also hydrolyzed the soybean isoflavonoid  $\beta$ -glucosides, genistin, daidzin and malonyl genistin (Chuankhayan *et al.*, 2005; 2007).

Alkaloid metabolism is mediated by a highly substrate specific strictosidine  $\beta$ -D-glucosidase, which activates the substrate strictosidine to form a reactive aglycone which enters multiple biochemical pathways to indole and quinoline alkaloid types in three plant families, including Apocynaceae, Rubiaceae, and Loganiaceae (Barleben *et al.*, 2007). Some of the products formed by the hydrolysis of strictosidine  $\beta$ -D-glucosidase include vindoline, reserpine, ajmaciline, strychnine camptothecin and quinine. In addition, they have also been assigned to play a role in damage induced defense system (Geerlings *et al.*, 2000).

$\beta$ -glucosidases are also involved in biosynthesis of lignins by hydrolysing the monolignol  $\beta$ -glucosides. A coniferin specific  $\beta$ -glucosidase from

*Pinus contorta* has been shown to perform a vital role in pine stem lignification by releasing the monolignol coniferol (Dharmavardhana *et al.*, 1995). In *Arabidopsis thaliana*, two  $\beta$ -glucosidases, BGLU45 and BGLU46, which cluster with the *Pinus contorta*  $\beta$ -glucosidase, also hydrolyze monolignol glucosides suggesting, their involvement in lignin biosynthesis (Escamilla-Treviño *et al.*, 2006).

$\beta$ -Glucosidase (BGQ60) from barley hydrolyzed the main cell wall polysaccharides present in barley grain endosperm tissues,  $\beta$ -(1 $\rightarrow$ 3, 1 $\rightarrow$ 4)-glucan, cellulose,  $\beta$ -(1 $\rightarrow$ 4)-glucan, and  $\beta$ -(1 $\rightarrow$ 3)-glucan and cell wall degradation is extremely important to provide a carbohydrate source during germination. Combined with the catalytic activities of the  $\beta$ -glucan endohydrolases (endo- $\beta$ -glucanases), barley  $\beta$ -glucosidase (BGQ60) appears to play an important role in cell wall metabolism (Leah *et al.*, 1995; Hrmova *et al.*, 1998).

In rice, a cell wall bound  $\beta$ -glucosidase showed an increase in activity during germination, highlighting its likely role in cell wall metabolism (Akiyama *et al.*, 1998). This enzyme displayed broad specificity toward  $\beta$ -linked oligosaccharides, further enhancing its role in hydrolyzing the oligosaccharides from cell wall bound  $\beta$ -glucans during the germination process.

In phytohormone regulation,  $\beta$ -glucosidases hydrolyze the phytohormone conjugates releasing the active phytohormones. A  $\beta$ -glucosidase from maize hydrolyzed the biologically inactive cytokinin-O-glucosides and kinetin-N3-glucoside to release active cytokinins, which might be important for embryo development (Brzobohaty *et al.*, 1993). A  $\beta$ -glucosidase from *Brassica napus* expressed in young tissues undergoing cell division hydrolyzed the cytokinin zeatin-O-glucoside, providing the young tissue cells with active cytokinins and thus playing

a role in seedling development (Falk and Rask, 1995). A  $\beta$ -glucosidase from barley leaves hydrolyzed abscisic acid (ABA) glucopyranoside in the apoplast releasing the physiologically active free ABA, which triggers stress responses (Dietz *et al.*, 2000). *Arabidopsis*  $\beta$ -glucosidase AtBG1, which is ER localized, hydrolyzes inactive glucose-conjugated ABA to active ABA (Lee *et al.*, 2006). The ABA-GE is stored separately, in vacuoles and apoplastic space in the cell, and AtBG1 its substrate come together when cells need to increase the ABA content, in response to abiotic stress conditions.

### 1.6.3 Substrate specificity and 3D structures

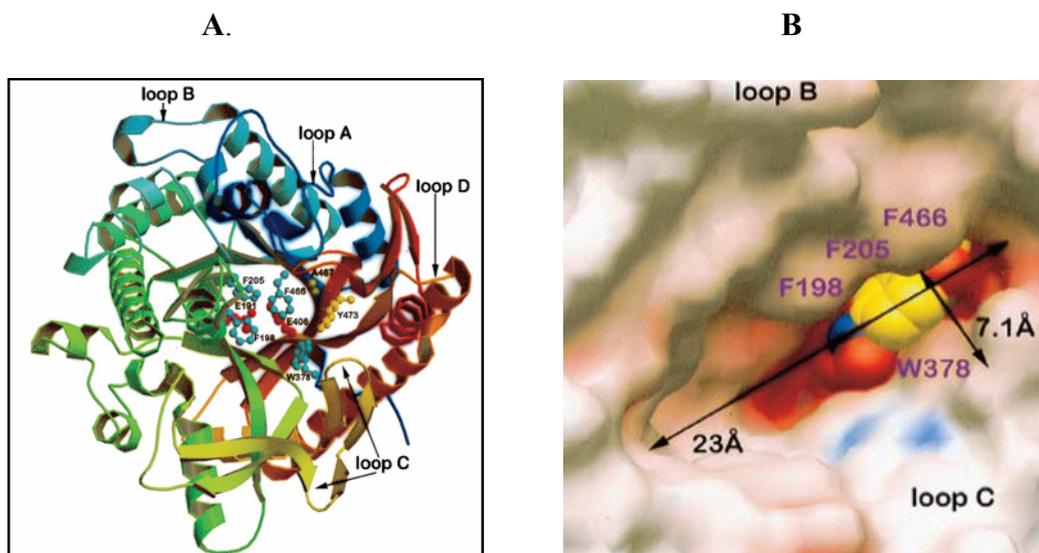
$\beta$ -glucosidases possess a wide range of substrate specificities with respect to the aglycone moiety, and the preferred aglycone moieties vary with each  $\beta$ -glucosidase. For example, the maize isoenzymes ZmGlu1 and ZmGlu2 hydrolyze a broad range of artificial and natural substrates, including DIMBOA-Glc (2-O- $\beta$ -D-glucopyranosyl-4-hydroxy-7-methoxy-1,4-benzoxazin-3-one) (Verdoucq *et al.*, 2004, Czjzek *et al.*, 2000). The substrate specificity of maize ZmGlu1 $\beta$ -glucosidase was studied by determination of the structure of the inactive mutant Glu191Asp with its substrate (DIMBOA-Glc), aglycone (DIMBOA), and the inhibitor *p*-hydroxy-(S)-mandelonitrile- $\beta$ -D-glucose (dhurrin) in the active site (Czjzek *et al.*, 2000). The active site was shown to be shaped like a flattened crater with two distinct regions. The first region, which is at half the depth of the active site, which is the aglycone binding pocket, is 23 Å long and 7.1-7.6 Å wide, with Phe198, Phe205, and Phe466 on one side and Trp378 on the other side (Figure 1.4). The aglycone moiety is found sandwiched by these four hydrophobic residues. In addition, other amino acid residues which stabilize the interaction between the aglycone and the aglycone binding pocket

were found, including Tyr473, Ala467, Cys210, Cys216, Thr334, and Met374. The second region is the glycone-binding pocket, which is the bottom half of the active site. This region is 11.6 by 9.6 Å and, as such, has a larger width than the aglycone binding site. The four key aromatic amino acids Trp378, Phe198, Phe205, and Phe466 and the shape of the aglycone binding pocket determine the substrate specificity in maize  $\beta$ -glucosidase. With the exception of Trp378, the hydrophobic amino acids at these positions are highly variable among  $\beta$ -glucosidases, which corresponds to their function in determining substrate specificity.

Sorghum  $\beta$ -glucosidase (Dhurrinase, SbDhr1) which shares a sequence identity of 72% with maize ZmGlu1  $\beta$ -glucosidase, acts specifically on the natural substrate dhurrin. In maize ZmGlu1, dhurrin is not a substrate and instead inhibits enzyme activity. The difference in substrate specificity between these two enzymes is attributed to the amino acid residues located in the aglycone binding site. The SbDhr1 residues that interact with dhurrin include Asn259, Phe261, and Ser462, which are different from the maize enzyme. The residues which determine the substrate specificity of maize ZmGlu1 have been mutated to the corresponding residues of SbDhr1 (Verdoucq *et al.*, 2004). The Phe198Val mutant had a major effect on the enzyme activity of maize ZmGlu1. The ZmGlu1 double mutant Glu191D-Phe198Val with the DIMBOA-Glc complex showed that Val198 had an influence on the positions of residues Phe205, Phe466, and Glu264. Consequently, rearrangement of these residues in the aglycone binding pocket caused the substrate to bind in a different manner when compared with the native enzyme (Verdoucq *et al.*, 2004). Mutation of Tyr473 to phenylalanine increased the catalytic efficiency of maize ZmGlu1 for hydrolysis of dhurrin to 300% of its level in wild type ZmGlu1. This mutation

destroyed an H-bond which formed between Tyr473 and Trp378, resulting in Trp378 being more flexible, so that it could adapt its orientation to bind the aglycone of dhurrin in a productive manner.

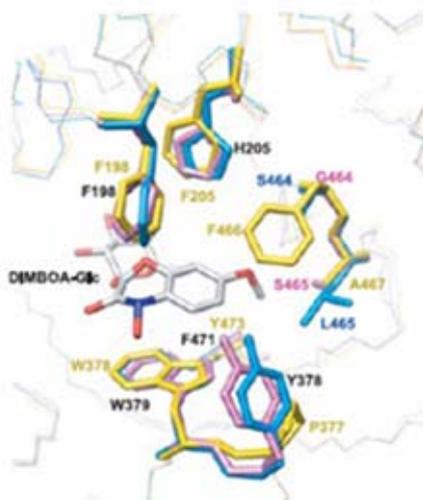
Wheat (TaGlu1a) and rye  $\beta$ -glucosidases (ScGlu) are also known to hydrolyze DMBOA-Glc and 2-O- $\beta$ -D-glucopyranosyl-4-hydroxy-1,4-benzoxazin-3-one (DIBOA-Glc), respectively. The preferred natural substrate for each enzyme is consistent with the predominant hydroxamic acids found in each of the plants: DIMBOA-Glc in TaGlu1a and DIBOA-Glc in ScGlu (Sue *et al.*, 2006). Rye ScGlu shares 70% sequence identity to maize ZmGlu1, for which the natural substrate is DIMBOA-Glc. Trp378 and Phe198 of ZmGlu1 are conserved in ScGlu at the corresponding positions, however Phe205 and Phe466 are substituted by a histidine and serine in TaGlu1a and histidine and glycine, in ScGlu, respectively. These substitutions allow the ScGlu to accept DIBOA-Glc as the preferred substrate (Nikus *et al.*, 2003). The sole aromatic residue (corresponding to Phe198 in ZmGlu1) is positioned opposite Trp379 in TaGlu1a and ScGlu, and is likely to play a significant role in substrate binding (Figure 1.5). The broader substrate specificity of ScGlu may be due to the wider aglycone binding site compared with those of ZmGlu1 and TaGlu1a.



**Figure 1.4** Structure of maize ZmGlu1  $\beta$ -glucosidase and its inactive Glu191Asp mutant bound to DIMBOA-glc.

(A) Ribbon diagram of maize ZmGlu1 and its inactive Glu191Asp mutant. The catalytic residues, Glu191 (Asp191 in the mutant) and Glu406, are shown in red. Four residues (Phe198, Phe205, Trp378, and Phe466), which form the aglycone-binding pocket are shown in blue and other residues probably important for aglycone recognition, Ala467 and Tyr473 are shown in yellow. The protein is colored in order of the visual spectrum with the N-terminus in dark blue and C-terminus in dark red;

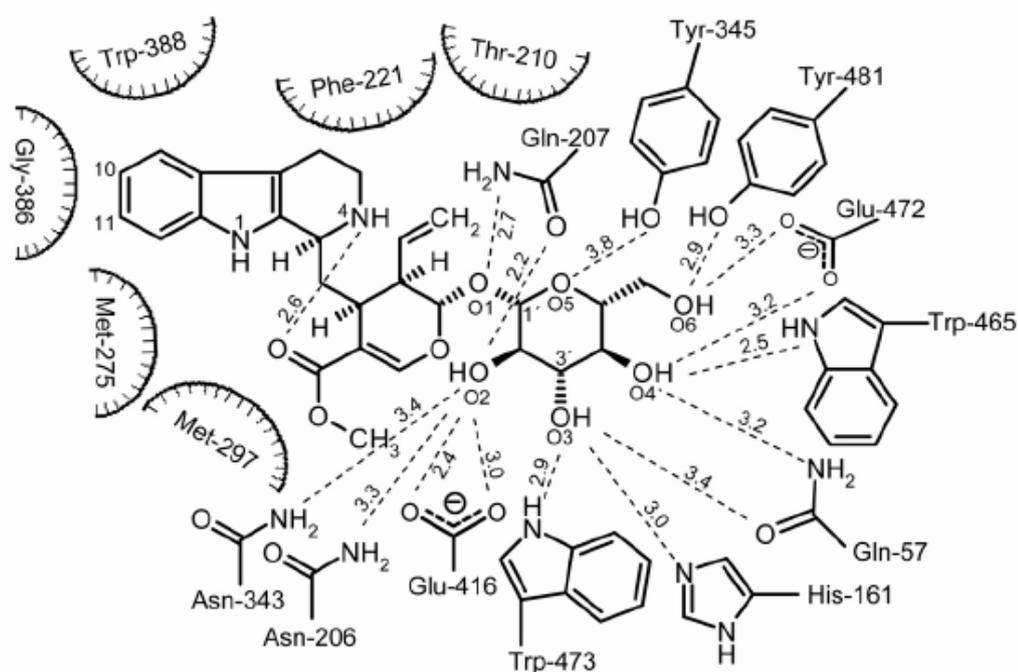
(B) The aglycone binding pocket of maize Glu1E191D showing positively charged regions (blue), negatively charged regions (red) and neutral regions (white). The slot-like active site, contains the natural substrate DIMBOAGlc in compact representation with standard atom-type colors. In the figure, only the aglycone moiety is visible in its binding site, as the glucose molecule is hidden below the aglycone (Czjzek *et al.*, 2000).



**Figure 1.5** Superimposition of the crystal structures of TaGlu1b (PDB code: 2DGA) and ZmGlu1-Glu191Asp (PDB code: 1E56) and the modeled structure of ScGlu. A natural substrate, DIMBOA-Glc, bound to the maize enzyme is also shown. Blue, TaGlu1b; magenta, ScGlu; yellow, ZmGlu1-Glu191Asp. Identical residues found in TaGlu1b and ScGlu are labeled in black (Sue *et al.*, 2006).

In plant cells, strictosidine  $\beta$ -glucosidase (SG) is known to hydrolyze its natural substrate strictosidine, the aglycone of which in turn undergoes further reactions of indole alkaloid metabolism (Barleban *et al.*, 2007). In the SG Glu207Gln mutant structure (PDB code, 2JF6), the aglycone part of the strictosidine substrate is surrounded by Phe221, Trp388, Gly386, Met275, Thr210, and Met297, most of which are hydrophobic (Figure 1.6). 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 larger residues like serine decreased the enzyme activity by 90%, as demonstrated for the Gly386Ser mutant. The glucose moiety interacts with a number of hydrophilic residues (Asn206, Gln207, Asn343, Tyr345, Glu416, Trp473,

His161, Gln57, Trp465, Glu472, and Tyr481). Gln207 and Glu416 of SG are located within the pocket near the sugar moiety, with a distance of 5.2 Å between their carboxyl carbons 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 C-atom of the glucose 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 the correct orientation for deglycosylation. Tyr481 corresponds to Tyr473 in maize (*Zea mays*)  $\beta$ -glucosidase, 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 Trp378 (Trp388 in SG) (Verdoucq *et al.*, 2003).



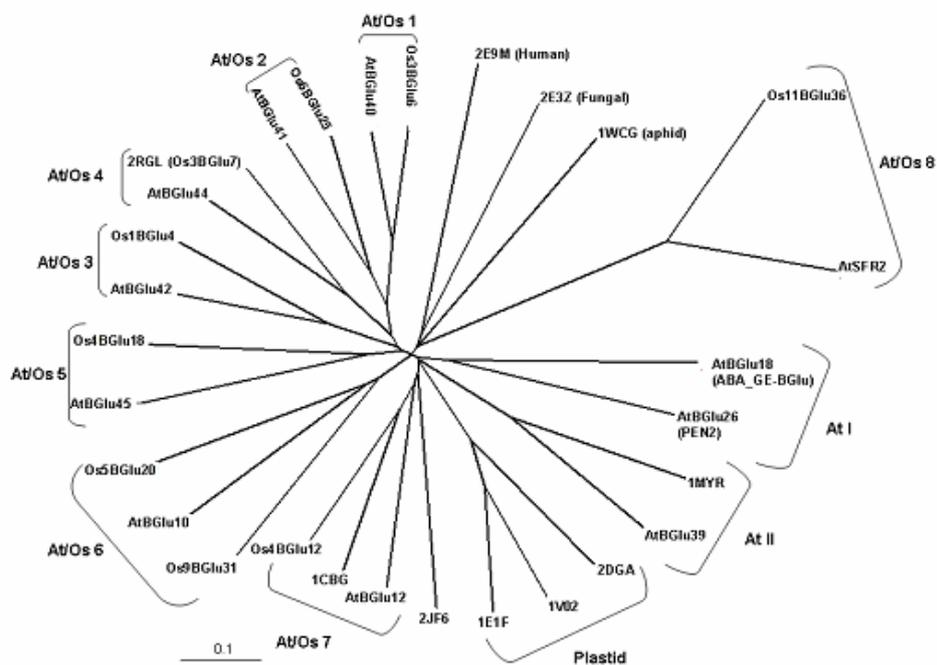
**Figure 1.6** Hydrogen bonding network between the glucosidic part of strictosidine and residues in the ligand structure of strictosidine  $\beta$ -glucosidase (SG) inactive mutant Glu207Gln (Barleben *et al.*, 2007).

## 1.7 Rice GH Family1 $\beta$ -glucosidases

In rice,  $\beta$ -glucosidases are implicated in functions such as germination (Palmiano and Juliano, 1973), and hydrolysis of gibberlin glucosides (Schliemann, 1984), cyanogenic  $\beta$ -D-glucosides (*in vitro*), such as amygdalin, prunasin and linamarin (Esashi *et al.*, 1991) and cell-wall-derived oligosaccharides (Akiyama *et al.*, 1998). Two cDNAs for rice  $\beta$ -glucosidase isoenzymes BGlu1 (Os3BGlu7) and BGlu2 (Os9BGlu30) were isolated from germinating rice (Opassiri *et al.*, 2003), and recently another rice isoenzyme Os4BGlu12 cDNA was cloned and expressed (Opassiri *et al.*, 2006) to test its substrate specificity. One isoenzyme, Os3BGlu7, was highly expressed in flower and germinating shoot, while Os9BGlu30 was expressed in seedling shoot. Os3BGlu7 catalyzed hydrolysis of gluco-oligosaccharides and pyridoxine-5'-*O*- $\beta$ -glucoside, a vitamin B<sub>6</sub> metabolite found in rice bran (Opassiri *et al.*, 2004). Os3BGlu7 shares an amino acid sequence identity of 66% with barley  $\beta$ -glucosidase (BGQ60), which prefers  $\beta$ -D-mannoside to  $\beta$ -D-glucoside and cellobiose to cellotriose (Hrmova *et al.*, 1998). In contrast, Os3BGlu7 hydrolyzes these same substrates with the opposite preferences. A thioredoxin fusion protein of Os3BGlu7 expressed in *E. coli* hydrolysed short  $\beta$ -(1 $\rightarrow$ 3)- and  $\beta$ -(1 $\rightarrow$ 4)-linked gluco-oligosaccharides, but preferred cellotriose to cellobiose. Os4BGlu12 also hydrolyzed  $\beta$ -(1 $\rightarrow$ 4)-linked cellooligosaccharides, similar to rice Os3BGlu7 and the cell wall bound  $\beta$ -glucosidase, but not  $\beta$ -mannoside, which was hydrolyzed by Os3BGlu7. In contrast to barley and rice cell wall-bound enzyme, Os4BGlu12 did not hydrolyze  $\beta$ -(1 $\rightarrow$ 3)-linked oligosaccharides longer than laminaribiose, but hydrolyzed various *p*NP-derivatives of monosaccharides, including  $\beta$ -D-fucoside,  $\beta$ -D-glucoside,  $\beta$ -D-galactoside,  $\beta$ -D-xyloside and  $\alpha$ -L-arabionoside. The crystal structure of rice

Os3BGlu7 has recently been determined (Cheunchor *et al.*, 2008) and the  $\beta$ -(1 $\rightarrow$ 4)-linked celotriose was docked into the active site to try to explain the differences in preferences of substrates between Os3BGlu7 and BGQ60. Four of the residues at the +1 and +2 subsites of rice were mutated to the corresponding barley residues, as follows: Ile179Val, Asn190His, Asn245Val, and Leu442Arg. Although the first three mutations were shown to have an effect on substrate binding and catalytic rate, no single mutation caused a reversal in the hydrolytic preferences for cellobiose and celotriose.

Forty rice GH1 genes, 34 of which appear to be functional  $\beta$ -glucosidase genes have been identified from rice genome sequences (Opassiri *et al.*, 2006) while 48 GH1 genes were found in *Arabidopsis thaliana* genome (Xu *et al.*, 2004). Database information was used to determine the tissue-specific expression and predict the localization of these rice GH1  $\beta$ -glucosidases. Phylogenetic analysis of the predicted GH1 proteins from the rice and *Arabidopsis* genomes, showed eight clusters containing both rice and *Arabidopsis* proteins that are more closely related to each other than they are to enzymes from the same plants outside the clusters (Opassiri *et al.*, 2006). In addition, there were two clusters of *Arabidopsis* enzymes without closely related rice counterparts, including a group of myrosinases (not found in rice) and a large group of  $\beta$ -glucosidases. Comparison to characterized GH1 enzymes from other plants showed other clusters of related enzymes not found in rice or *Arabidopsis*, including the chloroplastic enzymes from which the maize, sorghum and wheat structures are derived, and the cytoplasmic metabolic enzymes, such as strictosidine  $\beta$ -glucosidases. On the other hand, cyanogenic  $\beta$ -glucosidases, such as 1CBG, are rather close to cluster 7 of rice and *Arabidopsis* enzymes (Figure 1.7).



**Figure 1.7** Simplified phylogenetic tree of the amino acid sequences of eukaryotic GH1 proteins with known structures and those of rice and *Arabidopsis* GH1 gene products.

The protein sequences of the eukaryotic proteins with known structures are marked with 4-letter PDB codes for one of their structures, including *Trifolium repens* cyanogenic  $\beta$ -glucosidase (1CBG, Barrett *et al.*, 1995), *Sinapsis alba* myrosinase (1MYR, Burmeister *et al.*, 1997), *Zea mays* ZmGlu1  $\beta$ -glucosidase (1E1F, Czjzek *et al.*, 2000), *Sorghum bicolor* Dhr1 dhurrinase (1V02, Verdouqc *et al.*, 2004), *Triticum aestivum*  $\beta$ -glucosidase (2DGA, Sue *et al.*, 2006), *Rauwolfia serpentina* strictosidine  $\beta$ -glucosidase (2JF6, Barleben *et al.*, 2007), and *Oryza sativa* BGlu1  $\beta$ -glucosidase (2RGL, Cheunchor *et al.*, 2008) from plants, along with *Brevicoryne brassicae* myrosinase (1WCG, Husebye *et al.*, 2005), *Homo sapiens* cytoplasmic (Klotho)  $\beta$ -glucosidase (2E9M, Hayashi *et al.*, 2007), and *Phanerochaete*

*chrysochlorium*  $\beta$ -glucosidase (2E3Z, Nijekkin *et al.*, 2007), while those encoded in the *Arabidopsis thaliana* and *O. sativa* genomes are labeled with the systematic names given by Xu *et al.* (2004) and Opassiri *et al.* (2006), respectively. One or two example proteins from each plant are given for each of the eight clusters of genes shared by Arabidopsis (At) and rice (Os), At/Os 1 (which has 1 At and 2 Os genes in total), At/Os 2 (1 At, 1 Os), At/Os 3 (1 At, 1 Os), At/Os 4 (2 At, 5 Os), At/Os 5 (3 At, 3 Os), At/Os 6 (10 At, 11 Os, which fall in two clusters bracketing the At sequences) At/Os 7 (6 At, 12 Os), and At/Os 8 (1 At, designated Sfr2 or AtBGlu48, 1 Os), and the Arabidopsis-specific clusters At I (16 At, including the abscissic acid glucosyl ester  $\beta$ -glucosidase, ABA-GE BGlu, and microbial interaction-related PEN2  $\beta$ -glucosidase) and At II (6 At myrosinases). These sequences were aligned with all the At and Os sequences in Clustalx (Thompson *et al.*, 1997), the alignment manually edited, all but representative sequences were removed, and the tree was calculated by neighbor joining method, then drawn with TreeView (Page, 1996). The grass plastid  $\beta$ -glucosidases, which are not represented in *Arabidopsis* and rice, are marked in the group marked "Plastid."

Though the known structures provide good tools for molecular modeling of rice enzymes, most lack a close correspondence in sequence and functional evolution, suggesting the variable loops that determine the active site may be different. To improve on this situation, it would be useful to have the structure and known function of representative members of each of the 8 clusters seen in rice and *Arabidopsis*. To begin to elucidate this information and to understand the different substrate specificities of rice GH1  $\beta$ -glucosidases, which help to determine their physiological

functions in rice, their three-dimensional structures are necessary, since small differences in the binding site may lead to differences in specificity.

One such enzyme is Os3BGlu6, a representative member of cluster 1, which has not previously been functionally or structurally characterized. This study aimed at the recombinant expression of Os3BGlu6 as an N-terminal thioredoxin and polyhistidine-tagged fusion protein. Subsequently, the recombinant protein was purified and characterized to determine what substrate(s) it hydrolyzed. The 3D structure of native Os3BGlu6, and those of two inhibitor complexes were also determined to understand the structural basis for its glycone and aglycone binding specificity.

## 1.8 Research Objectives

1. To clone the rice Os3BGlu6 cDNA into an *E. coli* fusion protein expression system, optimize for suitable expression conditions, and purify it.

2. To characterize the biochemical properties and substrate specificity of Os3BGlu6 with *p*NP-glycosides, and other synthetic and natural glycosides and oligosaccharides.

3. To determine the three-dimensional structure of Os3BGlu6, a representative of a plant GH1 phylogenetic cluster (A1/OS1) from which a structure has yet to be determined, by x-ray crystallography.

4. To determine the interactions involved in the active site glycone and aglycone binding and specificity of Os3BGlu6 by determining its 3D structures in complexes with the inhibitors 2-deoxy-2-fluoroglucofuranoside and *n*-octyl- $\beta$ -D-thioglucofuranoside.

## CHAPTER II

### MATERIALS AND METHODS

#### 2.1 Materials

##### 2.1.1 Plasmids and bacterial strains

The Sapporo cold-treated rice (*Oryza sativa* L., cultivar Yukihihikari) seedling cDNA library clone for *Os3BGlu6* (Accession number AY129294) was provided by Dr. Takashi Akiyama. Other plasmids used for this work included the pENTR™/TEV/D-TOPO (Invitrogen) and pET32a(+)/DEST (Opassiri *et al.*, 2006) vectors. The *E. coli* host cell strains used included DH5 $\alpha$  and Mach 1 cells for cloning and Origami(DE3) for recombinant protein expression.

##### 2.1.2 Reagents and Chemicals

Ammonium persulfate, acrylamide, *N,N',N'',N'''*-tetramethylethylenediamine (TEMED), sodium dodecyl sulfate (SDS), *N,N'*-methylene-bis-acrylamide, Triton X-100, and lysozyme were purchased from Amersham Pharmacia Biotech (Uppsala, Sweden). Sodium acetate, sodium hydroxide, sodium chloride (NaCl), sodium carbonate, disodium ethylenediamine tetraacetate (EDTA), piperazine-1,4-bis(2-ethanesulfonic acid) (PIPES), bromophenol blue, sulfuric acid, methanol, HPLC- grade distilled water, ethanol, glacial acetic acid and ethyl acetate were purchased from Carlo ERBA (Rodano, Milano, Italy). Imidazole, chloramphenicol, Coomassie Brilliant Blue R250, phenylmethylsulfonylfluoride

(PMSF), calcium chloride, bovine serum albumin (BSA), dichlorodimethylsilane and ethyidium bromide were purchased from Fluka (Steinheim, Switzerland). Ampicillin and thin-layer chromatography silica gel 60 aluminum plates for oligosaccharide detection and F<sub>254</sub> plates were purchased from Merck (Darmstadt, Germany). 2, 2'-azinobis(3-ethylbenthaiazolinesulfonic acid) (ABTS), isopropyl thio-β-D-galactoside (IPTG), kanamycin, tetracycline, DNase I, soybean trypsin inhibitor, 2-deoxy-2-fluoro-β-D-glucoside, BisTris, *n*-octyl-β-D-thio-glucopyranoside, chloroform/isoamyl alcohol (24:1), Tris base, 2-mercaptoethanol, *p*-nitrophenol β-D-glucoside (*p*NPGlc), *p*NP-β-D-fucoside, *p*NP-β-D-arabinoside, *p*NP-β-D-galactoside, *p*NP-β-D-xyloside, *p*NP-β-D-mannoside, *p*NP-β-D-cellobioside, sophorose, gentiobiose and cellobiose were purchased from Sigma (St. Louis, MO, USA). Cellooligosaccharides of degree of polymerization (DP) 3-6 and laminarioligosaccharides of DP 2-5 were purchased from Seikagaku Kogyo Co. (Tokyo, Japan). *Taq* DNA polymerase, *Pfu* DNA polymerase, agarose and deoxynucleoside triphosphates (dNTPs) were purchased from Promega (Madison, WI, USA). Tryptone, glycine and yeast extract were purchased from Scharlau Chemie S.A. (Barcelona, Spain). Talon Co<sup>2+</sup> affinity resin was purchased from Clontech (San Francisco, CA, USA). Ultrafiltration membranes 30 kDa MW and 10 kDa MW cutoff, Ultrafree MC 0.22 μm and 0.45 μm filters were purchased from Millipore Corporation, (Bedford, MA, USA). High vacuum grease was purchased from Dow Corning, (PA, USA). Some of the kits used were Bio-RAD protein assay kit (Bio-RAD Corp., Hercules, CA, USA), QIA prep spin miniprep plasmid extraction kit (QIAGEN) and Perfectprep Gel Cleanup kit (QIAGEN, Hilden, Germany). Other chemicals and molecular reagents used but not listed here were purchased from a variety of suppliers.

## 2.2 General Methods

### 2.2.1 Preparation of *E. coli* strains DH5 $\alpha$ and Origami(DE3) for competent cells

Glycerol stocks of DH5 $\alpha$  and Origami(DE3) *E. coli* were streaked on a fresh LB plate and incubated overnight at 37°C for 16 to 18 hr. A well formed single colony was then picked and inoculated into 3 ml LB medium (containing 15  $\mu$ g/ml kanamycin and 12.5  $\mu$ g/ml tetracycline for Origami(DE3) and incubated with shaking at 200 rpm overnight at 37°C (14-16 hr). A 0.5 ml aliquot of the starter culture was subsequently transferred into 50 ml LB medium and shaken at 240-250 rpm at 37°C until the optical density at 600 nm (OD<sub>600</sub>) of the culture reached 0.4-0.6. The sterile bacterial culture was transferred into pre-cooled centrifuge tubes, placed on ice for 10 min and centrifuged at 3000 *g* for 10 min at 4°C. The cell pellets were resuspended with 10 ml of ice-cold calcium chloride solution (60 mM CaCl<sub>2</sub>, 10 mM PIPES, pH 7.0, 15% glycerol and sterile distilled water) gently and slowly until the suspension looked homogeneous. The cell suspensions were then centrifuged at 3000 *g* for 10 min at 4°C. The cell resuspension and centrifugation were repeated twice and the supernatant was discarded each time. Finally, an ice-cold calcium chloride solution (1.2 ml) containing 70  $\mu$ l dimethyl sulfoxide (DMSO) was added to the cell pellets, which were then resuspended and 100  $\mu$ l aliquots were stored in pre-cooled tubes at -80°C.

### 2.2.2 Isolation of recombinant plasmid by alkaline lysis

The recombinant plasmids were extracted by the alkaline lysis method (Sambrook *et al.*, 1989). A well formed colony was inoculated in 3 ml LB medium containing 50 µg/ml ampicillin and incubated at 37°C with shaking at 200 rpm for 9-14 hr. The cells were centrifuged at 5000 g at room temperature for 5 min and the medium discarded. The pellets were then be resuspended in 100 µl lysis buffer containing 0.05 M Tris-HCl, pH 8.0, 0.01 M EDTA, pH 8.0, and 0.05 M glucose. Two hundred microliters of freshly prepared 1% SDS/0.2 M NaOH solution was added to the cell lysate and it was incubated on ice for 3 min. One hundred and fifty microliters of ice cold 3 M potassium acetate, pH 4.8, was then added, the mixture was centrifuged at 14,000 g for 5 min at 4°C, and the supernatant transferred into a fresh tube. Four hundred microliters of 1:1 phenol/chloroform was added, inverted briefly and centrifuged at 14,000 g, at room temperature for 5 min. The supernatant was then transferred to a new tube and 900 µl of absolute ethanol was added and mixed thoroughly. The precipitated DNA was compacted by centrifugation at 12,000 g for 5 min. The pellets were washed with 70% ethanol, centrifuged at 14,000 g for 5 min, and air dried at 37°C for 15 to 20 min. The plasmid DNA pellet was resuspended in 50 µl TE buffer, pH 8.0, (10 mM Tris, 1 mM EDTA, pH 8.0) containing 1 µl of 2 mg/ml RNase A to degrade the RNA and incubated at 37°C for 15 min. Seventy microliters of ice cold 20% PEG 6000 solution containing 2.5 M NaCl was added and the mixture incubated for 1 hr on ice and then centrifuged at 14,000 g for 20 min at 4°C. The supernatant was carefully removed and the translucent pellet was rinsed with 500 µl of 70% ethanol and centrifuged at 14,000 g for 3 min at 4°C. Ethanol was removed

carefully and the DNA pellet dried at 37°C for 15 min and finally dissolved in TE buffer, pH 8.0, and stored at -20°C until further use.

### **2.2.3 Agarose gel electrophoresis**

The DNA bands were analyzed on a 1% agarose gel electrophoresis as described by Sambrook *et al.* (1989). One gram agarose was melted in 100 ml of 1X TAE buffer (0.04 M Tris-HCl, pH 8.0, 0.04 M acetic acid, 0.001 M EDTA, pH 8.0) by warming in a microwave oven for 2 to 3 minutes. The agarose gel was set on a gel tray supplied with Pharmacia gel electrophoresis apparatus GNA-300 (GE Healthcare Pharmacia Biotech, San Francisco, CA). The DNA sample was mixed with 5:1 6X loading dye (0.25% bromophenol blue, 0.25% xylene cyanol FF and 30% glycerol) and electrophoresis was conducted for 45 min at 90 V in 1X TAE buffer. The sample was then stained with 0.1 µg/ml ethidium bromide solution for 5 min and destained with distilled water. The gel was then visualized with the Flour-S<sup>TM</sup> Multimager UV light transilluminator (Bio-RAD Laboratories). The size of the DNA was compared with standard 1kB DNA ladder.

### **2.2.4 Purification of TEV protease**

Tobacco etch virus (TEV) protease was produced as described by Waugh *et al.* (2001). BL21-RIL cells containing pRK793 plasmids were grown at 37°C in LB agar containing 100 µg/ml ampicillin and 30 µg/ml chloramphenicol. A final concentration of 1 mM IPTG was added and the cells were induced at 30°C for 4 to 5 hr. The cell pellets were dissolved in lysis buffer (50 mM phosphate, pH 8.0, 100 mM sodium chloride, 10% glycerol and 25 mM imidazole). Five percent polyetheleneimine was added to a final concentration of 0.1% and the extract was loaded on a Ni-NTA column equilibrated with lysis buffer. TEV protease was eluted

in lysis buffer containing 200 mM imidazole. EDTA and DTT at final concentrations of 1 mM were added and TEV protease was purified further by gel filtration chromatography on a S-100 column equilibrated with 25 mM phosphate, pH 8.0, 200 mM sodium chloride, 2 mM EDTA, 10% glycerol and 10 mM DTT. The purified TEV protease was concentrated and stored at -80°C.

### **2.2.5 Os3BGlu6 activity assay**

The routine  $\beta$ -glucosidase activity assay was done with *p*-nitrophenyl- $\beta$ -D-glucopyranoside (*p*NPGlc) as the substrate in a reaction volume of 140  $\mu$ l, containing 60  $\mu$ l 100 mM sodium acetate, pH 5.0, 70  $\mu$ l of 2 mM *p*NPGlc and 10  $\mu$ l of the enzyme solution. The reaction was incubated on a microtiter plate for 10 min at 30°C. The reaction was stopped by the addition of 100  $\mu$ l of 2 M sodium carbonate. The amount of the product *p*-nitrophenol (*p*NP) released was determined by measuring the absorbance of *p*-nitrophenolate at 405 nm with an iEMS Reader MF microplate photometer (Labsystems iEMS Reader MF, Finland) and comparing it to that of a standard curve of 0 to 30 mM *p*NP in the same buffers.

### **2.2.6 Bio-RAD Protein Assay**

The concentration of purified protein was determined with the Bio-RAD protein assay kit (Hercules, CA, USA) using bovine serum albumin (BSA) as the standard (0-12  $\mu$ g BSA). The reaction volume was 1 ml, which contained 0.1 ml of diluted protein and 0.2 ml Bio-RAD protein assay solution. The rest of the volume was made up with deionized water. The reaction was mixed well and incubated for 10 min at room temperature. The absorbance at 595 nm was determined with a Genesys 10 UV spectrophotometer (Genesys, Rochester, NY, USA).

### 2.2.7 Determination of protein expression profiles by SDS-PAGE

The protein profile and the apparent molecular weight were determined by SDS-PAGE as described by Laemmli (1970). The SDS-PAGE 12% separating gel consisted of 12% (w/v) acrylamide, 0.375 M Tris-HCl, pH 8.8, 0.1% SDS, 0.05% ammonium persulfate and 0.05% TEMED and the 4% stacking gel consisted of 4% w/v acrylamide, 0.125 M Tris-HCl, pH 6.8, 0.1% SDS, 0.05% ammonium persulfate and 0.05% TEMED. Protein samples were mixed with 1X loading buffer (0.05 M Tris-HCl, pH 6.8, 10% SDS, 0.2 mg/ml bromophenol blue, 50% glycerol, 20% 2-mercaptoethanol) and boiled for 5 min to denature proteins. Aliquots of 8-12  $\mu$ l were loaded into sample wells, and electrophoresed at 120 V in 1X Tris-glycine electrode buffer (0.05 M Tris base, 0.125 M glycine and 0.1% SDS, pH 8.3) until the dye front reached the bottom of the gel plate. The gels were subsequently stained in staining solution containing 0.1% (w/v) Coomassie Brilliant Blue R250, 40% (v/v) methanol, 10% (v/v) acetic acid for 30 min and destained with destaining solution [40% (v/v) methanol and 10% (v/v) acetic acid] for 1-2 hr. The molecular masses of protein bands were determined by comparing to standard low molecular weight protein markers (GE Healthcare, Uppsala, Sweden), which consist of phosphorylase b (97.4 kDa), bovine serum albumin (66 kDa), ovalbumin (45 kDa), bovine carbonic anhydrase (31 kDa), trypsin inhibitor (21.5 kDa) and bovine  $\alpha$ -lactalbumin (14.0 kDa).

## 2.3 Cloning and Expression of Os3BGlu6

### 2.3.1 Cloning and PCR amplification of cDNA encoding mature Os3BGlu6

The full length cDNA encoding rice *Os3BGlu6* from a Sapporo cold treated library clone provided by Dr. Takashi Akiyama of the National Agricultural Research Center for the Hokkaido Region, Japan, was used as a template in PCR to amplify the cDNA encoding the predicted mature Os3BGlu6 rice  $\beta$ -glucosidase with an PCR 9700 thermocycler (PE Applied Biosystems, Foster City, CA). Amplification was performed with the specific primers SapporoMstrtF: CACCGCGCAGCAGAGCGGAG and SapporoStopR: GGAGTTCAGGTCTTCAGGAG by increasing the temperature to 94°C for 1 min followed by 30 cycles of 94°C for 1 min, 60°C for 45 s, 72°C for 1 min 30 s, with *Pfu* DNA polymerase. The PCR product was analyzed on a 1% agarose gel. The 1500 base pair PCR product was then purified with the Perfectprep Gel Cleanup Kit and was cloned into pENTR™/TEV/D-TOPO® Gateway® entry vector, according to the supplier's directions (Invitrogen). This vector has a TEV protease site and attL recombination sites for efficient transfer into Gateway® destination vectors. A molar ratio of 2:1 PCR product to TOPO vector was used for cloning. The cloning reaction was allowed to proceed for 5-10 min at 22-23°C and stored at -20°C overnight and then transformed into One Shot® chemically Competent Mach 1 *E.coli* by heat shock at 42°C (30 s) and selected on LB agar containing 50  $\mu$ g/mL kanamycin. The plasmids from selected clones were isolated by the alkaline lysis method (methodology described in section 2.2.2) and the presence of insert in the recombinant plasmid was confirmed by digesting with a specific restriction enzyme

(*EcoR* I), by PCR amplification with the SapporoMstrtF and SapporoStopR primers and by automated DNA sequencing at Macrogen Corp, (Seoul, Korea). The cDNA was subcloned into pET32a(+)/DEST expression vector (Opassiri *et al.*, 2006) by LR Clonase™ (Invitrogen) recombination and transformed into 70 µl of competent *E. coli* (DH5α) cells by heat shock at 42°C for 40 s, and the cells were spread on LB agar containing 50 µg/ml ampicillin. The recombinant pET32a(+)/DEST/Os3BGlu6 expression clone was verified by automated DNA sequencing with the T7 forward and reverse and the specific internal primers shown in Table 2.1.

**Table 2.1** Oligonucleotide primers used for sequencing of recombinant pET32a(+)/DEST/Os3BGlu6.

Primers	Primer Sequence	Tm (°C)
F427	CTTGACAGGCAGATAGTGGA	60
R600	GCAGTAGAGGTGGAGCAG	58
F969	GCACAACAACACCAACATC	56
R946	TCGTGTAGTAGG TGG TGTAGTG	66
T7 forward primer	TAATACGACTCACTATAGGG	-
T7 reverse primer	GCTAGTTATTGCTCAGCGG	-

### 2.3.2 Expression of Os3BGlu6 in *E. coli* Origami(DE3)

The pET32a(+)/DEST/Os3BGlu6 recombinant plasmid was transformed into 70 µl of competent *E. coli* strain Origami(DE3) cells by heat shock at 42°C for 40 s. Then 250 µl of LB medium was added and the cells were shaken at 37°C for one hour. The medium containing the transformed bacteria was spread on an

LB agar plate containing 50 µg/mL ampicillin, 15 µg/ml kanamycin and 12.5 µg/ml tetracycline.

Os3BGlu6 was expressed in *E. coli* strain Origami(DE3) as a fusion protein with N-terminal thioredoxin and His<sub>6</sub> tags. For recombinant Os3BGlu6 protein expression, a freshly transformed colony was grown in 5 ml LB medium containing 50 µg/ml ampicillin, 15 µg/ml kanamycin and 12.5 µg/ml tetracycline at 37°C, shaken at 200 rpm for 14-16 hr. The starter culture was diluted to a 1:100 ratio with LB medium containing the same antibiotics and grown at 37°C with shaking at 200 rpm for 3-5 hr, until an OD (at 600 nm) of 0.5-0.6 was obtained.

Optimal expression conditions for recombinant Os3BGlu6 β-glucosidase were screened by varying temperature, time and IPTG concentration. The conditions were as follows: time after IPTG induction: 0, 4, 8, 16, and 18 hr, temperature: 15°C, 20°C, 25°C, 30°C, and 37°C, and IPTG concentration: 0 mM, 0.1 mM, 0.2 mM, 0.3 mM, 0.4 mM, and 0.5 mM. All the variations were tested with 25 ml LB cultures. Aliquots of cell pellets were run on SDS-PAGE to observe the expression levels of the recombinant protein. Enzyme activity in the crude soluble fractions was assayed with *p*NPGlc as the substrate (Section 2.2.5).

For large scale expression, the induced cultures were transferred to 50 ml pre cooled centrifuge tubes and chilled on ice for 10 min, centrifuged at 5000 *g* for 15 min at 4°C. The weights of the cell pellets were noted and the pellets were stored at -80°C.

### **2.3.3 Os3BGlu6 recombinant protein extraction and purification**

The cell pellets were thawed, resuspended in freshly prepared extraction buffer (20 mM Tris-HCl, pH 8.0, 300 mM sodium chloride, 200 µg/ml

lysozyme, 1% Triton-X 100, 1 mM phenylmethylsulfonylfluoride (PMSF), 0.1 mg/ml soybean trypsin inhibitor and 0.25 mg/mL DNaseI, 5 ml/g of cell pellet) and incubated at room temperature for 30 min. Cells were lysed on ice by sonicating (Ultrasonic processor GE probe sonicator, Treadlitei, Woodstock, MA) for 15 s at 10 W output for 3-5 times with one minute cooling in between and the soluble crude protein was separated from the insoluble fraction by centrifugation (12,000 *g* at 4°C for 10 min). The soluble protein fractions were stored on ice for the next step of protein purification. Preliminary purification was done by immobilized metal affinity chromatography (IMAC) on immobilized Co<sup>2+</sup> affinity resin. Crude protein was added to IMAC resin (capacity of the resin: 40 mg protein/ml) pre-equilibrated with 10 column volumes (CV) of equilibration buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl), the IMAC column containing resin and crude protein was shaken for 30 min to 1 hr on ice to bind the His-tagged proteins to the resin. The resin/protein was centrifuged at 3,000 *g* for 2 min at 4°C, the supernatant removed, and the resin reloaded onto the column. The protein-bound resin was washed with 5 CV of equilibration buffer, 5 CV of wash buffer 1 (5 mM imidazole in equilibration buffer), and 5 CV volumes of wash buffer 2 (10 mM imidazole in equilibration buffer). The recombinant protein was finally eluted with elution buffer (250 mM imidazole in equilibration buffer).

The elution fractions containing *p*NPGlc hydrolysis activity were pooled and concentrated in a 30 kDa MW cutoff ultrafiltration membrane. The buffer was changed to 20 mM Tris-HCl, pH 8.0, in the same device. The N-terminal thioredoxin, His<sub>6</sub> and S-tags were then excised with tobacco etch virus (TEV) protease, produced in Section 2.2.4, at a ratio of 1:50 (w/w) TEV protease to fusion protein for 12 to 14 hr

at 4°C. The cleaved Os3BGlu6 protein was separated from the N-terminal fusion tag by adsorption of the tag and TEV protease to IMAC Co<sup>2+</sup> resin, as described in the preceding paragraph. The flow-through and wash fractions in 5 mM and 10 mM imidazole containing *p*NPGlc hydrolysis activity were pooled and concentrated with a 10 kDa MW cutoff ultrafiltration membrane in 20 mM Tris-HCl, pH 8.0.

## 2.4 Os3BGlu6 pH and temperature optima & stability studies

The optimum pH was determined by incubating 1 µg enzyme in a reaction volume of 140 µl containing 100 mM universal buffer (citric acid-disodium hydrogen phosphate), pH 2-11 at 0.5 pH unit increments with 1 mM *p*NPGlc for 10 minutes at 30°C. The reaction was stopped with 100 µl of 2 M sodium carbonate. The amount of *p*NP released was measured at 405 nm.

The enzyme's pH stability was studied by incubating 25 µg enzyme in 20 µl universal buffer at the same pH range for time periods of 10 min, 1, 2, 6, 12, and 24 h at 30°C. The enzyme was diluted 400 fold in 100 mM buffer at the optimum pH, and the activity was determined as described above.

To determine the temperature optimum, 1 µg enzyme was incubated in 100 mM sodium acetate, pH 5.0, over a temperature range of 5 to 100°C in 5° increments for 10 min and then *p*NPGlc was added at 1 mM final concentration to the reaction and incubated for 10 min. The amount of *p*NP released was determined as described above.

Temperature stability was determined by incubation of 1 µg enzyme in 100 mM sodium acetate, pH 5.0, over a temperature range of 5-70°C. At 10 min increments from 0-60 min, aliquots of sample containing 1 µg enzyme incubated at

each temperature were placed on ice and the enzyme activity was determined at 30°C by addition of *p*NP<sub>Glc</sub> to a 1 mM final concentration and measurement of the release of *p*NP in 10 min.

## 2.5 Os3BGlu6 substrate specificity and enzyme kinetics

Os3BGlu6 was tested for hydrolysis of glycosides of *p*NP, oligosaccharides, and commercially available natural and synthetic glycosides (listed in Tables 3.1 and 3.2) by Thin Layer Chromatography (TLC) measurement and of release of *p*NP or glucose from the glycosides and oligosaccharides. For TLC, 2 µg of purified enzyme was mixed with 5 mM concentrations of various substrates in 50 mM sodium acetate, pH 5.0, in a reaction volume of 70 µl and incubated at 30°C overnight. The reactions were stopped by boiling 5 min, and 2-5 µl of products from the reactions were spotted on a silica gel-coated F<sub>254</sub> aluminium plates, (Merck) and chromatographed vertically with solvents of 2:1:1 ethyl acetate:acetic acid:water for the oligosaccharides and 15:2:1:2 ethyl acetate:acetic acid:methanol:water for natural glycosides. Glucose was used as the standard. The products were visualized under UV light at 366 nm, and then by coating the plates with 10% (v/v) sulfuric acid in methanol and heating at 120°C until the spots were visible. The release of *p*NP from *p*NP glycosides was quantified as described in Section 2.2.5. Activity was assayed in triplicate in 100 mM sodium acetate, pH 5.0, at 30°C. For those substrates that were seen to be hydrolyzed on TLC, glucose release was measured by the glucose oxidase peroxidase (PGO) coupled assay. One capsule of PGO enzyme (Sigma) containing 500 units of glucose oxidase peroxidase (*Aspergillus niger*) was dissolved in 100 ml distilled water. The reactions containing 100 mM sodium acetate, pH 5.0, 5 mM of different substrates (glucosides

or 1 mM standard glucose) and 25  $\mu$ l of 1  $\mu$ g of the enzyme were incubated for 30 min at 30°C and stopped by boiling for 5 min at 100°C. Each reaction was then transferred into a microtiter plate containing 50  $\mu$ l of 1 mg/ml 2,2'-azinobis(3-ethylbenzthiazolinesulfonic acid) (ABTS), and 100  $\mu$ l of PGO enzyme was added just before incubating at 37°C for 30 minutes. The release of glucose was determined by measuring the absorbance of the reaction at 405 nm.

Kinetic parameters were determined for those substrates that showed significant rates of hydrolysis in the initial relative activity assays. The reactions were performed in triplicates at 30°C and the reaction time was set the moment enzyme was added. The initial velocity ( $V_0$ ) 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  $\mu$ g 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 1/5 to 5 times the apparent  $K_m$  were included. Michaelis-Menton and linear reciprocal plots (Hanes and Lineweaver-Burk plots) were obtained to determine the  $K_m$  and  $V_{max}$ . The kinetic parameters, including  $k_{cat}$ ,  $K_m$ , and  $k_{cat}/K_m$  were calculated by nonlinear regression of Michaelis-Menton plots with the Grafit 5.0 computer program (Erithacus Software, Horley, UK, 2001).

## 2.6 Tris Inhibition

Competitive  $K_i$  values were determined by incubating 1  $\mu$ g enzyme with eight different concentrations of the inhibitors (0 to 30 mM) in 100 mM sodium acetate, pH 5.0, in presence of 1, 2, 5, 10, 15, and 20 mM *p*NPGlc substrate under the reaction

conditions described above. Lineweaver-Burk and Kmapp/Vmaxapp Vs Tris concentration plots were used to calculate the inhibition constants.

## **2.7 Os3BGlu6 Crystallization**

### **2.7.1 Preparation of Os3BGlu6 protein for crystallization**

Purified Os3BGlu6 (10 mg/ml) in 20 mM Tris-HCl, pH 8.0, was used to screen for crystallization conditions by the microbatch method. Prior to screening, the protein solution was filtered with an Ultrafree MC 0.22  $\mu\text{m}$  filter by centrifugation at 3000  $g$  at 4°C for 2-5 min, to eliminate any dust particles and particulate matter that might interfere with the crystallization process. The filtered protein was stored at 4°C until crystallization trials began. The precipitant solutions were prepared with HPLC grade water, filtered with a 0.45  $\mu\text{m}$  filter and stored at 15°C.

### **2.7.2 Preliminary screening by micro-batch technique**

Preliminary screening by the microbatch method was conducted under 100% paraffin oil using 60 well microbatch plastic plates (Nunc, Denmark), with the following crystallization kits from the lab.

**Table 2.2** Screening kits for Microbatch method.

	Screening Kits	Company
1	Crystals Screen High Throughput HR2-130 and HR2-134	Hampton Research, Aliso Viejo, CA, USA
2	Wizard I <sup>TM</sup> and II	Emerald Biosystems, Seattle, WA, USA
3	JY Screen of salts in PEG	Mahidol University, Bangkok, Thailand

Ten microliters of 100% paraffin oil was pipetted into each well of the plate and 0.5  $\mu$ L precipitant solution was added. Finally, a 0.5 to 1  $\mu$ l of 10 mg/ml pure protein solution was dispensed under the oil with the precipitant drop and all the components were mixed. Since very small volumes of protein, precipitants and buffers were used, they were covered with oil to prevent any possible evaporation. Oil also protects protein from oxidation and any airborne contaminants. The microbatch plates were carefully placed on a moist sponge in a plastic box and equilibrated at 15°C and the quality and growth of the crystals were monitored by observing the drops under a Zeiss Stemi 2000-C stereo microscope (Zeiss Corp, NJ, USA) on a daily basis for a month and then weekly until 200 days and recorded. Once crystals were obtained from the microbatch method, the positive conditions in which the crystals were obtained were reproduced.

### 2.7.3 Optimization of crystals by hanging drop

To obtain well-formed single crystals, optimization was conducted by the hanging drop method in sterile 24 well plates (Greiner Bio-One, Frickenhausen, Germany), by varying the protein concentration from 1-10 mg/ml, protein to

precipitant volume ratio at 1:1 and 2:1 PEG precipitant concentrations from 11% to 24%, pH range of the buffer from 5.5 to 6.9 and type of PEG to obtain larger single crystals for x-ray diffraction. High vacuum grease was applied to the edges of each well of the TC plate, and 0.5 ml of reservoir solution containing appropriate amounts of precipitant was added each well. Then, 1  $\mu$ l of precipitant solution and 1 or 2  $\mu$ l of pure protein were pipetted into a mixed drop on a cover slip siliconized with dichlorodimethylsilane. The cover slip was carefully inverted, so that the crystallization drop faced the reservoir solution and the cover slip was sealed to the grease layer. This was further equilibrated against the reservoir solution. The plates were labeled and equilibrated in a 15°C incubator and the crystallization drops were monitored by observing them under the stereomicroscope every two days and the observations recorded.

The first step of optimization included a preliminary variation of the percentage of polyethylene glycol (PEG 5000 MME and 0.1 M Bis Tris pH 6.5) with 1:1 and 2:1 ratios of protein to precipitant, which included the condition in which the crystals were seen in the microbatch method (Table 2.3). Further optimization was conducted by varying the concentration of protein (2, 4, 6, 8, and 10 mg/ml on the same cover slide) with a wider range of PEG 5000 MME concentrations (Table 2.4). In addition, optimization was conducted with different types and concentrations of PEG (Table 2.5) and with different pH values of 0.1 M Bis Tris buffer and different percentages of PEG 5000 MME with 4 and 8 mg/ml protein (Table 2.6).

**Table 2.3** Optimization of Os3BGlu6 crystals in the condition 17 to 22% PEG 5000 MME and 0.1 M Bis Tris, pH 6.5.

17%	18%	19%	20%	21%	22%
PEG-204	PEG -216	PEG – 228	PEG – 240	PEG–252	PEG – 264
Buffer-60	Buffer-60	Buffer-60	Buffer-60	Buffer-60	Buffer-60
Water-336	Water-324	Water– 312	Water– 300	Water-288	Water– 276

The numbers indicate the volume of the precipitants in the reservoir in  $\mu\text{l}$ .

PEG- PEG 5000 MME

Buffer -0.1 M Bis Tris, pH 6.5

**Table 2.4** Optimization of Os3BGlu6 crystals with different concentrations of protein.

11%	12%	13%	14%	15%	16%
PEG – 132	PEG – 144	PEG – 156	PEG –168	PEG – 180	PEG – 192
Buffer-60	Buffer-60	Buffer-60	Buffer-60	Buffer-60	Buffer-60
Water –408	Water –396	Water– 384	Water– 372	Water–360	Water– 348
17%	18%	19%	20%	21%	22%
PEG – 204	PEG – 216	PEG – 228	PEG – 240	PEG – 252	PEG – 264
Buffer-60	Buffer-60	Buffer-60	Buffer-60	Buffer-60	Buffer-60
Water– 336	Water– 324	Water– 312	Water– 300	Water–288	Water– 276

The numbers indicate the volume of the precipitants in the reservoir in  $\mu\text{l}$ .

PEG- PEG 5000 MME

Buffer -0.1 M Bis Tris, pH 6.5

Os3BGlu6 crystallization was tested in 11 to 22% PEG 5000 MME and 0.1 M Bis Tris, pH 6.5, with different concentrations 2, 4, 6, 8, and 10 mg/ml of protein.

**Table 2.5** Optimization of Os3BGlu6 crystals with different types of PEG.

Type of PEG	13%	15%	17%	18%	20%	22%
PEG 400	PEG – 78	PEG – 90	PEG –102	PEG –108	PEG – 120	PEG–132
	Buffer-60	Buffer–60	Buffer–60	Buffer–60	Buffer – 60	Buffer–60
	Water-462	Water-450	Water-438	Water-432	Water– 420	Water-408
PEG 2000	PEG– 156	PEG– 180	PEG– 204	PEG– 216	PEG – 240	PEG–264
	Buffer–60	Buffer–60	Buffer–60	Buffer–60	Buffer – 60	Buffer–60
	Water-384	Water-360	Water-336	Water-324	Water– 300	Water-276
PEG 3350	PEG– 156	PEG– 180	PEG– 204	PEG– 216	PEG – 240	PEG–264
	Buffer–60	Buffer–60	Buffer–60	Buffer–60	Buffer – 60	Buffer–60
	Water-384	Water-360	Water-336	Water-324	Water– 300	Water-276
PEG 4000	PEG– 156	PEG– 180	PEG– 204	PEG– 216	PEG – 240	PEG–264
	Buffer–60	Buffer–60	Buffer–60	Buffer–60	Buffer – 60	Buffer–60
	Water-384	Water-360	Water-336	Water-324	Water– 300	Water-276

The numbers indicate the volume of the precipitants in the reservoir in  $\mu\text{l}$ .  
Buffer -0.1 M Bis Tris, pH 6.5

Os3BGlu6 optimized in 13 to 22% PEG 5000 MME and 0.1 M Bis Tris, pH 6.5, with different types of PEG with 4 and 8 mg/ml protein.

**Table 2.6** Optimization of Os3BGlu6 crystals with different pH and PEG 5000 MME concentrations.

%	5.5	6.2	6.4	6.5	6.7	6.9
PEG5000 MME						
13%	PEG– 156	PEG– 156	PEG– 156	PEG– 156	PEG- 156	PEG–156
	buffer – 80	buffer–80	buffer–80	buffer–80	buffer–80	buffer–80
	water- 384	water-384	water-384	water-384	water-384	water-384
15%	PEG – 180	PEG– 180	PEG– 180	PEG– 180	PEG– 180	PEG–180
	buffer – 60	buffer–60	buffer–60	buffer– 60	buffer– 60	buffer–60
	water– 360	water-360	water–360	water-360	water-360	water-360
19%	PEG– 228	PEG– 228	PEG– 228	PEG– 228	PEG– 228	PEG–228
	buffer – 60	buffer –60	buffer –60	buffer –60	buffer –60	buffer–60
	water – 312	water–312	water-312	water–312	water-312	water-312
22%	PEG– 264	PEG– 264	PEG– 264	PEG– 264	PEG– 264	PEG–264
	buffer – 60	buffer–60	buffer–60	buffer–60	buffer- 60	buffer–60
	water– 276	water–276	water–276	water- 276	water-276	water-276

The numbers indicate the volume of the precipitants in the reservoir in  $\mu\text{l}$ .  
PEG- PEG 5000 MME  
Buffer -0.1 M Bis Tris

Optimization of Os3BGlu6 crystallization with 13 to 22% PEG 5000 MME and 0.1 M Bis Tris in the pH range 5.5 to 6.9 with 4 and 8 mg/ml protein.

#### **2.7.4 Microseeding by the streak seeding method**

The quality of the crystals was improved by microseeding with crushed Os3BGlu6 crystal clusters. The clusters of Os3BGlu6 crystals obtained in 21% PEG5000 MME, 0.1 M Bis Tris, pH 6.5, were picked in a nylon loop, and washed 3 to 4 times in freshly prepared 100  $\mu$ l mother liquor containing the same precipitant solution in which the clusters were obtained. The clusters were then crushed with a cat whisker and the microcrystal stock was then stored at 4°C. The stock solution was further diluted 1:100 and 1:1000 fold for seeding. The diluted seeding stocks were then streaked with a cat whisker into hanging drops (Table 2.6) that had been preequilibrated for 2 hr at 15°C and further incubated at the same temperature.

#### **2.7.5 Soaking of crystals with Inhibitors**

To obtain Os3BGlu6 crystals with inhibitors, native crystals obtained by microseeding were picked from the drop with a nylon loop and washed in freshly prepared precipitant solution. The washed crystal was then soaked in 1, 2 or 5 mM 2-deoxy-2-fluoro- $\beta$ -D-glucoside (DNPG2F) or 2, 5 or 10 mM *n*-octyl- $\beta$ -D-thioglucopyranoside (prepared in the precipitant solution) overnight at 15°C in a sitting drop plate (Hampton Research). The reservoir contained the same precipitant solution as was added to the drop. The wells were sealed with high vacuum grease to a cover slip.

#### **2.7.6 Cryofreezing of crystals**

Prior to transfer of the crystals to the dry shipper for the transfer to the data collection location, the crystals were picked with the nylon loops of the

appropriate size, which had been premounted to cryo screw caps or magnetic caps (Hampton Research), and soaked in the cryoprotectant solution (containing each crystallization precipitant component at concentrations increased by 15% of its original concentration and 18% (v/v) glycerol), for 5-10 s just before flash freezing with liquid nitrogen (-180°C). For crystals with inhibitors, the cryoprotectant contained the inhibitors at the same concentrations in which the crystals had been soaked. The loop along with the crystal was then positioned into the cryovial, labeled and transferred to the cane and placed vertically inside the dry shipper containing liquid nitrogen.

### **2.7.7 Data collection, processing and structure refinement**

The datasets for native Os3BGlu6, the Os3BGlu6/G2F complex and the Os3BGlu6/*n*-octyl- $\beta$ -D-thio-glucopyranoside complex were collected on the BL13B1 beamline at the National Synchrotron Radiation Research Center, Hsinchu, Taiwan with a 1.0 Å wavelength x-ray beam and an ADSC Quantum 315 CCD detector. The crystals were maintained at 110K in a cold stream of nitrogen from an Oxford Instrument CryoJet throughout diffraction. The crystal was mounted on the goniometer head and positioned in the center of the x-ray beam, so that the crystal could rotate while retaining the centering. All datasets were indexed, integrated and scaled and with the HKL 2000 package (Otwinowski & Minor, 1997). Prior to data collection, the distance of the detector was set in the range of 160-300 nm. The data was collected over 180° rotation with 0.3 to 0.5° oscillations. The exposure time was set from 5-15 s. The maximum resolution limit which the crystal could be diffracted was judged based on the ratio of the intensity measured to the standard deviation, i.e.  $I/\sigma$ , which was kept  $>2$  in the outer shell, and the  $R_{merge}$  for the outer shell which was

generally kept at <40%. In addition, the overall quality of the data was assessed by the dataset completeness, average redundancy per shell and  $R_{merge}$ .

The crystal structure for native Os3BGlu6 was solved by molecular replacement. The Os3BGlu6 protein sequence was submitted to the GENO3D server to search for a phasing model from the database, which gave *Trifolium repens* L., cyanogenic  $\beta$ -glucosidase as the structural model (PDB code 1CBG, Barrett *et al.*, 1995). 1CBG shared approximately 54% sequence identity to Os3BGlu6. The model obtained from the Geno3D server was used in the Molrep program in CCP4 suite (Vagin & Teplyakov, 1997). A subset of 5% of the structure factor amplitudes were reserved for the free R-factor determination. The structures of Os3BGlu6/G2F complex and Os3BGlu6/*n*-octyl- $\beta$ -D-thioglucopyranoside were solved by molecular replacement with native Os3BGlu6 structure as the template model. Restrained refinement was performed for all the three structures with Refmac 5.0 (Murshudov *et al.*, 1999; CCP4 Suite) and the convergence values of the R and  $R_{free}$  factors were observed during refinement process. The R factor gave the measure of agreement between the calculated and observed structure factors, while  $R_{free}$  indicated the agreement between these values for reflections not used in the refinement, so it was a less biased assessment of model quality and improvements during refinement.

### 2.7.8 Model building and structure validation

The structure was built into the electron density map with the coot graphics program (Emsley & Cowtan, 2004), alternating with refinement with Refmac 5.0. The two Fourier maps  $2F_o-F_c$  and  $F_o-F_c$  were calculated after each round of refinement for rebuilding. Water molecules were added with Arp/wArp program (CCP4 suite) with 10 cycles of refinement for each cycle of waters added. At the final

stages of refinement, the ligands and water molecules were added or deleted based on the electron density ( $2F_o-F_c$  map sculpted at 1.5 sigma), B-factors and hydrogen bonding distances between the water molecules and the neighboring amino acid residues. The active site Tris ligand was built into the  $F_o-F_c$  density map in the native structure in two possible conformations, each, with an occupancy of 0.5. The 2-fluoroglucoside (G2F) inhibitor in the standard  ${}^4C_1$  chair conformation was fit to the electron density map of Os3BGlu6/G2F structure with an occupancy of 1. For the structure of Os3BGlu6 with the *n*-octyl- $\beta$ -D-thioglucopyranoside inhibitor, both the  ${}^4C_1$  chair and  ${}^1S_3$  skew boat conformations of the glucosyl moiety were separately built into the  $F_o-F_c$  density map. The *n*-octyl chain was then built manually with addition of single carbon atoms one at a time with occupancy of 1, and refinement of the structure after each addition.

The overall quality of the models was also evaluated based on the chemical reasonability with the program PROCHECK (Vaguine *et al.*, 1999). The dihedral  $\Phi$  and  $\Psi$  angles for each of the amino acid residues were plotted on a square matrix (Ramachandran plot). The root mean square deviations (R.M.S.) for bond lengths and bond angles were determined for each structure by PROCHECK. The criteria for any good quality model is that R.M.S. deviations of not more than 0.002 Å for bond length and 4° for bond angle deviations are acceptable. In addition, overall structural comparisons between Os3BGlu6 and other known structures of GH1  $\beta$ -glycosidases were executed with the Superpose program of the CCP4 suite (Krissinel and Henrick, 2004). Figures representing the structures were generated with Pymol (Delano Scientific).

### **2.7.9 N-terminal amino acid sequencing**

Automated Edman degradation was performed to confirm the N-terminus of the Os3BGlu6 protein in the crystals. Os3BGlu6 crystal clusters (containing 100 nmol protein) were dissolved in 1X SDS buffer and separated on a 12% polyacrylamide gel (as described in section 2.2.7). The gel was subsequently electroblotted onto a Biotrace™ polyvinylidene fluoride (PVDF) membrane (Pall Corporation, East Hills, NY, USA) of pore size 0.45  $\mu\text{m}$ . The membrane was prewet with 100% methanol for 10 s and immersed in HPLC grade water. Electro-blotting was conducted with 100 mM CAPS, buffer, pH 11.0, in 10% methanol. The blotted membrane was rinsed with HPLC water, to remove left over Tris and glycine from the SDS PAGE, and stained with Coomassie Brilliant Blue R250. The developed protein bands were excised, and Edman degradation sequencing was performed at the Biomolecular Unit, Scientific Equipment Center, Prince of Songkla University (Hat Yai, Thailand).

## CHAPTER III

### RESULTS

#### 3.1 Cloning and expression of *Os3BGlu6*

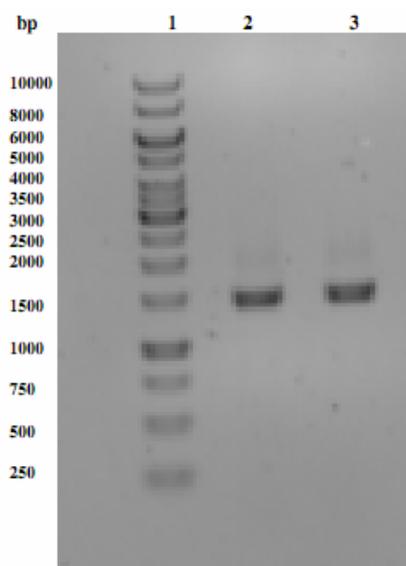
The cDNA encoding the *Os3BGlu6* protein was isolated from a cold-treated rice library, and its sequence determined by Dr. Takashi Akiyama of the National Agricultural Research Center for the Hokkaido Region, Japan (Genbank accession number AY129294). This full length cDNA contained an open reading frame of 1563 nucleotides encoding 521 amino acids, which were predicted to include an N-terminal secretory signal peptide of either 31 (neural net prediction) or 38 (hidden Markov model, HMM, prediction) residues and a mature protein of 484 or 477 residues, respectively, by SignalP (Bendtsen *et al.*, 2004). The full length *Os3BGlu6* cDNA, the primers used for PCR amplification and cDNA sequencing and the derived amino acid sequence derived are shown in Figure 3.1.

ATGGGGAGGATAAAGAGTAGTAGTGGGCGATGCAGCACGGCGAGGCTGGAGGCGGTCCG  
 M G R I K S S S G R C S T A R L E A V A  
 VTCTCTCGTCTGCTCTTCCGGCGTCGCGTCTGCTGCGAGGATGCATTGCCGAGCAG  
 V L V V V F G V A S S <sup>S</sup> L R G C I A <sup>Q</sup>  
 SapporoMstrtF  
 AGCGGAGGAGGGCTAACCAGGGGCAGCTTCCCCGAGGGGTTCTGCTTCCGGCACCGCTCC  
 S G G G L T R G S F P E G F V F G T A S  
 GCCGCGTACCAGTACGAGGGAGCTGTGAAGGAGGACGGGAGAGGGCAGACCATCTGGGAC  
 A A Y Q Y E G A V K E D G R G Q T I W D  
 ACGTTCCGCGCACACCTTTGGAAAGATCACCGACTTCAGCAATGTGATGTTGCAGTTGAT  
 T F A H T F G K I T D F S N A D V A V D  
 CAGTACCACCGTTTCGAGGAGGATATACAGCTCATGGCAGACATGGGGATGGATGCGTAT  
 Q Y H R F E E D I Q L M A D M G M D A Y  
 CGGTTCTCGATAGCATGGTCAAGAATCTACCCAAATGGTGTGGTCAAGTCAATCAAGCT  
 R F S I A W S R I Y P N G V G Q V N Q A  
 GGATATCGACCACTACAACAAGCTGATCGATGCACCTTCTAGCAAAAGGAATTGAGCCATAT  
 G I D H Y N K L I D A L L A K G I Q P Y  
 Os3Bglu6F427  
 GTGACACTCTACCACTGGGACCTCCCCAGGCCCTTGAAGACAAGTACAAGGGCTGGCTT  
 V T L Y H W D L P Q A L E D K Y K G W L  
 CACAGGCAGATAGTGGACGATTTTCGCGCGTACGCGGAGACGTGCTTTCAGGGAGTTCCGGG  
 D R Q I V D D F A A Y A E T C F R E F G  
 GACAGGGTGAAGCACTGGATCACGCTCAACGAGCCGACACGGTGGCCATCCAGGGCTAC  
 D R V K H W I T L N E\* P H T V A I Q G Y  
 Os3Bglu6R600  
 GACGAGGGCTCCAGGCCCGCGCTGCTCCGTGCTGCTCCACCTCTACTGCAAGGCC  
 D A G L Q A P G R C S V L L H L Y C K A  
 GGCAACTCCGGCACCGAGCCCTACGTCGTCGCCACCACTTCATCTCGCCACGCCCGCC  
 G N S G T E P Y V V A H H F I L A H A  
 GCCGCCAGCATCTACAGGACAAAATACAAGGCGACGAGACGAGGCGAGCTTGGGATAGCG  
 A A S I Y R T K Y K A T Q N G Q L G I A  
 TTCGACGTGATGTGGTTCGAGCCGATGTCCAACACCACGATCGACATCGAGGCGGCCAAG  
 F D V M W F E P M S N T T I D I E A A K  
 AGAGCGCAGGAGTTTCAGCTAGGATGGTTTGCTGATCCGTTCTTCTTCGGCGACTACCCG  
 R A Q E F Q L G W F A D P F F F G D Y P  
 GCACGATGAGGGCGAGGGTGGGGGAGAGGCTGCCGAGGTTACGGCGGATGAGGCCGCC  
 A T M R A R V G E R L P R F T A D E A A  
 Os3Bglu6R946  
 GTCGTCAAGGGGGCGCTGATTTCGTCGGCATAAACCCTACACCCTACTACAGGAGG  
 V V K G A L D F V G I N H Y T T Y Y T R  
 Os3Bglu6F969  
 CACAACAACCAACATCATCGGGACATTGCTCAACAACACCTTGGCAGACACCGGCACC  
 H N N T N I I G T L L N N T L A D T G T  
 GTCAGCCTCCATTCAAGAATGGGAAGCAATTGGAGATAGGGCAAATTCGATATGGCTG  
 V S L P F K N G K P I G D R A N S I W L  
 TACATTGTGCCCGAGGGATGAGGAGCCTGATGAACTATGTCAAGGAAAGGTACAACAGC  
 Y I V P R G M R S L M N Y V K E R Y N S  
 CCACCAGTGTACATCACTGAAAACGGGATGGATGACAGCAACAACCCGTTTCATTTCCATC  
 P P V Y I T E\* N G M D D S N N P F I S I  
 AAGGACGCCCTCAAGGACAGCAAGAGGATCAAATACCACAATGACTACCTACCAATCTG  
 K D A L K D S K R I K Y H N D Y L T N L  
 GCTGCTCCATCAAGGAGGACGGGTGCGACGTACGTGGGTACTTCGCGTGGTCTCTGCTG  
 A A S I K E D G C D V R G Y F A W S L L  
 GACAACCTGGGAGTGGGCGCCGGATACTCCTCGAGATTCGGGCTCTACTTCGTGGACTAC  
 D N W E W A A G Y S S R F G L Y F V D Y  
 SapporoStopR  
 AAGGATAACCTCAAGAGATACCCCAAGAACTCGGTGCGAGTGGTTCGAGCCCTCTGAAG  
 K D N L K R Y P K N S V Q W F K A L L K  
 ACCTGAACTCCAGCTG  
 T\*

**Figure 3.1** The full-length cDNA sequence and deduced protein sequence of rice Os3Bglu6 with SapporoMstrtF and SapporoStopR Primers.

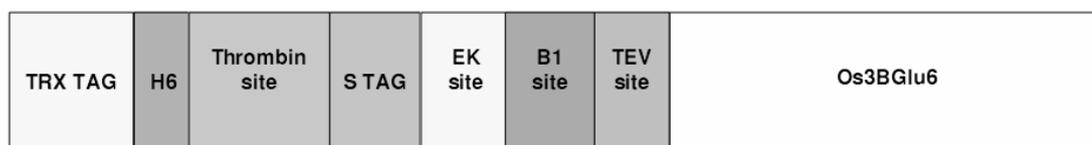
Underlined bold letters represent the regions of DNA sequence corresponding to primers PCR amplification and forward and reverse primers designed for DNA sequencing. Upward arrows represent predicted signal sequence cleavage sites. The stars indicate the catalytic residues.

The cDNA encoding the precursor residues 38 to 521 and to stop codon was amplified by PCR (Figure 3.2) cloned four codons behind the Tobacco-Etch Virus (TEV) protease site in the pENTR<sup>TM</sup>/TEV/D-TOPO<sup>®</sup> cloning vector. The cDNA insert was transferred to the pET32a(+)/DEST expression vector (Opassiri *et al.*, 2006), to produce an expression cartridge encoding an N-terminal tag containing thioredoxin and His<sub>6</sub>, followed by a linker peptide, including a thrombin site, an S-tag, an enterokinase site, a B1 recombination site, a TEV protease site and the Os3BGlu6  $\beta$ -glucosidase (Figure 3.3), which was confirmed by sequencing analysis. This allowed the Os3BGlu6  $\beta$ -glucosidase to be produced as a fusion protein in *E. coli* expression strain Origami(DE3).



**Figure 3.2** 1% agarose gel electrophoresis of PCR product of Os3BGlu6 cDNA after PCR amplification.

Lane 1, 1 kb DNA ladder (Bio-Rad); lanes 2 and 3, PCR product of Os3BGlu6.

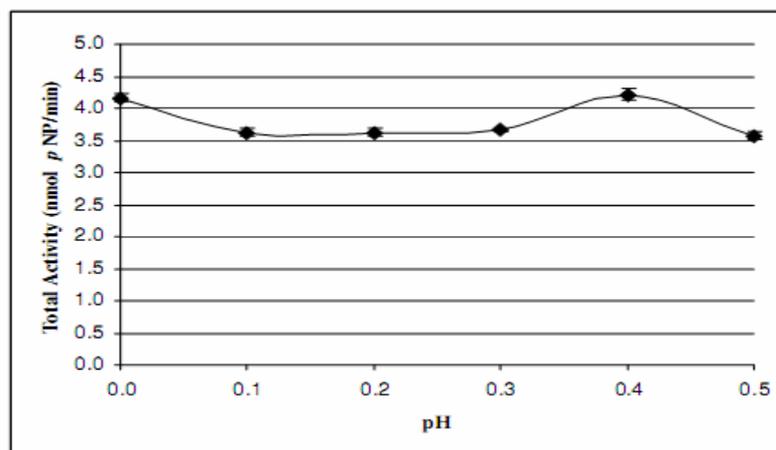


**Figure 3.3** Expression cartridge of the pET32a(+)/Os3BGlu6 plasmid.

The N-terminal tag containing thioredoxin (TRX) tag, a His<sub>6</sub> (H<sub>6</sub>) tag, a thrombin cleavage site, an S-tag, an enterokinase (EK) cleavage site, a B1 recombination site, a TEV protease cleavage site followed by the mature protein coding segment of the *Os3BGlu6* cDNA.

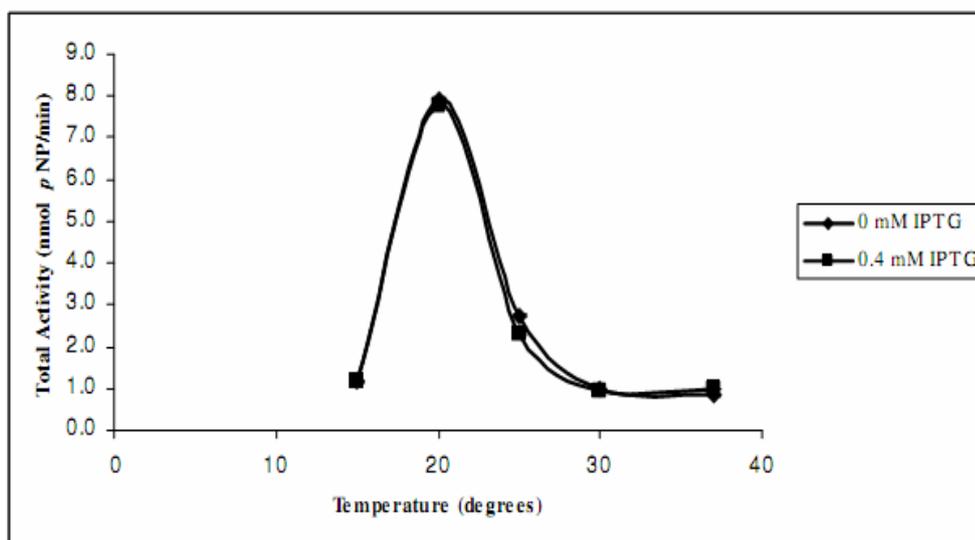
Os3BGlu6 was expressed as a soluble, active protein with the thioredoxin His<sub>6</sub> fusion tag in the *E. coli* strain Origami(DE3). Origami(DE3) is a redox-deficient strain, which allows for disulfide bond formation in cytoplasm. Optimization of the expression conditions yielded interesting results. At least 16 hrs was necessary for

optimal expression of this enzyme. After 16 hrs of incubation at 20°C, soluble extracts of induced cells with 0.4 mM IPTG and with no addition of IPTG had similar  $\beta$ -glucosidase activity (Figure 3.4). Comparison on the activities of Os3BGlu6 at 15°C, 20°C, 25°C, 30°C, and 37°C (Figure 3.5) at 0 and 0.4 mM IPTG indicated that Os3BGlu6 could be expressed at 20-25°C, even without addition of any IPTG inducer. Higher temperatures (30 and 37°C) did not favor production of higher active protein, while lower temperatures (20 and 25°C) favored production of the enzyme in the active form.



**Figure 3.4** Comparison of the  $\beta$ -glucosidase activity of Os3BGlu6 in *E. coli* extracts after induction at 20°C at different concentrations of IPTG.

The culture was harvested after expression for 16 hrs. One microgram of protein from the extracts was tested with pNPGlc in 100 mM sodium acetate, pH 5.0, for 10 min.

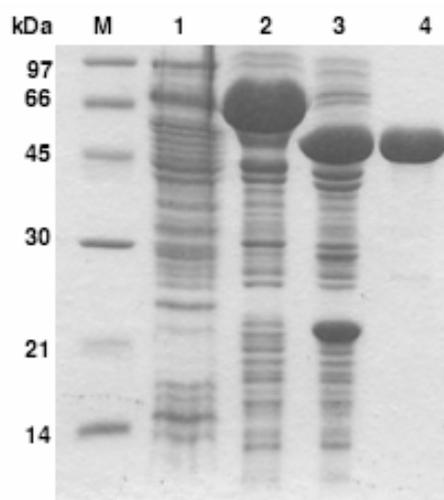


**Figure 3.5** The effect of temperature on the induction of Os3BGlu6.

Expression was induced at 15°C, 20°C, 25°C, 30°C and 37°C with 0 and 0.4 mM IPTG for 16 hrs. The soluble fractions (1 µg protein) obtained from different temperatures were tested with *p*NPGlc in 100 mM sodium acetate, pH 5.0, for 10 min.

### 3.2 Os3BGlu6 protein purification

Four liters of bacterial culture was used and the recombinant protein was purified as described in section 2.3.3. The recombinant protein was conveniently purified by binding the fusion protein to an immobilized metal affinity chromatography (IMAC) column, followed by release of the active Os3BGlu6 (tested against the hydrolysis of *p*NPGlc) from the tag by TEV protease digest and removal of the fusion tag and TEV protease by binding them to the IMAC column. A large scale purification of recombinant Os3BGlu6 was set up in order to obtain a higher amount of protein, which was required for crystallization techniques. The resulting soluble 55 kDa protein was active, >95% pure, based on SDS PAGE (Figure 3.6), and could be concentrated to a final concentration of 10-20 mg/ml.



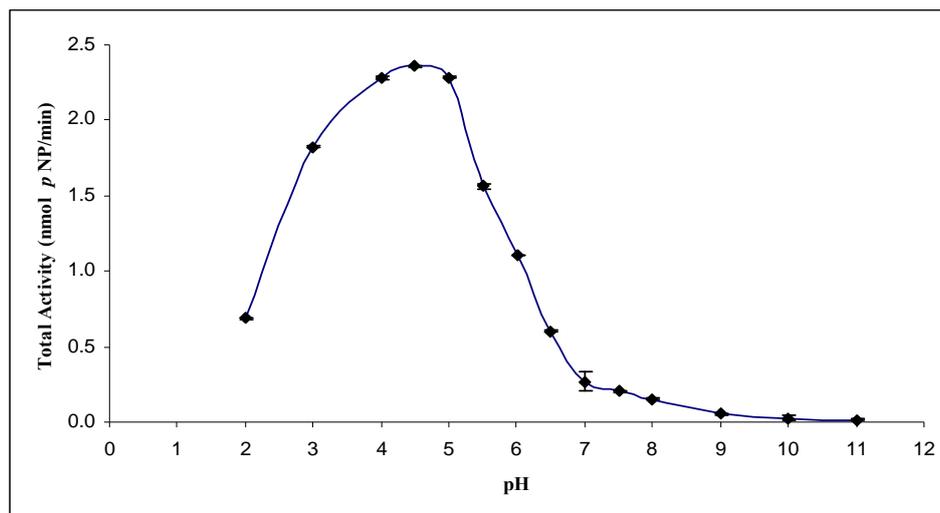
**Figure 3.6** SDS PAGE analysis of Os3BGlu6 purification.

Lane M, Bio-Rad low molecular weight markers; lane 1, crude protein extract of Origami(DE3) cells after Os3BGlu6 expression; lane 2, fusion protein from initial IMAC purification; lane 3, Os3BGlu6 protein after TEV protease digest; lane 4, wash fractions after 2nd IMAC, showing >95% pure Os3BGlu6 protein. Approximately 5  $\mu$ g of protein was loaded in each lane.

### **3.3 Effect of pH and temperature on the activity and stability of Os3BGlu6**

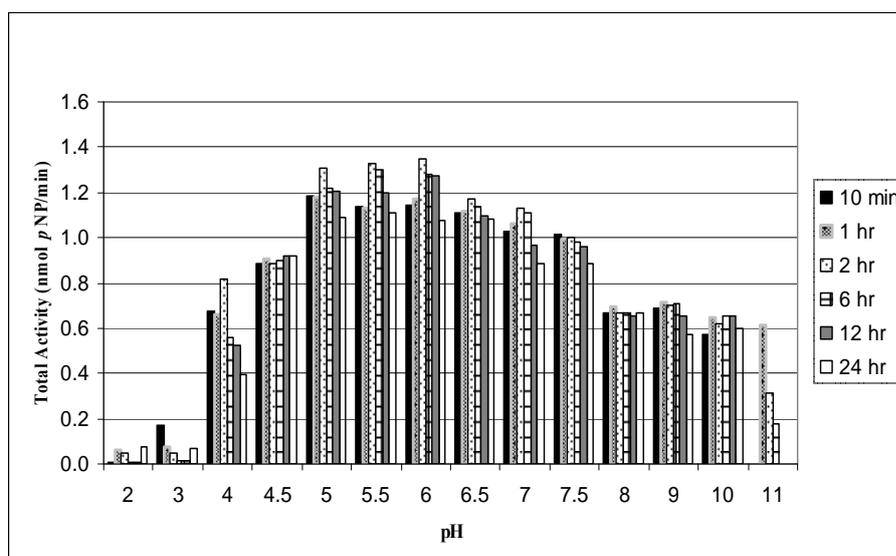
The optimum pH for Os3BGlu6 was found to be 4.5 (Figure 3.7). The activity dropped by 50% at pH 2.5 and 6.0 and was negligible from pH 7.0 upward. Results for the enzyme's stability over a period of 10 min up to 24 hours indicated that the enzyme was stable from pH 4.5 to pH 7.0 and slowly began losing its activity at pH values of 7.5 and higher (Figure 3.8). The enzyme was not stable at lower pH values (2 & 3). Although the enzyme was stable even at 24 hours at pH 5.0-7.5, there was further decrease in activity from pH 8.0 to pH 11.0. The temperature optimum for this enzyme showed a broad range of 40-55 $^{\circ}$  (Figure 3.9), but the enzyme was most stable

at 20-30°C and began to lose activity after 40 min at 40°C (Figure 3.10), or higher, hence the standard assays were conducted at 30°C, which is physiologically relevant for rice.



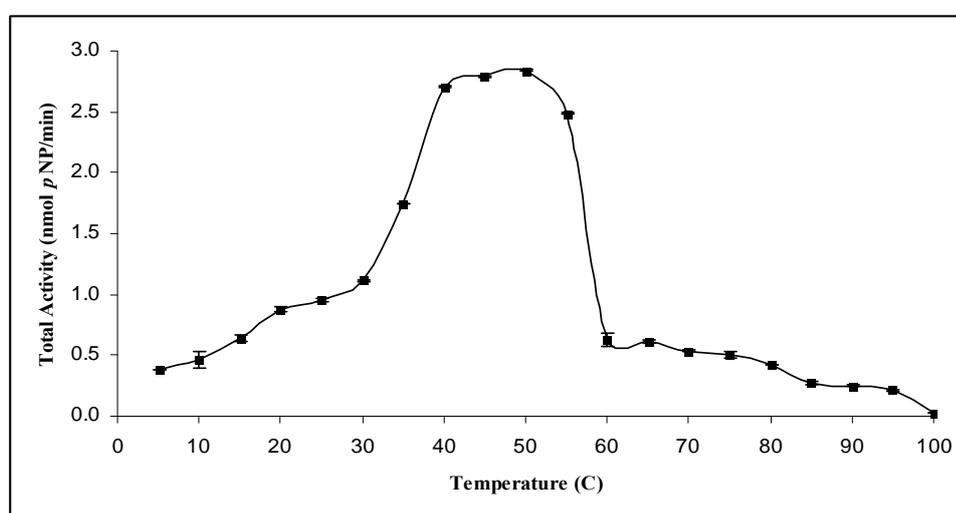
**Figure 3.7** pH dependence of Os3BGlu6 activity.

Os3BGlu6 was assayed with 1 mM *p*-nitrophenyl  $\beta$ -D-glucopyranoside (*p*NPGlc) substrate in universal citrate-phosphate buffer with overlapping pH (2-11) for 10 min at 30°C.



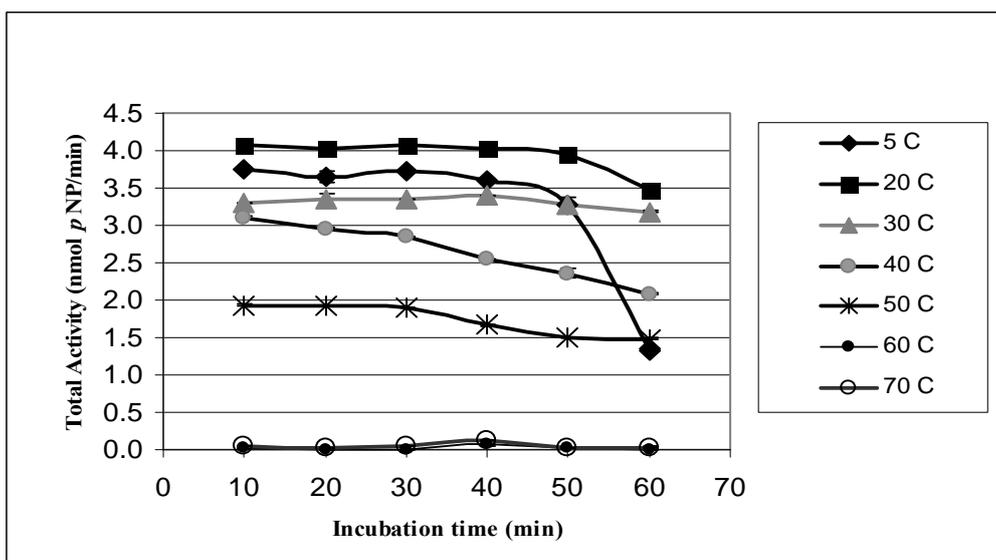
**Figure 3.8** pH stability of Os3BGlu6 activity.

Os3BGlu6 was incubated in universal buffer (pH 2-11) for time periods of 10 min, 1, 2, 6, 12, and 24 h at 30°C, then enzyme was diluted and assayed with *p*NPGlc for 10 min in 100 mM sodium acetate, pH 5.0, at 30°C.



**Figure 3.9** Temperature dependence of Os3BGlu6 activity.

Os3BGlu6 was incubated at various temperatures (5-100°C) for 10 min and assayed in 100 mM sodium acetate, pH 5.0, with 1 mM *p*NPGlc for 10 min.



**Figure 3.10** Thermostability of Os3BGlu6.

Os3BGlu6 was incubated at various temperatures (5-70°C) from 0 to 60 min in 100 mM sodium acetate, pH 5.0, and enzyme activity was determined at 30°C by adding 1 mM final concentration of *p*NPGlc and incubating for 10 min.

### 3.4 Substrate specificity and kinetic analysis of Os3BGlu6

The ability of Os3BGlu6 to hydrolyze *p*-nitrophenyl (*p*NP) glycosides was tested to assess its glycone specificity (Table 3.1). Among the *p*NP glycosides, Os3BGlu6 preferred *p*NP- $\beta$ -D-fucopyranoside (*p*NPFuc) with a  $k_{cat}/K_m$  of 66.8 mM<sup>-1</sup>s<sup>-1</sup>, due to its relatively low  $K_m$  of 0.50 mM, followed by *p*NP- $\beta$ -D-glucoside (*p*NPGlc) and *p*NP- $\beta$ -D-galactoside (*p*NPGal) with  $k_{cat}/K_m$  values of 6.21 and 1.61 mM<sup>-1</sup>s<sup>-1</sup>, respectively. Os3BGlu6 did not hydrolyze  $\alpha$ -D-glucoside, but did show a trace of activity toward  $\alpha$ -L-arabinoside, the structure of which is more similar to *p*NPGlc and *p*NPFuc. The glycone specificity is somewhat different from previously characterized rice  $\beta$ -glucosidases, such as rice BGlul (Os3BGlu7), which showed a slight preference for *p*NPFuc over *p*NPGlc due to a higher  $k_{cat}$  for *p*NPFuc, (Opassiri *et al.*,

2004), whereas for Os3BGlu6, the difference is larger due to a 10-fold lower  $K_m$  for *p*NP-Fuc than *p*NP-Glc. In contrast, Os3BGlu6 had a much lower efficiency for hydrolysis of *p*NP- $\beta$ -D-mannopyranoside than Os3BGlu7 ( $k_{cat}/K_m$  1.01 mM<sup>-1</sup>s<sup>-1</sup>).

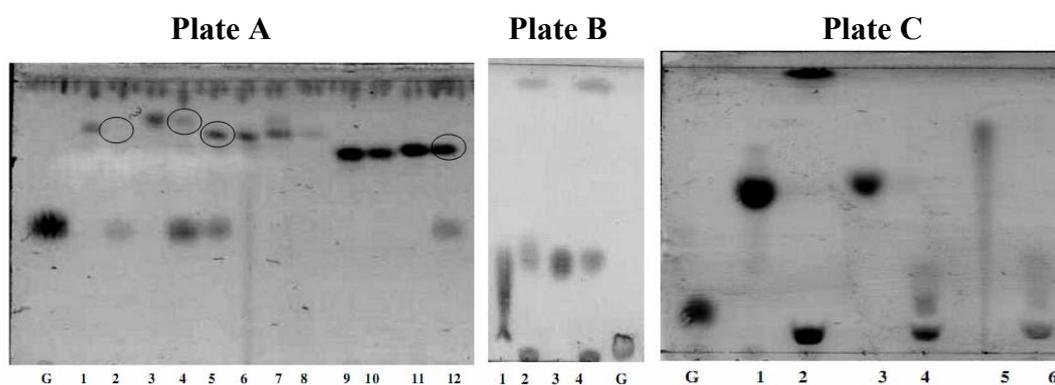
**Table 3.1** Kinetic parameters of Os3BGlu6 for the hydrolysis of *p*NP-glycosides.

Substrate	nmole <i>p</i> NP/min/mg protein	Percent Relative Activity	$K_m$ (mM)	$k_{cat}$ (s <sup>-1</sup> )	$k_{cat}/K_m$ (mM <sup>-1</sup> s <sup>-1</sup> )
<i>p</i> NP- $\beta$ -D-glucoside	5844 ± 1.5	100	6.3 ± 0.4	38.9 ± 0.9	6.2
<i>p</i> NP- $\beta$ -D-fucoside	24,470 ± 1.5	415 ± 0.03	0.50 ± 0.02	33.47 ± 0.7	66.8
<i>p</i> NP- $\beta$ -D-galactoside	1719 ± 1.4	29.4 ± 0.02	6 ± 0.5	9.75 ± 0.2	1.61
<i>p</i> NP- $\alpha$ -L-arabinoside	51.9 ± 0.3	0.88 ± 0.005	nd	nd	nd
<i>p</i> NP- $\beta$ -D-cellobioside	27.7 ± 0.2	0.46 ± 0.003	nd	nd	nd
<i>p</i> NP- $\beta$ -D-xyloside	6.3 ± 0.5	0.1 ± 0.008	nd	nd	nd
<i>p</i> NP- $\beta$ -D-mannoside	3.8 ± 0.5	0.06 ± 0.008	nd	nd	nd
<i>p</i> NP- $\beta$ -D-glucuronide	2.7 ± 0.3	0.04 ± 0.005	nd	nd	nd

nd indicates “not determined”

Hydrolysis of synthetic and natural glycosides and oligosaccharides were tested to assess the aglycone specificity of Os3BGlu6 (Table 3.2). Os3BGlu6 hydrolyzed *n*-octyl- $\beta$ -D-glucoside with the highest catalytic efficiency ( $k_{cat}/K_m = 2.7$  mM<sup>-1</sup>s<sup>-1</sup>), which was approximately 3-fold higher than that for *n*-heptyl- $\beta$ -D-glucoside (0.85 mM<sup>-1</sup>s<sup>-1</sup>). Among the disaccharides, Os3BGlu6 hydrolyzed the  $\beta$ -(1→3)-linked

laminaribiose best ( $k_{\text{cat}}/K_{\text{m}}=1.7 \text{ mM}^{-1}\text{s}^{-1}$ ), followed by  $\beta$ -(1 $\rightarrow$ 2)-linked sophorose ( $k_{\text{cat}}/K_{\text{m}} = 0.96 \text{ mM}^{-1}\text{s}^{-1}$ ),  $\beta$ -(1 $\rightarrow$ 6)-linked gentiobiose ( $k_{\text{cat}}/K_{\text{m}} = 0.011 \text{ mM}^{-1}\text{s}^{-1}$ ) and  $\beta$ -(1 $\rightarrow$ 4)-linked cellobiose ( $k_{\text{cat}}/K_{\text{m}}=0.009 \text{ mM}^{-1}\text{s}^{-1}$ ), in order of decreasing catalytic efficiency ( $k_{\text{cat}}/K_{\text{m}}$ ). Longer  $\beta$ -(1 $\rightarrow$ 3)- and  $\beta$ -(1 $\rightarrow$ 4)-linked oligosaccharides were hydrolyzed at decreasing rates. Os3BGlu6 also hydrolyzed natural glycosides, including apigenin-7-glucoside, glycitin, diadzin, genistin, esculin, arbutin, coumaryl alcohol  $\beta$ -D-glucoside, coniferin, and salicin, although at a longer time of incubation (overnight) as analyzed on TLC (Figures 3.11 A, B, and C). The glycosides amygdalin, linamarin, prunacin dalcochinin- $\beta$ -glucoside, gossypin, naringin, phlorizin, pyridoxine-5'-O- $\beta$ -D-glucoside, and quercitin 3- $\beta$ -D-glucoside were not hydrolyzed.



**Figure 3.11** TLC analyses of natural substrates hydrolyzed by Os3BGlu6.

Hydrolysis was assessed by overnight incubation of 2  $\mu\text{g}$  of purified enzyme with 5 mM substrates in 50 mM sodium acetate, pH 5.0, at 30°C. The reactions were analyzed on a normal silica gel 60 for plate A and F<sub>254</sub> for plates B and C with a solvent system of ethyl acetate:acetic acid:methanol:water (15:2:1:2). (Plate A) G, glucose standard; 1, diadzin control; 2, diadzin reaction; 3, genistin control; 4, genistin reaction; 5, glycitin reaction; 6, glycitin control; 7, linamarin control; 8, linamarin reaction; 9, amygdalin control; 10, amygdalin reaction; 11, coumaryl alcohol  $\beta$ -D-

glucoside control; 12, coumaryl alcohol  $\beta$ -D-glucoside reaction; (Plate B) 1, arbutin control; 2, arbutin reaction; 3, salicin control; 4, salicin reaction; G, glucose standard; (Plate C) G, glucose standard; 1, coniferin control; 2, coniferin reaction; 3, esculin control; 4, esculin reaction; 5, apigenin-7-O- $\beta$ -D-glucoside control; 6, apigenin-7-O- $\beta$ -D-glucoside reaction.

**Table 3.2** Hydrolysis of glycosides and oligosaccharides by Os3BGlu6.

Substrate	nmole <i>p</i> NP/min/mg protein	Percent Relative Activity	$K_m$ (mM)	$k_{cat}$ ( $s^{-1}$ )	$k_{cat}/K_m$ ( $mM^{-1}s^{-1}$ )
<i>n</i> -octyl - $\beta$ -D-glucoside	2282 $\pm$ 0.9	100 $\pm$ 0.04	4.54 $\pm$ 0.3	12.4 $\pm$ 0.4	2.7
<i>n</i> -heptyl- $\beta$ -D-glucoside	673.8 $\pm$ 0.9	29.5 $\pm$ 0.04	5.0 $\pm$ 0.3	4.25 $\pm$ 0.10	0.85
Laminaribiose	1272 $\pm$ 1.3	55.7 $\pm$ 0.05	3.6 $\pm$ 0.3	6.17 $\pm$ 0.2	1.7
Laminaritriose	169.6 $\pm$ 0.9	7.43 $\pm$ 0.04	8.7 $\pm$ 0.69	1.64 $\pm$ 0.06	0.18
Laminaritetraose	2.0 $\pm$ 0.01	0.08 $\pm$ 0.0004	nd	nd	nd
Laminaripentaose	1.4 $\pm$ 0.01	0.06 $\pm$ 0.001	nd	nd	nd
Cellobiose	5.1.0 $\pm$ 0.5	0.22 $\pm$ 0.002	15.3 $\pm$ 1.2	0.13 $\pm$ 0.001	0.009
Cellotriose	13.1 $\pm$ 0.6	0.57 $\pm$ 0.03	nd	nd	nd
Cellotetraose	12.9 $\pm$ 0.5	0.56 $\pm$ 0.02	nd	nd	nd
Cellopentaose	8.0 $\pm$ 0.3	0.35 $\pm$ 0.01	nd	nd	nd
Cellohexaose	6.38 $\pm$ 0.3	0.28 $\pm$ 0.01	nd	nd	nd
Sophorose	728.7 $\pm$ 1.3	31.9 $\pm$ 0.05	9.84 $\pm$ 0.84	9.84 $\pm$ 0.36	0.96
Gentiobiose	5.3 $\pm$ 0.17	0.23 $\pm$ 0.006	14.9 $\pm$ 1.1	0.11 $\pm$ 0.003	0.007
<sup>+</sup> Apigenin-7-glucoside	15.25 $\pm$ 0.22	0.66 $\pm$ 0.009	nd	nd	nd
<sup>+</sup> Glycitin	4.4 $\pm$ 0.07	0.19 $\pm$ 0.003	nd	nd	nd
<sup>+</sup> Diadzin	4.0 $\pm$ 0.04	0.17 $\pm$ 0.001	nd	nd	nd
<sup>+</sup> Genistin	2.9 $\pm$ 0.02	0.13 $\pm$ 0.001	nd	nd	nd
<sup>+</sup> Esculin	2.4 $\pm$ 0.01	0.10 $\pm$ 0.005	nd	nd	nd
<sup>+</sup> Arbutin	0.59 $\pm$ 0.05	0.02 $\pm$ 0.002	nd	nd	nd
<sup>+</sup> Coumaryl alcohol $\beta$ - D-glucoside	0.54 $\pm$ 0.01	0.02 $\pm$ 0.0005	nd	nd	nd
<sup>+</sup> Coniferin	0.46 $\pm$ 0.007	0.02 $\pm$ 0.0003	nd	nd	nd
<sup>+</sup> Salicin	0.18 $\pm$ 0.001	0.008 $\pm$ 0.003	nd	nd	nd

nd- not determined. <sup>+</sup> The natural glycosides for which relative activities were determined are those for which hydrolysis as detected on TLC.

### 3.5 Os3BGlu6 crystallization

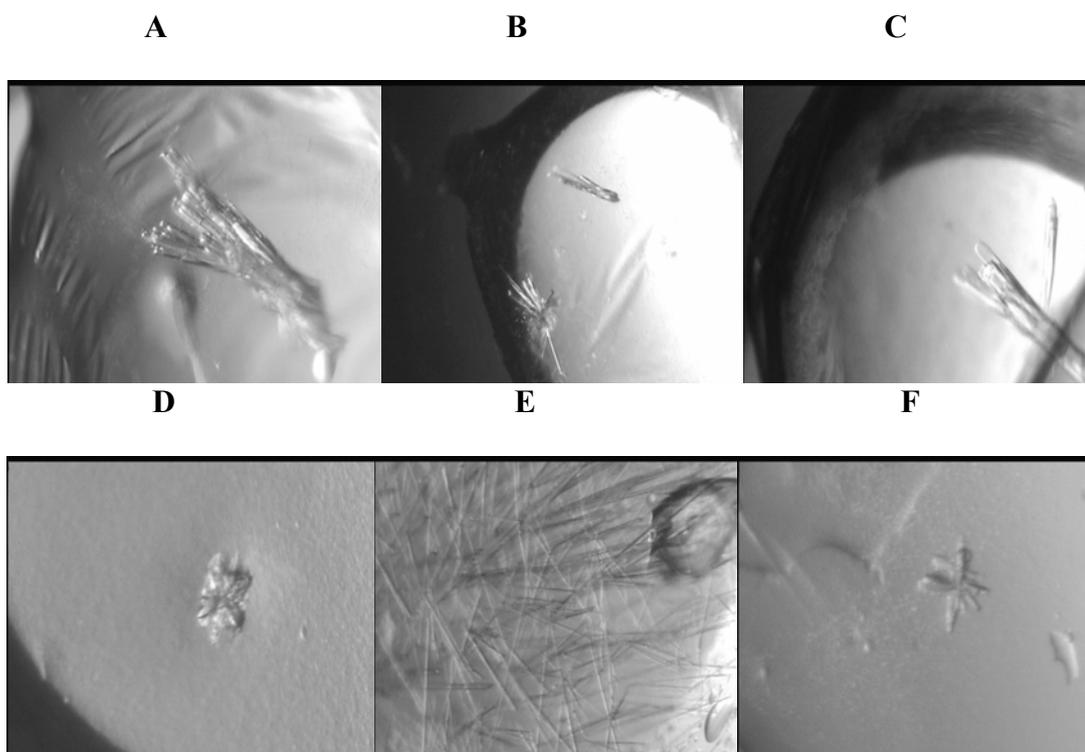
#### 3.5.1 Initial screening by the microbatch method

Four screening kits were used for preliminary screening of crystallization trials (as described in section 2.7.2). A solution of pure filtered protein of at least 10 mg/ml was used for initial screening with a protein: precipitant volume ratio of 1:1 and 2:1. Positive conditions in which various sizes of needle-type and crystals clusters appeared after 180-200 days (Figure 3.12) under several conditions, are summarized in Table 3.3.

**Table 3.3** Summary of positive conditions of native Os3BGlu6 obtained from the commercially available screening kits.

Screening Kit	Condition	Precipitant	Crystal morphology
HRII-134	D6	0.1 M Bis Tris, pH 6.5, 25% PEG 3,350	Cluster
	D10*	0.1 M Bis Tris, pH 6.5, 20% PEG 5000	Cluster
		MME	Cluster
	D11	0.1 M Bis Tris, pH 6.5, 28% PEG 2000 MME	
Wizard I <sup>TM</sup>	37	1 M K/Na Tartrate, Tris, pH 7.0, LiSO <sub>4</sub>	Cluster
	27	1.2 M Na <sub>2</sub> H <sub>2</sub> PO <sub>4</sub> /0.8 M K <sub>2</sub> HPO <sub>4</sub> , pH 10.5, LiSO <sub>4</sub>	Needle cluster
JY Screen	2F	PEG 4000 25%, Tris, pH 8.5, and 0.2 M NH <sub>4</sub> SO <sub>4</sub>	Clusters

\* indicates the condition chosen for further optimization



**Figure 3.12** Os3BGlu6 crystals obtained from microbatch screening after 180 days. (A) 0.1 M Bis Tris, pH 6.5, 25% PEG 3,350; (B) 0.1 M Bis Tris, pH 6.5, 20% PEG 5000 MME; (C) 0.1 M Bis Tris, pH 6.5, 28% PEG 2000 MME; (D) 1 M K/Na Tartrate, Tris, pH 7.0, LiSO<sub>4</sub>; (E) 1.2 M Na<sub>2</sub>H<sub>2</sub>PO<sub>4</sub>/0.8 M K<sub>2</sub>HPO<sub>4</sub>, pH 10.5, LiSO<sub>4</sub>; and (F) PEG 4000 25%, Tris, pH 8.5, and 0.2 M NH<sub>4</sub>SO<sub>4</sub>.

### 3.5.2 Optimization by hanging drop

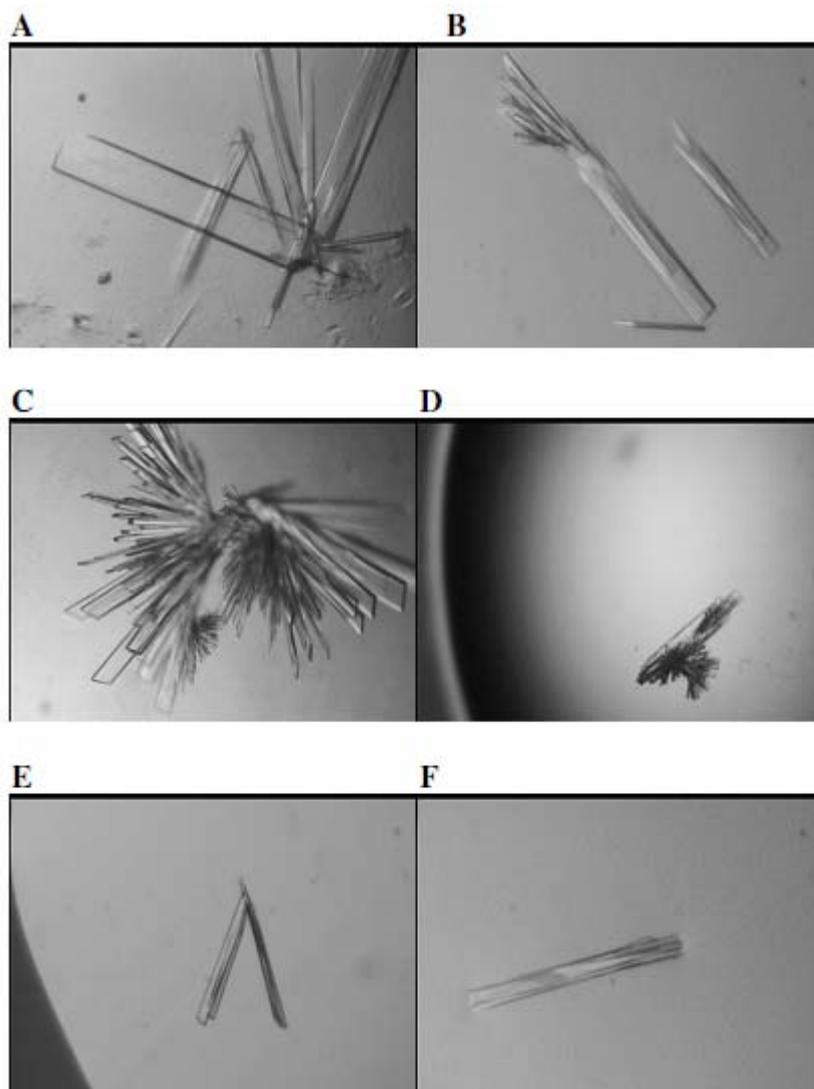
Of the positive conditions, condition D10 from HR2-134 was chosen for further optimization, based on the reproducibility of the crystals. To obtain single, well-formed larger Os3BGlu6 crystals for x-ray diffraction, optimization was conducted by varying the concentration of PEG, pH of the buffer used and type of PEG. In the first step of optimization, the D10 condition (0.1 M Bis Tris, pH 6.5, 20%

PEG 5000 MME) was the starting point for a preliminary coarse screening in which just the percent PEG was varied with 10 mg/ml protein (Table 3.4 and Figure 3.13).

**Table 3.4** Results of coarse screening with different concentrations of PEG 5000 MME obtained from condition D10.

% PEG 5000 MME (w/v)	D10	
	1:1 protein: precipitant	2:1 protein: precipitant
17%	clusters	-
18%	clusters	-
19%*	clusters	-
20%	clusters	-
21%	clusters	-
22%	clusters	-

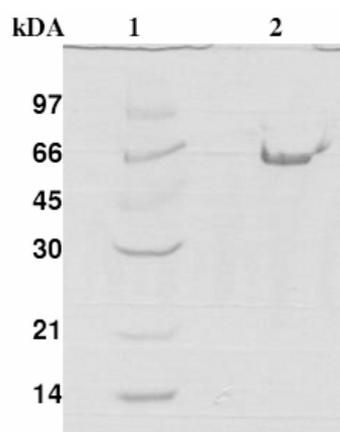
\* indicates the seeding stock for microseeding in subsequent experiments.  
 - indicates no crystal formed. Only precipitant formed.



**Figure 3.13** Clusters of Os3BGlu6 crystals obtained from optimization by varying the concentration of PEG 5000 MME. in 0.1 M Bis-Tris, pH 6.5.

(A), 17% PEG 5000 MME; (B), 18% PEG 5000 MME; (C), 19% PEG 5000 MME; (D), 20% PEG 5000 MME ; (E), 21% PEG 5000 MME and (F), 22% PEG 5000 MME.

Clusters of plates-like crystals were formed between 9-12 days in all the conditions with a 1:1 protein to precipitant volume ratio. The crystals of Os3BGlu6 obtained were dissolved in SDS-PAGE buffer and the protein run on SDS PAGE and the presence of Os3BGlu6 protein could be confirmed (Figure 3.14).



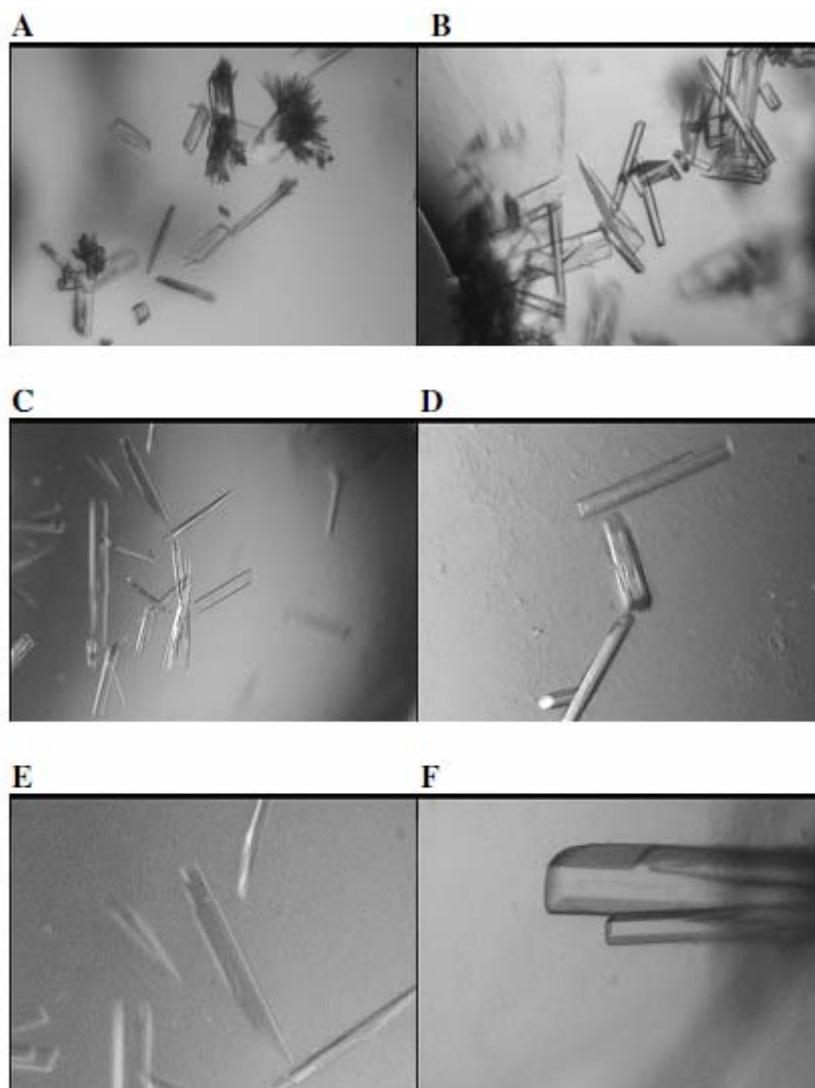
**Figure 3.14** Protein from Os3BGlu6 crystals run on SDS PAGE.

Lane 1, low molecular weight protein marker; lane 2, protein from native Os3BGlu6 crystals.

Further coarse screening was conducted by fixing this ratio of protein to precipitant and varying the protein concentration from 1 to 10 mg/ml and the percent PEG 5000 MME over a wider range to produce single crystals (Table 3.5). Both single crystals and clusters of better quality were seen, particularly in drops with 4 and 8 mg/ml protein in 13%, 15%, 17%, 19%, 20%, 21%, and 22% PEG 5000 MME and 0.1 M Bis Tris, pH 6.5 (Figure 3.15).

**Table 3.5** Results of coarse screening with different concentrations of PEG 5000 MME in 0.1 M Bis Tris, pH 6.5, with five different concentrations of protein (1 to 10 mg/ml).

% PEG 5000 MME (w/v)	D10 1:1 protein: precipitant	% PEG 5000 MME (w/v)	D10 1:1 protein: precipitant
11%	-	17%	Single/clusters
12%	-	18%	-
13%	Single/clusters	19%	Single/clusters
14%	-	20%	Single/clusters
15%	Single/clusters	21%	Single/clusters
16%	-	22%	Single/clusters



**Figure 3.15** Os3BGlu6 crystals obtained in different concentrations of PEG 5000 MME and protein in 0.1 M Bis-Tris pH, 6.5.

The conditions are (A), 13% PEG 5000 MME, 0.1 M Bis Tris, pH 6.5, and 4 mg/ml protein; (B), 15% PEG 5000 MME, 0.1 M Bis Tris, pH 6.5, and 4 mg/ml protein; (C), 17% PEG 5000 MME, 0.1 M Bis Tris, pH 6.5 and 8 mg/ml protein; (D), 19% PEG 5000 MME, 0.1 M Bis Tris, pH 6.5, and 4 mg/ml protein; (E), 21% PEG 5000 MME, 0.1 M Bis Tris, pH 6.5, and 4 mg/ml protein; (F), 22% PEG 5000 MME, 0.1 M Bis Tris, pH 6.5, and 8 mg/ml protein.

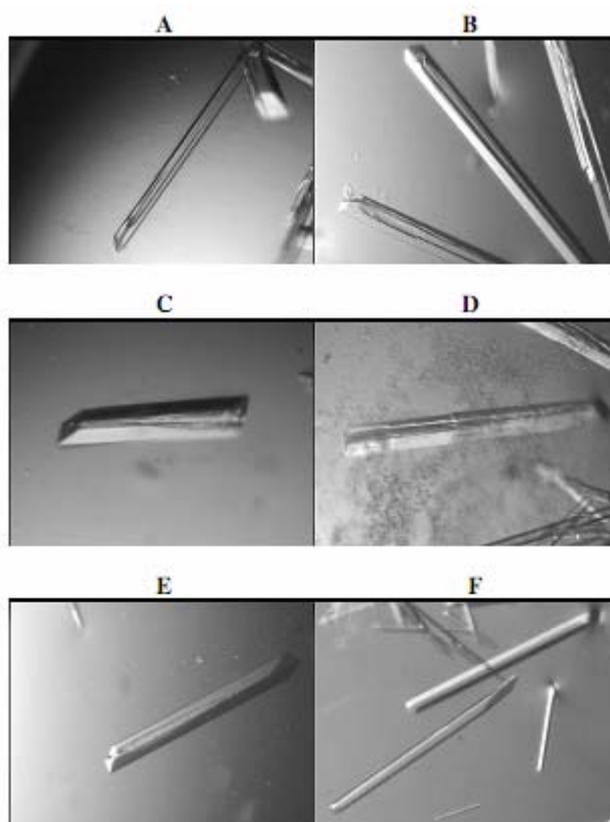
A refined screening was conducted by fixing the protein concentrations at 4 and 8 mg/ml, based on the quality of the crystals obtained from the previous screening step, and the ratio between the volumes of protein and precipitant and the pH of the 0.1 M Bis Tris buffer was varied (Table 3.6). Large single crystals were obtained at pH 6.5 within 12 days. In addition, large single crystals were obtained in 0.1 M Bis-Tris pH 6.4 and 6.7, (Figure 3.16). However, the crystals which were produced in the condition with 0.1 M Bis Tris, pH 6.5, and in 8 mg/ml protein had better shape and edges.

**Table 3.6** Results of refined screening with different concentrations of PEG 5000 MME and 0.1 M Bis-Tris pH range of 5.5 to 6.9.

pH of 0.1 M Bis Tris	5.5	6.2	6.4	6.5	6.7	6.9
13% PEG MME	-	-	-	-	-	-
15% PEG MME	-	-	-	-	-	-
19% PEG MME	-	-	+	+	+	-
21% PEG MME	-	-	+	+	+	-

+ indicates single crystals obtained in 12 days of optimization

- indicates no crystals obtained in 12 days

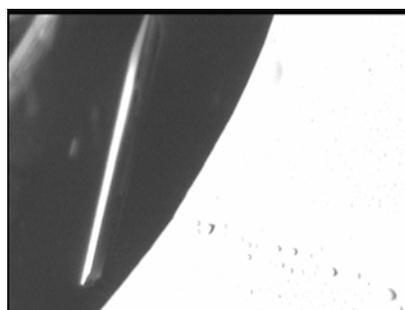


**Figure 3.16** Os3BGlu6 crystals obtained in different concentrations of PEG 5000 MME and 0.1 M Bis-Tris, pH 6.4, 6.5, and 6.7.

The conditions are (A), 19% PEG 5000 MME, 0.1 M Bis Tris, pH 6.4, and 4 mg/ml protein; (B), 21% PEG 5000 MME, 0.1 M Bis Tris, pH 6.4, and 4 mg/ml protein; (C), 19% PEG 5000 MME, 0.1 M Bis Tris, pH 6.7, and 4 mg/ml protein; (D), 21% PEG 5000 MME, 0.1 M Bis Tris, pH 6.7, and 4 mg/ml protein; (E), 19% PEG 5000 MME, 0.1 M Bis Tris, pH 6.7, and 4 mg/ml protein; (E), 19% PEG 5000 MME, 0.1 M Bis Tris, pH 6.5, and 8 mg/ml protein; (F), 21% PEG 5000 MME, 0.1 M Bis Tris, pH 6.5, and 8 mg/ml protein.

### 3.5.3 Optimization by microseeding

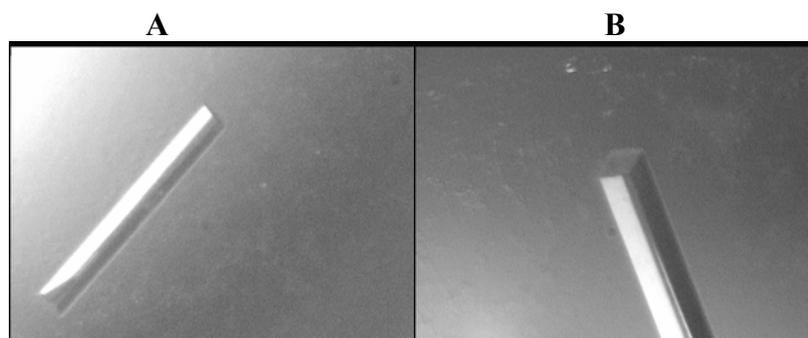
By the process of microseeding, larger, better quality, single crystals of Os3BGlu6 were obtained with 8 mg/ml protein in 21% PEG 5000 MME, 0.1 M Bis Tris, pH 6.5, (Figure 3.17). Crystals from this condition were soaked in the cryoprotectant (24.15% PEG 5000 MME, 0.115 M Bis Tris, pH 6.5, and 18% (v/v) glycerol), and flash-frozen in liquid nitrogen for subsequent diffraction.



**Figure 3.17** A crystal of native Os3BGlu6 of dimensions 210 x 20 x 15  $\mu\text{m}$  obtained by microseeding in the condition 21% PEG 5000 MME, 0.1 M Bis-Tris, pH 6.5.

### 3.5.4 Soaking of Os3BGlu6 crystals with inhibitors

Os3BGlu6 crystals with inhibitors were generated by soaking crystals obtained in 19% PEG 5000 MME, 0.1 M Bis Tris, pH 6.5, with 2-deoxy-2-fluoro- $\beta$ -D-glucoside (DNPG2F) or *n*-octyl- $\beta$ -D-thioglucopyranoside (Figure 3.18). The cryoprotectants contained inhibitors present in 21.85% PEG 5000 MME, 0.115 M Bis-Tris, pH 6.5, and 18% v/v glycerol.



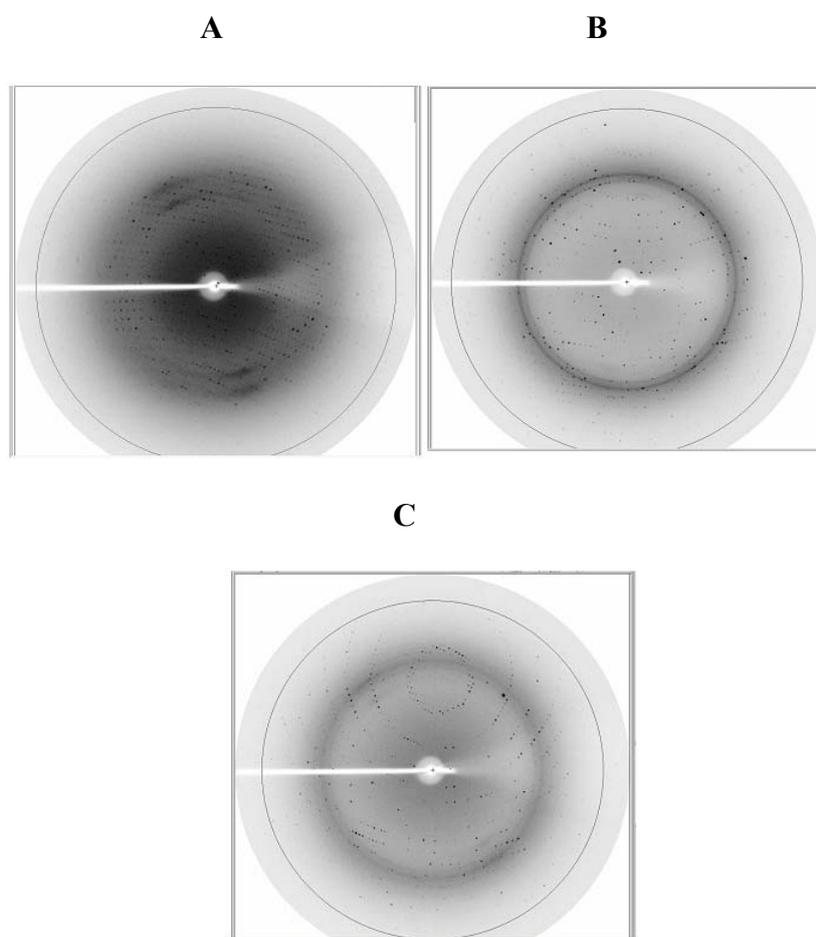
**Figure 3.18** Os3BGlu6 crystals soaked with inhibitors in 19% 5000 MME, 0.1 M Bis Tris, pH 6.5 and 8 mg/ml protein.

(A) Os3BGlu6 crystal with dimensions 210 x 20 x 30  $\mu\text{m}$  soaked in 5 mM 2-deoxy-2-fluoro- $\beta$ -D-glucoside (B) Os3BGlu6 crystal with dimensions 220 x 40 x 50  $\mu\text{m}$  soaked in 10 mM *n*-octyl- $\beta$ -D-thioglucopyranoside.

### 3.5.5 Diffraction, Data Collection, processing and refinement

All the crystals of Os3BGlu6 both native and with inhibitors were used to diffract x-ray beams from synchrotron radiation at the National Synchrotron Radiation Research Center, Taiwan. The data was collected with  $0.5^\circ$  oscillations for native Os3BGlu6 and Os3BGlu6/ *n*-octyl- $\beta$ -D-thioglucopyranoside, and  $0.3^\circ$  oscillations for Os3BGlu6/G2F which gave diffraction images as shown in Figure 3.19. Since Os3BGlu6/G2F dataset showed overlap of spots and gave  $\text{Chi}^2$  values greater than 1, the data was collected at  $0.3^\circ$  oscillations. Os3BGlu6 crystals diffraction data with and without inhibitors was indexed and scaled to the orthorhombic  $P2_12_12_1$  space group, when processed with HKL2000. The native crystals had the unit cell parameters  $a=56.80 \text{ \AA}$ ,  $b=90.48 \text{ \AA}$ ,  $c=101.71 \text{ \AA}$ , and those for the G2F complex were  $56.67 \text{ \AA}$ ,  $90.59 \text{ \AA}$ , and  $101.88 \text{ \AA}$ , while those of the *n*-octyl- $\beta$ -D-thioglucopyranoside complex were  $a=57.18 \text{ \AA}$ ,  $b=91.09 \text{ \AA}$  and  $c=111.35 \text{ \AA}$ . Each

asymmetric unit was found to contain one protein monomer. The Matthew's coefficient ( $V_M$ ) for native structure of Os3BGlu6 was  $2.38 \text{ \AA}^3/\text{Da}$  and the solvent content was 48.2% for the native structure. Solvent contents of 53.3% and 48.3%, with corresponding Matthew's coefficients of 2.64 and  $2.38 \text{ \AA}^3/\text{Da}$  were calculated for *n*-octyl- $\beta$ -D-thio-glucopyranoside and G2F complexes, respectively.



**Figure 3.19** Diffraction images of Os3BGlu6 crystals.

(A) native (B) *n*-octyl- $\beta$ -D-thioglucopyranoside and (C) 2-deoxy-2-fluoro- $\beta$ -D-glucopyranosideside.

The cyanogenic  $\beta$ -glucosidase from white clover (*Trifolium repens*; PDB code: 1CBG; Barrett *et al.*, 1995), which was used as a search model for molecular

replacement (described in section 2.7.7), had 54.4% amino acid sequence identity with Os3BGlu6. The native structure of Os3BGlu6 was refined at 1.83 Å resolution. The self rotation and translation functions during molecular replacement gave an initial solution of one molecule in the asymmetric unit with the crystallographic *R factor* of 53% in the data ranging from 24.80 to 2.01 Å. The initial 60 cycles of restrained refinement with Refmac 5.0 gave a reduced *R-factor* of 29.3% and *R<sub>free</sub>* of 33.3%. After addition of Tris ligand, waters and glycerol, the final refinement gave an *R-factor* of 16.39% and *R<sub>free</sub>* of 20.42%. The native Os3BGlu6 structure was used as the template model for solving the structures of Os3BGlu6/*n*-octyl-β-D-thioglucopyranoside and Os3BGlu6/G2F. The structure of the Os3BGlu6/*n*-octyl-β-D-thioglucopyranoside complex was solved at 1.80 Å and initial 7 cycles of restrained refinement with *R-factor* of 23.2% and *R<sub>free</sub>* of 25.2%. A final *R-factor* and *R<sub>free</sub>* of 17.85% and 20.66% were obtained after several cycles of restrained refinement and rebuilding. The structure of the G2F complex was refined at 1.81 Å and yielded *R factor* of 23.7% and *R<sub>free</sub>* of 25.7% during initial 3 rounds of refinement. The final *R-factor* and *R<sub>free</sub>* were 16.24% and 18.48%, respectively. The final models indicated a good fit of the electron density map with average B-factors of 12.76 Å<sup>2</sup>, 18.59 Å<sup>2</sup> and 16.07 Å<sup>2</sup> for the protein atoms of the native, and *n*-octyl-β-D-thioglucopyranoside and G2F complex structures, respectively. The refined structure of native Os3BGlu6 contained 466 water molecules with an average B-factor of 26.22 Å<sup>2</sup>, while *n*-octyl-β-D-thioglucopyranoside and G2F complexes contained 340 and 452 water molecules with average B-factors of 27.1 and 27 Å<sup>2</sup>, respectively. The data processing and refinement parameters are provided in Table 3.7.

**Table 3.7** Data collection, processing and refinement parameters.

Dataset	Native Os3BGlu6	Os3BGlu6/ <i>n</i> -octyl- $\beta$ -D-thioglucoside complex	Os3BGlu6/G2F complex
PDB Code	3GNO	3GNP	3GNR
Beamline	BL13B1	BL13B1	BL13B1
Wavelength (Å)	1.00	1.00	1.00
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Unit-cell parameters (Å)	a = 56.80 b = 90.48 c = 101.71	a = 57.18 b = 91.09 c = 111.35	a = 56.67 b = 90.59 c = 101.88
Resolution range (Å)	30.0 - 1.83	30.0 - 1.80	30.0 - 1.81
Resolution outer shell(Å)	1.90 -1.83	1.86-1.80	1.87-1.81
No. Unique reflections	46801	49796	45577
No. Observed reflections	326991	343111	170048
Completeness (%)	98.9 (96.8) <sup>a</sup>	91.1 (98.0) <sup>a</sup>	92.9 (95.1) <sup>a</sup>
Average redundancy per shell	7.0 (6.9)	6.9 (6.9)	3.7 (4.0)
$I/\sigma(I)$	18.9 (4.9) <sup>a</sup>	30.7 (6.4) <sup>a</sup>	23.9 (8.7) <sup>a</sup>
R <sub>(merge)</sub> (%)	9.5 (36.6) <sup>a</sup>	5.2 (29.3) <sup>a</sup>	4.5 (18.4) <sup>a</sup>
R <sub>factor</sub> (%)	16.39	17.85	16.24
R <sub>free</sub> (%)	20.42	20.66	18.48
No. of residues in protein	488	488	488
No. protein atoms	3859	3848	3871
No. ligand atoms	8	20	11
No. other hetero atoms	36	24	12
No. waters	466	340	452
<i>Mean B-factor</i>			
Protein	12.76	18.59	16.07
Ligand	28.81	41.25	14.85
Water	26.22	27.10	27.00
r.m.s. bond length deviations (Å)	0.013	0.015	0.012
r.m.s bond angle deviations (degrees)	1.321	1.441	1.287
Ramachandran plot			
-Residues in most favorable regions (%)	89.6	89.6	89.6
-Residues in additional allowed regions (%)	10	10	10
-Residues in generously allowed regions (%)	0.5	0.5	0.5
-Residues in disallowed regions (%)	0	0	0

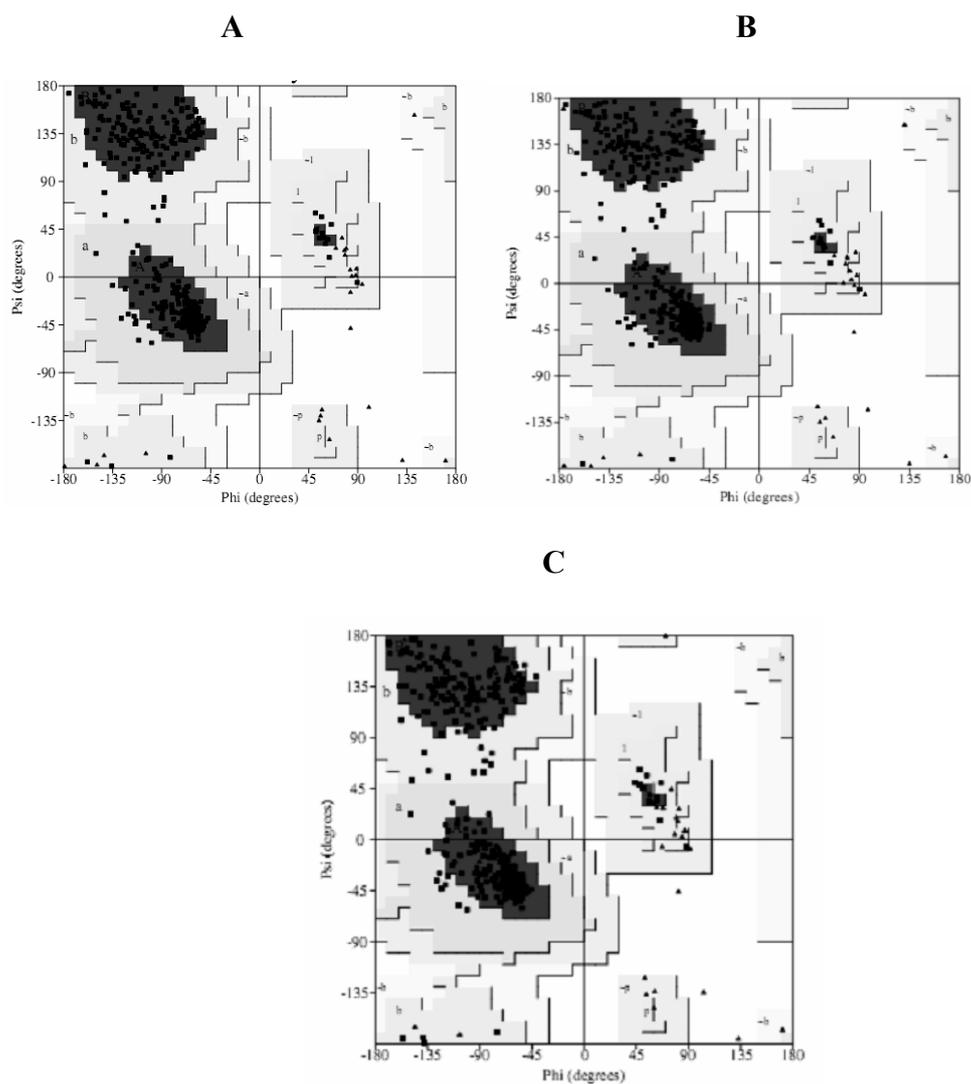
<sup>a</sup> Numbers in parentheses indicate values for the highest resolution shell.

## 3.6 3D structures of Os3BGlu6

### 3.6.1 Validation of structures

The final models were validated as described in section 2.7.8. The Ramachandran plots for the native Os3BGlu6, and *n*-octyl- $\beta$ -D-thioglucopyranoside and G2F complex (Figure 3.20) indicated that approximately 90% of non proline and non glycine residues were found in the most favored regions and a further 10% in the additional allowed regions. Two of the residues, Ala65 and Trp452, were found in the generously allowed regions in the all the structures. No residues were found in disallowed regions in any of the structures. The final model comprised 488 amino acid residues with 3859, 3848, and 3871 protein atoms for the native, *n*-octyl- $\beta$ -D-thioglucopyranoside and G2F crystal structures, respectively. These differences in the number of protein atoms were due to the occurrence of two conformations of amino acid residues.

The RMS deviations of 0.013, 0.015, and 0.012 Å for bond lengths and 1.321°, 1.441°, and 1.287° for bond angles in the native, *n*-octyl- $\beta$ -D-thioglucopyranoside and G2F complex structures, respectively, reflected the acceptable geometry observed for almost all the residues.



**Figure 3.20** Ramachandran plots obtained from PROCHECK validation of the refined Os3BGlu6 structure.

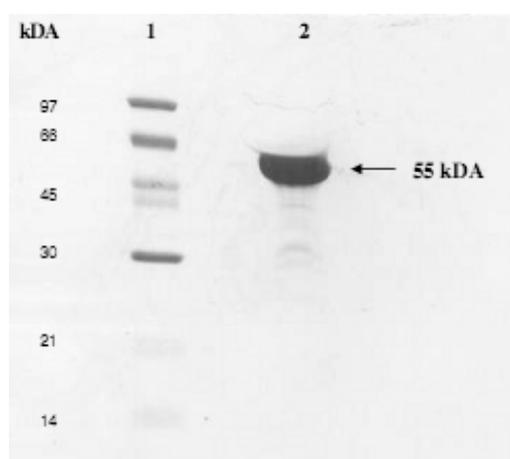
(A) Native Os3BGlu6, (B) Os3BGlu6/G2F complex, (C) Os3BGlu6/*n*-octyl- $\beta$ -D-thioglucopyranoside complex.

### 3.6.2 Overall structures of Os3BGlu6 and its complexes with G2F and *n*-octyl- $\beta$ -D-thioglucopyranoside

The overall structures were very similar to other known structures of GH1 enzymes, with the typical  $(\beta/\alpha)_8$  barrel (Figure 3.22). The highly conserved

catalytic residues, the acid/base Glu178 and the nucleophile Glu394, are located on  $\beta$ -strands four and seven of the  $\beta$ -barrel. Since the electron density for the first 10 residues at the N-terminus was lacking, protein recovered from crystals was blotted onto PVDF and submitted to Edman degradation to identify the N-terminus (Figure 3.21). N-terminal sequencing of the first 8 amino acid residues of the protein recovered from crystals gave a sequence consistent with that predicted from the TEV cleavage site, with 4 residues from the cloning vector and primers, Ser1, Phe2, Thr3, and Met4, followed by the first 4 residues included from the protein sequence, Ala5, Gln6, Gln7, and Ser8. However, these residues, together with Gly9 and Gly10, had no detectable electron density, likely due to high flexibility, and hence were not built in the structure. The C-terminus of the model was well defined, except for last two residues Lys487 and Thr488, which had poor density for their side chains. In addition, very poor density was observed for the side chains of residues 334 to 337 in loop C for all the three structures, although the  $C\alpha$  backbone was clearly visible for the native and Os3BGlu6/G2F complex structures. This loop is apparently flexible since it is exposed at the outer surface of the protein and the position of this loop was slightly different in the case of the Os3BGlu6/*n*-octyl- $\beta$ -D-thioglucopyranoside. No density was observed for six of the atoms of the surface residues which were  $N^Z$  in Lys114,  $O^{\epsilon 1}$  and  $O^{\epsilon 2}$  in Glu296,  $O^{\epsilon 1}$  and  $O^{\epsilon 2}$  in Glu383, and  $N^Z$  in Lys432, both in the native and Os3BGlu6/G2F structures, so they were deleted from the structure. Due to the high resolution of the model, multiple conformations could be observed for residues Arg233, Thr234, Asn332, and Glu451 in native Os3BGlu6 and Asn241, Ser231, Arg293, Asp361, Glu433, and Glu451 in the Os3BGlu6/G2F complex. One disulfide bond was seen linking Cys197 and Cys205 in loop B, a feature which is conserved in

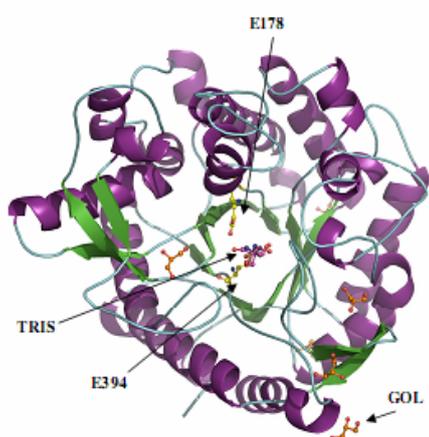
plant GH1 enzymes. Two *cis*-peptide bonds were found between Ala193 and Pro194 and between Trp444 and Ser445, which is also common in GH1. Six glycerol molecules from the cryoprotectant were also seen in the native Os3BGlu6 structure on the surface of the protein. Four of these were also seen in the Os3BGlu6/*n*-octyl- $\beta$ -D-thioglucopyranoside complex and two in the Os3BGlu6/G2F complex. The single protein molecule in the asymmetric unit lacked close crystal contacts with its neighbors, indicating that Os3BGlu6 is monomeric.



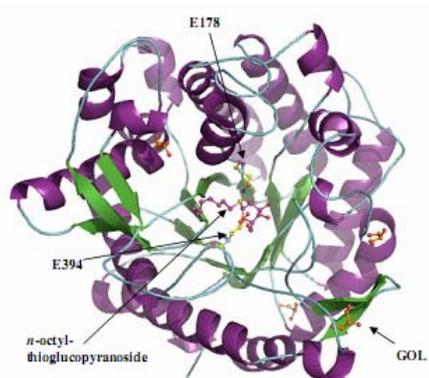
**Figure 3.21** Protein from Os3BGlu6 crystals electroblotted on PVDF (polyvinylidene fluoride) membrane for N-terminal sequencing.

Prior to electroblotting, Os3BGlu6 crystal clusters were dissolved in 1X SDS buffer and separated on a 12% polyacrylamide gel. Lane 1, Bio-Rad low molecular weight marker, lane 2, Dissolved Os3BGlu6 crystals obtained from the condition PEG 5000 MME, 0.1 M Bis Tris, pH 6.5.

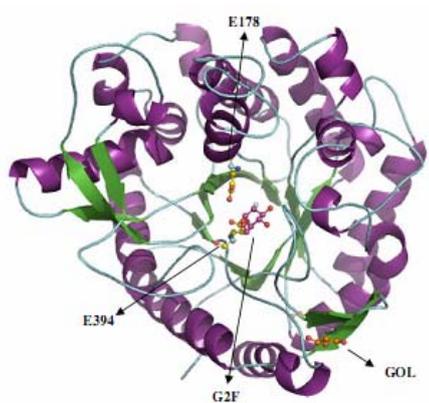
A



B



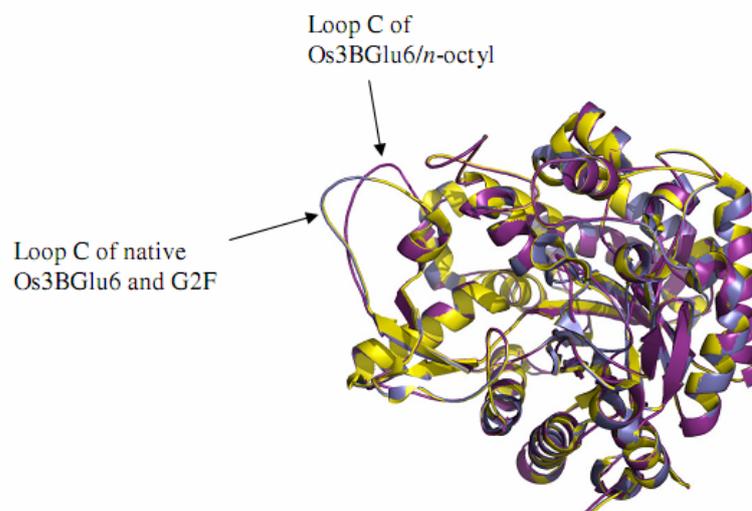
C



**Figure 3.22** Stereo cartoon diagrams of the Os3BGlu6 structures.

(A) Native Os3BGlu6. One molecule of Tris in two conformations is represented by balls and sticks with carbon atoms shown in pink, (B) Os3BGlu6/*n*-octyl- $\beta$ -D-thioglucopyranoside complex. The ligand is represented by balls and sticks with carbon in pink, oxygen in red, and sulphur in yellow, (C) Os3BGlu6/ G2F structure. The catalytic residues and G2F are shown as balls and sticks with carbon in pink, oxygen in red and fluoride in cyan. The alpha-helices are colored purple,  $\beta$ -strands green, and loops cyan. The catalytic residues Glu394 and Glu178 are shown as balls and sticks colored by atoms with carbon in yellow. Glycerol molecules bound to the surface of the protein in the native, Os3BGlu6/*n*-octyl- $\beta$ -D-thioglucopyranoside and Os3BGlu6/G2F structures are colored with orange carbons. The figure was generated by Pymol (Delano Scientific LLC).

The overall structures were very similar, with the typical  $(\beta/\alpha)_8$  barrel. Superimposition of all the three structures gave an RMSD of C-alpha atoms of 0.09 Å between the native and G2F complex structures, 0.16 Å between the native and *n*-octyl- $\beta$ -D-thioglucopyranoside complex structures and 0.18 Å between the G2F and *n*-octyl- $\beta$ -D-thioglucopyranoside complex structures. The differences between the three structures mainly occurred in one highly flexible part of the loop C which helps form the entrance to the active site, in the *n*-octyl- $\beta$ -D-thioglucopyranoside complex structure. Asn330 to Leu338, which are exposed to the solvent on the surface were in a different position when compared to the native and G2F complex structures (Figure 3.23).

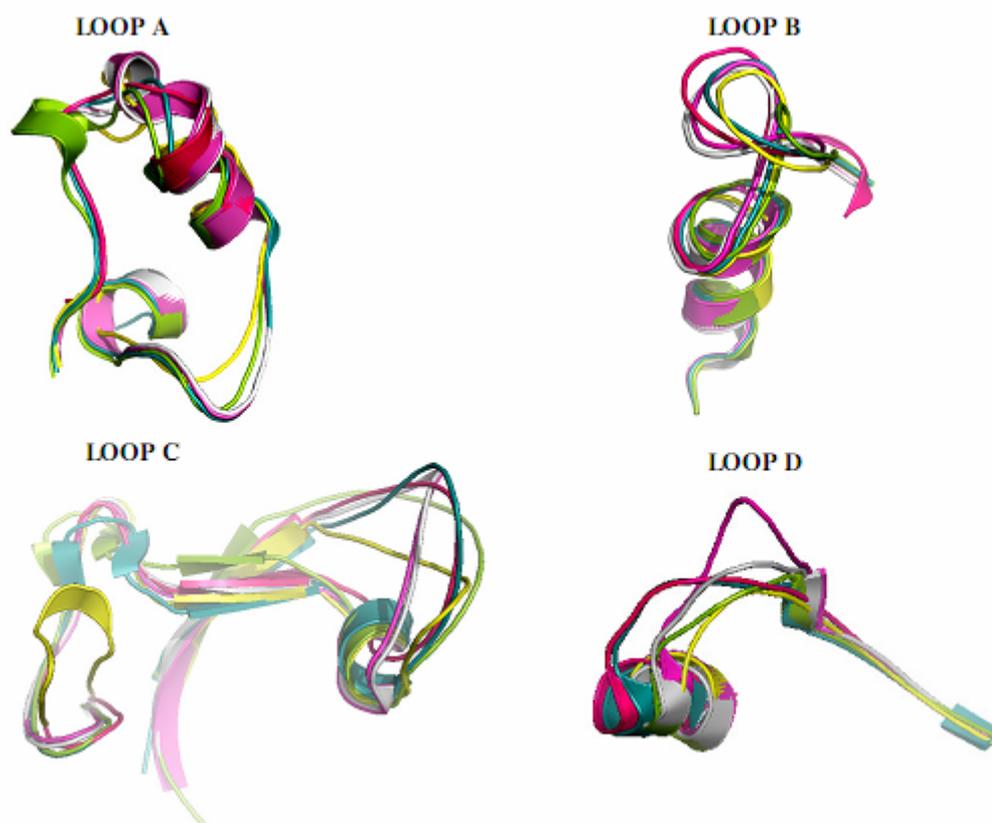


**Figure 3.23** Superimposition of the overall structures of native Os3BGlu6, and the Os3BGlu6/G2F and Os3BGlu6/*n*-octyl-β-D-thioglucopyranoside complexes.

Native Os3BGlu6 is colored blue, Os3BGlu6/G2F is yellow and Os3BGlu6/*n*-octyl-β-D-thioglucopyranoside is purple. The figure was generated by Pymol (Delano Scientific LLC).

In order to visualize the differences in the overall structure of Os3BGlu6 with other enzymes, the C $\alpha$  atoms of known GH1 three-dimensional structures from plants: white clover cyanogenic β-glucosidase, 1CBG (Barrett *et al.*, 1995); white mustard myrosinase, 1E4M (Burmeister *et al.*, 2000); rice Os3BGlu7 β-glucosidase, 2RGL (Chuenchor *et al.*, 2008); *R. serpentina* strictosidine β-glucosidase, 2JF7 (Barleben *et al.*, 2007); sorghum dhurrinase, 1VO2 (Verdoucq *et al.*, 2004); maize ZmGlu1, 1E1E (Czjzek *et al.*, 2001) and wheat β-glucosidase, 2DGA (Sue *et al.*, 2006); a mammal: human cytosolic β-glucosidase, 2ZOX (Noguchi *et al.*, 2008); and bacteria: *Bacillus polymxa* β-glucosidase, 1BGA (Aparicio *et al.*, 1998), and *Lactococcus lactis* 6-phospho-β-galactosidase, 1PBG (Wiesmann *et al.*, 1995), were superimposed on the

structure of Os3BGlu6. The highest sequence identity (approx. 54%) and structural similarity were observed for the plant secreted  $\beta$ -glucosidases, 1CBG and 2RGL, with the root mean squared deviation (RMSD) of their C $\alpha$  being 0.81 Å for both structures. The plant myrosinase (1E4M) and cytosolic (2JF7) and chloroplastic (2DGA, 1VO2, and 1E1E)  $\beta$ -glucosidases and mammalian  $\beta$ -glucosidase had the next highest sequence identities with Os3BGlu6 of 42-47% and C $\alpha$  RMSD of 0.83 Å, 0.87 Å, 0.87 Å, 0.90 Å, 0.96 Å, and 0.96 Å, respectively. Not surprisingly, the bacterial structures 1BGA and 1PBG showed the highest structural diversity with RMSD of 1.13 Å and 1.31 Å, respectively. Although all the above mentioned structures are similar overall and possess a typical  $(\beta/\alpha)_8$  barrel core, many differences reside in the loop regions (Loops A, B, C, and D) around their active sites where the glycone and aglycone binding sites are located (Figure 3.24). Loop A (29-69 in Os3BGlu6) shows difference in the number of residues in six of the structures that were superimposed below. This loop is involved in dimerization of monomers as seen in myrosinase Burmeister *et al.*, 2000). This loop is shorter by one residue in Os3BGlu6 than in Os3BGlu7. Loop B (189-220) in Os3BGlu6 contains the conserved disulfide bridge in GH 1 enzymes. Loop C (321-371) is the longest of the loops, and is four residues longer in Os3BGlu6 than in Os3BGlu7. It forms a major part of the wall to the active site and hence likely to account for differences in substrate recognition and binding. Loop D (397-415) is the shortest and this loop is located far from the active site in Os3BGlu6.



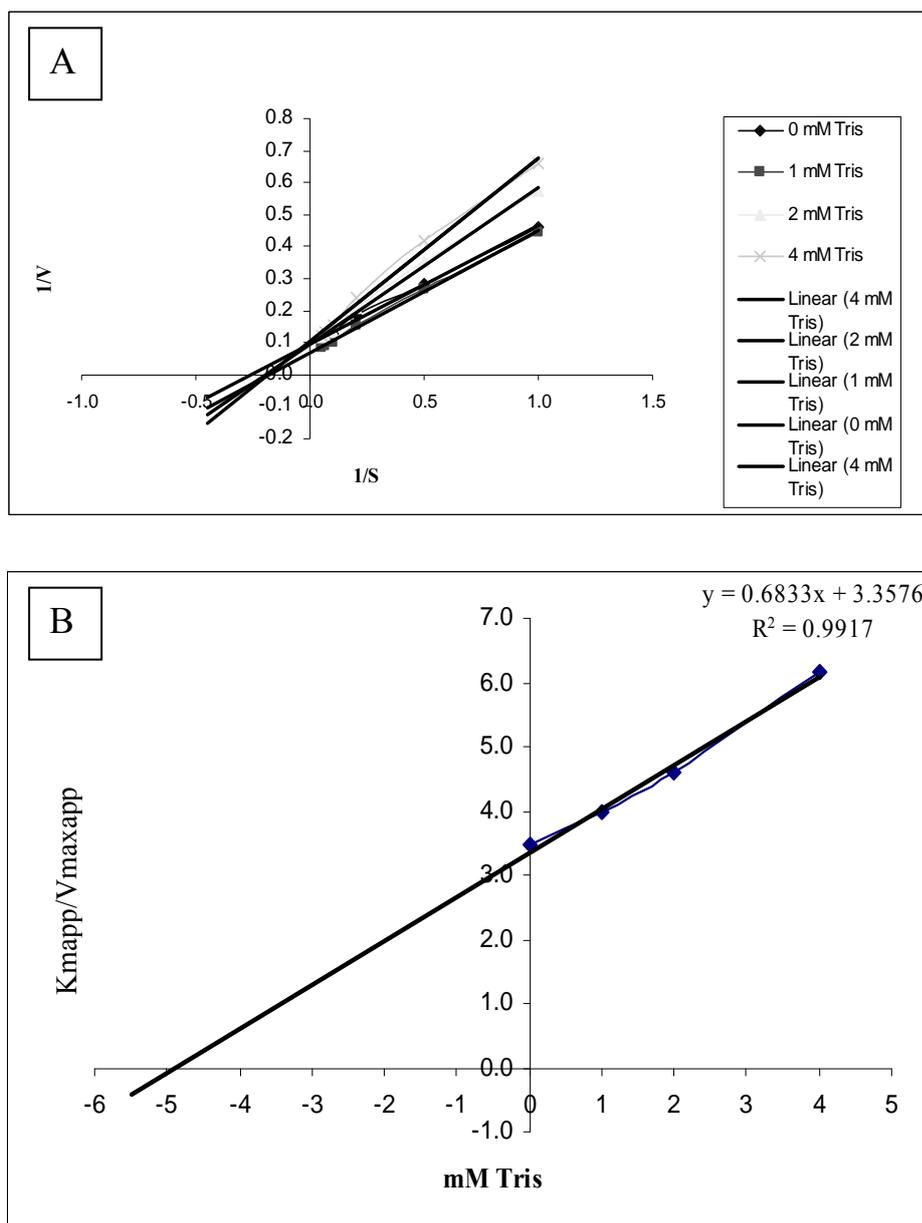
**Figure 3.24** Superimposition of Loops A, B, C, and D from six known GH1 structures.

The loop regions are coloured as follows; Rice Os3BGlu6 PDB code: 3GNO (blue), maize ZmGlu1; PDB code: 1E1E (magenta), white mustard myrosinase; PDB code: 1E4M (yellow), sorghum SbDhr1 dhurrinase; PDB code: 1V02 (grey), rice Os3BGlu7; PDB code: 2RGL (green) and white clover linamarase; PDB code: 1CBG (pink).

### 3.7 Tris inhibition by Os3BGlu6

The active site pocket of native Os3BGlu6 contained one molecule of Tris in two possible conformations. This molecule of Tris could have come from the Tris-HCl buffer in which the protein was stored for crystallization. In order to check the binding efficiency of Tris, the  $K_i$ , competitive inhibition constant, of Os3BGlu6 for Tris was

determined and was found to be  $5.1 \pm 0.2$  mM. The Lineweaver-Burk plots and  $K_{mapp}/V_{maxapp}$  versus Tris concentration are shown in Figure 3.25.

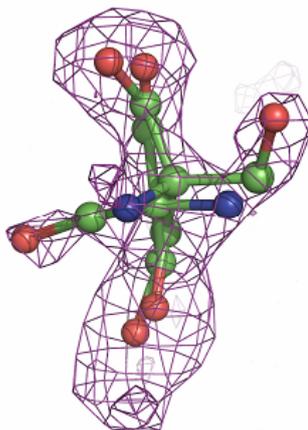


**Figure 3.25** Graphs indicating enzyme inhibition for Tris.

(A) Lineweaver-Burk plot, (B)  $K_{mapp}/V_{maxapp}$  versus Tris concentration.

### 3.8 Active site analysis

#### 3.8.1 Hydrogen bonding interactions of Tris

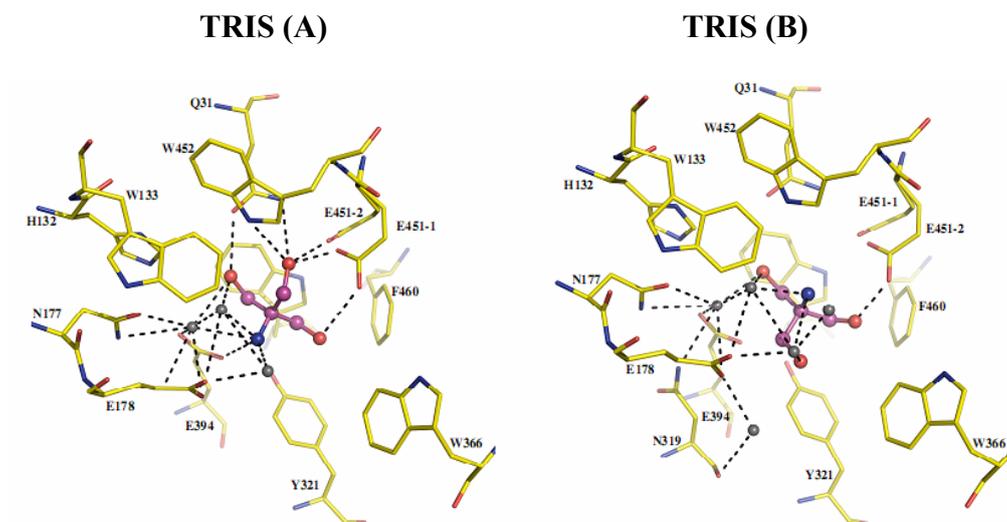


**Figure 3.26** The Fo-Fc omit map of Tris molecule in two conformations calculated from the final model of native structure contoured at 3.0 sigma.

The ligand is represented by balls and sticks with carbons in green, oxygen in red and nitrogen in blue.

The electron density map (Fo-Fc omit map) of the Tris ligand in two conformations is shown in Figure 3.26. The two conformations for Tris (A and B) had average B-factors of 26.91 and 30.28 Å<sup>2</sup>. They formed both direct and water-mediated hydrogen bonds to the neighboring amino acid residues, as shown in Figure 3.25. One of the two conformations of Glu451 makes additional hydrogen bonding interactions with one binding mode of Tris compared with the other. In conformation A of Tris (Figure 3.27 A), Glu451 O<sup>ε1</sup> and O3 of Tris are at distances of 2.61 Å and 2.41 Å in the conformations Glu451-1 and Glu451-2, respectively. Glu451-2 O<sup>ε2</sup> is hydrogen bonded to O1 of Tris at 3.17 Å. The O<sup>ε1</sup> of Gln31 makes a hydrogen bond with O2 of Tris at 3.03 Å and N<sup>ε2</sup> of Gln31 is at a distance of 3.18 Å from O3 of Tris (A). Trp452 N<sup>ε1</sup> is linked to O3 of Tris (A) at 3.11 Å. The catalytic residue Glu394 O<sup>ε1</sup> interacts

with the N of Tris at a distance of 3.18 Å. In addition, three waters (184, 250, and 460) are very close to the ligand and directly interact with both Tris and the Asn177 and Glu178 residues, contributing to hydrogen bonding interactions. A different bonding interaction was seen for the other binding mode of Tris (B) shown in Figure 3.27 B. Glu451-2 O<sup>ε2</sup> is bonded to O2 (B) of Tris at 2.42 Å. This is the only direct hydrogen bond seen between Tris (B) and Glu451. However, in binding mode B, four waters (184, 250, 258, and 460) make hydrogen bonding interactions with Tris and with the residues Asn177, Glu178, and Glu394. The residues binding to Tris, which included Glu451, Gln31, Trp452, Glu394, Glu178, and Asn177, were also involved in glucose binding in the Os3BGlu6/G2F and Os3BGlu6/*n*-octyl-β-D-thioglucopyranoside complex structures, however, in addition, Tyr321 and His132 also interacted with glucose in these structures. All the hydrogen bonds formed between Tris, amino acid residues and waters and their distances are summarized in the Table 3.8.



**Figure 3.27** Active site structure of Os3BGlu6 with a Tris molecule with two different binding modes of Tris, A and B.

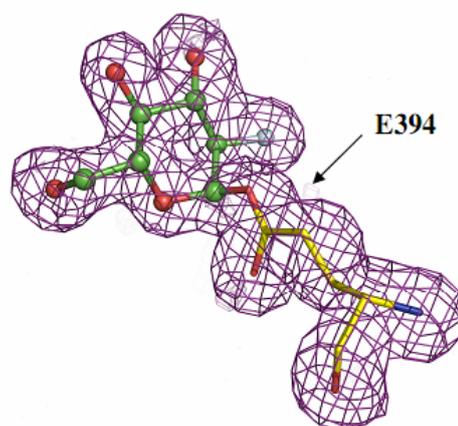
The amino acids surrounding Tris are represented by sticks with carbon in yellow, nitrogen in blue, and oxygen in red. The Tris ligand is represented by balls and sticks with carbon in pink, oxygen in red, and nitrogen in blue. Hydrogen bonding interactions between Tris and surrounding amino acids and waters are shown as black dotted lines. Waters are shown as grey colored spheres. The figure was generated by Pymol (Delano Scientific LLC).

**Table 3.8** Hydrogen bonds occurring between the Tris molecule, binding residues and waters in the enzyme's active site.

TRIS (A)			TRIS (B)		
Residue/Water	Tris/Water	H bonding distance (Å)	Residue/Water	Tris/Water	H bonding distance (Å)
Glu451-2 O <sup>e2</sup>	O1	3.17	Glu451-2 O <sup>e2</sup>	O2	2.42
Glu451-2 O <sup>e1</sup>	O3	2.41	Wat90	Wat250	2.65
Glu451-1 O <sup>e1</sup>	O3	2.61	Wat250	O3	2.47
Gln31 N <sup>e2</sup>	O3	3.18	Glu178 O <sup>e1</sup>	Wat250	2.69
Gln31 O <sup>e1</sup>	O2	3.03	Wat460	N	2.66
Trp452 N <sup>e1</sup>	O3	3.11	Wat460	Wat184	2.42
Glu394 O <sup>e1</sup>	N	3.18	Wat460	O1	2.65
Glu178 O <sup>e1</sup>	Wat250	2.69	Wat184	O1	2.65
Wat250	N	3.21	Glu178 O <sup>e2</sup>	Wat184	3.28
Wat460	N	2.62	Glu394 O <sup>e2</sup>	Wat184	2.54
Wat460	O2	2.84	Asn177 O <sup>D1</sup>	Wat184	2.71
Wat460	Wat184	2.42	Asn177 O	Wat258	3.13
Glu178 O <sup>e2</sup>	Wat184	3.28	Glu178 O <sup>e2</sup>	Wat258	2.77
Wat184	N	2.97			
Wat184	O2	2.92			
Glu394 O <sup>e2</sup>	Wat184	2.54			
Asn177 O <sup>D1</sup>	Wat184	2.71			

The interactions for the two conformations of Tris (A and B) and their distances are shown in Table 3.8.

### 3.8.2 Os3BGlu6/G2F covalent intermediate



**Figure 3.28** The Fo-Fc omit map of 2-fluoroglucoside covalently linked to Glu394 contoured at 3 sigma.

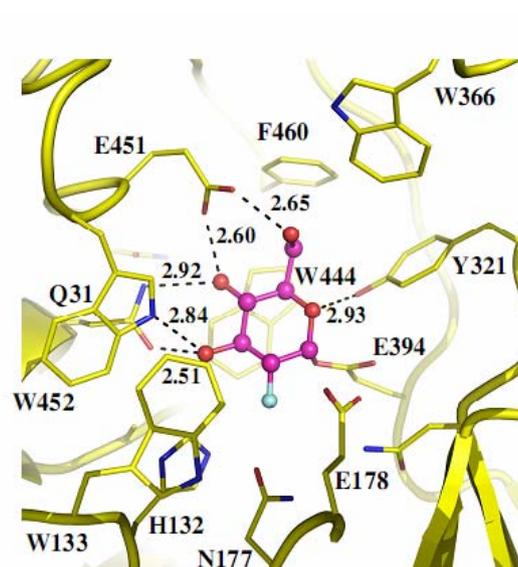
The ligands are represented by balls and sticks with carbon atoms in green, oxygen in red, nitrogen in blue, and fluoride in cyan. The catalytic nucleophile residue Glu394 is shown as sticks with carbon atoms in yellow. The figure was generated by Pymol (Delano Scientific LLC).

The density for the 2-deoxy-2-fluoroglucosyl residue in a relaxed  ${}^4C_1$  chair conformation, covalently bound to the catalytic nucleophile (Figure 3.28) was clearly evident in the glycone-binding subsite. The average B-factor for the 2-deoxy-2-fluoroglucoside was calculated to be  $14.85 \text{ \AA}^2$ . The glycone binding site is surrounded by the universally conserved plant GH1 residues Gln31, His132, Asn177, Glu178, Glu394, Glu451, and Trp452. The residues Glu451, Gln31, Tyr321, and Trp452 formed direct hydrogen bonds with G2F.  $O^{\epsilon 1}$  and  $O^{\epsilon 2}$  of Glu451 were linked to O4 and O6 of the glucose ring at 2.60 and 2.65  $\text{\AA}$ , while Gln31  $O^{\epsilon 1}$  and Gln31  $N^{\epsilon 2}$

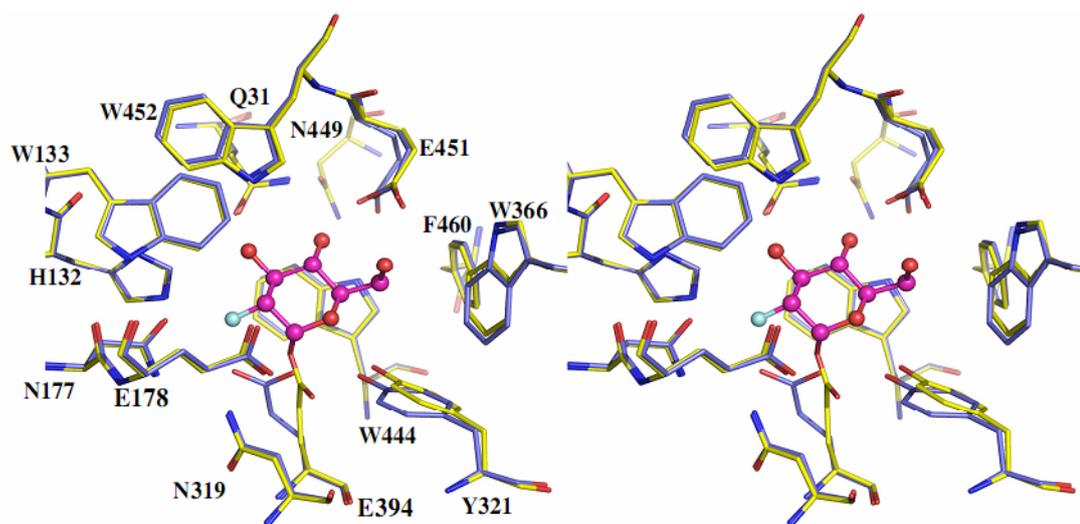
were hydrogen bonded to O3 and O4 at 2.51 and 2.92 Å, respectively. In addition, the hydroxyl group of Tyr321 was hydrogen bonded to O5 (2.93 Å) and Trp452 N<sup>ε1</sup> to O3 (2.84 Å). A typical hydrophobic stacking interaction was observed between the sugar and the indole ring of Trp444, which is also conserved in other GH family 1 glycosidases (Figure 3.29 A).

Superimposition of the G2F complex against the native enzyme (Figure 3.29 B) showed that the hydroxyl group of Tyr321, which is distanced at 2.75 Å from O<sup>ε1</sup> of the catalytic nucleophile Glu394 in the native structure, twists away in the G2F complex, to form a hydrogen bond to O5 of the sugar (2.93 Å). The nucleophile side chain also moves in producing the covalent intermediate, so that Glu394 O<sup>ε1</sup> and O<sup>ε2</sup>, which are positioned at 4.59 Å and 3.88 Å from O<sup>ε2</sup> of Glu178 in the native structure, are distanced at 3.99 and 3.25 Å from O<sup>ε2</sup> of Glu178 in the G2F complex. In addition, the indole ring of Trp452 is slightly shifted inward to allow its nitrogen to hydrogen bond with O3 of the sugar ring at 2.84 Å, in the G2F complex while it is positioned away in the native enzyme. Two possible conformations for Glu451 were seen in the native structure, while just one of the conformations was found in the G2F complex. Instead, one molecule of water was located where the second position of the sidechain was found to be in the native structure.

A



B



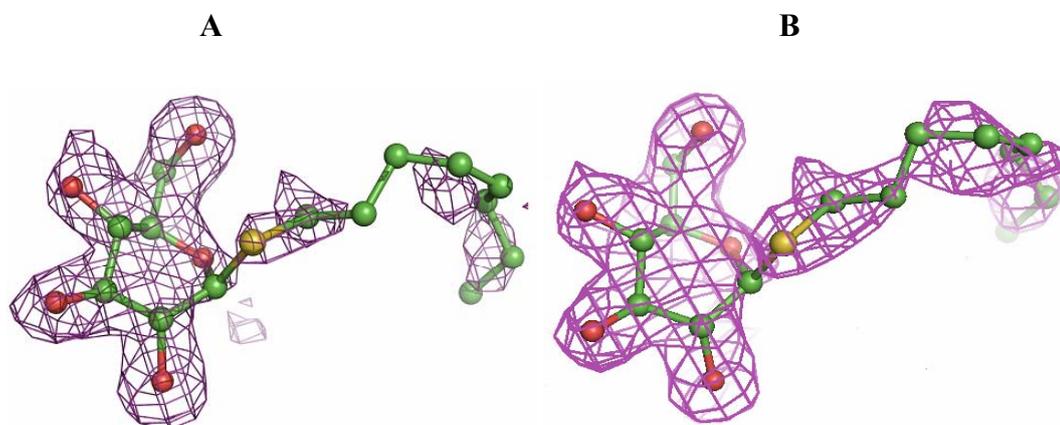
**Figure 3.29** Binding of 2-fluoroglucoside in the Os3BGlu6 active site.

(A) Active site of Os3BGlu6 with G2F covalently bound to the catalytic nucleophile E394 showing protein-ligand binding interactions. The amino acids surrounding the -1 subsite are represented by sticks with carbon in yellow, nitrogen in blue and oxygen in red. The covalently bound G2F ligand is shown in ball and stick representation with carbon in pink, oxygen in red and fluoride in cyan. Hydrogen bonding interactions

between G2F and amino acids are shown as black dotted lines labeled with the bond distances in Å. (B) Stereoview of the superimposition of active site residues at -1 subsite of the native Os3BGlu6 and Os3BGlu6/G2F complex structures. The residues surrounding the -1 subsite are represented by sticks colored with carbons in blue for native Os3BGlu6 and in yellow for Os3BGlu6/G2F. The G2F ligand is represented by balls and sticks with carbon in pink, oxygen in red, and fluoride in cyan. The figure was generated by Pymol (Delano Scientific LLC).

### 3.8.3 Os3BGlu6/*n*-octyl-β-D-thioglucopyranoside

Since *n*-octyl-β-D-glucopyranoside was among the best substrates for Os3BGlu6, the structure of Os3BGlu6 soaked with its nonhydrolysable substrate analogue *n*-octyl-β-D-thioglucopyranoside was determined as the first step toward understanding the mechanism of the aglycone affinity of Os3BGlu6. The density for the sugar was well defined. For the aglycone moiety, the density was strongest for C1', C4', and C8', and was lower for C2', C3', C5', C6', and C7', likely due to the flexibility of the carbon chain (Figure 3.30), which seemed to be able to bind in multiple positions, hence the average B-factor was high (41.25 Å<sup>2</sup>). The inhibitor was located deep in the active site slot, as shown in Figure 3.31 (A).



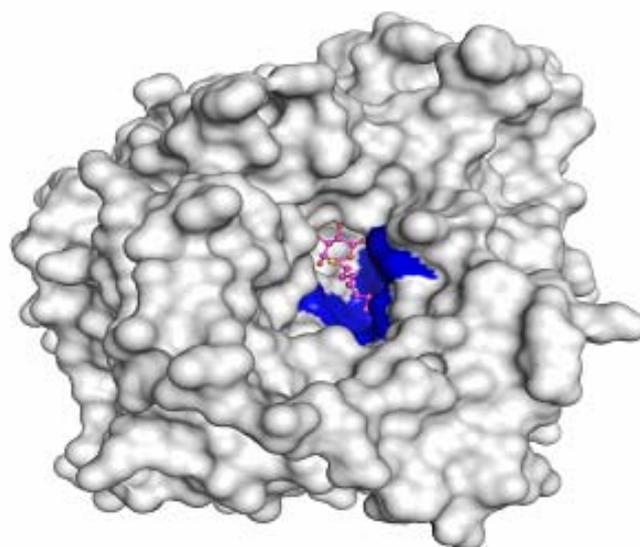
**Figure 3.30** Fo-Fc omit map and 2Fo-Fc map of ligand *n*-octyl- $\beta$ -D-thioglucopyranoside contoured at 3 sigma and 1 sigma.

(A) and (B) The ligand is represented by balls and sticks with carbons in green, oxygen in red, nitrogen in blue, and sulphur in yellow. The figure was generated by Pymol (Delano Scientific LLC).

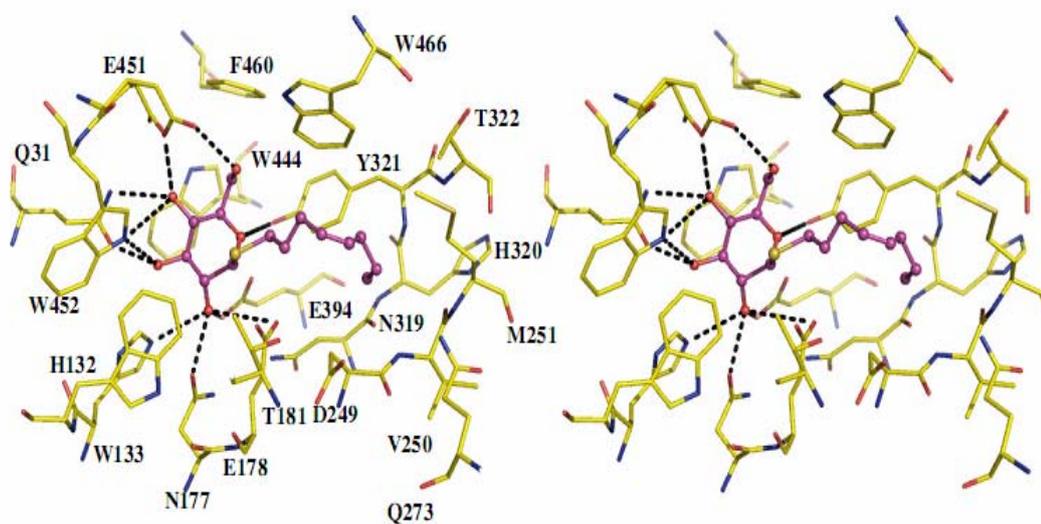
Although both the relaxed  ${}^4C_1$  chair and less energetically favorable  ${}^1S_3$  skew boat conformations fit the density well, the distance of O $^{\epsilon 1}$  of the acid-base catalyst Glu178 to the sulfur in  ${}^4C_1$  was 2.13 Å, which was unacceptably close. However, this distance was 3.07 Å in the  ${}^1S_3$  conformation, putting the sulfur within distance to accept a hydrogen bond and proton transfer from the catalytic acid. The glycone residue was in a similar position to that of G2F in its complex and formed similar hydrogen bonding interactions, but it had more possible H-bonding partners, and interactions between the O2 hydroxyl and His132, Asn177 and Glu178 were also evident (Figure 3.31 B). The binding interactions between the sugar ring and the surrounding amino acid residues are listed in the Table 3.9. The aglycone part is flanked by hydrophobic residues and a few polar residues. The *n*-octyl chain is

surrounded by hydrophobic contacts from Tyr321 and Trp366 on one side, and Val250 and Met251, which also form the entrance to the active site, on the other side of the chain. Tyr321 is also a part of the glycone recognition and binding site, and this residue is conserved in GH family 1 structures. Trp366, Val250, and Met251 seem to be most important for the interaction with the *n*-octyl chain, since they are at a distance of less than 4 Å.

A



B



**Figure 3.31** Os3BGlu6 interactions with *n*-octyl- $\beta$ -D-thioglucopyranoside.

(A) Surface view of the Os3BGlu6 structure showing the active site cleft with the ligand *n*-octyl- $\beta$ -D-thioglucopyranoside. The hydrophobic amino acid residues making contacts with the *n*-octyl chain are represented in blue; (B) Stereoview of Glycone and aglycone interactions of *n*-octyl- $\beta$ -D-thioglucopyranoside with Os3BGlu6 residues.

The amino acids are represented by sticks with carbon in yellow, nitrogen in blue, and oxygen in red. The ligand is shown in ball and stick representation with carbon in pink, oxygen in red, and sulphur in yellow. Hydrogen bonding interactions between the ligand and amino acid residues are shown as black dotted lines. The figure was generated by Pymol (Delano Scientific LLC).

**Table 3.9** Hydrogen bonds occurring between *n*-octyl-thioglucopyranoside and amino acid residues in the enzyme's active site.

<b>Residue</b>	<b><i>n</i>-octyl-<math>\beta</math>-D-thioglucopyranoside</b>	<b>H bonding distance (Å)</b>
Tyr 321 OH	O5	3.08
Glu451-1 O <sup><math>\epsilon</math>1</sup>	O6	2.83
Glu451-1 O <sup><math>\epsilon</math>1</sup>	O4	2.56
Glu451-2 O <sup><math>\epsilon</math>1</sup>	O4	2.62
Gln31 O <sup><math>\epsilon</math>1</sup>	O4	3.03
Trp452 N <sup><math>\epsilon</math>1</sup>	O3	2.92
Gln31 O <sup><math>\epsilon</math>1</sup>	O3	2.53
His132 N <sup><math>\epsilon</math>2</sup>	O2	3.32
Asn177 O <sup>D1</sup>	O2	2.83
Glu394 O <sup><math>\epsilon</math>2</sup>	O2	2.58

## CHAPTER IV

### DISCUSSION

#### 4.1 Cloning, expression, and purification of recombinant rice Os3BGlu6 $\beta$ -glucosidase

The *Os3BGlu6* cDNA was isolated from a cold-treated rice cDNA library by subtractive methods, but appeared to be a false positive. Results from northern blot analysis did not show any induction upon cold treatment, and the mRNA were present at low constitutive levels in rice seedlings and mature plant parts, which made analysis by this method difficult (Takashi Akiyama, unpublished data). A view of the UniGene entry Os.15799 for this gene (Os03g0212800) indicates low transcript levels of 20 to 30 transcripts per million in callus, leaf and stem, and no detection in flower, root, and seed. The GEO microarray profile indicates this gene is in the 98<sup>th</sup> percentile of rice genes for expression in 4-day coleoptiles, and expression is decreased over 10-fold in anoxic stress to put it in the 79<sup>th</sup> to 86<sup>th</sup> percentile (Lasanthi-Kudahettige *et al.*, 2007). The physiological relevance of this expression is still unclear, however.

The cDNA encoding glycosyl hydrolase family 1  $\beta$ -glucosidase Os3BGlu6 was successfully amplified by PCR. The forward primer for PCR amplification contained the additional four bases CACC, which in turn annealed to GTGG overhang in the pENTR<sup>TM</sup>/TEV/D-TOPO<sup>®</sup> vector to allow directional cloning of the cDNA. The lambda derived recombination sequences attL1 and attL2 in the pENTR<sup>TM</sup>/TEV/D-

TOPO<sup>®</sup> vector ensured for efficient subcloning of the *Os3BGlu6* cDNA into pET32a(+)/DEST expression vector by an LR clonase recombination reaction.

When expressed in *E. coli* strain Origami(DE3) cells, Os3BGlu6 was produced as a soluble N-terminal thioredoxin and hexahistidine tag fusion protein, which was catalytically active. During extraction, to prevent the protein from proteolysis, soybean trypsin inhibitor was added to extraction buffer. Os3BGlu6 was stable from proteolytic degradation when stored at 4°C or 30°C. This fusion tagged protein was purified by immobilized metal affinity chromatography with Co<sup>2+</sup> resin. Tobacco etch virus (TEV) protease in the form of a hexahistidine tagged, maltose binding protein (MBP) fusion protein (auto inactivation-resistant mutant S219V, Kapust *et al.*, 2001) was used to cut off the N-terminal tag, to yield mature Os3BGlu6 protein with four of the amino acid residues retained from the cloning vector. By removing the N-terminal tag and TEV protease in a second IMAC purification, high concentrations of Os3BGlu6 with >95% purity could be obtained for functional and structural characterization.

Optimization to identify suitable expression conditions for Os3BGlu6 yielded some interesting results. It was found that Os3BGlu6 was expressed in its soluble form at 20°C, even without any IPTG. Induction was detected at 4 hr, but reached its optimum at 16 hr. The activities of Os3BGlu6 expressed at 0 mM and 0.4 mM IPTG were almost the same when tested with *p*NPGlc substrate. This kind of expression could have occurred due to the leaky promoter. Origami(DE3) is a λDE3 lysogen carrying a T7 RNA polymerase gene needed for target protein expression in pET32a(+) in a lambda prophage under the control of a lac UV5 promoter (Novagen). The leaky lac UV5 promoter is likely to have expressed T7 RNA polymerase. *E. coli*

RNA polymerase produces small amounts of T7 RNA polymerase in the absence of any inducer due to the leaky nature of the lac UV5 promoter (Studier *et al.*, 1990). However, whether more Os3BGlu6 was produced at low temperature was not distinguishable from the SDS PAGE gels, since the expression patterns at higher temperatures (30°C and 37°C) in both soluble and insoluble fractions were not clear, so it could not be determined whether Os3BGlu6 was cold induced. A similar result, where protein was expressed under the leaky promoter in pET system, was observed in Shikimate kinase and 5-enolpyruvylshikimate3-phosphate synthase enzyme from *Mycobacterium tuberculosis*, although that the protein was expressed at 37°C. (Oliveira *et al.*, 2001).

Comparison with other cold induction vector systems would be useful to establish the recombinant protein expression in *E. coli*. A pCold expression vector, which is cold-shock induced vector, is routinely used for recombinant protein expression under control of the cspA promoter. This vector contains the lac operator downstream of a cspA promoter so that expression is strictly controlled and constitutive expression of the lacI gene prevents leaky expression of the cloned genes (Hayashia and Kojima, 2008).

Expression of recombinant proteins is favored at colder temperatures as they tend to yield active proteins by lessening the protein denaturation. Also expressing proteins at higher temperatures, proteins tend to form inclusion bodies with less production of soluble active protein. This is particularly true for eukaryotic proteins, since they are normally synthesized at a slower rate than *E. coli* proteins, and may not have time to fold properly. Os3BGlu7 (Opassirri *et al.*, 2003) and Os4BGlu12 (Opassirri *et al.*, 2006) had to be expressed at 20°C for optimal activity. This

phenomenon was also observed in beta-primeverosidase from tea leaves (Mizutani *et al.*, 2002) where the recombinant primeverosidase in *E. coli* was produced as inclusion bodies at 37°C, but were expressed in soluble form when expressed at 22°C.

## 4.2 Biochemical properties of Os3BGlu6

Os3BGlu6 had a molecular weight of approximately 66 kDa for the fusion protein and 55 kDa after the fusion tag was removed, as observed on SDS PAGE (Figure 3.9). These matched the molecular weights predicted from the sequence of 67,523 Da for the fusion protein and 58,538 Da for the free protein reasonably well. The rice cell wall bound  $\beta$ -glucosidase from rice also is similar to Os3BGlu6, with a molecular weight of 56 kDa (Akiyama *et al.*, 1998). The predicted pI of Os3BGlu6 was found to be 6.36, indicating that Os3BGlu6 is a slightly acidic protein, different from rice Os3BGlu7, which had a basic predicted pI of 9. The rice and barley beta glucosidases, which hydrolyzed oligosaccharides similar to Os3BGlu7 showed basic pI values (Leah *et al.*, 1995; Akiyama *et al.*, 1998).

Os3BGlu6 had three possible N-linked glycosylation sites at positions Asn258, Asn329 and Asn339. All the three glycosylation sites are located on the surface of the protein, where glycosylation is likely to occur. It is known that glycosylation of proteins is important in directing the proteins to their cellular location, in addition to maintaining the folding and stability. Terashima *et al.* (1994) showed that the removal of one glycosylation site in rice alpha-amylase had a drastic effect on the thermostability of the enzyme. A barley recombinant wild-type alpha-glycosidase expressed in *E. coli* had increased thermostability of the enzyme, indicating that the position of the glycosylation site, rather than the number are important (Olsen and Thomsen, 1991). The tea primiverosidase (Mizutani *et al.*, 1999) and linamarase from

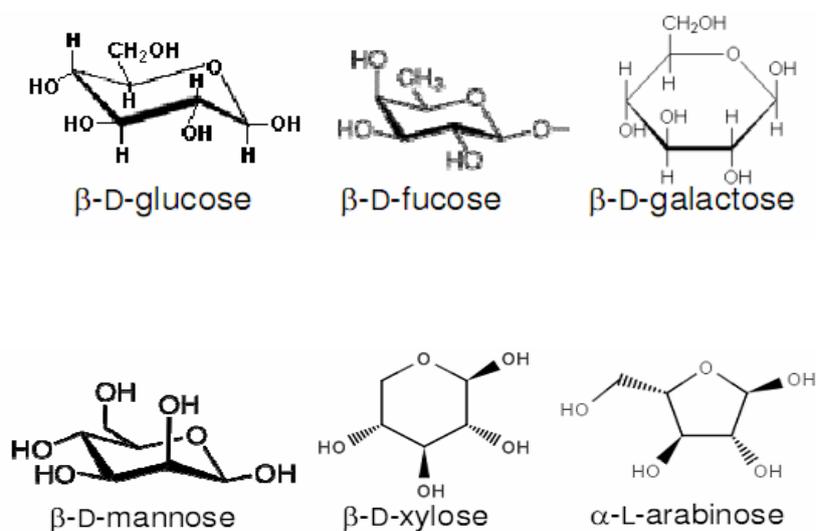
cassava (Keresztessy *et al.*, 1996) expressed in *E. coli* were produced as soluble active proteins, despite the fact they were not glycosylated, suggesting that glycosylation may not be required for the activity of these plant  $\beta$ -glucosidases.

The optimum pH for this enzyme was 4.5 (Figure 3.8), which is in the range commonly seen in plant  $\beta$ -glucosidases (Esen, 1993), and suggestive of a role in the acidic vacuole or apoplast. At the optimum pH, we would expect the acid/base catalyst to be protonated and the nucleophile to be deprotonated. The pH dependent activity for the enzyme is related to the ionization states of the catalytic residues, nucleophile Glu394 and the acid/base catalyst Glu178 involved in catalysis. In all retaining  $\beta$ -glycosyl hydrolases, the high pKa of the acid catalyst is expected to drop by two to three pH units during the hydrolysis (White and Rose, 1997). The high pKa of the acid catalyst is restored after the glycone product is released after the nucleophilic attack on the carbonyl carbon. High pKa of Glu178 is due to the contribution of electrostatic interactions with neighbouring carboxylate groups, seen in xylanases (Davoodi *et al.*, 1995), or dependent upon the local pronounced hydrophobic environment (Keresztessy *et al.*, 1994).

Os3BGlu6 showed optimum activity over a broad temperature range of 40-55° (Figure 3.10). Although some residual activity was seen at high temperature, boiling at 100°C completely inactivated the enzyme. Os3BGlu6 was most stable at 20-30°C (Figure 3.11). The temperature optimum is related to the enzyme's denaturation and catalytic reaction mechanism (Dixon and Webb, 1979). Temperature stability depends on individual enzyme and reflects the forces stabilizing the tertiary and quaternary structures.

### 4.3 Substrate specificity of Os3BGlu6

The glycone specificity of Os3BGlu6 is somewhat different from previously characterized rice  $\beta$ -glucosidases, such as Os3BGlu7 (rice BGlu1). The chemical structures of sugars are shown in Figure 4.1.



**Figure 4.1** Chemical structures of some sugars, used in determining the substrate specificity of Os3BGlu6 in their glycosidic forms.

Os3BGlu7 showed a slight preference for *p*NPFuc over *p*NPGlc due to a higher  $k_{\text{cat}}$  for *p*NPFuc ( $13.34 \text{ s}^{-1}$ ) (Opassiri *et al.*, 2004), whereas for Os3BGlu6, the difference is larger due to a 10-fold lower  $K_m$  for *p*NPFuc than *p*NPGlc. This appears to be primarily due to a higher tolerance for the equatorial position of the 4' hydroxyl group in Os3BGlu6, since it had a  $K_m$  for *p*NPGal similar to that for *p*NPGlc, whereas Os3BGlu7 had a  $K_m$  over 10-fold higher for *p*NPGal than *p*NPGlc. To understand why Os3BGlu6 showed high *p*NPFuc hydrolysis activity, a  $\beta$ -D-fucoside was modeled and superimposed with the Os3BGlu6/G2F complex structure in the active site of

Os3BGlu6 to see the binding with the residues. The 6<sup>th</sup> hydroxyl is hydrogen bonded to Glu451 in Os3BGlu6/G2F structure but is replaced by a CH<sub>3</sub> group in β-D-fucoside. The lack of formation of hydrogen bond at this position would make fucose flexible to move around in the active site making better binding interactions with the residues in the active site, but however, there were no obvious favorable interactions to be gained by replacing glucose with fucose when the structures were superimposed. High β-fucosidase activity was also seen in Thai rosewood (*Dalbergia cochinchinensis* Pierre) β-glucosidase with very similar  $K_m$  values (0.54 mM for *p*NPFC and 5 mM for *p*NPGLC) to Os3BGlu6 (Svasti *et al.*, 1999), but it hydrolyzed *p*NP-β-D-xyloside and α-L-arabinoside better than Os3BGlu6, which hydrolyzed these substrates poorly. Other plant enzymes were shown to have high β-fucosidase activity include rice Os4BGlu12, which had a higher relative activity (118%) when compared with β-glucosidase activity, and a β-glucosidase isolated from *Juglans regia* L. which preferred *p*NP-β-D-fucoside ( $K_m=5.7$  mM) to *p*NP-β-D-glucoside ( $K_m=19.9$  mM). The incidence of β-fucosidase activity in plants is unexpected, because no β-fucosides have been identified in plant tissues, although one report indicated a strict β-fucosidase from lettuce (*Lactuca sativa*) released fucose from cell-wall extracts (Giordani and Noat, 1998). Comparison of hydrolysis other *p*NP glycosides with other rice β-glucosidases show differences in their substrate specificities. Os3BGlu7 hydrolyzed *p*NP-β-D-mannopyranoside with relatively high catalytic efficiency ( $k_{cat}/K_m=1.01$  mM<sup>-1</sup>s<sup>-1</sup>) compared to Os3BGlu6, which showed slow hydrolysis of this substrate, indicating that the equatorial position of the 2' hydroxyl group of glycosyl moiety is more critical to Os3BGlu6. Os3BGlu6 is similar to Os4BGlu12 in this regard.

In hydrolysis of gluco-oligosaccharides, Os3BGlu6 showed a different pattern when compared with other isoenzymes in rice. The hydrolysis of  $\beta$ -(1 $\rightarrow$ 3)-laminaribiose ( $k_{cat}/K_m=1.7 \text{ mM}^{-1}\text{s}^{-1}$ ) and laminaritriose ( $k_{cat}/K_m=0.18 \text{ mM}^{-1}\text{s}^{-1}$ ), followed a similar pattern to that of previously characterized rice Os3BGlu7 and Os4BGlu12 enzymes (Opassiri *et al.*, 2004; Opassiri *et al.*, 2006), but Os3BGlu6 hydrolyzed the longer  $\beta$ -(1 $\rightarrow$ 4)-linked cellooligosaccharides with very low efficiency, while Os3BGlu7 and Os4BGlu12 efficiently hydrolyzed cellooligosaccharides with DP>2. Os3BGlu6 showed preference for hydrolysis of  $\beta$ -(1 $\rightarrow$ 2)-linked sophorose, however. The  $k_{cat}/K_m$  of Os3BGlu6 for sophorose was higher ( $0.96 \text{ mM}^{-1}\text{s}^{-1}$ ) when compared with Os3BGlu7 ( $0.42 \text{ mM}^{-1}\text{s}^{-1}$ ). Barley BGQ60  $\beta$ -glucosidase (Leah *et al.*, 1995) was also shown to hydrolyze sophorose at a  $K_m$  of 6.3 mM, in addition to hydrolyzing other disaccharides. These results indicate that Os3BGlu6 hydrolyzed  $\beta$ -(1 $\rightarrow$ 3)- and  $\beta$ -(1 $\rightarrow$ 2)- linked glucooligosaccharides well, and  $\beta$ -(1 $\rightarrow$ 6)- and  $\beta$ -(1 $\rightarrow$ 4)-linked oligosaccharides much more slowly, suggestive of its role in hydrolysis of  $\beta$ -(1 $\rightarrow$ 3)- and  $\beta$ -(1 $\rightarrow$ 2)-linked disaccharides in cell wall metabolism.

Alkyl glycosides have been known for their use as biodegradable detergents. Os3BGlu6 hydrolyzed *n*-octyl- $\beta$ -D-glucoside with high catalytic efficiency of  $2.7 \text{ mM}^{-1}\text{s}^{-1}$ . Os3BGlu7 also hydrolyzed this substrate, although relatively with lower efficiency ( $1.25 \text{ mM}^{-1}\text{s}^{-1}$ ). In addition to hydrolysis, some plant enzymes like Thai rosewood  $\beta$ -glucosidase (Lirdprapamongkol *et al.*, 2000) and almond  $\beta$ -glucosidase (Romeu *et al.*, 1994) can also be used to synthesize alkyl glycosides of various chain lengths, by reverse hydrolysis or transglycosylation, which suggests an application for Os3BGlu6. Os3BGlu6 might play a role in hydrolyzing hydrophobic glycosides in the plants. Among the natural glycosides, Os3BGlu6 hydrolyzed apigenin-7-O- $\beta$ -D-

glucoside, the isoflavones genistin, glycitin and diadzein, the phenolic compounds arbutin, salicin and esculin and the monolignols coumaryl-glucoside and coniferin. These substrates were however hydrolyzed relatively poorly, hence it is difficult to identify the functions of Os3BGlu6 based on the hydrolysis of these natural substrates. However, the relatively high rate of hydrolysis of the 7-O- $\beta$ -linked flavonoid apigenin and isoflavonoids (diadzin, genistin and glycitin), but lack of hydrolysis of the 8-O- $\beta$ -linked flavonoid gossipin, suggests a 7-O- $\beta$ -linked flavonoid could be a substrate.

#### **4.4 Os3BGlu6 crystallization**

In order to determine three-dimensional structures by x-ray crystallography, formidably high concentrations of pure proteins are required. In this study, several crystallization techniques were applied to produce good quality crystals of Os3BGlu6, such as microbatch and hanging drop vapor diffusion. During later stages of optimization, microseeding by streak seeding was employed to produce native crystals. By the streak seeding, single large crystals of native Os3BGlu6 were obtained at a faster rate, with improved quality by avoiding the spontaneous nucleation rates. Once the native crystals were obtained, 2-fluoroglucoside and *n*-octyl-thioglucopyranoside, were soaked into the native crystals. *n*-octyl-thioglucopyranoside is a nonionic detergent and possesses amphipathic properties, which means that the detergent monomers have limited solubility in water, and at concentrations above the critical micelle concentration, they partition into a micellar phase (Gilbert, 2007). Since the critical micelle concentration of *n*-octyl thioglucoside was found to be 9 mM (Tsuchiya and Saito, 1984), high concentrations could not be used to soak crystals of Os3BGlu6 with the ligand.

Os3BGlu6 was stored in 20 mM Tris-HCl, pH 8.0, after protein purification and protein in this buffer was used for crystallization. Unexpected electron density was observed in the native structure active site which happened to be a molecule of Tris in the active site, in two possible confirmations. In order to verify that Tris could bind to the active site of the enzyme, inhibition of Os3BGlu6 by Tris was examined. Tris was a weak inhibitor with a competitive  $K_i$  of 5.1 mM, indicating that, at the equilibration phase of crystallization experiment in hanging drop, when the concentration of Tris could be expected to be 15-20 mM, it could be bound to 75-80% of the protein molecules in the crystals.

## 4.5 Overall structure

### 4.5.1 Unit cell differences

The unit cell dimensions of crystal structure of Os3BGlu6/*n*-octyl-thioglucopyranoside had the unit cell dimension  $c=111.3 \text{ \AA}$ , which was greater than either native or G2F complex crystal structures ( $c=101.7 \text{ \AA}$  and  $101.88 \text{ \AA}$ ), though the other dimensions were similar. To test if the difference in the one dimension was indeed due to the ligand, 2 crystals of native Os3BGlu6, 3 crystals of Os3BGlu6/G2F complex and 6 crystals of Os3BGlu6/*n*-octyl-thioglucopyranoside were diffracted and all the crystals displayed similar unit cell dimensions as shown in Table 4.1.

**Table 4.1** Unit cell dimensions for Os3BGlu6 crystals.

Unit cell dimension	a (Å)	b (Å)	c (Å)
Native Os3BGlu6	56.8	90.4	101.7
	56.8	90.7	101.8
Os3BGlu6/G2F	56.6	90.7	101.8
	56.7	90.6	102.0
	56.6	90.5	101.8
Os3BGlu6/ <i>n</i> -octyl-thioglucopyranoside	57.2	91.9	111.4
	57.2	91.3	110.8
	57.2	91.1	110.4
	57.2	91.3	110.8
	57.2	91.3	110.8
	57.1	91.0	111.3

This difference in the unit cell dimensions appears to have been likely due to the presence of the *n*-octyl-thioglucopyranoside ligand. Similar results were observed with native protein crystals of concanavalin B and crystals grown with  $\beta$ -octyl-glucoside (McPherson *et al.*, 1985). The unit cell dimensions of native crystals in presence of the  $\beta$ -octyl glucoside had a greater dimension in one of the cell edges than that of native crystals. This happened in concanavalin B protein due to the presence of the detergent, where the unit cell expanded by about 1% on c and 0.5% on a and b.

Ramachandran plots for all three structures showed Ala65 and Trp452 fell in the generously allowed region. These two residues had well defined electron density in the structure. Ala65 is located along the surface of O3BGlu6 forming interactions with the neighboring waters at the solvent surface while Trp452 is located

in the active site and formed hydrogen bonding interactions with the ligand and the surrounding amino acid residues. The corresponding residues were also seen to have similar torsion angles in other GH family 1 structures including 1CBG (Barrett *et al.*, 1998) and 2DGA (Sue *et al.*, 2006), in which they fell in the same region of the Ramachandran plot.

#### **4.6 Crystal structures of Os3BGlu6/G2F intermediate and Os3BGlu6/*n*-octyl- $\beta$ -D-thioglucopyranoside**

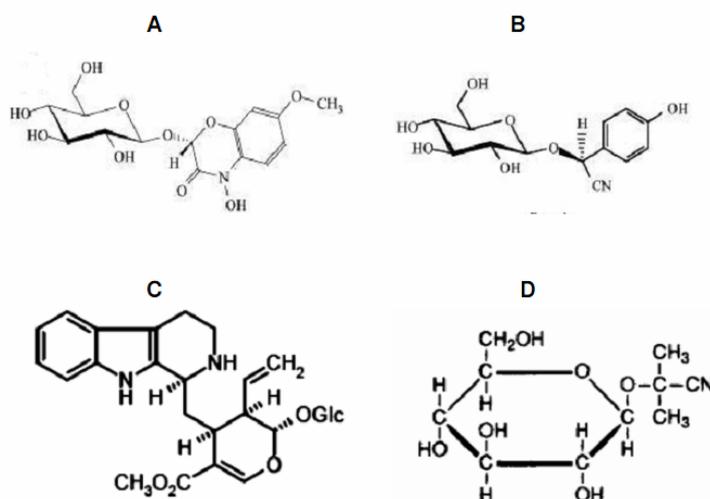
The crystal structure of 2-fluoroglucoside which is covalently bound to the catalytic nucleophile, Glu394, in the glycone binding pocket of Os3BGlu6 was obtained. The G2F complex corresponds to the covalent enzyme intermediate of the double displacement mechanism of GH1 enzymes (Withers *et al.*, 1987). The position of the glycosyl moiety was stabilized by 6 hydrogen bonds from the surrounding residues of the active site. The catalytic nucleophile, Glu394, was found to have moved to bind to the C1 of the glycosyl moiety and the residue Tyr321 also moved, as these movements may have been necessary for the G2F to form the covalent intermediate. This was similarly observed in other GH1 covalent intermediates. (Burmeister *et al.*, 1997; Zechel *et al.*, 2003 and Isorna *et al.*, 2007).

Os3BGlu6 hydrolyzed the nonionic surfactants, *n*-heptyl and *n*-octyl glucosides relatively well. In order to investigate the structural basis for the enzyme's preference for hydrolyzing hydrophobic substrates, a substrate analogue of *n*-octyl glycoside, *n*-octyl- $\beta$ -D-thioglucopyranoside was introduced into Os3BGlu6 crystals. The glycosyl group was stabilized by 10 hydrogen bonds contributed by seven residues making direct hydrogen bonds, while the *n*-octyl chain interacted via

hydrophobic contacts from Tyr321 and Trp366 on one side, and Val250 and Met251, which also forms the entrance to the active site, on the other side of the chain.

#### 4.7 Comparison of substrate binding and recognition between Os3BGlu6 and other GH family 1 enzymes

The aglycone recognition and binding between Os3BGlu6 and other GH1 enzymes have been compared to understand and provide a structural basis for the preferences for different substrates among these enzymes, which in turn determines the functions of these enzymes. The chemical structures of some of the natural substrates, hydrolyzed by GH1 enzymes, DIMBOA-Glc, dhurrin, strictosidine and linamarin are shown in Figure 4.2 below.



**Figure 4.2** Natural substrates hydrolyzed by plant GH1 enzymes. (A) DIMBOA-Glc; (B), dhurrin; (C), strictosidine and (D), linamarin.

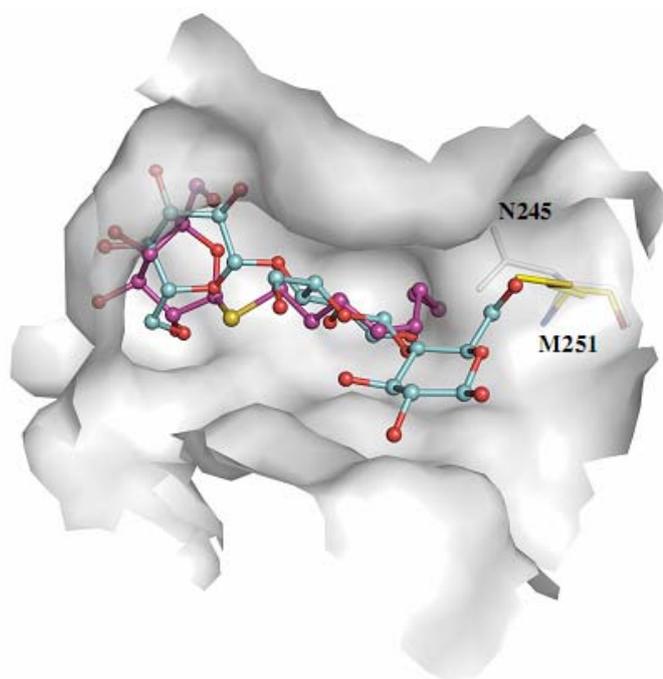
#### 4.7.1 Os3BGlu6 and Os3BGlu7

Os3BGlu6 hydrolyzed disaccharides better than longer chain oligosaccharides, while rice Os3BGlu7 hydrolyzed  $\beta$ -(1 $\rightarrow$ 4)-linked oligosaccharides with increasing efficiency, up to 6 glucosyl residues (Opassiri *et al.*, 2004). However, both enzymes hydrolyze cellobiose poorly. In order to understand these differences, the Os3BGlu6 structure was superimposed with those of a Os3BGlu7 docked with the  $\beta$ -(1 $\rightarrow$ 4)-linked substrate cellotriose (Chuenchor *et al.*, 2008) (Figure 4.3).

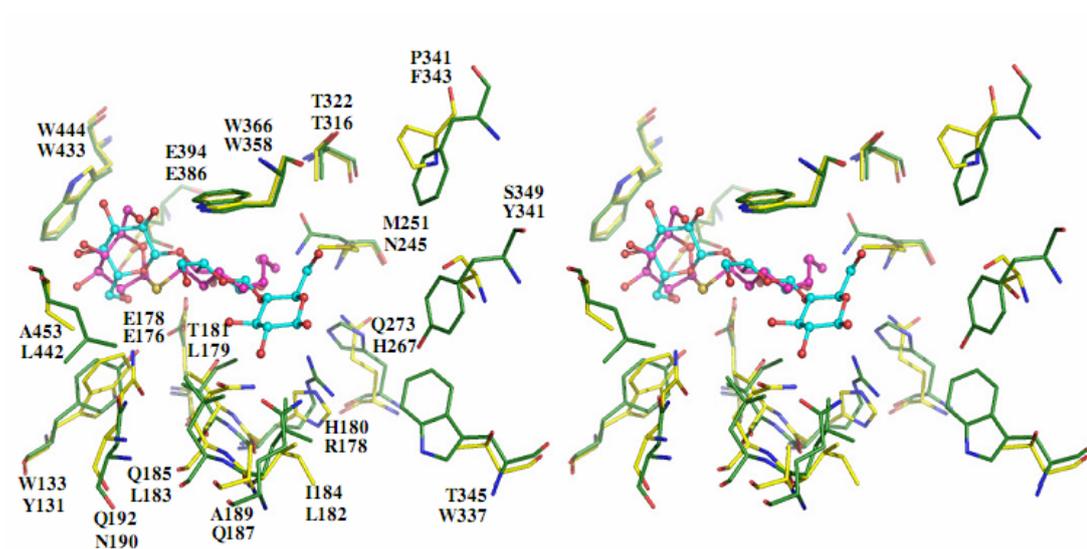
Superimposition of these two proteins' structures revealed differences in the shape of the active site and the residues lining the aglycone binding region. While Os3BGlu7 had a deep, narrow and a relatively straight binding cleft, that of Os3BGlu6 is shorter and narrower. The glycone binding pocket is 7.5 Å deep and 9 Å wide. The narrowest part of aglycone binding site, i.e. at the entrance to the active site is about 5.3 Å across, due to the extension of Met251 into the space corresponding to the +2 subsite in Os3BGlu7, while the broadest region is 7.6 Å. The distance from the deepest part of the active site slot to the surface entrance is 14 Å, which is shorter by 4 Å from that of Os3BGlu7 (18 Å). The difference may correlate to the Os3BGlu6 enzyme's preference for shorter oligosaccharides. The nonreducing end glucosyl residue was in nearly the same position in both structures and had similar hydrogen bonding to the amino acid residues surrounding the -1 subsite. However, many differences were observed with respect to the residues involved in binding the aglycone, which included Trp133, Thr181, Ile184, Gln185, Gln192, Asp249, Met251, and Ala453 in Os3BGlu6 corresponding to Tyr131, Ile179, Leu182, Leu183, Asn190, Asp243, Asn245, and Leu442 in Os3BGlu7, respectively. The tryptophan, Trp366 (Trp358 in Os3BGlu7), which appears to stack and orient the glucose residues at the

+1 and +2 subsites in Os3BGlu7 (Chuenchor *et al.*, 2008) and orient aglycones in other plant GH1 enzymes (Czczek *et al.*, 2000; Barleben *et al.*, 2007) is conserved. However, Asn245, which is thought to hydrogen bond to the third glucosyl residue of cellooligosaccharides in Os3BGlu7, is replaced with Met251 in Os3BGlu6. The Met251 at this position sterically blocks the third glucosyl residue of cellotriose from binding to Os3BGlu6 in the same position as in Os3BGlu7, which may explain the poor hydrolysis of cellotriose and longer  $\beta$ -(1 $\rightarrow$ 4)-linked oligosaccharides by Os3BGlu6. Met251 is, however, critical in making hydrophobic interactions with the *n*-octyl chain in Os3BGlu6. The residues which may interact by stacking interactions in Os3BGlu7 beyond the subsite +2 i.e., Trp337, Tyr341, and Phe343, are replaced by the much smaller residues Thr345, Ser349, and Pro351 in Os3BGlu6, reflecting the fact that cellooligosaccharides do not bind in this position. These differences between the two enzymes clearly reflect the poor hydrolysis of long cellooligosaccharides by Os3BGlu6. In addition, it can explain the fact that Os3BGlu6 prefers hydrophobic  $\beta$ -glucosides and short oligosaccharides, while Os3BGlu7 prefers longer  $\beta$ -(1 $\rightarrow$ 4)-linked oligosaccharides.

A



B



**Figure 4.3** Comparison of Os3BGlu6 and Os3BGlu7 ligand interactions.

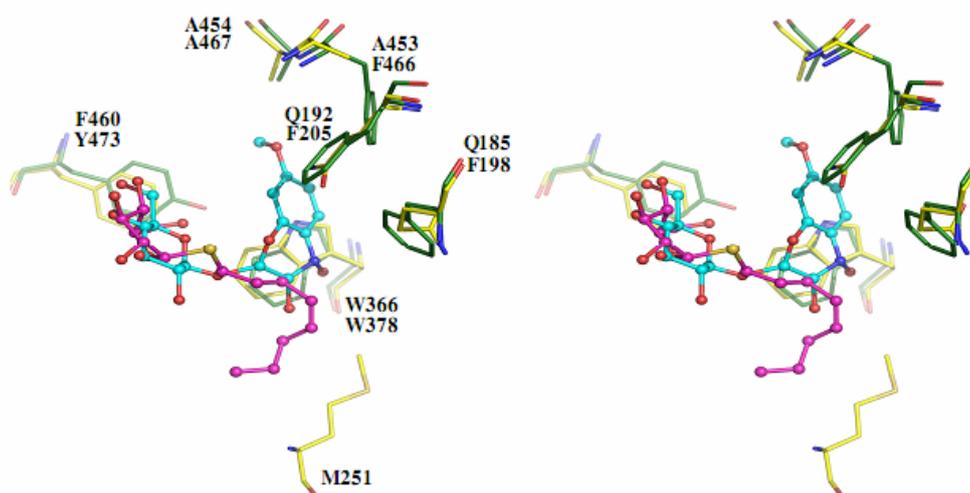
(A) Surface view of the active site of Os3BGlu6 and docked cellotriose from rice Os3BGlu7 superimposed on the *n*-octyl- $\beta$ -D-thioglucopyranoside bound to Os3BGlu6. Asn245 at the +2 subsite of Os3BGlu7 is shown in white and Met251 of

Os3BGlu6 in yellow, both shown in sticks representation. The celotriose ligand is represented by balls and sticks with carbon in cyan and oxygen in red. The *n*-octyl- $\beta$ -D-thioglucopyranoside is also shown in ball and stick representation with carbons in pink, oxygen in red and sulphur in yellow; (B) Stereo view superimposition of the active site residues of the Os3BGlu6/*n*-octyl- $\beta$ -D-thioglucopyranoside complex and rice Os3BGlu7 docked with celotriose. The *n*-octyl- $\beta$ -D-thioglucopyranoside ligand is shown as balls and sticks with carbons in cyan, oxygen in red and sulphur in yellow. The celotriose ligand is represented as balls and sticks with carbon in cyan and oxygen in red. Amino acid residues are shown as sticks with Os3BGlu6 carbons in yellow and rice Os3BGlu7 carbons in green with oxygen in red and nitrogen in blue. The figure was generated by Pymol (Delano Scientific LLC).

#### 4.7.2 Os3BGlu6 and Maize ZmGlu1

Among plant GH1 enzymes, the aglycone specificities of the chloroplastic  $\beta$ -glucosidases from maize and sorghum have been most extensively studied (Czjzek *et al.*, 2000; 2001; Verdoucq *et al.*, 2003; 2004). The crystal structures of maize ZmGlu1 with 2-O- $\beta$ -D-glucopyranosyl-4-hydroxy-7-methoxy-1,4-benzoaxazin-3-one (DIMBOAGlc) (PDB: 1E56, Czjzek *et al.*, 2000) and Os3BGlu6 with *n*-octyl- $\beta$ -D-thioglucopyranoside were superimposed (Figure 4.4). The residues that directly interact with the aglycone in ZmGlu1 include W378 on one side and three phenylalanines Phe198, Phe205 and Phe466 on the opposite side of the active-site slot (Czjzek *et al.*, 2000). Although the tryptophan was conserved in Os3BGlu6 (Trp366), the opposing amino acid residues were completely different in Os3BGlu6, in which they were replaced by Gln185, Gln192, and Ala453. Among the residues noted to indirectly affect aglycone binding in ZmGlu1, Ala467 is conserved (Ala454 in

Os3BGlu6), while Tyr473, which orients the plane of Trp378 for its stacking interaction with the DIMBOA substrate, is replaced by Phe460, which has no hydroxyl to hydrogen bond to the indole ring of Trp366 in Os3BGlu6. Thus, Trp366 is twisted away by 1.13Å for its interaction with the *n*-octyl chain of Os3BGlu6. The aglycone part of *n*-octyl-β-D-thioglucoside in Os3BGlu6 is on the opposite side of the active site compared to that of DIMBOA, where it makes hydrophobic contacts with different residues. These differences in aglycone binding may in turn affect the position and orientation of glucose, as noted by Verdoucq *et al.* (2004).

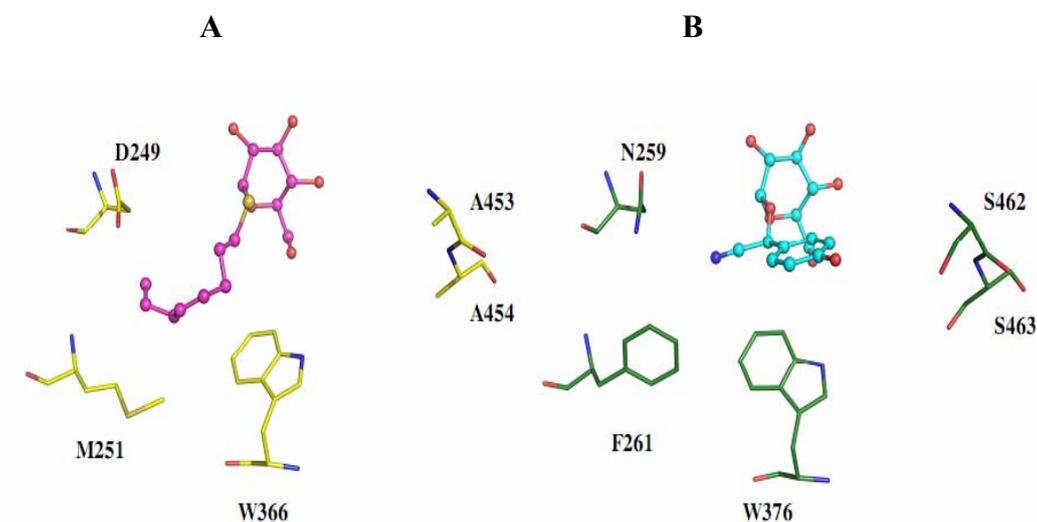


**Figure 4.4** Comparison of the aglycone binding pockets of Os3BGlu6/*n*-octyl-β-D-thioglucopyranoside and ZmGlu1/DIMBOA-Glc (PDB code: 1E56) in stereoview.

The ligands are shown in ball and stick representation with *n*-octyl-β-D-thioglucopyranoside carbons in pink and DIMBOAGlc carbons in cyan, oxygen in red and sulphur in yellow. Amino acid residues are shown as sticks with Os3BGlu6 carbons in yellow and ZmGlu1 carbons in green with oxygen in red and nitrogen in blue. The figure was generated by Pymol (Delano Scientific LLC).

### 4.7.3 Os3BGlu6 and Sorghum (SbDhr1)

Sorghum  $\beta$ -glucosidase (SbDhr1) hydrolyzes the natural substrate dhurrin (*p*-hydroxy-(S)-malendelonitrile- $\beta$ -D-glucoside), but this substrate acts as an inhibitor to maize ZmGlu1 (Verdoucq *et al.*, 2004). To compare the active site residues between Os3BGlu6 and SbDhr1 and in order to understand the two enzymes preferences for different substrates, the key residues involved in dhurrin hydrolysis were superimposed with the corresponding residues in the Os3BGlu6/*n*-octyl- $\beta$ -D-thioglucopyranoside complex. The glycosyl moieties of dhurrin and *n*-octyl- $\beta$ -D-thioglucopyranoside were both observed in  ${}^1S_3$  skew boat conformations in very similar positions, but the residues involved in aglycone binding were quite different (Figure 4.5). In SbDhr1, the key residues in dhurrin recognition were Asn259, Phe261, and Ser462, with the corresponding Os3BGlu6 residues being Asp249, Met251, and Ala453. Ser462 and Asn 259 bind to the hydroxyl group and cyano group of phenyl ring of dhurrin through water mediated hydrogen bonding. The replacement of Ser462 by a smaller hydrophobic Ala453 in Os3BGlu6, would eliminate the interaction with the phenolic group. Asp249 is located further away in Os3BGlu6 than Asn259 in SbDhr1 which would make weaker interactions with the cyano group. Phe261 is replaced by Met251 in Os3BGlu6 located at a distance of approximately 6.1 Å from the cyano group of dhurrin as opposed to 4.4 Å for Phe261. These two residues in Os3Glu6 (Asp249 and M251) are the in the same positions as those of Asp261 and M263 in maize ZmGlu1, which also did not hydrolyze dhurrin.



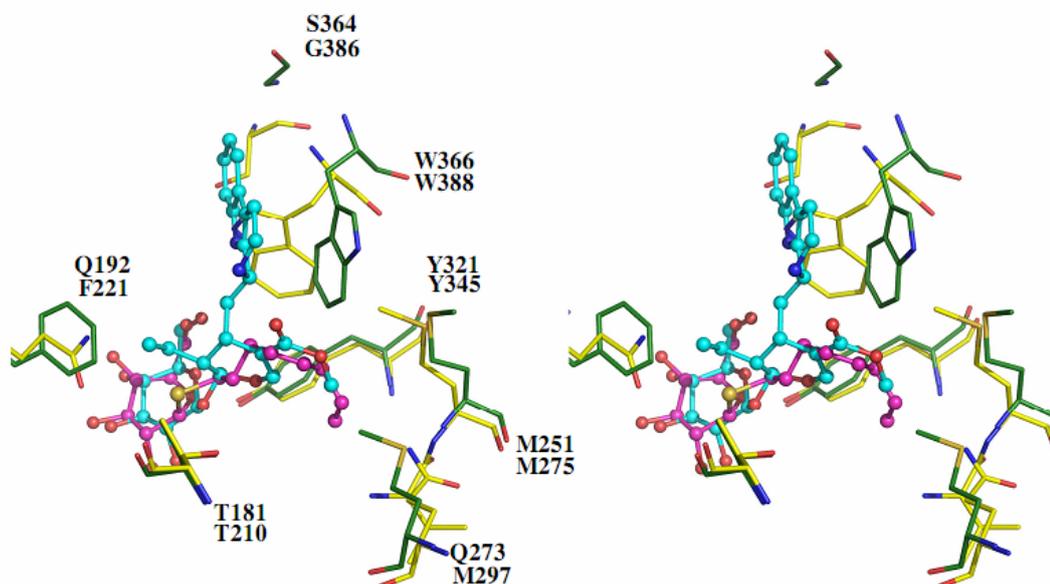
**Figure 4.5** Comparison of the aglycone binding pockets of the Os3BGlu6/*n*-octyl- $\beta$ -D-thioglucopyranoside complex and a SbDhr1 mutant bound to dhurrin (PDB code: 1V03).

(A) Active site of Os3BGlu6/*n*-octyl- $\beta$ -D-thioglucopyranoside. The ligand is shown in ball and stick representation with *n*-octyl- $\beta$ -D-thioglucopyranoside carbons in pink. Amino acid residues are shown as sticks with Os3BGlu6 carbons in yellow; (B) Active site of SbDhr1. The dhurrin ligand is shown in balls and sticks with carbons in cyan, oxygen in red and nitrogen in blue. Amino acid residues are shown with carbons in green. The figure was generated by Pymol (Delano Scientific LLC).

#### 4.7.4 Os3BGlu6 and Strictosidine- $\beta$ -glucosidase

Strictosidine- $\beta$ -glucosidase (SG) is a GH family one enzyme involved in indole alkaloid metabolism and its crystal structure and that of a mutant SG with the natural substrate strictosidine have been determined (Barleben *et al.*, 2007). Comparison of Os3BGlu6 with strictosidine- $\beta$ -glucosidase, both of which recognize hydrophobic substrates, would give insight on residues involved in aglycone recognition and the structural differences between these two enzymes. The aglycone

part of strictosidine is surrounded by mainly hydrophobic residues, including Thr210, Phe221, M275, M297, Gly386, and W388, while in Os3BGlu6, the corresponding residues are Thr181, Gln192, M251, Gln273, Ser364, and W366. Both Gly386 and W388 of SG are close to the indole system of the strictosidine substrate and are important for the recognition of this substrate (Figure 4.6). The equivalent Os3BGlu6 residues are Ser364, which is larger than glycine, and W366, which is in a different orientation in Os3BGlu6. These differences make the active site smaller in Os3BGlu6, thus it would be difficult for the large strictosidine substrate to enter the active site. Met251 (Met275 in SG) is conserved in both the enzymes and this residue is important in Os3BGlu6 to recognize the *n*-octyl chain. Although both enzymes contain similar hydrophobic residues in the active site, the shape of the aglycone binding pocket and the position of the residues in the active site in the two enzymes are different, which would likely allow them to bind different hydrophobic substrates.



**Figure 4.6** Comparison of the aglycone binding pockets of the Os3BGlu6/*n*-octyl- $\beta$ -D-thioglucopyranoside complex and strictosidine- $\beta$ -glucosidase (SG) (PDB code: 2JF6) in stereoview.

The ligands are shown in ball and stick representations with *n*-octyl- $\beta$ -D-thioglucopyranoside carbons in pink and SG carbons in cyan. Amino acid residues are shown as sticks with Os3BGlu6 carbons in yellow and SG carbons in green. All oxygen are shown in red and nitrogen in blue. The figure was generated by Pymol (Delano Scientific LLC).

Further, on this research could be extended by mutating the key residues around the aglycone binding site of Os3BGlu6 and determining their structures with the natural substrates extracted from the plant, which would throw more light on the physiological role of this enzyme.

## CHAPTER V

### CONCLUSION

In this study, a cDNA encoding rice Os3BGlu6, a GH1  $\beta$ -glucosidase, representing cluster one of eight phylogenetic clusters containing both rice and *Arabidopsis* proteins, was used to produce an Os3BGlu6 protein, which was functionally and structurally characterized. The cDNA sequence contained an open reading frame of 1563 nucleotides encoding 521 amino acids, which were predicted to include an N-terminal secretory signal peptide of 31 or 38 residues and a mature protein of 484 or 477 residues. The *Os3BGlu6* cDNA was successfully cloned into the pENTR™/TEV/D-TOPO® cloning vector, subcloned into pET32a(+)/DEST expression vector, and expressed as a fusion protein with N-terminal thioredoxin and His<sub>6</sub> tags, in *E. coli* strain Origami(DE3). Os3BGlu6 was expressed at 20°C without the addition of any IPTG. By two steps of IMAC purification, with removal of the N-terminal tags in between, the 55 kDa Os3BGlu6 could be purified to greater than 95% purity and concentrated to 10-20 mg/ml, which was essential for crystallization.

The purified Os3BGlu6 had a pH optimum of 4-5 and a temperature optimum of 40-55°C. Among *p*NP glycosides, Os3BGlu6 hydrolyzed *p*NP- $\beta$ -D-fucopyranoside, *p*NP- $\beta$ -D-glucoside and *p*NP- $\beta$ -D-galactoside, in order of efficiency of hydrolysis. Among glucose disaccharides, the  $\beta$ -(1→3)-linked laminaribiose was hydrolyzed the best ( $k_{cat}/K_m = 1.7 \text{ mM}^{-1}\text{s}^{-1}$ ), followed by the  $\beta$ -(1→2)-linked sophorose ( $k_{cat}/K_m = 0.96 \text{ mM}^{-1}\text{s}^{-1}$ ),  $\beta$ -(1→6)-gentiobiose ( $k_{cat}/K_m=0.07 \text{ mM}^{-1}\text{s}^{-1}$ ) and  $\beta$ -(1→4)-

linked cellobiose ( $k_{\text{cat}}/K_m = 0.009 \text{ mM}^{-1}\text{s}^{-1}$ ), in order of decreasing catalytic efficiency ( $k_{\text{cat}}/K_m$ ). Among alkyl glucosides, *n*-octyl- $\beta$ -D-glucoside was hydrolyzed with higher catalytic efficiency ( $k_{\text{cat}}/K_m = 2.7 \text{ mM}^{-1}\text{s}^{-1}$ ), than *n*-heptyl- $\beta$ -D-glucoside ( $k_{\text{cat}}/K_m = 0.85 \text{ mM}^{-1}\text{s}^{-1}$ ). In overnight incubation, Os3BGlu6 hydrolyzed the natural glycosides apigenin-7-glucoside, glycitin, diadzin, genistin, esculin, arbutin, coumaryl alcohol  $\beta$ -D-glucoside, coniferin, and salicin, but not amygdalin, linamarin, prunacin dalcochinin-glucoside, gossypin, naringin, phlorizin, pyridoxine-5'-O- $\beta$ -D-glucoside, and quercitin 3-  $\beta$ -D-glucoside.

Crystal clusters of Os3BGlu6 were obtained from microbatch crystallization in the condition with 20% PEG 5000 MME, 0.1 M Bis Tris, pH 6.5. Crystallization was further optimized in hanging drop vapor diffusion by varying the pH, protein to precipitant ratios, and protein concentrations. The qualities of the crystals were improved through microseeding by streak seeding. Single large native crystals of Os3BGlu6 were obtained that were suitable for x-ray diffraction. The native crystals were then soaked with 2-deoxy-2-fluoro- $\beta$ -D-glucopyranoside and *n*-octyl- $\beta$ -D-thioglucopyranoside. The crystals of native Os3BGlu6 alone and with the ligands were diffracted to obtain datasets with 1.83, 1.80, and 1.81 Å resolution, respectively. Os3BGlu6 crystals exhibited P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> space group symmetry with one molecule in the asymmetric unit. The native Os3BGlu6 structure was solved by molecular replacement with a phasing model derived from the structure for white clover cyanogenic  $\beta$ -glucosidase (PDB code: 1CBG). The structures of Os3BGlu6 in complexes with the ligands *n*-octyl- $\beta$ -D-thioglucopyranoside and G2F were solved by molecular replacement with the native Os3BGlu6 structure as the template model. With repeated cycles of model rebuilding and refinement, the models gave final R and

free R ( $R_{\text{free}}$ ) factors of 16.39% and 20.42% for native Os3BGlu6, 17.85% and 20.66% for Os3BGlu6 with *n*-octyl- $\beta$ -D-thioglucopyranoside, and 16.24% and 18.48% for Os3BGlu6 with G2F. The RMS deviations of 0.012 to 0.015 Å for bond lengths and 1.287-1.441° for bond angles in the native, and *n*-octyl- $\beta$ -D-thioglucopyranoside and G2F complex structures reflected the acceptable geometry for almost all the residues. The geometry was verified for all three structures in the Ramachandran plot, in which 90% of the residues fell in most favored region, 10% in additionally allowed region and 0.5% in generously allowed region.

Although the first 10 residues showed no electron density, N-terminal amino acid sequencing confirmed their presence in O3BGlu6 protein, suggesting these residues were disordered. Some residues in loop C (334 to 337) with solvent exposed side chains also showed poor density suggesting they were also mobile. All the three overall structures of Os3BGlu6 had the typical  $(\beta/\alpha)_8$  barrel of the known GH1 structures, with the conserved catalytic residues Glu178 and Glu394 on  $\beta$ -strands 4 and 7, respectively. The only significant difference in the backbone found through superimposition of the three Os3BGlu6 structures was the position of part of Loop C, which was exposed to the solvent surface and was different in *n*-octyl- $\beta$ -D-thioglucopyranoside complex. This was consistent with the comparison of Os3BGlu6 to other known GH1 structures, which showed that the differences in the GH1 structures mainly resided in the loops A, B, C, and D, which help determine the substrate specificities. The glycone binding pocket was 7.5 Å deep and 9 Å wide. The aglycone binding site at the entrance to the active site was about 5.3 Å across, while the broadest region was 7.6 Å. The native Os3BGlu6 had one molecule of Tris in its active site in two possible conformations, and Tris acted as a weak inhibitor of the

enzyme. The conserved residue Glu451, which was important for the ligand binding in Os3BGlu6, was found in two different conformations in the native and *n*-octyl- $\beta$ -D-thioglucopyranoside complex structures, but one of these structures was stabilized by binding to G2F in its complex with Os3BGlu6. The Os3BGlu6/*n*-octyl- $\beta$ -D-thioglucopyranoside structure had disordered density for the *n*-octyl chain, which could interact with various hydrophobic residues in the active site while the glycosyl group had well defined density, with the sugar in a  ${}^1S_3$  skew boat conformation. The Os3BGlu6/G2F structure showed the catalytic nucleophile Glu394 covalently bound to the C1 of the glucosyl moiety, representing the glycosyl-enzyme intermediate of the retaining mechanism of glycosyl hydrolases (Withers and Street, 1988).

The structure of Os3BGlu6 was compared with other known structures of plant GH 1 enzymes to analyze the differences in substrate interactions that determine their specificities. The hydrolysis pattern of oligosaccharide hydrolysis observed in Os3BGlu6 and Os3BGlu7 were different, hence a comparison of active site residues between the two enzymes revealed that the residues at the aglycone binding site were also different. Asn245 at subsite +2 for binding cellooligosaccharides in Os3BGlu7 is replaced with Met251 in Os3BGlu6. This Met251 appeared to block the third glucosyl residue of celotriose from binding to Os3BGlu6 in the same position as in Os3BGlu7. Beyond the +2 subsite, Trp337, Tyr341, and Phe343, which in Os3BGlu7 help bind oligosaccharides, were replaced by the smaller residues Thr345, Ser349, and Pro351 in Os3BGlu6, reflecting the fact that Os3BGlu6 could not bind longer cellooligosaccharides. Maize ZmGlu1, which hydrolyzes DIMBOAGlc, strictosidine- $\beta$ -glucosidase, which hydrolyzes the indole alkaloid strictosidine and sorghum (SbDhr1), which hydrolyzes dhurrin, were also compared to Os3BGlu6. In all these

enzymes, the differences among the active site residues and the shape of the active site contributed by these residues, appeared to reflect the differences in which specific substrates bind for hydrolysis to take place.

In summary, Os3BGlu6, one of the 34  $\beta$ -glucosidases belonging to GH family 1 predicted from the rice with a unique phylogeny compared to previously characterized GH1 enzymes, was expressed in recombinant *E. coli*, and its substrate specificity and 3D structure were determined. Since GH1 enzymes possess a broad range of specificities, in terms of the substrates they hydrolyze, which determines the functional role they play, this research provides critical information for assessing GH1 functions and their structural basis.

## **REFERENCES**

## 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  $\beta$ -glucosidase from germinated rice: Purification and properties. **Phytochemistry**. 48: 49-54.
- Alasdair, M., Lindhorst, M., Stephen, G., Withers, R., and Warren, A.J. (1994). The Acid/Base Catalyst in the Exoglucanase/Xylanase from *Cellulomonas fimi* is Glutamic Acid 127: Evidence from Detailed Kinetic Studies of Mutants. **Biochemistry**. 33: 6371-6376.
- Aleshin, A., Golubev, A., Firsov, L.M., and Honzatko, R.B. (1992). Crystal structure of glucoamylase from *Aspergillus awamori* var X100 to 2.2-Å resolution. **J. Biol. Chem.** 267: 19291-19298.
- Argandona, V.H., Corcuera, L.J., Niemeyer, H.M., and Cambell, B.C. (1983). Toxicity and feeding deterrency of hydroxamic acids from gramineae in synthetic diets against the green bug, *Schizaphis graminum*. **Ned. Entomol. Ver. Amsterdam**. 34: 134-138.

- 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.
- 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  $\beta$ -glucosidase from white clover, a family 1 glycosyl hydrolase. **Structure**. 3(9): 951-960.
- Béguin, P. (1990). Molecular biology of cellulose degradation. **Ann. Rev. Microbiol.** 44: 219-248.
- Brzobohaty, B., Moore, I., Kristoffersen, P., Bako, L., Campos, N., Schell, J., and Palme, K. (1993). Release of active cytokinin by a  $\beta$ -glucosidase localized to the maize root meristem. **Science**. 262: 1051-1054.
- 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.
- Chothia, C., and Lesk, A.M. (1986). The relation between the divergence of sequence and structure in proteins. **EMBO J.** 5: 823-826.
- Chuankhayan, P., Hua, Y., Svasti, J., Sakdarat, S., Sullivan, P.A., and Ketudat-Cairns JR. (2005). Purification of an isoflavonoid 7-O- $\beta$ -apiosyl-glucoside  $\beta$ -glycosidase and its substrates from *Dalbergia nigrescens* Kurz. **Phytochemistry.** 16(66): 1880-1889.
- Chuankhayan, P., Rimlumduan, T., Tantanuch, W., Mothong, N., Kongsaree, P.T., Metheenuku, P., Svasti, J., Jensen, O.N., and Ketudat-Cairns JR. (2007). Functional and structural differences between isoflavonoid  $\beta$ -glycosidases from *Dalbergia* sp. **Arch. Biochem. Biophys.** 468: 205-216.
- Chuenchor, W., Pengthaisong, S., Robinson, R.C., Yuvaniyama, J., Oonant, W., Bevan, D.R., Esen, A., Chen, C.J., Opasiri, R., Svasti, J., and Ketudat-Cairns, J.R., (2008). Structural insights into rice BGlu1  $\beta$ -glucosidase oligosaccharide hydrolysis and transglycosylation. **J. Mol. Biol.** 377: 1200-1215.
- Coutinho, P.M., and Henrissat, B. (1999). Carbohydrate-active enzymes: an integrated database approach. In Gilbert, H.J., Davies, G., Henrissat, B., and Svensson, B. (eds.). **Recent Advances in Carbohydrate Bioengineering** (pp. 3-12). Cambridge: The Royal Society of Chemistry.

- Czjzek, M., Cicek, M., Zamboni, V., Bevan, D.R., Henrissat, B., and Esen, A. (2000). The mechanism of substrate (aglycone) specificity in  $\beta$ -glucosidase is revealed by crystal structures of mutant maize  $\beta$ -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 ZMGluc1) beta-glucosidase and a model of its complex with *p*-nitrophenyl  $\beta$ -D-thioglucoside. **Biochem. J.** 354: 37-46.
- Dan, S., Marton, I., Dekel, M., Bravdo, B.A., He, S., Withers, S., and Shoseyov, O. (2000). Cloning, expression, characterization and nucleophile identification of family 3, *Aspergillus niger*  $\beta$ -glucosidase. **J. Biol. Chem.** 275: 4973- 4980.
- Davies, G., and Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. **Structure.** 3: 853-859.
- Day, A.J., Canada, F.J., Diaz, J.C., Kroon, P.A., Mclauchlan, R., and Faulds, C.B. (2000). Dietary flavonoid and isoflavone glycosides are hydrolysed by the lactase site of lactase phlorizin hydrolase. **FEBS Lett.** (166-170).
- Delano, W.L. (2002). **The Pymol Molecular Graphics System** [On-line]. Available: <http://www.pymol.org.html>.
- Dharmawardhana, D.P., Ellis, B.E., and Carlson, J.E., (1995). A  $\beta$ -glucosidase from lodgepole pine specific for the lignin precursor coniferin. **Plant Physiol.** 107: 331-339.

- Dietz, K.J., Wichert, K., Sauter, A., Messdaghi, D., and Hartung, W. (2000). Characterization of an extracellular  $\beta$ -glucosidase in barley involved in the hydrolysis of ABA glucose conjugate in leaves. **J. Exptl. Bot.** 25: 937-944.
- Dixon, M., and Webb, E.C. (1979). **Enzymes** (3<sup>rd</sup> ed). New York: Academic Press.
- Emsley, P., and Cowtan, K. (2004). Coot-Model-building tools for molecular graphics. **Acta. Cryst. D60**: 2126-2132.
- Escamilla-Trevino, L.L., Chen, W., Carda, M.L., Shiha, M.C., Chenga, C.L., and Poulton, J.E. (2006). *Arabidopsis thaliana*  $\beta$ -Glucosidases BGLU45 and BGLU46 hydrolyse monolignol glucosides. **Phytochemistry**. 67: 1651-1660.
- Esen, A. (1993).  **$\beta$ -Glucosidases: Biochemistry and Molecular Biology**, ACS Symposium Series 533. ACS Symposium Series 533, American Chemical Society, Washington, D.C.
- Falk, A., and Rask, L. (1995). Expression of a zeatin-O-glucoside-degrading  $\beta$ -glucosidase in *Brassica napus*. **Plant Physiol.** 108: 1369-1377.
- Fowler, T. (1993). Deletion of the *Trichoderma reesei*  $\beta$ -glucosidase gene, bgl1. In Esen, A. (ed.).  **$\beta$ -Glucosidases: Biochemistry and Molecular Biology** (56-65). ACS Symposium Series 533, American Chemical Society, Washington, D.C.
- Geerlings, A., Memelink, I.J., Heiden, R., and Verpoorte, R. (2000). Molecular cloning and analysis of strictosidine  $\beta$ -D-glucosidase, an enzyme in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. **J. Biol. Chem.** 275: 3051-3056.

- Gilbert G.P. (2007). Detergents for the stabilization and crystallization of membrane proteins. **Structural Biology of Membrane Proteins**. 41: 388-397.
- Giordani, R., and Noat, G. (1988). Isolation, molecular properties and kinetic studies of a strict beta-fucosidase from *Lactuca sativa* latex. **Eur. J. Biochem**. 175: 619-625.
- Haberer, G., and Kieber, J.J. (2002). Cytokinins. New insights into a classic phytohormone. **Plant Physiol**. 128: 354-362.
- Harvey, A.J., Hrmova, M., Gori, R., Varghese, J.N., and Fincher, G.B. (2000). Comparative modeling of the three-dimensional structures of family 3 glycoside hydrolases. **Proteins: Structure, Function, and Genetics**. 4: 257-269.
- Hayashi K., and Kojima, C. (2008). pCold-GST vector: A novel cold-shock vector containing GST tag for soluble protein production. **Prot. Express. Purif**. 62:120-127.
- Henrissat, B. (1990). Weak sequence homologies among chitinases detected by clustering analysis. **Pro. Seq. Dat. Anal**. 3: 523-526.
- 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 Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. **Biochem. J.** 316: 695-696.
- Henrissat, B., and Davies, G. (1997). Structural and sequence-based classification of glycoside hydrolases **Curr. Opin. Struct. Biol.** 7: 637-644.
- Henrissat, B., Callebaut, I., Fabrega, S., Lehn, S., Mornon, J.P., and Davies, G. (1995). Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. **Proc. Natl. Acad. Sci.** 92: 7090-7094.
- Henrissat, B., Claeysens, M., Tomme, P., Lemesle, L., and Mornon, J.P. (1989). Cellulase families revealed by hydrophobic cluster analysis. **Gene.** 81(1): 83-95.
- Hösel, W., and Barz, W. (1975).  $\beta$ -Glucosidases from *Cicer arietinum* L. **Eur. J. Biochem.** 57: 607-616.
- Hrmova, M., and Fincher, G.B. (2001). Plant enzyme structure. Explaining substrate specificity and the evolution of function. **Plant Physiol.** 125: 54-57.
- Hsieh, M.C., and Graham, T.L. (2001). Partial purification and characterization of a soybean  $\beta$ -glucosidase with high specific activity towards isoflavone conjugates. **Phytochemistry.** 58: 995-1005.
- 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-Garcia, L., Canada, F.J., Gonzalez, B., and Aparicio, S.J. (2007). Crystal structures of *Paenibacillus polymyxa* glucosidase B

complexes reveal the molecular basis of substrate specificity and give new insights into the catalytic machinery of family I glycosidases. **J. Mol. Bio.** 371: 1204-1218.

Jenkins, J., Leggio, L.L., Harris, G., and Pickersgill, R. (1995).  $\beta$ -glucosidase,  $\beta$ -galactosidase, family A cellulases, family F xylanases and two barley glucanases from a superfamily of enzymes with 8-fold  $\beta/\alpha$  architecture and with two conserved glutamates near the carboxy-terminal ends of  $\beta$  strands four and seven. **FEBS Lett.** 362: 281-285.

Kapust, R.B., Tözsér, J., Fox, J.D., Anderson, D.E., Cherry, S., Copeland, T.D., and Waugh, D.S. (2001). Tobacco etch virus protease: Mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. **Prot. Eng.** 14: 993-1000.

Krissinel, E., and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. **Acta. Cryst.** D60: 2256-2268.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature.** 227: 680-685.

Lasanthi-Kudahettige, R., Magneshi, L., Loreti, E., Gonzali, S., Licausi, F., Novi, G., Beretta, O., Vitulli, F., Alpi, A., and Perata, P. (2007). Transcript profiling of the anoxic rice coleoptile. **Plant Physiol.** 144: 218-231.

Leah, R., Kigel, J., Svendsen, I., and Mundy, J. (1995). Biochemical and molecular characterization of a barley seed  $\beta$ -glucosidase. **J. Biol. Chem.** 270: 15789-15797.

- Lee, K.H., Piao, H.L., Kim, H.Y., Choi, S.M., Jiang, F., Hartung, W., Hwang, I., Kwak, J.M., Lee, I.J., and Hwang, I. (2006). Activation of glucosidase via stress-induced polymerization rapidly increases active pools of abscisic acid. **Cell**. 126: 1109-1120.
- Li, Y.K., Chir, J., and Chen, F.Y. (2001). Catalytic mechanism of family 3  $\beta$ -glucosidase and mutagenesis study on residues Asp-247. **Biochem. J.** 355: 835-840.
- Lirdprapamongkol, K., and Svasti, J. (2000). Alkyl glucoside synthesis using Thai rosewood  $\beta$ -glucosidase. **Biotech. Lett.** 22: 1889-1894.
- Ljung, K., Ostin, A., Lioussanne, L., and Sandberg, G. (2001). Developmental regulation of indole-3-acetic acid turnover in Scots pine seedlings. **Plant Physiol.** 125: 464-475.
- MacGregor, E.A., and Svensson, B. (1989). A super-secondary structure predicted to be common to several  $\alpha$ -1,4-D-glucan-cleaving enzymes. **Biochem. J.** 259: 145-152.
- McCarter, J., and Withers, S.G. (1994). Mechanisms of enzymatic glycoside hydrolysis. **Curr. Opin. Struct. Biol.** 4: 885-892.
- McPherson, A., Koszelak, S., Axelrod, H., Day, J., Williams, R., Robinson, L., McGrath M., and Cascio D. (1986). An experiment regarding crystallization of soluble proteins in the presence of  $\beta$ -octyl-glucoside. **J. Biol. Chem.** 261: 1969-1975.
- Michel, G., Chantalat, L., Fanchon, E., Henrissat, B., Kloareg, B., and Dideberg, O. (2001). The  $\iota$ -carrageenase of *Alteromonas fortis*, a  $\beta$ -helix fold-containing

- enzyme for the degradation of a highly polyanionic polysaccharide. **J. Biol. Chem.** 276: 40202-40209.
- Michelin, K., Wajner, A., and Goulart, L.S. (2004). Biochemical study on  $\beta$ -glucosidase in individuals with Gaucher's disease and normal subjects. **Clin. Chem. Acta.** 343: 145-153.
- Mizutani, F.M., Nakanishi, H., Ema, J., Ma, S.J., Noguchi, N., Ochiai, M.I., Mizutani, M.F., Nakao, M., and Sakata, K. (2002). Cloning of  $\beta$ -primeverosidase from tea leaves, a key enzyme in tea aroma. **Plant Physiol.** 130: 2164-2176.
- Murdushov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum likelihood method. **Acta. Cryst. D53:** 240-255.
- Naim, H.Y. (2001). Molecular and cellular aspects and regulation of intestinal lactase-phlorizin hydrolase. **Histol. Histopathol.** 16: 553-561.
- Niemeyer, H.M. (1988). Hydroxamic acids (4-hydroxy-1,4-benzoxazin-3-ones), defense chemicals in the gramineae. **Phytochemistry.** 27: 3349-3358.
- Nijikken, Y., Tsukada, T., Igarashi, K., Samejima, M., Wakagi, T., Shoun, H., and Fushinobu, S. (2007). Crystal structure of intracellular family 1  $\beta$ -glucosidase BGL1A from the basidiomycete *Phanerochaete chrysosporium*. **FEBS. Lett.** 581: 1514-1520.
- Noguchi, J., Hayashi, Y., Baba, Y., Zkino, O., and Kimura, M. (2008). Crystal structure of the covalent intermediate of human cytosolic  $\beta$ -glucosidase. **Biochem. Biophys. Res. Comm.** 374: 549-552

- Oliveira, J.S., Pinto, C.A., Basso, L., and Santos, D.S. (2001). Cloning and overexpression in soluble form of functional Shikimate kinase and 5-enolpyruvylshikimate 3-phosphate synthase enzymes from *Mycobacterium tuberculosis*. **Prot. Express. Purif.** 22: 430-435.
- Opassiri, R., Hua, Y., Wara-Aswapati, O., Akiyama, T., Svasti, J., Esen, A., and Ketudat-Cairns, J.R. (2004).  $\beta$ -glucosidase, exo- $\beta$ -glucanase and pyridoxine transglucosylase activities of rice BGlu1. **Biochem. J.** 379: 125-131.
- Opassiri, R., Ketudat-Cairns J.R., Akiyama, T., Wara-Aswapati, O., Svasti, J., and Esen, A. (2003). Characterization of a rice  $\beta$ -glucosidases highly expressed in flower and germinating shoot. **Plant Science.** 165: 627-638.
- Opassiri, R., Pomthong, B., Onkoksoong, T., Akiyama, T., Esen, A., Ketudat Cairns, J.R. (2006). Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12  $\beta$ -glucosidase. **BMC Plant Biol.** 6: 33.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In Carter, C.W. and Sweet, R.M. (eds.). **Meth. Enzymol** (307-326). New York: Academic Press.
- Persans, M.W., Wang, J., and Schuler, M.A. (2001). Characterization of maize cytochrome P450 monooxygenases induced in response to safeners and bacterial pathogens. **Plant Physiol.** 125: 1126-1138.
- Raimbaud, E., Buleon, A., Perez, S., and Henrissat, B. (1989). Hydrophobic cluster analysis of the primary sequences of alpha-amylases. **Intl. J. Biol. Macromol.** 11: 217-225.

- Raychaudhuri, A., and Tipton, P.A. (2002). Cloning and expression of the gene for soybean hydroxyisourate hydrolase. Localization and implications for function and mechanism. **Plant Physiol.** 130: 2061-2068.
- 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. (2<sup>nd</sup> ed). New York: Cold spring harbor laboratory.
- Sanz-Aparicio, J., Hermoso, J.A., Martinez-Ripoll, M., Lequerica, J.L., and Polaina, J. (1998). Crystal structure of  $\beta$ -glucosidase A from *Bacillus polymyxa*: insights into the catalytic activity in family 1 glycosyl hydrolases. **J. Mol. Bio.** 275: 491-502.
- Schliemann, W. (1984). Hydrolysis of conjugated gibberellins by  $\beta$ -glucosidases from dwarf rice (*Oryza sativa* L. cv. Tan-ginbozu). **J. Plant. Physiol.** 116: 123-132.
- Srisomsap, C., Svasti, J., Surarit R., Champattanachai V., Sawangareetrakul, P., Boonpuan, K., Subhasitanont P., and Chokchaichamnankit, D. (1996). Isolation and characterization of an enzyme with  $\beta$ -glucosidase and  $\beta$ -fucosidase activities from *Dalbergia cochinchinensis* Pierre. **J. Biochem.** 119: 585-590.
- Studier, F.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. **Meth. Enzymol.** 185: 60-89.

- Sue, M., Yamazaki, K., Yajima, S., Nomura, T., Matsukawa, T., Iwamura, H., and Miyamoto, T. (2006). Molecular and structural characterization of hexameric  $\beta$ -D-glucosidases in wheat and rye. **Plant. Physiol.** 141: 1237-1247.
- Suzuki, H., Takahashi, S., Watanabe, R., Fukushima, Y., Fujita, N., Noguchi, A., Yokoyama, R., Nishitani, K., Nishino, T., and Nakayama, T. (2006). An isoflavone conjugate hydrolyzing  $\beta$ -glucosidase from the roots of soybean (*Glycine max*) seedlings. Purification, gene cloning, phylogenetics and cellular localization. **J. Biol. Chem.** 281: 30251-30259.
- Svensson, B. (1988). Regional distant sequence homology between amylases, alpha-glucosidases and transglucanoylases. **FEBS Lett.** 230: 72-76.
- Tews, I., Perrakis, A., Oppenheim, A., Dauter, Z., Wilson, K.S., and Vorgias, C.E. (1996). Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. **Nature Struct. Biol.** 3: 638-648.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Res.** 25: 4878-4882.
- Tsuchiya, T., and Saito, S. (1984). Use of n-octyl- $\beta$ -D-thioglucoside, a new nonionic detergent, for solubilization and reconstitution of membrane proteins. **J. Biochem.** 96: 1593-1597.
- Vagin, A., and Teplyakov, A. (1997). MOLREP: an automated program for molecular replacement. **J. Appl. Cryst.** 30: 1022-1025.

- Vaguine, A.A., Richelle, J., and Wodak, S.J. (1999). SFCHECK: a unified set of procedure for evaluating the quality of macromolecular structure-factor data and their agreement with atomic model. **Acta. Cryst.** D55: 191-205.
- Varghese, J.N., Hrmova, M., and Fincher, G.B. (1999). Three-dimensional structure of a barley  $\beta$ -D-glucan glucohydrolase; a family 3 glycosyl hydrolase. **Structure.** 7: 179-190.
- Verdoucq, L., Czjzek, M., Moriniere, J., Bevan, D.R., and Esen, A. (2003). Mutational and structural analysis of aglycone specificity in maize and sorghum  $\beta$ -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  $\beta$ -glucosidases: Novel insights from the crystal structure of sorghum dhurrinase-1, a plant  $\beta$ -glucosidase with strict specificity, in complex with its natural substrate. **J. Biol. Chem.** 279: 31796-31803.
- Wang, Q., Graham, R.W., Trimbur, D., Warren, R.A.J., and Withers, S.G. (1994). Changing enzymatic reaction mechanism of a retaining glucosidase to an inverting enzyme. **J. Am. Chem. Soc.** 116: 11594-11595.
- White, A., and Rose, D.R. (1997). Mechanism of catalysis by retaining  $\beta$ -glycosyl hydrolases. **Curr. Opin. Struct. Biol.** 7: 645-651.
- White, A., Withers, S.G., Gilkes, N.R., and Rose, D.R. (1994). Crystal structure of the catalytic domain of the  $\beta$ -1,4 glycanase Cex from *Cellulomonas fimi*. **Biochemistry.** 33: 546-552.

- Wiesmann, C., Beste, G., Hengstenberg, W., Schulz, G.E. (1995) The three-dimensional structure of 6-phospho-beta-galactosidase from *Lactococcus lactis*. **Structure**. 3: 961-968.
- Withers, S.G., and Aebersold, R. (1995). Approaches to labeling and identification of active site residues in glycosidases. **Prot. Sci.** 4: 361-371.
- Withers, S.G., and Street, I.P. (1988). Identification of a covalent  $\alpha$ -D-glucopyranosyl enzyme intermediate formed on a  $\beta$ -glucosidase. **J. Am. Chem. Soc.** 110: 8551- 8553.
- Withers, S.G., Street, I.P., Bird, P., and Dolphin, D.H., (1987). 2-Deoxy-2-fluoroglucosides: A novel class of mechanism-based glucosidase inhibitors. **J. Am. Chem. Soc.** 109: 7530-7531.
- Wolfenden, R., Lu, X., and Young, G. (1998). Spontaneous hydrolysis of glycosides. **J. Am. Chem. Soc.** 120: 6914-6815.
- Xu, Z., Escamilla-Treviño, L.L., Zeng, L., Lalgondar, M., Bevan, D.R., Winkel, B.S. J., Mohamed, A., Cheng, C., Shih, M., Poulton, J.E., and Esen, A. (2004). Functional genomic analysis of *Arabidopsis thaliana* glycoside hydrolase family 1. **Plant. Mol. Biol.** 55: 343-367.
- 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  $\alpha$ -glucosidases. **J. Am. Chem. Soc.** 125: 14313-14323.

Zhou, J., Hartmann, S., Brienne, K., Shepherd, and Poulton, J.E. (2002). Investigation of the microheterogeneity and aglycone specificity-conferring residues of black cherry prunasin hydrolases. **Plant Physiol.** 129: 1252-1264.

## **CURRICULUM VITAE**

**NAME** Miss Supriya Seshadri

**DATE OF BIRTH** 27 October 1978

**PLACE OF BIRTH** Mysore, India

### **INSTITUTION ATTENDED**

Manasagangothri, University of Mysore, Mysore,  
India, 1999-2001,

Master of Science, Food Science & Nutrition.

Yuvaraja's College, Mysore, India, 1996-1999,

Bachelor of Science, Biochemistry, Zoology,  
Food Science & Nutrition.

### **AWARDS AND SCHOLARSHIPS**

University of Mysore Topper in M.Sc., Food  
Science and Nutrition, (Gold Medalist).

National Merit Scholarship for securing the 22nd  
rank in the state in S.S.L.C.