Clendennen, S.K., and May, G.D. 1997. Differential gene expression in ripening banana fruit. *Plant Physiol.* **115**: 463-469.

Collinge, D.B., Kragh, K.M., Mikkelsen, J.D., Nielsen, K.K., Rasmussen, U., & Vad, K. 1993. "Plant chitinases" *The Plant J.* **3**: 31-40.

Dayhoff, M & R. Eck, 1968. *Atlas of Protein Sequence and Structure*, Chapter 4. Natl. Biomed. Res. Fund. Silver Spring, MD, USA.

Dey, P. 1984. "Biochemistry of the Multiple Forms of Glycosidases in Plants" In *Adv. Enzymol.* (Meister, A. ed) *vol 56*: 141-249 John Wiley & Sons Inc. New York.

Doolittle, R. F., *Of URFs and ORFs.* 1987. 103 pgs. University Science Books. Mill Valley, CA, USA

Dwek, R. 1995. "Glycobiology: towards understanding the function of sugars." *Biochem. Soc. Trans.* **23**: 1-25.

Dwek, R. C. Edge, D. Harvey, M. Wormald & Parekh. 1993 "Analysis of glycoprotein-associated oligosaccharides." *Ann. Rev. Biochem.* **62**:65-100

Estruch, J.J., Chriqui, D., Grossmann, K., Schell, J. and Spena, A. 1991a. The plant oncogene rolC is responsible for the release of cytokinins from glucoside congjugates. *EMBO J.* **10**: 2889-2895.

Estruch, J.J. Shell, J. and Spena, A. 1991b. The protein encoded by the rolB plant oncogene hydrolyses indole glucosides *EMBO J.* **10**: 3125-3128.

Facchini,P. J., & V. De Luca. 1994. "Differential and tissue-specific expression of a gene family for tyrosine/dopa decarboxylase in opium poppy." *J. Biol. Chem.* **269**: 266684-90.

Feng, D. F. and R. Doolittle, 1996 "Progressive alignment of Amino Acid Sequences and Construction of phylogenetic trees from them." *Meth. Enzymol.* **266**: 368-382

Guex, N. and Peitsch, M.C. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**: 2714-23.

Henrissat, B. 1991. "A classification of glycosyl hydrolases based on amino acid sequence similarities." *Biochem. J.* **280**: 309-16.

Henrissat, B. and Bairoch. A. 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **293**: 781-8.

Henrissat, B. and Bairoch, A. 1996. Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* **316**: 695-6.

Higgins, D.G., Thompson, J.D., and Gibson, T.J. 1996. Using CLUSTAL for Multiple Sequence Alignments. Meth. Enzymol. 76: 383-402.

Hofmann K., Bucher P., Falquet L., Bairoch A. 1999. "The PROSITE database, its status in 1999." *Nucleic Acids Res.* **27**:215-219

Ichikawa, Y., G. Look, & C. Wong. 1992. "Enzyme-catalyzed oligosaccharide synthesis." *Anal. Biochem.* **202**: 215-238

Jenkins, J. Leggio, L.L., Harris, G. and Pickersgill, R. (1995) β-Glucosidase, β-galactosidase, family A cellulases, family F xylanases and two barley glycanases form a superfamily of enzymes with 8-fold β/α architecture and with two conserved glutamates near the carboxy-terminal ends of β-strands four and seven. *FEBS Letters* **362**, 281-285

Johannson, E, L. Hedbys, P.E. Larsson, K. Mosbach, A. Gunnarsson & S. Svensson. 1986. "Synthesis of mannose oligosaccharides via reversal of α-mannosidase reaction." *Biotechnol. Letter* **8**: 421-424.

Juliano, B.O. 1985. "Polysaccharides, Proteins, and Lipids of Rice" in *Rice: Chemistry and Technology*. Juliano, B.O. Editor. The American Society of Cereal Chemists Inc. St Paul, USA. pp. 59-174

Ketudat-Cairns, J.R., Chantarangsee, M., Chaiwangrad, S., Phawong, J. (1999) Primary Structure-Based Screening for Glycosyl Hydrolases in Thai Plants. *Thai Journal of Biotechnology* **1**, 20-30.

Leah, R.H., Tommerup, I. Svendsen, & J. Mundy. 1991. "Biochemical and molecular characterization of three barley seed proteins with antifungal properties." *J. Biol. Chem.* **266**: 1564-73.

McPherson, M., K. Jones, & S. Gurr, 1992. "PCR with highly degenerate primers" in *PCR. A Practical Approach.* (M. McPherson, *et. al.* eds.), Oxford, UK. pp 171-186

Maier, J. K. Wilitter, M. Gutlich, I. Ziegler, T. Werner, & H. Ninnemann. 1995. "Homology cloning of GTP-cyclohydrolase I from various unrelated eukaryotes by reverse-transcription polymerase chain reaction using a general set of degenerate primers." *Biochem. Biophys. Res. Comm.* **212**: 705-11.

Malboobi, M.A. and Lefebvre, D.D. 1997. A phosphate-starvation inducible β-glucosidase gene (psr3.2) isolated from *Arabidopsis thaliana* is a member of a distinct subfamily of the BGA family. *Plant Molec. Biol.* **34**: 57-68.

Melchers, L., Apotheker-de Groot, M., van der Knapp, J., Ponstein, A., Sele-Buurlage, M., Bol, J., Cornelissen, B., van den Elzen, P., and Linthorst, H. 1994. A new class of tobacco chitinases homologous to bacterial exo-chitinases displays antifungal activity. The *Plant J.* **5**: 469-80.

Minic, Z., Brown, S., De Kouchkovsky, Y., Schultze, M. and Staehelin, C. (1998) "Purification and characterization of a novel chitinase-lysozyme, of another chitinase, both hydrolysing Rhizobium meliloti Nod factors, and of a pathogenesis-related protein from *Medicago sativa* roots." *Biochem. J.* **332**: 329-335

Nilsson, K. 1988. "Enzymatic Synthesis of Oligosaccharides, *Trends Biotechnol.* **6**: 256-264

Palmiano, E.P., and Juliano, B.O. 1973. Changes in the Activity of Some Hydrolases, Peroxidase, and Catalase in Rice Seed during Germination. *Plant Physiol.* **52**: 274-277.

Paulsen, H. 1984. *Chem. Soc. Rev.* **13**:15-45.

Pluthero, F.G. 1993. Rapid purification of high-activity Taq DNA polymerase. *Nucleic Acids Res* **21**:4850-1

Poulton, J.E. 1990. "Cyanogenesis in plants." Plant Physiol. 94: 401-405

Quigley,F., Dao,P., Cottet,A. and Mache,R. 1996. Sequence Analysis of an 81kb contig from *Arabidopsis thaliana* chromosome III. *Nucleic Acids Res.* **24**: 4313-4318.

Raghothama, K.G., Lawton, K.A., Golsbrough, P.B., & Woodson, W.R. 1991. "Characterization of an ethylene-regulated flower senescence-related gene from carnation." *Plant Mol. Biol.* **17**: 61-71.

Smith,D.L., Starrett,D.A. and Gross,K.C. 1998. "A gene coding for tomato fruit beta-galactosidase II is expressed during fruit ripening. Cloning, characterization, and expression pattern." *Plant Physiol.* **117**: 417-423.

Srisomsup C., Svasti, J. Surarit, R., Champattanachai, V., Boonpuan, K. & Sawangareetrakul, P. 1996. "Isolation and Characterization of β-D-Glucosidase/β-D-Fucosidase from *Dalbergia cochinchinensis* Pierre." *J. Biochem.* **119**: 111-116

Staehelin, C., M. Schultze, E. Kondorosi, R.B. Mellor, T. Boller and A. Kondorosi. 1994. Structural modificatoins in Rhizobium meliloti Nod factors influence their stability against hydrolysis by root chitinases. *The Plant J.* **5**: 319-330.

Suginta, W. & J. Svasti, 1995. "Purification and Properties of β-Galactosidase from *Hibiscus sabdariffa* L. var. *Altissima*" *J. Sci. Soc. Thailand* **21**: 183-185

Surarit, R., J. Svasti, C. Srisomsap, W. Suginta, S. Khunyosheng, S. Nilwarangkoon, P. Harnsakul, & E. Benjarongkulchai. 1995. "Screening of Glycohydrolase Enzymes in Thai Plant Seeds for Potential use in Oligosaccharide synthesis." *J. Sci. Soc. Thailand.* **21**: 293-303.

Surarit, R., Matsui, H., Chiba, S., Svasti, J., and Srisomsap, C. 1997. Bioscience, Biotechnology and Biochemistry 61, 93-

Svasti, J., Srisomsap, C., Techasakul, S., and Surarit, R. 1999. Dalcochinin-8′ -O-D-glucoside and its β-glucosidase enzyme from *Dalbergia cochinchinensis*. *Phytochemistry* **50**: 739-43.

Welply, J. 1989. "Sequencing methods for carbohydrates and their biological applications. *Trends Biotechnol.* **7**: 5-10.

Output:

International Paper Submitted:
Cairns, J.R.K., Champattanachai, V., Srisosap, C., Svasti, J. "Primary Structure of the Dalcochinin-8'O-β-glucoside β-glucosidase from the Thai Rosewood *Dalbergia cochinchinensis* Pierre. Submitted to the *Biochemical Journal*, Portland Press, UK (August, 1999).

Thai Paper
Ketudat-Cairns, J.R., Chantarangsee, M., Chaiwangrad, S., Phawong, J. (1999) Primary Structure-Based Screening for Glycosyl Hydrolases in Thai Plants. *Thai Journal of Biotechnology* 1, 20-30.

Papers Presented:
International Meetings:
Svasti, J., Srisomsap, C., Surarit, R., Techasakul, S., & Ketudat-Cairns, J. 1998. Characterization of a Novel Rotenoid-β-glucosidase Enzyme and its Natural Substrate from Thai Rosewood. *J. Pure Appl. Chem.*, **70:** 71
http://www.iupac.org/symposia/proceedings/puket97/svasti.html.

J. Svasti, C. Srisomsap, R. Surarit, and J. Ketudat-Cairns, 1997. Rotenone β-Glucosidase from *Dalbergia cochinchinensis* Pierre: Identification of the Substrate from Its Enzyme. "2nd European Protein Symposium," Cambridge, England.

International Meeting Presentations Accepted:
James R. Ketudat Cairns, Voraratt Champattanachai, Phimolphan Chuankayan, Mallika Chantarangsee, Chantragan Srisomsap, Kanokporn Boonpuan, Rudee Surarit Jisnuson Svasti. Primary Structure Determination of Plant Beta-Glucosidases and Modeling of Their 3-Dimensional Structures. The 8th Asian Chemical Congress. Taipei, Taiwan, Nov. 21-24, 1999.

Jisnuson Svasti, Chantragan Srisomsap, Rudee Surarit, James Ketudat-Cairns, and Supanna Techasakul. A Glycosidase Enzyme from Thai Rosewood with Substrate Specificity for Dalcochinin-β-D-Glucoside The 8th Asian Chemical Congress. Taipei, Taiwan, Nov. 21-24, 1999.

Thai Meetings:
Ketudat Cairns, J., M. Chantarangsee, M. Kaomek, J. Phawong, S. Chaiwangrad, & T. Suksongwon, 1998. Homology-Based Cloning of Thai Plant Glycosidases. 10th Annual Meeting of the Thai Society for Biotechnology & 1998 Annual Meeting of BIOTECH. Bangkok, Thailand 15-17 Nov. 1998.

Ketudat Cairns, J & M. Chantarangsee. Protein Structure-Based Cloning of Partial Beta-Glucosidase Genes or Pseudogenes from Mango. 24th Congress on Science and Technology of Thailand. Bangkok, Thailand October, 1998.

Kaomek, M. & J. Ketudat Cairns, 1998. Partial Primary Sequence of a Putative Class I Chitinase from *Leucaena leucocephala*. 24th Congress on Science and Technology of Thailand. Bangkok, Thailand October, 1998.

Champattanachai, V., J. Ketudat Cairns, C. Srisomsap, & J. Svasti. 1998. Classification of Thai Rosewood β-Glucosidase Based on Amino Acid Sequence Alignment. 24th Congress on Science and Technology of Thailand. Bangkok, Thailand October, 1998.

Ketudat Cairns, J.R., Chantarangsee, M., Chaiwangrad, S., Phawong, J. 1997. Primary Structure-Based Screening for Glycosyl Hydrolases in Thai Plants. The 9th Annual Meeting of the Thai Society of Biotechnology & The 2nd JSPS-NRCT-DOST-LIPI-VCC Seminar, "Biotechnology: An Essential Tool for Future Development?" Nakhon Ratchasima, Thailand, 19-22 Nov. 1997.

Ketudat-Cairns, J.R., Phawong, J., Chaiwangrad, S.& M. Chantarangsee. 1997. Application of Protein Sequence Analysis to Screening for Glycosyl Hydrolase Enzymes from Thai Plants. 23rd Annual Congress on Science and Technology of Thailand, Chiang Mai, Thailand 20-22 October, 1997.

Champattanachai, V., Ketudat-Cairns, J.R., Srisomsap, C., and Svasti, MR.J. 1997. Cloning of the cDNA for a β-Glucosidase/β-Fucosidase Gene from *Dalbergia cochinchinensis* Pierre (Thai Rosewood). 23rd Annual Congress on Science and Technology of Thailand, Chiang Mai, Thailand 20-22 October, 1997.

ภาคผนวก

Appendices:

1. Copy of the Thailand Journal of Biotechnology Paper:

Ketudat-Cairns, J.R., Chantarangsee, M., Chaiwangrad, S., Phawong, J. (1999) Primary Structure-Based Screening for Glycosyl Hydrolases in Thai Plants. *Thai Journal of Biotechnology* 1, 20-30.

2. Copy of Manuscript for submitted to Biochemical Journal:

Cairns, J.R.K., Champattanachai, V., Srisosap, C., Svasti, J. "Primary Structure of the Dalcochinin-8'O-β-glucoside β-glucosidase from the Thai Rosewood *Dalbergia cochinchinensis* Pierre. Submitted to the *Biochemical Journal*, Portland Press, UK (August, 1999).

Thai J. Biotechnol., September 1999, p.20-30 Vol. 1, No. 1

Primary Structure-Based Screening for Glycosyl Hydrolases in Thai Plants

JAMES R. KETUDAT CAIRNS[1,2], MALLIKA CHANTARANGSEE[1], SUKSABAI CHAIWANGRAD[1], and JITTRA PHAWONG[1]

1 School of Chemistry, Institute of Science, Suranaree University of Technology Nakhon Ratchasima 30000, Thailand

2 Laboratory of Biochemistry, Chulabhorn Research Institute, Bangkok, Thailand

Glycohydrolase enzymes have many present and potential applications in medical, agricultural, and industrial biotechnology. In order to identify novel glycohydrolases from Thai plants, a protein sequence-based approach was undertaken. Known protein sequences of α-mannosidases, β-galactosidases, and β-glucosidases were retrieved from on-line databases, and grouped into sequence families as previously described by Henrissat (1991). Multiple sequence alignments were performed and phylogenetic trees constructed with the computer programs of Feng and Doolittle (1996). Sequences likely to be conserved in a broad range of plant species were identified from the alignment and phylogenetic relationships. These sequences were used to design degenerate primers, which were used for PCR screening of plant DNA and reverse-transcribed RNA. Initial characterization of primer conditions were carried out with DNA and cDNA from rice as a monocot model, and Thai rosewood, and bean as dicot models. Specific PCR products of the proper size have been amplified for rice α-mannosidase and β-galactosidase and for β-galactosidase and β-glucosidase from bean and Thai rosewood. These products are currently being cloned and characterized.

Introduction

In order to find glycosyl hydrolases for potential use in oligosaccharide and glycoside synthesis, Surarit *et al.* (Surarit *et al.*, 1995) screened plants found in Thailand for 8 activities. Six different activities were found, with 38 of the 50 species containing at least one activity assayed by the hydrolysis of p-nitrophenol (pNP) glycosides. One plant, *Dalbergia cochinchinensis* Pierre (Thai Rosewood), produced the highest amount of product by far, using both pNP-fucoside and pNP-glucoside. Moderately high levels of α-mannosidase and β-galactosidase activities were found in other plants. Each of these activities were partially purified and shown to be useful for oligosaccharide synthesis by reverse hydrolysis. The *D. cochinchinensis* enzyme has since been purified and shown to be a single enzyme with glucose and fucose reacting at a single active site (Srisomsap, 1996; Surarit, 1997). Recently, the cDNA for this protein was cloned by polymerase chain reaction based on peptide sequence and shown to be a member of glycosyl hydrolase family 1 with closest relationship to black cherry prunasin hydrolase and white clover cyanogenic β-glucosidase (Champattana-chai, 1991).

Despite the success of the pNP-glycoside substrate screen, it has proved arduous to go from a recognized activity to a cloned cDNA and protein sequence. To date, only *D. cochinchinensis* β-glucosidase has been fully purified and sequenced. Moreover, enzymes with low activity in the pNP-glycoside assay could not be purified. A more direct cloning method could produce

sequence and cloned proteins faster with more certainty. Henrissat and Baroch have shown that glycosyl hydrolases with related activities can be grouped into families related by their amino acid sequence homologies (Henrissat, 1991; Henrissat and Bairoch, 1993; Henrissat and Bairoch, 1996). Therefore the use of these sequence relationships to develop degenerate oligonucleotides for PCR cloning from conserved sequences in lieu of direct protein sequence information was investigated. This strategy has previously been successfully used to screen for GTP-cyclohydrolase in divergent species (Maier *et al.*, 1995) and glutamate dehydrogenase from bacteria, fungi, and plants (McPherson *et al.*, 1991). Here, it was tested for screening of α-mannosidase, β-galactosidases, and β-glucosidases in plants.

Materials and Methods

Enzymes and Reagents

Taq polymerase was either purchased from Promega (Madison, USA) or prepared by the method of Pluthero. Reverse transcriptase was either Superscript II from GIBCO-BRL (Grand Island, USA) or Ready-to-Go from Pharmacia (Upsala, Sweden). Deoxyribonucleotides, 10X buffer for PCR, and 25 mM $MgCl_2$ were purchased from Promega. TriReugent and molecular biology grade chloroform used in RNA purification was purchased from Sigma (St. Louis, USA). All other chemicals were analytical grade or higher.

Sequence Retrieval and Analysis

Known protein sequences for β-glucosidases, β-galactosidases, and α-mannosidases were retrieved from GENBANK using the NCBI Entrez search engine (http://www.ncbi.nlm.nih.gov/Entrez) or from the SWISSPROT database using the Prosite utility of the expasy server (http://www.expasy./ch/). Sequences were retrieved from the GENBANK by E.C. number, name, and related proteins searches and from the SWISSPROT with the Prosite protein pattern search for the glycosyl hydrolase families of interest.

Protein sequences were formatted and analyzed using the Propack protein sequence analysis computer programs of Feng and Doolittle (1996) using a DEC alpha 600 workstation. Sequences were arranged in rough phylogenetic order using the arrange program and sequences with greater than 95% identity from the same species were reduced to a single representative. Multiple sequence alignments and phylogenetic trees were calculated using the tree program with the Dayhoff PAM250 matrix (Swartz and Dayhoff, 1978). After resolution of negative branch lengths, data was transferred to a personal computer and phylogenetic trees drawn with the Phylyp package drawgram program of Felsenstein (Felsenstein, 1989).

Design of Degenerate Oligonucleotides

Conserved sequences were identified by inspection of multiple sequence alignments and their phylogenetic extent was assessed by inspection of the phylogenetic trees. Protein sequences were back-translated to yield possible nucleotide sequences. Degeneracy (total number of nucleotide sequences and possible binding sites) was calculated and sequences with lowest degeneracy were used to make nested (inner and outer sets) primers for PCR screening. Oligonucleotides were designed to include inosine (I) at all positions of 4-fold degeneracy, except those within 4 nucleotides of the 3' end, where a mixture of all four nucleotides (N) was used instead. All oligonucleotides were designed to have conserved G or C residues at the 3' end. The derived sequences were used to order degenerate oligonucleotides were ordered from GIBCO-BRL.

RNA preparation and Reverse Transcription

RNA was produced either from *Dalbergia cochinchinensis* seed pods or rice or bean seeds that had been washed in 0.1% hypochlorite 10 min., washed in distilled water and germinated in distilled water 2-4 days. The soft-green tissues (approximately 0.1 g) were dissected from any hard tissues, frozen in liquid nitrogen, and crushed to a powder. TriReagent was added and the RNA extracted according to the manufacturer's instructions. The final RNA produced was dried under air and dissolved in 20 µl sterile DEPC-treated distilled water.

From 5-20 µl of RNA was transferred to a new microcentrifuge tube, 100-300 pmols of poly-T_{17} primer or specific reverse primers (3' reverse primer for nested PCR) were added, and the sample was melted at 85°C 10 min. The reverse transcription was then performed according to the reverse transcriptase supplier's directions. Final volume was 20 µl. For PCR, the cDNA product was diluted to 1/10 and 1 µl was used directly as template for the PCR reaction.

Polymerase Chain Reaction

PCR was conducted on a Perkin-Elmer-ABI 9700 thermocycler system. The typical reaction included 0.2 mM of each deoxyribonucleotide (dATP, dCTP,

B. Sequence Alignment

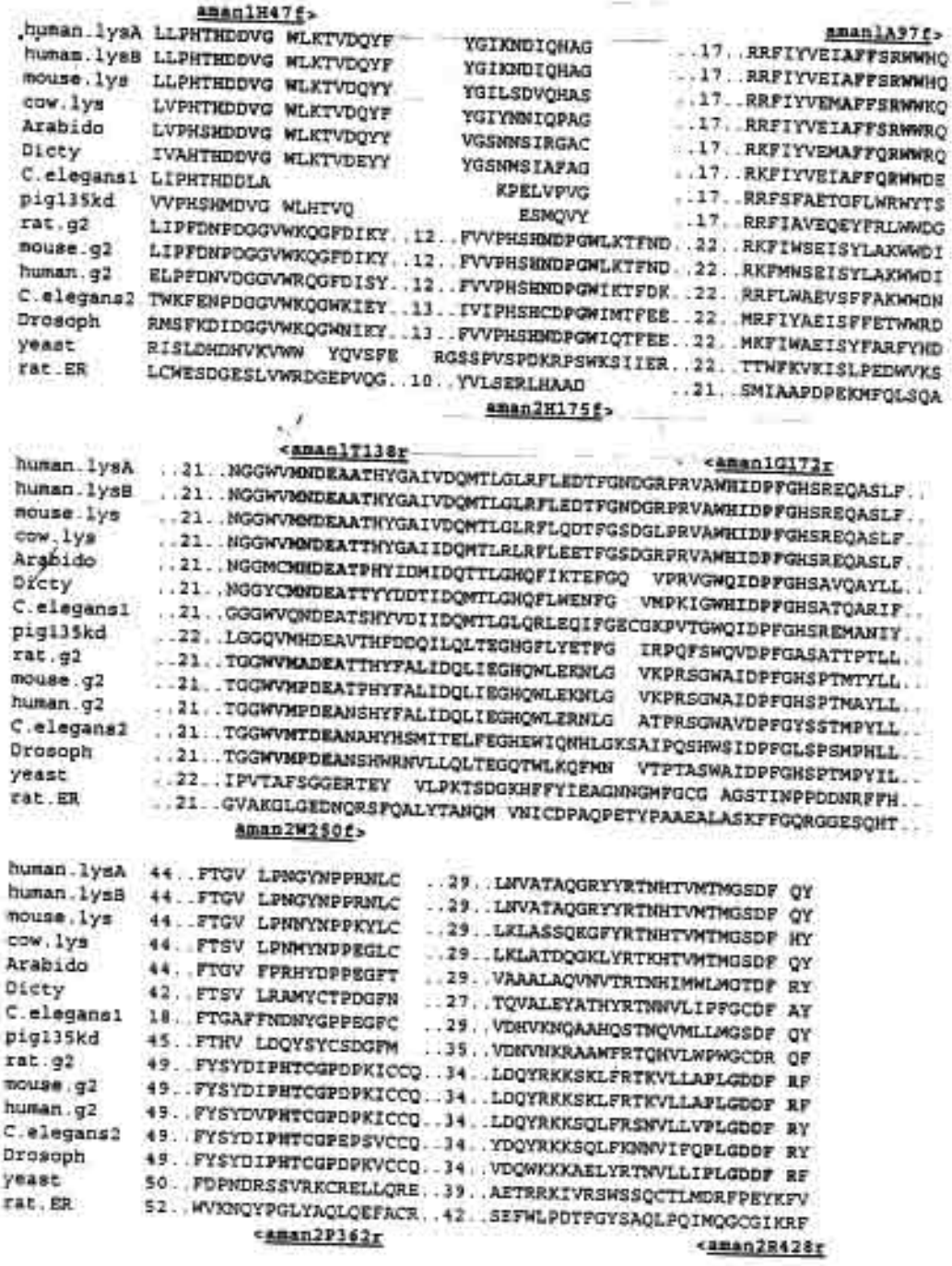
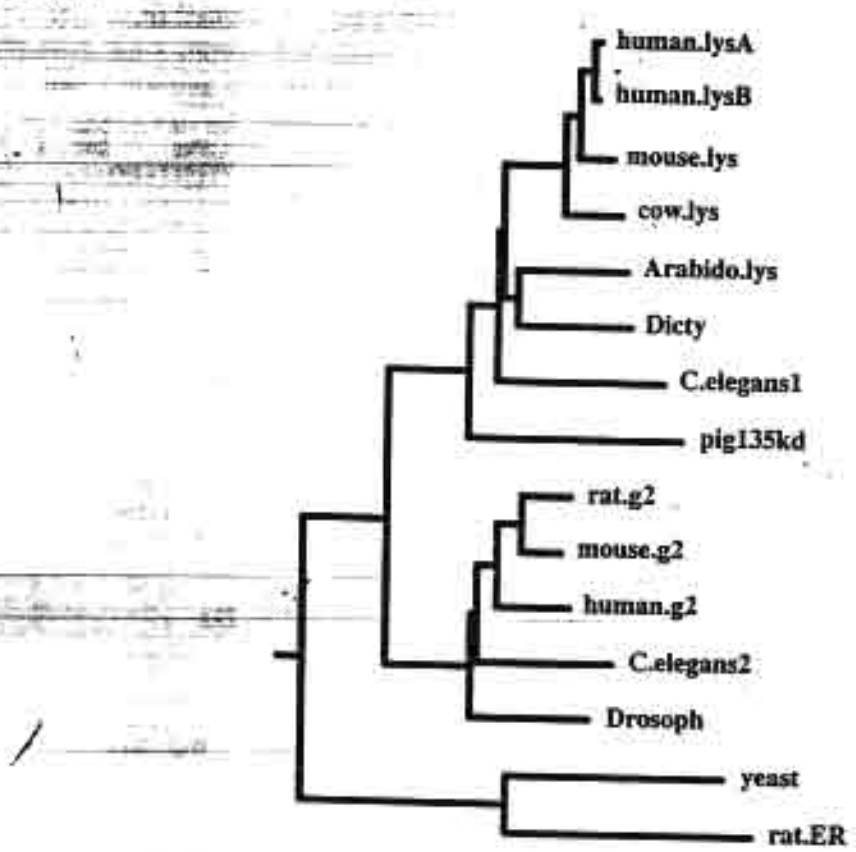


Fig 1. (continued)

plant sequence, from *Arabidopsis thaliani*, was found in family 38. Other sequences were from animals and fungi. A phylogenetic tree and partial sequence alignment for family 38 are shown in fig. 1. Two groups of proteins in this family, the lysosomal α-mannosidases and the golgi type II α-mannosidases contained separate sets of sequences conserved enough to make degenerate oligonucleotide primers. These sequences are indicated in bold in fig. 1.

A. Phylogenetic Tree



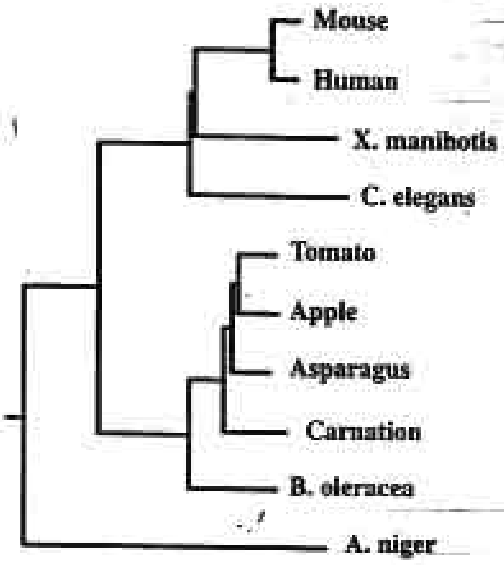
Sequences listed were extracted from the NCBI GENBANK database server and the SWISSPROT database and phylogenetic tree (A) and alignment (B) developed by the method of Feng and Doolittle (4). Sequences used to produce degenerate primers for PCR cloning are indicated in bold face. The designation of the primer derived from the sequences are indicated above the alignment for α-mannosidase type I (lysozomal isozyme) and below the alignment for α-mannosidase type II (golgi isozyme) primers. Numbers between sequence blocks indicate the number of amino acids between the represented blocks.
The sequences are α-mannosidases from: human.lysA: human lysozome (type I A) (Nebes and Schmidt, 1994); human.lysB: human lysozome (type I B) (Nebes and Schmidt, 1994); mouse.lys; mouse lysozome (Beccari *et al.*, 1997); cow.lys: bovine lysozome (Tollersund *et al.*, 1997); Arabido: *Arabidopsis thaliani* gene (Quigley *et al.*, 1996); Dicty: *D. discoidium* (Schatzle *et al.*, 1992); C.elegans1: gene from *Caenorhabditis elegans* (Wilson *et al.*, 1994); pig135kd: pig epididymis-secreted 135 kd protein (Okamura *et al.*, 1994); rat.g2: rat golgi (type II) (Moremen and Robins, 1991); mouse.g2: mouse golgi (type II) (Moremen, 1989); human.g2: human (type II) (Misago *et al.*, 1995); Drosoph: *D. melanogaster* golgi (type II) (Foster *et al.*, 1995); C.elegans2: translation product 1 from *C. elegans* gene F58H1 (Wilson *et al.*, 1994; yeast: *S. cervaseae* AMS1 gene (Yoshihisa and Anraku, 1989); rat.ER: rat endoplasmic reticulum (Bischott *et al.*, 1990). Note: sequences with >95% identity from the same species are included as only one representative sequence.

Fig 1 : Phylogenetic Tree and Partial Sequence Alignment of Glycosyl Hydrolase Family 38 α-Mannosidases

dGTP, and dTTP), 1-4.5 mM $MgCl_2$ (typically 3 mM), 50 mM KCl, 10 mM Tris (pH 9.0 at 25°C), 0.1% Triton X-100, 50 pmols forward and reverse oligonucleotides, 0.01-1 µl template cDNA or plasmid, and 1 unit of Taq polymerase in a 25 µl volume. Cycle conditions were generally as follows: 5 min. 94°C melt; then 30 cycles: 30 s 94°C melt, 60 s 40-55°C annealing, 60 s 72°C extension; then one 7 min. 72°C extension and hold the samples at 4°C.

For each enzyme cDNA fragment amplified, the amplification took the form of two nested reactions. First, the most 5′ forward primer and the most 3′ reverse primer were used to amplify the reverse-transcribed RNA. Then, the most 3′ forward primer and the most 5′ reverse primer were used to amplify a fragment from within the predicted product of the first amplification using 1 µl of 1/10 dilution of the first reaction as template. The nested PCR product was then analyzed by either 1% agarose gel or 8% polyacrylamide gel electrophoresis (Sambrook et al., 1989).

A. Phylogenetic Tree



Sequences:
Mouse (Namba, 1990)
Human (Oshima, 1988)
Xanthomonas manihotis (Tarron, 1995)
C. elegans (Wilson, 1994)
Tomato (Carey, 1995)
Apple (Ross, 1994)
Asparagus (King, 1995)
Carnation (Raghothama, 1991)
Brassica oleracea (Downs, 1995)
Aspergillis niger (Kumar, 1992)

B. Multiple Sequence Alignment

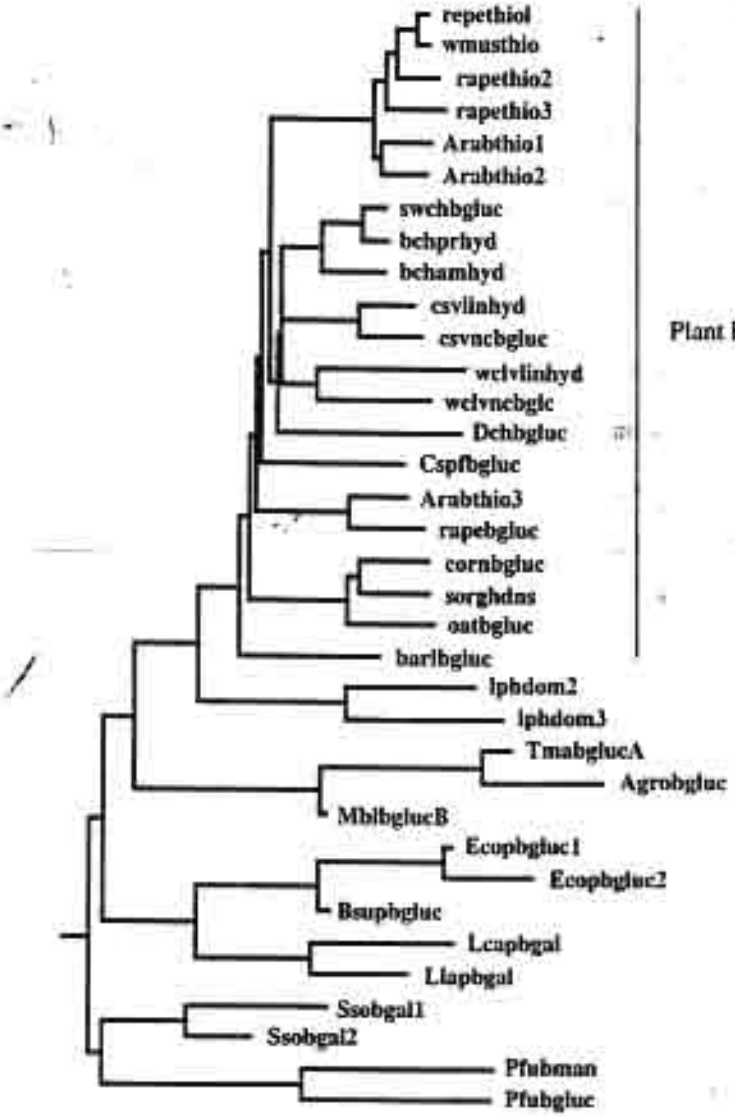
```
                    *         ** *   *           *    *           *       * **     **    ** *  **
Asparagus   WPDLIQKAKDGGLDVIQTYVFWNGHEPSPGQYYFGGRYDLVRFLKLV KQAGLYAHLRIGPYVCAEWNFGGFPVWLKYVPG
Tomato      WPDLIQKAKEGGVDVIQTYVFWNGHEPEEGKYYFEERYDLVKFIKVV QEAGLYVHLRIGPYACAEWNFGGFPVWLKYVPG
Apple       WPDLIQKAKDGGLDVIQTYVFWNGHEPSPGNYYFEERYDLVKFIKLV QQEGLFVNLRIGPYVCAEWNFGGFPVWLKYVPG
Carnation   WPDIIEKAKDSQLDVIQTYVFWNGHEPSEGKYYFEGRYDLVKFIKLI HQAGLFVHLRIGPFACAEWNFGGFPVWLKYVPG
B.oleracea  WPDLISKAKDGGLDTIETYVFWNAHEPSRRQYDFSGNLDLVRFIKTI QSAGLYSVLRIGPYVCAEWNYGGFPVWLHNMPD
Human       WKDRLLKMKMAGLNAIQTYVPWNFHEPWPGQYQFSEDHDVEYFLRLA HELGLLVILRPGPYICAEWEMGGLPAWLLEKES
Mouse       WEDRLLKMKMAGLNAIQMYVPWNFHEPQPGQYEFSGDRDVEHFIQLA HELGLLVILRPGPYICAEWDMGGLPAWLLEKQS
X.manihotis WKDRLQKARALGLNTVETYVFWNLVEPQQGQFDFSGNNDVAAFVKEA AAQGLNVILRPGPYACAEWEAGGYPAWLFGKGN
C.elegans   WDERLGKVRALGFNAIQYYIPWNMHELEEGNHDFSGLLDFAEFSMMAFNKYGLMTILRVGPYICGELENGGLPWWLLNKNV
A.niger     QLDIFQKVKALGFNCVSFYVDWALVEGKPGEYRADGIFDLEPFFDAA SEAGIYLLARPGPYINAESSGGGFPGWLQRVNG
PRIMERS:                    bgalI73f>                                                       bgalG130f>
                                                                                           bgalW127f>

                                                               *       *         *
Asparagus   ..93..CN        GFYC        DYFSPNK DMKPE   MWTEAWTGWFTGFGGAVPQRPAEDMAFAVAR FIQKGGSF
Tomato      ..93..CN        GFYC        DYFTPNK ANKPK   MWTEAWTAWFTEFGGPVPYRPAEDMAFAVAR FIQTGGSF
Apple       ..93..CN        GFYC        ENFKPNK DYKPK   MWTEVWTGWYTEFGGAVPTRPAEDVAFSVAR FIQSGGSF
Carnation   ..94..CN        GFYC        EGFVPKD KSKPK   MWTENWTGWYTEYGKPVPYRPAEDVAFSVAR FIQNGGSF
B.oleracea  ..93..CN        GFYC        DQYKPSN PSSPK   MWTENWTGWFKNWGGKHPYRTAEDLAFSVAR FFQTGGTF
Human       ..98..TG        SNIT        DAFLSQR KCEPKGPLINSEFYTGWLDHWGQPHSTIKTEAVASSLYD ILARGAS
Mouse       .159..TG        NNIT        QAFLVQR KFEPKGPLINSEFYTGWLDHWGKPHSTVKTKTLATSLYN LLARGAN
X.manihotis ..92..AP        GEAK        SAFDKLI KFRPDQPRMVGEYWAGWFDHWGKPHAATDARQQAEEFEW ILRQGHS
C.elegans   ..88..VEGVFPTVDFGPTDDAKEIENNFKLQRKFAPNGPLVNSEYYPGWLVLWGQKKQNLPSPQTIINGSQTMYSLGAS
A.niger     ..96..YGHDSYPL  GFDC..12..  TNFRTLHLEQSPTTPYAIVEFQGGSYDPWGGPGFAACSELLNNEFERVFYKNDFSF
PRIMERS:                                                <bgalT259r

                  * **  ***                                  *****      *
Asparagus         INYYMYHGGTNFG   RTA        GGPFI   STSYDYDAPIDEYGLLRQPKWGHLRDLHKAIKLCEPALVSGEPTI
Tomato            INYYMYHGGTNFG   RTS        GGPFI   ATSYDYDAPLDEFGSLRQPKWGHLKDLHRAIKLCEPALVSVDPTV
Apple             LNYYMYHGGTNFG   RTA        GGPFM   ATSYDYDAPLDEYGLPREPKWGHLRDLHKAIKSCESALVSVDPSV
Carnation         MNYYMFHGGTNFE    TT        AGRFV   STSYDYDAPLDEYGLPREPKYTHLKNLHKAIKMCEPALVSSDAKV
B.oleracea        QNYYMYHGGTNFG   RVA        GGPYI   TTSYDYDAPLDEYGNLNQPKWGHLKQLHTLLKSMEKPLTYGNIST
Human             VNLYMFIGGTNFAYWNGA         NSPYAAQPTSYDYDAPLSEAGDLTEKYFALRNIIQKFEKVPEGPIPPSTPK
Mouse             VNLYMFIGGTNFAYWNGA         NTPYEPQPTSYDYDAPLSEAGDLTKKYFALREVIQMFKEVPEGPIPPSTPK
X.manihotis       ANLYMFIGGTSFGFMNGANFQNNPSDHYAPQTTSYDYDAILDEAGH PTPKFALMRDAIARVTGVQPPALPAPIT
C.elegans         FNYYMIHGGTNFGFWNGAE         TE APCITSYDYDAPISESGDV TTKYLEIRKWIK  GLTDWPTPP
A.niger         QIAIMNLYMIFGGTNWG            NLGYPNGYTSYDYGSAVTESRNITREKYSELKLLGNFAKVSPGYLTASPGNL
PRIMERS:                <bgalF304r                          <bgalD323r
```

Fig 2: Phylogenetic Tree and Multiple Sequence Alignment of Glycosyl Hydrolase Family 35 β-Galactosidases.

A. Phylogenetic Tree of Glycosyl Hydrolase Family 1



Enzyme sequences included are: rapethio1-3, rape *(Brassica napus)* thioglucosidase (Falk *et al.*, 1992; Falk *et al.*, 1995; Thangstad, 1993); wmusthio, white mustard thioglucosidase (Xue *et al.*, 1992); Arabthio1-3, *Arabidopsis thaliani* thioglucosidases (Xue *et al.*, 1995; Schmidt *et al.*, 1995); swchbgluc, sweet cherry β-glucosidase (Weisma and Fils-Lycaon, 1996); bchprhyd, black cherry prunacin hydrolase (Poulton and Jurk, 1996); bchamhyd, black cherry amigdalin hydrolase (Li, 1992); csvlinhyd, cassava linamarin hydrolase (Hughes *et al.*, 1992); csvncbgluc, cassava non cyanogenic β-glucosidase (Liddle *et al.*, 1996); wclvlinhyd, white clover linamarin hydrolase *(Trifolium repens L.)* (Oxtoby *et al.*, 1991); wclvncbgluc, white clover non cyanogenic β-glucosidase (Oxtoby *et al.*, 1991); Dchbgluc, Thai Rosewood (*Dalbergia cochinchinensis* Pierre) β-glucosidase (Cairnsetal in preparation); Cspfbgluc, *Costus speciosis* furatanol β-glucosidase (Inoue *et al.*, 1996); rapebgluc, rape β-glucosidase (Falk, 1995); cornbgluc, maize β-glucosidase (Brzobohaty, *et al.*, 1993); sorghdns, sorghum durinase (Cicek and Esen, 1998); oatbgluc, oat β-glucosidase (Gus *et al.*, 1994); barlbgluc, barley β-glucosidase (Leah *et al.*, 1995); lphdom 2, 3, human lactase phlorizin hydrolase domains 2 & 3 (Manti *et al.*, 1998); TmabglucA, *Thermotoga maritima* β-glucosidase A (Liebl *et al.*, 1994); *Agrobgluc, Agrobacterium* sp. strain ATCC 21400 β-glucosidase (Wakarchuk *et al.*, 1988); MblbglucB, *Microbispora bispora* β-glucosidase B (Wright *et al.*, 1992); Ecopbgluc1, 2, *Eschericia coli* phospho-β-glucosidases 1 & 2 (Schnetz *et al.*, 1987; Burland *et al.*, 1993); Bsupbgluc, *Bacillus subtilis* phospho-β-glucosidase (Vosman, 1988); Lcapbgal, *Lactobacillus casei*, phospho-β-galactosidase (Porter and Chassy, 1988); Llapbgal, *Lactococcus lactus* phospho-β-galactosidase (Boizet *et al.*, 1988); Ssobgal1, 2, *Sulfolobus solfataricus* β-galactosidase 1 & 2 (Little *et al.*, 1989); Pfubman, *Pyrococcus furiosus* β-mannosidase (Bauer *et al.*, 1996); and Pfubgluc, *Pyrococcus furiosus* β-glucosidase (Bauer *et al.*, 1996).

Fig 3 : Phylogenetic Tree of Glycosyl Hydrolase Family 1 and Multiple Sequence Alignment of Plant β-Glucosidases and Thioglucosidases.

B. Family 1 Plant Glucosidase Sequence Alignment

```
                            **                     **         *          *   *
rapethio1   IEGG    RGRGVNVWDGFSHRYPEKAGSDLKNGDTTCESYTRWQKDVDVMGELNATGYRFSFAWS
wmusthio    IEGG    RGRGVNVWDGFSHRYPEKSGSDLKNGDTSCESYTRWKKDVEIMGELNATGYRFSFAWS
rapethio2   IEGG    RGRGLNIWDGFTHRFPEKGGSDLGNGDTTCESYTMWQKDIDIMDEMNATGYRFSFAWS
rapethio3    ACC    LGRGLNVWDGFTHRYPNKSGPDHGNGDTTCDSFSYWQKDIDVLDELNATGYRFSIAWS
Arabthio1   VEGG    RGRGLNVWDSFTHRFPEKGGADLGNGDTTCDSYTLWQKDIDVMDELNSTGYRFSIAWS
Arabthio2   IEGG    RGRGLNVWDGFTHRYPEKGGADLGNGDTTCDSYRTWQKDLDVMEELGVKGYRFSFAWS
scbgluc     LEGAANIDGRGPSIWDAFTHNHPEKI TDGSNGDVAIDQYHRYKEDVAIMKDMGLDAYRFSISWS
bcprhyd     LEGAANIDGRGPSVWDNFTHEHPEKI TDGSNGDVAIDQYHRYKEDVAIMKDMGLDAYRFSISWS
bcamhyd     FEGAAKEDGRGPSIWDTYTHNHSERI KDGSNGDVAVDQYHRYKEDVRIMKKMGFDAYRFSISWS
csvlinhyd   IEGEATAKGRAPSVWDIFSKETPDRI LDGSNGDVAVDFYNRYIQDIKNVKKMGFNAFRMSISWS
clvlinhyd   YEGAAFEDGKGPSIWDTFTHKYPEKI KDRTNGDVAIDEYHRYKEDIGIMKDMNLDAYRFSISWP
clvncbglc   FEGAVNEGGRGPSIWDTFTHKYPEKI RDGSNADITVDQYHRYKEDVGIMKDQNMDSYRFSISWP
Dcobgluc    YEG     EGRVPSIWDNFTHQYPEKI ADRSNGDVAVDQFHRYKKDIAIMKDMNLDAYRMSISWP
Cspfbgluc   VEGAWNEGGRGPSIWDTFTHDHPEKI ADHSNGDKATDSYKKYKEDVKLLKDLGLDSYRFSISWS
csvncbglc   IEGAANKFGRGASVWDTFTHQYPERI LDHSTGDVADGFYYRFKGDIQNVKNMGFNAFRFLISWP
Arabthio3   VEGAIMETCRGPALWDIYCRRYPERC  NNDNGDVAVDFFHRYKEDIQLMKNLNTDAFRMSIAWP
rapebgluc   VEGAVNEGCRGPSLWDIYTKKFPHRV  KNHNADVAVDFYHRFREDIKLMKKLNTDALRLSIAWP
cornbgluc   IEGAWNEDGKGESNWDHFCHNHPERI LDGSNSDIGANSYHMYKTDVRLLKEMGMDAYRFSISWP
sorghdns    IEGAWNEDGKGPSTWDHFCHNFPEWI VDRSNGDVAADSYHMYAEDVRLLKEMGMDAYRFSISWP
barlbgluc   VEGMARQGGRGPCIWDAFV AIQGMI AGNGTADVTVDEYHRYKEDVGIMKNMGFDAYRFSISWS
oatbgluc    IEGAWNEGGKGPSSWDNFCHSHPDRI MDKSNADVAANSYYMYKEDVRMLKEIGMDSYRFSISWP
PRIMERS:                bgluG62f>                                     bgluA109f>


                             * * *       *          * * *          *  **  ***
rapethio1   30..NITPFVTLFHWDLPQTLQDEYEGFLD..24..WITINQLYTVPTRGYAIGTDAPG RCSP
wmusthio    30..NITPFVTLFHWDLPQTLQDEYEGFLD..24..WITINQLYTVPTRGYALGTDAPG RCSP
rapethio2   30..NITPFVTLYHWDLPQTLQDEYEGFLN..24..WLTINQLYSVPTRGYSTGADAPV RCSP
rapethio3   30..GITPFVTLFHWDLPQVLQDEYEGFLD..24..WLTINQLYTVPTRGYGAGSDAPG RCSP
Arabthio1   30..NMTPFVTLFHWDLPQTLQDEYNGFLN..24..WITINQLYTVPTRGYALGTDAPG RCSP
Arabthio2   30..NITPFVTLFHWDLPQSLQDEYEGFLD..24..WITINQLFTVPTRGYALGTDAPG RCSQ
scbgluc     30..GIEPLVTLFHWDVPQALEEEYGGVLS..24..WTTLNEPYTISNHGYTIGIHAPG RCSS
bcprhyd     30..GIEPLVTLFHWDVPQALVDEYDGLLS..24..WTTLNEPYTVSNHGYTIGIHAPG RCSC
bcamhyd     30..GLKPFVTIYHWDLPQALEDEYGGFLS..24..WITLNEPYTFSSSGYAYGVHAPG RCSA
csvlinhyd   30..GLEPFVTIFHWDTPQALQDKYGGFLS..24..WMTFNEPSAYVGFAHDDGVFAPG RCSS
clvlinhyd   30..GMQPYVTLFHWDVPQALEDEYRGFLG..24..WITLNEPWGVSMNAYAYGTFAPG RCSD
clvncbglc   30..GIQPFVTLFHWDLPQVLEDEYGGFLN..24..WSTLNEPWVFSNSGYALGTNAPG RCSA
Dcobgluc    30..GITPFVTIFHWDLPQALEDETGGFLN..24..WITLNEPSIFTANGYAYGMFAPG RCSP
Cspfbgluc   30..GIRPMVTLFHWDVPQALEDSYKGFRS..24..WITLNEPWSLSTMGYAFGRHAPG RCST
csvncbglc   30..GMEPFVTIFHWDTPQAIEDKYGGFLS..24..WMTFNEPWSLSGFAYDDGVFAPG RCSS
Arabthio3   30..GITPFVTVFHWDTPQDLEDEYGGFLS..24..WITFNEPWVFLHAGYDVGKKAPG RCSS
rapebgluc   30..DLTPLVTIFHWDMPADLEDEYGGFLS..24..WITFNEPWVFSRSAYDVGKKAPG RCSP
cornbgluc   30..GIEPYVTIFHWDVPQALEEKYGGFLD..27..WLTFNEPQTFTSFSYGTGVFAPG RCSP
sorghdns    30..GIEPYITIFHWDTPQALVDAYGGFLD..24..WLTFNEPETFCSVSYGTGVLAPG RCSP
barlbgluc   28..GITPYANLYHYDLPLALHQQYLGWLS..24..WFTFNEPRVVAALGYDNGFHAPG RCSK
oatbgluc    30..GIKPYITLFHWDTPQALADEYKDFLD..24..WFTFNEPHSFCGLGYGTGLHAPGARCSA
PRIMERS         <bgluW158r                                  <bgluC222r
```

Protein sequence alignment and phylogenetic analysis of known β-galactosidases grouped all known plant β-galactosidases into glycosyl hydrolase family 35. A phylogenetic tree and partial multiple sequence alignment for family 35 are shown in fig. 2. All plant enzymes grouped together and conserved sequences back-translated to design degenerate oligonucleotides are indicated in bold print.

Fig 3. (continued)

Protein sequence analysis of the b-glucosidases indicated all known plant enzymes fall into glycosyl hydrolase family 1. As shown in fig. 3, this family also contained the plant thioglucosidases, which group very closely to the dicot and monocot b-glucosidases. Degenerate primers were designed from the sequences shown in bold in the partial multiple sequence alignment of the plant enzymes (fig. 3b). These sequences were based primarily on sequences from the group of dicot b-glucosidases. Note the primer sequences were less conserved than the b-galactosidase and b-mannosidase sequences.

Table 1 : Products of PCR with Degenerate Primers

Protein	Plant RNA	Predicted Product Size (bp)	Actual Product Size (bp)
α-mannosidase (lysozomal-type)	rice	110-120	80, 120, 160
α-mannosidase (Golgi - type II)	rice, bean	330	no specific bands
β-galactosidase	rice, bean, *D. cochinchinensis*	390	390
β-glucosidase	bean, rice, *D. cochinchinensis*	150	160 150

PCR was conducted as described in the methods. Single strand cDNA's were first amplified with the most N-terminal forward primer and the most C-terminal reverse primer (the outer primers). A 1 µl aliquot of the product or a 1/10 dilution of it was used for the second round of PCR with the C-terminal forward and N-terminal reverse primers (the nested primers). Nested PCR product sizes were predicted by taking the median length of the known cDNA sequences in the multiple sequence alignments.

Results

Protein Sequence Analysis

Protein sequence alignment and phylogenetic analysis indicated that most known α-mannosidases fall into either glycosyl hydrolase family 38 or 47. The only

PCR of cloned β-glucosidase and reverse-transcribed plant RNA

In order to test the ability of the β-glucosidase primers, which were the most divergent, to PCR amplify cDNA from β-glucosidase. A plasmid containing the cloned full length β-glucosidase cDNA from *D. cochichinensis* was used as a substrate for PCR. The cloned cDNA did not exactly match that for the oligonucleotide primers, since as seen in fig. 3, the cDNA had a valine instead of glycine in the 3rd amino acid of the first forward primer (bgluG62f), and a methionine instead of phenylalanine in the fourth amino acid position of the second forward primer (bgluA109f). These differences would result in nucleotide differences of T instead of G in the first case, and A instead of T and G instead of T or C in the second case. Nonetheless, the outer primers amplified the expected 484 bp product and the inner primers the expected 150 bp product from the plasmid. PCR of reverse-transcribed *D. cochinchinensis* seed pod RNA, germinated rice RNA, and germinated bean RNA produced similar size bands after nested PCR.

Table 1 indicates the size products observed in nested PCR amplifications of reverse-transcribed RNA's with α-mannosidase, β-galactosidase, and β-glucosidase primers. Specific, expected size bands were seen when nested PCR was performed with lysosomal α-mannosidase, β-galactosidase, and β-glucosidase primers, but not with golgi α-mannosidase primers. In all cases, controls in which either of the primers or the template was left out failed to produce the specific bands of the expected size.

Conclusion

Protein sequence analysis demonstrated that the known plant β-glucosidase, β-galactosidase, and α-mannosidase fall into discreet subgroups within glycosyl hydrolase families 1, 35, and 38, respectively. Within these subgroups, sequences were found that were sufficiently conserved for design of oligonucleotide primers. The most degenerate of the primers designed were from the plant β-glucosidases. These primers were tested for viability by amplifying correct size bands from a plasmid containing a β-glucosidase from *D. cochinchinensis*. Despite mismatching the sequence of the cDNA at one position each, the outer set of primers amplified the expected 482 bp major band from the cloned cDNA. The inner set of primers, amplified the expected 150 bp product as the most predominant band. When *D. cochinchinensis* RNA was reverse transcribed and used as template for nested PCR, the same size product was seen, and similar size products were seen

with rice and bean RNA. Thus, it appears that the primers designed should be useful for PCR amplification of β-glucosidase cDNA fragments from a variety of plants.

Similarly, β-galactosidase and lysosomal α-mannosidase primers were able to amplify the expected size bands from reverse transcribed plant RNA. In the case of the β-galactosidase, this was the only predominant band when proper conditions were discovered. However, in the case of the α-mannosidase, other bands larger and smaller than the expected size were encountered at similar intensities under conditions tried to date. These bands appear specific for both primers and the DNA template, so it is possible that they are all α-mannosidases or related proteins. If conditions cannot be found to produce a single size product, cloning and sequencing of these bands will be necessary to resolve whether they are actually α-mannosidases or related proteins.

The degenerate primers designed from the golgi α-mannosidases have yet to produce bands not also seen in the negative controls. This may not be surprising, since this group did not contain any sequences from plants, but only from diverse animal species. It is likely that plants also have this type of protein, since it appears to diverged from the lysosomal α-mannosidases before the plants and animals diverged based on the phylogenetic tree (fig. 1A). However, it is uncertain which of the conserved sequences found in the animals will also be conserved in the plants, so the oligonucleotides may not contain to plant sequences.

This screening approach based on primary structure of known proteins and nested PCR appears to be applicable to cloning of β-glucosidases, β-galactosidases, and α-mannosidases from plants. Cloning and sequencing of the RT-PCR products produced from plant RNA's will be necessary to confirm that these are indeed derived from the appropriate enzyme cDNA. However, at least for the β-glucosidases and β-galactosidases, the presence of a single major band with the expected size in electrophoresis of PCR products, provides strong evidence that the proper products are being produced. This should allow rapid cloning and sequencing of the proteins, which can then be expressed in recombinant systems to evaluate their activities. Though determination of the functional information may be slower than functional screens, the primary structure should be determined much more quickly. Thus, the primary-structure based PCR screen should provide a useful compliment to the substrate-based activity screening for useful glycosyl hydrolases[1].

Acknowledgment

This work was supported by grants from the Thailand Research Fund (RSAOII/39) and Suranaree University of Technology.

References

1. Bauer, M.W., Bylina, E.J., Sanson, R.V. and Kelly R.M. (1996) Comparison of a beta-glucosidase and a beta-mannosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. Purification, characterization, gene cloning, and sequence analysis. J. Biol. Chem. 271, 23749-23755.
2. Beccari, T., Appolloni, M.G., Costanzi, E., Stinchi, S., Stirling, J.L., Della Fazia, M.A., Servillo, G., Viola, M.P., Orlacchio, A. (1997) Lysosomal alpha-mannosidases of mouse tissues: characteristics of the isoenzymes, and cloning and expression of a full-length cDNA. Biochem J. , 45-49.
3. Bischoff, J., Moremen, K. and Lodish, H.F. (1990) Isolation, characterization, and expression of cDNA encoding a rat liver endoplasmic reticulum alpha-mannosidase. J. Biol. Chem. 265, 17110-17117.
4. Boizet, B., Villeval, D., Slos, P., Novel, M., Novel, G., Mercenier, A. (1988) Isolation and structural analysis of the phospho-beta-galactosidase gene from *Streptococcus lactis* Z268. Gene 62, 249-261.
5. Brzobohaty, B., Moore, I., Kristoffersen, P., Bako, L., Campos, N., Schell, J. and Palme, K. (1993) Release of active cytokinin by a β-glucosidase localized to the maize root meristem. Science 262, 1051-1053.
6. Burland, V.D., Plunkett, G., Daniels, D.L., Blattner, F.R. (1993) DNA sequence and analysis of 136 kilobases of the *Escherichia coli* genome: organizational symmetry around the origin of replication. Genomics 16, 249-261.
7. Cairns, J.R.K., Champattanachai, V., Srisomsap, C., and Svasti, J. (in preparation) Primary structure of the dalcochinin-8'-O-β-glucoside β-glucosidase from *Dalbergia cochinchinenzis* Pierre.
8. Carey, A.T., Holt, K., Picard, S., Wilde, R., Tucker, G.A., Bird, C.R., Schuch, W. and Seymour, G.G. (1995) Tomato exo-(1-4)-β-D-galactanase. Isolation, changes during ripening in normal and mutant tomato fruit, and characerization of a related cDNA clone. Plant Physiol. 108, 1099-1107.
9. Cicek, M. and Esen, A. (1998) Structure and expression of a dhurrinase (β-glucosidase) from sorghum. Plant Physiol. 116, 1469-1478.
10. Downs,C.G. and Almira,C. (1995) A beta-galactosidase (Genbank X84684) cDNA homolog from broccoli (*Brassica oleracea* L.) (PGR95-017). Plant Physiol. 108, 1342-1342.

11. Falk, A., Xue, J., Lenman, M. and Rask, L. Sequence of a cDNA clone encoding the enzyme myrosinase and expression of myrosinase in different tissues of *Brassica napus*. (1992) Plant Sci. 83, 181-186.
12. Falk, A., Ek, B. and Rask, L. (1995) Characterization of a new myrosinase in *Brassica napus*. Plant Mol. Biol. 27, 863-874.
13. Falk, A. and Rask, L. (1995) Expression of a zeotin-O-glucoside-degrading β-glucosidase in *Brassica napus*. Plant Physiol. 108, 1369-1377.
14. Felsenstein, J. (1989) PHYLIP Cladistics 5, 264.
15. Feng, D. F. and Doolittle, R.F. (1996) Progressive alignments of amino acid sequences and construction of phylogenetic trees from them. Meth. Enzymol. 266, 368-382.
16. Foster, J.M., Yudkin, B., Lockyer, A.E. and Roberts, D.B. (1995) Cloning and sequence analysis of GmII, a Drosophila melanogaster homologue of the cDNA encoding murine Golgi alpha-mannosidase II. Gene 154, 183-186.
17. Gus-Mayer, S., Brunner, H., Schneider-Poetsch, H.A. and Rudiger, W. (1994) Avenacosidase from oat: purification, sequence analysis and biochemical characterization of a new member of the BGA family of beta-glucosidases. Plant Mol. Biol. 26, 909-921.
18. Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 280, 309-316.
19. Henrissat, B. and Bairoch, A. (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 293, 781-788.
20. Henrissat, B. and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. Biochem. J. 316, 695-696.
21. Hughes, M.A., Brown, K., Pancoro, A., Murray, B.S., Oxtoby, E., and Hughes, J. (1992) A molecular and biochemical analysis of the structure of the cyanogenic β-glucosidase (linamarase) from cassava (*Manihot esculenta Crantz*). Arch. Biochem. Biophys. 295, 273-279.
22. Inoue, K., Shibuya, M., Yamamoto, K., and Ebizuka, Y. (1996) Molecular cloning and bacterial expression of a cDNA encoding furostanol glycoside 26-O-beta-glucosidase of *Costus speciosus*. FEBS Lett. 389, 273-277.
23. King, G.A. and Davies, K.M. (1995) Cloning of a harvest-induced beta-galactosidase from tips of harvested asparagus spears. Plant Physiol. 108, 419-420.
24. Kumar, V., Ramakrishnan, S., Teerit, T., Knowless, J.K., and Harley, B.S. (1992) Saccharomyces cerevisiae cells secreting an *Aspergillus niger* beta-galactosidase grow on whey permease. Biotechnology (N.Y.) 10, 82-85.
25. Leah, R., Kigel, J., Svendsen, I., and Mundy, J. (1995) Biochemical and molecular characterization of a barley seed β-glucosidase. J. Biol. Chem. 250, 15789-15797.
26. Li, P.L., Swain, E., and Poulton, J.E. (1992) *Prunus serotina* amygdalin hydrolase and prunasin hydrolase. Plant Physiol. 100, 282-290.
27. Liddle, S., Keresztessy, Z., Hughes, J. and Hughes, M.A. A genomic cyanogenic beta-glucosidase gene from Cassava (Accession No. X94986). (PGR98-148) Plant Physiol. 117, 1526-1526.
28. Liebl, W., Gabelsberger, J. and Schleifer, K.H. (1994) Comparative amino acid sequence analysis of Thermotoga maritima beta-glucosidase (BglA) deduced from the nucleotide sequence of the gene indicates distant relationship between beta-glucosidases of the BGA family and other families of beta-1,4-glycosyl hydrolases. Mol. Gen. Genet. 242, 111-115.
29. Little, S., Cartwright, P., Campbell, C., Preneta, A., Mochesney, J., Mountain, A. and Robinson, M. (1989) Nucleotide sequence of a thermostable beta-galactosidase from *Sulfolobus solfataricus*. Nucleic Acids Res. 17, 7980.
30. Maier, J., Wilitter, K., Gutlich, M., Ziegler, I., Werner, T. and Ninnemann, H. (1995) Homology cloning of GTP-cyclohydrolase I from various unrelated eukaryotes by reverse-transcription polymerase chain reaction using a general set of degenerate primers. Biochem. Biophys. Res. Commun. 212, 705-711.
31. Mantei, N., Villa, M., Enzler, T., Wacker, H., Boll, W., James, P., Hunziker, W. and Semenza, G. (1988) Complete primary structure of human and rabbit lactase-phlorizin hydrolase: implications for biosynthesis, membrane anchoring and evolution of the enzyme. EMBO J. 7, 2705-2713.
32. McPherson, M., Jones, K., and Gurr, S. (1992). PCR with highly degenernte primers, pp 171-186. In McPherson, M.J., Quirke, P., and Talylor, G.R. (eds.) PCR. A Practical Approach. Oxford University Press, Oxford, UK.
33. Misugo, M., Liao, Y.F., Kudo, S., Eto, S., Maitei, M.G., Moremen, K.W., and Fukuda, M.N.(1995) Molecular cloning and expression of cDNAs encoding human alpha-mannosidase II and a previously unrecognized alpha-mannosidase IIx isozyme. Proc. Natl. Acad. Sci. U S A. 92, 11766-11770.
34. Moremen, K.W. (1989) Isolation of a rat liver Golgi mannosidase II clone by mixed oligonucleotide-primed amplification of cDNA. Proc. Natl. Acad. Sci. U S A. 86, 5276-5280.
35. Moremen, K.W., Robbins, P.W. (1991) Isolation, characterization, and expression of cDNAs encoding murine alpha-mannosidase II, a Golgi enzyme that controls conversion of high mannose to complex N-glycans. J Cell Biol. 115, 1521-1534.
36. Nanba, E. and Suzuki, K. (1990) Molecular cloning of mouse acid beta-galactosidase cDNA: sequence, expression of catalytic activity and comparison with the human enzyme. Biochem. Biophys. Res. Commun. 173, 141-148.
37. Nebes, V.L. and Schmidt, M.C. (1994) Human lysosomal alpha-mannosidase: isolation and nucleotide sequence of the full-length cDNA. Biochem Biophys Res Commun. 200, 239-45.
38. Okamura, N., Iwaki, Y., Hiramoto, S., Tamba, M., Bannai, S., Sugita, Y., Syntin, P., Dacheux, F., Dacheux, J.L. (1997) Molecular cloning and characterization of the epididymis-specific glutathione peroxidase-like protein secreted in the porcine epididymal fluid. Biochim. Biophys. Acta 1336, 99-109
39. Oshima, A., Tsuji, A., Nagao, Y., Sakuraba, H. and Suzuki Y(1988) Cloning, sequencing, and expression of cDNA for human beta-galactosidase. Biochem. Biophys. Res. Commun. 157, 238-244.
40. Oxtoby, E., Dunn, J.A., Pancoro, A., and Hughes, M.A. (1991) Nucleotide and derived amino acid sequence of the cyanogenic β-glucosidase (linamarase) from white clover (*Trifolium repens L.*). Plant Mol. Biol. 17, 209-219.
41. Pluthero, F.G. (1993) Rapid purification of high-activity Taq DNA polymerase. Nucleic Acids Res. 21, 4850-4851.
42. Porter, E.V. and Chassy, B.M. (1988) Nucleotide sequence of the

beta-D-phosphogalactoside galactohydrolase gene of *Lactobacillus casei:* comparison to analogous pbg genes of other gram-positive organisms. Gene 62, 263-276.

43. Poulton, J.E. and Jurk, S. (1996) Direct submission. GENBANK AC: 1236961.
44. Quigley, F., Dao, P., Cottet, A. and Mache, R. (1996) Sequence analysis of an 81 kb contig from *Arabidopsis thaliana* chromosome III. Nucleic Acids Res. 24, 4313–4318.
45. Raghothama, K.G., Lawton, K.A., Goldsbrough, P.B. and Woodson, W.R. (1991) Characterization of an ethylene-regulated flower senescence-related gene from carnation. Plant Mol. Biol. 17, 61-71.
46. Ross, G.S., Wegrzyn, T., MacRae, E.A. and Redgwell, R.J. (1994) Apple β-galactosidase; Activity against cell wall polysaccharides and characterization of a related cDNA clone. Plant Physiol. 106, 521-528.
47. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, New York.
48. Schatzle, J., Busbe, J. and Cardelli, J. (1992) Molecular cloning and characterization of the structural gene coding for the developmentally regulated lysosomal enzyme, alpha-mannosidase, in *Dictyostelium discoideum.* J Biol Chem. 267, 4000-4007.
49. Schmidt, K.P., Burrows, P.R., Davies, K.G., Kammerloher, W., Schaeffner, A.R., Buck, F., Cai, D. and Grundler, F.M.W. (1995) A root specific myrosinase in Arabidopsis responding to cyst nematode infection. EMBL Data Library X89413.1.
50. Schnetz, K., Toloczyki, C. and Rak, B. (1987) Beta-glucoside (bgl) operon of *Escherichia coli* K-12: nucleotide sequence, genetic organization, and possible evolutionary relationship to regulatory components of two *Bacillus subtilis* genes. J. Bacteriol. 169, 2579-2590
51. Srisomsap, C., Svasti, J., Surarit, R., Champattanachai, V. Sawangarutrakul, P., Boonpuan, K., Subhasitanont, P., Chokechaichamnankit, D. (1996) Isolation and characterisation of an enzyme with β-glucosidase and β-fucosidase activities from *Dalbergia cochinchinensis* Pierre. J. Biochem. 119, 585-590.
52. Surarit, R. Svasti, J. Srisomsap, C., Suginta, W., Khunyoshyeng, S., Nilwarangkoon, S., Harnsakol, P., and Benjavongkulchai, E. (1995) Screening of glycoshydrolase enzymes in Thai plant seeds for potential use in oligosaccharide synthesis J. Sci. Soc. Thailand 21, 293-303.
53. Surarit, R. Chiaba, S. Matsui, H., Svasti, J., Srisomsap, C. (1996) Chemical modification of β-glucosidase/β-fucosidase from *Dalberia cochinchinensis* Pierre by conduritol B epoxide. Biosci. Biotech. Biochem. 60(8), 1265-1268.
54. Swartz, R.M. and Dayhoff, M.O. (1978) Pam 250 matrix of accepted point mutations, p. 353 *In* Dayhoff, M.O. (ed.) Atlas of protein structure and sequence 5, supplement 3. National Biomedical Research Foundation, Washington, D.C. USA
55. Taron, C.H., Benner, J.S., Hornstra, L.J. and Guthrie, E.P. (1995) A novel beta-galactosidase gene isolated from the bacterium Xanthomonas manihotis exhibits strong homology to several eukaryotic beta-galactosidases. Glycobiology 5, 603-610.
56. Thangstad, O.P., Winge, P., Husebye, H. and Bones, A. (1993) The myrosinase (thioglucoside glucohydrolase) gene family in Brassicaceae. Plant Mol. Biol. 23, 511-524.
57. Tollersrund, O.K. Berg, T., Healy, P., Evjen, G., Ramachandran U. and Nilssen, O. (1997) Purification of bovine lysosomal alpha-mannosidase, characterization of its gene and determination of two mutations that cause alpha-mannosidosis. Eur. J. Biochem. 246, 410-419.
58. Vosman, B., Kuiken, G., Kooistra, J. and Venema, G. (1988) Transformation in *Bacillus subtilis:* involvement of the 17-kilodalton DNA-entry nuclease and the competence-specific 18-kilodalton protein. J. Bacteriol. 170, 3703-3710.
59. Wakarchuk, W.W., Greenberg, N.M., Kilburn, D.G., Miller, R.C. and Warren, R.A. (1988) Structure and transcription analysis of the gene encoding a cellobiase from Agrobacterium sp. strain ATCC 21400. J. Bacteriol. 170, 301-307.
60. Wiersma, P.A. and Fils-Lycaon, B.R. (1996) A cherry protein and its gene, abundantly expressed in ripening fruit, have been identified as thaumatin-like. Plant Physiol. 111, 269-273.
61. Wilson, R., Ainscough, R., Anderson, K, Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fraser, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laisster, N., Latreille, P., Lightning, J., Lloyd, C., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A., Saunders, D., Shownkeen, R., Sims, M., Smaldon, N., Smith, A., Smith, M., Sonnhammer, E., staden, R., Sulston, J., Thierry-Mieg, J., Thomase, K., Vaudlin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkinson-Sproat, J., and Wohldman, P. (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans.* Nature 368, 32-38.
62. Wright, R.M., Yablonsky, M.D., Shalita, Z.P., Glyal, A.K. and Eveleigh, D.E. (1992) Cloning, characterization, and nucleotide sequence of a gene encoding *Microbispora bispora* BglB, a thermostable beta-glucosidase expressed in *Escherichia coli.* Appl. Environ. Microbiol. 58, 3455-3465.
63. Xue, J.P., Lenman, M., Falk, A. and Rask, L. (1992) The glucosinolate-degrading enzyme myrosinase in Brassicaceae is encoded by a gene family. Plant Mol. Biol. 18, 387-398.
64. Xue, J.P., Jorgensen, M., Pihlgren, U., Rask, L. (1995) The myrosinase gene family in *Arabidopsis thaliana:* gene organization, expression and evolution. Plant Mol. Biol. 27, 911-922.
65. Yoshihisa, T. and Anraku, Y. (1989) Nucleotide sequence of AMS1, the structure gene of vacuolar alpha-mannosidase of Saccharomyces cerevisiae. Biochem. Biophys. Res. Commun. 163, 908-915.

Primary Structure of the Dalcochinin-8'-O-β-glucoside β-glucosidase from the Thai Rosewood *Dalbergia cochinchinensis* Pierre

James R. Ketudat Cairns*, Voraratt Champattanachai*, Chantragan Srisomsap*, and Jisnuson Svasti*[1]

Laboratory of Biochemistry, Chulabhorn Research Institute, Vipavadee-Rangsit Highway. Bangkok, 10210, Department of Biochemistry, Faculty of Science, Mahidol University, Bangkok, 10400 Thailand, and Institute of Science, Suranaree University of Technology, Nakhon Ratchasima, 30000 Thailand.

[1]Corresponding Author:
Jisnuson Svasti, Ph.D.
Dept. of Biochemistry,
Faculty of Science,
Mahidol University,
Bangkok 10400, Thailand.
Fax: (66-2)-2480375, 2596634
Tel: (66-2)-2461360 ext 4305 email: scjsv@mahidol.ac.th.

Abbreviations: dalcochinase, dalcochinin 8'-O-β-glucoside β-glucosidase; endo F, endoglycosidase-F; endo Lys C, endoproteinase lysine-C; NCBI, National Center for Biotechnological Information; PAGE, polyacrylamide gel electrophoresis; PMSF, phenyl methyl sulfonyl fluoride; pNP, para-nitrophenol; RACE, rapid amplification of cDNA ends.

Synopsis:

The dalcochinin-8'-O-β-glucoside β-glucosidase (dalcochinase) from the Thai Rosewood (*Dalbergia cochinchinensis* Pierre) has aglycone specificity for isoflavonoids and can hydrolyze both β-glucosides and β-fucosides. In order to determine its structure and evolutionary lineage, the sequence of the enzyme has been determined. N-terminal and proteolytic peptide sequences were determined by Edman degradation and used to design primers to amplify the cDNA for the protein by polymerase chain reaction (PCR) and rapid amplification of cDNA ends (RACE). The cDNA included a reading frame coding for 547 amino acids including a 23 amino acid propeptide and a 524 amino acid mature protein. The sequences determined at peptide level were found in the cDNA sequence, indicating the sequence obtained was indeed the dalcochinase enzyme. Eight putative glycosylation sites were identified and one was confirmed to be glycosylated by Edman degradation and mass spectroscopy. The dalcochinase sequence was homologous to sequences in glycosyl hydrolase family 1, indicating that dalcochinase is a member of this family. The mature enzyme is 60% identical to the cyanogenic β-glucosidase from white clover, for which an x-ray crystal structure has been solved. Based on this homology, several structural features can be postulated.

Introduction:

Beta-glucosidases (3.2.1.21) play important roles in plants, including growth regulation, response to stress, cellulose degradation, and defense[1]. They are thought to regulate plant growth by releasing cytokinin growth factors from their glucosides and help hydrolyze cellobiose produced from cell wall degradation. β-Glucosidases involved in defense generally produce toxic compounds by deglycosylation of their substrates. Although these enzymes are closely related evolutionarily and perform the same basic reaction, their substrate specificities are distinct[1,2]. Prunasin hydrolase, for instance, is unable to hydrolyze amygdalin, linustatin, and neolinustatin [3,4]. Myrosinases (3.2.3.1) are closely related evolutionarily to β-glucosidases and also function in defense by hydrolyzing β-S-glycosidic bonds between glucose and a toxic aglycone[5]. Other related β-glucosidases which perform diverse roles in plant growth [6] and phosphate starvation response[7], most likely have different aglycone specificities as well. These homologous enzymes show different substrate specificity, so small sequence differences must affect this specificity.

The wide variety of β-glucosidases and other glycosidases in plants also have several potential applications. Notably, they have been applied to synthesis of oligosaccharides by reverse hydrolysis[8] and mapping of carbohydrate linkages, including those in glycoproteins and glycolipids [9]. For this reason, the discovery of new glycohydrolases with novel substrate specificities is of great interest.
During screening studies for glycohydrolases in Thai plants using p-nitrophenol (pNP) glycosides, extracts of the seeds of the Thai Rosewood, *Dalbergia cochinchinensis* Pierre, were distinguished by high levels of β-fucosidase and β-glucosidase activities[10]. The enzyme was purified to a single band on native gels,

having both β-glucosidase and β-fucosidase activities. A single band was also obtained in SDS-PAGE with an apparent molecular weight of 66,000 Da [11]. Native enzyme appears to consist of 4 subunits and a pI of approximately 5.5.

Both the apparent K_m and the k_{cat} for the enzyme were lower for the pNP-β-fucoside compared to pNP-β-glucoside, while the other pNP-glycosides had much lower k_{cat} values[11]. A series of inhibitors had similar affects on β-glucosidase and β-fucosidase activities, suggesting that the two activities are located at the same active site. Modification of the purified enzyme with conduritol β-epoxide (CBE) inactivated β-fucosidase and β-glucosidase activities at essentially the same rate, indicating the presence of a carboxylic acid in the common active site [12].

Further studies showed that the enzyme does not efficiently hydrolyze commercially available natural β-glucosides, so the natural substrate was isolated from extracts of *D. cochinchinensis* seeds based on its cleavage by the enzyme[11,13]. The natural substrate, which was present at over 3% (by wt) of the seeds, was found to be a novel rotenoid β-glucoside, which we named dalcochinin-8′-O-β-glucoside. Due to its structural similarity to rotenone, a natural insecticide and piscicide, the aglycone was postulated to play a role in plant defense against herbivores, such as insects. However, the molecule is distinct from the cyanogenic β-glucosides, thought to play similar roles, as it is not cyanogenic and the respective β-glucosidases show little ability to hydrolyze each other's substrates.

Here, the primary structure for the *D. cochinchinensis* dalcochinin-8′-O-β-glucoside β-glucosidase (dalcochinase) was determined by peptide sequencing followed by RT-PCR amplification, cloning, and sequencing of the cDNA. The cDNA-derived protein sequence is shown to be closely related to other known plant

defense-related β-glucosidases in glycosyl hydrolase family 1 [2,14]. Based on the crystal structure of the white clover cyanogenic β-glucosidase [15], some structure-function relationships can be postulated.

Experimental:

Materials and Reagents

Dalbergia cochinchinensis Pierre seeds were kindly provided by the ASEAN-Canada Forest Seed Center at Muek Lek, Saraburi, Thailand. Endo-F/Peptide-N-Glycosidase F, pNP-glucose and pNP-fucose were purchased from Sigma Chemical Co. (St. Louis, MO, USA). Oligonucleotides were obtained from GIBCO-BRL (Life Technologies, Grand Island, NY, USA), Midland Certified Reagent Co. (Midland, TX, USA), and the BioServices Unit of the National Science and Technology Development Agency (NSTDA, BIOTECH, Bangkok, Thailand). Superscript II reverse transcriptase and the 3' RACE kit were purchased from GIBCO-BRL. Trypsin, Taq polymerase, deoxyribonucleotides, other PCR reagents, pGEM-T cloning vector, and Wizard 373 plasmid preparaton kit were purchased from Promega (Madison, WI, USA). Endopeptidase lysine-C, and terminal deoxynucleotide transferase were acquired from Boehringer Mannheim (Mannheim, Germany). A PRISM cycle sequencing kit from Perkin Elmer (Roche Molecular Systems, Branchburg, NJ, USA) was used in DNA-Sequencing. All other reagents were analytical grade or better.

Peptide Mapping and Protein Sequencing

The β-glucosidase/β-fucosidase enzyme was purified from *Dalbergia cochinchinensis* Pierre seeds as previously described[11]. Deglycosylation was

accomplished by incubating 0.5 mg protein at 37° C 24 h with 2.5 units Endoglycosidase-F/peptide-N-glycosidase F in 200 µL 10 mM sodium phosphate (pH 7.5), 25 mM EDTA, and 0.01% sodium azide (final pH 6.7). SDS-PAGE electrophoresis was used to confirm deglycosylation. Following deglycosylation, the protein was purified over an Aquapore RP-300 (220 x 4.6 mm) C8 column, eluted with a linear gradient from 0 to 70% acetonitrile in 0.1% TFA, on a Waters 510 HPLC. Products were detected by monitoring absorbance at 230 nm and 280 nm. Portions of the enzyme preparation were digested with Trypsin or Endoproteinase Lys-C. Trypsin digestion was accomplished by incubation at an enzyme: substrate ratio of 1:50 by wt for 12 h at 37° C in 50 mM Tris pH 7.85, 1 mM $CaCl_2$, 10% acetonitrile. Endo Lys C cleavage was performed at an enzyme: substrate ratio of 1:45 by wt in 50 mM Tris pH 9.0, 0.5 mM EDTA, 10% acetonitrile at 36° C for 12 hours. The reaction was stopped by addition of 10% TFA to lower the pH to 2.0. The cleaved peptides from each digest were separated using a LichroCART 250-4 HPLC cartridge [Lichrospher 100 RP-8 encapped (250 x 4 mm)]. Purified peptide fractions were dried by speed vacuum.

From 200 pmoles to 1 nmole of deglycosylated protein or digested peptides were dissolved in 50% acetonitrile in water, applied to a TFA-treated glass-fiber filter, dried under argon, and sequenced on an automated protein sequencer (Model 473A and 477 from ABI, CA, USA). Some fractions were also studied by mass spectroscopy using matrix-assisted laser desorption/ionization (MALDI) on a VG TOFSpec (Fisons Instruments, Manchester, UK) at the Max-Delbruck-Centrum fur Molekulare Medizin (Berlin, Germany). A 0.8 µL sample was mixed with 12 µL of a

matrix with 50 mmol sinapinic acid in ethanol, applied to the probe, ionized with a 337 nm nitrogen laser, and accelerated over a voltage of 22 kV.

Determination of Enzyme Activities in Thai Rosewood Tissues

Crude enzyme was prepared based on the method of Bednar [16]. Seeds were surface sterilized in 0.1% bleach, soaked in distilled water overnight, and germinated on moist cheese cloth. Young leaves and stems were collected from *Dalbergia cochinchinensis* Pierre trees, surface sterilized as above, and extracted within 1 h of collection. Tissue samples (0.3-0.5 g) were weighed, quick frozen in liquid nitrogen, and ground in a cold mortar and pestle. They were then mixed with 2 mL/g tissue cold 0.05 M sodium acetate, pH 5.0, 1 mM PMSF, and 5% (w/v) polyvinyl polypyrrolidine (PVPP), and filtered through 4 layers of cheese cloth. The homogenate was then centrifuged at 8940 g (4° C) for 30 min. Dowex 2X-8 resin was added to the supernatant up to 25% (w/v) and stirred 1 h at 4° C to adsorb phenolic compounds. The resin was removed by centrifugation at 8940 g for 30 min, and the supernatant was either immediately assayed or stored at -20° C until assay. Activities toward pNP-β-glucoside and pNP-β-fucoside were determined as previously described[11], using a Shimadzu UV-2100S spectrophotometer (Shimadzu Co., Tokyo, Japan) to measure the p-nitrophenol absorbance at 400 nm.

PCR Amplification and cDNA cloning

Protein peptide sequences were compared to the nr database for related sequences, using the BLAST facility of the National Center for Biotechnology Information at the USA National Institute of Health [17,18,19]. Related β-glucosidase sequences were downloaded and aligned with the programs of Feng and Doolittle[20] to determine relative positions of peptides in the protein. Degenerate

oligonucleotide primers were designed from the N-terminal sequence and the sequences of two internal peptides (TP7 and Lys9). The N-terminal sequence was back-translated to the For. 1 primer, while the TP7 peptide sequence, was converted to Rev. 1 primer and the Lys9 peptide sequence, to the Rev. 2 primer, shown in Table 1.

D. cochinchinensis seeds were germinated as described, and the pericarp and endosperm were removed. The remainder of the seedlings were ground in liquid nitrogen and RNA was extracted with Trizol reagent, according to the recommended procedure [21]. The RNA obtained was dissolved in 0.100 mL DEPC-treated water and 1-5 μg total RNA was reverse transcribed into single strand cDNA using Superscript II reverse transcriptase, with 2.5 μM reverse 2 primer (50 pmoles per 20 μ L), as described by the manufacturer. A 1 μL aliquot of the product was used as template for polymerase chain reaction (PCR) with 1 μM each of the forward 1 and reverse 1 primers, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Triton X-100, 0.2 mM of each dideoxy-nucleotide (dATP, dCTP, dGTP, and dTTP) and 2.5 units of Taq polymerase in a 50 μL reaction. The thermocycling was done in a Perkin-Elmer 480 thermocycler using the following program: 94° C 5 min, then 5 cycles of : 94° C 30 s, 37° C 30 s, and extension at 72° C 60 s; followed by 30 cycles of: 94° C 30 s, 44° C 30 s, and 72° C 60 s; and finishing with 7 min at 72° C.

Four oligonucleotide primers, For. 2, For. 3, Rev. 3, and Rev. 4, were made from the sequence of the intitial PCR product for use in rapid amplification of cDNA ends (RACE). A 3' RACE system kit from GIBCO-BRL was used for amplification of the 3' end of the cDNA. The first strand cDNA was produced as described above, but 10 pmoles of AP (poly T) primer provided in the kit was used for priming the synthesis. The RNA was degraded by incubating with RNase H 20 min at 37° C,

according to the recommended protocol. The cDNA was first amplified with 0.25 μM each For. 2 and AUAP primers (Table 1) in a 50 μL reaction, containing the other reagents as described above. The amplification was done by incubating 5 min at 94° C, followed by 35 cycles of: 94° C 1 min, 60° C 1 min, and 72° C 3 min. The reaction was extended for an additional 7 min at 72° C after the final cycle. A second amplification used 5.0 μL of the first reaction and 10 pmol For. 3 and AUAP primer, under the same conditions as the first amplification. The product was evaluated by 1% agarose gel electrophoresis according to standard methods[22] and gel purified using the Geneclean III (Bio 101 Co., Vista, CA, USA) method.

Similarly, the 5′ end of the cDNA was amplified using the same system for 5′ RACE. Synthesis of the first strand cDNA was primed using 1 μM Rev. 2 primer under the same conditions as 3′ RACE. After heat inactivation of the reverse transcriptase, the RNA was digested by adding 1 μL RNase H to the reaction and incubating at 37° C for 20 min. The first strand cDNA was then purified with Geneclean III and eluted in 10 μL sterile, distilled water. The product was tailed with dGTP using terminal deoxy-transferase (TdT) in a reaction containing 0.5 M potassium cacodylate, 25 mM Tris-HCl, pH 6.6, 0.25 mg/mL BSA, 0.1 mM DTT, 1 mM $CoCl_2$, 25 μM dGTP and 25 units TdT in a 20 μL volume. The reaction was incubated for 20 min at 37° C and heat inactivated at 70° C for 15 min. 10% of the product was then amplified by PCR with 1 μM each Oligo dC14-Sal I and Rev. 3 primers and the following program. After 5 min at 94° C, 5 cycles of denaturation at 94° C 1 min, annealing at 37° C 1 min, and elongation at 72° C 1 min were performed followed by 30 cycles of 94° C 1 min, 60° C 1 min, and 72° C 1 min. The product was extended 7 min at 72° C after the final cycle, as usual. A second