ISOLATION AND CHARACTERIZATION

OF RICE β-GLUCOSIDASE

Miss Rodjana Opassiri

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in Environmental Biology

Suranaree University of Technology

Academic Year 2003

ISBN 974-533-263-1

การแยกและการศึกษาคุณสมบัติของเบต้ากลูโคซิเดสจากข้าว

นางสาวรจนา โอภาสศิริ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาชีววิทยาสิ่งแวดล้อม

มหาวิทยาลัยเทคโนโลยีสุรนารี

ปีการศึกษา 2546

ISBN 974-533-263-1

ISOLATION AND CHARACTERIZATION

OF RICE β-GLUCOSIDASE

Miss Rodjana Opassiri

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in Environmental Biology

Suranaree University of Technology

Academic Year 2003

ISBN 974-533-263-1

ISOLATION AND CHARACTERIZATION OF

RICE β-GLUCOSIDASE

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for a Doctoral Degree

Thesis Examining Committee

....
(Assoc. Prof. Dr. Korakod Indrapichate)
Chairperson

....
(Asst. Prof. Dr. James R. Ketudat-Cairns)
Member (Thesis Advisor)

....
(Prof. Dr. Asim Esen)
Member

....
(Prof. Dr. M.R. Jisnuson Svasti)
Member

....
(Assoc. Prof. Dr. Onnop Wara-Aswapati)
Member

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

....
(Assoc. Prof. Dr. Tawit Chitsomboon)
Vice Rector for Academic Affairs

....
(Assoc. Prof. Dr. Prasart Suebka)
Dean of the Institute of Science

รจนา โอภาสศิริ: การแยกและการศึกษาคุณสมบัติของเบต้ากลูโคซิเดสจากข้าว (ISOLATION AND CHARACTERIZATION OF RICE β-GLUCOSIDASE) อาจารย์ที่ปรึกษา: ผศ. ดร. เจมส์ เกตุทัต-คาร์นส์, 247 หน้า.

ISBN 974-533-263-1

สารพันธุกรรมที่บรรจุรหัสทางพันธุกรรม (cDNA) สำหรับเบต้ากลูโคซิเดสของข้าว 2 ไอโซไซม์ ได้แก่ *bglu1* และ *bglu2* ได้ถูกเพิ่มปริมาณและหาลำดับนิวคลีโอไทด์ Southern analysis บ่งชี้ว่า *bglu1* และ *bglu2* เป็นยีนที่มีหนึ่งชุดบนโครโมโซม Northern blot analysis บ่งชี้ว่า *bglu1* แสดงออกมากในใบของต้นอ่อนและดอก ส่วน *bglu 2* แสดงออกมากในใบของต้นอ่อน สภาพแวดล้อมมีผลต่อการแสดงออกของยีนทั้งสองในต้นอ่อนของข้าว เบต้ากลูโคซิเดสไอโซไซม์ที่ 1 ถูกผลิตขึ้นใน *E. coli* ในรูปของโปรตีนที่ต่ออยู่กับโปรตีนไทโอรีดอกซิน ในสภาพที่เร่งปฏิกิริยาได้ ส่วนเบต้ากลูโคซิเดสไอโซไซม์ที่ 2 อยู่ในสภาพของโปรตีนที่เกาะกลุ่มกัน เบต้ากลูโคซิเดสไอโซไซม์ที่ 1 ย่อยสลายโอลิโกแซคคาไรด์ที่มีพันธะ β-(1,2), -(1,3), -(1,4) และ -(1,6) ได้ เอนไซม์นี้ยังย่อยสลายพาราไนโตรฟีนอลเบต้าดีไกลโคไซด์ และกลูโคไซด์บางชนิดที่พบในธรรมชาติได้ เบต้ากลูโคซิเดสไอโซไซม์ที่ 1 เร่งปฏิกิริยาการนำกลูโคสไปเชื่อมต่อกับโมเลกุลของโอลิโกแซคคาไรด์ พาราไนโตรฟีนอลเบต้าดีกลูโคไซด์ เอทธิลแอลกอฮอล์ และไพริดอกซิน เอนไซม์นี้ถูกยับยั้งการเร่งปฏิกิริยาโดยกลูโคโนแลคโตน สารนี้ยังสามารถยับยั้งการงอกของเมล็ดข้าวได้ การวิเคราะห์ลำดับนิวคลีโอไทด์ของฐานข้อมูลจีโนมข้าวพบว่ามียีนของเบต้ากลูโคซิเดสจำนวน 49 ยีน จากการวิเคราะห์ ESTs พบว่ายีนที่มีการแสดงออกมีจำนวน 26 ยีน

สาขาวิชาชีววิทยา | ลายมือชื่อนักศึกษา...

ปีการศึกษา 2546 | ลายมือชื่ออาจารย์ที่ปรึกษา...................................

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม............................

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม............................

RODJANA OPASSIRI: ISOLATION AND CHARACTERIZATION OF RICE β-GLUCOSIDASE

THESIS ADVISOR: ASSISTANT PROFESSOR JAMES R. KETUDAT-CAIRNS, Ph.D. 247 PP. ISBN 974-533-263-1

The cDNAs for two β-glucosidase isozymes from rice (*Oryza sativa* L.), designated *bglu1* and *bglu2*, were cloned and sequenced. Southern analysis indicated that *bglu1* and *bglu2* were single copy genes. Northern blot analysis indicated that *bglu1* was highly expressed in germinating shoot and flower, while *bglu2* was highly expressed in germinating shoot. Environmental conditions had effects on both gene transcript levels in rice seedlings. BGlu1 was expressed as an active thioredoxin fusion protein in *E. coli*, while most of BGlu2 was in inclusion bodies. BGlu1 hydrolyzed β-(1,2), -(1,3), -(1,4) and -(1,6) -linked oligosaccharides. This enzyme also hydrolyzed p-nitrophenol β-D-glycosides and some natural glucosides. BGlu1 showed transglucosylation activity toward oligosaccharides, p-nitrophenol β-D-glucoside, ethyl alcohol and pyridoxine. This enzyme was inhibited by glucono-1,5-lactone, which also inhibited rice germination. Analysis of the rice genome database showed 49 β-glucosidase genes, 26 of which were expressed based on EST analysis.

School of Biology | Signature of Student..................................

Academic year 2003 | Signature of Advisor.................................

Signature of Co-Advisor............................

Signature of Co-Advisor............................

Acknowledgements

I would like to express the deepest gratitude to my advisor, Asst. Prof. Dr. James R. Ketudat-Cairns for kindly providing me an opportunity to work in plant molecular biology and biochemistry and his valuable advice. I am also deeply grateful to my co-advisors, Prof. Dr. Asim Esen and Assoc. Prof. Dr. Onnop Wara-Aswapati for their valuable advice and guidance along this project. I am very grateful to Prof. Dr. M.R. Jisnuson Svasti, for his valuable suggestions.

I wish to thank Dr. Takashi Akiyama for his help with the Southern and Northern blot analysis, Asst. Prof. Dr. Mariena Ketudat-Cairns and Voraratt Champattanachai for advice on research techniques, and Yanling Hna for NMR structural analysis of pyridoxine glucoside. Sukanda Wichitphan is acknowledged for first identifying the rice *bglu2* cDNA fragment. I am also grateful to Chulabhorn Research Institute, Biochemistry Laboratory for kindly providing facilities. Special thanks are extended to all friends in the School of Biochemistry and Biology for their help. I wish to express my special thanks to the Thailand Research Fund (TRF) for providing the Royal Golden Jubilee Ph.D Scholarship and BGJ Research Grant for my study in Suranaree University of Technology, and to the Institute for the Promotion of Teaching Science and Technology for offering the DPST Scholarship for my study in Khon Kaen University and Suranaree University of Technology.

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

Rodjana Opassiri

Contents

Page

Abstract in Thai I

Abstract in English II

Acknowledgements III

Contents IV

List of Tables XII

List of Figures XIV

List of Abbrevations XX

Chapters

I Introduction 1

1.1 Overview of β-glucosidases 1

1.2 Catalytic mechanism of β-glucosidases 4

1.3 Substrate specificity of plant glycosyl hydrolases 7

1.4 Plant β-glucosidase purification and characterization 8

1.5 Purification and substrate specificity of rice β-glucosidases 9

1.6 Environmental effects on plant β-glucosidase expression 11

1.7 Glycosides 12

1.8 Transglycosylation activity of β-glucosidases 13

1.9 The rice genome 15

1.10 Rice germination 17

1.11 Objectives 20

Contents (Continued)

Page

II Materials and Methods 21

2.1 Materials 21

2.1.1 Plant materials 21

2.1.2 Chemicals 21

2.1.3 PCR primers 22

2.2 General methods 23

2.2.1 RNA extraction 23

2.2.2 Determination of quantity and purity of RNA 24

2.2.3 Purification of Poly A^+ RNA 24

2.2.4 First-strand cDNA synthesis by ThermoScript RT-PCR system 25

2.2.5 DNA analysis by agarose gel electrophoresis 26

2.2.6 Purification of PCR products from agarose gels 27

2.2.7 Quantitation and expected yield of DNA 27

2.2.8 Vector preparation for ligation reaction 28

2.2.9 Ligation of DNA fragment into vector 29

2.2.10 Transformation and selection 29

2.2.11 Plasmid isolation by boiling miniprep 30

2.2.12 Restriction enzyme digestion for plasmid analysis 31

2.2.13 QIAGEN plasmid miniprep 32

2.2.14 DNA Sequencing 32

2.2.15 Analysis of DNA sequences 33

Contents (Continued)

Page

2.3 Amplification of rice β-glucosidase cDNAs 34

2.3.1 Cloning of rice β-glucosidase cDNA for isozyme 1 34

2.3.2 Cloning of rice β-glucosidase cDNA for isozyme 2 37

2.4 Southern blot analysis 40

2.4.1 Genomic DNA isolation and restriction enzyme digestion 40

2.4.2 Transferring genomic DNA from agarose gel to nylon membrane 41

2.4.3 DNA probe preparation 42

2.4.4 DNA probe labeling 43

2.4.5 DNA-DNA probe hybridization, post hybridization stringency wash, and autoradiography 44

2.5 Northern blot analysis to determine β-glucosidase gene expression 45

2.5.1 Growth conditions 45

2.5.2 RNA extraction 46

2.5.3 RNA gel electrophoresis and transferring RNA from agarose gels to nylon membranes 47

2.5.4 DNA probe preparation 48

2.5.5 RNA-DNA probe hybridization 48

2.5.6 Determination of signal intensity on autoradiography film by densitometry technique 49

Contents (Continued)

Page

2.6 Semiquantitative RT-PCR 49

2.7 Recombinant protein expression in *Escherichia coli* 51

2.7.1 Amplification of the cDNAs to insert in expression vectors 51

2.7.2 Construction of recombinant plasmids 54

2.7.3 Identification and isolation of recombinant expression plasmids 55

2.7.4 Transformation of recombinant plasmid into expression host cells 55

2.7.5 Protein expression in *Escherichia coli* 57

2.7.6 Extraction of recombinant protein from induced cells 58

2.7.7 Purification of recombinant protein by Ni-NTA superflow 59

2.7.8 Cleavage of recombinant protein 59

2.7.9 Protein determination 60

2.7.10 Determination of the subunit molecular weight of protein by SDS polyacrylamide gel electrophoresis 60

2.7.11 Estimation of the native molecular weight of protein by gel filtration chromatography.. 61

2.7.12 Western blot analysis 63

2.8 Biochemical characterization 64

2.8.1 β-glucosidase assays and kinetics study 64

2.8.2 pH and temperature profiles of activity and stability 65

2.8.3 TLC 66

Contents (Continued)

Page

2.8.4 Determination of the kinetic parameters of pyridoxine glucoside hydrolyzed by BGlu1 using high performance liquid chromatography (HPLC) 68

2.8.5 Non-denaturing gel electrophoresis for β-glucosidase gel activity staining 69

2.8.6 Inhibition studies 70

2.8.7 Effect of reducing and denaturing agents on BGlu1 activity ... 71

2.8.8 Effects of metal ions on enzyme activity 71

2.9 Transglycosylation reactions 72

2.10 The effect of glucono-1,5-lactone on rice germination 74

2.11 Genomic analysis 75

2.11.1 Database searching and sequence analysis 75

2.11.2 Digital northern analysis 76

III Results 78

3.1 β-Glucosidase cDNA cloning 78

3.2 Nucleic acid-blot analysis 95

3.3 Temporal and spatial expression and the environmental effects on β-glucosidase gene expression 97

3.3.1 Spatial expression pattern in seedling and flowering stage 97

3.3.2 Temporal expression in seedlings during germination stage ... 97

Contents (Continued)

Page

3.3.3 Effects of environmental conditions on rice β-glucosidase gene expression 100

3.4 Semiquantitative RT-PCR analysis 103

3.5 Expression of rice β-glucosidases in *E. coli* 107

3.5.1 Recombinant protein expression of rice BGlu1 107

3.5.2 Recombinant protein expression of rice BGlu2 113

3.6 Biochemical properties of rice BGlu1 117

3.6.1 Substrate specificity 117

3.6.2 Molecular weight (Mr) estimation 129

3.6.3 pH and temperature profile of activity and stability 129

3.6.4 Inhibition studies 133

3.6.5 Effect of reducing agents, denaturants and alcohol on enzyme activity 135

3.7 Transglycosylation activity of rice BGlu1 138

3.8 Inhibition of rice germination by a β-D-glucosidase inhibitor 147

3.9 Identification of rice β-glucosidase genes 150

IV Discussion 165

4.1 Cloning and expression of rice β-glucosidase cDNAs 165

4.2 Expression of rice β-glucosidase 167

4.3 Recombinant protein expression of rice β-glucosidase 170

Contents (Continued)

Page

4.4 Catalytic properties of rice BGlu1 173

4.5 Possible function of rice BGlu1 176

4.6 Deduced protein sequence similarity between rice and barley

β-glucosidases 177

4.7 Biochemical properties of rice BGlu1 178

4.7.1 Molecular weight and pI 178

4.7.2 pH and temperature profile of stability and activity 179

4.7.3 Inhibition studies 180

4.8 Transglycosylation activity of rice BGlu1 182

4.8.1 Transglycosylation of pNP β-D-glucoside 182

4.8.2 Transglycosylation of ethyl alcohol 183

4.8.3 Transglycosylation of pyridoxine 185

4.9 Inhibition of rice germination by β-glucosidase inhibitor 187

4.10 Rice β-glucosidase gene family 189

V **Conclusion** 192

References 196

Appendices 215

Appendix A Multiple sequence alignment and standard curve...................... 216

Appendix B Plasmid maps 224

Contents (Continued)

Page

Appendix C Competent cell preparation and antibody production 229

Appendix D Solution preparation 235

Curriculum Vitae 247

List of Tables

Table **Page**

2.1 Primers used for amplification of the *bglu1* cDNA 37

2.2 Primers used for sequencing the *bglu1* cDNA 37

2.3 Primers used for amplification of the *bglu2* cDNA 39

2.4 Primers used for sequencing the *bglu2* cDNA 39

2.5 Primers used for amplification of *bglu1*and *bglu2* probes 43

2.6 Primers used for amplification of the 3' UTR of *bglu1*, *bglu2* and β-actin cDNA fragments 51

2.7 Primers used for amplification of *bglu1* and *bglu2* cDNA 54

2.8 pET vector characteristics 56

2.9 pET system host strain characteristics 57

2.10 Concentration of protein solutions used for gel filtration 62

3.1 Predicted properties of rice β-glucosidases 90

3.2 Expression of rice β-glucosidase genes judged from corresponding ESTs ... 93

3.3 Comparison of transcript levels of *bglu1* and *bglu2* in rice seedlings treated with various environmental conditions 102

3.4 Comparison of specific activity between the recombinant BGlu1 expressed in various *E. coli* systems 111

3.5 Comparison of specific activity between the recombinant BGlu2 expressed in various *E. coli* systems 114

3.6 Kinetic parameters of BGlu1 during the hydrolysis of oligosaccharides 118

List of Tables (Continued)

Table **Page**

3.7 Kinetic parameters of BGlu1 during the hydrolysis of pNP-glycosides, alkyl glucosides and natural glucoside substrates 118

3.8 Glycosides which were not cleaved by BGlu1 122

3.9 Effect of inhibitors on activity of purified rice BGlu1 133

3.10 Effect of metal ions on activity of purified rice BGlu1 134

3.11 Yields of transfer products produced by BGlu1 in transglycosylation using pNPG as glucose donor and acceptor 140

3.12 Yields of transfer products produced by BGlu1 in transglycosylation using pNPG as glucose donor and ethyl alcohol as glucose acceptor 140

3.13 Yields of transfer products produced by BGlu1 in transglycosylation using pNPG as glucose donor and pyridoxine as glucose acceptor 143

3.14 Summary of identified genes homologous to glycosyl hydrolase family 1 glucosidase from the Beijing Genomic Institute (BGI), GenBank at NCBI of rice CGI (CGI/NCBI), the Monsanto Rice Genome Sequencing (MST) Project, and the Torrey Mesa Research Institute (TMRI) database 156

3.15 Predicted rice β-glucsidase protein properties and the gene patterns 159

3.16 Expression of rice beta-glucosidase genes judged from corresponding ESTs 162

List of Figures

Figure **Page**

1.1 Generally accepted endocyclic pathway of the double displacement mechanism proposed for retaining β-glucosidases 6

1.2 Kinetics of hydrolysis or glycosyltransfer reactions catalysed by a plant family 1 β-glucosidase 14

3.1 Rice EST clones corresponding to *bglu1* cDNA 79

3.2 Cloning strategies of rice *bglu1* from Orion rice cDNA 80

3.3 1% agarose gel electrophoresis of the 3′ RACE PCR product of *bglu1* cDNA fragment amplified with β-glu163 and oligo $(dT)_{17}$ primers 81

3.4 1% agarose gel electrophoresis of 5′ end PCR product of *bglu1* cDNA fragment amplified with the β-glu156 and β-glu178 primers 82

3.5 1% agarose gel electrophoresis of 5′ end PCR product of *bglu1* cDNA fragment amplified with the β-glu217 and β-glu178 primers 83

3.6 1% agarose gel electrophoresis of PCR product of full coding *bglu1* cDNA and its 5′ UTR amplified with the β-glu217 and Ro13pr primers 84

3.7 The full-length cDNA sequence and deduced amino acid sequences of rice *bglu1* 85

3.8 Cloning strategies of rice *bglu2* from Orion rice 86

3.9 1% agarose gel electrophoresis of 5′ end PCR product of the *bglu2* cDNA fragment amplified with the primers β-glu156 and β-glu41 87

List of Figures (Continued)

Figure **Page**

3.10 1% agarose gel electrophoresis of the *bglu2* 3′RACE product (amplified with the β-glu176 and oligo $(dT)_{17}$ primers) and full-length *bglu2* cDNA (amplified with the β-glu214 and Ro23pr primers) 88

3.11 The full-length cDNA sequence and deduced amino acid sequence of rice *bglu2* 89

3.12 Alignment of predicted protein sequences of rice β-glucosidases 91

3.13 Southern-blot analysis of rice *bglu1* and *bglu2* gene copy number and probe specificity 96

3.14 Northern-blot analysis of *bglu1* and *bglu2* transcript levels in various tissues from 7-day-old rice seedlings and 6-week-old mature plants 98

3.15 Northern-blot analysis of *bglu1* and *bglu2* transcript levels in rice seedlings at days 0 through 14 of germination 99

3.16 RNA gel-blot analysis of *bglu1* and *bglu2* transcript levels in rice seedlings after growing under various environmental conditions 101

3.17 1% agarose gel electrophoresis of RT-PCR products 104

3.18 1% agarose gel electrophoresis of RT-PCR products to determine the mRNA levels in various parts of rice seedlings 105

3.19 Semiquantitative RT-PCR to determine the *bglu1* and *bglu2* mRNA levels in various tissues of 5-day-old rice seedlings 106

3.20 SDS-PAGE of recombinant BGlu1 protein expressed in the pET23d (+) and pET32a(+) systems 109

List of Figures (Continued)

Figure **Page**

3.21 SDS-PAGE of recombinant BGlu2 protein expressed in the pET32a(+) system 110

3.22 Comparison of β-glucosidase activity of pET32a(+)-*bglu1* expressed in Origami(DE3) *E .coli* after incubation in the presence of various IPTG concentrations for 0 to 18 h 112

3.23 Comparison of β-glucosidase activity of pET32a(+)-*bglu1* expressed in Origami(DE3) *E. coli* after incubation in the presence of 0.4 mM IPTG, at different temperatures for 0 to 18 h 113

3.24 SDS-PAGE profiles of recombinant BGlu2 protein expressed in pET23d (+) and pET32a(+) systems 115

3.25 Immunoblots of recombinant proteins expressed in *E. coli* 116

3.26 Hydrolysis and transglycosylation of oligosaccharides by rice BGlu1 detected by TLC 119

3.27 Scheme of the hydrolysis and transglycosylation of laminari-oligosaccharides by rice BGlu1 120

3.28 Product specificity of rice BGlu1 as assayed with pNP-β- D-cellobioside and detected by TLC 123

3.29 HPLC elution profiles of the products of hydrolysis of pyridoxine glucoside by rice BGlu1 124

3.30 Hydrolysis of pyridoxine-5'-O-β-D-glucoside by rice BGlu1 detected by TLC 125

List of Figures (Continued)

Figure **Page**

3.31 Acidic native PAGE (5%) gel zymograms of thioredoxin fusion BGlu1 expressed in Origami(DE3) *E. coli* 126

3.32 Hydrolysis of 5-bromo-4chloro-3-indolyl-β-D-glucoside (X-glucoside) by rice BGlu1 126

3.33 Chemical structures of some artificial glycoside substrates used in determining substrate specificity of rice β-glucosidase 127

3.34 Chemical structures of some natural glycoside substrates used in determining substrate specificity of rice β-glucosidase 128

3.35 Native molecular weight of rice BGlu1 as estimated by gel filtration 130

3.36 pH profile of activity for BGlu1 over a pH range 3.0 to 11 assayed with 1 mM pNPG at 30°C for 10 min 131

3.37 The pH stability of BGlu1 after incubation for 10 min to 24 h at 25°C 131

3.38 Temperature profile of BGlu1 assayed with 1 mM pNPG for 10 min at temperatures ranging from 5 to 70°C 132

3.39 Thermostability of BGlu1 incubating at 20-60°C 132

3.40 Inhibition of BGlu1 by Hg^{2+} 135

3.41 Activity profile of BGlu1 with 1 mM pNPG (pH 5.0 at 30°C for 10 min, respectively) in the presence of various concentrations of SDS 136

3.42 Activity profile of BGlu1 in the presence of various concentrations of 2-mercaptoethanol assayed with 1 mM pNPG (pH 5.0) at 30°C for 10 min 136

List of Figures (Continued)

Figure **Page**

3.43 Activity profile of BGlu1 in the presence of various concentrations of ethyl alcohol assayed with 1 mM pNPG (pH 5.0) at 30°C for 10 min 137

3.44 Transglucosylation of pNPG using rice β-glucosidase (0.01 nmol) incubated with various concentration of pNPG in 50 mM NaOAc pH 5.0 buffer in 100 μl reactions at 30°C for 1 to 24 h 139

3.45 Transglucosylation of various concentrations of ethyl alcohol using rice β-glucosidase (0.01 nmol) incubated with various concentration of pNPG in 50 mM NaOAc pH 5.0 buffer at 30°C for 3 h in 100 μl reactions 141

3.46 Transglucosylation of 20 mM pyridoxine using rice β-glucosidase (0.01 nmol) incubated with various concentration of pNPG in 50 mM NaOAc pH 5.0 buffer in 100 μl reaction at 30°C for 1 to 24 h 144

3.47 HPLC elution profiles of the transglycosylation products of pyridoxine using rice BGlu1 145

3.48 Structure of pyridoxine-5'-O-β-D-glucoside 146

3.49 Scheme of the transfer of glucose from pNPG to pyridoxine to form pyridoxine-5'-O-β-D-glucoside by rice BGlu1 146

3.50 Effect of glucono-1,5-lactone (0-100 mM) on 0 to 4-day-old rice germination, growth and development 148

3.51 Percent of rice seeds and seedlings (at various stages) grown in the presence of 0-100 mM glucono-1,5-lactone from day 0 today 4 149

3.52 Predicted gene structure patterns for putative rice β-glucosidase genes 153

List of Figures (Continued)

Figure **Page**

3.53 Unrooted phylogenetic tree of rice β-glucosidase gene family together with other plant β-glucosidases 154

List of Abbrevations

A	Absorbance
Amp	Ampicillin
ATP	Adenosine triphosphate
bp	Base pairs
BSA	Bovine Serum Albumin
°C	Degree celsius
cDNA	Complementary deoxynucleic acid
CTP	Cytosine triphosphate
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
dNTPs	dATP, dCTP, dGTP and dTTP
DP	Degree of polymerization
EDTA	Ethylene diamine tetraacetic ccid
EST	Expressed Sequence Tag
g	Gravitational acceleration
GTP	Guanidine triphosphate
(m, n) g	(milli, nano) Gram
h	Hour
HPLC	High performance liquid chromaography
IPTG	Isopropyl-β-D-thiogalactopyranoside
kDa	Kilo Dalton

List of Abbrevations (Continued)

min	Minute
(m, μ, n) M	(milli, micro, nano) Molar
(m, μ) L	(milli, micro) Liter
(μ, n, pmol) mol	(micro, nano, pico) Mole
mRNA	Messenger ribonucleic acid
Mr	Molecular weight
4MUGlc	4-Methylumbellriferyl-β-D-glucoside
N-link	Number of asparagine-linked glycosylation sites
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pI	Isoelectric point
pNP	p-Nitrophenol
pNPG	p-Nitrophenol-β-D-glucoside
RNA	Ribonucleic acid
RNase	Ribonuclease
SDS	Sodium dodecyl sulfate
s	Second
SSC	Saline sodium citrate
TEMED	Tetramethylenediamine
Tris	Tris-(hydroxymethyl)-aminoethane
TTP	Thymidine triphosphate

List of Abbrevations (Continued)

UV	Ultraviolet
U	Unit, μmol/min
v/v	Volume/volume

Chapter I

Introduction

1.1 Overview of β-glucosidases

β-glucosidases (β-glucoside glucohydrolases; EC 3.2.1.21) catalyze the hydrolysis of alkyl- and aryl- β-glucosides, as well as diglucosides and oligosaccharides, to release glucose and aglycone (Reese, 1977). These enzymes are found widely in microorganisms, animals and plants, indicating their general importance to life. Therefore, β-glucosidase structures, catalytic properties, and biological functions are of interest. Recently, β-glucosidases have been studied by many researchers because of their important roles and medical, biotechnological, agricultural and industrial applications.

In microorganisms, β-glucosidases have an important role in biomass conversion (the hydrolysis of cellulose and cellobiose). This process has important applications in industry, such as the production of alcohol based fuels (Bothast and Saha, 1997) and sugar (Woodward and Wiseman, 1982). In addition, these enzymes have been used in degradation of polysaccharide wastes, development of novel carbohydrate foods, and release of aromatic compounds from flavorless glucoside precursors in wine and fruit juice processing (Riou *et al.*, 1998).

Mammals contain two β-glucosidases, lysosomal glucocerebrosidase and cytosolic β-glucosidase. Glucocerebrosidase, catalyzes the degradation of the

glycosphingolipid glucocerebroside to glucose and ceramide (Beutler, 1992). It has an acidic pH optimum. Deficiency of this enzyme in the lysosome causes Gaucher disease, which results from mutations in the gene encoding the enzyme that result in altered β-glucosidase forms which lack catalytic activity or are not properly targeted to the lysozyme. When catalytic activity of the enzyme is lacking, the accumulation of the glycosphingolipid glucosylceramide in the lysosome of reticuloendothelial cells occurs (Beutler, 1992). Studies have been conducted to investigate the structure-function relationships and mechanism of catalysis in order to understand the molecular pathology of Gaucher disease, and to use the enzyme or gene therapy for treatment of this disease. In contrast, cytosolic β-glucosidase, which belongs to the glycosyl hydrolase family 1, has a very broad substrate specificity and is most active at neutral pH. The physiological function of this enzyme is obscure, but it does hydrolyze toxic plant glycosides found in the diet.

In plants, β-glucosidases have been implicated in several important functions, including defense, phytohormone regulation, lignification, and release of glucose from oligosaccharides. In defense against pathogens and herbivores, β-glucosidases act by releasing toxic defense compounds, such as coumarins, saponins, thiocyanates, hydrogen cyanide, hydroxamic acids and terpenes (Bell, 1981; Conn, 1993; Niemeyer, 1988; Nisius, 1988; Poulton, 1990; Sahi *et al*., 1990), and by releasing induced factors, such as volatiles that attract parasitoids of the herbivores to attack them (Mattiacci *et al*., 1995). β-glucosidases regulate the biological activity of plant phytohormones, such as cytokinins, gibberellins, auxins and abscisic acids by releasing active forms from inactive hormone glucoside conjugates (Gaskin and MacMillian, 1975; Ganguly *et al*., 1974; Brzobohaty *et al*., 1994; Millborrow, 1970).

Leah *et al*. (1995) and Hrmova *et al*. (1998) found that barley β-glucosidase (BGQ 60) hydrolyzed cellobiose, laminaribiose and other short-chain polysaccharides produced by cellulases and β-glucanases. β-Glucosidases are also critical for activation of lignin precursors, release of plant volatiles from their glycoside storage forms, and metabolism of many other important natural products (Esen, 1993; Svasti *et al.*, 1999; Geerlings *et al.*, 2000).

Substrate specificity of β-glucosidases has been the subject of extensive research over the past decade. There are various natural substrates of β-glucosidases, such as steroid β-glucosides, the β-glucosyl ceramides of mammals, the cyanogenic β-glucosides of plant secondary metabolism and oligosaccharide products released from digestion of the cellulose of plant cell-walls (Clarke *et al*., 1993). In fact, β-glucosidases from different sources have similar specificity for glycone (glucose) and some non-physiological aglycone moieties, such as p-nitrophenol (Esen, 1993). However, in plants, the aglycone specificity of β-glucosidases has been well established by numerous studies (e.g. Babcock and Esen, 1994; Hösel and Conn, 1982; Hughes and Dunn, 1982). The aglycone moieties, which are important for the specificity of plant β-glucosidases, include plant hormones, hydroxamic acids, flavonols, cyanoglucosides and mandelonitrils (Smitt and Van Staden, 1978; Campos *et al*., 1993; Cuevas *et al*., 1992; Niemayer, 1988; Conn, 1993). Since, as mentioned above, these aglycone moieties are implicated in a number of functions in plant metabolism it should be emphasized that the aglycone is important in determining the specificity of β-glucosidases in plants (Hösel and Conn, 1982).

1.2 Catalytic mechanism of β-glucosidases

β-glucosidases fall into related groups called glycosyl hydrolase (GH) families, with most plant β-glucosidases characterized so far falling in GH Family 1 (Henrissat, 1991; Henrissat and Bairoch,1993, 1996). Several structures have been solved for GH family 1 enzymes, which are $(\beta/\alpha)_8$ barrels with the catalytic acid-base and catalytic nucleophile, used for hydrolysis with retention of anomeric configuration, on β-strands 4 and 7, respectively (Barrett *et al.*, 1995; Jenkins *et al.*, 1995; Henrissat *et al.*, 1995). The substrate binding site lies in a funnel-shaped pocket on the C-terminal end of the β-barrel, and residues found in this region dictate the substrate-specificity of the enzyme (Barrett *et al.*, 1995; Burmeister *et al.*, 1997; Czjzek *et al.*, 2000, 2001). This specificity may be quite broad or narrow both in terms of aglycone and glycone moieties, but β-D-glucose and β-D-fucose are generally the most active glycones (Esen, 1993).

This structural organization is also found in F-xylanases, β-galactosidases (family 2), bacterial chitinases (family 18), and β-glucanases (family 17) (Herrissat *et al.*, 1995; Perrakis *et al.*, 1994, Jenkin *et al.*, 1995). However, the catalytic region in both bacterial chitinase and barley glucanases (1-3, 1-4-β-glucanase isoenzyme EII and 1-3-β-glucanase isoenzyme GII) is different. Their active sites are in the form of grooves or channels located across the C-terminus of the barrel (Varghese *et al*., 1994, Perrakis *et al*., 1994). Retaining glycosidases catalyze a two-step reaction by employing a mechanism involving a covalent glycosyl-enzyme intermediate formed and hydrolyzed with acid/base catalytic assistance via oxocarbenium ion-like transition states (Withers, 2001). In the first part of the catalytic cycle, one Glu

residue on β-strand 7 acts as a nucleophile toward the anomeric carbon (C-1) of the substrate and forms a covalent glycosyl-enzyme intermediate, and the second Glu on β-strand 4 donates its proton to the leaving aglycon. To regenerate free enzyme, the second Glu acts as a base catalyst to activate a water molecule, which performs a nucleophilic attack on the covalent glycosyl-enzyme intermediate, releasing the glucose and regenerating the nucleophilic Glu carboxylate (Withers, 2001).

Three-dimensional structures of several of these glycosyl-enzyme complexes, along with those of Michaelis complexes, have been determined through X-ray crystallographic analysis, revealing the identities of important amino acid residues involved in catalysis (Davies and Henrissat, 1995, McCarter and Withers, 1994). They reveal the involvement of the carbonyl oxygen of the catalytic nucleophile in strong hydrogen bonding to the sugar 2-hydroxyl for the β-retainers or in interactions with the ring oxygen for α-retainers. The glucose ring in the –1 (cleavage) site in the intermediates formed on several cellulases and a β-glucosidase adopts a normal $^{4}C_{1}$ chair conformation (Figure 1.1). By contrast the xylose ring at this site in a xylanase is substantially distorted into a $^{2,5}B$ boat conformation (McCarter and Withers, 1994).

Although the mechanism of catalysis and the roles of the two catalytic glutamates has been extensively studied, the determination of the precise substrate specificity, including the site and mechanism of aglycone binding still remain. Until recently, there was little or no information as to how β -glucosidases recognize their substrates and interact with them, specifically the aglycone moiety (Verdoucq *et al.*, 2003).

The recently resolved crystal structures of maize β-glucosidase 1 (Czjzek *et al.*, 2000, 2001) identified the amino acids within the active site that are involved in

glycone recognition and binding (Q38, H142, E191, E406, E464 and W465). The major mechanism of aglycone recognition and binding appears to be aromatic stacking and π-interactions between aromatic aglycones and the above-mentioned amino acids. Several other amino acids (e.g., P377, D261, M263, A467 and Y473) were also identified to be potentially involved in substrate specificity (Czjzek *et al.*, 2000).

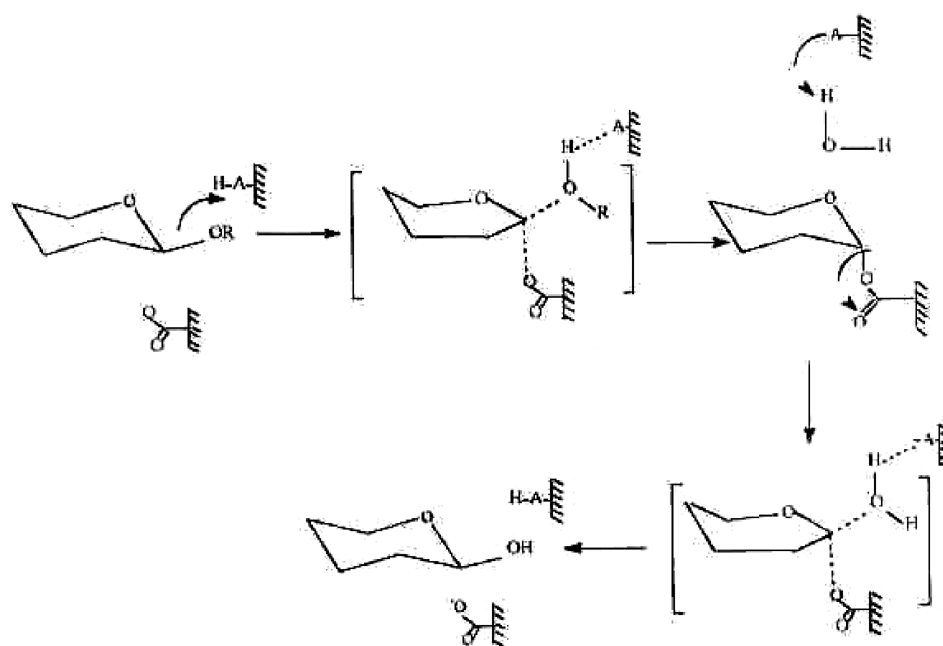


Figure 1.1 Generally accepted endocyclic pathway of the double displacement mechanism proposed for retaining β-glucosidases (Esen, 1993).

1.3 Substrate specificity of plant glycosyl hydrolases

The fundamental factors determining substrate specificity of enzymes are conformational and chemical complementarity between the substrate and its binding site on the enzyme (Hrmova and Fincher, 2001a). Hrmova and Fincher (2001a) compared the protein structures of plant glycosyl hydrolases and explained how their structures can be related to the substrate specificities. The difference of substrate specificity between endohydrolase, exohydrolase, and β-glucosidase is derived from the shape of their binding sites. Endohydrolases usually have substrate binding grooves or depressions that extend across their surface, and the catalytic amino acid residues are located in the substrate-binding cleft. As a result, the endohydrolase can essentially bind anywhere along the polymeric substrate and hydrolyze internal linkages. In contrast, an exohydrolase has a substrate binding site in a shape of a dead-end tunnel, slot, or funnel shape. Substrate specificity depends on how the substrates can fit into the substrate-binding pockets to contact the catalytic amino acid residues at the bottom of funnel (Hrmova and Fincher, 2001a, 2001b). The tight specificity of β-glucosidase may be due to its much longer binding site than exohydrolase. β-glucosidase has a substrate binding site in a shape of shallow coin slot in which only two glucosyl residues of the substrate can fit. So, most β-glucan substrates can penetrate to the bottom of the slot, while the majority of the polysaccharide substrate remains outside the enzyme. This tolerance of a wide range of substrate shapes is reflected in the broad substrate specificity of this enzyme (Hrmova and Fincher, 2001a).

1.4 Plant β-glucosidase purification and characterization

β-glucosidases have been isolated and characterized at the biochemical and DNA levels in many species of plants, both from dicots and monocots. Most β-glucosidases isolated from dicots so far, including *Trifolium repens* (Kakes, 1985), *Brassica napus* (Hoglund *et al*., 1992) and black cherry (Poulton and Li, 1994) are localized in the cell wall or vacuole, while the enzyme isolated from *Polygonum tinctorium* are localized in the plastid (Minami *et al*., 1997). In addition, most dicot β-glucosidases have been shown to be glycosylated (Esen, 1993). Most β-glucosidases isolated from monocots, including sorghum (Thayer and Conn, 1981), oat (Nisius, 1988) and maize (Esen, 1992) are localized in the plastid. The endosperm specific β-glucosidase from barley seems to be an exception which has a 24-amino acid-long signal peptide for ER targeting (Leah *et al*., 1995). Complete cDNA and deduced amino acid sequences of plant β-glucosidases have been determined from at least 14 different species (e.g. *Arabidopsis thaliana*, white clover, cabbage, *Brassica napus*, sweet cherry, black cherry, sweet almond, cassava, costus, maize, sorghum, barley, oat, Thai rosewood and rice). Comparison of amino acid sequences of these β-glucosidases has shown that their similarity ranged from 35-90% depending, in some cases, on evolutionary distances amoung taxa. For instance, the deduced amino acid sequence of maize β-glucosidase showed 70% identity with dhurrinase from sorghum (Cicek and Esen, 1998), 55% identity with oat, and 47% identity with *Prunus serotina,* respectively. The similarity between the sequences of plant β-glucosidases cannot be used to infer the specific functions of enzymes, because only small sequence differences affect substrate specificity of the enzymes (Esen, 1993).

1.5 Purification and substrate specificity of rice β-glucosidases

In rice, one of the first report of β-glucosidase activity was by Palmiano and Juliano (1973). They tested the crude extracts of germinated seedlings of rice (cultivar IR8) using the synthetic substrate, pNP β-D-glucoside, to quantitate β-glucosidase activity. The results showed that enzyme activity increased during germination of rice seeds and was highest in the fifth day after germination.

Rice β-glucosidase, hydrolyzed conjugated gibberellin from dwarf-rice *(O. sativa* cv. *Tan-ginbozu*), was partially purified by Schliemann (1984). The purification involved extraction, 90% ammonium sulfate precipitation, and CM-Sephadex C-50 chromatography. Schliemann found four different activity peaks with pNP β-D-glucoside. It was also found that β-glucosidase activity in dwarf rice increased with germination time. He stated that the highest activity hydrolyzing conjugated gibberellin β-D-glucoside was in ungerminated seed; this activity decreased as the seeds began to germinate. The enzyme hydrolyzed GA_8-O-2-glucoside eight times faster than it did GA_3-O-2-glucoside. In contrast, enzyme activity hydrolyzing pNPG increased during germination up to 10 days after imbibition, and then decreased, indicated for the presence of the other isozymes found during rice germination. His investigation suggested that the GA-glucoside β-D-glucosidase may be involved, along with α-amylases, in the breakdown of carbohydrates during the early stage of germination.

Iwami and Yasumoto (1986) found that rice bran β-glucosidases can hydrolyze pyridoxine-β-D-glucoside. First, the rice bran was defatted using cold ethanol and extracted with phosphate buffer pH 6.8. The purification steps involved

40-70% ammonium sulfate cut, DEAE Sephadex column (anion exchange chromatoghraphy), and CM-Sephadex column (cation exchange chromatography). The enzymes bound to the resins were eluted with an ionic strength gradient of acetate buffer pH 5.3. Iwami and Yasumoto found three peaks of rice β-glucosidase activity on the basis of their isoelectric point (pI) value, which were 5.6, 6.7, and 8.2, respectively.

β-glucosidase, Fr-1 and Fr-2, were partially purified from 4- to 5-day old rice seedlings by Muslim (1995). The purification comprised of extraction with 100 mM citric acid-200 mM sodium phosphate pH 5, cryoprecipitation and CM-Sephadex column chromatography using pNP-glucoside hydrolysis for detection. The pI value of Fr-1 was estimated to be 9.42, while Fr-2 was 10.08. The both enzyme fractions exhibit subtle differences with respect their broad substrate specificity with various glycosides, though their exact functions were not identified.

Akiyama *et al*. (1998) purified a cell wall-bound β-glucosidase, which had a pI of 10, from germinating rice and determined its N-terminal amino acid sequence (44 residues). The activity of β-glucosidase increased more than eight-fold within five days of germination. Akiyama *et al*. also tested the substrate specificity of the purified enzyme and showed that rice β-glucosidase had a relatively broad substrate specificity towards β-linked small oligosaccharides. So, it seems that one possible function of the enzyme might be further hydrolysis of oligosaccharides released from cell wall β-glucans during germination.

Based on the investigation of Esashi *et al*. (1991) one rice β-glucosidase appears *in vitro* to be a cyanogenic β-D-glucoside β-glucosidase. They showed that

the β-D-glucosidase activity of a crude enzyme fraction from rice seeds hydrolyzed cyanogenic β-D-glucosides such as amygdalin, prunasin, and linamarin; and the activity was increased dramatically as germination progressed. They stated that the cyanogenic glucoside may not be the only substrate. This observation needs to be confirmed, as Conn (1993) remarked that native substrate and homogeneous enzymes are required in order to obtain an unequivocal result.

Based on these works, it is obvious that rice β-glucosidase consists of more than one form. Although the occurrence of the enzyme in rice tissues is correlated with growth and development, fundamental information about its structure, physiological function, natural substrates and the mode of regulation of its expression is lacking.

1.6 Environmental effects on plant β-glucosidase expression

β-glucosidase has recently been found to be differentially regulated in response to various environmental conditions (Malboobi and Lefebvre, 1997, Kawasaki *et al*., 2001). Environmental conditions which have effects on plant β-glucosidase gene expression have been investigated by some researchers. Hösel *et al*. (1987) found that different isozymes of dhurrinase in sorghum seedling occurred in different growing conditions. Dhurrinase I occurred in shoots of seedlings grown in darkness, in contrast, dhurrinase II occurred in green shoots of young seedlings grown in light. Malboobi and Lefebvre (1997) identified through differential screening of a cDNA library groups of genes that were induced at high levels under phosphate limiting conditions in *Arabidopsis thaliana*. It was also found in the same study that

the expression of one β-glucosidase in *A. thaliana* was induced in phosphate starvation and was also repressed by heat shock and anaerobic conditions, but it was not significantly induced by high salinity or nitrogen or sulfur deprivation. Kawasaki *et al*. (2001) found that β-glucosidase genes transcript levels increased when rice seedlings were grown under salinity stress in a salt-tolerant rice cultivar (Pokkali), but not in a salt-sensitive cultivar (IR29).

1.7 Glycosides

A glycoside is a compound formed between a furanose or pyranose sugar and one or more nonsugar (aglycone) compounds through a glycosidic linkage (Noggle and Fritz, 1976). The glycoside structure is an ether-like compound containing a group or a molecule with a nucleophilic atom which is usually an oxygen atom. The linkage between the glycone and the aglycone through an oxygen atom yields O-glycosides, i. e., C-O-C. Less commonly, the nuclephilic atom is carbon, giving rise to the C-glycosides in which the sugar-aglycone linkage is C-C. When the nucleophilic atom is nitrogen or sulphur, N-glycosides and S-glycosides are obtained, giving rise to C-N-C or C-S-C, respectively.

The sugar component of glycosides is usually D-glucose, although D- and L-galactose, D-mannose, D-fucose, and L-rhamnose have also been found. Some pentoses, for example, D- and L-arabinose, and D-ribose, have also been detected. In some glycosides the glycone part is present as disaccharides or, in some case, as trisaccharides.

1.8 Transglycosylation activity of β-glucosidase

Like other glycosidase enzymes some β-glucosidases may catalyze reverse hydrolysis and transglycosylation, leading to synthesis of oligosaccharides, and alkyl glucosides (Makropoulou *et al*., 1998; Fischer *et al*., 1996; Svasti *et al*., 2003). Transglycosylation, which is kinetically controlled, is much faster than reversed hydrolysis which is thermodynamically controlled, and reactions generally take a few hours rather than days.

Transglycosylation involves the use of a reactive glycosyl donor to generate a steady state concentration of the glycosyl-enzyme intermediate, which can be intercepted with a second sugar acceptor molecule rather than water (Withers, 2001). The transglycosylation mechanism was described by van Rantwijk *et al*. (1999). Transglycosylation according to reaction (1) is based on monopolization of the catalyst by a reactive glycosyl donor.

$$\text{Glycosyl–OR}^1 + \text{R}^2\text{ OH} \rightarrow \text{Glycosyl–OR}^2 + \text{R}^1\text{ OH} \qquad (1)$$

Because the reaction is controlled kinetically, it becomes possible to overshoot the equilibrium conversion of reactant into product (van Rantwijk *et al*., 1999). As the reactant is consumed the concentration of the product will peak when its rates of synthesis and dealkylation become equal; at this point kinetic control is lost and the reaction should be stopped before thermodynamic control takes over and the product undergoes enzymatic hydrolysis (Figure 1.2).

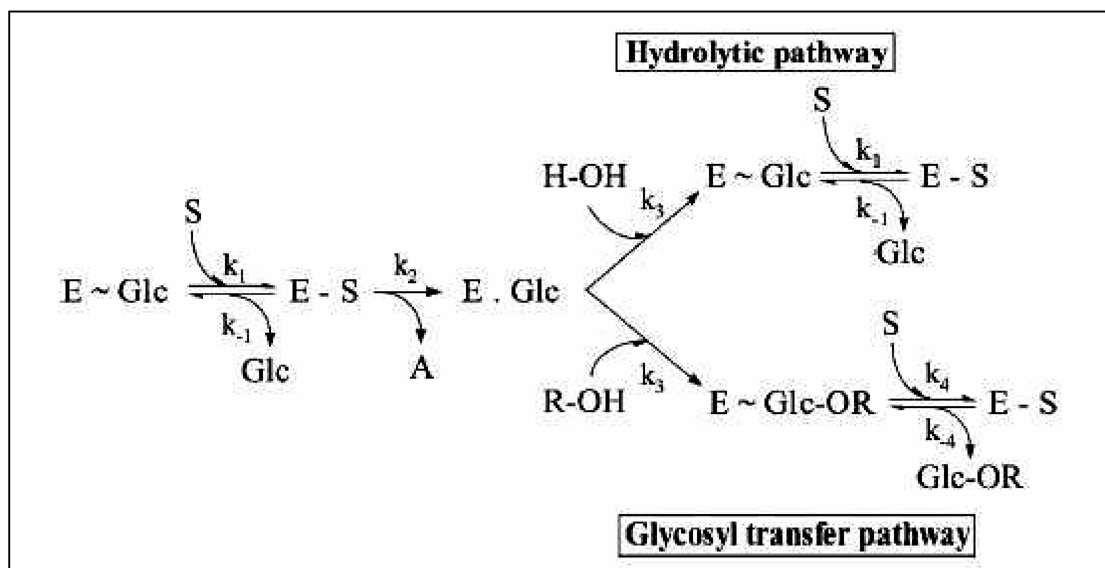


Figure 1.2 Kinetics of hydrolysis or glycosyltransfer reactions catalysed by a plant family 1 β-glucosidase (Hrmova *et al*., 2002). After the enzyme containing a noncovalently bound glucose (Glc) product in the active site (E ~ Glc) binds the first molecule of substrate (S), the Michaelis complex (E - S) is formed (K_1) and the Glc product of the previous reaction is released from the active site. In the second step, the glycosidic bond is cleaved (K_2), and the aglycone part of the substrate becomes attached covalently to the enzyme to produce a metastable covalent glycosyl-enzyme intermediate (E .Glc). At the same time, the aglycone part of the substrate (A) is released. In the third step, the covalent glycosyl enzyme intermediate is subjected (K_3) to cleavage by a water molecule (H-OH), and a noncovalent E ~ Glc product complex is formed, which is ready to interact (K_1) with the second substrate molecule (S) to generate the next Michaelis complex (E –S), and again, the Glc molecule (Glc) is released from the active site. Alternatively, in the third step, the covalent glycosyl-enzyme intermediate (E . Glc) can be cleaved by an activated substrate molecule (R-OH), leading to glycosyl transfer product (E ~ Glc-OR), which remains noncovalently

bound to the enzyme and is released (K_4) when a second substrate molecule approaches the active site and forms the next Michaelis complex (E –S) (Hrmova *et al*., 2002).

1.9 The rice genome

Rice (*Oryza sativa* L.) was the first grass species to have its genome sequenced and, thereby, has become a model system for genomics research. Factors contributing to the choice of rice include the comparatively small size of the rice genome, the synteny of its genome with those of other cereals, the availability of densely populated molecular maps containing more than 2,300 DNA markers, well-characterized YAC and BAC libraries, large-scale analysis of expressed sequence tags (ESTs), the vast amount of genetic resources (mutant markers, genetic stocks, wild species, mapping populations, introgression lines) and the comparative ease of transformation (Khush and Brar, 2001).

There are four draft genome sequences available. The first rice genome sequence of *Oryza sativa* subsp *japonica* cv Nipponbare generated by International Rice Genome Sequencing Project (IRGSP) headquartered in Tsukuba, Japan, which began in 1998. Using a bacterial artificial chromosome (BAC)-by BAC approach, the IRGSP has generated draft sequences of 3,083 BAC or P1 artificial chromosome (PAC) clones that are available through the public databases, GenBank, DNA data bank of Japan (DDBJ) and EMBL. These 3,083 BAC/PAC clones represent 426 Mb of sequence, and assuming an overlap of 15% between the clones, this would represent 362 Mb of unique sequence (Buell, 2002). With an estimated genome size of 430 Mb (Arumuganathan and Earle, 1991), this represents 84% of the rice genome.

Annotation for the IRGSP BAC/PAC clones is available for finished clones in GenBank/DDBJ/EMBL.

Draft sequences of the same *japonica* rice cultivar, Nipponbare, sequenced by the IRGSP have been made available from two separate private sources, Monsanto (Pharmacia) and the Torrey Mesa Research Institute (Syngenta). The Monsanto draft sequence was generated using a BAC-by-BAC approach and represents 259 Mb of sequence (Barry, 2001). Access to the sequence information was made available under an access agreement with Pharmacia (http://www.rice-research.org). In August 2000, Monsanto transferred its rice genome draft sequence data to IRGSP. Because this information is now largely in the public databases, public access to the private database is no longer supported.

The Syngenta draft sequence was generated using a whole-genome shotgun sequencing approach and provides 93% coverage of the genome (Goff *et al*., 2002). This draft sequence is available through a licensing agreement with Syngenta (http://www.tmri.org). Although the Syngenta draft sequence has been annotated, the annotation data is not available to the public. From this data, it was estimated that the rice genome encodes between 32,000 and 50,000 proteins (Goff *et al*., 2002). Comparative analyses with cereal showed a high degree of similarlity present between rice and other cereal genes which further highlighted the role of rice in cereal comparative genomics (Buell, 2002).

A draft sequence of the *O. sativa* subsp. i*ndica* cultivar (93-11) by the Beijing Genomics Institute (BGI) was reported by Yu *et al*. (2001, 2002). This draft, which was generated by a whole-genome shotgun sequencing approach, represents 360 Mb of assembled sequence. The BGI sequence is freely available in BGI website

(http://btn.genomics.org.cn/rice) and in the same public databases as RGP sequences. This sequence provides a resource for gene discovery in rice subsp *indica* and for rice comparative genomics.

Other analyses of the rice genome, such as alignment with expressed sequence tags from other monocot species, identification of synthetic sequences are available through several public sources (http://www.tigr.org/tdb/e2k1/osa1/; http://rgp.dna. affrc.go.jp/; http://www.gramene.org). Although these four draft sequences provide a rich resource for data mining, they have limitations in the incomplete nature of the sequence and the fact that they still contain many sequencing errors.

In comparing between rice and *Arabidopsis thal*iana, rice has a homolog for approximately 81% of the proteins encoded by the Arabidopsis genome (Yu *et al*., 2002), indicating many basic plant proteins that are found both in monocots and dicots. However, in the reciprocal comparison, a homolog in Arabidopsis could only be found for one-half of the rice proteins.

These sequencing efforts have also provided sequences for the intergenic DNA that plays an important, but poorly understood, role in gene expression, DNA replication, chromosome organization, recombination, specialization and evolution and is likely to contribute to significant progress in rice improvement.

1.10 Rice germination

In rice, a fruit (represented as pericarp) is combined together with a single seed to form a grain called the caryopsis. The caryopsis is surrounded by two modified leaves, the palea and the lemma. Both are very thick, lignified and fragile. The pericarp is originated from the mature ovary. The seed consists of an endosperm

and an embryo, both covered by a layer of nucellus. In the mature seed, the endosperm is differentiated into an aleurone layer at its outer most layer and contains mostly protein and lipid bodies; and a starchy endosperm whose cells are filled with mostly starch granules and protein bodies. During germination, the food storage in the endosperm tissue is degraded to soluble and transportable compounds to provide energy and building block molecules for the early stage of growth and development of the seedling (Mayer and Poljakoff-Mayber, 1989). The embryo consists of the plumule (primary leaves), radicle (root), short stem (mesocotyl), and seed leaf (scutellum). The leaf and root are covered by the coleoptile and coleorhiza, respectively.

According to Takahashi (1984), rice germination is characterized by two discrete events closely related to each other. The two events are imbibition and activation. Germination is the end of the activation stage in which the seed is physiologically ready to start to grow to further stages. Morphologically, germination is characterized by the protrusion of the radicle through the coleorhiza (De Datta, 1981).

Prior to germination, imbibition and oxygen uptake cause the activation of metabolism, i.e., activation of hydrolytic enzymes, release of hormones controlling germination, the synthesis of ribonucleic acid (mRNA) coding for enzymes and other proteins and the synthesis of protein. The metabolism is indicated by the increasing rate of respiration after the seeds are placed in water (Mayer and Poljakoff-Mayber, 1989). Typical chemical changes that follow the imbibition are the breakdown of storage material from the endosperm, the transport of the breakdown material from the endosperm to the embryo and actively growing parts, and the synthesis of new

materials. Palmiano and Juliano (1973) observed a decrease of starch and an increase of free sugar, soluble amino acids and protein during rice germination.

It is well known that during the activation stage prior to germination, the seed provides hydrolytic enzymes in two ways: (1) activation of latent enzymes that preexist in the mature seed and (2) *de novo* synthesis of enzymes (Ting, 1982). The latent enzymes are usually stored in the ungerminated seed as zymogens or enzyme precursors. Protease activity may activate the latent enzymes by cleavage of certain portion of the polypeptide. The *de novo* synthesis may be triggered by environmental factors, chemically (hormone action, e.g. gibberellin), or physically (light). β-amylase is an example of the preexisting enzymes, while α-amylase is a product of *de novo* synthesis (Ting, 1982).

1.11 Objectives

The objectives of this thesis include:

1) To isolate and characterize full-length rice β-glucosidase cDNAs and to deduce the complete primary structure of the enzyme they each encode; 2) To determine the multiplicity of the genes encoding rice β-glucosidases; 3) To study the expression pattern of β-glucosidase genes in different rice organs, and developmental stages; 4) To investigate the effects of environmental conditions on rice β-glucosidase gene expression during germination; 5) To determine the biochemical and catalytic properties of the β-glucosidase expressed from the rice cDNA in *E. coli*.

Chapter II

Materials and Methods

2.1 Materials

2.1.1 Plant materials: The rice (*Oryza sativa* L.) seeds cultivar 'Orion' were obtained from the Rice Reseaech and Extension Center, Stuttgart, AR, USA, whereas rice seeds cultivar 'Nipponbare' were obtained from Department of Low Temperature Science, Hokkaido National Agricultural Experiment Station, Sapporo, Japan.

2.1.2 Chemicals

Acrylamide, bis-*N,N′*-acrylamide, ammonium persulfate, *N,N′*, *N′′,N′′′*-tetramethylethyllenediamine (TEMED), lysozyme, sodium dedocyl sulfate (SDS), Sephacryl S-200, and Triton X-100 were purchased from Amersham Pharmacia Biotech (Uppsala, Sweden). Ammonium hydroxide, bromophenol blue, chloroform, copper sulfate, disodium ethylenediamine tetraacetate (EDTA), ethanol, Folin reagent, 37% formaldehyde, glacial acetic acid, glycerol, glycine, hydrochloric acid, HPLC water, methanol, PIPES (piperazine-1,4-bis(2-ethanesulfonic acid)), potassium chloride, potassium hydroxide, sodium acetate, sodium bicarbonate, sodium chloride, sodium hydroxide, sodium dihydrogen phosphate, and sodium hydrogen phosphate were purchased from Carlo ERBA (Rodano, Milano, Italy). Calcium chloride, chloramphenicol, Coomassie brillian blue R-250, imidazole, diethyl pyrocarbonate, ehtydium bromide, magnesium chloride, phenyl methylsulfonyl fluoride (PMSF), and

sodium citrate were purchased from Fluka (Steinheim, Swizerland). Bacto tryptone, yeast extract, and bacto agar were purchased from DIFCO (Grayson, USA). YM-10 Amicon filters were purchased from Millipore Corporation (Bedford, USA). Ampicillin, silica gel 60 F_{254} aluminium sheet (0.2 mm) and trifluoroacetic acid were purchased from Merck (Damsladt, Germany). β-alanine, 2,2'-azinobis(3-ethylbenzthiazolinesulfonic acid) (ABTS), chloroform/ isoamyl alcohol (24:1), isopropyl β-D-thiogalactoside (IPTG), 2-mercaptoethanol, RNase Zap, cellobiose, gentiobiose, p-nitrophenol β-D-glucoside (pNPG), pNP β-D-cellobioside, pNP β-D-cellotrioside and pyridoxine (Vitamin B_6), and sophorose and some natural and artificial glucosides were purchased from Sigma (St.Louis, MO, USA). Cellooligosaccharides of DP 3-6 and laminarioligosaccharides of DP 2-5 were from Seikagaku Kogyo CO. (Tokyo, Japan). Pyridoxine-5′-O-β-D-glucoside used as the standard was provided by Prof. J. F. Gregory. Torvoside A was provided by Prof. J. Svasti. Restriction endonucleases including *Bam*H I, *Eco*R I, *Eco*R V, *Nco* I, *Xho* I; RNase H, Taq DNA polymerase, agarose (molecular grade), deoxyribonucleotide (dATP, dCTP, dTTP, dGTP), and X-gal, were purchased from Promega (Madison, WI). Restriction endonucleases *Nco* I, *Xho* I, *Eco*R I, and SnaB I were purchased from New England Biolabs (Bevery, MA). Other chemicals and molecular reagents used but not listed here were purchased from a variety of suppliers.

2.1.3 PCR primers

Oligonucleotide primers for cDNA amplification are provided by Prof. Asim Esen, Virginia Tech, VA. These oligonucleotide primers cover essentially the entire β-glucosidase cDNA in both directions. Some are gene specific primers derived from

incomplete sequences of Orion rice β-glucosidase cDNAs which were determined in Prof. Asim Esen's laboratory. The others primers were ordered from Geneset (Singapore) and the BioServices Unit (BSU) of the National Science and Technology Development Agency (Thailand).

2.2. General methods

2.2.1 RNA extraction

Trizol Reagent, composed of phenol and guanidine isothiocyanate was used for the isolation of total RNA. During sample homogenization or lysis, Trizol Reagent maintains the integrity of the RNA, while disrupting cells and dissolving cell components. One hundred milligrams whole rice seedlings were homogenized with a cool mortar and pestle in liquid nitrogen to a fine powder. The frozen homogenates were transferred into new RNase-free microtube containing 1 ml of Trizol Reagent and mixed with the reagent by inverting the tube. The sample tube was incubated for 5 min at room temperature (~28°C) to permit the complete dissociation of nucleoprotein complexes. Next, 0.2 ml chloroform/isoamyl alcohol (24:1) was added, mixed vigorously by hand for 15 s and incubated at room temperature for 2-3 min. The tube was centrifuged at 12,000 x g for 15 min at 4°C. Chloroform was added to separate solution into an aqueous phase and organic phase. The upper aqueous phase, the RNA phase, was transferred into a new sterile microtube. The RNA sample was precipitated from the aqueous phase by mixing with 0.5 ml of isopropanol and incubated at room temperature, for 10 min and RNA centrifugated at 12,000 x g at 4° C, for 10 min. The supernatant was removed and the pellet was washed once with 1 ml 75% ethyl alcohol, centrifuged at 7,500 x g for 5 min at 4°C, and dried about 10-

15 min. The pellet was resuspended in 20 μl DEPC-treated distilled water or RNA Storage Solution (Ambion Inc., Austin, TX), and kept at -70°C.

2.2.2 Determination of quantity and purity of RNA

Two microliters total RNA solution was mixed with 998 μl DEPC-treated water and the A260/ A 280 measured with a Lamda Bio20 UV/VIS Spectrometer (Perkin Elmer, Shelton CT). Normally, the A_{260}/A_{280} ratio approximately 1.8-2.0 indicates the good purity of RNA. The RNA content was calculated using the A_{260} value of single strand RNA as following equation:

$$\mu g/ml \text{ of total RNA} = (A_{260} \text{ x dilution factor x } 40)$$

One A_{260} unit equals 40 μg of single strand RNA/ml (Sambrook *et al.*, 1989).

2.2.3 Purification of Poly A^+ RNA

Dynabeads Oligo $(dT)_{25}$ (Dynal, Oslo Norway) is used for rapid isolation of highly purified intact polyadenylated RNA (poly A^+ RNA) from eukaryotic total RNA. This step was sometimes used to isolate poly A^+ RNA, which was used as the template for reverse transcription reaction. Oligo$(dT)_{25}$ Dynabeads (beads) solution was resuspended well by shaking the vial gently to obtain a homogeneous dispersion of beads in solution. The beads suspension was transferred to an RNase-free microtube placed in a Dynal MPC-E magnetic tool. When the suspension was clear, the supernatant was removed. The beads were washed by resuspending in 100 μl 2 X binding buffer (20 mM Tris-HCl pH 7.5, 1 M LiCl, 2 mM EDTA), 2 X binding buffer was removed from the beads. The beads were resuspended in 100 μl 2 X binding

buffer. Seventy-five micrograms RNA was adjusted to 100 µl with DEPC-treated water or with elution solution (10 mM Tris-HCl pH 7.5). The RNA sample was heated at 65°C for 2 min to disrupt secondary structure. The RNA solution was added to the beads suspension and mixed gently and thoroughly. The vial was placed in the Dynal MPC-E for 30 s, the supernatant was removed. The beads were washed twice with 200 µl washing buffer (10 mM Tris-HCl pH 8.0, 0.15 M LiCl, 1 mM EDTA). Ten to twenty microliters elution buffer was added to mix with the beads, and the vial was incubated at 65°C for 2 min. To elute the mRNA, the tube was immediately placed in the Dynal MPC-E, and the supernatant was transferred to a new RNase-free tube and stored at -70°C until ready to use.

2.2.4 First-strand cDNA synthesis by ThermoScript RT-PCR system

ThermoScript RT-PCR System containing ThermoScript reverse transcriptase, an avian RNase H-minus reverse transcriptase, is engineered to have higher thermal stability in order to produce higher yields of cDNA and more full-length cDNA transcript. One microliter of 50 µM Oligo $(dT)_{20}$ primer and RNA (1 pg-5 µg) were combined and the volume adjusted to 10 µl with DEPC-treated water in a 0.5 ml tube. The tube was incubated at 65°C for 15 min and then placed on ice immediately to denature the secondary structure which may impede full-length cDNA synthesis. In the meantime, 5 X cDNA synthesis buffer was mixed by vortexing for 5 s just prior to use. A master reaction was mixed in a microtube on ice, which was composed of 4 µl 5X cDNA synthesis buffer, 1 µl 0.1 M DTT, 1 µl RNaseOUT (40 units/µl), 1 µl DEPC-treated water, 1 µl 10 mM dNTP mix, and 1 µl ThermoScript reverse

transcriptase (15 units). The master reaction was pipeted into the RNA-primer tube on ice which was then transferred to a GeneAmp® PCR System 9700 thermocycler (PE Applied Biosystems, Foster City, CA) preheated to 55°C and was incubated for 60 min. The reaction was terminated by heating the tube at 85°C for 5 min and then placed on ice. One microliter of RNase H (40 units) was added and the reaction was incubated at 37°C for 20 min to digest the complementary RNA which was base-paired with the first strand cDNA. RNase H digestion improves the RT-PCR signal of many targets and is required for the efficient and consistent amplification of long RT-PCR templates. The first strand cDNA was kept at -30°C ready for use in PCR reaction.

2.2.5 DNA analysis by agarose gel electrophoresis

The amplified PCR products were analyzed on 1% agarose gel electrophoresis in 1 X TAE buffer (0.04 M Tris-HCl pH 8.0, 0.04 M acetic acid, 0.001 M EDTA pH 8.0) as described by Sambrook *et al*. (1989). The agarose gel (molecular grade) was prepared using Pharmacia Gel Electrophoresis Apparatus GNA-100 (Pharmacia Biotech, San Francisco, CA). DNA sample was prepared by mixing 5:1 with 6 X loading dye (0.25% bromophenol blue, 0.25% xylene cyanol FF and 30% glycerol) and applied to the gel wells. Electrophoresis was performed at a constant 95 V for 45-50 min. After electrophoresis, the gel was stained in 0.1 μg/ml ethidium bromide solution for 5-10 min, and destained with distilled water for 30 min. The band in the gel was visualized by UV light transillumination, with Fluor-S™ MultiImager (Bio-RAD Laboratories, CA). The sizes of DNA fragments were estimated by comparing

with Hilo-DNA marker, or 1 kb ladders (New England Biolabs, Beverly, MA), or 100 bp and 1 kb EZ load DNA marker (Bio-RAD, Richmond, CA).

2.2.6 Purification of PCR products from agarose gels

The DNA bands in agarose gels were extracted with the QIAQuick gel purification kit (QIAGEN) by the vendor's recommended protocol. The PCR products from 100 µl reaction mix were separated on 1% TAE agarose gels. The DNA bands were excised, placed in a clean 1.5 ml microtube, and the gel was washed once with distilled water. Three volumes of QG buffer were added to 1 volume of gel and the tube was incubated at 50°C for 10 min with vortexing every 2-3 min. One gel volume of isopropanol was added to the sample and mixed (this step was used with the DNA fragments with sizes lower than 500 bp or higher than 4 kb to increase the yield of DNA fragments). The sample was applied to the QIA Quick column and centrifuged for 1 min at 13,000 x g, and the flow through was discarded. Seven hundred fifty microliters of PE buffer was added to the column, allowed to stand for 2–5 min and centrifuged at 13,000 x g for 1 min. The flow through was discarded and the column was centrifuged for an additional 1 min. The column was placed in a clean 1.5 ml microtube, and the DNA eluted from the column with 50 µl EB buffer (10 mM Tris-HCl pH 8.5) by centrifugation at 13,000 x g for 1 min and the DNA stored at -30°C.

2.2.7 Quantitation and expected yield of DNA

Five microliters DNA solution was mixed with 995 µl distilled water, the absorbance at 260 nm (A_{260}) and 280 nm (A_{280}) measured and the A_{260}/A_{280} ratio

calculated with a UV/VIS Spectrometer Lamda Bio20 (Perkin Elmer, Shelton CT). Normally, an A_{260}/A_{280} ratio of approximately 1.8-2.0 indicates good purity of DNA. The DNA content was calculated using the A260 value of double strand DNA with the following equation:

$$\mu g/ml \text{ of DNA} = (A_{260} \times \text{dilution factor} \times 50\ \mu g/ml)$$

$$= (A_{260} \times 200 \times 50\ \mu g/ml)$$

One A_{260} unit equals 50 μg of double strand DNA/ml (Sambrook *et al*., 1989).

2.2.8 Vector preparation for ligation reaction

pBlueScript II SK(+) (2.996 kb) (Stratagene, La Jolla, CA) isolated from the culture of DH5α cells by QIA miniprep kit as described in method 2.2.14. The plasmids were digested with the restriction enzyme *Eco*R V (Promega, Madison, WI) or *Sma* I (GIBCO, BRL, Invitrogen Corp.) to obtain blunt-ended vectors. The digestion reaction mix for the restriction enzyme *Eco*R V included 10 μg pBlueScript II SK(+) plasmid, 2 μl 10 X Buffer D, 0.2 μl 100 X BSA, 10 units of *Eco*R V and sterile distilled water to bring the volume up to 20 μl. The digestion reaction mix for the restriction enzyme *Sma* I included 10 μg pBlueScript II SK(+) plasmid, 2 μl 10 X React® 4 buffer, 0.2 μl 100 X BSA, 10 units of *Sma* I, and sterile distilled water to bring the volume up to 20 μl. The reaction mix was incubated at 37°C for 3-4 h. To remove the uncut plasmid, the digested reaction products were separated on 0.8% agarose gel and purified by QiaQuick gel extraction kit as previously described in method 2.2.6.

2.2.9 Ligation of DNA fragment into vector

Blunt-end ligation of purified PCR products into pBlueScript II SK(+) vector was done by using T4 DNA ligase (GIBCO-BRL, Invitrogen Corp., Carlsbad, CA) according to the supplier's directions. The amount of vector and purified PCR product (ng) was calculated according to the following equation:

$$\text{Amount of PCR product (ng)} = \frac{\text{(amount of vector, ng)} \times \text{(size of PCR product, kb)} \times \text{(molar ratio of PCR product: vector)}}{\text{size of vector (kb)}}$$

For good ligation efficiency, the amount of vector is about 60-100 ng, and the molar ratio between PCR product insert and vector is 3:1 to 5:1 for blunt-end ligation. Just prior to use, the 5 X DNA ligase reaction buffer was thawed at room temperature and mixed vigorously to dissolve any precipitated material. The reaction mix consisted of 4 µl 5 X ligase reaction buffer, 60-100 ng purified pBlueScript II SK(+) vector digested with *Eco*R V or *Sma* I, 60-120 ng purified PCR product, 1 unit T4 DNA ligase, and sterile distilled water to bring the volume up to 20 µl. The reaction mix was incubated at 14°C for 16-24 h. This reaction mix was used to transform directly into cloning host strain. It is noted that T4 DNA ligase is unstable on ice for long periods, so it was returned to -20°C within 5-10 min after taking out.

2.2.10 Transformation and selection

Transfer of ligation reactions or plasmid into *E. coli* was done as described by Sambrook *et al.* (1989). A 50 µl aliquot of frozen competent cells was thawed on ice for 5 min. The plasmids (10-100 ng) or 5 µl ligation reaction mix was added to

thawed competent cells, mixed gently by swirling the tube and incubated on ice for 30 min. The cells were transformed by using heat shock at 42°C for 60 s and immediately chilled on ice for 2 min. The transformed cells were grown in 450 μl of SOC medium or LB broth (Appendix D) by shaking at 200 rpm for 45 min. For blue-white colony selection of recombinant pBlueScript II SK(+) vector, 200 μl of cell culture was spread on an LB plate containing 100 μg/ml of ampicillin which was pre-spread with 25 μl of 50 mg/ml X-Gal and 50 μl of 0.1 M IPTG (Appendix D). For antibiotics resistant selection of recombinant pET clones, 200 μl of cell culture was spread on LB plate containing appropriate antibiotics for each vector and host system. After spreading, the plate was sat-upright at room temperature for 15-20 min to let the spread cells and medium absorb into the agar, then the plate was inverted and incubated at 37°C overnight (16-18 h). Among the recombinant clones identified by blue-white color screening, the white colonies generally contained inserts while the blue colonies had no inserts (self-ligated pBlueScript plasmid).

2.2.11 Plasmid isolation by boiling miniprep

Preparation of recombinant plasmid for preliminary analysis was done using the boiling miniprep method (Sambrook *et al.*, 1989). Normally, the size of recombinant plasmid is larger than nonrecombinant plasmid. Single colonies were picked with sterile toothpicks, restreaked on LB plate containing appropriate antibiotics, and grown at 37°C overnight. The fresh overnight colonies were inoculated in 5 ml LB broth containing appropriate antibiotics and grown overnight with 200 rpm shaking at 37°C. The cell cultures were pelleted in microtubes by

centrifugation at 3,000 x g for 3 min. The cell pellets were resuspended in 350 μl STET buffer (Appendix D) and mixed with 20 μl freshly prepared 10 mg/ml lysozyme. The suspensions were boiled at 100°C for 40 s to break the cells. The broken cells were precipitated by centrifugation at 12,000 x g for 10 min at room temperature, and the cell pellets were removed by sterile toothpick and discarded. To precipitate the plasmids in the supernatant, 400 μl isopropanol was added and the tubes were incubated at room temperature for 5 min, then centrifuged at 12,000 x g for 5 min at 4°C to pellet the plasmids. The pellets were washed with 500 μl 75% ethyl alcohol and dried for 15-20 min. The pellets was resuspended in 50 μl 10 mM Tris-HCl pH 8.0 containing 0.1 mg/ml RNase A, and incubated to degrade the RNA at 37°C for 30 min. The circular sizes of purified plasmids were compared to plasmid without insert by analysis on 1% agarose gel. The plasmids containing insert were stored at -30°C

2.2.12 Restriction enzyme digestion for plasmid analysis

To determine the size of DNA inserts in the pBlueScript II SK(+) plasmid, restriction endonucleases whose sites are present in the flanking region of *Eco*R V (*Eco*R I and *Xho* I) and *Sma* I (*Bam*H I and and *Xho* I) at the cloning site 5′ and 3′ ends, respectively, were used to cut the recombinant plasmids. The reaction mix was prepared including 3-5 μg plasmids, 2 μl 10 X Buffer D, 0.2 μl 100 X BSA, 5 units of each restriction enzyme and sterile distilled water to 20 μl. The reaction was incubated at 37°C overnight. One percent agarose gel electrophoresis was carried out to determine the DNA insert size as described previously in method 2.2.5.

2.2.13 QIAGEN plasmid miniprep

The QIA prep spin miniprep plasmid extraction kit (QIAGEN) for was used to purify recombinant plasmid for sequencing using the recommended protocol. A recombinant colony was inoculated in 5 ml LB broth with appropriate antibiotics and the cells were grown by shaking at 200 rpm (12-16 h) at 37°C. The cultured cells were pelleted by centrifugation at 3,000 x g for 3 min at room temperature. The cell pellets were resuspended completely in 250 μl P1 buffer containing RNase A. Two hundred fifty microliters of P2 buffer was added to the resuspended cells, and mixed by inverting the tube gently (to avoid genomic DNA shearing) 4-6 times until the solution became viscous and slightly clear. Three hundred fifty microliters of N3 buffer was added and mixed immediately, to avoid localized precipitation, by inverting the tube gently for 4-6 times. The tube was centrifuged at 13,000 x g for 10 min to compact the white pellet. The supernatant was applied to a QIA prep column by pipetting, centrifuged at 13,000 x g for 1 min, and the flow through was discarded. The column was washed by applying 0.75 ml PE buffer and centrifuged at 13,000 x g for 1 min. The flow through was discarded, and the column was centrifuged for an additional 1 min. The column was placed in a new 1.5 ml microtube and 25-50 μl distilled water or 10 mM Tris-HCl (pH 8.5) was added to the center of column. The column was allow to stand for 1 min, and centrifuged at 13,000 x g for 1 min to elute the plasmid DNA. The DNA was stored at -30°C.

2.2.14 DNA Sequencing

Purified DNA samples were sequenced in cycle-sequencing using the ABI PRISM dye labeled terminator kit (Big Dye) (PE Applied Biosystems, Foster City,

CA) with the recommended protocol. The reaction mix was composed of 4 μl Terminator Ready Reaction Mix, 2 μl purified plasmid DNA (100 μg/μl), 1 μl 3.2 pmol/μl primer and 3 μl distilled water. Amplification was done using the Gene Amp PCR system 9700 (PE Applied Biosystems). The program was conducted by rapidly increasing the temperature to 96°C, followed by 25 cycles of 96°C for 20 s, 50°C for 10 s and 72°C for 4 min. The amplified products were purified by ethanol precipitation as described in the kit manual. The purified DNA pellet was dissolved in 20 μl TSR solution, heated at 100°C for 2 min, quickly cooled on ice for 3 min and loaded onto an ABI 310 DNA automated sequencer.

2.2.15 Analysis of DNA sequences

The DNA sequence data obtained from DNA sequencing software of ABI 310 DNA automated sequencer was interpreted and converted to single letter code in text file format by the Chromas 1.56 program (Technelysium Pty. Ltd). The DNA sequences were also corrected by manual inspection of the chromatogram. The sequences were analyzed using the computer analysis programs in BCM Search Launcher (http://dot.imgen.bcm.tmc.edu:9331/). The sequence was translated to deduced amino acid sequence by the 6 frame translation program and antisense sequences converted to sense strand sequence with the Reverse complement program in Sequence Utilities. The sequences were confirmed to correspond to β-glucosidases by a local alignment search of the Genebank database by the BLAST program at the National Center for Biotechnological Information (NCBI) (http://www.ncbi.nlm. nih.gov). Comparison of the DNA fragment sequences was done with the ClustalX implementation of ClustalW (Jeanmougin *et al*., 1998; Thompson *et al*., 1994) or

BLAST2 at NCBI. The cDNA sequences of *bglu1* and *bglu2* were searched against the dbEST database using BLASTn to identify matching ESTs in rice, and these ESTs were used to search the Rice Genome Research program (http://rgp.dna. affrc.go.jp) mapped EST markers to identify the chromosomal locations of *bglu1* and *bglu2*.

Protein sequence alignments were done with the ClustalX implementation of ClustalW. Protein analyses were done at the Expasy proteomics server (http://www. expasy.org/), signal sequences were predicted by SignalP (Nielson *et al*., 1997) and cellular locations were predicted by PSORT (Nakai and Horton, 1999).

2.3 Amplification of rice β-glucosidase cDNAs

Rice (*Oryza sativa* L. cv. Orion and cv. Nipponbare) seeds were sterilized by soaking in 10% Clorox for 10 min, washed twice with sterile distilled water for 2-3 min, soaked in 95% ethyl alcohol for 3 min, and washed 4 times with sterile water for 2-3 min. The seeds were germinated in the dark from day 0 to day 3 and in the 12 h light-12 h dark from day 4 to day 5 at 28°C under sterile condition on germinating paper moistened with sterile distilled water. The whole seedlings were harvested, and the RNA was isolated and reverse transcribed as described in methods 2.2.1 and 2.2.4.

2.3.1 Cloning of rice β-glucosidase cDNA for isozyme 1

The first rice-β-glucosidase designated *bglu1* was cloned by homology and identification of corresponding sequences in databases. Two rice EST clones, R2847_1A (GenBank AC D24959) and R2336 (GenBank AC D24663) were identified as belonging to glycosyl hydrolase family 1 by a tBLASTn search (Altschul

et al., 1997) of the dbEST database. These two clones were obtained from the Japanese Ministry of Agriculture, Forestry and Fisheries (MAFF) Genome Research Program, Japan, and sequenced thoroughly by automated DNA sequencing as described in method 2.2.14. The sequence of the longest cDNA insert was determined and deposited in GenBank, accession number (AC) U28047.

A sense primer β-glu163 (Table 2.1) derived from the R2847_1A sequence was used for 3′ rapid amplification of cDNA ends (3′ RACE) using the first-strand cDNA from Orion rice as the template, essentially as described by Zhang and Frohman (1997). The PCR reaction mix included 2.5 μl 10 X Elongase buffer, 2 mM 2.5 mM dNTP mixed, 0.5 μl 10 μM β-glu163 primer, 0.5 μl 10 μM oligo $(dT)_{17}$ primer, 1.25 units Elongase (Life Technologies, Invitrogen, Carlsbad, CA), 1 μl first-strand cDNA reaction and sterile distilled water to a final volume of 25 μl. The PCR amplification was done with an initial denaturation step of 94°C for 5 min, followed by 35 cycles of 94°C for 30 s, 42°C for 30 s, and 68°C for 2 min, and a final extension of 5 min at 68°C after the last cycle. The PCR product was analyzed by 1% agarose gel electrophoresis, as described in method 2.2.5, purified with the QIA quick gel extraction kit, and the sequence determined by direct sequencing using β-glu163 primer.

EST clone R2847_1A was also selected to design the antisense primer β-glu178 for amplification of 5' end fragment. The degenerate sense primer β-glu156 designed from the conserved β-glucosidase protein sequence, GFPAGFVFG, and the antisense primer β-glu178 were used to amplify the approximately 300 bp 5' region cDNA fragment using first-strand cDNA as the template. The PCR product was

analyzed by 1% agarose gel electrophoresis, purified by QIA quick gel extraction kit, and its sequence determined by direct sequencing with the β-glu156 primer. The 300 bp 5' DNA sequence was BLASTn searched against dbEST to identify a corresponding cDNA fragment with a sequence which covered the start codon, methionine and 5' UTR of OSU28047. The EST clone c29108 was identified as the sequence containing the 5' end of the coding sequence. From this sequence, a 5' sense primer, β-glu217, was designed and used with the antisense primer β-glu178 to amplify the 5' end region. The PCR reaction mix included 1 µl first-strand cDNA, 2.5 µl of 10 X HotStar Taq buffer, 5 µl of Q-solution, 2 µl of 2.5 mM dNTP mix, 1 µl of 10 µM β-glu217 primer, 1 µl of 10 µM β-glu178 primer, 0.625 unit of HotStar Taq DNA polymerase (QIAGEN, Hilden, Germany), and sterile distilled water to bring the volume to 25 µl. The PCR was done at 94°C for 15 min as the initial denaturation step, followed by 35 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 2 min, and a final extension of 5 min at 72°C after the last cycle. The HotStar Taq system helps to overcome the difficulties to amplify the 5′ GC-rich region. The product was cloned into the pBlueScript II SK(+) *Eco*R V site and sequenced as described in methods 2.2.6-2.2.14.

After the sequence at the 5' end was obtained, the 5' sense primer β-glu217 was used with an antisense primer Ro13pr derived from the 3' end sequence of U28047 to amplify the full coding cDNA. The amplification was done as for the β-glu217/β-glu178 reaction above but with 3 min extension. The product was cloned into the pBlueScript II SK(+) *Sma* I site and sequenced as described in methods 2.2.6-2.2.14.

Table 2.1 Primers used for amplification of the *bglu1* cDNA

Primer name	Primer sequence
β-glu156	5′-GGTTCCCGSCKGGSTTCGTSTTYGG-3′ (S= G+C, K= T+G, Y= C+T)
β-glu163	5′-TGGTTTACATTTAATGAGCC-3′
β-glu178	5′-CTGCAGAAGGTAGTTTATAAGA -3′
β-glu217	5′-GTGCGTGCGTGGCTGCGTGTGCTT-3′
Ro13pr	5′-CACCAAGCCAAATCTCATCAAC-3′

Table 2.2 Primers used for sequencing the *bglu1* cDNA

Primer name	Primer sequence
β-glu156	5′-GGTTCCCGSCKGGSTTCGTSTTYGG-3′ (S= G+C, K= T+G, Y= C+T)
β-glu163	5′-TGGTTTACATTTAATGAGCC-3′
β-glu178	5′-CTGCAGAAGGTAGTTTATAAGA -3′
T7	5′-TAATACGACTCACTATAGGG-3′
T3	5′-AATTAACCCTCACTAAAGGG-3′

2.3.2 Cloning of rice β-glucosidase cDNA for isozyme 2

A second rice-β-glucosidase (*bglu2*) cDNA was isolated from Orion rice by RT-PCR using the sense primer β-glu156 and the antisense primer β-glu41 (Table 2.3), derived from the conserved peptide sequences GFPAGFVFG and WFTFNEP, respectively, to amplify an approximately 500 bp PCR product. The PCR reaction mix included 2.5 µl 10 X Elongase buffer, 2 mM 2.5 mM dNTP, 0.5 µl 10 µM β-glu156 primer, 0.5 µl 10 µM β-glu41 primer, 1.25 units Elongase, 1 µl first-strand cDNA reaction and sterile distilled water to give a final volume of 25 µl. The PCR

amplification was done at 94°C for 5 min as the initial denaturation step, followed by 35 cycles of 94°C for 30 s, 48°C for 30 s, and 68°C for 2 min, and a final extension of 5 min at 68°C after the last cycle. The product was cloned into the pBlueScript II SK (+) *Sma* I site and sequenced as described in methods 2.2.6-2.2.14.

An approximately 500 bp cDNA fragment sequence was used to design a sense primer, β-glu176, and this primer was used in 3′ RACE as described previously. The PCR reaction mix included 2.5 μl 10 X Elongase buffer, 2 μl 2.5 mM dNTP, 0.5 μl 10 μM β-glu177 primer, 0.5 μl 10 μM oligo $(dT)_{17}$ primer, 1.25 units Elongase, 1 μl first-strand cDNA reaction and sterile distilled water to give a volume to 25 μl. The amplification was done as above but with 40°C annealing temperature, and the approximately 1450 bp product was sequenced completely. An approximately 500 bp cDNA fragment sequence was used to identify matching ESTs in dbEST at NCBI. A sense primer β-glu214 was designed from the EST clone E60611_11A (GenBank AC AU065204) sequence to include the start codon.

Finally, the β-glu214 sense primer and Ro23pr primer derived from the 3'UTR of the cDNA, were used to amplify the full-length CDS cDNA. The PCR reaction mix included 1 μl first-strand cDNA, 2.5 μl 10 X HotStar Taq buffer, 5 μl Q-solution, 2 μl 2.5 mM dNTP mix, 1 μl 10 μM β-glu214 primer, 1 μl 10 μM Ro23pr primer, 0.625 unit HotStar Taq DNA polymerase (QIAGEN), and sterile distilled water bring to a volume 25 μl. The PCR condition was done at 94°C for 15 min as the initial denaturation step, followed by 35 cycles of 94°C for 30 s, 54°C for 30 s, and 72°C for 3 min, and a final extension of 5 min at 72°C after the last cycle. The full-length

product was cloned into the pBlueScript II SK(+) *Sma* I site and sequenced as described in methods 2.2.6-2.2.14.

Table 2.3 Primers used for amplification of the *bglu2* cDNA

Primer name	Primer sequence
β-glu41	5′-GGCTCATTAAATGTAAACCA-3′
β-glu156	5′-GGTTCCCGSCKGGSTTCGTSTTYGG-3′
β-glu176	5′-GATCCTTGAGTGGAGGAATAAACA-3′
β-glu177	5′-TGTTTATTCCTCCACTCAAGGATC-3′
β-glu214	5′-CATCTCTCGCTGGAATGG-3′
Ro23pr	5′-AACACTGCTCGACAGGGAGC-3′

Table 2.4 Primers used for sequencing the *bglu2* cDNA

Primer name	Primer sequence
RO1r	5′-TGAGTGGAACTCGATCCTG-3′
β-glu176	5′-GATCCTTGAGTGGAGGAATAAACA-3′
β-glu218	5′-AGCCTGAACCAGCCCAAC-3′
β-glu219	5′-TGACGAAACACAAATCTGGTG-3′
T7	5′-TAATACGACTCACTATAGGG-3′
T3	5′-AATTAACCCTCACTAAAGGG-3′

2.4 Southern blot analysis

Southern blot analysis was used to determine the specificity of the *bglu1* and *bglu2* probes and the copy number of each gene in the rice genome. For this, restriction enzyme digested genomic DNA was separated by agarose gel electrophoresis, transferred to nylon membrane, and the gene of interested detected by hybridization with radio-labeld DNA followed by autoradiography.

2.4.1 Genomic DNA isolation and restriction enzyme digestion

Genomic DNA was isolated from 4-day-old rice seedlings by the method described by Doyle and Doyle (1987). Two grams of whole rice seedlings were ground to a fine powder with a mortar and pestle cooled in liquid nitrogen. The frozen homogenate was resuspended in 5 ml of 2 X CTAB solution (Appendix D), incubated at 56°C for 5 min, 5 ml of 24:1 chloroform/isoamyl alcohol added, gently shaken and incubated at 56°C for 30 min. The sample was centrifuged at 12,000 x g at 4°C for 5 min. The supernatant was transferred into a new 50 ml tube and mixed with 1/10 volume of 10 X CTAB solution and 1 volume of 24:1 chloroform/isoamyl alcohol. The mixture was incubated at 56°C for 30 min, centrifugated at 12,000 x g for 5 min at 4°C and the supernatant transferred into a new tube. One volume of precipitation solution (1% CTAB, 50 mM Tris-HCl pH 8.0, 10 mM EDTA) was added, gently shaken and incubated at -20°C for 10-20 min. The DNA was pelleted by centrifugation at 12,000 x g at 4°C for 5 min and the pellet was dissolved in 750 µl TE buffer (10 mM Tris-HCl pH 8.5, 1 mM EDTA). One volume of isopropanol was added to DNA solution, incubated at -20°C for 5 min, and centrifuged at 12,000 x g

for 5 min at 4°C. The DNA pellet was washed with cold (0 to 4°C) 75% ethanol and dried at 37°C for 20 to 30 min. The DNA pellet was dissolved in 200 to 300 µl TE buffer. The DNA solution was treated with 1 µl of 10 mg/ml RNase A per 100 µl DNA solution and incubated at 37°C for 15 min. DNA was kept at -30°C or -70 °C for short or long term storage, respectively.

The DNA solution concentration and purity was estimated by measurement of the A_{260}/A_{280}, as described in method 2.2.7. Ten micrograms of genomic DNA was digested with 20 units of *Bam*H I, *Bgl* II, *Eco*R I, *Hin*d III, *Xho* I or *Xba* I for 24 h, loaded onto a 0.8% agarose gel and electrophoresed in Tris- acetate-EDTA (TAE) buffer at 1.5 V/cm for 18 h.

2.4.2 Transferring genomic DNA from agarose gel to nylon membrane

Genomic DNA was transferred onto nylon membranes by capillary transfer (Sambrook *et al*., 1989). After agaorse gel electrophoresis of digested genomic DNA, the gel was soaked in 25 ml of 0.5 M HCl, denaturing solution (0.5 M NaOH, 1.5 M NaCl) and neutralizing solution (3 M NaCl, 0.5 M Tris-HCl pH 7.4) for 7, 30 and 30 min, respectively, with gentle shaking on an orbital shaker and rinsed briefly with distilled water between each soaking. Then genomic DNA was transferred onto a nylon membrane. The nylon membrane was cut to the same size as the gel containing the DNA, and then was equilibrated in transfer buffer (20 X SSC) (Appendix D) for 10 min. A sheet of Whatman 3 MM paper was cut as a wick that will act as the contact between the transfer buffer reservoir and the gel. Next, the 3 MM paper was placed on a support placed in a baking dish. The baking dish was filled with enough transfer buffer such that a minimum of 2.5 cm of each end of the paper was

submerged. When the 3 MM paper was throughly wet, all air bubbles were removed out with a glass rod. The gel was placed on the support, and any air bubbles that were trapped under the wick were removed. Next, wet nylon membrane was placed on top of the gel and trapped air bubbles were removed. Another five sheets of Whatman 3 MM paper of the same size as the gel were placed on top of the filter membrane and air bubbles were removed. A stack of paper towels the same size as the gel was placed on top of the sheets of 3 MM paper. Finally, the entire stack was covered with a 500 g weight to compress the entire transfer set-up. The transfer was allowed to proceed overnight. Then, the nylon membrane containing DNA was UV irradiated while still damp to fix the DNA sample to the membrane.

2.4.3 DNA probe preparation

The specific DNA probes for the *bglu1* and *bglu2* genes were generated by PCR amplification. The *bglu1* and *bglu2* specific probes were generated by amplification of portions of *bglu1* and *bglu2* cDNA from the 3' UTR (284 and 267 bp, respectively) using Orion rice cDNA isolated as described in method 2.2.4 as the template. The *bglu1* (284 bp) DNA probe was amplified with the Ro13pf sense primer and Ro13pr antisense primer (Table 2.5). The *bglu2* (267 bp) DNA probe was amplified with the Ro23pf sense primer and Ro23pr antisense primer. The PCR reaction included 1 µl first-strand cDNA, 2.5 µl 10 X HotStar Taq buffer, 2 µl 2.5 mM dNTP mix, 1 µl 10 µM sense primer, 1 µl 10 µM antisense primer, 0.625 unit HotStar Taq DNA polymerase (QIAGEN), and sterile distilled water bring to a volume 25 µl. The amplification for both probes was done using 35 cycles of 94°C 30 s, 52°C 30 s and 72°C 1 min. The amplified DNA products were electrophoresed

through 1% agarose gel. The DNA band was cut and purified by the QIAGEN gel extraction kit as described in method 2.2.6.

Table 2.5 Primers used for amplification of *bglu1*and *bglu2* probes

Primer name	Primer sequence
Ro13pf	5′-TGAGAATGAAAGACACGCAGAC-3′
Ro13pr	5′-CACCAAGCCAAATCTCATCAAC-3′
Ro23pf	5′-TCAAGAGGCACTGATCCAGC-3′
Ro23pr	5′-AACACTGCTCGACAGGGAGC-3′

2.4.4 DNA probe labeling

cDNA probe was isotopicly labeled with [α-^{32}P]dCTP using Ready-To-Go DNA Labeling Beads (Amersham Pharmacia Biotech). Twenty to fifty nanograms of DNA were adjusted to a volume of 48 μl. The DNA solution was denatured by heating at 100°C for 5 min, then immediately put on ice for 5 min to fix the denatured form of DNA, then centrifuged briefly. Forty-eight milliliters of denatured DNA and 2 μl [α-^{32}P]dCTP (3000 Ci/mmol) were added into the microtube containing the DNA labeling beads, and gently mixed by pipetting up and down 4-5 times. The reaction was incubated at 37°C for 30 min. The unincorporated nucleotides in the reaction were removed by gel filtration on Sephadex G-50 (packed in a 1 ml plastic xylink column) with centrifugation at 3,400 x g for 4 min. Before hybridization, labeled DNA was denatured by heating at 100°C for 5 min, then placed on ice immediately.

2.4.5 DNA-DNA probe hybridization, post hybridization stringency wash, and autoradiography

DNA blots were prehybridized in 15-20 ml of hybridizing buffer (1 mM EDTA, 0.5 M Na_2HPO_4 pH 7.2, 1% BSA, 7% SDS) for 30-45 min at 50°C to equilibrate the membrane. The denatured α^{32}P-labeled specific *bglu1* probe (100-200 μl) was added to the buffer in a hybridization bottle containing the membrane and continued to hybridize at 50°C for 16 h. Then, the hybridization buffer was poured off and the hybridized blots were washed to remove all the probe molecules that did not form hybrids twice in 2 X SSC/0.05% SDS at room temperature for 10 min, and once in 1 X SSC/0.05% SDS at 55°C for 10 min.

After post-hybridization washing, the plastic-wrapped filter membranes were placed in direct contact with a piece of high speed medical x-ray film (Kodak). A Geiger counter probe was held about 2.5 cm above the surface of the membranes to estimate the required time for autoradiography. The plastic-wrapped membrane attached to a sheet of x-ray film was placed in a cassette under safelight conditions. The cassette was then stored at -80°C for the appropriate time (normally 16-72 h depending on signal intensity).

Photograph development was done in the dark. The exposed x-ray film was detached from the filter membrane, submerged into the film developer, and gently agitated in the developing tray about 10 s or until the bands began to appear. The film was quickly removed from the developer and placed in water to stop the reaction. Then, the film was transferred to a photographic fixer tray until the film began to clear. After fixation, the film was washed in water and then air-dried.

To reprobed with *bglu2* probe on the same membrane, the *bglu1* probe was stripped out from the membrane by boiling the membrane at 100°C in high stringency 0.1 X SSC/ 1%SDS wash buffer for 15 min. Then the blot was re-hybridized with α ^{32}P-labeled *bglu2* (267 bp) probe (100-200 μl) at 50°C for 16 h. The blot was washed to remove unbound probe and autoradiographed to detect the irradiation signal on the film as described above.

2.5 Northern blot analysis to determine β-glucosidase gene expression

Northern blot analysis was used to determine the size and amount of mRNA for the gene of interest in total RNA samples. For this, total RNA was separated by agarose gel electrophoresis, transferred to nylon membranes, and the mRNA of interest detected by hybridization with radio-labeled DNA followed by autoradiography.

2.5.1 Growth conditions

The seeds were germinated in the dark from day 0 to day 3 and in 12 h light-12 h dark from day 4 to day 14 at 28°C under sterile conditions on germinating paper moistened with sterile distilled water. The whole seedlings were harvested and some were divided to separate parts (shoot, root, and endosperm) and kept at -70°C. Some 14 day-old rice seedlings were transferred to soil and grown for an additional 5.5 weeks. Rice plants were harvested and separated to different parts (flower, leaf blade, leaf sheath, node, internode and root) and kept at -70°C for RNA isolation.

To determine the effect of environmental conditions on rice β-glucosidase gene expression, various growth conditions were applied during germination. Rice seeds were grown on germinating paper moistened with sterile distilled water at 28°C for 6 days in the dark from day 0 to day 3 and in 12 h light-12 h dark conditions starting from day 4. Then the rice seedlings were moved to various environmental conditions and grown for an additional 2 days, along with rice seedlings grown at 28° C. The abiotic conditions were: 1) drought, by supplying no water; 2) cold temperature, at 5°C; 3) salt stress, with 0.3 M NaCl solution; 4) flooding, by submerging the seedling roots in 0.5 cm high water; 5) osmotic stress, with 0.5 M manitol; 6) ethylene, (10 mg/ml Ethephon), and 7) 10^{-4} M abscissic acid.

2.5.2 RNA extraction

Total RNA was isolated from rice seedlings by the method of Bachem *et al.* (1996). Rice tissue material (1-2 g) was ground to a fine powder in a pre-cooled mortar and pestle in liquid N_2. Plant tissue powder was transferred into pre-cooled 50 ml plastic tubes. Next, 8 ml of a hot (80°C) 1:1 mixture of phenol and RNA extraction buffer (100 mM Tris-HCl, 100 mM LiCl, 10 mM EDTA, 1% SDS, pH 8.0) was added into the tubes. The sample was vortexed vigorously for at least 30 s. Four milliliters of chloroform were added to the mixture and shaken vigorously. The sample was centrifuged for 20 min at 12,000 x g to separate the phases. Chloroform extraction was repeated at least twice until no inter-phase could be detected. The resultant clear aqueous phase was transferred to a new 50 ml tube. To precipitate the RNA, 1/3 volume of ice cold 8 M LiCl was added and RNA precipitated for at least 3 h to overnight at 0°C. RNA was pelleted by centrifugation at 4°C for 20 min and then

resuspended in 500 µl DEPC-treated water, after extensive washing of the pellet with ice cold 75% ethanol. A small aliquot of this sample was analyzed for concentration, integrity and purity of the RNA as described in method 2.2.2.

2.5.3 RNA gel electrophoresis and transferring RNA from agarose gels to nylon membranes

RNA was denatured and electrophoresed on formaldehyde–agarose gel by standard procedures (Sambrook *et al.*, 1989). Twenty micrograms of total RNA samples were denatured by mixing with RNA premix (Appendix D) in a 1:1 ratio (v/v), and incubated at 65°C for 15 min and then put on ice. Before loading the RNA samples, the gel was prerun for 30 min at 5 V/cm. The denatured RNA solutions were mixed with 2 µl 5 X RNA loading dye (Appendix D) and applied to the gel wells. The RNA samples were electrophoresed on a 1.5% formaldehyde-agarose gel which was submerged in 1 X formaldehyde gel-running buffer (Appendix D) at a constant 80 V for 1 h. After running, the gel was stained in ethidium bromide and photographed on UV light to document the RNA bands appearance. Size separated RNA was transferred onto nylon membrane as described by Sambrook *et al*. (1989). After electrophoresis, the gel was rinsed in several changes of DEPC-treated water to remove formaldehyde, and then soaked in 20 X SSC (Appendix D) for 45 min. Then the RNA was transferred onto Hybond N+ nylon filter membrane (Amersham Pharmacia Biotech) in the same was as the typical set up used for capillary transfer of DNA described in method 2.4.2.

2.5.4 DNA probe preparation

The specific DNA probes for the *bglu1* and *bglu2* mRNAs were generated by PCR amplification as described in method 2.4.3. The 18S rRNA was used as control transcript for a housekeeping gene to show the amount of total RNA on the blots. The primers for amplification of rice 18S rRNA probe (342 bp), 18Sricef sense primer (5′-AAGTTTGAGGCAATAACAG-3′) and 18Sricer antisense primer (5′-CCTCTAAAT GATAAGGTTC-3′), were derived from the 3′ end region of the published sequence AC AF069218. The probe was amplified using Orion rice cDNA as template. The PCR reaction mix included 1 μl first-strand cDNA, 2.5 μl 10 X HotStar Taq buffer, 2 μl 2.5 mM dNTP mix, 1 μl 10 μM 18Sricef primer, 1 μl 10 μM 18Sricer primer, 0.625 unit HotStar Taq DNA polymerase (QIAGEN), and sterile distilled water to bring the volume to 25 μl. The amplification was done using 35 cycles of 94°C 30 s, 48°C 30 s and 72°C 1 min. The amplified DNA products were electrophoresed on a 1% agarose gel. The DNA band was cut and purified as described in method 2.2.6. and isotopicly labeled with [α-^{32}P]dCTP as described in method 2.4.4.

2.5.5 RNA-DNA probe hybridization

RNA blots were hybridized with *bglu1* and *bglu2* probes, post hybridization washed and autoradiographed as described in method 2.4.5. Finally, all blots which were previously hybridized with both *bglu1*and *bglu2* probes, were stripped of all hybridized probes on the membranes by boiling the membranes at 100°C in 0.1 X SSC/1%SDS wash buffer for 15 min. Then, the blots were reprobed with α^{32}P-labeled 18S rRNA (342 bp) probe (100-200 μl) at 55°C for 16 h. Then, the

hybridized membranes were washed from lower to higher stringency with the following steps: 2 X SSC/0.1% SDS, 1 X SSC/0.1% SDS, 0.5 X SSC/0.1% SDS, and 0.1 X SSC/0.1% SDS at 60°C for 15 min in each step. The membranes were then wrapped in plastic wrap and set up for autoradiography as described in method 2.4.5.

2.5.6 Determination of signal intensity on autoradiography film by densitometry technique

The autoradiography film was scanned to capture the picture with the Fluor-S^{tm} MultiImager and the intensity of the signal quantified using Quantify One software (Bio-RAD). The ratio between the *bglu1* and *blgu2* signals to 18S rRNA signal was plotted on graph with Excel Microsoft Office 97 version.

2.6 Semiquantitative RT-PCR

The expression of rice β-glucosidase mRNAs in different rice seedling parts was also studied by reverse transcription-polymerase chain reaction (RT-PCR). Rice (*Oryza sativa* L. cv. Orion) seeds were germinated in the dark from day 0 to day 3 and in the 12 h light-12 h dark from day 4 to day 5 at 28°C under sterile conditions on germinating paper moistened with sterile distilled water. The seedling parts were divided into mesocotyl lower and upper parts, coleoptile, root elongation zone, root tip, and endosperm. Seedling parts were soaked in the RNA Later Solution (Ambion Inc, Austin, TX) and then kept at -70°C until used for RNA isolation. Total RNA was isolated from rice seedling parts using the RNaeasy Plant Mini Kit (QIAGEN, Hilden, Germany) according to the manufacture's protocol. The RNA pellet was resuspended

in RNA Storage Solution (Ambion Inc.). First-strand cDNA synthesis was carried out using SuperScript Reverse Transcriptase II (GIBCO-BRL Invitrogen Corp.) according to the manufacturer's protocol, using 2.5 μg of total RNA. The following sense and antisense primers, respectively, were used in the PCR step: for *bglu1*, Ro13pf and Ro13pr; for *bglu2*, Ro23pf and Ro23pr and for rice β-actin, actin1f and actin1r (Table 2.6). The actin primers were derived from a published sequence (McElroy, 1990).

Two microliter aliquots of serial dilutions: 1, 1/5, 1/10, 1/20, and 1/40 X of the cDNA product from the RT reaction were used as template for PCR reactions with 0.2 mM each dNTP, 0.2 μM each primer for β-glucosidase and β-actin, 10 X PCR buffer, 2 mM $MgCl_2$, and 1.25 units of HotStar Taq DNA Polymerase in a final volume of 50 μl. The PCR was carried out in a DNA thermocycler under the following conditions: 15 min at 94°C, followed by 30 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 90 s, followed by a final extension of 10 min at 72°C. In this reaction, β-actin was co-amplified as an internal standard to quantify the PCR amplification of β-glucosidase mRNA. The number of thermocycles used was optimized to allow quantification without saturation. The fragments amplified by PCR were separated by electrophoresis on a 1% agarose gel containing 10 μg/ml ethidium bromide (Sambrook *et al.*, 1989) using 100 bp Molecular Mass Marker (New England Biolabs, Beverly, MA) as a standard to determine the concentraion of PCR products. The gel images were captured with a Fluor-S™ MultiImager and quantified using Quantify One software (Bio-RAD). The intensity ratio of β-glucosidase and β-actin cDNAs was used for semiquantitative analysis.

Table 2.6 Primers used for amplification of the 3' UTR of *bglu1*, *bglu2* and β-actin cDNA fragments

Primer name	Primer sequence
Ro13pf	5′-TGAGAATGAAAGACACGCAGAC-3′
Ro13pr	5′-CACCAAGCCAAATCTCATCAAC-3′
Ro23pf	5′-TCAAGAGGCACTGATCCAGC-3′
Ro23pr	5′-AACACTGCTCGACAGGGAGC-3′
Actin1f	5'-ACTCTGGTGATGGTGTCAGCC-3'
actin1r	5'-GTCAGCAA TGCCAGGGAACATA-3'

2.7 Recombinant protein expression in *Escherichia coli*

2.7.1 Amplification of the cDNAs to insert in expression vectors

The cDNA encoding the predicted mature protein products of *bglu1* and *bglu2* genes were cloned into pET vectors and expressed in *E. coli*. To predict the N-terminal cleavage site and start of the mature protein, the Signal P software (http://www.cbs.dtu.dk/services /SignalP/) was used. BGlu1 was predicted to have the cleavage site between amino acid residues 28 and 29, while the BGlu2 cleavage site was predicted to be between amino acid residues 25 and 26. The restriction enzymes that could not cut inside the cDNA but cut in the expression vector cloning site were selected. The primers specific for *bglu1* and *bglu2* (Table 2.7) were designed to amplify the cDNA encoding the mature proteins and introduce the restriction enzyme sites complementary to the cloning site of the expression vector to flank the cDNA, allowing inframe translation of the proteins.

For *bglu1*, cDNA encoding the mature protein product was cloned into pET23d(+) and pET32a(+) (Novagen, Madison, WI). The *bglu1* cDNA prepared for cloning into pET23d(+) was amplified by using the Orion rice cDNA as template with the Roex5f and β-glu135 primers to introduce *Bam*H I and *Eco*R I sites at the 5' and 3' ends, respectively. The *bglu1* cDNA prepared for cloning into pET32a(+) was PCR amplified using the Orion rice cDNA as template with the Roex8f and β-glu135 primers to introduce *Sna*B I and *Eco*R I sites at the 5' and 3' ends, respectively.

The first round of PCR to amplify the full coding sequence of *bglu1* cDNA was done with the β-glu217 sense primer and the Ro13pr antisense primer. The PCR reaction mix included 1 µl first-strand cDNA, 2 µl of 10 X HotStar Taq buffer (QIAGEN), 1 µl 10 X *Pfu* buffer (Promega), 5 µl of Q-solution (QIAGEN), 2 µl of 2.5 mM dNTP, 1 µl of 10 µM primer β-glu217, 1 µl of 10 µM primer Ro13pr, 0.625 unit of HotStar Taq DNA polymerase (QIAGEN), 0.75 unit *Pfu* DNA polymerase (Promega), and sterile distilled water to give a final volume of 25 µl. The PCR was performed by incubating at 94°C for 5 min as initial denaturation step, followed by 30 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 4 min, and a final extension of 10 min at 72°C after the last cycle.

One microliter of first round PCR product was amplified with the sense primer and the antisense primer as indicated above to introduce a *Bam*H I (by Roex5f primer) or *Sna*B I (by Roex8f primer) site at the 5′end and an *Eco*R I (by β-glu135) site at the 3′ end. The PCR reaction mix included 2.5 µl 10 X *Pfu* buffer, 2 µl 2.5 mM dNTP, 1 µl each 10 µM sense and antisense primers, 0.75 unit *Pfu* DNA Polymerase, 1 µl PCR product, and distilled water to bring the volume up to 25 µl. The PCR was done at

94°C for 5 min as the initial denaturation step, followed by 30 cycles of 94°C for 45 s, 52°C for 30 s, and 72°C for 4 min, and a final extension of 10 min at 72°C after the last cycle. *Pfu* DNA polymerase has proof reading ability to minimize mutations during PCR and decreasing the number of PCR cycles also helps. Mixing HotStar Taq DNA polymerase with *Pfu* DNA polymerase in the first round of amplification was necessary to amplify the GC rich 5′ end region, while allowing proof reading of the DNA synthesis.

The *bglu2* cDNA was amplified using similar strategies to those used for the *bglu1* cDNA. The first round of PCR to amplify the full-coding sequence was done with the sense β-glu215 primer and the antisense primer β-glu218 (Table 2.7). The PCR reaction mix included 2.5 µl 10 X *Pfu* buffer, 2 µl 2.5 mM dNTP, 1 µl 10 µM β-glu218 primer, 1 µl 10 µM β-glu215 primer, 0.75 µl *Pfu* DNA polymerase, 1 µl first strand cDNA, and distilled water to bring the volume to 25 µl. The PCR was performed at 94°C for 5 min as the initial denaturation step, followed by 30 cycles of 94°C for 45 s, 52°C for 30 s, and 72°C for 4 min, and a final extension of 10 min at 72°C after the last cycle. The second amplification was performed as above using one microliter of the PCR product from the first round amplification as template, and with the Roex3f sense primer and the Roex1r antisense primer to introduce an *Nco* I site at the 5′ end and an *Xho* I site at the 3′ end.

The PCR products were evaluated to confirm they were the correct size on 1% agarose gel and purified as described in methods 2.2.5-2.2.6.

Table 2.7 Primers used for amplification of *bglu1* and *bglu2* cDNA

Primer name	Primer sequence
β-glu217	5′-GTGCGTGCGTGGCTGCGTGTGCTT -3′
Ro13pr	5′-CACCAAGCCAAATCTCATCAAC -3′
β-glu215	5′-GATCAGAATGGGAAGGCGACTCC-3′
β-glu218	5′-AGCCTGAACCAGCCCAACACTGCTC-3′
Roex5f	5′-TGCCGGATCCGTGCCCAAGCCCAACTG GCTG-3′
Roex8f	5′-TGCCTACGTAGTGCCCAAGCCCAACTGGCTG-3′
β-glu135	5′-AGTCTGCGTGTCTTGAATTCTCAGTGC-3′
Roex3f	5′-ATGGCCCCATGGCCAAGTTTACCCGCTACAGCTTC-3′
Roex1r	5′-CTCGAGATCAGTGCCTCTTGAGGA-3′

2.7.2 Construction of recombinant plasmids

The PCR products were cloned into the *Eco*R V site of pBluScript II SK(+) and transformed into DH5α *E. coli*. The resultant recombinant plasmids were purified and digested with appropriate restriction enzymes, as described in methods 2.2.6-2.2.12. The gel-purified inserts for *bglu1* were cloned into the *Bam*H I and *Eco*R I sites of pET23d(+) and *Eco*R V and *Eco*R I sites of pET32a(+), *bglu2* were ligated to the *Nco* I and *Xho* I sites of pET23d(+), pET32a(+), pET39b(+), and pET40b(+) (Novagen), as described in method 2.2.9. The ligation reaction mix were transformed into DH5α *E. coli* and spread on LB agar plates containing selective antibiotic resistance of each plasmid type as indicated in Table 2.8. Transformation of DH5α *E. coli* with the ligation reaction mix of recombinant plasmid helped to multiply the number of recombinant plasmids, and these plasmids can be kept permanently (1-4 years) in this host cells.

2.7.3 Identification and isolation of recombinant expression plasmids

The recombinant pET plasmids were identified by a colony PCR method. The transformant colonies were randomly selected, restreaked on a new LB agar plate containing the appropriate antibiotics, and incubated overnight at 37°C. The restreaked colonies were picked with a sterile toothpick and resuspended in 70 µl sterile water in a 1.5 ml microtubes until the solution looked slightly turbid. The suspensions were boiled at 100°C for 5 min, and the cell debris was spun down at 13,000 x g for 3 min. Three microliters of supernatant (plasmid phase) were used as the template in a PCR reaction. The PCR reaction mix included 2.5 µl 10 X HotStar Taq buffer, 2 µl 2.5 mM dNTP, 1 µl each 10 µM sense and antisense primers, 0.625 unit HotStar Taq DNA polymerase (QIAGEN), 3 µl of the plasmid aqueous phase supernatant and distilled water to bring the volume to 25 µl. The forward and reverse primer used in the reaction were Roex8f and β-glu135 for *bglu1*; and Roex3f and Roex1r for *bglu2*. The PCR reaction was done as described above. The PCR product was separated on 1% agarose gel electrophoresis to identify the clones that contained recombinant plasmid. Then the recombinant colonies were grown in LB broth containing the appropriate antibiotics and plasmid isolated by QIA Prep Spin Miniprep kit (QIAGEN) and sequenced, as described in methods 2.2.13 and 2.2.14.

2.7.4 Transformation of recombinant plasmid into expression host cells

The recombinant pET23d(+)-*bglu1* was transformed into BL21(DE3) and pET32a(+)-*bglu1* were transformed into Origami(DE3) and BL21(DE3) *E. coli* (Novagen). While the recombinant pET23d(+)-*bglu2* was transformed into BL21

(DE3) and BL21(DE3) pLysS and pET32a(+)-*bglu2* constructs were transformed into Origami(DE3). The pET39b(+)- and pET40b(+)-*bglu2* constructs were transformed into BL21(DE3). The competent cells of these host strains were prepared as described in Appendix C. Plasmids (10-100 ng) were used to transform a 50 µl aliquot of the frozen competent cells, according to the standard method as described in method 2.2.10 and spread on LB agar plates containing selective antibiotic resistance of each plasmid type and host strains as indicated in Table 2.8 and 2.9. The recombinant clones were identified by colony PCR method as described above.

Table 2.8 pET vector characteristics

Vector	Antibiotic resistance	Promotor	Tag	Protease cleavage site
PET23d(+)	Ampicillin (50 µg/ml)	T7	N-terminal T7 tag, C-terminal His tag	None
PET32a(+)	Ampicillin (50 µg/ml)	T7*lac*	Internal and C-terminal His tag, Internal S tag, N-terminal Trx tag, Internal and C-terminal His tag,	Thrombin, enterokinase
pET39b(+), pET40b(+)	Kanamycin (30 µg/ml)	T7*lac*	Internal S tag, N-terminal DsbA (for pET39b), and DsbC (for pET40b) tag	Thrombin, enterokinase

Table 2.9 pET system host strain characteristics

Strain	Genotype	Description application	Antibiotic Resistance[1]
Bl21(DE3)	F^- *ompT* hsdS_B(r_B^- m_B^-) *gal dcm* (DE3)	General purpose expression host	None
BL21(DE3) pLysS	F^- *ompT* hsdS_B(r_B^- m_B^-) *gal dcm* (DE3) pLysS (Cm^R)	High-stringency expression host[2]	Chloramphe-nicol (34 µg /ml)
Origami(DE3)	*Δara-leu7697 ΔlacX74 ΔphoAPvuII phoR araD139 galE galK rspL* F' [*lac*+(*lacIq*)*pro*] *gor522* ::Tn*10* (TcR) *trxB*::kan (DE3)	General expression host; two mutations in redox gene to allow disulfide bond formation in *E. coli* cytoplasm	Tetracycline (12.5 µg/ml), Kanamycin (15 µg/ml)

1. The appropriate antibiotic to select for the target plasmid must also be added

2. High-stringency means that the strain carries pLysS, a pET-compatible plasmid that produces T7 lysozyme, thereby reducing basal expression of target genes, by inhibiting T7 polymerase.

2.7.5 Protein expression in *Escherichia coli*

For expression, the selected clones were grown overnight in LB broth containing appropriate antibiotics to select for the target plasmid and host strain as indicated in Table 2.8 and 2.9. The fresh inoculum cultures were then diluted to a ratio 1:100 with LB broth containing appropriate antibiotics. The cultures were grown in a shaking incubator at 37°C. After the optical density at a wavelength of

600 nm of the cultures reached 0.5-0.6, IPTG was added to a final concentration of 0.4 mM and the cultures were shaken at 200 rpm at 20°C for an additional 8 h. Induced cultures were transferred to 50 ml centrifuge tubes and chilled on ice for 10 min and then harvested by centrifugation at 5000 rpm at 4°C for 10 min, and the cell pellet kept at -70°C for 30 min or longer until used for protein extraction.

To optimize protein expression, the optimal IPTG concentration, temperature, time and medium type for growing cells during induction of expression from the cells were determined. During induction of protein expression, IPTG concentrations were varied between 0.1, 0.2, 0.4, 0.6, 0.8, and 1 mM. The induction temperatures tested were 15°, 20°, 27°, and 37°C with times of induction of 4, 8 and 18 h.

To determine the protein expression level, total protein profiles of host cells containing control-empty plasmids and recombinant plasmids were analyzed. For this, the cells were pelleted from 1 ml liquid culture in a microtube and the pellets resuspended in 40 µl of 2X SDS sample buffer (125 mM Tris pH 6.8, 15% glycerol, 1% SDS, 5.6% 2-ME, 0.05% bromophenol blue). The suspension was boiled for 5 min and 8-12 µl aliquots were loaded onto 12% SDS polyacrylamide gels (SDS-PAGE) as described in method 2.7.10.

2.7.6 Extraction of recombinant protein from induced cells

Bacterial culture pellets were thawed, suspended in 5 ml per gram freshly prepared extraction buffer (20 mM Tris-HCl pH 8.0, 200 µg/ml lysozyme, 1% Triton-X100, 1 mM phenylmethylsulfonylfluoride (PMSF), 40 µg/ml DNase I), and incubated at room temperature for 30 min. The cell suspensions were sonicated on

ice 5 times at 50 W output for 10 s, with 1 min cooling in between, with an Ultrasonic Processor GE 100 probe sonicator (Treadlitei, Woodstock, MA). The pellets were extracted twice with extraction buffer to increase the amount of soluble protein. The soluble protein fraction was recovered by centrifugation at 12,000 x g at 4°C for 10 min and subjected to electrophoresis by SDS-PAGE.

2.7.7 Purification of recombinant protein by Ni-NTA superflow

Purification of BGlu1 was done with Ni-NTA superflow (QIAGEN) at 4°C. Ten milliliters of the soluble extract of bacteria expressing BGlu1 protein from pET32a(+) was added to a 1 ml of Ni-NTA superflow column (QIAGEN) equilibrated with 8 ml (8 volumes) equilibration buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl). The column was washed with 4 ml (4 volumes) of equilibration buffer, 8 ml (8 volumes) of wash buffer 1 (5 mM imidazole in equilibration buffer), and 4 ml (4 volumes) of wash buffer 2 (10 mM imidazole in equilibration buffer). The recombinant protein was eluted with 4 ml (4 volumes) each of elution buffer (250 mM imidazole in equilibration buffer). The 0.5 ml fractions with β-glucosidase activity were pooled and concentrated 4 fold and exchanged with 20 mM Tris-HCl pH 8.0 in a 30 kDa cutoff centrifugal ultrafiltration membrane (YM-30, Amicon).

2.7.8 Cleavage of recombinant protein

One hundred micrograms of recombinant protein was cleaved with 1 unit of enterokinase in 50 μl of 20 mM Tris-HCl pH 7.2, 200 mM calcium chloride at 23°C for 16 h. The reaction was added to 50 μl of Ni-NTA superflow and incubated at 4°C for 60 min with gently shaking. The reaction mix was centrifuged at 4°C for 10 min

at 12,000 x *g,* and the supernatant containing purified BGlu1 was collected and evaluated by SDS PAGE and β-glucosidase activity determined.

2.7.9 Protein determination

Protein concentrations of the crude extracts of *E. coli* were estimated by the method of Lowry (Lowry *et al*., 1951) using bovine serum albumin (BSA) as standard (15-105 μg BSA). The proper dilution of protein sample in 1 ml was mixed with 1.0 ml alkaline copper solution (0.05% $CuSO_4$, 0.1% potassium-sodium tartate, 5% Na_2CO_3 and 0.4 M NaOH). The reaction was left to stand for 10 min at 28°C and 0.5 ml diluted Folin diluted reagent (2:1 to water) was then added. The mixture was vigorously mixed by vortexing and allowed to stand at 28°C for 30 min. The absorbance at 750 nm was measured using a Jenway 6405 UV/VIS Spectrophotometer (Essex, UK).

Purified recombinant protein concentration was determined by Bio-RAD Protein Assay kit (Bio-RAD) using BSA as the standard (2-10 μg). Appropriate dilutions of protein samples were prepared in 0.2 ml volume and mixed with 0.8 ml Bio-RAD Protein Assay Solution. The reaction was allowed to proceed for 10 min and the absorbance at 595 nm was measured using a Jenway 6405 UV/VIS Spectrophotometer (Essex, UK).

2.7.10 Determination of the subunit molecular weight of protein by SDS polyacrylamide gel electrophoresis

Denaturing polyacrylamide gel electrophoresis was performed according to the method of Laemmli (1970). The stock solution of polyacrylamide consisted of

29% (w/v) acrylamide and 1% (w/v) of *N, N′*- methyllene-bis-acrylamide. The separating gel consisted of 12% polyacrylamide, 0.375 M Tris-HCl pH 8.8, 0.1% SDS, 0.1% ammonium persulphate, 0.05% TEMED. The separating gel consisting of 5% polyacrylamide, 0.126 M Tris-HCl pH 6.8, 0.1% SDS, 0.1% ammonium persulphate, 0.05% TEMED. One part of protein solutions was mixed with 3 parts of 4 X sample buffer (250 mM Tris pH 6.8, 30% glycerol, 2% SDS, 11.2% 2-ME, 0.1% bromophenol blue). The suspension was boiled for 5 min, 8-12 μl aliquots were loaded onto 12% SDS polyacrylamide gel (SDS PAGE) using a Mini-protein 3 cell (Bio-RAD), and electrophoresed in 1 X Tris-glycine pH 8.3 SDS buffer (Appendix D) at a constant 150 V for 1 h from cathodic (-) to anodic (+) end. After electrophoresis, the gel was stained with Coomassie Brilliant Blue R 250 (Appendix D) for 30 min and destained with destaining solution (Appendix D) for 1 h or until the protein bands appeared clear. The sizes of protein bands were estimated by comparing with Bio-RAD protein Markers (Bio-RAD): phosphorylase b (97.4 kDa), bovine serum albumin (66 kDa), ovalbumin (45 kDa), bovine carbonic anhydrase (31 kDa), trypsin inhibitor (21.5 kDa) and bovine α-lactalbumin (14.0 kDa).

2.7.11 Estimation of the native molecular weight of protein by gel filtration chromatography

Gel filtration chromatography on a Sephacryl S-200 (Pharmacia) with gel filtration molecular weight markers 12.4-200 kDa (Sigma) was used to estimate the native molecular weight of thioredoxin fusion BGlu1 according to the supplier’s protocol. A 70 x 1.6 cm Sephacryl S-200 column was equilibrated with 50 mM Tris-HCl pH 7.5, 100 mM KCl at 4°C. To determine the void volume (V_0), 0.5 ml of 4

mg/ml blue dextran in equilibration buffer containing 5% glycerol was applied to the column. The column was eluted with equilibration buffer at a flow rate of 0.5 ml/min. As soon as the blue color of blue dextran reach the gel bed surface, collecting of fractions was started with 1 ml/min and the absorbance at 280 nm measured for plotting the effluent peak. The elution volume of blue dextran (V_0) was determined by measuring the volume of effluent collected from the point of the center of the effluent peak. To determine the elution volume (V_e) of the standard and purified thioredoxin fusion BGlu1, each protein was dissolved in elution buffer containing 5% glycerol as the concentration shown in Table 2.10. Each protein sample (0.5 ml) was applied to the column separately, and eluted with the same conditions used for blue dextran. The standards were measured by absorbance reading at 280 nm, and the BGlu1 was measured by pNPGase activity assay (see method 2.8.1). The V_e for each protein was determined in the same way as for blue dextran. The molecular weight and V_e/V_0 for each protein was plotted on semilog paper and the thioredoxin fusion BGlu1 molecular weight was estimated.

Table 2.10 Concentration of protein solutions used for gel filtration

Protein sample	Mr (kDa)	Concentration (mg/ml)
Albumin, Bovine	66	10
Alcohol dehydrolgenase	150	8
Carbonic anhydrase	29	5
Cytochrome c	12.4	2.5
Thioredoxine fusion BGlu1	-	1

2.7.12 Western blot analysis

Protein samples were mixed with one-fourth volume of the 4 X SDS-gel sample buffer and heated at 100°C for 5 min. The protein samples were loaded onto a 12% SDS-PAGE gel and electrophoresed. After electrophoresis, the gel was soaked in blotting buffer (25 mM Tris, 125 mM glycine, 0.25% SDS, 5% methanol), and the proteins transferred onto the nitrocellulose membrane (Hybond ECL, Amersham Pharmacia Biotech) on a Semidry gel blotting system (Bio-Rad). After transfer, the membrane was stained with FAST-Green BB dye (Appendix D) to confirm the presence of protein bands on the membrane. The membrane was washed twice with 10 ml PBS. The membrane was incubated with 5% Non-fat milk in PBST (blocking solution) for 1-4 h for non-specific blocking. Immunodetection was carried out by incubating the membrane in a 1:1000 dilution of an anti-rice β-glucosidase serum in blocking solution for 2-4 h. Anti-rice β-glucosidase sera used were 403, 404 and 406 and the production of these antisera was described in Appendix C. After incubation with antiserum, the membrane was washed 4 times with PBST, followed by incubation in a 1:5000 dilution of the secondary antibody, goat anti-rabbit coupled to horse radish peroxidase, in PBST for 2 h. The membrane was washed twice with PBST and twice with PBS, respectively with 10 min per each wash. Colorimetric detection was done with 4-chloro-1-napthol substrate solution (5.5 ml 3 mg/ml 4-chloronapthol (in 100% methanol), 10 µl 30% H_2O_2, 19.5 ml PBS). Immunoreactive bands yielded brown-violet color.

2.8 Biochemical characterization

2.8.1 β-glucosidase assays and kinetics study

β-glucosidase activity on various substrates was assayed by determining either (1) the nitrophenol liberated from the pNP derivatives of monosaccharides or disaccharides, or (2) glucose released from natural or artificial glucoside substrates by PGO assay as previously described (Babcock and Esen, 1994). All solutions, except those specified otherwise, were prepared in 50 mM NaOAc buffer, pH 5.0. The first assay, the pNP assay, was used to test sugar specificity. In this assay, 70 µl of various dilutions of substrates were placed into the wells of a microtiter plate. Next, 70 µl of appropriately diluted enzyme were added to the wells to start the reaction. An appropriate dilution of enzyme is one that gives a V_m equivalent to 0.1 to 1.0 absorbance units under the conditions used. The assays were done at 30°C for 10 to 60 min depending on each substrate to establish the initial velocity (V_0). The reaction was stopped by adding 70 µl of 0.4 M Na_2CO_3 prepared in water. The absorbance of the p-nitrophenol liberated was measured at 405 nm.

The second assay, the peroxidase/glucose oxidase-based assay (PGO assay), was used to test aglycone specificity. PGO enzymes (Sigma) come in capsules, each of which contains 500 units of glucose oxidase (*Aspergillus niger*), 100 Purpurogalin units of peroxidase, and buffer salts. One capsule was dissolved in 100 ml sterile distilled water. In the reaction, 25 µl of various dilutions of glucoside or glucose (as a standard) were placed in a 1.5 ml tubes. Next, 25 µl of appropriate dilutions of enzyme were added to start the reactions. After 10 min or longer (not more than 60 min), the reaction was stopped by heating at 100°C for 5 min and the reactions were

placed in the wells of a microtiter plate. Next, 50 μl of 1 mg/ml 2,2'-azinobis(3-ethylbenzthiazolinesulfonic acid) (ABTS) were added to each well. Lastly, 100 μl of PGO enzyme was added to each well and mixed. The reaction was allowed to proceed for 30 min at 37°C. The absorbance of the product from the PGO assay (green color) was measured at 405 nm.

One unit of β-glucosidase activity was defined as the amount of enzyme that produced 1 μmol of product per min. Note that the activity values for disaccharides were determined by dividing the amount of glucose released by two, since two glucose molecules are released per molecule of disaccharides hydrolyzed. The micromole of product definition for oligosaccharides is in terms of total glucose released, though oligosaccharides may also have more than one glucose released per substrate molecule due to sequential cleavage.

Kinetic parameters, K_m and V_{max} (at pH 5.0, at 30°C), of purified BGlu1 were calculated according to the method of Lineweaver and Burk using the Enzfitter computer program (Elsevier Biosoft, Cambridge, U.K.). For substrate, 3 replicates each of 5-7 substrate concentrations were used. Protein assays were performed colorimetrically by the Bio-RAD Protein Assay (Bio-RAD) using BSA as a standard as described in method 2.7.9.

2.8.2 pH and temperature profiles of activity and stability

To determine the pH optimum for enzyme activity, BGlu1 and BGlu2 (0.005 nmol) were prepared in different buffers ranging from pH 3 to 11 (formate pH 3-4; NaOAc pH 4.5-5.5; sodium phosphate pH 6-8.0; Tris pH 7.5-8.5; CAPS pH 9-9.5; CAPS pH 10-11), at 0.5 pH unit intervals at the same buffer concentration of 50 mM.

The activity at a given pH was measured by mixing the enzyme solution with 1 mM final concentration of p-nitrophenol β-D-glucoside (pNPG) in the same pH buffer. The reactions were allowed to proceed for 10 min at 30°C and the amount of product released determined as described in method 2.8.1.

The temperature optimum was determined by incubating the BGlu1 with 1 mM final concentration of pNPG in 50 mM NaOAc pH 5.0 over a temperature range of 5 to 70°C at 5° increments for 10 min. The amount of product released was determined as described in method 2.8.1.

The pH stability of BGlu1 was determined by incubating BGlu1 in buffers varying in pH 3 to 11 at increments of 1.0 pH unit for 10 min, 1, 3, 6, 12 and 24 h at 30°C. After the incubation, the enzyme was diluted 140 fold in 50 mM NaOAc buffer pH 5.0. All the samples were assayed for activity with 1 mM pNPG in 50 mM NaOAc pH 5.0 at 30°C for 10 min.

Thermostability of the enzyme was measured by incubating the enzyme in 50 mM NaOAc buffer pH 5.0 over a temperature range of 20° to 60°C at 10° intervals. At 10 min increments from 0 to 60 min, samples incubated at each temperature were placed on ice. All the samples were assayed for activity with 1 mM pNPG in 50 mM NaOAc pH 5.0 at 30°C for 10 min.

Note that all parameter testing reactions were done in triplicate. Enzyme molarity was calculated based on molecular weight of 66 kDa of BGlu1.

2.8.3 TLC

Some substrates hydrolyzed by BGlu1 were analyzed by TLC. Purified BGlu1 was incubated with a 5 mM final concentration of glucoside substrates

prepared in 50 mM NaOAc pH 5.0. In the reaction, 25 μl of 10 mM substrate was mixed with 25 μl purified BGlu1 (0.05 nmol), and incubated at 30°C for 1 h. Five microliters of the reaction mixture was spotted on a 7 x 4 cm F_{254} 0.2 mm silica-coated aluminium plate (Merck, Darmstadt, Germany) and chromatographed vertically at room temperature for 10 min. The solvent system was chloroform: methanol: NH_4OH, at a ratio (v/v/v) of 6:3.6:0.4 for pNP-cellobiose and 7:3:2 for pyridoxine glucoside. The products were detected under 356 nm UV light, then sprayed with developer solution (4% anisaldehyde, 4% H_2SO_4, 3% water in ethyl alcohol) and then baked at 100°C for 2-3 min to visualize the glucose. For each substrate, a control without enzyme was included in the assay.

The hydrolysis and transglycosylation activity towards disaccharides and oligosaccharides by BGlu1 was investigated by TLC. Purified BGlu1 was incubated in 5 mM glucoside substrates prepared in 50 mM NaOAc pH 5.0. In the reaction, 25 μl of 10 mM glucoside were mixed with 25 μl purified BGlu1 (0.05 nmol) in a 1.5 ml tube. The reactions were incubated at 30°C for 1-4 h and stopped by heating at 100°C for 5 min. Five microliters of the reaction mixture was spotted on the 7 x 10 cm F_{254} 0.2 mm of silica-coated aluminium plate and chromatographed vertically at room temperature for 20-30 min in (2:2:1) (v/v/v) ethyl acetate: acetic acid: water. The spots of the products were detected by spraying with developer solution and then baking at 100°C for 2-3 min to visualize the release of glucose (hydrolysis) or the synthesis of glucosides (transglycosylation).

Note that enzyme molarity was calculated based on molecular weight of 66 kDa of BGlu1.

2.8.4 Determination of the kinetic parameters of pyridoxine glucoside hydrolyzed by BGlu1 using high performance liquid chromatography (HPLC)

The release of glucose from pyridoxine glucoside by BGlu1 could not be quantified accurately by the glucose oxidase assay, because pyridoxine or its breakdown products interfered with the assay, producing a reddish color, rather than the usual green with ABTS. Therefore, an HPLC method was developed to quantitate the amount of pyridoxine (Vitamin B_6) glucoside hydrolyzed. Solutions of 0.004 to 0.04 μmol pyridoxine were prepared as standard. Various concentrations of pyridoxine-5′-O-β-D-glucoside were incubated with 0.01 nmol purified BGlu1, and 50 mM NaOAc pH 5.0 in 50 μl reactions at 30°C for 15 min. The reaction was terminated by heating at 100°C for 5 min. The reaction mixture was diluted 4 fold with 50 mM NaOAc pH 5.0 and filtered through a 0.45 micron filter and 100 μl volume was separated on an Eclipse XDB-C-18 column. (Agilent Corp, USA) with an HP Series 1100 HPLC (Hewlett Packard, Waldbronn, Germany). A gradient of methanol (0 to 80%) in 0.1% trifluoroacetic acid in water with a flow rate 0.8 ml/min was run to separate the products as follows:

Time	Percent methanol
0	0
20	10
25	20
30	80
35	0

Products were detected with an HP 1100 variable wavelength detector measuring absorbance at a wavelength of 290 nm. Kinetic constants were determined as

described in method 2.8.1. Note that enzyme molarity was calculated based on molecular weight of 66 kDa of BGlu1.

2.8.5 Non-denaturing gel electrophoresis for β-glucosidase gel activity staining

Non-denaturing gel electrophoresis, on both basic and acidic gels, was performed in duplicate, one for activity gel staining with glucoside substrate, and another for Coomassie Brilliant Blue R250 for protein profile analysis.

Basic non-denaturing gel electrophoresis was performed by a method adapted from Hames (1990). The separating gel consisted of 10% polyacrylamide, 0.375 M Tris-HCl pH 8.8, 0.1% ammonium persulphate, and 0.05% TEMED, while the separating gel consisted of 5% polyacrylamide, 0.126 M Tris-HCl pH 6.8, 0.1% ammonium persulphate, and 0.05% TEMED. One part of control nonrecombinant or BGlu1 protein solution was mixed with 3 parts of 4 X basic native sample buffer (250 mM Tris pH 6.8, 30% glycerol, 0.1% Bromophenol blue). Then 15 μl aliquots were loaded onto 10% native gels and electrophoresed in 1 X Tris-Glycine pH 8.3 native buffer (Appendix D) at a constant 30 V for stacking for 20-30 min and then at a constant 120 V for separation for 1.5 h from cathodic to anodic end at 4°C.

Acidic non-denaturing gel electrophoresis was performed by a method adapted from Hames (1990). The separating gel consisted of 10% polyacrylamide, acetic acid-KOH resolving gel buffer pH 4.3, freshly prepared 3 mM sodium sulfite, 0.075% ammonium persulphate, and 0.05% TEMED (Appendix D), while the stacking gel consisted of 5% polyacrylamide, acetic acid-KOH stacking gel buffer pH 6.8, 0.075% ammonium persulphate, and 0.05% TEMED (Appendix D). One part of control

nonrecombinant or BGlu1 protein solution was mixed with 3 parts 4 X acidic native sample buffer (2 X acidic stacking buffer pH 6.8, 30% glycerol, 0.1% methylene blue). The 15 µl aliquots were loaded onto the gel and electrophoresed in 1 X acetic acid-β-alanine running buffer pH 4.5 (Appendix D) at a constant 30 V for stacking for 20-30 min and at 120 V constant for separation for 3 h from positive pole to negative pole at 4°C.

After gel electrophoresis, the gels were equilibrated twice in 50 mM NaOAc pH 5.0 buffer on ice for 30 min each. Then, the gels were stained for β-glucosidase activity with 5 mM 4-methylumbelliferyl-β-glucoside (4MUGlc) for 30 min at 30°C and fluorogenic bands of activity was detected under UV light using a Fluor-S™ MultiImager (Bio-RAD).

2.8.6 Inhibition studies

For the inhibition studies, 1 mM pNPG was used as the substrate. At least three different inhibitor concentrations were used, each with 3 replicates. The inhibitors tested were: glucono-1,5-lactone, galacto-1,4-lactone, indole acetic acid (IAA), gibberellic acid 3 (GA_3), 4-methylumbelliferol, EDTA, Tris, ethanol, monosaccharide (glucose, galactose, mannose, xylose). In most cases, 70 µl of 2 mM pNPG dilutions were placed in the wells of a microtiter plate. Thirty-five microliters of various concentrations of inhibitor prepared in the same buffer were added to each well, followed by 35 µl of enzyme (0.005 nmol) diluted in the same buffer to start the reaction. The reactions were allowed to proceed to 10 min at 30°C, then the reaction was stopped by adding 70 µl 0.4 M Na_2CO_3, and the absorbance was read at 405 nm.

IAA and GA_3 were dissolved in 50% ethanol, and a control of 50% ethanol-50% buffer was used to correct for inhibition by ethanol alone.

2.8.7 Effect of reducing and denaturing agents on BGlu1 activity

The effect of SDS on BGlu1 activity was studied by incubating the BGlu1 with 1 mM pNPG in 50 mM NaOAc pH 5.0 in the presence of various concentrations of SDS (0.1, 0.05, 0.025, 0.01, 0.005, 0%). In the reaction, 70 μl of 2 mM pNPG dilutions were placed in the wells of a microtiter plate. Thirty-five microliters of various concentration of SDS prepared in the same buffer were added to each well, followed by 35 μl of enzyme (0.005 nmol), diluted in the same buffer, to start the reaction. The reactions were allowed to proceed to 10 min at 30°C, then the reaction was stopped by adding 70 μl 0.4 M Na_2CO_3, and the absorbance was read at 405 nm.

The effect of reducing agent 2-mercaptoethanol (2-ME) on BGlu1 activity was studied by incubating the BGlu1 with 1 mM pNPG in 50 mM NaOAc pH 5.0 in the presence of various concentrations of 2-ME (0-2000 mM). The activity assays were done as described above for the SDS effect experiment.

2.8.8 Effects of metal ions on enzyme activity

The effects of selected cations on enzyme activity were determined by adding the following metal ion salts to 10 mM final concentration: $HgCl_2$, $ZnSO_4$, $MgCl_2$, $MnCl_2$, $CaCl_2$, and $CuSO_4$. Purified BGlu1 solution was incubated in the presence of metal ions for 30 min. The enzyme was taken and diluted 70 times in 50 mM NaOAc pH 5.0 and assayed for activity with 1 mM pNPG as described above. The inhibition by Hg^{2+} was tested further 0 to 256 μM Hg^{2+} concentration.

2.9 Transglycosylation reactions

To determine the transglycosylation activity of BGlu1 using pNPG as the glucose donor molecule, various reaction conditions and glucose acceptor compounds were used. Glucosyl transfer reactions were studied with 3 different acceptor molecules, pNPG, ethanol and pyridoxine.

In the reactions using pNPG as a glucose acceptor molecule, 50 µl of various concentrations of pNPG substrate (5, 10, 20, or 40 mM) in 50 mM NaOAc pH 5.0 were mixed with 0.01 nmol purified BGlu1, and 50 mM NaOAc pH 5.0 to bring the volume to 100 µl. For each substrate concentration, a control without enzyme was included in the assay.

In the reactions using ethanol as a glucose acceptor molecule, 50 µl of 20 mM pNPG was mixed with various concentrations of ethanol (5, 10, 20, or 30%), 0.01 nmol purified BGlu1, and 50 mM NaOAc pH 5.0 to bring the volume to 100 µl. For each substrate concentration, a control without enzyme and without ethyl alcohol were included in the assay.

Glucosyl transfer reactions using pyridoxine (Vitamin B_6) as a glucose acceptor molecule, freshly prepared 5, 10, 20 and 40 mM pNPG and 1, 5, 10 and 20 mM pyridoxine in 50 mM NaOAc pH 5.0 was incubated with 0.01- 0.02 nmole purified rice BGlu1 in a 100 µl reaction. For each reaction, controls without enzyme and without pyridoxine were included in the assay.

All assay reactions were incubated at 30°C for 1, 3, 6, 12 or 24 h in the dark. The reactions were stopped by heating at 100°C for 5 min. Note that enzyme molarity was calculated based on molecular weight of 66 kDa of BGlu1.

TLC was carried out to visualize the reaction products. Five microliters of the reaction solution were spotted on a 10 x 10 cm F_{254} 0.2 mm of silica-coated aluminium plate. The silica plate was chromatographed vertically at room temperature in 7:2.8:0.2 (v/v/v) chloroform: methanol: NH_4OH for 40 min. The products were visualized by detecting with 356 nm UV light, then sprayed with developer solution (4% anisaldehyde, 4% H_2SO_4, 3% water in ethyl alcohol) and baked at 100°C for 2-3 min to visualize the compounds without UV absorbance.

Quantification of the intensity of compounds separated by TLC was performed by densitometry scanning of the TLC plates with the Fluor-S^{tm} MultiImager and the intensity of signal quantified using Quantify One software (Bio-RAD). The known standards purchased from Sigma were: pNPG, pNP-β-cellobioside, pNP-β-cellotrioside, pNP-β-gentiobiose and pyridoxine. Standard pyridoxine-5′-O-β-D-glucoside was provided by Prof. J. F. Gregory.

The reaction mix of transglucosylation of pyridoxine was filtered through a 0.45 micron filter and separated on a 7.6 x 250 mm C-18 column (GL Sciences Inc., Japan) with an HP Series 1100 HPLC (Hewlett Packard, Waldbronn, Germany). A gradient of methanol (0-80%) in 0.1% trifluoroacetic acid in water was run to separate the product compounds as described in method 2.8.4. Products were detected by absorbance at a wavelength of 290 nm.

For structural analysis of pyridoxine glucoside product, the reactions was scaled up to 15 ml reaction and the pyridoxine glucoside product was purified from a 7.6 x 250 mm C-18 column and subjected to NMR analysis. The ^{1}H, ^{13}C, ^{1}H-^{1}H COSY, HMQC, and HMBC spectra were recorded on a INOVA-300, operating at 299.9 MHz for proton, and 75.4 MHz for carbon. All samples for NMR analysis were

dissolved in D_2O. Determination of the anomeric configulation was based on observed chemical shifts, and the $J_{1,\ 2}$ coupling constants measured from one-dimentional ^1H spectra. ^1H-^1H shift correlation spectroscopy (COSY), heteronuclear multiple quantum coherence (HMQC) and heteronuclear multiple bond correlation (HMBC) experiments were made in order to assign the respective proton and carbon signals

2.10 The effect of glucono-1,5-lactone on rice germination

To study the effect of glucono-1,5-lactone which is a rice β-glucosidase inhibitor on rice germination, the rice seeds were soaked and germinated in the presence of this inhibitor. The experiment was done in duplicate with 110 seeds for each treatment. One hundred seeds were counted for the percentage of germination, while 10 seeds were subjected to 5-bromo-4-chloro-3-idolyl-β-D-glucoside (X-glucoside) staining assay to check for the occurrence of β-glucosidase activity in rice seed or tissues.

Orion rice seeds were sterilized by soaking in 10% Clorox for 10 min, followed by washing twice with sterile distilled water for 2-3 min each, soaking in 95% ethyl alcohol for 3 min, and washing 4 times with sterile water for 2-3 min each. Then, the seeds were soaked in the absence or in the presence of various concentrations (0, 25, 50, 100 mM) of glucono-1,5-lactone prepared in sterile distilled water for 12 h. The seeds were germinated in the dark at 28°C under sterile condition for 4 days on germinating paper in a beaker in the presence of the concentrations of glucono-1,5-lactone in which they were previously soaked. Each day, a number of

germinated seeds in each beaker were counted and the length of their shoots or roots measured.

Each day rice seeds from each beaker were sampled for β-glucosidase activity in rice tissues by staining with X-glucoside substrate. After sampling, the rice seeds or germinated rice seeds were soaked in 1 ml X-glucoside substrate solution (1 mM X-glucoside, 100 mM sodium phosphate buffer pH 7.0, 1 mM EDTA pH 8.0) for 30 min at 30°C. β-glucosidase activity was detected as the appearance of blue color in rice tissues.

2.11 Genomic analysis

2.11.1 Database searching and sequence analysis

Identification of rice genes homologous to glycosyl hydrolase family 1 β-glucosidase was done using the BLAST suite of programs (Altschul, *et al.*, 1990) in 4 databases: GenBank at NCBI (http://www.ncbi.nlm.nih.gov, Bethesda, USA), the Monsanto Rice Genome Draft Database (http://www.rice-research.org; Barry, 2001), the Beijing Genomic Institute (http://btn.genomics.org.cn/rice/) and the Syngenta Torrey Mesa Research Institute database (http://portal.tmri.org/). Identification of homologous genes and cDNA was done using tBLASTn and the known β-glucosidase protein sequences: maize *bglu1* (AC U33816), barley BGQ60 (AC L41869), and *Arabidopsis* psr3.2 (AC U72155), while BLASTn was used to identify sequences from the same gene. Coding regions of genes were identified by BLASTX searches against the NCBI nr protein database. Exact splice sites were predicted by identification of splice site consensus sequences near the ends of identified coding regions, which maintained the correct reading frame.

Protein sequence alignment and phylogenetic analysis was undertaken in order to form a rational evolutionary framework relating the members of the β-glucosidase gene family (Glycosyl hydrolase family 1) in plants. Translation of gene sequences was done using the 6-frame translation facility at the Baylor College of Medicine (BCM) search launcher site (http://dot.imgen.bcm. tmc.edu:9331/seq-util) (Smith *et al*., 1996). The ClustalX implementation of ClustalW was used for protein sequence alignments (Jeanmougin *et al*., 1998; Thompson *et al.*, 1994) and phylogenetic analyses done by the built in NJ-tree facility of this program with bootstrapping (1000 iterations).

Gene organization of the full-length and near full-length genes was also constructed from the conservation of introns and exons in rice β-glucosidase gene structures. The data from sequence and gene structure analyses were correlated to describe the evolutionary relationships between the genes. Each β-glucosidase gene sequence was searched against the GenBank at NCBI (http://www.ncbi.nlm.nih.gov, Bethesda, USA) using BLASTn to identify the chromosomal locations.

Protein parameters (Mr, pI) predictions were done at the Expasy proteomics server (http://www.expasy.org/), signal sequences were done by SignalP (Nielson *et al*., 1997) and cellular locations were predicted by PSORT (Nakai and Horton, 1999).

2.11.2 Digital northern analysis

A BLASTn search using the derived cDNA sequence for each gene was used to determine the relative abundance of ESTs derived from each rice β-glucosidase gene in libraries of various origins. All EST clone IDs were retrieved and collected in

the catalog to compare gene expression in various library sources. In addition, rice-specific tBLASTn searches using known β-glucosidase protein sequences were performed in the dbEST to identify all ESTs encoding β-glucosidase proteins from rice, as described for gene identification.

Chapter III

Results

3.1 β-Glucosidase cDNA cloning

Primary structures of two β-glucosidase mRNAs from rice were determined using PCR cloning and database searching methods. For sequence determination of the first rice β-glucosidase cDNA (*bglu1*), a partial cDNA sequence was previously identified in Prof. Dr. Asim Esen's laboratory by identifying rice EST sequences homologous to plant β-glucosidases in the dbEST database at NCBI. The partial sequences in the database of the corresponding clones, R2847_1A (GenBank AC D24959) and R2336 (GenBank AC D24663), had sequences encoding a glycosyl hydrolase family 1 β-glucosidase homologue and overlapped with each other (Figure 3.1). These two clones were then ordered from the Ministry of Agriculture, Forestry and Fisheries (MAFF), Japan, and sequenced completely. The overlapping sequences were the same, indicating both clones came from the same β-glucosidase gene. This sequence was deposited in GenBank, AC U28047.

In my study, the *bglu1* cDNA was amplified by RT-PCR using Orion rice cDNA as the template according to the strategy shown in Figure 3.2. To determine the 3'end sequence of the *bglu1* mRNA, the first strand cDNA was synthesized using mRNA extracted from 3-4 day old Orion rice seedling as template. The first strand cDNA was then used as a template in the 3'RACE reaction using the specific sense primer β-glu163, derived from the EST clone R2847_1A sequence, and oligo $(dT)_{17}$

as antisense primer to amplify a 3' end cDNA fragment (Figure 3.3). This PCR product was gel purified, and its sequence determined by direct sequencing. The sequence of this PCR fragment consisted of 1229 bp DNA, which included the stop codon and 3' untranslated region (3' UTR) of the mRNA, and was essentially identical to above 2 EST clones.

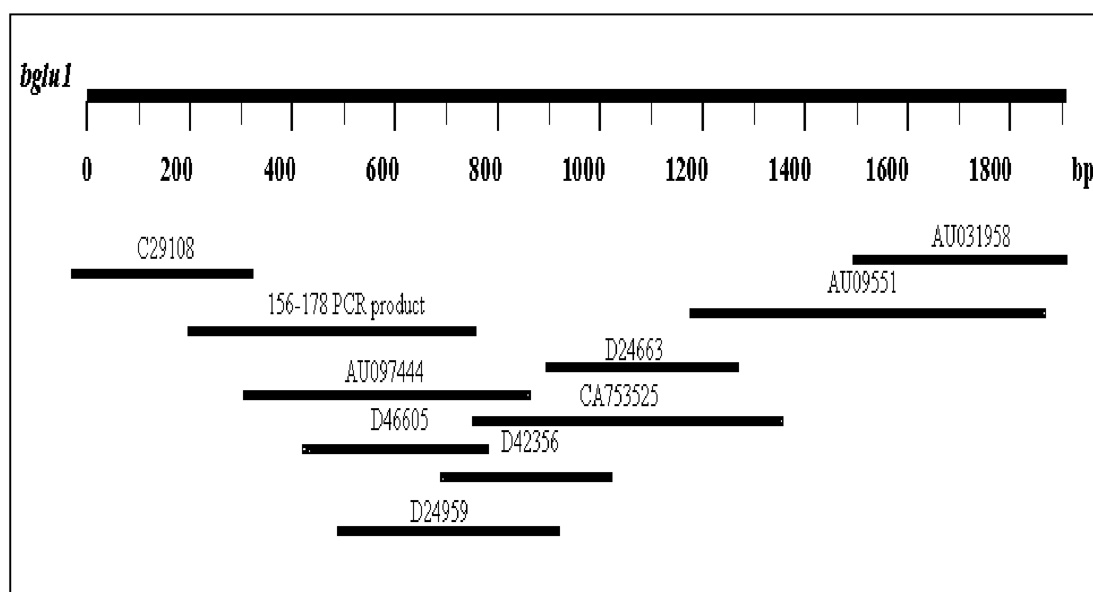


Figure 3.1 Rice EST clones corresponding to *bglu1* cDNA. Note that each EST clone was mapped to the region to which it is aligned on the *bglu1* (OSU28047) cDNA. The 156-178 PCR product is the cDNA fragment amplified by the β-glu156 and β-glu178 primers using Orion rice cDNA as the template.

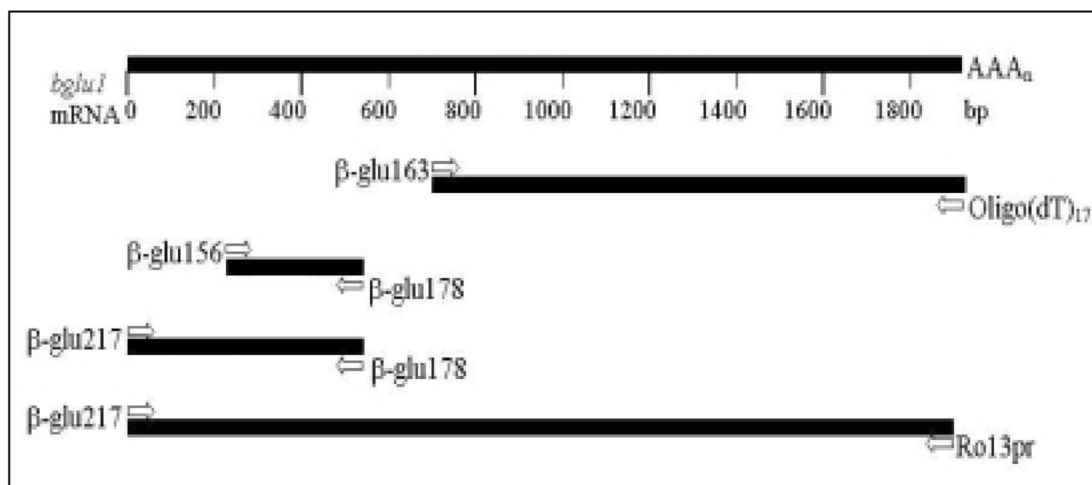


Figure 3.2 Cloning strategies of rice *bglu1*from Orion rice cDNA. The first line: the generation of a cDNA fragment encoding the 3′ end of *bglu1* by 3′RACE with β-glu163 and oligo $(dT)_{17}$ primers. The second line: the generation of cDNA fragment at the 5′ region of *bglu1* using the degenerate primers β-glu156 and β-glu178 derived from the EST clone R2847_1A sequence. The third line: the generation of cDNA fragment encoding the 5′ end of the *bglu1* mRNA using the specific primers β-glu217 (derived from EST clone c29108) and β-glu178. The fourth line: the generation of full-length cDNA using specific primers β-glu217 and Ro13pr.

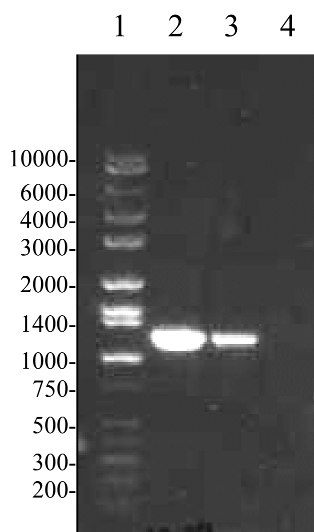


Figure 3.3 1% agarose gel electrophoresis of the 3′ RACE PCR product of *bglu1* cDNA fragment amplified with β-glu163 and oligo $(dT)_{17}$ primers. The PCR product was amplified by using Orion rice cDNA as template. Lane 1, Hi-low DNA marker; lanes 2 and 3, 3′ RACE PCR product; lane 4, no DNA template control PCR amplification with β-glu163 and oligo $(dT)_{17}$ primers.

To amplify the 5' end cDNA fragment, the antisense primer β-glu178 was designed from the 5' end region of EST clone R2847_1A. The antisense primer β-glu178 and the degenerate sense primer β-glu156 amplified a 317 bp 5' end region cDNA fragment using the first strand cDNA as the template (Figure 3.4). The result from direct sequencing using the β-glu156 primer showed that the sequence of this 5' end fragment started from the cDNA sequence encoding the peptide AFPKRFVFG, which is localized near the start codon (126 bp distance) and overlapped with other EST clones (Figure 3.1). The 317 bp 5' end sequence of this clone was BLASTn searched against dbEST, and the EST c29108, which coressponds to the *bglu1*

sequence and contains the start codon and at least part of the 5'UTR was identified. This EST was selected to design the sense primer β-glu217, the sequence of which is in the 5'UTR. The β-glu217 primer and the β-glu178 antisense primer were used to amplified a 544 bp cDNA fragment by RT-PCR (Figure 3.5). This PCR product was cloned into pBlueScript II SK(+) vector and the sequence of PCR product was determined to be identical to the c29108 sequence.

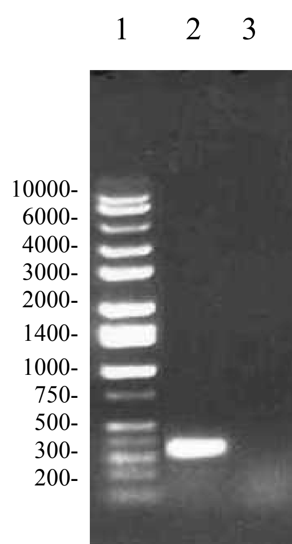


Figure 3.4 1% agarose gel electrophoresis of 5′ end PCR product of *bglu1* cDNA fragment amplified with the β-glu156 and β-glu178 primers. The PCR product was amplified by using Orion rice cDNA as template. Lane 1, Hi- low DNA marker; lane 2, 5′ end cDNA PCR product; lane 3, no DNA template control PCR amplification with β-glu156 and β-glu178 primers.

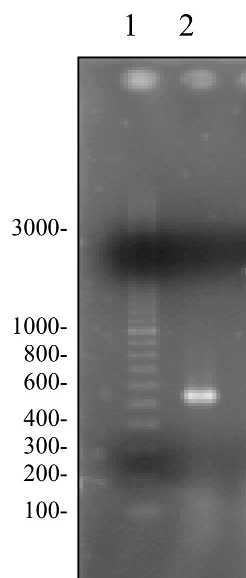


Figure 3.5 1% agarose gel electrophoresis of 5′ end PCR product of *bglu1* cDNA fragment amplified with the β-glu217 and β-glu178 primers. The PCR product was amplified by using Orion rice cDNA as template. Lane 1, 100 bp EZ load DNA marker (Bio-Rad); lane 2, 5′ end cDNA PCR product.

After the sequences at 3' and 5' ends of the β-glucosidase mRNA were known, their sequences were used to design the primers to construct a full-coding-region β-glucosidase cDNA using RT-PCR. The primers, β-glu217 from 5' end sequence and Ro13pr from 3' end sequence were used in PCR amplification. A specific PCR product of 1899 bp was produced (Figure 3.6), and its sequence was determined. The full-length *bglu1* cDNA sequence and the primers used to amplify it are summarized in Figure 3.7. The reconstructed cDNA sequence of *bglu1* consisted of 1909 nucleotides, which included a 1512 nucleotide open reading frame (ORF) encoding a 504 amino acid long precursor protein. The Signal P program (Nielson *et al*., 1997)

predicted the protein to contain a 28 amino acid signal sequence and a 476 amino acid mature protein.

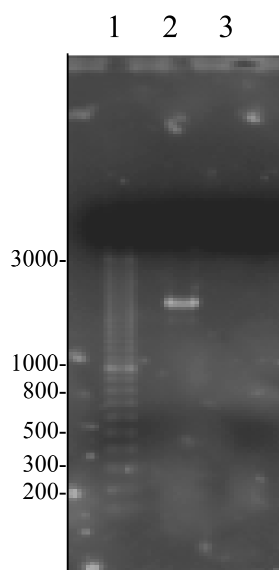


Figure 3.6 1% agarose gel electrophoresis of PCR product of full coding *bglu1* cDNA and its 5′ UTR amplified with the β-glu217 and Ro13pr primers. The PCR product was amplified by using Orion rice cDNA as template. Lane 1, 100 bp EZ load DNA marker (Bio-Rad); lane 2, full-length cDNA PCR product; lane 3, no DNA template control PCR amplification with β-glu217 and Ro13pr primers.

```
β-glu217 →
GTGCGTGCGTGGCTGCGTGTGCTTGCGCTACTGCTTGGTCTCACTGTGTGAACAACTGCTGCCGCCGCCTGCCGGCGGTG-80

TGCTCCGGCCGCGCGCGGTCGTGATGGCTGCACGGCGAGCGAACTGCGCGCTGGTGCTGGTGCTGGCGCTCGCGCTGTTG-160
                         M  A  A  R  R  A  N  C  A  L  V  L  V  L  A  L  A  L  L
                                                                        β-glu156 →
GCGGCGCGGGATGCTGGTGCCGCGGCGGTGCCCAAGCCCAACTGGCTGGGCGGGCTGAGCCGCGCCGCGTTCCCCAAGCG-240
A  A  R  D  A  G  A  A  A  V  P  K  P  N  W  L  G  G  L  S  R  A  A  F  P  K  R

GTTCGTGTTCGGGACGGCCACGTCCGCGTACCAGGTCGAGGGCATGGCGGCGTCCGGCGGCCGCGGGCCGTCCATCTGGG-320
 F  V  F  G  T  A  T  S  A  Y  Q  V  E  G  M  A  A  S  G  G  R  G  P  S  I  W  D

ACGCCTTCGCGCACACCCCGGGAAATGTTGCAGGAAATCAGAATGGAGATGTTGCGACAGATCAATATCATCGCTATAAG-400
  A  F  A  H  T  P  G  N  V  A  G  N  Q  N  G  D  V  A  T  D  Q  Y  H  R  Y  K

GAAGATGTTAATCTCATGAAAAGTTTGAATTTTGATGCCTACCGGTTTTCAATCTCATGGTCCAGGATCTTCCCAGATGG-480
E  D  V  N  L  M  K  S  L  N  F  D  A  Y  R  F  S  I  S  W  S  R  I  F  P  D  G
                                                       ← β-glu178
TGAGGGACGAGTTAACCAAGAAGGCGTAGATTATTACAACAATCTTATAAACTACCTTCTGCAGAAAGGTATCACTCCTT-560
 E  G  R  V  N  Q  E  G  V  D  Y  Y  N  N  L  I  N  Y  L  L  Q  K  G  I  T  P  Y

ATGTCAATCTTTACCACTACGATCTCCCTCTTGCGCTTGAGAAGAAGTACGGAGGCTGGTTGAATGCAAAAATGGCGGAT-640
  V  N  L  Y  H  Y  D  L  P  L  A  L  E  K  K  Y  G  G  W  L  N  A  K  M  A  D
                                                             β-glu163 →
CTATTCACGGAGTACGCGGACTTCTGTTTCAAGACCTTTGGCAATCGTGTCAAGCACTGGTTTACATTTAATGAGCCAAG-720
L  F  T  E  Y  A  D  F  C  F  K  T  F  G  N  R  V  K  H  W  F  T  F  N  E  P  R

GATAGTAGCACTTCTTGGTTATGACCAAGGAACAAATCCTCCTAAAAGGTGCACAAAATGCGCTGCTGGTGGGAACTCAG-800
 I  V  A  L  L  G  Y  D  Q  G  T  N  P  P  K  R  C  T  K  C  A  A  G  G  N  S  A

CAACAGAACCATACATTGTTGCTCATAATTTTCTCTTATCACATGCTGCAGCTGTTGCAAGATACCGGACTAAGTATCAG-880
  T  E  P  Y  I  V  A  H  N  F  L  L  S  H  A  A  A  V  A  R  Y  R  T  K  Y  Q

GCAGCTCAGCAAGGCAAAGTTGGAATAGTTCTGGACTTCAATTGGTATGAAGCTCTTTCCAACTCAACTGAAGACCAAGC-960
A  A  Q  Q  G  K  V  G  I  V  L  D  F  N  W  Y  E  A  L  S  N  S  T  E  D  Q  A

AGCGGCTCAAAGAGCCAGGGACTTCCACATTGGTTGGTATCTTGATCCATTGATAAATGGACATTACTCGCAGATAATGC-1040
 A  A  Q  R  A  R  D  F  H  I  G  W  Y  L  D  P  L  I  N  G  H  Y  S  Q  I  M  Q

AAGATCTTGTGAAAGATAGGCTGCCCAAGTTCACTCCTGAACAAGCTAGATTGGTCAAGGGCTCGGCTGACTATATTGGC-1120
  D  L  V  K  D  R  L  P  K  F  T  P  E  Q  A  R  L  V  K  G  S  A  D  Y  I  G

ATCAACCAATACACAGCTAGCTACATGAAGGGCCAGCAATTGATGCAGCAGACACCGACGAGTTACTCAGCCGATTGGCA-1200
I  N  Q  Y  T  A  S  Y  M  K  G  Q  Q  L  M  Q  Q  T  P  T  S  Y  S  A  D  W  Q

GGTTACCTATGTTTTTGCGAAAAACGGCAAACCAATTGGACCACAGGCAAATTCTAATTGGCTTTACATCGTGCCATGGG-1280
 V  T  Y  V  F  A  K  N  G  K  P  I  G  P  Q  A  N  S  N  W  L  Y  I  V  P  W  G

GGATGTACGGATGTGTGAACTACATCAAGCAGAAGTACGGGAATCCGACAGTCGTCATAACTGAAAACGGAATGGATCAA-1360
  M  Y  G  C  V  N  Y  I  K  Q  K  Y  G  N  P  T  V  V  I  T  E  N  G  M  D  Q

CCTGCGAACCTCAGCCGCGACCAGTACTTGCGCGACACCACGAGGGTTCATTTCTACAGGAGCTACCTCACCCAGCTGAA-1440
P  A  N  L  S  R  D  Q  Y  L  R  D  T  T  R  V  H  F  Y  R  S  Y  L  T  Q  L  K

GAAGGCCATCGACGAGGGCGCCAATGTGGCTGGCTACTTCGCCTGGTCTCTCCTGGACAACTTCGAGTGGCTGTCAGGTT-1520
 K  A  I  D  E  G  A  N  V  A  G  Y  F  A  W  S  L  L  D  N  F  E  W  L  S  G  Y

ACACGTCCAAGTTTGGTATCGTCTACGTCGACTTCAACACGCTCGAGCGCCACCCCAAGGCGTCGGCCTACTGGTTCAGG-1600
  T  S  K  F  G  I  V  Y  V  D  F  N  T  L  E  R  H  P  K  A  S  A  Y  W  F  R
             Ro13pf →
GACATGCTAAAGCACTGAGAATGAAAGACACGCAGACTGAAGTTGAAGTTGCCATCCATTGATCATCATGTTCAGTTCTA-1680
D  M  L  K  H  *

GGCTGCCGCTTGGATGGCATCAGAGCTGGCTATGAACTATGTGATGGAAATAAAAGGACTGCAAATTGATTCGATTATGC-1760

AATTCGTGTCTATGATCAGAATTGTGGCCGTCTAGTGAACTTGTAGAAGTGTACTCTTGATGTGTGATCGATCTTGAGTA-1840
                                                      ← Ro13pr
ATCAATGTTTGATCAGATTTCGCCTTATAATTTATTAGTTGATGAGATTTGGCTTGGTGATTTTATAGG-1909
```

Figure 3.7 The full-length cDNA sequence and deduced amino acid sequence of rice *bglu1*. Underlined letters represent the DNA sequence of PCR primers for cloning and sequencing.

The second rice β-glucosidase isozyme cDNA (*bglu2*) was cloned by RT-PCR with the strategy shown in Figure 3.8. The first strand cDNA was synthesized using mRNA extracted from 3-4 day old Orion rice seedling as template in reverse transcription with an oligo $(dT)_{17}$ primer. Amplification of *bglu2* was done with degenerate primers, β-glu156 (sense) and β-glu41 (antisense), based on conserved peptide sequences (GFPAGFVFG and WFTFNEP), using the first strand cDNA as the template. The expected size (496 nucleotides) cDNA fragment product was obtained and its sequence determined (Figure 3.9).

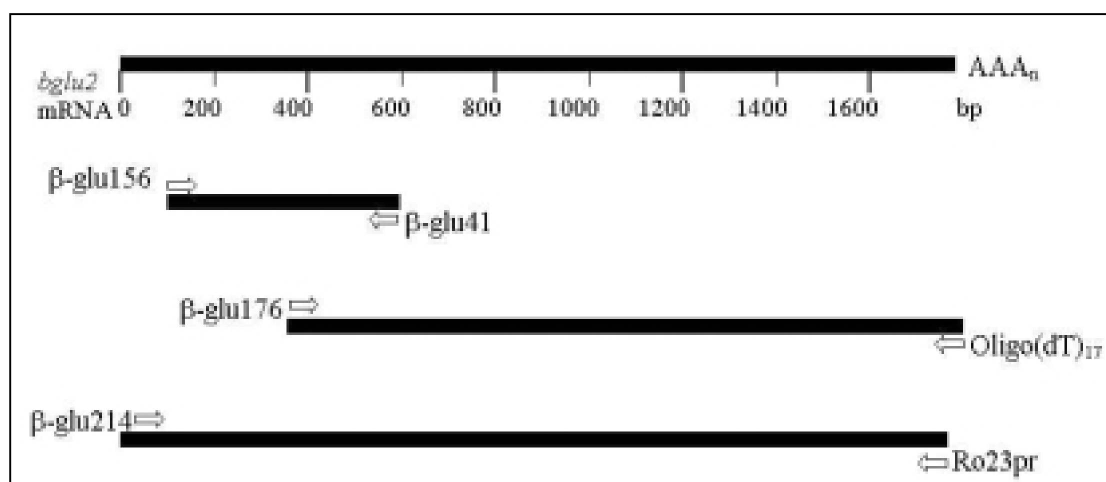


Figure 3.8 Cloning strategies of rice *bglu2* from Orion rice. First, the cDNA fragment at the 5′ region of *bglu2* was amplified using the degenerate primers β-glu156 and β-glu41. Second, the cDNA fragment encoding the 3′ end of *bglu2* was generated by 3′ RACE with primers β-glu176 and oligo $(dT)_{17}$. Third, the full-length cDNA was amplified using the specific primers β-glu214 and Ro23pr.

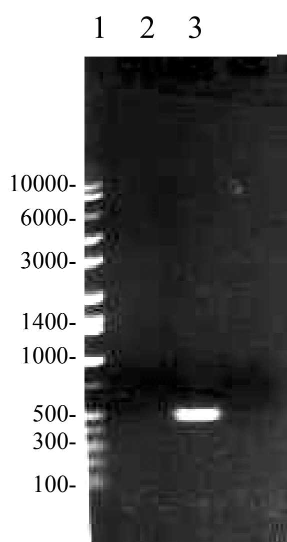


Figure 3.9 1% agarose gel electrophoresis of 5′ end PCR product of the *bglu2* cDNA fragment amplified with the primers β-glu156 and β-glu41. The PCR product was amplified by using Orion rice cDNA as template. Lane 1, Hi-low DNA marker; lane 2, no DNA template control PCR amplification with β-glu156 and β-glu41 primers; lane 3, the 5′ end PCR product using primers β-glu156 and β-glu41.

To amplify the 3' end cDNA fragment, the primer β-glu176, derived from the 496 bp cDNA fragment was used with oligo $(dT)_{17}$ in 3' RACE, and a PCR product of approximately 1400 bp was obtained (Figure 3.10). This PCR product was cloned into pBlueScript II SK (+) vector and the plasmids containing the expected insert size were sequenced. The sequence of the 1442-bp PCR fragment included the stop codon and 268 bp 3' UTR of the mRNA, in addition to 1500 bp of *bglu2* CDS. The sequence overlapped that of the AU065204 EST (clone number E60611_1A), which covered the 5' end of the CDS. This EST was used to design the sense primer β-glu214, the sequence of which included the predicted start codon. To obtain the full

CDS of *bglu*2, primers derived from both end sequences, β-glu214 from 5' end sequence and Ro23pr from 3' UTR, were used in PCR amplification. An approximately 1800 bp PCR product band was obtained (Figure 3.10) and its sequence was determined. The full-length *bglu2* cDNA sequence and the primers used in PCR amplification steps are summarized in Figure 3.11. The reconstructed cDNA sequence, consisted of 1797 nucleotides, including a 1500-nucleotide ORF encoding a 500 amino acid long precursor protein. The protein consisted of a 25 amino acid long prepeptide and a 475 amino acid long mature protein, as predicted by the Signal P program. The sequence was deposited in GenBank, AC AY056828.

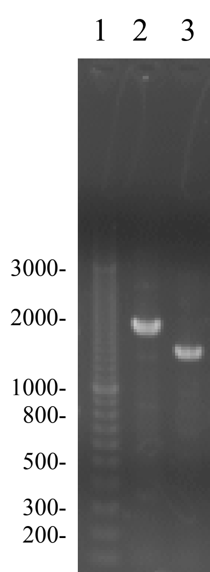


Figure 3.10 1% agarose gel electrophoresis of the *bglu2* 3′RACE product (amplified with the β-glu176 and oligo $(dT)_{17}$ primers) and full-length *bglu2* cDNA (amplified with the β-glu214 and Ro23pr primers). The PCR products were amplified by using Orion rice cDNA as template. Lane 1, 100 bp EZ load DNA marker (Bio-Rad); lane 2, PCR product of full-length cDNA; lane 3, PCR product of 3′RACE.

```
β-glu214 ──►
TCTCTCGCTGGAATGGGGATCAGAATGGGAAGGCGACTCCTCTTCACCTTGTTCTTGGGAGCTCTGTTCTGCAATGGCGT-80
            M  G  I  R  M  G  R  R  L  L  F  T  L  F  L  G  A  L  F  C  N  G  V
                      β-glu156 ──►
TTACGCCAAGTTTACCCGGTACAGCTTCCCCAAGGACTTCATCTTCGGCACAGGTTCAGCAGCTTATCAGTATGAGGGCG-160
 Y  A  K  F  T  R  Y  S  F  P  K  D  F  I  F  G  T  G  S  A  A  Y  Q  Y  E  G  A

CCTACAAAGAAGGGGGCAAAGGTCCTAGCGTCTGGGACAACTTCACTCACATTCCAGGTAAAATTTTAAACAATGATAAC-240
  Y  K  E  G  G  K  G  P  S  V  W  D  N  F  T  H  I  P  G  K  I  L  N  N  D  N

GGCGATGTGGCAAATGACTTCTATCACCGATACAAGGAGGATGTGAGCCTCCTGAAGGACATGAACATGGATGCTTTCCG-320
G  D  V  A  N  D  F  Y  H  R  Y  K  E  D  V  S  L  L  K  D  M  N  M  D  A  F  R
                                     β-glu176 ──►
GTTCTCCATTGCGTGGACCAGGATCCTGCCAAATGGATCCTTGAGTGGAGGAATAAACAAAGAAGGGGTTGCTTTCTACA-400
 F  S  I  A  W  T  R  I  L  P  N  G  S  L  S  G  G  I  N  K  E  G  V  A  F  Y  N

ACAGCTTGATCAATGATGTCATAGCAAAAGGGATGATCCCATTTGTCACTATCTTCCACTGGGACACCCCCTTGGCTCTG-480
  S  L  I  N  D  V  I  A  K  G  M  I  P  F  V  T  I  F  H  W  D  T  P  L  A  L

GAAAGCAAATACGGAGGATTCCTCAGTGAAGACATAGTGAAAGAATACGTGGACTTCGCGGAGGTGTGCTTTCGCGAGTT-560
E  S  K  Y  G  G  F  L  S  E  D  I  V  K  E  Y  V  D  F  A  E  V  C  F  R  E  F
                                ◄── β-glu41
CGGCGACCGTGTCAAGTACTGGTTTACATTTAATGAGCCATTCACATATAGCGCCTACGGCTACGGCAAGGGCGTGTTTG-640
 G  D  R  V  K  Y  W  F  T  F  N  E  P  F  T  Y  S  A  Y  G  Y  G  K  G  V  F  A

CGCCGGGACGATGCTCTTCATATGTTTCCAAGTCATGCGGCGTCGGTGACTCCAGTCGCGAGCCCTACCTCGTGGCACAC-720
  P  G  R  C  S  S  Y  V  S  K  S  C  G  V  G  D  S  S  R  E  P  Y  L  V  A  H

CACATCCACCTCTCCCACGCTGCCGCTGTCCAGCTTTACCGCACCAAGTACCAGCCAACACAGAAGGGACAGATCGGCAT-800
H  I  H  L  S  H  A  A  A  V  Q  L  Y  R  T  K  Y  Q  P  T  Q  K  G  Q  I  G  M

GGTGGTGGTCACCCACTGGTTCGTGCCGTACGACAACTCCGACGCTGACCGTGGCGCTGTGCAACGGAGCCTAGACTTCA-880
 V  V  V  T  H  W  F  V  P  Y  D  N  S  D  A  D  R  G  A  V  Q  R  S  L  D  F  I

TCTATGGGTGGTTCATGGACCCTATCGTGCATGGTGACTACCCAGGCACCATGAGAGGTTGGCTCGGCAATCGGCTGCCA-960
  Y  G  W  F  M  D  P  I  V  H  G  D  Y  P  G  T  M  R  G  W  L  G  N  R  L  P

GAGTTCACGCCTGAACAGTCGGCGATGGTGAAGGGCTCCTACGACTTCATCGGAGTTAATTACTACACCACCTACTACGC-1040
E  F  T  P  E  Q  S  A  M  V  K  G  S  Y  D  F  I  G  V  N  Y  Y  T  T  Y  Y  A

TAAGAGTATACCGCCGCCTAACTCCAACGAGCTATCCTACGACCTCGACAACCGCGCCAACACCACCGGCTTCCGTAATG-1120
 K  S  I  P  P  P  N  S  N  E  L  S  Y  D  L  D  N  R  A  N  T  T  G  F  R  N  G

GCAAACCCATCGGTCCACAGGAATTTACACCGATCTTCTTCAACTACCCTCCAGGTCTCCGTGAGCTCCTCCTCTACACC-1200
  K  P  I  G  P  Q  E  F  T  P  I  F  F  N  Y  P  P  G  L  R  E  L  L  L  Y  T

AAGAGGAGATACAACAACCCGACCATCTATGTTACAGAAAACGGCATCGATGAGGGTAACAACAGCACACTGCCAGAGGC-1280
K  R  R  Y  N  N  P  T  I  Y  V  T  E  N  G  I  D  E  G  N  N  S  T  L  P  E  A
              RO1r ──►
GCTCAAGGATGGACACAGGATCGAGTTCCACTCAAAGCACCTGCAGTTCGTCAACCACGCCATCAAGAATGGGGTGAACG-1360
 L  K  D  G  H  R  I  E  F  H  S  K  H  L  Q  F  V  N  H  A  I  K  N  G  V  N  V

TGAAGGGCTACTTCACGTGGACATTCATGGACTGCTTTGAGTGGGGTGACGGCTACCTTGACAGGTTCGGCCTCATCTAC-1440
  K  G  Y  F  T  W  T  F  M  D  C  F  E  W  G  D  G  Y  L  D  R  F  G  L  I  Y
                                                                  Ro23pf ──►
GTCGACCGCAAGACGCTCAAGCGCTACCGCAAGGAGTCCAGCTACTGGATCGAAGACTTCCTCAAGAGGCACTGATCCAG-1520
V  D  R  K  T  L  K  R  Y  R  K  E  S  S  Y  W  I  E  D  F  L  K  R  H  *
                                                ◄── β-glu219
CCCCACAAAATTTCGGATTGATCACGAGATCGATCGGCTGGCCACCAGATTTGTGTTTCGTCAGTCTAGCTTTCAGCTGG-1600

TTAATTTGTGATTTGTGCATCAGTGGTGATGACCGATATGATGGCGTGATGAATTATGCTGCATTACTATTGAATGTGTA-1680
                                                                          ◄──
CCTCAGTACCTGTGCATCATTGCTGGCTGATCCCATTGAGCCGGACAATAATAAGAAGAAACAATAATGCTCCCTGTCGA-1760
  Ro23pr   ◄──β-glu218
GCAGTGTTGGGCTGGTTCAGGCT-1783
```

Figure 3.11 The full-length cDNA sequence and deduced amino acid sequence of rice *bglu2*. Underlined letters represent the DNA sequence of PCR primers for cloning and sequencing.

The BGlu1 polypeptide contains Glu residues at positions 204 and 414 that lie within the sequences TFNEP and ITENG, respectively, while BGlu2 contains Glu residues at positions 195 and 409 that lie within the sequence TFNEP and VTENG, which closely resemble the consensus motifs in family 1 β-glycosidases. The BGlu1 protein sequence had the highest sequence identity (66%) with barley β-glucosidase BGQ60 (AC AAA87339_1), while BGlu2 protein had highest sequence identity (54%) with amygdalin hydrolase isoform AH I of *Prunus serotina* (AC AAA93234.1) and other cyanogenic β-glucosidases (Figure 3.12).

Both BGlu1 and BGlu2, unlike the previously characterized monocot β-glucosidases of maize, sorghum and oat, lacked a typical plastid targeting sequence, and the SignalP and PSORT programs predicted that they should be targeted instead to the secretory pathway, like the barley β-glucosidase BGQ60. The BGlu1 and BGlu2 protein sequences contained 3 and 6 possible N-linked glycosylation sites, respectively (Table 3.1).

Table 3.1 Predicted properties of rice β-glucosidases

Clone	Mr (kDa)	pI	N-Link	Target
bglu1	56.9	9.06	3	out/ microbody/ ER membrane/ ER lumen
bglu2	57.4	8.05	6	out/ vacuole/ microbody/ ER membrane

```
BGlu1        1  MAARRANCALVLVLALALLAARDAGAAA------------------------------VP
BGlu2        1  ME-IRMGRRLLFTLFILCALFCNGVYA--------------------------------
barley       1  M---RSSPVLLLVIALVAAAHLAPLEC----------------------------DGP
amygdalin    1  MA-TKLSSLLLCALLLAGFALTNSKA--------------------------------A
Zea mays     1  MAPLLAAAMNHAAAHPGLRSHLVGPNNESFSRHHLPSSSPQSSKRRCNLSFTTRSARVGS

BGlu1       31  KPNWL----GGL-SRAAFPK---RFVFGTATSAYQVEGMAASGGRGPSIWDAFAHTPGN-
BGlu2       26  ----------KF-TRYSFPK---DFIFGTGSAAYQYEGAYKEGGKGPSVWDNFTHIPGK-
barley      28  NPNPEIGNTGGL-SRQGFPA---GFVFGTAASAYQVEGMARQGGRGPCIWDAFVAIQGM-
amygdalin   27  KTDPPI-HCASL-NRSSFDALEPGFIFGTASAAYQFEGAAKEDGRGPSIWDTYTHNHSER
Zea mays    61  QNGVQMLSPSEIPQRDWFPS---DFTFGAATSAYQIEGAWNEDGKGESNWDHFCHNHPER

BGlu1       82  VAGNQNGDVATDQYHRYKEDVNLMKSLNFDAYRFSISWSRIFPDG---EGRVNQEGVAYYN
BGlu2       71  ILNNDNGDVANDFYHRYKEDVSLLKDMNMDAFRFSIAWTRILPNGSLSGGINKEGVAFYN
barley      83  IAGNGTADVTVDQYHRYKEDVGIMKNMGFDAYRFSISWSRIFPDG---TGKVNQEGVDYYN
amygdalin   85  IKDGSNGDVAVDQYHRYKEDVRIMKRMGFDAYRFSISWSRVLPNGKISGGVNEDGIKFYN
Zea mays   118  ILDGSNSDIGANSYHMYKTDVRLLKEMGMDAYRFSISWPRILPKGTKEGGINPDGIKYYR

BGlu1      140  NLINYLLDKGITPYVNLYHYDLPLALEKKYGGWLNA---KMADLFTEYADFCFKTFGNRV
BGlu2      131  SLINDVIAKGMIPFVTIFHWDTPLALESKYGGFLSE---DIVKEYVDFARVCFREFGDRV
barley     141  KLIDYMLQQGITPYANLYHYDLPLALHKQYLGWLSP---KIVGAFADYAEFCFKVFGDRV
amygdalin  145  NLINEILRNGLKPFVTIFHWDLPQALEDEYGGFLSP---NIVDHFRDYANLCFKKFGDRV
Zea mays   178  NLINLLLENGIEPYVTIFHWDVPQALEEKYGGFLDKSHKSIVEDYTYFAKVCFDNFGDKV

BGlu1      197  KNWFTFNEPRIVALLGYDQGTNPPKRCTKCA----AGGNSATEPYIVAHNLLLSHAAAVA
BGlu2      188  KYWFTFNEPFIYSAYGYGKGVFAPGRCSSYVSKSCGVGDSSREPYLVAHNIHLSHAAAVQ
barley     198  KNWFTFNEPRVVAALGYDNGFHAPGRCSKCP----AGGDSRTEPYIVTHNIILSHAAAVQ
amygdalin  202  KHWITLNEPYTFSSSGYAYGVHAPGRCSAWQKLNCTGGNSATEPYLVTHHQLLAHAAAVK
Zea mays   238  KNWLTFNEPQTFTSFSYGTGVFAPGRCSPGLDCAYPTGNSLVEPYTAGHNILLAHAEAVD

BGlu1      253  KYRTKYQAAQQGKVGIVLDFNWYEALSNSTEDQAAAQRARDFHIGWYLDPLINGHYSQIM
BGlu2      248  LYRTKYQPTQKGQIGKVVVTHWFVPYDNSDADRGAVQRSLDFIYGWFMDPIVHGDYPGEM
barley     254  RYREKYQPHQKGRIGILLDFVWYEPHSDTDADQAAAQRARDFHIGWFLDPITNGRYPSEM
amygdalin  262  LYKDEYQASQNGLIGITLVSPWFEPASEAEEDINAAFRSLDFIFGWFMDPLTNGNYPHIM
Zea mays   298  LYN-KHYKRDDTRIGLAFDVMGRVPYGTSFLDKQAEERSWDINLGWFLEPVVRGDYPFSM

BGlu1      313  QDLVKDRLPKFTPEQARLVKGSADYIGINQYTASYMKGQQL-MQQTPTSYSADWQVTYVF
BGlu2      308  RGWLGNRLPEFTPEQSAMVKGSYDFIGVNYYTTYYAKSIPP-PNSNELSYDLDNRANTTG
barley     314  LKIVGNRLPGFSADESRMVKGSIDYVGINQYTSYYMKDPGA-WNQTPVSYQDDWHVGFVY
amygdalin  322  RSIVGERLPNFTEEQSKLLKGSFDFIGLNYYTTRYASNAPK-ITSVHASYITDPQVNATA
Zea mays   357  RSLARERLPFFKDEQKEKLAGSYNMLGLNYYTSRFSKNIDISPNYSPVLNTDDAYASQEV

BGlu1      372  A-KNGKPIGPQANSNWLYIVPWGMYGCVNYIKQKYGNPTVVITENGMDQPAN----LSRD
BGlu2      367  F-RNGKPIGPQEFTPIFFNYPPGLRELLLYTKRRYNNPTIYVTENGIDEGNN----STLP
barley     373  E-RNGVPIGPRANSDWLYIVPWGMNKAVTYVKERYGNPTMILSENGMDQPGN----VSIA
amygdalin  381  E-LKGVPIGPMAASGWLYVYPKGIHDLVLYTKEKYNDPLIYITENGVDEFNDPK--LSME
Zea mays   417  NGPDGKPIGPPMGNPWIYMYPEGLKDLLMIMKNKYGNPPIYITENGIGDVDTKETPLPME

BGlu1      427  QYLRDTTRVHFYRSYLTQLKKAIDEGANVAGYFAWSLLDNFEWLSGYTSKFGIVYVDFN-
BGlu2      422  EALKDGHRIEFHSKHLQFVNHAIKNGVNVKGYFTWTFMDCFEWGDGYLDRFGLIYVDRK-
barley     428  DGVHDTVRIRYYRDYITELKKAIDNGARVAGYFAWSLLDNFEWRLGYTARFGIVYVDFN-
amygdalin  438  EALKDTNRIDFYYRHLCYLQAAIKKGSKVKGYFAWSFLDNFEWDAGYTVRFGINYVDYND
Zea mays   477  AALNDYKRLDYIQRHIATLKESIDLGSNVQGYFAWSLLDNFEWFAGFTERYGIVYVDRNN

BGlu1      486  TLERHPKASAYWFRDMLKH-----------------------------------
BGlu2      481  TLKRYRKESSYWIEDFLKRH----------------------------------
barley     487  TLKRYPKDSALWFKNMLSEKKRS-------------------------------
amygdalin  498  NLKRHSKLSTYWFTSFLKKYERSTKEIQMFVESKLEHQKFESQMMNKVQSSLAVVV
Zea mays   537  NCTRYMKESAKWLKEFNT-AKKPSKKILTPA-----------------------
```

Figure 3.12 Alignment of predicted protein sequences of full-length rice β-glucosidases. Rice cDNA derived sequences are labeled as BGlu1 (AC U28047) and BGlu2 (AC AY056828); zea mays is maize β-glucosidase 1 (AC U33816); barley, barley BGQ60 (AC L41869); amygdalin, *Prunus serotina* (AC AAA93234.1). Sequences corresponding to the catalytic acid/base and nucleophile consensus sequences are marked by a thick line over the top. Residues shown by Czjzek *et al.* (2000) to contact the DIMBOA aglycone in the maize β-glucosidase are indicated by: ▼, while those conserved residues making contacts with the sugar are marked by: ∆ above the column. Other residues shown by Keresztessy *et al.*, (2001) by site directed-mutagenesis or predicted by Ketudat Cairns *et al.* (2000) based on homology modeling to play a role in substrate binding and specificity of linamarase and Thai rosewood dalcochinase are marked by: ○. Shading represents the conservation of residues within the alignment (black for the most common residue in a column and grey for similar residues). The alignment was generated using the Clustal X implementation of Clustal W (Jeanmougin *et al.*, 1998; Thompson *et al.*, 1994); analyzed and manually adjusted by Genedoc (Nicholas and Nicholas, 1997) and shaded with the Boxshade 3.21 program of K. Hofman and M. Baron.

A BLAST search of the dbEST database showed that *bglu1* was the most highly represented glycosyl hydrolase family 1 gene transcript, with 60 matching ESTs. The source libraries for the *bglu1* ESTs included shoot, root, callus, panicle, leaf and stem, reflecting its expression in many tissues. *Bglu2* matched only 8 ESTs, all of which came from young growing plant parts: shoot, stem and young leaf (Table 3.2). As ESTs for both the *bglu1* and the *bglu2* genes had been mapped to chromosome location by the Rice Genome Research Program, their map location could be identified. The ESTs R2847, D24959 and AU031958, matching *bglu1*, have been mapped to rice chromosome 3 at position 122.8 cM (distance to the locus from the short arm end). Similarly, the ESTs AU033015 and D39581, matching *bglu2*, have been mapped to chromosome 9 at position 72.1 cM.

Table 3.2 Expression of rice β-glucosidase genes judged from corresponding ESTs[a].

Clone	EST match	Tissue library	GenBank accession no.
bglu1	3	Seedling (cold stress)	CB000637, CA999001, CA999930
	16	Seedling shoot	AU097444, AU33138,AU222984, AU162859, AU033196, AU066086, AU181125, AU097203 , AU097647, AU162860, AU096432, C25358, D46605, D39740, D46619, D39779,
	14	Seedling root	AU095510, CA753525, BE607353, D24959, AU174381, AU031958, AU165088, D24663, AU032064, AU070800, AU071104, AU176521, AU184447, AU165093
	11	Callus	AU092238, AU09227, AU069490, AU161415, AU092735, C26022, C97305, C27840, C29108, D42356, D43502

Table 3.2 (Continued)

Clone	EST match	Tissue library	GenBank accession no.
bglu1	2	35-day old leaf	BU673604, BI305771
	1	mature leaf (with *M.grisea*)	BI808817
	1	3-4 weeks old root (salt)	CA756464
	1	stem (3-4 leaf stage)	BI805640
	7	panicle less than 3 cm long	AU183651, U183650, AU183652, AU029514, AU029546, AU029665, AU064419
	4	panicle at flowering stage	AU093932, C72237, AU93914, C73533
Total	60		
bglu2	3	Shoot	AU033015, D39581, D39589
	4	Immature leaf	AU065170, AU065204, AU065193, AU058282
	1	Stem (3-4 leaf stage)	BI806911
Total	8		

[a]EST sequences were retrieved from the dbEST section of GenBank (http://www.ncbi.nlm.nih.gov/) by BLASTn search with full cDNA sequences. They were inspected to ensure they matched the gene-coding region and their full files retrieved to determine the cDNA library source tissue and clone number when necessary. The ESTs assigned to each gene had greater than 96% identity and had no higher similarity with another gene.

3.2 Nucleic acid-blot analysis

Specific probes were generated by amplification of portions from the 3' UTR of *bglu1* and *bglu2* cDNA (284 and 267 bp, respectively). To confirm hybridization to a single gene of each of these isozyme probes, Southern blots of rice genomic DNA were probed with each probe. For *bglu1* probe, only one strong band was seen for each digest with the sizes of approximately 9.4, 5.0, 4.7, 5.2, 6.5 and 7.0 kb in *Bam*H I, *Bgl* II, *Eco*R I, *Hin*d III, *Xba* I and *Xho* I digested genomic DNA lanes, respectively. For *bglu2* probe, only one strong band for each digest was detected with the sizes of approximately 7.0, 8.7, 3.8, 6.8, 9.3 and 9.1 kb in *Bam*H I, *Bgl* II, *Eco*R I, *Hin*d III, *Xba* I and *Xho* I digested genomic DNA lanes, respectively (Figure 3.13).

These results were consistent with a single gene copy gene corresponding to each cDNA, as specific hybridization with a single intense band was seen in most lanes. The patterns of hybridization were different, indicating that *bglu1* and *bglu2* probes do not hybridize to the same genes.

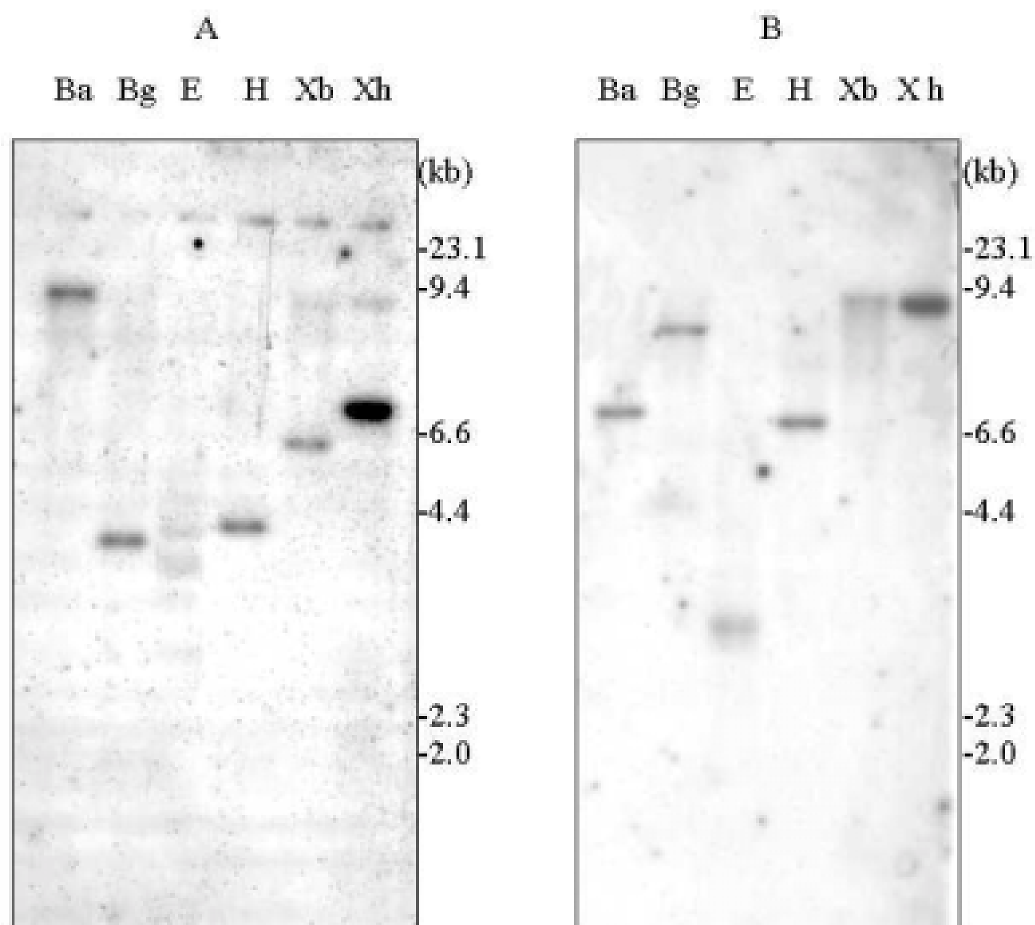


Figure 3.13 Southern-blot analysis of rice *bglu1* and *bglu2* gene copy number and probe specificity. Genomic DNA was digested with *Bam*H I (Ba), *Bgl* II (Bg), *Eco*R I (E), *Hin*d III (H), *Xba* I (Xb), and *Xho* I (Xh). A, the blot was probed with an α^{32}P-labeled 3' UTR 284 bp cDNA fragment of *bglu1*. B, the blot in A was stripped and reprobed with an α^{32}P-labeled 3'UTR 267 bp 3'UTR cDNA fragment of *bglu2*.

3.3 Temporal and spatial expression and the environmental effects on β-glucosidase gene expression

3.3.1 Spatial expression pattern in seedling and flowering stage

RNA gel-blot analysis was used to determine transcript levels in tissues of 7-day-old seedlings and 6-week-old mature plants, relative to an 18S rRNA control (Figure 3.14). In seedlings, both *bglu1* and *bglu2* transcripts were detected in high abundance in the shoot, and at low levels in the root and endosperm. In 6-week-old mature plants, the *bglu1* transcript was detected at highest levels in flower, with high level in node, moderate levels in leaf sheath, low level in inter node, and no or very low signal in root and leaf blade. In contrast to *bglu1*, *bglu2* was only detected at significant levels in node, while no or low signal was seen in other plant parts.

3.3.2 Temporal expression in seedlings during germination stage

The temporal expression pattern of *bglu1* and *bglu2* in seedling tissues during germination was examined using total RNA extracted from whole seedlings, relative to 18S rRNA control. As shown in Figure 3.15, the transcripts of *bglu1* and *bglu2* were not detectable in seeds before germination (germination at day 0). Both transcripts were detected on day 2 with a moderate signal for *bglu1* and a low signal for *bglu2*. Both signals increased on day 4 and *bglu1* maintained a similar high level of expression from day 4 to day 10, with decreasing signal at 12 and 14 days. The *bglu2* expression pattern was similar with the highest levels seen on day 8, then levels decreased on days 12 and 14.

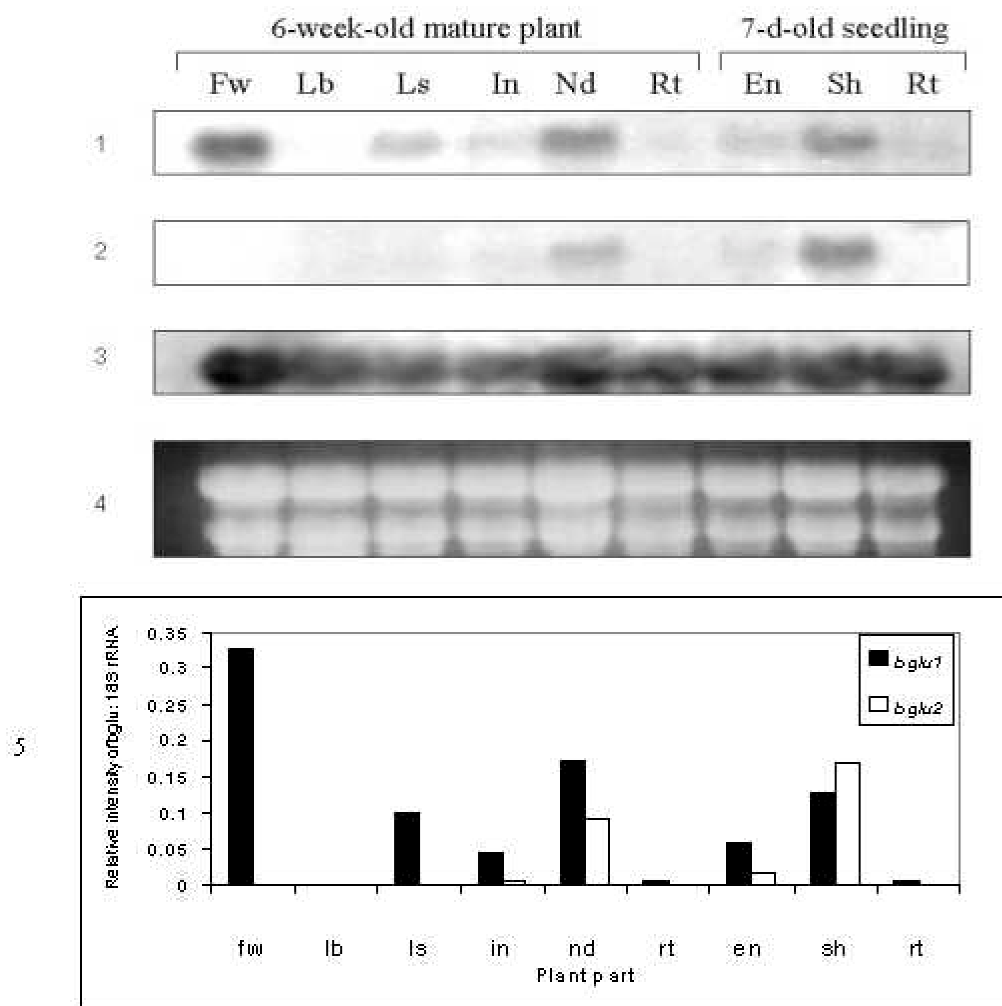


Figure 3.14. Northern-blot analysis of *bglu1* and *bglu2* transcript levels in various tissues from 7-day-old rice seedlings and 6-week-old mature plants. 1, RNA blot was hybridized with α^{32}P-labeled 3' UTR cDNA of *bglu1*. 2 and 3, Same blot probed with an α^{32}P-labeled 3' UTR cDNA of *bglu2* and an α^{32}P-labeled 18S rRNA cDNA, respectively. 4, The ethidium bromide-stained gel showed twenty micrograms of total RNA from the appropriate tissues loaded in each lane. 5, The graph shows quantification of *bglu1* and *bglu2* expression as a ratio relative to18S rRNA. Fw, flower; Lb, leaf blade; Ls, leaf sheath; In, inter node; Nd, node; Rt, root; En, endosperm; Sh, shoot.

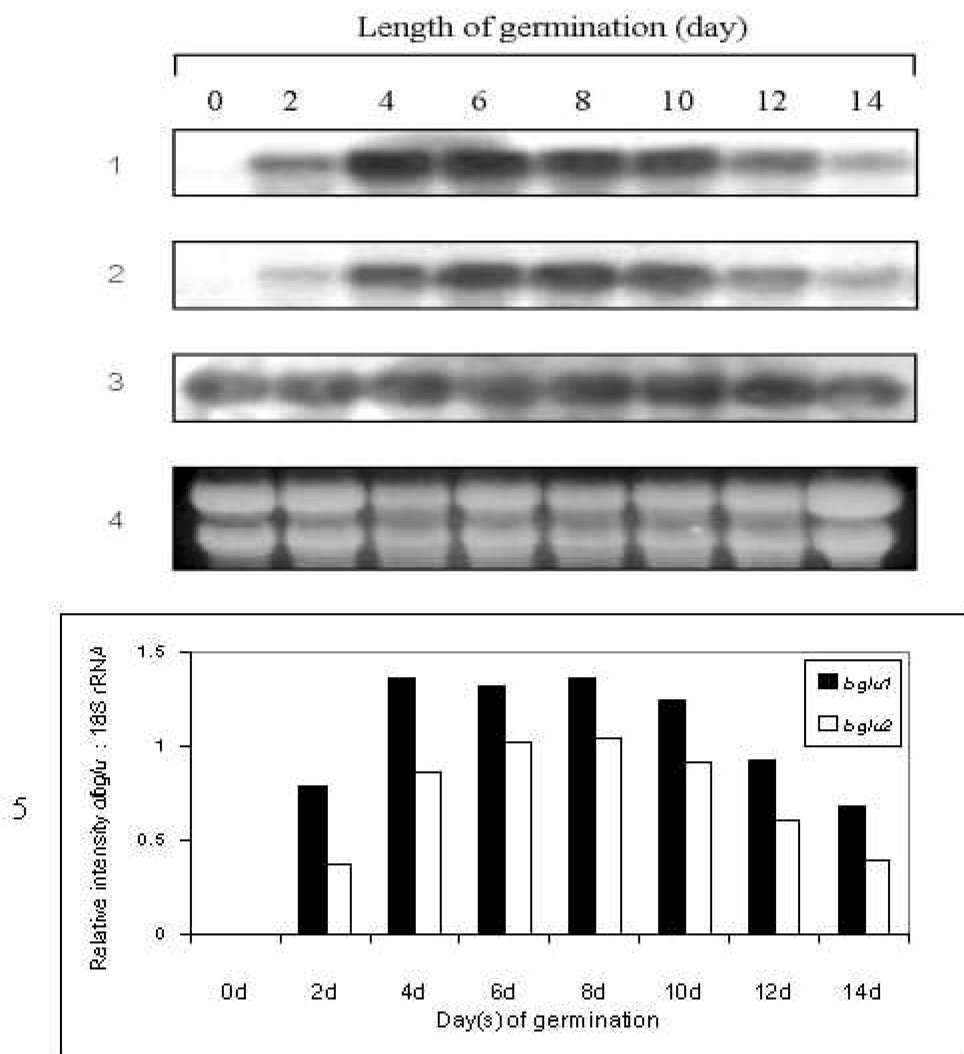


Figure 3.15. Northern-blot analysis of *bglu1* and *bglu2* transcript levels in rice seedlings at days 0 through 14 of germination. 1, RNA blot was hybridized with α ^{32}P-labeled 284 bp 3'UTR fragment of *bglu1*. 2, Same blot probed with an α^{32}P-labeled 267 bp 3'UTR cDNA fragment of *bglu2*. 3, Same blot probed with an α^{32}P-labeled 18S rRNA cDNA fragment. 4, The ethidium bromide-stained gel showed twenty micrograms of total RNA from the appropriate tissues loaded in each lane. 5, The graph shows quantification of *bglu1* and *bglu2* expression as a ratio relative to18S rRNA.

3.3.3 Effects of environmental conditions on rice β-glucosidase gene expression

To determine the effects of environmental conditions on rice β-glucosidase gene expression during germination, RNA gel-blot analysis was also used to determine transcript levels of *bglu1* and *bglu2* in seedlings relative to 18S rRNA. Various environmental conditions were applied after rice seedlings had grown. The transcript levels of *bglu1* and *bglu2* in rice seedlings were determined after they were treated with various environmental conditions for 2 days in comparison to rice seedlings grown at 28°C (control) (Figure 3.16). Comparison between signal intensity (relative to 18S rRNA) for each sample was presented as percent intensity relative of the control (Table 3.3).

For *bglu1*, the transcript was detected at high levels similar to the control when the seedlings were treated with cold temperature at 5°C. The transcript levels decreased approximately 15% when they were treated with osmotic stress (0.5 M manitol). The transcript levels decreased about 40% when the seedlings were treated with 10 mg/ml Ethephon and 10^{-4} M abscissic acid and approximately 60% under salt stress and flooding conditions. Transcript levels decreased dramatically after the seedlings were treated with drought stress. For *bglu*2, the transcript was aprroximately 30% higher than the control when the seedlings were treated with 0.5 M manitol. The transcript levels decreased slightly when the seedlings were treated with 10^{-4} M abscissic acid. In the presence of salt and 10 mg/ml Ethephon, the transcript levels decreased approximately 30%. The transcript level decreased up to 60% when seedlings were grown at low temperature (5°C) and decreased dramatically nearly 100%) under flooding and drought conditions.

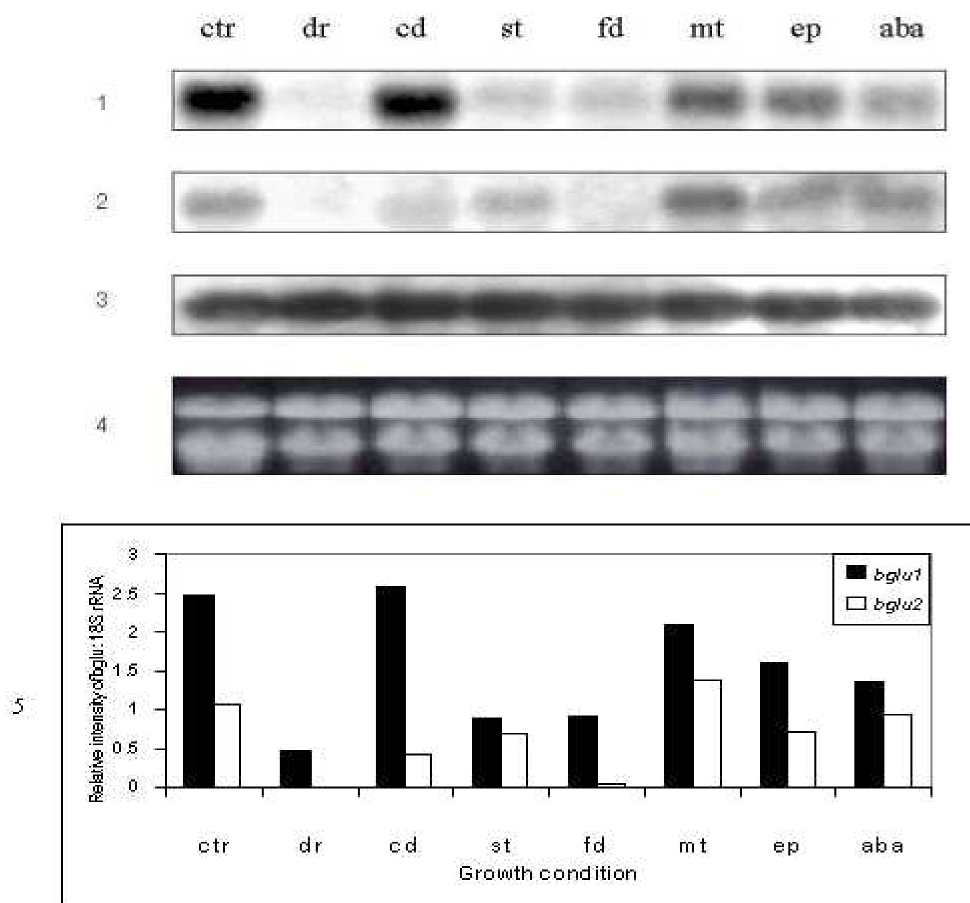


Figure 3.16 RNA gel-blot analysis of *bglu1* and *bglu2* transcript levels in rice seedlings after growing under various environmental conditions. 1, RNA blot was hybridized with α^{32}P-labeled 284 bp 3' UTR fragment of *bglu1*. 2 and 3, the same blot was probed with an α^{32}P-labeled 267 bp 3' UTR cDNA fragment of *bglu2* and an α^{32}P-labeled 18S rRNA cDNA fragment, respectively. 4, The ethidium bromide-stained gel showed twenty micrograms of total RNA from the appropriate tissues loaded in each lane. The graph below showed quantification of *bglu1* and *bglu2* expression as a ratio relative to18S rRNA. ctr, control at 28°C; dr, drought; cd, cold stress 5°C; st, salt stress with 0.3 M NaCl; fd, flooding; mt, 0.5 M manitol; ep, 10 mg/ml Etherphon; and aba, 10^{-4} M abscissic acid.

Table 3.3 Comparison of transcript levels of *bglu1* and *bglu2* in rice seedlings treated with various environmental conditions

Environmental condition	Relative signal intensity (%)	
	bglu1	*bglu2*
control (28°C)	100	100
drought	0.4	0.2
cold (5°C)	105	40
salt (0.3 M NaCl)	36	64
flooding (0.5 cm high water)	37	4
0.5 M manitol	85	130
10 mg/ml Ethephon	64	67
10^{-4} M abscissic acid.	55	88

3.4 Semiquantitative RT-PCR analysis

The mRNA level in various parts of 5-day-old rice seedling was also determined by using the more sensitive technique of reverse-transcription (RT-PCR) using first strand cDNA as the template to obtain a semiquantitative assessment of mRNA levels. Co-amplification of a conserved region of β-actin was used as an internal standard. In RT-PCR, the 3' UTR gene specific primers used for generating the specific probes used in Northern blot analysis were similarly used to generate the *bglu1* and *bglu2* PCR fragments with known product lengths, 287 and 264 bp, respectively. PCR products were analyzed by electrophoresis to show a single band of the expected product size was produced (Figure 3.17A). The number of thermocycles and template dilution concentrations used were optimized to allow quantification without saturation (Figure 3.17B).

The relative intensity of *bglu1* and *bglu2* 3' UTR cDNA to β-actin cDNA products was used to compare the transcript levels of β-glucosidase in different plant parts. The results in Figure 3.18 showed that *bglu1* and *bglu2* were expressed in all of the organs tested. The transcript levels for each was different for different organs of rice seedlings (Figure 3.19). The transcript levels of *bglu1* appeared not to show strong cell-type expression specificity, since expression differed only slightly between different organs, while *bglu2* exhibited higher mRNA levels in shoot tissues than in root tissues. These results should be considered only semiquantitative, since it is not known how actin gene expression varied in these organs.

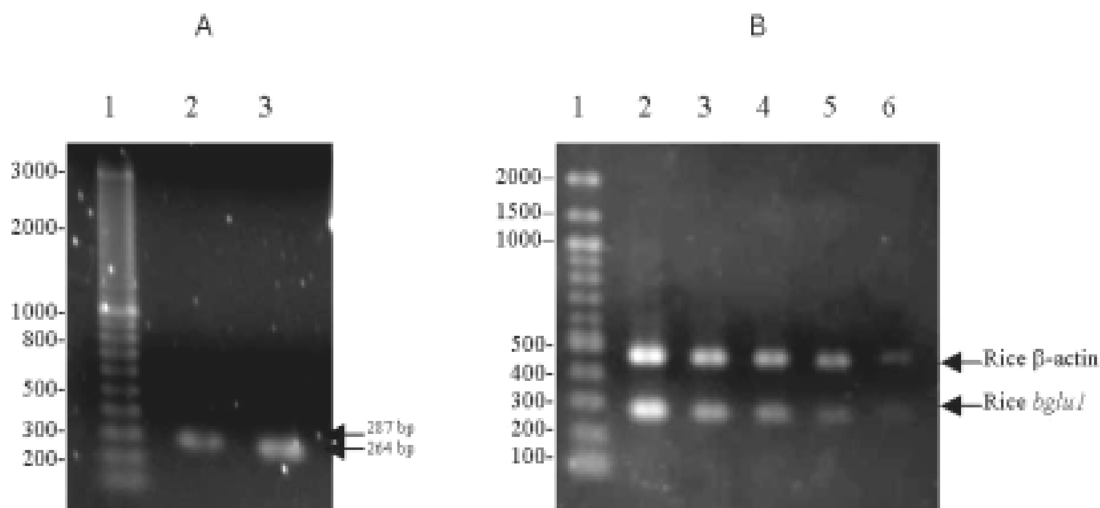


Figure 3.17 1% agarose gel electrophoresis of RT-PCR products. A, Examples of specificity of primers derived from 3'UTR of *bglu1* and *bglu2*. Lane 1, 100 bp EZ load DNA marker (Bio-Rad); lanes 2 and 3, 3'UTR cDNA fragment of *bglu1* and *bglu2*, respectively. B, RT-PCR with primers to co amplified rice β-actin and *bglu1* using serial dilutions of 1, 1/5, 1/10, 1/20, and 1/40 X of cDNA from rice seedlings (lanes 2, 3, 4, 5 and 6, respectively) as the template. Lane 1 is 1 kb ladder marker (New England Biolabs, Beverly, MA).

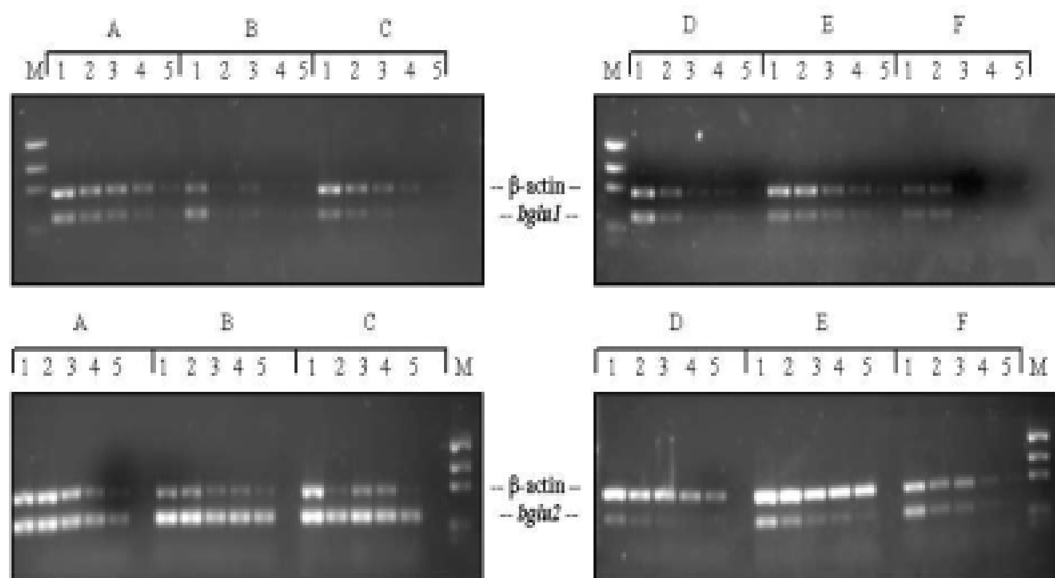


Figure 3.18 1% agarose gel electrophoresis of RT-PCR products to determine the mRNA levels in various parts of rice seedling. In RT-PCR, the gene specific 3'UTR primers were used to generate *bglu1* and *bglu2* PCR fragments with known product lengths, 287 and 264 bp, respectively, and co-amplification of the conserved region of β-actin was used as an internal standard. Column A, mesocotyl lower part; B, mesocotyl upper part; C, coleoptile; D, root elongation zone; E, root tip; and F endosperm. Lanes 1 to 5 of each column are dilutions 1, 1/5, 1/10, 1/20, and 1/40 X of cDNA from each rice seedling tissue in PCR reaction, respectively. M is 100 bp Molecular Mass marker (New England Biolabs, Beverly, MA); β-actin is rice β-actin PCR product; *bglu1* and *bglu2* are 3′ UTR PCR products of *bglu1* and *bglu2*, respectively.

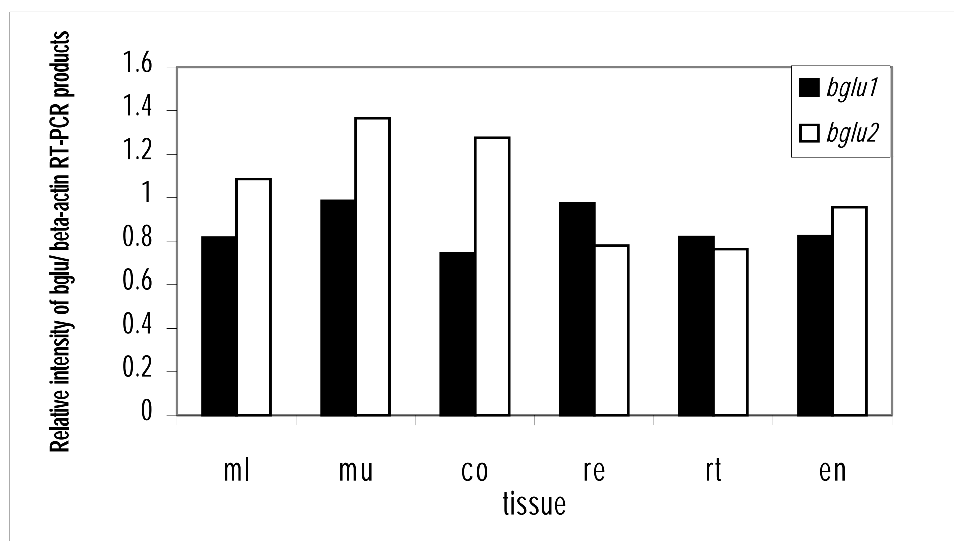


Figure 3.19 Semiquantitative RT-PCR to determine the *bglu1* and *bglu2* mRNA levels in various tissues of 5-day-old rice seedlings. Note that the plot showed the relative intensity of PCR products of *bglu1* or *bglu2* cDNA to β-actin cDNA (the co-amplified product as an internal control). Ml, mesocotyl lower part; mu, mesocotyl upper part; co, coleoptile; re, root elongation zone; rt, root tip and en, endosperm.

3.5 Expression of rice β-glucosidases in *E. coli*

In this study, cDNAs encoding mature β-glucosidase protein, BGlu1 and BGlu2, were amplified by the polymerase chain reaction (PCR) and cloned into various expression vectors in order to generate a highly effective system for producing recombinant protein for functional and structure analysis. Both proteins were produced in active form in *E. coli* and hydrolyzed p-nitrophenol β-D-glucoside and -D-fucoside at significant levels compared to control culture extracts. However, BGlu1 seemed to be found in more active form, while most of BGlu2 fusion protein appeared to be localized to inclusion bodies, though some active, soluble protein was produced.

3.5.1 Recombinant protein expression of rice BGlu1

To achieve soluble protein expression, the pET vector expression system and the *E. coli* host strain were used in this study. The *bglu1* cDNA including the *bglu1* stop codon was inserted into pET23d(+) (at the *Bam*H I and *Eco*R I sites) and pET32a (+) (at the *Sna*B I and *Eco*R I sites). The constructs were used to transform BL21 (DE3) *E. coli,* except for the pET32a(+) construct which was put into both BL21 (DE3) and Origami(DE3) *E. coli*. Origami(DE3) *E. coli* is the redox-deficient strain which allows cytoplasmic disulfide bond formation. BL21(DE3) *E. coli* is a general expression host strain with a normal reducing environment in the cytoplasm, so disulfide bonds will not generally form.

The expression profiles among these systems were different. The pET23d(+)-*bglu1* expressed in BL21(DE3) produced the non-fusion version of the mature BGlu1

protein, which was detected at 50 kDa on SDS-PAGE (Figure 3.20). However, all of the expressed protein was localized to inclusion bodies. Comparisons of the protein profiles of induced and uninduced cultures by SDS-PAGE showed that the induced cultures of pET 32a(+)-*bglu1*/Origami(DE3), the thioredoxin (TRX) fusion BGlu1 was highly expressed and an intense band in the 66-kDa region was seen on the SDS-PAGE (Figure 3.21). This band was absent from the profiles of uninduced cultures as well as from those of induced cultures containing pET32a(+) plasmid without a *bglu1* cDNA insert (negative control). Typically, the gel scanning densitometry indicated that the thioredoxin fusion protein produced was approximately 17% of the total protein in the lysate, and approximately 26% of the total expressed BGlu1 was in soluble form. The pET32a(+)-*bglu1* construct expressed in BL21(DE3), however, showed no overexpressed protein band present in any protein fractions on SDS-PAGE (Figure 3.20).

To check for the activity of β-glucosidases, both with pNPG and pNP β-D-fucoside, hydrolysis assays were done with the soluble crude protein extract fraction of all three expression systems and compared to pET control (Table 3.4). In the crude protein extract of pET32a(+)-*bglu1*/Origami(DE3) system, activities of 1.75 and 3.15 units/mg were measured with pNPG and pNP β-D-fucoside, respectiviely. β-glucosidase activity could not be detected in the soluble fraction of uninduced cultures. While the clone pET32a(+)-*bglu1*/Origami(DE3) showed high β-glucosidase activity, no activity with these two substrates could be measured in the enzyme extracts of pET32a(+)-*bglu1*/BL21(DE3) or pET23d(+)-*bglu1*/BL21(DE3) cultures.

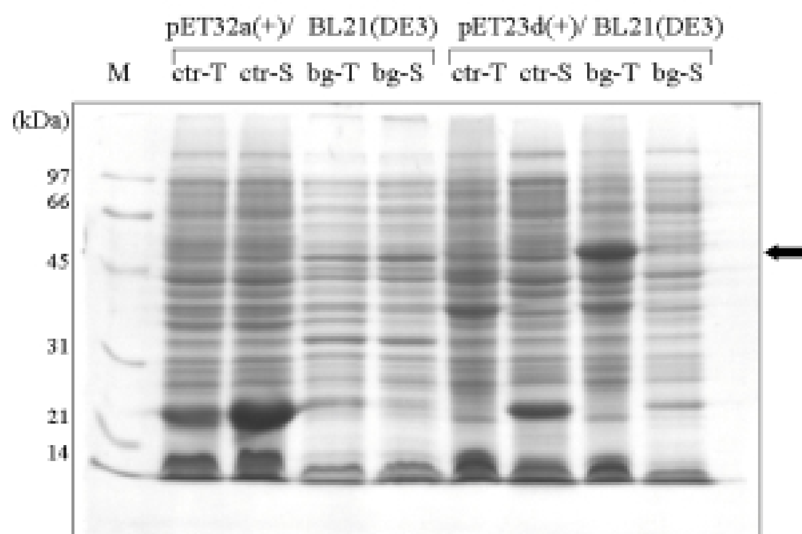


Figure 3.20 SDS-PAGE of recombinant BGlu1 protein expressed in the pET23d (+) and pET32a(+) systems. The profiles showing the total (T) and soluble (S) proteins expressed in BL21(DE3) *E. coli* transformed with pET23d(+), pET32a(+) negative control (ctr) and the recombinant plasmids, pET23d(+)-*bglu1,* pET32a(+)-*bglu1* (bg) after incubation in the presence of 0.4 mM IPTG, at 20°C for 8 h. Note that M represents protein standard marker (Bio-RAD) lane and the arrow points to the positions of mature BGlu1 monomer at 50 kDa.

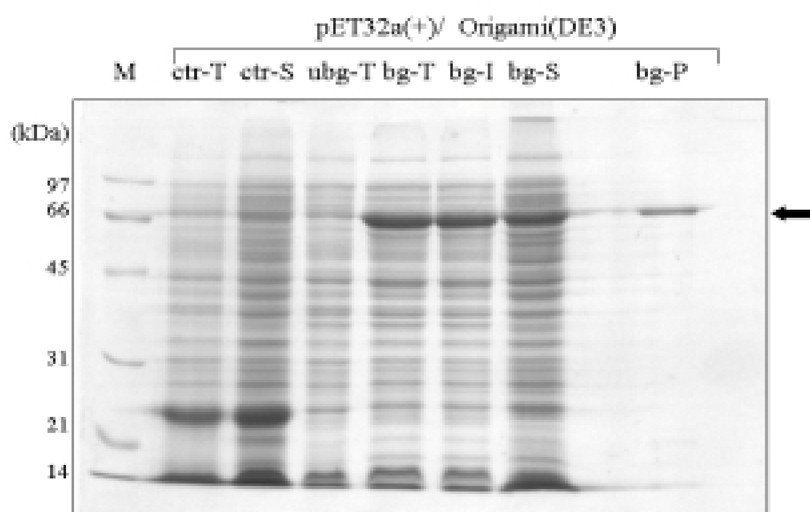


Figure 3.21 SDS-PAGE of recombinant BGlu2 protein expressed in the pET32a(+) system. The profiles of showing the total (T), insoluble (I), soluble (S), purified (P) proteins expressed in Origami(DE3) *E. coli* transformed with pET32a(+) negative control (ctr) and the recombinant plasmids, pET32a(+)-*bglu1* (bg1) after incubation in the presence of 0.4 mM IPTG, at 20°C for 8 h. Note that M represents protein standard marker (Bio-RAD) lane and the arrow points to the positions of thioredoxin fusion BGlu1 monomer at 66 kDa, and ubg-T represents total protein profile of uninduce culture of pET32a(+)-*bglu1* after incubation at 20°C for 8 h.

Table 3.4 Comparison of specific activity between the recombinant protein expressed in various *E. coli* systems

Construct clone	Specific activity (μmole min^{-1} mg^{-1})[a]	
	pNP β-D-glucoside	pNP β-D-fucoside
PET32a(+)-*bglu1*/ Origami(DE3)	1.75 ± 0.08	3.15 ± 0.05
Control-pET32a(+)/ Origami(DE3)	<0.001	<0.001
PET32a(+)-*bglu1*/ BL21(DE3)	<0.001	0.003 ± 0.05
Control-pET32a(+)/ BL21(DE3)	<0.001	0.003 ± 0.03
PET23d(+)-*bglu1*/ BL21(DE3)	<0.001	0.004 ± 0.06
Control-pET23d(+)/ BL21(DE3)	<0.001	0.004 ± 0.05

[a] Assays were done with 1 mM each substrate at 30° C for 10 min at pH 5.0.

The condition for biosynthesis of functional BGlu1 protein in pET32a(+)-*bglu1*/Origami(DE3) expression was optimized. In the experiment using 0 to 1 mM final concentration of IPTG, there was no constitutive expression of protein in the absence of the inducer, and the amount of BGlu1 in cell lysates was constant as the IPTG concentration increased from 0.4 to 1 mM (Figure 3.22). There were no significant difference in BGlu1 activity expressed in *E. coli* at 15, 20 and 27°C, while at 37°C protein was much less active (Figure 3.23).

The TRX fusion BGlu1 was purified by Ni-NTA agarose affinity chromato - graphy and concentrated with a 30 kDa molecular weight cut-off ultracentricon. The recombinant protein was eluted from the column as a single peak (approximately 95% pure) and a band corresponding to 66 kDa was observed in SDS-PAGE (Figure 3.21). Approximately 3 mg of purified protein could be obtained per litre of bacterial culture. The enzyme was stable for several months at 4°C. The enzyne activtity with pNPG still remained after the TRX fusion-BGlu1 was cleaved with enterokinase.

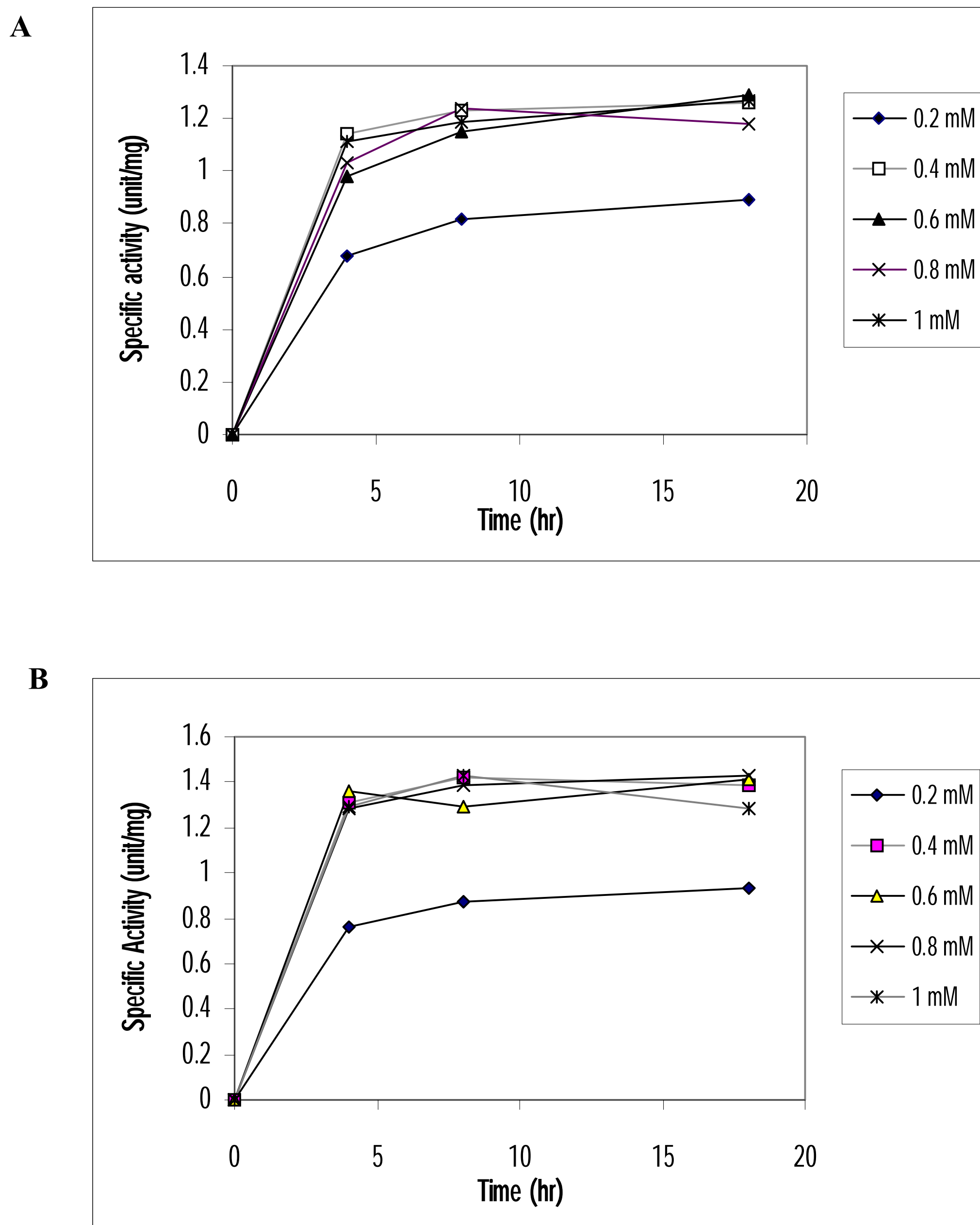


Figure 3.22 Comparison of β-glucosidase activity of pET32a(+)-*bglu1* expressed in Origami(DE3) *E .coli* after incubation in the presence of various IPTG concentrations and at times ranging from 0 to 18 h. A, at 20°C and B, at 27 °C.

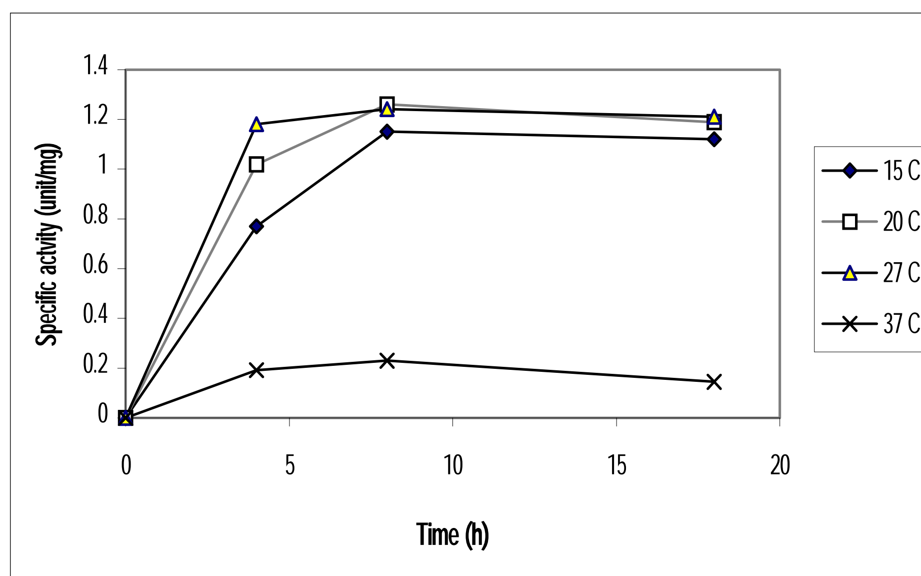


Figure 3.23 Comparison of β-glucosidase activity of pET32a(+)-*bglu1* expressed in Origami(DE3) *E. coli* after incubation in the presence of 0.4 mM IPTG, at different temperatures for 0 to 18 h.

3.5.2 Recombinant protein expression of rice BGlu2

Similar attempts to express protein from the *bglu2* cDNA were made in many pET systems. The *bglu2* cDNA including the *bglu2* stop codon was inserted into pET23d(+), pET32a(+), pET39b(+) and pET40b(+) at the *Nco* I and *Xho* I sites. The constructs were used to transform BL21(DE3) *E. coli,* except for the pET32a(+) construct which was put into both BL21(DE3) and Origami(DE3) *E. coli*.

The expression profiles among these systems were different. The pET23d(+)-*bglu2* expressed in BL21(DE3) produced the mature BGlu2 protein which was detected at approximately 50 kDa on SDS-PAGE (Figure 3.24A). However, all of the expressed proteins were insoluble and presumed to be in inclusion bodies at all conditions tested, which included type of medium, IPTG concentration, temperature

and time of induction. In induced cultures of pET32a(+)*bglu2*/Origami(DE3), the TRX fusion BGlu2 was highly expressed and an intense band in the 66-kDa region was seen on the SDS-PAGE (Figure 3.24B). This band was absent from extracts of induced cultures containing pET32a(+) plasmid without insert (negative control). Typically, gel scanning densitometry indicated that the TRX fusion protein expressed was approximately 42% of the total protein in the lysate but no obvious band was detected in the soluble fractions (Figure 3.24C). When the pET39b(+)- and pET40b (+)-*bglu2* constructs expressed in BL21(DE3), however, no overexpressed protein band was present in the total protein, the soluble protein and periplasmic fractions on SDS-PAGE. Enzyme activity assays were performed to check for β-glucosidase activity with pNPG and pNP β-D-fucoside in the soluble crude protein extract fractions of all three expression systems compared to no-insert-pET control. In the crude protein extract of pET32a(+)-*bglu2*/Origami(DE3) system, an activity with pNPG and pNP β-D-fucoside of 0.005 and 0.011 unit/ mg, respectively, was measured, but this was not much different from negative control activity (Table 3.5). However, no activity with these two substrates could be measured in the enzyme extracts of pET23d(+)-*bglu2*/BL21 or pET39b and pET40b systems.

Table 3.5 Comparison of specific activity between the recombinant protein expressed in various *E. coli* systems.

Construct clone	Specific activity (μmole min^{-1} mg^{-1})[a]	
	pNP β-D-glucoside	pNP β-D-fucoside
Bglu2-pET32a(+)/ Origami(DE3)	0.005	0.011
Control-pET32a(+)/ Origami(DE3)	<0.001	<0.001

[a] Assays were done with 1 mM each substrate at 30° C for 20 min at pH 5.0.

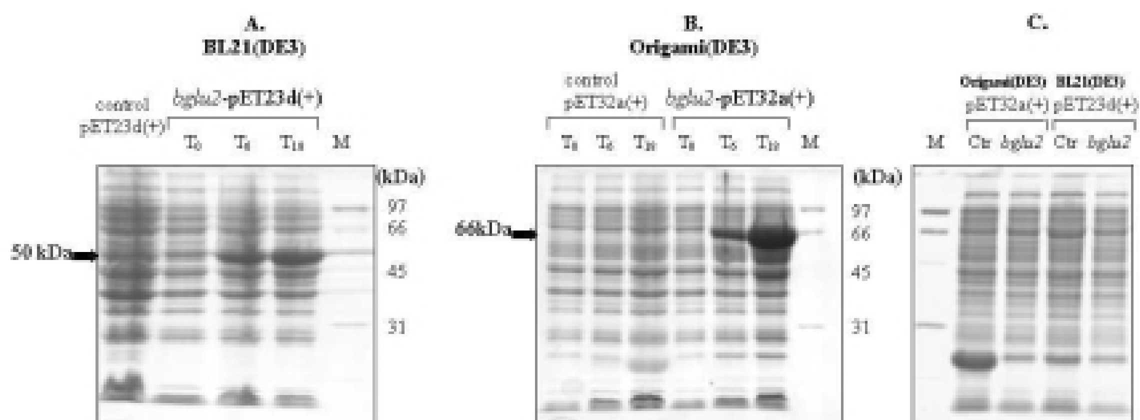


Figure 3.24 SDS-PAGE profiles of recombinant BGlu2 protein expressed in pET23d (+) and pET32a(+) systems. A and B, total protein profiles of cell lysates from *E. coli* transformed with nonrecombinant plasmid (control pET23d(+), control pET32a(+)), and the recombinant plasmids (pET23d(+)- and pET32a(+)-*bglu2*) before and after induction with 0.6 mM IPTG and grown at 15°C for 0 (T_0), 6 (T_6), and 18 (T_{18}) h. The arrow point to positions of recombinant β-glucosidase monomer. C, Soluble fractions profiles of protein extracted from A and B. Note that M represents protein standard marker (Bio-RAD) and the arrow points to the positions of mature and thioredoxin fusion BGlu2 protein monomer at 50 and 66 kDa, respectiviely.

Western blot analysis with peptide conjugate (406) and inclusion body (404) BGlu2 antisera antisera, confirmed that most of the reactive protein for Bglu2 was in the insoluble fraction, though a small amount was in the soluble fraction (Figure 3.25). The recombinant BGlu1 protein also cross-reacted to some degree with 404 antiserum, but no reaction was seen for proteins from the control *E. coli* lysate. Unfortunately, no specific antiserum could be produced for BGlu1.

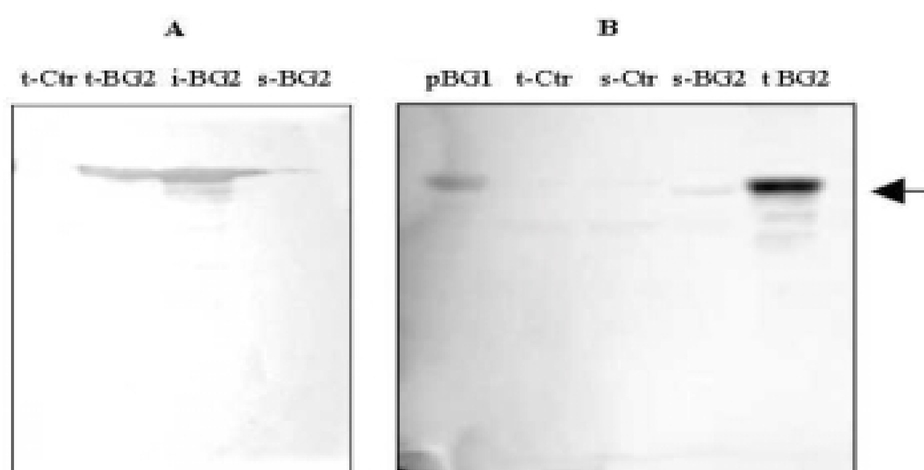


Figure 3.25 Immunoblots of recombinant proteins expressed in *E. coli*. (A) and (B) the blots detected with rice β-glucosidase antisera produced against BGlu2 peptide conjugate (R406) and BGlu2 inclusion bodies of mature protein (R404), respectively. The total (t-), soluble (s-), insoluble, (i-) and purified (p-) proteins of *E. coli* transformed with nonrecombinant plasmid pET32a(+), negative control (Ctr) and the recombinant plasmids pET32a(+)-*bglu1* (BG1) and pET32a(+)-*bglu2* (BG2) after incubation in the presence of 0.4 mM IPTG, at 20°C for 8 h. The arrow points to the positions of thioredoxin-fusion protein monomer at 66 kDa.

3.6 Biochemical properties of rice BGlu1

3.6.1 Substrate specificity

The activity of the purified rice BGlu1 β-glucosidase towards artificial and natural glycosides is summarized in Table 3.6 and Table 3.7. BGlu1 hydrolyzed the disaccharides, sophorose (β-1→ 2), laminaribiose (β-1 → 3), cellobiose (β-1 → 4), and gentiobiose (β-1 → 6). It also hydrolyzed laminaritriose and cello-oligosaccharides with degree of polymerization (DP) of 3-6. The relationship between the kinetic parameters (K_m, k_{cat}, and k_{cat}/K_m) and the DP of oligosaccharide substrates are shown in Table 3.6. BGlu1 hydrolyzed sophorose and gentiobiose with high K_m and low k_{cat} values similar to cellobiose. For cello-oligosaccharides, K_m decreased with increasing chain length of the substrate, while k_{cat} values appear to be relatively independent of DP. Catalytic efficiency k_{cat}/K_m values increase with increasing chain length of the cello-oligosaccharides in which k_{cat}/K_m of cellohexaose is 3058-fold higher than the values of cellobiose. In contrast, the hydrolysis of laminari-oligosaccharides the k_{cat}, and k_{cat}/K_m values decreased with increasing chain length from DP 2 to 3. The rate of hydrolysis of laminaritetraose and laminaripentaose were not detectable, since no product could be detected. In addition to hydrolysis activity, BGlu1 showed transglycosylation activity towards laminari-oligosaccharides with DP of 2 and 3, and cello-oligosaccharides with a DP of 3 and 4 at the relatively low substrate concentration of 5 mM, as detected by TLC (Figure 3.26 and Figure 3.27).

Table 3.6 Kinetic parameters of rice BGlu1 during the hydrolysis of oligosaccharides

Substrate	K_m (mM)	K_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} mM^{-1})
Cello-oligosaccharides (DP)			
2	31.50 ± 1.6	1.52 ± 0.13	0.05 ± 0.002
3	0.72 ± 0.02	18.13 ± 0.35	25.36 ± 0.37
4	0.28 ± 0.01	17.34 ± 0.63	61.06 ± 0.37
5	0.24 ± 0.01	16.90 ± 0.06	71.50 ± 2.2
6	0.11 ± 0.01	16.93 ± 0.32	152.9 ± 0.5
Laminari-oligosaccharides (DP)			
2	2.05 ± 0.1	31.91 ± 3.12	15.69 ± 1.87
3	1.92 ± 0.04	21.2 ± 0.17	11.04 ± 0.16
4	N. D.[a]	N. D.[a]	N. D.[a]
5	N. D.[a]	N. D.[a]	N. D.[a]
Sophorose	13.89 ± 0.92	5.87 ± 0.21	0.42 ± 0.02
Gentiobiose	38.33 ± 4.11	0.99 ± 0.08	0.03 ± 0.003

[a] means not detectable.

Table 3.7 Kinetic parameters of rice BGlu1 during the hydrolysis of pNP-glycosides, alkyl glucosides and natural glucoside substrates

Substrate	K_m (mM)	K_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} mM^{-1})
pNP-β-D-glucoside	0.23 ± 0.02	7.93 ± 0.37	34.7 ± 1.4
pNP-β-D-fucoside	0.23 ± 0.01	13.3 ± 0.2	57.5 ± 2.1
pNP-β-D-galactoside	3.16 ± 0.16	3.40 ± 0.18	1.08 ± 0.01
pNP-β-D-mannoside	1.27 ± 0.10	1.32 ± 0.05	1.01 ± 0.02
pNP-β-D-xyloside	1.29 ± 0.11	1.14 ± 0.01	0.88 ± 0.07
pNP-α- L-arabinoside	1.21 ± 0.13	1.24 ± 0.19	1.03 ± 0.05
pNP-β-D-cellobioside	0.78 ± 0.04	2.48 ± 0.04	3.18 ± 0.13
Methylumberiferyl-β-D-glucoside	0.28 ± 0.02	5.01 ± 0.23	18.0 ± 0.9
Methyl-β-D-glucoside	49.2 ± 5.1	0.65 ± 0.08	0.013 ± 0.0003
n-Heptyl-β-D-glucoside	1.41 ± 0.27	0.87 ± 0.22	0.62 ± 0.02
n-Octyl-β-D-glucoside	0.63 ± 0.01	0.78 ± 0.01	1.25 ± 0.03
Amygdalin	8.20 ± 0.44	0.44 ± 0.05	0.05 ± 0.004
Prunasin	5.83 ± 0.53	1.41 ± 0.09	0.24 ± 0.01
Dhurrin	0.62 ± 0.10	0.06 ± 0.006	0.09 ± 0.004
Torvoside A	0.99 ± 0.11	0.67 ± 0.05	0.68 ± 0.02
Pyridoxine-5'-O-β-D-glucoside	0.71 ± 0.10	1.64 ± 0.19	2.30 ± 0.09

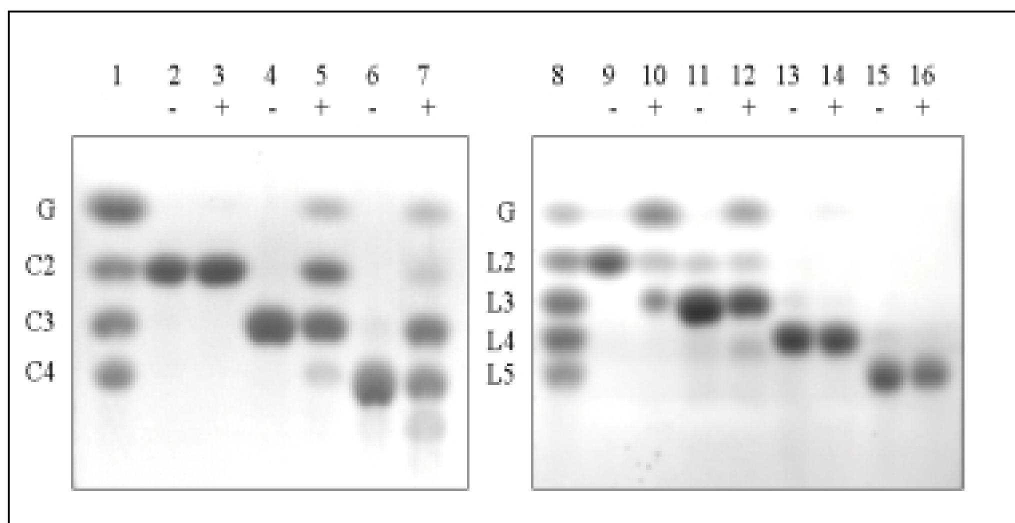


Figure 3.26 Hydrolysis and transglycosylation of oligosaccharides by rice BGlu1 detected by TLC. The BGlu1 (0.05 nmol) was incubated with 5 mM substrates for 1 h and the products were detected by the staining method. Samples were incubated with (+) or without (-) enzyme. Lane 1, markers: glucose (G) and cello-oligosaccharides of DP 2-4 (C2-C4); lane 2 and 3, cellobiose; lane 4 and 5, cellotriose; lane 6 and 7, cellotetraose; lane 8, marker: glucose (G) and laminari-oligosaccharides of DP 2-5 (L2-L5); lane 9 and 10, laminaribiose; lane 11-12 laminaritriose; lane 13-14, laminaritetraose; lane 15-16, laminaripentaose.

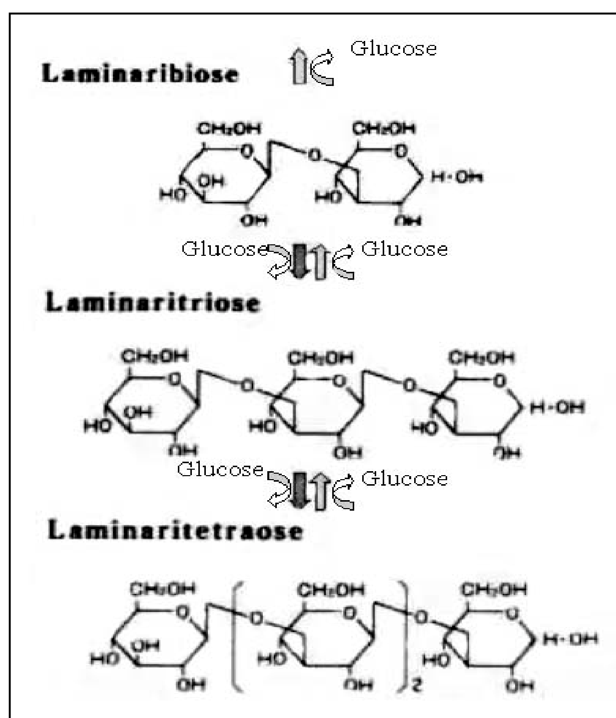


Figure 3.27 Scheme of the hydrolysis and transglycosylation of laminari-oligosaccharides by rice BGlu1.

The purified BGlu1 hydrolyzed the pNP β-glycosides of the monosaccharides, D-glucose, D-fucose, D-galactose, D-mannose, D-xylose, and the disaccharide D-cellobiose. The enzyme also hydrolyzed pNP-α-glycosidic linkage of L-arabinoside but not the β- L-arabinoside. BGlu1 did not hydrolyze the α-link pNP derivative of D-glucose or β-glycosidic linked pNP glycosides of L-fucoside, D-thioglucoside or D-thiofucoside. Kinetic studies of rice BGlu1 enzyme were performed to determine the K_m and the k_{cat}, and k_{cat}/K_m values of the enzyme for various pNP β-D-glycosides (Table 3.7). Interestingly, the enzyme showed both a low K_m value and high k_{cat} value for both pNP β-D-glucoside and pNP β-D-fucoside. Moreover, the k_{cat}/K_m ratio of the enzyme was highest for pNP β-D-fucoside. BGlu1 hydrolyzed pNP-β-D-

galactopyranoside, pNP β-D-mannoside, pNP β-D-xylose, pNP α- L-arabinoside and pNP β-D-cellobioside with much lower catalytic efficiency than pNPG.

The products of the hydrolysis of pNP β-D-cellobioside by BGlu1 identified by TLC included pNPG, pNP and glucose. The mechanism of the hydrolysis of pNP β-D-cellobioside by BGlu1 was further analyzed. There products of pNP β-D-cellobioside hydrolysis were seen by TLC after hydrolysis with different times of incubation ranging from 0 to 20 min with 2.5 min intervals and 1 h to be pNP, pNPG and glucose with no occurrence of cellobiose (Figure 3.28). This indicates the initial cleavage is between the two glucose residues.

Studies with several available natural substrates showed that these compounds were hydrolyzed poorly or not at all relative to pNPG. BGlu1 hydrolyzed cyanogenic glucosides such as prunasin from *Prunus serotina*; amygdalin from almond, and dhurrin from sorghum, but it did not hydrolyze linamarin, from cassava or DIMBOA-glucoside from maize. This enzyme also hydrolyzed steroid glucoside from *Solanum tovumm* (Torvoside A) but with lesser efficiency than pNPG. BGlu1 did not hydrolyze the phenolic glycosides tested (arbutin, salicin), thioglucoside (sinigrin), hydroxy coumarin glucoside (esculin), flavonoid glucoside (naringin), magniferin-C-glucoside or α-D-disaccharides (nigerose) (Table 3.8).

Pyridoxine-5'-O-β-D-glucoside was cleaved by BGlu1, but could not be quantified accurately by spectrophotometric measurement, because its breakdown products interfered with the PGO assay, yielding reddish color rather than green with ABTS. Determined by reverse phase HPLC using pyridoxine-5′-O-β-D-glucooside as standard (Figure 3.29), BGlu1 showed hydrolytic activity with pyridoxine-5′-O-β-D-

glucoside with estimated K_m and K_{cat} values, 0.71 mM and 1.64 s^{-1}, respectively. BGlu1 hydrolyzed this substrate to give glucose and pyridoxine, as detected by TLC (Figure 3.30).

BGlu1 hydrolyzed MUGlc with high catalytic efficiency relative to pNPG. For the studies with alkyl glucosides, these compounds were hydrolyzed but poorly relative to pNPG or not hydrolyzed. BGlu1 hydrolyzed methyl, n-heptyl, and n-octyl-β-D-glucosides, but did not hydrolyze, n-amyl, or phenyl-β-D-glucoside (Table 3.7 and 3.8). In addition, the zymogram assay of recombinant BGlu1 with 4MUGlc in acidic native gel showed that the enzyme hydrolyzed this substrate, while the negative control of non recombinant plasmid protein lysate did not (Figure 3.31). The enzyme also hydrolyzed 5-bromo-4-chloro-3-indolyl-β- D-glucoside (X-glucoside) yielding the blue precipitate products (Figure 3.32).

Table 3.8 Glycosides which were not cleaved by BGlu1

Substrate	mM	Time
pNP-α-D-glucopyranoside	5	1 h
pNP-β- L-arabinoside	5	1 h
pNP-β- L-fucopyranoside	10	1 h
pNP-β-D-thioglucopyranoside	10	1 h
pNP-β-D-thiofucopyranoside	10	1 h
n-Amyl-β-D-glucoside	10	12 h
Phenyl-β-D-glucoside	10	12 h
DIMBOA-glucoside	10	12 h
Linamarin	10	12 h
Salicin	10	12 h
Arbutin	10	12 h
Esculin	10	12 h
Sinigrin	10	12 h
Nigerose	10	12 h
Naringin	10	12 h
Magniferin-C-glucoside	10	12 h

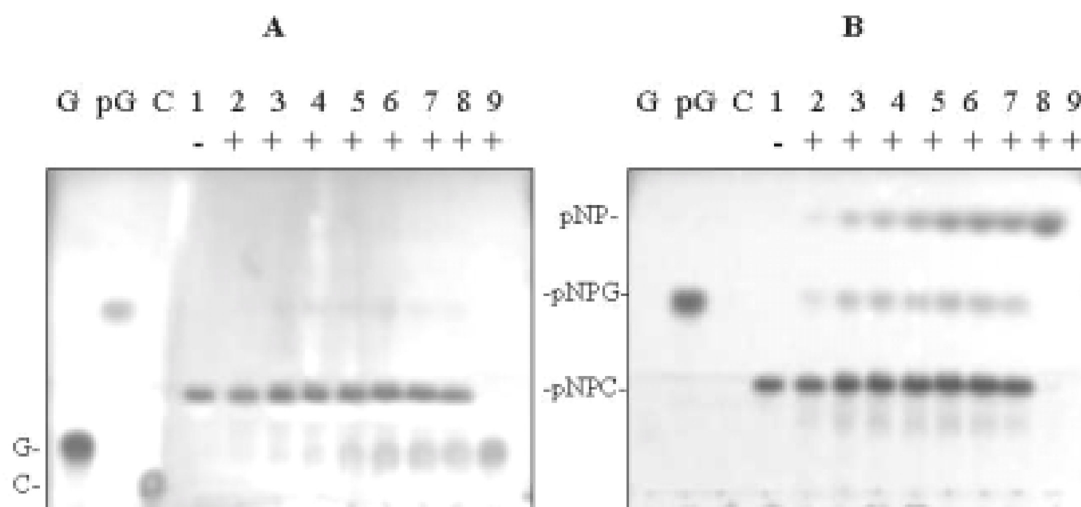


Figure 3.28 Product specificity of rice BGlu1 as assayed with pNP-β-D-cellobioside and detected by TLC. Purified BGlu1 (0.05 nmol) was incubated with 5 mM pNP-β-D-cellobioside at 30°C. for 1 hr. The plate in A was detected with staining method; B, was detected with 356 nm UV light. Samples were incubated with (+) and without (-) enzyme. Lane G is glucose marker; lane pG is pNPG marker; lane C is cellobiose marker. The letters beside the plates represent the these molecules: glucose (G); pNP-β-D-glucoside (pNPG); pNP-β-D-cellobioside (pNPC); pNP (pNP) and cellobiose (C). The incubation times from lanes 2-8 of E and F increase at 2.5 min intervals from 5 to 20 min, and lane 9 is 1 h.

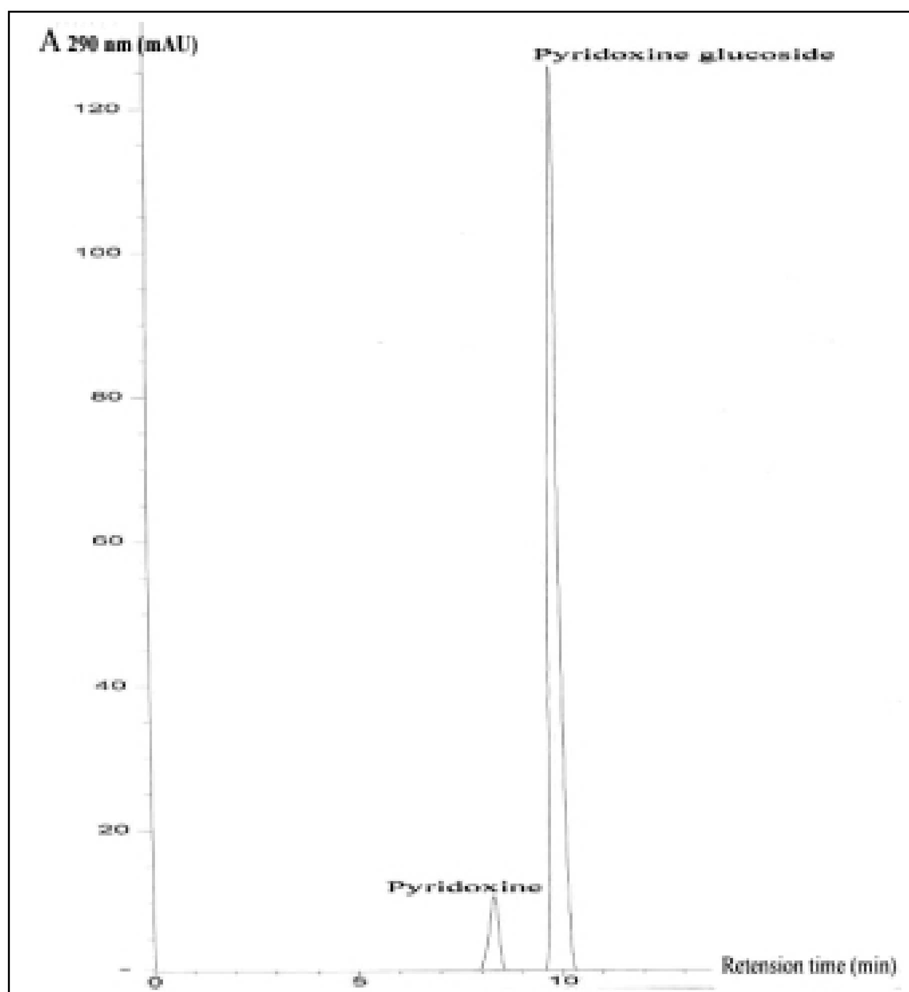


Figure 3.29 HPLC elution profiles of the products of hydrolysis of pyridoxine glucoside by rice BGlu1. The reaction contained BGlu1 (2.25 pmol) and 0.4 mM pyridoxine glucoside in 50 mM NaOAc buffer (pH 5.0) incubated at 30°C for 20 min.

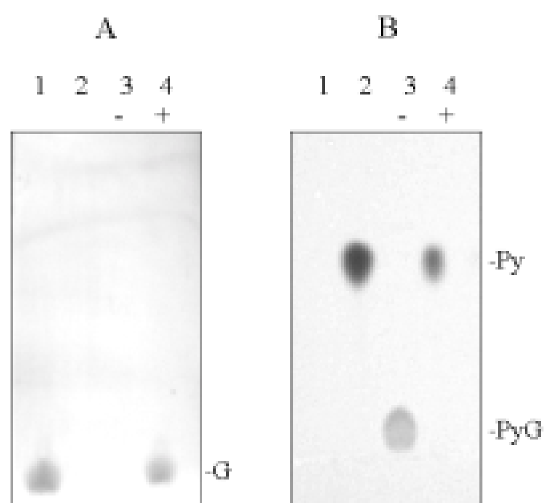


Figure 3.30 Hydrolysis of pyridoxine-5'-O-β-D-glucoside by rice BGlu1 detected by TLC. The substrate rice-pyridoxine-5'-O-β-D-glucoside was incubated at 5 mM final concentration with 0.05 nmol of purified BGlu1. A shows detection with staining, while B was detected with 365 nm UV light. Samples were incubated with (+) and without (-) enzyme. Lane 1, glucose (G) marker; lane 2, pyridoxine (Py) marker; lane 3 and 4, pyridoxine-5'-O-β-D-glucoside (PyG).

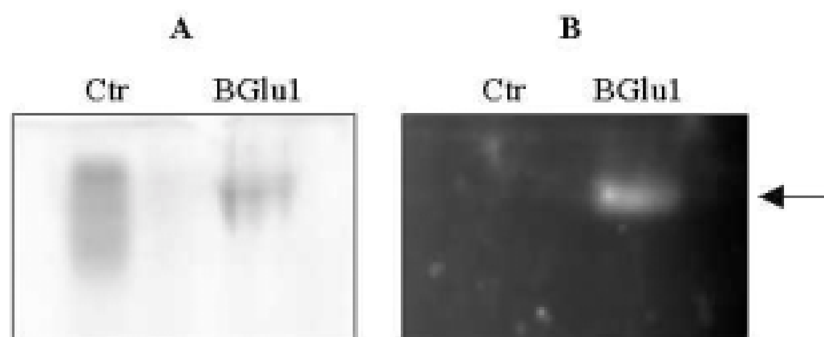


Figure 3.31 Acidic native PAGE (5%) gel zymograms of thioredoxin fusion BGlu1 expressed in Origami(DE3) *E. coli*. A, the acidic native PAGE staining with Coomassie blue. B, a zymogram developed with the fluorogenic substrate 4MUGlc. Note that the recombinant BGlu1 (BGlu1) hydrolyzed this substrate, while the lysate of *E coli* transformed with empty pET32a(+) (Ctr) had no enzyme activity detectable. An arrow points to the activity band.

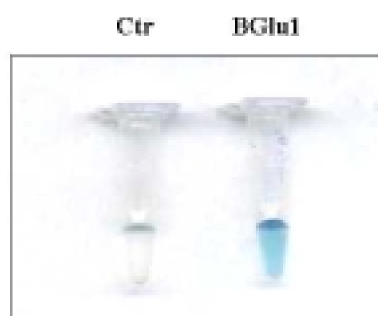


Figure 3.32 Hydrolysis of 5-bromo-4-chloro-3-indolyl-β-D-glucoside (X-glucoside) by rice BGlu1. Purified BGlu1 (BGlu1) (0.05 nmol) was incubated with 5 mM X-glucoside at 30°C for 30 min yielding the blue precipitate products, while the lysate of *E.coli* transformed with empty pET32a(+) (Ctr) had no enzyme activity detectable.

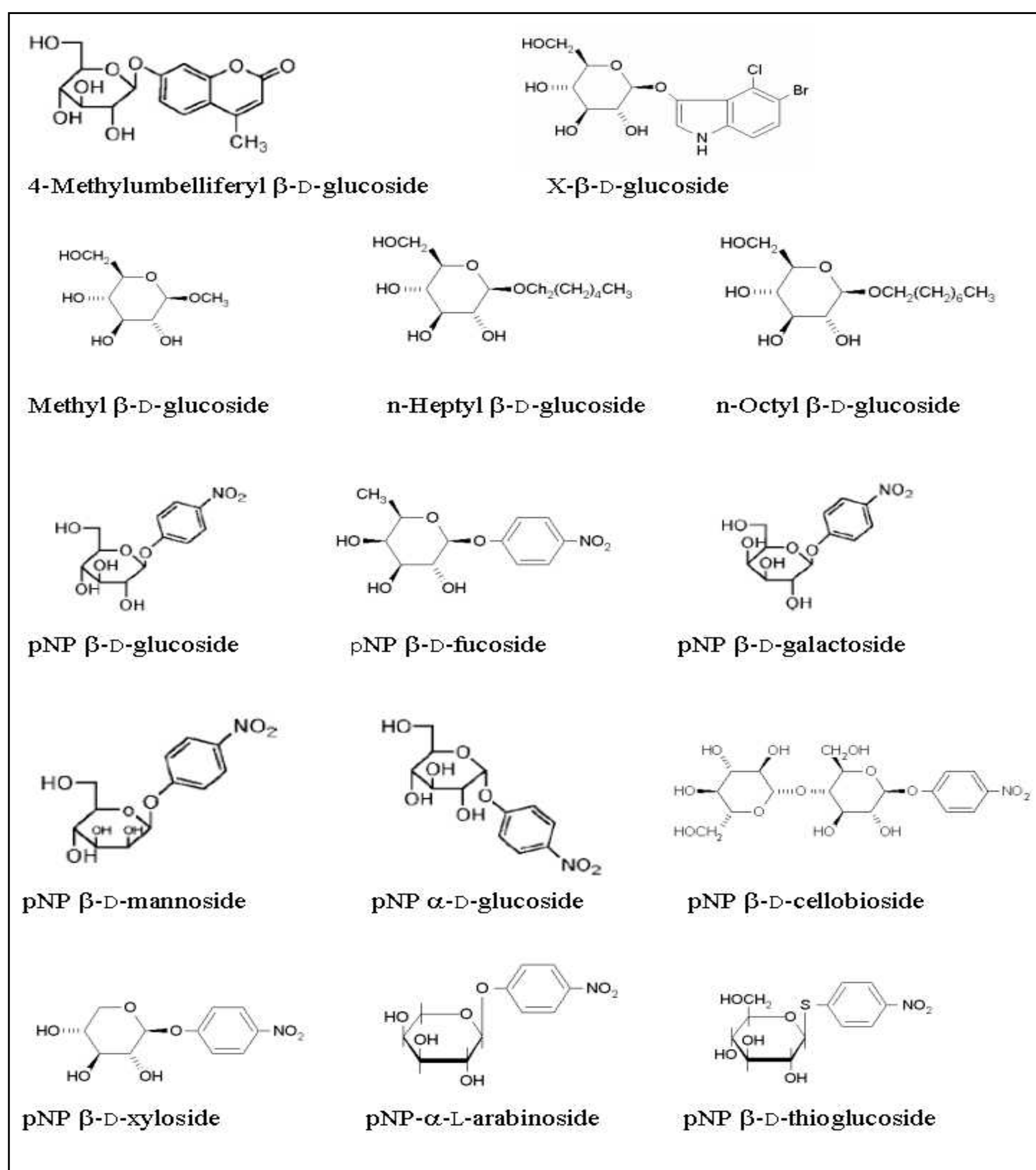


Figure 3.33 Chemical structures of some artificial glycoside substrates used in determining substrate specificity of rice β-glucosidase. Note that all structures came from Sigma.

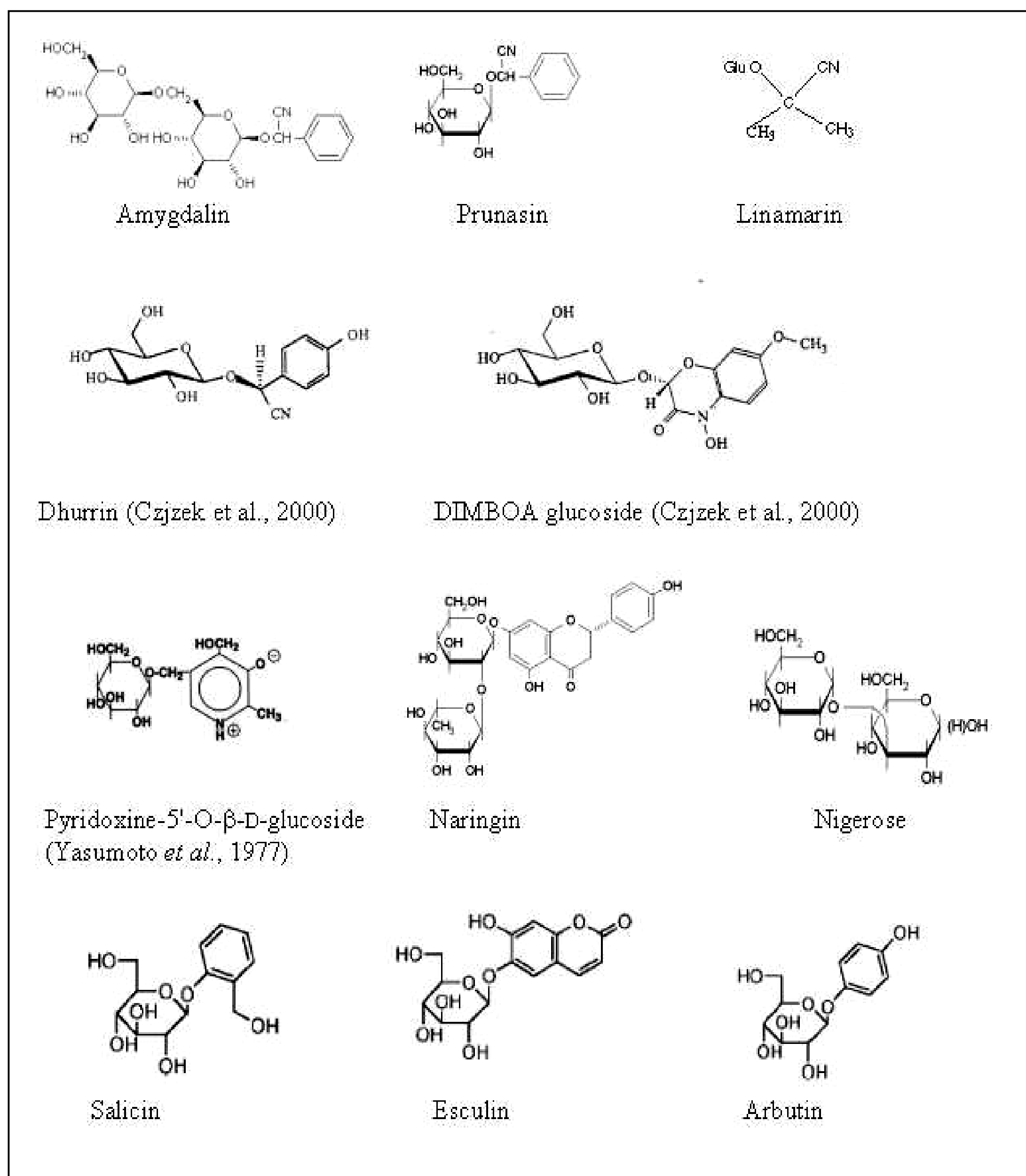


Figure 3.34 Chemical structures of some natural glycoside substrates used in determining substrate specificity of rice β-glucosidase. Note that all structures with no reference came from Sigma.

3.6.2 Molecular weight (Mr) estimation

BGlu1 migrated as a single band on SDS-PAGE and its apparent Mr of TRX fusion and non fusion protein (after cleavage with enterokinase) was estimated to be 66 and 54 kDa, respectively. The native molecular weight of the purified TRX fusion BGlu1 was 66 kDa and supports the idea that rice BGlu1 is a monomeric protein. This estimation is based on a gel-filtration. Figure 3.35 presents calibration curve obtained with protein standard and rice BGlu1 fusion protein as run on Sephacryl S-200. It was clear that rice BGlu1 eluted at the similar elution volume as Bovine albumin, (BSA), indicating that the native rice BGlu1 fusion protein is about the same as the Mr of BSA (66 kDa).

3.6.3 pH and temperature profile of activity and stability

BGlu1 reached optimum activity at pH 5.0 (Figure 3.36). The activity dropped dramatically at a pH greater than 7.5. The pH stability studies showed that BGlu1 was relatively stable over a wide range of pH (5.0-7.0), when incubated for up to 24 hr at the corresponding pH at room temperature (25°C) (Figure 3.37). At a pH above 7.0 the enzyme was less stable after incubations longer than 6 hr. The enzyme lost activity dramatically at acidic pH (3.0-4.0) within 10 min of incubation. The optimum temperature for activity of BGlu1 was a wide range of 15-35°C. The activity decreased dramatically above 40°C (Figure 3.38). The thermostability studies at pH 5.0 showed that the enzyme was stable at 20 and 30°C, when incubated for up to 60 min. (Figure 3.39). The enzyme was less stable at 40°C at longer incubation times. It was very unstable at 50 and 60°C, with complete loss of activity within the first 10 min of incubation.

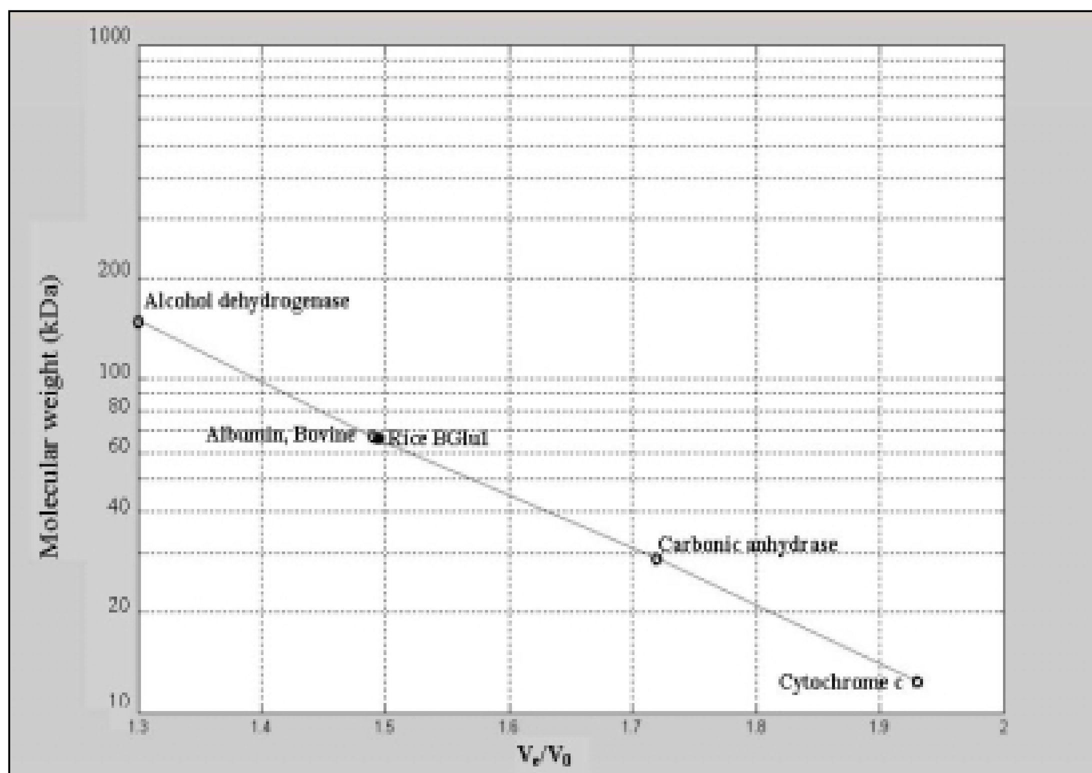


Figure 3.35 Native molecular weight of rice BGlu1 as estimated by gel filtration. Purified rice BGlu1 and protein standards were loaded onto a Sephacryl S-200 column. The fractions were collected and the protein measured by absorbance reading at 280 nm for protein standards, and pNPG activity assay for rice BGlu1. The elution volume of each protein (V_e), and of blue dextran (V_0) were calculated. The V_e/V_0 and molecular weight for each protein were plotted on semilog graph. Calibration curve obtained with protein standard was used to calculate the native molecular weight of rice BGlu1.

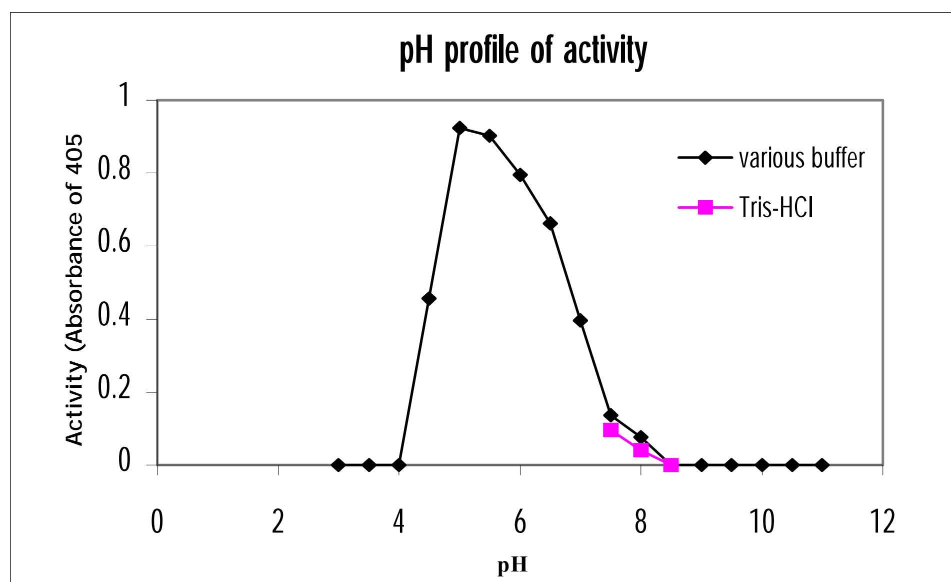


Figure 3.36 pH profile of activity for BGlu1 over a pH range 3.0 to 11 assayed with 1 mM pNPG at 30°C for 10 min.

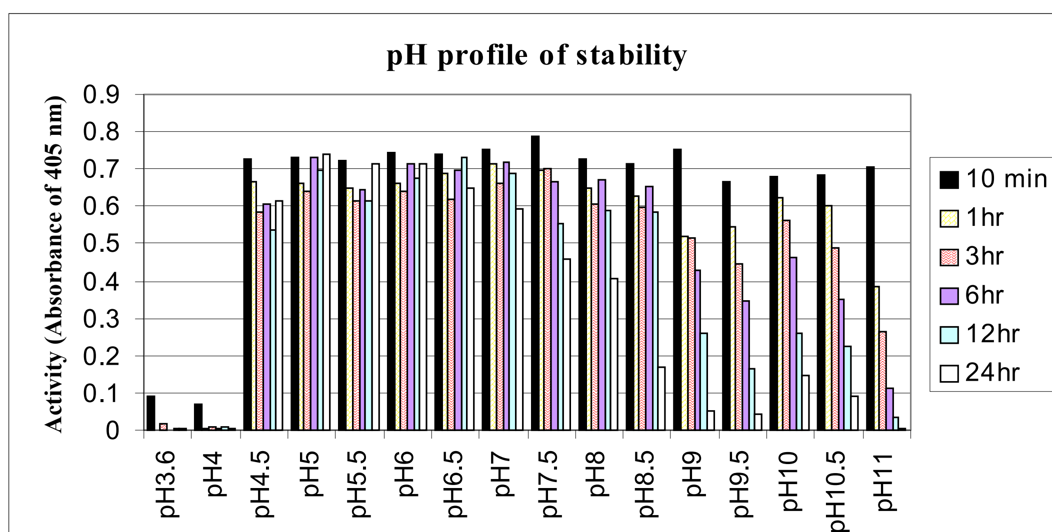


Figure 3.37 The pH stability of BGlu1 after incubation for 10 min to 24 h 25°C. The aliquots of enzyme in each pH buffer at different time course was diluted 140 fold and assayed in 50 mM NaOAc buffer pH 5.0, and incubated with 1 mM pNPG at 30°C for 10 min.

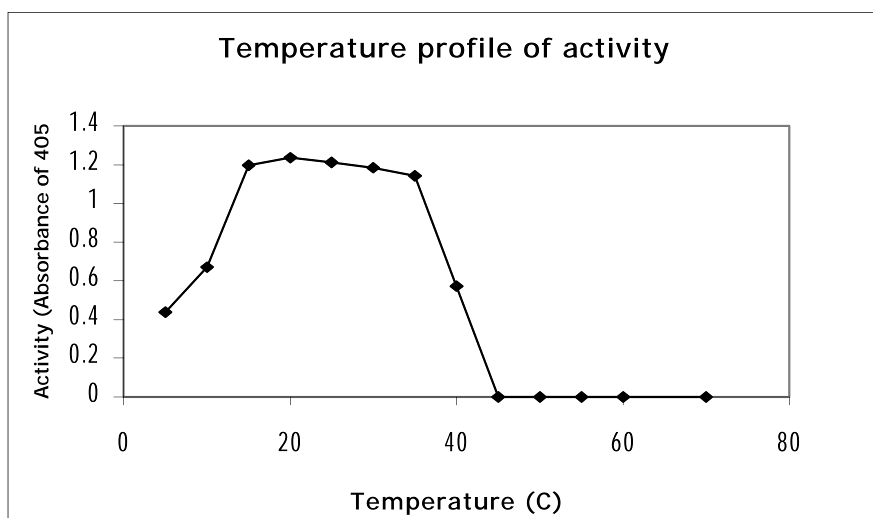


Figure 3.38 Temperature profile of BGlu1 assayed with 1 mM pNPG for 10 min at temperatures ranging from 5 to 70°C.

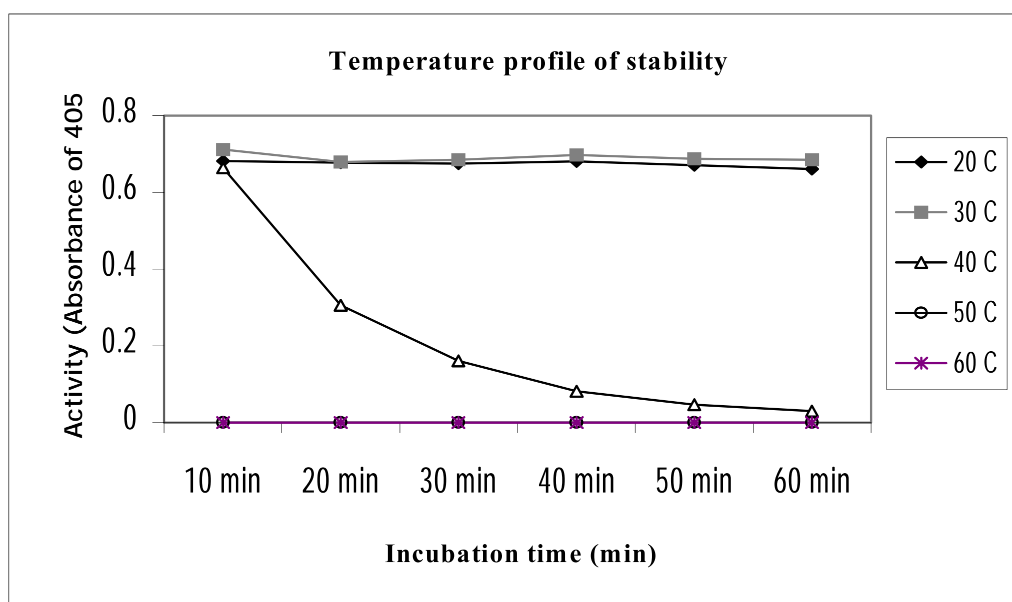


Figure 3.39 Thermostability of BGlu1 incubating at 20-60°C. Incubation time ranged from 0 to 60 min at 10-min increment. Aliquots of enzyme were diluted and assayed in 50 mM NaOAc buffer pH 5.0, and incubated with 1 mM pNPG at 30°C for 10 min.

3.6.4 Inhibition studies

The effect of various potentially inhibitory compounds was tested on rice BGlu1 activity. The inhibitory effects of some organic compounds toward BGlu1 are shown in Table 3.9. Rice BGlu1 was inhibited by many aglycones and aglycone analogues, but not inhibited by monosaccharides (10 mM D-glucose, D-manitol, D-xylose, D-mannose and D-galactose), or EDTA. Glucono-1,5-lactone, a strong inhibitor of many β-glucosidases, strongly inhibited BGlu1 at the level of 0.005 mM. Galactono-1,4-lactone, IAA, GA_3, pyridoxine and 4-methyl umbelliferone, imidazole and Tris base also inhibited BGlu1, but the degree of inhibition was much less.

Table 3.9 Effect of inhibitors on activity of purified rice BGlu1

Inhibitor	Concentration (mM)	Relative activity (%)
None[a]	0	100
D-glucono-1, 5-lactone	0.4	0
	0.1	5
	0.005	39
D-galactono-1, 4-lactone	5	45
	1	74
IAA	5	46
	1	67
GA_3	5	76
Pyridoxine	5	77
4-Methyl umbelliferone	5	81
Imidazole	5	83
D-glucose, D-galactose	10	100
D-xylose, D-mannose	10	100
EDTA	10	101
Tris base	10	75

[a]The standard pNP assay mixture contained 1 mM pNPG, 50 mM NaOAc pH 5.0, and 1 µg enzyme at 30°C

Of the various cations tested, Hg^{2+} and Cu^{2+} were potent inhibitors of enzyme activity, whereas other cations (Zn^{2+}, Mn^{2+}, Mg^{2+} and Ca^{2+}) had little or no effect on enzyme activity. The data on Hg^{2+} showed that this metal ion inhibited the enzyme activity completely at 1 mM $HgCl_2$ concentration. (Figure 3.40). When 2-mercaptoethanol (2-ME) was added to 25 mM final concentration to the enzyme previously exposed to Hg^{2+}, the activity did not recover but decreased comparative to the sample without 2-ME (Figure 3.40).

Table 3.10 Effect of metal ions on activity of purified rice BGlu1

Metal ions	Relative activity(%)[a]
None	100
$HgCl_2$	N.D.[b]
$CuSO_4$	32
$CaCl_2$	131
$MgSO_4$	121
$MnCl_2$	137
ZnSO4	108

[a] The enzyme was incubated with 10 mM final concentration of metal ions for 30 min. The enzyme was diluted in 50 mM NaOAc pH 5.0 and assayed with 1 mM pNPG according to the standard method. [b] means not detectable.

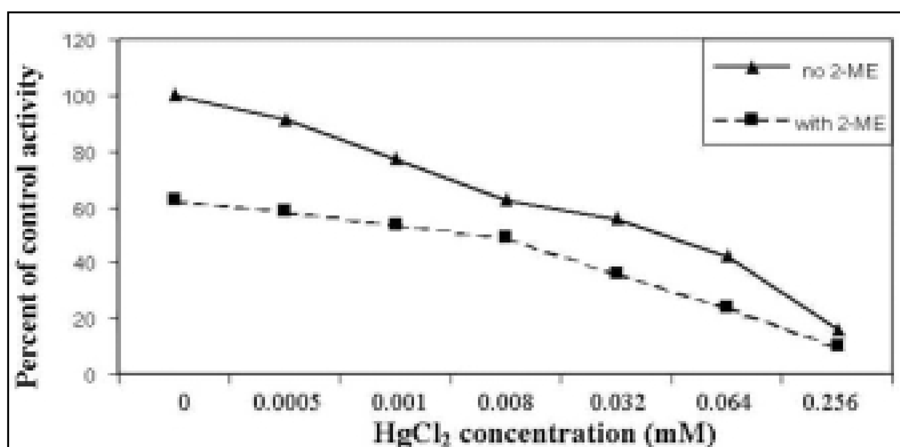


Figure 3.40 Inhibition of BGlu1 by Hg^{2+}. The enzyme was incubated in the presence of 0 to 256 μM $HgCl_2$ for 30 min and was then assayed for activity with 1 mM pNPG (pH 5.0) at 30°C for 10 min. After exposure to $HgCl_2$ for 30 min, half of each reaction solution was aliquoted and 25 mM 2-mercaptoethanol (2-ME) added and then assayed for activity with 1 mM pNPG (pH 5.0) at 30°C for 10 min.

3.6.5 Effect of reducing agents, denaturants and alcohol on enzyme activity

As shown in Figure 3.41, rice BGlu1 was lost activity in the presence of SDS. When the enzyme was assayed with 1 mM pNPG in the presence of 0-0.1% SDS, the nearly 100% loss of activity was observed at or above 0.04% SDS. The effect of 2-ME (0.064-2.048 M) on BGlu1 was also tested (Figure 3.42). The activity was lost up to 50% in the presence of 1 M 2-ME. At 0.064 M 2-ME, the activity was not significantly different from the control. In the presence of ethyl alcohol, rice BGlu1 activity decreased in proportional to the increasing of percent ethanol (Figure 3.43). At 30% ethanol concentration in the reaction assay, the enzyme activity was lost up to 50%. In all of these experiments, the enzyme was assayed with 1 mM pNPG at 30°C and pH 5.0 for 10 min.

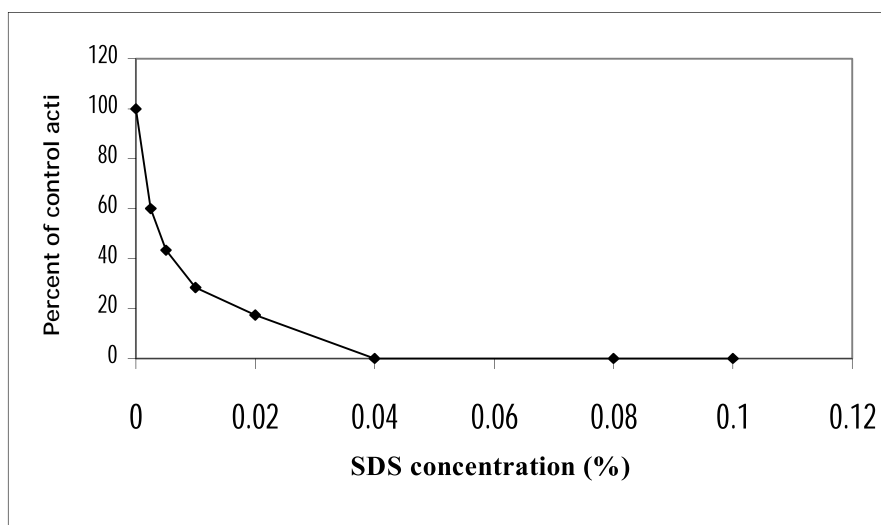


Figure 3.41 Activity profile of BGlu1 with 1 mM pNPG (pH 5.0 at 30°C for 10 min, respectively) in the presence of various concentrations of SDS.

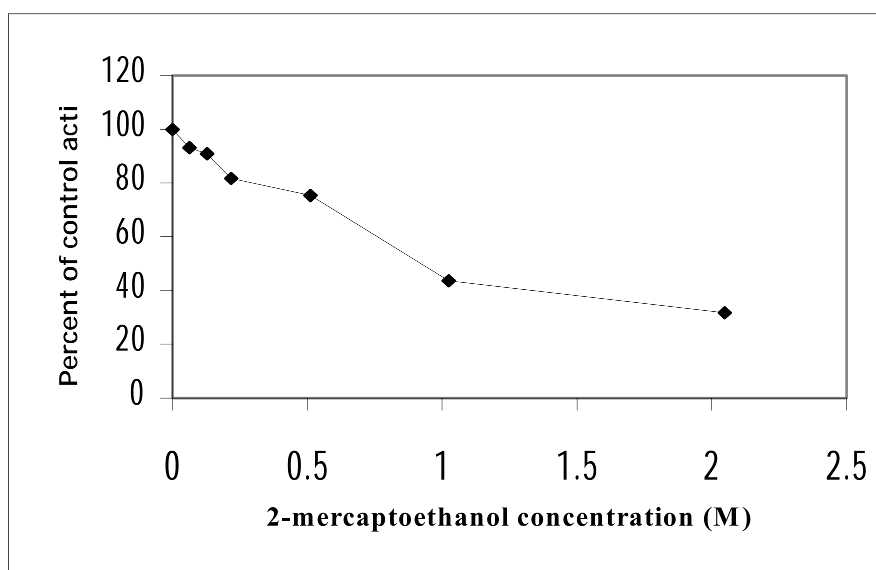


Figure 3.42 Activity profile of BGlu1 in the presence of various concentrations of 2-mercaptoethanol assayed with 1 mM pNPG (pH 5.0) at 30°C for 10 min.

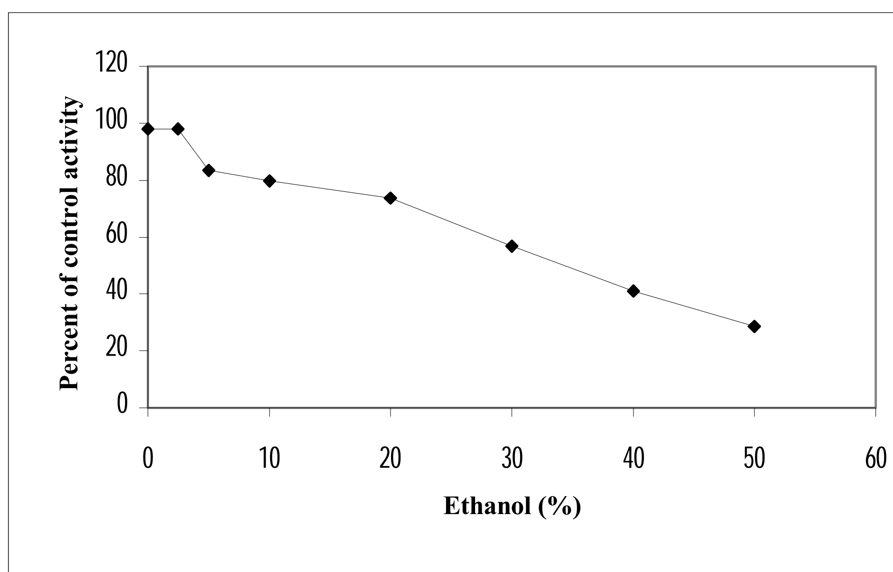


Figure 3.43 Activity profile of BGlu1 in the presence of various concentrations of ethyl alcohol assayed with 1 mM pNPG (pH 5.0) at 30°C for 10 min.

3.7 Transglycosylation activity of rice BGlu1

The preliminary study for the transglycosylation activity of rice β-glucosidase was performed using pNPG as glucose donor molecule and various glucose acceptor molecules. The purified rice β-glucosidase, BGlu1, catalyzed substrate transglycosylation with pNPG, which acts both as a glycosyl donor and acceptor. The substrate transglycosylation products analyzed by TLC demonstrated that a series of significant amounts of pNP derivatives with different hydrophobic properties and higher molecular weight were formed (Figure 3.44). When these pNP derivative products were compared with known standards from Sigma, one product had an R_f value equal to that of pNP β-D-cellobioside, and the other the same as pNP β-D-cellotrioside (Table 3.11). According to Fincher (1998), the R_f values of pNP-disaccharides higher and lower than pNP β-D-cellobiose were pNP β-D-laminaribiose, and pNP β-D-gentiobiose. At various substrate concentrations of pNPG, BGlu1 clearly had higher transglycosylation activity with the higher substrate concentrations, with respect to the major product spots on the TLC. The time course showed that the transglycosylation products visible on the TLC increased proportionally from 1 to 6 h and then remained constant until 24 h. The relative proportions of individual transglycosylation products are compared in Table 3.11. The pNP-disaccharides were the most abundant.

The transfer of a glycosyl group from pNPG to ethyl alcohol was also investigated. Significant amounts of product, ethyl glucoside, with lower hydrophobicity (R_f value lower than pNPG substrate) of the product was visible on the TLC. There was no significant difference in the amounts of transfer product between reactions containing 5%, 10%, and 20% ethyl alcohol, incubated at 3 h with 10 mM

pNPG, while there were little product produced in the presence of 30% ethyl alcohol (Figure 3.45 and Table 3.12).

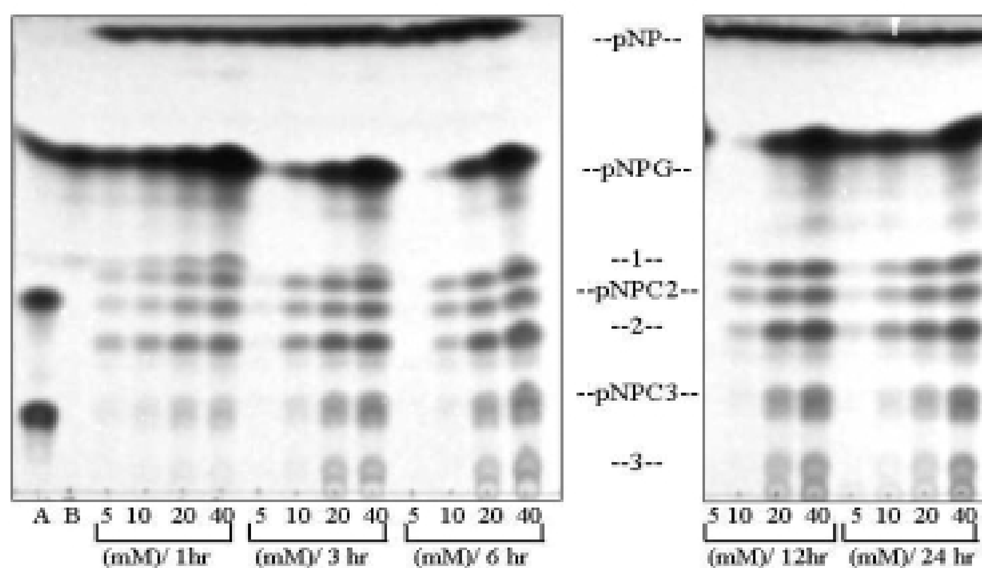


Figure 3.44 Transglucosylation of pNPG using rice β-glucosidase (0.01 nmol) incubated with various concentrations of pNPG in 50 mM NaOAc pH 5.0 buffer in 100 μl reactions at 30°C for 1 to 24 h. The numbers 5, 10, 20, 40 under each lane represent the concentration of pNPG as mM, respectively. The incubation time for each reaction sets were 1, 3, 6, 12 and 24 h. Reactions were analyzed by TLC on silica gel 60 F_{254}. Lane A, pNPG, pNPC2 and pNPC3 standard markers; lane B, blank without enzyme. pNP, p-nitrophenol; pNPG, pNP β-D-glucoside; pNPC2, pNP β-D-cellobioside ; pNPC3, pNP β-D-cellotrioside. 1, 2 and 3 represent unknown compounds.

Table 3.11 Yields of transfer products produced by BGlu1 in transglycosylation using pNPG as glucose donor and acceptor[a].

Compounds	Rf[b]	Percent of products obtained[c]
pNP-β-D-cellobiose	0.43	4.4
pNP-β-D-cellotriose	0.22	1.8
Product 1	0.48	5.2
Product 2	0.36	9.2
Product 3	0.07	0.5

[a] Enzyme (0.01 nmol) was incubated in 50 mM NaOAc pH 5.0 buffer containing 20 mM pNPG in 100 µl reaction for 6 h at 30°C. [b] Mobilities of each product are relative to pNP mobility. [c] Percent glycoside obtained represents the percent intensity of the spot relative to the total intensity present in the pNP, pNPG, and product spots.

Table 3.12 Yields of transfer products produced by BGlu1 in transglycosylation using pNPG as glucose donor and ethyl alcohol as glucose acceptor[a].

Compounds	Rf[b]	Percent of product obtained[c]			
		5%	10%	20%	30% ethyl alcohol
product 1 (ethyl glucoside)	0.34	24.2	46.8	50.7	0.8

[a] Enzyme (0.001 unit) was incubated in 50 mM NaOAc pH 5.0 buffer containing 10 mM pNPG and various percentage of ethyl alcohol for 6 h at 30°C. [b] Mobilities of each product are relative to pNP mobility. [c] Percent glycoside obtained represents the percent intensity of the spot relative to the total intensity present in the glucose, pNPG, and product spots.

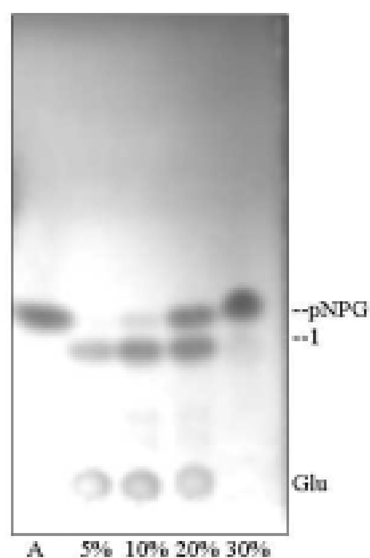


Figure 3.45 Transglucosylation of various concentrations of ethyl alcohol using rice β-glucosidase (0.01 nmol) incubated with various concentration of pNPG in 50 mM NaOAc pH 5.0 buffer at 30°C for 3 h in 100 μl reactions. The number 5%, 10%, 20%, 30% under each lane represents the concentration of ethyl alcohol in the reactions. Reactions were analyzed by TLC on silica gel 60 F_{254}. Lane A, blank without enzyme containing pNPG. pNPG, pNP β-D-glucoside; Glu, glucose; 1, transglycosylation products.

When pyridoxine (vitamin B_6) was used as the glucose acceptor in the transglycosylation reaction using pNPG as a glucose donor molecule, BGlu1 was able to transfer glucose from pNPG to pyridoxine. There were five major transglycosylation products visible on the TLC (Figure 3.46). Four of these spots were pNP derivatives, which had the same R_f values as the products in the above pNPG glucose acceptor reactions, and the R_f values of these products were the same as the products in the control reaction containing only pNPG and the enzyme. One

outstanding purple fluorescent spot under 356 nm wavelength UV light (in addition to the spot of pyridoxine) differed from the spot pattern of the control reaction. This purplish fluorescent spot had an R_f value equal to that of pyridoxine glucoside standard provided by Dr. G.F. Gregory. The tranglycosylation products were also analyzed by HPLC and the retension time of the fluorescent sensitive product was the same as pyridoxine glucoside standard (Figure 3.47).

The yield of pyridoxine glucoside product was optimized by varying enzyme concentration, time of incubation, concentration of pNPG glycosyl donor and pyridoxine acceptor. The products present on the TLC showed that pyridoxine glucoside amount increased in proportion to the concentrations of the pyridoxine and pNPG substrates (Figure 3.46). However, the yield of products tended to remain constant after incubation from 6 to 24 h. The yields of transfer products produced by BGlu1 in each reaction are compared in Table 3.13. Hydrolysis of transglucosylation products occurred at higher concentrations of enzyme and longer incubation times. Finally, suitable enzyme concentration (0.01 nmol) and incubation time (6 h), 20 mM pyridoxine and 20 mM pNPG in a 100 μl reaction were chosen to maximize the product yields.

Pyridoxine glucoside product was purified by reverse phase HPLC to greater than 95% purity and subjected to detailed structural investigations using ^{1}H and ^{13}C, COSY, HMBC and HMQC NMR techniques. The result showed that the product retrieved in the reaction was pyridoxine-5′-O-β-D-glucopyranoside as previously reported by Yasumoto *et al.* (1977) (Figure 3.48)

Table 3.13 Yields of transfer products produced by BGlu1 in transglycosylation using pNPG as glucose donor and pyridoxine as glucose acceptor[a].

Compounds	R_f [b]	Percent of products obtained [c]
Pyridoxine-5′-O-β-D-glucoside	0.30	14
pNP-β-D-cellobiose	0.44	4
Product 1	0.50	5
Product 2	0.36	10

[a] Enzyme (0.01 nmol) was incubated in 50 mM NaOAc pH 5.0 buffer containing 20 mM pNPG and 20 mM pyridoxine for 6 h at 30°C. [b] Mobilities of each product are relative to pNP mobility. [c] Percent glycoside obtained represents the percent intensity of the spot relative to the total intensity present in the pNP, pNPG, pyridoxine and product spots.

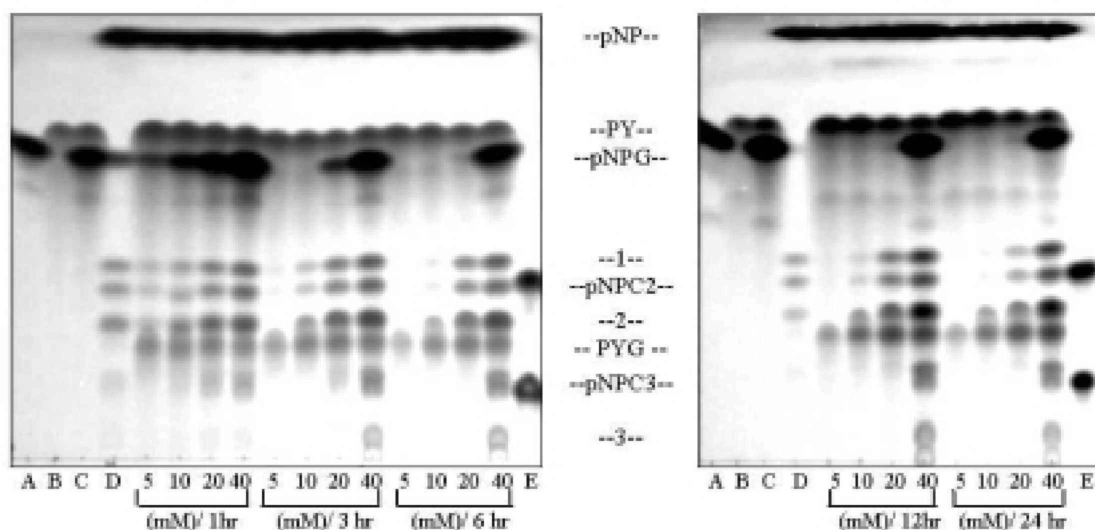


Figure 3.46 Transglucosylation of 20 mM pyridxine using rice β-glucosidase (0.01 nmol) incubated with various concentrations of pNPG in 50 mM NaOAc pH 5.0 buffer in 100 μl reactions at 30°C for 1 to 24 h. The number 5, 10, 20, 40 under each lane represent the concentration of pNPG as mM, respectively. The incubation times for each reaction set were 1, 3, 6, 12 and 24 h. Reactions were analyzed by TLC on silica gel 60 F_{254}. Lane A, pNPG standard marker; lane B, pyridoxine standard marker; lane C, blank without enzyme; lane D, control containing enzyme and 10 mM pNPG; lane E, pNP β-D-cellobiose and pNP β-D-cellotriose markers. pNP, p-nitrophenol; PY, pyridoxine; PYG, pyridoxine-5′-O-β-D-glucoside; pNPG, pNP β-D-glucoside; pNPC2, pNP β-D-cellobiooside; pNPC3, pNP β-D-cellotrioside. 1, 2 and 3 represent unknown compounds.

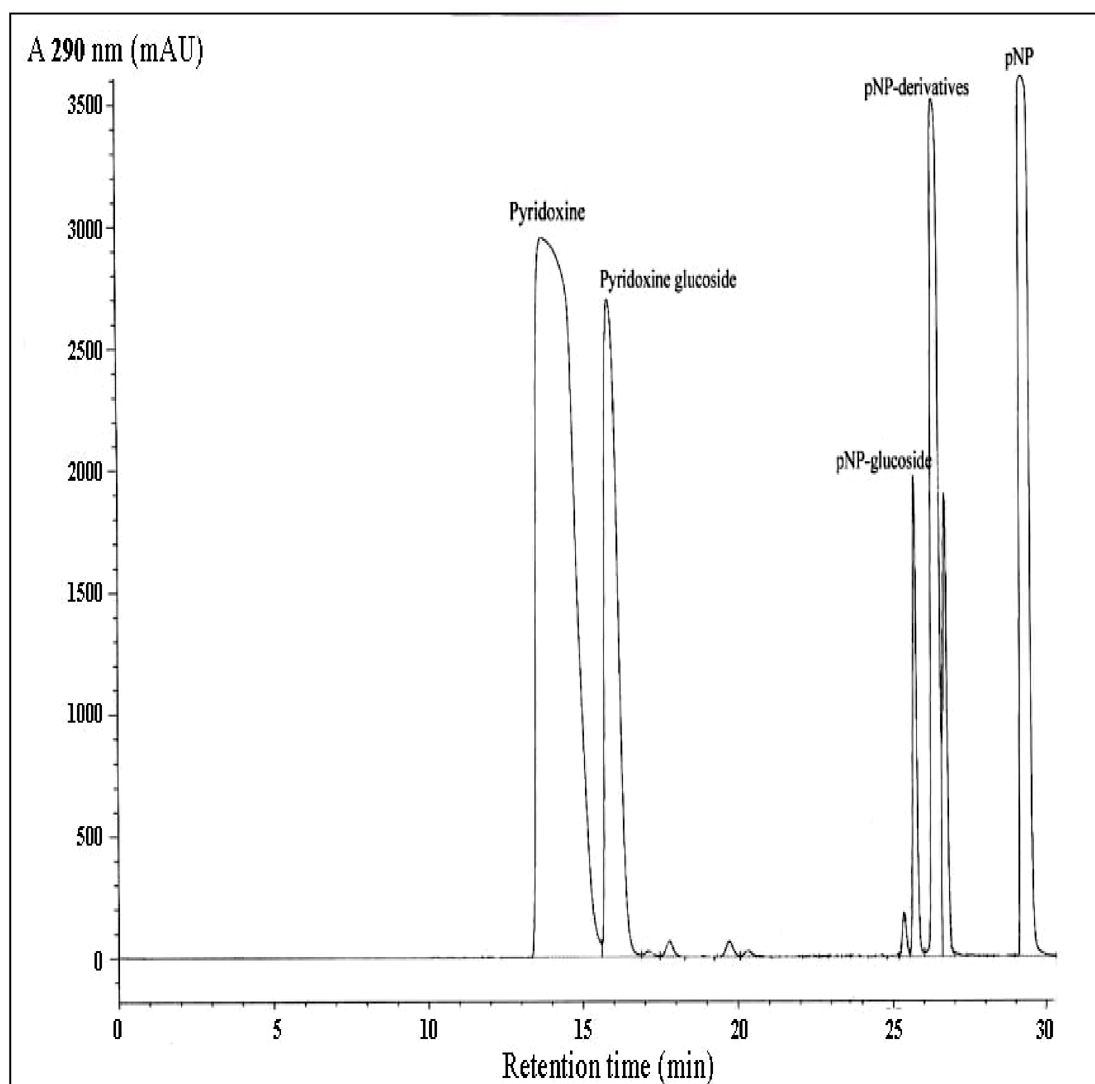


Figure 3.47 HPLC elution profiles of the transglycosylation products of pyridoxine using rice BGlu1. The enzyme (0.05 nmol) was incubated with 20 mM pyridoxine and 20 mM pNPG in 50 mM NaOAc buffer pH 5.0 in 500 μl reaction at 30°C for 3h. The reaction products were subjected to HPLC analysis.

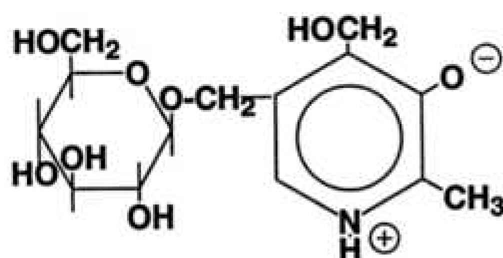


Figure 3.48 Structure of pyridoxine-5'-O-β-D-glucoside (Yasumoto *et al*., 1977)

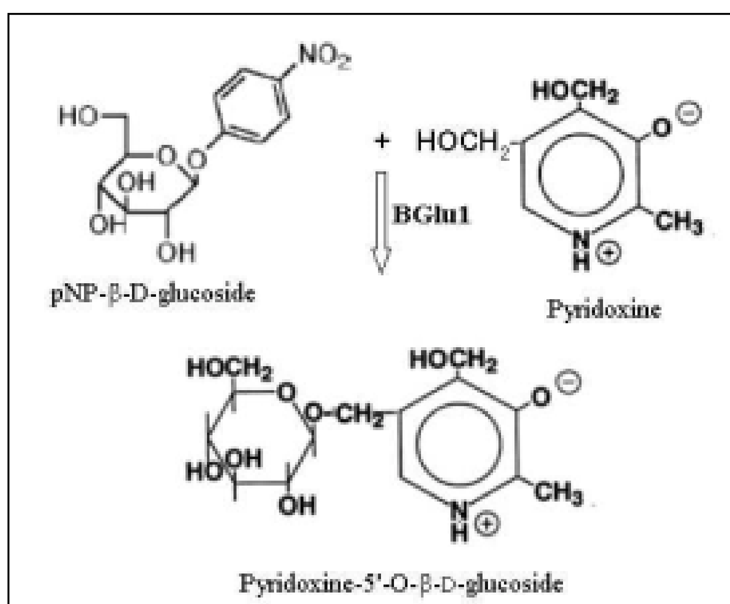


Figure 3.49 Scheme of the transfer of glucose from pNPG to pyridoxine to form pyridoxine-5'-O-β-D-glucoside by rice BGlu1.

3.9 Inhibition of rice germination by a β-D-glucosidase inhibitor

The effect of glucono-1,5-lactone, the β-D-glucosidase inhibitor, on rice germination was studied by soaking and germinating them in the presence of this inhibitor. The seeds were soaked and germinated in the absence and presence of various concentrations (0, 25, 50, 100 mM) of glucono-1,5-lactone. The results shown in Figure 3.51 and 3.52 clearly showed that the inhibition of rice germination by this inhibitor is concentration dependent. In each germinating day (from 1-4 days), the number of germinated seeds and also the length of their shoot or root were significantly different at each gluconolactone concentration. The control exhibited full and normal growth of rice seedlings (Figure 3.50). The results showed that percent of germination and lengths of seedling parts tended to decrease in the presence of the increasing concentration of gluconolactone (from 25 mM 50 mM to 100 mM). Note that at 25 and 50 mM δ-gluconolactone, the seeds still germinated but the lengthes of the shoots and roots were much shorter than the control in each germinating date. At 100 mM gluconolactone concentration, there was no germination of rice seeds. Note that the percent of germination in normal growth in this experiment is approximately 70%.

In each day of germination in the presence of various concentration of gluconolactone the seeds or germinating seeds were brought to check for the occurence of β-glucosidase activity in rice seed or tissues with X-glucoside substrate by staining assay. Figure 3.51 showed the appearance of blue color in rice tissues in all control samples, but the color in gluconolactone samples tended to decrease with length of germinating time and increase of inhibitor concentration.

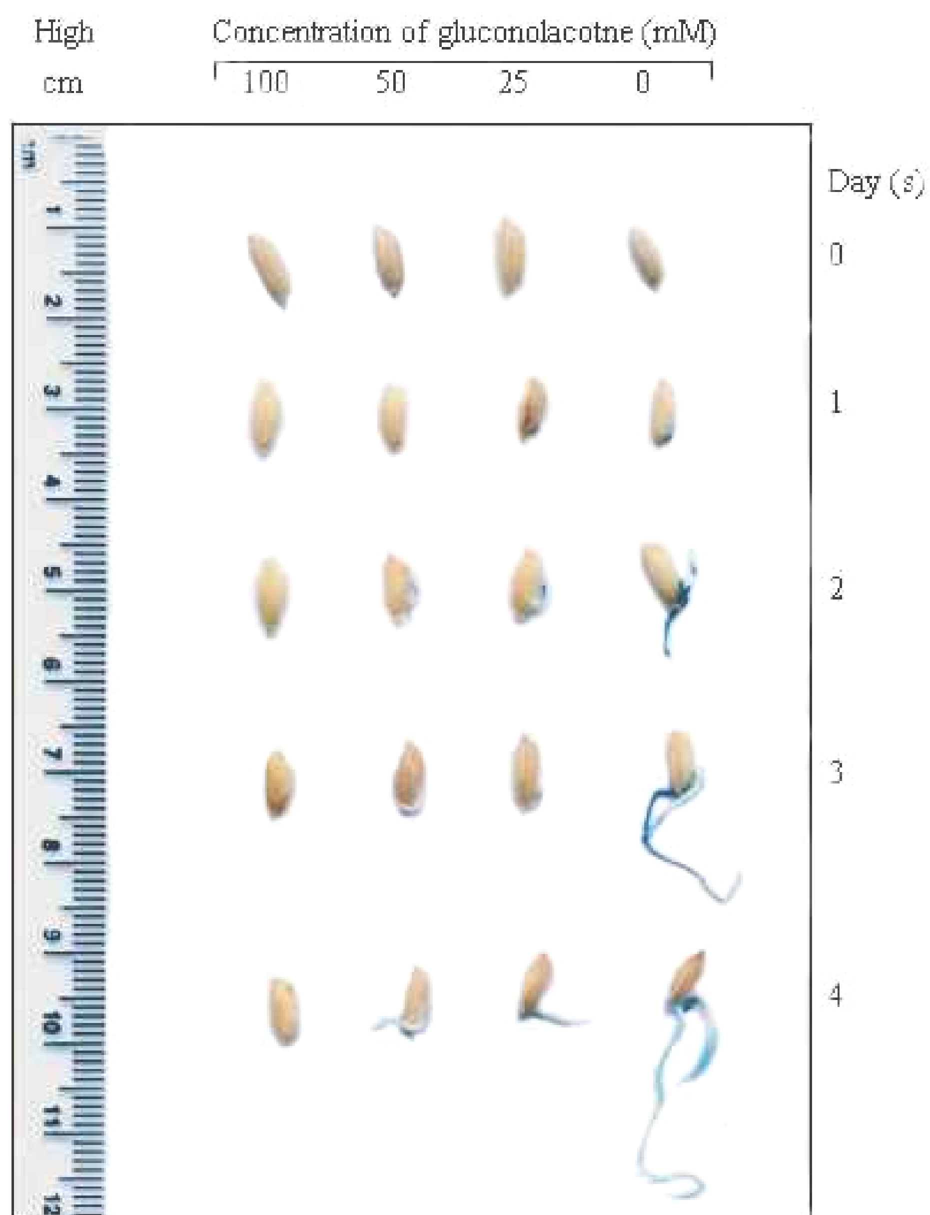


Figure 3.50 Effect of glucono-1,5-lactone (0-100 mM) on 0 to 4-day-old rice germination, growth and development.

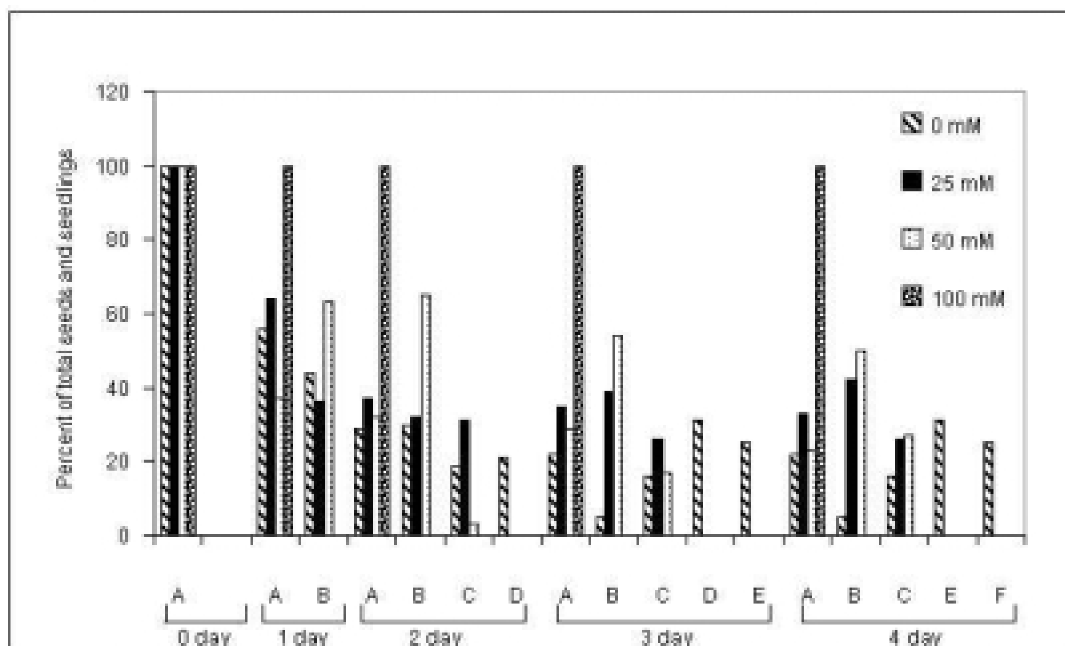


Figure 3.51 Percent of rice seeds and seedlings (at various stages) grown in the presence of 0-100 mM glucono-1,5-lactone from day 0 today 4. Rice seedlings at various developmental stages are separated respective to the length of roots. A, B, C, D, E, and F represent the lenght of roots. A, 0 cm; B; 0.1-0.2 cm; C, 0.3-0.5 cm; D, 1 cm; E, 2-3 cm; F, 4 cm. Note that at 25 and 50 mM δ-gluconolactone, the seeds still germinated but the length of shoot and root were much shorter than the control (0 mM) at each germination day. There were no germinated seeds in the presence of 100 mM δ-gluconolactone.

3.10 Identification of rice β-glucosidase genes

Rice genes homologous to glycosyl hydrolase family 1 β-glucosidase have been identified from Beijing Genomic Institute (BGI), GenBank at NCBI, Monsanto Rice Genome Sequencing Project (MST), Torrey Mesa Research Institute (TMRI) database. The retrieved gene sequences were searched against the EST database to scan the expression of each gene in rice. Forty-nine β-glucosidase genes (including the genes that matched *bglu1* and *bglu2*), including 32 full-length genes, 10-nearly full-length genes, and 7 intronless genes, were identified, as listed in Table 3.14. Several rice β-glucosidase genes were mapped to chromosomes 1, 4, 7, 8, 10, and 12 (Table 3.14).

The open reading frames (ORFs) of forty-five gene-derived cDNAs showed significant level of shared deduced amino acid sequence identity to each other and other known β-glucosidase sequences (see Appendix A). All deduced β-glucosidase protein sequences contained the putative catalytic acid/base and nucleophilic glutamate residues, except CLl029716.6 and contig16468 in which the nucleophilic glutamate was replaced with glutamine. The catalytic acid/base and nucleophile consensus sequences were: W-x-T/I/V-F/L/IV/S/M- N/AL/I/D/G/H-E-P/I/Q and V/I/L/P-x-E-N-G-x, (x can be various residues), respectively, similar to the consensus sequences previously derived from known β-glucosidase sequences (Czjzek *et al.*, 2000; PROSITE database, Hoffman *et al.*, 1999). The phylogenetic tree derived from this alignment shows that rice β-glucosidases clustered into groups, some of which were more closely related to proteins from other plants than to each other. For example, the *bglu1* OSM129364, OSM125591, and OSM11391 β-glucosidases

clearly grouped with barley BGQ60. All cDNAs also had unique 3'UTR sequences, which may therefore be used as unique probes for each gene. Almost all β-glucosidase ORFs were predicted to have a signal sequence targeting them to the secretory pathway (Table 3.15). Though assignment of cellular location was generally unclear using the PSORT program, none of the proteins are predicted to localize to the chloroplast, unlike the other grass β-glucosidases characterized to date (maize, oat, sorghum and rye), except for barley BGQ60 (Esen, 1992; Cicek and Esen, 1998a; Sue *et al*., 2000; Nisius, 1988; Leah *et al*., 1995). The deduced proteins were also analyzed for predicted molecular mass, pI, and potential N-linked glycosylation sites (Table 3.15).

Intron-exon boundaries and intron numbers are highly conserved among rice and other plant β-glucosidase genes, but some introns are missing in some genes. There are 5 patterns of gene organization in rice with either 13, 11, 12, 9, or 1 exons (Figure 3.52). Group one rice β-glucosidase genes, those with 13 exons like the cassava linamarase gene (AC U95258), included those with genome contig numbers OSM11258, 118774, 1731, 1729, contig 22112, 3224 etc. The exons 1 and 2 and 12 and 13 were combined to generate an 11 exon/10 intron gene pattern in group two, which included OSM13853 (*bglu1*), 11391, 125591, 129364, and 118061. In group three, exons 9 and 10 were combined to generate a 12 exon/11 intron pattern, which included contigs 16091, 20212, 54140, 4293 etc. In group four, group one exons 7, 8, 9 and 10 are combined, as are exons 12 and 13, to generate the 9 exon/8 intron pattern seen in OSM11159 (*bglu2*), and 11161. And group five without intron was seen in 7 genes identified in BGI sequence database. However in each case, remaining introns maintained the same splice sites. Intron sizes in these genes, however, are highly

variable. In most cases, very long introns contained retro-transposon-like sequences, while the orthologous short introns did not. The results from deduced amino acid sequence alignment and phylogenetic analysis (Figure 3.53) showed that the sequences in intron-exon pattern groups 2, 3, 4 and 5 (and some in group 1) are more closely related to each other within their groups than to the other groups.

In order to begin to analyze the expression of the β-glucosidase genes in rice, a search for ESTs corresponding to each of 49 different predicted genes was performed. As shown in Table 3.16, an initial homology search with β-glucosidase sequences, identified 143 ESTs and it was shown that 26 genes have corresponding ESTs. The *bglu1* gene was most highly represented in the dbEST database, with 60 matching ESTs. The source libraries for rice β-glucosidase gene matching ESTs included seedling (shoot and root), immature plant parts (stem, root, leaf), mature leaf, panicle at flowering and ripening stage, immature seeds, callus, reflecting its expression in many tissues during plant development. Some rice β-glucosidase genes have ESTs represented in stressed plant tissue libraries, including salt, drought, cold, and Blast/fungus infection. Because of the small number of ESTs identified for β-glucosidases other than *bglu1* and *bglu2*, no conclusions could be drawn on their expression patterns. Several genes were not represented in the EST database (i e. OSM 1731, 125591, 127811), so whether and where they are expressed remains unclear.

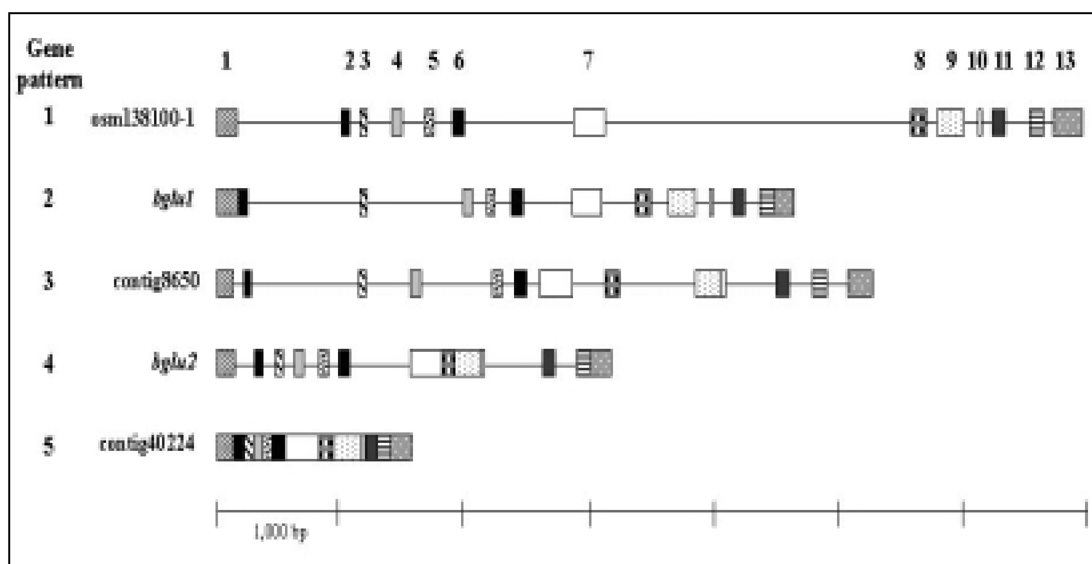


Figure 3.52 Predicted gene structure patterns for putative rice β-glucosidase genes. Exons are shown as boxes with corresponding exons having the same pattern. Introns, represented as simple lines, are drawn in proportion to their length. Note that 5 gene organization patterns can be seen in rice genes, those with 13, 12, 11, 9 exons and intronless patterns with the splice sites conserved in each group and between groups for common exons and introns.

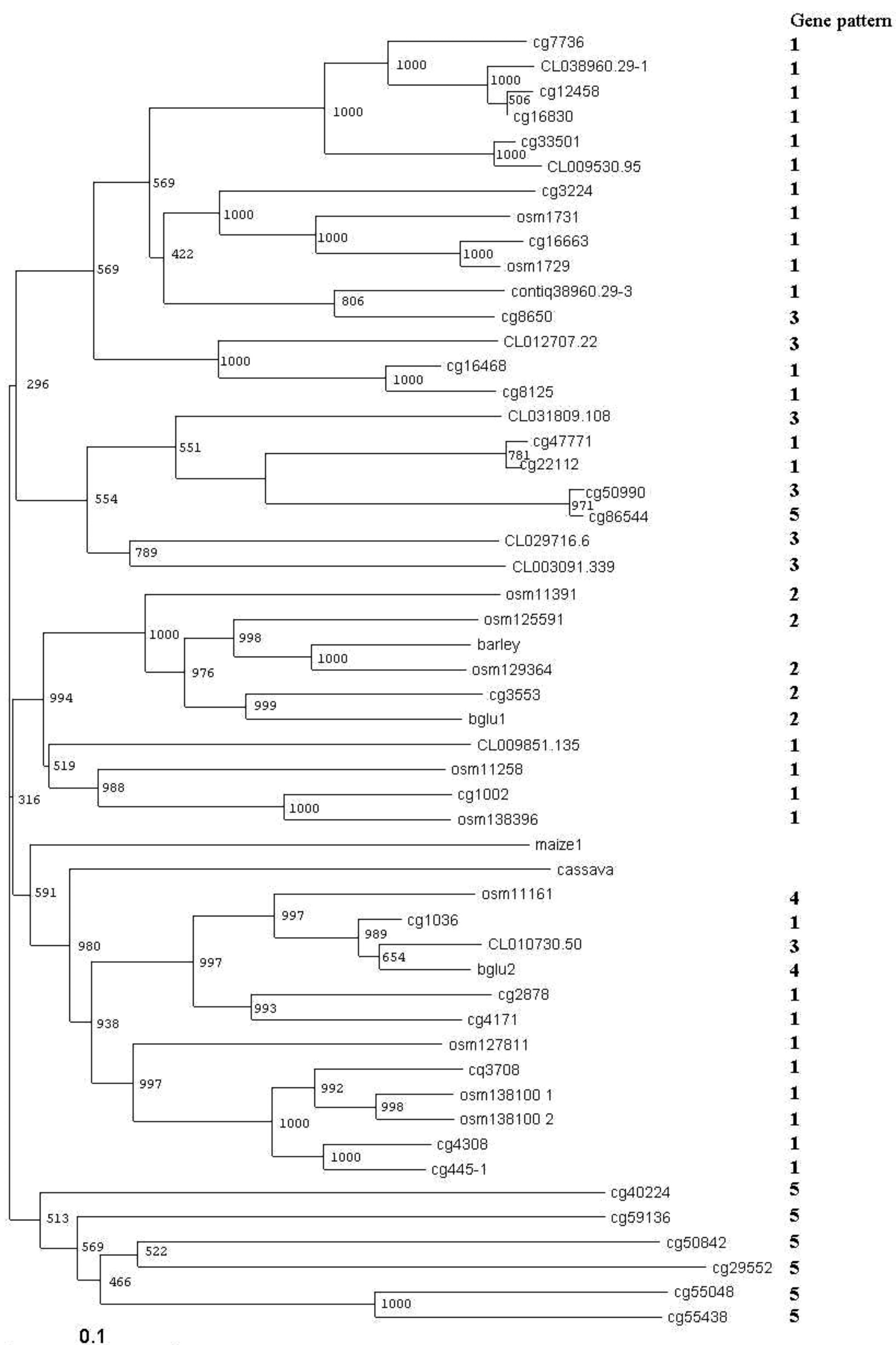

Gene pattern
cg7736 1
CL038960.29-1 1
cg12458 1
cg16830 1
cg33501 1
CL009530.95 1
cg3224 1
osm1731 1
cg16663 1
osm1729 1
contiq38960.29-3 1
cg8650 3
CL012707.22 3
cg16468 1
cg8125 1
CL031809.108 3
cg47771 1
cg22112 1
cg50990 3
cg86544 5
CL029716.6 3
CL003091.339 3
osm11391 2
osm125591 2
barley
osm129364 2
cg3553 2
bglu1 2
CL009851.135 1
osm11258 1
cg1002 1
osm138396 1
maize1
cassava
osm11161 4
cg1036 1
CL010730.50 3
bglu2 4
cg2878 1
cg4171 1
osm127811 1
cq3708 1
osm138100 1 1
osm138100 2 1
cg4308 1
cg445-1 1
cg40224 5
cg59136 5
cg50842 5
cg29552 5
cg55048 5
cg55438 5
0.1

Figure 3.53 Unrooted phylogenetic tree of rice β-glucosidase gene family together with other plant β-glucosidases. The tree was derived by the Neighbor joining method from the alignment in Appendix A and tested by bootstrapping. The bootstrap values for each node are shown. Note that the gene pattern refers to the gene organization patterns showed in Figure 3.52.

Table 3.14 Summary of identified genes homologous to glycosyl hydrolase family 1 glucosidase from the Beijing Genomic Institute (BGI), GenBank at NCBI of rice CGI (CGI/NCBI), the Monsanto Rice Genome Sequencing (MST) Project, and the Torrey Mesa Research Institute (TMRI) database

No	BGI Contig[1]	CGI/NCBI AAAA[2]	MST OSM[3]	TMRI	RGP GenBank at NCBI AC[4]	Full ?	Chro#
1	1036 (ex 1-7)	01001103	-	-	-	no	-
2	12173 (full)	01008811.1	-	CL038960.29-3 (full)	-	yes	-
3	12458 (ex 4-13)	01013412	-	-	-	no	-
4	16091 (ex 3-12)	01015706	-	CL003091.339 (ex 3-13)	OSJN00108	no	4
5	16468 (ex 1-9)	01015943	12635 (ex 1-10, 13)	CLC16907 (ex 1-10, 13)	-	yes[a]	-
6	16663 (ex 5-13)	01016051	-	-	-	no	-
7	16830, 5918-2 (ex 1-11)	01016177	-	-	-	no	-
8	17843 (ex 1-12)	01003642	-	CL009851.135 (full)	AP003349 (BAC66894 [b]) AP003418 (BAB86071 [b])	yes	1
9	20212 (ex 5-13)	01018204	-	CL031809.108 (full)	-	yes	-
10	22112 (ex 1-8)	01018346	-	CL019196.174 (ex 1, 3-8)	-	no	-
11	24140 (ex 1-11)	01011188	115156 (full)	CL029716.6 (full)	OSJN00059 No annotate seq	yes	4
12	2878 (ex 1-7)	01004189	-	-	AP006049(full) (BAC57391 [b])	yes	8
13	29552[c] (full)	01022924	-	-	-	yes	-
14	3224 (full)	01005593	-	CL028390.101 (full)	AP004330 (BAB90767 [b]) AP003272 (BAB67933 [b])	yes	1
15	33501 (ex 2-7), 5918-1 (ex 7-13)	01025681, 01008246	117680 (ex 6-13)	-	-	no	-
16	3708 (full)	01001056	-	CL002316.105 (ex 2-13)	-	yes	-
17	40224 [c]	01028472	-	-	-	yes	-
18	4171 (full)	01006396	-	CL004087.285 (ex 7-13)	AP006049 (BAC57391 [b])	yes	8
19	4293 (ex 2-13)	01003080	-	CL012707.22 (full)	-	yes	-

Table 3.14 (Continued)

No	BGI Contig[1]	CGI/NCBI AAAA[2]	MST OSM[3]	TMRI	RGP GenBank at NCBI AC[4]	Full ?	Chro#
20	4308 (full)	01003735	-	CL004649.229 (full)	-	yes	-
21	445-1 (full)	01001963	-	CL028257.55-1 (full)	-	yes	-
22	47771 (ex 9-13)	01008811	-	-	-	no	-
23	50842[c] (ex 1-9)	01036508	-	-	-	no	-
24	50990 (ex 9-13)	01032141	-	-	-	no	-
25	55048 [c] (full)	01037636	-	-	-	yes	-
26	55438[c] (ex 1-11)	01038534	-	-	-	no	-
27	59136[c] (ex 1-8)	01044943	-	-	-	no	-
28	7736 (full)	01011001	-	CL038960.29-2 (full)	-	yes	-
29	8125 (full)	01001106	115071 (ex 1-4)	CL018856.38 (ex 3-13)	-	yes	-
30	8650 (full)	01010730	19464 (ex 4-9)	CL030301.60 (ex 4-13)	-	yes	-
31	86544[c] (ex 1-7)	01065025	-	-	-	no	-
32	8974 (full)	01011181	11159 (full) (BGlu2)	CL002761.161 (full)	-	yes	9 [d]
33	4652 (full)	01007459	11161 (full)	HIC120900.B0 1.1.1 (full)	-	yes	9 [d]
34	146 (full)	01001022	11258 (full)	CL032728.120 (full)	-	yes	-
35	8992 (full)	01003736	11391 (full)	CL001465.454 (full)	CNS08C94 CNS08C89	yes	12
36	3553 (full)	01001283	118061 (ex 1-9)	CL02035.62 (full)	-	yes	-
37	1002 (full)	01001470	118774 (ex 2-13)	CLC10115 (full)	AC074354 (AAK92581[b])	yes	10
38	13072, 32318 (full)	01002424	125591 (full)	CL011680.129 (full)	AP003217 (BAC06877[b]) AP003435 (BAB68015 [b])	yes	1
39	1428 (full)	01000828	127811 (full)	CL022951.99 (full)	-	yes	-
40	24898, 14029 (full)	01002930	129364 (full)	CL021113.23 (full)	AP005184 (BAC16123 [b]), AP005182 (BAC208021 [b])	yes	7
41	445-2 (full)	-	138100-1 (full)	CL028257.55-2 (full)	-	yes	-
42	445-3 (full)	-	138100-2 (full)	CL028257.55-3 (full)	-	yes	-

Table 3.14 (Continued)

No	BGI contig	CGI/NCBI AAAA[2]	MST OSM[3]	TMRI	RGP GenBank at NCBI AC[4]	Full ?	Chro#
43	21522, 43796 (full)	01018976, 01001283	13853 (full) (BGlu1)	CL011484.96 (full)	-	yes	3[d]
44	4437, 32370 (full)	01005003, 0101827	139385, 139386 (full)	CL037869.83 (full)	-	yes	-
45	-	0101605	1729 (ex 1-8)	CL007813.109 (full)	AP003570(full) (BAB63224[b])	yes	1
46	7356 (full)	01009748	1731 (full)	CL007822.79 (full)	AP003570(full) BAB63225	yes	1
47		01035250 (full)	-	CL038960.29-1 (full)	-	yes	-
48		-		CL010730.50 (ex 1-11)	-	no	-
49		0102568.1 (full)		CL009530.95 (full)	-	yes	-

Full? means full-length or not. Chro# means chromosomal location (chromosome number) of the genes. Ex means exon. [1] contig number, [2] GenBank accession number 'AAAA..' for BGI rice genome templates, [3] Osm number, [4] GenBank accession number, [a] the gene does not contain exon 11 and 12, [b] annotated deduced β-glucosidase in GenBank, [c] intronless gene, [d] chromosome location was determined by mapping of corresponding ESTs on chromosomes by the Rice Genome Project (Japan).

Table 3.15 Predicted rice β-glucsidase protein properties and the gene patterns

No	Clone	aa length	Mr (kDa)	Signal peptide	pI	N-Link	Target	Gene pattern[a]
1	contig 1036	254	-	21	-	2	vacuole/outside/ microbody/	1
2	contig 12173	533	59.5	26	4.96	5	outside/ vacuole/ microbody	1
3	contig 12458	428	-	-	-	-	-	1
4	contig 16091	446	-	-	-	-	-	3
5	contig 16468 [b]	365	-	-	-	1	outside/ microbody	5
6	contig 16663	457	-	-	-	-	-	1
7	contig 16830, contig 5918-2	412	-	30	-	4	outside/ microbody	1
8	contig 17843	479	54.8	0	5.15	0	microbody/ cytoplasm/ mitochondria space	1
9	contig 20212	516	58.7	26	6.58	4	outside/ microbody	3
10	contig 22112	291	-	27	-	1	outside/microbody	1
11	contig 24140	499	56.4	23	6.96	6	outside/microbody	3
12	contig 2878	497	56.1	18	8.53	6	Mitochondria membrane/vacuole/ outside	1
13	contig 29552[c]	458	53.0	0	5.19	0	microbody/ cytoplasm/ mitochondria space	5
14	contig 3224	528	59.2	37	5.53	3	outside/ microbody	1
15	contig 33501, contig 5918-1	469	-	-	-	-	-	1
16	contig 3708	537	60.8	51	7.60	4	outside/ vacuole/ microbody	1
17	contig 40224 [c]	567	63.6	24	5.71	0	Plasma membrane/ ER membrane/ microbody	5
18	contig 4171	499	56.8	17	8.53	5	outside/ microbody	1
19	contig 4293	510	57.1	24	5.60	2	outside/ microbody	3
20	contig 4308	520	58.7	39	6.68	6	outside/ vacuole/microbody	1
21	contig 445-1	510	57.4	24	8.84	6	outside/ vacuole/ microbody	1
22	contig 47771	219	-	-	-	-	-	1
23	contig 50842 [c]	321	-	0	-	0	microbody/ cytoplasm	5
24	contig 50990	222	-	-	-	-	-	3
25	Contig 55048 [c]	425	48.0	0	5.37	0	microbody/ mitochondria space	5
26	Contig 55438 [c]	365	-	0	-	0	Microbody/ mitochondria space	5
27	Contig 59136 [c]	280	-	0	-	0	Microbody/ cytoplasm/ mitochondria space	5
28	Contig 7736	526	59.1	31	5.79	4	outside/ vacuole/ microbody	1
29	Contig 8125	510	57.0	30	5.58	2	outside/ microbody /vacuole	1

Table 3.15 (Continued)

No	Clone	aa length	Mr (kDa)	Signal peptide	pI	N-Link	Target	Gene pattern[a]
30	Contig 8650	523	58.5	23	5.15	3	outside/ vacuole/ microbody	3
31	Contig 86544[c]	181	-	0	-	0	microody/ cytoplasm/ mitochondria space	5
32	osm 11159 (BGlu2)	500	57.4	25	8.05	6	outside/ vacuole/ microbody	4
33	osm 11161	508	57.7	27	8.73	4	outside/ microbody	4
34	osm 11258	501	57.2	16	5.50	2	outside/ microbody	1
35	osm 11391	492	57.0	19	7.32	5	outside/ microbody	2
36	osm 118061	568	63.1	33	6.68	0	plasma membrane/ microbody	2
37	osm 118774	510	58.0	23	6.20	4	outside/ microbody	1
38	osm 125591	516	58.0	21	8.35	5	outside/ microbody	2
39	osm 127811	504	57.7	18	7.18	5	outside/ microbody	1
40	osm 129364	510	58.5	0	7.2	0	mitochondria inner membrane/ microbody/ plasma membrane/ mitochondria space	2
41	osm 138100-1	530	59.9	25	7.59	5	outside/ vacuole/ microbody	1
42	osm 138100-2	510	58.1	23	7.99	4	vacuole/ outside/ microbody	1
43	osm 13853 (BGlu1)	504	56.9	28	9.0	3	outside/ microbody	2
44	osm 139385, osm 139386	521	58.5	38	6.9	3	outside/ microbody	1
45	osm 1729	500	56.1	20	5.34	2	outside/ microbody / plasma membrane	1
46	osm 1731	514	57.5	22	7.19	3	outside/ microbody	1
47	CL038960.29-1	559	63.1	0	6.22	0	ER membrane/ microbody	1
48	CL010730.50	411	-	21	-	3	outside/ microbody / vacuole	3
49	CL009530.95	531	59.5	41	5.07	6	outside/ microbody	1

The cDNA nucleotide sequences, derived from exon mapping using BLASTx and identification of splice site consensus sequences were translated to protein sequences and analyzed by the proteomics tools at the Expasy proteomics site: http://www.expasy.ch, and PSORT. 'aa' means amino acid. Mr means molecular weight. N-link means number of N-glycosylation site. Contig clones are Beijing Genome Institute clone number, osm clones are Monsanto Rice Genome Project clone

number and CL.clones are Torrey Mesa Research Institute genome database clone numbers. [a] means gene organization pattern according to Figure 3.52. [b] the gene does not contain exons 11 and 12. [c] intronless gene.

Table 3.16 Expression of rice β-glucosidase genes judged from corresponding ESTs

No	Clone	Full ?	# EST match	GenBank accesssion number of ESTs	Tissue library
1	Contig 1036	no	0		
2	Contig 12173	yes	9	AU162707, AU096577, AU163618, AU056687, BI807518, BM420558, BQ907495, D48232, C72756,	Shoot, mature leaf, mature leaf induced by *M. grisea*, panicle (flowering stage)
3	Contig 12458	no	0		
4	Contig 16091	no	0		
5	Contig 16468 [a]	yes	0		
6	Contig 16663	no	1	BI306425	Drought stress leaf
7	Contig 16830, contig 5918-2	no	0		
8	Contig 17843	yes	4	CA762923, C98907, C98908, D41563	Shoot, drought stress panicle (flowering stage), panicle (flowering stage)
9	Contig 20212	yes	2	BI812350, CA755317	Salt stress root (3-4 wk old), mature leaf induced by *M. grisea*
10	Contig 22112	no	1	AU056344	Mature leaf
11	Contig 24140	yes	0		
12	Contig 2878	yes	1	D41997	Shoot
13	Contig 29552 [b]	yes	0		
14	Contig 3224	yes	1	BI805927	Young stem
15	Contig 33501, contig 5918-1	no	5	AU085894, CA762922, CA764111, CB096565, C72798,	Drought stress panicle (flowering stage), panicle (flowering stage)
16	Contig 3708	yes	5	AU183254, AU081347, AU081344, AU06214, CA755751	Salt stress root (3-4 wk old), immature leaf
17	Contig 40224 [b]	yes	0		
18	contig 4171	yes	2	AU089744, CA999441	Shoot, cold stress seedlings
19	contig 4293	yes	7	AU091520, AU174728, AA751550, AU092973, C19313, C28471, C28603,	Callus, leaf, panicle (ripening stage), immature seed
20	contig 4308	yes	5	AU163237, AU096011, BI811510, BI811418, D46905	Shoot, mature leaf induced by *M. grisea*,
21	contig 445-1	yes	7	BG101702, BU572330, CA759481, CA759482, CA763998, CA998998, CB000178	Shoot, drought stress panicle (flowering stage), cold stress seedlings
22	contig 47771	no	2	AU096577, BQ907495	Shoot, mature leaf induced by *M. grisea*,
23	contig 50842 [b]	no	0		
24	contig 50990	no	0		
25	contig 55048 [b]	yes	0		
26	contig 55438 [b]	no	0		
27	contig 59136 [b]	no	0		
28	contig 7736	yes	0		

Table 3.16 (Continued)

No	Clone	Full ?	#EST match	GenBank accesssion number of ESTs	Tissue library
29	Contig 8125	yes	5	AA751232, BI118724, BI812832, BQ909161, CA767064	Immature seed, leaf infected Blast/ Fungus, drought stress panicle, mature leaf induced by *M. grisea*,
30	contig 8650	yes	0		
31	contig 86544 [b]	no	0		
32	osm 11159 (BGlu2)	yes	8	AU033015, AU058282, AU065170, AU065204, AU065193, BI806911, D39581, D39589	Shoot, young stem, immature leaf
33	osm 11161	yes	2	AU032044, D28326	Seedling root
34	osm 11258	yes	0		
35	osm 11391	yes	3	AU182378, AU182379, CA765065	Drought stress panicle, panicle (flowering stage)
36	osm 118061	yes	2	CA998259, CA998261	Cold stress seedlings
37	osm 118774	yes	0		
38	osm 125591	yes	0		
39	osm 127811	yes	0		
40	osm 129364	yes	1	C19278	Panicle (ripening stage)
41	osm 138100-1	yes	2	CA999014, CA999016	Cold stress seedlings
42	osm 138100-2	yes	1	AU066064	Shoot
43	osm 13853 (BGlu1)	yes	60	CB000637, CA999001, CA999930, AU097444, AU33138,AU222984, AU033196, D46605, C25358, AU162859, D46619, etc.	Seedling root, shoot, callus, young leaf, cold stress seedlings, young root, young stem, young panicle, mature leaf
44	osm 139385, osm 139386	yes	2	AU070534, D41838	Shoot
45	osm1729	yes	1	BI306425	Drought stress leaf
46	osm1731	yes	0		
47	CL038960.29-1	yes	0		
48	CL010730.50	no	0		
49	CL009530.95	yes	4	AU085894, CA762922, CA764111, CB096565, C72798,	Drought stress panicle, panicle (flowering stage)

EST sequences were retrieved from the dbEST section of GenBank (http://www.ncbi. nlm.nih.gov/) by BLASTn search with full gene sequences. They were inspected to ensure they matched the gene-coding region over and their full files retrieved to determine cDNA library source tissue and clone number when necessary. The ESTs assigned to each gene had greater than 97% identity and no higher similarity with

another gene. Full? means full-length ans #EST match means number of EST match. Contig clones are Beijing Genome Institute clone number, osm clones are Monsanto Rice Genome Project clone number and CL.clones are Torrey Mesa Research Institute genome database clone numbers. [a] the gene does not contain exons 11 and 12. [b] intronless gene.

Chapter IV

Discussion

In this study, 2 cDNAs for rice β-glucosidases, *bglu1* and *bglu2*, were isolated and characterized. It was shown that both *bglu1* and *bglu2* are highly expressed in shoots during germination, but expressed only at lower levels in other germinating seed parts. In adult plants, *bglu1* expression is high in flower, while the expression of both isozymes is at lower levels in other plant parts, such as node and internode. The function of these genes as β-glucosidases was confirmed by demonstration of β-glucosidase activity in recombinant proteins produced after expressing the cDNA in *E. coli*. The biochemical and catalytic properties of rice BGlu1 were further characterized. In addition, genes for these two enzymes and other β-glucosidases were identified in the rice genome database and their gene structures and representation in the dbEST database were confirmed.

4.1 Cloning and expression of rice β-glucosidase cDNAs

Two cDNAs for rice β-glucosidases, *bglu1* and *bglu2*, were amplified by RT-PCR and their sequences were characterized. The cDNA sequences for *bglu1* and *bglu2* included open reading frames encoding 504 and 500 amino acid long precursor proteins, respectively. Both of these enzymes appeared to enter the secretory pathway, as judged by their N-terminal signal sequences. Sequence analysis of family 1 β-glycosidases has revealed that two Glu residues are highly conserved and appear

within the consensus sequences T(F/L/M)NEP and (I/V)TENG (Czjzek *et al*., 2001). The BGlu1 contains Glu residues at positions 204 and 414 that lie within the sequences TFNEP and ITENG, while BGlu2 contains Glu residues at positions 195 and 409 that lie within the sequence TFNEP and VTENG, which closely resemble the consensus motifs in family 1 β-glycosidases. The presence of the appropriate active site glutamic acids in the appropriate consensus sequence motifs suggests that both *bglu1* and *bglu2* genes at least have the potential to produce catalytically active β-glucosidases.

Additionally, as seen in Figure 3.13, the amino acids identified by Czjzek *et al.* (2000) as critical for glucose binding (Q38, H142, E191, E406, E464 and W465 in maize Bglu1), are highly conserved in retaining glycosidases. These residues are conserved in BGlu1 and BGlu2. Structural studies on maize β-glucosidase substrate binding, indicated that residues W378, F198, F205, and F466 that line the active site cleft interact with the substrate aglycone (Czjzek *et al*., 2000). Therefore, it was suggested that these residues, especially the variable residues F198, F205 and F466, are critical for determining substrate specificity. These residues are indeed quite variable in predicted rice β-glucosidases encoded by *bglu1* and *bglu2* genes, indicating that they are likely to hydrolyze different substrates. Since the maize W378 is well conserved, except in BGlu2 where it is changed to I, it would appear that the substrate binding of BGlu2 is somewhat different than maize Bglu1, where W378 is thought to provide π-interactions for substrate binding. Two other residues near the active site in the dicot cyanogenic β-glucosidases, corresponding to maize Bglu1 Q193 and D261, have been suggested to affect the substrate specificities of white clover linamarase and Thai rosewood dalcochinin β-glucosidase (Ketudat

Cairns *et al*., 2000). Keresztessy *et al*. (2001) showed that mutation of the second of these sites from F to N in cassava linamarase did indeed increase its K_m for its natural substrate linamarin 16 fold without severely affecting the hydrolysis of pNP-glucoside, and mutation of cassava linamarase A201 (corresponding to maize Bglu1 T194) had a similar, but smaller effect. These residues were also variable, though less so than those identified by Czjzek *et al.* (2000).

4.2 Expression of rice β-glucosidase

The presence of many related genes showed that very specific probes are critical to any study of the expression of these genes. Such probes were produced from the 3' UTR of rice *bglu1* and *bglu2* and shown to hybridize to just one gene by Southern analysis. Similarly, it was necessary to pay careful attention to look for exact matches or nearly exact matches of nucleotide sequences when analyzing database search results for ESTs corresponding to certain genes. Thus, it is not clear that such searches can be done with complete success by fully automated methods.

The number of likely isozymes also complicates the interpretation of traditional biochemical approaches, such as measuring activities in extracts. For instance, the pNP-β-glucoside hydrolase activity in germinating rice peaks at 6-7 days (Palmiano and Juliano, 1973, Akiyama *et al*., 1998), but our northern analysis suggested the mRNA peaked at 8-10 days for *bglu1* and *bglu2*. It is unclear whether this represents translational or post-translational control of the β-glucosidases or the dominance of other β-glucosidase isozymes in the extracts.

It was shown that both *bglu1* and *bglu2* are highly expressed in shoots during germination. In adult plants, *bglu1* expression is high in flower, while the expression

of both isozymes is at lower levels in other plant parts. Unlike rice β-glucosidase, the enzyme activity in barley has been shown to decrease during germination (Simos *et al*., 1994). Expression of the barley β-glucosidase is limited to maturing grains and no new synthesis was seen in germinating tissues (Simos *et al*., 1994; Leah *et al.*, 1995; Hrmova *et al*., 1998). Other grass β-glucosidases studied so far were present in seedling including maize sorghum rye and oat (Esen, 1992; Cicek and Esen, 1998a; Sue *et al*., 2000; Nisius, 1988). The transcript levels of durrinase from sorghum was highly elevated in the node and upper half of the mesocotyl but low in root of 4-day-old seedling (Cicek and Esen, 1998a). In dicots, hydroxyisourate hydrolase (β-glucosidase) transcripts of soybean were present in all the plant tissues but highest in nodes when compared with leaves, stems, and roots, where it was constitutively expressed at low levels (Raychaudhuri and Tipton, 2002). An activity in both cytosolic/apoplastic and in cell wall bound β-glucosidase purified from ripe fruit cherry increased during fruit ripening until ripe stage (Gerardi *et al*., 2001).

Some environmental conditions have an effect on the transcript levels of *bglu1* and *bglu2* in rice after 6-day-old seedlings were grown under biotic and abiotic stress for additional 2 days. The drought stress seemed to be the most severe effect on the expression of both genes. Cold and osmotic stresses have little or no effects on *bglu1* gene transcript levels, while salt, flooding, ethylene and abscissic acid have moderate effects. Differences between *bglu1* and *bglu2* were seen to some extent. For instance, under flooding condition, *bglu1* seemed to have higher transcript levels than *bglu2*. While *bglu1* transcript levels seemed to decrease more than *bglu2* in the presence of salt and abscissic acid. However, the time frame have to be considered, since perhaps 2 days response may be different from more immediate response.

Using microarrays, the transcript of clone 1342 matching *bglu1* gene was upregulated in response to high salinity stress in salt-torerant rice (var Pokkali), but not in the salt-sensitive cultivar IR29 (Kawasaki *et al*., 2001). The transcript was increased from 15 min to 6 h. The cDNA encoding ORF of *bglu1* was used in this experiment, hence this probe may have cross hybridized with other rice β-glucosidase mRNAs. However, this experiment implied that different time points and rice strains have to be considered more for these experiments.

The high expression of the rice *bglu1* and *bglu2* in rapidly growing shoots suggests a role in growth, possibly in activation of gibberellin or cytokinin, or in recycling of oligosaccharides generated during rapid cell wall expansion and cell division as well as defense against pests. The high expression of *bglu1* in flowers and presence of gibberellin glucosides in these tissues (Hasegawa *et al*., 1994) is consistent with a role in phytohormone activation. Cell wall recycling is likely to be necessary for flower expansion as well. Some role in defense for rice *bglu1* and *bglu2* also cannot be ruled out, and the enzymes may play more than one role. The roles of rice *bglu1* and *bglu2* are likely to be somewhat different, however, since their sequence differences, especially around the active site, are likely to be reflected in substrate specificity differences, and they have overlapping expression patterns.

It is not possible to predict what the *in vivo* substrate is and might be for a β-glucosidase based on results obtained with artificial substrates or natural substrates from other plant taxa. This is because many β-glucosidases hydrolyze certain artificial substrates with higher catalytic efficiency than they do their natural substrates due to fortuitously better shape and size complementarity at the aglycone binding site. Rice β-glucosidases may play functions similar to those attributed to

β-glucosidases from other plants, but any such function is speculative until it is demonstrated *in vivo*, the natural substrate(s) of each enzyme is isolated from rice tissues, and the conditions under which each substrate is hydrolyzed determined.

4.3 Recombinant protein expression of rice β-glucosidase

The functional production and purification of recombinant soluble proteins expressed in *E. coli* provide a basis for the future characterization of structure and function of rice β-glucosidases. Differences between the expression systems were investigated for the efficient production of rice BGlu1 protein in *E. coli*. Rice BGlu1 is prone to aggregate as inclusion bodies when expressed in the normal-expression system, pET23d(+)/BL21(DE3) as mature protein. However, soluble and catalytically functional rice BGlu1 was obtained when the protein was produced by N-terminal fusion with thioredoxin (TRX) and expressed in the special *E. coli* strain Origami (DE3), which is a *trxB/gor* double mutant. However, there was no protein produced when the recombinant pET32(+)-*bglu1* expressed in BL21(DE3). These results strongly suggested that the host strains as well as the plasmid types are important for obtaining the enzyme.

A higher percentage of the total expressed BGlu1 was soluble when cultures were grown and induced at 15°, 20° or 27°C than at 37°C. This suggested that the recombinant protein tends to form inclusion bodies when synthesis is induced at high temperature (37°C). The recombinant linamarase of cassava Keresztessy *et al*. (1996) and primeverosidase of tea leaves (Mizutani *et al*., 2002) in *E. coli* were also not soluble and directed to inclusion bodies when cultures were grown at 37°C.

To our knowledge, this is the first report of the recombinant expression in *E. coli* of catalytically active grass β-glucosidase which is targeted to secretion and likely be glycosylated in the plant. Expression in *E. coli* of other grass β-glucosidases has been successful in maize, oat, and sorghum in active and soluble form using pET21a(+)/BL21(DE3) pLys system (Cicek and Esen, 1998b). However, their protein products are not glycosylated and are targeted to the plastid, rather than to secretory pathway. There have been a few reports of catalytically active expression in *E. coli* of dicot β-glucosidases which are glycoproteins targeted to the ER via a signal peptide. The cassava linarmarase produced in fusion with glutathione S-transferase (GST) (Keresztessy *et al*., 1996) and the tea primeverosidase expressed in fusion with maltose-binding protein (MBP) (Mizutani *et al*., 2002) in *E. coli* were active, which demonstrated that a nonglycosylated, catalytically active form can be produced. In the case of indican hydrolase of indigo plant, the non-fusion enzyme was produced in active soluble form in pET19b/BL21(DE3) (Minami *et al*., 1999). This suggested that in some cases, glycosylation is not required for activity and stability.

Similar success with this fusion system was not achieved with heterologous expression of rice BGlu2 containing 6 possible glycosylation sites. Most of the recombinant BGlu2 was found to aggregate as inclusion bodies when produced as mature protein in the normal-expression system, pET23d(+)/BL21(DE3) or BL21 (DE3) pLys S. A very small amount of soluble and active rice BGlu2 was obtained when the protein was produced by fusion with TRX in the Origami(DE3) *E. coli* strain. TRX fusion protein may help to decrease aggregation of insoluble protein. However, the biologically active conformation of this protein may require additional post-translational modifications that cannot be done in *E. coli*. Unfortunately, the

recombinant clone which produced active protein lost expression of recombinant protein over time. Therefore, an in-depth characterization could not be performed.

Fusion of polypeptides to soluble fusion partners, such as TRX, GST and MBP, is thought to help to overcome the aggregation of insoluble protein forms by promoting the proper folding of the attached protein into its biologically active conformation. (Kapust and Waugh, 1999). Kapust and Waugh (1999) developed a model to explain how MBP promotes solubility and influences the folding of its fusion partners by comparing the solubility of 18 different fusion proteins in *E. coli*. The results indicated that many aggregation-prone polypeptides could be rendered soluble by fusing them to an appropriate partner. They suggested that MBP plays a passive role in protein folding, merely serving to prevent the off-pathway aggregation of intermediates in the folding process.

Does the recombinant enzyme maintain its activity even after proteolytic cleavage of the fusion moiety? There was no loss of β-glucosidase activity of BGlu1 after removal of TRX by enterokinase. Some fusion enzymes showed no loss of activity after the cleavage, such as tea leave primeverosidase (Mizutani *et al.*, 2002). The cleavage of the calmodulin-binding-protein (CBP) tag from glutamate 1-semialdehyde aminotransferase resulted in 60% loss of specific activity (Tsang *et al.*, 2003). This suggested possible interactions between the recombinant protein and the CBP tag. In contrast, Keresztessy *et al.* (1996) showed that fusion of mature cassava cyanogenic β-glucosidase with GST and expression in *E. coli* lowered the V_{max} values of enzyme activity about 100-200 fold, though it had the K_m value comparable with the native enzyme. The dissociation of the chaperonin GroEL and the cleavage of the GST from this fusion protein led to an increase of specific activity. This suggested

the correct folding of the recombinant protein after dissociation from the chaperonin GroEL and removal of GST.

The expression systems which were constructed sets the stage for a variety of investigations. First, it allows the production of the large amounts of enzyme for purification and further characterization including tertiary structure determination. Second, it confirmed the importance of the fusion tag and the suitable host strain in the production of active recombinant protein. It is significant in that some plant β-glucosidases, such as sorghum dhurrinase2 and rice β-glucosidases which are not abundant proteins, purification in sufficient quantities from their natural source is difficult for thorough characterization. Since multiple isozymes of β-glucosidase are found in rice, with up to 49 genes identified (see section 3.10), it may be a hard task to purify to homogeneity each isozyme from rice.

4.4 Catalytic properties of rice BGlu1

The active rice BGlu1, the protein product of *bglu1* cDNA, was produced as recombinant protein in *E. coli*, characterized for substrate specificity and biochemical properties. Rice BGlu1 showed high hydrolytic efficiency with β-linked oligosaccharides. The enzyme hydrolyzed β-(1,3), -(1,4), -(1, 5) and -(1,6) linkage of oligosaccharides at different rates. BGlu1 hydrolyzed β-(1,4) linked oligosaccharides with increasing catalytic efficiency (k_{cat}/K_m) values as the degree of polymerization (DP) increases from 2 to 6. In contrast, hydrolysis of β-(1,3)-linked oligosaccharides by this enzyme decreased with longer chain lengths, and longer polymers than DP of 3 were not hydrolyzed. BGlu1 also hydrolyzed sophorose and gentiobiose with high

K_m and low k_{cat} values similar to cellobiose. The increased hydrolytic rate with oligosaccharides is a characteristic often observed with polysaccharide exohydrolase, unlike the rate of hydrolysis of oligomeric substrates by β-glucosidase which tend to remain approximately constant or decrease with increasing DP (Reese *et al*., 1968).

Hrmova and colleagues (1996, 1998) have done extensive studies of the substrate binding and hydrolysis of the closely related barley β-glucosidase. They found similar preferences for short β-1,3-linked glucooligosaccharides and longer β-1,4-linked oligosaccharides. However, while the barley β-glucosidase was a good cellobiosidase and seemed to bind poorly at the third position, BGlu1 is a poor cellobiosidase and is better at hydrolyzing cellotriose.

The preference of the rice BGlu1 for (1,4)-β-oligosaccharides of increasing chain length suggested that it has an extensive subsite binding region which is dependent on substrate geometry, as has been described for barley (Hrmova *et al*., 1996, 1998, 2001). According to Hiromi *et al*. (1973) and Suganuma *et al*. (1978), the subsite-binding region of polysaccharide hydrolases consists of an array of tandemly arranged subsites, where each subsite binds a single glycosyl residue of the polymers. Such an array of subsites has now been demonstrated for barley β-glucan glucohydrolases and β-glucosidase/1,4-β-glucan exohydrolase exohydrolases (Hrmova *et al*., 1998).

Rice BGlu1 also hydrolyzed many kinds of aglycones conjugated with glucose, though rather poorly compared with pNPG. BGlu1 hydrolyzed rice pyridoxine-glucoside rapidly as seen by TLC and HPLC. This enzyme also hydrolyzed the cyanogenic glycosides prunasin, dhurrin, and amygdalin, but not linamarin or DIMBOA glucoside. However, BGlu1 did not hydrolyze phenolic

glycosides (arbutin, salicin), thioglucoside (sinigrin), hydroxy courmarin glucoside (esculin), flavonoid glucoside (naringin) or α-D-disaccharides (nigerose). Therefore, rice BGlu1 could not be classified as a phenolic glucosidase, myrosinase (thioglucosidase), flavonoid glucosidase, courmarin glucosidase or α-glucosidase.

The sugar moiety seemed to have less influence on binding to the rice BGlu1 enzyme than the aglycone moiety, like maize β-glucosidase (Babcock and Esen, 1994), but unlike the rice cell-wall bound β-glucosidase (Akiyama *et al*., 1998) and barley BGQ60 (Leah *et al*., 1995). BGlu1 hydrolyzed pNP-derivatives of various monosaccharides, but with specificity for the D-form of sugar residues and the β-O-glycosidic linkage, as it did not hydrolyze the α-O-glycosidic bond, β-S-glycosidic linkage (pNP β-D thioglucoside), nor the L-glycosides (except for pNP-α-L-arabinoside). Catalytic efficiency (k_{cat}/K_m) values varied 65 fold from pNP β-D-xylose to pNP β-D fucose, which suggests that the sugar moiety has a moderately high effect on the rate of hydrolysis. The k_{cat}/K_m ratio of the enzyme is highest for pNP β-D fucoside, indicating that this substrate is the most efficiently hydrolyzed. Since the pentosides have a H- atom at position C-5, while the fucoside has a -CH_3 group at C-6, and the other hexosides have a -CH_2OH at C-6, it is possible that the presence of a hydrophilic substituent at C-6 detracts from binding efficiency (Svasti *et al*., 1999).

The fact that pNP β-D-cellobioside was a substrate with relatively high activity suggested the enzyme might also cleave disaccharides. However, the TLC analysis clearly showed that the first cleavage is of the β-glucosidic bond between the 2 glucose residues, releasing pNPG and glucose. The pNPG is then rapidly hydrolyzed. This sequential pathway is similar to an exoglucanase or exocellulase. The rapid

hydrolysis of cellotriose to cellobiose and glucose may reflect binding to the active site in a manner similar to pNP β-D-cellobiose.

As was noted for barley β-glucosidase (Hrmova *et al.*, 1996), BGlu1's specificity for cello-oligosaccharides would allow it to be characterized as a 1,4-β-glucan exohydrolase (3.2.1.74). However, as it hydrolyses other oligosaccharides, and alkyl-glucosides, and showed structural similarity with family 1 glycosyl hydrolases, classifying BGlu1 as a β-glucosidase (E.C. 3.2.1.21) seems more appropriate.

4.5 Possible function of rice BGlu1

Rice BGlu1 hydrolyzed β-linked disaccharides and small oligosaccharides at relatively high rates and had high transglucosylation activity towards short chain laminari- and cello-oligosaccharides. This indicated that a possible role of the enzyme might be further hydrolysis of oligosaccharides released from cell wall β-glucans and generating new oligosaccharides by recycling the sugar residues released from the cell wall during germination or flower expansion. This is similar to barley enzyme (Leah *et al.*, 1995, Hrmova *et al.*, 1998) and rice cell wall enzyme (Akiyama *et al.*, 1998), which appear to be involved in recycling of cell wall-derived oligosaccharides produced during germination. However, the roles of rice BGlu1, cell-wall-bound rice β-glucosidase and barley β-glucosidase βII in plant tissue are likely to be somewhat different, since they have difference in expression patterns. The high expression of *bglu1* in flowers and presence of gibberellin glucosides in these tissues (Hasegawa *et al.*, 1994) is consistent with a role in phytohormone

activation. Interestingly, BGlu1 was also inhibited by IAA and GA_3, which might indicate these compounds bind to the active site, perhaps at the aglycone binding site. However, the glucosides of these and similar phytohormones were not available to test whether they are actually substrates for BGlu1. Some role in defense for rice *bglu1* and *bglu2* also cannot be ruled out, and the enzymes may play more than one role. For instance, maize BGlu1 has been identified by different labs as hydrolyzing DIMBOA-glucoside, a defense-related compound (Babcock and Esen, 1994), and phytohormone glucosides (Brzobohaty *et al*., 1993). The enzyme also hydrolyzed rice pyridoxine-5'-O-β-D-glucoside, a major metabolite in rice bran (Yasumoto *et al*., 1977). This substrate might implicate BGlu1 in the metabolism of the coenzyme pyridoxine (vitamin B_6).

4.6 Deduced protein sequence similarity between rice and barley β-glucosidases

Rice BGlu1, cell-wall-bound rice β-glucosidase (Akiyama *et al*., 1998) and barley β-glucosidase (Hrmova *et al*., 1998) are enzymes that hydrolyze β-linked oligosaccharides. However, the specificity for substrate chain lengths of these enzymes is different. In contrast to barley and cell-wall-bound rice enzyme, rice BGlu1 preferred to hydrolyze β (1,4) linked oligosaccharides of DP higher than 3, but did not hydrolyze disaccharides well (except laminaribiose). Thus, the comparison between this enzyme was considered at the protein sequence level. The deduced amino acid sequence of BGlu1 had the highest sequence identity (66%) with barley β-glucosidase BGQ60 (or βII), while cell-wall-bound rice (deduced from contig 445-1

see section 3.10) are closely related to cyanogenic β-glucosidases. The amino acids identified by Czjzek *et al*. (2000) as critical for glucose binding (Q38, H142, E191, E406, E464 and W465 in maize Bglu1) are conserved in BGlu1, cell wall bound rice and barley β-glucosidase. Interestingly, rice BGlu1 protein sequence was closest to barley BGQ60 at some of substrate binding residues that line the active site cleft and interact with the substrate aglycone of maize β-glucosidase (W378, F198, F205, and F466) (Czjzek *et al*., 2000), suggesting it may have a similar substrate-specificity. However, these above mentioned amino acid residues were different from those in the cell-wall-bound enzyme and may indicate for the different substrate specificties with some oligosaccharides and glycone specificity. Therefore, the tertiary structures of these enzymes are needed to elucidate the enzyme-substrate binding mechanism.

4.7 Biochemical properties of rice BGlu1

4.7.1 Molecular weight and pI

Rice BGlu1 is a monomeric protein with Mr of TRX fusion and nonfusion rice BGlu1 was estimated to be 66 and 54 kDa, respectively. The Mr of rice BGlu2 mature protein (produced as mature protein in *E. coli*) is 52 kDa. This is similar to the cell-wall-bound rice β-glucosidase's Mr (56 kDa) which is a monomeric protein (Akiyama *et al*., 1998). In contrast, both Fr-1 and Fr-2 rice β-glucosidases partially purified by Muslim (1995) have the denaturing form Mr of 60 kDa and native dimers of 120 kDa. The barley and maize β-glucosidases have Mr's ranging 53 to 62 kDa (Leah *et al*., 1995; Hrmova *et al*., 1996; Esen, 1992). So, the Mr of rice BGlu1 and BGlu2 are in this range. Many β-glucosidases contain identical subunits. For

instance, the furostanol glycoside 26-O-β-glucosidase from *Costus speciosus* is a 110 kDa protein with two subunits with Mr of 54 and 58 kDa (Inoue and Ebizuka,1996). Furthermore, the isoflavone hydrolyzing β-glucosidase from chickpea has apparent Mr of 125-135 kDa with two identical subunits of 68 kDa (Hosel an Barz, 1975).

The predicted pI of BGlu1 is 9 indicated that the enzyme is a basic protein. This is consistent with the fact that BGlu1 entered the acidic pH gel, but not basic pH gel. The enzyme also hydrolyzed 4MUGlc in acidic pH native gel. The basic pI of the rice BGlu1 is similar to that described for the other enzymes involved in oligosaccharide hydrolysis and purified from cell-wall of rice (pI 10) (Akiyama *et al*., 1998), germinated barley (pI 9-9.5) (Leah *et al*., 1995), and barley βI (8.9) and βII (9) (Hrmova *et al*., 1996). β-glucosidase partially purified from rice seedlings (Muslim, 1995) also has pI value of 9.83 and 9.42 for Fr-1 form , while Fr-2 was 10.08. While Iwami and Yasumoto (1986) found three different enzymes have pI 5.6, 6.7, and 8.2. In addition, a rice β-glucosidase which hydrolyzed conjugated gibberellin from dwarf-*rice (O. sativa* cv. *Tan-ginbozu*) was partially purified by using CM-Sephadex C-50 chromatography (Schliemann, 1984), which suggested that the enzyme seem to be a basic protein. From the above mentioned reports, it is obvious that rice β-glucosidases consist of more than one form, and both acidic and basic proteins.

4.7.2 pH and temperature profile of stability and activity

The pH optimum for activity of BGlu1 was 5.0 and the enzyme was stable at 25°C in the pH range 5.0 to 7.0 up to 24 hr. This indicated that the enzyme was stable for a long time at the pH of optimum activity. Esen (1993) indicated that almost all

β-glucosidase have an optimum pH in the 5 to 6 range. The pH optimum of the enzyme is related to the pH environment of the catalytic glutamic acids.

The temperature optimum of the enzyme was 15 to 35°C. Thermostability studies indicated that the enzyme was relatively stable at its temperature optimum of activity (20 and 30°C), when incubated for up to 60 min. Dixon and Webb (1979) stated that the optimum temperature for a given enzyme depends on a balance between the rate of catalytic reaction and denaturation. β-glucosidases from different organisms might have different optimum temperature and stability, which would indicate different interaction stabilizing enzyme tertiary and quaternary structures. However, the temperature optimum of rice BGlu1 is lower than that of other plant β-glucosidases; such as maize (50°C) (Esen, 1992). However, the recombinant enzyme may not be as stable as the native enzyme in the plant.

4.7.3 Inhibition studies

Glucono-1,5-lactone, an analogue of the oxocarbanium ion-like transition state, strongly inhibited rice BGlu1 like some other β-glucosidases (Akiyama *et al.,* 1998; Babcock and Esen, 1994, Svasti *et al.*, 1999). However, glucono-1,5-lactone is not a very effective inhibitor in some cases, such as soybean IC50 of this inhibitor was 24 mM (Hsieh *et al*., 2001).

Galactono-1,4-lactone, IAA, GA_3, pyridoxine, 4-methyl umbelliferone, Tris base and imidazole also showed some inhibitory effect, but the degree of inhibition was much less. On the other hand, glucose and other monosaccharides had no effect on enzyme activity at 10 mM, suggesting that sugars do not bind tightly to the active

site. This inhibition pattern is like that of cell-wall-bound (Akiyama *et al*., 1998) and Fr-1 and Fr-2 rice β-glucosidases (Muslim, 1995).

SDS, an amphiphatic anion detergent, is a strong denaturing agent. Rice BGlu1 is susceptible to denaturantion by this detergent which reduced pNPG hydrolase activity. When incubating the enzyme with 1 mM pNPG, the activity was totally lost in the presence of SDS at or above 0.04 percent. The degree of denaturing effect varies considerably, depending on concentration, pH, temperature, and the structure of a given β-glucosidase (Esen and Gungor, 1993).

BGlu1 showed activity loss when it was subjected to increasing concentration of ethanol. BGlu1 lost activity up to 50% after it was incubated with pNPG in the presence of 30% ethanol. Pure organic solvents are non-aqueous media; therefore, organic solvents reduce enzyme activity and stability by disrupting hydrogen-bonding interactions with the solvent. The ethanol acts as competitive inhibitors of almond β-glucosidase: at 6 M concentration, ethanol decreased the activity by 69.5% (Romeu *et al*., 1994).

Several metal ions have been shown to have an inhibitory effect on β-glucosidase activity; for example: Ag^{+}, Ca^{2+}, Co^{2+}, Cu^{2+}, $Fe3^{+}$, and Hg^{2+} (Zollner, 1993). Of all the chemicals tested for the inhibition of enzyme activity, only copper and mercury ions inhibited the enzyme. This sensitivity was similar to that of maize β-glucosidase (Esen, 1992). The addition of 2-mercaptoethanol (2-ME), however, caused an additional loss of activity of the enzyme. Mercuric compounds usually inhibit enzymes by reaction with sulfhydryl groups (SH). The importance of SH groups of cysteine residues may be involved in stability and activity of β-glucosidases (Hakulinen *et al*., 2000). However, the strongest inhibition of Thai rosewood β-

glucosidase with mercuric compounds, at the level of 10^{-7} M possibly results from chelation of catalytically active acidic amino acids in the enzyme, as suggested for other enzymes (Svasti *et al.*, 1999). BGlu1 is a monomeric enzyme without the presence of intermolecular disulfide bonds. This suggested that the loss of an activity by metal ion may not be due to the accessibility of the protein SH groups to these compounds as previously described for other β-glucosidases such as maize (Esen, 1992) or soybean (Hsieh *et al.*, 2001).

4.8 Transglycosylation activity of rice BGlu1

BGlu1 showed high efficiency in the synthesis of many glucoside compounds. Like cell wall-bound rice β-glucosidase (Akiyama *et al.*, 1998) and barley β-glucosidases from germinated seeds (Hrmova *et al.*, 1998), BGlu1 can use a variety of oligosaccharides, such as cello-oligosaccharides and laminari-oligosaccharides, which act as both glucose donor and glucose acceptor molecules to synthesize the longer oligosaccharides. Both rice BGlu1 and cell wall-bound β-glucosidases catalyze this reaction in the presence of low concentrations 5 and 2 mM, respectively, of oligosaccharides substrates, which indicates the high transglycosylation efficiency of these enzymes. So, The transglycosylation properties of rice BGlu1 with different glucose acceptor molecules were investigated.

4.8.1 Transglycosylation of pNP β-D-glucoside

The rice enzyme, likes barley β-glucosidase (Hrmova *et al.*, 1998), could use pNPG both as glucose donor and acceptor in the transglycosylation reaction leading to

the synthesis of high amounts of pNP-containing compounds including pNP β-D-cellobiose, pNP β-D-cellotriose and at least two unidentified compounds which likely to be pNP-oligosaccharides. These other two compounds maybe pNP-laminaribiose and pNP-gentiobiose, which were seen by Hrmova *et al*. (1998) with barley β-glucosidase. Different conformation of pNP-diglucosides generates by transglycosylation begs the question how the glucosyl molecule can be transferred from the donor to different hydroxyls on the acceptor molecule by the enzyme. Hrmova *et al*. (1998) suggested that this abilities derived from some flexibility in the binding of a glucosyl residue to subsite +1 of barley β-glucosidase, which allows the pNPG to move sufficiently in the active site to present hydroxyls on C atoms 3, 4 and 6 to the bound glycone prior to transfer of that bound glucosyl residue to the pNPG.

The transglycosylation reaction is controlled kinetically, hence the yield of glycoside is determined by a delicate balance between the rates of donor synthesis and hydrolysis, on the one hand, and product hydrolysis on the other (Rantwijk *et al*., 1991). The other β-glycosidases reported so far required high substrate concentrations to exhibit sufficient transglycosylation activities (Hrmova *et al*., 1998; Mala *et al*., 1999). The efficiency of BGlu1 in the pNPG transglycosylation also depends on the substrate concentration. However, at 10-20 mM pNPG, high amounts of products can be observed in a short incubation period.

4.8.2 Transglycosylation of ethyl alcohol

Rice BGlu1 also catalyzed the transfer of glucose from pNPG to ethyl alcohol, to form ethyl glucoside. High amounts of ethyl glucoside product were obtained in the presence of 5-20% alcohol. However, decreased product in the presence of a high

percentage of ethyl alcohol in the solution limits the efficiency of BGlu1 to catalyze this reaction suggesting that inhibition may occur. The synthesis of alkyl glucosides by many β-glucosidases in the presence of alcohol using glucose donors such as pNPG or cellobiose have been reported (Makropoulou *et al*., 1998; Fischer *et al*., 1996; Svasti *et al*., 2003). Instead of pNPG as glucose donor, almond β-glucosidase could use cellobiose and gentiobiose as glucosyl donors, while Thai rosewood β-glucosidase could use gentiobiose but the cassava enzyme could use neither of them (Svati *et al*., 2003). For rice β-glucosidase, some disaccharide and oligosaccharides such as cellotriose, cello tetraose, laminaribiose and laminaribiose, except cellobiose (because of low catalytic efficiency) may be used as glucosyl donor instead of pNPG to produce variety of alkyl glycosides with different degree of polymerization. However, the high costs of these oligosaccharide substrates are prohibitive for this use.

For application of this enzyme for synthesis of alkyl glycosides, which have potential use as biodegradable detergents, the enzyme might be tested with various alcohol chain-lengths. Interestingly, almond, Thai rosewood and cassava β-glucosidases not only have the ability to synthesize hexyl glycosides by transferring glucose from pNP-glycosides to hexanol, they also have the ability to cleave this substrate (Romeu *et al*., 1994; Srisomsap *et al*., 1996; Eksittikul *et al*., 1988). This suggested that the binding of the aglycone binding portion in the active sites of these enzymes with the hexyl group are important for both synthesis and hydrolysis reactions. Since rice BGlu1 cleaved the hexyl and octyl glucosides, with low efficiency but high binding strength (low K_m), this implied that the aglycone binding

portion in the active site of BGlu1 could bind octyl and heptyl molecule. Hence, the synthesis of alkyl glycosides from longer chain alcohol may be possible.

4.8.3 Transglycosylation of pyridoxine

In addition to the hydrolysis of pyridoxine-5'-O-β-D-glucoside, rice BGlu1 enzyme also showed considerably high transglycosylation activity towards pyridoxine to make pyridoxine-5'-O-β-D-glucoside in the presence of the glucose donor molecule p-nitrophenol β-D-glucoside. Though its low catalytic efficiency for the hydrolysis suggest this may not be a physiological function of the enzyme, the transglycosylation reaction implies the enzyme may play a role in the glycosylation of pyridoxine seen when germinating rice is incubated with pyridoxine (Suzuki *et al*., 1986). However, this evidence may imply another possible function of the enzyme in germinating shoot and panicle at flowering stage. This enzyme may be involved in hydrolysis of the coenzyme pyridoxine from its glucose-conjugated storage form, and the formation of pyridoxine glucoside via transglucosylation during flowering stage to accumulate the inactive vitamin form just prior to storage in the seeds.

This hypothesis is consistent with the data of Iwami and Yasumoto (1986). They found that three different rice bran β-glucosidases (based on their isoelectric point measurements: i.e., pH 5.6, 6.7, and 8.2) catalyzed the transfer of the glucosyl moiety of pNPG to pyridoxine to synthesize pyridoxine-β-D-glucoside. Conjugation of pyridoxine and glucose by incubation of cellobiose and pyridoxine in the presence of the wheat bran β-glucosidase was previously reported (Suzuki *et al.*, 1979). Although many papers reported the synthesis of pyridoxine glucoside, the mechanism of its formation in plants has not been determined (Gregory, 1993).

There have been no reports about the primary structure of plant β-glucosidases that catalyze the transglucosylation to synthesize pyridoxine glucosides. So, this is the first rice β-glucosidase for which a protein sequence has been deduced and shown to catalyze this reaction. This enzyme can be used for the synthesis of pyridoxine glucoside *in vitro* to use for vitamin research, such as screening for enzymes that hydrolyze this substrate efficiently.

In plants including rice, vitamin B_6 is found predominantly as pyridoxine-5'-O-β-D-glucoside or other complex conjugates such as 4'-O-(β-D-glucosyl)-5'-O-(β-cellobiosyl) pyridoxine, and 5'-O-(β-cellobiosyl) pyridoxine (Yasumoto *et al.*, 1977; Tadera *et al.*, 1988). Though the functions of vitamin glycosides are not completely understood, one possible function of pyridoxine glucoside in plants is to serve as a metabolically inactive storage form (Gregory, 1998). These conjugates are partially utilized in humans and have a competitive inhibitory effect on the absorption and utilization of pyridoxine. So treatment of plant material with β-glucosidases may improve the nutritional content of the plants (Gregory, 1993).

Some research groups have extensively studied the production of efficient catalysts for glycoside synthesis but not hydrolysis by mutation of amino acid residues (Withers, 2001). Wang *et al.* (1994) showed that mutation of the catalytic nucleophile glutamic residue, to a non-nucleophilic residue such as alanine, eliminated activity in a retaining glycosidase. The mutant had maximal hydrolysis rates at least 105 times lower than the wild type enzyme. Circular dichroism and X-ray crystallographic analyses in several cases revealed that such mutants fold correctly, thus have an active site capable of binding both the aglycone and glycone moieties. Since rice *bglu1* cDNA was cloned and the BGlu1 protein sequence was

thereby determined, the mutation of amino acid residue for efficient catalyst for glycoside synthesis but not hydrolysis of this enzyme might be investigated in the future. β-glucosidases from rice might also have the potential for deglycosylation and production of novel glycosides and oligosaccharides.

4.9 Inhibition of rice germination by β-glucosidase inhibitor

An examination of the physiological role of rice β-glucosidase by histochemical localization of the enzyme in dry seeds and application of inhibitors of the enzyme to the germinating seeds was previously done by Muslim (1995). Glucono-1,5-lactone is known to be strong inhibitor of β-glucosidases of various enzyme sources including rice BGlu1. This compound directly competes with a regular substrate for binding site in the enzyme. Muslim stated that rice β-glucosidases present in dry seeds were preexisting enzymes and was located histochemically in the embryo and in the aleurone layers. In his study, gluco-1,5-lactone inhibits the germination at the activation stage with 50% germination inhibition value at 63.5 mM.

The result in this study showed that in the presence of gluco-1,5-lactone, the seeds were able to germinate. At 25 and 50 mM glucolactone the seeds still germinated but with lengths of the shoot and root that were much shorter than the control (0 mM). Even when gibberellin was added (GA_3, 1 mM application), this could not overcome the inhibition. The enzyme inhibitor could block at a step after gibberellin; that is why it is not reversed by GA_3. Both results are consistent with Muslim's work.

Although gluconolactone is a potent inhibitor of rice germination, it requires that a high concentration be applied to the rice seeds. In aqueous solution, gluconolactone is hydrolyzed to gluconic acid by water. The inactivation and instability of gluconolactone might be the reason for the need for high concentration of this inhibitor. X-glucosidase activity in rice seeds and seedlings were also inhibited by gluconolactone. The inhibition was only temporary, since with longer incubation time in staining buffer (more than 3 h), the development of the staining eventually occured. These data indicate that the inhibition is reversible only when the inhibitor is diluted out.

Functionally, the aleuron layers serve as a storage tissue, containing mostly protein bodies and lipid bodies. During germination, aleuron layers serve as a site of gene expression and synthesis of several hydrolytic enzymes, (for example α-amylase isozymes) (Thomas and Rodriguez, 1994). The embryo is the most active part in regulating mobilization of storage materials from the endosperm to sites of growth and development during germination. The highly dynamic cell division in the embryo occurs along the embryo axis (Thomas and Rodriguez, 1994). The presence of rice β-glucosidases in dry seeds as preexisting enzymes located in the embryo and in the aleurone layer, indicated the enzyme may play an important role in germinating process. Some studies showed that β-glucosidases in rice regulate gibberellin (Schliemann, 1984), cytokinin (Saha *et al*., 1984), and pyridoxine (Iwami and Yasumoto, 1986) availability and activity. Therefore, gluconolactone may inhibit the activity of these hydrolytic enzymes, and thus blocks the metabolic processes needed for germination.

4.10 Rice β-glucosidase gene family

In the past, most studies have assumed a small number of β-glucosidase isozymes in plants because only a small number of enzymes are present in large amounts and hydrolyze the artificial substrates commonly used and thus can be detected. However, a search of rice genome databases identified 49 β-glucosidase genes at this point, at least 26 of which are expressed. Since cDNA or ESTs have been identified were from transcripts corresponding to at least 26 of these genes and the EST data are incomplete, the number of expressed genes is likely to be high.

This is not unexpected since a search of the *Arabidopsis thaliana* genome identified 48 glycosyl hydrolase family 1 homologues (A. Esen, unpublished). How many of these genes are functional β-glucosidase genes? The presence of the appropriate active site glutamic acids in the appropriate consensus sequences motifs suggests that all the full length genes identified in the rice genome database at least have the potential to produce catalytically active β-glucosidases. Additionally, as seen in multiple sequence alignment in Appendix A, the amino acids identified by Czjzek *et al*. (2000) as critical for glucose binding (Q38, H142, E191, E406, E464 except W465 in maize Bglu1) are well conserved in these predicted sequences. Only the predicted β-glucosidase from contig 3224 has H in stead of Q38 in maize. The predicted β-glucosidase including OSM12458, contig 7736, 33501, 3224, 8125, CL38960.29-1 and CL009350.95 have I/L/M/Y/F intead of W465 in maize. The residues that line the active site cleft and interact with the substrate aglycone of maize (Czjzek *et al*., 2000) are indeed quite variable in the predicted rice β-glucosidases

(Appendix A), indicating they are likely to hydrolyze a variety of substrates, assuming their substrate binding is similar to that of maize Bglu1.

The number of likely isozymes also complicates the interpretation of results from traditional biochemical approaches, such as measuring activities in extracts. It is unclear whether this represents translational or post-translational control of β-glucosidases or the dominance of other β-glucosidase isozymes *in* the extracts. Protein purification may also be difficult due to the similar sizes and isoelectric points (pI) of several predicted isozymes, as seen in table 2. For instance, 6 putative proteins have predicted isoelectric points in the 5.50 to 5.79 range, and Mr of proteins each range from 53 to 63.6 kDa within 0.2-1 kDa different from other isozymes. Though multimerization and glycosylation differences may exist, it is likely to be difficult to separate these isozymes, if they are expressed in the same tissues. Therefore, production in heterologous hosts and specific immunological approaches are likely necessary to assess the roles of individual isozymes.

Gene structural analysis of the β-glucosidases showed that intron/exon splice sites were conserved, but 5 patterns of gene structures were distinguished by the number of exons and introns. It was found that *Arabidopsis thaliana* has as many gene organization patterns as rice, though they are different to some extent (A. Esen unpublished). Genes with 13 exons (group 1) probably represent the ancestral gene organization and they appeared to be highly divergent by protein sequence analysis. Those genes with 11 exons clustered together in one group with barley BGQ60 (which is also missing the intron between group 1 exons 2 and 3), while those with 9 and 12 exons clustered in a separate group. This phylogeny is consistent with an ancestral plant β-glucosidase having 13 exons and 12 introns, with losses of introns in

groups 2, 3 and 4 since 13 exons are most divergent and gain of introns would require their insertion at the exact same splice site position. Six of the seven intronless genes found in BGI database had the highest sequence identity with microorganisms. These sequences may derive from the contamination of endophyte DNA in the rice genome sequences. However, these genes may be the retrotransposon on the rice genome.

The OSM129364 and *bglu1* gene appears to have diverged from the barley BGQ60 gene after it diverged from the other rice β-glucosidases, so the rice/barley ancestor species would have contained already diverged ancestors to these genes group. However, an unreliable evolution divergece was seen in the contig 445-1, cell-wall bound rice β-glucosidase (Akiyama *et al*., 1998). This gene was grouped together with cyanogenic glucosidase, though its function seems to be in oligosaccharides hydrolysis. As recently noted by Henrissat and Davies (2000), it is not generally possible to assign glycosyl hydrolase function based on sequence similarity scores alone.

Although these four draft sequences provide a rich resource for data mining, this analysis was somewhat limited by the use of the rice genome database, which seemed to contain about 93% coverage of the genome.

Chapter V

Conclusion

In this study, 2 cDNAs for rice β-glucosidases, *bglu1* and *bglu2*, were isolated and characterized. The cDNA sequences for *bglu1* and *bglu2* included open reading frames encoding 504 and 500 amino acid precursor proteins, respectively. Both of these enzymes appeared to enter the secretory pathway, as judged by their N-terminal signal sequences. The BGlu1 protein sequence had the highest sequence identity (66%) with barley β-glucosidase BGQ60, while BGlu2 protein had highest sequence identity (54%) with amygdalin hydrolase of *Prunus serotina.* The presence of the appropriate active site glutamic acids in the appropriate consensus sequence motifs suggested that both *bglu1* and *bglu2* genes at least have the potential to produce catalytically active β-glucosidases.

Southern blots using gene-specific probes indicated that *bglu1* and *bglu2* were single copy genes. The patterns of hybridization were different, indicating that *bglu1* and *bglu2* probes do not hybridize to the same genes. It was shown that both *bglu1* and *bglu2* are highly expressed in shoots during germination, but expressed only at lower levels in other germinating seed parts. In adult plants, *bglu1* expression is high in flower, while the expression of both isozymes is at lower levels in other plant parts, such as node and internode. Subjecting seedlings to various stress conditions for 2 days resulted in similar or lower *bglu1* and *bglu2* expression levels.

cDNAs encoding the mature β-glucosidase proteins, BGlu1 and BGlu2, were amplified by the polymerase chain reaction (PCR) and cloned into various expression vectors in order to generate an effective systems for producing recombinant protein for functional and structure analysis. BGlu1 and BGlu2 were produced in redox-deficient *E. coli* as N-terminal thioredoxin fusion proteins. However, BGlu1 seemed to be found more in an active form, while most of BGlu2 fusion protein appeared to be localized to inclusion bodies, though some active, soluble protein was produced.

Thioredoxin fusion BGlu1 was purified by chelated-metal affinity chromatography. Purified rice BGlu1 was further analysed for its biochemical properties and substrate specificity. The apparent Mr of the TRX BGlu1 fusion and non fusion protein (after cleavage with enterokinase) was estimated to be 66 and 54 kDa, respectively. Native rice BGlu1 fusion protein is about 66 kDa indicating that the enzyme is a monomeric protein. The enzyme activity had pH and temperature optima at 5.0 and 15-35°C, respectively.

Rice BGlu1 hydrolyzed β-oligosaccharides with the highest catalytic efficiency among the substrates tested. The enzyme hydrolyzed β-(1,3), -(1,4), -(1, 5) and -(1,6) linkage of oligosacchirdes at different rates. For cello-oligosaccharides, the catalytic efficiency (k_{cat}/K_m) values increased with increasing chain length of the cello-oligosaccharides. In contrast, hydrolysis of β-(1,3)-linked oligosaccharides by this enzyme decreased with longer chain lengths and longer polymers than DP of 3 was not hydrolyzed. In addition to hydrolysis activity, BGlu1 showed transglycosylation activity towards laminari-oligosaccharides with DP of 2 and 3, and cello-oligosaccharides with a DP of 3 and 4. BGlu1 also hydrolyzed pyridoxine-5′-O-β-D-glucooside. Studies with several other available natural substrates showed that these

compounds were hydrolyzed poorly or not at all. In addition, BGlu1 hydrolyzed the pNP β-glycosides of the monosacharides, D-glucose, D-fucose, D-galactose, D-mannose, D-xylose, and the disaccharide D-cellobiose and pNP-α- L-arabinoside.

The transglycosylation activity of rice β-glucosidase was also studied. Rice BGlu1, catalyzed substrate transglycosylation with pNPG, which acts both as a glycosyl donor and acceptor to produce high amounts of larger pNP-derivatives. BGlu1 also catalyzed the transfer of glucose from pNPG to ethyl alcohol to form ethyl glucoside) and to pyridoxine to synthesize pyridoxine-5'-O-β-D-glucoside.

Rice BGlu1 was inhibited by many aglycones and aglycone analogues for known plant glycosides, but not inhibited by monosaccharides at 10 mM. Glucono-1,5-lactone, a strong inhibitor of many β-glucosidases, strongly inhibited BGlu1. Galactono-1,4-lactone, IAA, GA_3, pyridoxine and 4-methyl umbelliferone also inhibited BGlu1, but the degree of inhibition was much less.

In the presence of the β-glucosidase inhibitor glucono-1,5-lactone, rice seeds were able to germinate. At 25 and 50 mM glucono-1,5-lactone, the seeds still germinated but with lengths of shoots and roots averaged much shorter than the control. Even when gibberellin was added (GA_3, 1 mM application), however, the inhibition could not be overcome. X-glucoside hydrolysis activity of β-glucosidase in rice seeds and seedlings was also inhibited by this inhibitor.

Genes homologous to glycosyl hydrolase family 1 β-glucosidase have been identified from the Beijing Genomic Institute, GenBank at NCBI, the Monsanto Rice Genome Sequencing Project, Torrey Mesa Research Institute rice genome databases. Forty-nine β-glucosidase genes, including 32 full-length genes, 10-nearly full-length

genes, and 7 intronless genes, were identified. The open reading frames (ORFs) of forty-nine gene-derived cDNAs showed significant shared deduced amino acid sequence identity to each other and other known β-glucosidase sequences. Intron-exon boundaries and intron numbers are highly conserved among rice and other plant β-glucosidase genes. At least 26 rice β-glucosidase genes have corresponding ESTs, indicating their expression, and these ESTs came from many tissues, including seedling (shoot and root), immature plant parts (stem, root, leaf), mature leaf, panicle at flowering and ripening stage, immature seeds, and callus. The expression of β-glucosidases in many tissues and plant developmental stages suggests their general importance to plants.

In conclusion, the fundamental knowledge about rice β-glucosidases, *bglu1* and *bglu2*, including structure, gene organization, temporal and spatial expression, expression levels in response to various environmental conditions were characterized. In addition, the biochemical properties, substrate specificity, and transglucosylation activity of rice BGlu1 were characterized these studies suggested possible functions of the enzyme. This work is a good start toward determining the roles of the Glycosyl hydrolase family 1 β-glucosidases in rice, which presents an opportunity to investigate the molecular basis for differences in substrate specificity, which may be correlated with the evolution of enzyme functions. In the future, the tertiary structure of rice BGlu1 should be studied to understand the enzyme-substrate binding mechanism and gain a thorough understanding of the basis for the different substrate specificity as compared with other enzymes in this family. Knowledge gained through these studies may be applied to use in biotechnological and industrial processes.

References

References

Akiyama, T., Kaku, H., and Shibuya, N. (1998). A cell wall-bound β-glucosidase from germinated rice: purification and properties. **Phytochemistry.** 48: 49-54.

Arumuganathan, K., and Earle, E. D. (1991). Nuclear DNA content of some important plant species. **Plant Mol. Biol. Rep.** 9: 208-218.

Babcock, G. W., and Esen, A. (1994). Substrate specificity of maize β-glucosidase. **Plant Sci.** 101: 31-39.

Bachem, C. W., van der Hoeven, R. S., de Bruijn, S. M., Vreugdenhil, D., Zabeau, M., and Visser, R. G. (1996). Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: analysis of gene expression during potato tuber development. **Plant J.** 9: 745-753.

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.

Barry, G. F. (2001). The use of the Monsanto draft rice genome sequence in research. **Plant Physiol.** 125: 1164–1165.

Bell, A. (1981). Biochemical mechanisms of disease resistance. **Annu. Rev. Plant Physiol.** 32: 21.

Beutler, E. (1992). Gaucher disease: new molecular approaches to diagnosis and treatment. **Science.** 256: 794-798.

Bothast, R. J., and Saha, B. C. (1997). Ethanol production from agricultural biomass

substrates. **Adv. Appl. Microbiol.** 44: 261-286.

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.

Brzobohaty, B., Moore, I., and Palme, K. (1994). Cytokinin metabolism: implications for regulation of plant growth and development. **Plant Mol. Biol.** 26: 1483-1497.

Buell, C. R. (2002). Current status of the sequence of the rice genome and prospects for finishing the first monocot genome. **Plant Physiol.** 130: 1567-1568.

Burmeister, W.P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S., and Henrissat, B. (1997). The crystal structures of *Sinapsis alba* myrosinase and a covalent glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinery of an S-glycosidase. **Structure.** 5: 663-675.

Campos. N. L., Bako, L., Brzobohaty, B., Feldwisch, J., Zettl, R., Boland, W., and Palme, K. (1993). Identification and characterization of a novel phytohormone conjugate specific β-glucosidases activity from maize. In: A. Esen (ed.). **β-glucosidase: Biochemistry and molecular biology. ACS Symposium Series 533** (pp 204-212). Washington, D. C.: American Chemical Society.

Cicek, M., and Esen, A. (1998a). Structure and expression of a dhurrinase (β-glucosidase) from sorghum. **Plant Physiol.** 116: 1469-1478.

Cicek, M., and Esen, A. (1998b). Expression of soluble and catalytically active plant β-glucosidases in *E. coli*. **Biotech. Bioengineer.** 63: 392-400.

Clarke, A. J., Bray, M. R., and Strating, H. (1993). Specific β-glucosidases, specific

β-glucanases, and xylanases. In: A. Esen (ed.). **β-glucosidase: Biochemistry and molecular biology. ACS Symposium Series 533** (pp 27-41). Washington, D. C.: American Chemical Society.

Conn, E. E. (1993). Cyanogenic glycosides. In: A. Esen (ed.). **β-glucosidase: Biochemistry and molecular biology. ACS Symposium Series 533** (pp 15-26). Washington, D. C.: American Chemical Society.

Cuevas, L., Niemeyer, H. M., and Jonson, L. M. V. (1992). Partial purification and characterization of a hydroxamic acid glucoside β-glucosidase from maize. **Phytochemistry.** 31: 2609-2612.

Czjzek, M., Cicek, M., Zamboni, V., Burmeister, W. P., Bevan, D. R., Henrissat, B., and Esen, A. (2000). The mechanism of substrate (aglycone) specificity in β-glucosidases is revealed by crystal structures of mutant maize β-glucosidase-DIMBOA, -DIMBOAGlc, and -dhurrin complexes. **Proc. Natl. Acad. Sci. USA.** 97: 13555-13560.

Czjzek, M., Cicek, M., Zamboni, V., Burmeister, W. P., Bevan, D. R., Henrissat, B., and Esen, A. (2001). Crystal structure of a monocotyledon (maize ZMGlu1) β-glucosidase and a model of its complex with p-nitrophenyl β-D-thioglucoside. **Biochem. J.** 354: 37-46

Davies, G., and Henrissat, B. (1995). Structure and mechanisms of glycosylase. **Structure.** 3: 853-859.

De Detta, S. K. (1981). **Principles and practices of rice production.** NewYork: John Wiley and Sons.

Dixon, M., and Webb, E. C. (1979). **Enzymes** (3nd ed.). New York: Academic Press.

Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. **Phytochem. Bull.** 19: 11-15.

Eksittikul, T., and Chulavatnatol, M. (1988). Characterization of cyanogenic β-glucosidase (linamarase) from cassava (Manihot esculenta Crantz). **Arch. Biochem. Biophys.** 266: 263-269.

Esen, A. (1992). Purification and partial characterization of maize (*Zea mays* L.) β-glucosidase. **Plant Physiol.** 98: 174-182.

Esen, A (1993). **β-glucosidase: Biochemical and molecular biology.** Washington DC: American chemical Society.

Esen, A., and Gungor, G. (1993). Stability and activity of plant and fungal β-glucosidase under denaturing conditions. In A. Esen, (ed.). **β-glucosidase: Biochemistry and molecular biology** (pp 214-239). Washington DC: American Chemical Society.

Esashi, Y., Matsuyama, S., Ashino, H., Ogaswara, M., and Hasegawa, R. (1991). β-glucosidase activities and HCN liberation unimbibed and imbibed seed and the induction of cocklebur seed germination by cyanogenic glycosides. **Physiol. Plantarum.** 83: 43-50. Quoted in Muslim, C. (1995). **Isolation and characterization of rice (Oryza sativa L.) β-glucosidases.** Ph.D. Dissertation, Virginia Polytechnic Institute and State University, Virginia.

Fischer, L., Bromann, R., Kengen, S. W. M., de Vos, W. M.., and Wagner, F. (1996). Catalytical potency of β-glucosidase from extremophile *Pyrococcus furiosus* in glycoconjugate synthesis. **Biotechnology.** 14: 88-91.

Ganguly, T., Ganguly, S., Sircar, P., and Sircar, S. (1974). Rhamnose bound indole-3-acetic acid in the floral parts of *Peltophorum Ferrugineum*. **Physiol. Plant.** 31: 330-334.

Gaskin, P., and MacMillian, J. (1975). Polyoxygenated ent-kauranes and water-soluble conjugates in seed of *Phaseolus coccineus*. **Phytochemistry.** 14: 1575-1578.

Geerlings, A., Matinez-Lozano Ibañez, M., Memlink, J., van der Heijden, R., and Verpoorte, R. (2000). Molecular cloning and analysis of strictosidine β-D-glucosidase, an enzyme in terpeoid indole alkaloid biosynthesis in *Catharanthus roseous*. **J. Biol. Chem.** 275: 3051-3056.

Gerardi, C., Blando, F., Santino, A., and Zacheo, G. (2001). Purification and characterization of a β-glucosidase abundantly expressed in ripe sweet cherry (*Prunus avium* L.) fruit. **Plant Sci.** 160: 795-805.

Gregory, J. F. (1993). Nutrition Properties of pyridoxine-β-glucosides. In A. Esen, (ed.). **β-glucosidase: Biochemistry and molecular biology** (pp 113-131). Washington DC: American Chemical Society.

Gregory, J. F. (1998). Nutrition properties and significance of vitamin glycosides. **Annu. Rev. Nutr.** 18: 277-296.

Goff, S. A., Ricke, D., Lan, T. H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H. *et al*. (2002). A draft sequence of the rice genome *(Oryza sativa* L. ssp. *japonic*a). **Science.** 296: 92–100.

Hakulinen, N., Paavilainen, S., Korpela, T., and Rouvinen, J. (2000). The crystal structure of β-glucosidase from *Bacillus circulans* sp.*alkalophilus*: crystal

structure of β-glucosidase from *Bacillus circulans* sp.*alkalophilus*: ability to form long polymeric assemblies. **J. Struct. Biol.** 129: 69-79.

Hames, B. D. (1990). One dimensional polyacrylamide gel electrophoresis. In B. D. Hames and D. Rickwood (eds). **Gel electrophoresis of proteins-A practical approach** (2nd ed., pp 35-36). New York: Oxford University Press.

Hasegawa, M., Nakajima, M., Takeda, K., Yamaguchi, I., and Murofushi, N. (1994). A novel gibberellin glucoside 16α,17-dihydroxy-16,17-dihydrogibberellin A4-17-O-β-D-glucopyranoside, from rice anthers. **Phytochemistry.** 37: 629-634.

Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. **Biochem. J.** 280: 309-316.

Henrissat, B., and Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. **Biochem. J.** 293: 781-788.

Henrissat, B., and Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. **Biochem. J.** 316: 695-696.

Henrissat, B., 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.

Hiromi, 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. Quoted in 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.

Hoglund, A. S., Lenman, M., and Rask, L. (1992). Myrosinase is localized to the interior of myrosin grains and is not associated to the surrounding tonoplast membrane. **Plant Sci.** 85: 165-170.

Hösel, W., and Barz, W. (1975). β-Glucosidases from *Cicer arietinum* L.: purification and properties of isoflavone-7-O -glucoside-specific-β-glucosidases. **Eur. J. Biochem.** 57: 607-616.

Hösel, W., and Conn, E. E. (1982). The aglycone specificity of plant β-glucosidases . **Trends Biochem. Sci.** 7: 219-221. Quoted in Cicek,M., and Esen., A. (1998). Structure and expression of a dhurrinase (β-glucosidase) from sorghum. **Plant Physiol.** 116: 1469-1478.

Hösel, W., Tober, I., Eklund, S. H., and Conn, E. E. (1987). Characterization of β-glucosidases with high specificity for the cyanogenic glucoside dhurrin in *Sorghum bicolor* L. Moench seedlings. **Arch. Biochem. Biophys.** 252: 152-162.

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.

Hrmova, M., and Fincher, G. B. (2001a). Plant enzyme structure. Explaining substrate specificity and the evolution of function. **Plant Physiol.** 125: 54-57.

Hrmova, M. and Fincher, G. B. (2001b). Structure-function relationship of β-glucan exohydrolases from higher plants. **Plant Mol. Biol.** 47: 73-91.

Hrmova, M., De Gori, R., Smith, J. B., Fairweather, J. K., Driguez, H., Varghese, J. N., and Fincher, G. B. (2002). Structural basis broad substrate specificity in higher plant β-glucan glucohydrolases. **Plant Cell.** 14: 1033-1052.

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.

Hsieh, M., and Graham, T. L. (2001). Partial purification and characterization of a soybean β-glucosidase with high specific activity towards isoflavone conjugates. **Phytochemistry.** 58: 995-1005.

Inoue, K., and Ebizuka, Y. (1996). Purification and characterization of furostanol glycoside 26-O-β-glucosidase from *Costus speciosus* rhizomes. **FEBS Lett.** 378: 157-160.

Iwami, K., and Yasumoto, K. (1986). Synthesis of pyridoxine β-glucosidase and its in situ absorbtion in rat small intestine. **Nutr. Res.** 6: 407-414.

Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G., and Gibson, T. J. (1998). Multiple sequence alignment with Clustal X. **Trends Biochem. Sci.** 23: 403-405.

Jenkins, J., Leggio, L. 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.

Kakes, P. (1985). Linamarase and other β-glucosidases are present in cell walls of *Trifolium repens* L. leaves. **Planta.** 166: 156-160. Quoted in Cicek,M., and Esen., A. (1998). Structure and expression of a dhurrinase (β-glucosidase) from sorghum. **Plant Physiol.** 116: 1469-1478.

Kapust, R. B., and Waugh, D. S. (1999). *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. **Protein Sci**. 8: 1668-1674.

Kawasaki, S., Borchert, C., Deyholos, M., Wang, H., Brazille, S., Kawai, K., Galbraith D., and Bohnert, H. J. (2001). Gene expression profiles during the initial phase of salt stress in rice. **Plant Cell.** 13: 889-905.

Keresztessy, Z., Hughes, M. A., and Liptak, A. (1996). Co-purification from *Escherichia coli* of a plant plant β-glucosidase-glutathione S-transferase fusion protein and the bacterial GroEL. **Biochem. J.** 314: 41-47.

Keresztessy, Z., Brown, K., Dunn, M. A., and Hughes, M. A. (2001). Identification of essential active-site residues in the cyanogenic β-glucosidase (linamarase) from cassava (*Manihot esculenta* Crantz) by site-directed mutagenesis. **Biochem. J.** 353: 199-205.

Ketudat Cairns, J. R., Champattanachai, V., Srisomsap, C., Thiede, B., Wittman-Liebold, B., and Svasti, J. (2000). Sequence and expression of Thai rosewood

β-glucosidase/β-fucosidase, a family 1 glycosyl hydrolase glycoprotein. **J. Biochem.** 128: 999-1008.

Khush, G. S., and Brar, D. S. (2001). Rice genetics from model to functional In: G.S. genomics. Khush, D.S. Brar, and B. Hardy (eds). **Rice Genentics IV** (pp 3-25). Enfield, NH: Science Publishers Inc.

Laemmli, U. K. (1970). Cleavage of structural proteins during assembly of head of bacteriophage-T_4. **Nature.** 227: 680-685.

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.

Lowry, O. H., Rosenbrough, M. J., Farr, A. I., and Ramdall, R. L. (1951). Protein measurement with the folin reagent. **J. Biol. Chem.** 1951: 528-544.

Mala, S., Dvorakova, H., Hrabal, R., and Kralova, B. (1999). Towards regioselective synthesis of oligosaccharides by use of α-glucosidases with different substrate specificity. **Carbohydr. Res.** 322: 209-218.

Malboobi, M. A., and Lefebvre, D. D. (1997). A phosphate-starvation inducible β-glucosidase gene (*psr 3.2*) isolated from *Arabidopsis thaliana* is a member of a distinct subfamily of BGA family. **Plant Mol. Biol.** 34: 57-68.

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.

Mattiacci, L., Dicke, M., and Posthumus, M. A. (1995). β-glucosidase: an elicitor of herbivore-induced plant odor that attracts host-searching parasitic wasps. **Plant Biol.** 92: 2036-2040.

Mayer, A. M., and Poljakoff-Mayber, A. (1989). **The germination of seeds.** New York: Pergamon Press. pp. 111-200.

McCarter, J., and Withers, S. G. (1994). Mechanisms of enzymatic glycoside hydrolysis. **Curr. Opin. Struc. Bio.** 4: 885-892.

Millborrow, B. (1970). The metabolism of abscisic acid. **J. Exp. Bot.** 21: 17-31.

Minami, Y., Takao, H., Kanafuji, T., Miura, K., Kondo, M., Hara-Nishimura, I., Nishimura, M., and Matsubara, H. (1997). β-glucosidase in the indigo plant: intracellular localization and tissue specific expression in leaves. **Plant Cell Physiol.** 38: 1069-1074.

Minami, Y., Shigeta, Y., Tokumoto, U., Tanaka, Y., Yonekura-Sakakibara, H., Oh-oka, H., and Matsubara, H. (1999). Cloning, sequencing, characterization, and expression of a β-glucosidase cDNA from indigo plant. **Plant Sci.** 142: 219-226.

Mizutani, M., Nakanishi, H., Ema, J., Ma, S., Noguchi, E., Inohara-Ochiai, M., Fukuchi-Mizutani, M., Nakao, M., and Sakata, K. (2002). Cloning of β-primeverosidase from tea leaves, a key enzyme in tea aroma formation. **Plant Physiol.** 130: 2164-2176.

Muslim, C. (1995). **Isolation and characterization of Rice (Oryza sativa L.) β-glucosidases.** Ph.D. Dissertation, Virginia Polytechnic Institute and State University, Virginia.

Nakai, K., and Horton, P. (1999). PSORT: a program for detecting the sorting signals of proteins and predicting their subcellular localization. **Trends Biochem. Sci.** 24: 34-35.

Nicholas, K. B. and Nicholas, H. B. (1997). **Genedoc: a tool for editing and annotating multiple sequence alignments** [Online]. Available: http://www.psc.edu/biomed / gendoc.

Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. **Protein Engineer.** 10: 1-6.

Niemeyer, H. M. (1988). Hydroxamic acids (4-hydroxy-1,4-benzoxazin-3-ones), defense chemicals in the Gramineace. **Phytochemistry.** 27: 3349-3358.

Nisius, A. (1988). The stromacenter in Avena plastids: an aggregation of a β-glucosidase responsible for the activation of oat-leaf saponins. **Planta.** 173: 474-481.

Noggle, G. R., and Fritz, G. J. (1976). **Introduction plant physiology.** New Delhi: Prentice Hall of India.

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.

Perrakis, A., *et al*. (1994). Crystal structure of a bacterial chitinase at 2.3 A resolution. **Structure.** 2: 1169-1180.

Poulton, J. E., (1990). Cyanogenesis in plants. **Plant physiol.** 94: 401-405.

Poulton, J. E., and Li, C. P. (1994). Tissue level compartmentation of (R)-amygdalin

hydrolase prevent large-scale cyanogenesis in undamaged *Prunus* seeds. **Plant Physiol.** 104: 29-35.

van Rantwijk, F. Woudenberg-van Oosterom, M., and Sheldon, R. A. (1999). Glycosidase-catalysed synthesis of alkyl glycosides**. J. Mol. Catal. B-Enzym.** 6: 511–532.

Raychaudhuri, A., and Tipton P. A. (2002). Cloning and expression of the gene for soybean hydroxyisourate hydrolase. localization and implications for function and mechanism. **Plant Physiol.** 130: 2061–2068.

Reese, E. T., Maguire, A. H., and Parrish, F. W. (1968). Glucosidases and exo-glucanases. **Can. J. Biochem.** 46: 25-34. Quoted in 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.

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.

Riou, C., Salmon, J-M., Vallier, M-J., Gunata, Z., and Barre, P. (1998). Purification, characterization, and substrate specificity of a novel highly glucose-tolerant β-glucosidase from *Aspergillus oryzae*. **Appl. Env. Microb**. 64: 3607-3614.

Romeu, A., Montero, M. A., Ruiz, J. A., De Maria, M. J., Jimenez, J. A., Rojas, A., and Arola, L. (1994). Effects of glucose, ethanol, Hg(II) and Cu(II) on almond β-glucosidase. **Biochem. Mol. Biol. Int.** 33: 939-946.

Saha, S., Nagar, P. K., and Sirkar, P. K. (1984). Changes in cytokinin activity during seed germination in rice (*Oryza sativa* L.). **Ann. Bot.** 54: 1-5.

Sahi, S. V., Chilton, M-D., and Chilton, W. S. (1990). Corn metabolites after growth and virulence of *Agrobacterium tumefaciens*. **Proc. Natl. Acad. Sci.** 87: 3879.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). **Molecular cloning: a laboratory manual** (2nd ed.). New York: Cold Spring Harbor Laboratory Press.

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.

Smitt, A. R., and Van Staden, J. (1978). Changes in endogenous cytokinin levels in kernels of *Zea mays* L. durring inhibition and germination. **J. Exp. Bot.** 29: 1067-1073.

Smith, R. F., Wiese, B. A., Wojzynski, M. K., Davison, D. B., and Worley, K. C. (1996). BCM Search Launcher-An integrated interface to molecular biology database search and analysis services available on the world wide web. **Genome Res.** 6: 454-462.

Srisomsap, C., Svasti, J., Surarit, R., Champattanachai, V., Sawangareetrakul, P., Boonpuan, K., Subhasitanont, P., and Chokchaichamnankit, D. (1996). Isolation and characterization of an enzyme with β-glucosidase and β-

fucosidase activities from *Dalbergia cochinchinensis* Pierre. **J. Biochem.** 119: 585-590.

Sue, M., Ishihara, A., and Iwamura, H. (2000). Purification and characterization of a β-glucosidase from rye (*Secale cereale* L.) seedlings. **Plant Sci.** 155: 67-74.

Suganuma, T., Matsuno, R., Ohnishi, M., and Hiromi, K. (1978). A study of the mechanismof action of Taka-amylase A1 on linear oligosaccharides by product analysis and computer simulation. **J. Biochem. (Tokyo).** 84: 293-316. Quoted in 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.

Suzuki, Y., Uchida, K., and Tsuboi, A. (1979). **J. Agric. Chem. Soc. Japan.** 53: 189.

Suzuki, Y., Inada, Y., and Uchida, K. (1986). β-Glucosylpyridoxines in germinating seeds cultured in the presence of pyridoxine. **Phytochemistry.** 25: 2049-2051.

Svasti, J., Srisomsap, C., Techasakul, S., and Surarit, R. (1999). Dalcochinin-8'-O-β-D-glucoside and its β-glucosidase enzyme from *Dalbergia cochinchinensis*. **Phytochemistry.** 50: 739-743.

Svasti, J., Phongsak, T., and Sarnthima, R. (2003). Transglucosylation of tertiary alcohols using cassava β-glucosidase. **Biochem. Bioph. Res. Co.** 305: 470-475.

Tadera, K., Kaneko, T., and Yagi, F. (1988). Isolation and structural elucidation of three new pyridoxine-glycosides in rice bran. **J. Nutr. Sci. Vitaminol. (Tokyo)**. 34: 167-177.

Takahashi, N. (1984). Seed germination and seddling growth. In: S. Tsonuda, and N. Takahashi (eds). **Biology of Rice** (pp. 71-88). Oxford: Japan Scientific Societies Press.

Thayer, S. S., and Conn, E. E. (1981). Subcellular localization of dhurrin β-glucosidase and hydroxynitrile lyase in the mesophyll cells of sorghum leaf blades. **Plant Physiol.** 67: 617-622.

Ting, I. P. (1992). **Plant physiology.** Philipines: Addison-Wesley Publishing Co., Inc. pp. 585-590.

Thomas, B. R., and Rodriguez, R. L. (1994). Metabolite signals regulate gene expression and source/sink relations in cereal seedlings. **Plant Physiol.** 106: 1235-1239.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions- specific gap penalties and weight matrix choice. **Nucleic Acids Res.** 22: 4673-4680.

Tsang, E. W., Hu, Z., Chang, Q., McGregor, D. I., and Keller, W. A. (2003). Expression of a brassica napus glutamate 1-semialdehyde aminotransferase in *Escherichia coli* and characterization of the recombinant protein. **Protein Expr. Purif.** 29: 193-201.

Varghese, J. N., Garrett, T. P. J., Colman, P. M., Chen, L., Hoj, P. B., and Fincher G. (1994). The three-dimentional structures of two plant β-glycan endohydrolases with distinct substrate specificities. **Proc. Natl. Acad. Sci. USA.** 91: 2785-2789.

Verdoucq, L., Czjzek, M., Moriniere, J., Bevan, D. R., and Esen, A. (2003). Mutation and structural analysis of aglycone specificity in maize and sorghum β-glucosidases. **J. Biol. Chem.** 278: 25055-25062.

Withers, S. G. (2001). Mechanisms of glycosyl transferases and hydrolases. **Carbohyd. Polymers.** 44: 325-337.

Wang, Q., Graham, R. W., Trimbur, D., Warren, R. A. J., and Withers, S. G. (1994). **J. American Chem. Soc.** 166: 11594-11595.

Withers, S. G. (2001). Mechanisms of glycosyl transferases and hydrolases. **Carbohyd. Polymers**. 44: 325-337.

Woodward, J., and Wiseman, A. (1982). Fungal and other β-D-glucosidases: their properties and applications. **Enzyme Microb. Technol.** 4: 73.

Yasumoto, K., Tsuji, H., Iwami, K., and Mitsuda, H. (1977). Isolation from rice bran of a bound form of vitamin B6 and its identification as 5'-O-β-D-glucopyranosyl-pyridoxine. **Agric. Biol. Chem.** 41: 1061-1067.

Yu, J., Hu, S., Wang, J., Li, S., Wong, K. S. G., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X. *et al*. (2001). A draft sequence of the rice *(Oryza sativa* ssp. *indic*a) .genome. **Chin. Sci. Bull.** 46: 1937–18941.

Yu, J., Hu, S., Wang, J., Wong, G. K., Li, S., Liu, B., Deng, Y., Dai, L., Zhou, Y., Zhang, X. *et al*. (2002). A draft sequence of the rice genome *(Oryza sativa* L. ssp. *indic*a). **Science.** 296: 79–92.

Zhang, Y., and Frohman, M. A. (1997). Using rapid amplification of cDNA ends (RACE) to obtain full-length cDNAs. In I. G. Cowell and C.A. Austin (eds). **cDNA Library Protocols** (pp 61-87). New Jersey: Humana Press.

Zollner, H. (1993). **Handbook of enzyme inhibitors** (2nd ed.). New York: Part A. VCH. Quoted in Muslim, C. (1995). **Isolation and characterization of Rice (Oryza sativa L.) β-glucosidases.** Ph.D. Dissertation, Virginia Polytechnic Institute and State University, Virginia.

Appendices

Appendix A

Multiple sequence alignment and standard curve

1. Multiple sequence alignment of predicted protein sequences from rice β-glucosidase homologues found in rice genome databases

Rice cDNA derived sequences are labeled as *bglu1* and *bglu2*. Other plant β-glucosidases are labeled as: maize1 (AC AAD10503), barley (BGQ60) (AC AAA87339), and cassava (AAB71381). The number represent proteins predicted from the Rice Genome database contigs with these numbers: cg is contig clones of Beijing Genome Institute clone number, osm is the Monsanto Rice Genome Project clone number and CL.is Torrey Mesa Research Institute clone number. Sequences corresponding to the catalytic acid/base and nucleophile consensus sequences are marked by a thick line over the top. Residues shown by Czjzek *et al*. (2000) to contact the DIMBOA aglycone in the maize β-glucosidase are indicated by: ▼, while those conserved residues making contacts with the sugar are marked by: Δ above the column. Other residues shown by Keresztessy *et al*. (2001) by site directed-mutagenesis or predicted by Ketudat Cairns *et al*. (2000) based on homology modeling to play a role in substrate binding and specificity of linamarase and Thai rosewood dalcochinase are marked by: ○. Shading represents the conservation of residues within the alignment (black for the most common residue in a column and grey for similar residues). The alignment was generated using the Clustal X implementation of Clustal W (Jeanmougin *et al*., 1998; Thompson *et al*., 1994); analyzed and manually adjusted by Genedoc (Nicholas and Nicholas, 1997).

```
cg12458      : ------------------------------------------------------------------------------- :  -
CL038960.2   : MIVPLTQIFVLLTSYSTYFDRSWRFRRGRRRRPRTTRSQRTQWRGVLCICISSSS------SRRGLLLLLL---------- : 65
cg16830      : ---------------------------------------MERRPLHLHLLLLFF------SAWLMLLPLL---------- : 25
cg7736       : --------------------------------------METRRPLHLLLVFLSS------PWLLLLLLLLLVV------- : 29
cg33501      : ------------------------------------------------------------------------------- :  -
CL009530.9   : ----------------------------MCGCASGEDAMETRRPLHPLLLFFSP------WLLLLLLLVV---------- : 36
cg8650       : ------------------------------------------MAAYTSSLVS------LLLLLLLLLLVAG------- : 23
contig3896   : ----------------------------------------MAVSSSTSTCSS------FSLLLLLLLLAAASWRSGEA : 32
cg16663      : ------------------------------------------------------------------------------- :  -
osm1729      : -------------------------------------------MGAAAAAGF------FFVLLFLSVQGG-------- : 21
osm1731      : ---------------------------------------------MAAAAAF------FCALLFISVQHG-------- : 19
cg3224       : --------------------------------MKLSIKPPQPHVLTRMAAAIAV------VYLSLLLLLLHGA------- : 35
cg8125       : ----------------------------------------MAAGARALVPSP------FIVVVFLLLAAAARDAS--- : 29
CL012707.2   : ----------------------------------------MTP-ARVV----------FICCVVLLAAAAAAASSS-- : 25
bglu2        : --------------------------------------MGIRMGRR-LLFT-----------LFLGALFCN------G : 22
CL010730.5   : --------------------------------------MGIRMGRRLLLIT-----------LLLGALLCNN-----V : 24
osm11161     : -----------------------------------MANLGIGMGRQ-IVPV-----------LVFVAVLCS------G : 25
cg2878       : --------------------------------------------MDRRLLLSA-----------LLFIALACSS---N-R : 21
cg4171       : ----------------------------------------------MRWLLLA-----------LFLVALVSNG---A-A : 19
osm138100_   : -------------------------------------MAVAGAMVMSGALLLLH-----------LLAFTCVACN---GGS : 30
osm138100_   : -------------------------------------MAVAGAMVMSGGVLLL------------LLAFTCAAYN---DAG : 29
cg3708       : --------------MLSLVLSNSSLFLLQNCRSSFVQMAVAGAVAMSGGILLL------------LLLLLAAACV---EAG : 52
cg4308       : -----------------------MSLFQPLGHTHILPMAAAGSVVMLGGILLPL-----------LLVVAVS--------G : 39
cg445-1      : -------------------------------------MAAAG--AMPGGLLLTP-----------LLLAVVASSAYN-SAG : 30
osm127811    : -------------------------------------MSLLN-------LLLLL-----------LMASSTSSRSEN-KAG : 25
cassava      : -------------------------------------------------MMIF-----------LTILAVN-------- : 11
maize1       : ----------------MAPLLAAAMNHAAAHPGLRSHLVGPNNESFSRHHLPSSSPQSSKRRCNLSFTTRSARVGS-QNG : 63
barley       : -----------------------------------MRS---SPVLLLVIALVAAAH----------LAPLECDGSNS------ : 29
osm129364    : -----------------------------------MR----KFIAALRLALAAAAHLLLT-----LPPAQCYWLN------- : 31
osm125591    : -----------------------------------MGR---RLVVVLLLALLVAGAARAA-----SQAAGEDGIRGGAGADH : 39
cg3553       : -----------------------MGCARAAHYLPGGGRRRLLVVVVALVVLDR-------------AGARVRA------- : 37
bglu1        : -----------------------MAARRARCAL-------VLVLALALLAARD-------------AGAAAVP------- : 30
osm11391     : -----------------------------------MN-----MPLLLLIAIVVVSLS-------------HG-NGEQ------- : 23
cg1002       : -------------------------MGNGGRC--------MVEVVILLVLMAMS-------------QGCDAQNTT----- : 30
osm138396    : ---------------------MGRIKSSSGRCSTARLSAVAVLVVVFGVASSSL-------------RGCIAQGSG----- : 42
osm11258     : ----------------------------------------MSLLTLVHILVSF-------------SACVEAIS------ : 21
CL009851.1   : --------------------------------------------------MGS-------------TGRDAEVT------ : 11
CL003091.3   : ------------------------------------------------------------------------------- :  -
CL029716.6   : ----------------------------------------MAAAWLVVLLTVHR----------LLHLSGVS-------- : 22
CL031809.1   : -------------------------------------MAXAAATRIAVVVVVLA----------LAVLAPAAR------- : 26
cg55048      : ------------------------------------------------------------------------------- :  -
cg55438      : ------------------------------------------------------------------------------- :  -
cg40224      : -------------------------------------MRFSNWFATAALGAMA----------LFSASVSAET------ : 26
cg29552      : ------------------------------------------------------------------------------- :  -

                            Δ
cg12458      : ------------------------------------------------------------[illegible]K-[illegible]TE :   8
CL038960.2   : --QGVSSLQFR[illegible]TH--SGRHPEDGTG[illegible]K-[illegible]TE : 141
cg16830      : --QGVSSLQFR[illegible]TH--SGRHPEDGTG[illegible]K-[illegible]TE : 101
cg7736       : --QGMSSLQFT[illegible]TH--SGRHPKDETG[illegible]K-[illegible]KE : 105
cg33501      : ------------------------------Y[illegible]AH--SWRNPGGKT[illegible]M-[illegible]NE :  49
CL009530.9   : --QGVRSLQFT[illegible]AH--SWRNPGGET[illegible]NE : 113
cg8650       : EATAEAALNFT[illegible]TH--SGRMADNSTG[illegible]K-[illegible]D : 101
contig3896   : AAAAARALNFT[illegible]TH--AGKMPDKSTG[illegible]K-[illegible]D : 110
cg16663      : ---------------------------------------------------KTKDKGTG[illegible]K-[illegible]AE :  27
osm1729      : -----AVG-YT[illegible]AH--EGKTKDKGTG[illegible]K-[illegible]AE :  93
osm1731      : -----VLGGYT[illegible]TH--AGKMKDKSTG[illegible]K-[illegible]TE :  92
cg3224       : --APAVLG-YT[illegible]TH--QGKMPGKSTA[illegible]K-[illegible]VD : 110
cg8125       : ---------ALT[illegible]TH--QGVSPGGAIA[illegible]K-[illegible]VD :  99
CL012707.2   : -----RRRGIT[illegible]SH--SGYSVDGATG[illegible]K-[illegible]QD :  98
bglu2        : VYAK-----FT[illegible]TH-IPGKILANDNG[illegible]G-[illegible]KD :  96
CL010730.5   : AYAK-----FS[illegible]TH-IPGKILANDTG[illegible]N-[illegible]KD :  98
osm11161     : VDAS-----FN[illegible]SH-IPGKILAGDTG[illegible]N-[illegible]KD :  99
cg2878       : VHGA-----LN[illegible]SH-IPGKIADGSNA[illegible]N-[illegible]TA :  93
cg4171       : VHGA-----FN[illegible]AH-IPGKVKDGSNG[illegible]N-[illegible]TD :  93
osm138100_   : ELPP-----IS[illegible]THQSPDKITDKSNG[illegible]R-[illegible]KE : 105
osm138100_   : ELPP-----IS[illegible]THQYPDKITDKSNG[illegible]R-[illegible]KE : 104
cg3708       : ELPP-----IS[illegible]TYGHPDKIADKSNG[illegible]N-[illegible]KE : 127
cg4308       : EPPP-----IS[illegible]THGHPDKIADKSNG[illegible]R-[illegible]KD : 114
cg445-1      : EPP------VS[illegible]THGHPEKIADKSNG[illegible]R-[illegible]KD : 104
osm127811    : EV-------IR[illegible]THNHPEKIANGSNG[illegible]G-[illegible]KG :  98
cassava      : ---I-----FRMTS[illegible]FP--RNRIIDGRNG[illegible]K-[illegible]KK :  75
maize1       : VQMLSPREIPQ[illegible]CNNHPERILDGSNS[illegible]R-[illegible]KE : 143
barley       : NPEIGNPGGLS[illegible]VAIQG-MIAGNGTA[illegible]G-[illegible]KN : 108
osm129364    : -PEIYDAGGLS[illegible]IEKPG-TIP[illegible]NATA[illegible]N-[illegible]KN : 109
osm125591    : QEAAGITGGLS[illegible]VKTPG-EIA[illegible]NATA[illegible]N-[illegible]KE : 118
cg3553       : --ADDDTGGLS[illegible]VHTPG-MIAGNGNA[illegible]D-[illegible]KE : 114
bglu1        : --KIWWLGGLS[illegible]AHTPG-MVAGNQNG[illegible]N-[illegible]KE : 107
osm11391     : -----TD--LT[illegible]LMQPG-VTP[illegible]NTA[illegible]D-[illegible]VR :  95
cg1002       : -----GG--LT[illegible]AHTPG-KIIDPSNA[illegible]Q-[illegible]AD : 102
osm138396    : -----GG--LT[illegible]AHTPG-KITDPSNA[illegible]Q-[illegible]AD : 114
osm11258     : ------------[illegible]LTKRPG-RVIDPSNA[illegible]E-[illegible]ND :  89
CL009851.1   : ------------[illegible]TE--N-K--DGSSG[illegible]E-[illegible]AS :  75
CL003091.3   : ------------------------------[illegible]THLPG-MIKDGSNG[illegible]E-[illegible]NS :  49
CL029716.6   : ---------AVD[illegible]THKGG-TIKDGSNG[illegible]E-[illegible]NS :  93
CL031809.1   : ---------GLR[illegible]THTQAGRIKDGSNG[illegible]D-[illegible]NN :  98
cg55048      : -----------M[illegible]PARRPNRFHHGVGPQQTSTP[illegible]Q-[illegible]KQ :  66
cg55438      : -----------MEYQ[illegible]AACI[illegible]RN[illegible]YKIAPERFHDRVGPASASTP[illegible]G-[illegible]KT :  70
cg40224      : ----------KGTRC[illegible]CR----EKPGVMCA[illegible]K-[illegible]VE :  93
cg29552      : -----------MNTSL[illegible]IREAG-----EMISDWK[illegible]YD-[illegible]QD :  65
```

Sequence	End position
cg12458	223
CL038960.2	356
cg16830	316
cg7736	320
cg33501	265
CL009530.9	329
cg8650	324
contig3896	325
cg16663	243
osm1729	309
osm1731	308
cg3224	326
cg8125	314
CL012707.2	314
bglu2	313
CL010730.5	315
osm11161	316
cg2878	312
cg4171	310
osm138100_	322
osm138100_	321
cg3708	344
cg4308	331
cg445-1	321
osm127811	315
cassava	292
maize1	362
barley	319
osm129364	320
osm125591	328
cg3553	325
bglu1	318
osm11391	306
cg1002	317
osm138396	329
osm11258	304
CL009851.1	281
CL003091.3	255
CL029716.6	308
CL031809.1	313
cg55048	265
cg55438	269
cg40224	307
cg29552	266

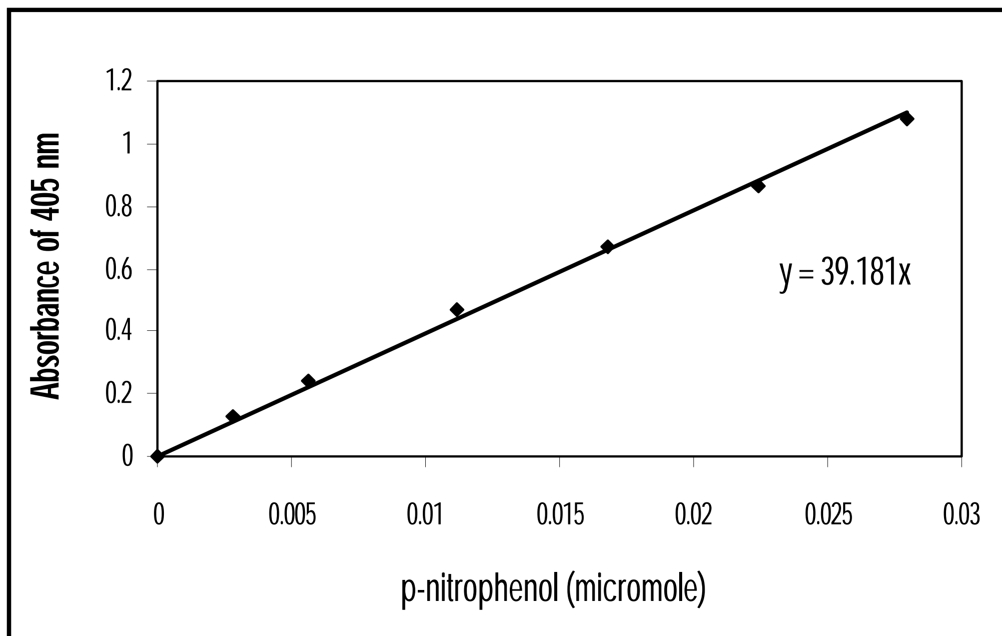
Sequence	End position
cg12458	290
CL038960.2	423
cg16830	383
cg7736	387
cg33501	332
CL009530.9	398
cg8650	384
contig3896	392
cg16663	310
osm1729	374
osm1731	376
cg3224	391
cg8125	375
CL012707.2	371
bglu2	380
CL010730.5	381
osm11161	383
cg2878	375
cg4171	378
osm138100_	389
osm138100_	388
cg3708	411
cg4308	398
cg445-1	388
osm127811	382
cassava	360
maize1	431
barley	386
osm129364	387
osm125591	395
cg3553	392
bglu1	385
osm11391	371
cg1002	386
osm138396	398
osm11258	373
CL009851.1	347
CL003091.3	322
CL029716.6	376
CL031809.1	381
cg55048	332
cg55438	336
cg40224	385
cg29552	329

```
cg12458       : 357
CL038960.2    : 490
cg16830       : 412
cg7736        : 457
cg33501       : 402
CL009530.9    : 464
cg8650        : 454
contig3896    : 463
cg16663       : 378
osm1729       : 442
osm1731       : 456
cg3224        : 459
cg8125        : 448
CL012707.2    : 444
bglu2         : 453
CL010730.5    : 411
osm11161      : 458
cg2878        : 450
cg4171        : 453
osm138100_    : 464
osm138100_    : 463
cg3708        : 486
cg4308        : 473
cg445-1       : 463
osm127811     : 457
cassava       : 436
maize1        : 508
barley        : 459
osm129364     : 460
osm125591     : 468
cg3553        : 465
bglu1         : 458
osm11391      : 444
cg1002        : 462
osm138396     : 474
osm11258      : 449
CL009851.1    : 422
CL003091.3    : 397
CL029716.6    : 451
CL031809.1    : 454
cg55048       : 406
cg55438       : 365
cg40224       : 453
cg29552       : 402
```

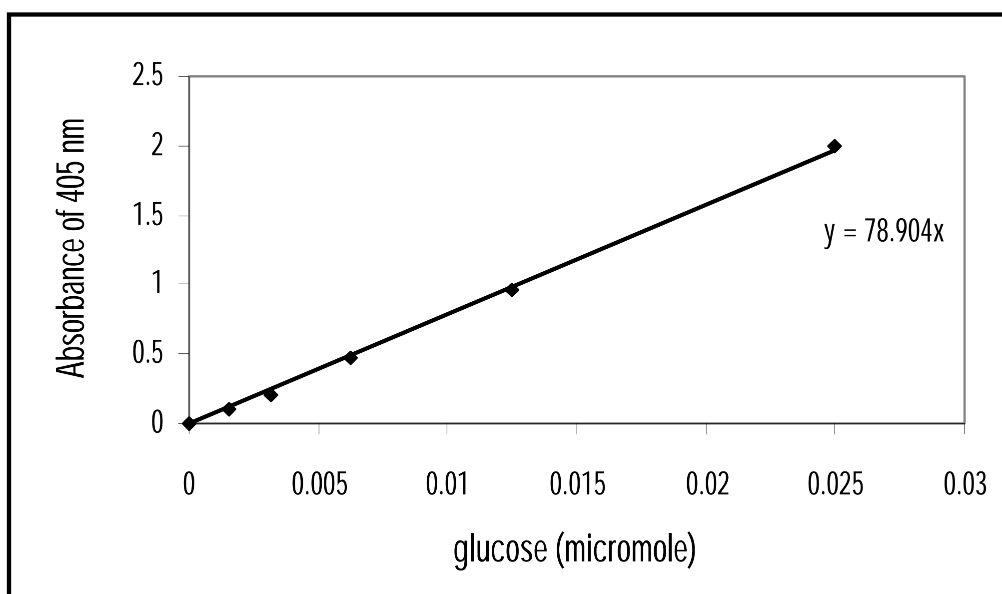
```
cg12458       : 414
CL038960.2    : 547
cg16830       :  -
cg7736        : 514
cg33501       : 457
CL009530.9    : 519
cg8650        : 510
contig3896    : 519
cg16663       : 437
osm1729       : 492
osm1731       : 506
cg3224        : 512
cg8125        : 501
CL012707.2    : 497
bglu2         : 500
CL010730.5    :  -
osm11161      : 508
cg2878        : 496
cg4171        : 499
osm138100_    : 520
osm138100_    : 510
cg3708        : 537
cg4308        : 520
cg445-1       : 510
osm127811     : 504
cassava       : 492
maize1        : 564
barley        : 509
osm129364     : 510
osm125591     : 516
cg3553        : 520
bglu1         : 504
osm11391      : 492
cg1002        : 510
osm138396     : 521
osm11258      : 501
CL009851.1    : 478
CL003091.3    : 445
CL029716.6    : 499
CL031809.1    : 516
cg55048       : 416
cg55438       :  -
cg40224       : 523
cg29552       : 458
```

```
cg12458      : APLSLQPLMLNSE----------------------------------- : 427
CL038960.2   : GSFVSAASHAQL------------------------------------ : 559
cg16830      : ------------------------------------------------ :   -
cg7736       : DGFVSAASHAQL------------------------------------ : 526
cg33501      : DGPVSTAFHAQL------------------------------------ : 469
CL009530.9   : DGSVSTAFHAQL------------------------------------ : 531
cg8658       : MISPDEHEHAYYQ----------------------------------- : 523
contig3896   : TISPDEHEHEHADQ---------------------------------- : 533
cg16663      : SAVNRSLALANQIAQEIER----------------------------- : 456
osm1729      : ----DSVLIASQ------------------------------------ : 500
osm1731      : ----QSLQIAQQ------------------------------------ : 514
cg3224       : DQSSLAMAAAKSRAQQ-------------------------------- : 528
cg8125       : VKPPRSQ--RY------------------------------------- : 510
CL012707.2   : AAAALAGGGAYSQ----------------------------------- : 510
bglu2        : ------------------------------------------------ :   -
CL010730.5   : ------------------------------------------------ :   -
osm11161     : ------------------------------------------------ :   -
cg2878       : ------------------------------------------------ :   -
cg4171       : ------------------------------------------------ :   -
osm138100_   : KRVAYNAFSN-------------------------------------- : 530
osm138100_   : ------------------------------------------------ :   -
cg3708       : ------------------------------------------------ :   -
cg4308       : ------------------------------------------------ :   -
cg445-1      : ------------------------------------------------ :   -
osm127811    : ------------------------------------------------ :   -
cassava      : CPSKDSKKVGKFYVM--------------------------------- : 507
maize1       : SA---------------------------------------------- : 566
barley       : ------------------------------------------------ :   -
osm129364    : ------------------------------------------------ :   -
osm125591    : ------------------------------------------------ :   -
cg3553       : GTSDSQVGSATSASHFVGSAISSSHHLLLPLLLSLHFLFPSFFMFLSL : 568
bglu1        : ------------------------------------------------ :   -
osm11391     : ------------------------------------------------ :   -
cg1002       : ------------------------------------------------ :   -
osm138396    : ------------------------------------------------ :   -
osm11258     : ------------------------------------------------ :   -
CL009851.1   : N----------------------------------------------- : 479
CL003891.3   : ------------------------------------------------ :   -
CL029716.6   : ------------------------------------------------ :   -
CL031809.1   : ------------------------------------------------ :   -
cg55048      : ------------------------------------------------ :   -
cg55438      : ------------------------------------------------ :   -
cg40224      : ASSADASSVSYVGYGIGGIAAVGVIVAAAVVHQRRRGYTSIPQN---- : 567
cg29552      : ------------------------------------------------ :   -
```

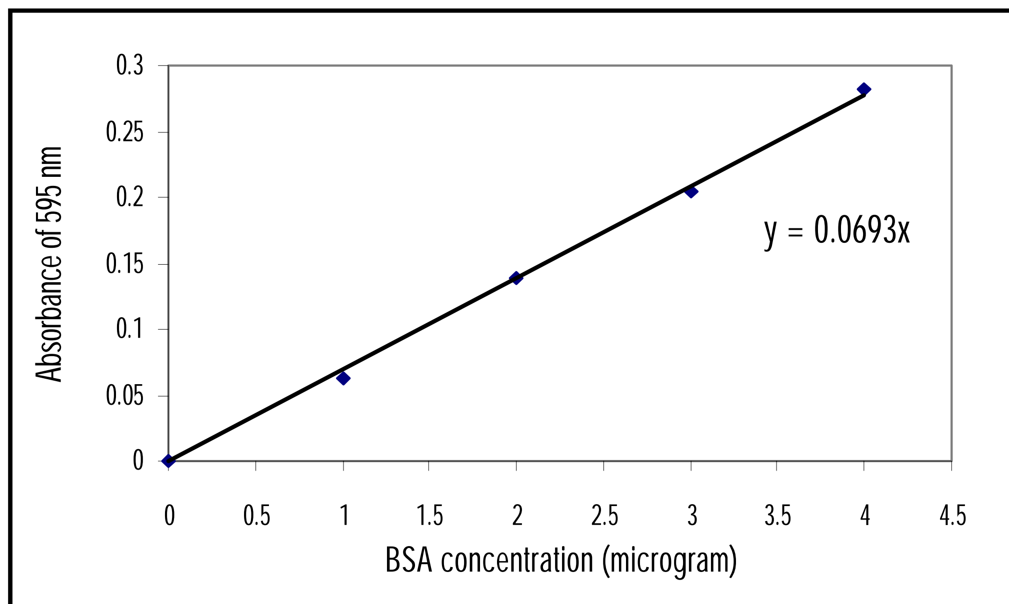
2. Standard curve of p-nitrophenol



3. Standard curve of glucose



4. Stand curve of BSA by Bio-Rad Protein Assay

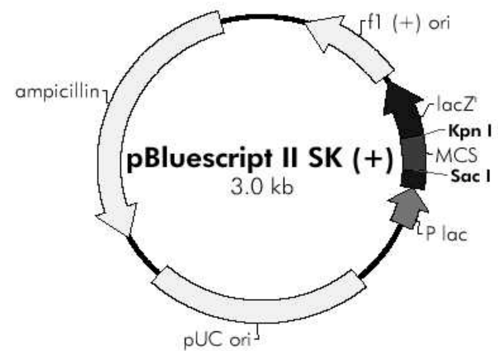


Appendix B

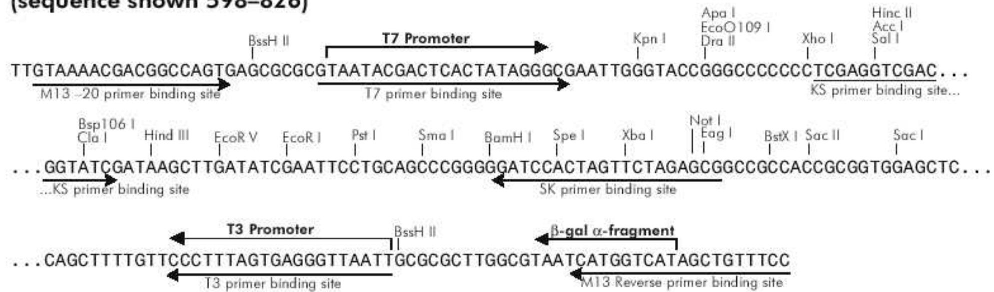
Plasmid maps

1. pBlueScript II SK(+)

f1 (+) origin 135–441
β-galactosidase α-fragment 460–816
multiple cloning site 653–760
***lac* promoter** 817–938
pUC origin 1158–1825
ampicillin resistance (*bla*) ORF 1976–2833

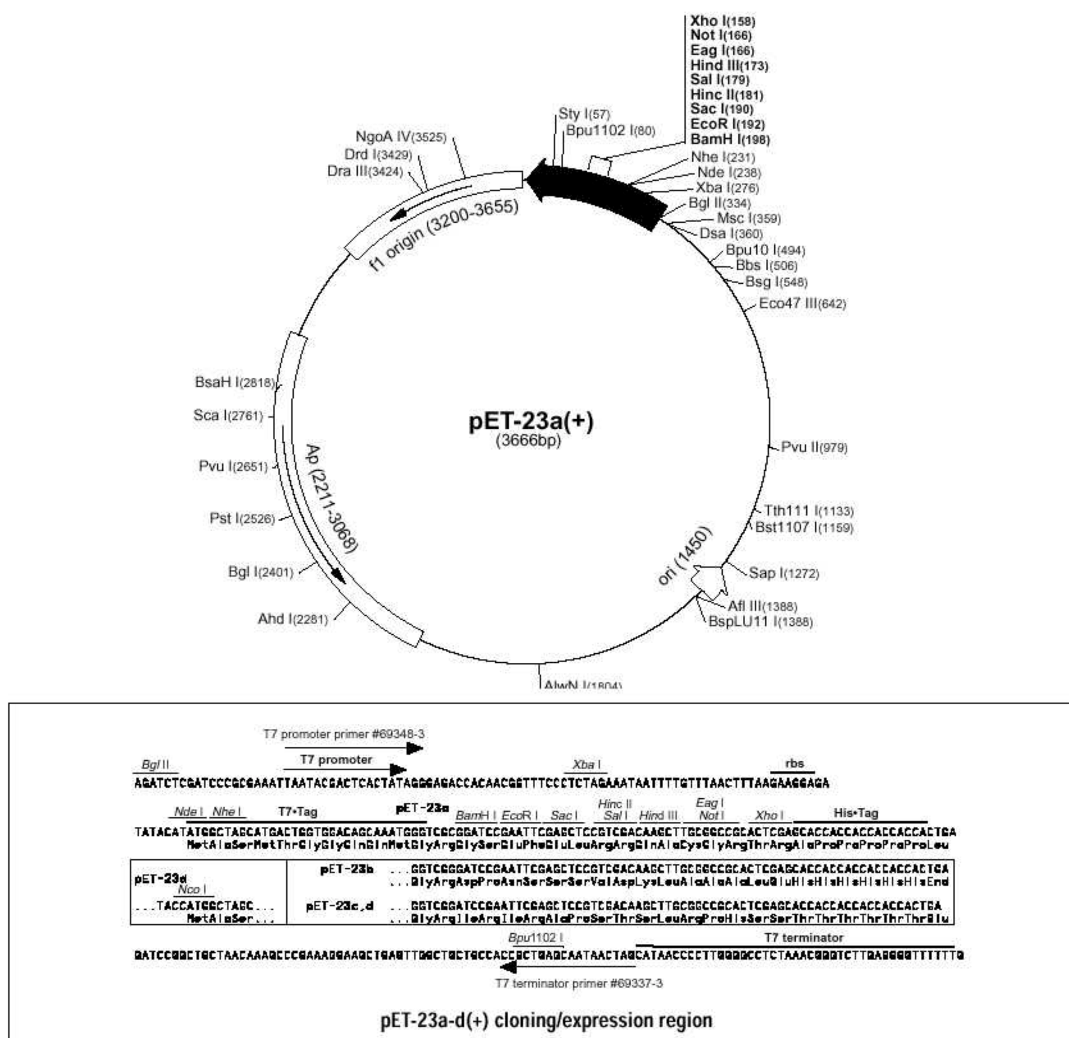


pBluescript II SK (+/–) Multiple Cloning Site Region
(sequence shown 598–826)



2. pET-23d(+) Vector

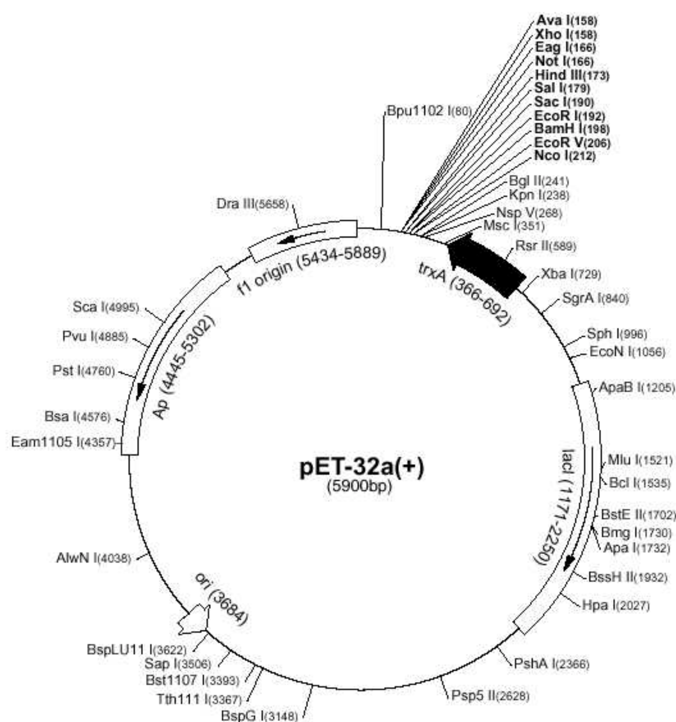
The pET-23d(+) vector carry an N-terminal T7•Tag sequence plus an optional C-terminal His•Tag sequence. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown below. The f1 origin is oriented so that infection with helper phage will produce virions containing single-stranded DNA that corresponds to the coding strand. Therefore, single-stranded sequencing should be performed using the T7 terminator primer.

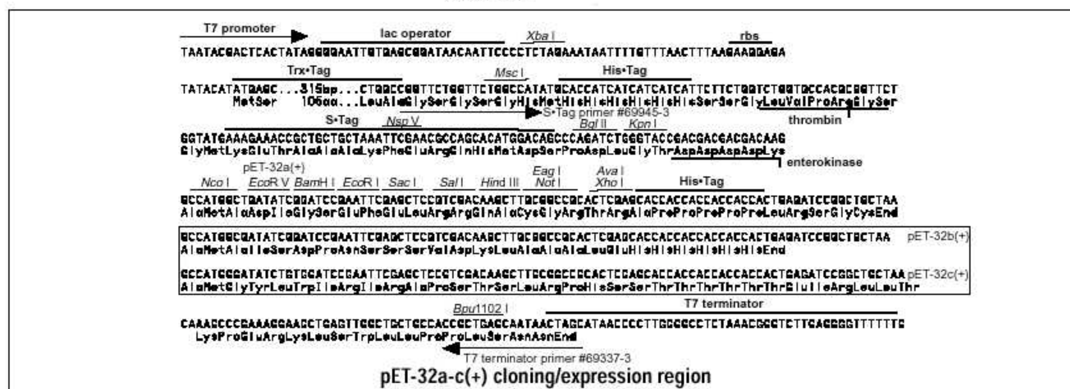


pET-23a-d(+) cloning/expression region

3. pET32a(+) Vector

The pET-32a(+) vector is designed for cloning and high-level expression of peptide sequences fused with the 109 amino acids of thioredoxin protein (Trx·Tag). Cloning sites areavailable for producing fusion proteins also containing cleavable His6·Tag and S·Tag sequences for detection and purification. The sequence is numbered by the pBR322 convention, so the T7 expression region is reserved on the circle map. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown below.

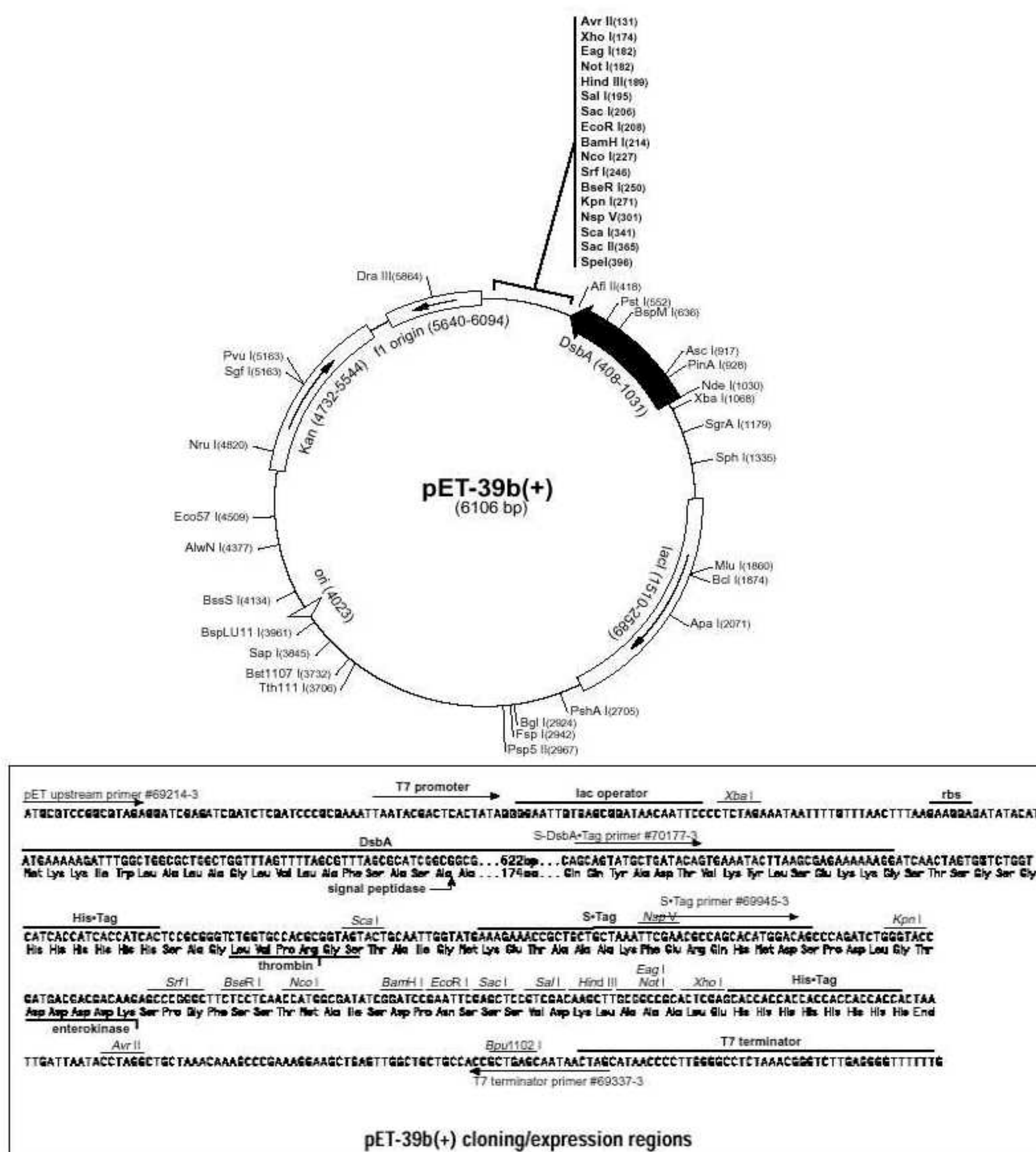




pET-32a-c(+) cloning/expression region

4. pET39b(+) vector

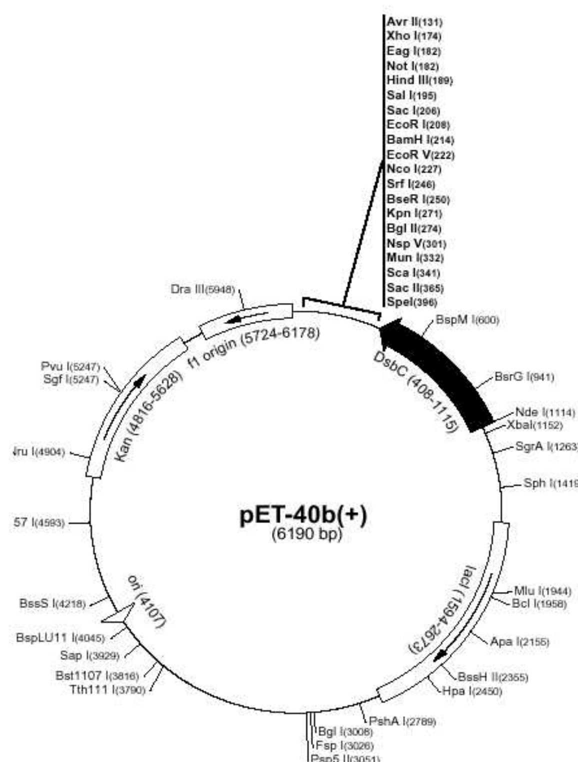
The pET39b(+) vector is designed for expression of DsbA fusion proteins. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown below. The f1 origin is oriented so that infection with helper phage will produce virions containing single stranded DNA that corresponds to the coding strand. Therefore, single stranded sequencing should be performed using the T7 terminator primer.

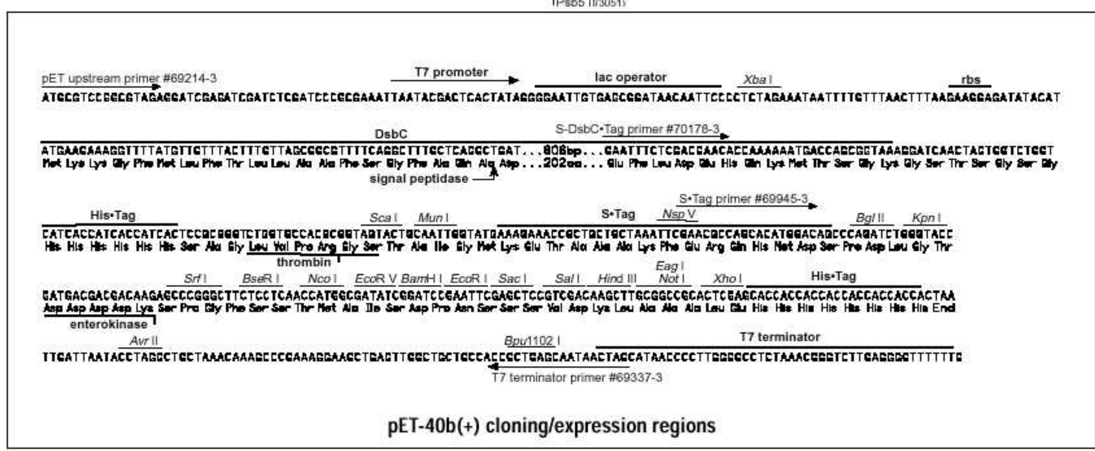


pET-39b(+) cloning/expression regions

5. pET40b(+)

The pET40b(+) vector is designed for expression of DsbC fusion proteins. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown below. The f1 origin is oriented so that infection with helper phage will produce virions containing single stranded DNA that corresponds to the coding strand. Therefore, single stranded sequencing should be performed using the T7 terminator primer.





pET-40b(+) cloning/expression regions

Appendix C

Competent cell preparation and antibody production

1. Competent cell preparation

E. coli DH5α strain, a cloning host strain, was used to prepare competent cells by using $CaCl_2$ methods as described by Sambrook *et al*. (1989). A single colony of DH5α *E. coli* was inoculated into 5 ml of LB broth (Appendix D) and grown at 37°C overnight (12-14 h) with 160 rpm shaking. The overnight culture was used as inoculum for preparation of competent cells. One milliliter of inoculum culture (1%) was added into a 500 ml flask containing 100 ml of LB broth and grown at 37°C with 240 rpm shaking until the optical density (OD) at 600 nm was about 0.3-0.4 (about 3 h). The cell culture was transferred into a pre-chilled polypropylene tube, chilled on ice for 10 min, and pelleted by centrifugation at 3000 x g at 4°C, 10 min. The cell pellets were resuspended completely and gently with 20 ml of ice-cold $CaCl_2$ Solution (60 mM $CaCl_2$, 10 mM PIPES pH 7.0, 15% glycerol) (Appendix D). The resupended cells were pelleted by centrifugation as described above. The cell pellets were resuspended completely and gently again with 20 ml of ice-cold $CaCl_2$ Solution and chilled on ice for 30 min. The pellets were centrifuged as above, resuspended in 2 ml ice-cold $CaCl_2$ solution containing 7% dimethyl sulfoxide (DMSO), and aliquoted into microtube (100 μl/tube). These competent cells were immediately kept at -70°C in a freezer until transformation.

2. Antibody production

Polyclonal antibodies are produced in rabbits with specific peptides derived from BGlu1 and BGlu2 protein sequences and BGlu2 recombinant protein produced in *E. coli*.

2.1 Peptide synthesis as antigens

Short synthetic peptides were used to generate specific antibodies for BGlu1 and BGlu2 protein. The peptide sequences were derived from deduced amino acid sequences which were unique to the BGlu1 and BGlu2 proteins. The peptide sequences were selected from a variable loop region among the deduced amino acid sequences of rice β-glucosidase genes identified from the Monsanto Rice Genome Database (see method 2.11). Peptides were synthesized by Genemed Synthesis, Inc., South San Francisco, CA. The selected BGlu1 peptide sequence was NH_3-CGKGQALMQQTPTSYS-$CONH_2$, and BGlu2 was NH_3-CGKSIPPPNSNELSYD-$CONH_2$.

2.2 Peptide-carrier conjugation

Two milligrams peptide was dissolved in 300 μl conjugation buffer (83 mM sodium phosphate buffer, 0.1 M EDTA, 0.9 M NaCl, 0.02% sodium azide, pH 7.2). Next, 2 mg pre-activated carrier protein, Imject supercarrier BSA or mcKLH (PIERCE, Rockford, IL) was dissolved in 200 μl distilled water. Then, the peptide and activated carrier solutions were mixed immediately by pipeting, and the mixture was allowed to react for 2 h at room temperature. To remove EDTA, an anti-coagulant which should not be injected into laboratory animals, the conjugate was

purified by gel filtration. The conjugates were applied to Dextran gel filtration columns (with molecular weight cut-off 5,000), and eluted with 0.083 M sodium phosphate, 0.9 M NaCl, pH 7.2. The absorbance at 280 nm of the collected fractions was measured to find the fractions containing the conjugate. The conjugate was kept frozen at -20°C until immunoinjection.

2.3 Purification of inclusion bodies as an antigen

One gram cell pellet of BL21(DE3) *E. coli*, which expressed BGlu2 as inclusion bodies from pET23d(+) as described in method 2.7, was extracted to remove soluble proteins 10 times with 5 ml of 20 mM Tris-HCl pH 8.0, 1% Triton X-100 buffer in each time and centrifuged to pellet the cells. The cell pellet was resuspended in 500 µl 2 X SDS sample buffer. The suspension was boiled for 5 min and loaded onto 12% SDS-PAGE gels. After electrophoresis, proteins were visualized by staining with Coomassie Brilliant Blue R 250. The expected size protein band was cut from the gels and ground in a mortar and pestle to fine pieces. The gel was boiled in 2 ml of 100 mM sodium phosphate buffer pH 7.4 containing 0.5% SDS and 5% 2-mercaptoethanol at 75°C, for 15 min. The gel was centrifuged and the pellet discarded. The supernatant was kept at
-20°C until immunoinjection.

2.4 Antigen injection

BGlu1 and BGlu2 conjugated peptides and BGlu2 inclusion body protein immunogen were injected into rabbits coded 403, 406 and 404 respectively. The injection and bleed procedure was done as follows:

Day	Procedure	Amounts (v/v) per rabbit
0	Pre-bleed Primary injection	5 ml serum for 403-0,404-0, 406-0 (1:1) 400 µg protein or 200 µg conjugated peptide solution in Freund's Complete adjuvant
14	First boost injection	(1:1) 400 µg protein or 200 µg conjugated peptide solution in Freund's Incomplete adjuvant
21	First immune bleed (1)	5 ml serum for 403-1, 404-1, 406-1
28	Second boost injection	(1:1) 400 µg protein or 100 µg conjugated peptide solution in Freund's Incomplete adjuvant
35	Second immune bleed (2)	5 ml serum for 403-2, 404-2, 100 ml for 406-2 (Last bleed)
42	Third boost injection	(1:1) 300 µg protein or 100 µg conjugated peptide solution in Freund's Incomplete adjuvant
49	Third immune bleed (3)	5 ml serum for 403-3, 100 ml for 404-3 (last bleed)
56	Fourth boost injection	(1:1) 100 µg conjugated peptide solution in Freund's Incomplete adjuvant
63	Fourth immune bleed (4)	100 ml for 403-4 (last bleed)

After the immune blood samples were taken from each rabbit, the blood samples were put in refrigerator until clotting (about 6-12 h). The serum (upper phase) was collected by centrifugation at 3,000 x g for 10 min. The lytic complement activity of serum was abolished by heating at 56°C for 30 min. For long-term storage, antibodies were aliquoted in to many tubes and kept at -30°C. Some aliquots were stored in 50% glycerol at -30°C.

2.5 ELISA Titration of Anti-Serum

The serum was taken to check the production of specific antibodies after the second injection by ELISA. The ELISA test was performed to determine the interaction between the antigen and the antibody by calculating the titer of a specific polyclonal antiserum. 96-well microtiter plates coated with 100 µl peptide (5 µg/ml) or purified recombinant protein (5 µg/ml) prepared in coating buffer (50 mM $NaHCO_3/Na_2CO_3$ buffer, pH 9.5). The plates were incubated at 4°C overnight on an orbital shaker. Unbound peptides or proteins were removed by washing 4 times with 200 µl PBST buffer (Appendix) and padded dry on paper towels. Serum dilutions were prepared by diluting the serum with PBST. Serial dilutions were made as follow:

A. sample preparation for pre-bleed serum (control) and test bleed

B. 1:100 (100 µl A + 990 µl PBST) G. 1:3,200 (500 µl F + 500 µl PBST)

C. 1:200 (500 µl B + 500 µl PBST) H. 1:6,400 (500 µl G + 500 µl PBST)

D. 1:400 (500 µl C + 500 µl PBST) I. 1:12,800 (500 µl H + 500 µl PBST)

E. 1:800 (500 µl D + 500 µl PBST) J. 1:25,600 (500 µl I + 500 µl PBST)

F. 1:1,600 (500 µl E + 500 µl PBST)

One hundred microliter of serum samples were added to wells (Starting from B), run in duplicate and incubated at room temperature for 2 h. Then, the serum dilutions were removed and the wells were washed immediately and thoroughly 4 times with 200 µl PBST. One hundred microliters of 1:2000 dilution in PBST of the secondary

antibody, goat anti-rabbit coupled to horse radish peroxidase, was added to each well, and incubated at 37°C for 2 h. The solutio0n was discarded and the wells were washed immediately and thoroughly 4 times with 200 µl PBST. One hundred microliters of 1 mg/ml 2,2′-azinobis(3-ethylbenzthiazolinesulfonic acid) (ABTS) in citric acid pH4.0 substrate plus 0.006% hydrogen peroxide was added to the wells and the green product color of oxidized ABTS was developed at room temperature for 5-10 min. One hundred microliters of 10 mM sodium azide was added to each well to stop the color development and the absorbance read immediately at 405 nm. The titer value is estimated to be the dilution that yields a A_{405} value similar to that of the lowest dilution when tested against a fixed (the same) dilution of the test antigen.

Appendix D

Solution preparation

1. Reagents for baterial culture and competent cell transformation

1.1 LB broth containing antibiotics (1 L)

Dissolve 10 g Bacto Tryptone, 5 g Bacto Yeast Extract and 5 g NaCl in 800 ml distilled water. Adjust pH to 7.2 with NaOH and the volume to 1 L with distilled water. Autoclave the solution at 121°C for 15 min. Allow the medium to cool to 50°C before adding antibiotics with concentration recommended for each cloning system and store at 4°C.

1.2 LB plate with 100 μg/ml of ampicillin (1 L)

Dissolve 10 g Bacto Tryptone, 5 g Bacto Yeast Extract, 5 g NaCl and 15 g Bacto agar in 800 ml distilled water. Adjust pH to 7.2 with NaOH and the volume to 1 L with distilled water. Sterilize by autoclaving at 121°C for 20 min. Allow the medium to cool to 50°C, then add ampicillin to a final concentration 100 μg/ml. Pour medium into Petri-dishes. Allow the agar to harden, and keep at 4°C.

1.3 LB agar plate with 100 μg/ml of ampicillin, 15 μg/ml kanamycin, 12.5 μg/ml tetracyclin (1 L)

Dissolve 10 g Bacto Tryptone, 5 g Bacto Yeast Extract, 5 g NaCl and 15 g Bacto agar in 800 ml distilled water. Adjust pH to 7.2 with NaOH and the volume to

1 L with distilled water. Sterilize by autoclaving at 121°C for 20 min. Allow the medium to cool to 50°C, then add ampicillin to a final concentration 50 μg/ml, kanamycin 15 μg/ml, 12.5 μg/ml tetracyclin. Pour medium into Petri-dishes, allow the agar to harden, and keep at 4°C.

1.4 LB plate with 100 μg/ml of ampicillin/IPTG/X-Gal

Make the LB plates with ampicillin as above. Then spread 50 μl of 100 mM IPTG and 25 μl of 50 mg/ml X-Gal over the surface of the plates and allow to absorb for 10 min before use.

1.5 Antibiotics solution stock

Ampicillin (50 mg/ml): dissolve 50 mg ampicillin in 1 ml sterile distilled water.

Kanamycin (30 mg/ml): dissolve 30 mg kanamycin in 1 ml sterile distilled water.

Tetracyclin (12.5 mg/ml): dissolve 12.5 mg tetracyclin in 1 ml absolute alcohol.

Chloramphenicol (34 mg/ml): dissolve 34 mg chloramphenicol in 1 ml absolute alcohol.

Filter sterile all antibiotics solution and keep at -20°C.

1.6 IPTG stock solution (100 mM)

Dissolve 0.12 g IPTG (isopropyl thio-β-D-galactoside) in distilled water and make to 5 ml final volume. Sterilize by filter sterilize and store at -20°C.

1.7 X-gal stock solution

Dissolve 50 mg X-gal in DMF and store in the dark bottle at -20°C.

2. Reagent for competent cell preparation

2.1 $CaCl_2$ Solution (60 mM $CaCl_2$, 10 mM PIPES pH 7.0, 15% glycerol)

To prepare 100 ml solution, mix the stock solution as follow:

- 6 ml of 1 M $CaCl_2$ (14.7 g/100 ml, filter sterile)
- 10 ml of 100 mM PIPES (piperazine-1,4-bis(2-ethanesulfonic acid)) pH 7.0 (3.02 g/100 ml adjust pH with KOH, filter sterile)
- 15 ml of 100% glycerol (autoclave at 121°C, 15 min)

Add sterile distilled water to bring a volume up to 100 ml. Store the solution at 4°C.

2.2 SOC media (1 L)

Dissolve 20 g Bacto Tryptone, 5 g Bacto Yeast Extract, 10 ml 1M of NaCl (5.85 g/100 ml), 2.5 ml 1 M KCl (7.44 g/100 ml) in distilled water and make to 980 ml final volume. Sterilize by autoclaving at 121°C for 20 min. Allow the medium cool to room temperature, then add 5 ml of 1 M $MgCl_2 \cdot 6H_2O$ (20.33 g/100 ml), 5 ml of 1 M $MgSO_4 \cdot 7 H_2O$ (12.30 g/100 ml), 10 ml of 2 M of glucose (36 g/100 ml) which were filter sterilized. Store the solution at 4°C.

3. Reagents for isolation plasmid DNA (boiling prep)

3.1 STET (100 mM NaCl, 10 mM Tris pH 8.0, 1 mM EDTA, 5% Triton X-100) (100 ml)

Mix 20 ml of 1 M NaCl (sterile), 2 ml of 1 M Tris-HCl pH 8.0 (sterile), 0.4 ml of 0.5 M EDTA pH 8.0 (sterile), 5 ml Triton X-100 and adjust the volume to 100 ml with sterile distilled water. Store at room temperature.

3.2 3 M Sodium acatate pH 4.8 (100 ml)

Dissolve 24.6 g sodium acetate in 80 ml distilled water. Adjust pH to 4.8 with glacial acetic acid and the volume to 100 ml with distilled water. Sterilize by autoclaving at 121°C for 20 min. Store at room temperature.

3.3 RNase A (10 mg/ml)

Dissolve 10 mg RNase A in 10 mM Tris-HCl pH 7.4, 15 mM NaCl buffer (sterile). Store at -20°C.

3.4 1 M Tris-HCl pH 7.4 and pH 8.0 (100 ml)

Dissolve 12.11 g of Tris Base in 80 ml distilled water. Adjust pH with HCl to pH 7.4 or 8.0 and a volume to 100 ml with distilled water and autoclave at 121°C for 20 min. Store at room temperature.

3.5 0.5 M EDTA (pH 8.0) (100 ml)

Dissolve 18.61 g EDTA (disodium ethylene diamine tetraacetate $\cdot 2H_2O$) in 70 ml distilled water. Adjust pH to 8.0 with NaOH (about 20 g) and the volume to 100 ml with distilled water. Sterilize by autoclaving at 121°C for 20 min. Store at room temperature.

4. Reagent for agarose gel electrophoresis

4.1 50 X TAE for agarose gel electrophoresis (1 L)

Mix 242 g Tris base, 57.1 ml glacial acetic acid, 100 ml 0.5 M EDTA (pH 8.0) and adjust the volume to 1 L with distilled water. Store at room temperature.

4.2 6 X DNA loading dye (10 ml)

Mix 0.025 g Bromophenol blue, 0.025 g xylene cyanol and 3 ml of 100% glycerol in distilled water to a 10 ml final volume and store at 4°C.

5. Reagents for total RNA isolation by using CTAB

5.1 DEPC treated water (1 L)

Add 1 ml of diethyl pyrocarbonate (DEPC) in distilled water and make to 1L final volume. Shake vigorously overnight and autoclave the solution at 121°C for 15 min to inactivate the remaining DEPC. Store at room temperature.

5.2 2 X CTAB solution (100 ml) (2% Cetyltrimethylamonium bromide (CTAB), 0.1 M Tris-HCl pH 9.5, 20 mM EDTA, 1.4 M NaCl, 5% 2-mercabtoethanol)

Dissolve 2 ml CTAB, 1.21 g Tris base, 7.44 g EDTA and 8.19 g NaCl in distilled water and adjust the volume to 95 ml. Sterilize by autoclaving at 121°C for 40 min and add 1 ml 2-mercaptoethanol just prior to use.

5.3 10 M LiCl (100 ml)

Dissolve 42.39 g LiCl in distilled water to 100 ml final volume. Sterilize by autoclaving at 121°C for 20 min. Store at room temperature.

5.4 3 M Sodium acetate pH 5.2 (100 ml)

Dissolve 24.6 g sodium acetate in 80 ml distilled water. Adjust pH 5.2 with glacial acetic acid and the volume to 100 ml. Sterilize by autoclaving at 121°C for 20 min. Store at room temperature.

5.5 TE saturated phenol (pH 9.0)

Mix 300 ml TE (pH 8.0), 300 ml melted phenol at 50°C, and 4.5 g 8-hydroxyquinoline. Incubate at 4°C overnight. Remove supernatant. Store at 4°C.

5.6 Chloroform solution (Chloroform: Isoamylalcohol = 24:1)

Mix 960 ml chloroform with 40 ml isoamyalcohol. Store at room temperature.

6. Reagents for northern blot analysis

6.1 10 X MOPS buffer (200 mM MOPS, 50 mM sodium acetate, 10 mM EDTA) (1 L)

Dissolve 41.7 g MOPS free acid, 6.8 g sodium acetate, and 3.7 g EDTA in 700 ml sterile DEPC treated water. Adjust pH to 7.0 with NaOH, and a volume to 1 L with DEPC treated water. Store at 4°C.

6.2 1 X MOPS buffer (Running buffer) (1 L)

Mix 100 ml of 10 X MOPS buffer, 20 ml 37% (=12.3 M) formaldehyde and 880 ml DEPC treated water. Store at 4°C.

6.3 1.5% Formaldehyde agarose gel (100 ml)

Dissolve 1.5 g agarose in 100 ml 1 X MOPS buffer and heat in microwave.

6.4 RNA premix (500 µl)

Mix 62.5 µl 10X MOPS, 109 µl 37% formaldehyde and 312.5 µl formalmide. Mix this solution with RNA solution in ratio 1:1 (v/v).

6.5 5 X RNA loading buffer (10 ml)

Mix 16 µl saturated bromophenol blue, 80 µl of 0.5 M EDTA (pH 8.0), 720 µl 37% formaldehyde, 3084 µl formalmide, 4 ml 10 X MOPS buffer, and 2 ml of 100% glycerol. Add DEPC treated water to final volume 10 ml (Stability 3 months at 4°C)

6.6 20 X SSC (3 M NaCl, 0.3 M Sodium citrate) (1 L)

Dissolve 175.3 g NaCl, 88.2 g sodium citrate in 800 ml distilled water. Adjust pH to 7.0 with NaOH and the volume to 1 L. Sterilize by autoclaving at 121°C for 20 min. Store at room temperature.

6.7 Hybridization buffer (1 mM EDTA, 0.5 M phosphate buffer, 7% SDS, 1% BSA) (250 ml)

Mix 0.5 ml of 0.5 M EDTA (pH 8.0), 85.5 ml of 1 M Na_2HPO_4, 39.5 ml of 1 M NaH_2PO_4, and 17.5 g SDS and adjust a volume to 250 ml with distilled water. Sterilize by autoclaving at 121°C for 20 min. Sit to cool down to 60°C and then add 2.5 g BSA. Store at room temperature.

7. Solutions for protein

7.1 SDS-gel loading buffer (5 X stock) (2.5 M Tris, 10% SDS, 0.5% Bromophenol blue, 50% glycerol)

Dissolve 0.30 g Tris Base, 1 g SDS, 0.05 g Bromophenol blue, 5 ml glycerol and adjust pH to 6.8 with HCl and the volume to 8 ml with distilled water. Before use add 20 µl of 2-mercapthoethanol to 80 µl of solution mixture. Store at room temperature.

7.2 1.5 M Tris pH 8.8 (100 ml)

Dissolve 18.17 g Tris Base in 80 ml distilled water. Adjust pH to 8.0 with HCl and the volume to 100 ml with distilled water. Store at 4°C.

7.3 0.5 M Tris pH 6.8 (100 ml)

Dissolve 6.06 g Tris Base in 80 ml distilled water. Adjust pH to 6.8 with HCl and the volume to 100 ml with distilled water. Store at 4°C.

7.4 30% Acrylamide solution (100 ml)

Dissolve 29 g acrylamide and 1 g *N, N′*-methylene-bis-acrylamide in distilled water to a volume 100 ml. Mix the solution by stirring for 1 h to be homogenious and filter through Whatman membrane No. 1. Store in the dark bottle at 4°C.

7.5 Tris-Glycine electrode buffer (5 X stock) (1 L)

Dissolve 30.29 g Tris Base, 144 g glycine, 5 g SDS in distilled water. Adjust pH to 8.3 with HCl and the volume to 1 L with distilled water.

7.6 Staining solution with Coomassie brillant blue for protein

Mix 1 g Coomassie brilliant blue R-250, 400 ml methanol, 500 ml distilled water and 100 ml glacial acetic acid and filter through a Whatman No. 1.

7.7 Destaining solution for Coomassie Stain

Mix 400 ml methanol, 100 ml glacial acetic acid, and add distilled water to a final volume of 1000 ml.

7.8 10% (w/v) Ammonium persulfate (1 ml)

Dissolve 100 mg ammonium persulfate in 1 ml distilled water. Store at -20°C.

7.9 12% Separating gel SDS-PAGE (10 ml)

Mix the solution as follow:

1.5 M Tris (pH 8.8)	2.5 ml
Distilled water	3.3 ml
10% (w/v) SDS	0.1 ml
30% acrylamide solution	4.0 ml
10% (w/v) ammonium persulfate	0.1 ml
TEMED	0.004 ml

7.10 5% Stacking gel SDS-PAGE (5 ml)

Mix the solution as follow:

0.5 M Tris (pH 6.8)	1.26 ml
Distilled water	2.77 ml
10% (w/v) SDS	0.05 ml
30% acrylamide solution	0.83 ml
10% (w/v) ammonium persulfate	0.05 ml
TEMED	0.005 ml

7.11 1 X Acetic acid-β-alanine running buffer pH 4.5 (1 L)

Dissolve 31.2 g β-alanine and 8 ml glacial acetic acid in distilled water and adjust to 1 L final volume. Store at 4°C.

7.12 4 X Acetic-KOH resolving gel buffer pH 4.3 (100 ml)

Mix 48 ml 1 M KOH and 17.2 ml glacial acetic acid and add distilled water to 100 final volume.

7.13 4 X Acetic-KOH stacking gel buffer pH 6.8 (100 ml)

Mix 48 ml 1 M KOH and 2.9 ml glacial acetic acid and add distilled water to 100 final volume.

7.14 10% Separating gel acidic native PAGE (15 ml)

Mix the solution as follow:

4 X acetic-KOH resolving gel buffer pH 4.3	5 ml
Distilled water	7.99 ml
30% acrylamide solution	6.65 ml
300 mM sodium sulfile	0.2 ml
10% (w/v) ammonium persulfate	0.15 ml
TEMED	0.01 ml

7.15 5% Stacking gel acidic native PAGE (5 ml)

Mix the solution as follow:

4 X acetic-KOH stacking gel buffer pH 6.8	1.26 ml
Distilled water	2.86 ml
30% acrylamide solution	0.83 ml
10% (w/v) ammonium persulfate	0.05 ml
TEMED	0.005 ml

8. Solution for western blotting

8.1 1 X PBS (1.59 mM KH_2PO_4, 8.4 mM Na_2HPO_4, 2.68 mM KCl, 137 mM NaCl)

Dissolve 0.216 g KH_2PO_4, 1.192 g Na_2HPO_4, 0.199 g KCl, 8 g NaCl in distilled water and make to 1 L final volume.

8.2 1 X PBST (1.59 mM KH_2PO_4, 8.4 mM Na_2HPO_4, 2.68 mM KCl, 137 mM NaCl, 0.05% Tween 20)

Dissolve 0.216 g KH_2PO_4, 1.192 g Na_2HPO_4, 0.199 g KCl, 8 g NaCl, 0.5 ml Tween 20 in distilled water and make to 1 L final volume.

8.3 Fast green staining solution

Dissolve 0.05 g Fast green with 20 ml methanol, 5 ml glacial acetic acid, 75 ml distilled water.

8.4 3 mg/ml 4-Chloronapthol (20 ml)

Dissolve 60 mg 4-chloronapthol in 100% methanol and adjust to 20 ml final volume. Store in the dark bottle at 4°C.

9. Buffers and Reagents for enzyme assay

9.1 10 mM p-Nitrophenol (10 ml)

Dissolve 0.0139 p-nitrophenol in 50 mM NaOAc pH 5.0 buffer and make to 10 ml final volume.

9.2 50 mM NaOAc pH 5.0 buffer (100 ml)

Dissolve 0.41 g NaOAc in 80 ml distilled water, and adjust pH to 5.0 with glacila acetic acid and the volume to 100 ml with distilled water.

9.3 20 mM Tris-HCl pH 8.0 buffer (100 ml)

Dissolve 0.24 g Tris base in 80 ml distilled water, and adjust pH to 8.0 with HCl and the volume to 100 ml with distilled water.

9.4 0.4 M Na_2CO_3 (100 ml)

Dissolve 0.42 g Na_2CO_3 in distilled water, and adjust the volume to 100 ml with distilled water.

Curriculum Vitae

NAME Miss Rodjana Opassiri

DATE OF BIRTH 8 December 1975

PLACE OF BIRTH Khon Kaen, Thailand

INSTITUTION ATTENDED

Khon Kaen University, Khon Kaen, Thailand, 1994-1998, Bachelor of Science (Biology)

Suranaree University of Technology, Nakhon Ratchasima, Thailand, 1998-2003,
Doctor of Philosophy (Environmental Biology)

AWARDS AND SCHOLARSHIPS

Certificate for Outstanding in Biology, the Science Society of Thailand, 1995.

First Class Honors, Khon Kaen University, 1998.

The Development and Promotion of Science and Technology Talents Project of Thailand (DPST) Scholarship, 1994-1999.

The Royal Golden Jubilee Ph.D. Scholarship, from the Thailand Research Fund, 1999-2003.

BGJ Research Grant from the Thailand Research Fund, 2000-2003.