STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF β-GLUCOSIDASE FROM *DALBERGIA NIGRESCENS* KURZ.

Miss Phimonphan Chuankhayan

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Biochemistry

Suranaree University of Technology

Academic Year 2004

ISBN 974-533-392-1

การศึกษาโครงสร้างและคุณสมบัติของเบต้า-กลูโคซิเดส ที่สกัดจากเมล็ดฉนวน

นางสาวพิมลพรรณ ชวนขยัน

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาชีวเคมี
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2547

ISBN 974-533-392-1

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF β-GLUCOSIDASE FROM *DALBERGIA NIGRESCENS* KURZ. SEEDS.

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

Thesis Examining Committee

(Asst. Prof. Dr. Malee Tangsathitkulchai)

Chairperson

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

Member (Thesis Advisor)

(Prof. Dr. MR. Jisnuson Svasti)

Member

(Asst. Prof. Dr. Wipa Suginta)

Member

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

Member

(Assoc. Prof. Dr. Sarawut Sujitjorn)
Vice Rector for Academic Affairs

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

พิมลพรรณ ชวนขยัน : การศึกษาโครงสร้างและคุณสมบัติของบีตา-กลูโคซิเดสที่สกัดจากเมล็ดฉนวน (STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF β-GLUCOSIDASE FROM *DALBERGIA NIGRESCENS* KURZ. SEEDS.) อาจารย์ที่ปรึกษา : ผู้ช่วยศาตราจารย์ ดร.เจมส์ เกตุทัต-คาร์นส์, 209 หน้า.
ISBN 974-533-392-1

ได้ทำให้เอนไซม์บีตา-กลูโคซิเดสจากเมล็ดฉนวน (*Dalbergia nigrescens* Kurz.) บริสุทธิ์และศึกษาสมบัติของเอนไซม์ ได้เพิ่มปริมาณสารพันธุกรรมของฉนวน คือ *Dnbglu1* และ *Dnbglu2* และหาลำดับของนิวคลีโอไทด์ พบว่าลำดับของกรดอะมิโนที่ได้จาก Dnbglu1 และ Dnbglu2 เหมือนกับเอนไซม์บีตา-กลูโคซิเดสจาก *D. cochinchinensis* มากกว่า 80 เปอร์เซ็นต์ ได้ทำให้ซับสเตรตธรรมชาติของเอนไซม์กลูโคซิเดสจากเมล็ดฉนวน 2 ชนิด คือ S1, dalpatein 7-*O*-[β-D-apiofuranosyl-(1→6)-β-D-glucopyranoside] และ S2, dalnigrein 7-*O*-[β-D-apiofuranosyl-(1→6)-β-D-glucopyranoside] บริสุทธิ์ และได้หาโครงสร้าง พบว่าเอนไซม์นี้สามารถตัดน้ำตาลออกจาก S1และ S2 ซับสเตรต ได้น้ำตาลไดแซคคาไรด์ ได้ทำให้สารพันธุกรรม *Dnbglu2* แสดงออกใน *Pichia pastoris* พบว่าค่า K_m ของเอนไซม์จากธรรมชาติและลูกผสมต่อ *p*NP-β-D-glucoside และ *p*NP-β-D-fucoside มีค่าใกล้เคียงกัน และค่า K_m ต่อ S1 และ S2 เท่ากัน คือ 0.5 mM และ 0.7 mM ตามลำดับ ในขณะที่บีตา-กลูโคซิเดสจากต้นพยูงสามารถย่อยสลายซับสเตรตทั้งสองได้เพียงเล็กน้อย

สาขาวิชาชีวเคมี
ปีการศึกษา 2547

ลายมือชื่อนักศึกษา พิมลพรรณ ชวนขยัน
ลายมือชื่ออาจารย์ที่ปรึกษา
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม

PHIMONPHAN CHUANKHAYAN : STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF β-GLUCOSIDASE FROM *DALBERGIA NIGRESCENS* KURZ. SEEDS. THESIS ADVISOR : ASST. PROF. JAMES R. KETUDAT-CAIRNS, Ph.D. 209 PP. ISBN 974-533-392-1

β-GLUCOSIDASE AND *DALBERGIA NIGRESCENS* KURZ.

A β-glucosidase from seeds of *Dalbergia nigrescens* Kurz. was purified and characterized. *Dnbglu1* and *Dnbglu2* cDNAs which encode *D. nigrescens* glycosidases, were cloned and sequenced. The derived amino acid sequences of Dnbglu1 and Dnbglu2 were over 80% identical to *D. cochinchinensis* β-glucosidase. The natural substrates of the glycosidase were isolated from seeds of *D. nigrescens* Kurz. and their structures determined as compound S1, dalpatein 7-*O*-[β-D-apiofuranosyl-(1→6)-β-D-glucopyranoside] and compound S2, dalnigrein 7-*O*-[β-D-apiofuranosyl-(1→6)-β-D-glucopyranoside]. The enzyme was found to cleave the sugar from these substrates as a disaccharide. The *Dnbglu2* cDNA was expressed in *Pichia pastoris*. The native enzyme and the recombinant Dnbglu2 have similar K_m values for *p*NP-β-D-glucoside, and *p*NP-β-D-fucoside and the same K_m values of 0.5 mM for S1 and 0.7 mM for S2, respectively, while *D. cochinchinensis* β-glucosidase showed little activity towards these substrates.

School of Biochemistry

Academic Year 2004

Student's Signature _____

Advisor's Signature _____

Co-advisor's Signature _____

Co-advisor's Signature _____

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Asst. Prof. Dr. James R. Ketudat-Cairns for kind supervision, valuable advice, supporting and giving me encouragement throughout this project and providing me with opportunity to get a scholarship. I am deeply grateful to co-advisor, Prof. Dr. MR.Jisnuson Svasti for his valuable suggestion and providing facilities. I am also very grateful to Prof. Dr. Patrick A. Sullivan for his kind advice and greatly providing facilities at Institute of Molecular BioSciences, NZ. In addition, I am greatly indebted to Yanling Hua, for her kind, valuable advice and help for NMR structural analysis of glycoside substrates.

I am very grateful to Dr.Chantragan Phiphobmongkol and members of Biochemistry laboratory, Chulabhorn Research Institute (CRI) for helping and advice on research techniques. I would like to thank co-advisors, Asst. Prof. Dr.Mariena Ketudat-Cairns and Assoc. Prof. Dr.Poonsook Sriyotha for their guidance and supporting. I would like to thank Asst. Prof. Dr.Santi Sakdarat for providing facilities in his Chemistry Lab. Special thanks are extended to all friends in the Biochemistry Lab and the Thailand Research Fund (TRF) for providing the Royal Golden Jubilee Ph.D. Scholarship for my study at Suranaree University of Technology.

Finally, I would like to express my deepest gratitude to my families for their love, understanding and encouragement.

Phimonphan Chuankhayan

CONTENTS

CONTENTS (Continued)

Page

CONTENTS (Continued)

Page

CONTENTS (Continued)

Page

CONTENTS (Continued)

Page

CONTENTS (Continued)

Page

CONTENTS (Continued)

Page

CONTENTS (Continued)

Page

LIST OF TABLES

LIST OF TABLES (Continued)

Table **Page**

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

Figure **Page**

LIST OF FIGURES (Continued)

Figure **Page**

LIST OF ABBREVIATIONS

A	Absorbance
Amp	Ampicillin
ATP	Adenosine triphosphate
bp	Base pairs
BSA	Bovine Serum Albumin
°C	Degrees celsius
cDNA	Complementary deoxynucleic acid
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
dNTPs	dATP, dCTP, dGTP and dTTP
EDTA	Ethylene diamine tetraacetic ccid
g	Gravitational acceleration
(m, n) g	(milli, nano) Gram
hr	Hour
HPLC	High performance liquid chromaography
IPTG	Isopropyl-β-D-thiogalactopyranoside
kDa	Kilo Dalton
min	Minute
(m, μ, n) M	(milli, micro, nano) Molar
(m, μ) L	(milli, micro) Liter

LIST OF ABBREVIATIONS (Continued)

(μ, n, pmol) mol	(micro, nano, pico) Mole
mRNA	Messenger ribonucleic acid
Mr	Molecular weight
4MUGlc	4-Methylumbellriferyl-β-D-glucoside
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pI	Isoelectric point
p-NP	*p*-Nitrophenol
RNA	Ribonucleic acid
RNase	Ribonuclease
SDS	Sodium dodecyl sulfate
sec	Second
TEMED	Tetramethylenediamine
UV	Ultraviolet
U	Unit, μmol/min
v/v	Volume/vol

CHAPTER I

INTRODUCTION

1.1 Overview of β-glucosidases

β-glucosidases (EC 3.2.1.21) catalyze the hydrolysis β-O-glucosidic linkages formed between D-glucose and an aglycone or another sugar, releasing glucose and an aglycone. β-glucosidases are found widely in living organisms such as microorganisms, plants and animals. Because β-glucosides and β-glucosidases are ubiquitous in the living world, one might expect to find structural and catalytic properties shared by all β-glucosidases. Almost all β-glucosidases have subunit molecular weights of 55-65 kDa, acidic pH optima (5-6) and need β-glucosides as substrates (Esen, 1993). β-glucosidases have a variety of functions in different organisms, including cellulose and cellobiose catabolism, biomass conversion, chemical defense and regulating biological activity.

β-glucosidase enzymes have been identified in humans. The lysosomal glucocerebrosidase hydrolyzes glucocerebrosides (glycosphingolipids derived from endogenous membrane glycolipids) present in the lysosomal membranes, and deficiency of this enzyme results in Gaucher disease, one of the hereditary lysosomal storage disorders (Neufeld, 1991). Lactase-phlorizin hydrolase (LPH), which is found in the brush-border of the small intestine, hydrolyzes

lactose in milk and glucosylceramide, so a lack of this enzyme causes lactose intolerance. Another β-glucosidase is the broad-specificity cytosolic enzyme found in abundance in the liver, kidney, spleen, and small intestine of mammals, which has been found to efficiently hydrolyze a variety of plant glycosides (Day *et al*., 1998; Gopalan *et al*., 1992). Berrin *et al*., (2002) demonstrated that the recombinant cytosolic β-glucosidase had broad specificity to various aryl-glycosides (β-D-fucosides, α-L-arabinosides, β-D-glucosides, β-D-galactoside, β-L-xylosides, β-D-arabinosides), as does the native enzyme. It also has significant activity towards many dietary xenobiotics including glycosides of phytoestrogens, flavonoids, simple phenolics and cyanogens (Berrin *et a.l*, 2002; Day *et al*., 1998; Gopalan *et al*., 1992).

β-glucosidases from various sources also have potential applications for sequence determination of oligosacharides (Wemply, 1989), as well as for oligosacharide synthesis by reversal of their hydrolytic action (Srisomsap *et al*, 1996). In general, β-glucosidases display a broad range of specificities for the aglycone moiety of their substrates, but narrow specificity for the glycone moiety. In fact, β-glucosidases from every source have similar specificity for the glycone (glucose) portion of the substrate, but some enzymes, especially those from plants, have different specificity for the aglycone portion (Sanz-Aparicio *et al*., 1998 and Cicek and Esen, 1998). β-glucosidases may be divided into three groups according to their substrate specificity: (1) aryl β-glucosidases with a strong affinity towards aryl β-glucosides; (2) oligosaccharidases, which only hydrolyze oligosaccharides (such as cellobiose); and (3) β-glucosidases, which are active with both types of substrate and which are the most commonly observed group in cellulolytic microbes (Rojas *et al*.,

1995). Differences in substrate specificity reflect the different biological roles played by these enzymes and their evolutionary origins (Sanz-Aparicio *et al*., 1998).

1.2 Plant β-glucosidases

In plants, β-glucosidases have been implicated in a variety of processes, such as defense, growth related responses and development. One of the important functions of plant β-glucosidases is defense against pathogens and herbivores, catalyzing the cleavage of defensive glucosides, thereby releasing toxic aglycones, such as cyanide, hydroxamic acid, coumarins, thiocyanates, terpenes and saponins (Sue *et al*., 2000 and Cicek and Esen *et al*., 1999). The cyanogenesis mechanism has been well studied in model systems, such as linamarase and linamarin in white clover, in which it was shown that the β-glucosidase and the substrates (cyanogenic glucosides) are present in different cellular compartments. When the cells are injured by herbivores or pathogens, the enzyme and substrate come into contact, which leads to the hydrolysis of cyanogenic glucosides and the release of HCN, which is the toxic product (Esen, 1993). Selma (1993) proposed that in all plants with apoplastic β-glucosidases, stored monoglucosides, such as linamarin and lotaustralin, which are mobilized and converted to the corresponding diglucosides (e.g. linustatin and neolinustatin) for transport to other plant organs, where they are converted back to cyanogenic glucosides and other substances. This would suggest that β-glucosides are metabolized to other compounds and have other functions in the plant besides cyanogenesis.

In addition to cyanogenesis, β-glucosidases have been implicated in the activation of plant phytohormones, such as cytokinin, gibberellin and auxin, by releasing active forms from inactive hormone-glucose conjugates (Leah *et al*., 1995). Inactive phytohormone conjugates are abundant in plant tissues, and the apparent physiological activity of any particular conjugate correlates with its rate of hydrolysis in plant tissues. The *rolB* and *rolC* genes of the bacterial pathogen *Agrobacterium rhizogenes* in dicotyledonous plants may promote abnormal development in plants by encoding cytosolic β-glucosidases that can release the free form of cytokinin from their inactive conjugates (Estruch *et al*., 1991a). In the young seedling, the appearance of free forms of auxin, cytokinin, and gibberllin is correlated with marked decrease in the abundance of their conjugates in the endosperm. In maize seedlings, auxin and cytokinin have been hypothesized to be transported as conjugates from the endosperm to the embryo, where they are activated by hydrolysis to the free phytohormones by a β-glucosidase, Zm-p60 (Brzobohaty *et al*., 1993).

1.2.1 Cellulose metabolism

Cellulose is the most abundant biological polymer, and is found as the major structural component of the cell wall of terrestrial plants and marine algae but is also produced by other organisms such as some bacteria and marine animals. Cellulose is a polymer of glucose units, which are linked in a linear manner with β-1,4-glycosidic bonds of over 10,000 glucose residues. These chains are bound together with hydrogen bonding and van der Waals forces. Although an individual hydrogen bond is relatively weak, many such bonds acting together can impart great stability to certain conformations of large molecules (Teeri, 1997). Most animals

cannot digest cellulose as a food, and in the diets of humans this part of our vegetable intake functions as roughage and is eliminated largely unchanged. Some animals (such as cows and termites) harbor intestinal microorganisms that breakdown cellulose into glucose nutrient using of β-glycosidase enzymes. Cellulose is biodegradable by the combined activities of many different enzymes. The microbial degradation of cellulosic biomass requires the action of at least three groups of enzymes; cellulases (EC 3.2.1.4), cellobiohydrolases (EC 3.2.1.91) and β-glucosidases (EC 3.2.1.21) (Clarke *et al*., 1993). The different modes of action of cellulolytic enzymes on the cellulose are commonly described as endo- and exo-type of attack (Holm and Sander, 1994). A typical endoglucanase (cellulase) cleaves bonds along the length of the cellulose chains, resulting in a decrease in the degree of polymerization of cellulose, the exoglucanases (cellobiohydrolase) will initiate their action from the ends of the cellulose, producing cellobiose and cello-oligosaccharide with a degree polymerization of up to six glucose units, which are soluble in water. Then, β-glucosidases bind soluble cello-oligosaccharides in a productive mode and hydrolyze glucosyl residues sequentially from the non reducing end, such as the β-glucosidase from *Agrobacterium faecalis*, which has high specific activity and affinity for cellobiose.

1.2.2 Role of β-glucosidase in lignification

β-glucosidase isoenzymes from a crude cell wall preparation from *Cicer arietinum L*. cell suspension cultures showed high activity for the hydrolysis of coniferyl alcohol β-D-glucoside, which has been found to be one of the best substrates and might be involved in lignification of these plant cell cultures (Hösel *et al*., 1978).

In plants, lignin is a major component of cell walls of woody plants that provides the strength and rigidity that is characteristic of wood and of vessels in plants that transport water from roots to leaves, and also plays an important role in defense of plants against attack by pests. It is a complex organic polymer composed of three cinnamyl alcohols, *p*-coumaryl, coniferyl and synapyl. The relative proportions of the different monomers in lignin can vary greatly, depending on plant species or tissue. In gymnosperms, lignin has significant proportions of coniferyl alcohol monomer unit. Before being polymerized to lignin, coniferin (the 4-*O*-glucoside of coniferyl alcohol) must be hydrolyzed to free the phenolic hydroxyl by a cinnamyl alcohol glucoside/β-glucosidase system, which has been implicated in lignification (Dharmawardhana *et al*., 1995). Coniferin hydrolysis to release coniferyl alcohol is one of the final steps in the reaction sequence of lignin biosynthesis, which makes coniferin β-glucosidase a particularly suitable target for biotechnological modulation of lignin in trees.

1.3 Classification of glycosyl hydrolases

There are many systems to classify the glycoside hydrolases, including those based on substrate specificity, mode of attack and stereo-chemical mechanism. Since 1991, a classification of glycoside hydrolases based upon amino acid sequence similarities was introduced by Henrissat (1991).

1.3.1 Classification on the basis of substrate specificity

The classification of the glycoside hydrolases by substrate specificity is the simplest. *O*-glycoside hydrolases are given the code EC 3.2.1.X, where X

represents the substrate specificity. For instance, the differentiation between β-glucosidase (EC 3.2.1.21) and β-galactosidase (EC 3.2.1.23) is on the basis of their substrate preference for either a β-glucosides or β-galactosides. But the disadvantage of this classification is that it is not appropriate with enzymes which act on several substrates, especially, on enzymes which work on highly complex polysaccharides and display broad specificities. For example, cellulases are active on xylan, xyloglucan, β-glucan and various artificial substrates. Also, the classification based on substrate specificity fails to reflect the 3D structural features. For instance, myrosinase, which is an enzyme hydrolyzing a particular series of S-glucosides, is classified as EC 3.2.3.1 (thioglucosidase) and yet has a sequence, molecular mechanism and 3 D structure strikingly similar to O-β-glucosidases (EC 3.2.1.21) (Burmeister *et al*., 1997). Conversely, many structurally unrelated enzymes display similar substrate specificity and have identical International Union of classification Biochemistry and Molecular Biology (IUBMB) EC numbers.

1.3.2 Classification on the basis of mode of action

"*Exo*" and "*endo*" relates to glycoside hydrolase's action on a polysaccharide. Davies and Henrissat (1995) described how the shape of the active site cleft of glycoside hydrolases is reflected in their mode of action, that is, the catalytic residues of these enzymes are normally found in one of three locations, pocket, cleft or tunnel. The *exo*-enzymes tends to have their active site located within a pocket. The depth and shape of this pocket reflects the number of subsites that contribute to binding and to the length of the leaving group. *Endo*-enzymes have their catalytic site in an open cleft that allows a random binding and action within the

polymer chain (Davies and Henrissat, 1995; Davies *et al*., 1997). "*Exo*" and "*endo*" distinction is a powerful system, but it is difficult to catagorize the enzyme which displays properties between *exo* and *endo*, such as cellobiohydrolases, which bind to a polysaccharide substrates and are unlikely to release the polymer chain substrate before they perform a number of catalytic events. Further confusion often arises when the differentiation between *exo* and *endo* is made upon the hydrolysis of an inappropriate substrate, as is often the case.

1.3.3 Classification on the basis amino acid sequence similarity

The sequence and structure are related, so useful structural and mechanistic information can be derived based upon amino acid sequence alone. In addition, the molecular mechanism within each family is generally conserved and the members of a sequence-related family will have similar folds. Many of the sequence-based families are polyspecific, which means they contain enzymes of different substrate specificities, suggesting an evolutionary divergence. Sometime, the enzymes with similar specificities are found in different families, which raises the possibility of convergent evolution. The scope for a sequence and structure-based classification is not limited to glycoside hydrolases, but has been applied to glycosyltransferases, which show a similar family structure due to the corresponding diversity of the oligo- and polysaccharides that they synthesize (Campbell *et al*., 1997).

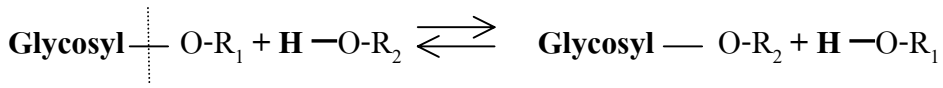
On the basis of similarities at the level of tertiary structure, families of glycosyl hydrolases are classified in groups designated superfamilies or clans (A clan is a group of families that are thought to have a common ancestry and are recognized by significant similarities in tertiary structure, together with conservation of the

catalytic residues and catalytic mechanism, Henrissat and Bairoch, 1996). Family 1 is included in clan GH-A (also named superfamily 4/7) characterized by and 8 fold α/β barrel structures, in which the two amino acids of the active site directly involved in catalysis, the acid/base catalyst and the nucleophile, are located close to the carboxy-terminus of β-strands four and seven, respectively (Jenkens *et al*., 1995 and Sanz-Aparicio *et al*., 1998). Studies of glycosyl hydrolase family 1 tertiary structures reveal similarities at the active center, including many polar and aromatic residues located toward the end of the carbohydrate-recognition site pockets (Barrett *et al*., 1995; Burmeister *et al*., 1997 and Sanz-Aparicioi *et al*., 1998). Also, most catalytic sites in family 1 are in the form of a pocket or tunnel, in contrast with the catalytic domains of barley glucanase and bacterial chitinase, which are in the form of grooves or channels (Perrakis *et al*., 1994, and Varghese *et al*., 1994).

From the primary and tertiary structures, active-site residues can be determined based on the identification of catalytic amino acid residues. If a catalytic residue in a sequence of the same family is known, it can be predicted in related sequences by sequence alignment (Doolittle, 1987). Proteins, which can be related by amino acid sequence, almost always have similar three-dimensional structures, allowing amino acid residues near the active site to be identified.

1.4 Glycosyl hydrolase Mechanism

Most glycosyl enzymes can catalyze the hydrolysis of glycosidic bonds and transfer reactions such as tranglycosylation between oligosaccharides. They do so with either retention or inversion of anomeric configuration in hydrolysis and transglycosylation. Retaining mechanisms involve a covalent glycosyl-enzyme intermediate formed and cleaved with acid/base catalytic assistance via oxocarbenium ion like transition states. The overall reactions can be simply described as breaking the bond between the anomeric carbon of the glycosyl residue and the glycosidic oxygen (C_1-O) and forming a new bond with a new hydroxyl oxygen (or other nucleophile) as follows:



where R_1 is the reducing side glycosyl residue or aglycone; **H** –O-R_2 is the acceptor and, in the case of hydrolysis, **H** –O-R_2 is water; and the dotted line is the cleavage point. Both hydrolysis and transglycosylation are exchange reactions between glycosyl residues and protons of water or another acceptor (Chiba, 1995). Glycosidases and transglycosylases are responsible for transferring of glycosyl residues from a donor sugar to an acceptor. In hydrolysis, the acceptor for glycosidases is water, but for the transferases, the acceptor is typically an alcohol on a sugar, though it can be an alcohol on a lipid, an aryl moiety, or other components of glycoconjugates (Ly and Withers, 1999).

Studying the mechanism of these enzyme and of the carbohydrate/protein interactions present in their complexes to identify active site residues has allowed

designing of new inhibitors, which are not only academically, but also commercially, interesting. For instance, the potential applications of the rapeuticuses include the control of blood glucose levels through inhibition of digestive glycosidase, control of viral diseases such as AIDS via interference with the correct processing of viral glycoproteins and other viral recognition events, and the control of cancer metastasis through inhibition of mannosidase II (Ellen *et al*., 1996).

1.4.1 Retaining mechanism

The reaction mechanism of retaining enzymes proposed since 1953 by Koshland is still generally accepted. A key feature of this type of mechanism is that it occurs via a two step, double displacement mechanism, through the formation of a covalent glycosyl-enzyme intermediate involving two key active site carboxylic acid residues which are approximately 5.5 Å apart. In the first step, one of the carboxyl groups acts as a general acid catalyst to protonate the glycosidic oxygen, while at the same time, the other carboxylic acid residue acts as a nucleophile to attack at the anomeric carbon, thereby forming a covalent glycosyl-enzyme intermediate and cleaving the C-O bond at the anomeric carbon to displace the aglycone group (leaving group). In the second step, the acid-base carboxylate deprotonates the incoming water molecule or alcohol, which attacks the anomeric carbon from the β-position and cleaves the glycosidic ester intermediate, leading to the net retention of the anomeric configuration of the substrate (Figure 1.1).

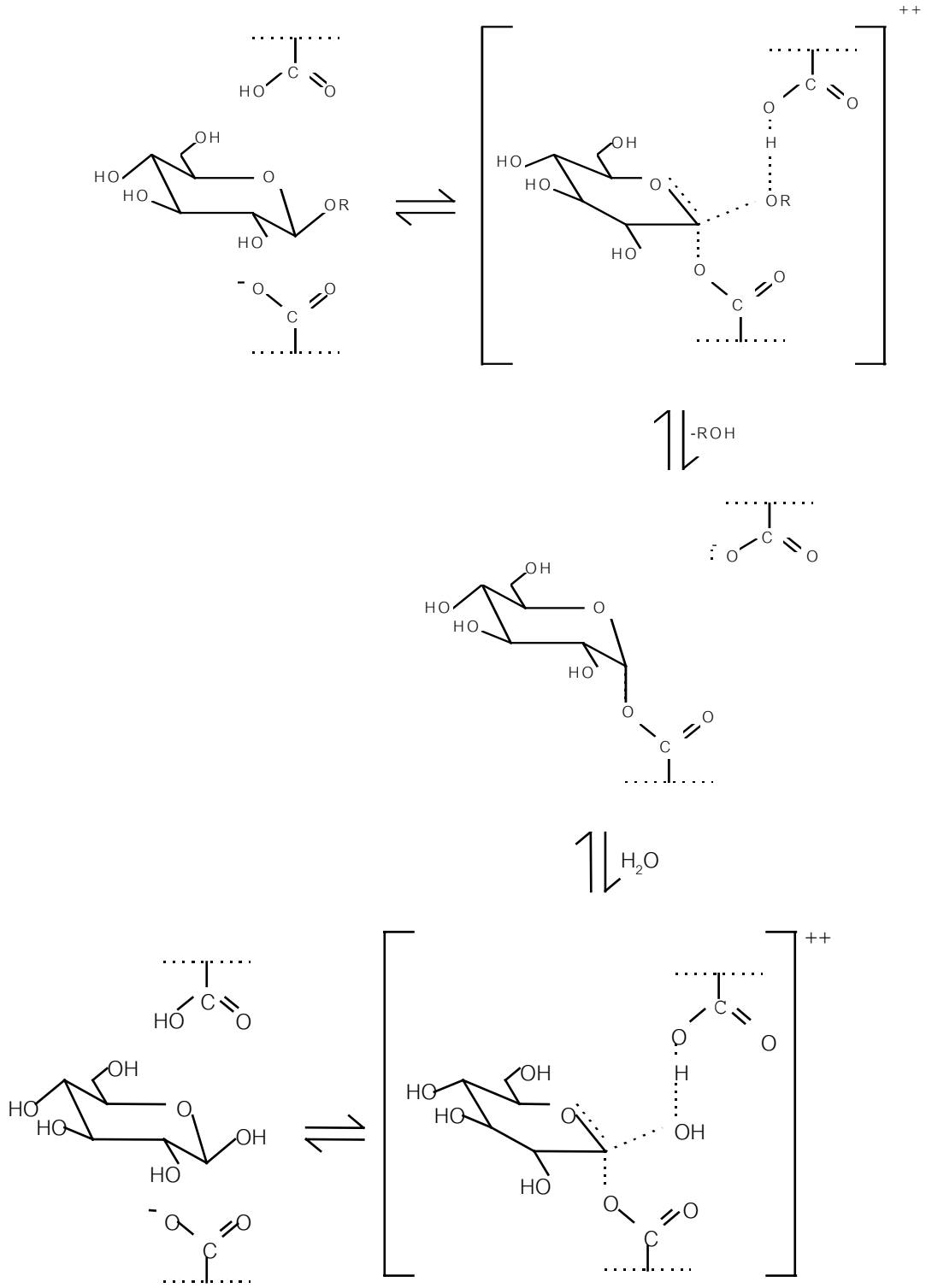


Figure 1.1 The retaining mechanism of glycosidase enzymes, such as β-glucosidases.

1.4.2 The inverting mechanism

The inverting mechanism occurs via a single displacement of the leaving group by water through an oxocarbenium ion-like transition state, as in the retaining mechanism. The two carboxyl groups which act as general acid and general base catalysts at the active site in this mechanism are 11 Å apart from each other. One provides basic catalytic assistance to the attack of the water or alcohol while the other provides acidic catalytic assistance in cleavage of the glycosidic bond, which is different from the retaining mechanism. Both mechanisms show significant similarities in that they have a pair of carboxylic acids at the active site and both mechanisms proceed via transition states with substantial oxocarbenium ion character (Zechel and Withers, 2000 and Withers, 2001).

1.4.3 Mechanistic studies on β-glucosidases

Most plant β-glucosidases have been classified in glycosyl hydrolase family 1, based on their amino acid sequence similarity (Henrissat, 1991). A lot of mechanistic information on the enzymatic action of retaining glycosidases has been produced through mutagenesis, enzyme kinetics, inhibition studies, and x-ray crystallography studies. Studies of covalent glycosyl-enzyme intermediates have used 2-deoxy-2-fluoroglucoside inhibitors with reactive leaving groups, which trap intermediates involved in the normal catalytic mechanism of the β-glucosidase, thereby effectively inactivating the enzyme. The reason for such trapping is that the fluorine at C-2 destabilizes the transition states for glycosyl enzyme formation and hydrolysis by withdrawing electrons, thus slowing both steps, while use of a reactive leaving group activates the glycosylation step enough that the enzyme accumulates as

its glycosyl enzyme adduct. This intermediate is sufficiently stable to be investigated. However, a simple glucoside such as *p*-nitrophenyl β-D-glucoside, isopropylthio β-D-glucoside, or, preferably, the C-linked glucoside β-D-glucopyranosylbenzene in the mixture resulted in reactivation of the enzyme. Upon incubation of the isolated 2-deoxy-2-fluoroglucosyl enzyme intermediate with a high concentration of β-D-glucopyranosylbenzene (200 mM), 80-90% of the enzyme activity had returned to the enzyme within a few hours, but the rate of reactivation was shown to depend on the concentration, which was due to an efficient trans-glycosylation reaction of the 2-deoxy-2-fluoro-D-glucose incorporated into disaccharide derivative of glucosyl benzene (Withers *et al*., 1987 and Withers *et al*., 1990).

Withers *et al*., 1990, found that Glu 358 from *Agrobacterium feacalis* β-glucosidase (Abg) is the attacking nucleophile by inactivating the enzyme with [1-^{3}H]-2′-4′-dinitrophenyl-2-deoxy-2-fluoro-β-D-glucopyranoside, cleaving the inactivated enzyme into peptides using pepsin, separating of the peptides by reverse phase HPLC and Edman sequencing the labeled peptide. The labeled amino acid was glutamate in the peptide sequence Y-I-T-E-N-G, which includes the (Ile/Val)-Thr-Glu-Asn-Gly ((I/V)TENG)) motif highly conserved in all family 1 glycosyl hydrolases. Nine different mutations of E358, showed it had extreme effects in decreasing the activity of *Agrobacterium* β-glucosidase (Abg) at least 10,000-fold (Trimbur *et al*., 1992). These effects on the enzyme activity of mutations at E358 supported its role as the nucleophile in glycoside hydrolysis. Since glutamate must form a covalent intermediate during catalysis, the spacing between its carboxyl and the catalytic acid-base and the ability to act as a nucleophile would be expected to be critical to catalysis. But, the mutation E358D, which has the same chemical group but

shortens the side chain, retained measurable activity. It would be expected that the shorter aspartic acid side chain length would strain any covalent bond formed during catalysis. In addition, N359 and G360 are conserved in all members of the family, suggesting that they may play a role in positioning the nucleophile. The mutation of the asparagine residue (N359S) retained essentially wild type activity, which, although it is a very conservative mutation, suggests that this residue is not absolutely critical to activity. But, replacement of G360 by serine or cysteine resulted in a large reduction in activity. This suggested that glycine may well be required to allow a turn. So, the mutations, which insert a larger side chain may not be tolerated. However, all mutations at this residue did retain some activity, which contrasts to mutations at E358. It was suggested that the importance of this residue could be to serve a structural role in maintaining the required conformation, or it might have a role in binding the substrate, or more critically in binding and stabilizing the transition state. In contrast, no critical role in catalysis was indicated by the mutations at R377 and Y380, which had only small effects on enzyme activity. One possible reason for the small effects is they may be involved in binding the second sugar moiety of cellobiose (Trimbur *et al*., 1992).

Several β-glucosidases have also been found to have a carboxylate residue as the acid catalyst involved in stabilizing the positive charge of the intermediate (Keresztessy *et al*., 1994 and Rojas *et al*., 1995). To identifying the amino acid residues which play the role of the acid/base catalysts, the kinetics of the mutants which replace the conserved glutamic and aspartic residues with sterically conservative and nonionizable amino acids such as asparagine and glutamine, or alanine have been observed. The activity of such mutants should be reduced,

especially the hydrolysis rate of substrates with poor leaving groups which require protonic assistance. With retaining glycosidases, poor leaving group substrates should be hydrolyzed very slowly, at least 10^5–fold slower than wild type, whereas hydrolysis of very good substrates should be less affected. Also, comparison of the pH dependence between mutant and wild type enzyme can demonstrate the acid/base mutant.

Normally, glycosidases have a bell shaped pH dependence curve because of the ionization of the two active-site carboxylic acids, but if the amino acid involved in acid catalyst is removed, the basic limb of the profile should be removed. Keresztessy (1994) showed the pH dependence of the V_{max} and the enzyme efficiency constant (V_{max}/K_m) for hydrolysis of both linamarin and *p*NP-glucoside (*p*NP-Glc) for cassava-β-D-glucosidase (linamarase). This dependence showed that there were two ionizable amino acid residues involved in the hydrolysis of both substrates. The theoretical pH-dependence curves for V_{max} and V_{max}/K_m could be characterized by limiting slopes of +1 and -1 using a double logarithmic scale, which demonstrated that the decrease in the parameter values was a result of the change in protonation state of single amino acid side chains both at low and at high pH. pH profiles for the V_{max} of linamarin hydrolysis showed best fit when using $pK^{ES}_1 = 4.1$ and $pK^{ES}_2 = 9.3$, while in the case of *p*NP-glucoside from the same curve only the value pK^{ES}_1 of 3.7 could be determined, since V_{max} did not show dependence on pH in the basic range. From comparison of pK^{E}_1 (4.5 and 4.6 for linamarin and *p*NP-glucoside, respectively) and pK^{E}_2 (7.1 and 7.27 for linamarin and *p*NP-glucoside, respectively) of free enzyme, it showed that the anionic and the protonated groups seem to play an essential role in the hydrolysis of linamarin, while the protonated acid catalyst does not have critical

function in the cleavage of *p*NP-glucoside. The shift in the acidity of the anionic group observed on binding *p*NP-glucoside was larger than in the case of linamarin, which might be because the glycosidic O atom in *p*NP-Glc has a relatively large δ (+) charge because of the strong electron-withdrawing para substituent of the phenyl aglycone, and the stronger polar environment might cause a higher dissociation ratio of the carboxylate group involved in catalysis. In the case of linamarin, the supposed proton-donor group, which had a high pK^{E}_{2} value, became an even weaker acid in the ES complex, which might be caused by a very strong shielding of this group from the aqueous environment by the bound substrate (Keresztessy *et al*., 1994 and Clarke *et al*., 1993).

Besides for kinetic analysis and pH profile observation, another method used to identify the mutant is "rescue" of the activity of the mutant by adding suitably nucleophilic anions, such as, azide, formate, or acetate (Ly and Withers, 1999). The principle is that the rate-limiting deglycosylation step for cleavage of good substrates by acid-base catalyst mutants can be accelerated substantially by the nucleophilic anions in the reaction mixture. These nucleophilic anions bind in the vacant anion-binding site created by the mutation and react more rapidly with the glycosyl enzyme intermediate than water in the absence of general-base-catalytic assistance. The result is increasing the k_{cat} value and the formation of a glycosyl azide or ester product. For instance, glycosyl formate and acetate are formed when formate and acetate are used.

1.5 Substrate specificity

The substrate specificity is important for the function of plant β-glucosidases, where the natural substrates may not be single glucosides but include glucosides of a vast array of aglycone groups, most of which are linked to glucose through a β-glycosidic bond. Hösel and Conn (1981) emphasized that the role of the aglycone or non carbohydrate moiety in determining the specificities of glycosidases for their natural substrates was largely unappreciated and suggested that more attention be directed to this matter.

β-glucosidases from every source have similar specificity for the glycone (Glc) portion of the substrate, but some enzymes, especially from plants, have different specificity for the aglycone portion (Cicek and Esen, 1998). Babcock and Esen (1994) proposed that a hydrophobic aglycone group is required for cleaving the β-glycosidic linkaged between the glycone and the aglycone in maize (*Zea mays* L.) β-glucosidase.

1.5.1 Aglycone specificity

The substrate specificities of several plant β-glucosidases have been examined. Grover and Cushley (1977) found that almond β-glucosidase could hydrolyze both *p*-nitrophenyl-β-D-glucopyranoside (*p*NP-glucoside) and *p*-nitrophenyl-β-D-galactopyranoside (*p*NP-Galactoside), although the K_m for the glucoside was ten fold lower. Yeoh and Woo (1992) reported the substrate specificity of cassava β-glucosidase (linamarase). They found that *p*NP-glucoside (K_m = 0.39 mM) and *p*NP-fucoside (K_m = 0.57 mM) were the best substrates, while the natural substrate linamarin has K_m = 2.08 mM. In addition to almond and cassava, Selma *et al*.

(1987) found another cyanogenic β-glucosidase from rubber tree (*Hevea brasiliensis*) that accepts a wide variety of substrates. It was thought to be the only β-glucosidase found in rubber tree and therefore was thought to not only be involved in cyanogenesis, but also lignification and other processes involving the hydrolysis of β-glucosides.

Czjzek *et al*. (2000) identified the aglycone recognition and binding site in maize β-glucosidase by crystallization of an inactive mutant (Glu1E191D) in complexes with the natural substrate 2-O-β-D-glucopyranosyl-4-hydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOAGlc), the free aglycone (DIMBOA) and the competitive inhibitor para-hydroxy-S-mandelonitrile β-glucoside (dhurrin). Maize β-glucosidase (DIMBOAGlc-hydrolase) has two known isozymes, Glu1 and Glu2, which share 90% sequence identity. Similarly, sorghum also has two β-glucosidase (dhurrinase) isozymes, Dhr1 and Dhr2, which share approximately 70% sequence identity with each other and with each of the two maize isozymes (Hösel *et al*., 1987). The primary structures of maize and sorghum β-glucosidases contain the highly conserved peptide motifs TFNEP and ITENG, which is a common characteristic for all family 1 β-glucosidases. Glutamic acids E191 and E406 in ZMGlu1, which are the acid-base catalyst and nucleophile residues, respectively, are located in the motifs TFNEP and I/VTENG, respectively. Because maize β-glucosidase isozymes, Glu1 and Glu2, and sorghum β-glucosidases, Dhr1 and Dhr2, have shown high substrate specificity towards their natural substrates, DIMBOAGlc and dhurrin, respectively, and also have different specificity to artificial substrates, they were proposed to be a good enzymatic model to determine the basis of substrate specificity. For instance,

Glu2 could not hydrolyze 6-bromo-2-naphthyl-β-D-glucoside, unlike Glu1, and it could hydrolyze other artificial substrates such as nitrophenyl glucosides about five to six times less well than Glu1. In addition to the maize β-glucosidase isozymes, sorghum Dhr1 β-glucosidase could hydrolyze only its natural substrate, whereas Dhr2 hydrolyzes certain artificial substrates in addition to the natural substrate dhurrin.

To investigate the substrate binding site and specificity in maize Glu1 β-glucosidase, the inactive Glu1 mutant (Glu1E191D) was complexed with the natural substrate (DIMBOAGlc), the free aglycone DIMBOA, and competitive inhibitor para-hydroxy-S-mandelonitrile β-glucoside (dhurrin) and crystallized. X-ray crystallography demonstrated that the aglycone part of the substrate is bound between W378 on one side and F198, F205 and F466 on another side. Thus, these four amino acid residues determined aglycone recognition and substrate specificity in Glu1. From the alignment of family 1 β-glucosidase, three of the amino acid residues (F198, F205, and F466) which are involved in DIMBOAGlc aglycone binding, are not conserved, while the fourth amino acid (W378) is highly conserved. It is the specific conformations of these four key amino acids (W378, F198, F205, and F466) and the shape of the aglycone-binding pocket they form that determines aglycone recognition and substrate specificity in maize Glu1. In addition to these four residues, only A467 interacts directly with DIMBOA. Comparison with other β-glucosidases supports the conclusion that they are involved in aglycone recognition and binding in all family 1 enzymes, but their specific conformations may vary depending on the nature of neighboring amino acids with which they have contact. Apparently, W378 interacts with the aglycone moieties, such as *p*- or o-nitrophenyl, 6-bromonaphthyl, indoxyl, and methylumbelliferyl, of all substrates hydrolyzed by Glu1, as well as those of

unhydrolyzed competitive inhibitors (e.g., dhurrin). However, stacking may not be as perfect with other aglycones as with its natural substrate. The A467 residue may also be critical to DIMBOAGlc specificity because only the maize isozymes and noncyanogenic β-glucosidase of white clover have this residue in this position among 61 plants β-glucosidases compared by Czjzek *et al*. Glu2 has a Y at position F466 of Glu1 and Dhr1 and Dhr2 have V/L, L, and S at the positions of F198, F205 and F466 of Glu1, respectively. Since these enzyme show different aglycone specificities, these differences are consistant with the importance of W378, F198, F205 and F466 in determining aglycone substrate specificity and may explain the basis of substrate specificity differences among β-glucosidases.

The crystallographic data of the inactive maize β-glucosidase-dhurrin complex, showed that the interaction of the para-hydroxy-S-mandelonitrile moiety of dhurrin with the aglycone binding pocket forces the glucose ring to take a position different from the productive position seen in DIMBOAGlc. The major consequence is that the C-1 and O-1 atoms of dhurrin's sugar, which are critical for hydrolysis, are positioned too far from the catalytic residues, which is most likely why dhurrin acts as a competitive inhibitor and is not hydrolyzed.

Keresztessy *et al*. (2001) identified essential amino acid residues in the active site in cyanogenic β-glucosidase (linamarase) from cassava (*Manihot esculenta* Crantz) on the basis of the white clover linamarase structure. They proved that the E413G mutant had no hydrolytic activity against either linamarin or *p*NP-glucoside, indicating an essential function for E413 as the catalytic nucleophile, which is located in the I/VTENG motif. Also, when Q339, a residue which is near E413 in the cassava β-glucosidase, was mutated to glutamate (E) to test if it has any influence on the pH-

activity profile, the activities toward of linamarin and *p*NP-glucoside were 300 times lower than wild type. An explanation was derived from molecular modeling, which predicts that E339 forms a salt bridge with R106, so the mutation would disrupt the essential salt bridge between R106 and E413 and an H-bond between Q339 and E198. In addition to Q339, A201 in the cassava enzyme was mutated to valine (V) to probe the extent of such a water sheltering effect. They found that the Michalelis constant (K_m) for the hydrolysis of linamarin was five times higher than the wild type, whereas the K_m for the hydrolysis of *p*NP-glucoside remained unchanged. The effect of the Ala→Val mutation suggests that the steric hindrance introduced by V201 may selectively inhibit docking of the tetrahedral acetone cyanohydrin moiety of linamarin, but not the planar *p*-nitrophenyl aglycone of *p*NP-glucoside.

F269 in cassava linamarase is equivalent to V254 of the clover linamarase, which is located in a competent position for a contact with the substrate aglycone, and is replaced by asparagine (N) in the disaccharide-cleaving bacterial β-glucosidase. The result of a mutation between the two was that F269N had no detectable cellobiase activity; however, and it was still active against linamarin and *p*NP-glucoside. However, the K_m of linamarin hydrolysis was increased 16 fold by the mutation while the K_m for *p*NP-glucoside hydrolysis was increased only 2.5 fold. In contrast, the catalytic constant (k_{cat}) measured with the linamarin substrate was decreased only 1.9 times, whereas the k_{cat} for *p*NP-glucoside is 1.5 times higher in the F269N mutation. So, replacing F269 with a polar residue confers a significant effect on the K_m of linamarin cleavage without a similar change in the k_{cat}, which can only be attributed to selective alterations in the rate of binding/dissociation of this substrate to and/or from the enzyme, since not much effect was seen on the K_m for *p*NP-glucoside hydrolysis.

This result established that F269 in the wild type enzyme has a more important function in binding the natural substrate, linamarin, than *p*NP-glucoside.

1.5.2 Glycone specificity

Sanz-Aparicio *et al*. (1998) have identified the interaction of the substrate with specific protein residues by co-crystallized BglA from *Bacillus polymyxa* with D-glucono-1,5-lactone. The location and interactions of the inhibitor in the active site confirmed the roles proposed for the catalytic residues in family 1 enzyme. The acidic residues E166 and E352 (catalytic acid/base and nucleophile residues, respectively) are positioned at a similar distance from the anomeric sugar C1 atom. E352 was pointed to the C1 atom and also made hydrogen bonding with O2 of the glucanate ligand, while H121, which is also conserved in family 1, made a hydrogen bond to O2. In addition, the nitrogen atom of W406, another conserved residue among β-glucosidases, is hydrogen bonded to O3. Q20 and W398, which are strictly conserved in family 1, make bidentate hydrogen bonds with the O3 and O4 hydroxy groups and make stacking interactions with the plane defined by the ring-like gluconate chain, respectively. Vyas (1991) showed that forming bidentate hydrogen bonds can occur both equatorially or equatorially and axially, which was seen in the interactions to glucose and galactose (O4 equatorial and axial, respectively). Also, E405 can bond bidentate to O4 and O6, as well, it is able to interact with O4 either in either axial or equatorial orientations, but the phosphorylated O6 cannot be accommodated. Glutamic acid is predicted to be located at the substrate phosphoryl site (Wiesmann *et al*., 1995), so these interactions might indicate the enzyme specification against phosphorylated substrates and ability to display both

glucosidases and galactosidase activities. In addition, Y296 is hydrogen bonded to O5 and also to E352, for which it was suspected to stabilize the oxocarbonium.

1.6 Plant glycosides

Glycosides are derivatives formed when monosaccharides react with alcohols in the presence of an acid catalyst. In naming of glycosides, the "ose" suffix of the sugar name is replaced by "oside", and the alcohol group name is placed first. Generally, for most aldols, glycoside formation involves the loss of an equivalent of water. Glycosides are abundant in biological systems, and may be formed by attaching a sugar moiety to a lipid or benzenoid structure, in which cases, the solubility and other properties of the compound may be changed substantially. So, many glycosidases have evolved for the attachment and removal of sugars from alcohols, phenols and amines. Examples of glycosides include salicin, one of the oldest herbal remedies, which was the model for the synthetic analgesic aspirin, and anthocyanins, a large class of hydroxylated, aromatic oxonium cations, which provide the red, purple and blue colors of many flowers, fruits and some vegetables. Also, amino derivatives of ribose, such as cytidine which play important roles in biological phosphorylating agents, coenzymes and information transport and storage.

1.6.1 Classification

The classification of glycosides depends on the property and structure of the aglycone rather than the sugar type. Plant glycosides are classified into 13 groups (สรศักดิ์ เหลี่ยวไชยพันธ์, 2531) that are shown below.

1. Steroid cardioactive glycosides
2. Saponin glycosides
3. Flavonoid glycosides
4. Anthracene glycosides
5. Cyanogenic glycosides
6. Isothiocyanate glycosides
7. Coumarin glycosides
8. Alcohol glycosides
9. Aldehyde glycosides
10. Iridoid glycosides
11. Stilbene glycosides
12. Lignan glycosides
13. Xanthone glycosides

6.1.2 Glycoside properties and roles in plants

Plant glycosides are secondary metabolites. They are widely dispersed in the flowering plants, and can be found in many parts of plants such as tuber, root, bark, leaves, seed and fruit. The glycosides usually occur within plant vacuoles, whereas their specific glycosidases may be cytosolic soluble enzymes in the same or different cells. Glycosides in plants have two main properties differing from its aglycone (Hosel, 1981), which are enhanced water solubility that helps transport of the compound within the plant, and decreased biological activity (inactivation of the compound). The physiological role of a plant glycosides may be predicted from the aglycone structure, its biological activity, location and timing of the glycoside's

hydrolysis. Some plant glycosides act as inactive toxins in plant defense mechanisms, such as cyanogenic glucoside and saponin. Another role is regulation of phytohormones, where the glycosides of phytohormones serve as the inactive forms for regulation of active hormone concentrations during plant development and environmental response (Campos *et al*., 1993).

Saponins consist of triterpenoid or steroid glycoalkaloid molecules bearing one or more sugar chains, which are characterized by their surfactant properties, and give stable foam when shaken with water (Osbourn, 1996). They are widely found in plants, including some foods such as beans, spinach, tomatoes, potatoes, onions, garlic and asparagus. Potentially medicinal and health-promoting saponins are contained in foodstuffs, such as sugar beet and kidney beans, and folk medicines. The distribution of steroidal saponin is limited to several families in the plant kingdom, while triterpene saponin appears in a wide variety of plants (Yang *et al*., 1999). The primary mode of action of saponins towards fungi involves the formation of complexes with membrane sterols, which results in pore formation and loss of membrane integrity. The precise way in which saponins become incorporated into membranes is unclear, since many models have been described. The nature of the aglycone and the oligosaccharide moiety of the saponin molecule may both be likely to contribute to the membraneolytic properties of individual saponins, or the sugar part may be a solubilizing group aiding delivery of the saponin to the membrane, where membrane-bound glycosyl hydrolases are then presumed to activate the saponin by converting it to aglycone.

Flavonoids are polyphenolic compounds that occur ubiquitously in foods of plant origin. Variations in the heterocyclic ring C give rise to flavonols, flavones,

cathechins, flavonones, anthocyanidins and isoflavonoids (Figure 1.2). In addition, the basic structure of flavonoids allows a multitude of substitution patterns in the benzene rings A and B within each class of flavonoids: phenolic hydroxyls, *O*-sugars, methoxy group, sulfates and glucuronides (Middleton and Kandaswami, 1994). Flavonoids are common substances in the daily diet as shown in Table 1.1.

Table 1.1 Occurrence of flavonoids in common foods

Flavonoid subclass	**Major food sources**
Flavonols	onions, kale, broccoli apple, cherries, berries tea, red wine
Flavones	parsley, thyme
Flavonones	citrsu
Catechins	apples, tea
Anthocyanidins	cherries, grapes
Isoflavones	soy beans, legumes

From many studies has shown that flavonoids can inhibit and induce a large variety of mammalian enzyme systems, which are involved in important pathways that regulate cell division and proliferation, platelet aggregation, detoxification, and inflammatory and immune response (Hollman and Katan, 1999). Kandaswami and Middleton, (1994) hypothesized the antioxidant properties of flavonoids that may protect tissue against oxygen free radicals and lipid peroxidation, might be involved in several pathological conditions such as atherosclerosis, cancer, and chronic inflammation. It was suggested that the chemical nature of flavonoids, which are

antioxidants, might allow them to be involved in the prevention of atherosclerosis, cancer, and chronic inflammation (Halliwell, 1994).

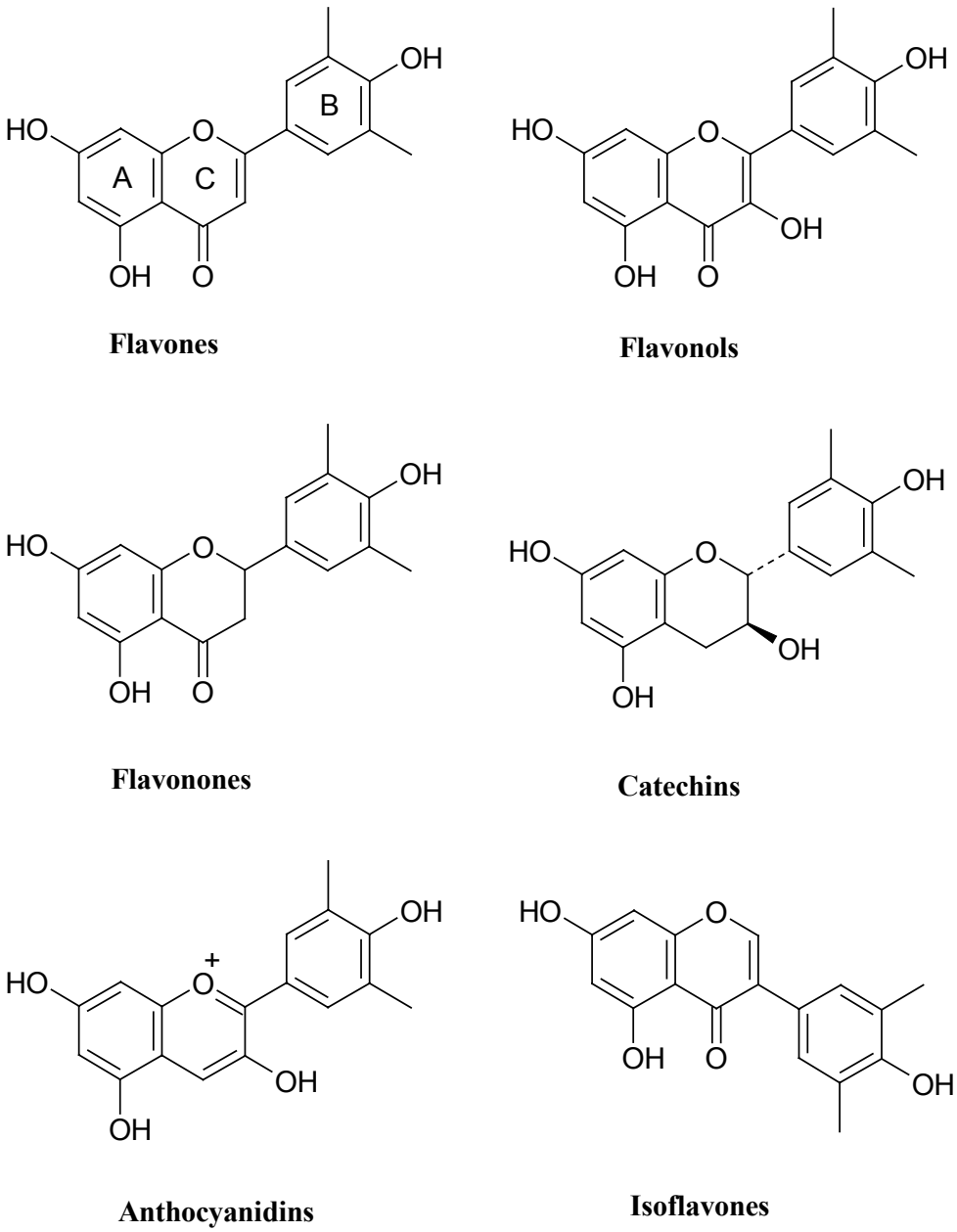


Figure 2.1 Subclasses of flavonoids. Classification is based on variations in the heterocyclic C-ring.

1.7 Thai rosewood β-glucosidase/β-Fucosidase

In order to find glycosyl hydrolases for potential use in oligosaccharide and glycoside synthesis, Surarit *et al*. (1995) screened seeds of 50 plants species found in Thailand. *Dalbergia cochinchinensis* Pierre (Thai Rosewood) had the highest hydrolytic activity towards both *p*NP-fucoside and *p*NP-glucoside. The *D. cochinchinensis* enzyme has been purified and characterized to show both activities reside in the same protein (Srisomsap, 1996; Surarit, 1997). To determine its structure and evolutionary lineage, the sequence of the enzyme was determined by peptide sequencing, followed by PCR cloning (Ketudat Cairns *et al*., 2000). The cDNA includes a reading frame coding for 547 amino acids, including a 23 amino acid propeptide and a 524 amino acid mature protein. From the sequence alignment, it was shown to be a member of glycosyl hydrolase family 1 and had the highest percent identity with β-glucosidase from sweet cherry (56.7%), cyanogenic β-glucosidase (linamarase) from white clover (56.3%), prunasin hydrolase precursor (54.9%), amygdalin hydrolase isoform AH I precursor (54.3%) from black cherry, and non-cyanogenic β-glucosidase precursor from white clover (53.9%). From the multiple alignment, the catalytic acid/base and nucleophile residues of the active site of the enzyme are conserved, E182 and E396 of the mature protein, respectively. Glutamate 396 was deduced to be the catalytic nucleophile due to its homology to E358 in the β-glucosidase from *Agrobacterium*, which has been shown to act as the catalytic nucleophile (Withers and Street, 1988). Similarly, E182 was homologous to E198 of the cassava cyanogenic β-glucosidase, which was shown to be the catalytic proton donor in that enzyme (Keresztessy *et al*., 1994). The phylogenetic tree indicated that

the Thai rosewood β-glucosidase has the closest relationship with the group of the enzymes from dicotyledon plants which includes the cyanogenic β-glucosidases, such as sweet cherry prunasin hydrolase precursor and the linamarase and β-glucosidase enzymes of cassava and white clover. The next most closely related group of enzymes is the plant thioglucosidases, such as thioglucosidase of arabidopsis, rape and white mustard and a diverse group of β-glucosidases, including those from monocots (rape, oat, barley, sorghum and maize) is also closely related (Champattanachai, 1998 and Ketudat Cairns *et al*., 2000).

Kinetic studies of the enzyme using *p*NP-β-D-glucoside and *p*NP-β-D-fucoside demonstrated that it has higher K_m (5.39 mM) and k_{cat} (307 s^{-1}) for β-glucoside than for β-fucosidase (0.54 mM and 151 s^{-1}). The purified enzyme also has activity toward *p*NP-β-D-galactoside and *p*NP-β-D-arabinoside (Surarit *et al*., 1995 and Srisomsap *et al*., 1996).

The natural substrate of Thai rosewood was isolated and identified by NMR and mass spectroscopy and shown to be a novel isoflavonoid glucoside, 12-dihydroamorphigenin-8′-O-β-D-glucoside, the aglycone of which was named "dalcochinin" (Svasti, J *et al*., 1999). The rosewood β-glucosidase could hydrolyze this substrate with the rate of 157% compared with *p*NP-β-D-glucoside, which suggested that the isolated natural substrate is a true natural substrate for Thai Rosewood β-glucosidase.

1.8 *Dalbergia nigrescens* Kurz

Dalbergia nigrescens Kurz is one of twenty-six species of the genus *Dalbergia*, which are found in Thailand (THAI FOREST BULLETIN (BOTANY), 2002). This species has 3 varieties, which are anomala (Pierre), nigrescens and aigonensis (Pierre). *Dalbergia nigrescens* Kurz is the variety nigrescens, which has trees 10-30 meters high and seeds with dark brown to black color when dry. This varity can be found in northern and north-eastern Thailand, such as Chiang Mai, Lampang, Phayao, Phrae, Tak, Loei and Nakhon Ratchasima. Thai names include kraphi, cha-nuan and sa-nuan.

When Surarit *et al*. (1995) screened the glycohydrolase enzymes in many Thai plants, they found that *D. nigrescens* Kurz was the plant source with the second highest β-glucosidase and β-fucosidase activities but it had approximately 10 times lower levels of these than *D. cochinchinensis*.

1.9 Research objectives

1) To determine the nucleotides and amino acids sequence of *D. nigrescens* β-glucosidase and compare it with that of *D. cochinchinensis* dalcochinase;

2) To characterize *D. nigrescens* β-glucosidase activity and substrate specificity;

3) To determine the structure of natural substrates of *D. nigrescens* β-glucosidase; and

4) To study and compare the catalytic region between *D. nigrescens* and *D. cochinchinensis* β-glucosidase.

CHAPTER II

MATERIALS AND METHODS

2.1 Materials

2.1.1 Plant material

Dalbergia nigrescens Kurz. seeds were collected from trees within Suranaree University of Technology campus, Nakhon Ratchasima, Thailand.

2.1.2 Chemicals

Calcium chloride ($CaCl_2$), Coomassie brilliant blue R-250, imidazole, diethyl pyrocarbonate (DEPC), magnesium chloride ($MgCl_2$), phenyl methylsulfonyl fluoride (PMSF), sodium citrate, Polyvinylpolypyrrolidone (PVPP), DOWEX 2X-8, chloroform/isoamyl alcohol (24:1), isopropyl β-D-thiogalactoside (IPTG), 2-mercaptoethanol, RNase Zap, *p*-nitrophenyl-β-D-glucopyranoside, *p*-nitrophenyl- β-D-fucopyranoside, *p*-nitrophenyl-β-D-galactopyranoside, 4-methyl-umbelliferyl-β-D-glucopyranoside, 4-methyl-umbelliferyl-β-D-fucopyranoside, 4-methyl-umbelliferyl-β-D-galactopyranoside and other artificial and natural glycoside substrates, which are not listed here were purchased from Sigma (St. Louis, MO, USA) and Fluka (Steinheim, Switzerland). Washed sea sand (O_2Si) QP, EU was purchased from Panreac (Barcelona, Spain). Ammonium hydroxide (NH_4OH), bromophenol

blue, chloroform ($CHCl_3$), copper sulphate ($CuSO_4$), disodium ethylenediamine tetraacetate (EDTA), ethanol (EtOH), Folin reagent, glacial acetic acid (CH_3COOH), glycerol, glycine, hydrochloric acid, water (HPLC grade), methanol (HPLC grade), trifluoroacetic acid (TFA), potassium chloride (KCl), potassium hydroxide (KOH), sodium acetate (NaOAc), sodium bicarbonate (NaH_2CO_3), sodium chloride (NaCl), sodium hydroxide (NaOH), sodium dihydrogen phosphate (NaH_2PO_4), sodium hydrogen phosphate (Na_2HPO_4), ammonium sulphate $(NH_4)_2SO_4$, sodium acetate (NaOAc, anhydrous), and sodium carbonate (Na_2CO_3, anhydrous) were purchased from Carlo Erba (Rodano, Milano, Italy). Tris (hydroxymethyl) aminomethane was purchased from Acros Organics and Fisher Scientific (Belgium). Acrylamide, bis-*N,N'*-acrylamide, ammonium persulphate, *N,N',N'',N'''*-tetramethylethyllenediamine (TEMED), lysozyme, sodium dodecyl sulphate (SDS), Sephacryl S-300 and Triton X-100 were purchased from Amersham Pharmacia Biotech (Uppsala, Sweden). Bacto tryptone, yeast extract and bacto agar were purchased from Difco (BD, NJ, USA). Trizol Reagent was purchased from Invitrogen (Carlsbad, CA, USA). Restriction endonucleases, including SmaI, EcoRI, PstI, PmeI and SacII, were purchased from New England Biolabs, (Beverly, MA, USA). RNase H, *Taq* DNA polymerase, *Pfu* DNA polymerase, agarose (molecular grade), deoxyribonucleotides (dNTP) and X-gal were purchased from Promega (Madison, WI, USA). The ABI PRISM Dye terminator cycle sequencing kit was obtained from Applied Biosystems Inc., (Foster City, CA, U.S.A). Thin-layer chromatography and preparative thin-layer chromatography silica gel 60 F_{254} were purchased from Merck (Darmstadt, Germany). 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 initial DNA amplification, which were derived from the cDNA sequence of Thai rosewood β-glucosidase, were kindly provided by Prof. Jisnuson Svasti. Gene specific primers derived from the sequence of the *D. nigrescens* cDNA and other cloning primers shown in Table 2.1 were obtained from the BioService Unit (BSU) of the National Science and Technology Development Agency (Bangkok, Thailand) and Proligo/Genset (Singapore).

Table 2.1 The oligonucleotide primers for initial DNA amplification

primer	Nucleotide sequence
Q_T	CCA GTG AGC AGA GTG ACG AGG ACT CGA GCT CAA GCT T_{16}
Q_I	GAG GAC TCG AGC TCA AGC
Q_O	CCA GTG AGC AGA GTG ACG
$PolyT_{17}$	TTT TTT TTT TTT TTT TT
For.2	AGG TTC CTC CAT TCA ACC GAA G
For.3	CCA CCA ATA TCC AGA AAA GAT AGC G
For.4	AGC CTT GGA GGA TGA GTA CGG T
For.5	TCA AAA TGA TAC CCA GCG ATA TCT
For.6	CGC CAT CTC TTT TAT ATT CGA TAT GC
For.7	GGC AAG ATG GAG CTT ATC AAC G
Rev.3	AAC ATC TCC GTT GCT TCT ATC CGC
Rev.4	GCT GTC CCA AAA ATG AAA TCT GA
Rev.5	TTT TGT CTT TTC CTT GAT CAT CG
Rev.7	CAT CCA CAT GTG AAG TCA AGA TAT CG
Rev.8	AAT TGG AAG CAA AGA TCC GCA TA
PCF	CTT TCA TCT CAT GCA GC
Fi	GGA GAT GTT ACA ATT GAC
Ri	TAG TGA CAT AAA TAC AAT C
Fo1	TCC TTG TAT CTG TCT TAC C
Ro1	AGG CAA ATA GAA TAA ATA AGG
PBGF1	CAC CAT TTG TAA CCA TTT TTC ATT G
PBGR1	GTA TGA TTT AAG AAG CCA CCA TAC
PBGF2	GTC CAG TGA CTC CCT CAG GAT G
PBGR2	GTA AAG CAA CAA ATC TCG AAG TCC
BseRIF	CCT TCA GAT TTC ATT TTT GGG
BseRIR	GGG AGT ATT CTA GGC CAG GAG ATG G
NdeIF	CCA TCT CCT GGC CTA GAA TAG TCC C
NdeIR	TGC AAA TAT ACC ATA TGC ATA GCC
EcoRIF	CCG AAT TCG AGG TTC CTC CG
EcoRIR	CAC CAC CTG TGC AAG TTG AAT TCC G
InsertF	GGG TGA GAA GAG ATA GAA TAT G
EcoRIMF	CGG CTT CTA TTG AAT TCG C
EcoRIMR	GCG AAT TCA ATA GAA GCC G
DNBgluR2SacII	GCC GCG GCT TTT CTT CCT TGA TCA TTG TTG
DNBgluR1SacII	CCG CGG AAA GCC TTC AAT GCC CCT CTT GGG
PstIgluF102	GCT GCA GGC ATT GCC TTT GCA AAA GAA GTC
PTDnF1	GGT GGC TTC TTA GAT CGT AG
PTDnR1	GCC ATT TGT GGT GAA GAC TTG
PTDnR2	CGA AAA TCA TTT ACA ACC CTA C
Dn2NTERMPstIf	ATT CCT GCA GTT CCT CCA TTC AAT CGA AG
Dn2CTERMXbaIr	ATC AAA ATG CTT GAA TGG CCC ACT T
3′ terminus	AAA TGT ACC AAA GCC ACA AAC
5′ terminus	TCC TTT CTT TCA TCT CAT GAT TG

2.1.4 Bacteria and yeast strains

Origami(DE3) and BL21(DE3) host strains (Novagen, Corp, Madison, WI, USA) were used to express protein in *E. coli* system. The Origami (DE3) strain has mutations in both the thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes, which greatly enhances disulfide bond formation in the cytoplasm to allow disulfide-containing proteins to achieve their active and properly folded conformations. A DE3 lysogen of strain BL21 is the most widely used host for gene expression for general purposes. It contains T7 polymerase gene under control of a *tac* promoter, which is induced upon IPTG induction. This strain is deficient of *lon* and *omp*-t proteases and is suitable for expression of non-toxic genes.

The *Pichia pastoris* strain GS115 (Invitrogen) was used for protein expression in yeast. It is a methylotrophic yeast, which can utilize methanol as its sole carbon source. The first step in the metabolism of methanol is the oxidation of methanol to formaldehyde using oxygen catalyzed by the enzyme alcohol oxidase. *Pichia pastoris* has the advantages of higher eukaryotic expression systems, such as protein processing, folding and posttranslational modificaion. There are two genes in *Pichia pastoris* that code for alcohol oxidases, *AOX1* and *AOX2*. The *AOX1* gene is responsible for the vast majority of alcohol oxidase activity in the cell, and is tightly regulated and induced by methanol to very high expression levels. The *AOX1* promoter is used to drive expression of the gene encoding the desired heterologous protein (Ellis *et al*., 1985; Koutz *et al*., 1989 and Tschopp *et al*., 1987a). *AOX2* is about 97% homologous to *AOX1*, but growth on methanol with *AOX2* alone is much slower than with *AOX1*. The *Pichia pastoris* strain GS115 has a mutation in the histidinol dehydrogenase gene (*his4*), which prevents it from synthesizing histidine.

Expression plasmids carry the *HIS4* gene, which complements *his4* in the host, so transformants are selected for their ability to grow on histidine deficient medium. GS115 will grow on complex medium such as YPD and on minimal media supplemented with histidine, until transformed; it cannot grow on minimal medium alone as it is His^-.

2.1.5 Plasmid vectors

The pET system is the most powerful system yet developed for the cloning and expression of recombinant proteins in *E. coli*, and is based on the T7 promoter-driven system. In pET vectors, target genes are cloned under control of strong bacteriophage T7 transcription and translation signals, and expression is induced by providing a source of T7 RNA polymerase in the host cell (Moffatt and Studier, 1986; Rosenberg *et al*., 1987 and Studier, 1990).

pET-32a and pET-40b vector were selected for expression of recombinant proteins in *E. coli*. The pET-32a vector produces a thioredoxin fusion protein, which may increase the yield of soluble protein in the cytoplasm (LaVallie *et al*., 1993). If the target protein contains one or more essential disulfide bonds, the combination of a pET-32 vector and a *trxB* or *trxB*/gor host may prove to be optimal, because disulfide bond formation in the cytoplasm appears to be dependent on the presence of thiredoxin (Stewart *et al*., 1998).

pET- 40b was designed to create fusions to an enzyme that catalyzes the isomerization of disulfide bonds in the periplasm (DsbC) for export and periplasmic folding of target proteins.

The pPICZαB-thrombin vector was selected to express and secrete recombinant proteins in *Pichia pastoris*. Proteins are expressed as fusions to an N-terminal peptide encoding the *Saccharomyces cerevisiae* α-factor secretion signal and a C-terminal peptide containing the *myc* epitope for detection and a polyhistidine tag for purification on metal chelating resin. In addition, the thrombin site was introduced before the *myc* epitope, to eliminate the C-terminal tag from the target protein, as may be useful for crystallization of the protein (Mariena Ketudat-Cairns, personal communication).

2.2 Methods

2.2.1 Enzymatic Characterization

2.2.1.1 Crude enzyme extraction

Seeds were decontaminated with 0.1% hypochlorite for 10 min, then washed with distilled water and soaked overnight. After this, all the procedures were carried out at 4^{o}C. Seeds were homogenized with buffer 1 (0.025 M Tris-HCl, pH 8.0, containing 70% $(NH_4)_2SO_4$, 10 mM ascorbic acid, 0.025 M β-mercaptoethanol, and 1 mM PMSF) at the ratio of 1 g seed: 2 ml buffer with a blender. Then, the solids were filtered with 4 layers of cheese clothes, and extracted 2 more times. After that, the solution was extracted with buffer 2 (0.2 M NaCl, 10 mM ascorbic acid, 0.025 M β-mercaptoethanol, 1 mM PMSF in 0.025 M Tris-HCl buffer, pH 8.0), at a ratio of 1 g seed: 4 ml buffer. The homogenate was centrifuged at 12,000 rpm for 15 minutes at 4^{o}C, and the precipitate discarded. To eliminate polyphenolic

compounds, PVPP was added to 2% and stirred at 4°C for 1 hr. The homogenate was centrifuged at 15,000 rpm for 20 min at 4°C, and the pellet was discarded. Then, 25% (w/v) activated Dowex 2X8 resin was added to the supernatant and stirred for 1 hr to eliminate phenolic compounds, which interfere with both glycosidase activity and protein determination. Next, the supernatant was centrifuged at 15,000 rpm for 20 min at 4°C to remove the Dowex 2X8 resin. The protein was precipitated from the crude extract with $(NH_4)_2SO_4$ in the range of 35-75% saturation. For the first 0-35% $(NH_4)_2SO_4$ fraction, $(NH_4)_2SO_4$ was added to 35% saturation of the final volume, stirred for 3 hr, centrifuged at 15,000 rpm 20 min, and the pellet discarded. The protein of interest was precipitated from the supernatant again with 75% $(NH_4)_2SO_4$ by stirring overnight at 4°C, then centrifuged at 15,000 rpm 20 min at 4°C to collect the protein pellet. The protein pellet was resuspended with as small a volume as possible of 25 mM Tris-HCl buffer, pH 7.0. The resuspended sample was dialyzed in a 12-14K molecular weight cut off dialysis bag (Spectrum, Texas, USA.) in five changes of 500 mL of 25 mM Tris-HCl buffer, pH 7.0 to remove $(NH_4)_2SO_4$ salt.

2.2.1.2 Protein purification

A DEAE-sepharose anion exchange column (1.25 cm × 23.5 cm, 115 cm^3) was equilibrated with 3 bed volumes of 25 mM Tris-HCl buffer, pH 7.0. After loading the samples into the column, the column was washed with 2 bed volumes 25 mM Tris-HCl buffer, pH 7.0, and then eluted with a linear gradient of 0 to 0.5 M sodium chloride (NaCl) in 25 mM Tris-HCl buffer, pH 7.0, over 2 bed volumes at a flow rate of 0.65 mL/min. Five milliliter fractions were collected and assayed for the enzyme activity by hydrolysis of 1 mM *p*-nitrophenyl-β-D-glucopyranoside and *p*-

nitrophenyl-β-D-fucopyranoside, as described below, and for amount of protein by measuring absorbance at 280 nm. Fractions that had activity were pooled and concentrated by ultrafiltration (Amicon, Bedford, MA, USA; 50K molecular weight cut off) at 1,500 g for 20 min and applied to a Sephacryl S-300 gel filtration column (Amersham Phamarcia, Uppsala, Sweden).

A Sephacryl S-300 gel filtration column (1.5 cm × 25 cm, 150 cm^3) was equilibrated with 3 bed volumes of 0.3 M NaCl in 50 mM Tris-HCl buffer, pH 7.0. After loading the 2 mL sample into the column, the column was eluted with the same buffer at a flow rate of 0.8 mL/min. Two milliliter fractions were collected, assayed for hydrolysis of 1 mM *p*-nitrophenyl-β-D-glucopyranoside and *p*-nitrophenyl-β-D-fucopyranoside, as described below, and the amount of protein was determined by measuring absorbance at 280 nm. Fractions that had activity were pooled and concentrated by ultrafiltration (Amicon, 50K molecular weight cut off) at 1,500 g for 20 min.

2.2.1.3 Determination of β-glucosidase and β-fucosidase activity

The reactions were performed in 1.0 mL of 0.1 M sodium acetate buffer, pH 5.0, containing 1 mM *p*-nitrophenyl-β-D-glucoside or *p*-nitrophenyl-β-D-fucoside and the properly diluted enzyme. The reaction was incubated at 30°C for 10 min and 2.0 mL of 2 M sodium carbonate was added to stop the reaction. Under alkali conditions, the released *p*-nitrophenol was determined by measuring absorbance on a UV spectrophotometer at 405 nm. The amount of released *p*-nitrophenol which was produced from the enzymatic reaction was calculated using a *p*-nitrophenol

standard curve ranging from 0-100 μmol. One unit of enzyme is defined as the amount of enzyme releasing 1 μmol *p*-nitrophenol per minute at 30^{o}C.

2.2.1.4 Protein determination by Lowry method

Protein concentration was determined using bovine serum albumin as standard (20-100 μg BSA). The sample was mixed with water (1 mL), then the 1 mL reaction was mixed with 1 mL of alkali copper solution, which was composed of 0.05% copper sulfate, 0.1% potassium-sodium tartrate, 5% sodium carbonate and 0.4 M sodium hydroxide. The reaction was allowed to stand for 10 min at room temperature and then 0.5 mL of 50% v/v Folin reagent and water was added. The mixture was vigorously mixed and allowed to stand at room temperature for 30 min. The blue color that developed was measured spectrophotometrically at 750 nm with a UV/Vis spectrophotometer (Lowry *et al*., 1951)

2.2.1.5 Denaturing polyacrylamide gel electrophoresis

Denaturing polyacrylamide gel electrophoresis was performed according to Laemmli (1970). The polyacrylamide gel was prepared from a stock solution of polyacrylamide consisting of 30% (w/v) acrylamide and 1% (w/v) of *N,N'*-methylene-bis-acrylamide. The separating gel contained 10% acrylamide and 0.1% SDS in 0.375 M Tris-HCl buffer, pH 8.8, while the stacking gel contained 5% acrylamide and 0.1% SDS in 0.065 M Tris-HCl buffer, pH 6.8. The gel was chemically polymerized by addition of TEMED to 0.05% and ammonium persulfate to 0.1%. The samples were mixed 5:1 with 5x loading dye sample buffer containing 0.5 M Tris-HCl buffer, pH 6.8, 10% glycerol (v/v), 2% SDS, 0.5% 2-mercaptoethanol

(2-ME) and 0.025% bromophenol blue. The sample was then boiled for 5 min and spun down at 12,000 rpm for 1 min before loading. Electrophoresis was run under a constant voltage of 120V for 1 hr 30 min with 1X Tris-glycine buffer, pH 8.3, (0.3% (w/v) Tris-base, 1.4% (w/v) glycine) with 1% SDS. After electrophoresis, the gel was stained with staining solution (0.025% Coomassie Brilliant Blue R-250, 40% methanol and 7% acetic acid) for 30 min and destained with destaining solution (40% methanol and 7% acetic acid) until the protein bands appeared. The protein bands were compared with the Low-range protein standard markers (Bio-Rad, Hercules, CA, USA): phosphoryase 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 kDa) to estimate protein sizes.

2.2.1.6 Non-denaturing polyacrylamide gel electrophoresis

Native gel was prepared from a stock solution of polyacrylamide consisting of 30% (w/v) acrylamide and 1% (w/v) of *N,N'*-methylene-bis-acrylamide. The separating gel contained 7% acrylamide in 0.375 M Tris-HCl buffer, pH 8.8, while the stacking gel contained 3% acrylamide in 0.065 M Tris-HCl buffer, pH 6.8. The gel was chemically polymerized by addition of TEMED to 0.05% and ammonium persulfate to 0.5%. The samples were mixed 1:4 with 5x loading dye sample buffer containing 0.5 M Tris-HCl buffer, pH 6.8, 10% (v/v) glycerol and 0.025% bromophenol blue. Electrophoresis was run under a constant voltage of 120 V for 1 hr 30 min in 1X Tris-glycine buffer, pH 8.3. The gel was washed with 1% Triton X-100 (v/v) 3 times, 15 min/time, and then washed and equilibrated in 0.1 M sodium acetate buffer, pH 5.0, for 10 min. Then, the gel was stained separately for β-

glucosidase, β-fucosidase and β-galactosidase activities using 1 mM 4-methylumbelliferyl-β-glucopyranoside, 4-methylumbelliferyl-β-fucopyranoside and 4-methylumbelliferyl-β-galactopyranoside, respectively, for 5 min in the dark at room temperature, and fluorogenic bands of activity were detected with a Fluor-S^{TM} MultiImager (Bio-RAD) with UV transillumination.

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

The native molecular weight of protein was estimated by using Sephacryl S-300 (Amersham Pharmacia) gel filtration chromatography. The Sephacryl S-300 column (1.5 cm × 25 cm, 150 cm^3) was equilibrated with 50 mM Tris-HCl buffer, pH 7.0, 0.3 M NaCl at 4^oC. Blue dextran (2000 kDa, Sigma) was used to determine the void volume (V_0) of the column by loading 0.5 ml of 10 mg/mL blue dextran onto the column. The blue dextran was eluted in the same buffer at a flow rate of 0.8 mL/min and the absorbance was measured at A_{280} for plotting the effluent peak. The elution volume of blue dextran (V_0) was determined by measuring the volume from when it was loaded until the effluent peak. To determine the elution volume (V_e) of the standard from the molecular weight marker kit (Sigma Aldrich, St. Louis, MO, USA) and purified protein, each protein was dissolved in the equilibration buffer to the final concentration of 10 mg/ml β-amylase (200 kDa), 8 mg/mL alcohol dehydrogenase (150 kDal), 5 mg/mL carbonic anhydrase (29 kDa), 2.5 mg/mL Cytochrome C (12.4 kDa) and 5 mg/mL *D. nigrescens* β-glucosidase. Five hundred microliters of each protein was loaded to the column and eluted with the same conditions as blue dextran. The V_e for each protein was determined the same way as

blue dextran. The V_e/V_0 for each protein was plotted against its molecular weight on semilog paper and the molecular weight of *D. nigrescens* was estimated from comparing its elution volume to a straight line through this plot.

2.2.1.8 N-terminal amino acid sequencing and amino acid compositon analysis

The protein was further purified using a Vydac-C4 (4.6 mm × 250 mm (5 μm)) reverse phase column (Grace Vydac, Columbia, MD, USA) with a gradient between buffer A: 0.1% TFA/ H_2O, B: 0.1% TFA/ 60% ACN and buffer D: ACN. The gradient used for separation is shown in Table 2.2. The collected protein fraction was dried using microcentrifuge evaporator. Three hundred and fifty micrograms of dried protein was resuspended in 17.5 μL of dH_2O to 20 mg/mL final concentration and an equal volume of 2% (w/v) NH_4HCO_3 to give a final concentration of 10 mg/mL. Then, 3.5 μL of 1X trypsin (Promega) (1 μg/μL) was added to give a ratio 1:100 for enzyme: substrate and the reaction incubated at 37^{o}C for 4 hrs. Seven microliters of trypsin inhibitor (1 μg/μL) was added to stop the reaction with a ratio of 2:1 trypsin inhibitor: trypsin. The triptic digested protein was separated using Vydac-C18 (4.6 mm × 250 mm (5 μm)) reverse phase column with buffer A: 0.1% TFA/ H_2O, B: 0.1% TFA/ 60% ACN and buffer D: ACN. The gradient used for separation is shown in Table 2.2.

Table 2.2 HPLC Gradient for separation of protein and tryptic digest

Gradient (min)	%A	%B	%D
0	100	0	0
5	100	0	0
60	0	100	0
65	20	0	80
70	20	0	80
71	100	0	0
81	100	0	0

The N-terminus of the purified protein and tryptic peptides were sequenced on an ABI 473A Protein Sequencer (Applied BioSystems Inc., Foster City, CA, USA).

The peptide peaks were collected in 1.5 mL microcentrifuge tubes and dried in a microcentrifuge evaporator, then they were hydrolyzed using 0.2 mL of 6 N HCl (with 1% phenol by volume) at 110^{o}C for 24 hours. The hydrolyzed amino acids were derivatized to their PTC-derivatives by adding 20 μL of PITC: ethanol: water: triethylamine (1:7:1:1 by vol.) and incubated at room temperature for 10 min. The PTC-derivatives were dried in a speed vacuum and analyzed using C18 reverse phase column with buffer A: 0.1% TFA/ H_2O and B: 0.1% TFA/ 60% ACN. The gradient used for amino acid separation is shown in Table 2.3.

Table 2.3 HPLC Gradient for amino acid composition analysis

Gradient (min)	%A	%B
0	100	0
10.0	54	46
10.5	0	100
11.5	0	100
12.0	0	100
12.5	100	0
20.0	100	0
20.5	100	0

2.2.1.9 pH and temperature optimum

To determine the pH optimum for enzyme activity, the reactions were performed in different buffers from pH 3-8 (citrate, pH 3-4; NaOAc pH 4-5; potassium phosphate, pH 6-8), in a buffer concentration of 0.1 M. The activity at various pH values was measured by mixing the enzyme solution with 1 mM final concentration of *p*NP-β-D-glucoside or *p*NP-β-D-fucoside. The reaction was allowed to proceed for 10 min at 30°C and the amount of released *p*-nitrophenol were determined as described in method 2.2.1.3.

The temperature optimum was determined by incubating the enzyme with 1 mM *p*NP-β-D-glucoside or *p*NP-β-D-fucoside in 0.1 M NaOAc buffer, pH 5.0 over a temperature range of 35-85°C for 10 min. The amounts of released *p*-nitrophenol were determined as described in method 2.2.1.3.

2.2.1.10 Hydrolysis of synthetic substrates

The reactions were performed by incubating the enzyme with 5 mM final concentration of *p*NP-β-D-glucoside, *p*NP-β-D-fucoside, *p*NP-β-D-galactoside, *p*NP-β-D-xyloside, *p*NP-α-L-arabinoside, *p*NP-β-D-thioglucoside, *p*NP-α-L-arabinoside and *p*NP-β-D-mannoside in 0.1 M NaOAc buffer, pH 5.0. In the reaction, 20 μL of 0.1 M NaOAc buffer, pH 5.0, 25 μL of 10 mM solution of each substrate and 5 μL of enzyme dilution (1.9×10^{-12} mols) were placed in the wells of a microtiter plate. The reactions were incubated at 30°C for 10 min, then the reaction was stopped by adding 100 μL of 2 M Na_2CO_3, and the absorbance was measured at 405 nm.

2.2.1.11 Hydrolysis of natural substrates and oligosaccharides

The peroxidase/glucose oxidase-based assay (PGO assay), was used to determine glucose. 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. The reactions to test hydrolysis of natural substrates were performed by incubating the enzyme with 2.5 mM substrates including laminaribiose, laminaritriose (β1→3 linked glucose), Cellobiose, cellotriose (β1→4), sophorose (β1→2), β-gentiobiose (β1→6), mangiferin, D-amygdalin, salicin, dhurrin, 2-O-β-D-glucopyranosyl-4-hydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOGlc), Thai rosewood dalcochinin 8′-*O*-β-D-glucoside and both of *D. nigrescens* β-D-glucoside substrates in 0.1 M NaOAc, pH 5.0. In the reactions, 70 μL of 0.1 M NaOAc buffer, pH 5.0, 25 μL of 10 mM substrate and 5 μL of enzyme dilution (1.9×10^{-12} mols) were mixed in 0.5 mL microfuge tubes. The reactions were incubated at 30°C for 30 min, stopped by boiling for 5 min and spun down. 100 μL of PGO and 50 μL of ABTS (1 mg/mL in NaOAc buffer, pH 5.0) were added to the stopped reaction, mixed and incubated at 37°C for 30 min. The reactions were transferred to microtiter plates and the absorbance of the developed green color was measured at 405 nm.

2.2.1.12 Effects of various substances on the enzyme activity

The effects of some substances on activity were determined by using 1 mM final concentration of $FeSO_4$, $MnCl_2$, $CaCl_2$, $ZnSO_4$, $HgCl_2$, EDTA, δ-glucono-1,5-lactone and 2,4 dinitrophenol-2-deoxy-2-fluoro glucoside. The purified

enzyme was preincubated with the various substances for 10 min at 30°C in 0.1 M NaOAc, pH 5.0. The glucono-1,5-lactone was also tested without preincubation. Because the activity toward *p*NP-β-D-glucoside was less than the activity toward *p*NP-β-D-fucoside, 1 mM and 2 mM *p*NP-β-D-fucoside and *p*NP-β-D-glucoside, respectively, were added and the reactions were incubated for 10 min at 30°C, stopped and quantitated, as described in method 2.2.1.10.

2.2.2 Determination of kinetic parameters of S1 glycoside and S2 glycoside hydrolyzed by *D. nigrescens* using high performance liquid chromatography (HPLC)

Glucose could not be detected after hydrolysis of either of the natural substrate glycosides with *D. nigrescens* enzyme, so the released aglycone part was quantitated by HPLC. The standard curve of products (aglycone) was prepared by injecting 20 μL of various concentrations of the aglycones of the S1 and S2 glycosides (0.125-1 mM) in methanol and measuring their absorbance peaks at 320 nm by HPLC. Various concentrations of each glycoside were incubated with *D. nigrescens* enzyme (1.9×10^{-14} mols for native enzyme and 3.9×10^{-15} mols for recombinant enzyme) in 0.1 M NaOAc, pH 5.0, at 30°C, for 10 min. The reaction was stopped by boiling at 100°C for 5 min. Then, the reaction was dried by speed vacuum and the pellet was resuspended with 100 μL methanol, and the suspension was centrifuged at 12,000 rpm for 10 min. Then, 20 μL of reactions and standards were injected onto an Eclipse XDB-C18 (4.6 mm ID*250 mm (5 μm)) reverse phase column (Agilent Corp, USA) with an HP-Series 1100 HPLC (Agilent) and eluted with

a linear gradient of 0-100% methanol in 0.1%TFA/water from 0-20 min. The products were detected by absorbance at 320 nm on a multiple wavelength spectrophotometric detector. Kinetic parameters, K_m and V_{max}, of the purified enzyme at pH 5.0 at 30^oC were calculated according to the method of Lineweaver and Burk (linear regression) and Michalis Menten (nonlinear regression) using Microsoft Excel and the Enzfitter computer program (Elsevier Biosoft, Cambridge, U.K.), respectively.

2.3 Determination of *D. nigrescens* glycoside and its aglycone by TLC

2.3.1 Purification of *D. nigrescens* glycoside

Dried seeds were removed from their pods and partially ground in a blender to allow removal of the seed coats. After the seed coats were removed, the seeds were blended to a powder, which was extracted with absolute methanol in a ratio of 1:3 (w/v) overnight at room temperature. The supernatant was centrifuged at 12,000 rpm for 15 min, the precipitate was discarded and the supernatant partitioned with ethyl acetate. The methanol phase was dried to remove the solvent with a rotary evaporator at room temperature. The dried weight was measured and the residue was redissolved with methanol. The crude extract was chromatographed over a Sephadex LH-20 chromatography column (1.3 × 87 cm, 451 cm^3), with methanol as solvent. Fractions were analyzed on analytical silica gel 60 F_{254} aluminum TLC sheets with ethyl acetate/acetic acid/methanol/water (15:2:1:2) (v/v) and UV absorbant spots visualized under a UV box at 254 and 366 nm. The fractions containing *D. nigrescens* glycoside were pooled and concentrated by speed vacuum, yielding a yellow powder. The residues were weighed, then dissolved with absolute methanol to

a final concentration of approximately 100 mg/mL. Then, the pooled glucoside fractions were chromatographed on preparative silica gel 60 F_{254} TLC using the solvent system, ethyl acetate/acetic acid/methanol/water (15:2:1:2) (v/v). The *D. nigrescens* glucoside band was detected under UV light as with analytical TLC, scraped off the plate and eluted from the silica gel with 30 mL of absolute methanol by stirring for 1 hr, then the gel was spun down at 12,000 rpm for 10 min to remove the silica material. The supernatant was concentrated by rotary evaporation with a tared round-bottom flask. The dried extract was dissolved with methanol to a final concentration of approximately 1 g/mL.

2.3.2 TLC analysis of *D. nigrescens* glucoside and *D. nigrescens* aglycone

Identification of substrates was performed by digesting the *D. nigrescens* seed extracts or purified glucosides with *D. nigrescens* β-glucosidase enzyme purified from the seeds and evaluating the product by TLC. The digested product aglycone had changed mobility on the TLC plate. Digestion reactions were performed on 10% of glycoside methanol extract in 0.1 M sodium acetate buffer, pH 5.0, with 0.01 unit of *D. nigrescens* β-glucosidase, and were mixed and incubated at 37^{o}C overnight. After the incubation, the reactions were spotted on an analytical silica gel 60 F_{254} aluminium TLC sheet. Developing solvent was prepared by mixing chloroform/methanol/water (12:1:5) (v/v) in a separation funnel, allowing for the phases to separate and using the lower phase for development. The development chamber was equilibrated with the developing solvent for 20 min to saturate the atmosphere of the chamber with solvent vapor. The TLC plate was developed in the chamber for a distance of about 5.5 cm. The TLC plate was dried and developed again

in the same developing solvent. The TLC results were visualized under a UV box at 254 and 366 nm. To visualize the sugars spots, the TLC was sprayed with 10% sulfuric acid in absolute methanol and heating until the brown spots appeared.

2.3.3 Structural analysis of glycoside substrates and their aglycones

The purified glycosides and their aglycones were sent for structural determination by NMR and mass spectrometry. The S1 and S2 compounds were dissolved in DMSO-d_6 and CD_3OD, respectively, for NMR analysis. For other experiments the samples were dissolved in absolute methanol. The ^{1}H, ^{13}C, COSY, NOESY, gHMQC, gHMBC and DEPT spectra were determined on an 300 MHz Varian Unity Inova spectrometer (Bruker, Germany). Mass spectrometry was done on an Esquire LC-Mass spectrometer (Bruker Daltonics, Beverly, MA, USA) using methanol as solvent.

2.3.4 Sugar composition analysis

The component sugars of the glycosides were determined using sulfuric acid (H_2SO_4) hydrolysis. The reaction was performed by boiling 5.5 mg of substrate with 0.7 mL 0.1 N H_2SO_4 in a 1.5 mL microcentrifuge tube for 25 min. The reaction was extracted with n-butanol and then neutralized with $NaHCO_3$ until it reached pH 7 by paper pH indicator, followed by evaporation with microcentrifuge speed vacuum. The residue obtained was dissolved in water and analyzed on TLC with a 14:6:1 $CHCl_3$:MeOH:H_2O solvent system. To visualize the sugar spots, the TLC was sprayed with 10% sulfuric acid in absolute methanol and heated until the brown spots appeared.

To analyze the sugars using GC-MS, 5.5 mg of compound S2 were boiled with 0.7 mL 0.1 N H_2SO_4 in a 1.5 mL microcentrifuge tube for 25 min. Then, the reaction was cooled and extracted with n-butanol, then neutralized with $NaHCO_3$ and dried by rotary evaporator. The trimethylsilylation (TMS) of sugar was done by dissolving three milligrams of dried residue in 3 μL dH_2O and then adding a mixture of 0.1 mL of trimethylsilyimidazole and 0.9 mL of anhydrous pyridine, which had been dried previously over anhydrous Na_2SO_4. Five microliters of the mixture was injected into a gas chromatograph-mass spectrometer (Varian-CP-3800 Gas Chromatograph and MS-1200L Quadrupole Spectrometer; Varian, CA, USA.) equipped with a Varian-FactorFour Capillary Column, VF-5ms 30 M × 0.25 MM I.D. The operating conditions were as follows: injector at 250^{o}C; split ratio 20:1; and helium carrier gas at 1.2 mL/min. Mass spectra were collected with electron ionization (EI) and a source temperature of 200^{o}C, a source voltage of 70 eV, and 3.2 × 10^{-7} torr. The data was collected in full-scan acquisition mode. The column and operating conditions were the same as used for GC analysis.

The retention time was compared with the glucose, apiose, ribose, galactose and fructose standards, which were prepared by the same method. The conditions for TMS-sugar separation by gas chromatography are shown in Table 2.4.

Table 2.4 Gas chromatography conditions for separation of TMS-sugars from acid hydrolysis (monosaccharides)

Temp(^{o}C)	Rate (^{o}C/min)	Hold (min)	Total (min)
150	0.0	5.00	5.00
190	20.0	12.00	19.00

Also, fifty milligrams of compound S2 were digested with the purified enzyme in 0.1 M NaOAc pH 5.0 at 30^{o}C for 2 hrs. Then, the aqueous layer was extracted with n-butanol to remove the aglycone part. The aqueous phase was dried by speed vacuum. The sugar part was analyzed on TLC, as described above. In addition, the sugar was analyzed by GC-MS with trimethylsilylation, as described above with the conditions for TMS-sugar separation (see Table 2.5). The operating conditions were otherwise as described for trimethylsilylated monosaccharides.

Table 2.5 Gas chromatography conditions for enzyme digested glycoside (disaccharide) separation

Temp(^{o}C)	**Rate (^{o}C/min)**	**Hold (min)**	**Total (min)**
150	0.0	5.00	5.00
190	20.0	12.00	19.00
240	40.0	40.00	60.25

2.4 Cloning and sequencing the cDNA of *D. nigrescens* β-glucosidase

2.4.1 Oligonucleotide primer design

The oligonucleotide primers which were used for DNA amplification of 3' RACE, 5' RACE, and partial cDNA were derived from the *D. cochinchinensis* β-glycosidase cDNA (Ketudat Cairns *et al*., 2000). Specific primers which were used for amplification of full length cDNA were designed from the nucleotide sequence of the *D. nigrescens* 3'RACE and 5'RACE PCR products.

2.4.2 Total RNA isolation

Isolation of total RNA was done using Trizol reagent. One hundred milligrams of immature seeds were ground in liquid N_2 using a cold mortar and pestle and transferred into an RNase-free microtube containing 1 mL of Trizol reagent. The homogenized sample was incubated for 5 min at room temperature, then 0.2 mL chloroform/isoamyl alcohol (24:1) was added. The mixture was mixed vigorously by inverting for 15 sec and incubated at room temperature for 5 min, then centrifuged at 12,000 rpm at 4^{o}C for 15 min. The colorless upper aqueous phase, the RNA phase, was transferred into a new microtube. The RNA was precipitated from the solution by mixing with 0.45 mL of isopropanol, incubating at room temperature for 10 min and centrifugation at 12,000 rpm for 10 min at 4^{o}C. The RNA pellet was washed with 75% (v/v) ethanol in DEPC-treated water and dried briefly at 37^{o}C. The dried RNA pellet was dissolved in 20 μL of DEPC-treated water. The RNA solution was immediately used to synthesize the first strand cDNA or kept at -70^{o}C. RNA concentration was determined by measuring the absorbance at 260 and 280 nm with a

Lamda Bio20 UV/VIS Spectrometer (Perkin Elmer, Shelton, CT, USA). One absorbance A_{260} unit was assumed to be equivalent to 40 μg/mL of RNA (Sambrook *et al*., 1989). Reasonable purity of RNA was indicated by a ratio of absorbance of A_{260}/A_{280} between 1.80-2.00. The RNA content was calculated from the A_{260} value of single strand RNA by the following equation:

μg/mL of RNA = (A_{260}×dilution factor×40 μg/mL)

= {A_{260}×[1000 μL/mL/(x microliter of RNA solution)]×40 μg/mL}

2.4.3 First-strand cDNA synthesis

First-strand cDNA was synthesized by using the Superscript Reverse Transcriptase II. The reaction mixture was constituted in a 0.5 mL tube by combining 5 μg of RNA and 500 ng of Q_T primer and adding enough DEPC treated water to adjust the volume to 12 μL. The solution was incubated at 70°C for 10 min and then immediately placed on ice. Then, 4 μL 5x first strand buffer (250 mM Tris-HCl, pH 8.3, containing 375 mM KCl and 15 mM $MgCl_2$), 2 μL 0.1 M DTT, 0.25 μL RNasein (40 units/μL), and 1 μL 10 mM dNTP (10 mM each dATP, dCTP, dGTP and dTTP) and 0.75 μL DEPC-treated water were added and mixed. The reaction was incubated at 42°C for 2 min, then 1 μL of Superscript Reverse Transcriptase II (200 units/μL) was added, mixed by pipetting and then incubated at 42°C for 50 minutes. The reaction was terminated by heating at 70°C for 15 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. The first-strand cDNA was kept at -30^{o}C until used in PCR reactions.

2.4.4 Amplification of partial cDNA

The cDNA was used as a template to amplify the partial regions of cDNA by using various primers, including For.2, For.3, For.4, For.5, For.6, Rev.4, Rev.5, Rev.6, Rev.7, and Rev.8 (Table 2.1). The PCR products were cloned and sequenced. PCR amplification was done with 1 μL cDNA, 1.5 μL of 10 mM $MgCl_2$, 2.5 μL of 2.5 mM dNTP, 2.5 μL of 10x PCR buffer w/o $MgCl_2$ (500 mM KCl, 100 mM Tris-HCl, pH 9.0), 0.25 μL of *Taq* polymerase (5 U/μL) and distilled water up to 25 μL. The amplifications were done with 35 cycles of 94^{o}C for 30 sec, 50^{o}C for 30 sec, and 72^{o}C for 1 min with premelting and post extension at 94^{o}C and 72^{o}C for 3 min, respectively.

2.4.5 3′ Rapid amplification of cDNA ends (RACE)

The cDNA made as described in secton 2.4.3 was used as template for 3′ RACE to amplify the 3′ end of the *D. nigrescens* β-glucosidase cDNA. The first PCR amplification was set up with the cDNA as a template, and Q_O and For.5 (Table 2.1) as primers by mixing 0.5 μL of each 10 mM of Q_O and For.5 primer, 1 μL cDNA, 1.5 μL of 10 mM $MgCl_2$, 2.5 μL of 2.5 mM dNTP, 2.5 μL of 10x PCR buffer w/o $MgCl_2$ (500 mM KCl, 100 mM Tris-HCl, pH 9.0), 0.25 μL of *Taq* polymerase (5 U/μL) and distilled water up to 25 μL. The amplifications were done with 35 cycles of 94^{o}C for 1 min, 60^{o}C for 1 min, and 72^{o}C for 3 min with premelting and post

extension at 94°C and 72°C for 5 min, respectively. The product of the first amplification was used as a template with Q_i and For.6 as primers in the second PCR amplification with the rest of the reaction the same as in the first step. The amplifications were done with 35 cycles of 94°C for 1 min, 60°C for 1 min, and 72°C for 3 min.

Amplification of the 3′ end of the *Dnbglu2* cDNA was achieved by PCR using the PTDnF1 primer as the sense primer and the Q_T primer as the antisense primer, to generate a product of approximately 1,250 bp. Both 3' RACE PCR products of *Dnbglu1* and *Dnbglu2* were cloned and sequenced, as described in methods 2.4.8-2.4.20.

2.4.6 Rapid amplification of the 5′ cDNA ends (5' RACE)

The first strand cDNA prepared as described in 2.4.3 was precipitated with seeDNA by adding 2 μL of seeDNA (Amersham) to the 22 μL of cDNA solution to be precipitated from method 2.4.3. Then, 0.1 volumes of 3 M NaOAc, pH 5.2 and 2 volumes of absolute ethanol were added and mixed briefly. The reaction was incubated at room temperature for 2 minutes and centrifuged at 12,000 rpm for 5 min. The supernatant was removed and the pink cDNA pellet was rinsed with 0.5 mL of 70% (v/v) ethanol. Then, the tube was centrifuged again at 12,000 rpm for 5 min, the residual ethanol was removed by air drying. The pellet was resuspended in 10 μL of dH_2O.

Terminal Deoxynucleotidyl transferase (TdT) is an enzyme that catalyzes the repetitive addition of mononucleotides from dNTPs to the terminal 3'-OH of a cDNA, accompanied by the release of inorganic phosphate (Kato *et al.*,

1967). To add the anchoring poly A tail sequence at the 3' end of the cDNA, 9.5 μL of purified cDNA was mixed with 4 μL of 5x Tailing buffer (50 mM Tris-HCl, pH 8.4, 125 mM KCl and 7.5 mM $MgCl_2$), 1 μL 2 mM DTT, 4 μL 5 mM $CoCl_2$, 0.5 μL 1 mM dGTP, and 1 μL of TdT (25 units/μL) (Promega) and incubated at 37°C for 20 min. The reaction was stopped by incubating at 70°C for 15 min. After that, the anchored cDNA was used as template in a PCR reaction with contents as described in method 2.4.5. For the first PCR amplification, the PCR product was amplified with Q_T and Rev.3 primers (Table 2.1). The amplification was done by cycles of 94°C for 1 min, 37°C for 1 min, and then 72°C for 1 min, and 30 cycles of 94°C for 1 min, 60°C for 1 min, and 72°C for 1 min using *Taq* polymerase. Then, the second and third PCR amplifications were performed using the Q_T and Rev.4 primers and Q_O and Rev.4 primers, respectively. The amplifications were done by 35 cycles of 94°C for 1 min, 60°C for 1 min, and 72°C for 3 min.

Amplification of the 5′ end of the *Dnbglu2* cDNA was carried out by synthesis of the first strand cDNA using PTDnR1 primer (Table 2.1), as described in section 2.4.3. After synthesis, the RNA was digested with RNaseH. The first strand cDNA was then purified by precipitating with 2 volumes of absolute ethanol and 0.1 volumes of 3 M NaOAc, pH 5.2. The product was tailed with dATP using terminal deoxy-transferase (TdT) as described above. The product was amplified with the Q_T and PTDnR1 primers, as described above with an annealing temperature of 60°C. A second amplification was performed using PTDnR2 and Q_o primers (approximately 695 bp). The 5' RACE PCR products of both *Dnbglu1* and *Dnbglu2* were cloned and sequenced as described in methods 2.4.8-2.4.20.

The sequences of the 3′ and 5′ RACE products were used to design the 3′ terminus and 5′ terminus primers (Table 2.1).

2.4.7 Amplification of full length cDNA

The full length *Dnbglu1* cDNA was amplified from first strand cDNA which was prepared as described in method 2.4.3. The first strand cDNA was used as a template in the first PCR amplification with the PCF (5′ terminus) and R_{O1} (5′ terminus) primers (Table 2.1), and then the product of this reaction was used as a template in the second PCR amplification with the PCF and R_i primers. The amplifications were done with 30 cycles of 94^{o}C for 1 min, 50^{o}C for 1 min, and 72^{o}C for 3 min. The PCR product was cloned and sequenced.

The full length *Dnbglu2* cDNA was amplified using the Dn2_3′UTRr1 (3′ terminus) and Dn2_5′UTRf1 (5′ terminus) primers. The full length cDNA was amplified using the PCR hybridization method (Horton *et al*., 1989). The 5′ cDNA end PCR product was amplified with a single strand cDNA, which was reverse transcribed from the $PTDnR_2$ primer and polyadenylated at its 3′ end (the RNA 5′ end) as template. The first amplifications were done by 10 cycles of 94^{o}C for 45 sec, 37^{o}C for 45 sec, and 72^{o}C for 90 sec, then 25 cycles of 94^{o}C for 45 sec, 60^{o}C for 45 sec, and 72^{o}C for 90 sec, with *Pfu* polymerase using Q_O and $PTDnR_2$ as forward and reverse primers. After that, the second amplification was done with the Dn2_5′UTRf1 forward and $PTDnR_2$ reverse primers with the first PCR product as template with 30 cycles of 94^{o}C for 45 sec, 60^{o}C for 45 sec, and 72^{o}C for 90 sec. The 3′ cDNA end PCR product was amplified using Dn2_3′UTRr1 reverse and $PTDnF_1$ forward

primers with a cDNA which was reverse transcribed from the Q_T primer as the template. The amplification was done with *Taq* polymerase by 30 cycles of 95°C for 1 min, 55°C for 30 sec, and 72°C for 90 sec, and the PCR product was reamplified again with the same conditions using *Pfu* polymerase. Then, both 3′ end and 5′ end PCR products were gel purified and reamplified with Dn2_3′UTRr1 reverse and $PTDnF_1$ forward primers and Dn2_5′UTRf1 forward and $PTDnR_2$ reverse primers, respectively. The amplifications were done by 30 cycles of 94°C for 30 sec, 50°C for 30 sec, and 72°C for 90 sec. Both new PCR products were gel purified. Then, the PCR hybridization was set up for amplification of full-length cDNA of *Dnbglu2*. The fifty microliter PCR reaction was composed of 10 μL Q solution, 4 μL of 2.5 mM dNTP, 4 μL of 10X Hot Star *Taq* buffer, 2 μL of 10X *Pfu* buffer, 2 μL each of 3′ end and 5′ end PCR products (250 ng each), 2 μL of Dn2_5′UTRf1 forward primer (10 μM), 2 μL of Dn2_3′UTRr1 reverse primer (10 μM), 0.25 μL of Hot Star *Taq* polymerase (1 unit), 0.4 μL of *Pfu* polymerase (1 unit) and made up to 50 μL with sterile dH_20. The amplification was done by 35 cycles of 94°C for 1 min, 60°C for 45 sec, and 72°C for 90 sec. Then, the PCR product was gel purified and reamplified again using *Pfu* polymerase at the same conditions. The fifty microliter PCR reaction was composed of 4 μL 2.5 mM dNTP, 5 μL 10X *Pfu* buffer, 1 μL Dn2_5′UTRf1 forward primer, 1 μL Dn2_3′UTRr1 reverse primer, 1 μL gel purified PCR product, 0.4 *Pfu* polymerase (1 unit) and dH_20 up to 50 μL total.

2.4.8 DNA analysis by agarose gel electrophoresis

The amplified PCR products were analyzed on 1% agarose gel electrophoresis in 1X 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 was prepared in a Pharmacia GNA-100 Gel Electrophoresis Apparatus (Amerscham Pharmacia Biotech). DNA samples were mixed 4:1 with 5X loading dye (0.15% bromophenol blue, 60% sucrose in TE, pH 8.0) and applied to the gel wells. Electrophoresis was performed at a constant 100 V for 40 min. After electrophoresis, the gel was stained in 0.1 μg/mL ethidium bromide solution for 5 min and then destained with water for 10 min. The bands in the gel were visualized and recorded by UV light transillumination with a Fluor-S^{TM} MultiImager (Bio-RAD). The sizes of DNA fragments were estimated by comparing with 1 kb ladder, 2-log ladder (New England Biolabs, MA), or 100 bp or 1 kb EZ load DNA markers (Bio-Rad, Richmond, CA, USA).

2.4.9 Purification of PCR products from agarose gel

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 reactions were separated on 1% TAE agarose gels. The DNA bands were excised and placed in 1.5 mL microtubes. The 3 volumes of QG buffer were added to 1 volume of gel (by weight) and the tube was incubated at 50^{o}C with shaking for 10 min or until the gel was completely dissolved. The samples were applied to the QIAQuick columns, which were centrifuged for 1 min at 12,000 rpm, and the flow through was discarded. 750 μL of PE buffer was added to the column

and centrifuged at 12,000 rpm for 1 min. The flow through was discarded and the column was centrifuged for an additional 1 min. The column was placed into a new 1.5 mL microtube, and 40 μL of distilled water was added, incubated for 1 min and centrifuged at 12,000 rpm for 1 min. The eluted DNA products were combined and kept at -30^{o}C.

2.4.10 Quantification and expected yield of DNA

A few microliters of DNA solution was mixed with distilled water to give 1 mL final volume. The absorbance at A_{260} and A_{280} nm of the 1 mL dilution solution was measured with a Lamda Bio20 UV/VIS Spectrophotometer (Perkin Elmer). The A_{260}/A_{280} ratio of 1.6-1.8 indicated good purity of DNA. One unit of absorbance at A_{260} nm is equivalent to 50 μg/mL of DNA (Sambrook *et al*., 1989). The DNA concentration was calculated with the following equation:

μg/mL of DNA = (A_{260}×dilution factor×50 μg/mL)

= {A_{260}×[1000 μL/mL/(x microliter of DNA solution)]×50 μg/mL}

2.4.11 Ligation of DNA fragments into vectors

Ligation of purified PCR products into pGEM-T, pGEM-T Easy or linearized pBlueScript SK(+) or pPICZαB-Thrombin plasmid expression vector was done according to the pGEMT plasmid supplier's recommendation (Promega). The reaction mixture (10 μL) was composed of 1 μL of pGEM-T vector or linearized pBlueScript SK(+) or pPICZαB-Thrombin vector (50 ng/μL), 5 μL of eluted DNA

fragment (100-200 ng), 1 μL of 10X T_4 DNA ligase buffer (30 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 10 mM DTT and 1 mM ATP) and 1μL T_4 DNA ligase. The reaction solution was incubated at 16°C overnight. Ligation into Zero Blunt® TOPO PCR cloning Kit vector was done according to the plasmid supplier's recommendation (Invitrogen). The reaction was composed of 3 μL DNA fragment (250 ng), 1 μL salt solution (1.2 M NaCl, 0.06 M $MgCl_2$), 1 μL TOPO vector (10 ng) and 1 μL H_2O. The reaction was mixed gently and incubated at room temperature for 5 min.

2.4.12 Preparation of competent *Escherichia coli* cells

Competent *E. coli* strain DH5α, Origami(DE3) or BL21(DE3) cells were prepared by the $CaCl_2$ method. A single colony of *E. coli* was inoculated into 2 mL of LB broth and grown at 37°C overnight with shaking at 200 rpm. Fifty microliters of inoculum culture was inoculated into a 250 mL flask containing 50 mL of LB broth and grown at 30°C with 250 rpm shaking until the optical density (OD) at 600 nm was about 0.4. The culture was transferred into the pre-chilled sterile polypropylene tube, chilled on ice for 5 min and centrifuged at 4,000 rpm, 4°C for 10 min to collect cell pellets. The cell pellets were resuspended with 10 mL of ice-cold 0.1 M $CaCl_2$. The resuspended cells were pelleted by centrifugation at 4,000 rpm, 4°C for 10 min. The pellets were resuspended with 2 mL of ice-cold 0.1 M $CaCl_2$, then 0.15 mL of DMSO was added and the tubes were chilled on ice for 10 min. The competent cells were aliquoted into microtubes (200 μL/tube) and used immediately or kept at -80°C.

2.4.13 Transformation and selection

For transformation, an aliquot of frozen competent cells were thawed 5 min on ice. The ligation reaction or plasmid (10-200 ng) was added to fresh or thawed competent cells, mixed by pipetting up and down and incubated on ice for 30 min. The plasmid was transformed by heat shocking the cells at 42^{o}C for 90 sec and quickly chilling on ice for 2 min. 0.8 mL of LB broth was added into the transformed competent cells and they were incubated at 37^{o}C with shaking at 200 rpm for 1 hr. The cells were collected by centrifugation at 12,000 rpm for 1 min, and the cell pellets were resuspended with 200 μL of LB broth. For blue/white colony selection of ligation into pBlueScript vector, the transformed DH5α cells were spread on LB plates containing 100 μg/mL ampicillin which was spread with 25 μL of 50 mg/ml X-gal and 100 μL of 0.1 M IPTG and incubated at 37^{o}C overnight. For antibiotic resistant selection of recombinant pET-40 clones, the transformed BL21(DE3) cells were spread on LB plates containing 15 μg/mL kanamycin and for the recombinant pET-32 clones, the transformed Origami(DE3) cells were spread on LB plates containing 15 μg/mL kanamycin, 12.5 μg/mL tetracycline and 50 μg/mL ampicillin.

2.4.14 Preparation of *Pichia* for Electroporation

A single colony of *Pichia pastoris* strain GS115 was inoculated in 50 mL YPD (1% yeast extract 2% peptone 2% glucose) in a 250 mL Erlermeyer flask, which was then incubated at 30^{o}C overnight with shaking at 250 rpm, until the OD_{600} of the culture reached 1.3-1.5. The cells were collected by centrifugation at 1,500 × g for 5 min at 4^{o}C. Then, the pellet was resuspended with 50 mL of ice-cold (0^{o}C),

sterile water and centrifuged again at 1,500 × g for 5 min at 4°C. The pellet was again resuspended with 25 mL ice-cold (0°C), sterile water, and centrifuged, as described above, then resuspended in 2 mL of ice-cold (0°C) 1 M sorbitol. The cell pellet was collected by centrifugation as before, then resuspended in 1 mL of ice-cold (0°C) 1 M sorbitol. The cells were kept on ice and used on the same day.

2.4.15 Transformation of *Pichia pastoris* by Electroporation

Eighty microliters of the cells from method 2.4.14 were mixed with 5-10 μg of linearized DNA and transferred to an ice-cold (0°C) 0.2 cm electroporation cuvette. The cells in the cuvette were incubated on ice for 5 min. The cells were pulsed using a GenePulser electroporator (BioRad) which generated a pulse length of approximately 5-10 ms with 1,500 V charging voltage, 25 μF capacitance and 200 Ω resistance. Immediately after that, 1 mL of ice-cold (0°C) 1 M sorbitol was added to the cuvette and the pulsed cells were transferred to a sterile 1.5 mL tube, and incubated at 30°C with shaking for 1 hr. The cells were collected by centrifugation at 5,000 rpm for 1 min, the pellet was resuspended with 0.2 mL YPD media and the cells spread on YPD agar plates containing 100 μg/mL zeocin. The plates were incubated at 30°C for 2-3 days until colonies formed. Single colonies were streaked on fresh YPD plates containing 100 μg/mL zeocin.

2.4.16 Plasmid isolation by boiling miniprep

Preparation of recombinant plasmid for preliminary analysis was done using the boiling prep method (Sambrook *et al*., 1989). Single *E. coli* colonies were picked with sterilized toothpicks and inoculated into 2 mL of LB medium containing appropriate antibiotics and incubated at 37^{o}C overnight with shaking at 220 rpm. The cell cultures were pelleted in 1.5 mL microtubes with a microfuge at 12,000 rpm for 1 min, the supernatant was poured off and the tubes drained well. The cell pellet was resuspended in 350 μL of STET, pH 8.0 (8% sucrose, 50 mM EDTA, 50 mM Tris-HCl and 0.5% of Triton X-100) by pipetting up and down with a P1000 micropipetter. Then, 25 μL freshly prepared lysozyme (10 mg/mL in water) was added and vortexed to mix. The samples were boiled for 40 seconds to break the cell walls and centrifuged immediately in a microfuge at 12,000 rpm for 10 min to pellet the broken cells. The gelatinous precipitate was removed with a tooth pick. The plasmids were precipitated by adding an equal volume of isopropanol (350 μL) to the supernatant and incubating at -20^{o}C or -80^{o}C for 30 min. Then, the DNA pellet was collected by centrifugation at 12,000 rpm for 15 min at 4^{o}C. The pellets were washed with 75% ethanol and dried for 10 min at 37^{o}C. The pellet was resuspended in 40 μL of TE buffer, pH 8.0 (10 mM Tris, 1 mM EDTA), and 2 μL of 1 mg/mL RNase A was added and incubated at 37^{o}C for 20 min. The plasmids were analyzed on 1% agarose gels, as described in method 2.4.8. The plasmids containing insert were kept at -30^{o}C.

2.4.17 Plasmid isolation by alkaline lysis with SDS

Preparation of recombinant plasmid for analysis and sequencing was done using the alkaline lysis method followed by polyethylene glycol precipitation, which was originally devised by Treisman (Sambrook *et al*., 1989). In this method, a solution of PEG in high salt is used to selectively precipitate the large plasmid DNA and leaved short RNA and DNA fragments in the supernatant. Single colonies were picked and inoculated into 33.3 mL of LB medium containing the appropriate antibiotic and incubated overnight at 37^{o}C with 200 rpm shaking. The pellet was collected by centrifugation at 6,000 rpm for 10 min, all the supernatant was removed and the tubes were drained well. The cell pellets were resuspended in 1.67 mL of Glucose/Tris/EDTA (50 mM glucose, 25 mM Tris-HCl, pH 8.0 and 10 mM EDTA) and allowed to stand at room temperature for 5 min. Then 3.33 mL of freshly prepared NaOH/SDS (0.2 M NaOH and 1% SDS) was added, mixed by gentle inversion, and the tubes were incubated on ice for 10 min. Then, 2.5 mL of ice cold solution III (3 M KOAc and 11.5% acetic acid in water) was added to the solution, mixed gently by inversion and left to stand on ice for 10 min. Removal of the lysed cell wall was done by centrifugation at 12,000 rpm for 20 min and then the supernatant was filtered with the Whatman No.1 filterpaper. The DNA was precipitated by adding 2.5 volumes of absolute ethanol and incubated at -20^{o}C for 30 min. The DNA was collected by centrifugation at 12,000 rpm for 20 min, the supernatant removed and the DNA pellet dried for 10 min at 37^{o}C. The pellet was resuspended in 0.7 mL of TE and transferred to a 1.5 mL microtube. The resuspended DNA was incubated at 65^{o}C for 20 min to completely dissolve the DNA, 2 μL of 10 mg/mL RNaseA was added and the reaction was incubated at 37^{o}C for 30 min. The

DNA solution was extracted twice with TE saturated phenol: 0.7 mL of phenol was added and mixed with vigorous shaking, then centrifuged at maximum speed in a microcentrifuge for 5 min. The upper phase was transferred to a new 1.5 mL microtube and extracted one more time. Then, the DNA phase was extracted by adding chloroform/isoamyl alcohol (24:1), mixing with vigorous shaking and centrifuging with maximum speed in a microfuge for 5 min. The upper phase was transferred to a new 1.5 mL microtube, 2.5 volumes of absolute ethanol and 0.1 volumes of 3 M NaOAc, pH 3.2, was added to the solution, and it was kept at -20^{o}C for 20 min. The DNA pellet was collected by centrifugation at 7,000 rpm for 15 min, dried for 10 min at 37^{o}C and dissolved with 100 μL of TE. Then, 25 μL of 4 M NaCl and 125 μL of 16.8% PEG_{6000} were added to the suspended DNA, vortexed briefly and placed on ice for 20 min. The purified DNA was recovered by centrifugation at maximum speed for 15 min at 4^{o}C. The DNA was washed with 70% ethanol and dried for 10 min at 37^{o}C. The plasmids were analyzed on 1% agarose gel and kept at -30^{o}C.

2.4.18 Extraction of genomic DNA from yeast

A single colony was inoculated into 2 mL YPD containing 100 μg/mL zeocin and incubated with shaking 200 rpm at 30^{o}C until the culture reached log phase (about 20 hrs). The pellet was collected by centrifugation at 3,000 rpm for 3 min. Then the cells were resuspended in 500 μL dH_20 by vortexing until suspended. Two hundred microliters of solution A (2% Triton X-100, 1% SDS, 0.1 M NaCl, 0.01 M Tris, pH 8.0, 1 mM EDTA), 0.3 g acid washed glass beads, and 200 μL of

phenol:chloroform:isoamyl alcohol (25:24:1) were added to the resuspended pellet, it was vortexed for 3 min, 200 μL of TE, pH 8.0, was added and the tube was spun down at maximum speed for 5 min. After that, the supernatant was carefully transferred to a new tube, 1 mL of absolute ethanol was added and the tube was mixed by inversion. The pellet was collected by centrifugation at maximum speed for 2 min. After the pellet was resuspended in 0.4 mL TE, pH 8.0, and 30 μg/mL of RNaseA, it was incubated at 37^{o}C for 5 min. Then, 18 μL of 5 M ammonium acetate (NH_4OAc) and 1 mL of absolute ethanol were added and mixed gently before the tube was stored at -20^{o}C for 3 hrs. The DNA pellet was collected by centrifugation at 12,000 rpm for 10 min, then the DNA was dried and resuspended in 40 μL of TE, pH 8.0.

2.4.19 Restriction enzyme digestion reactions

For the digestion of plasmids, including, pBlueScript SK(+), pET-32a, pET-40b and pPICZαB-thrombin vector, and plasmids with inserts or PCR product, reactions were performed in 20 μL total volume. DNA template (0.5-1 μg) was mixed with 0.1 volumes (2 μL) of 10x restriction enzyme buffer, 5-10 U of each restriction enzyme, 0.1 volumes (2 μL) of 1 mg/mL BSA (optional) and dH_2O up to 20 μL. The reaction was incubated at the 37^{o}C overnight. For the double digestion reaction, after the first digestion reaction was done and analyzed with agarose gel electrophoresis, the second digestion reaction was done by precipitating the DNA product from the first reaction with seeDNA, as described in 2.4.6, and then the reaction was set up by using the second restriction enzyme in a 20 μL total reaction

volume as above. After digestion was completed, the entire linearized sample was purified, as described in method 2.4.9.

2.4.20 DNA sequencing

Purified DNA samples prepared by the method in 2.4.17 were sequenced using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems, Foster City, CA) with the recommended protocol. The reaction mix was composed of 2 μL Dye Terminator Ready Reaction Mix, 400 ng of purified plasmid DNA, 1.6 pmol of primer and distilled water to bring the volumn up to 10 μL. Amplification was done with the Gene Amp PCR system 9700 (Applied Biosystems) and the following conditions for 25 cycles: 96^{o}C for 20 sec, 50^{o}C for 10 sec and 60^{o}C for 4 min. After the last cycle, the temperature was rapidly decreased to 4^{o}C and held until the sample could be purified.

The amplified products were purified by ethanol/sodium acetate precipitation in microcentrifuge tubes. For each sequencing reaction, the entire contents of each extension reaction were pipetted into a 1.5 mL microcentrifuge tube, which contained 2 μL of 3 M NaOAc, pH 4.6, and 50 μL of 95% ethanol, and mixed thoroughly. The tubes were vortexed and left at room temperature for 15 min to precipitate the extension products, then centrifuged at 12,000 rpm for 20 min. The supernatant was completely removed with a pipette and the pellet rinsed with 250 μL of 70% ethanol vortexed briefly. The tubes were centrifuged 5 min in a microcentrifuge at 12,000 rpm and the supernatant was carefully aspirated away. The pellets were air dried and each sample was resuspended in 15 μL of template

suppression reagent (TSR) by vortexing and the sample spun down. The samples were heated at 95^{o}C for 2 min to denature, then chilled on ice. The samples, were again centrifuge to collect them in the bottoms of the tubes and sequenced on an ABI PRISM 310 Genetic Analyzer or other automated sequencer (Applied Biosystems).

2.4.21 Analysis of DNA sequences

The DNA sequence chromatograms from the ABI PRISM 310 Genetic Analyzer were inspected and corrected with the BioEdit, Biological sequence alignment editor program (http://www.mbio.ncsu.edu/BioEdit/bioedit.html) and analyzed using the computer analysis programs at the BCM Search Launcher (http://searchlauncher.bcm.tmc.edu/). The sequence was translated to deduced amino acid sequence by the six frame translation program and antisense sequences converted to sense strand sequence with Reverse complement program in the BCM Sequence Utilities. The sequences were confirmed to be related to Thai rosewood β-glucosidase and other related β-glucosidases by the BLAST program at the National Center for Biotechnological Information (NCBI, http://www.ncbi.nlm.nih.gov (Altschul *et al.*, 1997). Comparison of DNA fragment sequences was done with the GENETYX-WIN V.3.2 genetic information software or BLAST2 at NCBI. The protein signal sequence was predicted by the Signal P program (http://www.cbs.dtu.dk/services/SignalP/, Bendtsen *et al*., 2004), and the mature protein MW and pI predicted by the protein prediction program at EXPASY (http://www.expasy.org).

2.5 Expression of protein in *E. coli*

2.5.1 Target DNA insert preparation

The full length cDNA which was cloned in pBlueScript SK vector was used as template to introduce an *Eco*R I site at the initial codon of the predicted mature protein to eliminate the signal peptide sequence by PCR amplification. The 2 pieces of PCR product were amplified using 2 sets of primers, T7 promoter forward and EcoRIMR reverse primers (Table 2.1) and EcoRIMF forward and BseRIR reverse primers (Table 2.1), respectively. The amplification was done with 30 cycles of 94^{o}C for 15 sec, 50^{o}C for 30 sec, and 72^{o}C for 70 sec. The two PCR products were purified as described in method 2.4.9. After that, both PCR products (188 bp and 307 bp) were used as primers to amplify the piece with the introduced *Eco*R I site using the full length cDNA as a template, the amplification was done using *Taq* polymerase with 30 cycles of 94^{o}C for 15 sec, 50^{o}C for 30 sec, and 72^{o}C for 70 sec. The expected sized 474 bp PCR product was purified as described in method 2.4.9. The purified PCR product was digested with *Pst* I and *BseR* I enzyme (New England Biolabs). Then, the digested DNA was ligated back into the full length cDNA in pBlueScript SK plasmid, which was cut with the same restriction enzymes, as described in method 2.4.11. The ligation reaction was transformed into *E. coli* strain DH5α and screened as described in methods 2.4.12-2.4.13. After confirming the *Eco*RI site which was introduced at the initial codon of the mature protein, both by digestion with restriction enzymes (*Eco*R I, *Pst* I and *BseR* I) and DNA sequencing, the mature cDNA encoding Dnbglu1 was cut from the full length cDNA construct in pBlueScript SK(+) (with the introduced *Eco*RI site) by digestion with *Eco*RI and *Not* I enzyme and gel purified.

2.5.2 Expression vector preparation

Each pET-32a and pET-40b vector was cut with the same restriction enzymes (*Eco*RI and *Not* I) and gel purified, as described in method 2.4.9.

2.5.3 Ligation and transformation

The ligation reaction and transformation were done as described in methods 2.4.11-2.4.13, respectively. The colony screening was done by preparing plasmids by the boiling miniprep as described in method 2.4.16 and digestion with *Eco*RI and *Not* I restriction enzymes, as described in method 2.4.11.

2.5.4 Expression of β-glucosidase fusion proteins in *E. coli*

A single colony was picked from a freshly streaked plate and innoculated into 50 mL LB containing 50 μg/mL ampicillin and 15 μg/mL kanamycin for transformants of pET-40b in BL21(DE3), and 50 μg/mL ampicillin, 15 μg/mL kanamycin and 12.5 μg/mL tetracycline for transformants of pET-32a in Origami(DE3). The culture was incubated in a 250 mL Erlenmeyer flask with shaking at 37^{o}C, 220 rpm until the OD_{600} reached 0.6. The IPTG (100 mM) was added to 0.4 mM final concentration in the 50 mL LB media, and the incubation continued for 16 hrs at 20^{o}C. Induced cultures were harvested by centrifugation at 8,000 × g at 4^{o}C for 10 min, the pellets were washed with distilled water and the cells kept at -80^{o}C for 1 hr. The cell pellets were resuspended in 5 mL/gram freshly prepared extraction buffer (20 mM Tris-HCl, pH 8.0, 200 μg/mL lysozyme, 1% Triton-X 100, 1 mM phenylmethylsulfonylfluoride (PMSF)), 40 μg/mL Dnase I), and incubated at room

temperature for 30 min. Then, cells were incubated on ice for 10 min and soluble and insoluble fractions were separated by centrifugation at 12,000 rpm at 4°C for 15 min. The insoluble pellet was resuspended in 5X sample buffer and stored at -20°C until SDS-PAGE analysis. The soluble fraction was kept at 4°C.

2.6 Expression of protein in *Pichia pastoris*

2.6.1 Expression vector preparation

The pPICZαB-Thrombin plasmid was prepared by digestion with *Pst* I and *Sac* II restriction enzyme for construction of *Dnbglu1* and *Pst* I and *Xba* I restriction enzyme for construction of *Dnbglu*2. Then, the linearized plasmid vector was gel purified as described in method 2.4.9.

For Dnbglu1, the full length cDNA construct in pBlueScript SK(+) vector was used as template to introduce a *Pst* I site at the initial codon of the mature protein and eliminate the signal peptide sequence, and a *Sac* II site at the 3′ end of the coding region. The amplification was done by 30 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 3 min using *Pfu* polymerase and PstIgluF102 forward and DNBgluR2SacII reverse primers (Table 2.1). The PCR product with the introduced *Pst* I and *Sac* II sites was purified as described in method 2.4.9, and ligated into pBlueScript SK(+) plasmid which had been linearized with *EcoR* V restriction enzyme, as described in method 2.4.11, and transformed into DH5α competent cells, as described in method 2.4.12-2.4.13. The colonies were checked for insert by digestion with *Pst* I and *Sac* II restriction enzymes. The insert was cut out from

pBlueScrip SK(+) plasmid by digestion with *Pst* I and *Sac* II restriction enzymes and gel purified, as described in method 2.4.9. Then the insert was ligated into the digested pPICZα-Thrombin plasmid and transformed into DH5α competent cells, as described in method 2.4.12-2.4.13, except that the colonies were selected on low salt LB agar plates (1% tryptone, 0.5% NaCl, and 0.5% yeast extract) containing 25 μg/mL Zeocin.

For Dnbglu2 expression, the cDNA encoding the Dnbglu2 protein was amplified using primers to introduce a *Pst* I site at the amino terminus of the mature protein and *Xba* I site at the 3′ end of the coding region. The PCR hybridization was set up for amplification of mature cDNA of *Dnbglu2* using the same method as 2.4.7 but with different primers. Fifty microliters of PCR reaction composed of 10 μL Q solution, 4 μL of 2.5 mM dNTP, 4 μL of 10X Hot Star *Taq* buffer, 2 μL of 10X *Pfu* buffer, 2 μL of each PCR product (250 ng each), 2 μL of Dn2NTERMPstIf forward primer, 2 μL of Dn2CTERMXbaIr reverse primer, 0.25 μL of Hot Star *Taq* polymerase, 0.4 of *Pfu* polymerase (1 unit) and made up to 50 μL with dH_20. The amplification was done by 35 cycles of 94^oC for 1 min, 60^oC for 45 sec, and 72^oC for 90 sec. Then, the PCR product was gel purified and reamplified again using *Pfu* polymerase at under the same conditions. The fifty microliter PCR reaction was composed of 4 μL of 2.5 mM dNTP, 5 μL of 10X *Pfu* buffer, 1 μL of Dn2NTERMPSstIf forward primer, 1 μL of Dn2CTERMXbaIr reverse primer, 1 μL of gel purified PCR product, 0.4 of *Pfu* polymerase (1 unit) and made up to 50 μL with dH_20. The PCR product was gel purified and ligated into Zero Blunt® TOPO PCR cloning kit vector (Invitrogen), as described in 2.4.11.

The clones containing insert were digested with *Pst* I and *Xba* I, and the insert was gel purified and ligated into the digested pPICZα-Thrombin plasmid. Then, the plasmid was transformed into DH5α competent cells, as described in method 2.4.12-2.4.13, but the colonies were selected on low salt LB agar plates with zeocin, as described above.

2.6.2 Expression of Dnbglu protein in transformed *P. pastoris*

The DNA sample from method 2.6.1 was transformed into *Pichia pastoris* as described in methods 2.4.14 and 2.4.15.

The selected clones were inoculated in 25 mL BMGY medium (1% yeast extract, 2% peptone, 100 mM potassium phosphate buffer, pH 6.0, 1.34% YNB, 4×10^{-5}% biotin, and 1% glycerol) at 30°C, 250 rpm until the OD_{600} reached 2-3, then the cell pellet was collected by centrifugation at 3,000 × g at room temperature for 5 min and resuspended in 5 mL of BMMY medium (1% yeast extract, 2% peptone, 100 mM potassium phosphate buffer, pH 6.0, 1.34% YNB, 4×10^{-5}% biotin, and 1% MeOH). Then, methanol was added to 1% final concentration every 24 hrs and the cell culture was incubated with shaking at 250 rpm at 30°C for 3-5 days. Induced cultures were harvested by centrifugation at 12,000 × g at 4°C for 5 min. The media were tested for β-fucosidase expression using *p*NP-β-fucoside hydrolysis. The clone with the highest expression of β-fucosidase activity was chosen and used to inoculate in 25 mL of BMGY medium as starter culture, which was used to innoculate in a 1 liter of BMGY. The cells were grown for 16-18 hrs at 30°C, 250 rpm until the culture reached log phase growth (OD_{600} = 2-6). Cells were harvested by centrifugation at

3,000 × g at room temperature for 5 min and washed twice with BMMY medium, then resupended the in 200 mL of BMMY medium which was supplemented with methanol to 1% final concentration every 24 hrs, and the media supernatant was collected after 5 days by centrifugation, as described above. The expressed recombinant enzyme was partially purified using Sephacryl S-300 gel filtration with 0.3 M NaCl in 50 mM Tris-HCl buffer, pH 7.0, at a flow rate of 0.8 mL/min. Kinetic parameters (K_m) were determined using pNP-β-D-glucoside, *p*NP-β-D-fucoside and natural substrates, as described previously.

CHAPTER III

RESULTS

3.1 Purification of β-glucosidase/β-fucosidase from *Dalbergia nigrescens* Kurz.

The enzyme was extracted from seeds and purified by 35-75% ammonium sulfate fractionation, DEAE anion exchange chromatography (Figure 3.1) and Sephacryl S-300 gel filtration chromatography (Figure 3.2). Both *p*NP-β-D-glucoside and *p*NP-β-D-fucoside hydrolysis activities were determined during the purification. The yields for β-glucosidase and β-fucosidase were similar, but total activities and specific activities of β-fucosidase were 2 times higher than those of β-glucosidase at each step, as shown in Table 3.1. The enzyme was purified 33 and 49 fold over the crude extract for β-glucosidase and β-fucosidase, respectively, with 2-3% yield of both total activities of about 1.8 mg of enzyme per 50 g of seed.

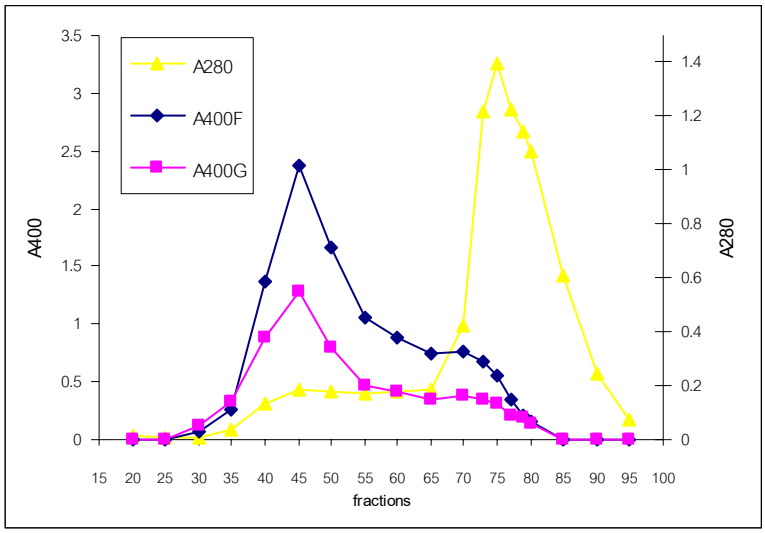


Figure 3.1 DEAE sepharose chromatography of the dialyzed 35-75% ammonium sulfate precipitate of *D. nigrescens* seed extract. The column was equilibrated with 0.025 M Tris-HCl, pH 7.0, then, eluted with 0.025 M Tris-HCl, pH 7.0, with a gradient from 0 to 0.5 M NaCl. Fractions 35-47 were pooled for gel filtration. (A400F; *p*NP-β-D-fucoside activity and A400G; *p*NP-β-D-glucoside activity in terms of absorbance at 400 nm for 0.005 mL of fractions in the standard assay).

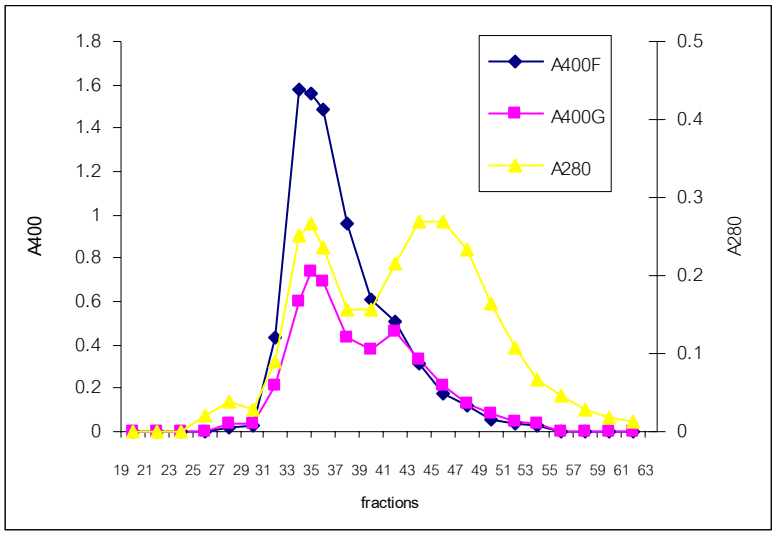


Figure 3.2 The purification profile on Sephacryl S-300 filtration of pooled peak fractions from the DEAE column. Fractions containing enzyme activity were pooled and concentrated with Centriprep-50 ultrafiltration unit and subjected to gel filtration chromatography on a S-300 column (1.5 cm × 62 cm), in 0.05 M Tris-HCl buffer, pH 7.0, with 0.3 M NaCl, collecting 2 mL fractions.

Table 3.1 Purification of β-glucosidase/β-fucosidase from *Dalbergia nigrescens* Kurz. Fifty grams of seeds were used for purification. The assays were performed with 1 mM *p*NP-β-D-glucoside or 1 mM *p*NP-β-D-fucoside.

Fraction	Total activity (nkat)	Total protein (mg)	Specific activity (nkat/mg)	Purification (fold)	Yield (%)
Crude extract		3040			
β-Glucosidase	6125		2.02	1.00	100
β-fucosidase	9467		3.11	1.00	100
35-75% $(NH_4)_2SO_4$		1386			
β-Glucosidase	4220		3.04	1.50	68.9
β-fucosidase	8,675		6.26	2.01	91.6
DEAE		12.27			
β-Glucosidase	574		46.8	23.1	9.37
β-fucosidase	1008		82.2	26.4	10.6
S 300 gel filtration		1.79			
β-Glucosidase	119		66.6	33.0	1.94
β-fucosidase	271		152	48.9	2.86

3.1.1 Molecular weight estimation

Native *D. nigrescens* β-glucosidase appeared as a single major band on SDS-PAGE, which was estimated to be 62-63 kDa (Figure 3.3) by comparison to low range molecular weight protein standards. Figure 3.4 illustrates the calibration curve of denatured molecular weight standard proteins obtained with *D. nigrescens* β-glucosidase purified from seeds. The native molecular weight of the purified enzyme was estimated to be 240 kDa using Sephacryl S-300 gel filtration chromatography, which suggests that *D. nigrescens* β-glucosidase is a tetrameric protein. Figure 3.5 illustrates the calibration curve of molecular weight standards obtained for gel filtration of native *D. nigrescens* β-glucosidase protein from seeds.

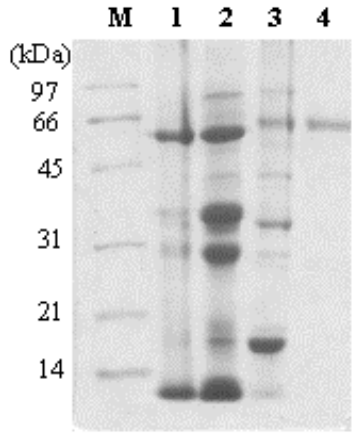


Figure 3.3 SDS-PAGE analysis of purified β-glucosidase. Lane M, Low-range protein markers**;** 1, crude extract; 2, 35-75% $(NH_4)_2SO_4$ fractionated; 3, DEAE pooled fractions; 4, S 300 gel filtration.

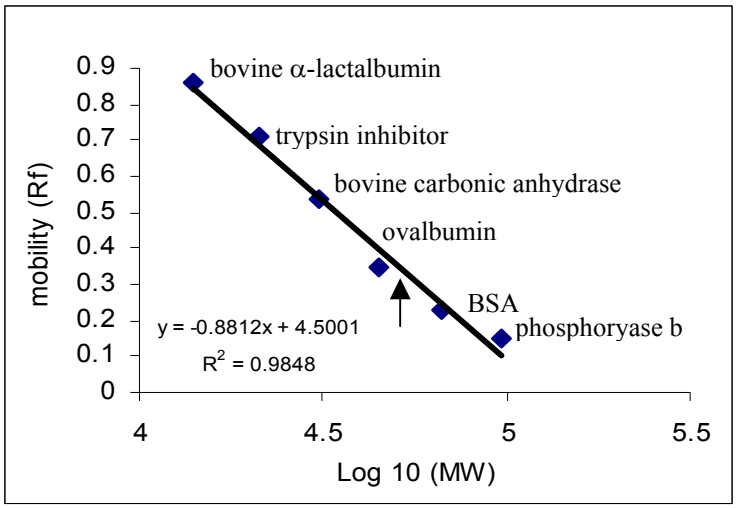


Figure 3.4 Mobilities (R_f) of standard molecular weight proteins plotted against Log_{10} of their molecular weights. Purified *D. nigrescens* β-glucosidase and protein standards were separated on 10% SDS-PAGE. The distances migrated by protein bands divided by that of the dye front were calculated as R_f values and the graph was plotted vs Log_{10} of molecular weight standard proteins. The calibration curve obtained with the standard proteins was used to calculate the molecular weight of *D. nigrescens* β-glucosidase (↑).

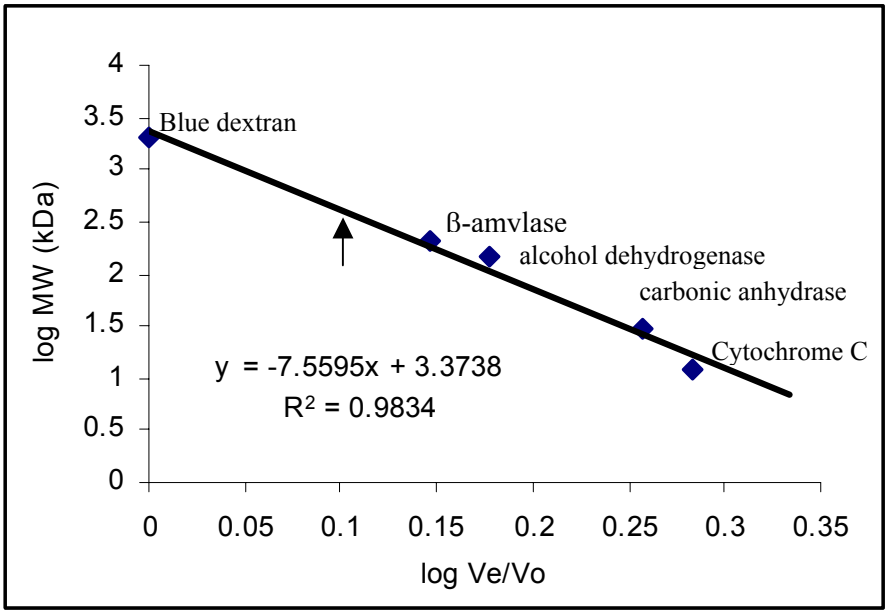


Figure 3.5 Native molecular weight of *D. nigrescens* β-glucosidase estimation by size exclusion column chromatography. Purified *D. nigrescens* β-glucosidase and protein standards were loaded onto a Sephacryl S-300 column. The fractions were collected and the protein absorbance measured at A_{280} for protein standards, and *p*NP-Glc activity assayed for *D. nigrescens* β-glucosidase. The elution volume of each protein (V_e), and of blue dextran (V_0) were calculated. The log V_e/V_0 and log molecular weight (kDa) for each protein were plotted. The calibration curve obtained with the standard proteins was used to calculate the native molecular weight of *D. nigrescens* β-glucosidase (↑).

3.1.2 Profile of pH and temperature optimum activity

D. nigrescens β-glucosidase has optimum activity in the pH range 5.5-6.0 (Figure 3.6) and at a pH 5.0, the enzyme has an optimum temperature at 65°C with 10 min incubation time (Figure 3.7). The enzyme showed similar activity toward both substrates at every pH, but the *p*NP-fucoside activity was significantly higher than *p*NP-glucoside when the temperature was increased from 55-75°C. The enzyme still had significant activity until the temperature reached 80°C.

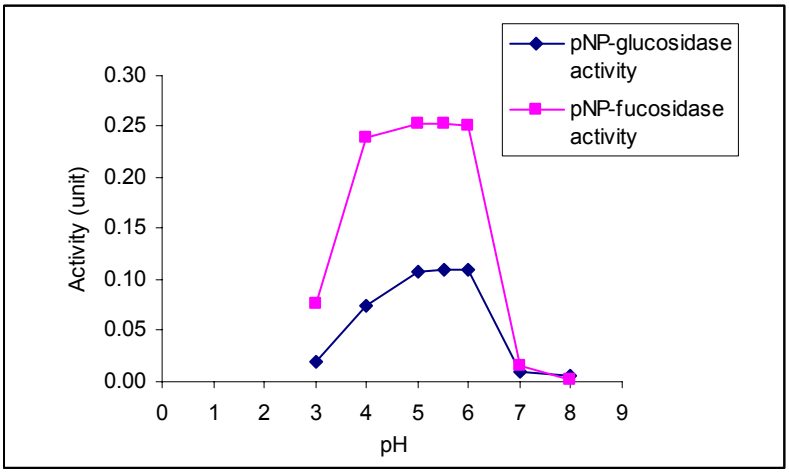


Figure 3.6 pH profile of activity for *D. nigrescens* β-glucosidase over a pH range 3-8 assayed with 1 mM *p*NP-glucoside and *p*NP-fucoside at 30°C for 10 min.

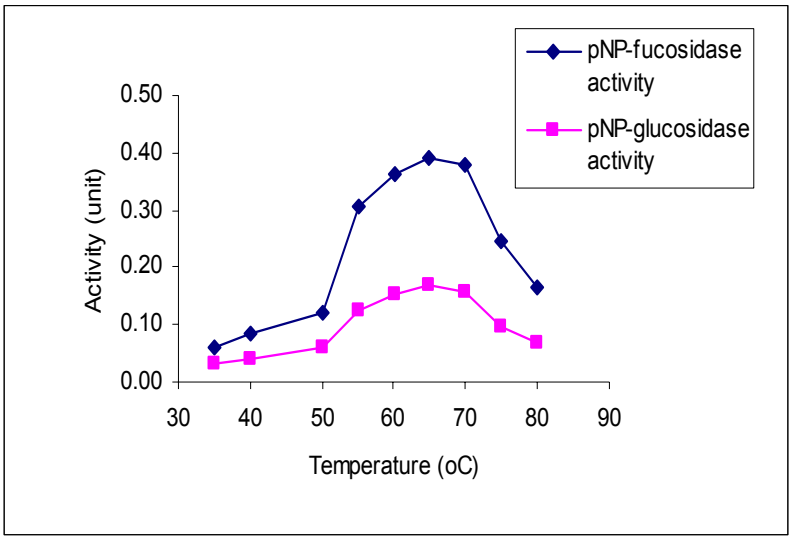


Figure 3.7 Temperature profile of *D. nigrescens* β-glucosidase at temperatures ranging from 35-80°C assayed with 1 mM *p*NP-glucoside and *p*NP-fucoside at pH 5.0 for 10 min.

3.1.3 Characterization of *Dalbergia nigrescens* β-glucosidase/β-fucosidase activity

The purified β-glucosidase was run on a native acrylamide gel and stained separately with 1 mM synthetic substrates, 4-methyl-umbelliferyl-β-D-fucopyranoside, 4-methyl-umbelliferyl-β-D-glucopyranoside and 4-methyl-umbelliferyl-β-D-galactopyranoside. As shown in Figure 3.8, the enzyme had activity toward all these substrates, but appeared to have less activity toward 4-methyl-umbelliferyl-β-D-galactopyranoside.

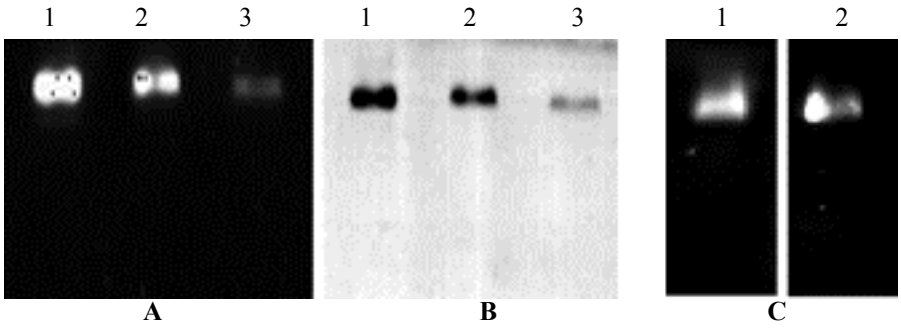


Figure 3.8 Activity staining gel of non denaturing electrophoresis of *D. nigrescens* β-glucosidase/β-fucosidase. A. Activity staining gel; lane 1, 4.8 μg; lane 2, 2.4 μg; lane 3, 1.2 μg of enzyme stained with 1 mM 4-MU-β-D-glucoside. B. lane 1, 4.8 μg; lane 2, 2.4 μg; lane 3, 1.2 μg of enzyme stained with Coomassie blue R-250. C. 10 μg of enzyme stained with (1) 1 mM 4-MU-β-D-fucoside and (2) 1 mM 4-MU-β-D-galactoside.

Studies of the relative activity of the enzyme with several available synthetic substrates and natural substrates showed that these compounds were hydrolyzed poorly or not at all relative to *p*NP-glucoside and *p*NP-fucoside (Table 3.2 and Table 3.3). Comparison of hydrolysis of synthetic substrates by *D. nigrescens* and *D. cochinchinensis* β-glucosidase indicated both enzymes have similar patterns of relative activity, except for *p*NP-β-D-galactoside and *p*NP-β-D-xyloside, for which the data was reversed. *D. nigrescens* β-glucosidase had higher relative activity toward *p*NP-β-D-xyloside than *D. cochinchinensis* β-glucosidase and it had less relative

activity toward *p*NP-α-L-arabinoside than *D. cochinchinensis* β-glucosidase. *D. nigrescens* β-glucosidase could hydrolyze linamarin and the disaccharides, laminaribiose and cellobiose, but not the corresponding trisaccharides, though the activity against these substrates was very low.

Table 3.2 Comparison of hydrolysis of synthetic substrates by *D. nigrescens* and *D. cochinchinensis* β-glucosidase. The reaction was performed with 5 mM glycosides in 0.1 M sodium acetate, pH 5.0 at 30^{o}C for 10 min, and reported relative to *p*NP-β-D-glucoside hydrolysis**.**

Substrate	***D. nigrescens*** **% Relative activity**	***D. cochinchinensis*** **%Relative activity**
*p*NP-β-D-glucoside	100	100
*p*NP-β-D-fucoside	124	124
*p*NP-β-D-galactoside	3.97	8.95
*p*NP-β-D-xyloside	7.55	3.91
*p*NP-α-L-arabinoside	1.91	4.89
*p*NP-β-D-thioglucoside	0.68	0.02
*p*NP-β-L-arabinoside	0.47	-
*p*NP-β-D-mannoside	0.40	0.26

- means not determined

Table 3.3 Relative hydrolytic activity toward natural glucosides and oligosaccharides. The reaction was performed with 2 mM various substrates in 0.1 M sodium acetate, pH 5.0 at 30°C, for 10 min, and reported relative to *p*NP-β-D-glucoside hydrolysis**.**

Substrates	% Relative activity
*p*NP-β-D-glucoside	100
Laminaribiose (β1→3)	1.01
Laminaritriose (β1→3)	nd
Cellobiose (β1→4)	0.33
Cellotriose (β1→4)	nd
Linamarin	0.37
Salicin	nd
Dhurrin	nd
DIMBOA glucoside	nd
Sophorose	nd
Amygdalin	nd

nd means not-detectable

3.1.4 Inhibition studies

The effect of various metal ions and δ-gluconolactone was tested on *D. nigrescens* enzyme β-glucosidase and β-fucosidase activities. The results are shown in Table 3.4. The effect of various compounds was tested on the hydrolysis of *p*NP-β-D-glucoside and *p*NP-β-D-fucoside. Most of these substances had little effect on either β-glucosidase or β-fucosidase, except for δ-gluconolactone and 2,4-Dinitrophenyl-2-deoxy-2-fluoroglucoside.

Table 3.4 Effect of metal ions and δ-gluconolactone on *D. nigrescens* enzyme β-glucosidase and β-fucosidase activities. Substances were preincubated with the enzyme 10 min before assaying the hydrolysis of *p*NP-β-D-glucoside (2 mM) and *p*NP-β-D-fucoside (1 mM) in 0.1 M sodium acetate buffer, pH 5.0.

Substrate	Final concentration	% β-D-glucosidase activity remaining	% β-D-fucosidase activity remaining
control	1 mM	100	100
$FeSO_4$	1 mM	115	109
$CaCl_2$	1 mM	115	109
EDTA	1 mM	118	115
$MnCl_2$	1 mM	113	114
$ZnSO_4$	1 mM	112	116
$HgCl_2$	1 mM	90	94
δ-gluconolactone	1 mM	66	55
	3 mM	55	40
2,4-Dinitrophenyl-2-deoxy-2-fluoroglucoside	10 μM	26	53
	40 μM	nd	nd

nd means not detecable

3.1.5 Determination of kinetic parameters for *p*NP -β-D-glucoside and *p*NP-β-D fucoside hydrolysis by *D. nigrescens* enzyme

Kinetic parameters (the K_m, k_{cat}, and k_{cat}/K_m values) of purified native *D. nigrescens* β-glucosidases were determined for release of *p*-nitrophenol from *p*NP-β-D-glucoside and *p*NP-β-D-fucoside. The native enzymes showed K_m values toward *p*NP-β-D-glucoside and *p*NP-β-D-fucoside of 14.7 mM and 1.8 mM, respectively. The k_{cat} and k_{cat}/K_m values of native *D. nigrescens* enzyme on *p*NP-β-D-glucoside were 10.4 s^{-1} and 876 $M^{-1}s^{-1}$, respectively, and on *p*NP-β-D-fucoside were 7.0 s^{-1} and

4,020 $M^{-1}s^{-1}$, respectively. The k_{cat} values were calculated assuming a subunit molecular weight of 61,858 Da. The kinetic data are summarized in Table 3.5.

Table 3.5 The kinetic properties of purified native *D. nigrescens* β-glucosidase enzyme. Assays were performed in 0.1 M sodium acetate buffer, pH 5, at 30°C. $k_{cat\ and}$ k_{cat} /K_m of native enzyme were estimated assuming a subunit molecular weight of 61,858 Da.

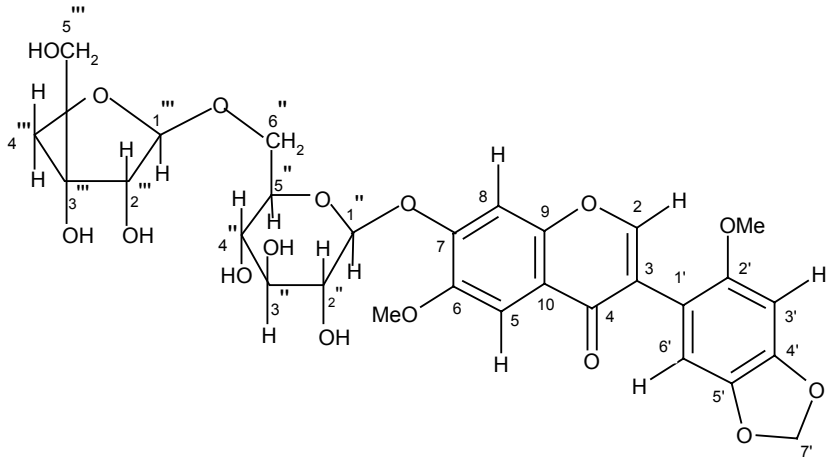
Substrate	K_m (mM)	k_{cat} (s^{-1})	k_{cat} /K_m ($M^{-1}s^{-1}$)
*p*NP-β-D-glucoside	14.7 ± 2.9	10.4 ± 2.3	876 ± 86
*p*NP-β-D-fucoside	1.8 ± 0.2	7.0 ± 0.3	4,020 ± 220

3.1.6 Structural analysis of glycoside substrates and their aglycone

The glycoside substrates of *D. nigrescens* β-glucosidase were identified on TLC by comparison of the crude glycoside extract before and after digestion with *D. nigrescens* β-glucosidase. The glycoside substrate spots would disappear after digestion. These substrates were purified from the crude extract by LH-20 column chromatography using methanol as eluent solvent and preparative TLC. The aglycones were also isolated by digesting the substrates with *D. nigrescens* β-glucosidase and then purifying the aglycones by preparative TLC.

The purified glycoside and aglycone structures of both compounds were determined by NMR, mass spectrometry and infrared spectroscopy. For the NMR determination, both the glycoside and aglycone of the S1 compound were dissolved in DMSO-$_{d6}$ and S2 glycoside and S2 aglycone were dissolved in D_2O and CD_3OD,

respectively. The structures of the S1 and S2 compound are shown in Figure 3.9 and Figure 3.10, respectively, and the ^{1}H and ^{13}C NMR spectral data of the S1 and S2 compounds and S2 aglycone are summarized in Table 3.6.



Compound 1

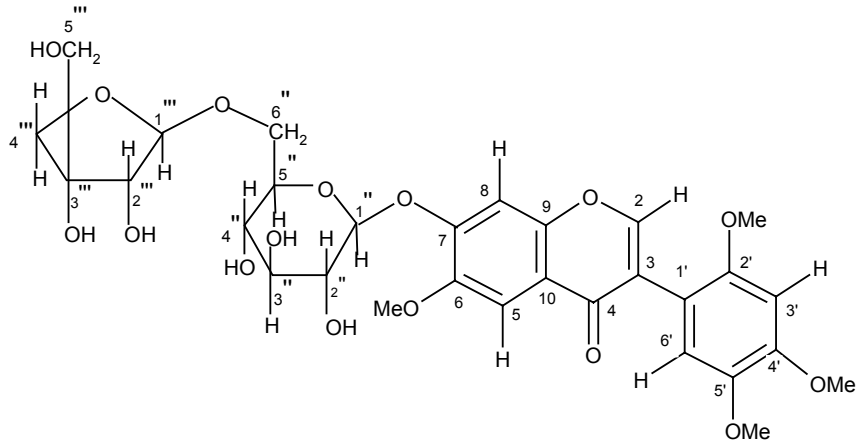
Figure 3.9 The structure of the S1 substrate.

Compound 1 glycoside (S1) was obtained as a light yellow powder. The UV spectrum showed the characteristic absorption of an isoflavone compound (λ^{MeOH}_{max} 260 and 315 nm). From FTMS, this compound showed a molecular related ion peak at m/z 659 $[M+Na]^{+}$, which corresponded to the molecular formula $C_{29}O_{16}H_{32}$ and to the 29 peaks in the ^{13}C NMR spectrum (Figure 3.15). Further fragmentation of this signal gave a peak at m/z 365 $[M+Na]^{+}$ (Figure 3.11), which

corresponded to the mass spectrum of the aglycone part after it was digested with *D. nigrescens* β-glucosidase and the molecular formula $C_{18}O_7H_{14}$ (Figure 3.13). The formula $C_{18}O_7H_{13}$ of the aglycone is consistent with an isoflavone (C_6-C_3-C_6) which contains two methoxy groups and one O-CH_2-O at C-4′ and C-5′. The 294 amu derived from mass subtraction indicated 2 sugar rings of five and six carbon, which corresponded to the 11 carbon signals obtained by subtraction of the aglycone carbon signals from the total number of glycoside ^{13}C signals observed in the ^{13}C NMR. After hydrolysis with enzyme or with acid, the spectrum of the purified aglycone, compound 1a, lacked these signals. The HMBC spectrum of compound 1 revealed cross peaks between two methoxy methyl protons at δ_H 3.66 and δ_H 3.93 with the C-6 aromatic carbon position at δ_C 147.31 and C-2′ position at δ_C 152.77. The ^1H NMR spectrum (Figure 3.16) showed five singlet signals at δ8.19, δ7.44, δ7.32, δ6.87, and δ6.83, which are attributable to hydrogens at H-2, H-5, H-8, H-3′ and H-6′ of an isoflavone derivative, as shown in Table 3.6. The COSY experiment showed no proton cross peak, which confirmed the para position of the protons at C-8 and C-5. The ^1H NMR signals for the aglycone in DMSO-*d6* matched those previously reported for dalpatein from *Dalbergia paniculata* (Adinarayana and Rao, 1975, Tahara *et al*., 1992 and Radhakrishniah, 1973)

A D-apio-β-D-furanoside moiety was determined by the chemical shift at $\delta_{C1'''}$ 109.48, which is ascribable to the anomeric dioxymethine carbon on an apiose ring, along with the carbonic signals of one quaternary carbon at $\delta_{C3'''}$ 78.60, two oxymethylenes at $\delta_{C4'''}$ 73.30 and $\delta_{C5'''}$ 63.32, and one oxymethine at $\delta_{C2'''}$ 72.98. The remaining ^{13}C signals which corresponded to carbonic carbons were used to elucidate

a β-D-glucopyranoside moiety with one anomeric dioxymethine $\delta_{C1''}$ at 99.90, and four monooxymethine carbons at $\delta_{C2''}$ 72.98, $\delta_{C3''}$ 76.67, $\delta_{C4'''}$ 69.80 and $\delta_{C5'''}$ 75.65. The chemical shift at $\delta_{C6''}$ 67.85 was used to define the disaccharide linkage as apiofuranosyl (C1‴ → C6″) because the chemical shift of a monosaccharide oxymethylene group appears at about δ_C 62 (Mathias *et al.,* 1998). Also, the chemical shift at $\delta_{C1'''}$ 109.48 indicated the C-1 anomeric orientation of apiose was a β-configuration because the ^{13}C NMR of α-D-apiose would appear at about δ_C 104.5 (Takahashi *et al.*, 2001). The NMR spectral data of the sugar part are in agreement with those of tectorigenin 7-O-[β-D-apiofuranosyl-(1→6)-β-D-glucopyranoside] from *Dalbergia sissoo* (Farag *et al.*, 1999).



Compound 2

Figure 3.10 The structure of S2 substrate.

Compound 2 glycoside was also obtained as a light yellow powder and had a molecular related ion peak $[M+Na]^+$ at m/z 675 in MS, which corresponded to the molecular formula $C_{30}O_{16}H_{36}$ and to the 30 peaks in the ^{13}C NMR spectrum (Figure 3.17). The future fragmentation of this signal gave an ion at m/z 381 (Figure 3.12), which corresponded to the mass spectrum of the aglycone released when S2 was digested with *D. nigrescens* β-glucosidase and molecular formula $C_{19}O_7H_{18}$ (Figure 3.14). The UV spectrum showed the characteristic absorption of an isoflavone compound (λ^{MeOH}_{max} 260 and 315 nm). The ^1H NMR spectrum (Figure 3.18) revealed 5 singlet signals at δ 7.93, δ 7.39, δ 6.58, δ 6.80 and δ 6.96, attributable to hydrogens H-2, H-5, H-8, H-3′ and H-6′ of an isoflavone derivative. The gHMBC spectrum had cross peaks between 4 methoxy groups which had signals at $\delta_{H-2'}$ 3.68(s), δ_{H-6} 3.71(s), $\delta_{H-4'}$ 3.85(s) and $\delta_{H-5'}$ 3.82 and $\delta_{C-2'}$ 152.16, δ_{C-6} 151.43, $\delta_{C-4'}$ 149.75 and $\delta_{C-5'}$ 142.35. The 11 carbon signals in the sugar region corresponded to one pentose and one hexose sugar ring, with nearly the same spectrum as in compound 1, though the compound 2 structure was solved in D_2O rather than DMSO-*d6* as solvent. From this comparison, the two compounds appeared to have the same sugar moiety structure. Again, these sugars were removed by hydrolysis with *D. nigrescens* β-glucosidase to give compound 2a. The compound 2a structure was solved in CD_3OD because it was less soluble in D_2O than compound 2. The ^1H NMR spectrum showed 5 singlet signals at δ 7.87, δ 6.50, δ 7.33, δ 6.73 and δ 6.93, indicating the hydrogen atoms at δ_{C-2} 156.24, δ_{C-5} 102.72, δ_{C-8} 101.97, δ_{C-3} 98.76 and $\delta_{C-6'}$ 116.90. The gHMBC spectrum had cross peaks between 4 methoxy groups which had signals at $\delta_{H-2'}$ 3.75(s), $\delta_{H-4'}$ 3.95(s), $\delta_{H-5'}$ 3.90(s) and $\delta_{H-6'}$ 3.80 and $\delta_{C-2'}$ 153.30, $\delta_{C-4'}$ 152.06,

$\delta_{C\text{-}5'}$ 143.10 and $\delta_{C\text{-}6}$ 152.92. The DEPT spectrum of the aglycone indicated the appearance of 7 methyl carbon peaks and 3 methylene carbon peaks which were in the sugar chemical shift region of compound 2.

Table 3.6 ^{1}H and ^{13}C NMR spectral data of compounds 1, 2 and 2a.

	Compound 1(DMSO-d_6)			Compound 2 (D_2O)			Compound 2a (CD_3OD)		
position	^{13}C	^{1}H	HMBC	^{13}C	^{1}H	HMBC	^{13}C	^{1}H	HMBC
2	154.42 CH	8.19 (s)		156.40 CH	8.03 (s)		156.24 CH	7.87 (s)	
3	121.03 C		2, 3′, 6′	121.05 C		2, 3′, 6′	120.52 C		2, 3′, 6′
4	174.66 C		2, 5, 8	177.47 C		2, 5, 8	176.48 C		2, 5, 8
5	104.63 CH	7.44 (s)	8	105.21 CH	7.17(s)	8	102.72 CH	6.50 (s)	8
6	147.31 C		5, 8, OMe	151.43 C		5, 8, OMe	152.92 C		5, 8, OMe
7	151.41 C		5, 8	152.29 C		5, 8	153.30 C		5, 8
8	103.72 CH	7.32 (s)	5	104.10 CH	7.35 (s)	5	101.97 CH	7.33 (s)	5
9	151.19 C		2, 5, 8	147.58 C		2, 5, 8	150.18 C		2, 5, 8
10	117.68 C		2, 5, 8	111.92 C		2, 5, 8	111.02 C		2, 5, 8
1′	112.92 C		2, 3′, 6′	115.60 C		2, 3′, 6′	113.95 C		2, 3′, 6′
2′	152.77 C		3′, 6′, OMe	152.16 C		3′, 6′, OMe	153.30 C		3′, 6′, OMe
3′	95.45 CH	6.87 (s)	6′	98.72 CH	6.71 (s)	6′	98.76 CH	6.73 (s)	6′
4′	147.79 C		3′, 6′, 7′	149.75 C		3′, 6′,OMe	152.06 C		3′, 6′,OMe
5′	140.26 C		3′, 6′, 7′	142.35 C		3′, 6′,OMe	143.10 CH		3′, 6′,OMe
6′	110.90 CH	6.83 (s)	3′	118.11 CH	6.86(s)	3′	116.90 CH	6.93 (s)	3′
7′	101.08 CH	6.0 (s)							
1″	99.90 CH	5.08(d) J=6.84		100.26 CH	5.21(d) J=7.00				
2″	72.98 CH	3.33		72.94 CH	3.65				
3″	76.67 CH	3.34		77.09 CH	3.86				
4″	69.80 CH	3.15		69.57 CH	3.59				
5″	75.65 CH	3.60		75.41 CH	3.76				
6″	67.85 CH_2	3.88(d) J=9.19		67.76 CH_2	4.02(d) J=9.15				
1‴	109.48 CH	4.80(d) J=2.94		109.38 CH	5.02(d) J=2.87				
2‴	75.85 CH	3.77		75.69 CH	3.69				
3‴	78.60 C			79.51 C					
4‴	73.30 CH_2	3.93		73.87 CH_2	3.92				
5‴	63.32 CH_2	3.37		64.02 CH_2	3.53				
6-OMe	55.75	3.93		6-OMe 56.67	3.73		6-OMe 55.98	3.80	
2′-OMe	56.63	3.66		2′-OMe 56.85	3.68		2′-OMe 56.35	3.75	
				4′-OMe 56.17	3.85		4′-OMe 54.68	3.95	
				5′-OMe 56.38	3.82		5′-OMe 55.65	3.90	

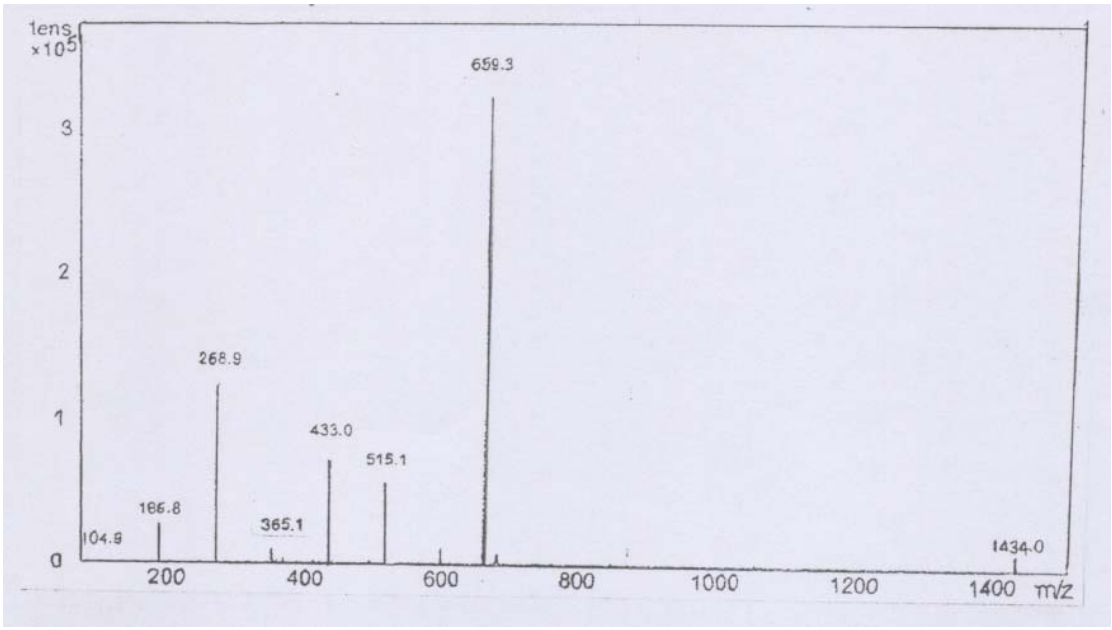


Figure 3.11 Mass spectrum of the S1 substrate in positive mode

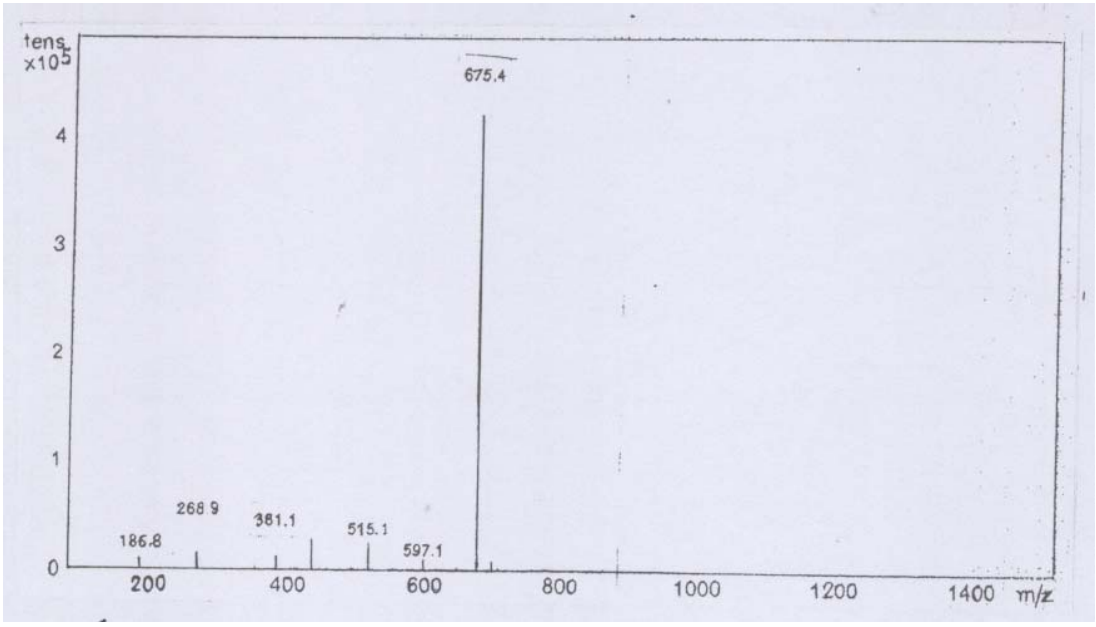


Figure 3.12 Mass spectrum of S2 substrate in positive mode

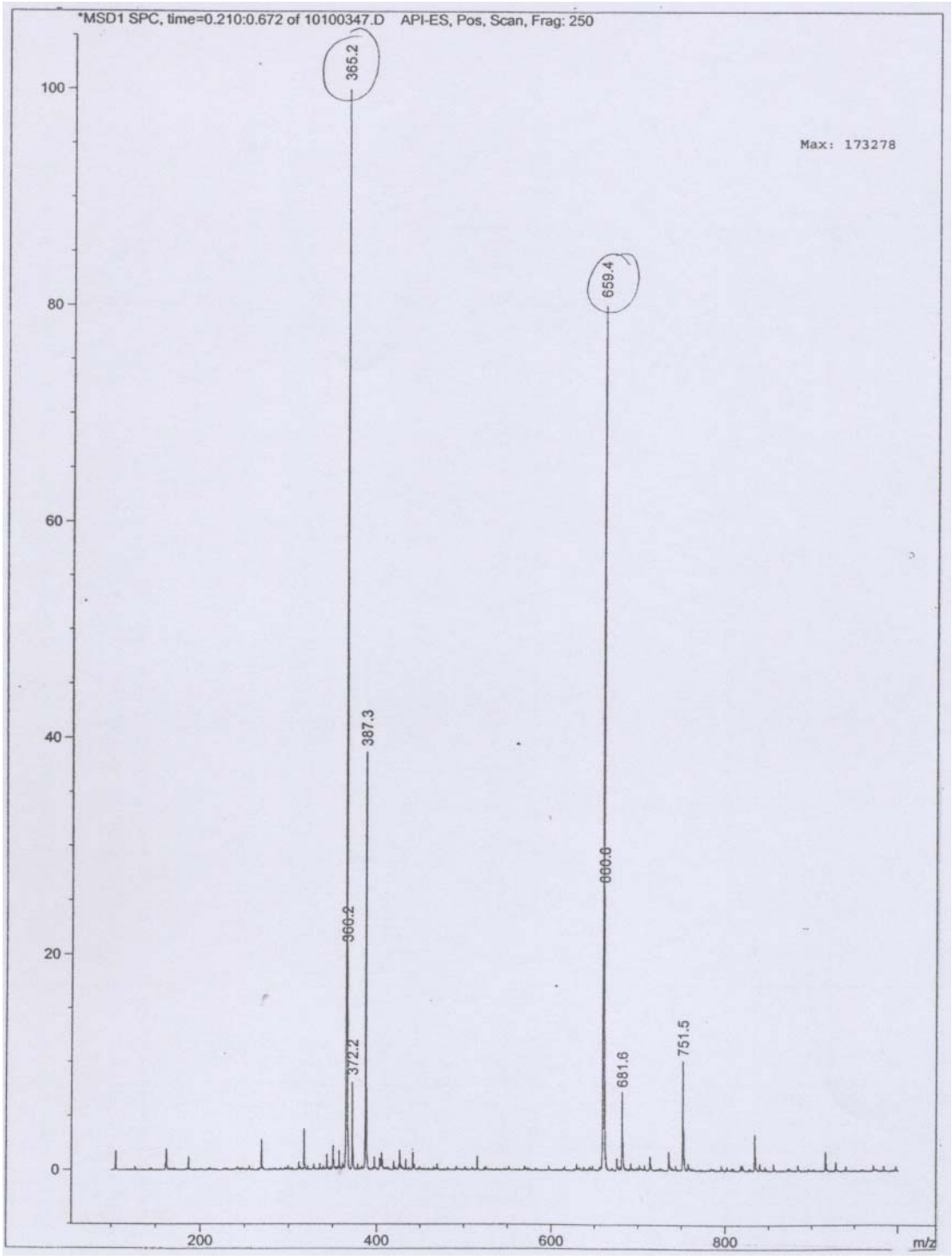


Figure 3.13 The mass spectrum of S1 substrate partially digested with *D. nigrescens* β-glycosidase.

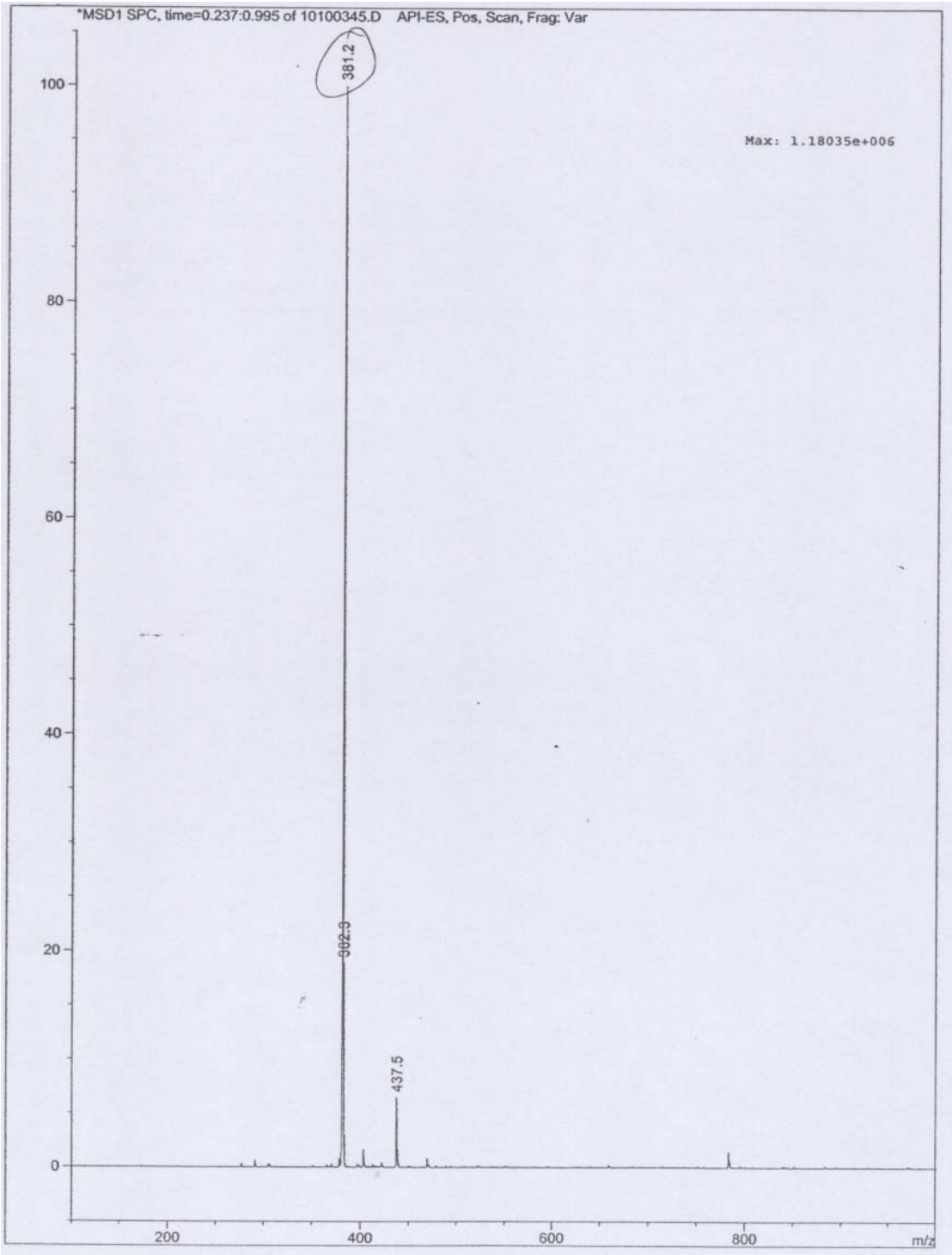


Figure 3.14 The mass spectrum of S2 substrate digested with *D. nigrescens* β-glycosidase.

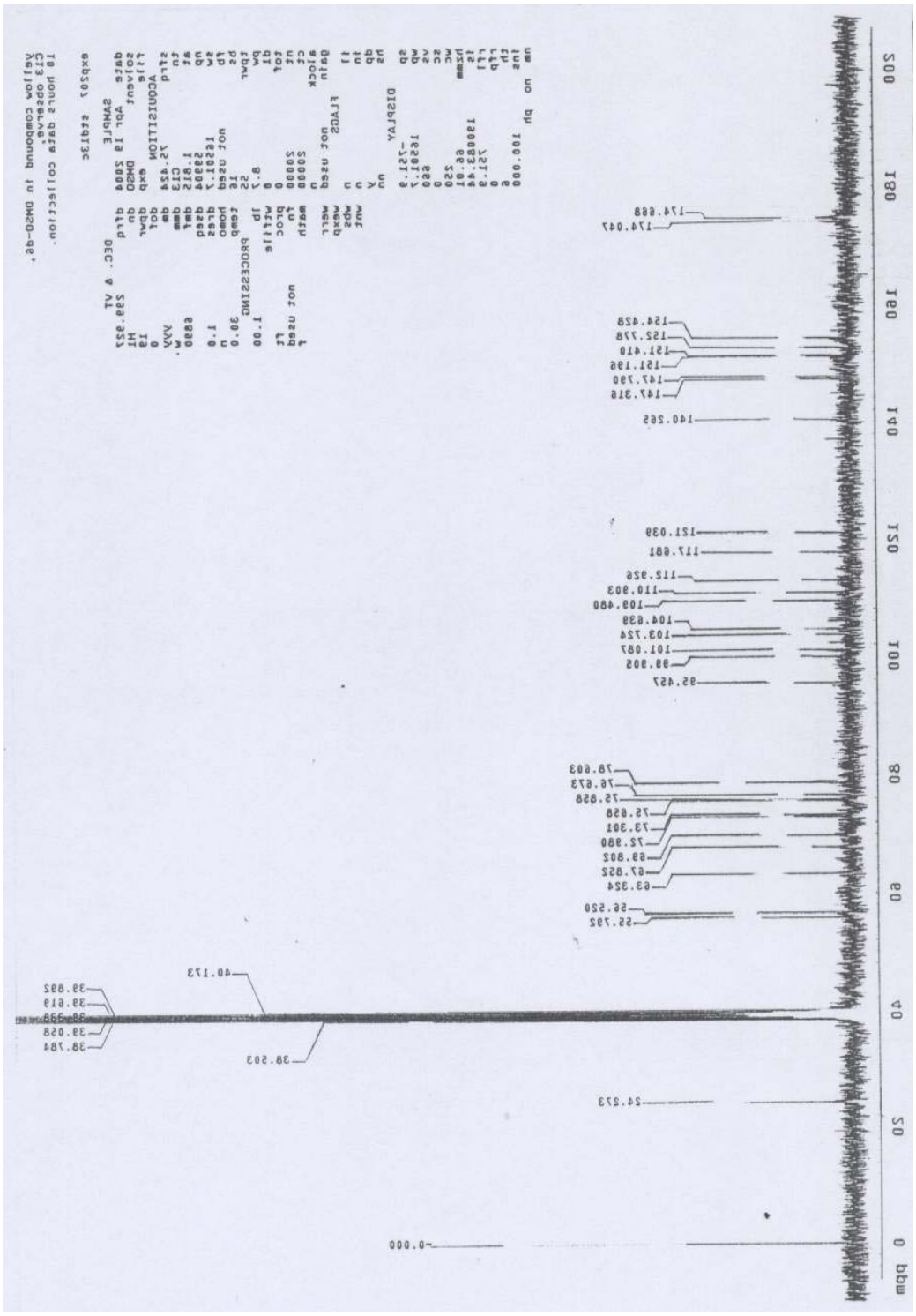


Figure 3.15 ^{13}C NMR Spectrum of the S1 substrate in DMSO$_{-d6}$

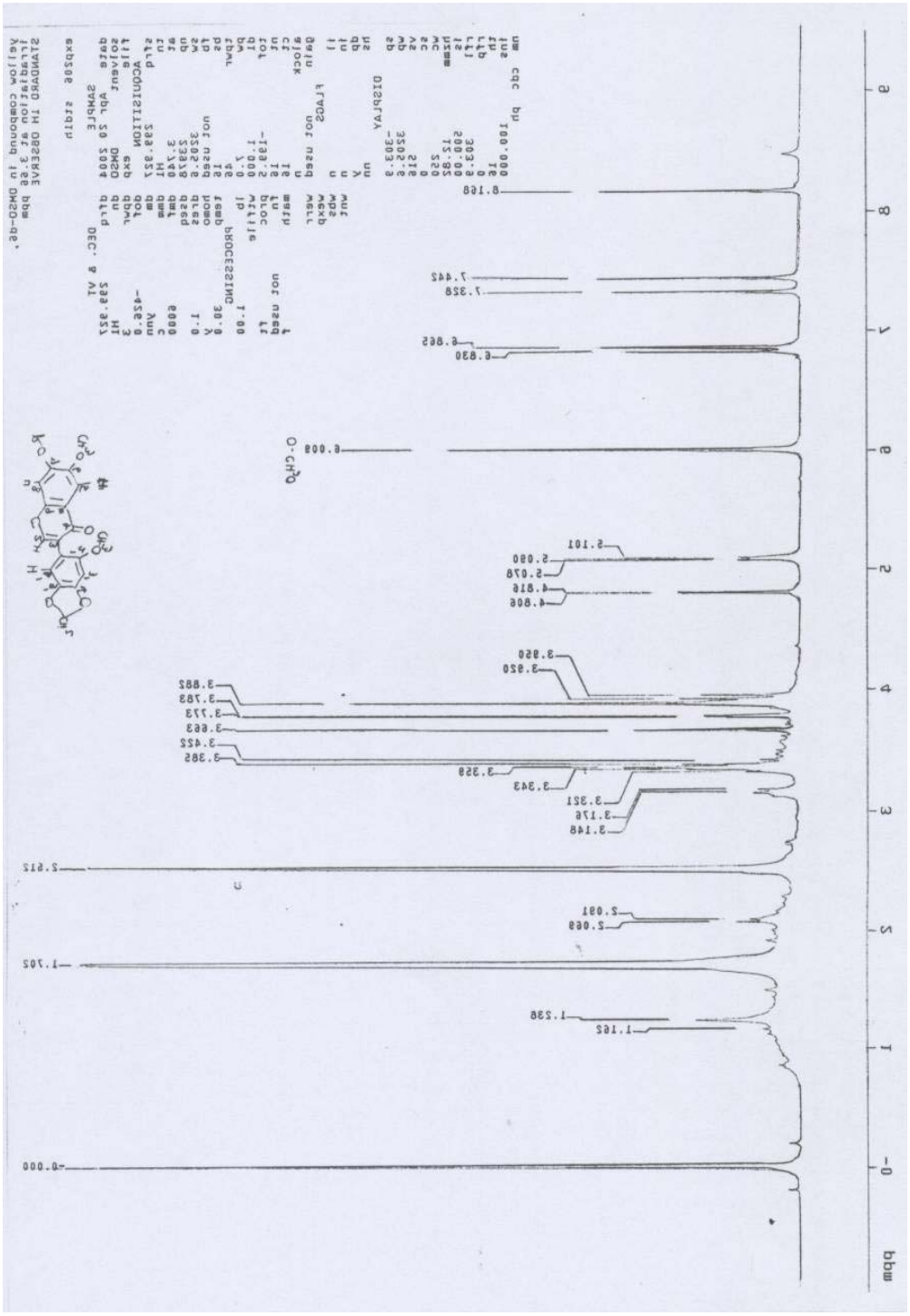


Figure 3.16 ^{1}H NMR Spectrum of the S1 substrate in DMSO$_{-d6}$

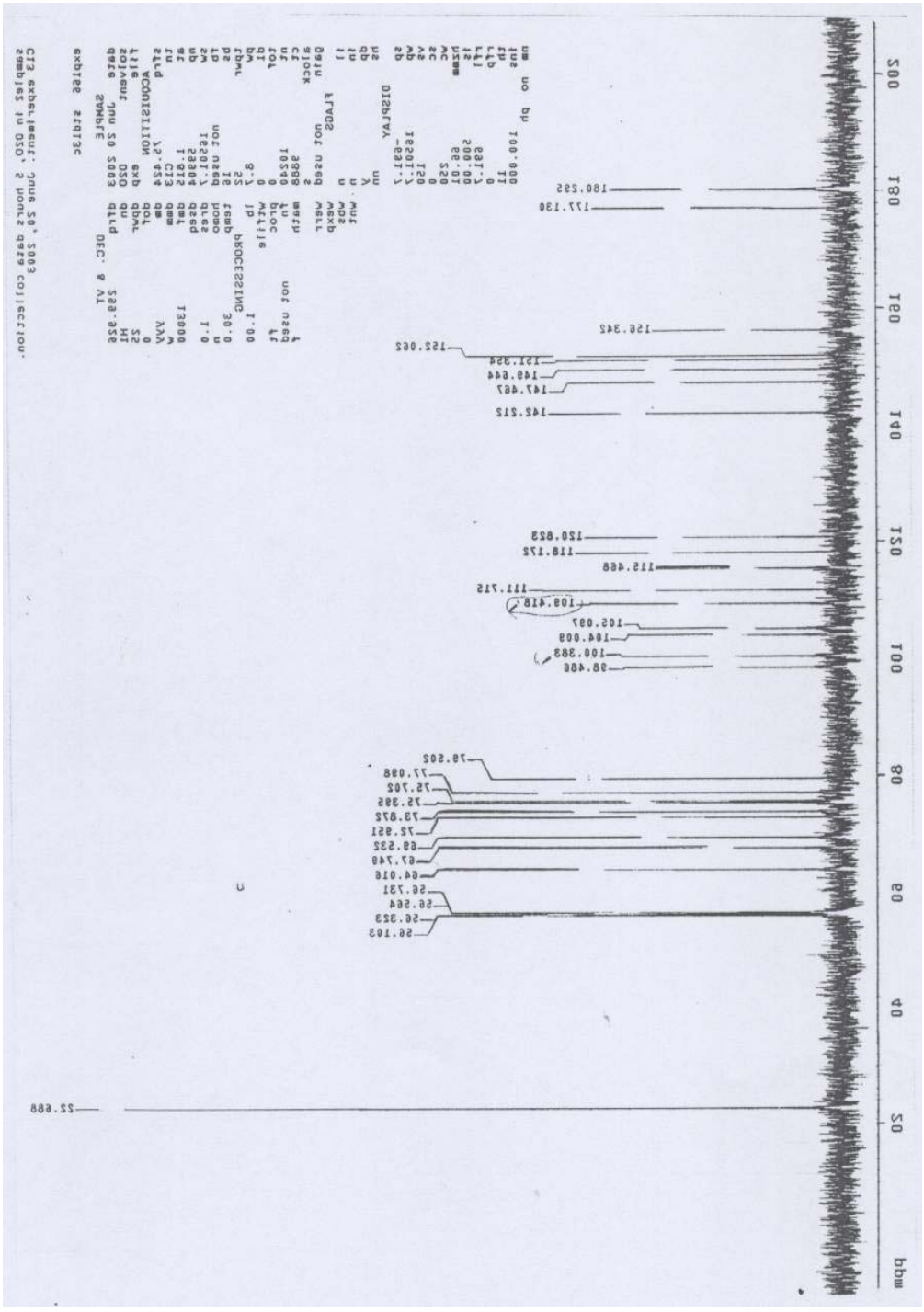


Figure 3.17 ^{13}C NMR Spectrum of the S2 substrate in D_2O

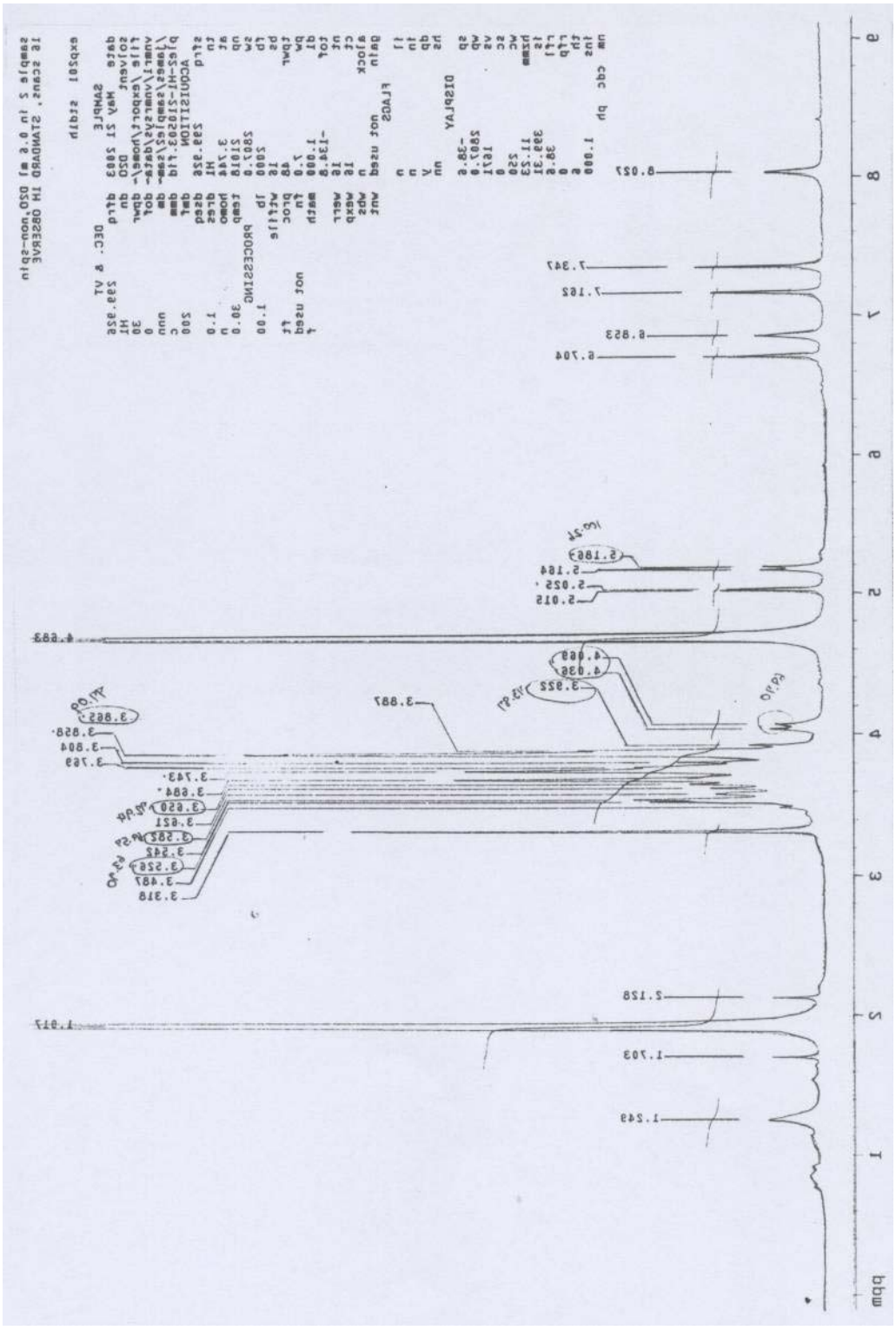


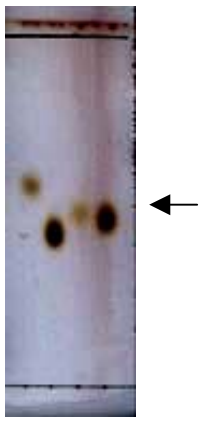
Figure 3.18 ^{1}H NMR Spectrum of the S2 substrate in D_2O

3.1.7 Sugar analysis

The sugar released from the S2 natural substrate after hydrolysis by native and recombinant *D. nigrescens* β-glucosidase was separated by TLC silica gel F_{254} with ($CHCl_3$/MeOH/H_2O; 14:6:1) and only one spot of sugar, which migrated similarly to disaccharide standards was seen (Figure 3.19). To confirm the result, the substrate was hydrolyzed with H_2SO_4 and the resulting sugars identified by TLC and GC-MS after trimethyl silylation. The acid hydrolysis reaction was run on TLC using galactose, fructose, glucose, apiose and ribose as standard sugars. The TLC analysis found 2 spots of sugar which matched with apiose and glucose (Figure 3.20). Also, the gas chromatogram (Figure 3.21) showed that the sugar peaks in the samples had the same retention times as those of the apiose and glucose standards. The mass fragmentation patterns of each peak matched those of the standard sugars, glucose (Figure 3.22) and apiose (Figure 3.23), which were treated in the same manner. Peaks a, b, c, d (Figure 3.21) exhibited *m*/*z* 103, 147, and 191 fragments seen in the mass spectra of the apiose standard. The fragmentation patterns of peaks e and f were the same as glucose which exhibited three main mass fragments of *m*/*z* 103, 147 and 204, and were found to match α- and β- D-glucopyranose when compared with the literature library. In addition, a *m*/*z* 73 fragment which observed in every peak does suggest that it is silylated. The g and h peaks might be some contamination from the experimental technique.

In addition, the disaccharide from the *D. nigrescens* β-glucosidase digested reaction was analyzed by GC-MS with the same derivative method as analysis of the sugar from acid hydrolysis. The result showed two peaks of TMS-

sugar sample which may indicate the β- and α- forms of the disaccharide (β(α)-apiosyl-β(1,6)-D-glucopyranoside) (Figure 3.24). Also, the disaccharide was confirmed using ^{13}C NMR and 19 peaks representing 22 ^{13}C signals were found in the chemical shift region of sugar, which indicated the 2 forms (α and β) of the 11 carbon disaccharide (Figure 3.25).



1 2 3 4

Figure 3.19 TLC chromatogram of hydrolysis of the S2 compound with *D. nigrescens* β-glucosidase. The reaction was extracted with n-butanol and the aqueous phase spotted onto an F_{254} silica plate and developed with $CHCl_3$/MeOH/H_2O (14:6:1) as developing solvent. Lane 1; glucose, lane 2; cellobiose, lane 3; sample and lane 4; laminaribiose.

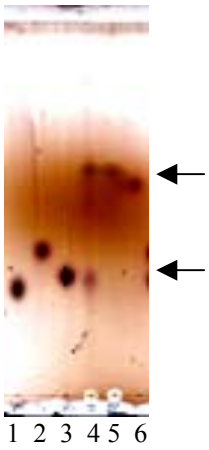


Figure 3.20 TLC chromatogram of acid hydrolysis of the S2 compound. After hydrolysis, the reaction was extracted with n-butanol, and the aqueous phase was separated on silica gel F_{254} TLC using $CHCl_3$/MeOH/H_2O (14:6:1) as developing solvent. Lane 1, galactose; lane 2, fructose; lane 3, glucose; lane 4, sample; lane 5, apiose and lane 6, ribose.

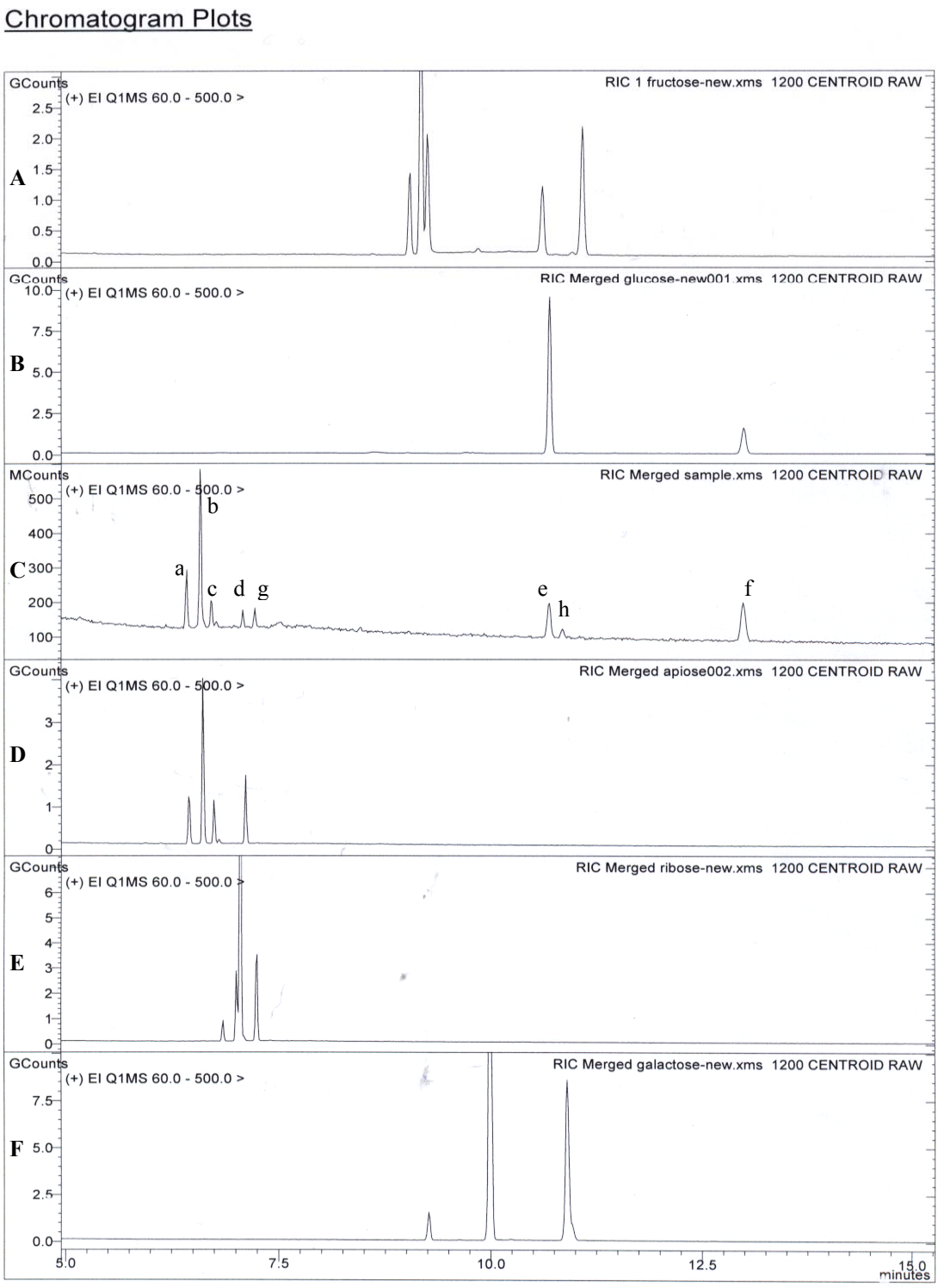


Figure 3.21 The gas chromatogram of derivatized sugar. The sample was compared with (A) fructose, (B) glucose, (C) sample, (D) apiose, (E) ribose and (F) galactose as standard sugars.

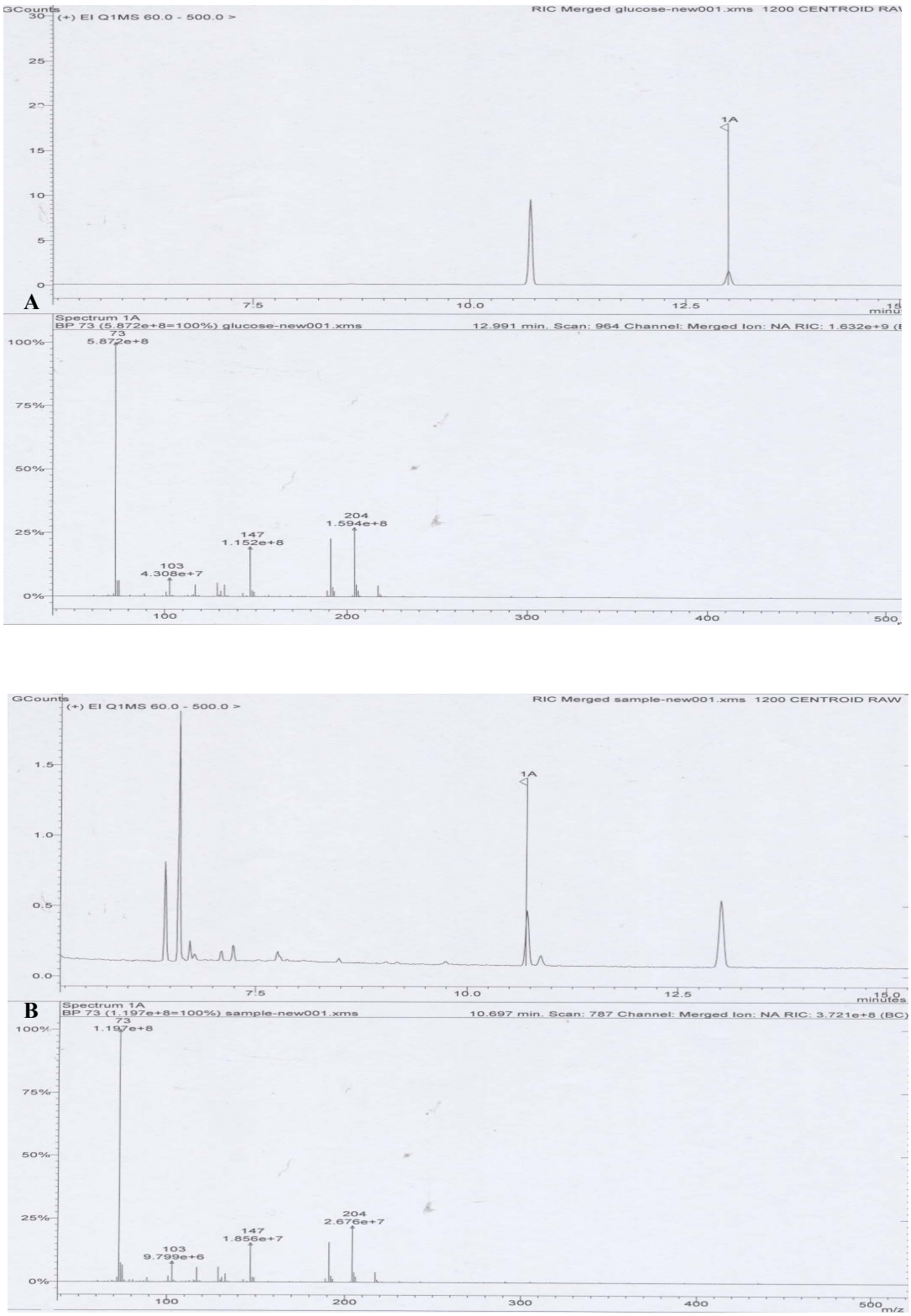


Figure 3.22 The Mass spectra of derivatized sample sugar and glucose standard; **A**, derivatived glucose standard; **B**, sample glucose peak.

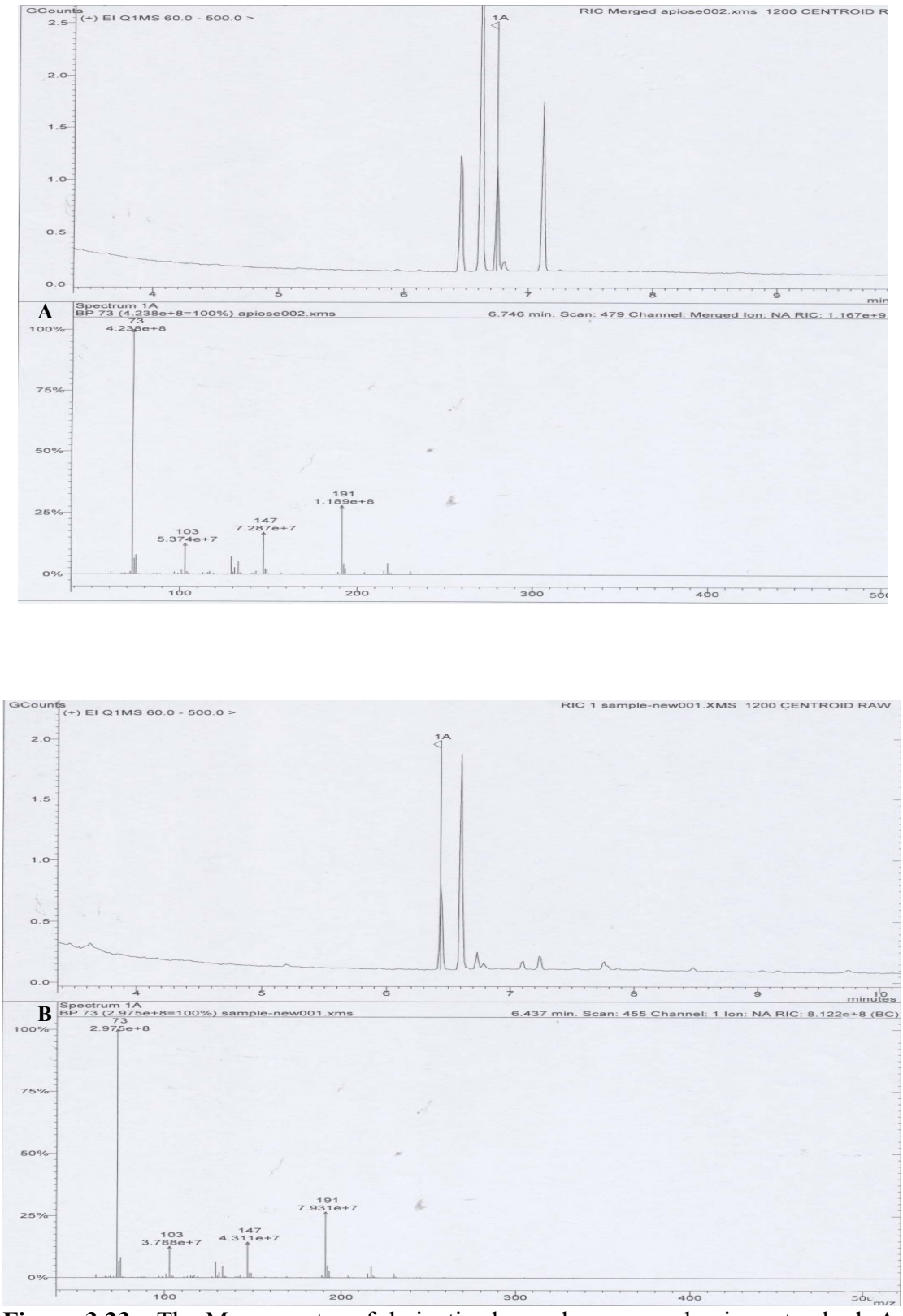


Figure 3.23 The Mass spectra of derivatized sample sugar and apiose standard; **A,** derivatived apiose standard**; B,** sample apiose peak.

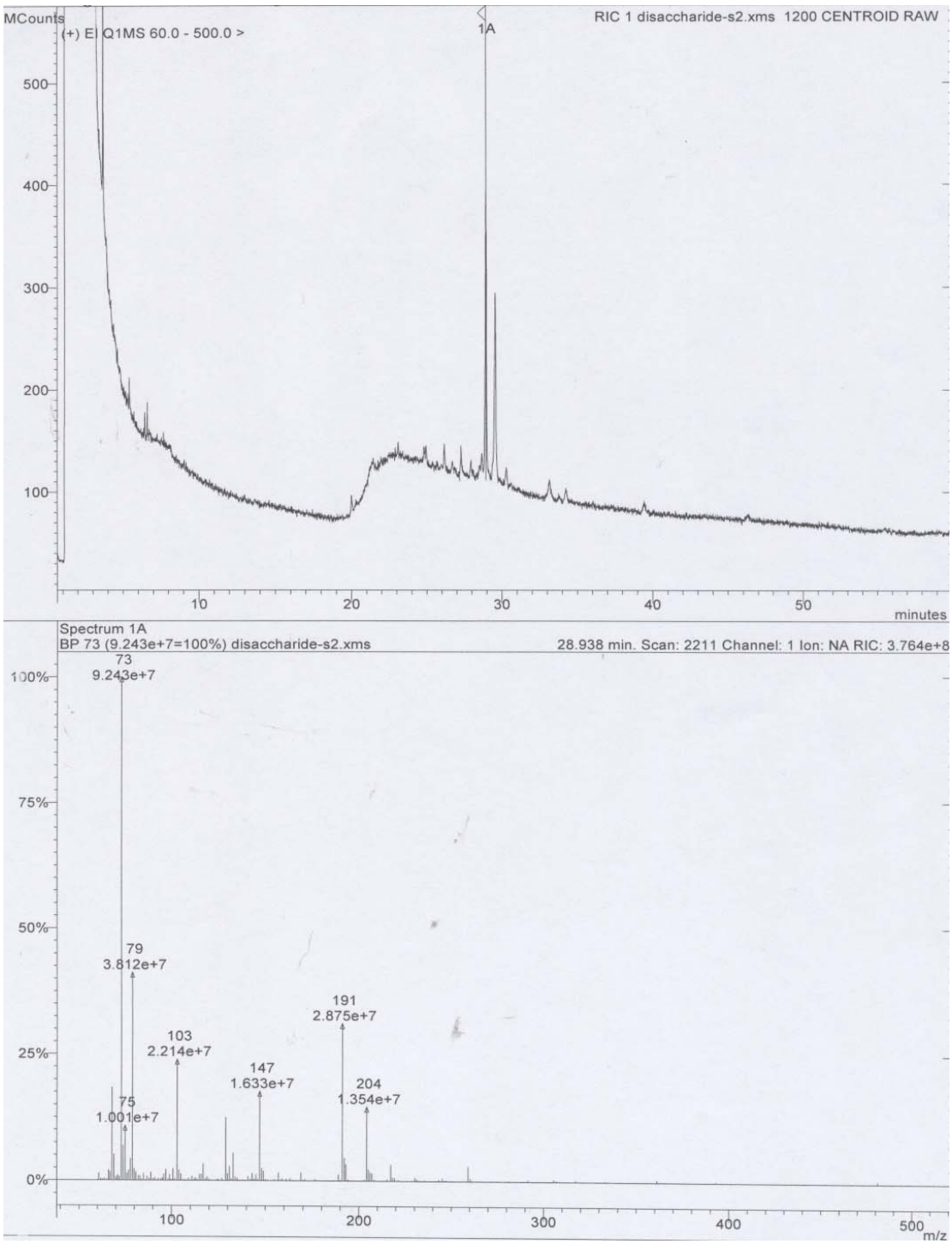


Figure 3.24 The GC chromatograph and mass spectrum of derivatized disaccharide sample**.**

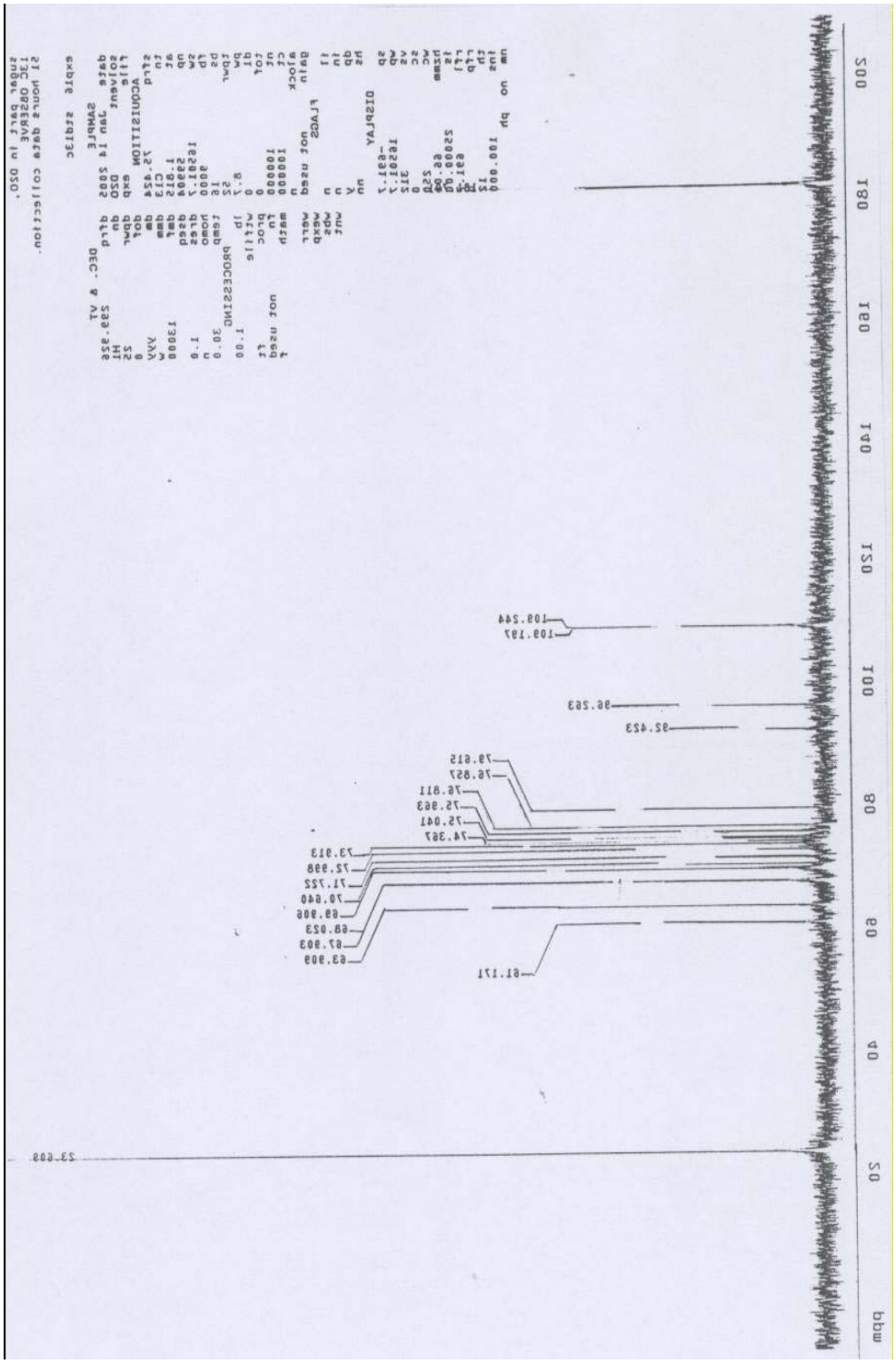


Figure 3.25 ^{13}C NMR spectrum of disaccharide from S2 substrate digestion in D_2O

3.1.8 Determination of kinetic parameters for hydrolysis of S1 and S2 glycoside natural substrates by *D. nigrescens* enzyme using high performance liquid chromatography (HPLC)

Because glucose could not be detected after hydrolysis of either of the natural substrate glycosides with native and recombinant *D. nigrescens* enzyme, the released aglycone part was quantitated by HPLC. The native β-glucosidase enzyme showed similar hydrolytic activity with S1 and S2 glucosides with estimated K_m values of 0.5 mM and 0.7 mM, respectively. The k_{cat} and k_{cat}/K_m values of native *D. nigrescens* enzyme on the natural substrates were 465 s^{-1} and 9.9 x 10^5 $M^{-1}s^{-1}$ for S1 and 334 s^{-1} and 4.4 x 10^5 $M^{-1}s^{-1}$ for S2, assuming a subunit molecular weight of 61,858 Da. The kinetic data are summarized in Table 3.7.

Table 3.7 The kinetic properties of purified native *D. nigrescens* β-glucosidase enzyme toward S1 and S2 glycosides. Assays were performed in 0.1 M sodium acetate buffer, pH 5, at 30°C. k_{cat} and k_{cat}/K_m of native enzyme were estimated assuming a subunit molecular weight of 61,858 Da.

Substrates	K_m (mM) Native	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
D. nigrescens S1 glycoside	0.5 ± 0.1	465 ± 59	9.9 x 10^5 ± 8.7 × 10^4
D. nigrescens S2 glycoside	0.7 ± 0.1	334 ± 19	4.4 x 10^5 ± 1.6 × 10^4

In addition to the *D. nigrescens* enzyme, hydrolysis of both natural substrates by Thai rosewood β-glucosidase was also tested. Thai rosewood β-glucosidase could not hydrolyze either natural substrate very well. Figure 3.26 demonstrates the hydrolyzed products separated after digestion with Thai rosewood β-glucosidase.

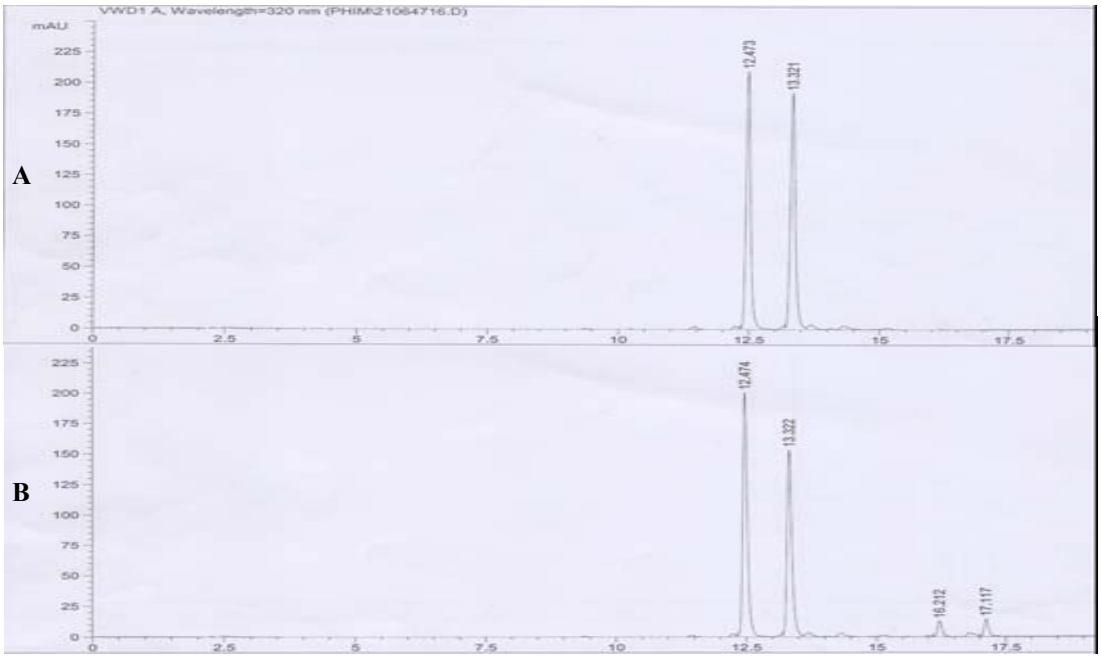


Figure 3.26 HPLC chromatogram of mixed S1 and S2 substrates with Thai rosewood β-glucosidase. **A**; the chromatogram of mixed 1 mM of S1 and S2 compound as a control, **B**; the chromatogram of mixed reaction of 1 mM of S1 and S2 compounds after incubating with 0.1 unit of Thai rosewood β-glucosidase in 0.1 M NaOAc, pH 5.0 for 10 min. The HPLC analysis was done as described in the methods with detection at 320 nm absorbance.

3.2 β-glucosidase cDNA cloning

3.2.1 *Dnbglu1* cDNA cloning

The *Dnbglu1* cDNA was amplified by RT-PCR using *D. nigrescens* cDNA as the template according to the strategy shown in Figure 3.27. To determine the 3′ end sequence of the *Dnbglu1* mRNA, the first strand cDNA was synthesized from the Q_T poly T primer using mRNA extracted from immature *D. nigrescens* seeds as template. The first strand cDNA was then used as a template in the 3′ RACE reaction using the For.6 sense primer, from the Thai rosewood β-glucosidase cDNA sequence, and Q_T as antisense primer to amplify a 3′ end cDNA fragment. Then, the PCR product was used as template and amplified with nested PCR using Q_O and For.6, then Q_i and For.6 primers. The PCR product of For.6 and Q_O was also reamplified using the For.6 and Rev.5 primers. The PCR products are shown in Figures 3.28 and 3.29. These PCR products were gel purified, cloned into pGEMT-easy vector and sequenced.

To amplify the 5′ end cDNA fragment, the poly A tail was added to the 3'-OH of a first strand cDNA which was synthesized from the Rev.8 primer, using Terminal Deoxynucleotidyl transferase (TdT). After that, the anchored cDNA was used as the template in a PCR reaction using Q_T and Rev.3 primers and then the second and third PCR amplifications were performed using the Q_T and Rev.4 primers and Q_O and Rev.4 primers, respectively (Figure 3.30). The PCR product of Rev.8 and For.2 was also amplified using single strand cDNA reverse-transcribed from the Rev.8 primer as template (Figure 3.31). The partial DNA thus amplified were cloned and sequenced using the primers from Thai Rosewood (Figure 3.32).

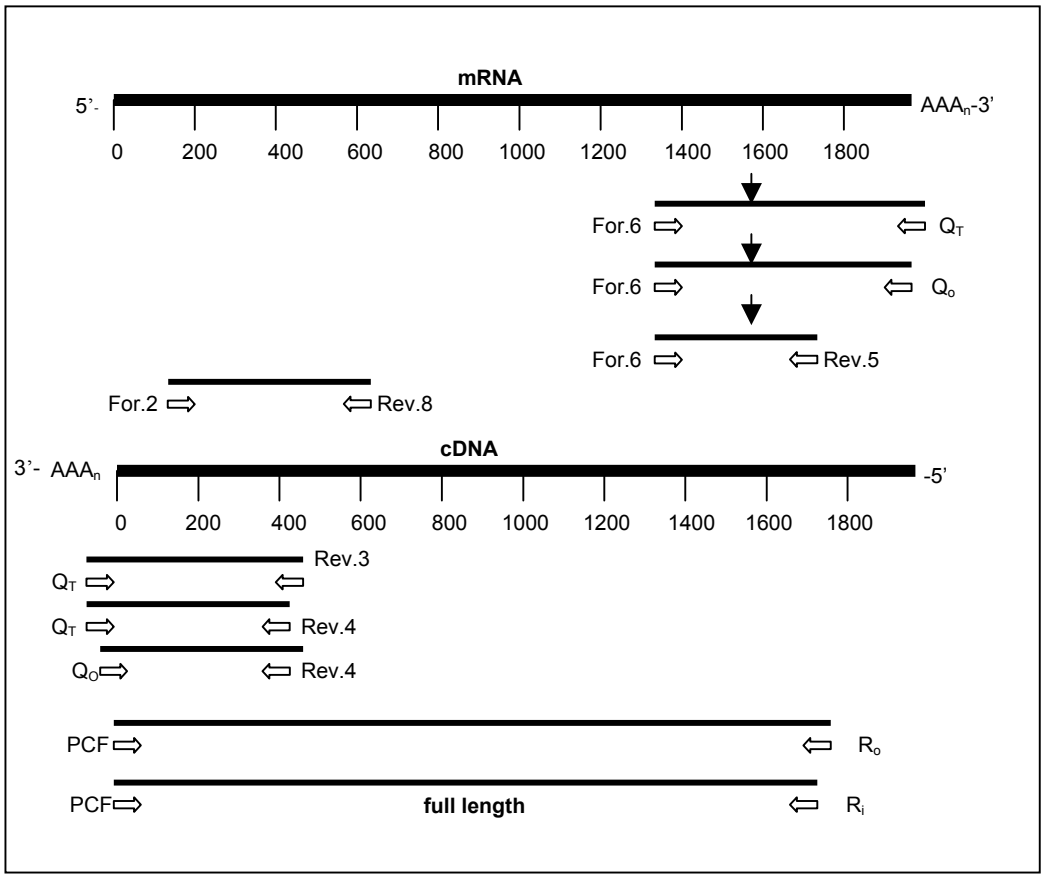


Figure 3.27 Cloning strategies of *Dnbglu1* cDNA. The 3′ end of *Dnbglu1* was amplified using For.6 with Q_T, Q_O and Q_i primers in order. The 5′ end of *Dnbglu1* was amplified using Q_T and Q_O as sense primers with Rev.3 and Rev.4 as antisense primers, respectively to amplify a polyadenylated first-strand cDNA template. The full-length cDNA was generated using the specific primers PCF, R_O and R_i.

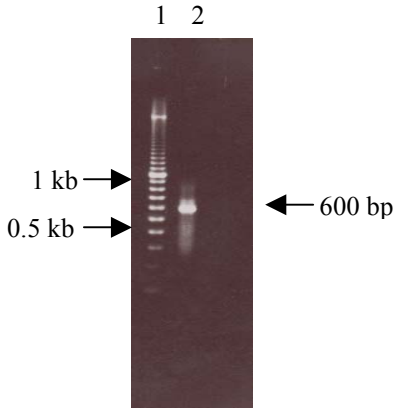


Figure 3.28 1% agarose gel electrophoresis of the 3′ RACE PCR product of the *Dnbglu1* cDNA amplified with For.6 and Q_i primers. The PCR product was amplified using the PCR product of the For.6 and Q_O primers as template. Lane 1, 100 bp DNA marker (Promega); lane 2, 3′ RACE PCR product.

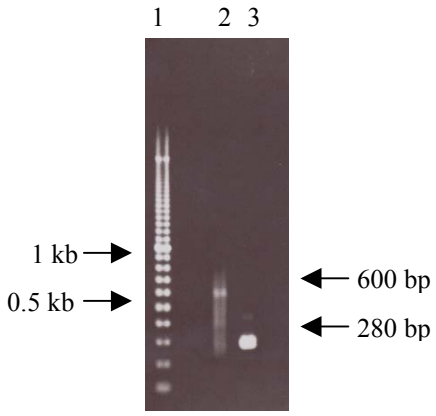


Figure 3.29 1% agarose gel electrophoresis of the *Dnbglu1* cDNA fragment PCR amplified with For.6 and Rev.5 primers from the 3′ RACE PCR product. The PCR product was amplified using the PCR product of the For.6 and Q_O primers as template. Lane 1, 100 bp DNA marker (Promega); lane 2, 3′ RACE PCR product amplified with For.6 and Q_O primers using For.6 and Q_T PCR product as template; lane.3, PCR product amplified with For.6 and Rev.5 primers using the PCR product of the For.6 and Q_o primers as template.

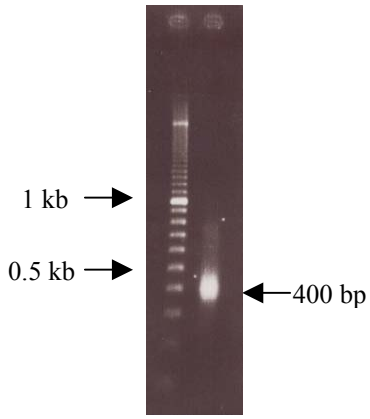


Figure 3.30 1% agarose gel electrophoresis of the 5′ RACE PCR product of *Dnbglu1* cDNA fragment amplified with Q_O and Rev.4 primers. The PCR product was amplified using a poly A tailed first-strand cDNA as template. Lane 1, 100 bp DNA marker (Promega); lane 2, 5′ RACE PCR product amplified with Q_O and Rev.4 primers using the PCR product of For Q_T and Rev.3 as template.

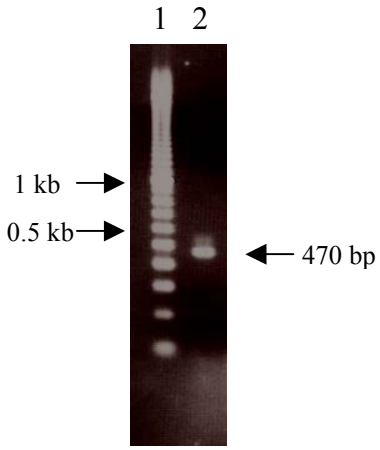


Figure 3.31 1% agarose gel electrophoresis of the *Dnbglu1* cDNA fragment product amplified with For.2 and Rev.8 primers. The PCR product was amplified using a first-strand cDNA with a poly A tail added to its 3′ OH as template. Lane 1, 100 bp DNA marker (Promega); lane 2, PCR product of For.2 and Rev.8 primers.

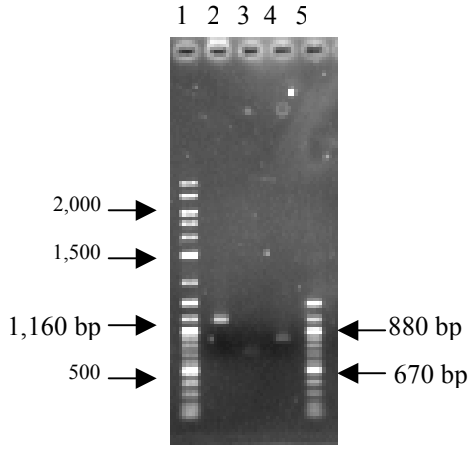


Figure 3.32 0.8% agarose gel electrophoresis of the *Dnbglu1* cDNA fragment PCR products amplified with Thai rosewood primers. Lane 1, 2-log DNA marker (New England Biolabs); lane 2, PCR product of For.3 and Rev.6 primers (1,160 bp); lane 3, PCR product of For.3 and Rev.7 primers (670 bp); lane 4, PCR product of For.4 and Rev.6 primers (880 bp); lane 5, 100 bp DNA marker.

After the sequences at the 3′ and 5′ ends of the β-glucosidase mRNA were known, they were used to design the primers to construct a full-coding-region β-glucosidase cDNA using RT-PCR. The primers, PCF from 5′ end sequence and R_O then R_i from 3′ end sequence were used in Semi-nested PCR amplification with a first-strand cDNA reverse transcribed from a poly T primer as template. The full-length coding region (CDS) cDNA PCR product was produced (Figure 3.33), and its sequence was determined and assembled with those of the 3′ and 5′ RACE products

(Figure 3.34). The full-length cDNA sequence of *Dnbglu1* consisted of 1,886 nucleotides, which includes a 1,638 nucleotides open reading frame (ORF) encoding a 547 amino acid precursor protein.

The Signal P program (http://www.cbs.dtu.dk/services/SignalP/, Bendtsen *et al*., 2004) predicted the protein to contain a 23 amino acid signal sequence and a 524 amino acid mature protein. The predicted N-terminus corresponded to the mature N-terminus of Thai Rosewood dalcochinase (Srisomsap *et al*., 1996). The mature protein had a predicted MW of 61858 Da and predicted pI of 5.60.

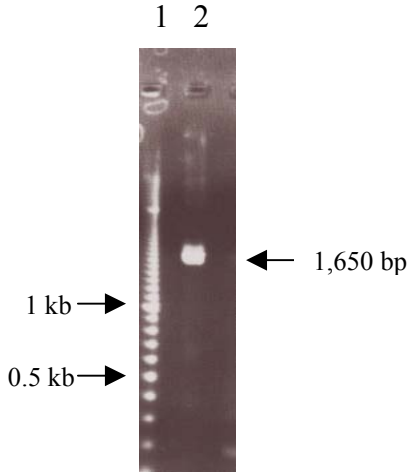


Figure 3.33 1% agarose gel electrophoresis of the full-length *Dnbglu1* CDS cDNA PCR product amplified with PCF and R_i primers. The PCR product was amplified using first-strand cDNA as template. Lane 1, 100 bp DNA marker (Promega); lane 2, PCR product of full-length cDNA.

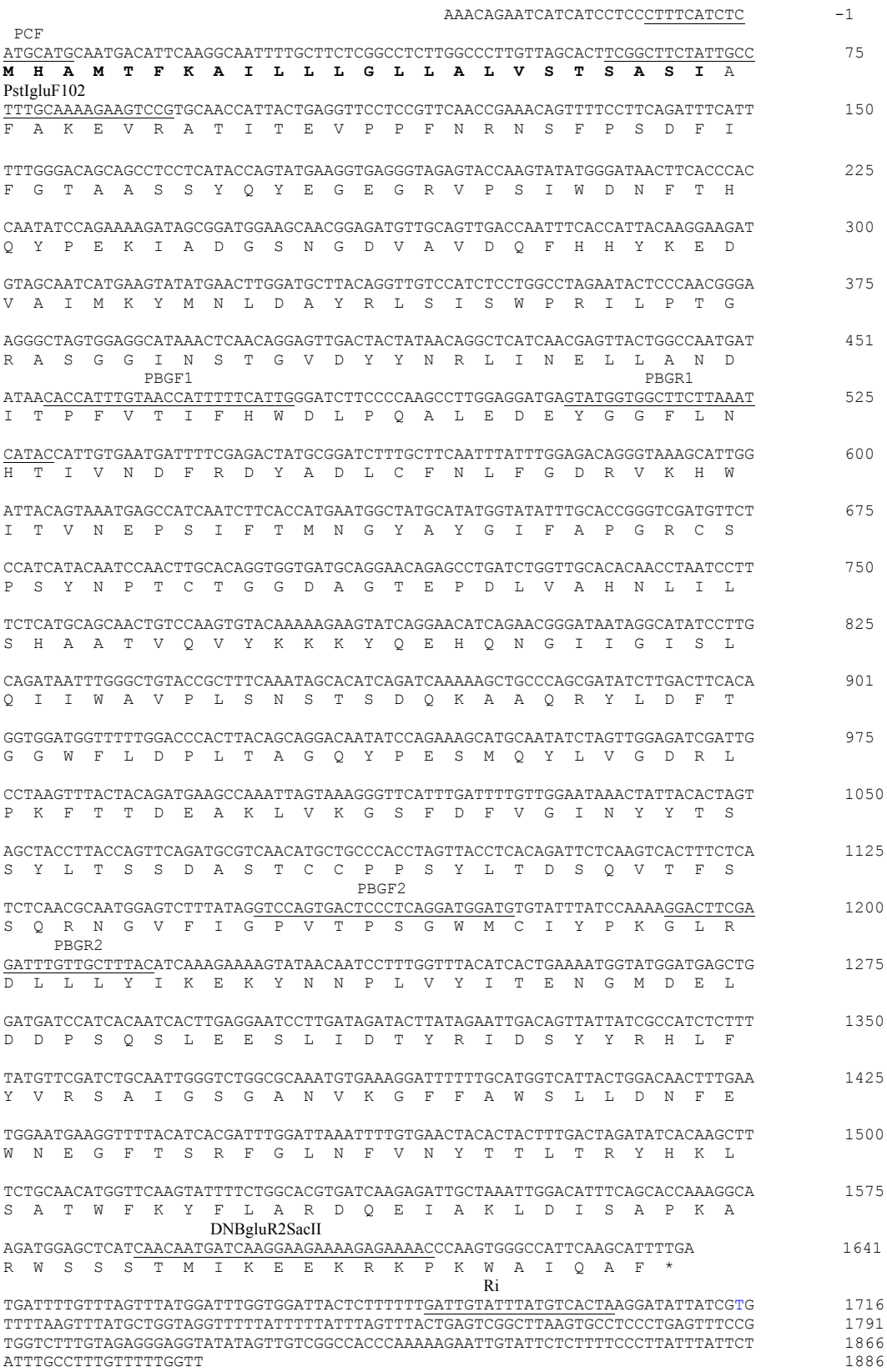


Figure 3.34 The full-length cDNA sequence and deduced amino acid sequence of

Dnbglu1. The bold letters indicate the prepeptide amino acid sequence and the sequence positions of the primers which were used for *Dnbglu1* amplification are underlined.

3.2.2 *Dnbglu2* cDNA cloning

Due to the presence of a stop codon in some *Dnbglu1* cDNA clones, an attempt was made to reamplify the *Dnbglu1* cDNA using the PCF and R_i primers (derived from *Dnbglu1* sequence) by RT-PCR from RNA extracted from *D. nigrescens* seeds, which was used to amplify the first *Dnbglu1* full-length cDNA. A short internal fragment of cDNA was produced. The synthesized PCR fragment was subsequently cloned, and sequenced. Then the sequence was aligned with the *Dnbglu1* sequence. The sequence alignment showed that there were some regions different from *Dnbglu1*. So, new primers (PTDnF1**,** PTDnR1 and PTDnR2**)** were designed from the sequence of the new cDNA fragment.

The second *D. nigrescens* β-glucosidase isozyme cDNA (*Dnbglu2*) was cloned by RT-PCR with the strategy summarized in Figure 3.35. First-strand cDNA synthesis and amplifications of *Dnbglu2* specific cDNA fragments were carried out by 3′ and 5′ RACE. The first strand cDNA was synthesized using mRNA extracted from 2-3 days *D. nigrescens* seedlings as template. The first amplification of the 3′ region of the *Dnbglu2* cDNA was achieved by PCR using the $PTDnF_1$ primer as the sense primer, and the Q_T primer as the antisense primer. Amplification of the 5′ region of the *Dnbglu2* cDNA was carried out with the 5′ RACE system. Single strand cDNA was synthesized using the PTDnR1 gene specific primer to prime reverse

transcription and then adding a poly A tail at the 5′ end of the first strand cDNA. Then, the first PCR reaction was done using the Q_T and $PTDnR_1$ primers (Figure 3.36). The product was reamplified using the $PTDnR_2$ and Q_O primers. RACE products were then gel purified, cloned and sequenced, as described in Method 2.4. The sequences of the 3′ and 5′ RACE products were used to design the Dn2_3′UTRr1 3′ terminus and Dn2_5′UTRf1 5′ terminus primers, which were used to amplify the full length coding region of the cDNA. A specific PCR product of 1,719 bp was produced (Figure 3.37) and the sequence was determined. The full-length cDNA sequence of *Dnbglu2*, including regions of the 5′ and 3′ RACE products out side the PCR primers, consisted of 1,964 nucleotides, which included a 1,593 nucleotide open reading frame (ORF) encoding a 531 amino acid precursor protein (Figure 3.38). The protein consisted of a 27 amino acid sequence prepeptide and 504 amino acid mature protein, as predicted by the Signal P program. The mature protein had a predicted MW of 60509 Da and predicted pI of 6.11.

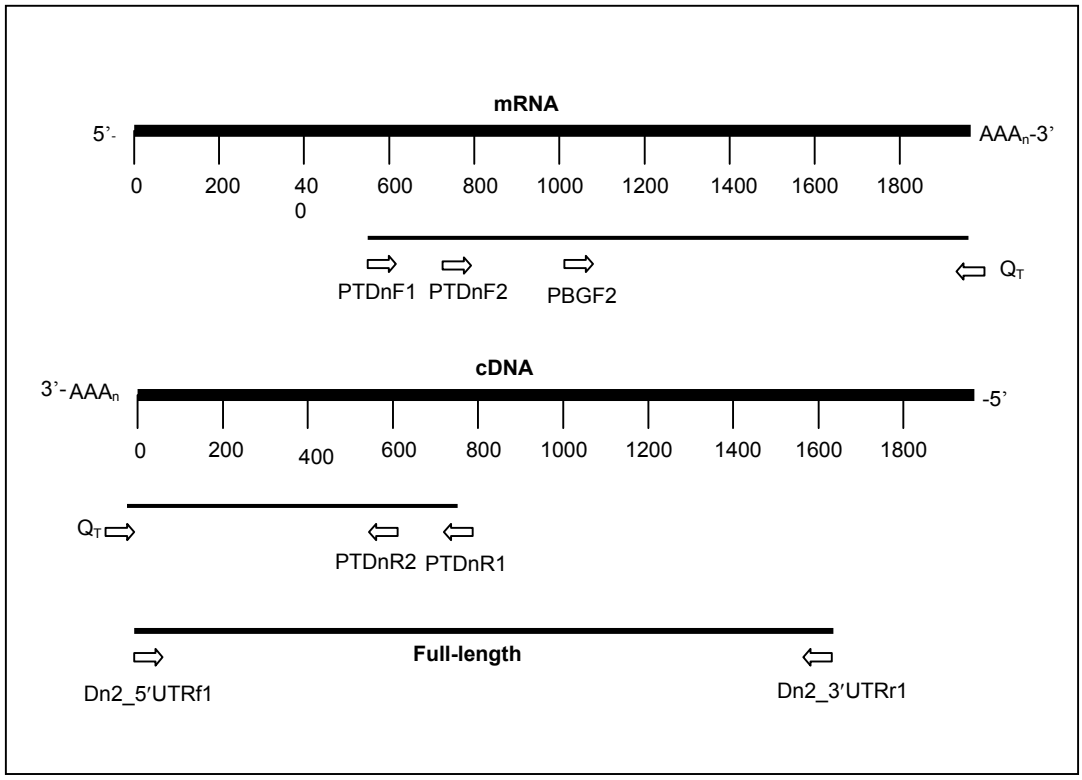


Figure 3.35 Cloning strategies for amplification of *Dnbglu2* from *D.nigrescens* Kurz. The cDNA fragment encoding the 3′ end of *Dnbglu2* was generated by 3′ RACE with primers PTDnF1 and Q_T, and the cDNA fragment encoding the 5′ end of *Dnbglu2* was generated by 5′ RACE with primers Q_T and PTDnR1. Then, the full-length cDNA was amplified using the specific Dn2_5′UTRf1 (5′ terminus) and Dn2_3′ UTRr1 (3′ terminus) primers.

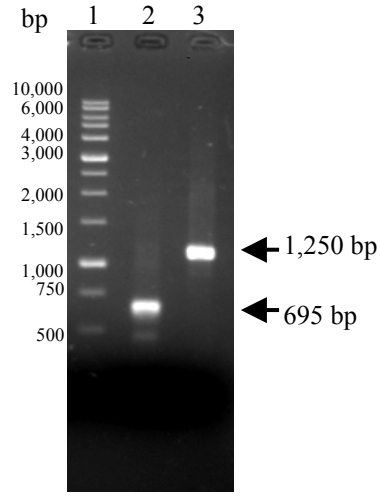


Figure 3.36 1% agarose gel electrophoresis of *Dnbglu2* 5′ RACE PCR product (amplified with the PTDnR$_1$ and Q$_T$ primers) and *Dnbglu2* 3′ RACE PCR product (amplified with the PTDnF$_1$ and Q$_T$ primers). Lane 1, 1 kb DNA marker (SibEnzyme); lane 2, 5′ RACE PCR product; lane 3, 3′ RACE PCR product.

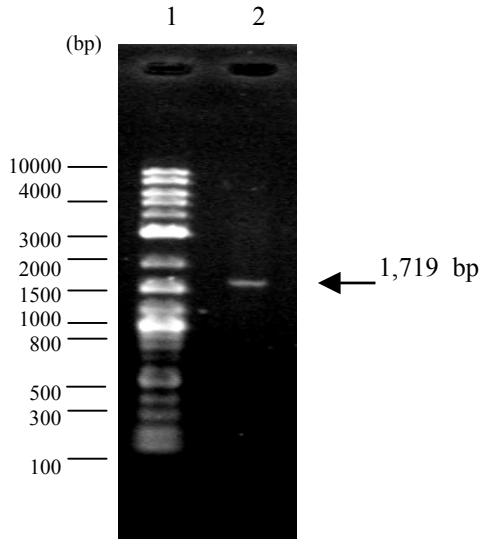


Figure 3.37 1% agarose gel electrophoresis of the PCR product of the full-length CDS *Dnbglu2* cDNA which was amplified with the Dn2_5′ UTRf1 (5′ terminus) and Dn2_3′ UTRr1 (3′ terminus) primers. Lane 1, 2- log ladder DNA marker (New England BioLabs); lane 2, PCR product of full-length *Dnbglu2* cDNA.

```
                                                  CTGCAGAATTCGCCCTTTCCTTTCTTTCATCTC        -1
5′ terminus
ATGATTGCAATGACATTCAAGGTGATTTTGCTTCTTGGCCTCTTGGCCCTTATTAGCACTTCGACTTCTATTGCC        75
M  I  A  M  T  F  K  V  I  L  L  L  G  L  L  A  L  I  S  T  S  T  S  I  A
                                 Dn2NTERMPstIf
TTTCCAAAAGAAGTCCGTGCAACCATTACTGAGGTTCCTCCATTCAATCGAAGTTGTTTTCCTTCAGATTTCATC       150
F  P  K  E  V  R  A  T  I  T  E  V  P  P  F  N  R  S  C  F  P  S  D  F  I

TTTGGGGCATCATCCTCCGCGTACCAGTATGAAGGTGAGGGTAGAGTACCAAGTATATGGGATAACTTCACCCAC       225
F  G  A  S  S  S  A  Y  Q  Y  E  G  E  G  R  V  P  S  I  W  D  N  F  T  H

CAATATCCAGAAAAGATAGCGGATGGAAGCAATGGAGATGTTACAATTGACCAATTTCACCGCTACAAGGAAGAT       300
Q  Y  P  E  K  I  A  D  G  S  N  G  D  V  T  I  D  Q  F  H  R  Y  K  E  D

GTTGCAATCATGAAGTATATGAACTTAGATGCTTATAGATTGTCCATCTCCTGGCCTAGAATACTCCCAACGGGA       375
V  A  I  M  K  Y  M  N  L  D  A  Y  R  L  S  I  S  W  P  R  I  L  P  T  G

AGGGCTAGTGGAGGCATAAACTCAACAGGAGTTGACTACTACAACAGGCTCATCAATGAGACACTACACAATGGC       450
R  A  S  G  G  I  N  S  T  G  V  D  Y  Y  N  R  L  I  N  E  T  L  H  N  G
                                                                PTDnF1
ATAACACCATATGTTACCATTTTTCATTGGGATCTTCCCCAAGCCTTGGAGGATGAGTATGGTGGCTTCTTAGAT       525
I  T  P  Y  V  T  I  F  H  W  D  L  P  Q  A  L  E  D  E  Y  G  G  F  L  D
            PTDnR2
CGTAGGGTTGTAAATGATTTTCGAGACTATGCGGATCTTTGCTTCAAATTTTTTGGAGATCGGGTAAAGCATTGG       600
R  R  V  V  N  D  F  R  D  Y  A  D  L  C  F  K  F  F  G  D  R  V  K  H  W
                        PTDnR1
ATTACAATAAATGAGCCACAAGTCTTCACCACAAATGGCTATACATATGGCATGTTTGCACCAGGTCGATGTTCT       675
I  T  T  I  N  E  P  Q  V  F  T  T  N  G  Y  T  Y  G  M  F  A  P  G  R  C  S

CCATCATATGATCCAACTTGCACAGGTGGCGATGCAGGAACAGAGCCTTATAAGGTTGCACATAACCTGCTCCTT       750
P  S  Y  D  P  T  C  T  G  G  D  A  G  T  E  P  Y  K  V  A  H  N  L  L  L

TCTCATGCAGCAACTGTCCAAGTGTACAAAGAGAAGTATCAGAAAGATCAAAATGGTAAAATAGGCATAACCTTG       825
S  H  A  A  T  V  Q  V  Y  K  E  K  Y  Q  K  D  Q  N  G  K  I  G  I  T  L

GATCAGCGCTGGGTTATACCACTTTCAAATAGCACATCAGATAAAAAAAGCTGCCCAGCGATATCTTGACTTCACA       900
D  Q  R  W  V  I  P  L  S  N  S  T  S  D  K  K  A  A  Q  R  Y  L  D  F  T

TTTGGATGGTTTATGGACCCACTTACAGTAGGAAGGTATCCAGATAGCATGCAATATCTAGTTGGAAATCGATTG       975
F  G  W  F  M  D  P  L  T  V  G  R  Y  P  D  S  M  Q  Y  L  V  G  N  R  L

CCTAAGTTTACTACATATGAAGCCAAATTAGTCAAGGGTTCATTTGATTTTATTGGAATAAACTATTACACCAGT      1050
P  K  F  T  T  Y  E  A  K  L  V  K  G  S  F  D  F  I  G  I  N  Y  Y  T  S

AACTACGCTACCAAATCAGATGCGTCGACATGCTGCCCACCTAGCTACCTCACAGATCCTCAAGTCACTCTCTCA      1125
N  Y  A  T  K  S  D  A  S  T  C  C  P  P  S  Y  L  T  D  P  Q  V  T  L  S

AGTCAACGTAATGGGGTCTTTATAGGTCCAATGACTCCCTCAGGATGGATATGTATTTATCCAAAAGGACTTCGA      1200
S  Q  R  N  G  V  F  I  G  P  M  T  P  S  G  W  I  C  I  Y  P  K  G  L  R

GATTTGTTGCTTTACATCAAGGAAAATTATAACAATCCTTTGGTTTACATTACTGAAAATGGTATGGATGAGACA      1275
D  L  L  L  Y  I  K  E  N  Y  N  N  P  L  V  Y  I  T  E  N  G  M  D  E  T

AATGATCCATCACTATCACTTGAGGAATCCTTGATGGATACTTATAGAATTGATAGTTATTATCGCCATCTCTTT      1350
N  D  P  S  L  S  L  E  E  S  L  M  D  T  Y  R  I  D  S  Y  Y  R  H  L  F

TATGTTCTATCTGCAATCAAGTCTGGCGCAAATGTGAAAGGATTTTTTGCATGGACATTAATGGACGACTTCGAA      1425
Y  V  L  S  A  I  K  S  G  A  N  V  K  G  F  F  A  W  T  L  M  D  D  F  E

TGGTCTGGAGGTTTTACATCACGATTTGGATTAAATTTTGTGGACTATAATACCTTGAATAGATATCCTAAGCTC      1500
W  S  G  G  F  T  S  R  F  G  L  N  F  V  D  Y  N  T  L  N  R  Y  P  K  L

TCTGCAAAATGGTTCAAGTATTTTCTGACACGTGATCAAGAGAGTGCTAAACTGGACATTTCAACACCAAAGGCA      1575
S  A  K  W  F  K  Y  F  L  T  R  D  Q  E  S  A  K  L  D  I  S  T  P  K  A
                                                    Dn2CTERMXbaIr
AGTGCAGCTTATCAACGATGATCAAGGAAGAAAAGACAAAACCCAAGTGGGCCATTCAAGCATTTTGAATTGTGTT      1650
S  A  A  Y  Q  R  *
                     3′ terminus
TAGTTTGTTTAGTTTGTGGCTTTGGTACATTTCTTTGTTTCGAATGTATTTGTGTCAATAAGGATATTATTGTGT      1725
TCTAAGTTTAAGGTGTATACCCTTATGGGTTTGCTTAGTTCATATTTATGTTATTAAGTTTTTGTTTTTATTTAG      1800
GTTAGTGAGTCGGCTTAAGTGCCTCCCTGAGTTTCGCGGTCTTTGTAGGGGGATGTATATAATTGTCGGCCGCCC      1875
AAAAAGATTTGTATTCTCTTTTCTTTTATTTATTCTACTTGCTTTTGTTTTCGTTT                          1931
```

Figure 3.38 The full-length cDNA sequence and deduced amino acid sequence of

Dnbglu2. The bold letters indicate the prepeptide amino acid sequence and the sequence positions of the primers which were used for *Dnbglu2* amplification are underlined.

The Dnbglu1 and Dnbglu2 proteins, which are encoded by *Dnbglu1* and *Dnbglu2*, respectively, matched peptide sequences from the purified protein, and had 83% identity with each other as shown in Figure 3.39. Dnbglu1 and Dnbglu2 had 85% and 81% identity, respectively, with the Thai rosewood β-glucosidase. Both proteins contain glutamic (E) residues, which are located in the catalytic acid/base and consensus motifs in family 1 β-glucosidases at position 182 within the sequences TVN**E**P and TIN**E**P for Dnbglu1 and Dnbglu2, respectively, and the nucleophile consensus motif at position 396 within the sequence YIT**E**NG for both proteins. Both proteins had highest sequence identity with cyanogenic β-glucosidase from white clover (1CBG) among proteins with known structures. From the NetNGly prediction program (NetNGlyc 1.0 server, http://www.cbs.dtu.dk/cgi-bin/), both Dnbglu1 and Dnbglu2 protein sequences contained 5 possible N-linked glycosylation sites.

A stop codon (TGA) was found in Dnbglu1 cDNA at amino acid position 95 but there was not any stop codon in partial DNA amplification sequences. Protein sequence alignment with Thai rosewood and other β-glucosidase, showed that there is a conserved tryptophan residue in this position. So, the stop codon in Dnbglu1 cDNA was mutated to TGG, which encodes a tryptophan residue.

```
>N-terminal seq                                        ATITEV                    •
Dnbglu1_          -23 MHAMTFKAILLLGLLALVSTSASIAFAKEVRATITEVPPFNRNSFPSDFIFGTAASSYQY 37
Dnbglu2           -23 MIAMTFKVILLLGLLALISTSTSIAFPKEVRATITEVPPFNRSCFPSDFIFGASSSAYQY 37
Thai_rosewood     -23 MLAMTSKAILLLGLLALVSTSASIDFAKEVRETITEVPPFNRSCFPSDFIFGTASSSYQY 37
1CBG                1 -------------------------FKPLPISFDDFSDLNRSCFAPGFVFGTASSAFQY 34
ZM_Glu1             1 ----------------------SARVGSQNGVQMLSPSEIPQ-RDWFPSDFTFGAATSAYQI 39

>peptide Tryp 1                                                          YMNLDAYR
Dnbglu1_          38  EG----EGRVPSIWDNFTHQYPEKIADGSNGDVAVDQFHHYKEDVAIMKYMNLDAYRLSI 93
Dnbglu2           38  EG----EGRVPSIWDNFTHQYPEKIADGSNGDVTIDQFHRYKEDVAIMKYMNLDAYRLSI 93
Thai_rosewood     38  EG----EGRVPSIWDNFTHQYPEKIADRSNGDVAVDQFHRYKKDIAIMKDMNLDAYRMSI 93
1CBG              35  EGAAFEDGKGPSIWDTFTHKYPEKIKDRTNGDVAIDEYHRYKEDIGIMKDMNLDAYRFSI 94
ZM_Glu1           40  EGAWNEDGKGESNWDHFCHNHPERILDGSNSDIGANSYHMYKTDVRLLKEMGMDAYRFSI 99

>peptides Tryp 2 & 3           ASGGIISTGVD    LINETLANGI       •
Dnbglu1_          94  SWPRILPTGRASGGINSTGVDYYNRLINELLANDITPFVTIFHWDLPQALEDEYGGFLNH 153
Dnbglu2           94  SWPRILPTGRASGGINSTGVDYYNRLINETLHNGITPYVTIFHWDLPQALEDEYGGFLDR 153
Thai_rosewood     94  SWPRILPTGRVSGGINQTGVDYYNRLINESLANGITPFVTIFHWDLPQALEDEYGGFLNH 153
1CBG              95  SWPRVLPKGKLSGGVNREGINYYNNLINEVLANGMQPYVTLFHWDVPQALEDEYRGFLGR 154
ZM_Glu1          100  SWPRILPKGTKEGGINPDGIKYYRNLINLLLENGIEPYVTIFHWDVPQALEEKYGGFLDK 159

                                                  •(Q)    ∇     (M)∇
>peptide Tryp 4                                  HWITVNEPSIFTMNGYAYGIFAPGR
Dnbglu1_         154  ---TIVNDFRDYADLCFNLFGDRVKHWITVNEPSIFTMNGYAYGIFAPGRCSPSYNPTCT 210
Dnbglu2          154  ---RVVNDFRDYADLCFKFFGDRVKHWITINEPQVFTTNGYTYGMFAPGRCSPSYDPTCT 210
Thai_rosewood    154  ---SVVNDFQDYADLCFQLFGDRVKHWITLNEPSIFTANGYAYGMFAPGRCSPSYNPTCT 210
1CBG             155  ---NIVDDFRDYAELCFKEFGDRVKHWITLNEPWGVSMNAYAYGTFAPGRCSDWLKLNCT 211
ZM_Glu1          160  SHKSIVEDYTYFAKVCFDNFGDKVKNWLTFNEPQTFTSFSYGTGVFAPGRCSPGLDCAYP 219

Dnbglu1_         211  GGDAGTEPDLVAHNLILSHAATVQVYKKKYQEHQNGIIGISLQIIWAVPLSNSTSDQKAA 270
Dnbglu2          211  GGDAGTEPYKVAHNLLLSHAATVQVYKEKYQKDQNGKIGITLDQRWVIPLSNSTSDKKAA 270
Thai_rosewood    211  GGDAGTETYLVAHNLILSHAATVQVYKRKYQEHQKGTIGISLHVVWVIPLSNSTSDQNAT 270
1CBG             212  GGDSGREPYLAAHYQLLAHAAAARLYKTKYQASQNGIIGITLVSHWFEPASKEKADVDAA 271
ZM_Glu1          220  TGNSLVEPYTAGHNILLAHAEAVDLYNKHYKRDDTR-IGLAFDVMGRVPYGTSFLDKQAE 278

                                                                          ∇
Dnbglu1_         271  QRYLDFTGGWFLDPLTAGQYPESMQYLVGDRLPKFTTDEAKLVKGSFDFVGINYYTSSYL 330
Dnbglu2          271  QRYLDFTFGWFMDPLTVGRYPDSMQYLVGNRLPKFTTYEAKLVKGSFDFIGINYYTSNYA 330
Thai_rosewood    271  QRYLDFTCGWFMDPLTAGRYPDSMQYLVGDRLPKFTTDQAKLVKGSFDFIGLNYYTTNYA 330
1CBG             272  KRGLDFMLGWFMHPLTKGRYPESMRYLVRKRLPKFSTEESKELTGSFDFLGLNYYSSYYA 331
ZM_Glu1          279  ERSWDINLGWFLEPVVRGDYPFSMRSLARERLPFFKDEQKEKLAGSYNMLGLNYYTSRFS 338

                                                      ∇    ∇
Dnbglu1_         331  TSSDASTCCPPSYLTDSQVTFSSQR--NGVFIGPVTPSGWMCIYPKGLRDLLLYIKEKYN 388
Dnbglu2          331  TKSDASTCCPPSYLTDPQVTLSSQR--NGVFIGPMTPSGWICIYPKGLRDLLLYIKENYN 388
Thai_rosewood    331  TKSDASTCCPPSYLTDPQVTLLQQR--NGVFIGPVTPSGWMCIYPKGLRDLLLYFKEKYN 388
1CBG             332  AKAPRIPNARPAIQTDSLINATFEH--NGKPLGPMAASSWLCIYPQGIRKLLLYVKNHYN 389
ZM_Glu1          339  KNIDISPNYSPVLNTDDAYASQEVNGPDGKPIGPPMGNPWIYMYPEGLKDLLMIMKNKYG 398

                      •                                                        •
Dnbglu1_         389  NPLVYITENGMDELD--DPSQSLEESLIDTYRIDSYYRHLFYVRSAIGSGANVKGFFAWS 446
Dnbglu2          389  NPLVYITENGMDETN--DPSLSLEESLMDTYRIDSYYRHLFYVLSAIKSGANVKGFFAWT 446
Thai_rosewood    389  NPLVYITENGIDEKN--DASLSLEESLIDTYRIDSYYRHLFYVRYAIRSGANVKGFFAWS 446
1CBG             390  NPVIYITENGRNEFN--DPTLSLQESLLDTPRIDYYYRHLYYVLTAIGDGVNVKGYFAWS 447
ZM_Glu1          399  NPPIYITENGIGDVDTKETPLPMEAALNDYKRLDYIQRHIATLKESIDLGSNVQGYFAWS 458

                       ••∇∇      ∇
Dnbglu1_         447  LLDNFEWNEGFTSRFGLNFVNYT-TLTRYHKLSATWFKYFLARDQEIAKLDISAPKARWS 505
Dnbglu2          447  LMDDFEWSGGFTSRFGLNFVDYN-TLNRYPKLSAKWFKYFLTRDQESAKLDISTPKASAA 505
Thai_rosewood    447  LLDNFEWAEGYTSRFGLYFVNYT-TLNRYPKLSATWFKYFLARDQESAKLEILAPKARWS 505
1CBG             448  LFDNMEWDSGYTVRFGLVFVDFKNNLKRHPKLSAHWFKSFLKK----------------- 490
ZM_Glu1          459  LLDNFEWFAGFTERYGIVYVDRNNNCTRYMKESAKWLKEFNTAKKPSKKILTPA------ 512

Dnbglu1_         506  SSTMIKEEKRKPKWAIQAF 524
Dnbglu2          506  YQR                 508
Thai_rosewood    506  LSTMIKEEKTKPKRGIEGF 524
```

Figure 3.39 Protein sequence alignment of β-glucosidase from *D. nigrescens*;

Dnbglu1, Dnbglu2, *D. cochinchinensis* (Thai rosewood), cyanogenic β-glucosidase from white clover, (1CBG) and *Zea maiz* β-glucosidase1 (ZMGlu1). N-terminal and tryptic peptide sequences obtained by Edman degradation of the purified protein are aligned above the full sequences with secondary signals in parentheses above the peptide. The sites indicated by Czjzek *et al*. as interacting with the glycone (•) or aglycone (∇) of maize β-glucosidase are marked and the catalytic acid/base and nucleophile consensus sequences are underlined with the corresponding glutamates in bold. The red letters indicate the N-glycosylation sites of *Dalbergia* β-glucosidases.

3.2.3 Amino acid sequencing

Amino acid sequencing of the purified enzyme showed 5 peptides which match to *D. nigrescens* proteins as in Table 3.8. The N-terminus sequenced from the non digested, native protein had a low sequencing signal, which corresponded to a sequence 8 residues after the N-terminus of *D. cochinchinensis* β-glucosidase.

Table 3.8 Peptides sequences of native *D. nigrescens* β-glucosidase

peptides	Amino acid sequences
N-terminus	ATITEV
1.	YMNLDAYR
2.	ASGGIISTGVD
3.	LINETLANGI
4.	HWITVNEPSIFTMNGYAYGIFAPGR

3.3 Expression of *Dalbergia nigrescens* β-glucosidases in *E. coli*

3.3.1 Construction of cDNA encoding mature β-glucosidase protein, Dnbglu1 in pET-32a and pET-40b expression plasmid vectors

The full-length cDNA of *Dnbglu1* was used as template to introduce an *Eco*R I site at the initial codon of the mature protein to eliminate the signal peptide sequence by PCR amplification with 2 sets of primers, as is summarized in Figure 3.40. The PCR products of the two fragments were produced using the T7 promoter and EcoRIMR primers (188 bp) and EcoRIMF and BseRIR primers (307 bp), respectively. Then, the PCR product containing the introduced *Eco*R I site was produced, with the expected size of 474 bp, with the 188 and 307 bp PCR products as primers and full-length *Dnbglu1* cDNA of as a template. Then, the 474 bp PCR product was re-ligated back into the full-length cDNA at the *Pst* I and *BseR* I sites. After confirming the introduced *EcoR* I site by sequencing, the mature DNA sequence was cut with *EcoR* I and *Not* I enzymes and ligated into the pET-32a and pET-40b expression vectors.

Dnbglu1 protein which was expressed in Origami (DE3) *E. coli* transformed with pET32a(+) vector seemed to be inactive, because cell extract could not hydrolyze *p*NP-β-D-glucoside and *p*NP-β-D-fucoside and the Dnbglu1 fusion protein appeared in the insoluble cytoplasmic fraction (inclusion bodies) (Figure 3.41). When the soluble protein fractions were analyzed on SDS-PAGE, no Dnbglu1 protein could be identified (Figure 3.42). So, the inclusion bodies of Dnbglu1 were used for antibody production. However, cell extracts of BL21(DE3) *E. coli* transformed with

the pET40b construct and induced had neither obvious Dnbglu1 protein bands nor detectable activity in either the soluble or insoluble fraction.

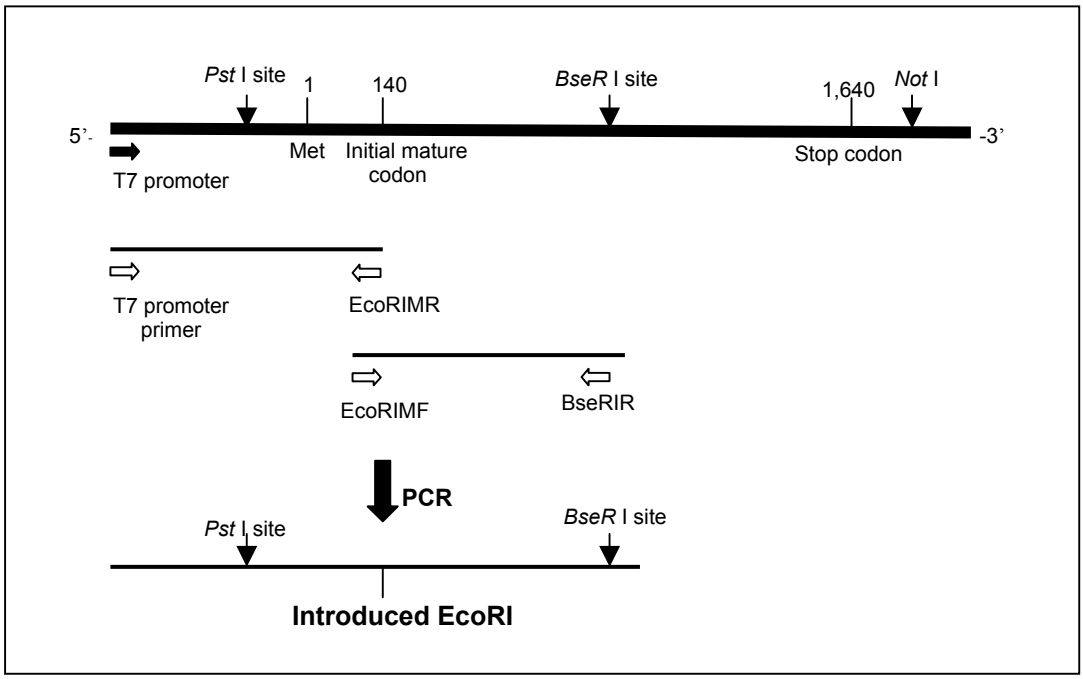


Figure 3.40 Cloning strategies for introducing on *EcoR* I site at initial mature codon of *Dnbglu1* before construction of pET expression vectors. The two PCR product pieces were produced using the T7 promoter and EcoRIMR primers to produce a 188 bp PCR product, and the EcoRIMF forward and BseRIR primers to produce a 307 bp PCR product. Then, the full N-terminal cDNA fragment PCR product with the introduced *Eco*R I site was produced with the expected size 474 bp, with 188 and 307 bp PCR products as primers and the full-length cDNA of *Dnbglu1* as a template. The 474 bp PCR product was cut with *Pst* I and *BseR* I and ligated back into the full-length cDNA in pBlueScript SK(+) construct which had been cut at the same sites. The cDNA encoding the mature protein was

then excised and cloned into the pET expression vector at the *EcoR* I and *Not* I sites.

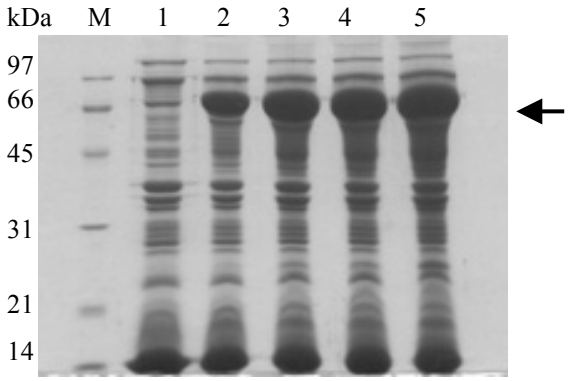


Figure 3.41 SDS-PAGE of insoluble fractions from recombinant Dnbglu1 protein expression in Origami (DE3) *E. coli* transformed with pET32a(+) after incubation in the presence of 0.4 mM IPTG**,** at 20^{o}C for 0-16 hrs. Lane M, Low-range molecular weight protein marker (Bio-RAD); lane 1, induced for 0 hr; lane 2, induced for 6 hrs; lane 3, induced for 9 hrs; lane 4**,** induced for 12 hrs; lane 5**,** induced for 16 hrs**.** The arrow indicates the positions of mature Dnbglu1 monomer at approximately 66 kDa.

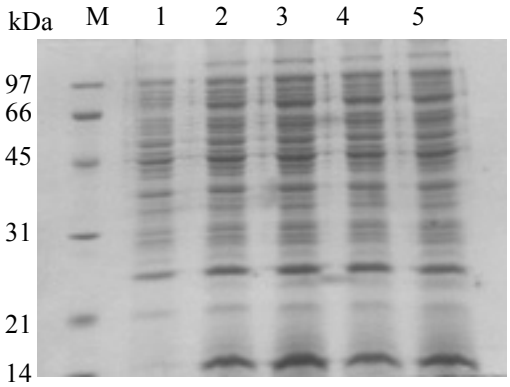


Figure 3.42 SDS-PAGE of soluble proteins from cell lysates of Origami (DE3) *E. coli* transformed with pET32a(+) with Dnbglu1 after incubation in the presence of 0.4 mM IPTG, at 20^{o}C for 0-16 hrs**.** Lane M, Low-range molecular weight protein marker (Bio-RAD); lane 1, induced for 0 hr; lane 2, induced for 6 hrs; lane 3, induced for 9 hrs; lane 4**,** induced for 12 hrs; lane 5**,** induced for 16 hrs**.**

3.4 Expression of *Dalbergia nigrescens* β-glucosidases in *Pichia pastoris*

3.4.1 Construction of pPICZαB-Thrombin expression vector plasmids containing the cDNA encoding mature Dnbglu1 and Dnbglu2 β-glucosidase proteins

The full length *Dnbglu1* cDNA in pBlueScript KS(+) vector construct was used as template to introduce a *Pst* I site at the initial codon of the predicted mature protein to eliminate the signal peptide and a *Sac* II site at the 3′ end by PCR amplification with PstgluF102 forward and DnbgluR2SacII reverse primers, which produced a 1,600 bp product (Figure 3.44). Then, the PCR product was cloned into the *Eco*RV site of pBlueScript KS(+), digested, and cloned into pPICZαB-Thrombin at the *Pst* I and *Sac* II sites, and the plasmid insert was confirmed by sequencing. The construction of the expression vector with the cDNA is summarized in Figure 3.43.

For Dnbglu2 expression, the cDNA encoding mature protein of Dnbglu2 was amplified using primers to introduce a *Pst* I site at the amino terminal of the predicted mature protein and an *Xba* I site at the 3′ end of the coding region with Dn2NTERMPstIf and Dn2CTERMXbaIr primers, respectively, to produce a 1,650 bp product. Then the PCR was cloned into the pPICZαB-Thrombin at the *Pst* I and *Xba* I sites, and the plasmid insert was confirmed by sequencing.

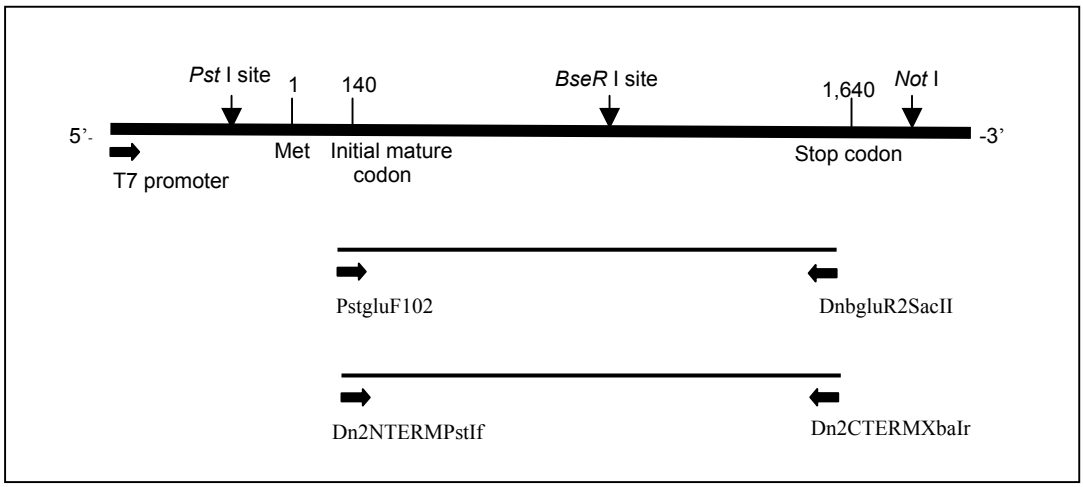


Figure 3.43 Production of *Pichia* expression vector inserts. The full length *Dnbglu1* cDNA in the pBlueScript KS(+) vector construct was used as template to introduce a *Pst* I site at the initial codon of the mature protein to eliminate the signal peptide and a *Sac* II site at the 3′ end by PCR amplification with PstgluF102 and DnbgluR2SacII primers. After that, the PCR product was cloned into pPICZαB-Thrombin at the *Pst* I and *Sac* II sites and expressed in *P. pastoris* strain GS115. *Dnbglu2* cDNA in the TOPO vector construct was used as template to introduce a *Pst* I and an Xba I site with Dn2NTERMPstIf and Dn2NTERMPXbaIr primers. After that, the PCR was cloned into the pPICZαB-Thrombin at the *Pst* I and Xba I sites and expressed in *P. pastoris* strain GS115.

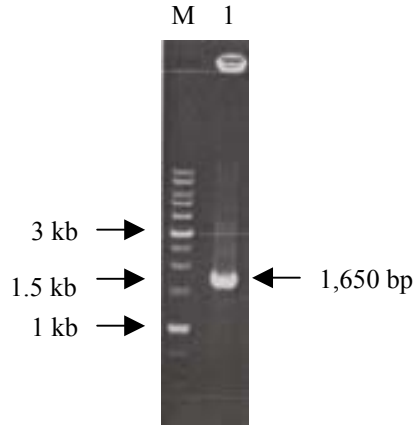


Figure 3.44 1% agarose gel electrophoresis of the PCR product of the *Dnbglu2* cDNA which was amplified with Dn2NTERMPstIf forward and Dn2CTERMXbaIr reverse primers, to produce a 1,650 bp product. Lane M, 1 kb DNA marker; lane 1, PCR product of cDNA for expression vector.

3.4.2 Recombinant protein expression of Dnbglu1 and Dnbglu2

After the cDNA encoding Dnbglu1 and Dnbglu2 were inserted in pPICZαB-Thrombin, both constructs were used to transform *Pichia pastoris* strain GS115. However, the pPICZαB-Thrombin-*Dnbglu1* which was expressed in *P. pastoris* produced no active protein, because media and cell lysates from induced cultures could not hydrolyze any of the expected substrates, such as *p*NP β-D-glucoside, *p*NP β-D-fucoside, and the both natural substrates.

Dnbglu2 protein was expressed as an active protein, then the clones which had highest activity toward *p*NP-β-D-fucoside were screened and compared

with that expressed from pPICZαB-Thrombin containing Thai rosewood β-glucosidase cDNA (M. Ketudat-Cairns, unpublished) which was expressed in the same experiment, as shown in Figure 3.45. After that, the selected clone was expressed on a large scale and partially purified using Sephacryl S-300 gel filtration (Figure 3.46). The fractions with activity were pooled, concentrated and characterization. However, no obvious band could be seen on SDS-PAGE of media from cultures expressing Dnbglu2.

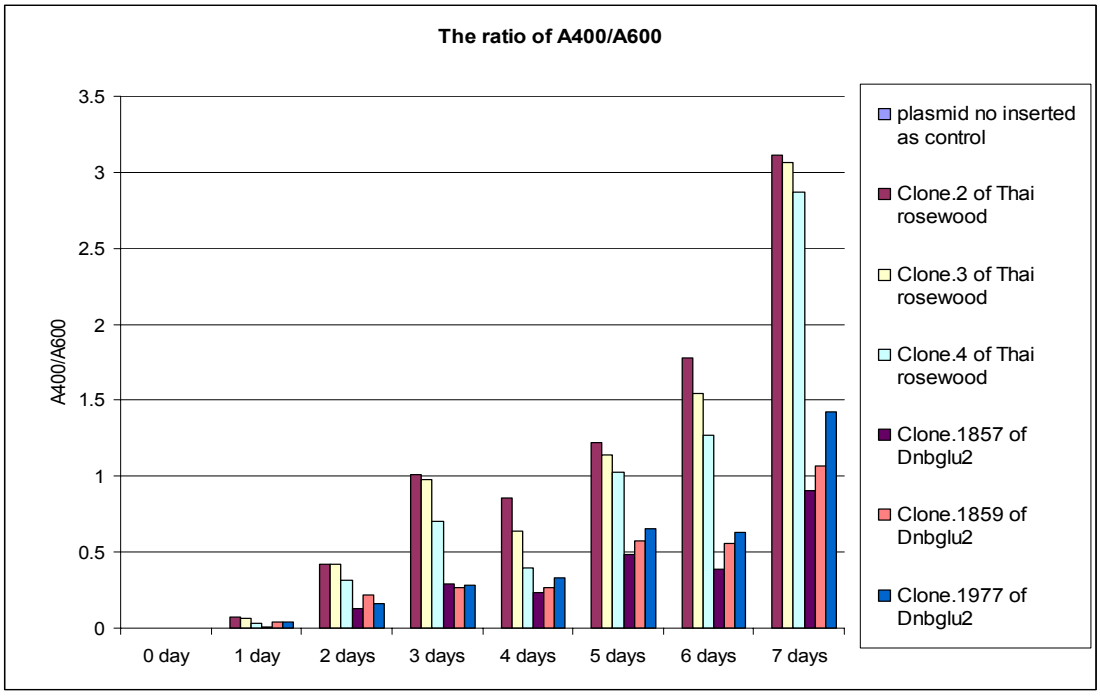


Figure 3.45 The ratio of activity to cell density A_{405}/A_{600} of expressed Dnbglu2 protein in BMMY medium using pPICZαB-thrombin without insert as a negative control and Thai rosewood as a positive control. The

activities were determined from 0 to 7 days using 1 mM *p*NP-β-D-fucoside as substrate with incubation at 30°C for 10 min.

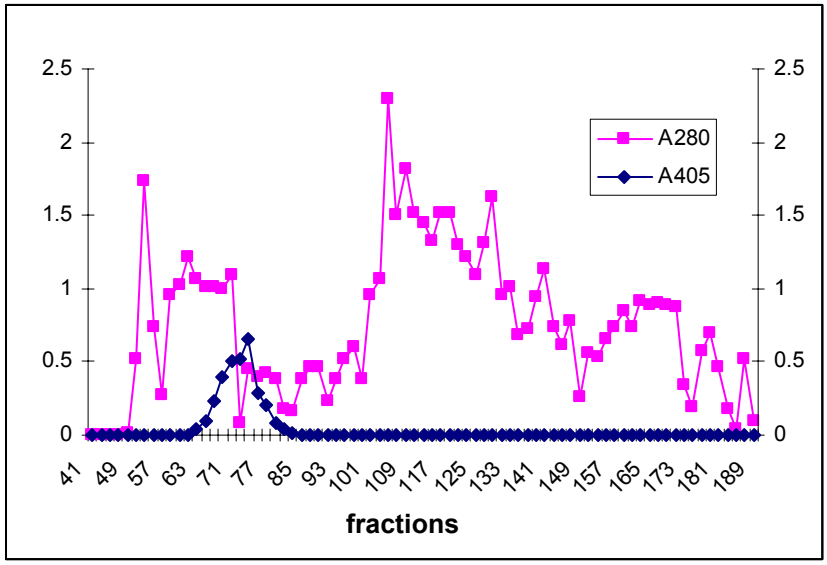


Figure 3.46 The chromatograph of purification of expressed Dnbglu2 protein by Sephacryl S-300 gel filtration. The red line shows the absorbance of protein at 280 nm and the blue lines show the activity with 1 mM *p*NP-β-D-fucoside as substrate and the amount of *p*-nitrophenol released measured by absorbance at 405 nm. The reaction was performed by incubating 5 µL of each sample fraction with 1 mM *p*NP-β-D-fucoside in the 140 µL total reaction at 30°C for 10 min and the reaction was stopped with 70 µL of 0.4 M Na_2CO_3.

Activity staining of a non denaturing electrophoresis gel of recombinant Dnbglu2 protein showed that it had activity toward 4-methyl-umbelliferyl-β-D-fucopyranoside, 4-methyl-umbelliferyl-β-D-glucopyranoside and 4-methyl-umbelliferyl-β-D-galactopyranoside, as did the native *D. nigrescens* β-glucosidase, but migrated somewhat faster (Figure 3.47).

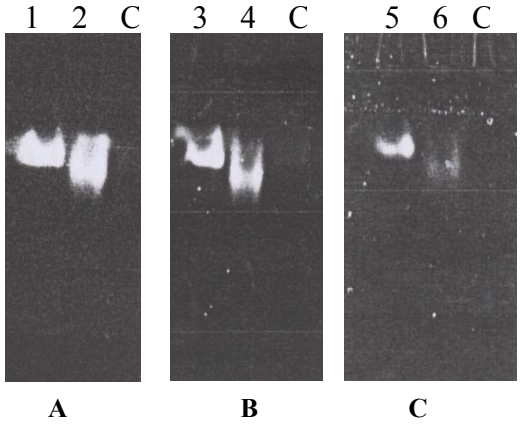


Figure 3.47 Activity staining of non denaturing electrophoresis of recombinant *D. nigrescens* Dnbglu2 β-glucosidase/β-fucosidase expressed in *P. pastoris*. Activity staining gel; lanes 1, 3, 5, native *D. nigrescens* enzyme; lanes 2, 4, 6, recombinant Dnbglu2 enzyme stained with **A**, 1 mM 4-MU-β-D-glucoside, **B**, 1 mM 4-MU-β-D-fucoside and **C**, 1 mM 4-MU-β-D-galactoside. Medium from control *Pichia* cultures without Dnbglu2 is shown in lane C on each gel.

The hydrolytic activities of purified native *D. nigrescens*, recombinant *D. nigrescens* and Thai rosewood β-glucosidase toward the S1 and S2 natural substrates were also confirmed by TLC (Figure 3.48).

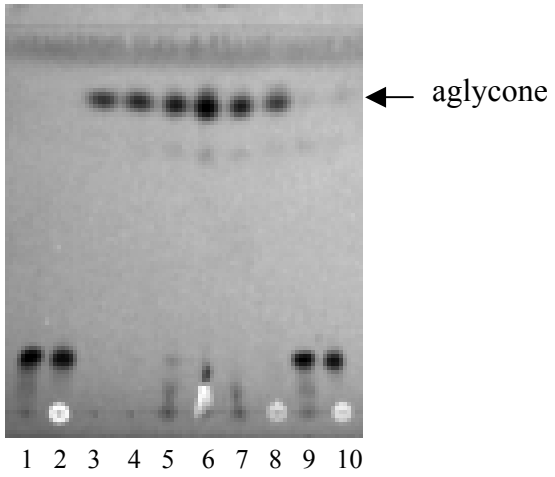


Figure 3.48 A TLC chromatogram representing hydrolysis of *D. nigrescens* substrates with purified *D. nigrescens* and *D. cochinchinensis* β-glucosidases. Purified *D. nigrescens* glycosides were hydrolyzed 10 min with equal amounts of the enzyme (based on 1 mM *p*NP-glucoside hydrolysis) *D. nigrescens* or *D. cochinchinensis* β-glucosidase, and its hydrolytic products were separated on silica gel 60 F_{254} with $CHCl_3$/MeOH/H_2O (15:3:1) as developing solvent. Lanes 1 and 2, S2 and S1 glycoside standards, respectively; 3 and 4, aglycones of S1 and S2, respectively; 5 and 6, S2 and S1, respectively, hydrolyzed with expressed recombinant β-glucosidase; 7 and 8, S2 and S1, respectively, hydrolyzed with native *D. nigrescens* β-glucosidase;

9 and 10, S2 and S1, respectively, hydrolyzed with *D. cochinchinensis* β-glucosidase.

3.4.3 Determination of K_m for hydrolysis of *p*NP-β-D-glucoside, *p*NP-β-D fucoside and natural *D. nigrescens* substrates (S1 and S2) by recombinant *D. nigrescens* Dnbglu2.

The K_m of recombinant *D. nigrescens* Dnbglu2 β-glucosidase/β-fucosidase was determined for *p*NP-β-D-glucoside, *p*NP-β-D-fucoside, S1 and S2 substrates. The recombinant *D. nigrescens* β-glucosidase had similar K_m values to those of the *D. nigrescens* β-glucosidase enzyme purified from seeds, for all four substrates, as summarized in Table 3.9.

Table 3.9 The kinetic properties of recombinant *D. nigrescens* β-glucosidase enzyme partially purified by Sephacryl S-300 gel filtration. Assays were performed in 0.1 M sodium acetate buffer, pH 5, at 30°C.

Substrate	K_m (mM)	
	Recombinant	Native
*p*NP-β-D-glucoside	12.1 ± 2.5	14.7 ± 2.9
*p*NP-β-D-fucoside	1.4 ± 0.05	1.8 ± 0.2
D. nigrescens S1 glycoside	0.5 ± 0.1	0.5 ± 0.1
D. nigrescens S2 glycoside	0.7 ± 0.1	0.7 ± 0.1

CHAPTER IV

DISCUSSION

The native *Dalbergia nigrescens* β-glucosidase and two natural substrates, compound S1, dalpatein 7-*O*-[β-D-apiofuranosyl-(1→6)-β-D-glucopyranoside] and compound S2, dalnigrein 7-*O*-[β-D-apiofuranosyl-(1→6)-β-D-glucopyranoside] were purified from seeds of *Dalbergia nigrescens* Kurz and characterized. *Dnbglu1* and *Dnbglu2* cDNA which encode *D. nigrescens* β-glycosidases, were also cloned and sequenced. The *Dnbglu2* cDNA was expressed in *Pichia pastoris,* characterized and compared with the native β-glucosidase. The native enzyme and the recombinant Dnbglu2 have similar K_m values for *p*NP-β-D-glucoside, and *p*NP-β-D-fucoside and the same K_m values of 0.5 mM for S1 and 0.7 mM for S2, respectively. However, recombinant Dnbglu1 could not be expressed in active form.

4.1 Purification of β-glucosidase/β-fucosidase from *Dalbergia nigrescens* Kurz.

During the extraction of enzyme based on the Thai rosewood purification procedure (Srisomsap *et al*., 1996), the crude *D. nigrescens* seed extract was found to darken over time. So, a new method was adapted and used for purification to inhibit phenol oxidase activity and eliminate phenolic compounds as much as possible. In

addition to PMSF and PVPP, β-mercaptoethanol and ascorbic acid were added in order to protect against oxidation reactions. Also, to remove phenolic compounds, the plant material was extracted with the high salt first (70% $(NH_4)_2SO_4$), since the phenolic compounds would be eliminated due to their dissolving, while the enzyme would precipitate in the solid material. Although, using buffers at alkaline pH, rapid extraction at 4^{o}C and using reducing agents in the extraction solution has been seen to be effective against SH proteinases, under these conditions phenol oxidase activity is seen, which may lead to browning and alteration of the proteins of interested by oxidized phenolics (Esen and Cokums, 1991).

*p*NP-β-D-glucoside and *p*NP-β-D-fucoside were used as substrates to determine the enzyme activity because they have been used to screen for glycosyl hydrolases in Thai plants and characterize the Thai rosewood enzyme during purification (Surarit *et al*., 1995 and Srisomsap *et al*., 1996). Also, the enzyme from *D. nigrescens* was similar to Thai rosewood β-glucosidase, which has both activities. Throughout the enzyme purification steps, the ratio of β-glucosidase to β-fucosidase was always approximately two. So, it might be the same as the *D. cochinchinensis* enzyme, which has been shown to be a single enzyme with glucosides and fucosides reacting at a single active site (Surarit *et al*., 1996).

The total protein dramatically decreased (approximated 100 times) after DEAE-anion exchange chromatography, which suggested that the decrease in the amount protein (the absorbance) might come from the absorbance of some phenolic compounds in the extract (browning colour). Dialysis of the $(NH_4)_2SO_4$ fraction against buffer A (for ion exchange chromatography) might have reactivated the

inactive phenolic oxidase which would oxidize phenolic compounds because the crude extract seemed to turn brown again.

The molecular weight of *D. nigrescens* β-glucosidase estimated from SDS-PAGE (62.30 kDa) was similar to the molecular weight from the predicted mature protein sequence (61.858 Da) The native molecular weight of purified enzyme was estimated to be 240 kDa using Sephacryl S-300 gel filtration chromatography, which supports the idea that *D. nigrescens* β-glucosidase is composed of 4 subunits, as is the Thai rosewood β-glucosidase (Srisomsap *et al*., 1996).

4.2 pH and temperature optimum

The pH and temperature optima of the purified enzyme were pH 5-5.5 and 65°C, respectively, for both β-glucosidase and β-fucosidase activities. The pH optimum of this enzyme was similar to other β-glucosidases (pH 5-6) (Hrmova and Fincher, 1998; Esen, 1993 and Esen, 1992). But, the temperature optimum of this enzyme was pretty high (65^{o}C) compared to other β-glucosidases, which are often approximately at 50^{o}C (Akiyama *et al*., 1998; Konno *et al*., 1996 and Riou *et al*., 1998). But, some β-glucosidases have higher optimal temperatures, such as the Thai rosewood β-glucosidase which has a temperature optimum of 60^{o}C (Srisomsap *et al*., 1996). However, the rate of chemical reactions increase with temperature, thus the temperature optimum in an enzyme-catalyzed reaction does not necessarily mean that the temperature is optimum for enzyme activity, but the one beyond which the enzyme begins to denature in the time frame of the assay (Esen, 1992).

4.3 Inhibition studies

There were some β-glucosidases which have been reported to be strongly inhibited by glucono-1,5-lactone (δ-gluconolactone), an analogue of the oxocarbenium ion-like transition state, such as β-glucosidase enzymes from rice, maize and Thai rosewood (Akiyama *et al*., 1998; Opassiri *et al*., 2003; Babcock and Esen, 1994 and Srisomsap *et al*., 1996). However, other β-glucosidases have been reported to be resistant to this inhibitor, such as β-glucosidases from the cell wall of carrot and an isoflavone-7-*O*-glucoside-specific β-glucosidases from *Cicer arietinum* L. (Konno *et al*., 1996 and Hösel and Barz., 1975). Both native and recombinant β-glucosidase enzymes were inhibited approximately 50% by 1 mM Glucono-1,5-lactone, which is similar to the *Cicer arietinum* isoflavone-7-*O*-glucoside-specific β-glucosidases. Glucono-1,5-lactone inhibited the enzyme more effectively at higher concentrations and incubation times seemed to effect the inhibition. The inhibition of *D. nigrescence* β-glucosidase by δ-gluconolactone was weaker than *D. cochinchinensis* β-glucosidase. However, aldonolactones are unstable and interact with water to produce complex conformational changes that depend on the pH, temperature and age. δ-gluconolactone in solution can transform into D-gluconic acid and D-glucono-1,4-lactone, which are less inhibitory (Sanz-Aparicio *et al*., 1998).

Many metal ions such as Zn^{2+}, Co^{2+}, Ba^{2+}, Mg^{2+}, Ca^{2+}, Na^{+}, K^{+}, Mn^{2+} and Cu^{2+} have little effect on β-glucosidase enzymes, but there are 2 metal ions, Ag^{2+}, Hg^{2+}, which have been shown to effectively inhibit the enzyme activity (Konno *et al*., 1996; Srisomsap *et al*., 1996 and Esen, 1992). The loss of enzyme activity due to Ag^{2+} and Hg^{2+} at low concentration may be caused by their forming insoluble

complexes with cysteine residues (-SH group) while other metals formed as soluble complexes. But, the enzyme activity can be recovered by adding the 2-mercaptoethanol, which will displace the metal ions from the sulfhydryl group of cysteine. Additional of 2-mercaptoethanol at low concentration can maintain the activity of the enzyme because it will reduced the oxidized sulhydryl groups which cause disfunction of the enzyme. However, at high concentration (ie 14% V/V), 2-mercaptoethanol completely inhibits the activity because it destroys the disulfide bonds located in the interior of the enzyme by reduction (Esen, 1992). But both native and recombinant *D. nigrescens* β-glucosidase enzyme were not affected much by Hg^{2+}, which may suggest that any cysteines that are involved in stability and activity at the enzyme are located in the interior of the enzyme, where it is difficult for the metal to access.

2,4-dinitrophynyl-2-deoxy-2-fluoroglucoside had the most effect on the activity of both native and recombinant β-glucosidase enzymes compared with other inhibitors but these enzymes seemed to be much more resistant than Thai rosewood β-glucosidase (Srisomsap *et al*., 1996), which was effectively inactivated at 10 μM (no detectable activity remaining).

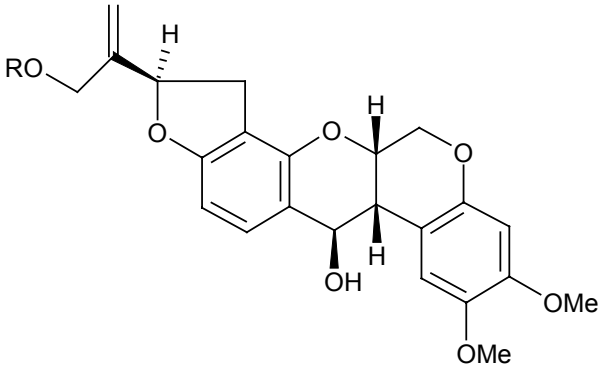
4.4 Substrate specificity

Of the various *p*NP-glycosides tested as substrates, the enzyme showed the highest activity toward *p*NP-β-D-fucoside and *p*NP-β-D-glucoside, which was similar to *D. cochinchinensis* β-glucosidase (Srisomsap *et al*., 1996). In contrast, the *D. nigrescens* enzyme hydrolyzed *p*NP-β-D-xyloside and *p*NP-β-D-thioglucoside with relative activities 2 and 34 times those of *D. cochinchinensis* β-glucosidase, and *p*NP-β-D-galactoside and *p*NP-α-L-arabinoside with lower relative activities.

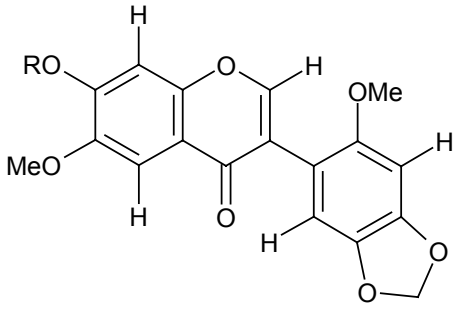
Release of glucose from commercially available natural glycosides and oligosaccharides was also tested. The rates of hydrolysis were extremely low compared to the hydrolysis of *p*NP-β-D-glucoside, with laminaribiose (β1→3), linamarin and cellobiose being the only substrates from which release of glucose could be detected. No hydrolysis of laminaritriose, cellotriose, salicin, dhurrin, DIMBOA glucoside, sophorose, torvoside A or amygdalin could be observed. This suggests that the aglycone structure is critical for glucose cleavage, such as has been seen for some cyanogenic glucosidases that display narrow specificities for the aglycone moiety (Hösel and Conn, 1982). The *D. nigrescens* enzyme also cleaved dalcochinin-8'-*O*-β-D-glucoside, the natural substrate of *D. cochinchinensis* β-glucosidase, as judged by TLC analysis. However, glucose could not be detected after dalcochinin glucoside hydrolysis by the PGO assay with *D. nigrescens* β-glucosidase, while it could when it was hydrolyzed with *D. cochinchinensis* β-glucosidase. But, *D. nigrescens* β-glucosidase showed transglycosylation activiy in the reaction which has 10% dalcochinin-8'-*O*-β-D-glucoside in methanol in the hydrolysis reaction.

When the *D. nigrescens* enzyme hydrolyzed this substrate in the reaction which had no methanol, glucose was detectable with the PGO assay.

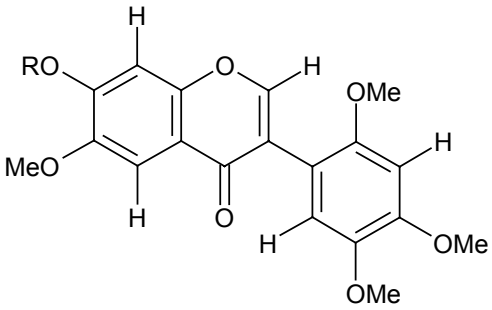
The *D. nigrescens* enzyme efficiently hydrolyzed two natural isoflavonoid glycosides isolated from *D. nigrescens* seeds, which appeared to have glycones composed of glucose and apiose, based on the NMR data and GC-MS results after the glycoside (S2) was hydrolyzed with H_2SO_4, which indicated peaks the same retention time of as derivatized glucose and apiose standards. This data was also confirmed by comparison with references of known structures. Both glycosides were determined to be isoflavonoid derivatives which are glycosylated in the 7 position, similar to other classes of flavonoids, such as the flavones, flavonones which are found mainly glycosylated in the 7 position (Harborne *et al*., 1975). *D. nigrescens* enzyme hydrolyzed both glycosides as disaccharide (apioglucoside) which was confirmed by hydrolyzing the glycoside with enzyme, purifying the sugar and then checking it on TLC. This experiment resulted in only one spot of sugar produced by the hydrolysis reaction, which migrated similar to standard disaccharides. The purified sugar product was also characterized by GC-MS and NMR to show it was the disaccharide. Previously, Hösel and Barz (1975) extracted β-glucosidases from *Cicer arietinum* L. which were specific to isoflavone-7-*O*-β-glucosides, found these enzymes had greatest specificity for the 7-position of the aglycone (isoflavone-7-*O*-β-D-glucosides), and, in the case of biochanin A (7-*O*-β-D-glucoapioside), the enzymes removed the glucoapiose as a disaccharide, which was not further hydrolyzed. Similarly, the *D. nigrescens* β-glucosidase enzyme could not completely hydrolyze the disaccharide to monosaccharides either during or after the deglycosylation step, so no glucose was released.



dalcochinin-8'-*O*-beta-D-glucoside



S1- glycoside



S2- glycoside

Figure 4.1 Structure of *D. nigrescens* and Thai rosewood glycoside substrates. In the case of dalcochinin-8′-*O*-β-D-glucoside, R is β-glucopyranose, while for S1 and S2, R is β-D-apiofuranosyl-(1→6)-β-D-glucopyranose.

The kinetic data of both native and recombinant enzyme toward the *p*NP-β-D-glucoside, *p*NP-β-D-fucoside and both natural *D. nigrescens* glycosides indicated that S1 (compound 1) and S2 (compound 2) were the best substrates. The K_m values of native and recombinant enzymes were the same, 0.5 mM for S1 and 0.7 mM for S2, respectively. Their K_m values for these substrates were about 3 and 30 times lower

than *p*NP-β-D-fucoside (1.8 mM) and *p*NP-β-D-glucoside (14.7 mM), respectively. The enzyme also showed much higher k_{cat} values for S1 and S2 than *p*NP-β-D-glucoside and *p*NP-β-D-fucoside. So, it seems likely that S1 and S2 are the natural substrates of this enzyme. The K_m and k_{cat} values for S1 and S2 are similar but if the k_{cat} /K_m of the enzyme is compared between these substrates, it may be seen that the k_{cat} /K_m of S1 is two times higher than S2, which suggests that the aglycone moiety of these natural glycosides has some influence on hydrolysis by the enzyme.

The kinetic data also indicated that native and recombinant β-glucosidases do not differ from each other with respect to substrate specificity, as is evident from their similar K_m values. So, the cloned enzyme appears to be the same as the enzyme purified from seeds.

The ability of *D. nigrescens* enzyme hydrolyzes the dalcochinin-8'-*O*-β-D-glucoside, the natural substrate of *D. cochinchinensis* β-glucosidase, suggests it might bind to the enzyme in a way analogous to substrates S1 and S2. S2 is more like S1 than it is like dalcochinin-8'-*O*-β-D-glucoside, and it is also more like dalcochinin-8'-*O*-β-D-glucoside than is S1, in that both S2 and dalcochinin have 4v′ and 5′-O-methyl groups , as shown in Figure 4.1.

4.5 Cloning the β-glucosidase and protein sequence analysis

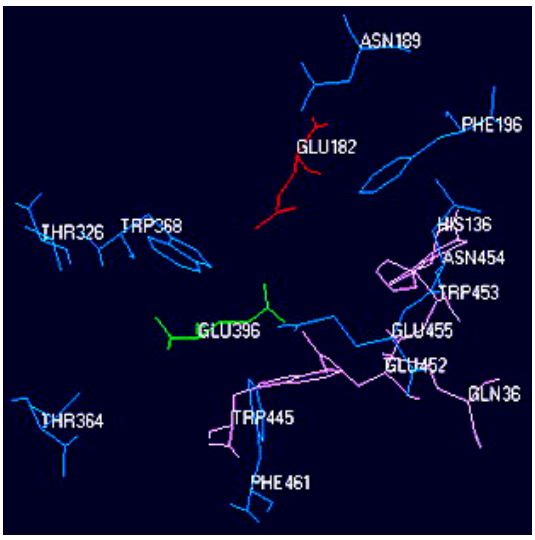
Sequencing of the N-terminus and 4 tryptic peptides of the native *D. nigrescens* enzyme protein which was purified from seeds found that the N-terminus, the tryp1 and the tryp2 peptides matched both Dnbglu1 and Dnbglu2 sequences exactly, except for the absence of N at a likely glycosylation site in tryp2 peptide. The tryp3 sequence was very similar to both Dnbglu1 and Dnbglu2, and the tryp4 peptide matched the Dnbglu1 sequence exactly, though some secondary signals appeared to match Dnbglu2. This may suggests that there may be a mixture of peptide sequences due to incompletely separated peptides in the peptide map or may come from a mixture of isozymes in the protein preparation.

Two cDNAs, *Dnbglu1* and *Dnbglu2*, of *D. nigrescens* β-glucosidase were amplified by RT-PCR and their sequences were characterized. The cDNA sequence for *Dnbglu1* included an open reading frame (ORF) encoding a 547 amino acid long precursor protein which corresponded to 23 amino acids of signal sequence and a 524 amino acid mature protein. The ORF sequence of *Dnbglu2* encoded a 531 amino acid precursor protein including a 27 amino acid of signal sequence and 504 amino acid mature proteins. Amino acid sequence analysis showed that these 2 β-glucosidases are classified as members of glycosyl hydrolase family 1, based on amino acid sequence similarities (Henrissat, 1991; Henrissat and Bairoch, 1996). The Dnbglu1 and Dnbglu2 protein sequences, which are encoded by *Dnbglu1 and Dnbglu2* had 83% identity with each other and 85% and 81% identity, respectively, with Thai rosewood β-glucosidase, which is most closely related to β-glucosidase from sweet cherry (*Prunus serotina*) (56.7%) and cyanogenic β-glucosidase (linamarase) from

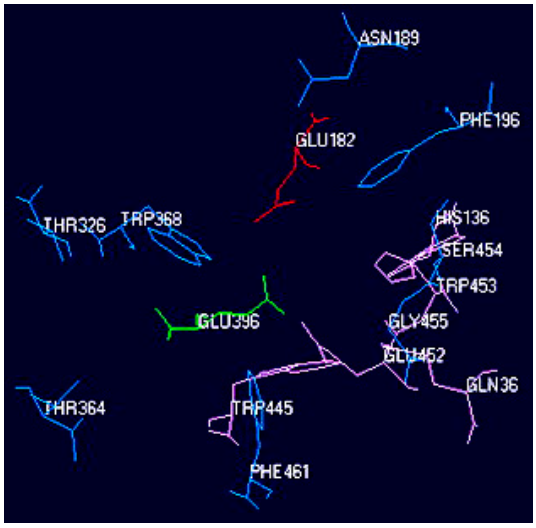
white clover (*Trifolium repens*) (56.3%) (Ketudat-Cairns *et al*., 2000). Both Dnbglu1 and Dnbglu2 protein contained the conserved amino acids regions which are found in all β-glucosidases in family 1 (Figure 3.36). Amino acid sequence alignment with the cyanogenic β-glucosidase (linamarase) from white clover, a protein of known three dimensional structure (1CBG) belonging to glycosyl hydrolase family 1 (Barett *et al*., 1995), revealed that two glutamic acids residues, E182 and E396 lie in the catalytic acid/base TXNEP (where X is a hydrophobic amino acid residue) and nucleophile ITENG motifs characteristic of family 1 members, respectively. Based on the known three dimensional structure of cyanogenic β-glucosidase (1CBG), the conserved E182 and E396 were found in the loop regions close to the carboxy terminal ends of β-strands 4 and 7, respectively. In addition, studying the active site in *Agrobacterium feacalis* β-glucosidase by labeling with 2-F-glucoside and site directed mutagenesis demonstrated that the residues homologous to E182 and E396 served as the general acid/base catalyst and the nucleophile (Withers *et al*.,1990; Withers *et al*., 1992 and Wang *et al*., 1995). So, the presence of both Glu182 and Glu396 located in these highly conserved regions suggest that both Dnbglu1 and Dnbglu2 could have the same catalytic mechanism of enzymatic hydrolysis of the glycosidic linkage via a double displacement reaction which involves general acid/base catalysis.

The aglycone binding site of maize β-glucosidase was identified by Czjzek *et al*. (2000), and found to contain 4 important amino acid residues involved in aglycone substrate binding, W378, F198, F205 and F466. Of these, only W378 is highly conserved residue in plant family 1 β-glucosidases, so the other 3 residues were proposed to provide the determination of substrate binding and specificity in this

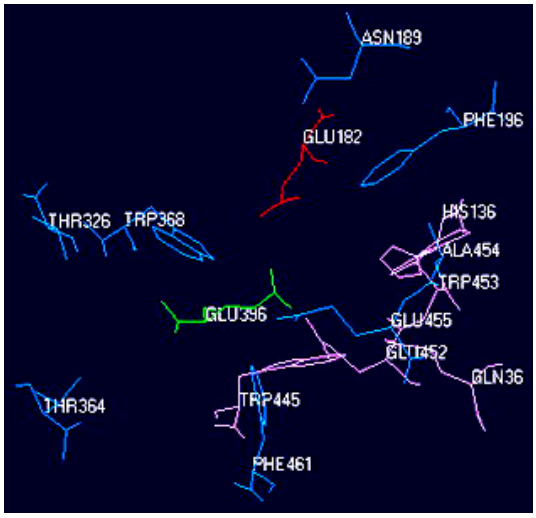
family. The equivalent residues in Dnbglu1, Dnbglu2 and Thai rosewood β-glucosidase are W368, N189, F196 and N454, S454, and A454 in Dnbglu1, Dnbglu2 and Thai rosewood protein, respectively. In addition, there were 4 other residues that may be involved in aglycone binding, T334, M374, Y473 and A467 which are equivalent to T326, T364, F461 and E455, G455, and E455 in Dnbglu1, Dnbglu2 and Thai rosewood protein, respectively. So, the different amino acids residues at positions 454 and 455 may cause the difference of substrate specificity between these enzymes, especially in *D. nigrescens* and *D. cochinchinensis* β-glucosidase hydrolysis of natural glycosides, as showed on the TLC (Figure 3.11), although unfortunately the Dnbglu1 could not be expressed and examined. Also, they identified the amino acid residues that are critical for glucose binding, Q38, H142, E191, E406, E464 and W465, which are highly conserved in retaining glycosidases, and are found in both Dnbglu1 and Dnbglu2 protein sequences. In addition, the residue corresponding to W445, which is highly conserved among family 1, has been reported to make stacking interactions with the plane defined by the ring-like gluconate chain in *Bacillus polymyxa* β-glucosidase (Sanz-Aparicio *et al*., 1998). The interaction of the sugars by hydrogen bonds between the sugar hydroxyls and polar side chains at the binding site of proteins, and aromatic residues oriented against the hydrophobic faces of sugar has been noted to give binding affinity and specificity to the interaction. (Vyas, 1991).



A: Dnbglu1



B: Dnbglu2



C: Thai rosewood

Figure 4.2 Homology models of the active sites of Dnbglu1, Dnbglu2 and Thai rosewood β-glucosidases. The models illustrate some differences between the *Dalbergia* β-glucosidases. At residue 454, it is Asn in Dnbglu1, Ser in Dnbglu2 and Ala in Thai rosewood and it is Glu in Dnbglu1 and Thai rosewood at residue 455 but Gly in Dnbglu2 β-glucosidase.

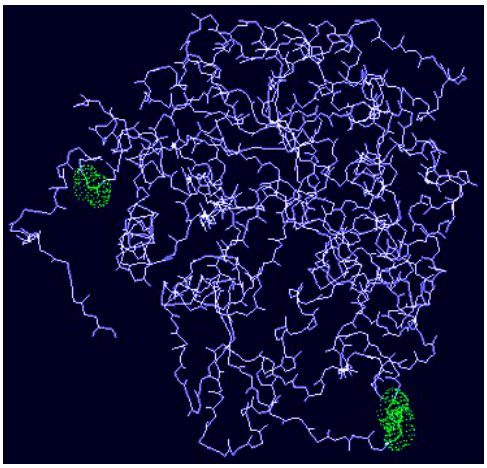
There are many reports that have revealed at least two isozymes in plant tissues, such as the cyanogenic β-glucosidases dhurrinase-Dhr1 and Dhr2 from sorghum (Cicek and Esen, 1998), Glu1 and Glu2, from maize (Bandaranayake and Esen, 1996) and 5 isozymes of Prunasin hydrolase (PH1 to PH5) from Black Cherry (Zhou *et al*., 2002). These isozymes may have different substrate specificities and binding affinities for their substrates.

4.6 Expression of *D. nigrescens* β-glucosidase in *Pichia pastoris*

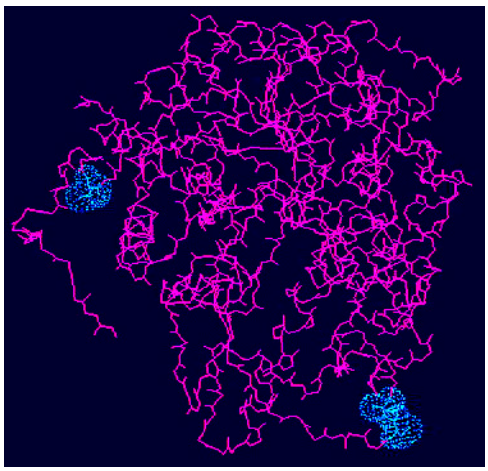
The *Dnbglu1* and *Dnbglu2* mature cDNA were cloned, sequenced and constructed into pPICZαB-Thrombin vectors in ordered to produce the *D. nigrescens* β-glucosidase protein in a recombinant *Pichia pastoris* system. Unlike bacteria, this system is capable of producing complex proteins with eukaryotic posttranslational modifications, for instance, correct folding, glycosylation, and proteolytic maturation (Sherma *et al*., 1996 and White *et al*., 1994). Only the Dnbglu2 protein could be expressed as an active protein in the culture media. The deduced amino acid sequences of *Dnbglu1* and *Dnbglu2* are very similar to each other (83% identity) and the deduced amino acid sequence of Dnbglu1 is closer to Thai rosewood β-glucosidase than is Dnbglu2 (85% vs. 81%, respectively). In addition, every critical amino acid residues among these three related β-glucosidases, such as the catalytic amino acid residues which are involved in cleavage and the substrate binding site are conserved with other family 1 β-glucosidases. So, it is unclear why Dnbglu1 could not be expressed in active form. However, there was at least one amino acid which

may be involved in the stabilization of the protein complex. As the results have shown that *D. nigrescens* β-glucosidase composed of 4 subunits, so cysteine residues may be important to stabilize each subunit and hold subunits together by disulfide bonding for proper function. Homology modeling showed that there were 3 cysteine residues (C21, C339 and C340) which are located at the surface of Thai rosewood and Dnbglu2 subunit. Dnbglu1 lacked C21, which is at the protein surface in Dnbglu2 and Thai rosewood protein. This may cause Dnglu1 to lack a disulfide bond between subunits in the tetrameric enzyme or to form incorrect disulfide bonding (Figure 4.3).

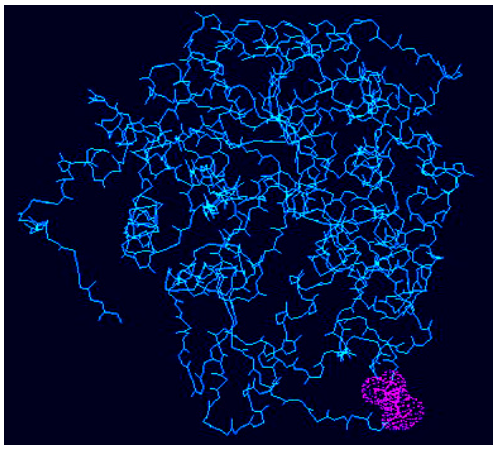
Also, examination of the active site of Dnbglu1 and Dnbglu2 from the homology modeling (Figure 4.2) found that there is a glutamic acid at residue 455 in Dnbglu1, but a less bulky glycine in Dnbglu2 at this residue, which is its in a position to affect the binding of the substrate moiety to the active site. But the active site of Thai rosewood also has a glutamic acid at this residue, which suggests it should not affect the activity of the enzyme, though it may affect the specificity for the substrate.



A. Thai rosewood



B. Dnbglu2



C. Dnbglu1

Figure 4.3 Illustration of the cysteine residues at the surface of *Dalbergia* β-glucosidases. The backbones of the proteins are show without sidechains with surface cysteine residues surrounded by dot surfaces.

Originally, there was one stop codon at W95X when the full-length Dnbglu1 cDNA was cloned, but it was changed to Trp by mutagenesis based on previous DNA sequence data, in which there was not any stop codon in the partial cDNA sequencing at this position. When compared with Thai rosewood β-glucosidase and other plant β-glucosidases, W95 was found to be one of the conserved residues in family 1. So, we presumed that it should be W95 in this position. So, it might come from the error of PCR step, or it may really occur at the RNA level, which would make it so this gene cannot produce functional protein. In addition, we have tried to mutate an aspartic to glycine residue (D127G) which glycine is a conserve residue among the most β-glucosidase to make an active Dnbglu1 protein. Because the aspartic amino acid residue has a larger side chain group which may disturb the structure, even though it was not included in the binding site region, as shown in Figure 3.40. However, it was found that this mutation could not turn the inactive Dnbglu1 to active protein like Dnbglu2.

Koganesawa *et al*. (2001) revealed the effect of selection of signal sequence on level of expression in the *Pichia pastoris* expression system. They found that silkworm lysozyme with the α-factor leader is so unstable that it could be easily attacked by some proteases. So, they suggested that the level of expression of heterologous protein with signal peptides and its stability are strongly affected by the selection of the appropriate signal sequence, because they showed that the secretion of silkworm lysozyme with its native signal was more efficient than the α-factor leader sequence in the *P. pastoris* system. Yeast α-factor is synthesized initially as a larger precursor, prepro- α-factor (Brake *et al*., 1983). During the translocation through the pathway, the propro- α-factor is digested by three kinds of peptidase, endopeptidase,

dipeptidyl aminopeptidase and carboxypeptidase. The Silk worm lysozyme molecules were incompletely processed at the C-terminus of the α-factor signal after secretion, because KEK2 protease cannot digest the recombinant protein entirely. The results clearly showed that silkworm lysozyme with some amino acids (EAEA sequence) added to its N-terminus becomes less stable than the wild type so the N-terminus seems to be critical to its conformation. However the production level of the yeast expression system is influenced not only by stability of the heterologously expressed protein. There are various factors, such as the frequency of transcription start, the "strength" of the promoter, the stability of mRNA, the toxicity of expressed protein and codon bias that should be considered (Dani *et al*., 1985 and Kane, 1995).

The post-translational processing might have an effect on the enzyme activity. Comparison of the cleavage site that was designed for cleaving the signal peptide, which was based on the signal peptide of Thai rosewood sequence before construction into the yeast expression vector with that predicted by the Signal P program indicated that the designed cleavage site that was used for Dnbglu2 was more close to the predicted site than that of Dnbglu1 (Figure 4.4). However, the sequenced N-terminus was later than these, possibly due to proteolytic processing after signal peptidase cleavage. Also, there are some differences in glycosylation sites between Dnbglu1 and Dnbglu2, which might have some effect.

In addition, the differences at C-terminus between the deduced amino acid sequences of Dnbglu1 and Dnbglu2 may have an effect. Dnbglu2 amino acid sequence has shorter a C-terminus compared with Dnbglu1, which may make Dnbglu1 inactive in some way.

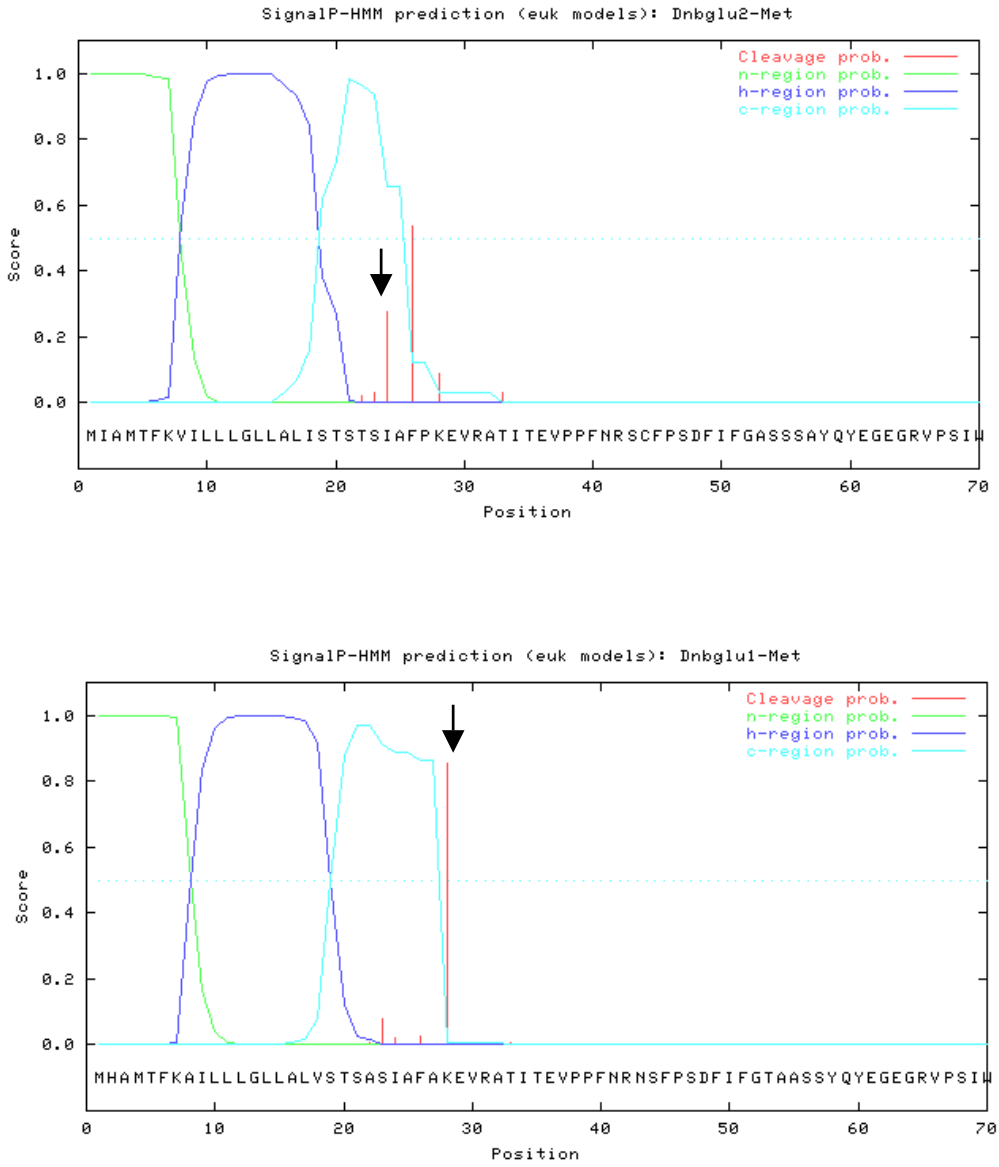


Figure 4.4 The prediction of signal peptides of Dnbglu1 and Dnbglu2 from the Signal P program. Vertical lines indicate the likely cleavage site position.

CHAPTER V

CONCLUSION

D. nigrescens β-glucosidase was purified from dry seeds of *Dalbergia nigrescens* Kurz by 35-75% ammonium sulfate fractionation, DEAE anion exchange chromatography and Sephacryl-300 size exclusion chromatography. The enzyme showed high activity toward both *p*NP-β-D-glucoside and *p*NP-β-D-fucoside, but the β-fucosidase activity was approximately 2 time higher than β-glucosidase activity throughout the purification steps. Determination of native molecular weight by Sephacryl-S300 gel filtration chromatography indicated it has a native molecular mass of approximately 240 kDa, and the SDS-PAGE showed a single protein band at about 62-63 kDa. Comparison of the native and denatured molecular weights indicated that this enzyme is likely composed of four subunits. The enzyme has pH and temperature optima of pH 5-5.5 and 65°C, respectively, in a 10 min assay. The enzyme cleaved *p*NP-β-D-xyloside, *p*NP-β-D-galactoside and *p*NP-α-L-arabinoside, though rather slowly compared with *p*NP-β-D-glucoside and *p*NP-β-D-fucoside, a property similar to that of *D. cochinchinensis* β-glucosidase from Thai rosewood. The enzyme could hydrolyze laminaribiose, cellobiose and linamarin slightly and it could hydrolyze dalcochinin 8'-O-β-D-glucoside, which is a natural substrate of *D. cochinchinensis* β-glucosidase. However, the Thai rosewood β-glucosidase could not hydrolyze the glycoside substrates of *D. nigrescens* β-glucosidase (S1 and S2) which were extracted

from dry *D. nigrescens* seeds very well.

The S1 and S2 glycosides were extracted and purified from *D. nigrescens* seeds. The structure of both substrates were solved by NMR and mass spectrometry and both were found to be isoflavonoid derivatives which were composed of an isoflavone and two sugar moieties, one five carbon ring and one six carbon ring. The NMR data, GC-MS and comparison to the literature indicated that the two sugars are apiose and glucose moieties. But, after hydrolysis with the *D. nigrescens* enzyme, glucose could not be detected by the PGO assay because the sugars were removed as disaccharide, so it was neccessary to characterize the kinetics of the enzyme toward these substrates by measurement of released aglycone by HPLC. This kinetic data indicated that S1 and S2 were the best substrates of the *D. nigrescens* enzymes in terms of K_m and k_{cat} values.

Two cDNA clones encoding *D. nigrescens* glycosidases, *Dnbglu1* and *Dnbglu2*, were amplified by polymerase chain reaction (PCR), based on homology with the coding sequence of *D. cochinchinensis* β-glucosidase. The proteins encoded by *Dnbglu1* and *Dnbglu2* matched peptide sequences from the purified protein and had 83% identity with each other and 85% and 81% identity, respectively, with *D. cochinchinensis* β-glucosidase. Thus, the family 1 *D. nigrescens* glycosidase purified here is likely to be a β-glucosidase, as judged by the similarities in amino acid sequence. The *Dnbglu1* and *Dnbglu2* cDNA were cloned into the pPICZαB expression vector and expressed in *Pichia pastoris*. Expression of *Dnbglu1* did not produce active β-glucosidase, expression of Dnbglu2 did. So, the Dnbglu2 recombinant protein was expressed, partially purified with S-300 gel filtration

chromatography and compared with the purified native *D. nigrescens* enzyme. The protein from both sources had similar enzymatic properties. The K_m values of the native enzyme and the recombinant Dnbglu2 were 14.7 and 12.1 mM for *p*NP-β-D-glucoside, and 1.8 and 1.4 mM for *p*NP-β-D-fucoside, respectively. Both the native and recombinant enzymes hydrolyzed two glycosides isolated from the seeds of *D. nigrescens* with the same K_m values of 0.5 mM for S1 and 0.7 mM for S2, while *D. cochinchinensis* β-glucosidase showed little activity toward these substrates. Differences of some amino acid residues between the *Dalbergia* β-glucosidases were identified which may have some effect toward enzyme activity and substrate specificity, which can account for differences in the catalytic properties.

REFERENCES

REFERENCES

Adinaraya, D. and Rajasekhara Rao, J. (1972). Isoflavonoid glycosides of *Dalbergia paniculata*. The constitutions of dalpanitin and dalpatin. **Tetrahedron.** 28: 5377-5384.

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

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

Bandaranayake, H. and Esen, A. (1996). Nucleotide sequence of a beta-glucosidase (glu2) cDNA from maize. **Plant Physiol**. 110: 1048.

Barrett, T., Suresh, C. G., Tolley, S. P., Dodson, E. J. and Hughes, M. A. (1995). The crystal structure of cyanogenic β-glucosidase from white clover (*Trifolium repens* L.) a family 1 glycosyl hydrolase. **Structure.** 3: 951-960.

Berrin, J. G., McLauchian W. R., Needs, P., Williamson, G., Puigserver, A., Kroon, P. A., and Juge, N. (2002). Functional expression of human liver cytosolic β-glucosidase *in Pichia pastoris*: Insights into its role in the metabolism of dietary glucosides. **European Journal of Biochemistry.** 269: 249-258.

Brake, A. J., Julius, D. J., and Thorner, J. (1983). A functional prepro-alpha-factor gene in Saccharomyces yeasts can contain three, four, or five repeats of the

mature pheromone sequence. **Molecular and Cellular Biology.** 3: 1440-1450.

Brzobohaty, B., Moore, L., Kristoffersen, P., Bako, L., Campos, N., Schell. J., and Palmet, K. (1993). Release of active cytokinin by a β-glucosidase localized to the maize root meristem. **Science.** 262: 1051-1053.

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

Campbell, J. A., Davies, G. J., Bulone, V., and Henrissat, B. (1997). A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino-acid similarities. **Biochemical Journal.** 15: 929-939.

Champattanachai, V. (1998). Thesis: Cloning and sequence analysis of β-glucosidase/β-fucosidase enzyme from Thai rosewood. Mahidol University.

Chiba, S. (1995). Enzyme Chemistry and Molecular Biology of Amylases and related enzymes (Yamamoto, T., ed). CRC Press, Inc. pp. 68.

Cicek, M. and Esen, A. (1999). Expression of soluble and catalytically active plant (monocot) beta-glucosidase in *E.coli*. **Biotechnology bioengineering.** 63: 392-400.

Cicek, M. and Esen, A. (1998). Structure and Expression of a Dhurrinase (β-glucosidase) from Sorghum. **Plant Physiology.** 116: 1469-1478.

Clarke, A. J., Bray, R. M., and Strating, H. (1993). β-glucosidases, β-glucanases and xylanases: Their mechanism of catalysis in *β-Glucosidases:Biochemistry and Molecular biology* (Esen, A., ed) pp.27-42, American Chemical Society, Washington, DC

Campos, N., Bako, L., Brzobohaty, B., Feldwisch, J., Zettl, R., Boland, W., and Palme K. Identification and characterization of a novel phytohormone conjugate specific β-glucosidase activity from maize:Esen A, editor. (1993). β-Glucosidase Biochemistry and Molecular Biology. Washington DC:American Chemical Society, pp. 205-213.

Combes, C. L. and Birch, G. G. (1988). Interaction of δ-glucono-1,5-lactone with water. **Food Chemistry.** 27: 283-298.

Czjzek, M., Cicek, M., Zamboni, V., Bevan, D. R., Henrissat, B. and Esen, A. (2000). The mechanism of substrate (aglycone) specificity in beta -glucosidases is revealed by crystal structures of mutant maize beta-glucosidase-DIMBOA, -DIMBOAGlc, and -dhurrin complexes. **Proceedings of the National Academy of Sciences, U S A.** 97: 13555-13560.

Dan, S., Marton, I., Dekel, M., Bravdo, B. A., He, S., Withers, S. G., and Shoseyov, O. (2000). Cloning, Expression, Characterization, and Nucleophile Identification of Family 3, *Aspergillus niger* β-Glucosidase. **Journal of Biological Chemistry.** 275: 4973-4980.

Dani, C. H., Mechti, N., Piechaczyk, M., Lebleu, B., Jean teur, P., and Blanchard, J. M. (1985). Increased rate of degradation of c-myc mRNA in interferon-

treated Daudi cells. **Proceedings of the National Academy of Sciences, U S A.** 82: 4896-9.

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

Davies, G. J., Wilson, K. S., and Henrissat, B. (1997). Nomenclature for sugar-binding subsites in glycosyl hydrolases. **Biochemical Journal.** 321: 557-559.

Dharmawardhana, D. P., Ellis, B. E., and Carlson, J. E. (1995). A β-glucosidase from lodgepole pine xylem specific for the lignin precursor coniferin. **Plant Physiology.** 107: 331-339.

Ellis, S. B., Brust, P. F., Koutz, P. J., Waters, A. F., Harpold, M. M., and Gingeras, T. R. (1985). Isolation of alcohol oxidase and two other methanol regulatable genes from the yeast, *Pichia pastoris*. **Molecular and Cellular Biology.** 5: 1111-1121.

Esen, A. (1992). Purification and Partial Characterization of Maize (Zea mays L.) β-glucosidase. **Plant Physiology.** 98: 174-182.

Esen, A. (1993). β–Glucosidase: Overview in *β-Glucosidases:Biochemistry and molecular biology* (Esen, A., ed) pp.1-26, American Chemical Society, Washington, DC

Esen, A. and Cokums, C. (1991). pH-and temperature-dependent β-glucosidase multiplicity in maize (*Zea mays* L.) is a proteolysis artifact. **Plant Science.** 74: 17-26.

Estruch, J. J., Chriqui, D., Grossmann, K., Schell, J., and Spena, A. (1991a). The plant oncogene *rolC* is responsible for the release of cytokinin from glucoside conjugates. **Journal of the European Molecular Biology Organization.** 10: 2889-2895.

Farag, S. F., Ahmed, A. S., Terashima, K., Takaya, Y., Niwa, M. (1999). Isoflavonoid glycosides from *Dalbergia sissoo*. **Phytochemistry.** 50: 739-743.

Frohman, M. A. (1993). Rapid amplification of complementary DNA ends from generation of full-length complementary DNAs: Thermal RACE. **Methods in Enzymology**. 218: 340-356.

Gopalan, V., Pastuszyn, A. P., Galey, W. R. and Glew, R. H. (1992*)*. Exolytic hydrolysis of toxic plant glucosides by guinea pig liver cytosolic beta-glucosidase. **Journal of Biological Chemistry.** 267: 14027-14032.

Grover, A. K. and Cushley, R. J. (1977). Studies on almond emulsion β-D-glucosidase:II. Kinetic evidence for independent glucosidase and and galctosidase sites. **Biochimica et Biophysica Acta.** 482: 109-124.

Harborne, J. B., Mabry, J. C., and Mabry, H. (1975). The flavonoids. Chapman and Hall, London.

Halliwell, B. (1994). Free radicals, antioxidants, and human disease: curiosity, cause, or consequence?. **Lancet.** 344: 721-724.

Henrissat, B. (1991). A classification of glycoside hydrolases based on amino acid sequence similarities. **Biochemical Journal.** 280: 309-316.

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

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

Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K., and Pease, L. R. (1989). Engineering hybrid genes without the use of restriction enzyme. **Gene.** 77: 61-68.

Hösel, W. and Conn, E. E. (1982). The aglycone specificity of plant β-glucosidase. **Trends in Biochemical Sciences.** 6: 219-221.

Hösel, W., Surholt, E., and Borgmann, E. (1978). Characterization of beta-glucosidase isoenzymes possibly involved in lignification from chick pea (*Cicer arietinum* L.) cell suspension cultures. **European Journal of Biochemistry.** 84: 487-492.

Hösel, W., Tober, I., Eklund, S. H., and Conn, E. E. (1987). Characterization of β-glucosidase with high specificity for the cyanogenic glucoside dhurrin in sorghum bicolor (L.) Moench seedling. **Archives of Biochemistry and Biophysica.** 252: 152-162.

Hösel, W., Glycosylation and glycosidases. In: Conn, E. E., editor. The Biochemistry of Plants: A Comprehensive Treatise, vol. 7: Secondary plant products. New York: Academic Press; 1981. pp.725-753.

Hösel, W. and Barz, W. (1975). β-glucosidases from *Cicer arietinum* L.: purificaton and properties of isoflavone-7-*O*-glucoside-specific β-glucosidases. **European Journal of Biochemistry.** 57: 607-616.

Holm, L. and Sander, C. (1994). Structural similarities of plant chitinase and lysozyme from animals and phage. **Federation of European Biochemical Societies Letter.** 340: 129-132.

Hollman, P. C. H. and Katan, M. B. (1999). Dietary Flavonoids: Intake, Health Effects and Bioavailability. **Food and Chemical Toxinology.** 37: 937-942.

Hrmova, M. and Fincher, G. B. (1998). Barley β-D-glucan exohydrolases, substrate specificity and kinetic properties. **Carbohydrate Research**. 305: 209-221.

Kandaswami, C., and Middleton, E., Jr (1994). Free radical scarvenging and antioxidant activity of plant flavonoids. **Advance in Experimental Biological medicine.** 366: 351-376.

Kane, J. F. (1995). Effects of rare codon cluster on high-level expression of heterologous proteins in *Escherichia coli*. **Current Opinion in Biotechnology.** 6: 494-500.

Ketudat Cairns, J. R., Champattanachai, V., Srisomsap, C., Liebold, B. W., Thiede, B., and Svasti, J. (2000). Sequence and Expression of Thai Rosewood β-glucosidase/β-Fucosidase, a Family 1 Glycosyl Hydrolase Glycoprotein. **Journal of Biochemistry.** 128: 999-1008.

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. **Biochemical Journal.** 353: 199-205.

Keresztessy, Z., Kiss, L., and Hughes, M. A. (1994). Investigation of the active site of the cyanogenic β-D-glucosidase (linamarase) from *Manihot esculenta* Crantz (cassava). I. Evidence for an essential carboxylate and a reactive histidine residue in a single catalytic center. **Archives of Biochemistry and Biophysics.** 314: 142-152.

Keresztessy, Z., Kiss, L., and Hughes, M.A. (1994). Investigation of the active site of the cyanogenic β-D-glucosidase (linamarase) from *Manihot esculenta* Crantz (cassava). II. Identification of Glu198 as an active site carboxylate group with acid catalytic function. **Archives of Biochemistry and Biophysics.** 315: 323-330.

Koganesawa, N., Aizawa, T., Masaki, K., Matsuura, A., Nimori, T, Bando, H., Kawano, K., and Nitta, K. (2001). Construction of an expression system of insect lysozyme lacking thermal stability: the effect of selection of signal sequence on level of expression system. **Protein Engineering.** 14: 705-710.

Konno, H., Yamasaki, Y., and Katoh, K. (1996). A β-glucosidase associated with cell walls from cell suspension cultures of carrot. **Phytochemistry.** 43: 1157-1161.

Koutz, P. J., Davis, G. R., Stillman, C., Barringer, K., Cregg, J. M., and Thill, G. (1989). Structural comparison of the *Pichia pastoris* alcohol oxidase genes. **Yeast.** 5**:** 167-177.

Lai, E. C. K., Morris, S. A., Street, I .P., and Withers, S.G. (1996). Substituted glycols as probes of glycosidase mechanism. **Bioorganic and Medicinal Chemistry.** 4: 1929-1937.

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

LaVallie, E. R., DiBrasio, E. A., Kovasic, S., Grant, K. L., Schendel, P. F., and McCoy, J.M. (1993). A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. **Biotechnology.** 11: 187-193.

Lineweaver, H., Burk, D. (1934). The determination of enzyme dissociation constants. **Journal of the American Chemical Society.** 56: 658-666.

Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951). Protein measurement with the Folin phenol reagent. **Journal of Biological Chemistry.** 193: 265-275.

Ly, H. D. and Withers, S. G. (1999). Mutagenesis of glycosidases. **Annual Review of Biochemistry.** 68: 487-522.

Middleton E. and Kandaswami C. (1994). The impact of plant flavonoids on mammalian biology: implications for immunity, inflammation and cancer. In *The Flavonoids: Advances in research since 1986*, ed. J. B. Harborne, 1st edn, pp. 619-652. Chapman & Hall, London.

Mathias, L., Vieira, I.J.C., Braz-Filho, R., and Rodrigues- Filho, E. (1998). A new isoflavone glycoside from *Dalbergia nigra*. **Journal of Natural Products.** 61: 1158-1161.

Moffatt, B.A. and Studier, F.W. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high level expression of cloned genes. **Journal of Molecular Biology.** 189: 113–130.

Neufeld, E. F. (1991). Lysosomal storage diseases. **Annual Review of Biochemistry.** 60: 257-280.

Osbourn, A. (1996). Saponins and plant defence a soap story. **Trends in Plant Science.** 1: 4-9.

Perrakis, A. and Vorgias, C. E. (1994). Crystal structure of a bacteria chitinase at 2.3 Å solution. **Structure.** 2: 1169-1180.

Radhakrishniah, M. (1973). Isoflavonoids of *Dalbergia paniculata* seeds. **Phytochemistry.** 12: 3003-3004.

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*. **Applied and Environmental Microbiology.** 64: 3607-3614.

Rosenberg, A. H., Lade, B. N., Chui, D., Lin, S., Dunn, J. J., and Studier, F. W. (1987). Vectors for selective expression of cloned DNAs by T7 RNA polymerase. **Gene.** 56: 125–135.

Sanz-Aparicio, J., Hermoso, J. A., Martinez-Ripoll, M., Lequerica, J. L., and Polaina, J. (1998). Crystal structure of β-glucosidase A from *Bacillus polymyxa*: Insights into the catalytic activity in family 1 glycosyl hydrolases. **Journal of Molecular Biology.** 275: 491-502.

Selma, D., Lieberei, B., Biehl, B., and Voigt, J. (1987). *Hevea* linamarase-α nonspecific β-glycosidase. **Plant Physiology.** 83: 557-563.

Sherna, G., Cappai, R., Henry, A., and Small, D. H. (1996). Advantages of the methylotrophic yeast *Pichia pastoris* for high level expression and purification of heterologous protein. **Australasian Biotechnology.** 6: 282-287.

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. **Journal of Biochemistry.** 119: 585-590.

Stewart, E. J., Aslund, F., and Beckwith, J. (1998). Disulfide bond formation in the *Escherichia coli* cytoplasm: an *in vivo* role reversal for the thioredoxins. **Journal of the European Molecular Biology Organization.** 17: 5543-5550.

Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. **Methods in Enzymology.** 185: 60–89.

Sue, M., Ishihara, A. and Iwamura, H. (2000). Purification and Characterization of hydroxamic acid glucoside β-glucosidase from wheat (*Triticum aestivum* L.) seedlings. **Planta.** 210: 432-438.

Surarit, R., Svasti, J., Srisomsap, C., Suginta, W., Khunyosyheng, S., Nilwarangkoon, S., Harnsakol, P., and Benjavongkulchai, E. (1995). Screening of glycosyl

hydrolase enzymes in Thai plant seeds for opotential use in oligosaccharide synthesis. **Journal of Science Society of Thailand.** 21: 293-303.

Surarit, R., Chiaba, S., Matsui, H., Svasti, J., and Srisomsap, C. (1996). Chemical modification of β-glucosidase/β-fucosidase from *Dalbergia cochinchinensis* Pierre by conduriol B epoxide. **Bioscience, Biotechnology, and Biochemistry.** 60: 1265-1268.

Teeri, T. T. (1997). Crystalline cellulose degradation new insight into the function of cellobiohydrolases. **Trends in Biotechnology.** 15: 160-167.

THAI FOREST BULLETIN (BOTANY) No.30. (2002).

Tahara, S., Moriyama, M., Ingham, J. L., and Mzutani, J. (1993). Diverse oxygenated isoflavonoids from *Piscida erythrina*. **Phytochemistry.** 34: 303-315.

Takahashi, H., Hirata, S., Minami, H., and Fukuyama, Y. (2001). Triterpene and flavanone glycoside from *Rhododendron simsii*. **Phytochemistry**. 56: 875-879.

Trimbur, D. E., Warren, R. A. J., and Withers, S. G. (1992). Region-directed mutagenesis of residues surrounding the active site nucleophile in β-glucosidase from *Agrobacterium feacalis*. **Journal of Biological Chemistry.** 267: 10248-10251.

Tschopp, J. F., Brust, P. F., Cregg, J. M., Stillman, C., and Gingeras, T. R. (1987). Expression of the *lacZ* gene from two methanol regulated promotors in *Pichia pastoris.* **Nucleic Acids Research.** 15: 3859-3876.

Vallmitjana, M., Ferrer-Nvarro, M., Planell, R., Abel, M., Ausin, C., Querol, E., Planas, A., and Pons, P. (2001). Mechanism of the family 1 β-Glucosidase from Streptomyces sp: Catalytic Residues and Kinetic Studies. **Biochemistry.** 40: 5975-5982.

Varghese, J. N., Garret, T. P. J., Colman, P. M., Chen, L., Hoj P. B., and Fincher, G. B. (1994). The three-dimensional structures of two plant-β-glycan endohydrolases with distinct substrate specificities. **Proceedings of the National Academy of Sciences, U S A.** 91: 2785-2789.

Vevodova, J., Marek, J., Zouhar, J., Brzobohaty, B., and Su, X-D. (2001). Purification, crystallization and preliminary X-ray analysis of a maize cytokinin glucoside specific β-glucosidase. **Acta Crystallographica Section D**. 57: 140-142.

Vyas, N. K. (1991). Atomic features of protein-carbohydrate interactions. **Current Opinion in Structural Biology.** 1: 732-740.

White, C. E., Kempi, N. M., and Komives, E. A. (1994). Expression of highly disulfide-bonded proteins in *Pichia pastoris*. **Structure.** 2: 1003-1005.

Wiesmann, C., Beste, G., Hengstenberg, W., and Schulz, G. E. (1995). The three dimensional structure of 6-phospho-β-galactosidase from *Lactococcus* lactis. **Structure.** 3: 961-968.

Withers, S. G., Street, I. P., Bird, P., and Dolphin, D. H. (1987). 2-deoxy-2-fluoroglucosides: A novel class of mechanism-based glucosidase inhibitors. **Journal of the American Chemical Society.** 109: 7530-7531.

Withers, S. G. and Street, I. P. (1988). Identification of a covalenta-D-glucopyranosylenzyme intermediate formed on a β-glucosidase. **Journal of the American Chemical Society.** 110: 8551-8553.

Withers, S. G., Warren, R. A. J., Street, I. P., Repitz, K., Kempton, J. B., and Aebersold, R. (1990). Unequivocal demonstration of the involvement of a glutamate residue as a nucleophile in the mechanism of a "retaining" glycosidase. **Journal of the American Chemical Society.** 112: 5887-5889.

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

Withers, S. G., Rupitz, K., Trimber, D., and Warren, R. A. J. (1992). Mechanistic consequences of mutation of the active site nucleophile Glu-358 in *Agrobacterium* β-glucosidase. **Biochemistry.** 31: 9970-9985.

Wang, Q., Trimber, D., Graham, R., Warren, R. A. J., and Withers, S.G. (1995). Identification of the acid/base catalyst in *Agrobacterium feacalis* β-glucosidase by kinetic analysis of mutants. **Biochemistry.** 34: 14554-14562.

Yang, C. R. and Tanaka, O. (1999). Studies in Plant Science, Vol. 6; Advances in Plant Glycosides, Chemistry and Biology. **Phytochemistry.** 52: 1177-1180.

Yeoh, H-H. and Woo, H-C. (1992). Cassava leaf β-glucosidase specificity and inhibition. **Phytochemistry.** 31: 2263-2265.

Zechel, D. L. and Withers, S. G. (2000). Glycosidase mechanism: anatomy of finely tuned catalyst. **Accounts of Chemical Research.** 33: 11-18.

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

สรศักดิ์ เหลี่ยวไชยพันธุ์. บทนำ. ใน: ตำราเภสัชเวช พฤกษธาตุ:ไกลโคไซด์. กรุงเทพมหานคร: โอเดียนสโตร์; 2531. หน้า 1-20.

APPENDICES

APPENDIX A

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 1,000 mL distilled water. Autoclave the solution at 121°C for 20 min. Allow the medium to cool 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 10 g Bacto agar in 1000 mL distilled water. Sterilize by autoclaving at 121°C for 20 min. Allow the medium to cool to 50°C, then add 1 mL of ampicillin (100 mg/mL) 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 plates with 50 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 1,000 mL distilled water. Sterilize by autoclaving at 121°C for 20 min.

Allow the medium to cool to 50°C, then add 0.5 mL of ampicillin (100 mg/mL) to final concentration 50 μg/ml, 1 mL of kanamycin (15mg/mL) to final concentration 15 μg/ml, 1 mL of tetracyclin (12.5 mg/mL) to final concentration 12.5 μg/mL. 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 100 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 (100 mg/mL):

Dissolve 100 mg ampicillin in 1 mL sterile distilled water.

Kanamycin (15 mg/mL):

Dissolve 15 mg kanamycin in 1 mL sterile distilled water.

Tetracyclin (12.5 mg/mL):

Dissolve 12.5 mg tetracyclin in 1 mL absolute ethanol.

Filter sterile all antibiotic solutions 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 1 mL *N, N*-dimethylformamide (DMF) and store in the dark bottle at -20°C.

2. Reagents for competent cell preparation

2.1 0.1 M $CaCl_2$ Solution (100 mL)

Dissolve 1.47 g $CaCl_2$ in 100 mL water and filter sterile. Store the solution at 4°C.

3. Reagents for isolation plasmid DNA

3.1 STET (8% sucrose, 50 mM Tris pH 8.0, 50 mM EDTA, 0.5% Triton X-100) (100 mL)

Dissolve 8 g sucrose, 1.86 g EDTA and 0.157 g Tris in 99.5 mL distilled water. Autoclave at 121°C for 20 min. Add 0.5 g Triton X-100. Store at room temperature.

3.2 Glucose/Tris/EDTA (50 mM glucose/25 mM Tris/10 mM EDTA) (100 mL)

Mix 5 mL 1 M glucose, 2.5 mL 1 M Tris pH 8.0, 2 mL 0.5 M EDTA pH 8.0 and adjust to 100 mL with distilled water. Autoclave at 121°C for 20 min. Store at 4°C.

3.3 Solution III (3 M K/5 M acetate solution) (100 mL)

Dissolve 29.442 g KOAc in 11.5 mL glacial acetic and 88.5 mL in distilled water. Autoclave at 121°C for 20 min. Store at 4°C.

3.4 TE (10 mM Tris, 1 mM EDTA) (100 mL)

Mix 1 mL 1 M Tris pH 8.0 and 0.2 mL 0.5 M EDTA pH 8.0, adjust to 100 mL with sterilized distilled water.

3.5 RNase A (10 mg/mL)

Dissolve 10 mg RNase A in 10 mM Tris-HCl, pH 7.4, 15 mM NaCl buffer (steriled). Boil for 10 min and store at -20°C.

3.6 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, adjust the volume to 100 mL with distilled water and autoclave at 121°C for 20 min. Store at room temperature.

3.7 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 and the volume to 100 mL with distilled water. Sterilize by autoclaving at 121°C for 20 min. Store at room temperature.

4. Reagents 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 5X loading dye (0.15% bromophenol blue, 60% sucrose in TE, pH 8.0) (10 mL)

Mix 0.015 g Bromophenol blue, 6 g sucrose in distilled water to a 10 mL final volume and store at 4°C

5. Reagents for total RNA isolation by using Trizol reagent

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 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.

6. Solutions for protein

6.1 SDS-gel loading buffer (5X stock) 0.5 M Tris-HCl buffer, pH 6.8, 10% glycerol (v/v), 2% SDS, 0.5% 2-mercaptoethanol and 0.025% bromophenol blue (10 mL)

Mixed 1.2 mL 0.5 M Tris-HCl, pH 6.8, 2 ml 10% SDS, 1 mL glycerol, 0.0025 g bromophenol blue and adjust to 10 mL with distilled water. Store at 4°C. Add 0.050 mL of 2-mercaptoethanol (2-ME) to 0.95 mL of stock solution before use.

6.2 1.5 M Tris pH 8.8 (100 mL)

Dissolve 18.17 g Tris base in 80 mL distilled water. Adjust the pH to 8.0 with HCl and the volume to 100 mL with distilled water. Store at room temperature.

6.3 0.5 M Tris pH 6.8 (100 mL)

Dissolve 6.06 g Tris base in 80 mL distilled water. Adjust the pH to 6.8 with HCl and the volume to 100 mL with distilled water. Store at room temperature.

6.4 30% Acrylamide solution (100 ml)

Dissolve 29.2 g acrylamide and 0.8 g *N,N′*-methylene-bis-acrylamide in distilled water to a volume 100 mL. Mix the solution by stirring for 1 hr to be homogeneous and filter through Whatman No. 1 filter paper. Store in a dark bottle at 4°C.

6.5 Tris-Glycine electrode buffer (1 X stock) (1 L)

Dissolve 3 g Tris base, 14 g glycine, 10 mL 10% SDS in distilled water. Adjust the pH to 8.3 with HCl and the volume to 1 L with distilled water.

6.6 Coomassie brilliant R-250 Staining solution (1L)

Mix 0.25 g Coomassie brilliant blue R-250 with 400 mL methanol, Stir until dissolved; while stirring, gradually add 70 mL glacial acetic acid**,** then adjust volume to 1 L with distilled water**.** Store at room temperature**.**

6.7 Destaining solution for Coomassie Stain (1L)

Mix 400 mL methanol, 70 mL glacial acetic acid, and add distilled water to a final volume of 1 L. Store at room temperature.

6.8 10% (w/v) Ammonium persulfate (1 mL)

Dissolve 100 mg ammonium persulfate in 1 mL distilled water. Store at -20°C.

6.9 10% Separating gel SDS-PAGE (10 mL)

Mix the solution as follows:

1.5 M Tris (pH 8.8)	2.50 mL
10% (w/v) SDS	0.10 mL
30% acrylamide solution	3.32 mL
10% (w/v) ammonium persulfate	0.10 mL
TEMED	0.004 mL

Distilled water	3.96	mL

Add TEMED last, mix solution well and immediately pour the gel.

6.10 5% Stacking gel SDS-PAGE (5 mL)

Mix the solution as follows:

0.5 M Tris (pH 6.8)	1.26	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
Distilled water	2.77	mL

Add TEMED last, mix solution well and immediately pour the gel.

6.11 7% Separating gel (Native-PAGE) (10 mL)

Mix the solution as follows:

1.5 M Tris (pH 8.8)	2.50	mL
30% acrylamide solution	2.32	mL
10% (w/v) ammonium persulfate	0.10	mL
TEMED	0.004	mL
Distilled water	5.07	mL

Add TEMED last, mix solution well and immediately pour the gel.

6.12 3% Stacking gel (Native-PAGE) (5 mL)

Mix the solution as follows:

0.5 M Tris (pH 6.8)	1.26 mL
30% acrylamide solution	0.49 mL
10% (w/v) ammonium persulfate	0.05 mL
TEMED	0.005 mL
Distilled water	3.20 mL

Add TEMED last, mix solution well and immediately pour the gel.

7. Buffers and Reagents for enzyme assay

7.1 10 mM *p*-Nitrophenol (10 ml)

Dissolve 0.0139 g *p*-nitrophenol in 100 mM NaOAc pH 5.0 buffer and make to 10 mL final volume.

7.2 100 mM NaOAc pH 5.0 buffer (500 mL)

Dissolve 4.10 g Sodium acetate in 450 mL distilled water, and adjust pH to 5.0 with glacial acetic acid and the volume to 500 mL with distilled water.

7.3 2 M Na_2CO_3 (100 mL)

Dissolve 21.198 g Na_2CO_3 in distilled water, and adjust the volume to 100 mL with distilled water.

8. Solutions for protein purification

8.1 Buffer 1 (500 mL)

Dissolve 1.51 g Tris, 218 g $(NH_4)_2SO_4$ in 350 mL distilled water, then adjust to pH 8.0 with HCl and added distilled water to 490 mL final volume. Add 5 mL 0.1 M PMSF, 5 mL 1 M ascorbic acid and 0.97 g β-mercaptothanol just before use.

8.2 Buffer 2 (200 mL)

Dissolve 0.605 g Tris and 2.33 g NaCl in 200 mL distilled water, then adjust to pH 8.0 with HCl and added water to 196 mL final volume. Add 2 mL 0.1 M PMSF, 2 mL 1 M ascorbic acid and 0.39 g β-mercaptothanol just before use.

8.3 0.1 M PMSF

Dissolve 0.0871 g PMSF in 5 mL isopropanol. Freshly prepare.

8.4 1 M Ascorbic acid

Dissolve 1.76 g ascorbic acid in 5 mL distilled water. Freshly prepare.

8.5 0.025 M Tris-HCl pH 7.0 (1,500 mL)

Dissolve 3.53 g Tris in 1,400 mL distilled water and adjust to pH 7.0 with HCl and add distilled water up to 1,500 mL.

8.6 0.025 M Tris-HCl pH 7.0 + 0.5 M NaCl (500 mL)

Dissolve 1.51 g Tris and 14.61 g NaCl in 400 mL distilled water and adjust to pH 7.0 with HCl and add distilled water up to 500 mL.

8.7 0.050 M Tris-HCl pH 7.0 + 0.3 M NaCl (1,000 mL)

Dissolve 6.057 g Tris and 17.53 g NaCl in 900 mL distilled water and adjust to pH 7.0 with HCl and the volume to 1,000 mL with distilled water.

APPENDIX B

PLASMID MAPS

pET-32a(+) sequence landmarks	
T7 promoter	764-780
T7 transcription start	763
Trx•Tag coding sequence	366-692
His•Tag coding sequence	327-344
S•Tag coding sequence	249-293
Multiple cloning sites	
(*Nco* I - *Xho* I)	158-217
His•Tag coding sequence	140-157
T7 terminator	26-72
lacI coding sequence	1171-2250
pBR322 origin	3684
bla coding sequence	4445-5302
f1 origin	5434-5889

The maps for pET-32b(+) and pET-32c(+) are the same as pET-32a(+) (shown) with the following exceptions: pET-32b(+) is a 5899bp plasmid; subtract 1bp from each site beyond *Bam*H I at 198. pET-32c(+) is a 5901bp plasmid; add 1bp to each site beyond *Bam*H I at 198 except for *Eco*R V, which cuts at 209.

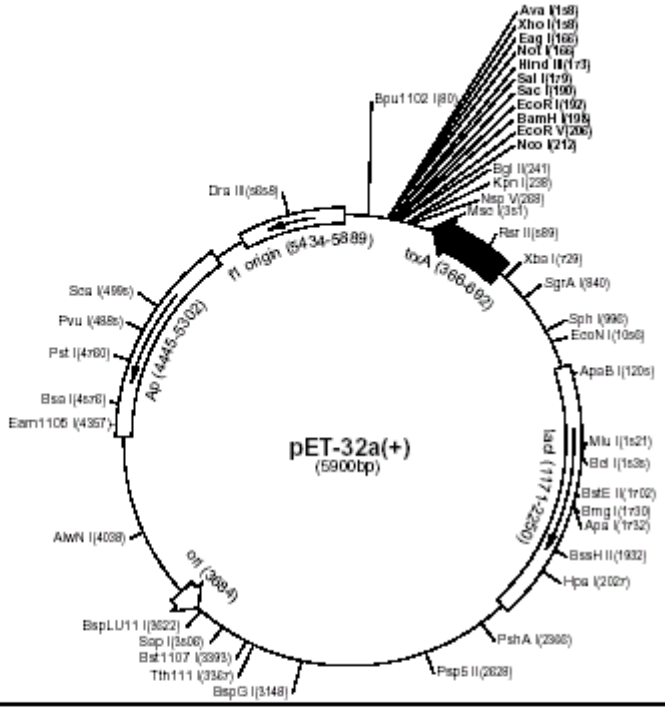


Figure 1 The pET-32a(+) vector map

(Novagen, http://www.emdbiosciences.com/docs/docs/PROT/TB122.pdf).

pET-40b(+) sequence landmarks	
T7 promoter	1187-1203
T7 transcription start	1186
DsbC•Tag™ coding seq	408-1115
His•Tag® coding sequence	369-386
S•Tag™ coding sequence	282-326
Multiple cloning sites	
(*Srf*I - *Xho* I)	174-250
His•Tag coding sequence	150-173
T7 terminator	26-72
lacI coding sequence	1594-2673
pBR322 origin	4107
Kan coding sequence	4816-5628
f1 origin	5724-6178

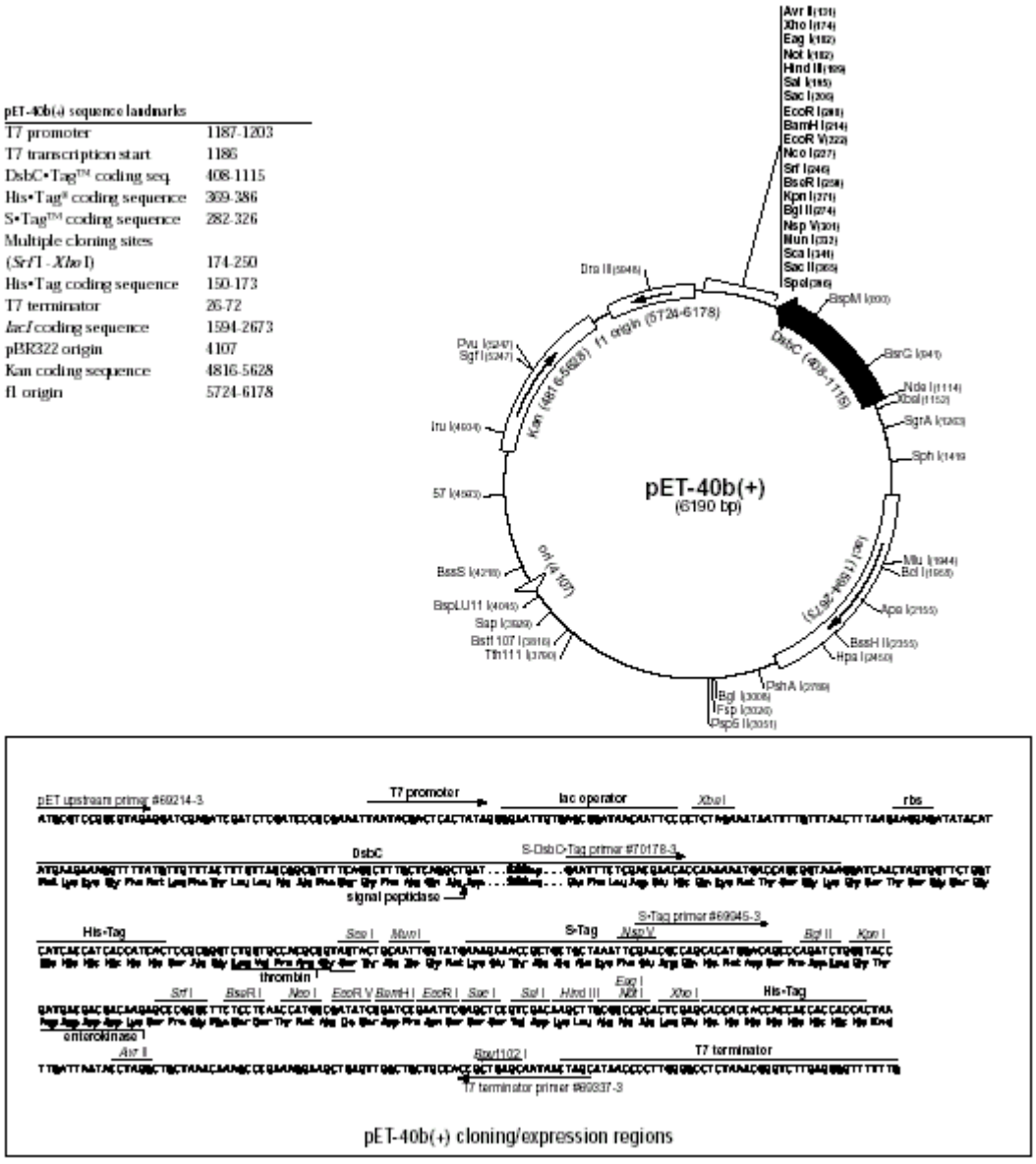


Figure 2 The pET-40b(+) vector map

(Novagen, http://www.emdbiosciences.com/docs/docs/PROT/TB229.pdf).

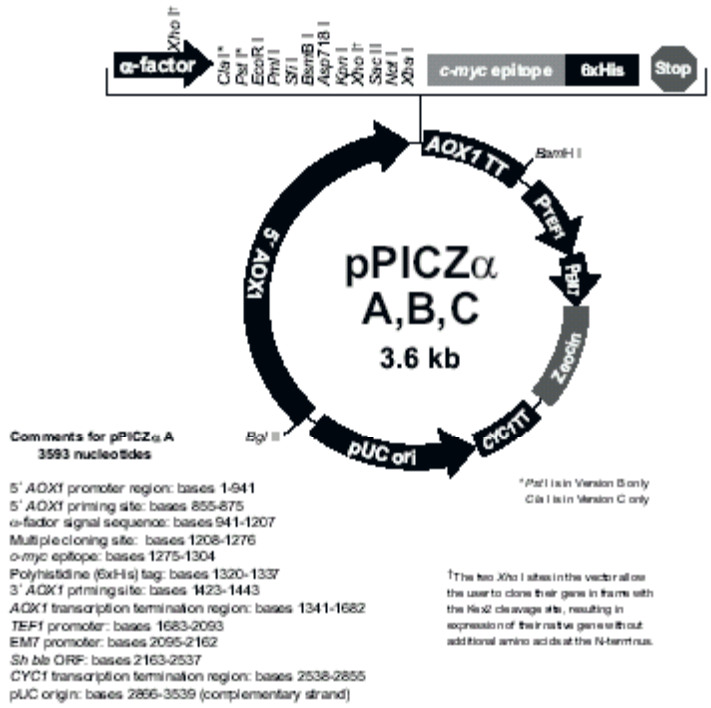


Figure 3 The pPICZαB map vector

(https://www.invitrogen.com/content/sfs/manuals/easyselect_man.pdf).

APPENDIX C

STANDARD CURVES AND REACTION TIME COURSE

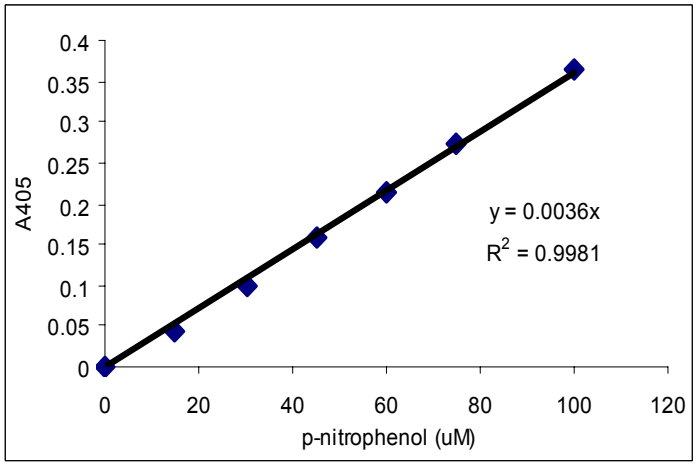


Figure 1 Standard curve of *p*-nitrophenol.

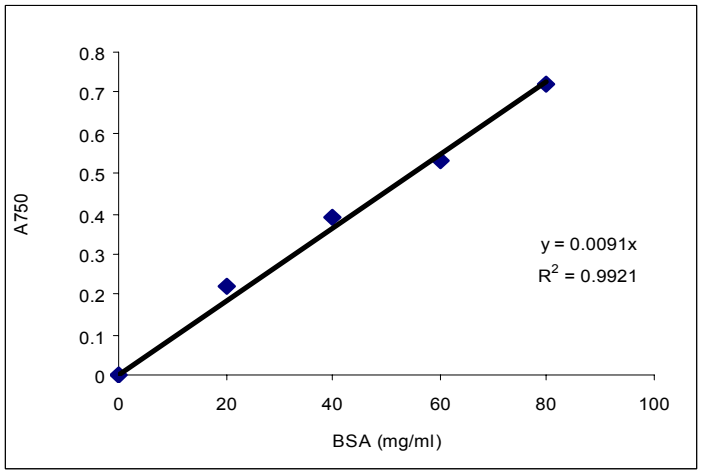


Figure 2 Standard curve of protein assay with BSA.

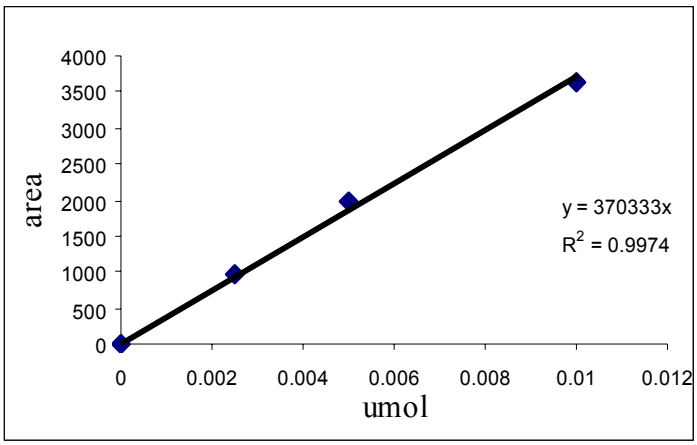


Figure 3 Standard curve of S1 aglycone peak area in HPLC.

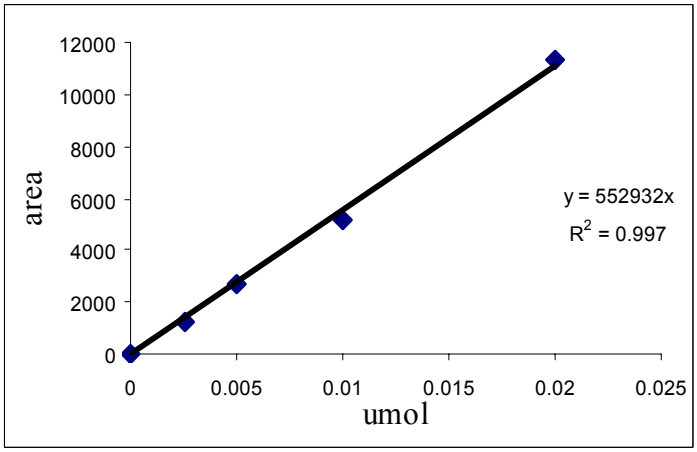


Figure 4 Standard curve of S2 aglycone peak area in HPLC.

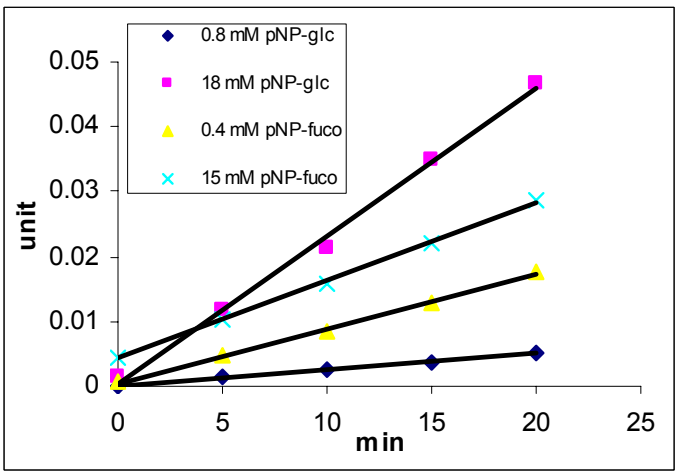
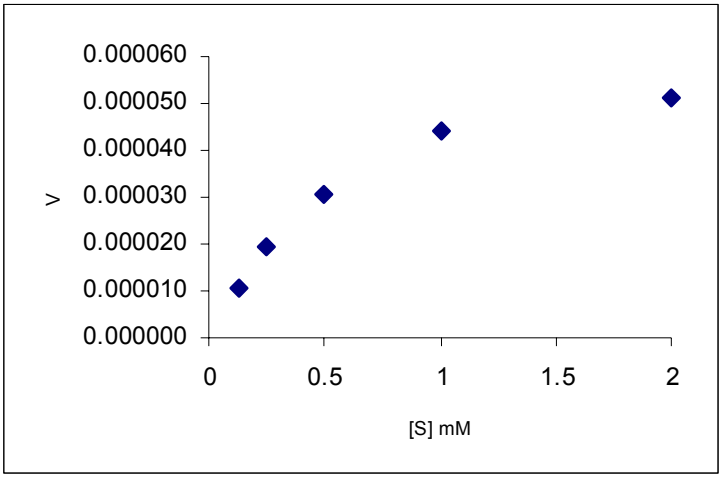
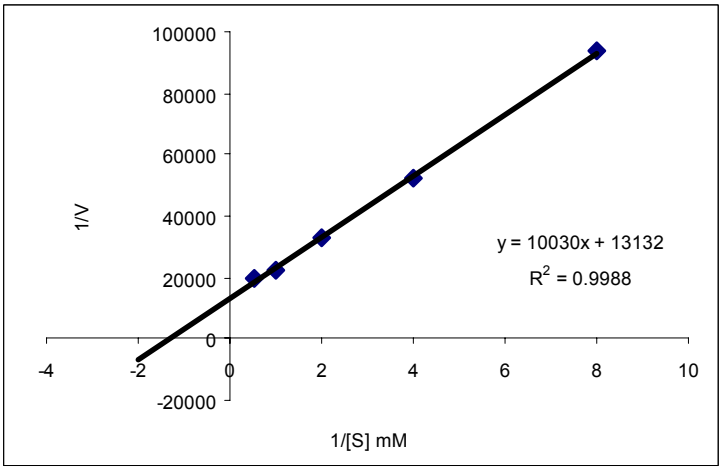


Figure 5 The time course of *p*NP-glycoside and *p*NP-fucoside substrates.



A



B

Figure 6 Example of curves for the determining K_m of S2 glycoside substrate of *D. nigrescens* β-glucosidase. **A**: Plotting V vs. [S] mM ; **B**: Plotting 1/V vs. 1/[S] mM.

APPENDIX D

DNA SEQUENCING CHROMATOGRAM

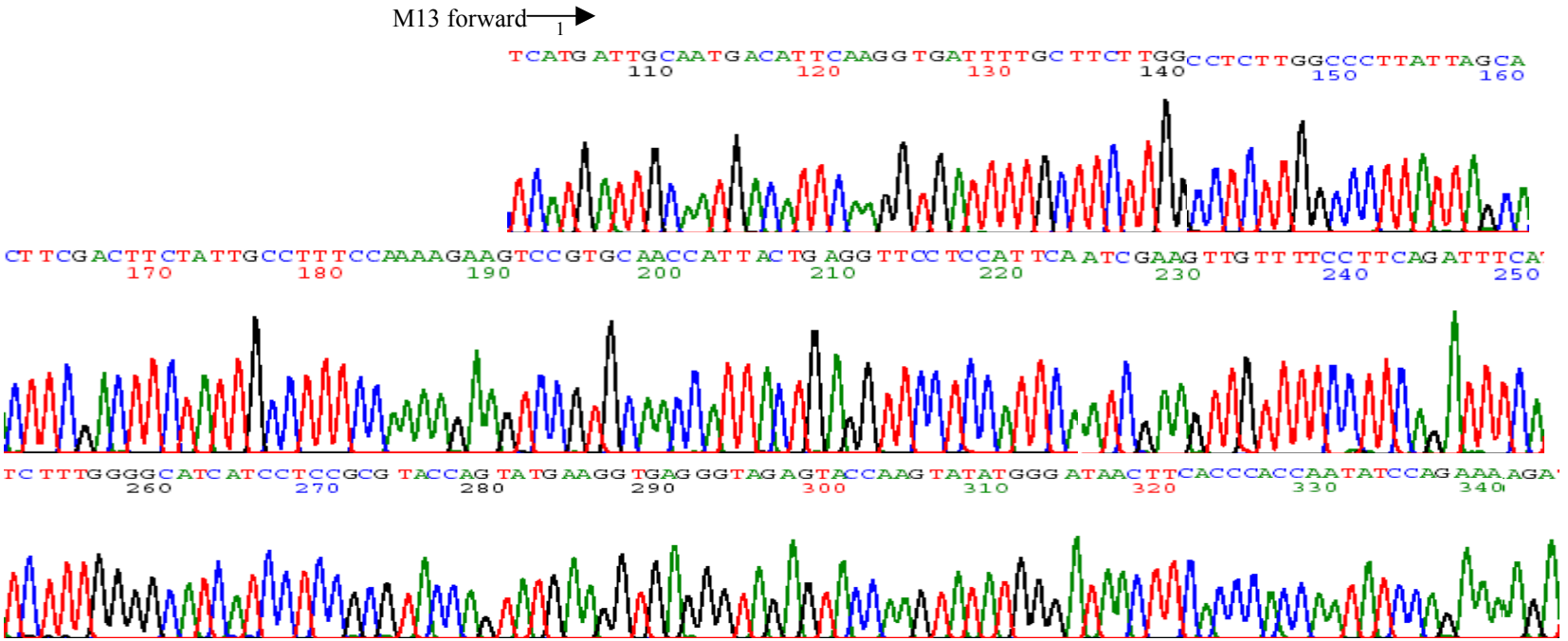


Figure 1 The chromatogram of cDNA sequence of *Dnbglu2*, using M13 forward, Dnbglu2_rev.4, PDnbglu2_for1 and Dnbglu2_for.4 as sequencing primers. The numbers labeled on the sequence are based on the cDNA sequence of *Dnbglu2* in Figure 3.38

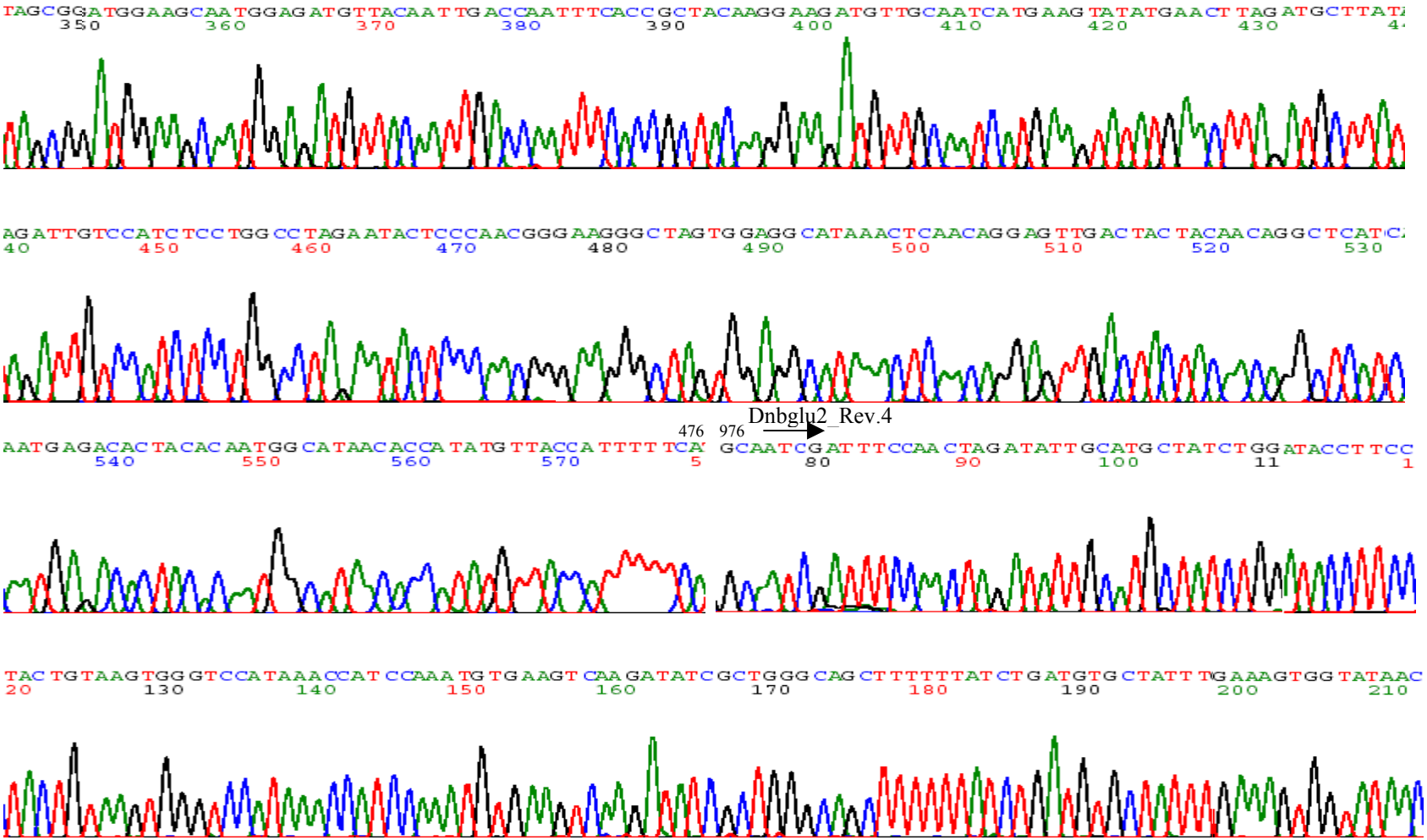


Figure 1 Continued

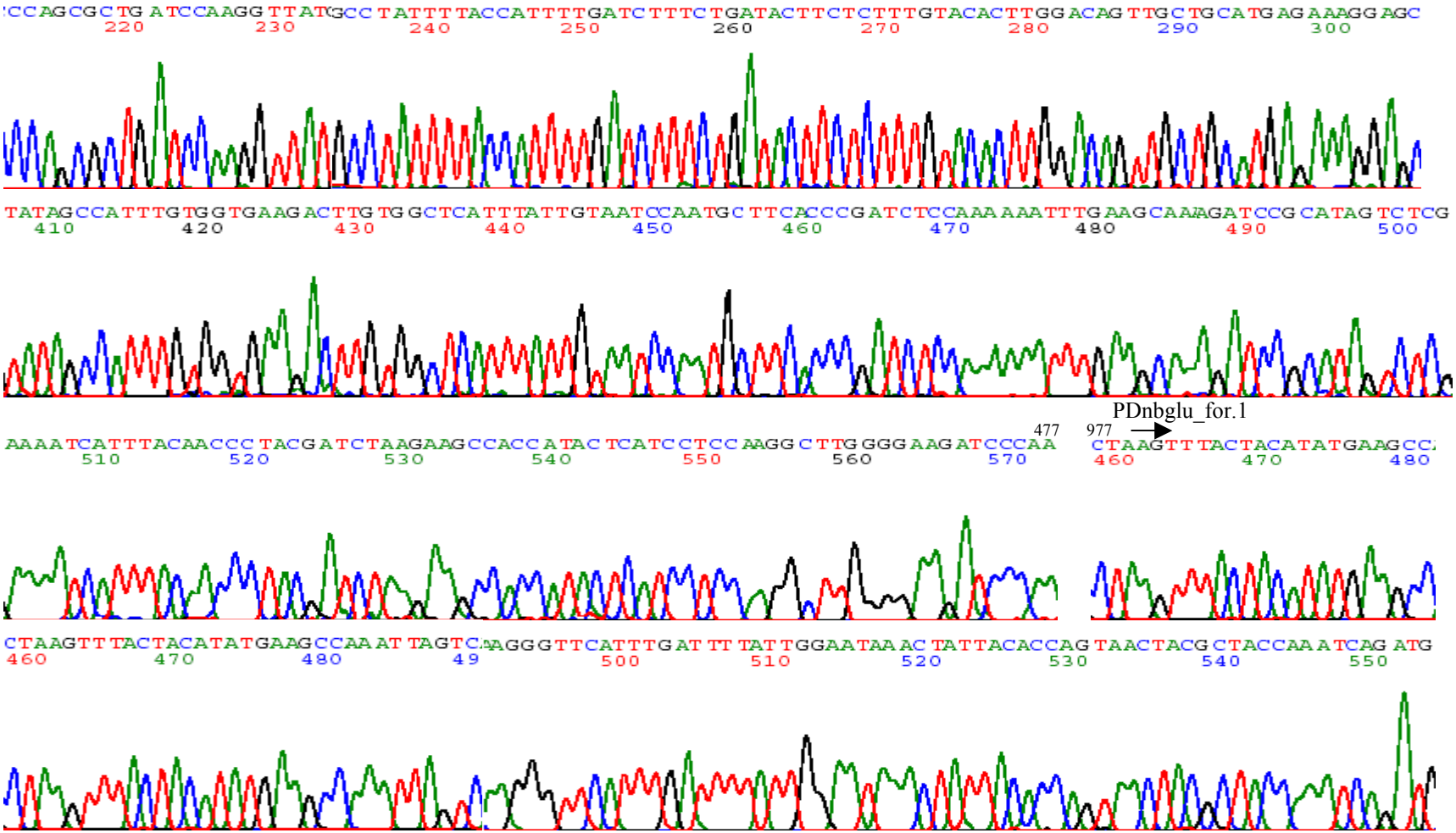


Figure 1 Continued

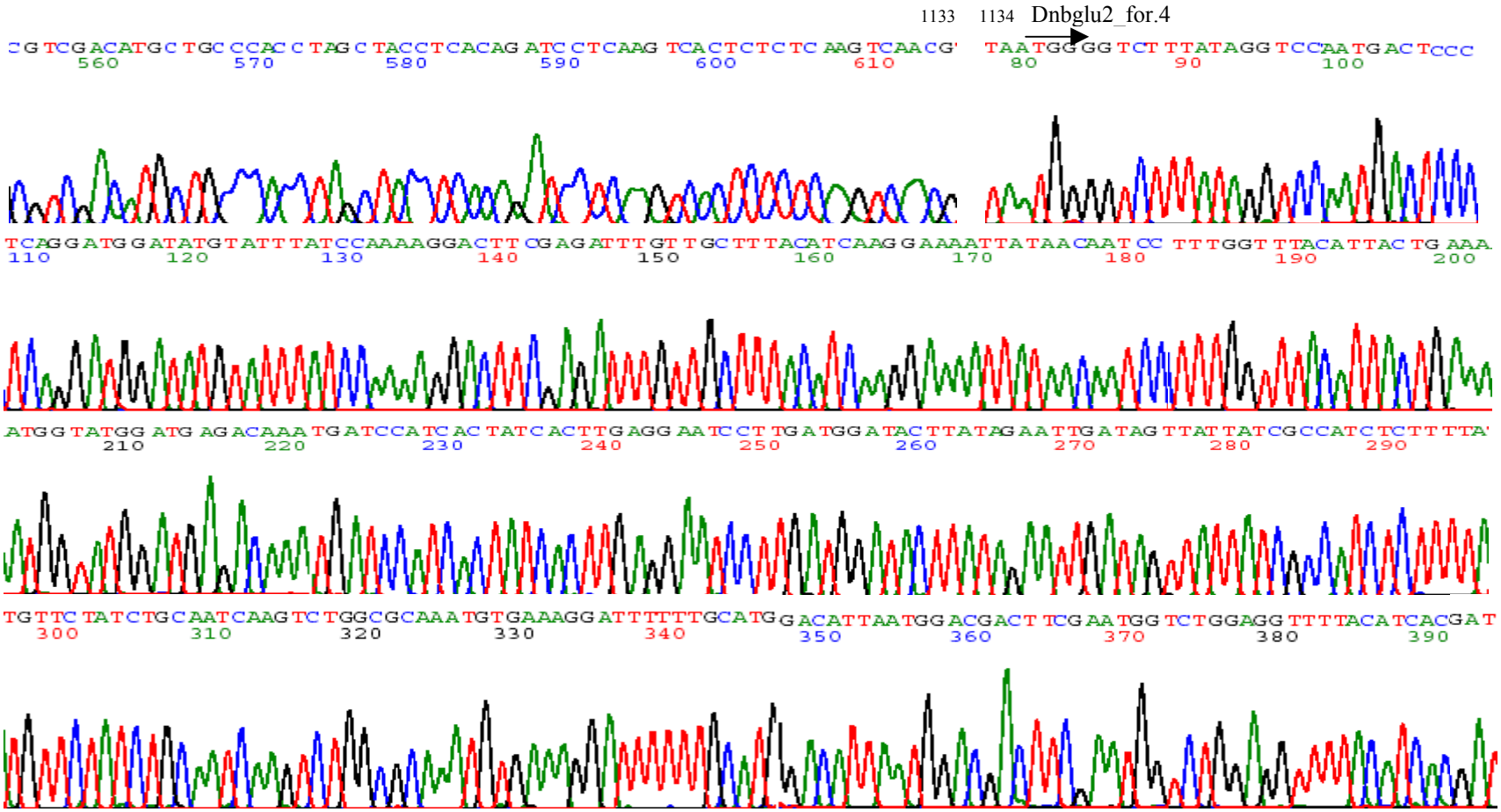


Figure 1 Continued

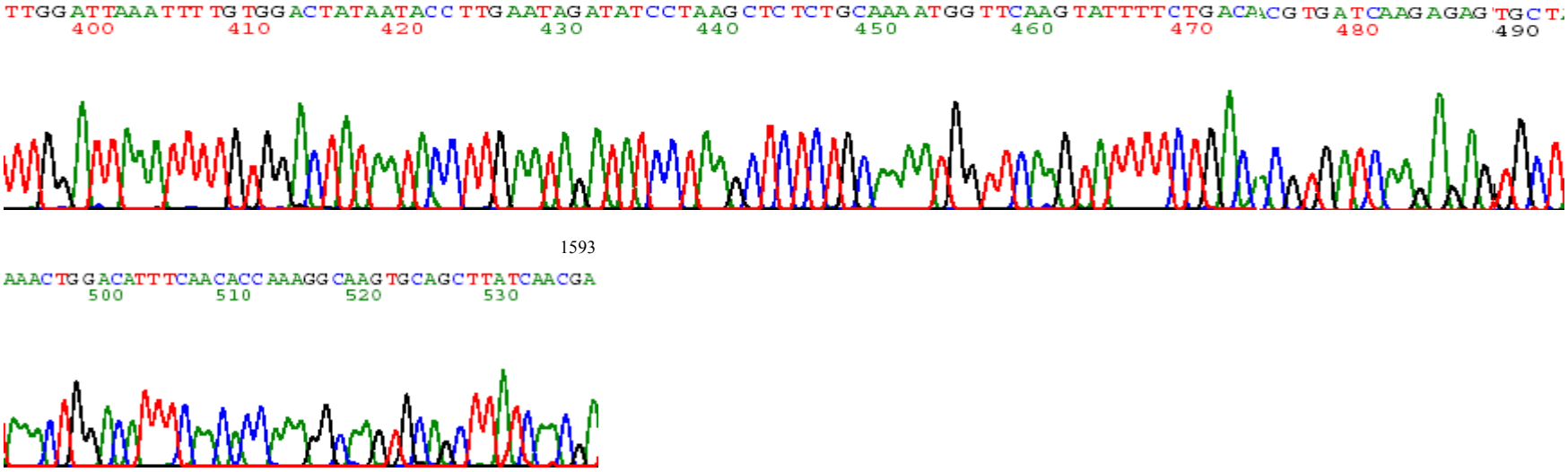


Figure 1 Continued

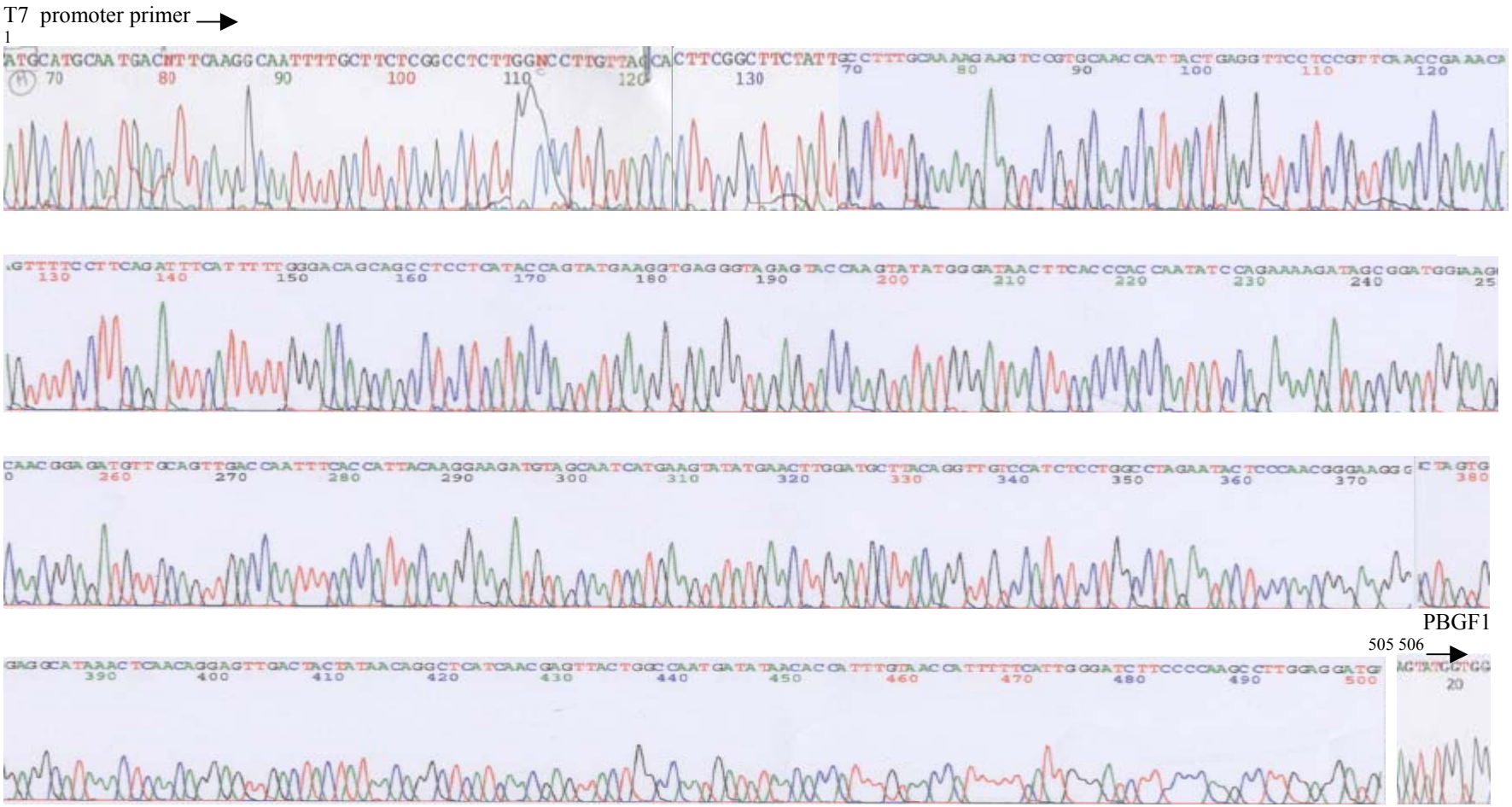


Figure 2 The chromatogram of cDNA sequence of *Dnbglu1*, using T7 promoter, PBGF1, T3 promoter and PBGF2 as sequencing primers. The numbers labeled on the sequence are based on the cDNA sequence of *Dnbglu1* in Figure 3.34

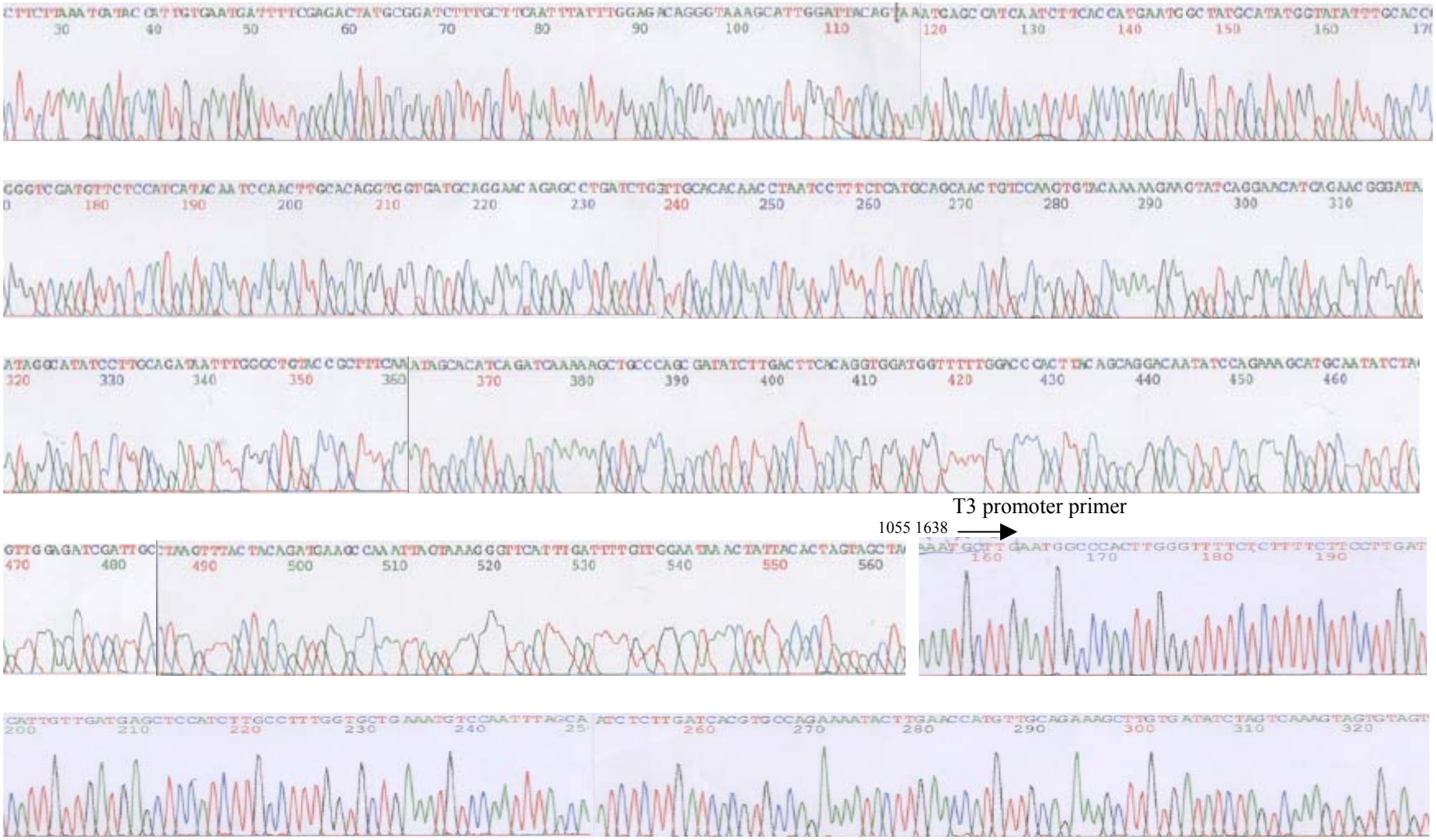


Figure 2 Continued

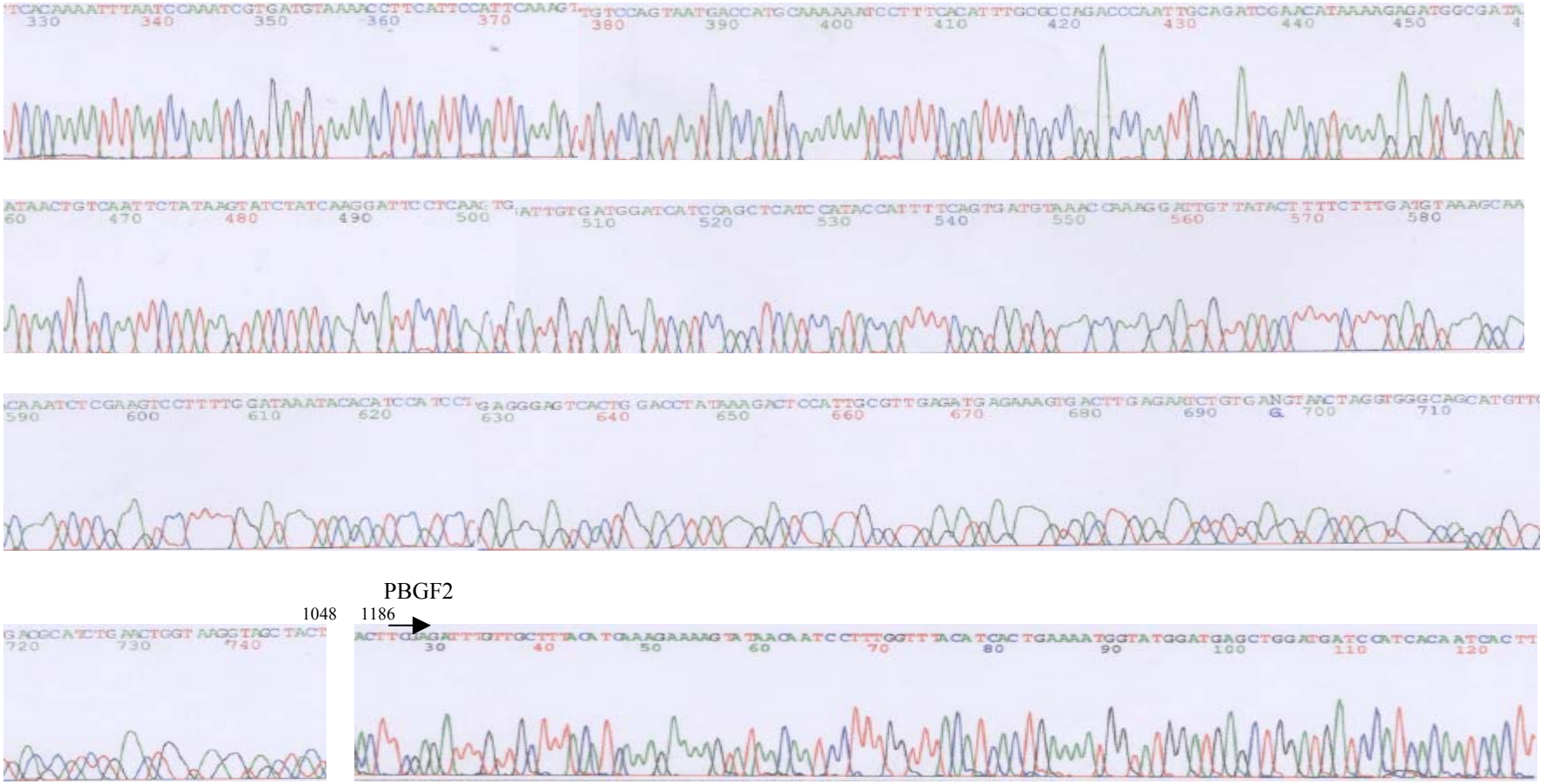


Figure 2 Continued

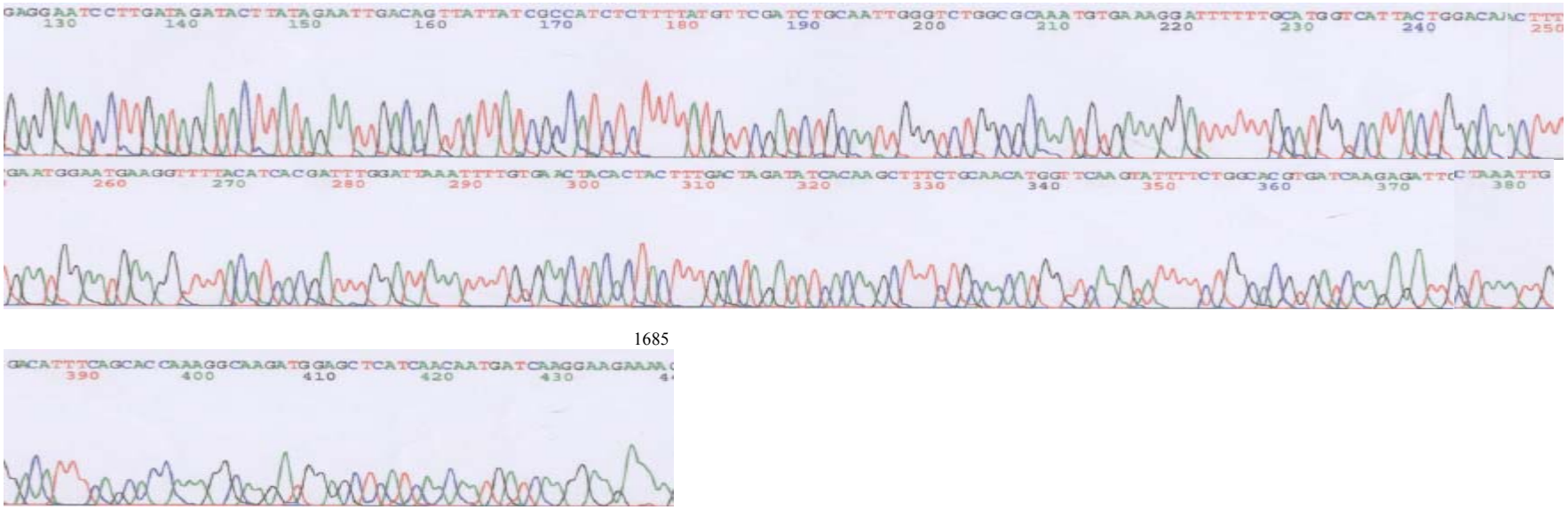


Figure 2 Continued

CURRICULUM VITAE

NAME Miss Phimonphan Chuankhayan

DATE OF BIRTH 14 December 1975

PLACE OF BIRTH Surin, Thailand

INSTITUTION ATTENDED

Mahasarakham University, Mahasarakham, Thailand, 1994- 1998, Bachelor of Science (Chemistry)

Suranaree University of Technology, Nakhon Ratchasima, Thailand, 1998-2004, Doctor of Philosophy (Biochemistry)

AWARDS AND SCHOLARSHIPS

The Royal Golden Jubilee Ph.D. Scholarship, from the Thailand Research Fund, 1999