

ความสัมพันธ์ระหว่างโครงสร้างและหน้าที่ของเอนไซม์เบตาไกลูโคซิเดส
BGLU1 ในข้าว

นางสาววัลลี ชื่นช่อ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาชีวเคมี
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2550

STRUCTURE-FUNCTION RELATIONSHIPS IN RICE
BGLU1 β -GLUCOSIDASE

Watchalee Chuenchor

A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Biochemistry
Suranaree University of Technology
Academic Year 2007

**STRUCTURE-FUNCTION RELATIONSHIPS IN RICE BGLU1 β -
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)

(Asst. Prof. Dr. Robert C. Robinson)

Member

(Asst. Prof. Dr. Jirundon Yuvaniyama)

Member

(Prof. Dr. Asim Esen)

Member

(Dr. Rodjana Opassiri)

Member

(Assoc. Prof. Dr. Saowanee Rattanaphani)

Vice Rector for Academic Affairs

(Assoc. Prof. Dr. Sompong Thammathaworn)

Dean of Institute of Agricultural Technology

วิชา ชื่อ : ความสัมพันธ์ระหว่างโครงสร้างและหน้าที่ของเอนไซม์เบตาไกลูโคซิเดส
BGLU1 ในข้าว (STRUCTURE-FUNCTION RELATIONSHIPS IN RICE
BGLU1 β -GLUCOSIDASE) อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร.เจมส์ เกตุทัต-
คาร์นส์, 189 หน้า.

เอนไซม์ BGlul ของข้าวพบมากในช่วงที่ข้าวออกดอกและกำลังงอกเมล็ด ทำหน้าที่ย่อย
โอลิโกแซคคาไรด์ซึ่งเป็นส่วนประกอบของผนังเซลล์พืช เพื่อศึกษาความจำเพาะในการย่อย
สับสเตรตโดยเอนไซม์เบตาไกลูโคซิเดสในตระกูล Glycosyl hydrolase 1 (GH1) โดยเฉพาะ
เอนไซม์ในพืชต่อการย่อยโอลิโกแซคคาไรด์ ผลิตภัณฑ์ชนิดคือ ผลิตภัณฑ์ดั้งเดิม ผลิตภัณฑ์
ดั้งเดิมกับสารยับยั้ง 2-deoxy-2-fluoroglucoside ผลิตภัณฑ์กลายพันธุ์ E176Q และผลิตภัณฑ์
กลายพันธุ์กับสับสเตรตโอลิโกแซคคาไรด์ cellopentaose ถูกนำมาวิเคราะห์หาโครงสร้างสามมิติ
ของเอนไซม์ โดยผลิตภัณฑ์ทั้ง 4 ชนิดสามารถหักเหรังสีเอ็กซ์ได้ความละเอียดถึง 2.20 1.52 1.35 และ
1.80 อังสตรอม ตามลำดับ จากการเปรียบเทียบโครงสร้างสามมิติกับเอนไซม์ชนิดอื่นในตระกูล
เดียวกันพบว่า โครงสร้างโดยรวมเหมือนกันแต่แตกต่างกันเฉพาะบริเวณโครงสร้างภายนอกที่เป็น
บริเวณทางเข้าของตำแหน่งเร่งปฏิกิริยา ซึ่งตำแหน่งดังกล่าวมีความหลากหลายของลำดับกรด
อะมิโนในเอนไซม์ตระกูลเดียวกัน แสดงให้เห็นว่าบริเวณดังกล่าวเกี่ยวข้องกับการกำหนด
ความจำเพาะต่อสับสเตรต จากโครงสร้างสามมิติของเอนไซม์ BGlul พบว่าบริเวณเร่งของเอนไซม์
เหมาะสมต่อการย่อยของโอลิโกแซคคาไรด์ที่ต่อกันด้วยพันธะ β -1,4 เนื่องจากมีลักษณะเป็นร่องลึก
ขนาดยาวที่ใหญ่บริเวณทางเข้าและแคบที่บริเวณด้านใน โครงสร้างสามมิติของโปรตีนดั้งเดิมกับตัว
ยับยั้ง แสดงให้เห็นภาพการเกิดปฏิกิริยา transglycosylation ซึ่งเกิดได้ดีโดยการเร่งของเอนไซม์
BGlul โดยมีโมเลกุลกลีเซอรอลเข้ามาทำปฏิกิริยาเป็นตัวรับกลูโคสที่เกิดจากการเร่งปฏิกิริยา
การศึกษาโครงสร้างสามมิติของโปรตีนกลายพันธุ์ E176Q กับสับสเตรต cellopentaose พบการ
เปลี่ยนแปลงที่บริเวณผิวนอกของ loop C ที่ตำแหน่งกรดอะมิโน Y341 V342 และ F343 ซึ่ง
พบว่ากรดอะมิโน Y341 มีหน้าที่ในการจับวงแหวนกลูโคสที่ 4 และ 5 ของสับสเตรต

เนื่องจากเอนไซม์เบตาไกลูโคซิเดส BGlul ในข้าว และ BGQ60 ในบาร์เลย์มีความจำเพาะ
ต่อการย่อยสับสเตรต cellobiose และ cellotriose แตกต่างกัน ดังนั้นกรดอะมิโนที่บริเวณตำแหน่งที่
ทำหน้าที่จับกับสับสเตรตด้านปลาย (aglycone binding site) ของเอนไซม์ในข้าว จึงถูกเปลี่ยนให้
เป็นกรดอะมิโนของเอนไซม์ในบาร์เลย์ แม้การกลายพันธุ์ที่ตำแหน่งกรดอะมิโนเพียงตัวเดียวนั้น ไม่
สามารถเปลี่ยนแปลงความจำเพาะต่อสับสเตรตได้ แต่จากการศึกษาจลศาสตร์การเร่งปฏิกิริยาของ
เอนไซม์กลายพันธุ์พบว่า กรดอะมิโน I179 N190 และ N245 จับกับสับสเตรตและเกี่ยวข้องกับการ
เร่งปฏิกิริยาของเอนไซม์ การศึกษาจลศาสตร์การเร่งของเอนไซม์กลายพันธุ์ที่ตำแหน่งกรดอะมิโนที่

ทำหน้าที่ในการเร่งปฏิกิริยาของเอนไซม์ BGlu1 (nucleophile และ acid/base residues) สามารถยืนยันได้ว่ากรดอะมิโนทั้งสอง (E386 และ E176) เป็นตัวเร่งปฏิกิริยาของเอนไซม์ BGlu1 นอกจากนั้นตัวเร่งที่ทำหน้าที่ให้และรับโปรตอน (acid/base residue) ยังสามารถกลับมาเร่งปฏิกิริยาได้ใหม่ในสภาวะที่มีโมเลกุลขนาดเล็กที่มีประจุลบเข้ามาทำหน้าที่ในการรับโปรตอนแทน และเอนไซม์กลายพันธุ์ E176Q มีความสามารถในการย่อยสลาย *p*-nitrophenol glucoside ได้ดีกว่าเอนไซม์กลายพันธุ์ตัวอื่นที่เคยรายงานมาก่อนหน้านี้

WATCHALEE CHUENCHOR : STRUCTURE-FUNCTION

RELATIONSHIPS IN RICE BGLU1 β -GLUCOSIDASE. THESIS

ADVISOR : ASSOC. PROF. JAMES R. KETUDAT-CAIRNS, Ph.D. 189 PP.

STRUCTURE-FUNCTION/ β -GLUCOSIDASE/X-RAY CRYSTALLOGRAPHY/
CELLOOLIGOSACCHARIDES/ENZYME-SUBSTRATE INTERACTIONS

Rice BGlu1, a plant β -glucosidase highly expressed in flower and germinating shoot, hydrolyzes cell wall-derived oligosaccharides. To understand the basis of substrate specificity in glycosyl hydrolase family 1 (GH1) β -glucosidases, especially plant enzymes acting on oligosaccharides, crystals were generated and the crystal structures of free wild type BGlu1, its covalently bound complex with 2-deoxy-2-fluoroglucoside, E176Q mutant BGlu1, and E176Q BGlu1 in complex with cellopentaose were solved at 2.20, 1.52, 1.35, and 1.80 Å resolution, respectively. The structures were similar to previous GH1 β -glucosidases, but showed several differences in the loops around the active site, which lead to an open active site with a narrow slot at the bottom, compatible with the hydrolysis of long β -1,4-linked oligosaccharides. The complex with the 2-fluoroglucoside provides a snap shot of the transglycosylation reaction, which is efficiently catalyzed by rice BGlu1, with a glycerol from the cryoprotectant playing the role of the incoming acceptor sugar. Two alternative conformations were found for the flexible loop C at residues Y341, V342, and F343, which are located at the entrance to the active site, in the E176Q mutant structure, though Y341 appeared to orient Glc4 and Glc5 of cellopentaose bound in the active site of E176Q mutant.

As the rice and barley enzymes have different preferences for cellobiose and cellotriose, putative aglycone binding residues were mutated to those of barley enzyme. Though no single residue appeared to be responsible for these differences, I179, N190 and N245 did appear to interact with the substrates and play a role in catalysis. Two catalytic residues (E386 and E176) were confirmed by kinetic analysis of nucleophile and acid/base mutants. The role of the acid/base catalyst was assessed by chemical rescue of the E176Q mutant with addition of anionic nucleophiles, which showed that this residue is indeed the catalytic acid/base and this mutant was especially active with *p*-nitrophenol glucoside compared to previously studied acid/base mutants.

School of Biochemistry

Academic Year 2007

Student's Signature _____

Advisor's Signature _____

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to my advisor Assoc. Prof. Dr. James R. Ketudat Cairns for providing me the opportunity to study towards my Ph.D. degree in Biochemistry, and for his patient and continuous guidance, unconditional support, and encouragement. I would also express my deepest gratitude to Dr. Jirundon Yuvaniyama and Assoc. Prof. Dr. Robert C. Robinson for providing me the chance to learn x-ray crystallography and valuable technical skills, and their precious advice. I gratefully recognize the contribution of Prof. Dr. Asim Esen in supporting my work at Virginia Tech, scientific discussions, and excellent guidance. I would like to acknowledge Assoc. Prof. Dr. David R. Bevan for valuable instruction on computational docking and useful discussions.

I would like to express special thanks for my colleague Salila Pengthaisong for kindly assisting in crystallization and sharing chemical stocks. I would like to thank Dr. Rodjana Opassiri for advice on biochemical techniques and providing me excellent data of rice BGlul and the construct for protein production. I would like to thank Asst. Prof. Dr. Chun-Jung Chen for his kindly help on data collection at the synchrotron and his valuable discussions. I would like to thank Hong-Hsiang Guan for assistance with diffraction and dynamic light scattering data collection. I would like to thank Apichai Boachookan and Asst. Prof. Dr. Chartchai Krittanai for help with the circular dichroism experiment. I would like to thank Asst. Prof. Dr. Wipa Suginta and Assoc. Prof. Dr. Malee Tangsathitkulchai for patiently reading through this dissertation and providing helpful comments.

I would like to give special thanks to Catleya Rojviriyaya, who is my lovely friend, for continuous help, friendly discussion, and advice with protein crystallography software. These thanks are also extended to Worrapoj Oonant for his tremendous help on the installation of crystallography softwares and his friendly encouragement. I would like to express my sincere gratitude to Asst. Prof. Dr. Mariena Ketudat-Cairns for her love and support and also to her daughters for playing with me. I would like to thank my colleagues and staff in Biochemistry department for their friendly encouragement and help.

I acknowledge the Thailand Research Fund (TRF) for providing me the Royal Golden Jubilee (RGJ) Ph.D. Scholarship for financial support. I acknowledge the protein crystallography beamline staff of the National Synchrotron Radiation Research Center (NSRRC), Hsinchu, Taiwan, for data collection and processing assistance and the Center for Excellence in Protein Structure and Function of the Faculty of Science, Mahidol University, for providing in-house x-ray diffraction.

I would like to thank Yingyos Jittayasothorn for his love, special care, understanding, unconditional support and encouragement.

Finally, my sincerest gratitude is to my family. Without their love, encouragement, and support, I would not have been able to complete this dissertation and obtain the doctorate degree.

Watchalee Chuenchor

CONTENT

	Page
ABSTRACT IN THAI.....	I
ABSTRACT IN ENGLISH.....	III
ACKNOWLEDGEMENT.....	V
CONTENTS.....	VII
LIST OF TABLES.....	XV
LIST OF FIGURES.....	XVII
LIST OF ABBREVIATIONS.....	XXIII
CHAPTER	
I INTRODUCTION.....	1
1.1 Glycosyl hydrolase (GH) clan A and GH family 1.....	1
1.1.1 Glycosyl hydrolase emzymes.....	1
1.1.2 GH clan A.....	1
1.1.3 GH family 1.....	2
1.1.4 β -glucosidase.....	3
1.2 Enzyme mechanism.....	4
1.2.1 Inverting mechanism.....	4
1.2.2 Retaining mechanism.....	5
1.2.3 Catalytic residues.....	7
1.2.4 Acid/base and nucleophile mutants with rescued activity.....	8
1.3 Crystal structures of GH1 enzymes.....	10

CONTENTS (Continued)

	Page
1.4 Substrate specificity.....	16
1.4.1 Glycone specificity.....	16
1.4.2 Aglycone specificity.....	16
1.5 Rice BGlu1 β -glucosidase.....	18
1.5.1 Enzyme activity of rice and barley β -glucosidases.....	19
1.5.2 Subsite mapping of rice and barley β -glucosidases.....	20
1.6 Transglycosylation activity.....	21
1.7 Research objectives.....	24
II MATERIALS AND METHODS.....	25
2.1 Materials.....	25
2.1.1 Recombinant plasmid DNA and bacterial strains.....	25
2.1.2 Oligonucleotides and mutagenic primers.....	27
2.2 General methods.....	30
2.2.1 Overlap extension polymerase chain reaction (PCR) method.....	30
2.2.2 Mutagenesis with the QuikChange® Site-Directed Mutagenesis Kit.....	31
2.2.3 Transformation of recombinant expression plasmid into expression host cells.....	33
2.3 Protein expression of wild type and mutant BGlu1.....	34
2.4 Protein purification.....	34
2.4.1 Thioredoxin fusion protein.....	34

CONTENTS (Continued)

	Page
2.4.1.1 Mutants of the acid/base and nucleophile residues.....	34
2.4.1.2 Mutants of putative aglycone binding residues.....	35
2.4.2 Non-fusion protein.....	36
2.4.2.1 Wild type BGlu1.....	36
2.2.4.2 E176Q mutant.....	37
2.5 Protein determination.....	38
2.5.1 β -Glucosidase activity assay.....	38
2.5.2 Bio-Rad protein assay.....	38
2.5.3 Denaturing gel electrophoresis (SDS-PAGE).....	38
2.5.4 Circular dichroism (CD).....	39
2.5.5 N-terminal amino acid sequencing.....	40
2.2.6 Dynamic light scattering.....	40
2.6 Protein crystallization.....	40
2.6.1 Preliminary sample preparation.....	40
2.6.2 Initial screening for crystallization conditions.....	41
2.6.3 Optimization of crystallization conditions.....	42
2.6.4 Microseeding by streak seeding.....	42
2.6.5 Cocrystallization of complexes.....	43
2.6.6 Crystal soaking.....	43
2.6.7 Crystal viewing.....	43
2.7 Data collection and processing.....	44

CONTENTS (Continued)

	Page
2.7.1 Crystal shipping in sitting drop plates.....	44
2.7.2 Cryofreezing technique.....	44
2.7.3 Crystal shipping by dry shipper.....	44
2.7.4 In-house x-ray diffraction.....	45
2.7.5 Synchrotron x-ray diffraction.....	45
2.8 Molecular replacement.....	47
2.9 Model building and refinement.....	47
2.10 Automated docking of cellobiose, laminaribiose, cellotriose into the BGlu1 structure.....	49
2.11 Enzyme kinetics.....	51
2.11.1 Kinetic parameters analysis.....	51
2.11.2 pH profile.....	52
2.11.3 Effect of anionic nucleophile on rescue activity of acid/base mutants.....	52
2.11.4 Nucleophile concentration dependence of the effects on the activity of wild type and E176Q mutant BGlu1.....	53
2.11.5 Kinetics of activity rescued by external anionic nucleophiles.....	53
2.11.6 TLC analysis.....	54
III RESULTS.....	55
3.1 Protein purification.....	55

CONTENTS (Continued)

	Page
3.1.1 Thioredoxin fusion protein.....	55
3.1.2 Nonfusion protein of wild type BGlu1	60
3.1.3 Nonfusion protein of the E176Q mutant.....	65
3.2 Cicular dichroism.....	66
3.3 Dynamic light scattering (DSL).....	67
3.4 Protein crystallization.....	68
3.4.1 Wild type BGlu1 without inhibitor.....	68
3.4.1.1 Initial screening.....	68
3.4.1.2 Coarse screen crystallization trials of promising conditions.....	73
3.4.1.3 Fine screen crystallization trails.....	73
3.4.2 Wild type BGlu1 with inhibitor.....	78
3.4.2.1 Initial screening.....	78
3.4.2.2 Optimization with seeding.....	78
3.4.3 E176Q mutant.....	79
3.4.3.1 Initial screening.....	79
3.4.3.2 Optimization without micro-seeding.....	80
3.4.3.3 Optimization with micro-seeding.....	80
3.4.4 E176Q mutant with cellopentaose.....	82
3.4.4.1 Optimization with micro-seeding.....	82
3.5 N-terminal amino acid sequencing.....	84

CONTENTS (Continued)

	Page
3.6 Data collection.....	85
3.6.1 Crystal mounting for x-ray diffraction.....	85
3.6.2 XAS scan of BGlu1 crystal at zinc edge.....	92
3.7 Molecular replacement.....	93
3.8 Structure determination of wild type BGlu1 with and without inhibitor.....	95
3.8.1 Overall structures and quality of models.....	95
3.8.2 Overall structure comparison of rice BGlu1 with other structures in GH1 family.....	99
3.8.3 Enzyme active site.....	104
3.8.4 Docking studies.....	110
3.9 Determination of the structure of the BGlu1 E176Q mutant and its complex with cellopentaose.....	114
3.9.1 Free E176Q mutant structure.....	114
3.9.2 Two conformational changes of loop C.....	116
3.9.3 E176Q mutant with its complex of cellopentaose.....	117
3.9.4 Cellopentaose binding and interactions involved in the extended binding site.....	118
3.10 Studies of substrate specificity by mutation at putative aglycone binding residues.....	121
3.10.1 Rice BGlu1 and ZmGlu1.....	121

CONTENTS (Continued)

	Page
3.10.2 Analysis of rice BGlu1 and barley BGQ60 β -glucosidase substrate binding residues.....	124
3.10.3 Kinetic analysis of substrate binding pocket mutants.....	126
3.11 Studies of the catalytic mechanism by mutation at the acid/base and nucleophile residues.....	129
3.11.1 β -glucosidase activity for <i>p</i> NPG hydrolysis of acid/base and nucleophile mutants.....	129
3.11.2 Effect of pH on β -glucosidase activity.....	130
3.11.3 The ability of various anionic nucleophiles to rescue the activity of acid/base mutants.....	134
3.11.4 pH dependence in the presence of azide and ascorbate.....	136
3.11.5 Rate enhancement by external nucleophile anions.....	139
3.11.6 Characterization of the products from rescued activity.....	142
VI DISCUSSION.....	145
4.1 Protein expression and purification of rice BGlu1 β -glucosidase.....	145
4.2 Protein crystallization.....	146
4.3 Crystal structures of rice BGlu1.....	148
4.3.1 Overall structure.....	148
4.3.2 Monomer and dimerization by zinc ion.....	148
4.3.3 Active site structure.....	150

CONTENTS (Continued)

	Page
4.3.4 Computational docking of cellobiose, laminariniose and cellotriose into the crystal structure of rice BGlu1	154
4.3.5 Crystal structures of the E176Q mutant soaked in laminaribiose and its complex with cellopentaose	155
4.4 Substrate specificity	157
4.5 Enzyme mechanism and rescue activity	161
IV CONCLUSION	165
REFERENCES	170
CURRICULUM VITAE	189

LIST OF TABLES

Table	Page
1.1 Available crystal structures of GH family enzymes.....	11
1.2 Comparison of the kinetic parameters of rice BGlu1 and barley β -glucosidase.....	20
2.1 Recombinant plasmid DNA used for this experiment.....	26
2.2 Bacterial strains.....	26
2.3 Oligonucleotide primers used in mutagenesis.....	28
2.4 Internal oligonucleotide primers used for DNA sequencing of full-length rice <i>bglu1</i>	29
2.5 PCR conditions for the classic PCR site-directed mutagenesis method.....	31
2.6 PCR condition for the QuikChange site-directed mutagenesis method.....	32
2.7 Details of BGlu1 mutations and internal primers used for sequencing the full length mutant cDNAs.....	33
3.1 Crystallization conditions of free wild type BGlu1 from screening kits that yielded small crystals in the microbatch method.....	70
3.2 The crystallization conditions obtained from optimization for the BGlu1 inhibitor complex.....	79
3.3 Summary of drop ingredients, crystallization conditions, sizes of diffracted crystals, cryoprotectants, and obtained resolution.....	87
3.4 Data collection statistics for BGlu1, BGlu1 with 2-fluoroglucoside, E176Q mutant, and E176Q mutant with cellopentaose.....	90

LIST OF TABLES (Continued)

Table	Page
3.5 Refinement statistics for BGlu1, BGlu1 with inhibitor, inactive E176Q mutant, and inactive E176Q mutant with cellopentaose.....	96
3.6 Hydrogen bonding interactions with conserved residues in the free-BGlu1 and G2F complex active sites.....	106
3.7 Kinetic parameters of wild type rice BGlu1, putative aglycone binding residue mutants and barley BGQ60 β -glucosidase for hydrolysis on <i>p</i> NPG, cellobiose, and cellotriose.....	128
3.8 Effects of mutations of the putative nucleophile and acid/base residues on <i>p</i> NPG hydrolysis.....	130
3.9 Apparent K_m , k_{cat} , and k_{cat}/K_m of E176Q mutant BGlu1 for <i>p</i> NPG hydrolysis in the presence and absence of various concentrations of acetate and azide	140
3.10 Apparent K_m , k_{cat} , and k_{cat}/K_m of E176Q mutant BGlu1 for 2,4-DNPG hydrolysis in the presence and absence of 0.2 M acetate and azide.....	140

LIST OF FIGURES

Figure	Page
1.1 Double displacement mechanism of retaining β -glucosidases.....	6
1.2 Mechanisms of azide rescue of the activity of a nucleophile mutant and an acid/base mutant.....	9
1.3 A natural version of chemical rescue by <i>Sinapis alba</i> myrosinase which utilizes ascorbate as a general base catalyst.....	10
1.4 Maize β -glucosidase Glu1 structure, which has the common $(\beta/\alpha)_8$ barrel structure found in all glycosyl hydrolase family 1 enzymes.....	15
1.5 Comparison of the active sites of ZmGlu1 β -glucosidase in complex with DIMBOA-Glc and SbDhr1 dhurrinase in complex with dhurrin.....	18
1.6 Comparison of the binding affinities of rice BGlu1 and barley BGQ60 for cello-oligosaccharide glucosyl residues.....	21
1.7 Mechanism of hydrolytic or glycosyl transfer reactions catalyzed by a plant β -glucosidase.....	22
2.1 Construct of the protein-coding sequence of recombinant pET32a(+) with the <i>bglu1</i> cDNA, inserted after the enterokinase cleavage site.....	25
2.2 Crystal mounting for x-ray diffraction	46
3.1 SDS-PAGE of thioredoxin fusion proteins of wild type rice BGlu1 and E176A, D243V, E176Q, E176D, E386Q, and E386D BGlu1 mutant proteins purified with a single step of Ni ²⁺ -NTA IMAC.....	56

LIST OF FIGURES (Continued)

Figure	Page
3.2 Purification of thioredoxin fusion protein for wild type BGlu1 using Co ²⁺ metal ion instead of the Ni ²⁺ -based resin.....	58
3.3 Elution profile of thioredoxin fusion protein of wild type rice BGlu1 obtained from S200 gel filtration chromatography and SDS-PAGE profile.....	59
3.4 Rice BGlu1 aglycone-binding residue mutant purification.....	60
3.5 Elution profile of the wild type BGlu1 nonfusion protein passed through an S200 gel filtration column and SDS-PAGE of fractions at the protein peak	62
3.6 Elution profile of nonfusion protein for BGlu1 obtained from SP sepharose cation exchange chromatography with a two-step salt gradient.....	63
3.7 Elution profile of the final purification step of nonfusion protein for BGlu1 enzyme and SDS-PAGE of protein fractions obtained from a single peak on gel filtration chromatography.....	64
3.8 SDS-PAGE of purified BGlu1 fractions from the steps of purification of BGlu1 for crystallization.....	65
3.9 SDS-PAGE analysis of the E176Q mutant from the steps of the purification for crystallization.....	66
3.10 CD spectra in the peptide bond region of wild type and all mutants.....	67
3.11 Dynamic Light Scattering (DLS) spectrum of rice BGlu1 fusion protein.....	68
3.12 Crystals of free wild type BGlu1 in microbatch plates from the initial	

LIST OF FIGURES (Continued)

Figure	Page
screening with the Hampton HR2-130 screening kit.....	71
3.13 Crystals of free wild type BGlu1 in a microbatch plate from the initial screening with a Hampton HR2-134 screening kit.....	72
3.14 Grid screen I with variation of (NH ₄) ₂ SO ₄ , PEG 5000 MME, and pH.....	73
3.15 Grid screen II with variation of (NH ₄) ₂ SO ₄ , PEG 5000 MME, and pH.....	75
3.16 Grid screen III for crystallization of free BGlu1 and single crystals used for data collection.....	77
3.17 Crystals of BGlu1 with the inhibitor DNP2FG.....	79
3.18 Optimization of crystallization for free E176Q mutant BGlu1 and a single crystal used for x-ray diffraction.....	81
3.19 Optimization of crystallization of E176Q mutant with cellopentaose and single crystal used for x-ray diffraction.....	83
3.20 Protein from dissolved crystals of rice BGlu1 blotted onto PVDF membrane before protein bands were cut for sequencing.....	84
3.21 X-ray diffraction patterns of BGlu1 with DNP2FG inhibitor and E176Q mutant, which diffracted to 1.55 and 1.37 Å resolution.....	86
3.22 XAS scan of the zinc edge for a BGlu1 crystal with plots of f' and f''	92
3.23 The improvement of the solved structure indicated by the $2 F_{\text{obs}} - F_{\text{calc}} $ map contoured at the 1σ level at amino acid residues 206-215 of rice BGlu1.....	94
3.24 A ribbon diagram representing the overall structure of the	

LIST OF FIGURES (Continued)

Figure	Page
BGlu1/G2F inhibitor complex asymmetric unit.....	97
3.25 Superimposition of the structures of rice BGlu1 with other enzymes in GH family 1.....	100
3.26 Amino acid sequence alignment of rice BGlu1 β -glucosidase with related enzymes.....	103
3.27 Molecular surface of the G2F inhibitor complex and free structure with a docked celotriose.....	105
3.28 Stereo view of protein-ligand interactions in the active site of the BGlu1-G2F complex.....	109
3.29 Superimposition of subsite -1 of the active site of free rice BGlu1 and the covalently bound G2F complex.....	110
3.30 Crystal structure of free BGlu1 with cellobiose, laminaribiose, and celotriose from docking.....	113
3.31 Electron density of the $2 F_{\text{obs}} - F_{\text{calc}} $ map at the active site of E176Q structure (soaked with laminaribiose).....	115
3.32 Electron density map of two conformations of loops C in the free E176Q mutant structure.....	116
3.33 The conformation of cellopentaose bound to E176Q BGlu1 and its electron density.....	118
3.34 Comparison of cellopentaose alignment from the crystal structure of the E176Q mutant and celltriase from automated docking.....	119

LIST OF FIGURES (Continued)

Figure	Page
3.35 The subsite positions indicated by cellopentaose binding in the rice BGlu1 active site	120
3.36 Superimposition of rice BGlu1 with ZMGlu1 maize β -glucosidase in complex with the natural substrate DIMBOAGlc.....	123
3.37 Differences in the active site geometry of rice BGlu1 and ZMGlu1 with celotriose and DIMBOAGlc.....	124
3.38 pH profiles of wild type BGlu1 and BGlu1 with mutations at substrate binding and acid/base residues with a series of overlapping pH buffer.....	132
3.39 Comparison of two different buffer systems: a series of overlapping pH buffers and universal pH buffer.....	133
3.40 Effects of various anionic nucleophiles on acid/base mutants and wild type BGlu1.....	135
3.41 Comparison of pH dependence of the wild type, E176Q, E176A, and E176D enzymes in 50 mM azide in universal pH buffer.....	137
3.42 Comparison of the pH dependence of the wild type, E176Q, E176A, and E176D enzymes in 50 mM ascorbate in universal pH buffer.....	138
3.43 The concentration dependence of azide, acetate, and ascorbate for <i>p</i> NPG hydrolysis by the BGlu1 E176Q mutant.....	139
3.44 Detection of the products of <i>p</i> NPG reactions with E176Q and wild type BGlu1 by TLC.....	143

LIST OF FIGURES (Continued)

Figure	Page
4.1 Crystallographic dimerization of rice BGlu1 by Zn ²⁺	150
4.2 View of the transglycosylation intermediate in the crystal structure of the G2F complex.....	152
4.3 Molecular surface of the E176Q mutant illustrating the alternative conformations of loop C at residues Y341, V342, and F343.....	156
4.4 Schematic representation of rice BGlu1 active site and the binding modes of cellobiose binding and a histogram of subsite affinities of rice BGlu1 for cellooligosaccharide hydrolysis.....	160

LIST OF ABBREVIATIONS

2,4-DNPG	2,4-dinitrophenyl β -D-glucopyranoside
2F-DNPG	2,4-dinitrophenyl 2-deoxy-2-fluoro- β -D-glucopyranoside
A	absorbance
Å	Angstrom
APS	Ammonium persulfate
AU	asymmetric unit
°C	degrees Celsius
BSA	bovine serum albumin
bis-acrylamide	N,N-methylene-bis-acrylamide
CaCl ₂	calcium chloride
CC	correlation coefficient
CD	Circular dichroism
cDNA	complementary deoxynucleic acid
DNP	2,4-dinitrophenolate
DNA	deoxyribonucleic acid
DNaseI	deoxyribonuclease I
DP	degree of polymerization
dsDNA	double strand deoxyribonucleic acid
EDTA	ethylene diamine tetraacetic acid
FPLC	fast purification liquid chromatography

LIST OF ABBREVIATIONS (Continued)

g	gravitational acceleration
(μ /m)g	(micro, milli) gram
h	hour
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IMAC	immobilized metal-affinity chromatography
IPTG	isopropyl β -D-thioglucoopyranoside
K	Kelvin
(k)bp	(kilo) base pair
kDa	kilodalton
kV	kilovolt
(μ /m)l	(micro, milli) liter
L2	laminaribiose
Li ₂ SO ₄	lithium sulfate
LGA	Lamarckian genetic algorithm
(n/ μ /m)m	(nano, micro, milli) miter
(μ /m)M	(micro, milli) molar
M_r	molecular weight
MES	2-(<i>N</i> -morpholino)ethanesulfonic acid
min	minute
MgCl	magnesium chloride
Mg OAc	magnesium acetate
MME	monomethylether

LIST OF ABBREVIATIONS (Continued)

MWCO	molecular weight cut-off
Na	sodium
NaCl	sodium chloride
NaN ₃	sodium azide
NCS	noncrystallographic symmetry
NH ₄ OAc	ammonium acetate
(NH ₄) ₂ SO ₄	ammonium sulfate
NSRRC	National Synchrotron Radiation Research Center
PCR	Polymerase chain reaction
PGO	peroxidase-glucose oxidase
PMSF	phenylmethylsulfonylfluoride
PNPG	<i>p</i> -nitrophenyl β-D-glucopyranoside
PNP	<i>p</i> -nitrophenol
PAGE	polyacrylamide gel electrophoresis
PEG	polyethyleneglycol
RMSD	root-mean-square deviation
rpm	revolutions per minute
S200	Superdex 200
SDS	Sodium dodecyl sulfate
sec	second
Tris	Tris-(hydroxymethyl)-aminomethane
T _m	melting temperature

LIST OF ABBREVIATIONS (Continued)

TEMED	N, N, N', N'-Tetramethylethylenediamine
TLC	Thin layer chromatography
V_M	Matthew's coefficient
V_0	initial velocity
v/v	volume by volume
w/v	weight by volume
ZnOAc	zinc acetate

CHAPTER I

INTRODUCTION

1.1 Glycosyl hydrolase (GH) clan A and GH family 1

1.1.1 Glycosyl hydrolase enzymes

Glycosyl hydrolases (GH) (EC 3.2.1.-) are widespread enzymes that hydrolyze the glycosidic bonds between carbohydrates or between a carbohydrate and a noncarbohydrate moiety (Durand et al., 1997). The enzymatic hydrolysis occurs by one of two major mechanisms leading to overall retention or inversion of anomeric configuration. Based on sequence similarities, glycosyl hydrolases were classified into 110 families (Henrissat, 1991; el Hassouni et al., 1992; Henrissat and Bairoch, 1993; 1996; Davies and Henrissat, 1995). This classification is available and updated on the CAZy (Carbohydrate-Active EnZymes) web site; www.cazy.org/CAZY/ (Coutinho and Henrissat, 1999).

1.1.2 GH clan A

The GH clan A (or superfamily 4/7) is grouped based on the three-dimensional structure and comprises more than 1000 proteins from 17 different GH families, 1, 2, 5, 10, 17, 26, 30, 35, 39, 42, 50, 51, 53, 59, 72, 79, and 86, that act with different substrate specificities but evolved from a common ancestry (Henrissat et al., 1995; Jenkins et al., 1995; Bolam et al., 1996; Henrissat and Bairoch, 1996). Despite low levels of sequence identity, all the proteins in this clan share a similar (α/β)₈ barrel

structure with two conserved carboxylate groups located at the ends of β -strands 4 and 7 (though the former is a glutamine in plant myrosinases). These enzymes hydrolyze the glycosidic bond with retention of the anomeric configuration.

1.1.3 GH family 1

Despite their similar sequences, GH family 1 (GH1) enzymes have a wide range of activities, including β -glucosidases (EC 3.2.1.21), β -galactosidase (EC 3.2.1.23), β -mannosidase (EC 3.2.1.25), β -glucuronidase (EC 3.2.1.31), β -D-fucosidase (EC 3.2.1.38), phlorizin hydrolase (EC 3.2.1.62), 6-phospho- β -galactosidase (EC 3.2.1.85), 6-phospho- β -glucosidase (EC 3.2.1.86), strictosidine β -glucosidase (EC 3.2.1.105), lactase (EC 3.2.1.108), amygdalin β -glucosidase (EC 3.2.1.117), prunasin β -glucosidase (EC 3.2.1.118), raucassicine β -glucosidase (EC 3.2.1.125), thioglucosidase (EC 3.2.1.147, EC 3.2.3.1), diglycosidases like β -primeverosidase (EC 3.2.1.149), furcatin hydrolase and isoflavone 7-*O*- β -apiosyl- β -1,6-glucosidase (EC 3.2.1.161), and hydroxyisourate hydrolase, which hydrolyzes a purine ring rather than a glycosidic bond (Coutinho and Henrissat, 1999; CAZY: www.cazy.org/CAZY/). GH1 glycosyl hydrolases also display a range of glycone and aglycone specificities, with some showing nearly absolute specificity for one sugar and one aglycone, while others accept a range of either glycone or aglycone or both. As such, GH1 provides a useful model for the study of structure-function relationships in enzyme specificity.

1.1.4 β -glucosidase

β -glucosidases (EC 3.2.1.21) catalyse the hydrolysis of β -*O*-glycosidic bonds at the nonreducing ends of a broad variety of glycosides, including aryl- and alkyl- β -D-glycosides, releasing β -D-glucose and an aglycone (Reese, 1977). They belong predominantly to the GH families 1, 3, and 5, with family 1 enzymes being more numerous in plants (Henrissat, 1991; Opassiri et al., 2006). These enzymes occur in all types of living organisms, including plants, fungi, bacteria, archaea, and animals (Woodward and Wiseman, 1982). They have uses in a variety of biotechnological processes, such as biomass conversion (Wattenberg, 1971), food detoxification and beverage quality enhancement (Fenwick et al., 1983).

Plant β -glucosidases are primarily classified into GH family 1 according to their primary structure (Henrissat and Bairoch, 1993; Henrissat et al., 1995). They have been widely found in a variety of plants, including *Arabidopsis* (Xu et al., 2004), rice (Palmiano and Juliano, 1973; Schliemann, 1984; Akiyama et al., 1998; Opassiri et al., 2003), sorghum (Hosel et al., 1987), maize (Bandaranayake and Esen, 1996), barley (Simos et al., 1994; Leah et al., 1995; Hrmova et al., 1996), white clover (Hughes and Dunn, 1982), Thai rosewood (Surarit et al., 1995), cassava (Hughes et al., 1992), and many others. They have been implicated in key roles in biological processes, such as hydrolysis of cell wall-derived oligosaccharides during germination (Simos et al., 1994; Leah et al., 1995; Hrmova et al., 1996), chemical defense against the plant pathogens and herbivores by release of bitter and toxic aglycones, such as thiocyanates, isothiocyanates, nitrites, terpenoid alkaloids, saponins, hydroxymates, benzaldehydes, and HCN (Fenwick et al., 1983; Jones, 1988; Niemeyer, 1988; Poulton, 1990; Oxtoby, 1991), control of active phytohormone levels (Smith and Van

Staden, 1978; Schliemann, 1984; Matsuzaki and Koiwai, 1986; Wiese and Grambow, 1986; Brzobohaty et al., 1993; Falk and Rask, 1995), lignification (Dharmawardharna et al., 1995), and vitamin B₆ metabolism (Opassiri et al., 2003).

1.2 Enzyme mechanism

Glycosyl hydrolase enzymes have two classes of elementary mechanisms, consisting of inverting and retaining mechanisms (Sinnott, 1990; McCarter and Withers, 1994; Ly and Withers, 1999). Both of these use a pair of carboxylate groups to play key roles in the mechanism. However, the distance between two carboxylate residues is different for the two mechanisms, with $\sim 9.5 \text{ \AA}$ for the inverting and a shorter distance ($\sim 5.5 \text{ \AA}$) for the retaining mechanism (Sinnott, 1990; McCarter and Withers, 1994; Wang et al., 1994; White et al., 1994; Davies et al., 1998; Withers, 1999; Zechel and Withers, 1999).

1.2.1 Inverting mechanism

The inverting glycosides use a direct displacement mechanism in which the two carboxylate groups are positioned approximately 9.5 \AA apart. This distance is suitable for one to provide general base i.e. assistance to activate water as a nucleophile, while the other provides a general acid residue to protonate the leaving group in cleavage of the glycosidic bond. The large separation between two carboxylate groups allows the water and the substrate to bind simultaneously.

1.2.2 Retaining mechanism

GH family 1 members perform catalysis with retention of configuration at the anomeric carbon via a double-displacement mechanism. The catalytic machinery of these enzymes involves two catalytic carboxylates on the opposite sides of the sugar plane at a distance of 5.0-5.5 Å apart, as mentioned above, which perform the catalysis in two separate chemical steps. In the first step (glycosylation), the oxygen at the glycosidic bond is protonated by the first carboxylate, residue which acts as the general acid catalyst, while the second carboxylate group is the nucleophile, which attacks at the anomeric carbon, resulting in the formation of a covalent glycosyl-enzyme intermediate. This intermediate is then hydrolyzed in a second step (deglycosylation), in which the first catalytic residue acts as a general base to activate the incoming nucleophile (a water molecule in the case of hydrolysis, and an alcohol in the case of transglycosylation), which cleaves the covalent glycosyl-enzyme intermediate (Figure 1.1).

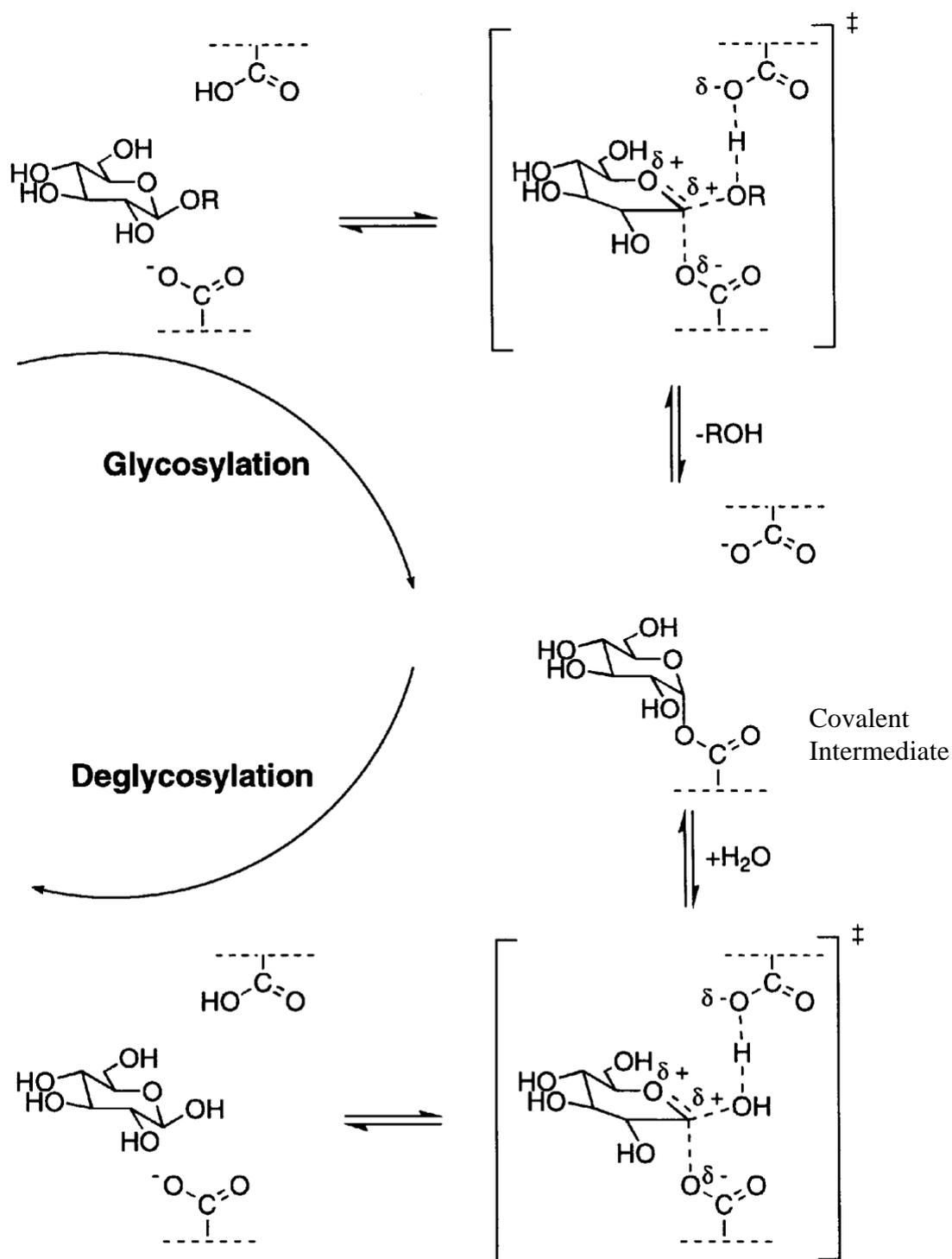


Figure 1.1 Double displacement mechanism of retaining β -glucosidases. R represents an alkyl or aryl group, which serves as the leaving group in the glycosylation step (Lawson et al., 1998).

1.2.3 Catalytic residues

The two catalytic amino acids in clan A β -glycosidases have been extensively investigated by mechanism-based labeling of catalytic amino acid residues, crystallography, and kinetic analysis of mutant enzymes. A precise identification of the catalytic nucleophile and the acid/base are available through the work of Withers and his collaborators (White et al., 1994). The 2-deoxy-2-fluoro- β -D-glycosides have been used successfully in identification of the catalytic nucleophile of several β -glucosidases (Withers and Street, 1988; Street et al., 1992; MacLeod et al., 1994; Withers and Aebersold, 1995; He and Withers, 1997; Dan et al., 2000). Since the substitution of the 2-hydroxy group of the substrate with fluorine causes electron withdrawal, it destabilizes the accumulation of a partial positive charge on C1 and O5 in the transition states. Therefore, the presence of the 2-fluoride leads to formation of stable covalent glycosyl-enzyme intermediates (similar to the covalent intermediate in Figure 1.1, but with a 2-F group in place of the 2-OH) from which the glycone product is released very slowly, when a reactive leaving group, such as dinitrophenolate, is incorporated into the 2-fluoroglucoside, thereby accelerating the glycosylation step. Trapping the intermediate in this way is very close to the natural mechanism (McCarter and Withers, 1994; Withers and Aebersold, 1995; Withers, 2001). In contrast, epoxide compounds, which had previously been used as labeling agents for the catalytic nucleophile, form irreversible covalent attachments to the nucleophile residue, and therefore they do not provide the information about the catalytic mechanism as provided by 2-fluoroglucoside (White and Rose, 1997). Site-directed mutagenesis has been widely used to verify the acid/base and nucleophile residues

tentatively identified based on homology, especially for *Agrobacterium faecalis* β -glucosidase, Abg (Trimbur et al., 1992; Wang et al., 1994; 1995).

1.2.4 Acid/base and nucleophile mutants with rescued activity

Rescue of activity by azide has been used as methodology to identify the acid/base and nucleophile residues. As shown in Figure 1.2A, an Ala mutant of the nucleophile of the Abg enzyme, which completely lacks activity, can be reactivated for hydrolysis of dinitrophenyl glycoside by azide, formate, or acetate with formation of a glycoside with inverted anomeric configuration, i.e. α -glucosyl azide (Wang et al., 1994). In contrast, the same anomeric configuration as the substrate, β -glucosyl azide, was obtained when the activity of an Abg acid/base mutant was rescued (Figure 1.2B). In this case, the good leaving group, which has a low pK_a and therefore does not require protonation, eliminates the requirement for acid catalysis in the glycosylation step, while the small nucleophile (azide) does not require basic assistance in the deglycosylation step (Wang et al., 1995). These results have been observed with mutant species of other retaining β -glycosidases, such as *Cellulomonas fimi* β -1,4-glycanase, *Geobacillus stearothermophilus* T6 β -xylosidase and *Streptomyces* sp. β -glucosidase (MacLeod et al., 1996; Bravman et al., 2003; Vallmitjana et al., 2001).

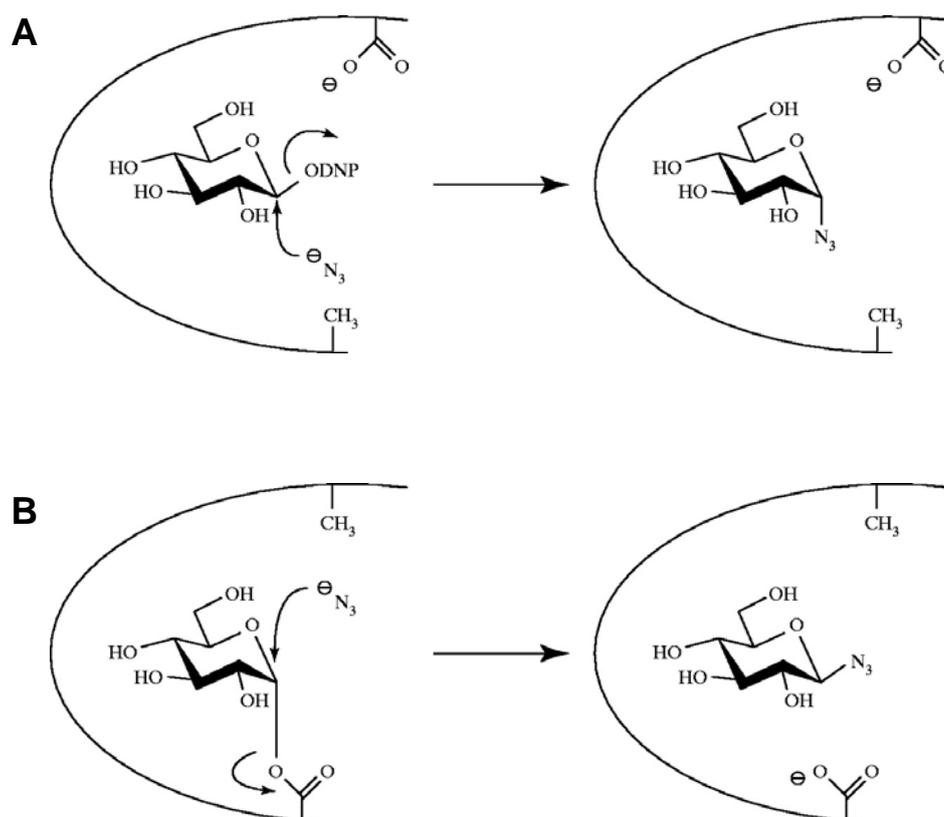


Figure 1.2 Mechanisms of azide rescue of the activity of a nucleophile mutant (A) and an acid/base mutant (B) (Ly and Withers, 1999).

In one catalytic residue variation found in nature, the acid/base glutamate residue, which is highly conserved in *O*-glycosidase, is replaced by glutamine in *Sinapis alba* myrosinase, which is *S*-glycosidase (thioglucosidase, Burmeister et al., 2000). Myrosinase, while lacking an intrinsic acid/base residue, could hydrolyze its highly reactive sinigrin and other glucosinolate substrates to perform the glycosylation step without assistance of an acid catalyst, while ascorbate acts as a base catalyst in the deglycosylation step after departure of the aglycone (Figure 1.3).

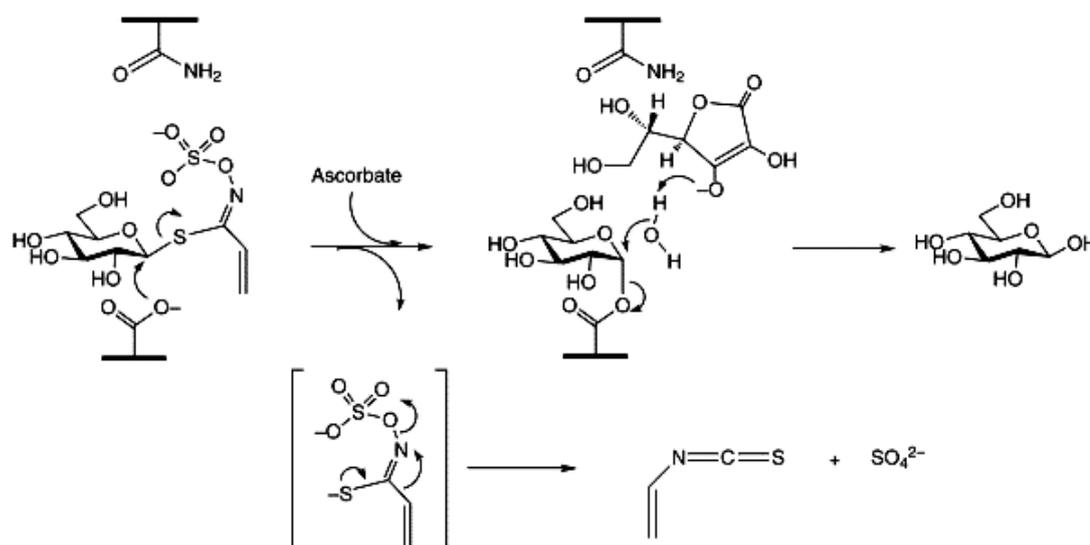


Figure 1.3 A natural version of chemical rescue by *Sinapis alba* myrosinase, which utilizes ascorbate as a general base catalyst (Burmeister et al., 2000; Zechel and Withers, 2001).

1.3 Crystal structures of GH1 enzymes

To date, coordinates for the structures of 20 GH family 1 proteins have been deposited in the Protein DataBank (PDB), 3 from archaea, 8 from bacteria, 2 from animals, 1 from fungus and 6 from plants, representing β -glycosidases, a 6-phospho- β -galactosidase, and β -glucosidases with various aglycone specificities and myrosinases (CAZY). In addition, several complexes with inhibitors and mutants are available, which provide insights into the mechanisms of both catalysis and substrate recognition (Table 1.1). These all have a common $(\beta/\alpha)_8$ or TIM barrel fold, which consists of a core of eight twisted parallel β -strands, connected by eight α -helices, which form the outer layer of the barrel (Figure 1.4). The two catalytic residues are located in the

highly conserved T(F/L/M)NEP and I/VTENG motifs positioned at the end of the active site pocket on opposite sides of the glycosidic bond (Davies and Henrissat, 1995; Henrissat and Bairoch, 1996).

Table 1.1 Available crystal structures of GH family1 enzymes. The structures were identified from the CAZY web site (www.CAZY.org/CAZY/)

Organism	Activity (activities)	PDB	Sequence, other	References
Archaea				
- <i>Pyrococcus horikoshii</i> OT3	alkyl β -glycosidase (membrane-bound) (BGPh;PH0366)	1VFF	BAA29440.1	Akiba et al., 2004
- <i>Sulfolobus solfataricus</i> P2	β -glycosidase S (LacS;Ss- β -gly;SSO3019)	1GOW 1UWI 1UWQ 1UWR 1UWS 1UWT 1UWU 2CEQ 2CER	AAA72843.1	Aguilar et al., 1997; Gloster et al., 2004a; 2004b; 2006b.
- <i>Thermosphaera aggregans</i> M11TL	β -glycosidase (T α - β -gly)	1QVB	AAD43138.1	Chi et al., 1999
Bacteria				
- <i>Bacillus circulans</i> subsp. <i>Alkalophilus</i>	β -glucosidase (BglA)	1QOX	AAA22266.1	Hakulinen et al., 2000
- <i>Lactococcus lactis</i> Z268	6-phospho- β -galactosidase	1PBG 2PBG 3PBG 4PBG	AAA25183.1	Wiesmann et al., 1995; 1997
- <i>Paenibacillus polymyxa</i>	β -glucosidase A (BglA)	1BGA 1BGG 1E4I 1TR1 1UYQ	AAA22263.1	Sanz-Aparicio et al., 1998

Table 1.1 (continued).

Organism	Activity (activities)	PDB	Sequence, other	References
- <i>Paenibacillus polymyxa</i>	β -glucosidase B (BglB)	2JIE 2O9P 2O9R 2O9T 2Z1S	AAA22264.1	Isorna et al., 2007
- <i>Streptomyces sp. QM-B814</i>	β -glucosidase (Bgl3)	1GNX 1GON	CAA82733.1 1.68 Å, 2.2	Guasch et al., to be published.
- <i>Thermotoga maritima MSB8</i>	β -glucosidase A (BglA)	1OD0 1OIF 1OIM 1OIN 1UZ1 1W3J 2CBU 2CBV 2CES 2CET 2J75 2J77 2J78 2J79 2J7B 2J7C 2J7D 2J7E 2J7F 2J7G 2J7H 2JAL	CAA52276.1	Zechel et al., 2003; Vincent et al., 2004; Gloster et al., 2004a; 2006a; 2006b; 2007.
- <i>Thermus nonproteolyticus HG102</i>	β -glycosidase	1NP2	AAF36392.	Wang et al., 2003
- <i>Thermus thermophilus HB8</i>	β -glycosidase (TTHB087)	1UG6	AAN05439.1 0.99 Å	Lokanath et al., to be published.

Table 1.1 (continued).

Organism	Activity (activities)	PDB	Sequence, other	References
Animal				
- <i>Brevicoryne brassicae</i>	myrosinase (BMY1)	1WCG	AAL25999.1 1.1 Å	Husebye et al., 2005
- <i>Homo sapiens</i>	neutral β -glycosylceramidase / β -glucosidase (Gba3; CbgCBgl1; GluC; KLrP) (cytosolic)	2E9L 2E9M 2JFE	CAC08178.1	Tribolo et al., 2007
Fungus				
- <i>Phanerochaete chrysosporium K-3</i>	β -glucosidase (Bg11A)	2E3Z 2E40	BAE87008.1 1.50 Å,	Nijikken et al., 2007
Plant				
- <i>Rauvolfia serpentina</i>	strictosidine β -glucosidase (Sgr1)	2JF6 2JF7	CAC83098.1 Not released	Barleben et al., to be published.
- <i>Sinapis alba</i>	myrosinase	1DWA 1DWF 1DWG 1DWH 1DWI 1DWJ 1E4M 1E6Q 1E6S 1E6X 1E70 1E71 1E72 1E73 1MYR 1W9B 1W9D 2MYR	Q7SIB0 1.2 Å (1E4M)	Burmeister et al., 1997; 2000; Burmeister, 2000; Bourderieux et al., 2005
- <i>Sorghum bicolor P721N</i>	cyanogenic β -glucosidase (dhurrinase 1)	1V02 1V03	Q41290	Verdoucq et al., 2004

Table 1.1 (continued).

Organism	Activity (activities)	PDB	Sequence, other	References
- <i>Trifolium repens</i>	β -glucosidase 2 (cyanogenic)	1CBG	CAA40057.1	Barrett et al., 1995
- <i>Triticum aestivum</i>	β -glucosidase (TaGlu1b)	2DGA	BAE92259.1	Sue et al., 2006
- <i>Zea mays</i>	β -glucosidase 1	1E1E 1E1F 1E4L 1E4N 1E55 1E56 1H49 1V08	AAA65946.1	Czjzek et al., 2000; 2001; Verdoucq et al., 2004.
- <i>Zea mays</i>	β -glucosidase p60.1	1HXJ	CAA52293.1	Zouhar et al., 2001
- <i>Rauvolfia serpentina</i>	raucaffricine β -glucosidase	Crystals	AAF03675.1	

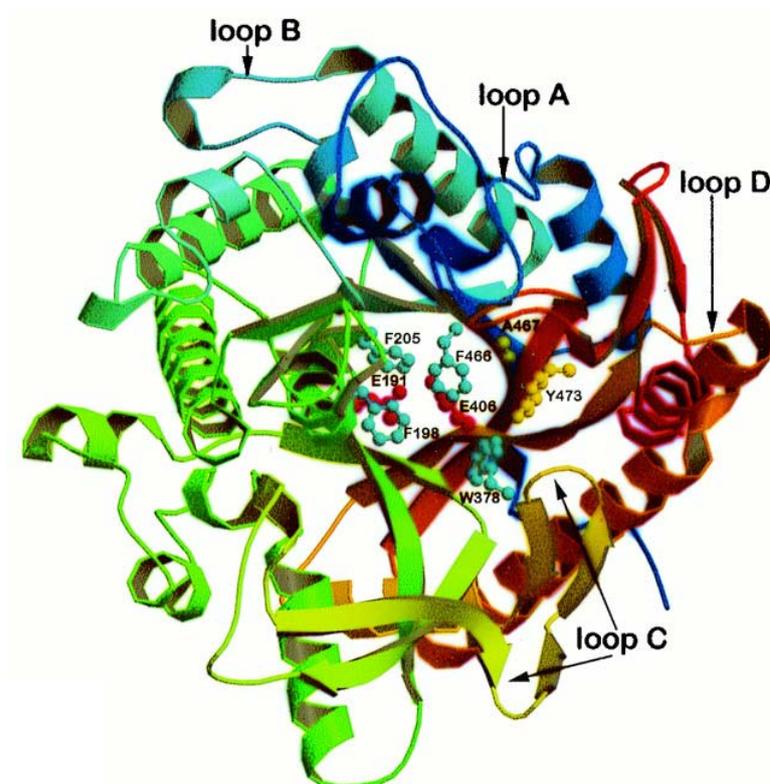


Figure 1.4 Maize β -glucosidase Glu1 structure, which has the common $(\beta/\alpha)_8$ barrel structure found in all glycosyl hydrolase family 1 enzymes. The residues shown in ball-and-stick in the model are the catalytic residues, E191 and E406, and other residues which are important for substrate specificity (Czjzek et al., 2000).

However, the loops at the C-terminal ends of the β -strands vary greatly in both length and sequences, resulting in active sites that range from narrow tunnels to wide pockets, and the solution derived from one enzyme for binding of a certain substrate may not be the same as that of an enzyme from another organism that has evolved separately. Thus, many GH1 enzymes are still under investigation and will yield new insights into glycone and aglycone binding.

1.4 Substrate specificity

1.4.1 Glycone specificity

The hydrogen bonding network at subsite -1 has been reported extensively in several GH1 structures with covalently bound inhibitor, or other ligands (Burmeister et al., 1997; Wiesmann et al., 1997; Sanz-Aparicio et al., 1998; Czjzek et al., 2000; 2001; Zechel et al., 2003; Gloster et al., 2004b; Verdoucq et al., 2004; Isorna et al., 2007). All hydrogen bonds between the glucose hydroxyl groups and the enzyme active site were conserved in GH family 1, although the glutamate residue interacting with OH6 is replaced by serine (S428), lysine (K435), and tyrosine (Y437) for hydrogen bonding with the 6-phosphate group in 6-phospho- β -galactosidase (Wiesmann et al., 1997). However, differences are still seen in the glycone specificity in this family, which relate to the different contacts with the OH in the sugar substrate and transition state (Namchuck and Withers, 1995).

1.4.2 Aglycone specificity

Since the glycone moiety is conserved among GH family 1 β -glucosidases, substrate specificity is mainly determined by high variation in the aglycone part. To date, two plant β -glucosidases have been crystallized in complex with several ligands in the aglycone binding pocket (Czjzek et al., 2000; 2001; Verdoucq et al., 2004). The crystal structures of ZmGlu1 and SbDhr1 with their natural substrates DIMBOA-Glc and dhurrin showed distinctly different aglycone binding pockets. The strict substrate specificity of SbDhr1, which hydrolyzes only dhurrin, is due to tighter binding of the aglycone moiety than ZmGlu1. Two residues, S462 and N259, form indirect hydrogen bonds with the phenol hydroxyl group of dhurrin via water molecules, which act

together with hydrophobic interaction from V196 and L203 (Verdoucq et al., 2004). In contrast, the aglycone moiety was bound with only hydrophobic interactions with W378 on one side and F198, F205, and F466 on the other side in ZmGlu1 (Czjzek et al., 2000). More flexibility for the aglycone moiety for the latter enzyme results in ZmGlu1 hydrolyzing a broader range of substrates, including several artificial substrates. Moreover, the three bulky amino acids (F198, F205, and F466) that form the aglycone binding pocket in ZmGlu1 are replaced by smaller residues (V196, L203, and S462) in SbDhr1. Consequently, the active site of ZmGlu1 is narrow like a slot, while that of SbDhr1 is wider (Figure 1.5).

Changing the substrate specificity between these two enzymes was extensively investigated with both creation of mosaic enzymes by swapping segments between them and site-directed mutagenesis (Cicek et al., 2000; Verdoucq et al., 2003). A little dhurrinase activity was obtained for some mutants of ZmGlu1, while no replacement of SbDhr1 amino acids with those of ZmGlu1 at particular residues allowed hydrolysis of maize substrates. Interestingly, replacement of residues which affect the orientation of an aglycone platform residue (W378 for ZmGlu1; W376 for SbDhr1) yielded changes of substrate specificity. For example, the Y473F change in ZmGlu1, which destroyed a hydrogen bonding with W378, thereby increasing its flexibility, apparently allowed it to bind with the aglycone moiety of dhurrin. Another mutation that influenced the flexibility of that W378 was changing P377, which generated the kink at loop C, to alanine. The latter mutation yielded higher flexibility of the loop carrying W378. The mutant had decreased K_m for the artificial substrates *o*NPG, *p*NPG, and 4MUGlc, though lower k_{cat} values were also obtained, likely due to poor positioning of the glycosidic bond (Verdoucq et al., 2003).

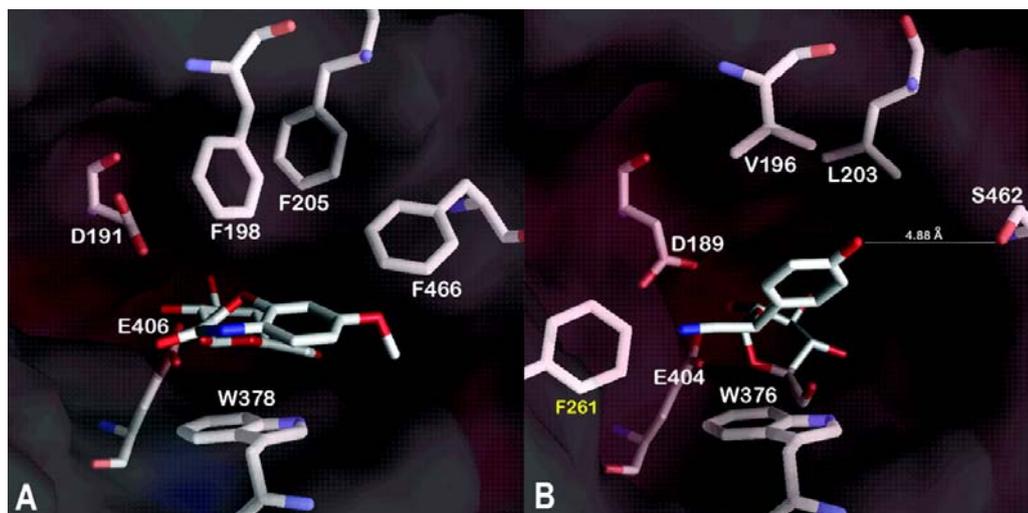


Figure 1.5 Comparison of the active sites of ZmGlu1 β -glucosidase in complex with DIMBOA-Glc and SbDhr1 dhurrinase in complex with dhurrin. (Verdoucq et al., 2004).

1.5 Rice BGlu1 β -glucosidase

In rice, β -glucosidase activity was found to be highest during germination (Palmiano and Juliano, 1973). Two rice β -glucosidases have been purified from germinating rice, each appearing to be active toward a different substrate. The first one hydrolyzed gibberellin glucoconjugates (Schliemann, 1984), while the second one hydrolyzed β -(1,3)- and β -(1,4)-linked oligosaccharides with DP 2-4 (Akiyama et al., 1998). The latter isozyme also hydrolyzed (1,2) and (1,6)- β -linked glucooligosaccharides. Recently, Opassiri et al. (2003) cloned cDNA of other rice β -glucosidase isozymes from germinating rice and characterized the expression level in germinating and mature rice tissues. One enzyme, which was highly expressed in flower and seedling shoot, was β -glucosidase isozyme 1 (BGlu1 or Os3bglu7; Opassiri

et al., 2006), while β -glucosidase isozyme 2 (BGlu2 or Os9bglu30) was expressed mainly in seedling shoot.

For rice BGlu1 protein, an N-terminal thioredoxin fusion protein was expressed in Origami (DE3) *E. coli* with pET32a (+) as a vector (Opassiri et al., 2003; 2004). The molecular weight of the fusion protein was 66 kDa on SDS-PAGE, and approximately 50 kDa after removal of the thioredoxin fusion tag by enterokinase. Fusion protein with approximately 95% purity was obtained from a single step of Ni-NTA chromatography for small scale expression and purification. Rice BGlu1 β -glucosidase exo- β -glucanase hydrolyzed various aglycones conjugated with glucose including β -(1,3)-, -(1,4)-, and -(1,6)-linked gluco-oligosaccharides, and pyridoxine-5'-*O*- β -glucoside, which is a major vitamin B₆ metabolite in rice bran.

1.5.1 Enzyme activity of rice BGlu1 and barley β -glucosidases

Rice BGlu1 shares 66% amino acid sequence identity with the barley BGQ60 β -glucosidase/ β -mannosidase and falls in the same phylogenetic cluster, but has significantly different substrate specificity. While both enzymes hydrolyze short β -1,3-linked oligosaccharides with degrees of polymerization (DP) of 2 or 3, and long (DP 4-6) cellooligosaccharides (β -1,4-linked glucose polymers) well, indicating at least 6 apparent β -1,4-linked glucosyl residue binding sites, they show distinct preferences for short cellooligosaccharides, as shown in Table 1.2 (Hrmova et al., 1998; Opassiri et al., 2003; 2004).

Table 1.2 Comparison of the kinetics parameters of rice BGlu1 and barley β -glucosidase.

Substrate DP	Rice BGlu1			Barley β II		
	K_m (mM)	k_{cat} (S-1)	k_{cat}/K_m (S-1.mM-1)	K_m (mM)	k_{cat} (S-1)	k_{cat}/K_m (S-1.mM-1)
Cello-oligosaccharides (β -1,4)						
2	31.50	1.52	0.05	2.67	11.58	4.34
3	0.72	18.13	25.40	0.97	1.95	2.01
4	0.28	17.34	61.10	0.89	8.88	9.98
5	0.24	16.90	71.50	0.41	11.66	28.44
6	0.11	16.93	152.90	0.29	11.80	40.69
Laminari-oligosaccharide (β -1,3)						
2	2.05	31.90	15.70	5.37	14.14	2.63
3	1.92	21.20	11.00	2.77	1.44	0.52
4	ND	ND	ND	0.52	0.01	0.02
5	ND	ND	ND			

Table from Opassiri et al., 2004 and Hrmova et al., 1998.
DP means degree of polymerization.

1.5.2 Subsite mapping of rice and barley β -glucosidases

Subsite affinities of rice BGlu1 and barley β -glucosidase were calculated from the Michaelis constants (K_m) and catalytic rate constants (k_{cat}) for hydrolysis of β -(1,4)-oligoglucosides of DP 2-6 (Hrmova et al., 1998; Opassiri et al., 2004). Six subsites were determined for each enzyme according to the equations described by Hiromi et al. (1973). However, there were different binding affinities at different subsites. Barley β -glucosidase prefers to hydrolyze cellobiose, with an apparent disruptive interaction at the third glucosyl residue of cellotriose in the +2 subsite, while rice BGlu1 hydrolyzes cellotriose much better than cellobiose and has high affinity for the third glucosyl residue at the +2 subsite (Figure 1.6). Thus, despite their 66% amino acid sequence identity and common hydrolysis of oligosaccharides, both

enzymes may have somewhat different substrate binding energy profiles and active-site geometries.

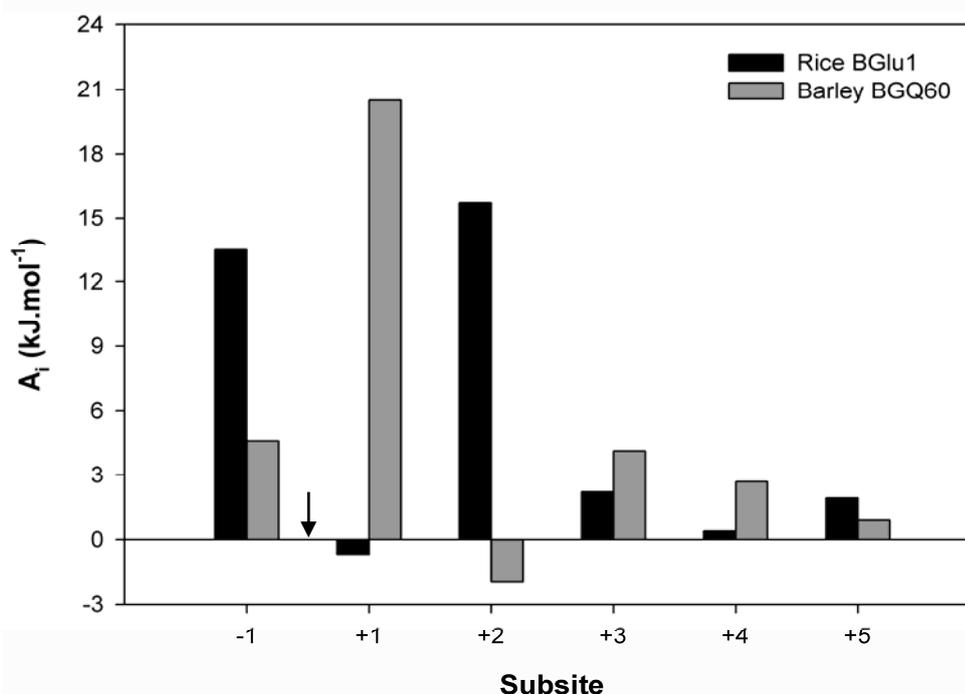


Figure 1.6 Comparison of the binding affinities of rice BGluc1 and barley BGQ60 for cello-oligosaccharide glucosyl residues.

1.6 Transglycosylation activity

Retaining glycosyl hydrolase enzymes are able to catalyse transglycosylation reactions to form new glycosidic bonds between donor and acceptor moieties, leading to synthesis of oligosaccharides, and alkyl glucosides (Makropoulou et al., 1998; Svasti et al., 2003; Opasiri et al., 2003; 2004; Hommalai et al., 2005). This occurs through a double-displacement mechanism, that is, an enzyme–substrate complex is first formed when the glycosidic linkage of the substrate is cleaved via protonation of the leaving aglycone moiety. The glycosyl-enzyme intermediate (E-Glc) is able to go

through either a hydrolysis or a transglycosylation pathway, depending on substrate concentrations, as shown in Figure 1.7 (Kempton and Withers, 1992; Hrmova and Fincher, 1998). An attack of water on the glycosyl-enzyme intermediate releases a hydrolyzed sugar, or a ternary transglycosylation complex will form if the water molecule is replaced by a competing acceptor (carbohydrates, alcohols, etc.). When a second substrate molecule acts as an acceptor, substrate transglycosylation occurs. The rate of substrate transglycosylation thus increases, while the rate of hydrolysis decreases, in proportion to the concentration of the acceptor (Figure 1.7).

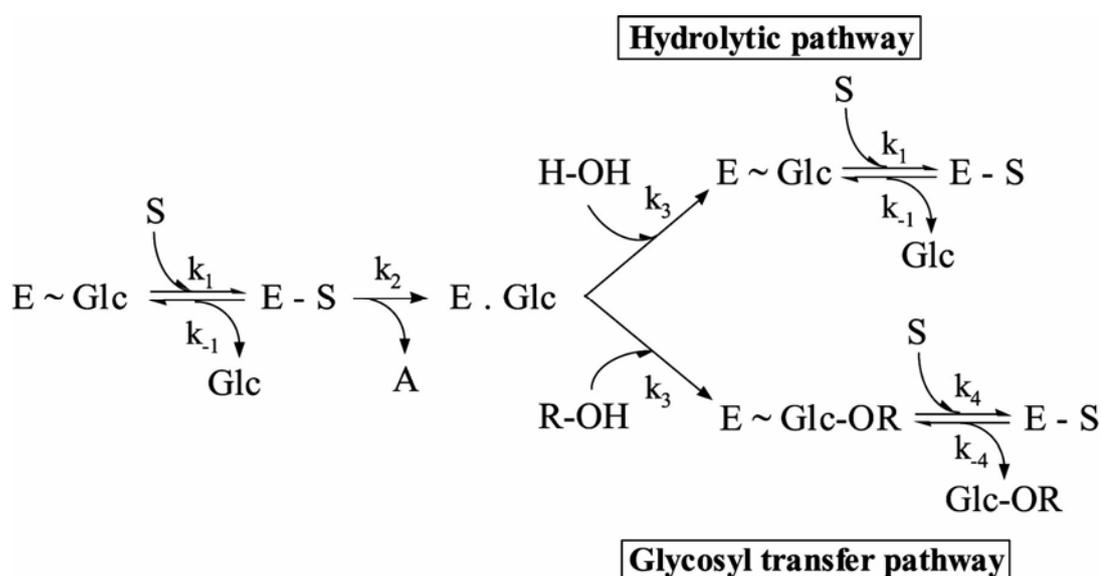


Figure 1.7 Mechanism of hydrolytic and glycosyl transfer reactions catalyzed by a plant β -glucosidase (Hrmova et al., 2002).

Rice BGlu1 β -Glucosidase can also catalyze transglycosylation reactions, such as transferring glucose between oligosaccharide molecules or to pyridoxine (vitamin B₆) to produce pyridoxine glucoside (Opassiri et al., 2003; 2004). Interestingly, a glycosyl synthase mutant of BGlu1 could synthesize long 1,4-linked oligosaccharides, but only with acceptors that occupied at least 3 subsites, such as *p*NP-cellobioside, cellotriose, and cellotetraose but not with cellobiose or *p*NP-Glc as acceptor, which is consistent with stronger binding at subsites +2 and +3 compared to +1 (Hommalai et al., 2007). Furthermore, the presence of β -glycosidic linkages and 2-hydroxyl groups in the acceptor was necessary for productive binding in the mutant BGlu1 active site for glycosynthase activity.

In order to understand the enzyme-substrate binding mechanism and the basis for its different substrate specificity compared to other family 1 glycosyl hydrolases, which may determine its biological function in rice, the three-dimensional structure of BGlu1 has been studied. As a basis for protein engineering, 3-D structural studies were initiated using X-ray crystallography. To facilitate these studies, large-scale expression and refined protein purification methods were developed for crystallization for x-ray diffraction. In addition, the putative catalytic residues were mutated to validate their assignment as the catalytic acid/base and nucleophile. Residues predicted to be involved in substrate binding based on the derived structures and computational docking were also mutated to see if substrate specificity could be changed.

1.7 Research objectives

1. To develop large-scale recombinant expression of BGlu1 protein in *E. coli* and refine protein purification methods to obtain highest purity for crystallization.
2. To determine the three-dimensional structure of rice BGlu1 and its complex with inhibitor and substrate by X-ray crystallography.
3. To clarify the basis of differential productive binding for cellobiose and cellotriose in the active sites of the rice BGlu1 and barley BGQ60 enzymes.
4. To confirm the identity of the key active site nucleophile and acid/base catalytic residues of rice BGlu1.

CHAPTER II

MATERIALS AND METHODS

2.1 Materials

2.1.1 Recombinant plasmid DNA and bacterial strains

BGlu1 was produced as an N-terminal thioredoxin/ His tag fusion protein by expression of the *bglu1* cDNA (accession no. U28047) cloned in pET32a(+) in *Escherichia coli* strain Origami (DE3) (Opassiri et al., 2003) as shown in Figure 2.1. All recombinant plasmids and bacterial cells used for this chapter are listed in Tables 2.1 and 2.2.

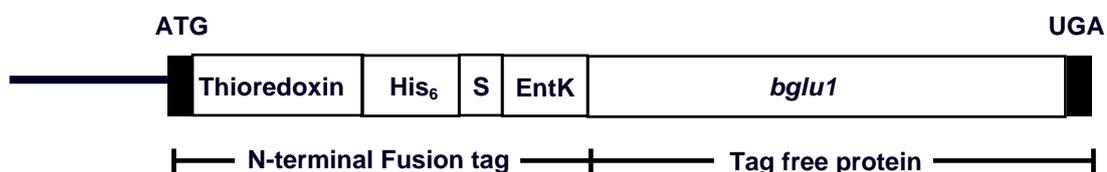


Figure 2.1 Construct of the protein-coding sequence of recombinant pET32a(+) with the *bglu1* cDNA inserted after the enterokinase cleavage site.

Table 2.1 Recombinant plasmid DNA used for this experiment

Recombinant Plasmid DNA	Antibiotic resistance	Insertion site of <i>bglu1</i>	Insertion site in vector	Total size (kb)
pBluescript- <i>bglu1</i>	Ampicillin (50 µg/ml)	Blunt end ligation	<i>EcoR</i> V	~4.5
pET32a(+)- <i>bglu1</i>	Ampicillin (50 µg/ml)	<i>SnaB</i> I, <i>EcoR</i> I	<i>EcoR</i> V, <i>EcoR</i> I	~7.5

Table 2.2 Bacterial strains

Bacterial strain	Antibiotic Resistance	Genotype	Features
DH5α	None	F ⁻ , φ80 <i>dlacZ</i> ΔM15, Δ(<i>lacZYA-argF</i>)U169, <i>deoR</i> , <i>recA1</i> , <i>endA1</i> , <i>hsdR17</i> (rk ⁻ , mk ⁺), <i>phoA</i> , <i>supE44</i> , λ ⁻ , <i>thi-1</i> , <i>gyrA96</i> , <i>relA1</i>	A high copy number strain used for DNA manipulation
XL10-Gold	Tetracyclin Chloramphenicol (<40 µg/ml)	Ter ^R Δ(<i>mcrA</i>)183 Δ(<i>mcrCB-hsdSMR-mrr</i>)173 <i>endA1 supE44 thi-1 recA1 gyrA96 relA1 lac</i> Hte [F' <i>proAB lacI</i> ^q ZΔM15 Tn10 (Ter ^R) Amy Cam ^R]	5-fold higher transformation efficiency than XL1-Blue and increases the transformation efficiency for larger DNA plasmids
Origami (DE3)	Kanamycin (30 µg/ml), Tetracyclin (12.5 µg/ml)	Δ(<i>ara-leu</i>)7697 Δ <i>lacX74</i> Δ <i>phoA PvuII phoR araD139 ahpC galE galK rpsL</i> F'[<i>lac</i> ⁺ <i>lacI</i> ^q <i>pro</i>] (DE3) <i>gor522::Tn10 trxB pLacI</i> (Cam ^R , Kan ^R , Str ^R , Ter ^R)	Enhances the formation of disulfide bonds in the cytoplasm for greater yield of active protein

2.1.2 Oligonucleotides and mutagenic primers

All mutagenic primers (Table 2.3) used for site-directed mutagenesis were ordered from Proligo Singapore Pty Ltd. (Singapore). Oligonucleotide primers used for DNA sequencing (Table 2.4) were provided by Prof. Asim Esen, Virginia Tech, VA., USA.

Table 2.3 Oligonucleotide primers used in mutagenesis

Mutation	Primer name	Sequence	Length (bp)	T_m (°C)
E176Q	T7p	5'-TAATACGACTCACTATAGGG-3'	20	48.0
	<i>Eco</i> RI_r	5'-TTCGATGAATTCTCAGTGCTT-3'	21	50.0
	E176Q_f	5'-TTTAAT <u>C</u> AGCCAAGGATAGTA-3'	21	49.0
	E176Q_r	5'-CCTTGGCT <u>G</u> ATTAAATGTAAA-3'	21	49.0
E176D	E176D_f	5'-GCACTGGTTTACATTTAATGAC <u>C</u> CAAGGATAGTAGCACT-3'	39	78.7
	E176D_r	5'-AGTGCTACTATCCTTGGG <u>T</u> CATTAATGTAAACCAGTGC-3'	39	78.7
E176A	E176A_f	5'-CACTGGTTTACATTTAATG <u>C</u> GCCAAGGATAGTAGCAC-3'	37	78.3
	E176A_r	5'-GTGCTACTATCCTTGGC <u>G</u> CATTAATGTAAACCAGTG-3'	37	78.3
E386Q	E386Q_f	5'-CCGACAGTCGTCATAACT <u>C</u> AGAACGGAATGGATCAACCT-3'	39	79.0
	E386Q_r	5'-AGGTTGATCCATTCCGTT <u>C</u> TGAGTTATGACGACTGTCCG-3'	39	79.0
E386D	E386D_f	5'-GACAGTCGTCATAACTGAC <u>A</u> ACGGAATGGATCAAC-3'	35	78.1
	E386D_r	5'-GTTGATCCATTCCGTT <u>G</u> TCAGTTATGACGACTGTC-3'	35	78.1
I179V	I179V_f	5'- TACATTTAATGAGCCAAGGGTAGTAGCACTTCTTGGTTATG -3'	41	78.6
	I179V_r	5'- CATAACCAAGAAGTGCTACTAC <u>C</u> CTTGGCTCATTAAATGTA -3'	41	78.6
N190H	N190H_f	5'- GTTATGACCAAGGAACA <u>C</u> ATCCTCCTAAAAGGTGC -3'	35	78.1
	N190H_r	5'- GCACCTTTTAGGAGGATG <u>T</u> GTTCCCTTGGTCATAAC -3'	35	78.1
N245V	N245V_f	5'- AAGTTGGAATAGTTCTGGACTTC <u>G</u> TATGGTATGAAGCTCTTCCAACCTC -3'	49	80.4
	N245V_r	5'- GAGTTGGAAAGAGCTTCATACCA <u>T</u> ACGAAGTCCAGAACTATTCCAACCTT -3'	49	80.4
L442R	L442R_f	5'- ACAACTTCGAGTGGCGGTCAGGTTACACGTC -3'	31	79.0
	L442R_r	5'- GACGTGTAACCTGAC <u>C</u> GCCACTCGAAGTTGT -3'	31	79.0

* The mutated nucleotides are underlined.

Table 2.4 Internal oligonucleotide primers used for DNA sequencing of full-length rice *bglu1*

Primer name	Direction	Sequence	Length (bp)	T _m
213R	Reverse	5'-CCAATTGACCAGGGAGTGGTAG -3'	22	56.7
40F	Forward	5'-CTTGAGAAGAAGTACGG-3'	17	44.6
163F	Forward	5'-TGGTTTACATTTAATGAGCC-3'	20	45.6
41R	Reverse	5'-CCGAGTAATTTACATTTGGT-3'	20	45.6
52R	Reverse	5'-CCTAGTTCTATGGTTGGTTA-3'	20	47.7
42F	Forward	5'-ACGGCAAACCAATTGGACC-3'	19	51.1
128R	Reverse	5'-GGTGAGCTTCAACAGGTCCTCTCTGGT-3'	27	62.8

2.2 General methods

2.2.1 Overlap extension polymerase chain reaction (PCR) method

Only the E176Q mutant was obtained by an overlap extension PCR method. The following formula was used for calculation of the primer melting temperatures (T_m):

$$T_m = 4(G+C) + 2(A+T)$$

Pfu DNA polymerase was used to amplify two overlapping PCR products using pBluescript-*bglu1* as a parental template. The primer sets of E176Q_f with *EcoRI_r* and E176Q_r with T7p were used in separate reactions. These two fragments containing the target mutation with terminal complementary sequences were purified by the QIAQuick gel purification kit (QIAGEN) by the vender's recommended protocol, and used as co-templates in the second PCR reaction. The latter reaction was amplified with the T7p and *EcoRI_r* primers to obtain the full-length mutant cDNA. Every reaction was performed using a GeneAmp PCR 9700 automated thermal cycler (AB Applied Biosystems, Singapore) according to the conditions in Table 2.5. The PCR product containing the mutation was digested with *HindIII* and *EcoRI* restriction enzymes and ligated to pBluescript cut at the same restriction sites with T4 ligase. DNA sequencing was performed to confirm the intended mutation. For protein expression, the full-length cDNA of BGLu1 containing the mutation was digested with *SnaBI* and *EcoRI* restriction enzymes and ligated to the pET32a+ vector at the *EcoRV* and *EcoRI* restriction sites.

Table 2.5 PCR conditions for the classic PCR site-directed mutagenesis method

Reaction	Primer	Template	PCR condition
Fragment1	T7p	<i>bgluI</i> -	94°C 5 min.
	E176Q_r	pBluescript	94°C 30 sec., 56°C 30 sec., 72°C 30 sec; ×32 72°C 5 min.
Fragment 2	E176Q_f	<i>bgluI</i> -	94°C 5 min.
	EcoRI_r	pBluescript	94°C 30 sec., 56°C 30 sec., 72°C 30 sec; ×32 72°C 5 min.
Full-length cDNA	T7p	Fragment 1	94°C 5 min.
	EcoRI_r	Fragment 2	94°C 30 sec., 56°C 30 sec., 72°C 30 sec; ×32 72°C 5 min.

2.2.2 Mutagenesis with the QuikChange® Site-Directed Mutagenesis Kit

To obtain high mutation efficiency and avoid several steps of DNA cloning, the QuikChange® Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) was used in place of the classical method. Nucleophile mutants (E386Q, E386D) and acid/base mutants (E176D, E176A) were constructed using the QuikChange site-directed mutagenesis kit. Briefly, the pBluescriptKS/*bgluI* or pET32a/*bgluI* vector containing the *bgluI* cDNA was used as a template for amplification a full-length plasmid strand using two overlapping oligonucleotide primers containing the desired mutation (Table 2.4). *PfuTurbo* DNA polymerase (with proofreading activity) was used to polymerize the mutated plasmid DNA during the temperature cycling (Table 2.6). The PCR products were treated with *DpnI* endonuclease to eliminate methylated and hemimethylated DNA of the parental DNA template. Repair of the nicked circular dsDNA products was performed by transformation into competent *E. coli* cells. More than 80% of transformants selected on agar plates containing ampicillin were expected

to contain the target mutation. The mutagenic oligonucleotide primers used for the kit were specifically designed using the criteria of the manual to have lengths of 25-45 bases with $T_m \geq 78^\circ\text{C}$. The T_m was calculated from following formula:

$$T_m = 81.5 + 0.41 (\%GC) - 675/N - \%mismatch$$

- N is the primer length in bases
- %GC and % mismatch are whole numbers

Table 2.6 PCR condition for the QuikChange site-directed mutagenesis method

Template	PCR condition	
pBluescript- <i>bglu1</i>	95°C 30 sec.	
	95°C 30 sec.	
	55°C 1 min.	} 12 ^a or 16 ^b cycles
	68°C 4.5 min.	
pET32a+- <i>bglu1</i>	95°C 30 sec.	
	95°C 30 sec.	
	55°C 1 min.	} 12 ^a or 16 ^b cycles
	68°C 7.5 min.	

^a a single nucleotide changes

^b a double or triple nucleotide changes at a single amino acid.

All mutations were introduced by site-directed mutagenesis, and confirmed by DNA sequencing. The results of sequencing are summarized in the Table 2.7.

Table 2.7 Details of BGlul mutations and internal primers used for sequencing the full length mutant cDNAs.

Group	Mutant	Sequencing primer	Nucleotide sequence	
			Before	After
Acid/base mutant	E176Q	213R, 52R , 163F, 42F	<u>GAG</u>	<u>CAG</u>
	E176D	52R , 40F, 42F	<u>GAG</u>	<u>GAT</u>
	E176A	213R, 52R , 163F, 42F	<u>GAG</u>	<u>GCG</u>
Nucleophile mutants	E386Q	41R, 52R, 163F , 42F	<u>GAA</u>	<u>CAG</u>
	E386D	213R, 52R, 128R , 42F	<u>GAA</u>	<u>GAC</u>
Substrate binding pocket mutant	I179V	213R, 52R , 163F, 42F	<u>ATA</u>	<u>GTA</u>
	N190H	213R, 52R , 163F, 42F	<u>AAT</u>	<u>CAT</u>
	D243V	213R, 52R, 163F , 42F	<u>GAC</u>	<u>GTC</u>
	N245V	41R, 52R , 40F, 42F	<u>AAT</u>	<u>GTA</u>
	L442R	41R, 40F, 42F	<u>CTG</u>	<u>CGG</u>

Primers in bold indicate the reaction that covered the sequence of the target mutation. Underlined nucleotides indicate that those were changed by site-directed mutagenesis.

2.2.3 Transformation of recombinant expression plasmid into expression host cells

Each expression experiment was initially started with a fresh transformation of pET32a+ containing the native or mutant *bglul* cDNA to competent cells of the expression host. The recombinant pET32a(+)-*bglul* was transformed into *E.coli* strain Origami (DE3) competent cells. 10-100 ng plasmid DNA was transformed into 70 µl of frozen competent Origami(DE3) cells by heat shock at 42°C for 90 sec. The cells were then quickly transferred to ice and incubated there for 3 min. Pre-warmed LB broth at 37°C was added to the transformed cells and they were incubated at 37°C for 45 min. Then, 70 µl of cells were spread on an LB agar plate containing 15 µg/ml

kanamycin, 12.5 µg/ml tetracycline, and 100 µg/ml ampicillin, and incubated at 37°C overnight.

2.3 Protein expression of wild type and mutant BGlu1

A 1% overnight culture of starter culture was inoculated in LB broth containing 15 µg/ml kanamycin, 12.5 µg/ml tetracycline, and 50 µg/ml ampicillin for 6-7 hr., at 37°C, until the OD₆₀₀ reached 0.6-0.8. Then, IPTG was added to a final concentration of 0.4 mM for induction, and the culture was grown for a further 18 hr. at 20°C. Pre-cooled bacterial cells were collected by centrifugation at 5,000 rpm for 15 min., and then were kept at -80°C. Frozen pellets were thawed and extracted on ice with extraction buffer (20 mM Tris-HCl pH 8.0, 200 µg/ml lysozyme, 1% Triton-X 100 mM, 1 mM PMSF, 4 µg/ml DNase I). The cell suspension was then incubated at room temperature for 30 min. Cells were completely broken using a Sonopuls Ultrasonic homogenizer with a 6 mm. diameter KE76 probe (Sigma-Aldrich, St Louis, MO, USA). Sonication was done on ice for 3 times with 30% power output, for 20 sec. and with 1 min. for cooling in between. The soluble protein was harvested by centrifugation at 12,000 g for 15 min. at 4°C.

2.4 Protein purification

2.4.1 Thioredoxin fusion protein

2.4.1.1 Mutants of the acid/base and nucleophile residues

The fusion protein was purified by immobilized metal-affinity chromatography (IMAC) on Ni-NTA superflow resin (QIAGEN) at 4°C. To bind the soluble protein, one milliliter of Ni-resin was added to thirty-five milliliters of soluble

protein extract, and then gently shaken upside down (60 rpm) for 20 min. The resin-bound protein was centrifuged at 4000 *g* for 4 min., and the supernatant was added to another one milliliter of Ni-resin to bind remaining soluble protein. The protein bound resin was equilibrated with 5 column volumes (CV) of equilibration buffer (150 mM NaCl, 20 mM Tris-HCl, pH 8.0), then washed with 10 CV of 5 mM imidazole and 5 CV of 10 mM imidazole. Finally, the bound protein was eluted with 150 mM imidazole. Fractions were assayed for β -glucosidase activity with *p*-nitrophenol β -D-glucopyranoside (*p*NPG) and those containing activity were pooled, as described by Opassiri et al. (2003). Imidazole was removed and the buffer changed to 20 mM Tris-HCl, pH 8.0, in a 30 kDa molecular weight cut-off (MWCO) Centricon centrifugal filter (Millipore, Beverly, MA, USA).

2.4.1.2 Mutants of putative aglycone binding residues

Rice BGlu1 proteins with mutations of putative aglycone binding residues were initially purified by IMAC on Co^{2+} (Talon resin, Clontech, Palo Alto, CA, USA) instead of Ni^{2+} to decrease contamination of non-specific binding proteins and protease. One milliliter of Co^{2+} resin equilibrated in phosphate buffer, pH 7.0, was applied to 5 ml soluble protein. The mixture was gently resuspended with end-over-end rotation for 20 min at 4°C, and transferred to an empty 10 ml column (BioRad) for further steps. Unbound protein was allowed to drain out by gravity flow. The column was initially equilibrated with 10 ml of phosphate buffer, pH 7.0. To eliminate nonspecifically bound proteins, 10 ml of 5 mM imidazole in phosphate buffer and then 20 ml of 10 mM imidazole in phosphate buffer were used to wash the column. Purified protein was eluted with 4 ml of 150 mM imidazole in phosphate buffer. To prevent the

proteolysis by metalloproteases, 20 μ l of 0.2 M EDTA, pH 8.0, was added to the purified protein (1 mM final concentration). The protein was concentrated to 0.5 ml using a 30 kDa MWCO centricon membrane by centrifugation at 4°C, 4000 g. Then, the new buffer, 20 mM Tris-HCl, pH 8.0, 500 mM NaCl, was added up to 20 ml, and concentrated to a final volume of 0.5 ml (2.5 μ l 0.2 M EDTA, pH 8.0, was added before keeping the enzyme at 4°C). Next, the protein was further purified by S200 gel filtration on an ÄKTA Protein Purifier (GE Healthcare, Sweden). The column was equilibrated and eluted with 20 mM Tris-HCl, pH 8.0, 500 mM NaCl with a flow rate of 0.250 ml/min. The protein sample was centrifuged at high speed for 15 min. to remove dust, then 0.5 ml was injected to the system, and 0.5 ml fractions were collected. Fractions containing β -glucosidase activity were pooled and concentrated to about 300 μ l in a 30 kDa MWCO centricon by centrifugation at 4000 g, 4°C.

2.4.2 Non-fusion protein

2.4.2.1 Wild type BGlu1

Since thioredoxin fusion tag might interfere with crystallization, only the non-fusion protein was used for crystallization. Initially, protein purification was started with IMAC on Ni-NTA as in the purification of the fusion protein. To remove the N-terminal fusion tag, BGlu1 fusion protein was cleaved with 0.006 μ g enterokinase (New England Biolabs, Beverly, MA, USA) per milligram of protein in 20 mM Tris-HCl, pH 7.2, 200 mM calcium chloride at 23°C for 16 hr. To remove the cleaved fusion tag and uncleaved fusion protein from the BGlu1 lacking fusion tag (tag-free BGlu1), the digest was incubated with 1 ml Ni-NTA Superflow resin per milligram of protein for 1 hr. at 4°C with gentle shaking. The solution of unbound

protein was collected and the resin was washed with 5 and 10 mM imidazole in equilibration buffer (150 mM NaCl, 20 mM Tris HCl, pH 8.0). The unbound and wash fractions containing tag-free BGlu1 were combined, the buffer was changed to 50 mM phosphate, pH 7.0, and the protein was concentrated using a 10 kDa MWCO Centricon centrifugal filter. The tag-free protein was loaded onto a 5 ml Hitrap SP Sepharose Fast Flow cation-exchange column (GE Healthcare, Sweden) equilibrated with five volumes of 50 mM phosphate buffer, pH 7.0. Tag-free BGlu1 was eluted with a 0–150 mM NaCl gradient in the same buffer at a flow rate of 1.5 ml/min. Fractions containing β -glucosidase activity were pooled and concentrated to 10 mg/ml in a 10 kDa MWCO Centricon centrifugal filter. The protein was then passed through a Superdex 200 gel-filtration chromatography column (HR10/30, GE Healthcare) equilibrated and eluted with 150 mM NaCl, 20 mM Tris-HCl, pH 8.0, at a flow rate of 0.25 ml/min. Later, to obtain more pure protein and decrease protein degradation by proteases, the immobilized metal affinity chromatography (IMAC) was performed on Talon Co⁺ resin (Clontech) instead of nickel resin for BGlu1 to be crystallized in complex with 2-deoxy-2-fluoro- β -D-glucoside (DNP2FG), and BGlu1 E176Q and E176D mutants to be crystallized in complex with cello-oligosaccharides.

2.4.2.2 E176Q mutant

The purification protocol for the E176Q mutant was similar to native BGlu1, except that the SP Sepharose cation-exchange step was omitted. The tag-free protein was loaded directly onto the Superdex 200 gel filtration column under the same conditions as the wild type protein.

2.5 Protein determination

2.5.1 β -Glucosidase activity assay

β -glucosidase activity was assayed in a reaction mixture containing purified enzyme, 1 mM *p*-nitrophenyl β -D-glucopyraonside (*p*NPG) dissolved in 50 mM sodium acetate, pH 5.0, and 50 mM sodium acetate buffer, pH 5.0, up to a final volume of 140 μ l. The reaction was incubated at 30°C for 10 min, and the reaction was stopped by adding 70 μ l of 0.4 M sodium carbonate. In this alkali condition the liberated *p*-nitrophenolate (*p*NP) was measured by its 405 nm absorbance on an iEMS Reader MF microplate photometer (Labsystems iEMS Reader MF, Finland). The amount of *p*NP was quantified (μ mol) by comparison to a *p*NP standard curve in the same buffer with 9 different *p*NP concentrations in the range of 0.001 to 0.1 μ mol.

2.5.2 Bio-Rad protein assay

Protein concentrations were determined according to the Bradford method (Bradford, 1976) with a Bio-Rad kit (Hercules, CA, USA) and bovine serum albumin as a standard (1-9 μ g). The reaction mixture contained suitably diluted enzyme and distilled water in a total volume of 800 μ l. To start the reaction, 200 μ l of concentrated Bio-Rad reagent was added and the tubes were vigorously mixed. The reaction was incubated at room temperature for 10 min, and the absorbance at 595 nm measured using a Genesys 10 UV spectrophotometer (Rochester, NY, USA).

2.5.3 Denaturing Gel Electrophoresis (SDS-PAGE)

SDS-PAGE was performed on 12% polyacrylamide gels mixed according to the method of Laemmli (Laemmli, 1970). Each protein sample (~5 μ g) was mixed

with 4 volumes of 4X SDS sample buffer (0.125 M Tris-HCl, pH 6.8, 8% SDS, 40% (v/v) glycerol, 0.01% bromophenol blue, and 20% (v/v) β -mercaptoethanol), and boiled for 5 min. before loading to the gel. The upper and lower reservoirs of the electrophoresis apparatus were filled with electrophoresis buffer (0.025 M Tris, 0.129 M glycine, and 0.1% SDS, pH 8.3). Electrophoresis was carried out at a constant voltage of 150 V until the tracking dye reached the bottom of gel. Protein bands were visualized by staining in staining solution containing 0.1% (w/v) Coomassie brilliant blue R-250, 40% (v/v) methanol and 10% (v/v) acetic acid for 30 min to 1 hr., followed by several washes with destaining solution containing 40% (v/v) methanol and 10% (v/v) acetic acid. The relative molecular weight (M_r) of the protein was estimated by comparison to a series of molecular weight markers (GE Healthcare, Uppsala, Sweden); phosphorylase b (97 kDa), bovine serum albumin (66 kDa), ovalbumin (45 kDa), glyceraldehyde-3-P-dehydrogenase (36 kDa), bovine carbonic anhydrase (29 kDa), bovine pancreatic trypsinogen (24 kDa), soy bean trypsin inhibitor (21 kDa), and bovine milk α -lactalbumin (14 kDa).

2.5.4 Circular dichroism (CD)

To check for secondary structure changes in mutant enzymes, the CD spectra of wild type and mutant (I179V, N190H, L442R, E176Q, E176D, E176A, E386Q, and E386D) BGlu1 were compared over the range of 260 to 190 nm using a Jasco J-715 spectropolarimeter. CD spectra were obtained with an enzyme concentration of approximately 0.5 mg/ml in 20 mM Tris-HCl, pH 8.0, at 50 nm/min, 1 nm resolution, 2 nm band width, and 1 sec. response time with correction for the buffer background.

2.5.5 N-terminal amino acid sequencing

N-terminal amino acid sequencing was performed at the Prince of Songkla University Protein Analysis Center (Hatyai, Thailand). Protein from dissolved crystals was separated by SDS-PAGE, then electro-blotted onto a polyvinylidene fluoride transfer membrane with 0.45 μm pore size (Pall Corporation, FL, USA). The blotted polypeptide bands were excised from the membrane and arranged into the sequencer cartridge and sequenced on a Applied Biosystems (Foster City, CA, USA) automated protein sequencer.

2.5.6 Dynamic light scattering

A protein solution (1 ml) with >95% purity was used for dynamic light scattering (DLS) measurement at a concentration of 1 mg/ml in 20 mM Tris-HCl, pH 8.0, using a Zetasizer nano series spectroscopy (Malvern Instruments Ltd, UK). Twenty measurements each for blank and protein sample were performed. The estimated molecular weight was calculated using an empirical mass vs. size calibration curve by DTS (Nano) software suit, provided by the supplier.

2.6 Protein crystallization

2.6.1 Preliminary sample preparation

The purified protein was concentrated using a 10 kDa MWCO Centricon centrifugal filter at 4000 *g*, adjusted to 10–20 mg/ml with 150 mM NaCl, 20 mM Tris-HCl, pH 8.0, and kept at 4°C until crystallization. Before crystallization, the protein solution was filtered through an Ultrafree-MC 0.22 μm filter (Millipore) at 4000 *g* for 5 min. to eliminate microbial contamination, dust, micro-crystals, and precipitated

protein. By this step, the nucleation rate was reduced so that larger crystals could be obtained.

2.6.2 Initial screening for crystallization conditions

Screening for crystallization conditions was done with the JBScreen HTS I and II (Jena Bioscience, Germany), Crystal Screen High Throughput HR2-130 (Hampton Research, CA, USA) kits, as well as a systematic screen of mono- and di-valent salts (0.2 M) and pH (in the range of 4.5–8.5 at 0.1 M buffer concentration) in the presence of 25% (w/v) PEG 4000. Microbatch screening was performed with 96 well plastic plates (Hampton Research) or 60 well plates (Nunc, Denmark). Ten microliters of 100% Paraffin oil (Paraffin, highly liquid, Merck) was first pipetted into each well, then 1 μ l of precipitant solution was added into the cone shaped depression well. Then, 1 μ l of 10 mg/ml protein was mixed by pipetting under an oil layer. To achieve a single drop of protein and precipitant, each well was carefully checked under a Zeiss Stemi 2000-C stereo microscope (Zeiss Corp, NJ, USA). If a single drop was not obtained, a cat whisker was used to push the separated drops together under oil. The crystallization plate was covered with the plate cover to prevent dust and debris from outside and placed on a moist sponge in a plastic box for incubation at 15°C. In these experiments, small aliquots of protein solution were taken out from 4°C and kept on ice while setting up crystallization at room temperature.

2.6.3 Optimization of crystallization conditions

Promising conditions that yielded small crystals and/or crystal clusters were systematically optimized by hanging drop vapor diffusion in a 24 well TC-plate (Hampton Research or Greiner Bio-One, Germany). Primary variables were precipitant concentration, pH, protein concentration, and ratios of protein:inhibitor and protein:precipitant, while secondary variables included addition of coprecipitant, salt, glycerol and buffer type. Twenty-two millimeter siliconized circular glass (Hampton Research) or square glass cover slides (Menzel-Glaser, Germany) were firstly siliconized using Dichlorodimethylsilane (Fluka, Switzerland) which reduces surface tension and prevents nonspecific binding. Then, high vacuum grease (Dow Corning, USA) was applied around the top edge of each well using a syringe. Five hundred microliters of appropriate precipitant was pipetted into each reservoir. Then, 1 μ l of reservoir solution was pipetted onto the center of a siliconized cover slide, and 1 μ l of protein solution (5–10 mg/ml) was then placed on the precipitant drop without mixing. The cover slide containing the crystallization drop was inverted over the reservoir and the grease seal on the top was twisted a few degrees to complete sealing. Multiple drops were placed over a single reservoir for parallel screening of several parameters for crystallization. Also, different drop sizes were screened when the ratio of protein:precipitant was optimized.

2.6.4 Microseeding by streak seeding

Crystal clusters or small crystals were removed from their initial drop, repeatedly washed with fresh mother liquor, and then crushed using cat whisker in a microcentrifuge tube containing an appropriate volume of fresh mother liquor. This

microcrystal stock was serially diluted with mother liquor and dilutions streaked with a cat whisker into pre-equilibrated hanging drops. Pre-equilibration time before seeding was further optimized with trials of 0, 1, 2, 5 hr., and overnight.

2.6.5 Cocrystallization of complexes

Complexes of native BGlu1 with 2,4-dinitrophenyl 2-deoxy-2-fluoro- β -D-glucopyranoside (DNP2FG, Sigma Chemical Corp.) as a mechanism-based inhibitor and E176Q mutant with cellopentaose (Seikagaku Kogyo Co., Tokyo, Japan) as a natural substrate were co-crystallized. Ligands were mixed at a 5:1 molar ratio with BGlu1 for initial cocrystallization trials. The protein solution was pre-incubated with ligands, then filtered through a Ultrafree-MC 0.22 μ m filter (Millipore), as mentioned above.

2.6.6 Crystal soaking

A free crystal of E176Q mutant BGlu1 was removed from the drop with a nylon loop and cleaned in a series of fresh precipitant solution in separate microbatch wells. The washed crystal was then transferred to a new drop of precipitant solution containing 1 mM laminaribiose in a sitting drop plate (Hampton Research) with the reservoir filled with precipitant solution. A non-siliconized glass cover slide was sealed on the top with high vacuum grease as in the hanging drop method.

2.6.7 Crystal viewing

Crystallization plates were observed by microscope on a Zeiss Stemi 2000-C stereoscope, which has a big and smooth platform to support the plate, at room

temperature (25°C). Normally, each crystallization plate was examined immediately after setup, once a week for one month, and then once a month and the status recorded.

2.7 Data collection and processing

2.7.1 Crystal shipping in sitting drop plates

Crystals were pipetted from the original drop to a fresh drop placed in a sitting drop well. A volume of 200 µl precipitant solution was added into the reservoir to prevent dehydration of the crystal along the trip. Each well was greased around the top and sealed with a non-siliconized glass cover slip.

2.7.2 Cryofreezing technique

All crystals were mounted in nylon loops (Hampton Research), briefly soaked in cryoprotectant solution containing 18% (v/v) glycerol in the same ingredients as used for crystallization with the concentrations of the original precipitant components increased by 15%.

2.7.3 Crystal shipping by dry shipper

A loop with suitable size was chosen and used to fish a crystal out from its original drop. The crystal was soaked in cryoprotectant for a few seconds. When freezing the crystal, cryoTongs (Hampton Research) was used to plunge the loop containing the crystal into liquid nitrogen in a dewar. The vitrified crystal mounted in the loop was capped under liquid nitrogen and transferred to the cane. All canes were carefully labeled and kept in a dry shipper containing liquid nitrogen, which was poured out before shipping.

2.7.4 In-house x-ray diffraction

Crystals in nylon loops soaked in cryoprotectant were flash-cooled in a 105 K nitrogen stream generated by an X-Stream 2000 low temperature system (Rigaku/MSK). X-ray diffraction data (wild type BGlu1 and wild type BGlu1 with inhibitor) were collected in-house at the Center of Excellence for Protein Structure and Function, Faculty of Science, Mahidol University, using an RU-H3R rotating anode X-ray generator (Rigaku/MSK) running at 50 kV and 100 mA. Diffraction from a single crystal was recorded on an R-AXIS IV⁺⁺ image plate system (Rigaku/MSK) and processed with the CrystalClear/d*TREK program suite (Pflugrath, 1999).

2.7.5 Synchrotron x-ray diffraction

Data were collected on beamline BL13B1 at the National Synchrotron Radiation Research Center (NSRRC, Hsinchu, Taiwan) with the ADSC Quantum 315 CCD detector. The wavelength was set at 0.98 Å and crystals were kept at 105K during diffraction with a nitrogen cold stream. The crystals were either flash cooled in liquid nitrogen or in the cold nitrogen cooling stream (Figure 2.2A). The mounted crystal was translated and rotated using the goniometer head to center crystal in the x-ray beam (Figures 2.2B and C). Typically, a crystal was collected for 180° rotations with 0.50° oscillations, and an exposure time of 5-30 sec. until the data were complete. The distance between the crystal and the detector was in the range of 160-300 mm, depending on the highest resolution of the data and spot overlap. Data was processed and scaled with the *HKL-2000* package (Otwinowski and Minor, 1997). The maximum resolution for each data set was determined using the ratio of measured intensity to its standard deviation ($I/\sigma(I)$), which gave >2-fold in the last shell. Moreover, the

completeness, redundancy, and R_{factor} after merging (R_{merg}) were considered for evaluating the quality of data set. The cut off for the linear R_{merg} was generally ≤ 0.4 (or 0.35) in the outer shell, depending on redundancy.

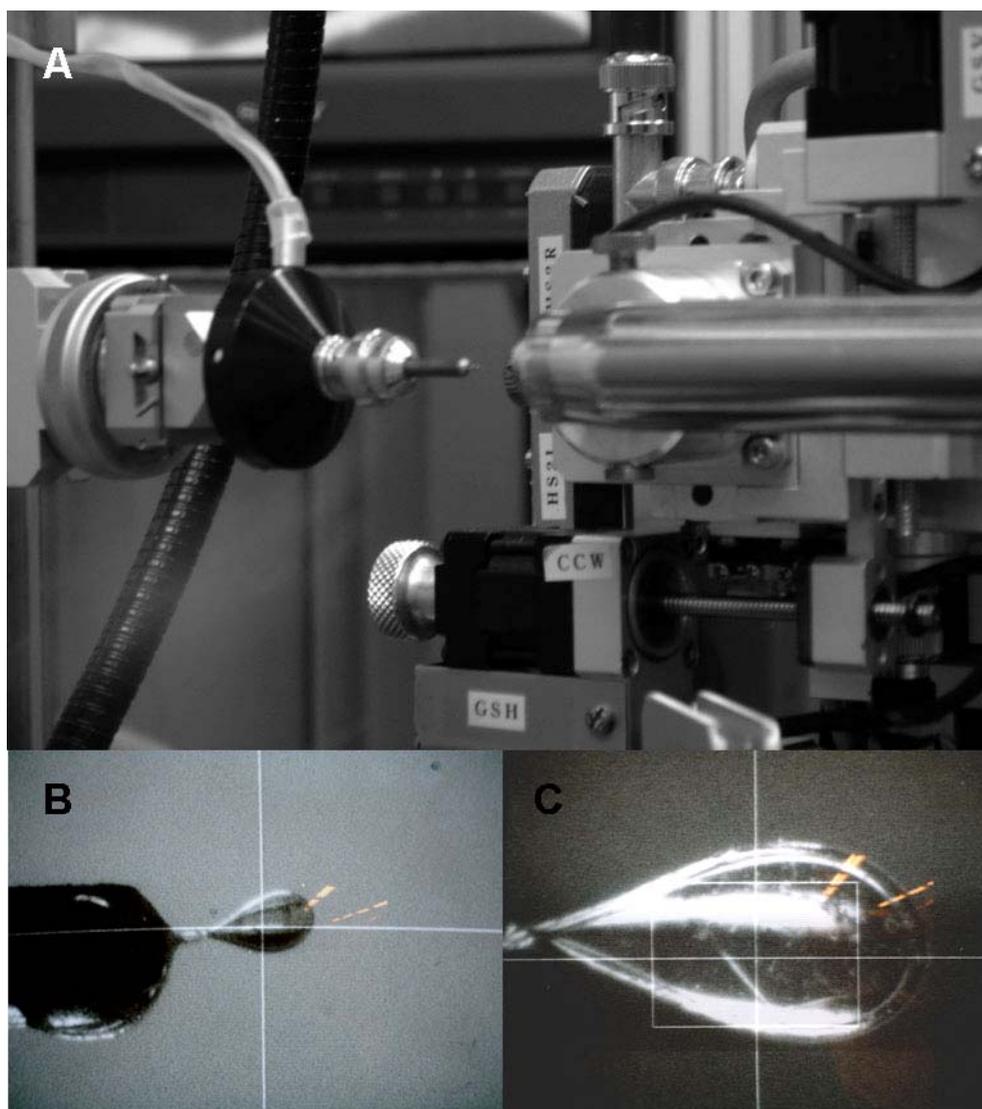


Figure 2.2 Crystal mounting for x-ray diffraction. A-B) A crystal mounted in a 0.1-0.2 mm nylon loop in a nitrogen cold stream. C) A nylon loop fixed to a metal base containing magnetic plate at the bottom to attach it quickly to the goniometer head where the crystal was flash cooled in a cold nitrogen stream.

2.8 Molecular replacement

The molecular replacement method (Navaza, 1994) was used to solve the phase problem using the phase of a known structure which is homologous to our unknown structure. In this work, the first model of free BGlu1 was solved by molecular replacement with the cyanogenic β -glucosidase from white clover (PDB code 1CBG; Barrett et al., 1995) as a search model in the AMoRe program (Navaza, 1994), based on an initial 2.75 Å resolution data set (Chuenchor et al., 2006). The rotation and translation function solutions were solved for two molecules per asymmetric unit with an initial correlation coefficient (CC) of 56.21 and R_{factor} of 45.235% in the data range of 10.00-4.00 Å.

2.9 Model building and refinement

The first model which obtained its initial phases from search model was further rebuilt in the O program. The Fourier map of $2|F_{\text{obs}}|-|F_{\text{calc}}|$ and $|F_{\text{obs}}|-|F_{\text{calc}}|$ were calculated from the observed structure factor amplitudes and the calculated structure factor amplitudes. The model rebuilding was carried out based on those 2 maps which were re-calculated for each cycle of rebuilding and refinement process. The atomic positions and B -factors of all atoms were refined to fit the observed diffraction data. The agreement is measured by R_{factor} defined as:

$$R = \frac{\sum |F_{\text{obs}}|-|F_{\text{calc}}|}{\sum |F_{\text{obs}}|}$$

A similar quality criterion is R_{free} , which is calculated from a subset (~5-10%) of reflections that were not included in the structure refinement. Both the R_{factor} and R_{free} are used to describe the quality of a model, depending on the resolution of the

data. The R_{factor} usually ranges between 0.6 (a random set of reflections with given model) and 0.2 (a well-refined model at a resolution of 2 Å). The R_{free} should be approximately the resolution in Ångströms divided by 10, thus, a data set with 2 Å resolution should yield a final R_{free} of roughly 0.2.

For free BGlu1 structure, the $2|F_{\text{obs}}|-|F_{\text{calc}}|$ and $|F_{\text{obs}}|-|F_{\text{calc}}|$ electron density of maps were first calculated at 2.8 Å resolution using the CNS program (Brunger et al., 1998). The amino acid sequence was subsequently changed from that of 1CBG to that of rice BGlu1 and manually adjusted into the electron density using O (Jones et al., 1991). The preliminary structure was refined using CNS with the 2.8 Å resolution free BGlu1 dataset. The residues V1-A14, K197-P200, R291-P293, G322-T329, and F343-K348 were deleted from the model, since no electron density was observed at main chain atoms in these loops. Other residues with poor side chain electron density were initially mutated to alanine.

The derived model, which had an R value of 39.5% and R_{free} of 41.8%, was used to solve the structure of the 2-fluoroglucoside (G2F) complex by rigid body refinement with the same R_{free} test set as the initial free model. The G2F inhibitor was built in the standard ${}^4\text{C}_1$ chair conformation and its torsion angles defined with ChemOffice (CambridgeSoft, Cambridge, MA). This covalently bound inhibitor was modeled into the $|F_{\text{obs}}|-|F_{\text{calc}}|$ electron density map at $\text{O}\epsilon 2$ of E386, the catalytic nucleophile. During refinement, water molecules positioned within hydrogen bonding distance of the model were identified from electron density peaks in the $|F_{\text{obs}}|-|F_{\text{calc}}|$ map and included in further refinement steps. After several cycles of rebuilding and refinement, the model of the G2F complex gave an R of 25.9% and R_{free} of 30.0% at 2.3 Å resolution.

The latter model was placed into the G2F complex and free enzyme datasets collected at the NSRRC synchrotron (Hsinchu, Taiwan) to 1.52 and 2.2 Å resolution, respectively, by rigid body refinement with the two molecules in the asymmetric unit refined as independent domains. The refinement was done with REFMAC5 in the CCP4 suite (Murshudov et al., 1999). Further refinement was initially done with tight noncrystallographic symmetry (NCS) restraints for both structures. These NCS restraints were maintained throughout the refinement of the free BGlu1 structure, but loose main chain and side chain NCS were applied at the final stages of the G2F complex refinement.

The final model of the G2F complex was further used as a template for subsequent structures of E176Q mutant and its complex with cellopentaose using rigid body refinement, as described above. Model rebuilding at the mutation site and some loop regions were performed and a few cycles of refinement with addition of water molecules were done with the ARP/WAP program (Lamzin et al., 1993). Cellopentaose was finally added into the structure until the electron density was shown clearly. Final refinement values are listed in Table 1. The stereochemical quality of the final models was assessed with PROCHECK (Laskowski et al., 1993).

2.10 Automated docking of cellobiose, laminaribiose, cellotriose into the BGlu1 structure

The free rice BGlu1 structure and oligosaccharide ligands were prepared for docking studies with the Sybyl 7.1 program (Tripos, Inc., St. Louis, MO). All hetero atoms and molecule B residues were removed from the structure. The N- and C-termini were fixed by adding blocking groups to the charged N-terminal amine (AMN)

and C-terminal carboxylate (CXL) at V6 and H476, respectively, and polar hydrogens were added. Charges were assigned by the Kollman uni-atom and Gasteiger method (Weiner et al., 1984), and then, the energy of the hydrogen atoms was minimized. The ligands used for docking were retrieved from the ligand-bound protein complexes of PDB database files 1FAE, 1FBW, and 2BN0, for cellobiose, cellotriose (Parsiegla et al., 2000), and laminaribiose (Meagher et al., 2005), respectively. All hydrogen atoms were added to ligands and charges were assigned by the Gasteiger-Huckel method (Gasteiger and Marsili, 1980) followed by energy minimization. The ligands were checked for accuracy and their nonrotatable and rotatable bonds were defined, then converted from pdb to mol2 format in the Autodock program. For disaccharides (cellobiose and laminaribiose), all of the bonds of the hydroxyl groups and glycosidic linkages, but not those of the rigid glucose pyranose ring were defined as rotatable bonds. A skew boat conformation was applied to the glucose residue at the nonreducing end of cellotriose with the Spartan program (Wavefunction Inc., Irvine, CA), by which the dihedral angle defined by C2, C1, O5, and C5 ring atoms was changed from -61.31° to $+45.00^\circ$, as previously reported (Biarnes et al., 2006) and 18 rotatable bonds were defined for cellotriose.

Autodock 3.0.5 was executed with the Lamarckian genetic algorithm (LGA) (Morris et al., 1998), a population size of 100, maximal mutations of 0.2 \AA in translation and 0.5 \AA in rotation and orientation, elitism set at 1, mutation rate at 0.02, and crossover rate of 0.80. Simulations were performed with a maximum of 1,500,000 energy evaluations and a maximum of 27,000 generations. Docking results were grouped by the criteria of falling within 1 \AA in root-mean-square deviation (RMSD).

2.11 Enzyme kinetics

2.11.1 Kinetic parameters analysis

All mutant enzyme activities were done concurrently with native enzyme for data comparison. All substrates and chemicals used in all reactions were prepared in 50 mM sodium acetate, pH 5.0, (apart from pH dependence and chemical rescue) and reactions were performed in triplicate at 30°C. Reactions containing buffer and 5-7 different substrate concentrations ranging from ~ 0.1 to 7 fold of K_m value (when practical) were pre-incubated at 30°C, for 10 min., and then the reaction was started by addition of enzyme solution.

The initial velocity (v_0) of hydrolysis for each substrate was initially determined using various protein concentrations and incubation times (5-60 min.), depending on the individual protein, to find conditions that yield a reasonable absorbance. The rate versus time was plotted and an appropriate time and enzyme amount were chosen for kinetic studies. For *p*NPG, 70 μ l of each substrate concentration was mixed with 70 μ l of enzyme (1.5 nmol) in microtiter plate wells. The reactions were incubated at 30°C for 20 min., then stopped by adding 70 μ l of 0.4 M NaCO₃. The 405 nm absorbance was measured and the liberated *p*NP quantified by comparison to the *p*-nitrophenolate (*p*NP) standard.

Glucose released from oligosaccharide substrates was assayed with a coupled peroxidase-glucose oxidase reaction. Twenty-five microliters each of substrate (1.5 nmol of cellotriose, or 30 nmol of cellobiose) and protein were mixed well in 1.5 ml tubes and incubated at 30°C for 20 min. The reaction was stopped by heating at 80°C for 5 min., then transferred to a microtiter plate for measurement of the glucose. One hundred microliters of peroxidase-glucose-oxidase enzyme and 50 μ l of ABTS were

added to each microtiter plate well and the plate was incubated at 37°C for 30 min. The 405 nm absorbance was measured and compared to a glucose standard curve to quantify the amount of glucose released by the β -glucosidase reaction.

A suitable blank containing buffer and substrate were set up at every substrate concentration to correct for spontaneous hydrolysis. The 405 nm absorbance was measured using an iEMS Reader MF microtiterplate photometer (Labsystems iEMS Reader MF, Finland) and quantified by comparison to standards. The kinetic parameters were calculated by nonlinear regression of Michaelis-Menten plots with the GraFit 5.0 program (Leatherbarrow et al., 2001). Linear regression of Lineweaver-Burk plots was used for some mutants that had very high K_m , for which a saturation curve could not be obtained due to limitations of substrate solubility (>200 mM).

2.11.2 pH profile of activity

The pH dependence of the enzyme activity was determined using *p*-nitrophenol β -D-glucopyranoside (*p*NPG) as a substrate. Two buffer systems of overlapping pH buffers (acetate pH 4.0-5.5, MES pH 5.5-6.6, phosphate pH 6.5-7.5, Tris-HCl pH 7.5-9.0) and universal pH buffer (solution A: 0.2 M boric acid + 0.05 M citric acid, solution B: 0.05 M tri-sodium phosphate with pH ranging from 2-12) were used for comparison.

2.11.3 Effect of anionic nucleophile on rescue activity of acid/base mutants

Various anionic nucleophiles such as acetate, azide, ascorbate, formate, citrate, trifluoroacetate (TFA), fluoride (KF), cyanate and glycine were tested for their effects on the enzyme activity of wild type and E176Q, E176A, and E176D mutant BGlu1.

Wild type BGlu1 and all mutants were assayed in 50 mM of anionic nucleophile in 50 mM MES buffer, pH 5.0, and 1 mM *p*NPG as a substrate at 30°C for 10 min.

2.11.4 Nucleophile concentration dependence of the effects on the activity of wild type and E176Q mutant BGlu1

The dependence of enzyme activity on nucleophile concentration was determined for azide, ascorbate and acetate at a single saturated substrate concentration of *p*NPG (20-100 times the K_m in the presence of the nucleophile). The enzyme activity for *p*NPG hydrolysis at saturating concentration (10 mM) was determined by using a range of 5- 100 mM nucleophile prepared in 50 mM MES buffer, pH 5.0.

2.11.5 Kinetics of activity rescued by external anionic nucleophiles

Enzyme kinetics for hydrolysis of *p*NPG and 2,4-dinitrophenyl β -D-glucopyranoside (2,4-DNP) (provided by Dr. Stephen Withers, Vancouver, Canada) was determined as described above, but in the presence of azide and acetate in the reaction mixture and a continuous assay was applied to the latter substrate. For *p*NPG, kinetics of the E176Q mutant enzyme in the presence of acetate (0.05 and 0.32 mM) and azide (0.05, 0.32, 1.0 M) in 50 mM MES buffer, pH 5.0, were evaluated at 30°C. Six *p*NPG concentrations, which ranged from 0.5 to 7 fold of K_m were used, with a fixed concentration of the anionic nucleophile.

For 2,4-DNPG, the kinetics of wild type and E176Q mutant with and without external anionic nucleophiles, including 200 mM azide and 200 mM acetate, were measured in 50 mM MES buffer, pH 5.0, at 30°C. Reactions, which contained buffer

and seven different substrate concentrations, were pre-incubated at 30°C, for 10 min., and initiated by adding enzyme solution. The plate was further incubated at 30°C and the release of 2,4-dinitrophenol (DNP) was continuously measured at 405 nm absorbance and quantified with a standard curve.

2.11.6 TLC analysis

The products from the reaction of the E176Q mutant with *p*NPG in the presence of various anion nucleophiles were separated by TLC, and compared to wild type reactions. Reaction mixtures containing 5.5 µg enzyme, 5 mM *p*NPG, and 0.05 M anionic nucleophile in 50 mM MES buffer, pH 5.0, were incubated at 30°C, for 1 hr. The reaction was terminated with boiling for 5 min. and centrifugation at 13,000 rpm for 5 min. The supernatants were then spotted on a TLC aluminum plate coated with silica gel 60 F₂₅₄ (Merck, Germany) with glucose as a standard. The TLC plate was developed with 1:1:2 (v/v/v) acetic acid: water: ethyl acetate as a running solvent, and the products detected by applying 10% (v/v) sulfuric acid in ethyl alcohol and heating at 120°C for 5 min. or until the products were visible.

CHAPTER III

RESULTS

3.1 Protein purification

3.1.1 Thioredoxin fusion protein

Wild type BGlu1 and individual mutants were expressed as a N-terminal thioredoxin and hexahistidine tag fusion protein in *E. coli* Origami (DE3) with similar amounts of total bacterial protein and soluble protein, except for the D243V mutant, which was expressed as an inactive protein and was mainly aggregated in the inclusion body fraction (Figure 3.1). This residue is conserved in rice BGlu1, barley β -glucosidase isozyme II: BGQ60, and ZmGlu1 maize β -glucosidase, and when the charged polar amino acid (aspartate) was replaced with the nonpolar amino acid (valine) found in linamarase, it could not be expressed as an active protein at all (Figure 3.1).

The presence of the histidine tag facilitates the protein purification by one step of IMAC on a Ni^{2+} . To minimize nonspecific binding and reduce the amount of contaminating protein, washing by a step gradient of imidazole was used for protein elution, including 5, 10, 20, 50, 100, 150, 250 mM imidazole in Tris-HCl, pH 8.0. Most bacterial host proteins came off in the unbound fraction and 5 mM imidazole wash. Histidine-tag-containing protein was gradually washed out with 10 to 100 mM imidazole. No protein was observed on SDS-PAGE in the 150 and 250 mM imidazole fractions. This result showed that our protein was not strongly bound to the resin

since the tagged protein could be washed out with a relatively low imidazole concentration. Thus, washing with 5 and 10 mM imidazole concentrations followed by elution with 150 mM imidazole was used in subsequent experiments.

Approximately 80-90% pure protein was obtained by IMAC on Ni^{2+} , which yielded 9–17 mg protein per liter of *E. coli* culture. The major band of fusion protein appeared at 66 kDa, with other prominent protein bands at approximately 50, 21 and 15 kDa on SDS-PAGE. The buffer was changed to 20 mM Tris-HCl, 150 mM NaCl, pH 8.0, to eliminate imidazole since it inhibits rice BGlu1 activity, as previously reported (Opassiri et al., 2003). Generally, approximately 85-90% pure protein was homogeneous enough for enzyme characterization and we used a single purification step with the E176Q, E176D, E176A, E386Q, and E386D mutant proteins (Figure 3.1).

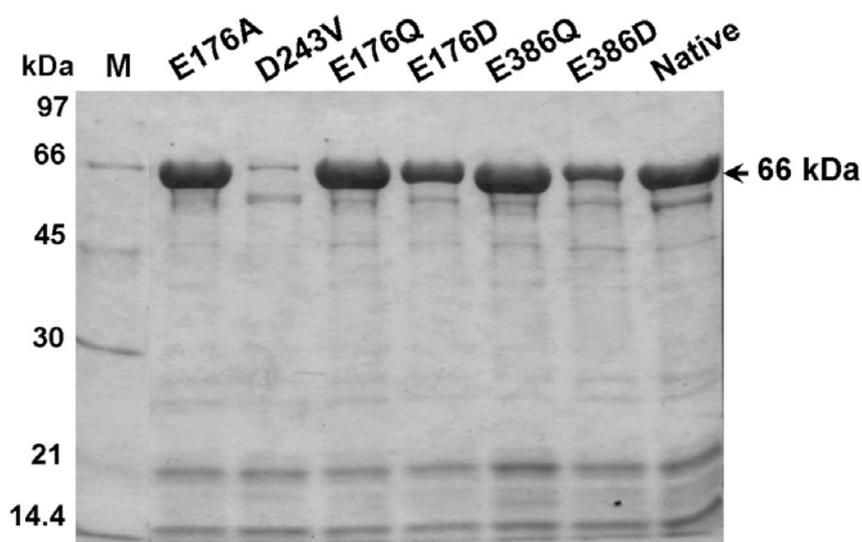


Figure 3.1 SDS-PAGE of thioredoxin fusion proteins of wild type rice BGlu1 and E176A, D243V, E176Q, E176D, E386Q, and E386D BGlu1 mutant proteins purified with a single step of Ni^{2+} -NTA IMAC.

It appears that the protein bands at approximately 50 kDa on SDS-PAGE came from the proteolysis of BGlu1 by a protease in the enzyme solution. This protein appeared as a faint band when the protein was first purified. Protein degradation was increased when purified protein was kept at 4°C for a few days. To decrease the proteolysis, several strategies had been tried, such as bacterial host cell change to OrigamiB (DE3), addition of sorbitol and betaine in LB media, or addition of a protease inhibitor cocktail to the extraction buffer. The proteolysis was not significantly reduced compared to the control as indicated by SDS-PAGE.

Using of Co²⁺ metal ion instead of the Ni²⁺-based resin was the most successful method to reduce the contamination of protease. The Co²⁺ metal ion binds to his-tagged proteins with enhanced selectivity, therefore, nonspecific binding was minimized. Furthermore, the pH of the buffer affects the strength of binding between the protein and metal ion, so purification with more stringent conditions using slightly lower pH was desirable. The purification using Talon Co²⁺ resin with pH 7.0 phosphate buffer (Tris-HCl, pH 8.0 was used for Ni²⁺-NTA resin) eliminated non-specifically bound proteins as shown in SDS-PAGE (Figure 3.2). No major protein bands appeared at the molecular weight range of 20 to 50 kDa, as were previously obtained with Ni²⁺ IMAC, only faint bands at 50, 21, 15 kDa were observed. Recently, Hommalai et al. (2007) reported that addition of soybean trypsin inhibitor reduced proteolysis of BGlu1 mutant enzymes. Therefore, the inhibitor was included in extraction buffer for the latest protein purifications.

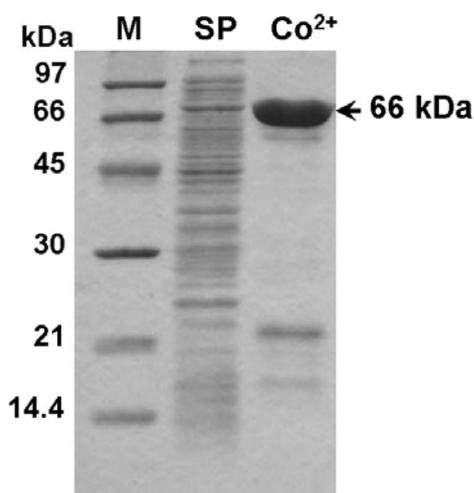


Figure 3.2 Purification of thioredoxin fusion protein of wild type BGlu1 using Co^{2+} metal ion instead of the Ni^{2+} -based resin.

A further step of purification by gel filtration was carried out for the proteins with mutations at putative aglycone binding pocket residues in order to eliminate the fusion tag seen at 21 and 15 kDa, which co-purified with our protein from IMAC. The gel filtration chromatogram showed several broad peaks with the major peak observed in fraction no. 30. All fractions from each peak were examined on SDS-PAGE (Figure 3.3), which showed that two small overlapping peaks, that came out before the major peak containing our protein. This suggested that dimers or larger complexes might be formed in solution, though the main peak was monomeric. Most of our protein was monomer and came out in fractions 30 to 37 with some overlap with the fusion tag peak (in fractions no. 33 to 40). Consequently, only pure fractions from no 30 to 32 were pooled and concentrated for further experiments.

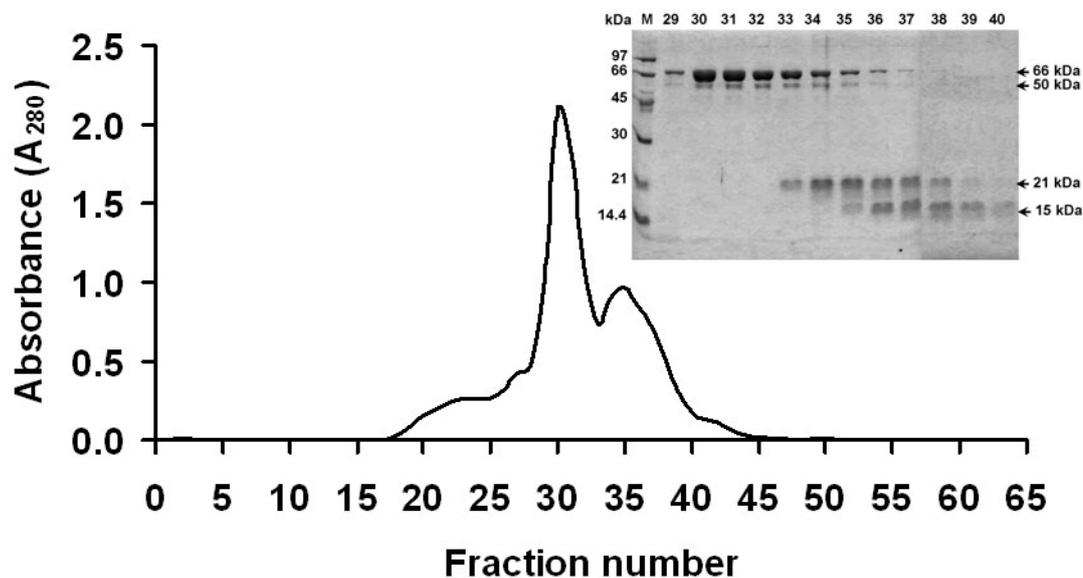


Figure 3.3 Elution profile of thioredoxin fusion protein of wild type rice BGlu1 obtained from S200 gel filtration chromatography and SDS-PAGE profile.

After S200 gel filtration, thioredoxin fusion protein was obtained with > 90% purity, with a yield of about 4-6 mg protein/l cell culture (Figure 3.4). A contaminating protein, likely to be a degradation product, appeared as a faint band on SDS-PAGE at approximately 50 kDa. This two step purification was applied for mutants of putative aglycone binding residues (I179V, N190H, N245V, and L442R) used to study the determinants of substrate specificity in rice BGlu1 and barley BGQ60 β -glucosidases.

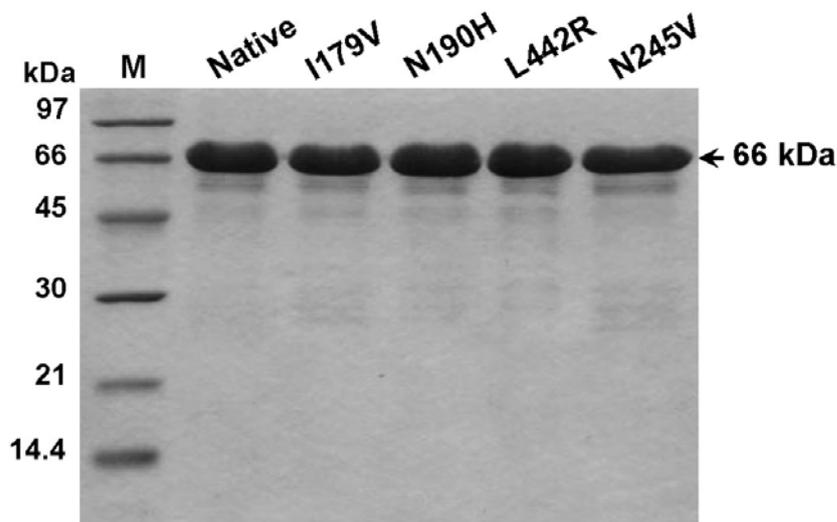


Figure 3.4 Rice BGLu1 aglycone-binding residue mutant purification. SDS-PAGE profiles of wild type and I179V, N190H, L442R, and N245V mutants of BGLu1 after S200 gel filtration column.

3.1.2 Nonfusion protein of wild type BGLu1

A large amount of protein that is as pure as possible is needed for crystallization trials, therefore the purification methods used for kinetic studies were further developed to obtain high quality and quantity. The fusion tag at the N-terminus was removed, since it might interfere with crystallization. Overnight digestion with enterokinase to remove the thioredoxin fusion tag was followed by adsorption of the fusion tag to Ni^{2+} -bound resin. Different fractions of purified proteins obtained from the second IMAC with different amounts of contamination with the free tag protein were separately applied onto an S200 gel filtration column. The protein fractions with high purity passed through gel filtration as a single peak on the chromatogram (in fractions no. 31 to 38, Figure 3.5A) and a single band of protein was detected at approximately 50 kDa on SDS-PAGE. Conversely, protein with less purity yielded free tag proteins that eluted simultaneously with nonfusion

protein in fractions 34 to 39 (Figure 3.5B). This suggested that we could not use gel filtration to separate free tag proteins apart from nonfusion protein, as was the strategy for the thioredoxin fusion protein, though better separation might be obtained if we used a longer column.

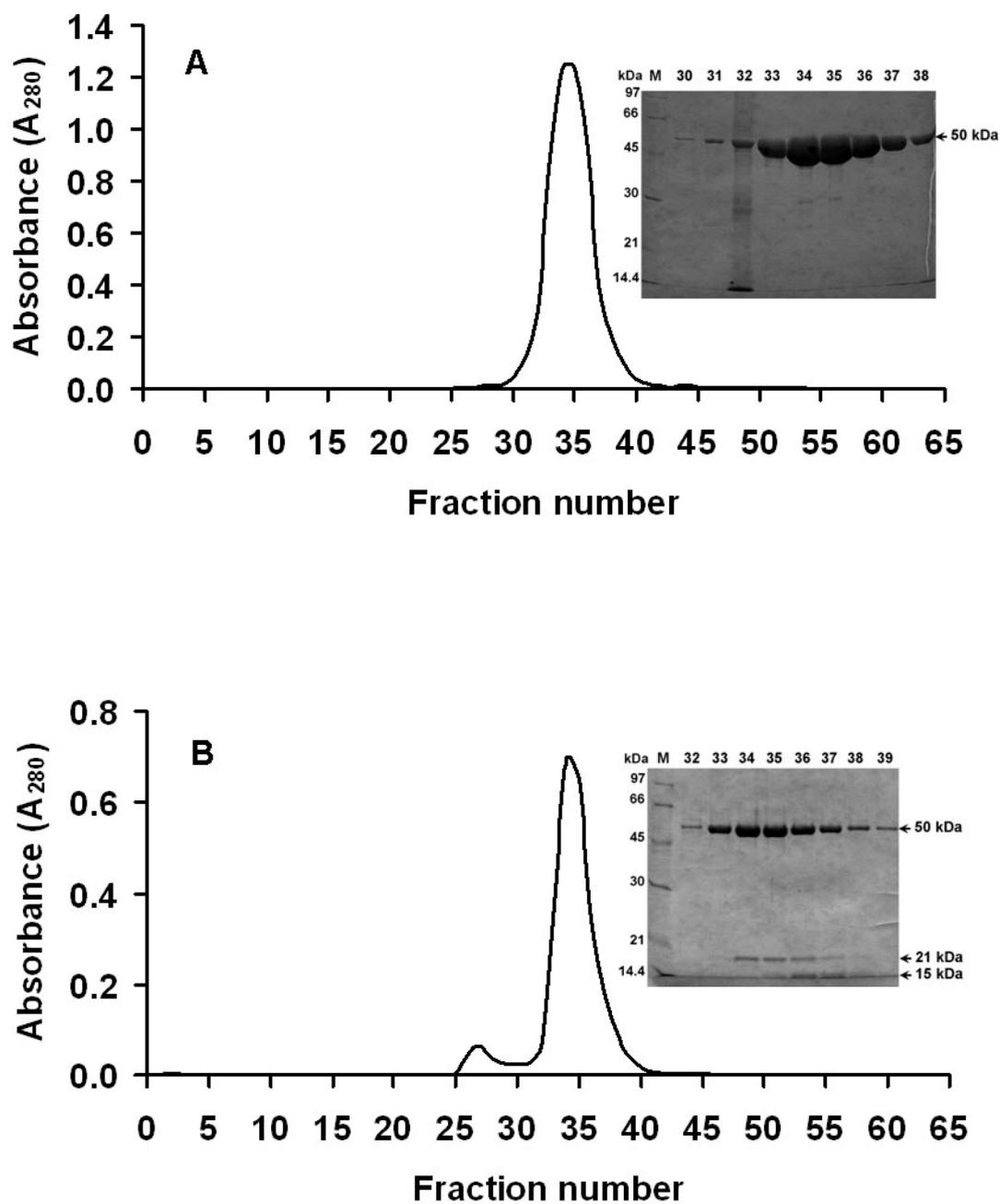


Figure 3.5 Elution profile of the wild type BGlul1 nonfusion protein passed through an S200 gel filtration column and SDS-PAGE of fractions at the protein peak. A) Highly pure protein injected onto the column. B) Less pure protein injected onto the column.

In later experiments, the protein was further purified by SP sepharose before gel filtration chromatography to eliminate free tag protein, which came off in the unbound fraction (0% NaCl). Some faint bands which appeared larger than the nonfusion protein (50 kDa) were detected in the unbound fraction of SP sepharose, as well, as shown in fractions no. 12-14 of Figure 3.6. Moreover, another contaminating protein could be eliminated, since a small peak was detected at 100% buffer B (150 mM NaCl in Tris-HCl, pH 8.0), while our target protein was eluted at 5-10% buffer B (in fractions no. 38-42).

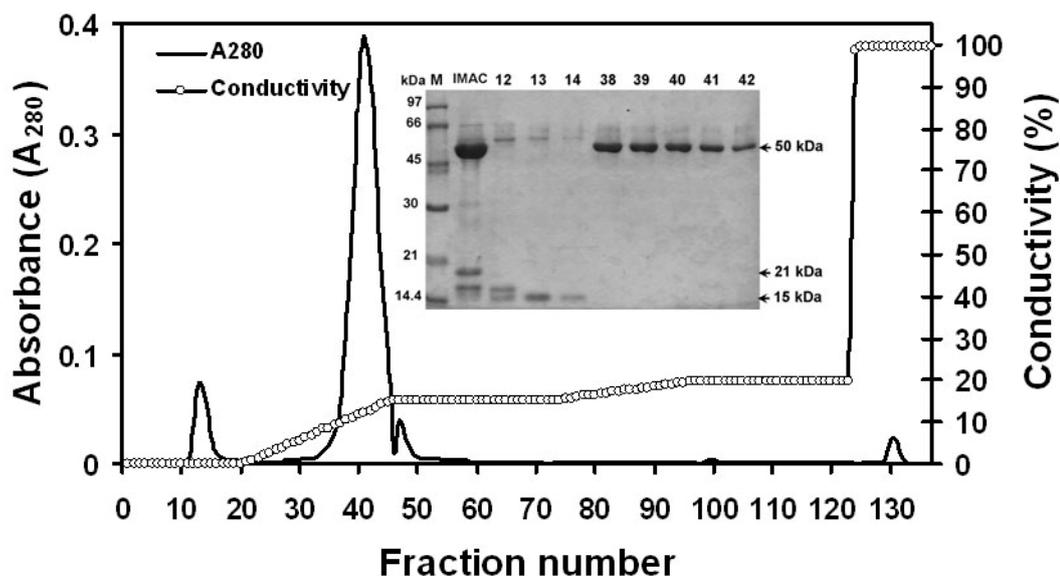


Figure 3.6 Elution profile of nonfusion protein for BGlu1 obtained from SP sepharose cation exchange chromatography with a two-step salt gradient (0-10% and 10-20%). Two distinct peaks were washed out at 0%, and 5-10% solution B (150 mM NaCl in 20 mM Tris-HCl, pH 8.0). The SDS-PAGE profile shows protein content in the first two peaks.

Homogeneous protein was obtained after the final purification step of S200 gel filtration, with >95% purity, with a yield of about 2.6 mg protein/l cell culture, as shown in Figure 3.7. The protein at the steps of the purification of nonfusion protein is compared in series in Figure 3.8.

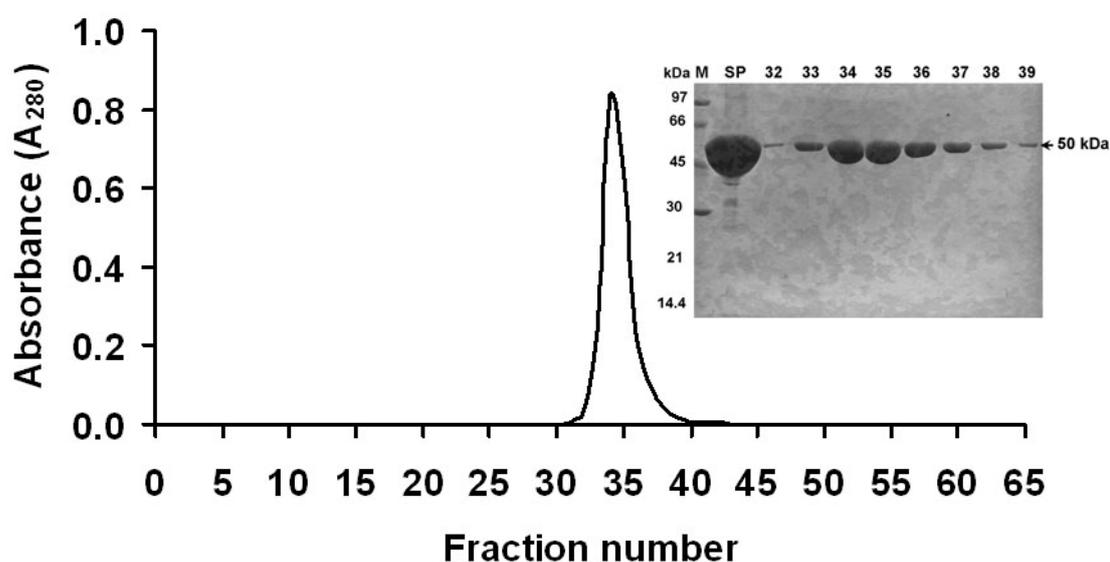


Figure 3.7 Elution profile of the final purification step of nonfusion protein for BGlu1 enzyme and SDS-PAGE of protein fractions obtained from a single peak on gel filtration chromatography.

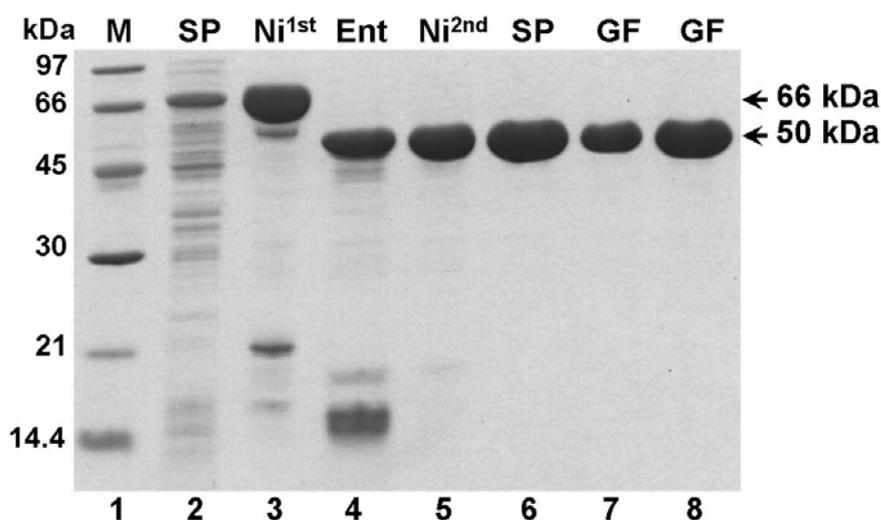


Figure 3.8 SDS-PAGE of purified BGLu1 fractions from the steps of purification of BGLu1 for crystallization. Lane 1 is Bio-Rad Low M_r markers; lane 2, soluble extract of *E. coli* cells; lane 3, fusion protein after initial IMAC, lane 4, enterokinase digest; lane 5, BGLu1 after subtractive IMAC; lane 6, BGLu1 after SP-sepharose, lanes 7 and 8, BGLu1 after the final S200. Five micrograms of protein is loaded in all lanes, except lane 7, which has 2.5 μg .

3.1.3 Nonfusion protein of the E176Q mutant

The purification for the E176Q mutant was slightly different from wild type BGLu1. Inconsistent enterokinase cleavage of this mutant was found, so the nonfusion protein could not be purified by SP cation exchange chromatography. The nonfusion protein came out in the unbound fraction, while wild type BGLu1 was normally eluted at 5-10% of 150 mM NaCl. Thus, the step of SP cation exchange chromatography was skipped for purification of the E176Q mutant protein, and protein from the 2nd IMAC was directly subjected to S200 gel filtration (Figure 3.9). However, the purity and homogeneity of protein was still acceptable and it was used for crystallization.

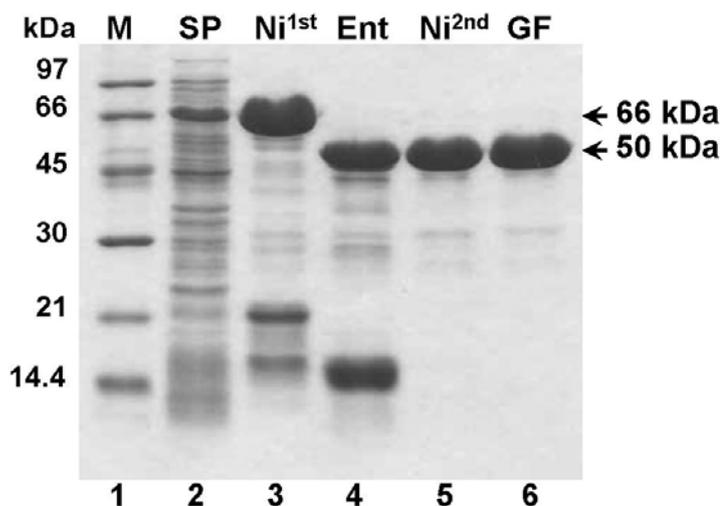


Figure 3.9 SDS-PAGE analysis of the E176Q mutant from the steps of the purification for crystallization. Lane 1 is Bio-Rad Low M_r markers; lane 2, soluble extract of *E. coli* cells; lane 3, fusion protein after initial Ni^{2+} IMAC, lane 4, enterokinase digest; lane 5, BGlu1 after subtractive Ni^{2+} IMAC; lane 6, BGlu1 after the final S200. Five micrograms of protein was loaded in each lane.

3.2 Circular dichroism

The CD spectra of 0.5 mg/ml protein solutions of the wild type and mutant BGlu1 enzymes were compared to check for secondary structure changes. All mutants yielded CD spectra that were similar to that of wild type BGlu1 (Figure 3.10), suggesting that the mutations has no significant effect on the conformation of the protein. CD was not performed with N245V, but it could be expressed as an active soluble protein. Therefore, it is not likely that any significant structural change occurs with this mutant.

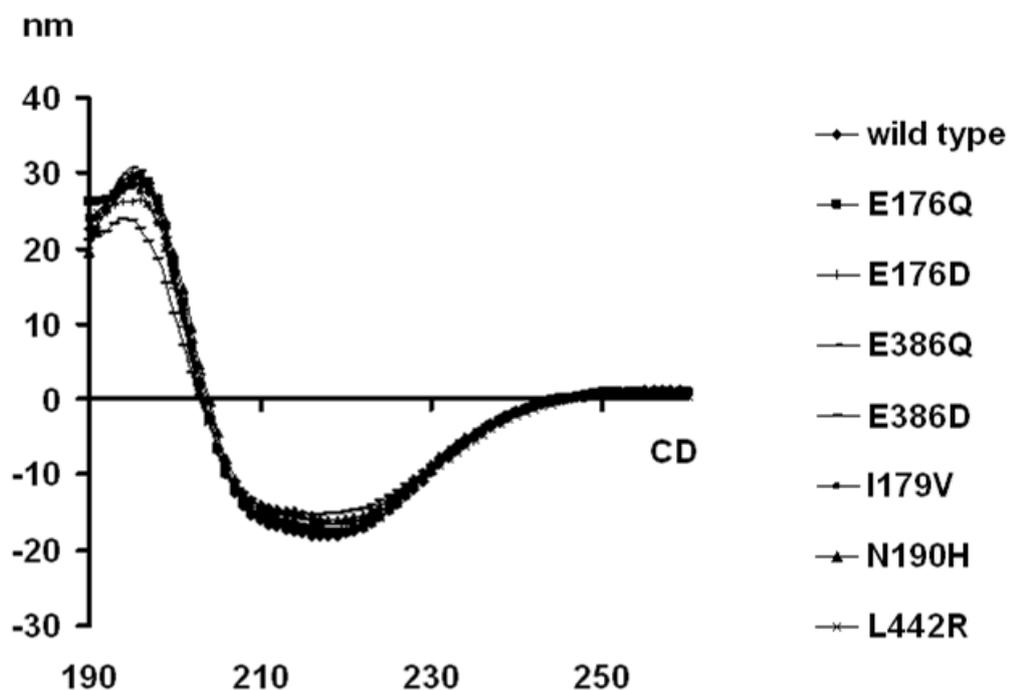


Figure 3.10 CD spectra in the peptide bond region (190-260 nm) of wild type and all mutants (except for N245V). The protein concentration was 0.5 mg/ml for all eight samples and each was purified by single step of Ni-NTA IMAC, followed by changing the buffer to 20 mM Tris-HCl, pH 8.0.

3.3 Dynamic light scattering (DSL)

DLS was used to estimate the molecular weight of the fusion protein in solution (20 mM Tris-HCl, pH 8.0), and gave a single peak (100% of scattering) at 7.115 nm diameter, which corresponds to an estimated molecular weight of 65.6 kDa (Figure 3.11). This result confirmed that BGlul forms a functional monomeric enzyme in solution.

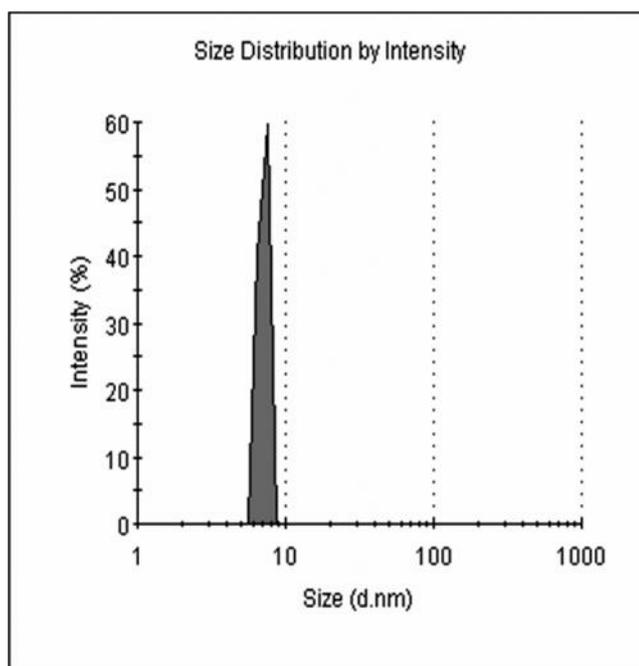


Figure 3.11 Dynamic Light Scattering (DLS) spectrum of rice BGlu1 fusion protein. DLS was used to determine the mean particle size and size distribution in a buffer solution (20 mM Tris-HCl, pH 8.0). The size distribution was plotted as intensity (%) versus size as the diameter of particle (d.nm). The protein gave a single peak with a mean diameter of 7.115 nm.

3.4 Protein crystallization

3.4.1 Wild type BGlu1 without inhibitor

3.4.1.1 Initial screening

Two screening kits from Hampton Research, HR2-130 and HR2-134, were used for the first crystallization trails. Ten milligrams per milliliter of protein solution was used for initial screening with a protein:precipitant ratio of 1:1. Small tetragonal crystals with various sizes appeared after 3–5 days under several conditions, as shown in Figures 3.12 and 3.13. The most successful precipitant was

PEG ranging in molecular weight from 3350 to 8000. Di-valent salts, like $(\text{NH}_4)_2\text{SO}_4$, and Li_2SO_4 , and mono-valent salts, like magnesium acetate (MgOAc), zinc acetate (ZnOAc), NaCl , and MgCl_2 were all successful for crystallization of BGlu1. The pH between 6.5-8.5 with Na cacodylate, MES, HEPES, and Tris-HCl as buffers was suitable to crystallize this protein (Table 3.1). After 2 months, larger crystals developed under some conditions and preliminary x-ray data collection was performed at an in-house x-ray diffraction system at Mahidol University.

Table 3.1 Crystallization conditions of free wild type BGlu1 from screening kits that yielded small crystals in the microbatch method.

Condition no.	Precipitant	Salt	Buffer
B6 (HR2-130)	20% (w/v) PEG 8000	0.2 M MgOAc	0.1 M Na cacodylate, pH 6.5
D9 (HR2-130)	18% (w/v) PEG 8000	0.2 M ZnOAc	0.1 M Na cacodylate, pH 6.5
G2 (HR2-130)	30% (w/v) PEG MME 5000	0.2 M (NH ₄) ₂ SO ₄	0.1 M MES, pH 6.5
G6 (HR2-130)	10% (w/v) PEG 6000, 5% (v/v) MPD	None	0.1 M HEPES, pH 7.5
F8 (HR2-134)	25% (w/v) PEG 3350	0.2 M (NH ₄) ₂ SO ₄	0.1 M HEPES, pH 7.5
F9 (HR2-134)	25% (w/v) PEG 3350	0.2 M (NH ₄) ₂ SO ₄	0.1 M Tris-HCl, pH 8.5
F12 (HR2-134)	25% (w/v) PEG 3350	0.2 M NaCl	0.1 M HEPES, pH 7.5
G4 (HR2-134)	25% (w/v) PEG 3350	0.2 M Li ₂ SO ₄	0.1 M HEPES, pH 7.5
G5 (HR2-134)	25% (w/v) PEG 3350	0.2 M Li ₂ SO ₄	0.1 M Tris-HCl, pH 8.5
G12 (HR2-134)	25% (w/v) PEG 3350	0.2 M MgCl	0.1 M HEPES, pH 7.5

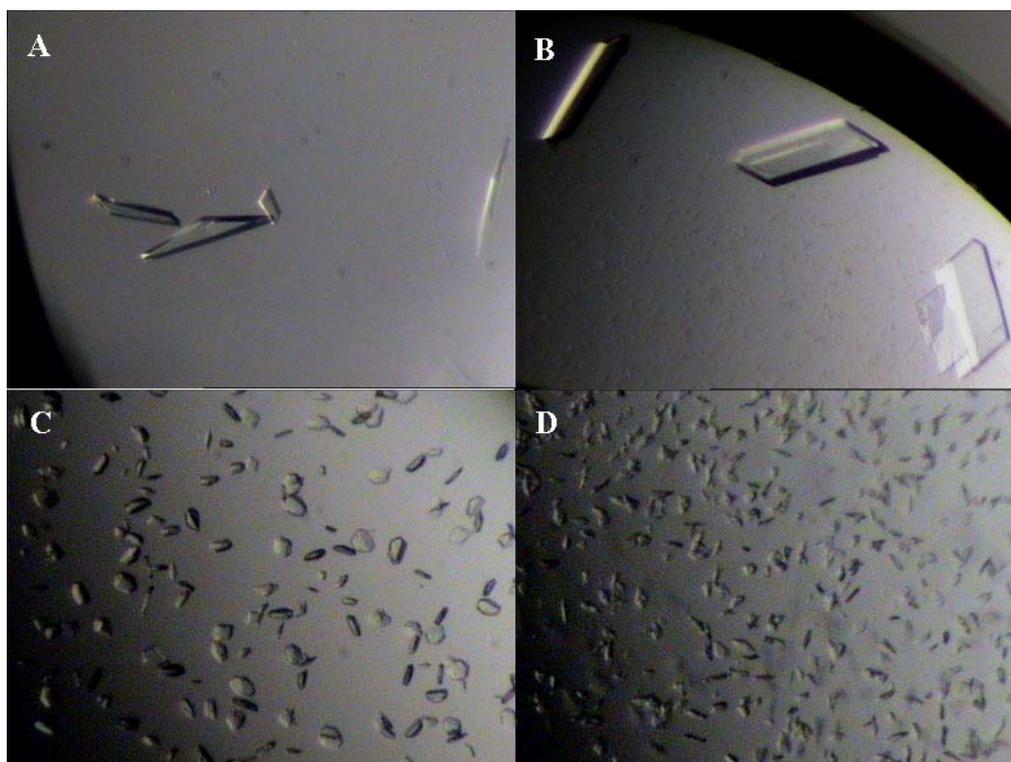


Figure 3.12 Crystals of free wild type BGlu1 in microbatch plates from the initial screening with the Hampton HR2-130 screening kit. Conditions are A: 20% (w/v) PEG 8000, 0.2 M MgOAc, 0.1 M Na cacodylate, pH 6.5 (HR2-130 B6); B: 18% (w/v) PEG 8000, 0.2 M ZnOAc, 0.1 M Na cacodylate, pH 6.5 (HR2-130 D9); C: 30% (w/v) PEG MME 5000, 0.2 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES, pH 6.5 (HR2-130 G2); D: 10% (w/v) PEG 6000, 5% (v/v) MPD, 0.1 M HEPES, pH 7.5 (HR2-130 G6).

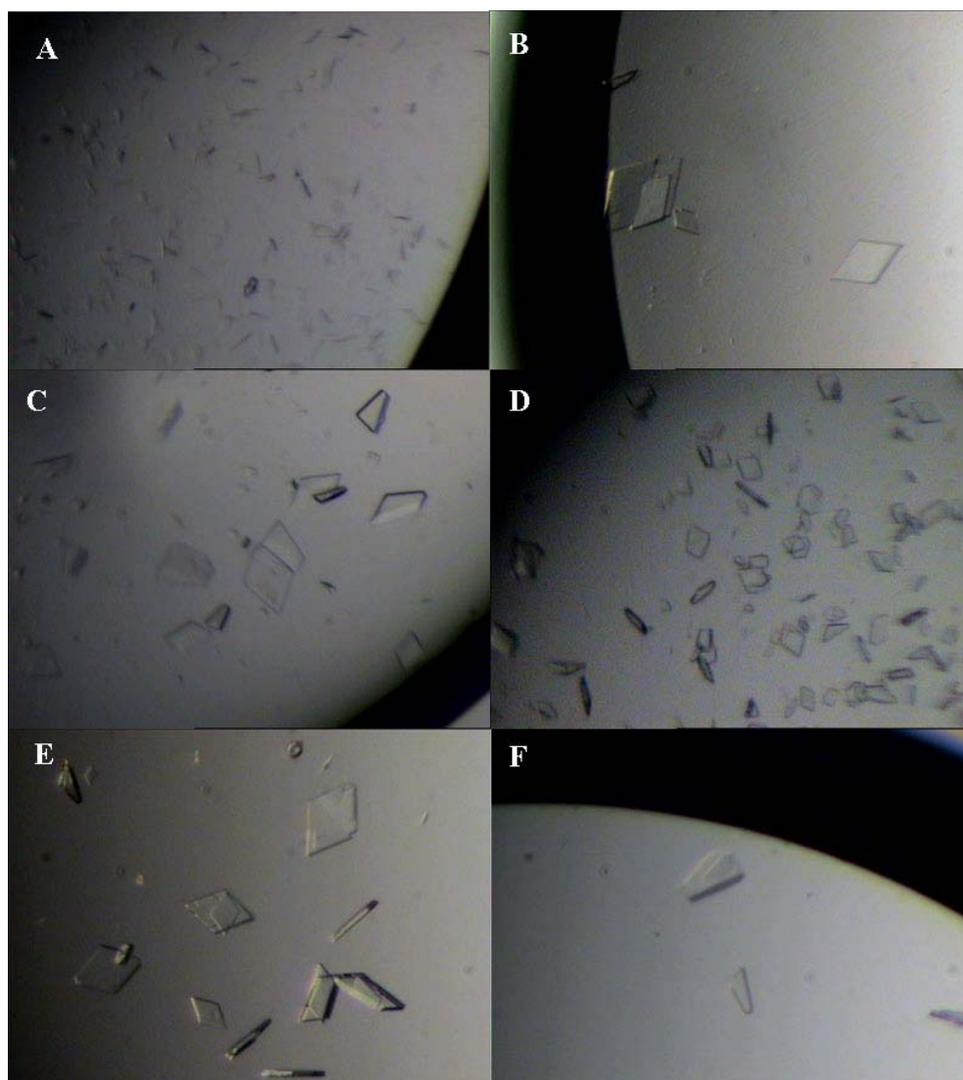


Figure 3.13 Crystals of free wild type BGlul1 in a microbatch plate from the initial screening with a Hampton HR2-134 screening kit. Conditions are A: 25% (w/v) PEG 3350, 0.2 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M HEPES, pH 7.5; B: 25% (w/v) PEG 3350, 0.2 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M Tris-HCl, pH 8.5 (HR2-134 F9); C: 25% (w/v) PEG 3350, 0.2 M NaCl, 0.1 M HEPES, pH 7.5 (HR2-134 F12); D: 25% (w/v) PEG 3350, 0.2 M Li_2SO_4 , 0.1 M HEPES, pH 7.5 (HR2-134 G4); E: 25% (w/v) PEG 3350, 0.2 M Li_2SO_4 , 0.1 M Tris-HCl, pH 8.5 (HR2-134 G5). F: 25% (w/v) PEG 3350, 0.2 M MgCl_2 , 0.1 M HEPES, pH 7.5 (HR2-134 G12).

3.4.1.2 Coarse screen crystallization trials of promising conditions

Three conditions from microbatch were optimized using chemical stocks from Hampton Research. Crystals were mainly reproduced from optimization of the G2 condition (30% (w/v) PEG 5000 MME, 0.2 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES, pH 6.5) using the grid screen shown in Figure 3.14.

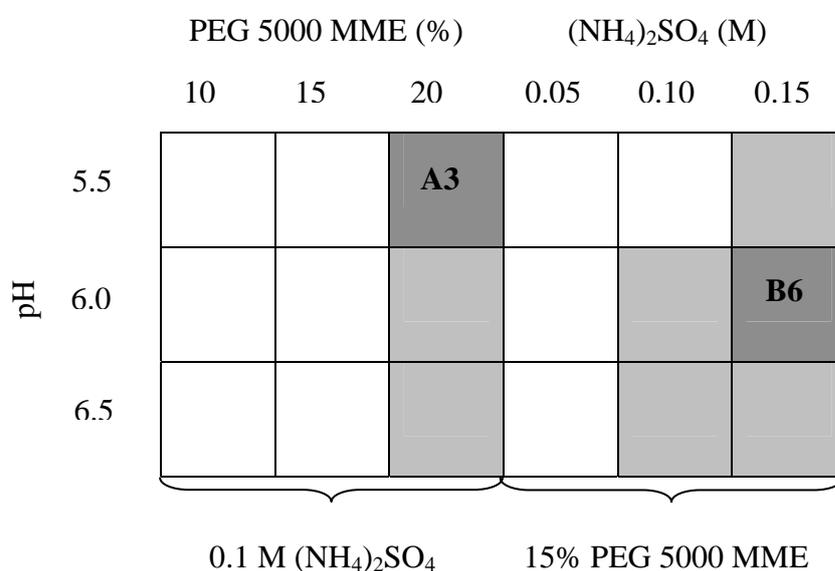


Figure 3.14 Grid screen I with variation of $(\text{NH}_4)_2\text{SO}_4$, PEG 5000 MME, and pH. The shading shows the conditions that yielded promising crystals and the darker ones for conditions where good crystals were obtained.

3.4.1.3 Fine screen crystallization trials

From grid screen I, good size and shape crystals were obtained in condition A3 (20% PEG 5000 MME, 0.1 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES, pH 5.5). Smaller, but thicker crystals, were observed in B6, which contained 15% PEG 5000

MME, 0.15 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES, pH 6.0). However, these crystals appeared after the crystallization was set up for about 8 months. Conditions were further optimized by increasing the concentrations of PEG 5000 MME up to 30% and $(\text{NH}_4)_2\text{SO}_4$ up to 0.25 M with the same pH range as in grid screen I.

As shown in grid screen II (Figure 3.15), crystals grew over a pH range of 5.5 to 6.5, but pH 6.5 seemed to be better than lower pH. At pH 6.5, BGlu1 crystallized in PEG concentrations from 20-25% and $(\text{NH}_4)_2\text{SO}_4$ concentrations from 0.10-0.20 M with the highest quality crystals in condition B10 (20% PEG 5000 MME, 0.15 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES, pH 6.5). Smaller size crystals were observed in C10, which contained a higher $(\text{NH}_4)_2\text{SO}_4$ concentration. This optimization gave crystals faster than the previous batch, in only 2-3 weeks.

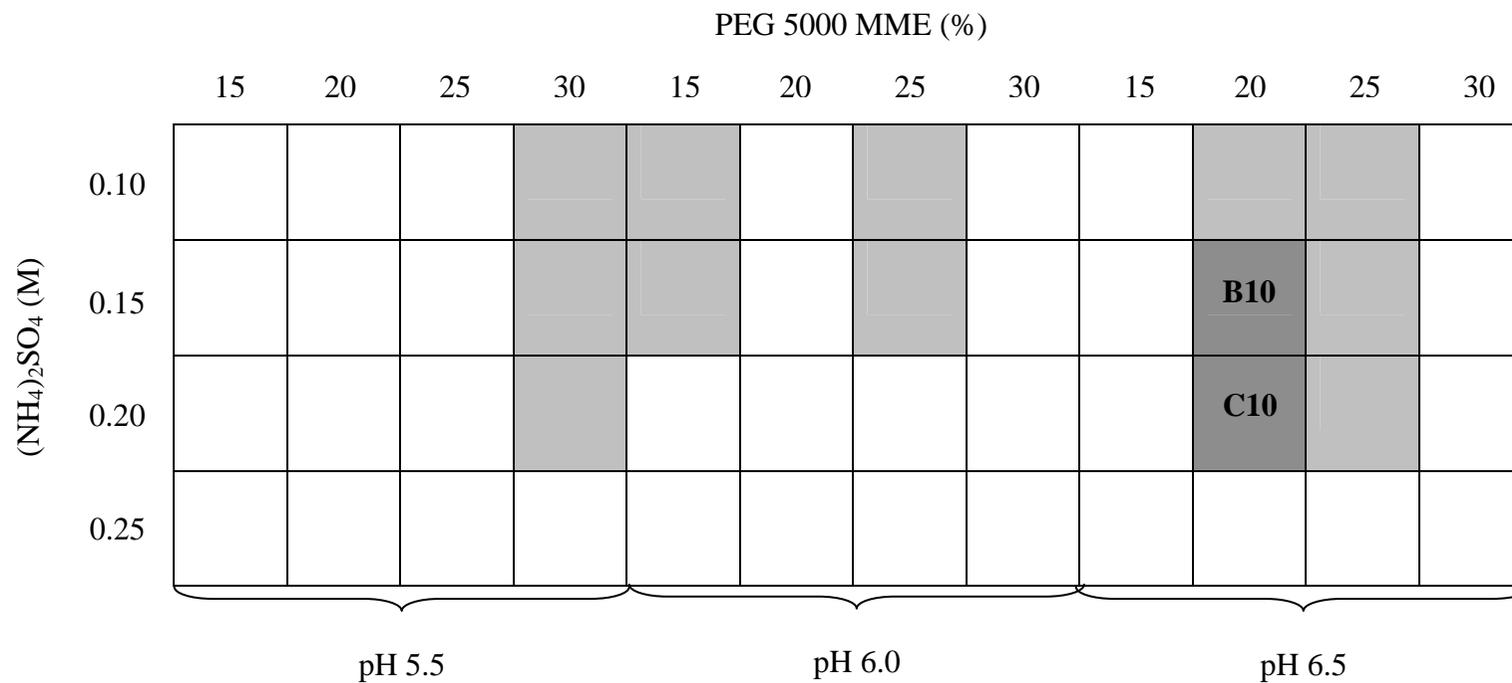


Figure 3.15 Grid screen II with variation of $(\text{NH}_4)_2\text{SO}_4$, PEG 5000 MME, and pH. The shading shows the conditions that yielded promising crystals and the darker ones (B10 and C10) indicate conditions where good crystals were obtained.

Optimization was further refined the concentrations of salt and PEG, and the pH around the successful conditions from grid screen II. In addition, the protein:precipitant ratio was also optimized this time. Multiple drops were set up over the same reservoir. For this optimization, good crystals were obtained from several conditions (C5, C6, D2, and D3) containing 23-24% PEG 5000 MME, 0.15-0.17 M $(\text{NH}_4)_2\text{SO}_4$, in MES, pH 6.5-6.7 (Figure 3.16). The best ratio of protein:precipitant was 3:1 and crystals first appeared about 2 weeks after set up.

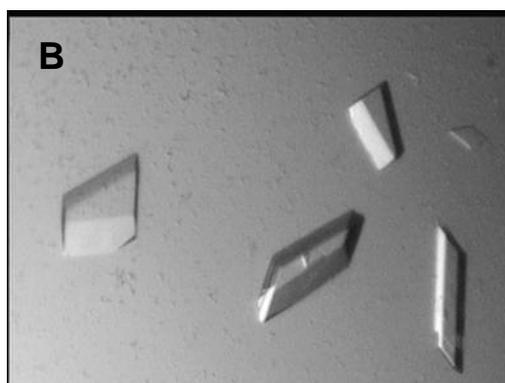
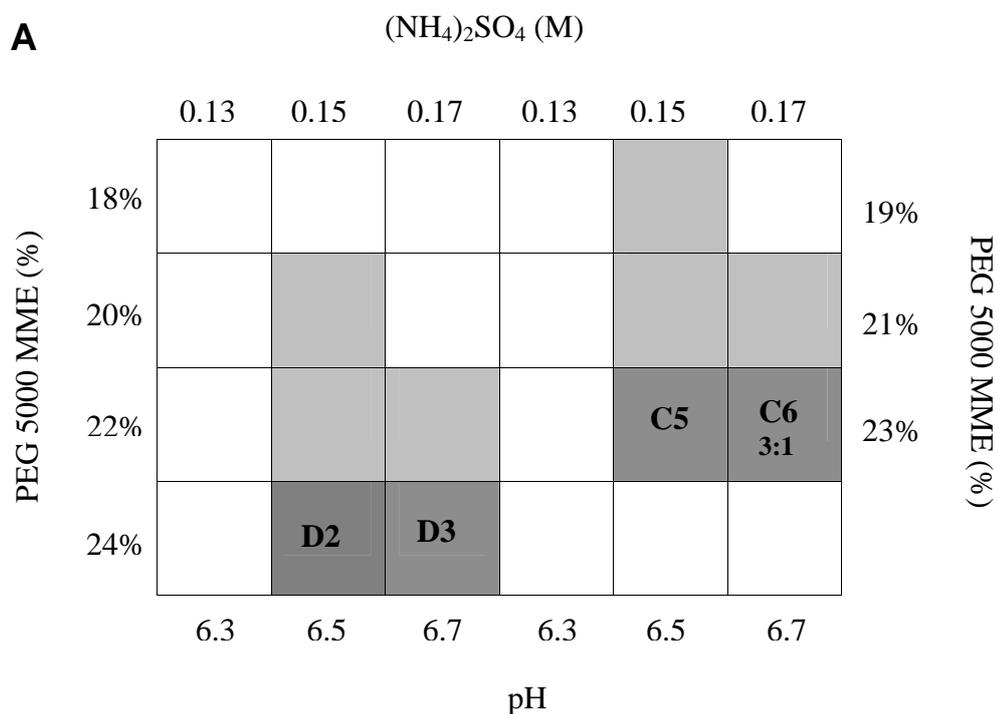


Figure 3.16 Grid screen III for crystallization of free BGlu1 and single crystals used for data collection. A) Grid screen III with variation of (NH₄)₂SO₄, PEG 5000 MME, pH, and protein: precipitant ratio (1:1, 1:2, 1:3, 2:1, 3:1). The shading shows the conditions that yielded small crystals and the darker shading conditions where good crystals were obtained. B) Free BGlu1 crystallized as tetragonal crystal with dimensions ca. 100 x 60 x 15 μm grown in the C6 condition: 23% (w/v) PEG MME 5000, 0.17 M (NH₄)₂SO₄, and 0.1 M MES, pH 6.7, by hanging drop method.

3.4.2 Wild type BGlu1 with inhibitor

3.4.2.1 Initial screening

BGlu1 10 mg/ml was first co-crystallized with 1 mM DNP2GF (1:5 molar ratio) with screening kits by the microbatch method. The crystals of BGlu1 with inhibitor crystallized within one week as clusters of thin plates in microbatch screening under the same conditions as crystals of free BGlu1. The condition that yielded crystals in the microbatch method was initially optimized with precipitant concentrations and pH. BGlu1 in complex with inhibitor still crystallized as clusters under all conditions tested.

3.4.2.2 Optimization with seeding

Cluster crystals obtained from initial optimization were crushed and used to streak seed into the new drops, which were equilibrated to be a metastable solution. Initially, fewer and larger single crystals appeared within 5 days, but they later became clusters on the tenth day. After adjustment of several parameters, seeding with optimization finally yielded single crystals. The three successful crystallization conditions listed in Table 3.2 were used in optimization of the protein concentration (1, 2, 3, 4, 5 mg/ml) with streak-seeding. Precipitant solutions no. 1 and 3 yielded several single crystals at two protein concentrations, 4 and 5 mg/ml. In contrast, crystals grew singly at all protein concentrations (1-5 mg/ml) in crystallization condition no. 3, and fewer and larger single crystals with diffraction quality were obtained from the 2 mg/ml drop (Figure 3.17).

Table 3.2 The crystallization conditions obtained from optimization for the BGlu1 inhibitor complex

No.	Precipitant condition
I	15% PEG 8000, 0.15 M ZnOAc, 0.1 M Na cacodylate, pH 6.7, 15% glycerol
II	21% PEG 8000, 0.18 M ZnOAc, 0.1 M Na cacodylate, pH 6.7
III	23% PEG MME 5000, 0.20 M (NH ₄) ₂ SO ₄ , 0.1 M MES, pH 6.7

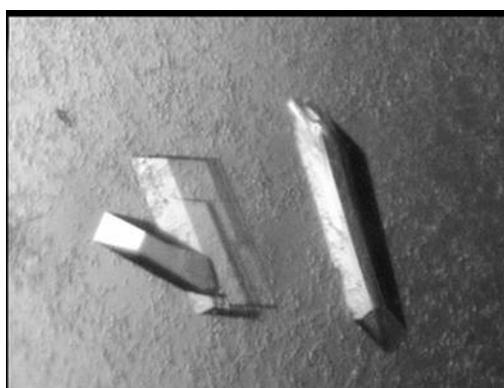


Figure 3.17 Crystals of BGlu1 with the inhibitor DNP2FG. Tetragonal crystals with sizes up to 320 x 140 x 20 μm formed in a hanging drop vapor diffusion plate. The precipitant was 23% (w/v) PEG MME 5000, 0.2 M ammonium sulfate, 0.1 M MES, pH 6.7.

3.4.3 E176Q mutant

3.4.3.1 Initial screening

Crystallization screening of the E176Q mutant was initially started like that of the wild type BGlu1 protein with microbatch screening using two screening kits, HR2-130 (Hampton Research) and a systematic screen of mono- and di-valent salts and pH in the presence of 25% (w/v) PEG 4000 at 10 mg/ml protein concentration. The same condition in which wild type BGlu1 crystallized yielded

E176Q mutant crystals (G2_HR2-130: 30% PEG MME 5000, 0.2 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES, pH 6.5).

3.4.3.2 Optimization without micro-seeding

Variable parameters, including the concentrations of PEG MME 5000 and $(\text{NH}_4)_2\text{SO}_4$, and the drop ratio of protein:precipitant were optimized in 0.1 M MES, pH 6.7 (obtained from the thorough optimization of wild type BGlu1 protein). Crystals of the E176Q mutant at an initial protein concentration of 10 mg/ml appeared after setting up for 1 week as poorly shaped single crystals.

3.4.3.3 Optimization with micro-seeding

The irregularly shaped single crystals produced in 20% PEG MME 5000, 0.17 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES, pH 6.7, were crushed to prepare a seed stock, which was diluted 1:10 for further optimization using the grid screen in Figure 3.18A. For this screening, protein concentration was also optimized by testing 2, 4, 6, and 8 mg/ml protein solutions. Good single crystals were produced in conditions A2 and D3, and a thicker one from A2 was chosen for x-ray diffraction (Figure 3.18B).

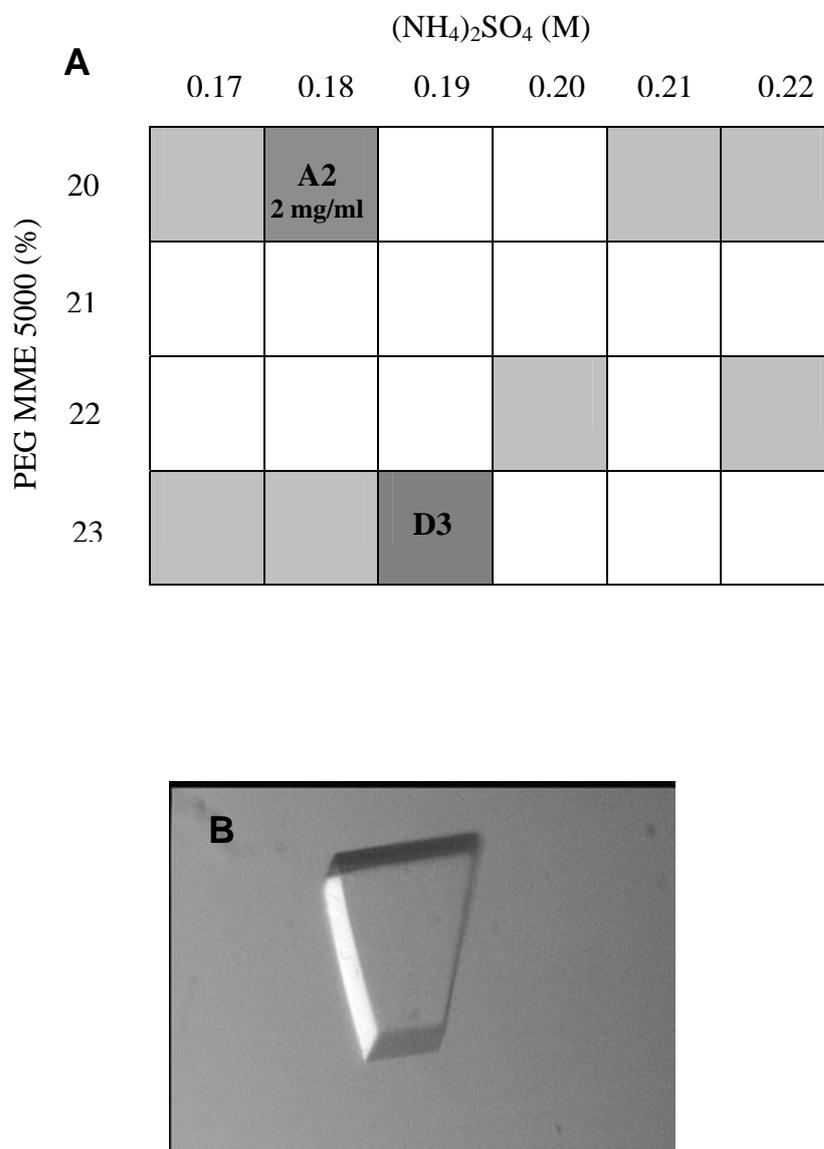


Figure 3.18 Optimization of crystallization for the free E176Q mutant BGlu1 and a single crystal used for x-ray diffraction. A) Grid screen for optimization of E176Q mutant BGlu1 crystallization. The shading shows the conditions that yielded small crystals and the darker shading indicates conditions where x-ray diffracted crystals were obtained. B) An E176Q mutant crystal with dimensions of 160 x 110 x 40 μm produced in the A2 condition: 20% PEG MME 5000, 0.18 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES, pH 6.7, by the hanging drop method.

3.4.4 E176Q mutant with cellopentaose

3.4.4.1 Optimization with micro-seeding

Crystallization trials of the E176Q mutant protein with cellopentaose were initiated at optimization with micro-seeding. Good single crystals of the free E176Q protein (22% PEG MME 5000, 0.20 mM $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES, pH 6.7) were used as seed stock. Optimization with micro-seeding was performed with refinement of $(\text{NH}_4)_2\text{SO}_4$ and PEG MME 5000 concentrations, in 0.1 M MES, pH 6.7, with low protein concentrations (2, 4, 6 mg/ml) as shown in the grid screen in Figure 3.19A. Clustered crystals appeared under all conditions after a week. Therefore, further optimization was performed with lower protein concentrations (2, 3, and 4 mg/ml) and optimization of pre-equilibration time before streak seeding (1 and 2 hr.) using the same grid. Decreasing the protein concentration and pre-equilibration time yielded single crystals in the A3 and B6 conditions (22% PEG 5000 MME, 0.19 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES, pH 6.7, and 23% PEG 5000 MME, 0.22 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES, pH 6.7) with 2 mg/ml protein and 1 hr pre-equilibration time. One crystal from the latter condition was used for diffraction in the x-ray beam (Figure 3.19B).

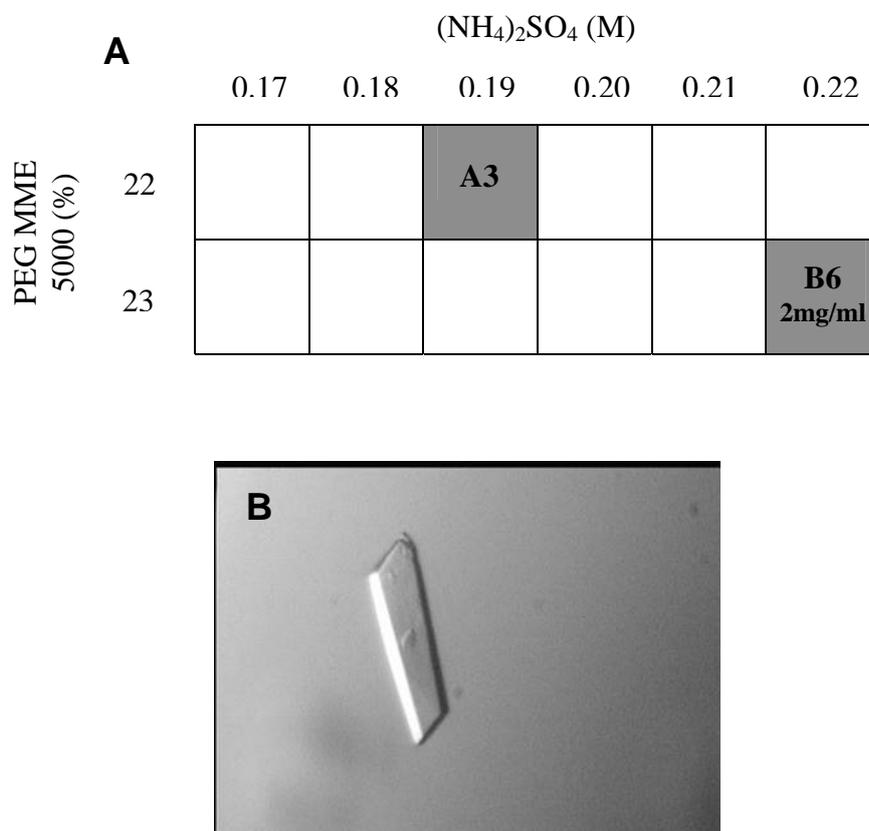


Figure 3.19 Optimization of crystallization of E176Q mutant with cellopentaose and single crystal used for x-ray diffraction. A) Grid screen with small variations in PEG MME 5000, (NH₄)₂SO₄, and protein concentrations. B) Single crystals with dimensions of 170 x 50 x 30 μm obtained in the B6 condition (23% PEG 5000 MME, 0.22 M (NH₄)₂SO₄, 0.1 M MES, pH 6.7) by the hanging drop method.

3.5 N-terminal amino acid sequencing

Since no electron density of amino acids at the N-terminus was observed until Trp6, the N-terminus of rice BGlu1 was sequenced to check the sequence of the mature protein after enterokinase cleavage. From N-terminal amino acid sequencing, the sequence AMADVVPK was obtained. The first five amino acids are from the linker between the enterokinase site and the mature protein N-terminus, after removal of the N-terminal thioredoxin-, His₆- and S-tags. The next 3 amino acid residues are the N-terminal sequence of the mature BGlu1 protein (Genbank Accession AAA84906). Therefore, the first 10 amino acids appear to be missing in the structure due to flexibility rather than proteolytic loss from the protein.

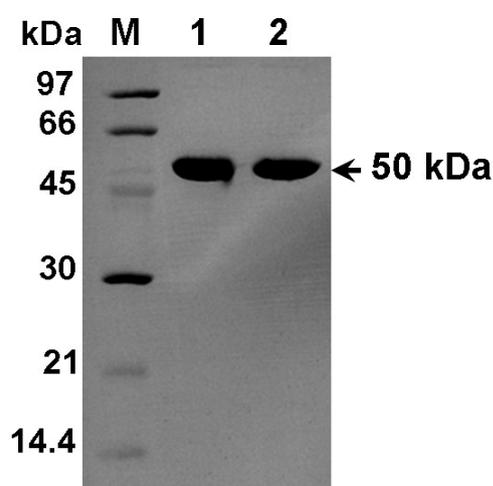


Figure 3.20 Protein from dissolved crystals of rice BGlu1 blotted onto PVDF membrane before the protein bands were cut from the membrane for sequencing.

3.6 Data collection

3.6.1 Crystal mounting for x-ray diffraction

We collected the data using x-ray beams from both an in-house rotating anode and from synchrotron radiation. A 0.5-1° rotation image taken from data collection at a synchrotron source is shown in Figure 3.21. All data collection from in-house and synchrotron sources of individual diffracted crystals are summarized in Table 3.3. After data integration and scaling, it was clear that all BGlu1 crystals with and without inhibitor or ligands belong to the $P2_12_12_1$ space group with almost the same unit cell parameters indicating they were isomorphous, as summarized in Table 3.4 with other data collection statistics.

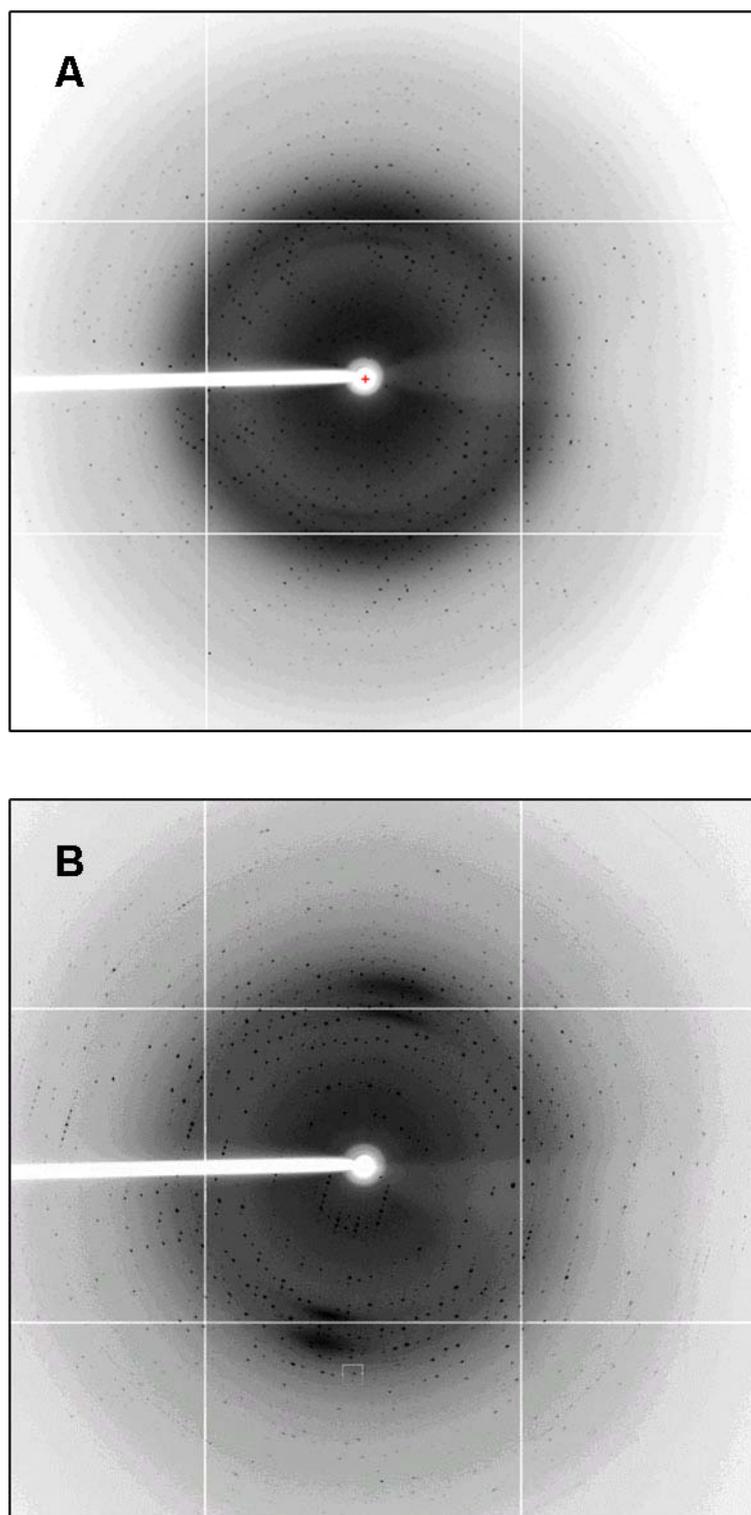


Figure 3.21 X-ray diffraction patterns of BGlul with DNP2FG inhibitor (A) and E176Q mutant (B), which diffracted to 1.55 and 1.37 Å resolution, respectively.

Table 3.3 Summary of drop ingredients, crystallization conditions, sizes of diffracted crystals, cryoprotectants used for data collection, and resolution obtained for each condition.

In-house diffraction

Crystals	Drop	Precipitant	Ligand	Size (μm^3)	Cryoprotectant	Resolution (\AA)
Wild type BGlu1	1 μl 10 mg/ml protein +1 μl precipitant	18% (w/v) PEG 8000, 0.2 M ZnOAc, 0.1 M Na cacodylate, pH 6.5	None	120×60×16	18% (w/v) PEG 8000, 0.2 M ZnOAc, 0.1 M Na cacodylate pH 6.5, 18% (v/v) glycerol	2.75
Wild type BGlu1 + DNP2FG	2 μl 2 mg/ml protein with 0.2 mM DNP2FG +1 μl precipitant	23% (w/v) PEG MME 5000, 0.2 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES, pH 6.7	2-fluoroglucoside (co-crystallized with 0.2 mM DNP2FG)	300×140×20 (140×140×20 fragment of crystal diffracted)	27% (w/v) PEG MME 5000, 0.238 M $(\text{NH}_4)_2\text{SO}_4$, 0.115 M MES, pH 6.7	2.15

Table 3.3 (continued).**Synchrotron diffraction**

Crystals	Drop	Precipitant	Ligand	Size (μm^3)	Cryoprotectant	Resolution (\AA)
Wild type BGlu1	3 μl 11.5 mg/ml protein +1 μl precipitant	23% (w/v) PEG MME 5000, 0.17 M (NH ₄) ₂ SO ₄ , 0.1 M MES pH 6.7	None	100×60×15	27% (w/v) PEG MME 5000, 0.2 M (NH ₄) ₂ SO ₄ , 0.115 M MES, pH 6.7, 18% (v/v) glycerol	2.20
Wild type BGlu1 + DNP2FG	2 μl 2 mg/ml protein with 0.2 mM DNP2FG +1 μl precipitant	23% (w/v) PEG MME 5000, 0.2 M (NH ₄) ₂ SO ₄ , 0.1 M MES, pH 6.7	2- fluoroglucoside (co-crystallized with 0.2 mM DNP2FG)	320×140×20	27% (w/v) PEG MME 5000, 0.238 M (NH ₄) ₂ SO ₄ , 0.115 M MES, pH 6.7	1.55

Table 3.3 (continued).

Crystals	Drop	Precipitant	Ligand	Size (μm^3)	Cryoprotectant	Resolution (\AA)
E176Q + soaked L2	2 μl 2 mg/ml protein + 1 μl precipitant	20% (w/v) PEG MME 5000, 0.18 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES, pH 6.7	1 mM L2 ^a (soaked)	160×110×40	23.6% (w/v) PEG MME 5000, 0.212 M $(\text{NH}_4)_2\text{SO}_4$, 0.118 M MES, pH 6.7, 18% (v/v) glycerol	1.33
E176Q + C5	2 μl 2 mg/ml protein with 1.0 mM C5 + 1 μl precipitant	23% (w/v) PEG MME 5000, 0.22 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES pH 6.7	1 mM C5 ^b (co-crystallized)	140×120×30	27.1% (w/v) PEG MME 5000, 0.26 M $(\text{NH}_4)_2\text{SO}_4$, 0.115 M MES, pH 6.7, 18% (v/v) glycerol and 1 mM cellopentaose	1.80

^a 1 mM laminaribiose was added into precipitant solution for soaking with E176Q crystal, before transferring to cryoprotectant solution. Since laminaribiose concentration was not high enough or the substrate was not added into cryoprotectant, no strong laminaribiose density was seen for this crystal. ^b 1 mM cellopentaose was co-crystallized with the E176Q mutant protein. The crystal of cellopentaose complex was cryoprotected by cryoprotectant solution in the presence of 1 mM cellopentaose.

Table 3.4 Data collection statistics for BGlu1, BGlu1 with 2-fluoroglucoside, E176Q mutant, and E176Q mutant with cellopentaose.

	Free BGlu1	BGlu1 with DNP2FG	Free BGlu1	BGlu1 with DNP2FG	E176Q mutant	E176Q mutant with cellopentaose
Wavelength (Å)	1.54	1.54	0.98	0.98	1.00	1.00
Resolution (Å)	55.7–2.75 (2.85–2.75)	53.8–2.15 (2.23–2.15)	30.0–2.20 (2.28–2.20)	30.0–1.52 (1.57–1.52)	30.0–1.35 (1.40–1.35)	30.0–1.80 (1.86–1.80)
Completeness (%)	99.5 (99.9)	96.5 (83.1)	99.9 (99.5)	98.8 (96.3)	97.3 (96.1)	97.6 (87.3)
R_{merge}^a (%)	14.6 (35.4)	11.0 (30.7)	13.2 (40.4)	5.2 (40.3)	5.1 (37.8)	7.5 (41.4)
$\langle I/\sigma(I) \rangle$	3.6 (1.6)	4.3 (1.8)	15.5 (5.2)	17.5 (2.29)	18.2 (2.0)	13.0 (2.1)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 78.85$ $b = 100.34$ $c = 127.10$	$a = 79.37$ $b = 100.92$ $c = 127.23$	$a = 80.05$ $b = 100.67$ $c = 127.02$	$a = 80.29$ $b = 101.29$ $c = 127.25$	$a = 79.62$ $b = 100.80$ $c = 127.63$	$a = 79.85$ $b = 101.25$ $c = 127.44$

Table 3.4 (continued).

	Free BGlu1	BGlu1 with DNP2FG	Free BGlu1	BGlu1 with DNP2FG	E176Q mutant	E176Q mutant with cellopentaose
No. of unique reflections	50,449	103,691	53,089	157,799	224,942	95,916
No. of observed reflections	97,594	213,255	363,672	491,141	2,161,078	2,010,315
Unit cell volume (Å ³)	1,005,589.7	1,019,114.5	-	-	-	-
V _M (Å ³ Da ⁻¹)	2.5	2.5	-	-	-	-
Solvent content (%)	50.7	51.3	-	-	-	-
No. of molecules per AU	2	2	2	2	2	2

Numbers in parentheses are outer shell parameters.

^a $R_{\text{merge}} = (\sum |I - \langle I \rangle| / \sum \langle I \rangle)$

3.6.2 XAS scan of BGlu1 crystal at zinc edge

The presence of zinc bound to protein in the crystal packing was confirmed by an XAS scan. The wavelength was scanned through the range expected for the zinc absorbance for the edge. The spectrum was separated into the real and imaginary components to find $\Delta f'$ and $\Delta f''$ a function of wavelength. The wavelengths that would be chosen for anomalous diffraction are at the f'' maximum and f' minimum, which are found at the absorption peak and edge, respectively. As shown in Figure 3.22, the sharp change was observed at the adsorption edge of zinc, which indicates that zinc was bound in crystal packing.

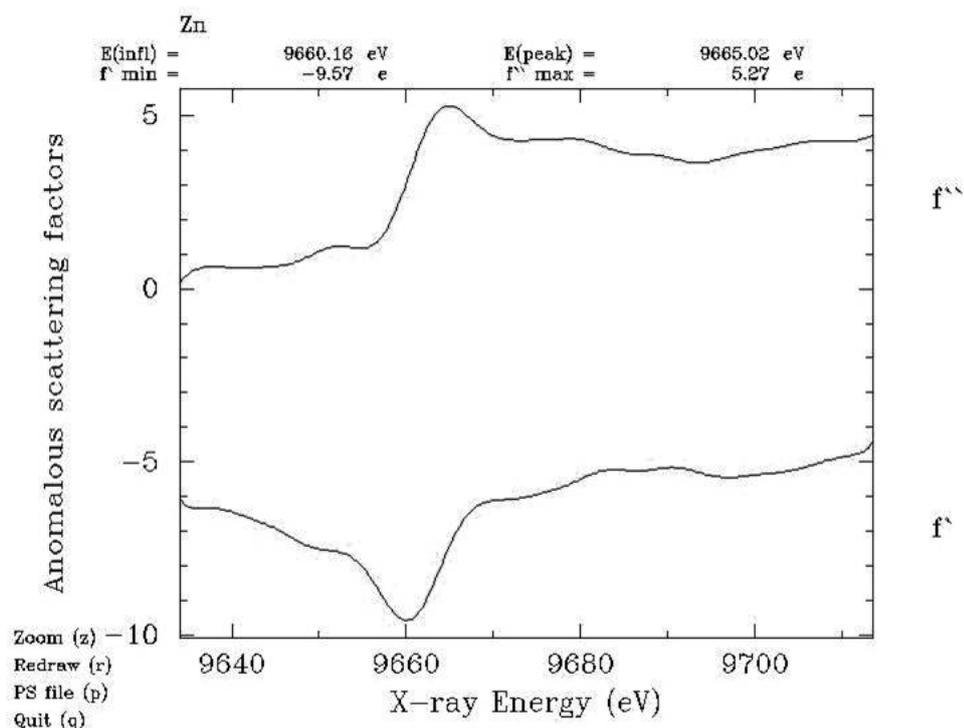


Figure 3.22 XAS scan of the zinc edge for a BGlu1 crystal with plots of f' and f'' .

3.7 Molecular replacement

The BGlu1 structure was solved by molecular replacement using the *AMoRe* program (Navaza, 1994). The cyanogenic β -glucosidase from white clover (*Trifolium repens*) 1CBG; Barrett et al., 1995) was used as a search model, since it had the highest sequence similarity (46% identity) of proteins with known structures. Two molecules were found per asymmetric unit. Model building and refinement during the refinement process are shown in Figure 3.23.

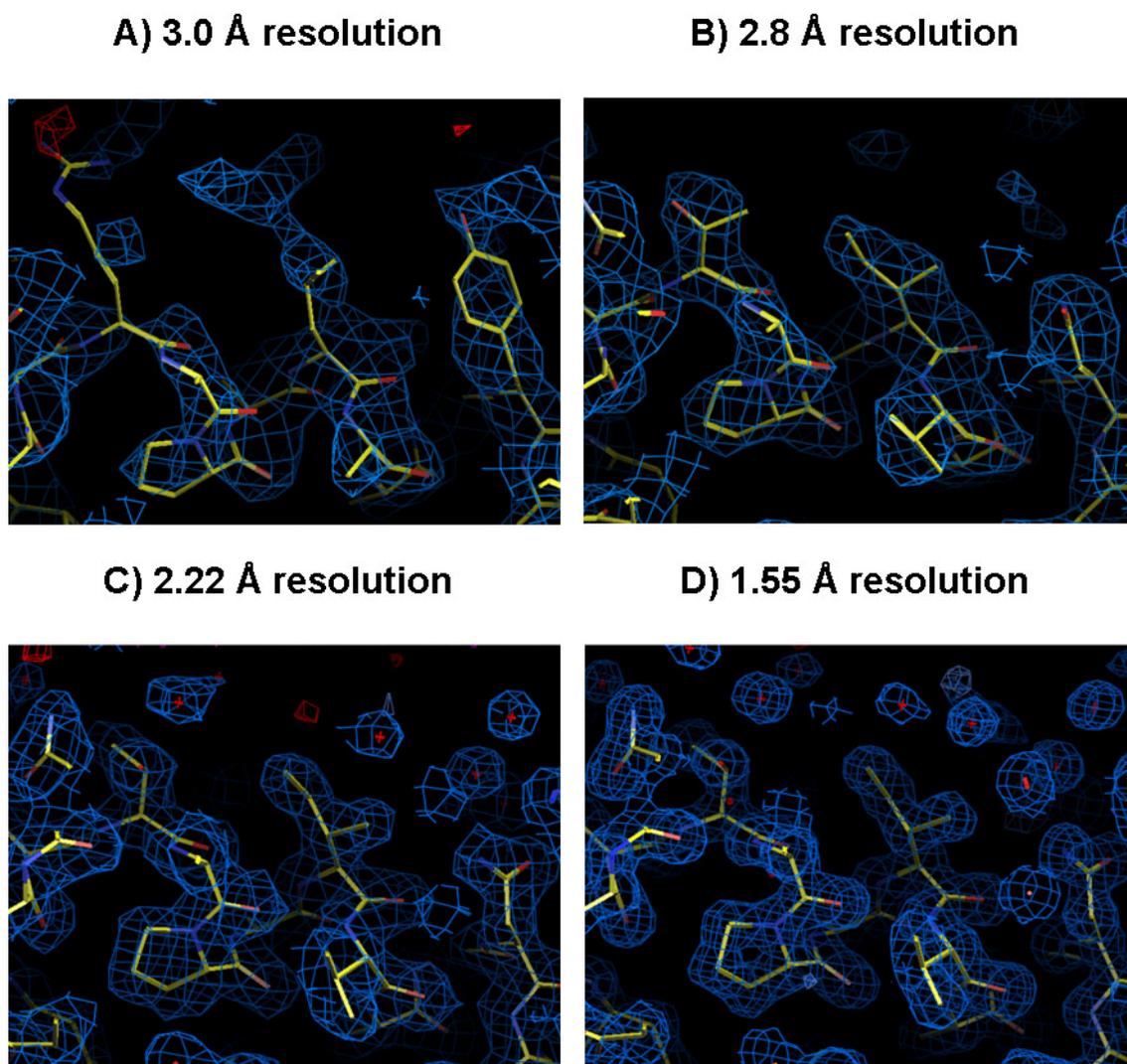


Figure 3.23 The improvement of the solved structure indicated by the $2|F_{\text{obs}}| - |F_{\text{calc}}|$ map (blue) contoured at the 1σ level at amino acid residues 206-215 of rice BGlu1. A) First model obtained from molecular replacement at 3.0 Å before changing residues from 1CBG as a search model to rice BGlu1. B) Subsequently refined structure and $2|F_{\text{obs}}| - |F_{\text{calc}}|$ maps at 2.8 Å resolution. C) Well-refined structure of free BGlu1 at 2.2 Å resolution. D) Final structure of rice BGlu1 with 2-fluoroglucose at 1.55 Å resolution. The figures were generated with program O (Jones et al., 1991).

3.8 Structure determination of wild type BGlu1 with and without inhibitor

3.8.1 Overall structure and quality of models

As previously noted, crystals of free BGlu1 enzyme and BGlu1 bound to G2F are isomorphous, and belong to the orthorhombic $P2_12_12_1$ space group. The final model of free enzyme was refined at 2.20 Å resolution to an R -factor of 16.2% and R_{free} of 20.5% for 7602 protein atoms (471 amino acid residues for each of two molecules in the asymmetric unit) and 1178 water molecules. The structure of the G2F complex was finally refined at 1.55 Å to an R -factor of 17.1% and R_{free} of 19.7% with the presence of 11 2-fluoroglucoside atoms per molecule, and 1532 water molecules in the asymmetric unit (Table 3.5).

Table 3.5 Refinement statistics for BGlu1, BGlu1 with inhibitor, inactive E176Q mutant, and inactive E176Q mutant with cellopentaose.

	Free BGlu1	BGlu with DNP2FG	E176Q mutant	E176Q mutant with cellopentaose
Resolution used in refinement (Å)	29.7-2.20	28.2-1.55	28.1-1.37	27.9-1.80
No. reflections				
Working set	50,374	141,569	198,814	90,013
Test set	2,652	7,487	10,551	4,669
$R_{\text{factor}}^{\text{a}}$ (%)	16.2	17.1	16.6	18.1
$R_{\text{free}}^{\text{b}}$ (%)	20.5	19.7	18.7	21.1
No. of residues in protein	952	952	952	952
No. of protein atoms	7602	7602	7602	7602
No. of water molecules	1187	1352	2085	712
No. of ligand atoms	none	22	none	112
No. of other hetero-atoms	47	47	94	94
Mean B -factor protein (Å ²)	10.43	10.74	10.26	14.75
Mean B -factor ligand (Å ²)	24.81	14.21	17.83	25.38
Mean B -factor solvent (Å ²)	23.40	27.77	36.96	31.69
R.m.s. bond deviation (Å)	0.010	0.011	0.008	0.015
R.m.s. angle deviations (°)	1.308	1.402	1.283	1.524
Residues in most favorable regions (%)	89.0	89.0	89.4	89.5

Numbers in parentheses are outer shell parameters.

^a $R_{\text{factor}} = (\sum |F_o| - |F_c|) / \sum |F_o|$.

^b Based on 5% of the data

^d The protein structures cover residues 6-476 of the mature rice BGlu1 protein sequence.

The overall structure of rice BGlu1 was similar to other GH1 structures, with a $(\beta/\alpha)_8$ or TIM barrel consisting of a core of eight twisted parallel β -strands connected by long loop structures and eight α -helices that form the outer layer of the core structure (Barrett et al., 1995). The catalytic acid/base E176 and nucleophile E386 are located at the ends of β -strands 4 and 7 on opposite sides of a cleft at the bottom of the active site (Figure 3.24). The two molecules in the asymmetric unit are related by two-fold symmetry, so they appear to be a dimer linked by a single metal ion.

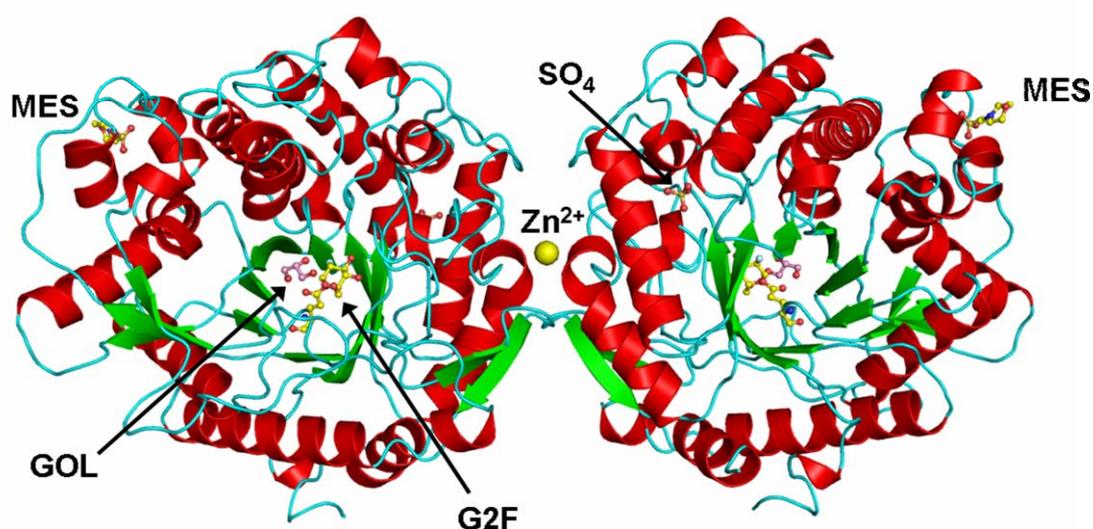


Figure 3.24 A ribbon diagram representing the overall structure of the BGlu1/G2F inhibitor complex asymmetric unit. Beta-strands are colored green, α -helices red and loops cyan. The nucleophilic catalytic residue E386 that is covalently bound to the G2F inhibitor is shown as ball and stick and colored by atoms with carbons in yellow. Hetero atoms found in crystal structure: GOL (glycerol), MES, and SO_4 , are drawn as ball and stick (colored by atoms) and a yellow sphere represents Zn^{2+} . The figure was produced with Pymol (DeLano, 1991).

The Ramachandran plots of the free BGlu1 enzyme and its G2F complex were calculated with PROCHECK (Laskowski et al., 1993). In both the free enzyme and the G2F complex structures, 89.0% of non-proline and non-glycine residues were found in the most favored regions and a further 10.2% in additional allowed regions. Four residues (Ala63, Lys289 of molecules A and B) were found in generously allowed regions in both structures. Only the Trp441 residues in both molecules A and B were found in the disallowed region for the free enzyme, though they were in the generously allowed region for the G2F complex. The homologous Trp454 of the 1CBG structure used as template also has phi and psi values very close to those of the G2F complex and free BGlu1, which fall into the generously allowed regions of the Ramachandran plot. No residues were found in disallowed regions for the G2F complex structure.

The quality of electron density at the four loop regions (residues T25-D65 of loop A, P177-T206 of loop B, Q314-P363 of loop C, and N387- D403 of loop D) was almost as good as in the core region for both main chain and side chains. However, the densities of the side chains of a few residues that are fully exposed to the solvent were still ambiguous. Poorest density was observed in loop C, where some residues, especially Q323, Q324, and Q327 had ambiguous side chain densities, but the main chain density was still clear and continuous. Linking C195 and C198 in loop B is a disulfide bond that is strictly conserved in plant GH1 enzymes. Two *cis*-peptide bonds were found between P191 and P192 and W433 and S434. The latter non-proline *cis*-peptide bond occurred after W433, which is a crucial residue to bind the glycone at subsite -1. *Cis*-peptide bonds are found at the conserved amino acids in the 1CBG template at these same positions. One SO_4^{2-} ion and one MES molecule from

the precipitant solution were observed to be associated with each monomer in the asymmetric unit, and a single Zn^{2+} was shared between molecules A and B of the asymmetric unit.

3.8.2 Overall structure comparison of rice BGlu1 with other structures in GH1 family

The known 3D structures of the GH family 1 linamarase from *Trifolium repens* (1CBG, Barrett et al., 1995); *Zea mays* β -glucosidase isozyme I (1E1E, Czjzek et al., 2001); dhurrinase isozyme I from *Sorghum bicolor* (1V02, Verdouqc et al., 2004); myrosinase from *Synapis alba* (1MYR, Burmeister et al., 1997); and *Paenibacillus polymyxa* β -glucosidase BglB (2JIE, Isorna et al., 2007) were superimposed on the BGlu1 structure for structural comparison (Figure 3.25).

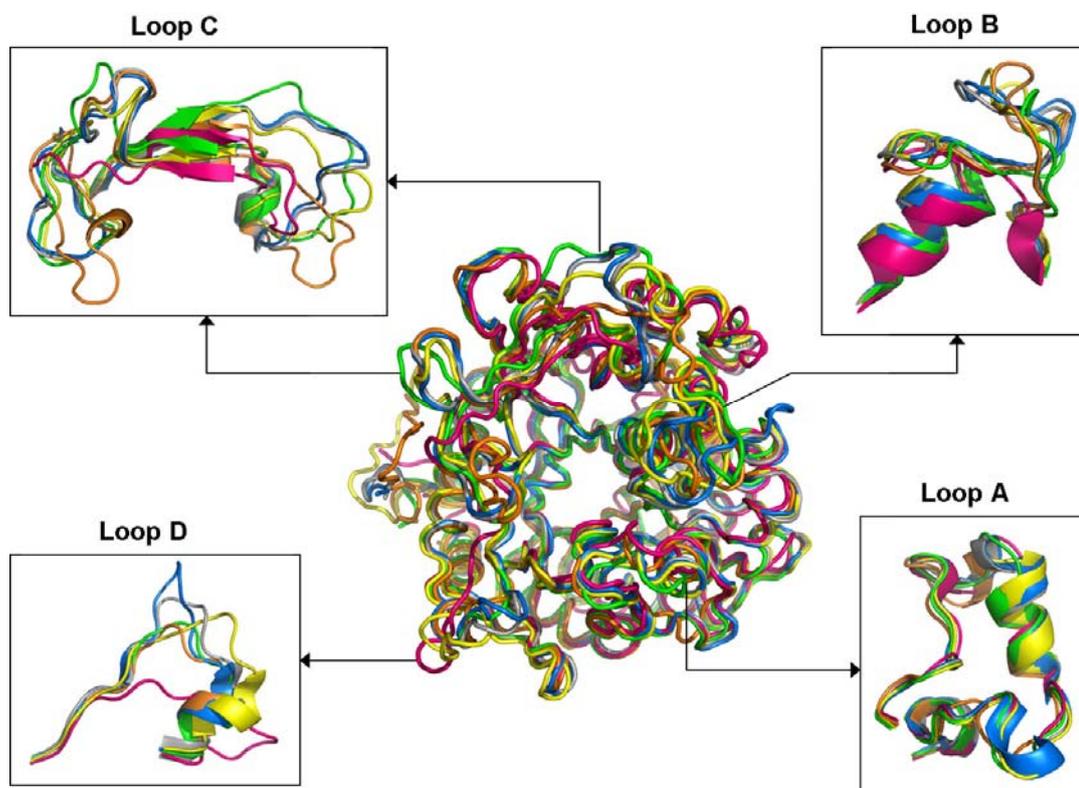


Figure 3.25 Superimposition of the structures of rice BGlu1 (green), with other enzymes in GH family 1; linamarase from *Trifolium repens* (1CBG, yellow), *Zea mays* β -glucosidase isozyme I (1E1E, blue), *Sorghum bicolor* dhurrinase isozyme I (1V02, gray), *Synapis alba* myrosinase (1MYR, orange) and *Paenibacillus polymyxa* β -glucosidase BglB (2JIE, pink). Loops A-D, which constitute the doorway to the active site are expanded to the side, as indicated by the arrows, to show the differences in loop structures. The figures were produced with the Pymol program (DeLano, 1991).

The 1CBG, 1E1E and 1V02 structures represent plant *O*-glucosidases, 1MYR is a plant *S*-glucosidase and 2JIE is bacterial enzyme that hydrolyzes cellooligosaccharides with high degrees of polymerization as substrates, similar to BGlu1. These structures share amino acid sequence identity of 37-48%. The C_α atoms of molecule A from the 1CBG, 1E1E, 1V02, 1MYR, and 2JIE structures superimposed to rice BGlu1 with rmsd of 0.80, 0.96, 0.91, 0.96, and 1.44 Å, respectively. The 1CBG structure, which was used as the search model had highest sequence identity (48%) and the lowest rmsd values of the overall structure and loops. Inversely, 2JIE, the only β-glucosidase from bacteria, had the largest differences from our structure.

The structural superimposition showed that the core (β/α)₈ structures of the GH1 structures are similar in shape and size, while the largest differences are seen in loops surrounding the active site. Four variable loop regions have been described that connect the β-strands to α-helices of the (β/α)₈-barrel: Loop A (loop 1) between β1 and α1, Loop B (loop 4) between β4 and α4, Loop C (loop 6) between β6 and α6, and Loop D (loop 7) between β7 and α7 (Barrett et al., 1995; Burmeister et al., 1997; Sanz-Aparicio et al., 1998). Loop A has little difference in length and structural alignment for the six structures, though it contains the residues involved in dimerization of rice BGlu1 and myrosinase. Loop B contains the conserved disulfide bond in plant GH1 enzymes and contributes part of the aglycone binding pocket. The hairpin structure closed by this disulfide bridge is several residues shorter in rice BGlu1 than other plant enzymes. The largest differences are in loop C, as seen in the sequence (Figure 3.26) and structural comparisons. This is the largest loop and it

covers almost a half of the entrance to the active site, and forms a part of the aglycone binding site. In comparison to rice BGlu1, two extra residues were observed in maize, sorghum, and myrosinase, and another insertion of six extra amino acid residues are found only in myrosinase. Consequently, this region of rice BGlu1 is less bulky and its active site has a more extended entrance way than other plant enzymes, which is suitable for binding of long oligosaccharide substrates of up to six oligomers (Opassiri et al., 2004). This loop (D342-Y357) was also reported to form a dimer interface in the maize β -glucosidase structure (Czjzek et al., 2001). Finally, Loop D, which is the smallest loop (12-16 amino acids) also showed high variation, but it is not close to the substrate-binding site. This loop functions in the dimerization of the 1CBG linamarase (Barrett et al., 1995).

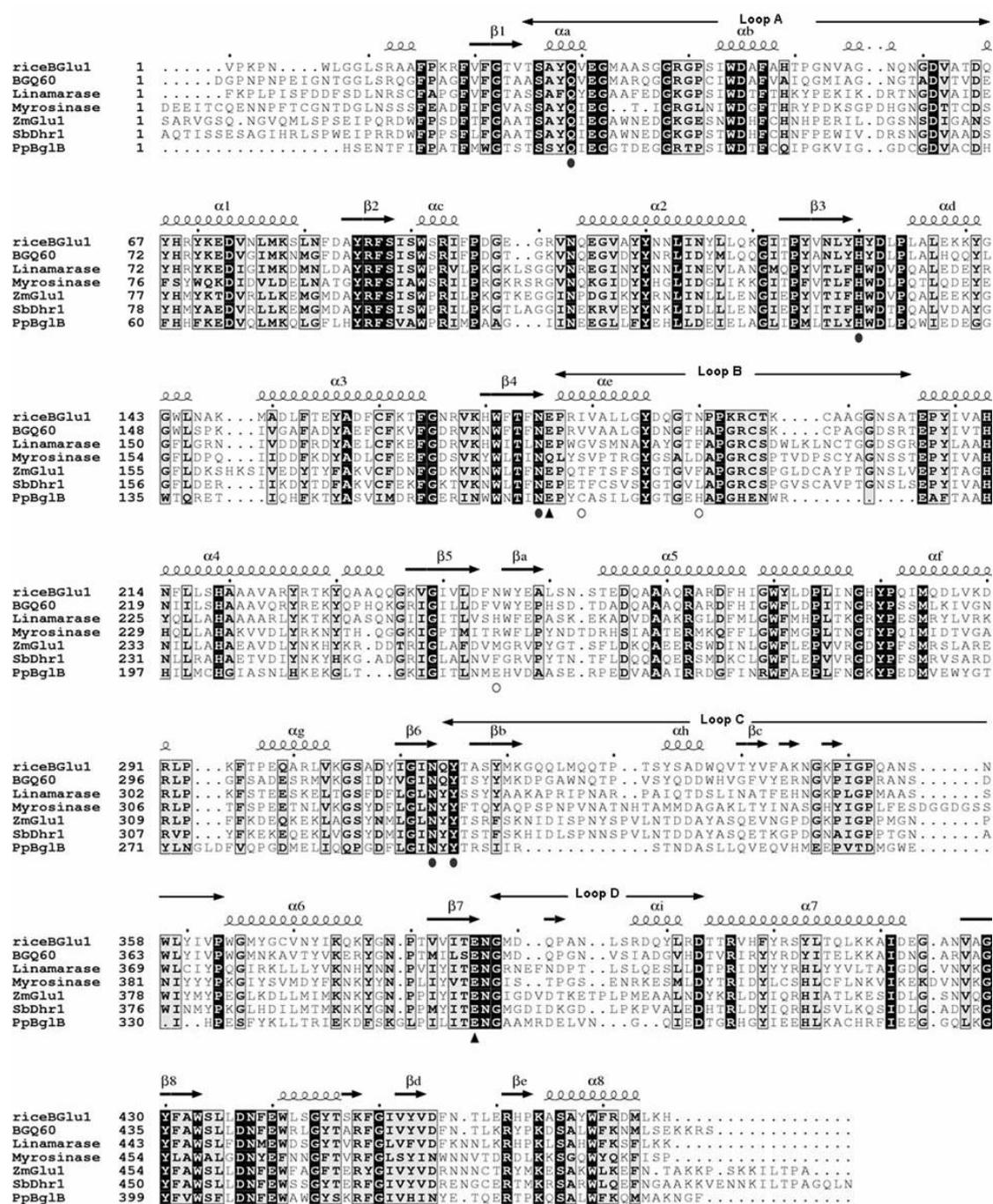


Figure 3.26 Amino acid sequence alignment of rice BGlul1 β -glucosidase with related enzymes. BGQ60, barley β -glucosidase BGQ60; linamarase, cyanogenic β -glucosidase from *Trifolium repens* (1CBG); myrosinase, *Sinapis alba* myrosinase (1MYR); ZmGlu1, *Zea mays* β -glucosidase isozyme I (1E1E); SbDhr1, dhurrinase

isozyme I from *Sorghum bicolor* (1V02); and PpBglB, *Paenibacillus polymyxa* β -glucosidase BglB (2JIE). The alignment was generated with ClustalW, and the secondary structure elements of rice BGlu1 were aligned at the top with the ESPript program. Black arrowheads indicate acid/base and nucleophilic catalytic residues, black circles marked the amino acids in close contact with the glucose residue at subsite -1, and white circles highlight the amino acids predicted as putative aglycone binding residues at subsites +1 and +2.

The loop structure of PpBglB is extremely different from the GH1 plant enzymes. A disulfide linkage is found at loop A instead of loop B. Both loops B and C were much shorter than the plant enzymes, but their folds are similar, except for the absence of the long coiled structures linking the short β -strands and α -helices. The shorter loops B and C in rice Bglu1 and PpBglB provide for an open active site, which is compatible with their activity on long cellooligosaccharides. However, the structure of loop D in PpBglB is different from BGlu1 and the other plant enzymes.

3.8.3 Enzyme active site

Rice BGlu1 has a deep, narrow and straight binding cleft, which is fully circled with negatively charged residues around the deepest and narrowest part, with a broader gate on the top which has positively charged residues along one side (Figure 3.27A). These features are obviously suitable for binding of long oligosaccharide chains, the large number of hydroxyl groups of which can hydrogen bond to the polar residues and water molecules positioned by them. Several aromatic residues are also aligned along the tunnel from the bottom to the entrance way to interact with the

carbon and hydrogen atoms of the glucopyranoside rings of oligosaccharides binding in the active site (Figure 3.27B). The active site geometry is consistent with previous kinetic data that rice BGlu1 could hydrolyze β -1,4-linked oligosaccharides of 2-6 glucose moieties but not β 1,3-linked oligosaccharides with DP more than 3, since their more twisted shape will not fit in the narrow cleft at the inside of the active site.

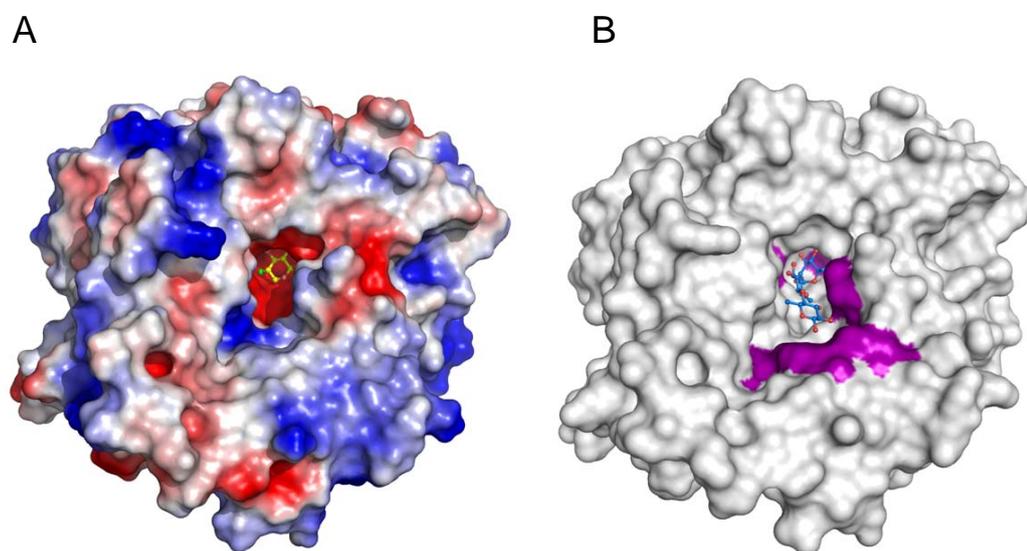


Figure 3.27 Molecular surface of the G2F inhibitor complex and free structure with a docked cellotriose. A) The G2F complex structure is shown as an electrostatic surface with positively charged, negatively charged, and neutral regions colored in blue, red and white, respectively. The G2F inhibitor is bound at the deepest part of the active site pocket. B) The surface structure of the free BGlu1 structure with the aromatic amino acids that can form sugar-binding platforms lining the active site colored in purple, and the docked cellotriose is drawn inside the pocket. The figures were generated with Pymol (DeLano, 1991).

The covalently bound 2-deoxy-2-fluoro-glycoside inhibitor (DNP2FG) complex showed strong electron density, similar to that found at surrounding active site residues. A standard relaxed 4C_1 chair conformation of the sugar ring fit well into the density at subsite -1, which is formed by Q29, H130, Y131, E175, E176, Y315, E386, W433, E440, W441, and W433. The overall conformation of free BGlu1 and its G2F complex are very similar with rmsd of 0.093 Å in molecule A, 0.096 Å in molecule B and 0.121 Å for both molecules, and the few changes that are seen are mainly at the active site pocket. Table 3.6 lists the hydrogen bonds that occur at the active site in both structures, including those between the covalently bound G2F with highly conserved residues in GH1.

Table 3.6 Hydrogen bonding interactions with conserved residues in the free-BGlu1 and G2F complex active sites. Interacting residues and atoms as well as interatomic distances observed in the free BGlu1 and G2F complex structures are listed together for comparison as Free/G2F complex. The asterisk indicates a covalent link between E386 and G2F in the complex structure.

Residue	Atom	Residue	Atom	Distance (Å)
E386	O-ε1	Y315/N313	O/N-δ2	2.60/3.18
	O-ε2	Wat354/G2F	O/C1	2.61/1.43*
		None/Wat368	None/O	None/2.91
E176	O-ε1	Y131/Y131	O/O	2.53/2.18
	O-ε2	N313/N313	N-δ2/N-δ2	2.78/2.70
N175	O-δ1	R86/R86	N/N	2.87/2.85
		None/Wat368	None/O	None/2.78
	N-δ2	Wat354/ G2f	O/F2	3.01//2.99
Q29	O-ε1	Wat181/G2F	O/O3	2.76/2.65
	N-ε2	Glycerol/G2F	O3/O4	3.08/2.90
W441	N-ε1	Wat181/G2F	O/O3	3.21/2.89
E440	O-ε1	Glycerol/G2F	O3/O4	2.61/2.63
	O-ε2	Glycerol/G2F	O2/O6	3.15/2.61

In total, five hydrogen bonds were formed between G2F and amino acids (Figure 3.28). The O3 atom of G2F is hydrogen bonded to W441 Nε1 (2.89 Å), and Q29 Oε1 (2.65 Å), while the O4 atom forms hydrogen bonds with Q29 Nε2 (2.90 Å) and E440 Oε1 (2.63 Å), and O6 interacts with Oε2 of E440 (2.61 Å). A hydrophobic stacking interaction between glucose and the indole ring of W433 was conserved with other GH1 β-glucosidases. The nucleophile, E386, is rotated ~60° around the C_α to

C_β bond with respect to the free enzyme in order to form the glycosyl-enzyme complex. A slight rotation of the 2-fluoroglucose-bound nucleophile was also observed in the *Thermotoga maritima* β-glucosidase, where it was noted to avoid a clash between the nucleophile carboxyl and the 2-fluoro group (Zechel et al., 2003). The distances of the E386 Oε1 and Oε2 to the E176 Oε2 are 3.27 and 4.22 Å, respectively, which are closer than in the free enzyme (4.04 and 4.67 Å). The Oε2 of E386, which is found within hydrogen bonding distance of Y315 (3.32 Å) in the free enzyme is shifted away in the G2F complex, in order for its hydroxyl group to form a hydrogen bond with O5 of G2F instead. The E386 Oε2 instead hydrogen bonds to a water molecule, Wat386 (2.90 Å), which is also hydrogen bonded to N175 Oδ1 (2.78 Å).

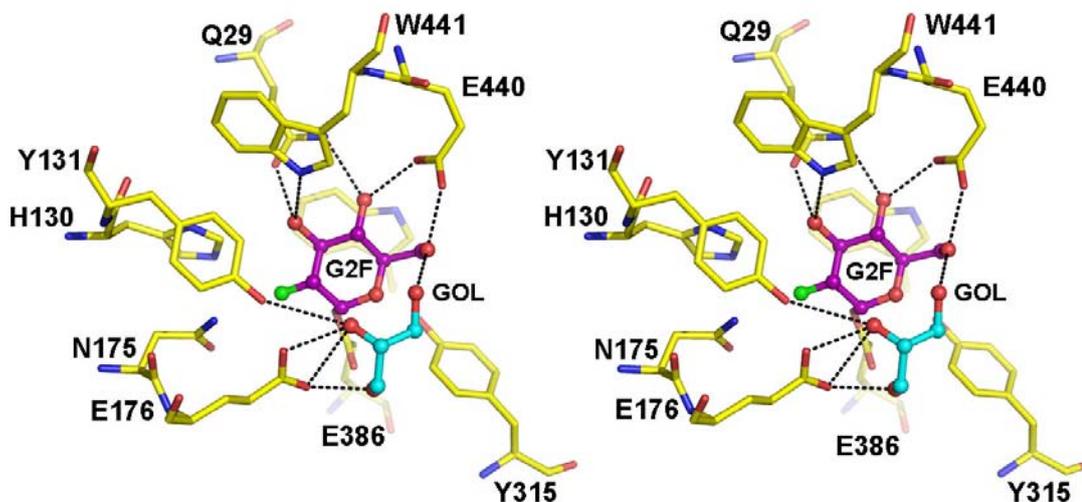


Figure 3.28 Stereo view of protein-ligand interactions in the active site of the BGluc1-G2F complex. The amino acid residues surrounding G2F are presented in stick representation, with carbon, nitrogen, oxygen, and fluoride atoms colored in yellow, blue, red and green, respectively. G2F and glycerol (GOL) are drawn in ball and stick representation in the same colors, except carbon atoms are purple and cyan, respectively. Hydrogen bonds between the protein and the glycone at subsite -1 and glycerol at subsite +1 are drawn as dashed black lines. The figure was produced with Pymol (DeLano, 1991).

A glycerol molecule was observed with its 2 OH at a distance of 3.41 Å from the anomeric carbon, 2.68 Å from O ϵ 1 and 3.36 Å from O ϵ 2 of E176, the catalytic acid/base residue, and 3.42 Å from the Tyr131 OH oxygen, which is twisted slightly compared to its position in the free enzyme (Figures 3.29). The glycerol 1 OH was also within hydrogen bonding distance of both E176 and E386, but was further (3.85 Å) from the anomeric carbon.

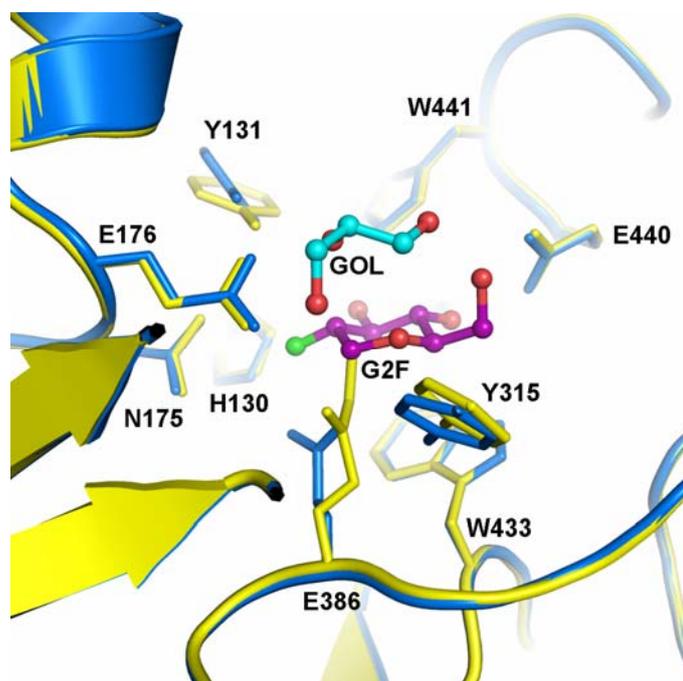


Figure 3.29 Superimposition of subsite -1 of the active site of free rice BGluc1 (colored in blue) and the covalently bound G2F complex (colored in yellow). The covalently bound G2F and solvent glycerol from the G2F complex are drawn as ball and stick and colored purple and cyan, respectively. The figure was generated with Pymol (DeLano, 1991).

3.8.4 Docking studies

Up to now, the crystal structure of a plant GH family 1 β -glucan exo- β -glucosidase in complex with β -1,3- and β -1,4-linked oligosaccharides has not been characterized. To determine the structural basis of substrate specificity of this enzyme, automated docking was used to predict the binding position of oligosaccharides with rice BGluc1. Cellobiose and laminaribiose (1,4- and 1,3- β -linked glucose disaccharides) were chosen for initial computational molecular docking experiments.

As shown in (Figure 3.30A), cellobiose and laminaribiose were directionally docked into the rice BGlu1 active site by the Autodock 3.0.5 program with reasonable distances between the enzyme and substrate at subsite -1 of the crystal structure. The total docking energy of the lowest-energy cellobiose complex -10.86 kcal/mol, while the total docking energy of the lowest-energy complex for laminaribiose is -9.81 kcal/mol. Both docking results show the non-reducing glucose ring in a similar position to that of the G2F in the complex crystal structure. Though the superposition of the nonreducing sugars on the G2F suggests both docking results are reasonable, it is not clear what differences in the binding of cellobiose and laminaribiose make the latter a much better substrate (Opassiri et al., 2003; 2004, Figure 3.30A).

Previously, subsite mapping studies showed that binding of cellooligosaccharides appears to be strongest at subsite +2, and less strong at subsites -1 and +1 (Opassiri et al., 2004). Therefore, cellobiose might be bound more tightly at subsite +2 and subsite +3, where it would block other substrate molecules from entering the active site, than at the productive position at subsites -1 and +1. However, the docking studies could not identify the higher affinity binding site for cellobiose, but only showed the apparently productive position.

Extra units of β -1,4-linked glucose monomer should contribute to productive binding of substrate in the active site of rice BGlu1, since its catalytic efficiency for hydrolysis of cello-oligosaccharide increases with increasing chain length (Opassiri et al., 2004). However, when cellotriose and longer β -1,4-linked oligomers with standard relaxed chair conformations (1C_4) at the non-reducing end were docked to active site of rice BGlu1, the ligands were found to dock in the wrong orientation, with the reducing end in the -1 subsite rather than the nonreducing end.

The structure of sorghum dhurrinase E189D (acid/base mutant) bound to dhurrin shows the glucose residue in the 1S_3 skew boat conformation in the Michaelis complex (Verdoucq et al., 2004). Molecular mechanics simulations suggest this adaptation of sugar conformation at subsite -1 from relaxed 1C_4 chair to 1S_3 skew boat conformation is likely to occur upon binding to the enzyme (Biarnes et al., 2006). A chair conformation is then recovered after release of the aglycone. Thus, celotriose with a skew boat conformation at the nonreducing end was docked to the active site, and it docked with the nonreducing glucose residue at subsite -1 with a lowest total docking energy of -4.77 kcal/mol, thereby indicating the likely position of binding an additional glucose residue at subsite +2 (Figure 3.30B).

Surprisingly few hydrogen bonds were observed between the enzyme and the docked celotriose at subsites +1 and +2, while subsite -1 was strongly stabilized by the five hydrogen bonds seen in the BGlu1-G2F complex, and additional interactions at O2, not seen due to the F in that position (Zechel et al., 2003). At subsite +1, Glc2: O3 was within hydrogen bonding distance of E440 O ϵ 2 and W441 N ϵ 1. There were 3 hydrogen bonding interactions between Glc3 and residues at subsite +2, including O6 with E176 O ϵ 2, and N245 N δ 2 with both Glc3 O2, and Glc3 O3. Some polar amino acids, such as, R178, Q187, N190, T316, and H267 line the wall of the aglycone binding site, but they are beyond direct hydrogen bonding distance. Conversely, several nonpolar interactions were found with amino acids closely surrounding these sugar residues, including I179, L182, L183, L442 and D243 (in which the fully charged side chain is turned away from substrate). Moreover, both Glc2 and Glc3 at subsites +1 and +2 had stacking interactions with the aromatic ring of W358, which is similar to the stacking interactions of the first sugar residue with W433 at subsite -1.

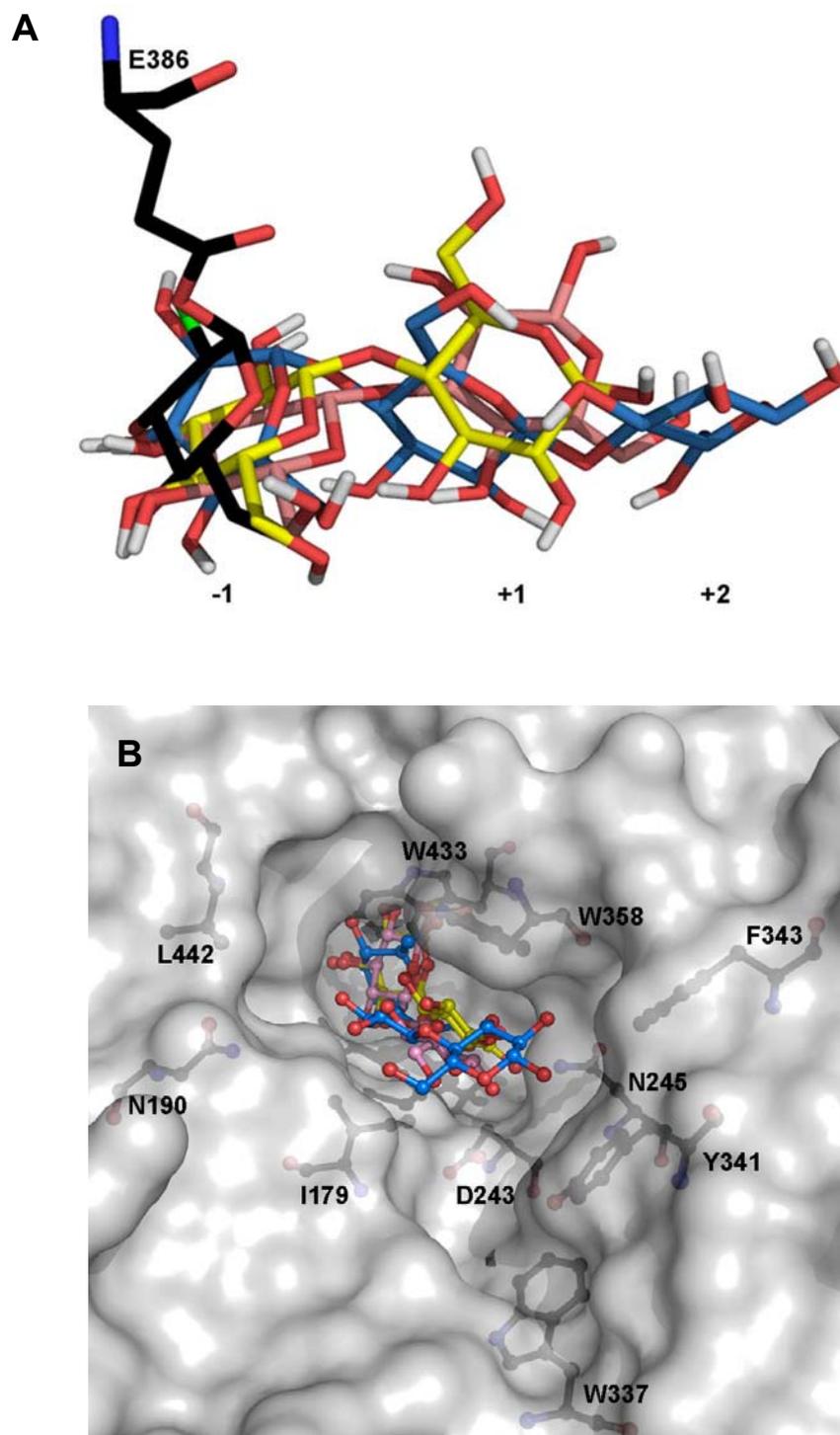


Figure 3.30 Crystal structure of free BGlu1 with cellobiose, laminaribiose, and cellotriose from docking. (A) Molecular surface structure of the free rice BGlu1 enzyme with overlaid structures of cellobiose (yellow), laminaribiose (pink),

cellotriose (blue) from automated docking. Amino acid residues at subsite -1 and putative aglycone binding residues at subsite +1 and +2, and likely positions of more distal subsites are shown as stick diagrams. (B) Overlay structure of covalently bound G2F inhibitor which occupies in its 4C_1 chair conformation only subsite -1 from G2F inhibitor complex structure, with cellobiose, laminaribiose, cellotriose from automated docking. Carbon atoms are black, yellow, pink, and blue in G2F, cellobiose, laminaribiose, and cellotriose, respectively. All oxygens and hydrogens are colored with red and white for all ligands. Numbering of binding subsites is indicated by -1, +1, and +2 at the bottom of the picture. The figures were produced with Pymol (DeLano, 1991).

3.9 Determination of the structure of the BGlu1 E176Q mutant and its complex with cellopentaose

3.9.1 Free E176Q mutant structure

The structure of the free E176Q mutant and its complex with cellopentaose was not significantly different from the wild type structure. The crystal of free E176Q mutant was diffracted to 1.35 Å, and final model was refined at 1.37 Å gave R_{factor} and R_{free} of 16.6 and 18.7%, respectively (Table 3.5). Actually, laminaribiose was soaked into the crystal overnight before data collection, but no clear density of that substrate was observed. Only some density of half a glucose ring appeared at subsite -1, and ambiguous density was also seen at subsite +1, so the laminaribiose was not built into that density, but it was instead filled with water molecules. However, an omit map showed better density, especially at subsite +1, though the density of a whole glucose

residue at subsite -1 was not observed (Figure 3.31). The overlaid structure of the single glucosyl residue at the nonreducing end of cellopentaose (in the E176Q with cellopentaose structure) indicates that the observed density at the active site of free E176Q probably belongs to laminaribiose, but low occupancy was obtained.

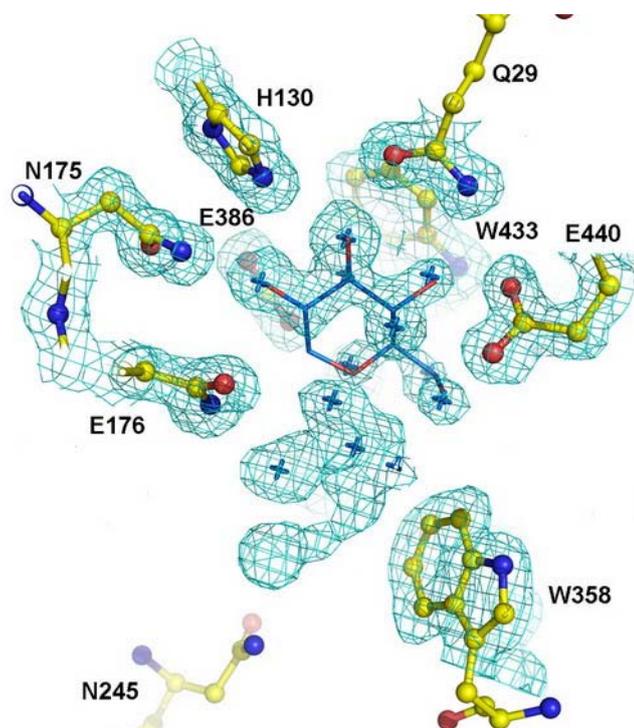


Figure 3.31 Electron density of the $2|F_{\text{obs}}| - |F_{\text{calc}}|$ map (contoured at 1σ) at the active site of E176Q structure (soaked with laminaribiose). The glucose at the nonreducing end of cellopentaose retrieved from the structure of E176Q with cellopentaose was superimposed on the density (blue stick). The blue crosses indicate the positions of water molecules in free E176Q mutant structure model. The figure was produced with Pymol (DeLano, 1991).

3.9.2 Two conformational changes of loop C

As mentioned above, the longest loop, loop C, has the highest variation in the GH family 1 structures and forms a part of the aglycone binding pocket. Interestingly, we observed two conformational changes in this loop. Both the $2|F_{\text{obs}}|-|F_{\text{calc}}|$ and omit maps showed two alternative conformations of three amino acid residues; Y341, V342, and F343 (Figure 3.32). This phenomenon was observed in both molecules A and B that are found in the asymmetric unit.

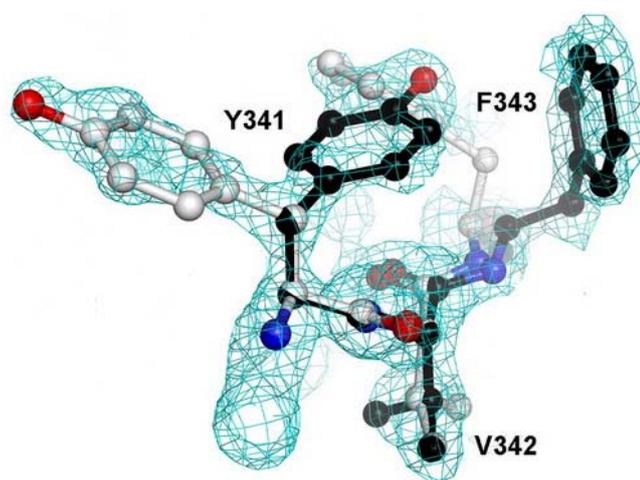


Figure 3.32 Electron density map of two conformations of loops C in the free E176Q mutant structure. The electron density was from an annealed omit map, contoured at the 1σ level. The black and white structures show alternative conformations of the 3 residues, Y341, V342, and F343, located in loop C. The figure was produced with Pymol (DeLano, 1991).

3.9.3 E176Q mutant with its complex of cellopentaose

In early stages of refinement, strong positive density in the $|F_{\text{obs}}| - |F_{\text{calc}}|$ map contoured at 3σ in the +1 to +4 subsites region clearly showed the presence of the sugar molecule bound to the active site, while subsite -1 appeared to have the density only half of a glucose moiety. The density of the first glucose improved along the steps of refinement with addition of water molecules. The latest model gave an R_{factor} and R_{free} of 18.1% and 21.1%, respectively, at 1.80 Å resolution (Table 3.5). The density of cellopentaose at molecule B was better than molecule A, and the 1,4 β -glycosidic linkage was observed between Glc1 and Glc2 in molecule B (Figure 3.33). Each glucosyl residue was numbered Glc1 to Glc5 from the nonreducing end to the reducing end, corresponding to residues that occupied subsites -1, +1, +2, +3, and +4, respectively. All glucose residues were fit to the density with a standard chair conformation (4C_1), though the SbDhr1 with dhurrin structure reported a 1S_3 skew boat conformation of Glc1 (Verdoucq et al., 2004). However, the skew boat conformation of Glc1 was not built into the model because the densities were ambiguous and not sufficient to justify an exact conformation. Glc5 had notably weaker density than the three internal glucosyl residues (Glc2, Glc3, and Glc4), indicating that this residue was more disordered, likely due to the fact that it appeared to have few direct interactions with the protein.

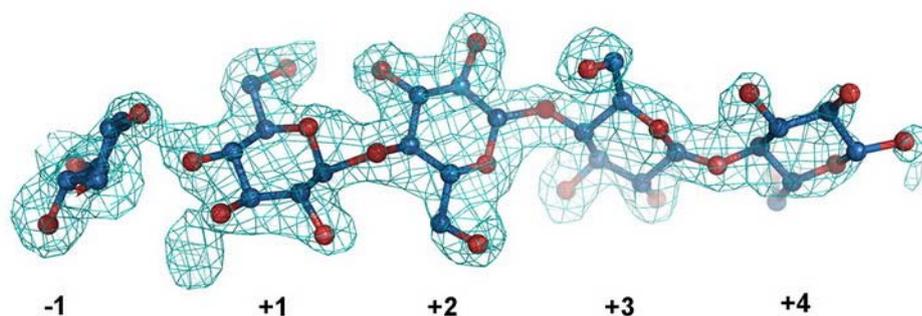


Figure 3.33 The conformation of cellopentaose bound to E176Q BGluc1 and its electron density. The nonreducing end of cellopentaose is at subsite -1 and reducing end at subsite +4. Relatively weak electron density is seen at subsites -1 and +4, which are the residues cleaved from the oligosaccharide and most distal residue, respectively. The figure was produced with Pymol (DeLano, 1991).

3.9.4 Cellopentaose binding and interactions involved in the extended binding site

The position of cellotriose that was obtained from automated docking was similar to that of the first 3 residues in the crystal structure of cellopentaose bound to the E176Q mutant (Figure 3.34).

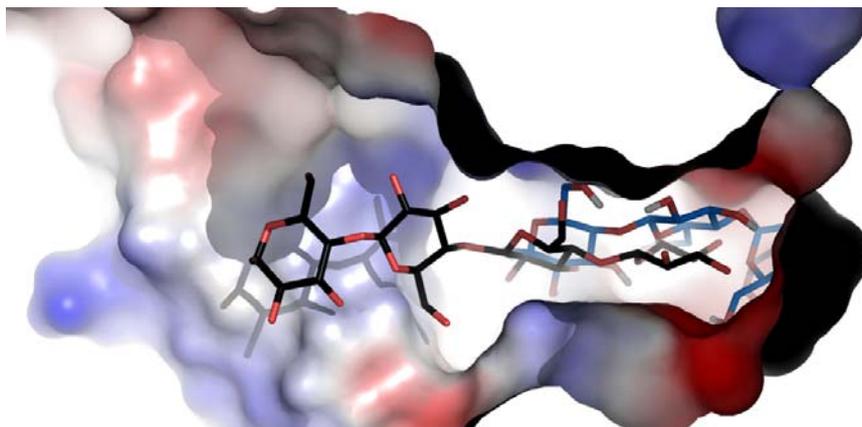


Figure 3.34 Comparison of cellopentaose (black) alignment from the crystal structure of the E176Q mutant and celltriose (blue) from automated docking. The figure was produced with Pymol (DeLano, 1991).

Both hydrophobic and hydrogen bond interactions appeared to be involved in cellopentaose binding. Interestingly, the aromatic side chain of Y341, which interacted with the hydrophobic faces of Glc4 and Glc5 in the crystal complex with cellopentaose, had two conformations that were observed in the free E176Q structure. As shown in Figure 3.35, the three aromatic side chains of W433, W358, and Y341 appeared to serve as platforms onto which the faces of individual glucose residues stack by hydrophobic and dispersive interactions. In addition, hydrophobic interactions via nonpolar amino acid side chains were extensively found at subsites +1 and +2, which included I179, L182, and L183. Aside from the hydrogen bond interactions at subsite -1, one direct hydrogen bond was seen at subsite +2, N245 with O2-Glc3, as was also obtained in the docking results. Another hydrogen bond was observed a suitable subsite +3, where Q187 bound with O3-Glc4 via water molecules. Additional water molecules may be added at promising densities around the substrate,

so several hydrogen bonds mediated by water molecules should be found, when this model is complete.

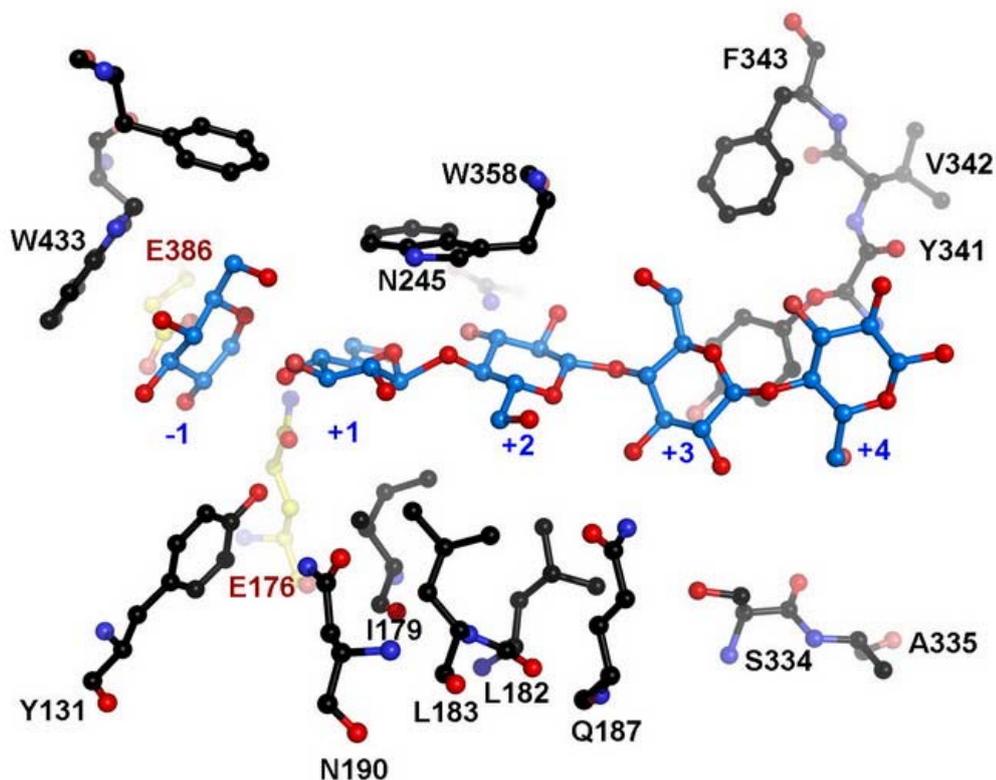


Figure 3.35 The subsite positions indicated by cellopentaose binding in the rice BGl1 active site. The cellopentaose molecule (blue) covers subsites -1 to +4. The figure was produced with Pymol (DeLano, 1991).

3.10 Studies of substrate specificity by mutation at putative aglycone binding residues.

GH1 enzymes hydrolyze substrates with high variation in their aglycone, so the subsite occupied by the glycone part is remarkable similar in the family, while the aglycone binding pocket is different for each enzyme. To explain the substrate specificity, the putative substrate binding residues at the aglycone binding pocket were carefully inspected.

3.10.1 Rice BGlu1 and ZmGlu1

The substrate specificity of maize Glu1 β -glucosidase is well characterized, and has been structurally investigated with crystal structures of wild type and inactive mutant enzymes in complex with several ligands (Czjzek et al., 2000; 2001). Four amino acid residues were observed to be important for substrate recognition, W378, which is strictly conserved in *O*-glycosidases, and three phenylalanine residues, F198, F205, and F466, which are located on the opposite side of the active site cleft from W378 and vary in GH1. A comparison of the rice BGlu1 and maize β -glucosidase active sites is shown in Figures 3.36 and 3.37, DIMBOAGlc and celotriose are stacked onto the hydrophobic face of the conserved tryptophan, which is in different orientations in the two structures. The aromatic rings of the three opposing phenylalanine residues in maize are replaced by L183, N190, and L442, respectively, in rice BGlu1.

The superimposition of the two substrate complex structures clearly showed that the reducing end of cellotriose occupied a different aglycone pocket site from DIMBOAGlc. The aglycone part of DIMBOAGlc was bound in a pocket adjacent to the three aromatic phenylalanine residues, while the nonreducing end of the more hydrophilic cellotriose appears to bind another site in order to form one hydrogen bond with the side chain of N245. The highly conserved tryptophan (W358 in rice BGlu1 and W378 in maize) forms a platform for binding both the second and third glucoses and DIMBOA, but it is tilted differently in the two enzymes to allow suitable substrate alignment with other interacting sidechains (Figure 3.37).

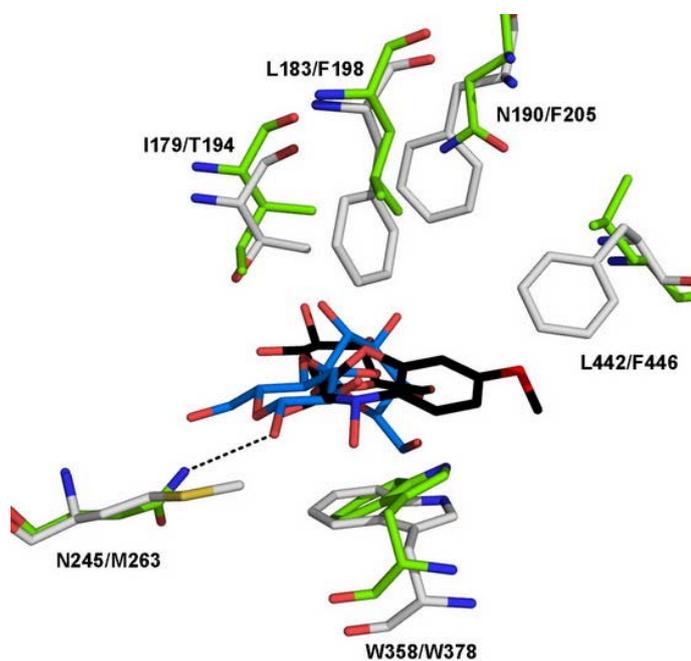


Figure 3.36 Superimposition of rice BGlu1 (green carbons) with ZMGlu1 maize β -glucosidase (gray carbons) in complex with the natural substrate DIMBOAGlc (black) (PDB code; 1E56). Cellotriose is shown docked into the active site of free rice BGlu1 with its carbons colored blue. The dashed line shows a hydrogen bond between cellotriose and N245 in the rice BGlu1 structure. The figure was produced with Pymol (DeLano, 1991).

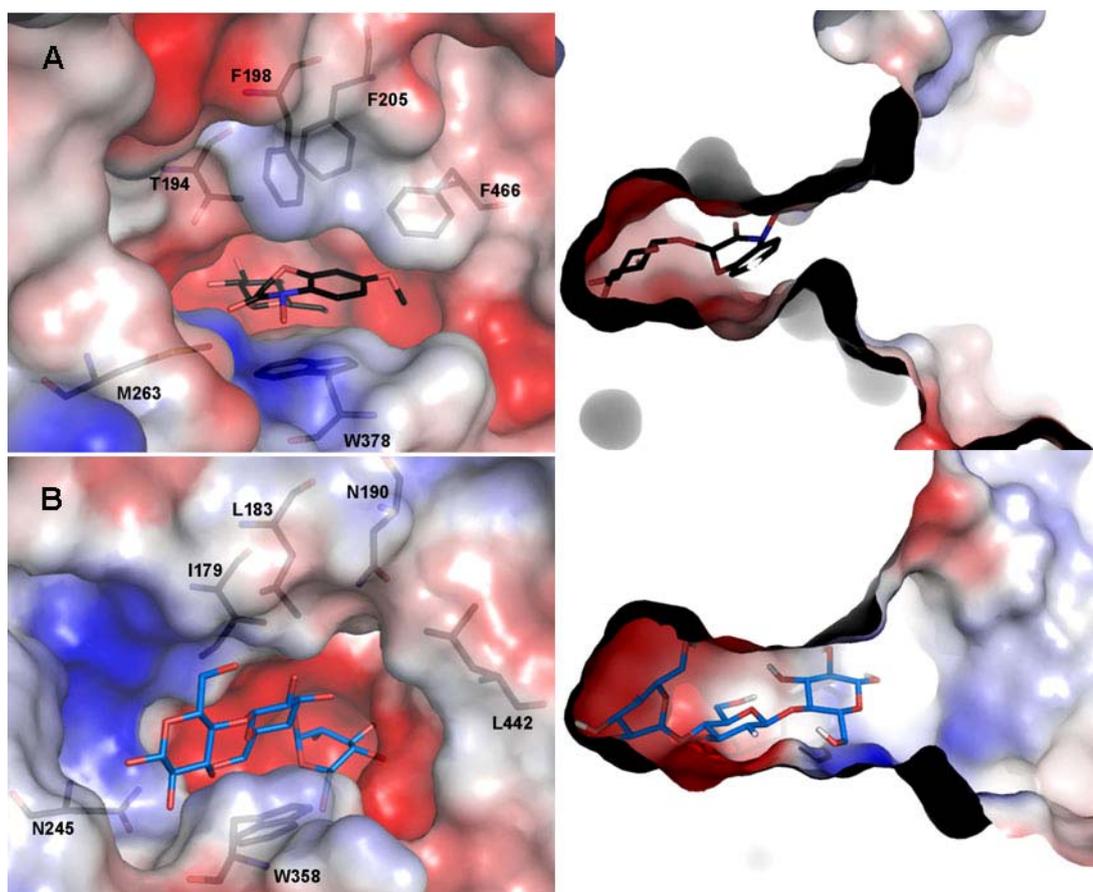


Figure 3.37 Differences in the active site geometry of rice BGluc1 (A) and ZMGluc1 (B) with bound cellobiose and DIMBOAGlc, respectively. The amino acid residues drawn as stick structures underneath the electrostatic surface formed the shape of active site. The figure was produced with Pymol (DeLano, 1991).

3.10.2 Analysis of rice BGluc1 and barley BGQ60 β -glucosidase substrate binding residues

Of the well characterized glycosyl hydrolase family 1 enzymes, rice BGluc1 is most similar to barley BGQ60 β -glucosidase (Hrmova et al., 1996; 1998; Opassiri et al., 2003; 2004), and both hydrolyze short β -1,3- and longer β -1,4-oligosaccharides.

However, while the barley enzyme prefers cellobiose to cellotriose, rice BGlu1 shows the opposite preference, and differences in the affinities of their +1 and +2 sites may explain their differences in activities toward oligosaccharides (Hrmova et al., 1998; Opassiri et al., 2004). The docking of cellotriose into the rice BGlu1 active site indicated that the amino acid residues involved in the hydrogen bonding and hydrophobic interactions with cellotriose were mainly located in loop B and loop C and none of them were conserved among GH1 family, so a comparison of these residues in rice and barley should be informative regarding substrate binding.

The rice BGlu1 residues Y131 (subsite +1), R178 (+2), L183 (+1), D243 (+1), H267 (+2), and W337 (+3) might be crucial for substrate recognition of oligosaccharides, since the corresponding residues are the same in the barley BGQ60 enzyme, but not in other β -glucosidases that do not hydrolyze oligosaccharides well. However, the equivalent residues in PpBglB (W123, Y169, L196, N223, L247, and S410) were almost completely different, indicating the bacterial β -glucosidase has evolved a different set of interactions for binding oligosaccharides. The residues I179, L182, N190, N245, and L442 in rice BGlu1 are replaced by V184, A187, H195, V250, and R447, respectively, in barley, so they might contribute to substrate specificity differences between the rice and barley enzymes. In order to probe the basis of differential binding of cellobiose and cellotriose between rice BGlu1 and barley BGQ60 β -glucosidases, four residues found in the putative aglycone binding region (I179, N190, and L442 at subsite +1 and N245 at subsite +2) of rice BGlu1 and different from those at the same positions in the barley enzyme were mutated.

3.10.3 Kinetic analysis of substrate binding pocket mutants

Contrary to expectations, all mutants show the same pattern for all substrates as shown in Table 3.7. The N190H mutation had the greatest effect and decreased the K_m and increased the k_{cat} for all substrates (*p*NPG, cellobiose and cellotriose), while the L442R mutant enzyme retained the same activities as wild type. The I179V and N245V mutant enzymes exhibited almost the same differences from wild type with higher values of K_m and k_{cat} for almost all substrates, except that N245V showed a similar k_{cat} for *p*NPG as wild type.

In fact, the mutations were expected to change the activity of rice BGl1 to be more similar to that of barley β -glucosidase with improvement in the ratio of $k_{cat}/K_m(\text{cellobiose})$ to $k_{cat}/K_m(\text{cellotriose})$ and a decreased rate of hydrolysis of *p*NPG, as previously reported and shown in Table 3.7 (Hrmova et al., 1996; 1998). None of the mutants had kinetics for *p*NPG hydrolysis similar to those of the barley enzyme. Although, the I179V and N245V enzymes showed approximately 5-fold and 10-fold increases in K_m , respectively, their k_{cat} remained high like wild type rice BGl1. In contrast, a slight improvement in *p*NPG hydrolysis was seen in the N190H mutant with about a 1.5-fold decrease in K_m , and a 1.4-fold increase k_{cat} .

The expected improvement in cellobiase activity was not detectable for any of the mutants (the k_{cat}/K_m of the mutants was very low compared to the barley enzyme). The N245V mutation gave a 6-fold increase in k_{cat} to a value similar to barley, but the K_m also increased over 3-fold (while that of barley is 8-fold lower than wild type BGl1), so the catalytic efficiency (k_{cat}/K_m) was only 2-fold higher than wild type. Since the N245V mutant also has a high K_m for cellotriose and its k_{cat}/K_m for cellotriose is 15-fold lower than wild type BGl1 and even lower than barley, its ratio

of k_{cat}/K_m for cellobiose and cellotriose is 0.14, much higher than the wild type ratio of 0.005. However, this is still well below the cellobiose/cellotriose hydrolysis efficiency ratio of 2.15 for the barley enzyme and comes at a high cost of substrate affinity for both oligosaccharides. The effects of I179V were similar, and its ratio of k_{cat}/K_m for hydrolysis of cellobiose/cellotriose was about 0.0167, about 3-fold higher than wild type. On the other hand, the N190H mutant showed a 1.5-fold decrease in K_m and almost 2-fold increase in k_{cat} for cellobiose, while it also increased the efficiency of hydrolysis of cellotriose, but not as much. Therefore, compared to wild type its ratio of k_{cat}/K_m for cellobiose to cellotriose increased slightly to 0.008.

Table 3.7 Kinetic parameters of wild type rice BGl1, putative aglycone binding residue mutants and barley β -glucosidase isozyme II BGQ60 for hydrolysis of *p*NPG, cellobiose, and cellotriose.

Substrate	Kinetic parameters	Wild type	L442R	N190H	I179V	N245V	BGQ60 ^a
<i>p</i> NPG	K_m (mM)	0.201 \pm 0.006	0.201 \pm 0.008	0.139 \pm 0.005	0.95 \pm 0.08	2.10 \pm 0.15	0.50 ^a
	k_{cat} (s ⁻¹)	4.69 \pm 0.30	4.57 \pm 0.22	6.76 \pm 0.32	6.08 \pm 0.46	4.70 \pm 0.34	0.50 ^a
	k_{cat} / K_m (s ⁻¹ mM ⁻¹)	23.33	22.74	48.63	6.40	2.24	1 ^a
Cellobiose	K_m (mM)	22.0 \pm 2.6	21.9 \pm 1.9	14.3 \pm 0.6	61.1 \pm 2.6	74.4 \pm 6.4	2.67 \pm 0.19 ^b
	k_{cat} (s ⁻¹)	1.16 \pm 0.07	1.12 \pm 0.03	2.09 \pm 0.06	3.42 \pm 0.29	7.17 \pm 0.11	11.58 \pm 0.63 ^b
	k_{cat} / K_m (s ⁻¹ mM ⁻¹)	0.05	0.05	0.15	0.06	0.10	4.34 ^b
Cellotriose	K_m (mM)	0.22 \pm 0.02	0.23 \pm 0.01	0.160 \pm 0.004	1.08 \pm 0.06	9.25 \pm 0.83	0.97 \pm 0.06 ^b
	k_{cat} (s ⁻¹)	2.35 \pm 0.29	2.58 \pm 0.14	3.05 \pm 0.29	3.30 \pm 0.30	6.58 \pm 0.25	1.95 \pm 0.12 ^b
	k_{cat} / K_m (s ⁻¹ mM ⁻¹)	10.7	11.2	19.1	3.06	0.71	2.01 ^b

^a kinetic data obtained from Hrmova et al., 1996, and ^b from Hrmova et al., 1998.

3.11 Studies of the catalytic mechanism by mutation at the acid/base and nucleophile residues.

3.11.1 β -glucosidase activity for *p*NPG hydrolysis of acid/base and nucleophile mutants.

Replacement of the putative catalytic nucleophilic glutamate (E386) with glutamine (E386Q) or aspartate (E386D) caused drops in activity by approximately 60,000 and 3,000 fold, respectively, supporting its identification as the nucleophile. In addition, changing the putative catalytic acid/base (E176) to glutamine (E176Q), aspartate (E176D) or alanine (E176A), caused approximately 50-, 180-, and 360-fold drops in *p*NPG hydrolysis activity in MES buffer, as shown in Table 3.8. This confirmed that E386, and E176 are important to enzyme catalysis, and likely act as the nucleophile and acid/base residues, respectively, for rice BGlu1. No significant activity for hydrolysis of oligosaccharide substrates (cellobiose and cellotriose) by nucleophile and acid/base mutants BGlu1.

Table 3.8 Effects of mutations of the putative nucleophile and acid/base residues on *p*NPG hydrolysis.

Group	Mutant	Specific activity ($\mu\text{mol}/\text{min}/\text{mg}$)	Fold decrease in specific activity
Wild type	None	3.7	1
	E176Q	0.1	50
Acid/base mutant ^a	E176D	0.01	320
	E176A	0.009	420
Nucleophile mutant ^b	E386Q	0.00006	60,000
	E386D	0.001	3,000

^a The activity was assayed in 50 mM MES buffer, pH 5.0 to prevent rescue activity.

^b The activity was assayed in 50 mM acetate buffer, pH 5.0.

3.11.2 Effect of pH on β -glucosidase activity

The pH profiles of wild type and mutant BGlu1 β -glucosidases for *p*NPG hydrolysis were determined at 30°C in series of overlapping buffers: acetate (pH 4.0-5.5), MES (pH 5.5-6.5), phosphate (pH 6.5-7.5), and Tris-HCl (pH 7.5-9.0). The wild type enzyme showed maximum activity at pH 5.0 for overlapping buffer and no activity in Tris-HCl, pH 8.0 (Figure 3.38A). Enzymes with mutations at putative substrate binding residues showed pH profiles similar to wild type BGlu1, and also had optimal pH of 5.0 (Figure 3.38A). Enzyme activity was hardly detectable for the acid/base and nucleophile mutants, except for E176Q which showed rescued activity in NaOAc buffer at pH 4.5 to 5.5, but not in MES buffer at pH 5.5 (Figure 3.38B). Like E176Q, the E176A mutant was partially rescued by acetate in the same range, but with lower activity.

Universal pH buffer (a series mixtures of solution A: 0.2 M boric acid + 0.05 M citric acid and solution B: 0.05 M tri-sodium phosphate) was used in place of the

series of overlapping pH buffers to eliminate the difference of chemical components at each pH point. As shown in Figure 3.39A, the pH profiles of the wild type enzyme were not remarkably different for the two systems, pH optimum was around pH 5.0-5.5 in the universal buffer. Conversely, the E176Q mutant had barely detectable activity for all pH values in universal buffer (Figure 3.39B). This suggested that the activity of this mutant was restored by acetate acting as an external acid/base catalytic group or nucleophile.

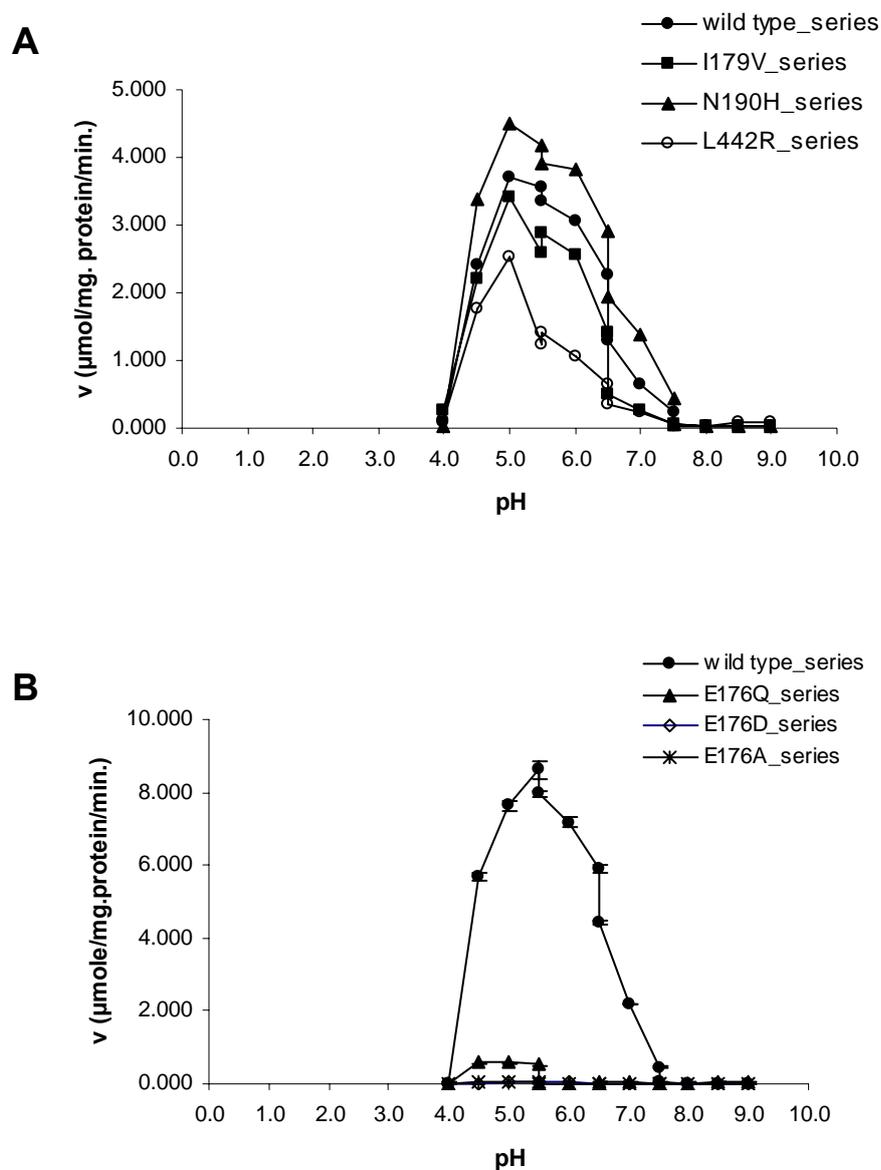


Figure 3.38 pH profiles of wild type BGlul and BGlul with mutations at (A) substrate binding and (B) acid/base residues with a series of overlapping pH buffers (acetate pH 4.0-5.5, MES pH 5.5-6.5, phosphate pH 6.5-7.5, and Tris-HCl pH 7.5-9.0). Series in the legend is short for “series of overlapping buffers”.

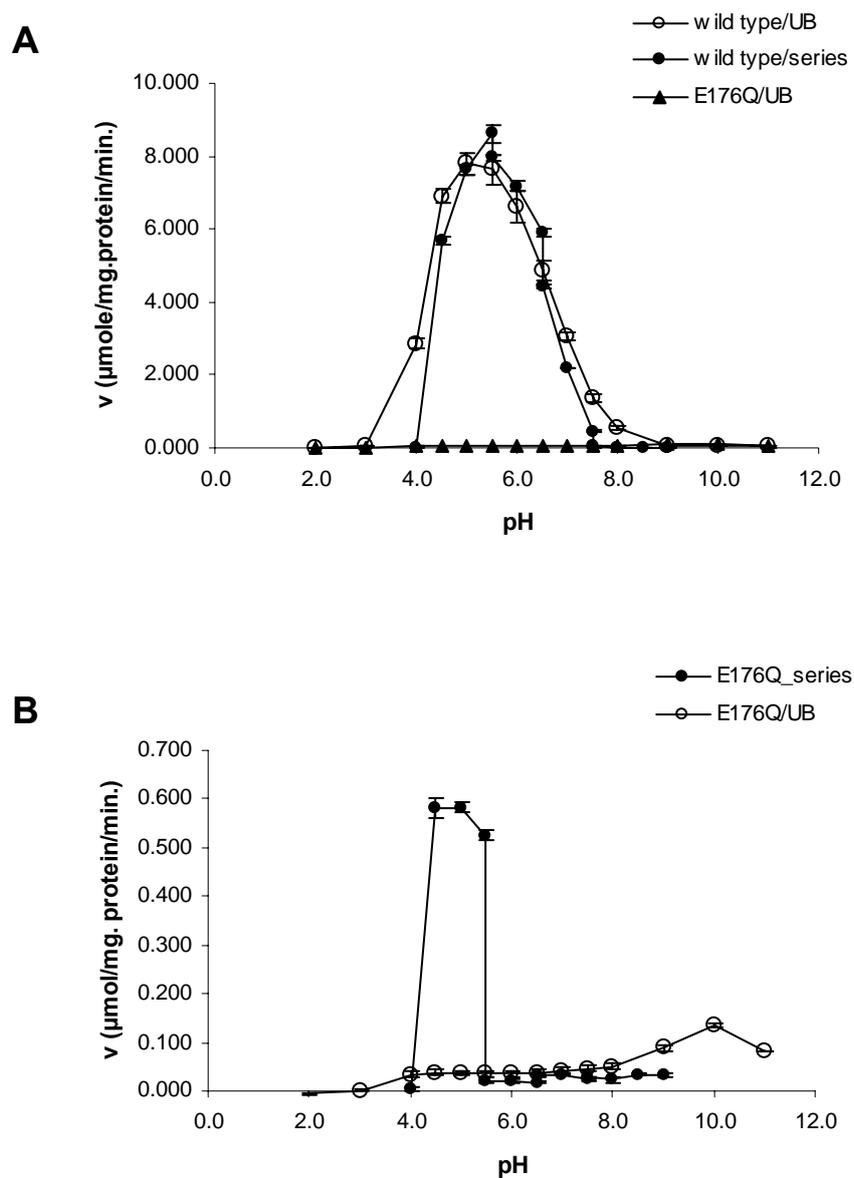


Figure 3.39 Comparison of two different buffer systems: a series of overlapping pH buffers (series) and universal pH buffer (UB) for hydrolysis of *p*NPG by wild type and E176Q BGLu1 enzymes. A shows a comparison of wild type and E176Q mutant BGLu1 in universal buffer vs. wild type in the series of buffers, while B shows a comparison on the activity of the E176Q mutant in the series of buffers vs. the universal buffer.

3.11.3 The ability of various anionic nucleophiles to rescue the activity of acid/base mutants

According to Wang et al., 1995, several anions, such as azide, carboxylates and thiolates restored the enzyme activity of an acid/base mutant of *Agrobacterium* β -glucosidase with different rate enhancements. A variety of anionic nucleophiles were also tested with the E176Q, E176A and E176D mutants, including azide, cyanate, fluoride, carboxylates (acetate, formate, TFA, citrate) and also ascorbate, which acts to replace the acid/base residue for myrosinase (Burmeister et al., 2000). The highest activities for both the E176Q and E176A mutants were obtained with relative order of ascorbate > azide > acetate (Figure 3.40A). On the other hand, formate, citrate, TFA, fluoride, and cyanate yielded very low rescued activity and modest activities were found with formate and citrate for the E176A mutant (higher than E176Q). Conversely, the activity of E176D and wild type enzymes were not significantly affected by any kind of the anionic nucleophiles at the concentration tested (50 mM, Figure 3.40B).

In addition, MES and universal pH buffers gave small differences in enzyme activities, with the latter buffer giving slightly lower rates. Further characterization of pH dependence and rate enhancement was performed with azide, ascorbate, and acetate, because they yielded the highest rescued activity.

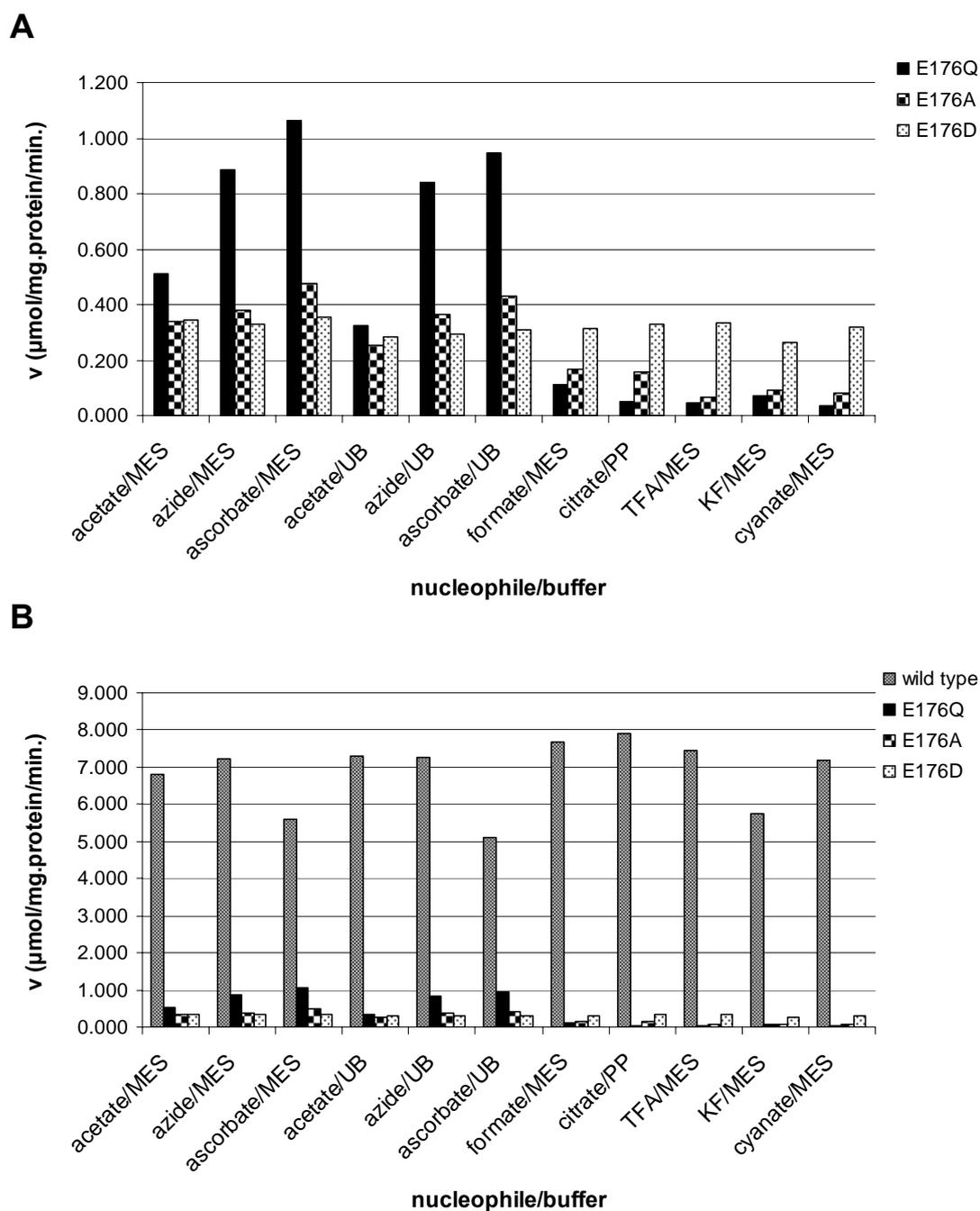


Figure 3.40 Effects of various anionic nucleophiles on acid/base mutants and wild type BGlu1. In (A), E176Q, E176D, and E176A are compared, and in (B), wild type enzyme is compared to them. Each reaction contained on anionic nucleophile at 50 mM concentration in 50 mM MES buffer, pH 5.0, or universal buffer (UB) or phosphate buffer (PP).

3.11.4 pH dependence in the presence of azide and ascorbate

The pH profile of wild type BGlu1 in the presence of 50 mM nucleophile (azide and ascorbate) showed strong pH dependence with pK_a values of approximately 4.5 and 6.5 for the two ionizable groups of the nucleophile and acid/base residues, respectively (Figures 3.41A and 3.42A). These pK_a values agree well with those previously reported using the ^{13}C -NMR titrations spectroscopy in *Bacillus circulans* xylanase (McIntosh et al., 1996). In the presence of azide, a similar profile as wild type was obtained for the E176D mutant, since two ionizable groups remained (Figure 3.41B), though they are further apart. On the other hand, the replacement with non-ionizable groups in the E176Q and E176A mutants completely abolished the pH dependence over the range pH 5-10. However, the dependence of the k_{cat} upon ionization of the nucleophile was still observed at pH 4.5 (Figure 3.41A). A similar profile was obtained when ascorbate was added as the nucleophile (Figures 3.42A, and B). The wild type enzyme had greatly reduced activity at pH 8-9, while enzyme activity for the E176Q and E176A mutants in the presence of azide was greatly reduced at pH ≥ 11 . However, the activity of wild type and mutants remained in the presence of ascorbate at pH ≥ 10 .

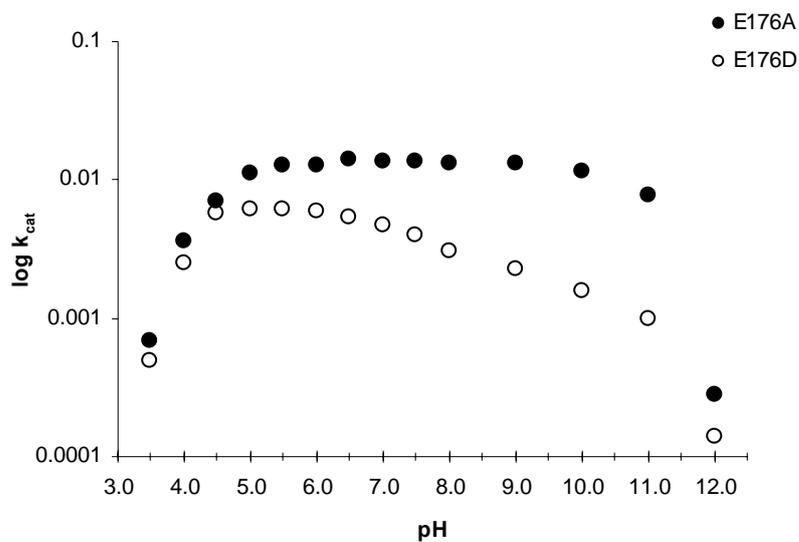
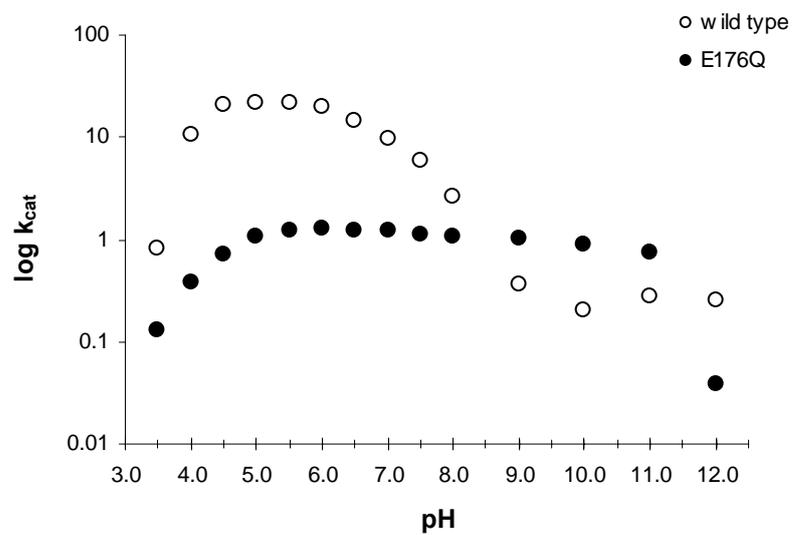


Figure 3.41 Comparison of the pH dependence of the wild type, E176Q, E176A, and E176D enzymes in 50 mM azide in universal pH buffer.

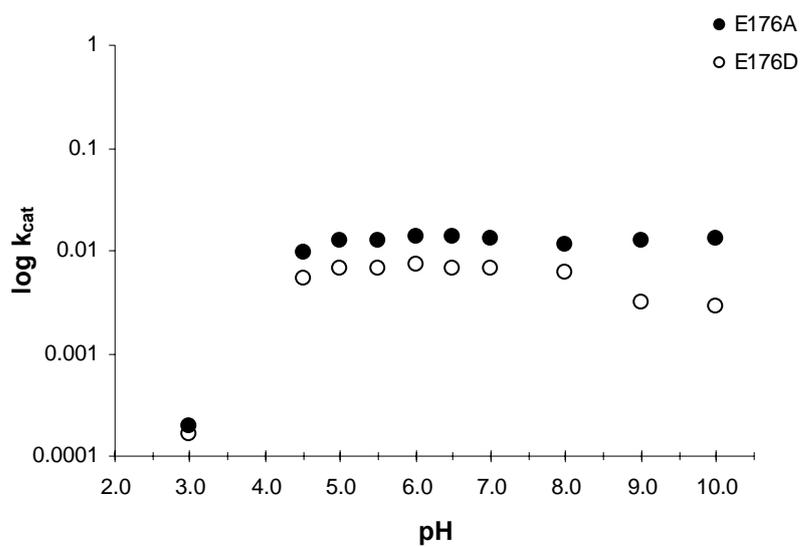
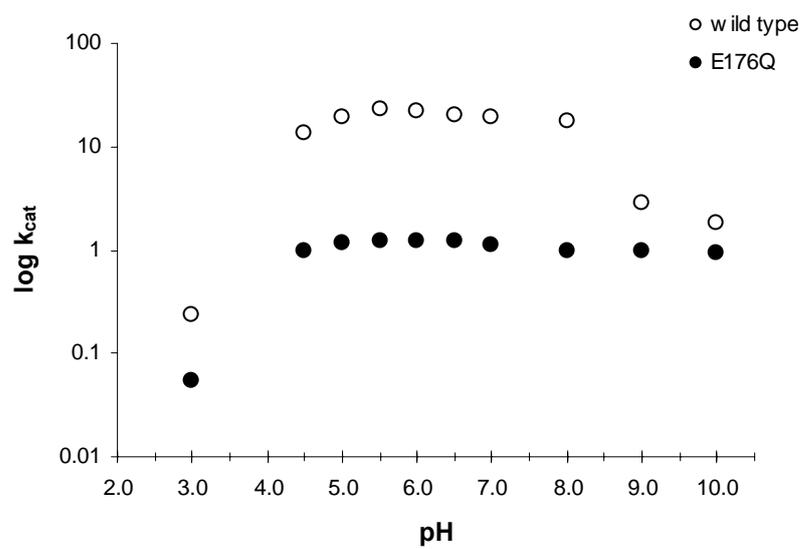


Figure 3.42 Comparison of pH dependence of the wild type, E176Q, E176A, and E176D enzymes in 50 mM ascorbate in universal pH buffer.

3.11.5 Rate enhancement by external nucleophile anions

An increase in the rate of hydrolysis of *p*NPG by the E176Q mutant was found with increasing nucleophile concentrations (ascorbate, azide, and acetate), but inhibition appeared at high concentrations of nucleophiles (Figure 3.43). The highest increase in rate enhancement was found in the presence of 5-40 mM ascorbate and the rate slightly increased from 40 mM until 160 mM, then dropped significantly at 320 mM. Azide had a little bit lower rate than ascorbate from 5-160 mM, however, the rate still increased with increasing azide concentration until 640 mM. A drop in the rate due to inhibition by azide was first found at 1280 mM and enzyme completely lost activity at 2560 mM. The same pattern with an approximately 4-fold lower rate enhancement was obtained for acetate. For wild type enzyme, low nucleophile concentrations could inhibit activity (data was not shown).

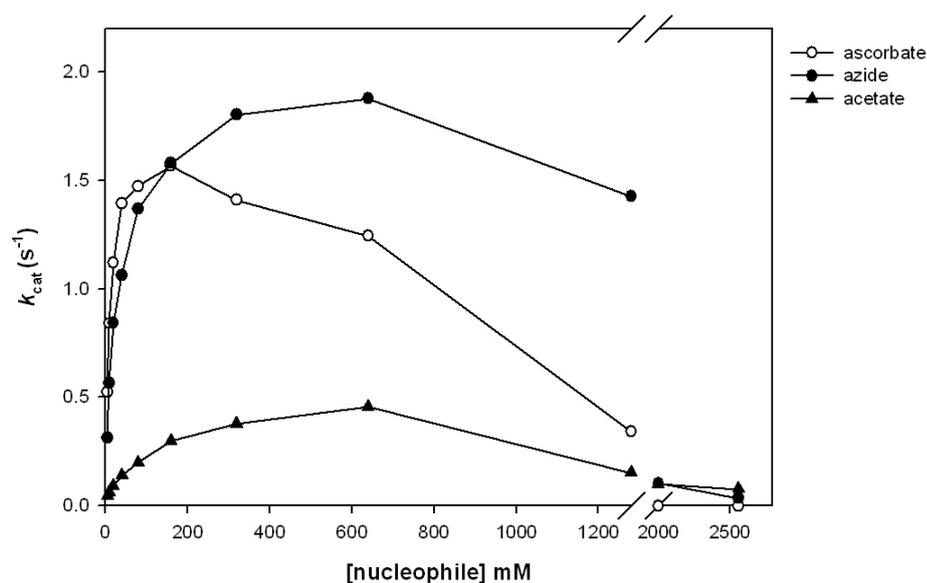


Figure 3.43 The concentration dependence of azide, acetate, and ascorbate for *p*NPG hydrolysis by the BGLu1 E176Q mutant.

Table 3.9 Apparent K_m , k_{cat} , and k_{cat}/K_m of E176Q mutant BGlu1 for *p*NPG

hydrolysis in the presence and absence of various concentrations of acetate and azide.

Substrate	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (s ⁻¹ mM ⁻¹)
Wild type ^a			
+ none	0.20 ± 0.006	4.69 ± 0.3	23.45
E176Q ^b			
+ none	NA ^c	NA ^c	NA ^c
+ 0.05 M acetate	0.046 ± 0.009	0.45 ± 0.03	9.83
+ 0.32 M acetate	0.11 ± 0.02	0.97 ± 0.06	9.07
+ 0.05 M azide	0.136 ± 0.009	1.3 ± 0.1	9.68
+ 0.32 M azide	0.21 ± 0.02	1.86 ± 0.01	9.01
+ 1 M azide	0.438 ± 0.007	1.84 ± 0.05	4.19

^a Activity was assayed in 50 mM acetate buffer, pH 5.0.^b Activity was assayed in 50 mM MES buffer, pH 5.0.^c NA, no activity detected.**Table 3.10** Apparent K_m , k_{cat} , and k_{cat}/K_m of E176Q mutant BGlu1 for 2,4-DNPG

hydrolysis in the presence and absence of 0.2 M acetate and azide.

	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (s ⁻¹ mM ⁻¹)
Wild type			
+ none	0.032 ± 0.003	9.1 ± 0.7	284.91
+ 0.2 M acetate	0.020 ± 0.002	2.3 ± 0.1	113.75
+ 0.2 M azide	0.021 ± 0.002	5.4 ± 0.3	257.05
E176Q			
+ none	NA ^a	NA ^a	NA ^a
+ 0.2 M acetate	0.032 ± 0.004	0.26 ± 0.02	8.03
+ 0.2 M azide	0.11 ± 0.02	1.32 ± 0.01	11.54

^a NA, no activity detected.

The mutation of the acid/base residue should affect both the rates of the glycosylation and deglycosylation steps. The effect on the glycosylation step strictly depends on the leaving group ability of the aglycone part of the substrate. Substrates with poor leaving groups (high pK_a) are more affected than those with good leaving groups (low pK_a). Therefore, if glycosylation is the rate limiting step, the rate should vary with substrate reactivity. Deglycosylation becomes a rate limiting step for substrates with good leaving groups, which do not require acid assistance in the first step. The mutation affected substrates with different leaving groups equally in the deglycosylation step, since the same glycosyl-enzyme complex formed from all substrates.

Interestingly, *p*NPG (which has a relatively poor leaving group, $pK_a = 7.18$) was hydrolyzed more than expected by the E176Q mutant in the presence of azide or acetate. Moreover, both the apparent K_m and k_{cat} increased with increasing azide concentrations, whereas the k_{cat}/K_m value was not significantly changed (9.01 vs. 9.98 $s^{-1} mM^{-1}$). However, at a very high concentration of azide (1 M), the K_m increased more but not the k_{cat} , which yielded about a 2-fold decrease in the k_{cat}/K_m value. The same trend was obtained for acetate, although no data were collected at 1 M acetate (Table 3.9). For 2,4-DNPG ($pK_a = 3.96$) hydrolysis, the k_{cat} in the presence of 0.2 M azide or acetate was not obviously different from that of *p*NPG, though acetate showed some inhibition of 2,4-DNPG hydrolysis by the wild type enzyme, with a 4-fold decrease in the k_{cat} value (Table 3.10).

3.11.6 Characterization of the products from rescued activity

TLC analysis of the reaction mixture containing different nucleophiles at 50 mM *p*NPG, and the E176Q acid/base mutant revealed 2 main products of glucose ($R_f = 0.39$) and an unknown compound ($R_f = 0.71$). Surprisingly, the R_f values of the unknown spot from reactions with different nucleophiles were not distinguishable, but azide yielded a much more intense spot (Figure 3.45). Furthermore, the UV light could detect all unknown spots, except for that from the azide reaction, which appeared later after staining for sugar. The results suggest that the weaker spots from the reactions of acetate, formate, cyanate, glycine, and MES might be different from the intense spot for the azide reaction. Moreover, a faint spot with a similar R_f was obtained with the wild type enzyme, which suggests that it might not be product from the activity rescued by the nucleophiles, except for in the azide reaction, but is more likely due to transglycosylation of the *p*NPG substrate.

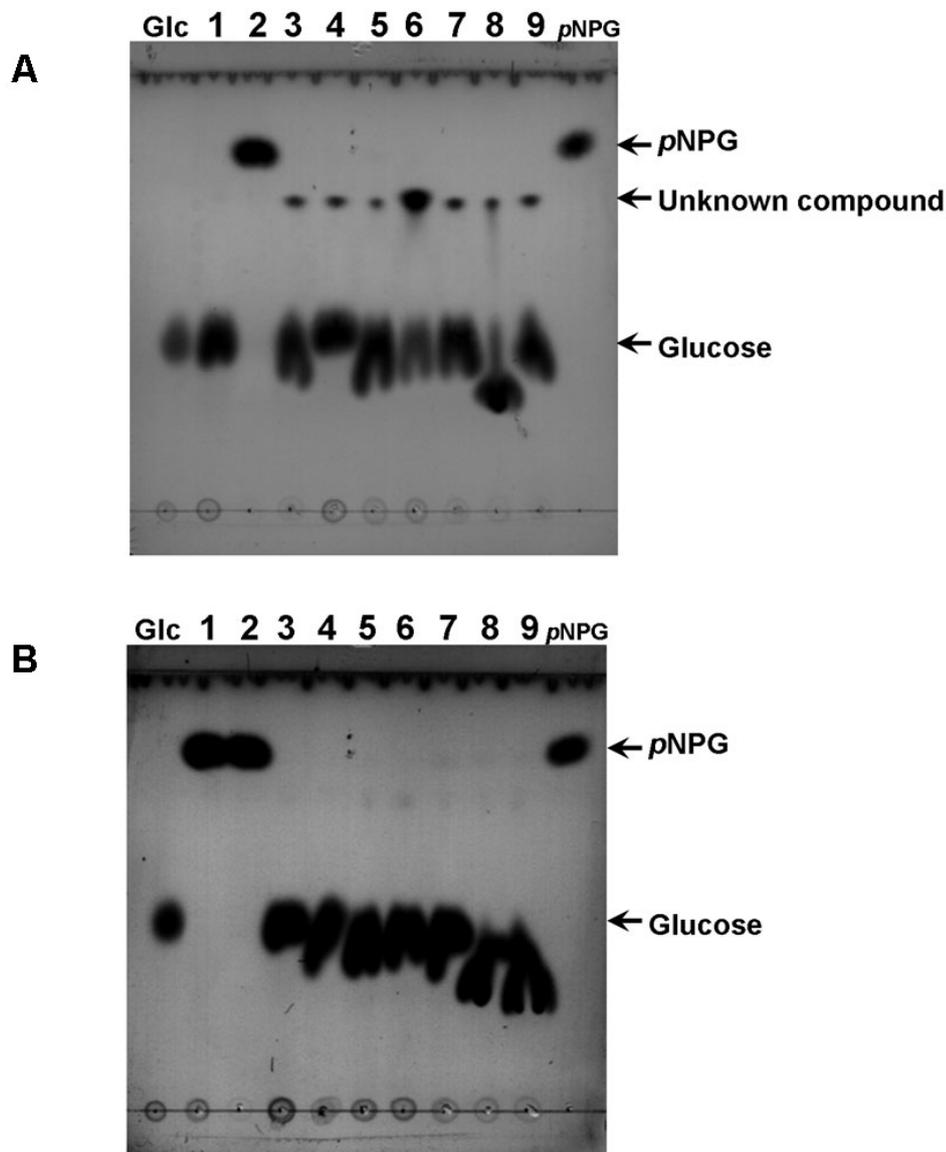


Figure 3.44 Detection of the products of *p*NPG reactions with E176Q and wild type BGlu1 by TLC. A) Rescued activity of E176Q using *p*NPG as a substrate, lane 1: wild type enzyme/acetate buffer, lane 2: MES buffer without enzyme, lane 3: E176Q/MES buffer, lane 4: E176Q/acetate buffer, lane 5: E176Q/acetate/MES buffer, lane 6: E176Q/azide/MES buffer, lane 7: E176Q/formate/MES buffer, lane 8: E176Q/cyanate/MES buffer, lane 9: E176Q/glycine/MES buffer. B) The same nucleophiles and buffer in each lane as A were used, but with the wild type enzyme

(for lanes 3-9). Lane 1: acetate buffer without enzyme, lane 2: MES buffer without enzyme. Glucose and *p*NPG standard spots are marked, as are the unknown product spots. All conditions contained 5 mM *p*NPG as a substrate. All nucleophiles were prepared at 50 mM in 50 mM MES, or acetate buffer, pH 5.0.

CHAPTER IV

DISCUSSION

4.1 Protein expression and purification of rice BGlu1 β -glucosidase

Rice BGlu1 was expressed in the *E. coli* strain Origami (DE3) as an N-terminal thioredoxin and hexahistidine tag fusion protein. Thioredoxin fusion protein was purified by 1 to 4 purification steps depending on the desired purity. A single step of IMAC was used to eliminate most of the contaminating host proteins by 5 and 10 mM imidazole washes. The fusion protein did not bind tightly to Ni²⁺-bound resin and gradually came out at 10 mM imidazole wash, suggesting that hexahistidine tag (His₆) might not be fully exposed to the solvent. Approximately 85-90% pure protein was obtained from a single step of purification by IMAC. However, proteolysis occurred when protein was stored at 4°C. Two intense bands of M_r 66 kDa and ~50 kDa were observed on SDS-PAGE, corresponding to the thioredoxin fusion protein and a likely proteolysis product. To reduce the contamination of proteases, Co²⁺ bound resin was used in place of Ni²⁺ resin and soybean trypsin inhibitor was added to extraction buffer. In addition, S200 gel filtration chromatography was used to eliminate 21 and 15 kDa protein bands derived from the fusion tag, which co-purified with the fusion protein in IMAC. This procedure gave protein of sufficient purity for kinetic analysis.

The thioredoxin fusion tag was cleaved off from BGlu1 proteins used for crystallization, which were further purified by a second round of IMAC, SP sepharose cation exchange chromatography, and S200 gel filtration column. After enterokinase cleavage, nonfusion protein was eluted from Ni²⁺- and Co²⁺-based resins differently.

For Ni²⁺-bound resin, some of the nonfusion protein came out in the unbound fraction, but most of it remained in the column. Most of the remaining BGlu1 was released from the resin in a 5 mM imidazole (10 cv) wash and a trace amount came off in a later 10 mM imidazole (5cv) wash. This indicates that nontagged protein can bind to the resin by amino acid residues other than the his-tag, or possibly the sepharose matrix of the resin, since our protein is a carbohydrate active protein. After this step, a trace amount of free tag was observed on SDS-PAGE, therefore BGlu1 protein was further purified by SP sepharose cation chromatography. In contrast, all of the nonfusion protein came off only in the unbound fraction by washing with equilibration buffer, and free tag protein did not come out until 150 mM imidazole wash on the Co²⁺ column. This implies that the difference in the metal ion affects the binding of protein to the resin and the binding of the Co²⁺ resin is much more specific and selective than Ni²⁺. When the tag was cleaved off, the protein could not bind to the column by any interaction, unlike what appeared to happen with Ni²⁺. However, the protein was still passed through SP sepharose to eliminate any trace amount of impurity that was not obvious on SDS-PAGE. Finally, >95% pure protein with a yield of 2.6 mg protein/l cell culture was used for crystallization. Later batches of protein gave similar or better purity and yield, as did purification of the BGlu1 E176Q mutant after the first batch, which displayed strange behavior in cation exchange chromatography. The pure protein obtained was critical for crystallization.

4.2 Protein crystallization

Purity and age of protein were crucial for protein crystallization. For rice BGlu1, highly pure protein yielded reproducible crystals, while no crystal was

obtained from protein with lesser purity. In addition, freshly prepared protein produced good crystals, whereas old protein produced only poor crystals. Moreover, when the protein batch, precipitant, or other reagent was changed, the crystallization conditions were finely tuned around the previous successful condition to reproduce crystals.

Microseeding was performed for crystallization of the BGlu1-G2F complex, since crystal clusters formed when optimization without seeding was performed. Any aggregated protein, dust, or particles in the protein solution may have caused excess nuclei in the streaking drop (Bergfors, 1999), which apparently resulted in cluster formation. Therefore, protein filtration was needed for seeding, although it was generally not necessary for crystallization without seeding. To reach a metastable drop where crystal growth but not nucleation takes place, the low protein concentration of 2 mg/ml was equilibrated with precipitant solution for 1-2 hr. Single crystals appeared along the seed line within a few days after the setup, and larger size was obtained after a month. These results suggest that low protein concentration is useful for crystallization with seeding, while high protein concentration is needed for nucleation of initial crystals. Moreover, seeding yielded larger crystals, allowed formation of single crystals, reduced the wait required for spontaneous nucleation and improved diffraction quality.

Seed improvement was also performed for crystallization of BGlu1 with 2-F-glucoside inhibitor. The first seeds were prepared from clusters and streak-seeded into a new drop. Single crystals with poor shape crystallized in this second drop, and one of them was used to produce seeds for a new drop. Finally, single crystals with good shape were obtained when a single crystal was used as seed, suggesting that improved seeds prepared from single crystals resulted in the best crystals at the end.

4.3 Crystal structures of rice BGlu1

4.3.1 Overall structure

The derived crystal structure of rice BGlu1 was compared with known GH family 1 structures from plant enzymes and one from bacteria. The core barrel fold was similar to all of known structures. The most similar enzyme was white clover cyanogenic β -glucosidase (1CBG, Barrett et al., 1995), which was used as a search model, while the bacterial β -glucosidase BglB from *P. polymyxa* (Isorna et al., 2007) had the lowest similarity. The main structural differences were in four loops (A-D), which constitute a part of the aglycone binding site. This agrees with previous reports that these loops are less conserved and contribute to the substrate specificity among GH family 1 enzymes (Burmeister et al., 1997; Wiesman et al., 1997; Sans-Aparicio et al., 1998; Czjzek et al., 2000; 2001; Husebye et al., 2005; Isorna et al., 2007).

4.3.2 Monomer and dimerization by zinc ion

Rice BGlu1 is a monomeric protein in solution (Opassiri et al., 2003), whereas other plant GH1 enzymes for which structures have been solved are dimeric (1CBG, cyanogenic β -glucosidase from white clover, Barrett et al., 1995; 1MYR, *Sinapsis alba* myrosinase, Burmeister et al., 1997; 2000; *Zea mays* BGlu1, Czjzek et al., 2000; 2001; sorghum, Verdoucq et al., 2004) or hexameric (wheat, Sue et al., 2006). However, in the BGlu1 crystal structures, two molecules related by a noncrystallographic 2-fold symmetry axis were linked through a single Zn^{2+} ion, which is symmetrically bound between the D65 and H68 residues of the two molecules (Figure 4.1). A peak was observed at the Zn^{2+} absorbance edge was observed in an x-ray absorbance/fluorescence scan, while no peak was seen at the

edge of Co^{2+} (which might be left from IMAC purification), which confirmed the identification of the metal ion. The Zn^{2+} was evidently left from seeding crystal conditions, or contamination in the chemicals or water used for purification or crystallization, since no Zn^{2+} was used in the current crystal growth conditions. The linkage through this Zn^{2+} suggests that BGlu1 could form a dimer in the presence of metal ions, though the small shared surface area between the monomers suggests this may not be stable in solution. The presence of Zn^{2+} was also reported to stabilize a dimer in the *Sinapsis alba* myrosinase structure, in which zinc was similarly bound between His and Asp residues, H56 and D70 from each myrosinase molecule (Burmeister et al., 1997). Though these residues are the same types of amino acids as in rice BGlu1 and fulfill the same role, they are in the opposite order in the protein sequence and neither is homologous to the corresponding residue in BGlu1. However, 95% of the rice BGlu1 protein eluted as monomer during gel filtration, dynamic light scattering (DLS) gave a single peak at a diameter of 7.115 nm, corresponding to the expected size of the monomer, and the enzyme is neither inhibited by EDTA nor stimulated by metal ions (Opassiri et al., 2003), so the monomer appears to be the primary biological unit, and the significance of the dimer, outside of crystallization, is unclear. This confirms that the monomeric enzyme is the functional form found in solution, despite the appearance of a crystallographic dimer.

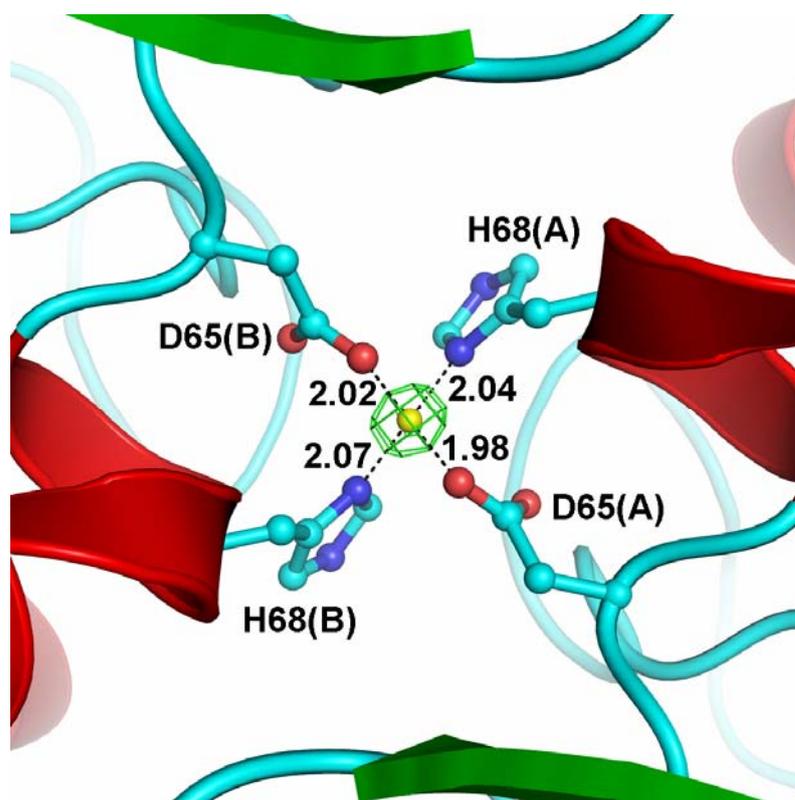


Figure 4.1 Crystallographic contacts mediated by Zn^{2+} . The two protein molecules in the asymmetric unit are linked by D65 and H68 (ball and stick, with blue nitrogen and red oxygen atoms) from molecule A and B via Zn^{2+} (yellow sphere). Dashed lines represent the hydrogen bonding interactions with their distances given in Ångströms. The figure was produced with Pymol (DeLano, 1991).

4.3.3 Active site structure

Rice BGlu1 has a long tunnel-shaped active site pocket, as does PpBGIB β -glucosidase, which hydrolyzes similar substrates. In both of these enzymes, the long active site pocket includes additional subsites for binding additional glucosyl residues, showing that they are adapted for hydrolysis of their cell wall-derived oligosaccharide substrates. In the innermost part of the active site in the crystal structure of the G2F

complex, the 2-F-glucosyl moiety is covalently bound to E386 nucleophile residue in the glycone binding pocket. This inhibitor is actually a slowly hydrolyzed substrate, so the structure represents the glycosylated enzyme intermediate of the double displacement mechanism used by GH1 enzymes (Withers et al., 1990). The glucosyl unit at this site formed extensive hydrogen bonds with the surrounding conserved amino acid residues, and a stacking interaction with W433, as seen in other GH1 enzymes.

The overall structures of the substrate-free and covalent G2F complex were very similar in terms of the global conformations of both molecules in the asymmetric units. The only differences between these structures were observed at the active site center. The movement of E386, the catalytic nucleophile, to link with glucose at C1 and of Y315 were prerequisite for formation of the covalent G2F intermediate, as observed in other GH1 covalent intermediate complexes (Burmeister et al., 1997, Zechel et al., 2003; Gloster et al., 2004b; Isorna et al., 2007). Interestingly, Y131, which is unique to rice BGlu1 and barley BGQ60 β -glucosidases, was twisted slightly to hydrogen bond with a glycerol that bound in subsite +1 next to the G2F moiety. This implies that this residue may be involved in the substrate specificity or the high transglycosylation activity of these two enzymes.

Interestingly, the covalent 2-F-glucoside-bound complex also provides a snapshot of the transglycosylation reaction (Figure 4.2). The glycerol molecule seems to be positioned for deprotonation by the acid/base residue for nucleophilic attack at the anomeric carbon in the deglycosylation step, which would lead to transglycosylation. This is consistent with the high transglycosylation activity of the enzyme with

oligosaccharides at relatively low concentrations (Opassiri et al., 2004), because glycerol is equivalent to half a hexose sugar and the coordination of the 1 OH positions the 2 OH in a way that could act in hexopyranose sugars as well. A water molecule is in a similar position between the anomeric carbon and an ascorbate that takes the place of the catalytic acid/base in the 1E73 myrosinase structure, which is also a 2-F-glucoside intermediate (Burmeister, 2000). In that case, the enzyme was trapped in the process of hydrolysis, instead in transglycosylation with the glycerol in the rice G2F complex structure.

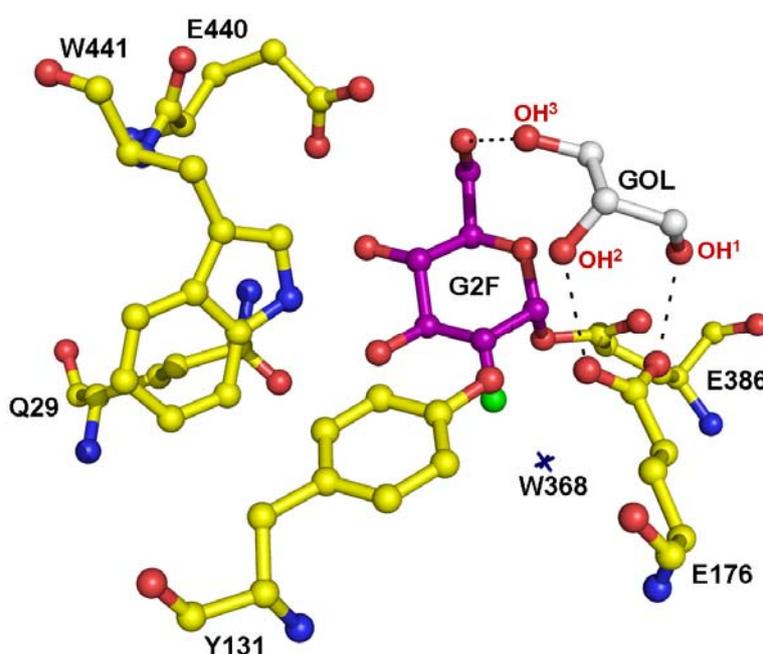


Figure 4.2 View of the transglycosylation intermediate in the crystal structure of the G2F complex. Hydrogen bonds involved in glycerol binding are shown as dashed lines. One water molecule observed in the active site center; W368, which appears to be hydrogen bonded to N175 and the catalytic nucleophile (E386), is drawn as a cross. The figure was produced with Pymol (DeLano, 1991).

Since glycerol molecules were found in active site of the free enzyme and inhibitor complex at subsites -1 and +1, respectively, they might also be found in the two structures of the E176Q mutant soaked with laminaribiose and co-crystallized with cellopentaose. Poor density at the -1 site in these structures might be due to competition with glycerol, which may occupy the site in some molecules in the crystal.

Although, a half density at subsite -1, which could be a glycerol molecule, was observed in initial inspection of the density maps of both mutant structures, a continuous electron density from subsite +1 to +4 appeared after rigid body refinement of the cellopentaose complex data with a ligand-free BGlu1 structure. Furthermore, stronger density that was more extensive than a glycerol molecule appeared at subsite -1 after several refinement cycles, especially when the omit map was calculated for the E176Q mutant soaked with laminaribiose. This suggests that the density at subsite -1 of mutant structures might not belong to glycerol, but to the substrate. As noted above, it may also be that some of the protein molecules had glycerol and some had substrate. Weak density of the sugar ring at subsite -1 was reported in the crystal structure of ZmGlu1-DIMBOAGlc (Czjzek et al., 2001), whereas the SbDhr1-dhurrin structure was reported to have well-defined density of a glucosyl residue in a skew boat conformation at that subsite (Verdouqc et al., 2004). In the structure of the E176Q mutant BGlu1 with cellopentaose, a chair conformation was built into that density, though the electron density was not sufficient to define the exact conformation.

4.3.4 Computational docking of cellobiose, laminaribiose and cellotriose into the crystal structure of rice BGlu1

Automated docking was able to provide information about enzyme-substrate interactions of cellobiose, laminaribiose and cellotriose, which have not currently been obtained from x-ray crystallography study. Cellobiose and laminaribiose were docked into the ligand-free structure of rice BGlu1 using Autodock 3.0.5, which yielded a similar position for the nonreducing terminal glucosyl residue as that of the G2F in the covalent complex crystal structure and reasonable distances between the enzyme and substrates. However, the differences in substrate binding between cellobiose and laminaribiose were not distinguishable by Autodock program. In addition, cellotriose with the nonreducing end glucosyl residue in a skew boat conformation was docked into the active site in a position similar to cellobiose and laminaribiose, and this revealed an additional binding site for third glucosyl unit at subsite +2.

Only one direct hydrogen bond between the enzyme and substrate was observed at subsite +2 and no direct hydrogen bonds were observed at subsite +1 for all three docking results. It should be noted that many of the hydrogen bonding interactions with the sugar hydroxyls may be mediated by water molecules, such as seen between the aromatic hydroxyl of dhurrin and a serine in the dhurrinase complex (Verdouqc et al., 2004), but waters were not included in the docking experiment. Nonetheless, the docking is likely to give a good approximation for the positions of glucose residues in subsites +1 and +2, since the nonreducing glucose appears to be close to its correct position in subsite -1. However, longer oligomers, such as cellohexaose, could not be successfully docked into the active site of rice BGlu1 with or without the skew boat conformation at nonreducing end, as was done for

cellotriase. Nonetheless, the position of the cellotriase chain suggested that an array of aromatic amino acids around the entrance way, including F343, Y341, and W337, may act as platforms for additional glucosyl residues.

4.3.5 Crystal structures of the E176Q mutant soaked in laminaribiose and its complex with cellopentaose

The crystal structure of the E176Q mutant soaked with laminaribiose was solved, but ligand electron density at subsites -1 and +1 was weak. It was not possible to model glucosyl residues into the density, therefore water molecules were added into those places. However, the electron density of laminaribiose was more clearly defined in the omit map, though there was density missing at the C1 of the scissile bond and the connection between the two glucosyl residues. Surprisingly, this structure showed two alternative conformations at loop C for both molecules in the asymmetric unit. The conformational ambiguity was seen in the positions of Y341, V342, and F343, which are located on the surface at the entrance to the active site. The side chains of Y341 and F343 swing over in a position to bind the glucosyl residues at the entrance of the active site for the first conformation (Figure 4.3A), whereas in the second conformation they have moved away from this binding position (Figure 4.3B). Conformational flexibility also occurred at this position in the wild type structures, since partial densities of the second conformation were observed, though with much lower occupancy.

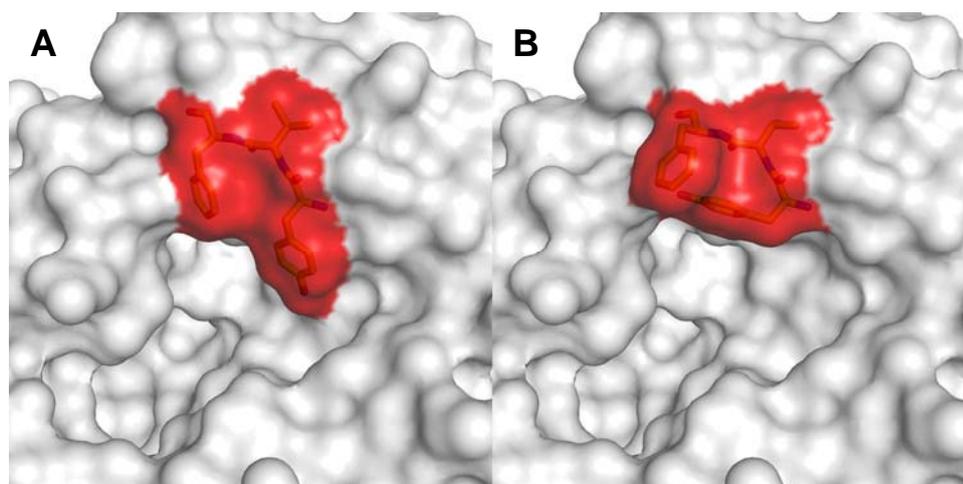


Figure 4.3 Molecular surface of the E176Q mutant illustrating the alternative conformations of loop C at residues Y341, V342, and F343. The structure on the left was predominantly observed in the wildtype and cellopentaose structures, while the second conformation on the right was observed at nearly half occupancy in the E386Q mutant structure from a laminaribiose-soaked crystal. The figures were produced with Pymol (DeLano, 1991).

The structure of E176Q mutant BGlul with cellopentaose has been solved. Cellopentaose density was as strong as surrounding active site atoms, except around C1 of Glc1 and the outer atoms of Glc5 at subsites -1 and +4, respectively. Glc1, Glc2 and Glc3 were buried in the narrowed tunnel part, while further glucosyl units were bound in the open groove. This is in good agreement with the calculation of subsite affinity by Opassiri et al. (2004) in that the subsite affinities of subsites +3, +4, and +5 were much smaller than those of subsites -1 and +2 in the tunnel section (except for the fact that subsite +1 had lower affinity). The enzyme and substrate make both hydrophobic and polar hydrogen-bonding contacts, many of which are mediated by water, as one would expect. The hydrophobic face of each glucosyl unit interacts with

one of the aromatic side chains lining the active site tunnel. These interactions included Glc1 with W433, Glc2 and Glc3 with W358, and Glc4 and Glc5 with Y341. Interestingly, Y341 must be in one of the two alternative conformations in the loop C for substrate recognition at subsites +3 and +4. Thus, electron density was only seen for that conformation in both molecules in the asymmetric unit. This is the same conformation as the one that appeared to predominate in the wild type BGlu1 structures. Aside from the stacking interactions with aromatic side chains, other nonpolar amino acids were extensively observed at subsites +1 and +2. Interactions of the array of polar amino acids lining the active site with the glucosyl residues beyond the -1 site could be mainly mediated by water molecules. Moreover, one directed hydrogen bond between N245 and O2-Glc3 at subsite +2, supported the highest subsite affinity in rice BGlu1, while barley BGQ60 β -glucosidase, where N245 is replaced by valine, showed unfavorable binding energy at this site. The comparison of cellopentaose obtained by crystallography and cellotriose from computational docking yielded similar positions. Celltriose with a skew boat conformation for the glucosyl residue at the nonreducing end yielded the stacking interaction between Glc1 and W433 at subsite -1 very close to those obtained by crystallography (~ 1.3 Å apart), while cellobiose and laminaribiose showed larger differences. This suggests that the docking with the skew boat conformation, as likely occurs in the real mechanism, yielded a more accurate docking.

4.4 Substrate specificity

Rice BGlu1 and barley BGQ60, which share 66% sequence identity, are different in substrate specificity. Barley BGQ60 prefers to hydrolyze cellobiose with

an apparent disruptive interaction at the third glucosyl residue in the +2 subsite. In contrast, rice BGlu1 hydrolyzes cellobiose very poorly and has high affinity for the third glucosyl residue at the +2 site. These two enzymes were used as a model to determine substrate specificity for enzymes which hydrolyze cell wall-derived oligosaccharides. To transform rice BGlu1 into barley BGQ60 (cellobiase activity), four residues found in the aglycone binding pocket at subsite +1 and +2 of rice BGlu1 were mutated to barley residues. Surprisingly, all mutants, I179V, N190H, N245V, L442R, had changes in K_m and k_{cat} in the same direction compared to wild type BGlu1 for hydrolysis of all the substrates tested (*p*NPG, cellobiose, and cellotriose). No distinct increase in cellobiase activity was obtained for any of the rice BGlu1 mutants and L442R exhibited the same activities as wild type. However, N190H and N245V showed increased catalytic efficiency for cellobiose by 3- and 2- fold, respectively, and the I179V, N190H and N245V mutants each showed an increase in relative catalytic efficiency for cellobiose hydrolysis vs. cellotriose ($k_{cat}/K_m^{cellobiose}/k_{cat}/K_m^{cellotriose}$), but not to the level of the barley enzyme. Though these are small differences, it is possible that if these changes were combined, they might yield together the more efficient cellobiase activity seen in the barley β -glucosidase. As noted by Mildvan et al. (1992), the effects of such mutations on the K_m and k_{cat} can be additive, cooperative or synergistic, so it should be informative to investigate the double mutations N190H/N245V and I179V/N245V, or triple mutation of I179V/N190H/N245V in the future to better understand the basis of substrate specificity between rice BGlu1 and barley enzyme.

The relative subsite affinities of β -glucosidases with cellobiase activity are very similar, that is, subsite affinity of subsite +1, +2, and +3 are much larger than

those of subsites +4, +5, and +6 (Yazaki et al., 1997; Hrmova et al., 1998; Alexandre et al., 2001; Fukuda et al., 2002). In addition, the subsite affinity of subsite -1 is larger than subsite +2, indicating that the cellobiose is bound mainly in the productive mode and the probability of productive binding of cellooligosaccharides is certainly high. Rice BGlu1 exhibits a different pattern of subsite affinity in that the affinity at subsite -1 is smaller than at subsite +2. It should be noted that only one direct hydrogen bond at subsite +2, found in crystal complex, contributed to its affinity being the highest, while other reported β -glucosidases showed highest affinity at subsite +1. Although more extensive hydrogen bonding was found at subsite -1, it still showed lower subsite affinity than subsite +2, which contains only one direct hydrogen bond. This implies that most of the energy of binding might be used for the distortion of glucose ring to the skew boat, which is between the half chair transition state and the relaxed chair conformation of the initial substrate and covalent complex (White and Rose, 1997).

The subsite mapping data suggests that rice BGlu1 appears to bind the second glucose from the nonreducing end of cello-oligosaccharides (1,4-linked) poorly, but the third glucose well. In contrast, cellobiases, such as barley BGQ60 (Hrmova et al., 1998), β -glucosidase from *Aspergillus niger* (Yazaki et al., 1997), β -glycosidase from *Tenebrio molitor* L. (Alexandre et al., 2001; 2003), and β -glucosidase from *Streptomyces* sp. (Fukuda et al., 2002) bind the second glucose well, but not the third, as noted above. The energy of binding at the different sites implies that cellobiose may mainly bind the nonproductive site in rice BGlu1, while productive binding can compete more effectively with nonproductive sites at high substrate concentration (Figure 4.4), which is consistent with some of our results that show no obvious

saturation of the enzyme at high cellobiose concentrations.

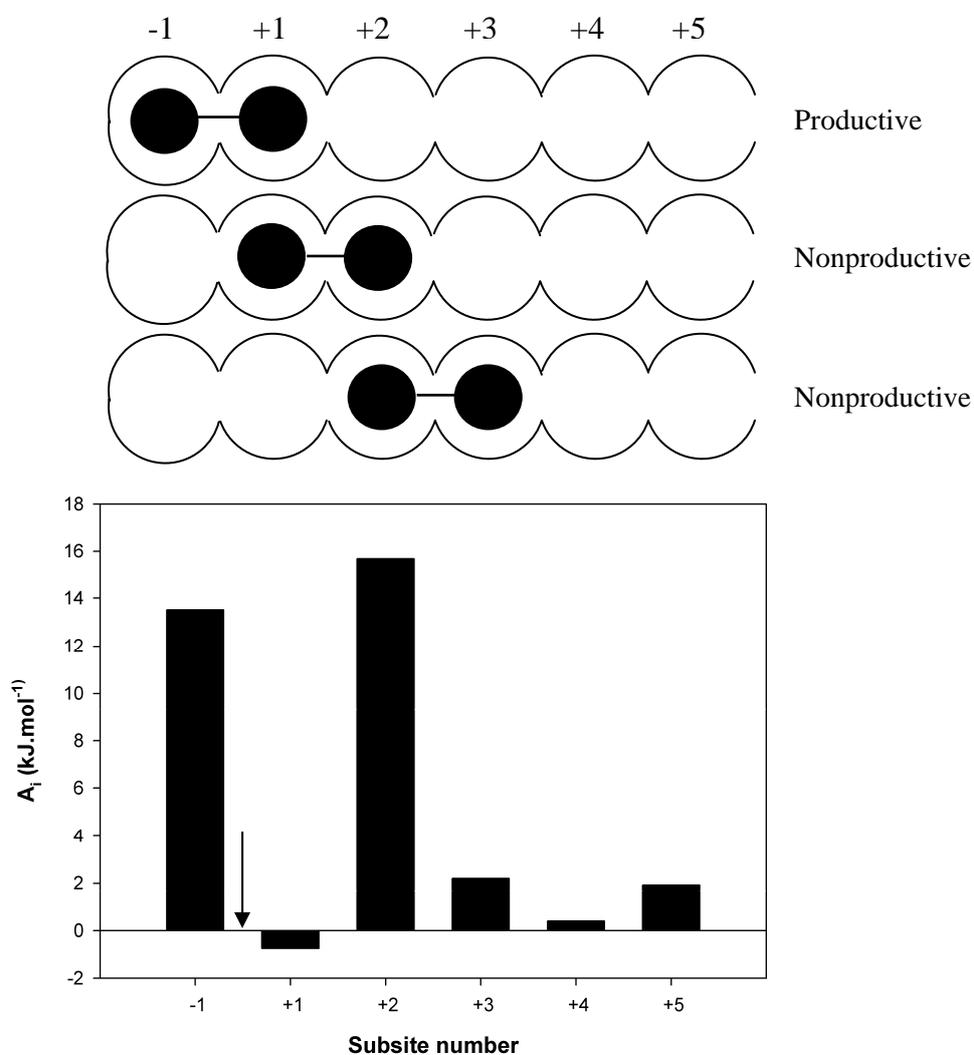


Figure 4.4 Schematic representation of rice BGlu1 active site and the binding modes of cellobiose binding and a histogram of subsite affinities of rice BGlu1 for celooligosaccharide hydrolysis. The arrow head indicates the catalytic site of the enzyme.

4.5 Enzyme mechanism and rescue activity

The nucleophile and acid/base catalyst in rice BGlu1 were identified by kinetic analysis of mutants. The functional role of both catalytic residues could be probed by mutagenesis and chemical rescue methodology based on activation of nucleophile or acid/base mutants by azide or other small anionic nucleophiles, such as formate, fluoride, and cyanate as exogenous nucleophile (Wang et al., 1995; MacLeod et al., 1996; Vallmitjana et al., 2001; Bravman et al., 2003).

The nucleophile attacks the anomeric carbon to form a covalent intermediate and permit the release of the aglycone leaving group. Therefore, replacing the nucleophile with a non-ionizing residue (E386Q) and shortening the carboxylic residue (E386D) severely affected enzyme catalysis with 60,000 and 3,000-fold decreases in enzyme activity. As previously reported, the activity of nucleophile mutants (E386G, E386S, and E386A) of rice BGlu1 were rescued by formate and azide (Hommalai et al., 2007). In addition, these latter nucleophile mutants showed glycosynthase activity when α -glucosyl fluoride was used as donor and *p*NP-cellobioside as acceptor. All these kinetic data indicate that E386 is the catalytic nucleophile for rice BGlu1, which agrees well with the 2-fluoroglucoside being covalently bound to E386 in the covalent intermediate crystal structure.

For the acid/base mutations, the replacements of E176 with glutamine, aspartate, and alanine affect the nucleophilicity, length, or both of these properties. Mutation at this putative acid/base residue caused approximately 50-, 320-, and 420-fold drops in activity for E176Q, E176D, and E176A, respectively. These results indicate that the shortened side chain of aspartate cannot effectively deprotonate the coming water molecule, since the carboxylate group of side chain was too far away

from them. However, elimination of nucleophilicity and polarity and shortening of the side chain in the alanine mutant had the largest effect on enzyme activity.

Studies of the pH dependence of wild type BGluc1 and mutants using overlapping buffer over the ranges of pH 4.0-9.0, yielded activity rescued by acetate in two mutations of the acid/base catalyst (E176Q and E176A). In contrast, no detectable activity was observed for acid/base mutants in universal buffer, because no effective nucleophile was present. High rescue activity of a catalytic acid/base glutamine mutant (E170Q) was also reported in *Agrobacterium* sp. β -glucosidase. The glutamine mutant yielded the highest catalytic efficiency for thioglycosylase activity compared to other mutants, such as E170G, E170N, E170S, E170A, and E170T (Mullegger et al., 2005).

The ability to restore the enzyme activity of acid/base mutants was tested for various anionic nucleophiles. Ascorbate and azide showed the highest efficiency, while acetate gave modest rescue of activity. This is the first report of rescue activity of acid/base mutant (replacing of glutamate by glutamine) by ascorbate, except for *Sinapis alba* myrosinase, which displays these phenomena in the nature (Burmeister et al., 2000). This activity rescue indicates that not only a small anionic nucleophile like azide could greatly increase the rate, but a larger nucleophile like ascorbate could also apparently penetrate into the space of the active site, even though no more space was created by replacing glutamate with glutamine. The structure of myrosinase in complex with ascorbate and covalently bound substrate clearly revealed that ascorbate is involved in the deglycosylation step and bound at the aglycone binding site after release of the aglycone (Burmeister et al., 2000). In that case, it was proposed that the ascorbate acts as a catalytic base in place of glutamate. In the presence of azide or

ascorbate, E176Q and E176A lost pH dependence over the range 5.0-10.0, indicating that azide and ascorbate were not substituting for the missing acid/base catalytic residue, but appear to be acting as nucleophiles in place of water, which cannot be activated in the absence of the catalytic base, though they could also act to activate water in some reactions. The shorter side chain of alanine probably allows more space to be occupied by those external nucleophiles than aspartate, whose negative charge would also repel these negatively charged nucleophiles. Replacement of glutamate by glutamine, which is the same size, so no more space was generated by this mutation as in the alanine mutant. It has been suggested that the amide group might form a hydrogen bond with the incoming nucleophile, thereby positioning it for attack on the anomeric carbon (Mulegger et al., 2005).

Effects on the glycosylation step due to removal of the acid catalyst depend on the aglycone leaving group ability, with minimal effects on substrates requiring no general acid catalyst such as 2,4-DNPG, but large effects on substrates with poorer leaving groups such as *p*NPG or cello-oligosaccharides. (MacLeod et al., 1996). The rate acceleration for 2,4-DNPG hydrolysis by external nucleophile anions such as azide has been extensively observed in other enzymes (Wang et al., 1995; Vallmitjana et al., 2001; Bravman et al., 2003) but was not observed for *p*NPG. Wang et al., 1995, reported that the *Agrobacterium* Abg β -glucosidase E170G acid/base mutant showed no significant rate enhancement by azide for *p*NPG, for which glycosylation was rate limiting. Later, Vallmitjana, et al., 2001 reported that the corresponding mutant E178A of Bgl3 β -glucosidase from *Streptomyces* had a very low k_{cat} for *p*NPG hydrolysis, but was significantly reactivated by azide. In this case deglycosylation was partially rate limiting, but glycosylation became fully rate limiting at a higher

degree of reactivation by azide. In rice BGlu1, the kinetics of the E176Q mutant *p*NPG hydrolysis reactivation by anionic nucleophiles showed that it appeared more reactive than in most other reported glycosidases, indicating that there are subtle differences in the fine-tuning of the enzyme activity. Perhaps the glutamine also forms a hydrogen bond with the departing *p*NP group, thereby stabilizing negative charge formation on this group in the glycosylation step.

As previously reported, varieties of α - and β - configurations were produced by catalysis of nucleophile and acid/base mutants, respectively. Various products, including glucosyl-azide, xylosyl-azide, cellobiosyl-azide, and mannosyl-fluoride (Wang et al., 1994; MacLeod et al., 1996; Vallmitjana et al., 2001; Zechel et al., 2001; Bravman et al., 2003), were detected by either TLC and or ^1H NMR. Based on these studies, the product released by E176Q catalysis in the presence of azide was expected to be β -glucosyl azide. However, similar sizes of unknown spots from different anionic nucleophiles were observed on TLC. Some spots might be *p*NP β -D-cellobioside produced by transglycosylation reaction, except for the product of the azide reaction, which did not show the UV absorbance expected for compounds containing *p*NP. If sufficient amounts can be purified, these unknown compounds products should be characterized and reconfirmed by NMR or Mass spectrometry or run on TLC with a suitable standard, such as the expected glucosyl nucleophile derivative or *p*NP β -D-cellobioside. However, it is expected that the transglycosylation products with the nucleophiles other than azide were not stable, as previously reported (Wang et al., 1995), which would explain the large amount of glucose released in these reactions.

CHAPTER V

CONCLUSION

Rice BGlu1 β -glucosidase was expressed as a N-terminal thioredoxin and hexahistidine (His₆) tag fusion protein in *E. coli* Origami (DE3) and purified for kinetic and structural studies. For kinetic analysis, the 66 kDa thioredoxin fusion protein was purified by IMAC to remove most of contaminating proteins, then by S200 gel filtration chromatography to remove 21 and 15 kDa proteins derived from the fusion tag, and obtain >90% purity in a yield of about 4-6 mg protein/l cell culture. To obtain higher purity and homogeneity for crystallization, the thioredoxin fusion tag was removed, followed by another round of IMAC, SP cation exchange chromatography, and S200 gel filtration chromatography. Approximately 2.6 mg of homogenous 50 kDa BGlu1 protein was obtained per liter of cell culture, with >95% purity.

Small BGlu1 crystals without ligand formed under several conditions containing PEG ranging in molecular weight from 3350 to 8000, and di-valent salts, like (NH₄)₂SO₄, and Li₂SO₄, and mono-valent salts, like MgOAc, ZnOAc, NaCl and MgCl₂, in the pH range 6.5-8.5 of Na cacodylate, MES, HEPES, and Tris-HCl. The systematic optimization of precipitant concentrations, pH, protein concentrations, and the ratios of protein:precipitant yielded larger crystals that allowed x-ray diffraction studies. BGlu1 was also crystallized with 2,4-dinitrophenyl 2-deoxy-2-fluoroglucoside (DNP2FG), which produces a stable covalent 2-F-glucosyl

intermediate. Crystal clusters were produced under the same conditions as free BGlu1 crystals. Microseeding and optimization were used to obtain single and larger crystals of BGlu1 with this inhibitor. A similar optimization method was used to produce diffraction quality crystals of the E176Q catalytic acid/base mutant of BGlu1 and its complex with cellopentaose. Crystals produced in 20-23% (w/v) PEG MME 5000, 0.17-0.22 M $(\text{NH}_4)_2\text{SO}_4$, 0.1 M MES, pH 6.7, using the hanging drop method, with the sizes of up to 350 μm x 140 μm x 15-40 μm . The crystals diffracted x-rays from in-house and synchrotron sources up to 1.33-2.75 \AA resolutions. All the crystals for wild type and mutant BGlu1 in free and complexed forms were isomorphic with 2 molecules in the asymmetric unit of the $\text{P}2_12_12_1$ space group and unit cell dimensions of approximately 76 x 101 x 126 \AA .

Rice BGlu1 has the classical $(\beta/\alpha)_8$ -barrel fold as other GH1 β -glucosidases do. The known structures of plant GH1 enzymes have been reported for oligomeric (dimer to hexamer) proteins, whereas rice BGlu1 is a monomeric enzyme in solution. However, a crystal dimer linked by a zinc ion occurred in the crystal packing, suggesting a dimer could occur in the presence of zinc. Rice BGlu1 has a long tunnel-shaped active site pocket, as does PpBGIB β -glucosidase, which also hydrolyzes long cell wall-derived oligosaccharide substrates, whereas the flattened and shorted slots were observed in ZmGlu1 and SbDhr1. The structural differences were mainly at loop regions (A-D), and loops B and C were mainly responsible for formation of the aglycone binding pocket. Moreover, the sequence and structure of loop C differed significantly from that of other GH1 structures.

The main differences in the free BGlu1 and G2F complex were the movement of the catalytic nucleophile, E386, and Y315, both of which are conserved in other

GH1 enzymes, and of Y131. The latter residue, which is conserved in barley BGQ60 but not other GH1 enzymes. It might be involved in the high transglycosylation activity of these two enzymes, because it was within hydrogen bonding distance with a glycerol molecule found in the G2F complex. This glycerol molecule was positioned to act as the incoming acceptor sugar in transglycosylation.

Molecular docking of cellobiose and laminaribiose and cellotriose was used to predict the binding position of oligosaccharides in rice BGlu1 active site. The binding of productive and nonproductive substrates, laminaribiose and cellobiose, could not be distinguished from their docking positions and calculated docking energies. The docking of cellotriose with a skew boat conformation at the nonreducing end glucosyl residue allowed it to be docked in a similar position as cellobiose and laminaribiose and showed the likely position of subsite +2, where one direct hydrogen bond was found between N245 and O2-Glc3.

Although no global structural changes were seen among the solved structures, two alternative conformations of loop C at residues Y341, V342, and F343 was observed in the high resolution structure of the E176Q mutant. The structural change in this region occurred at the entrance to the active site, and only one of two conformations could be found binding to the Glc4 and Glc5 in the structure of E176Q with cellopentaose. The latter structure showed both hydrophobic and polar hydrogen-bonding contacts, which were mainly mediated by water, between the enzyme and the cellopentaose substrate. Since no water molecules were included in molecular docking, the hydrogen bonding between enzyme and substrate (mediated by water) was not included for calculation. Therefore, there were some minor differences in the positions of the docked cellotriose and the cellopentaose observed

by x-ray diffraction. However, a direct hydrogen bond between N245 and O2-Glc3 at subsite +2 was observed in the results of both docking and cocrystallization. The stacking interaction between the hydrophobic face of each glucosyl unit and the aromatic side chains determined the orientations of each glucosyl unit. Glc1 was stacked by W422, which is located at the end of β -strand 8, while W358 and Y341 which appeared to position Glc2 to Glc5 are located at the end and middle of loop C, respectively.

Four residues located in the aglycone binding site of rice BGlu1 were mutated to barley BGQ60 residues. None of these single mutants showed marked changes in oligosaccharide preference from cellotriose to cellobiose hydrolysis, as had been expected. The L442R mutant still exhibited enzyme activity very similar to wild type BGlu1. As shown in the cellopentaose complex, L442 is located too far away for any interaction with cellopentaose. The I179V and N245V mutants, which have residues that form a nonpolar interaction and a direct hydrogen bond with the substrate, respectively, had K_m and k_{cat} increases for almost all substrates, resulting in lower catalytic efficiency than wild type BGlu1. The N190 mutation had the greatest effect with decreased K_m and increased k_{cat} values for all substrates tested (*p*NPG, cellobiose and cellotriose). However, the increased efficiency for cellotriose hydrolysis was not as high as that of barley BGQ60. Although, we cannot yet explain the difference in oligosaccharide preferences between rice BGlu1 and barley BGQ60 β -glucosidases, the residues I179, N190 and N245 were shown to have effects both on substrate binding and the catalytic rate, confirming their interactions with the substrate and suggesting no single change can fully account for these differences.

As a member of GH family1, rice BGlu1 shares a similar mechanism with other family 1 β -glucosidases. The identification of the catalytic nucleophile and acid/base residues was confirmed by site-directed mutagenesis. The role of the acid/base catalyst was assessed by chemical rescue of the E176Q mutant by anionic nucleophiles. Azide and ascorbate showed highest ability to rescue activity of the E176Q and E176A mutants, while acetate had modest ability. Rice BGlu1 E176Q showed high reactivation activity for *p*NPG hydrolysis by azide compared to other reported glycosidases, indicating a high reactivity even with a less reactive substrate.

The structural information gained about rice BGlu1 and its interaction with oligosaccharide substrates in these studies serve as the basis for understanding the mechanism of substrate recognition and binding and catalysis by this enzyme. Future studies may further investigate the determinants for oligosaccharide selectivity of this enzyme, as well as its preference for glucoside and ability to hydrolyze and transglycosylate other substrates. This may allow this enzyme or related enzymes to be engineered for efficient degradation of cellulosic waste or other applications in the future, in addition to providing more basic information on glycosidase-substrate interactions and selectivity.

REFERENCES

REFERENCES

- Aguilar, C.F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M., and Pearl, L.H. (1997). Crystal structure of the β -glycosidase from the hyperthermophilic archeon *Sulfolobus solfataricus*: resilience as a key factor in thermostability. **J. Mol. Biol.** 271: 789-802.
- Akiba, T., Nishio, M., Matsui, I., Harata, K. (2004). X-ray structure of a membrane-bound β -glycosidase from the hyperthermophilic archaeon *Pyrococcus horikoshii*. **Proteins.** 57: 422-431.
- Akiyama, T., Kaku, H., and Shibuya, N. (1998). A cell wall-bound β -glucosidase from germinated rice: purification and properties. **Phytochemistry.** 48: 49-54.
- Bandaranayake, H., Esen, A. (1996). Nucleotide sequence of a cDNA corresponding to a second β -glucosidase gene in maize (*Zea mays* L.). **Plant Physiol.** 110: 1048.
- Barrett, T., Suresh, C.G., Tolley, S.P., Dodson, E.J., and Hughes, M.A. (1995). The crystal structure of a cyanogenic β -glucosidase from white clover, a family 1 glycosyl hydrolase. **Structure.** 3: 951-960.
- Bergfors, T.M. (1999). Buffers and pH. Protein crystallization: **Techniques, strategies, and tips: A laboratory manual.** pp 54-61.
- Biarnes, X., Nieto, J., Planas, A., and Rovira, C. (2006). Substrate distortion in the Michaelis complex of *Bacillus* 1,3-1,4- β -glucanase: Insight from first principles molecular dynamics simulations. **J. Biol. Chem.** 281: 1432-1441.

- Bolam, D.N., Hughes, N., Virden, R., Lakey, J.H., Hazlewood, G.P., Henrissat, B., Braithwaite, K.L., and Gilbert, H.J. (1996). Mannanase A from *Pseudomonas fluorescens* subsp. *cellulosa* is a glycosyl hydrolase in which E212 and E320 are the putative catalytic residues. **Biochemistry**. 35: 16195-16204.
- Bourderioux, A., Lefoix, M., Gueyrard, D., Tatibouet, A., Cottaz, S., Arzt, S., Burmeister, W.P., and Rollin, P. (2005). The glucosinolate-myrosinase system. New insights into enzyme-substrate interactions by use of simplified inhibitors. **Org. Biomol. Chem.** 3: 1872-1879.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. **Anal. Biochem.** 72: 264-273.
- Bravman, T., Belakhov, V., Solomon, D., Shoham, G., Henrissat, B., Baasov, T., and Shoham, Y. (2003). Identification of the catalytic residues in family 52 glycoside hydrolase, a β -xylosidase from *Geobacillus stearothermophilus* T-6. **J. Biol. Chem.** 278: 26742-26749.
- Brunger, A.T., Adam, P.D., Clore, G.M., DeLano, W.L., Gros, P., and Grosse-Kunstleve, R.W. (1998). Crystallography and NMR system (CNS): a new software system for macromolecular structure determination. **Acta Crystallog. Sect. D.** 54: 905-921.
- Brzobohaty, B., Moore, I., Kristoffersen, P., Bako, L., Campos, N., Schell, J., and Palme, K. (1993). Release of active cytokinin by a β -glucosidase localized to the maize root meristem. **Science**. 262: 1051-1054.
- Burmeister, W.P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S., and Henrissat B. (1997). The crystal structures of *Sinapis alba* myrosinase and a covalent

- glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinery of an *S*-glycosidase. **Structure**. 5: 663-676.
- Burmeister, W.P. (2000). Structural changes in a cryo-cooled protein crystal owing to radiation damage. **Acta Crystallogr. Sect. D**. 56: 328-341.
- 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.
- Chi, Y.I., Martinez-Cruz, L.A., Jancarik, J., Swanson, R.V., Robertson, D.E., and Kim, S.H. (1999). Crystal structure of the β -glucosidase from hyperthermophile *Thermosphaera aggregans*: insights into its activity and thermostability. **FEBS Lett.** 445: 375-383.
- Chuenchor, W., Pengthaisong, S., Yavaniyama, J., Opassiri, R., Svasti, J., and Ketudat Cairns, J.R. (2006). Purification, crystallization and preliminary X-ray analysis of rice BGlu1 β -glucosidase with and without 2-deoxy-2-fluoro- β -D-glucoside. **Acta Crystallog. Sect. F**. 62: 798-801.
- Cicek, M., Blanchard, D., Bevan, D.R., and Esen, A. (2000). The aglycone specificity-determining sites are different in 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA)-glucosidase (Maize β -glucosidase) and dhurrinase (sorghum β -glucosidase). **J. Biol. Chem.** 275: 20002-20011.
- Coutinho, P.M., and Henrissat, B. (1999). Carbohydrate-active enzymes: an integrated database approach. In "Recent Advances in Carbohydrate Bioengineering", H.J. Gilbert, G. Davies, B. Henrissat and B. Svensson eds., pp. 3-12, The Royal Society of Chemistry, Cambridge, UK.

- Czjzek, M., Cicek, M., Zamboni, V., Bevan, D.R., Henrissat, B., and Esen, A. (2000). The mechanism of substrate (aglycone) specificity in β -glucosidases is revealed by crystal structures of mutant maize β -glucosidase-DIMBOA, -DIMBOAGlc, and -dhurrin complexes. **Proc. Natl. Acad. Sci.** 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) β -glucosidase and a model of its complex with *p*-nitrophenyl β -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* β -glucosidase. **J. Biol. Chem.** 275: 4973-4980.
- Davies, G., and Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolase. **Structure.** 3: 853-859.
- Davies, G.J., Mackenzie, L.F., Varrot, A., Dauter, M., Brzozowski, A.M., Schulein, M., and Withers, S.G. (1998). Snapshots along enzymatic reaction coordinate: Analysis of a retaining β -glycoside hydrolase. **Biochemistry.** 37: 11707-11713.
- DeLano, W.L. (2002). **The PyMOL Molecular Graphics System.** DeLano Scientific, Palo Alto, CA, USA. <http://www.pymol.org>
- Dharmawardhana, D.P., Ellis, B.E., and Carlson, J.E. (1995). A β -glucosidase from lodgepole pine xylem specific for the lignin precursor coniferin. **Plant Physiol.** 107: 331-339.

- Durand, P., Lehn, P., Callebaut, I., Fabrega, S., Henrissat, B. and Mornon, J.P. (1997). Active-site motifs of lysosomal acid hydrolases: invariant features of clan GH-A glycosyl hydrolases deduced from hydrophobic cluster analysis. **Glycobiology**. 7: 277–284.
- el Hassouni, M., Henrissat, B., Chippaux, M., and Barras, F. (1992). Nucleotide sequences of the *arb* genes, which control β -glucoside utilization in *Erwinia chrysanthemi*: comparison with the *E. coli bgl* operon and evidence for a new β -glycohydrolase family including enzymes from eubacteria, archeobacteria, and humans. **J. Bacteriol.** 174: 765-777.
- Falk, A., and Rask, L. (1995). Expression of a zeatin-*O*-glucoside-degrading β -glucoside in *Brassica napus*. **Plant Physiol.** 108: 1369-1377.
- Fenwick, G.R., Heaney, R.K., and Mullin, W.J., (1983). Glucosinolates and their breakdown products in food and food plants. **CRC Crit. Rev. Food Sci. Nutr.** 18:123-201. Günata, Z., (in press). Flavor enhancement in fruit juices and derived beverages by exogenous glycosidases and consequences of the use of enzyme preparations. In J. R. Whitaker, A. G. J. Voragen, and D. Wong (eds.) "**Handbook of Food Enzymology**". Marcel Dekker, Inc.
- Fukuda, K., Mori, H., Okuyama, M., Kimura, A., Ozaki, H., Yoneyama, M., and Chiba, S. (2002). Identification of essential ionizable groups and evaluation of subsite affinities in the active site of β -D-glucosidase F₁ from a *Streptomyces* sp. **Biosci. Biotechnol. Biochem.** 66: 2060-2067.
- Gasteiger, J., and Marsili, M (1980). Interactive partial equalization of orbital electronegativity- a rapid access to atomic charges. **Tetrahedron.** 36: 3219-3228.

- Gloster, T.M., Macdonald, J.M., Tarling, C.A., Stick, R.V., Withers, S.G., and Davies, G.J. (2004a). Structural, thermodynamic, and kinetic analyses of tetrahydrooxazine-derived inhibitors bound to β -glucosidases. **J. Biol. Chem.** 279: 49236-49242.
- Gloster, T.M., Roberts, S., Ducros, V.M., Perugino, G., Rossi, M., Hoos, R., Moracci, M., Vasella, A., and Davies, G.J. (2004b). Structural studies of the β -glucosidase from *Sulfolobus solfataricus* in complex with covalently and noncovalently bound inhibitors. **Biochemistry.** 43: 6101-6109.
- Gloster, T.M., Madsen, R., and Davies, G.J. (2006a). Dissection of conformationally restricted inhibitors binding to a β -glucosidase. **ChemBiochem.** 7: 738-742.
- Gloster, T.M., Roberts, S., Perugino, G., Rossi, M., Moracci, M., Panday, N., Terinek, M., Vasella, A., and Davies, G.J. (2006b). Structural, kinetic, and thermodynamic analysis of glucoimidazole-derived glycosidase inhibitors. **Biochemistry.** 45: 11879-11884.
- Gloster, T.M., Madsen, R., and Davies, G.J. (2007). Structural basis for cyclophellitol inhibition of a β -glucosidase. **Org. Biomol. Chem.** 5: 444-446.
- Hakulinen, N., Paavilainen, S., Korpela, T., and Rouvinen, J. (2000). The crystal structure of β -glucosidase from *Bacillus circulans* sp. *alkalophilus*: ability to form long polymeric assemblies. **J. Struct. Biol.** 129: 69-79.
- He, S., and Withers, S.G. (1997). Assignment of sweet almond β -glucosidase as a family 1 glycosidase and identification of its active site nucleophile. **J. Biol. Chem.** 272: 24864-24867.
- 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 in 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., Callebaut, I., Fabrega, S., Lehn, P., 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. USA.** 92: 7090-7094.
- Hiroimi, K., Nitta, Y., Namura, C., and Ono, S. (1973). Subsite affinities of glucoamylase: examination of the validity of the subsite theory. **Biochem. Biophys. Acta.** 302: 362-375.
- Hommelai, G., Chaiyen, P., and Svasti, J. (2005). Studies on the transglycosylation reactions of cassava and Thai rosewood β -glucosidases using 2-deoxy-2-fluoro-glycosyl-enzyme intermediates. *Arch. Biochem. Biophys.* 442: 11-20.
- Hommelai, G., Withers, S.G., Chuenchor, W., Ketudat Cairns, J., and Svasti, J. (2007). Enzymatic synthesis of cello-oligosaccharides by rice BGlu1 β -glucosidase glycosynthase mutants. **Glycobiology.** 17: 744-753.
- Hosel, W., Tober, I., Eklund, S.,H., and Conn, E.E. (1987). Characterization of β -glucosidases with high specificity for the cyanogeni glucoside dhurrin in *Sorghum bicolor* L. Moench seedlings. **Arch. Biochem. Biophys.** 252: 152-162.
- Hrmova, M., and Fincher, G.B. (1998). Barley β -D-glucan exohydrolases. Substrate specificity and kinetic properties. **Carbohydr. Res.** 305: 209-221.

- Hrmova, M., Gori, R.D., Smith, B.J., Fairweather, J.K., Driguez, H., Varghese, J.N., and Fincher, G.B. (2002). Structural basis for broad substrate specificity in higher plant β -D-glucan glucohydrolases. **The Plant Cell**. 14: 1033-1052.
- Hrmova, M., Harvey, A.J., Wang, J., Shirley, N.J., Jones, G.P., Stone, B.A., Hoj, P.B., and Fincher, G.B. (1996). Barley β -D-glucan exohydrolases with β -D-glucosidase activity. **J. Biol. Chem.** 271: 5277-5286.
- Hrmova, M., MacGregor, E.A., Biely, P., Stewart, R.J., and Fincher, G.B. (1998). Substrate binding and catalytic mechanism of a barley β -D-glucosidase/ (1, 4)- β -D-glucan exohydrolase. **J. Biol. Chem.** 273: 11134-11143.
- Hughes, M.A., Brown, K., Pancoro, A., Murray, B.S., Oxtoby, E., and Hughes, J. (1992). A molecular and biochemical analysis of the structure of the cyanogenic β -glucosidase (linamarase) from cassava (*Manihot esculenta* Cranz). **Arch. Biochem. Biophys.** 95: 273-279.
- Hughes, M.A., and Dunn, M.A. (1982). Biochemical characterization of the *Li* locus, which controls the activity of the cyanogenic β -glucosidase in *Trifolium repens* L. **Plant Mol. Biol.** 1: 169-181.
- Husebye, H., Arzt, S., Burmeister, W.P., Hartel, F.V., Brandt, A., Rossiter, J.T., and Bones, A.M. (2005). Crystal structure at 1.1 Angstroms resolution of an insect myrosinase from *Brevicoryne brassicae* shows its close relationship to β -glucosidases. **Insect Biochem. Mol. Biol.** 35: 1311-1320.
- Isorna, P., Polaina, J., Latorre-Garcia, L., Canada, F.J., Gonzalez, B., and Sanz-Aparicio, J. (2007). Crystal structures of *Paenibacillus polymyxa* β -glucosidase B complexes reveal the molecular basis of substrate specificity

- and give new insights into the catalytic machinery of family I glycosidases. **J. Mol. Bio.** 371: 1204-1218.
- Jenkins, J., Lo Leggio, L., Harris, G., and Pickersgill, R. (1995). β -glucosidase, β -galactosidase, family A cellulases, family F xylanases and two barley glycanases form a superfamily of enzymes with 8-fold β/α architecture and with two conserved glutamates near the carboxy-terminal ends of β -strands four and seven. **FEBS Lett.** 362: 281-285.
- Jones, D.A. (1988). Cyanogenesis in animal-plant interactions. In: D. Evered and S. Harnett (Editors), Cyanide compounds in biology. **Ciba Foundation Symposia.** 140: 151-170 New York.
- Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. **Acta Crystallog. Sect. A.** 47: 110-119.
- Kempton, J.B., and Withers, S.G. (1992). Mechanism of *Agrobacterium* β -glucosidase: Kinetic studies. **Biochemistry.** 31: 9961-9969.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature.** 227: 680-685.
- Lamzin, V.S., and Wilson, K.S. (1993). Automated refinement of protein models. **Acta Crystallog. Sect. D.** 49: 129-147.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. **J. Appl. Crystallog.** 26: 283-291.

- Lawson, S.L., Warren, R.A.J., and Withers, S.G. (1998). Mechanistic consequences of replacing the active-site nucleophile Glu-358 in *Agrobacterium* sp. β -glucosidase with a cysteine residue. **Biochemical J.** 330: 203-209.
- Leah, R., Kigel, J., Svedsen, I., and Mundy, J. (1995). Biochemical and molecular characterization of a barley seed β -glucosidase. **J. Biol. Chem.** 270: 15789-15797.
- Leatherbarrow, R. J. (2001). *GraFit*, 5 Ed., Erithacus Software Ltd., Horley, UK.
- Ly, H.D. and Withers, S.G. (1999). Mutagenesis of glycosidases. **Annu. Rev. Biochem.** 68: 487-522.
- MacLeod, A.M., Lindhorst, T., Withers, S.G., and Warren, R.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.
- MacLeod, A.M., Tull, D., Rupitz, K., Warren, R.A.J., and Withers, S.G. (1996). Mechanistic consequences of mutation of active site carboxylates in a retaining β -1,4-glycanase from *Cellulomonas fimi*. **Biochemistry.** 35: 13165-13172.
- Makropoulou, M., Christakopoulos, P., Tsitsimpikou, C., Kekos, D., Kolis, F.N., and Macris, B.J. (1998). Factors affecting the specificity of β -glucosidase from *Fusarium oxysporum* in enzymatic synthesis of alkyl- β -D-glucosides. **Int. J. Biol. Macromol.** 22: 97-101.
- Matsuzaki, T., and Koiwai, A. (1986). Germination inhibition in stigma extracts of tobacco and identification of MeABA, ABA and ABA- β -D-glucopyranoside. **Agric. Biol. Chem.** 50: 2193-2199.

- McCarter, J., and Withers, S.G. (1994). Mechanisms of enzymatic glycoside hydrolysis. **Curr. Opin. Struc. Biol.** 4: 885-892.
- McIntosh, L.P., Hand, G., Johnson, P.E., Joshi, M.D., Körner, M., Plesniak, L.A., Ziser, L., Wakarchuk, W.W., and Withers, S.G. (1996). The pK_a of the general acid/base carboxyl group of a glycosidase cycles during catalysis: A ¹³C-NMR study of *Bacillus circulans* xylanase. **Biochemistry.** 35: 9958-9966.
- Meagher, J.L., Winter, H.C., Ezell, P., Goldstein, I.J., and Stuckey, J.A. (2005). Crystal structure of banana lectin reveals a novel second sugar binding site. **Glycobiology.** 15: 1033-1042.
- Mildvan, A.S., Weber, D.J., and Kuliopulos, A. (1992). Quantitative Interpretations of Double Mutations of Enzymes. **Arch. Biochem. Biophys.** 294: 327-340.
- Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., and Olson, A J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. **J. Comp. Chem.** 19: 1639-1662.
- Müllegger, J., Jahn, M., Chen, H.M., Warren, R.A.J., and Withers, S.G. (2005). Engineering of a thioglycoligase: randomized mutagenesis of the acid-base residue leads to the identification of improved catalysis. **Protein Eng. Des. Sel.** 18: 33-40.
- Murshudov, G.N., Lebedev, A., Vagin, A.A., Wilson, K.S., and Dodson, E.J. (1999). Efficient anisotropic refinement of macromolecular structures using FFT. **Acta. Crystallog. Sect. D.** 55: 247-255.

- Namchuck, N.M. and Withers, S.G. (1995). Mechanism of *Agrobacterium* β -glucosidase: kinetic analysis of the role of noncovalent enzyme/substrate interactions. **Biochemistry**. 34: 16194-16202.
- Navaza, J. (1994). AMoRe: an automated package for molecular replacement. **Acta Crystallog. Sect. A**. 50: 157-163.
- Niemeyer, H.M. (1988). Hydroxamic acids (4-hydroxy- 1, 4-benzoxazin-3-ones). Defense chemicals in the Gramineace. **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 β -glucosidase BGL1A from the basidiomycete *Phanerochaete chrysosporium*. **FEBS Lett**. 581: 1514-1520.
- Opassiri, R., Hua, Y., Wara-Aswapati, O., Akiyama, T., Svasti, J., Esen, A., and Ketudat Cairns, J.R. (2004). β -glucosidase, exo- β -glucanase and pyridoxine transglucosylase activities of rice BGlu1. **Biochem. J**. 379: 125-131.
- Opassiri, R., Ketudat Cairns, J.R., Akiyama, T., Wara-Aswapati, O., Svasti, J., and Esen, A. (2003). Characterization of rice β -glucosidase highly expressed in flower and germinating shoot. **Plant Sci**. 165: 627-638.
- Opassiri, R., Pomthong, B., Onkoksoong, T., Akiyama, T., Esen, A., and Ketudat Cairns, J.R. (2006). Analysis of rice glycosyl hydrolase family I and expression of Os4bglu12 β -glucosidase. **BMC Plant Biol**. 6: 33.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. In **Methods in Enzymology, Volume 276: Macromolecular Crystallography, part A** (Carter, C.W. & Sweet, R.M., eds), pp. 307-326, Academic Press, New York.

- Oxtoby, E., Dunn, M.A., Pancoro, A., and Hughes, M.A. (1991). Nucleotide and derived amino acid sequence of the cyanogenic β -glucosidase (linamarase) from white clover (*Trifolium repens* L.) **Plant Mol. Bio.** 17: 209-219.
- Palmiano, E.P., and Juliano, B.O. (1973). Changes in the activity of some hydrolases, peroxidase and catalase in rice seed during germination. **Plant Physiol.** 52: 274-277.
- Parsiegla, G., Reverbel-Leroy, C., Tardif, C., Belaich, J.P., Driguez, H., and Haser, R. (2000). Crystal structures of the Cellulase Cel48F in complex with inhibitors and substrates give insights into its processive action. **Biochemistry.** 39: 11238-11246.
- Pflugrath, J.W. (1999). The finer things in X-ray diffraction data collection. **Acta Crystallogr. Sect. D.** 55: 1718-1725.
- Poulton, J.E. (1990). Cyanogenesis in plants. **Plant Physiol.** 94: 401-405.
- Reese, E.T. (1977). Degradation of polymeric carbohydrates by microbial enzymes. **Rec. Adv. Phytochemistry.** 11: 311- 364. Quoted in Cicek, M., and Esen, A. (1998). Structure and expression of a dhurrinase (β -glucosidase) from sorghum. **Plant Physiol.** 116: 1469-1478.
- Sanz-Aparicio, J., Hermoso, J.A., Martinez-Ripoll, M., Lequerica, J.L., and Polaina, J. (1998). Crystal structure of β -glucosidase A from *Bacillus polymyxa*: insights into the catalytic activity in family 1 glycosyl hydrolases. **J. Mol. Bio.** 275: 491-502.
- Schliemann, W. (1984). Hydrolysis of conjugated gibberellins by β -glucosidases from dwarf rice (*Oryza sativa* L. cv. Tan-ginbozu). **J. Plant Physiol.** 116: 123-132.

- Simos, G. Panagiotidis, C.A., Skoumbas, A., Choli, D., Ouzounis, C., and Georgatsos, G.J. (1994). Barley β -glucosidase: Expression during seed germination and maturation and partial amino acid sequences. **Biochem. Biophys. Acta.** 1199: 52-58.
- Sinnott, M.L. (1990). Catalytic mechanism of enzymatic glycosyl transfer. **Chemical Reviews.** 90: 1171-1202.
- Smith, A.R., and Van Staden, J. (1978). Changes in endogenous cytokinin levels in kernels of *Zea mays* L. during inhibition and germination. **J. Exp. Bot.** 29: 1067-1075.
- Street, I.P., Kempton, J.B., and Withers, S.G. (1992). Inactivation of a β -glucosidase through the accumulation of a stable 2-deoxy-2-fluoro- α -D-glucopyranosyl-enzyme intermediate: a detailed investigation. **Biochemistry.** 31: 9970-9978.
- Sue, M., Yamazaki, K., Yajima, S., Nomura, T., Matsukawa, T., Iwamura, H., and Miyamoto, T. (2006). Molecular and structural characterization of hexameric β -D-glucosidases in wheat and rye. **Plant Physiol.** 141: 1237-1247.
- Surarit, R., Svasti, J., Srisomsap, C., Suginta, W., Khunyoshyeng, S., Nilwarangkoon, S., Harnsakol, P., and Benjavongkulchai, E. (1995). Screening of glycosylhydrolase enzymes in Thai plant seeds for potential use in oligosaccharide synthesis. **J. Sci. Soc. Thailand.** 21: 293-303.
- Svasti, J., Phongsak, T., and Saranthima, R. (2003). Transglycosylation of tertiary alcohols using cassava β -glucosidase. **Biochem. Bioph. Res. Co.** 305: 470-475.

- Tribolo, S., Berrin, J.G., Kroon, P.A., Czjzek, M., and Juge, N. (2007). The Crystal structure of human cytosolic β -glucosidase unravels the substrate aglycone specificity of a family 1 glycoside hydrolase. **J. Mol. Biol.** 370: 964.
- Trimbur, D.E., Warren, R.A.J., and Withers, S.G. (1992). Region-directed mutagenesis of residues surrounding the active site nucleophile in β -glucosidase from *Agrobacterium faecalis*. **J. Biol. Chem.** 267: 10248-10251.
- Vallmitjana, M., Ferrer-Navarro, M., Planell, R., Abel, M., Ausín, C., Querol, E., Planas, A., and Pérez-Pons, J.A. (2001). Mechanism of the family 1 β -glucosidase from *Streptomyces* sp: Catalytic residues and kinetic studies. **Biochemistry.** 40: 5975-5982.
- Verdoucq, L., Czjzek, M., Moriniere, J., Bevan, D.R., and Esen, A. (2003). Mutational and structural analysis of aglycone specificity in maize and sorghum β -glucosidases. **J. Biol. Chem.** 278: 25055-25062.
- Verdoucq, L., Moriniere, J., Bevan, D.R., Esen, A., Vasella, A., Henrissat, B., and Czjzek, M. (2004). Structural determinants of substrate specificity in family 1 β -glucosidases: Novel insights from the crystal structure of sorghum dhurrinase-1, a plant β -glucosidase with strict specificity, in complex with its natural substrate. **J. Biol. Chem.** 279: 31796-31803.
- Vincent, F., Gloster, T.M., Macdonald, J., Morland, C., Stick, R.V., Dias, F.M., Prates, J.A., Fontes, C.M., Gilbert, H.J., and Davies, G.J. (2004). Common inhibition of both β -glucosidases and β -mannosidases by isofagomine lactam reflects different conformational itineraries for pyranoside hydrolysis. **ChemBiochem.** 5: 1596-1599.

- Wang, Q., Graham, R.W., Trimbur, D., Warren, R.A.J., and Withers, S.G. (1994). Changing enzymatic reaction mechanisms by mutagenesis: Conversion of a retaining glucosidase to an inverting enzyme. **J. Am. Chem. Soc.** 116: 11594-1595.
- Wang, X., He, X., Yang, S., An, X., Chang, W., and Liang, D. (2003). Structural basis for Thermostability of β -glycosidase from the *Thermophilic Eubacterium Thermus nonproteolyticus* HG102. **J. Bacteriol.** 185: 4248-4255.
- Wang, Q., Trimber, D., Grahan, R., Warren, R.A.J. and Withers, S.G. (1995). Identification of the acid/base catalyst in *Agrobacterium faecalis* β -glucosidase by kinetic analysis of mutants. **Biochemistry.** 34: 14554-14562.
- Wattenberg, L.W., (1971). Studies on polycyclic hydrocarbon hydroxylases of the intestine possibly related to cancer. Effect of diet on benzopyrene hydroxylase activity. **Cancer.** 28: 99-102.
- Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S., and Weiner, P. (1984). A new force field for molecular mechanical simulation of nucleic acids and proteins. **J. Am. Chem. Soc.** 106: 765-784.
- White, A. and Rose, D. (1997). Mechanism of catalysis by retaining β -glycosyl hydrolased. **Curr. Opin. Struc. 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 β -1, 4-glycanase Cex from *Cellulomonas fimi*. **Biochemistry.** 33: 12546-12552.

- Wiese, G., and Grambow, H. (1986). Indole-3-methanol- β -D-glucoside and indole-3-carboxylic acid β -D-glucoside are products of indole-3-acetic acid degradation in wheat leaf segments. **Phytochemistry**. 25: 2451-2455.
- Wiesmann, C., Beste, G., Hengstenberg, W., Schulz, G.E. (1995). The three-dimensional structure of 6-phospho- β -galactosidase from *Lactococcus lactis*. **Structure**. 3: 961-968.
- Wiesmann, C., Hengstenberg, W., and Schuz, G.E. (1997). Crystal structures and mechanism of 6-phospho- β -galactosidase from *Lacococcus lactis*. **J. Mol. Biol.** 269: 851-860.
- Withers, S.G. (1999). Understanding and exploiting glycosidases. **Canadian J. Chem.** 77: 1-11.
- Withers, S.G. (2001). Mechanisms of glycosyl transferases and hydrolases. **Carbohydr. Pol.** 44: 325-337.
- Withers, S.G., and Aebersold, R. (1995). Approaches to labeling and identification of active site residues in glycosidases. **Protein Sci.** 4: 361-372.
- Withers, S.G., and Street, I.P. (1988). Identification of a covalent α -D-glucopyranosyl enzyme intermediate formed on a β -glucosidase. **J. American Chem. Soc.** 110: 8551-8553.
- Withers, S.G., Warren, R.A.J., Street, I.P., Rupitz, K., Kempton, J.B., and Aebersold, R. (1990). Unequivocal demonstration of the involvement of a glutamate residue as a nucleophile in the mechanism of a "retaining" glycosidase. **J. American Chem. Soc.** 112: 5887-5889.
- Woodward, J., and Wiseman, A. (1982). Fungal and other β -D-glucosidases: their properties and applications. **Enzyme Microb. Technol.** 4: 73.

- Xu, Z., Escamilla-Trevino, L., Zeng, L., Lalgondar, M., Bevan, D.R., Winkel, B.S.J., Mohamed, A., Cheng, C., Shih, M., Poulton, J.E., Esen, A. (2004). Functional genomic analysis of *Arabidopsis thaliana* Glycoside hydrolase family 1. **Plant Mol. Biol.** 55: 343-367.
- Yazaki, T., Ohnishi, M., Rokushika, S., and Okada, G. (1997). Subsite structure of the β -glucosidase from *Aspergillus niger*, evaluated by steady-state kinetics with cello-oligosaccharides as substrates. **Carbohydr. Res.** 298: 51-57.
- Zechel, D.L., Boraston, A.B., Gloster, T., Boraston, C.M., Macdonald, J.M., Tilbrook, D.M., Stick, R.V., and Davies, G.J. (2003). Iminosugar glycosidase inhibitors: structural and thermodynamic dissection of the binding of isofagomine and 1-deoxynojirimycin to β -glucosidases. **J. Am. Chem. Soc.** 125: 14313-14323.
- Zechel, D.L., Reid, S.P., Nashiru, O., Mayer, C., Stoll, D., Jakeman, D.L., Warren, R.A.J., and Withers, S.G. (2001). Enzymatic synthesis of carbon-fluorine bonds. **J. Am. Chem. Soc.** 123: 4350-4351.
- Zechel, D.L., and Withers, S.G. (1999). Mechanisms of glycosyl transfer. In: Barton, D., Nakanishi, K. and Poulter, C.D. (Eds.), **Comprehensive Natural Products Chemistry**. Vol. 5. Elsevier, Chapter 12, 279-314.
- Zechel, D.L. and Withers, S.G. (2001). Dissection of nucleophilic and acide-base catalysis in glycosidases. **Curr. Opin. Chem. Biol.** 5: 643-649.
- Zouhar, J., Vevodova, J., Marek, J., Damborsky, J., Su, X.-D., and Brzobohaty, B. (2001). Insights into the functional architecture of the catalytic center of a maize β -glucosidase Zm-p60.1 **Plant Physiol.** 127: 973-985.

CURRICULUM VITAE

Name: Miss Watchalee CHUENCHOR

Date of Birth: September 10th, 1977

Place of Birth: Nakhon Ratchasima, Thailand

Education:

2001- 2007: School of Biochemistry, Suranaree University of Technology,
Nakhon Ratchasima, Thailand

Ph.D. in Biochemistry

1995-1999: Biochemistry Department, Khon Kaen University, Khon Kaen,
Thailand

B.Sc. in Biochemistry

Fellowships:

2003- 2007 The Royal Golden Jubilee Ph.D. Scholarship, from Thailand
Research Fund, Thailand.