Comparative characterization of *Arabidopsis* Subfamily III β-galactosidases

Dashzeveg Gantulga

Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirement for the degree of

Doctor of Philosophy

In

Biological Sciences

Dr. Brenda S.J. Winkel	Committee Chair
Dr. David R. Bevan	Committee member
Dr. Richard A. Walker	Committee member
Dr. Zhaomin Yang	Committee member

December 5, 2008

Blacksburg, Virginia

Keywords: *Arabidopsis*, β-galactosidase, cell wall

Comparative characterization of *Arabidopsis* Subfamily III β-galactosidases

Dashzeveg Gantulga

Abstract

The *Arabidopsis* genome encodes 17 putative β−galactosidases belonging to Glycosyl Hydrolase (GH) family 35, which have been classified into seven subfamilies based on sequence homology. The largest of these, Subfamily III, consists of six genes, *Gal-1* (At3g13750), *Gal-2* (At3g52840), *Gal-3* (At4g36360), *Gal-4* (At5g56870), *Gal-5* (At1g45130), and *Gal-12* (At4g26140) that share 60-81% sequence identity at the amino acid level. All six proteins have a signal peptide that may target them to the cell exterior.

We report purification and biochemical characterization of all six members of Subfamily III, each expressed as a recombinant protein in *Pichia pastoris* and one also in native form, purified from *Arabidopsis* leaves, with a special emphasis on substrate specificities. Organ specific expression of the six *Gal* genes was examined by analysis of the microarray databases and by semi-quantitative RT-PCR. The relative abundance and size of the Gal-1, Gal-2, Gal-5, and Gal-12 proteins was studied by immunoblotting using isoform-specific anti-peptide antibodies. The protein expression patterns of the *Gal* genes were generally consistent with microarray and RT-PCR data, though some discrepancies were observed suggesting distinct mechanisms of regulation for transcription and translation. Localization of total β-galactosidase activity was visualized using the substrate, 5-bromo-4-chloro-3-indolyl-β-D-galatopyranoside (X-Gal), to stain whole plants. Subcellular localization of the four isoforms examined by immuno-dotblotting and western blotting showed that Gal-1, Gal-2, Gal-5 and Gal-12 are present in apoplastic and cell wall bound protein extracts. Immuno-EM analysis of Gal-1 and Gal-12 showed

that these proteins are localized in the cell walls of vascular and epidermal tissues in mature root. Taken together, the biochemical properties, expression patterns, and subcellular localization of these isozymes indicate that the Subfamily III β-galactosidases all have potential functions in restructuring the cell wall during plant growth and development.

Acknowledgement

First of all, I would like to thank Dr. Winkel for everything she had done for me. She has encouraged me throughout the graduate study and helped me reach the point where I am now. I would like to thank my committee members, Dr. Walker, Dr. Bevan, and Dr. Yang for valuable advices and critical comments over the course of my study here at Virginia Tech. I would like to thank members of the Esen's lab, past and present, for providing me help on various levels throughout the course of this degree program. I wish to extend my thanks to many people from Department of Biological Science and other departments who have helped me technical side of the research. I could not have come this far without the tremendous support and constant love of my husband and our wonderful children. This was a great adventure and certainly life changing experience. Thank you. I am also grateful to my parents, brothers, sister and brother and sisters in-law for their love and support.

LIST OF ABBREVIATIONS

6BNG	6-bromo-naphthyl-β-D-galactopyranoside
CM	Accell plus carboxymethyl
ConA	Conconavalin A
Gal	β-galactosidase
GH	Glycosyl hydrolase
IPTG	isopropyl-β-D-galactopyranoside
BMGY	Buffered complex glycerol media
BMM(H)	Buffered minimal methanol media (supplied with histidine)
MALDI-TOF	Matrix-assisted laser-desorption-ionization–time-offlight
4-MUGal	4-methylumbellriferyl-β-D-galactopyranoside
nGal-2	native Gal-2
N-link	Asparagine linked glycosylation site
*o*NPGal	*ortho*-Nitrophenyl-β-D-galactopyranoside
PBS	phosphate buffered saline
PBST	phosphate buffered saline with tween-20
TBST	tris buffered saline with tween-20
TBSTh	tris buffered slaine with tween and high salt
*p*NPAra	*para*-Nitrophenyl-α-L-arabinopyranoside
*p*NPGal	*para*-Nitrophenyl-β-D-galactopyranoside
*p*NPGlc	*para*-Nitrophenyl-β-D-glucopyranoside
*p*NPFuc	*para*-Nitrophenyl-β-D-fucopyranoside
*p*NPMan	*para*-Nitrophenyl-β-D-mannopyranoside

*p*NPXyl	*para*-Nitrophenyl-β-D-xylofuranoside
TBG	Tomato β-galactosidase
TEM	Transmitting electron microscopy
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galatopyranoside

Table of Contents

Summary of Figures

Chapter 1

Chapter 2

Chapter 3

Summary of Tables

Chapter 2

Chapter 3

Chapter 1

Literature Review

Figure 1 reprinted from Structure 3, Davies, G., and Henrissat, B., 1995a. "Structures and mechanisms of glycosyl hydrolases," pages 853-859, with permission from Cell Press.

Figure 3 reprinted from Nature Review Molecular & Cell Biology 6, 2005, Cosgrove, D. J. "Growth of the plant cell wall," pages 850-861, with permission from Nature.

1.1. Classification of β-galactosidases and their catalytic mechanism

Carbohydrates play many important biological and physiological roles in living organisms, and they occur in different forms such as glycoproteins, glycolipids, glycosides, and oligo- and polysaccharides. Selective hydrolysis of glycosidic bonds specifies the breakdown of these glyco-conjugates and their function in subsequent processes such as activation of secondary metobolites, cell wall expansion and degradation, and turnover of signaling molecules. Living organisms have evolved a great diversity of glycosyl hydrolases (GH) with different substrate preferences and kinetic parameters, operating under different optimal conditions (Davies and Henrissat, 1995; Henrissat, 1995). GH enzymes are found in all animals, plants and microorganisms and have been classified into 113 families on the basis of amino acid sequence similarity (CAZY database, http://www.cazy.org; (H. J. Gilbert, 1999; Henrissat, 1991a, b; Henrissat and Bairoch, 1993, 1996; Henrissat et al., 2001; Henrissat and Davies, 1997).

GHs play many important roles in biological processes such as chemical defense, regulation, growth and development. Among the members of the different GH families, there is a group of enzymes with β-galactosidase activity (enzyme classification code: EC 3.2.1.23) that hydrolyze the β-D-galactosidic bonds of lactose, oligosaccharides, polysaccharides, and galactosides conjugated with aglycone moieties, as well as those of artificial substrates conjugated with chromogens such as *p*NPGal and X-Gal. This group is distributed among four distinct GH families, GH-1, GH-2, GH-35 and GH-42. Phylogenetic analysis shows that these families evolved independently via separate lineages (Javeri and Uhlenbruck, 1984). The β-galactosidases found predominantly in microorganisms belong to GH-1, GH-2, and GH-42,

whereas most plant β-galactosidases belong to GH-35.

Enzymatic hydrolysis of the glycosidic bond by GH enzymes takes place via general acid catalysis that requires two critical residues, a proton donor and nucleophile base. Hydrolysis occurs via either an inverting or a retaining mechanism. With the former, a product with opposite stereochemistry to the substrate is generated. β-galactosidases hydrolyze glycosidic bonds by the retaining mechanism (Figure 1), where the substrate and the product have identical stereochemical forms. All β-galactosidases fall into GH Clan A where catalytic acid and nucleophilic base are found at the carboxyl ends of β strands 4 and 7, respectively (Davies and Henrissat, 1995). The catalytic nucleophiles have been identified for the β-galactosidases from, *E. coli, Xanthamonas manihotas*, *Bacillus circulans, and* human lysosome (Blanchard et al., 2001; Juers et al., 2003; McCarter et al., 1997). Both the acid/base and nucleophile have been putatively identified in the crystal structure of the *Penicillium* sp. β-galactosidase (Rojas et al., 2004).

In addition to their critical biological functions in cells, β-galactosidases also have numerous important practical applications, including serving as a reporter gene in genetic engineering and for the hydrolysis of lactose or production of probiotics in the dairy and food industries (Chantarangsee etal. 2007). Because of their importance in medicine, agriculture, and the biotechnology industry, interest in β-galactosidases is steadily growing. For instance, in the food industry, β-galactosidases are widely used for prolonging the post harvest shelf life of fruits (Smith and Gross, 2000).

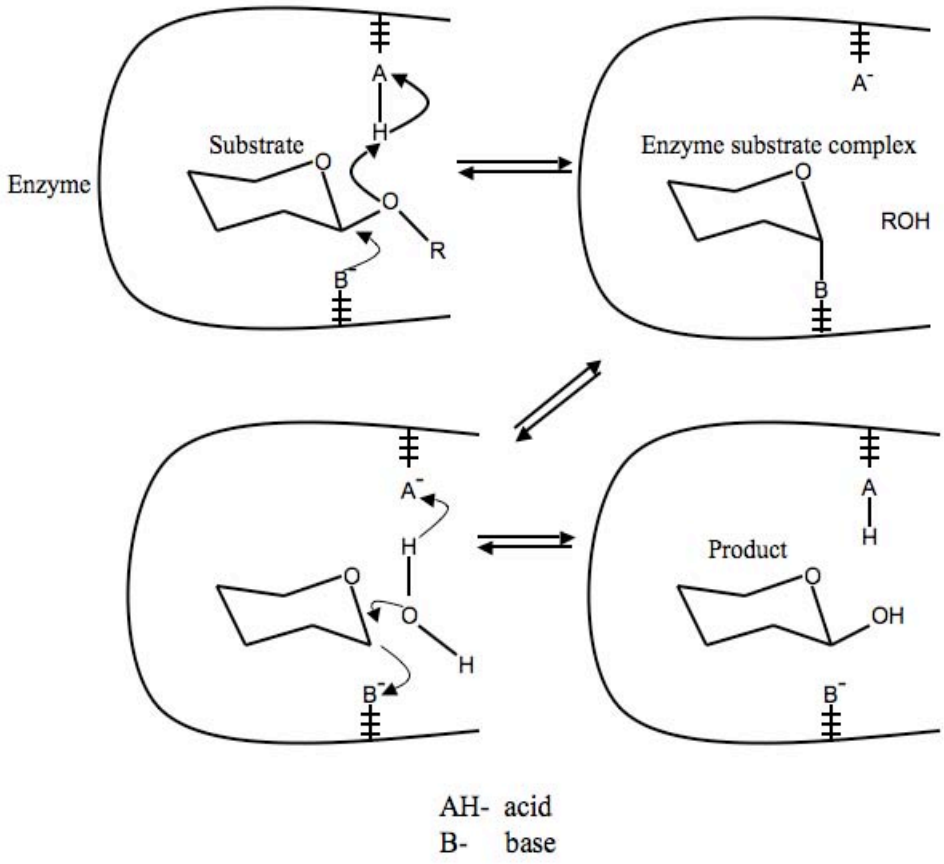


Fig. 1 Retaining mechanism of enzymatic hydrolysis of the glycosidic bond. Glycosidic oxygen is protonated by the acid catalyst (AH) and nucleophile assistance to algycone departure is provided by the base B^-. The resulting glycosyl-enzyme (enzyme substrate complex) is hydrolyzed by a water molecule and the second nucleophile substitution at the anomeric carbon generates a product with the same stereochemistry as the substrate. Reprinted with permission from Davies and Henrissat (1995).

1.2. β-Galactosidases in different organisms

Microbial β-galactosidase. Among microbial β-galactosidases, the most extensively studied enzymes are the bacterial β-galactosidases. Of these, *E. coli* β-galactosidase LacZ has the long and distinguished history in molecular biology and biotechnology. LacZ is a member of GH-2 and does not share high sequence similarity with other β-galactosidases. Detailed studies of the catalytic mechanism and three-dimensional structure of LacZ have been carried out (Juers et al., 2003). *E. coli* LacZ is a retaining β-glycosidase with a native homotetrameric structure. This enzyme allows the bacterum to hydrolyze lactose for use as an energy and carbon source. It is believed that other bacterial and fungal β-galactosidases may have roles similar to LacZ (Brady et al., 1995; Rojas et al., 2004). Besides its natural substrate, LacZ can hydrolyze artificial chromogenic substrates such as X-Gal and *o*NPGal (Juers et al., 2003). Another well-characterized bacterial enzyme is the thermostable enzyme GH-42 β-galactosidase (A4-β-Gal) from *Thermus thermophilus* A4 which is an extreme thermophilic organism (Hidaka et al., 2002). This enzyme is an attractive tool for biotechnology as it can be used for lactose processing in dairy products. It is desirable to have enzymes with greater stability that will improve the overall efficiency and reduce the cost of such process.

Mammalian β-galactosidase. Mammalian β-galactosidases have also been extensively studied, largely because many diseases are caused by deficiencies in these enzymes in humans and other mammals. There are two types of mammalian β-galactosidases, including GH-35 lysosomal β-galactosidase and GH-1 lactase phlorizin hydrolase (LPH). Lysosomal β-galactosidases are involved in degradation of glycolipids, glycoproteins, and proteoglycans, while intestinal lactase

in complex with phlorizin hydrolase (LPH) plays a role in digestion of milk lactose.

Diseases such as gangliosidosis and Morquio syndrome B, which lead to the accumulation of G_{M1} ganglioside and its derivative asialo-G_{M1} ganglioside (GA1), glycoprotein-derived oligosaccharides, and keratan sulfate, are caused by inherited deficiencies of lysosomal β-galactosidases (Callahan, 1999; Hoogeveen et al., 1983). Because of the high frequency of these genetic disorders in the human population, lysosomal β-galactosidases have been extensively studied (Callahan, 1999; Galjaard et al., 1987; Gerich, 1969; Hoogeveen et al., 1983; Nishigaki and Okada, 1998; Okada and O'Brien, 1968; Suzuki, 2006). Lysosomal β-galactosidase is synthesized as a precursor glycoprotein. After proteolytic cleavage, a mature protein is localized to the lysosome by endocytosis. The mature protein has the same acidic pH optimum and kinetic parameters as the precursor. Some mutations are responsible for defective enzymes with incorrect folding and/or susceptibilities for proteolytic degradation of mature proteins. Other mutations affect the actual active site residues, leading to inactive lysosomal β-galactosidases (Hoogeveen et al., 1983). Besides their natural substrates such as G_{M1} ganglioside, lactosyl ceramide, sialyloligosaccharide, glycopeptides, and mucopolysaccharides, these enzymes exhibit a wide range of activities toward artificial substrates such as *o*NPGal and X-Gal (DiCioccio et al., 1984; Fraser et al., 1983; Hiraiwa et al., 1986; Hiraiwa and Uda, 1986; Javeri and Uhlenbruck, 1984; Mutoh et al., 1988).

Small intestinal lactase-phlorizin hydrolase (LPH) hydrolyzes lactose, glycosyl-L-ceramides, and a number of aryl-L-glycosides including phlorizin (Zecca et al., 1998). In addition to hydrolyzing lactose, phlorizin, and a number of synthetic substrates, human and pig β-galactosidases have been found to have considerable activity with cellotriose and cellotetraose, and a low but significant activity against cellulose (Skovbjerg et al., 1982). Dietary flavonoid

and isoflavone glycosides are hydrolysed by lactase phlorizin hydrolase (Day et al., 2000). It is possible that this enzyme participates in the metabolism of these plant products when consumed by animals.

Plant β-galactosidases. β-galactosidases are found throughout the plant kingdom and play key roles in fruit ripening, flower senescence, mobilization of carbohydrate reserves, and galactolipid turnover, among other functions (Pressey, 1983, Bhalla and Dalling, 1984, Carey et al, 1995, Smith et al, 1998). β-galactosidases have been purified and characterized from a variety of plant sources, including ripening fruits and germinating seedlings of kiwifruit, tomato, apple, strawberry and mung bean (Ross et al. 1994, Li et al., 2001, Esteban et al., 2003, Trainotti et al., 2001). These studies also showed that β-galactosidases release stored energy for growth, release free galactose during normal metabolic recycling of galactolipids, glycoproteins, and the cell wall, and degrade the cell wall during senescence (Bhalla et al., 1984, Buckeridge et al 1994, Hall et al., 1998, Smith et al., 2000,). β-galactosidases also participiate in cell wall modification during elongation and differentiation of plant cells (Iglesias et al., 2006; Martin et al., 2008).

GH family 35 includes plant β-galactosidases that are thought to be involved in cell wall biogenesis and modification (Jamet et al., 2006, Vervelen and Vissenberg, 2007). In plants, β-galactosidase isozymes are encoded by multigene families. Smith et al. (2000) showed that a family of at least seven β-galactosidases is expressed during tomato fruit development. Most of these enzymes are large glycoproteins ranging in size from 89 to 97 kDa, with the exception of TBG4 whose molecular mass is ~77 kDa. All seven of the β-galactosidases purified from tomato were able to hydrolyze the synthetic substrate, *p*NPGal. Recombinant TBG4 expressed in yeast was also able to hydrolyze *p*NPGal, lactose, although poorly, as well as chelator soluble pectin, alkali soluble pectin, hemicellulose fractions from tomato cell walls, and commercially-prepared

galactan to different extents (Smith et al. 2000). This suggests that galactan side chains in the tomato cell wall are natural substrates of this enzyme and that TBG4 participates in controlling the dynamics of these polysaccharides.

Li et al (2001) purified five isoforms of β-galactosidases from mung bean and showed that these isoforms differ in physical-chemical and enzymatic characteristics. The mung bean enzymes did not hydrolyze lactose; instead they were partially inhibited by lactose and D-galactose. This indicates that these enzymes are highly specific for plant-derived substrates. An extensive study has been carried out on strawberry β-galactosidases. Trainotti et al. (2001) isolated three cDNAs from strawberry fruits encoding putative β-galactosidases. The enzymes were expressed as recombinant proteins in *Pichia pastoris* and shown to exhibit β-galactosidase activity with *p*NPGal and *o*NPGal. This study also showed that β-galactosidase activities in strawberry fruits are not correlated with the expression patterns of these enzymes at the transcriptional level (Trainotti et al., 2001). It had also been found that the enzyme activity of tomato isoform II (TBG4) increased while mRNA level declined in fruit (Carey et al., 1995; Smith and Gross, 2000). A possible explanation is that at given developmental stages the enzyme becomes more stable but the expression level decreases.

Amino acid sequence analysis of plant β-galactosidases revealed that these enzymes are divided into two subgroups with different polypeptide sizes. The larger β-galactosidases possess well-conserved C-terminal lectin-like SEUL (sea urchin egg lectin) type domains that are approximately 100 amino acids long and contain seven highly conserved cysteine residues. A 100-amino acid long C-terminal domain is common to many β-galactosidases. These lectin-like domains show homology to animal lectin proteins rather than plant lectins. SEUL lectins from sea urchin eggs have been shown to be L-rhamnose- and D-galactose-specific homodimers

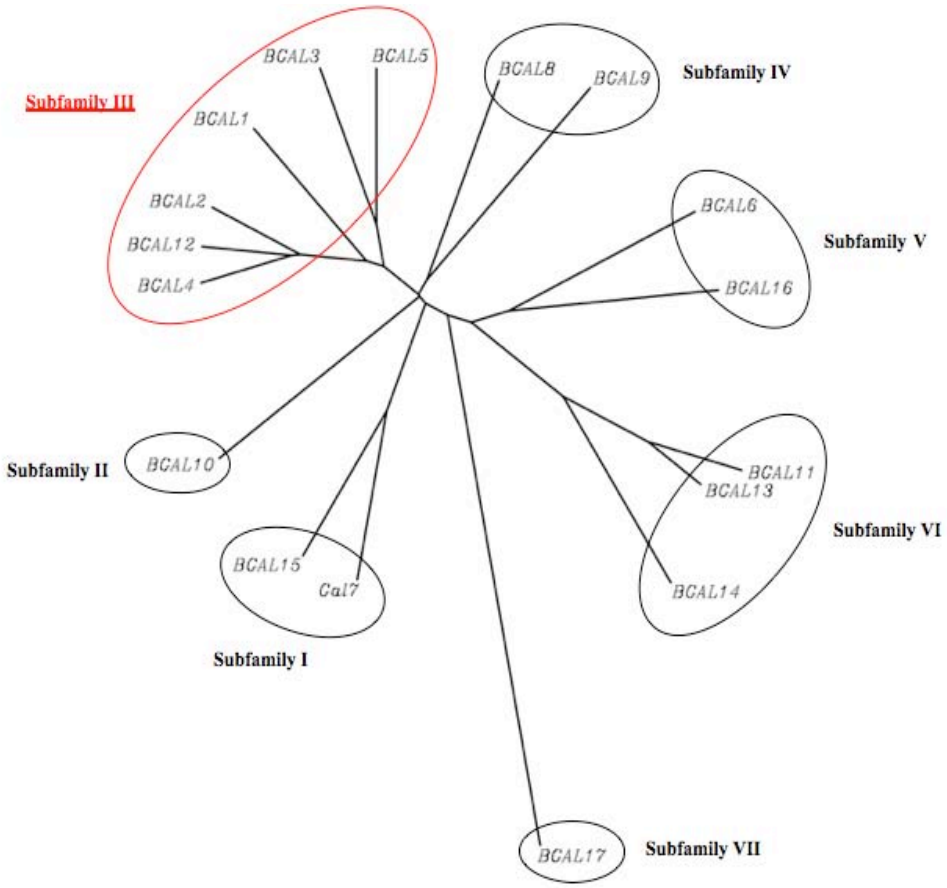
(Ozeki et al., 1995). Although there are no experimental data on the specific function of this domain in plants (Tanthanuch et al., 2008), it has been proposed that the lectin-like domain could enhance the affinity of the enzymes for their substrates, thereby increasing catalytic efficiency (Ahn et al., 2007) and possibly also enzyme stability (Trainotti et al. 2001).

1.3. *Arabidopsis thaliana* β-galactosidase

The National Scientific Foundation launched the Arabidopsis 2010 Program in 2001, one of the objectives was to determine the functions of all predicted genes in the recently-completed genome sequence. As a part of this effort, the "Functional genomics of the Arabidopsis β-galactosidase gene family" project aimed to determine the functions of all β-galactosidase genes in this model plant.

A family of 17 putative β-galactosidase genes has been identified in *Arabidopsis thaliana*, clustering into seven subfamilies based on sequence similarity (Figure 2) (Ahn et al., 2007; Gantulga et al., 2008). All of the *Arabidopsis* β-galactosidase proteins except Gal-14 have an N-terminal signal peptide, which may target the enzymes through the vesicular pathway to the outside of the cell, consistent with a proposed role in cell wall modification. Most of these enzymes appear to be glycoproteins with 1-9 potential N-glycosylation sites. Ten of the β-galactosidases have the C-terminal SUEL type lectin-like domain described above.

Among the four subfamilies that comprise the β-galactosidase genes in *Arabidopsis*



A)

B)

```
BGAL2           -----MSMHFRNKAWIILAILCFSSLIHSTEAVTYDHKALIINGQRRILISGSIHYPRS 55
BGAL12          -----MGLNFREKAWILLGILCCSSLICSVKAIVTYDRKAVIINGQRRILLSGSIHYPRS 55
BGAL4           -----MVLNFRDKSCIFLAILCCLSLSCIVKASVSYDRKAVIINGQRRILLSGSIHYPRS 55
BGAL1           MGSKPNAMKNVVAMAAVSALFLLGFLVCSVSGSVSYDSRAITINGKRRILISGSIHYPRS 60
BGAL3           ---MGTGDSASRLILWFCLGFLILGVGFVQCG-VTYDRKALLINGQRRILFSGSIHYPRS 56
BGAL5           ---MGTTILVLSKILTFLLTTMLIGSSVIQCSSVTYDKKAIVINGHRRILLSGSIHYPRS 57
                                .             . *:** :*: ***:****:*********

BGAL2           TPEMWPDLIKKAKEGGLDVIQTYVFWNGHEPSPGNYYFQDRYDLVKFTKLVHQAGLYLDL 115
BGAL12          TPEMWPDLIQKAKDGGLDVIQTYVFWNGHEPSPGQYYFEDRYDLVKFIKVVQQAGLYVHL 115
BGAL4           TPEMWPGLIQKAKEGGLDVIETYVFWNGHEPSPGQYYFGDRYDLVKFIKLVHQAGLYVNL 115
BGAL1           TPEMWPDLIRKAKEGGLDVIQTYVFWNGHEPSPGKYYFEGNYDLVKFVKLVQQSGLYLHL 120
BGAL3           TPDMWEDLIQKAKDGGIDVIETYVFWNLHEPSPGKYDFEGRNDLVRFVKTIHKAGLYAHL 116
BGAL5           TPEMWEDLIKKAKDGGLDVIDTYVFWNGHEPSPGTYNFEGRYDLVRFIKTIQEVGLYVHL 117
                **:** .**:***:**:***:****** ****** * * .. ***:* * ::: *** .*

BGAL2           RIGPYVCAEWNFGGFPVWLKYVPGMVFRTDNEPFKIAMQKFTKKIVDMMKEEKLFETQGG 175
BGAL12          RIGPYVCAEWNFGGFPVWLKYVPGMVFRTDNEPFKAAMQKFTEKIVRMMKEEKLFETQGG 175
BGAL4           RIGPYVCAEWNFGGFPVWLKFVPGMAFRTDNEPFKAAMKKFTEKIVWMMKAEKLFQTQGG 175
BGAL1           RIGPYVCAEWNFGGFPVWLKYIPGISFRTDNGPFKAQMQRFTTKIVNMMKAERLFESQGG 180
BGAL3           RIGPYVCAEWNFGGFPVWLKYVPGISFRTDNEPFKRAMKGFTERIVELMKSENLFESQGG 176
BGAL5           RIGPYVCAEWNFGGFPVWLKYVDGISFRTDNGPFKSAMQGFTEKIVQMMKEHRFFASQGG 177
                ********************:: *: ***** ***  *: ** :** :** ..:* :***
```

```
BGAL2           PIILSQIENEYGPMQWEMGAAGKAYSKWTAEMALGLSTGVPWIMSKQEDAPYPIIDTCNG 235
BGAL12          PIILSQIENEYGPIEWEIGAPGKAYTKWVAEMAQGLSTGVPWIMCKQDDAPNSIINTCNG 235
BGAL4           PIILAQIENEYGPVEWEIGAPGKAYTKWVAQMALGLSTGVPWIMCKQEDAPGPIIDTCNG 235
BGAL1           PIILSQIENEYGPMEYELGAPGRSYTNWAAKMAVGLGTGVPWVMCKQDDAPDPIINACNG 240
BGAL3           PIILSQIENEYGRQGQLLGAEGHNYMTWAAKMAIATETGVPWVMCKEDDAPDPVINTCNG 236
BGAL5           PIILSQIENEFEPDLKGLGPAGHSYVNWAAKMAVGLNTGVPWVMCKEDDAPDPIINTCNG 237
                ****:*****:      :*. *: * .*.*:** .  *****:*.*::*** .:*::***

BGAL2           FYCEGFKPNSDNKPKLWTENWTGWFTEFGGAIPNRPVEDIAFSVARFIQNGGSFMNYYMY 295
BGAL12          FYCENFKPNSDNKPKMWTENWTGWFTEFGGAVPYRPAEDIALSVARFIQNGGSFINYYMY 295
BGAL4           YYCEDFKPNSINKPKMWTENWTGWYTDFGGAVPYRPVEDIAYSVARFIQKGGSLVNYYMY 295
BGAL1           FYCDYFSPNKAYKPKMWTEAWTGWFTKFGGPVPYRPAEDMAFSVARFIQKGGSFINYYMY 300
BGAL3           FYCDSFAPNKPYKPLIWTEAWSGWFTEFGGPMHHRPVQDLAFGVARFIQKGGSFVNYYMY 296
BGAL5           FYCDYFTPNKPYKPTMWTEAWSGWFTEFGGTVPKRPVEDLAFGVARFIQKGGSYINYYMY 297
                :**: * **.  ** :*** *:**:*.***.:  **.:*:* .******:*** :*****

BGAL2           YGGTNFDRTAG-VFIATSYDYDAPIDEYGLLREPKYSHLKELHKVIKLCEPALVSVDPTI 354
BGAL12          HGGTNFDRTAG-EFIATSYDYDAPLDEYGLPREPKYSHLKRLHKVIKLCEPALVSADPTV 354
BGAL4           HGGTNFDRTAG-EFMASSYDYDAPLDEYGLPREPKYSHLKALHKAIKLSEPALLSADATV 354
BGAL1           HGGTNFGRTAGGPFIATSYDYDAPLDEYGLERQPKWGHLKDLHRAIKLCEPALVSGEPTR 360
BGAL3           HGGTNFGRTAGGPFVTTSYDYDAPIDEYGLIRQPKYGHLKELHRAIKMCEKALVSADPVV 356
BGAL5           HGGTNFGRTAGGPFITTSYDYDAPIDEYGLVQEPKYSHLKQLHQAIKQCEAALVSSDPHV 357
                :*****.****  *:::*******:***** ::**:.*** **:.** .* **:* :.  .

BGAL2           TSLGDKQEIHVFKSKT-SCAAFLSNYDTSSAARVMFRGFPYDLPPWSVSILPDCKTEYYN 413
BGAL12          TSLGDKQEAHVFKSKS-SCAAFLSNYNTSSAARVLFGGSTYDLPPWSVSILPDCKTEYYN 413
BGAL4           TSLGAKQEAYVFWSKS-SCAAFLSNKDENSAARVLFRGFPYDLPPWSVSILPDCKTEVYN 413
BGAL1           MPLGNYQEAHVYKSKSGACSAFLANYNPKSYAKVSFGNNHYNLPPWSISILPDCKNTVYN 420
BGAL3           TSIGNKQQAHVYSAESGDCSAFLANYDTESAARVLFNNVHYNLPPWSISILPDCRNAVFN 416
BGAL5           TKLGNYEEAHVFTAGKGSCVAFLTNYHMNAPAKVVFNNRHYTLPAWSISILPDCRNVVFN 417
                  :*  :: :*: : .  * ***:* . .: *:* * .  * **.**:******:.  :*

BGAL2           TAKIRAPT--ILMKMIPTSTK-FSWESYNEGSPSSNEAGTFVKDGLVEQISMTRDKTDYF 470
BGAL12          TAKVQVRTSSIHMKMVPTNTP-FSWGSYNEEIPSANDNGTFSQDGLVEQISITRDKTDYF 472
BGAL4           TAKVNAPS--VHRNMVPTGTK-FSWGSFNEATPTANEAGTFARNGLVEQISMTWDKSDYF 470
BGAL1           TARVGAQTSRMKMVRVPVHGG-LSWQAYNED-PSTYIDESFTMVGLVEQINTTRDTSDYL 478
BGAL3           TAKVGVQTS--QMEMLPTDTKNFQWESYLEDLSSLDDSSTFTTHGLLEQINVTRDTSDYL 474
BGAL5           TATVAAKTS--HVQMVPSGSILYSVARYDEDIATYGNRGTITARGLLEQVNVTRDTTDYL 475
                ** : . :       :*      .   : *  .:     ::   **:**:. * *.:**:

BGAL2           WYFTDITIGSDESFLKTGDNPLLTIFSAGHALHVFVNGLLAGTSYGALSNSKLTFSQNIK 530
BGAL12          WYLTDITISPDEKFL-TGEDPLLTIGSAGHALHVFVNGQLAGTAYGSLEKPKLTFSQKIK 531
BGAL4           WYITDITIGSGETFLKTGDSPLLTVMSAGHALHVFVNGQLSGTAYGGLDHPKLTFSQKIK 530
BGAL1           WYMTDVKVDANEGFLRNGDLPTLTVLSAGHAMHVFINGQLSGSAYGSLDSPKLTFRKGVN 538
BGAL3           WYMTSVDIGDSESFLHGGELPTLIIQSTGHAVHIFVNGQLSGSAFGTRQNRRFTYQGKIN 534
BGAL5           WYTTSVDIKASESFLRGGKWPTLTVDSAGHAVHVFVNGHFYGSAFGTRENRKFSFSSQVN 535
                ** *.: :  .* **  *. * * : *:***:*:*:** : *:::*  .  ::::   ::

BGAL2           LSVGINKLALLSTAVGLPNAGVHYETWNTGILGPVTLKGVNSGTWDMSKWKWSYK-IGLR 589
BGAL12          LHAGVNKLALLSTAAGLPNVGVHYETWNTGVLGPVTLNGVNSGTWDMTKWKWSYKQIGTK 591
BGAL4           LHAGVNKIALLSVAVGLPNVGTHFEQWNKGVLGPVTLKGVNSGTWDMSKWKWSYK-IGVK 589
BGAL1           LRAGFNKIAILSIAVGLPNVGPHFETWNAGVLGPVSLNGLNGGRRDLSWQKWTYK-VGLK 597
BGAL3           LHSGTNRIALLSVAVGLPNVGGHFESWNTGILGPVALHGLSQGKMDLSWQKWTYQ-VGLK 593
BGAL5           LRGGANKIALLSVAVGLPNVGPHFETWATGIVGSVVLHGLDEGNKDLSWQKWTYQ-AGLR 594
                *  * *::*:** *.****.* *:* *  *::*.* *:*:. *  *::  **:*:  * :

BGAL2           GEAMSLHTLAGSSAVKWWIKGFVVKK-QPLTWYKSSFDTPRGNEPLALDMNTMGKGQVWV 648
BGAL12          GEALSVHTLAGSSTVEWKEGSLVAKK-QPLTWYKSTFDSPTGNEPLALDMNTMGKGQMWI 650
BGAL4           GEALSLHTNTESSGVRWTQGSFVAKK-QPLTWYKSTFATPAGNEPLALDMNTMGKGQVWI 648
BGAL1           GESLSLHSLSGSSSVEWAEGAFVAQK-QPLTWYKTTFSAPAGDSPLAVDMGSMGKGQIWI 656
BGAL3           GEAMNLAFPTNTPSIGWMDASLTVQKPQPLTWHKTYFDAPEGNEPLALDMEGMGKGQIWV 653
BGAL5           GESMNLVSPTEDSSVDWIKGSLAKQNKQPLTWYKAYFDAPRGNEPLALDLKSMGKGQAWI 654
                **::.:   :  . : *   .:. :: *****:*: * :* *:.***:*:  ***** *:

BGAL2           NGHNIGRHWPAYTARGNCGRCNYAGIYNEKKCLSHCGEPSQRWYHVPRSWLKPFGNLLVI 708
BGAL12          NGQNIGRHWPAYTARGKCERCSYAGTFTEKKCLSNCGEASQRWYHVPRSWLKPTNNLVIV 710
BGAL4           NGRNIGRHWPAYKAQGSCGRCNYAGTFDAKKCLSNCGEASQRWYHVPRSWLKSQ-NLIVV 707
BGAL1           NGQSLGRHWPAYKAVGSCSECSYTGTFREDKCLRNCGEASQRWYHVPRSWLKPSGNLLVV 716
BGAL3           NGESIGRYWTAF-ATGDCSHCSYTGTYKPNKCQTGCGQPTQRWYHVPRAWLKPSQNLLVI 712
BGAL5           NGQSIGRYWMAF-AKGDCGSCNYAGTYRQNKCQSGCGEPTQRWYHVPRSWLKPKGNLLVL 713
                **..:**:* *: * *.*  *.*:* :  .**   **:.:********:***.  **:::
```

```
BGAL2           FEEWGGDPSGISLVKRTAK---------------------------------------- 727
BGAL12          LEEWGGEPNGISLVKRTAK---------------------------------------- 729
BGAL4           FEELGGDPNGISLVKRT------------------------------------------ 724
BGAL1           FEEWGGDPNGITLVRREVDSVCADIYEWQSTLVNYQLHASGKVNKPLHPKAHLQCGPGQK 776
BGAL3           FEELGGNPSTVSLVKRSVSGVCAEVSEYHPNIKNWQIESYGKGQTFHRPKVHLKCSPGQA 772
BGAL5           FEELGGDISKVSVVKRSVN---------------------------------------- 732
                :** **: . :::*:*

BGAL2           ------------------------------------------------------------
BGAL12          ------------------------------------------------------------
BGAL4           ------------------------------------------------------------
BGAL1           ITTVKFASFGTPEGTCGSYRQGSCHAHHSYDAFNKLCVGQNWCSVTVAPEMFGGDPCPNV 836
BGAL3           IASIKFASFGTPLGTCGSYQQGECHAATSYAILERKCVGKARCAVTISNSNFGKDPCPNV 832
BGAL5           ------------------------------------------------------------


BGAL2           ---------------------
BGAL12          ---------------------
BGAL4           ---------------------
BGAL1           MKKLAVEAVCA---------- 847
BGAL3           LKRLTVEAVCAPETSVSTWRP 853
BGAL5           ---------------------
```

Fig. 2 Phylogenetic analysis of *Arabidopsis thaliana* β-galactosidases. A) Phylogenetic tree of the β−galactosidase family derived from an alignment of the amino acid sequences of 17 members. The family is divided into seven subfamilies, of which Subfamily III, which includes Gal-1 through Gal-5 and Gal-12, is the largest. **B)** Alignment of Subfamily III enzymes that share 60-81% sequence identity with each other. Red color indicates putative active site residues and yellow color indicates conserved amino acid residues.

thaliana , subfamily III, which includes Gal-1, Gal-2, Gal-3, Gal-4, Gal-5 and Gal-12, is the largest. Most of the research on β-galactosidases has been focused on this subfamily because the genes are believed to be the most redundant and because the homologous genes have been studied for other organisms, which are believed to be involved in cell wall modification.

Plant cell walls are composed of independent but interacting networks of carbohydrates along with proteins, and aromatic substances (Carpita and Gibeaut, 1993; McCann and Roberts, 1991) (Fig. 3). Interacting with this complex matrix are several hundred enzymes and other proteins that carry out many functions, from wall assembly and disassembly to defense against pathogens. In recent years, several cell-wall structural proteins and enzymes, and their respective genes, have been identified (Henrissat et al., 2001; Richmond and Somerville, 2001). According to a recent microarray study, β-galactosidase genes are among these genes specific to cell wall dynamics (Imoto et al., 2005).

Analysis of several *Arabidopsis* β-galactosidase genes using promoter-GUS fusions uncovered differences in temporal and spatial expression among subfamily members (Perez, 2004). This study also showed that the galactosyl content of the cell wall is increased and total β-galactosidase activity is decreased in a *Gal-1* knockout mutant relative to wild-type plants. Iglesias et al. (2006) reported β-galactosidase activity on xyloglucan oligosaccharides in the apoplastic fluid in *Arabidopsis*. Proteomics analysis of apoplastic fluid from intact *Arabidopsis* rosettes identified numerous GH enzymes, among them Gal-1 (Boudart et al., 2005). All these studies indicate that β-galactosidases function in cell wall dynamics by helping to reorganize polysaccharide molecules during development. β-galactosidases also appear to function in degrading the plant cell wall during senescence (O'Donoghue et al., 2002). Polysaccharide

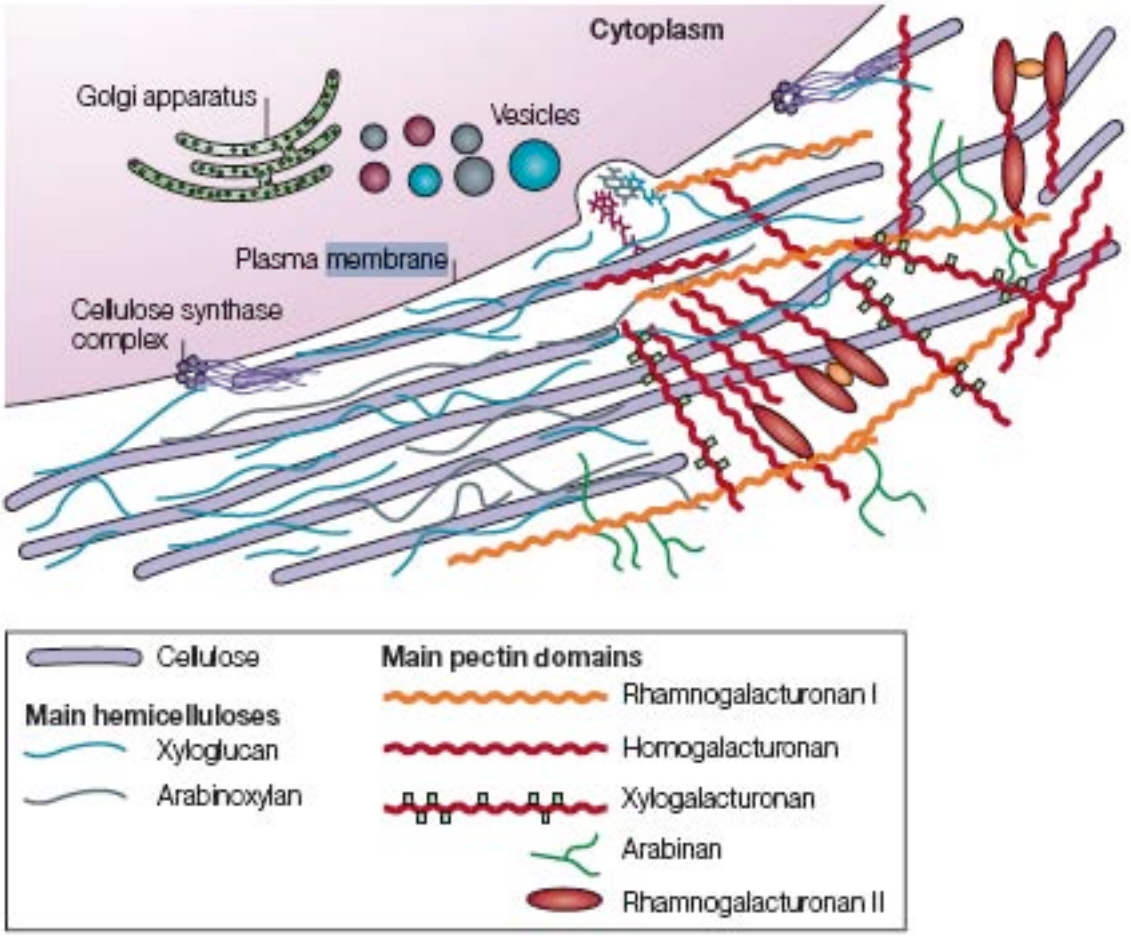


Fig. 3 A model of plant cell wall structure. The major components of plant cell walls include cellulose, hemicellulose, cross-linking glycans, pectin and proteins. Cellulose microfibrils are cross-linked through hemicellulose and pectin. Hemicellulosic xyloglucan and pectic rhamnogalacturonan contain side chains that are rich in galactan. Galactan is a linear polymer of galactose that is susceptible to hydrolysis by β-galactosidases. Reprinted with permission from Cosgrove (2005).

metabolism within the extracellular matrix allows flexibility in cell wall structure even in mature cell walls.

Although past experiments have advanced our understanding of β-galactosidases and suggested involvement of the subfamily III genes in cell wall dynamics, key questions remain unanswered, including the natural substrate specificities and subcellular localization of the gene products. Isolation and purification of native enzymes and recombinant proteins would greatly facilitate answering these questions. It is also desirable to study not a single gene but the entire subfamily to show whether the sequence similarities among the genes translate into similarities in function, and whether there is specialization among the enzymes for certain organs and tissues. The goal of the work described in this dissertation is to address these questions by systematic characterization of the *Arabidopsis* Subfamily III β-galactosidases.

References

Ahn, Y. O., Zheng, M., Bevan, D. R., Esen, A., Shiu, S. H., Benson, J., Peng, H. P., Miller, J. T., Cheng, C. L., Poulton, J. E., Shih, M. C., 2007. Functional genomic analysis of *Arabidopsis thaliana* glycoside hydrolase family 35. Phytochemistry 68, 1510-1520.

Blanchard, J. E., Gal, L., He, S., Foisy, J., Warren, R. A., Withers, S. G., 2001. The identification of the catalytic nucleophiles of two β-galactosidases from glycoside hydrolase family 35. Carbohydr Res 333, 7-17.

Boudart, G., Jamet, E., Rossignol, M., Lafitte, C., Borderies, G., Jauneau, A., Esquerre-Tugaye, M. T., Pont-Lezica, R., 2005. Cell wall proteins in apoplastic fluids of *Arabidopsis thaliana* rosettes: identification by mass spectrometry and bioinformatics. Proteomics 5, 212-221.

Brady, D., Marchant, R., McHale, L., McHale, A. P., 1995. Isolation and partial characterization of β-galactosidase activity produced by a thermotolerant strain of *Kluyveromyces marxianus* during growth on lactose-containing media. Enzyme Microb. Technol. 17, 696-699.

Callahan, J. W., 1999. Molecular basis of GM1 gangliosidosis and Morquio disease, type B. Structure-function studies of lysosomal beta-galactosidase and the non-lysosomal beta-galactosidase-like protein. Biochimica Et Biophysica Acta-Molecular Basis of Disease 1455, 85-103.

Carey, A. T., Holt, K., Picard, S., Wilde, R., Tucker, G. A., Bird, C. R., Schuch, W., Seymour, G. B., 1995. Tomato exo-(1-]4)-b-D-galactanase - isolation, changes during ripening in normal and mutant tomato fruit, and characterization of a related cDNA clone. Plant Physiol. 108, 1099-1107.

Carpita, N. C., Gibeaut, D. M., 1993. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. Plant J 3, 1-30.

Cosgrove, D. J., 2005. Growth of the plant cell wall. Nat Rev Mol Cell Biol 6, 850-861.

Davies, G., Henrissat, B., 1995. Structures and mechanisms of glycosyl hydrolases. Structure 3, 853-859.

Day, A. J., Canada, F. J., Diaz, J. C., Kroon, P. A., McLauchlan, R., Faulds, C. B., Plumb, G. W., Morgan, M. R., Williamson, G., 2000. Dietary flavonoid and isoflavone glycosides are hydrolysed by the lactase site of lactase phlorizin hydrolase. FEBS Lett 468, 166-170.

DiCioccio, R. A., Barlow, J. J., Matta, K. L., 1984. Purification of a β-D-galactosidase from bovine liver by affinity chromatography. Carbohydr Res 127, 109-120.

Fraser, N. L., Mahuran, D., Freeman, S. J., Callahan, J. W., 1983. Improved purification of β-galactosidase from rabbit brain: two separable fractions share kinetic and structural properties. Can J Biochem Cell Biol 61, 313-322.

Galjaard, H., Willemsen, R., Hoogeveen, A. T., Mancini, G. M., Palmeri, S., Verheijen, F. W., D'Azzo, A., 1987. Molecular heterogeneity in human beta-galactosidase and neuraminidase deficiency. Enzyme 38, 132-143.

Gantulga, D., Turan, Y., Bevan, D. R., Esen, A., 2008. The *Arabidopsis* At1g45130 and At3g52840 genes encode beta-galactosidases with activity toward cell wall polysaccharides. Phytochemistry 69, 1661-1670.

Gerich, J. E., 1969. Hunter's syndrome: beta-galactosidase deficiency in skin. N Engl J Med 280, 799-802.

H. J. Gilbert, D. G., B. Henrissat and B. Svensson, 1999. Carbohydrate-active enzymes: an integrated database approach. In: P. M. Coutinho & B. Henrissat (Ed.), Recent Advances in Carbohydrate Bioengineering. Cambridge, pp. 3-12.

Henrissat, B., 1991a. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J 280 (Pt 2), 309-316.

Henrissat, B., 1991b. Sequence homology between a beta-galactosidase and some beta-glucosidases. Protein Seq Data Anal 4, 61-62.

Henrissat, B., Bairoch, A., 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem J 293 (Pt 3), 781-788.

Henrissat, B., Bairoch, A., 1996. Updating the sequence-based classification of glycosyl hydrolases. Biochem J 316 (Pt 2), 695-696.

Henrissat, B., Coutinho, P. M., Davies, G. J., 2001. A census of carbohydrate-active enzymes in the genome of *Arabidopsis thaliana*. Plant Mol Biol 47, 55-72.

Henrissat, B., Davies, G., 1997. Structural and sequence-based classification of glycoside hydrolases. Curr Opin Struct Biol 7, 637-644.

Henrissat, G. D. a. B., 1995. Strucutures and mechanism of glycosyl hydrolases. Structure 3, 853-859.

Hidaka, M., Fushinobu, S., Ohtsu, N., Motoshima, H., Matsuzawa, H., Shoun, H., Wakagi, T., 2002. Trimeric crystal structure of the glycoside hydrolase family 42 beta-galactosidase from *Thermus thermophilus* A4 and the structure of its complex with galactose. J Mol Biol 322, 79-91.

Hiraiwa, M., Shiraishi, T., Uda, Y., 1986. Multiplicity of bovine liver GM1 ganglioside beta-galactosidase. J Biochem 100, 255-258.

Hiraiwa, M., Uda, Y., 1986. Purification and properties of GM1 ganglioside beta-galactosidases from bovine brain. J Biochem 100, 707-715.

Hoogeveen, A. T., Verheijen, F. W., Galjaard, H., 1983. The relation between human lysosomal β-galactosidase and its protective protein. J Biol Chem 258, 12143-12146.

Iglesias, N., Abelenda, J. A., Rodino, M., Sampedro, J., Revilla, G., Zarra, I., 2006. Apoplastic glycosidases active against xyloglucan oligosaccharides of *Arabidopsis thaliana*. Plant and Cell Physiology 47, 55-63.

Javeri, S., Uhlenbruck, G., 1984. Isolation of a β-galactosidase from chicken liver. J Clin Chem Clin Biochem 22, 735-739.

Juers, D. H., Hakda, S., Matthews, B. W., Huber, R. E., 2003. Structural basis for the altered activity of Gly794 variants of *Escherichia coli* β-galactosidase. Biochemistry 42, 13505-13511.

Martin, I., Jimenez, T., Esteban, R., Dopico, B., Labrador, E., 2008. Immunolocalization of a cell wall β-galactosidase reveals its developmentally regulated expression in *Cicer arietinum* and its relationship to vascular tissue. J. Plant Growth Regul. 27, 181-191.

McCann, M. C., Roberts, K., 1991. Architecture of the primary cell wall. In: C.W., L. (Ed.), Cytoskeletel basis of plant growth and form. Academic Press, San-Diego, pp. 109-129.

McCarter, J. D., Burgoyne, D. L., Miao, S., Zhang, S., Callahan, J. W., Withers, S. G., 1997. Identification of Glu-268 as the catalytic nucleophile of human lysosomal beta-galactosidase precursor by mass spectrometry. J Biol Chem 272, 396-400.

Mutoh, T., Naoi, M., Nagatsu, T., Takahashi, A., Matsuoka, Y., Hashizume, Y., Fujiki, N., 1988. Purification and characterization of human liver β-galactosidase from a patient with the

adult form of GM1 gangliosidosis and a normal control. Biochim Biophys Acta 964, 244-253.

Nishigaki, T., Okada, S., 1998. [beta-Galactosidosis (acid beta-galactosidase deficiency: GM1 gangliosidosis, Morquio B disease)]. Ryoikibetsu Shokogun Shirizu, 403-406.

O'Donoghue, E. M., Somerfield, S. D., Heyes, J. A., 2002. Organization of cell walls in *Sandersonia aurantiaca* floral tissue. J. Exp. Bot. 53, 513-523.

Okada, S., O'Brien, J. S., 1968. Generalized gangliosidosis: beta-galactosidase deficiency. Science 160, 1002-1004.

Ozeki, Y., Yokota, Y., Kato, K. H., Titani, K., Matsui, T., 1995. Developmental expression of D-galactoside-binding lectin in sea urchin (*Anthocidaris crassispina*) eggs. Exp Cell Res 216, 318-324.

Perez, A. I. B., 2004. *Arabidopsis* cell wall b-galactosidase gene family: expression, catalytic activities and biological function in galactose dynamics., Botany and Plant Pathology, vol. Ph.D. Purdue, west Lafayette, p. 141.

Richmond, T. A., Somerville, C. R., 2001. Integrative approaches to determining Csl function. Plant Mol Biol 47, 131-143.

Rojas, A. L., Nagem, R. A., Neustroev, K. N., Arand, M., Adamska, M., Eneyskaya, E. V., Kulminskaya, A. A., Garratt, R. C., Golubev, A. M., Polikarpov, I., 2004. Crystal structures of β-galactosidase from *Penicillium* sp. and its complex with galactose. J Mol Biol 343, 1281-1292.

Skovbjerg, H., Noren, O., Sjostrom, H., Danielsen, E. M., Enevoldsen, B. S., 1982. Further characterization of intestinal lactase/phlorizin hydrolase. Biochim Biophys Acta 707, 89-97.

Smith, D. L., Gross, K. C., 2000. A family of at least seven β-galactosidase genes is expressed during tomato fruit development. Plant Physiol. 123, 1173-1183.

Suzuki, Y., 2006. Beta-galactosidase deficiency: an approach to chaperone therapy. J Inherit Metab Dis 29, 471-476.

Tanthanuch, W., Chantarangsee, M., Maneesan, J., Ketudat-Cairns, J., 2008. Genomic and expression analysis of glycosyl hydrolase family 35 genes from rice (*Oryza sativa* L.). BMC Plant Biol 8, 84.

Trainotti, L., Spinello, R., Piovan, A., Spolaore, S., Casadoro, G., 2001. β-Galactosidases with a lectin-like domain are expressed in strawberry. J Exp Bot 52, 1635-1645.

Zecca, L., Mesonero, J. E., Stutz, A., Poiree, J. C., Giudicelli, J., Cursio, R., Gloor, S. M., Semenza, G., 1998. Intestinal lactase-phlorizin hydrolase (LPH): the two catalytic sites; the role of the pancreas in pro-LPH maturation. FEBS Lett 435, 225-228.

Chapter 2

The *Arabidopsis* At1g45130 and At3g52840 genes encode β-galactosidases with activity toward cell wall polysaccharides.

This work was published previously and reprinted with kind permission from Elsevier: Phytochemistry, "The *Arabidopsis* At1g45130 and At3g52840 genes encode β-galactosidases with activity toward cell wall polysaccharides." Vol. 69, 2008, 1661-1670, D. Gantulga, Y. Turan, D. R. Bevan, and A. Esen.

Keywords: Arabidopsis; Cruciferae; Cell wall; Glycosyl hydrolase family 35; β-Galactosidase; β-Galactosides; Galactan

Abbreviations: Arabidopsis thaliana, Arabidopsis; Gal-5, (AtBGal5)–β–galactosidase-5; Gal-2, (AtBGal2)–β-galactosidase-2; *p*NP, para-nitrophenyl; GH, glycosyl hydrolase.

Abstract

The *Arabidopsis* genes At1g45130 and At3g52840 encode the β-galactosidase isozymes Gal-5 and Gal-2 that belong to Glycosyl Hydrolase Family 35 (GH 35). The two enzymes share 60% sequence identity with each other and 38-81% with other plant β-galactosidases that are reported to be involved in cell wall modification. We studied organ-specific expression of the two isozymes. According to our western blot analysis using peptide-specific antibodies, Gal-5 and Gal-2 are most highly expressed in stem and rosette leaves. We show by dot-immunoblotting that Gal-5 and Gal-2 are associated with the cell wall in *Arabidopsis*. We also report expression of the recombinant enzymes in *P. pastoris* and describe their substrate specificities. Both enzymes hydrolyze the synthetic substrate *para*-nitrophenyl-β-D-galactopyranoside and display optimal enzyme activity between pH 4.0 and 4.5, similar to the pH optimum reported for other well-characterized plant β-galactosidases. Both Gal-5 and Gal-2 show a broad specificity for the aglycone moiety and a strict specificity for the glycone moiety in that they prefer galactose and its 6-deoxy analogue, fucose. Both enzymes cleave β-(1, 4) and β-(1, 3) linkages in galacto-oligosaccharides and hydrolyze the pectic fraction of *Arabidopsis* cell wall. These findings suggest that Gal-5 and Gal-2 could be involved in the modification of cell wall polysaccharides**.**

1. Introduction

Carbohydrates play many important biological and physiological roles in living organisms, and they occur in different forms such as glycoproteins, glycolipids, glycosides, and oligo- and polysaccharides. Selective hydrolysis of glycosidic bonds specifies the breakdown of these glycoconjugates and their function in subsequent biochemical processes. Living organisms have evolved a great diversity of glycosyl hydrolases (GH) with different substrate specificity and kinetic parameters, operating under different optimal conditions (Henrissat, 1995). Glycosyl hydrolases are found in all animals, plants and microorganisms, and they are classified into 111 families on the basis of amino acid sequence similarity (CAZY database, http://www.cazy.org; (H. J. Gilbert, 1999). The GH family 35 includes plant β-galactosidases (EC 3.2.1.23) that are thought to be involved in cell wall biogenesis and modification (Jamet et al., 2006, Vervelen and Vissenberg, 2007). In plants, β-galactosidase isozymes are encoded by a multigene family. The presence of β-galactosidases as a moderately large enzyme family is indicative of the structural complexity and heterogeneity of cell wall polysaccharides. Depending on developmental stages and cell differentiation, the different isozymes may act on the same or different cell wall polysaccharides (Smith and Gross, 2000; Perez, 2004; Kotake et al., 2005; Ishimaru et al., 2005; Jamet et al., 2006; Chantarangsee et al., 2007).

A growing body of evidence suggests that GH Family 35 β-galactosidases participate in loosening of the cell wall during growth and degrade cell wall components during ripening of fruits and senescence. Recently, many β-galactosidases have been identified and purified from different plant sources. For example, Smith and Gross (2000) showed that a family of at least seven β−galactosidases is expressed during tomato fruit development. Furthermore, these authors

expressed the tomato β−galactosidase isozyme 4 (TBG4) in yeast and studied its natural substrate specificity. TBG4 hydrolyzed chelator-soluble pectin, alkali-soluble pectin, hemicellulose from tomato cell walls, and commercially prepared galactan. Li et al. (2001) purified five isoforms of β−galactosidases from mung bean and showed that they differ with respect to enzymatic characteristics and substrate specificities. Kotake et al. (2004) isolated a β−galactosidase from radish (RsbGal), which specifically hydrolyzed β-(1, 6) and β-(1, 3) linkages in arabinogalactan protein. β-Galactosidase I purified from ripe carambola (*A. carambola*) was active in hydrolyzing β-(1, 4)-linked spruce and β-(1, 6)/ β-(1, 3)-linked gum arabic galactan, and alkali-soluble hemicelluloses of carambola cell wall (Balasubramaniam et al, 2005). Furthermore, papaya β-galactosidase isoforms differentially hydrolyzed cell wall during fruit ripening (Lazan et al., 2004).

Despite the status of *Arabidopsis thaliana* as a model plant with a completely sequenced genome, not much is known about its β-galactosidases. There are 17 putative GH family 35 β-galactosidases (17 BGALs) in *Arabidopsis* (Ahn et al., 2007). According to a recent microarray study, the β-galactosidase genes are among 765 genes specific to cell wall dynamics (Imoto et al., 2005). Perez (2004) studied expression of β-galactosidases by promoter-GUS fusion and found that temporal and spatial expression of β-galactosidase genes are differentially regulated. She also reported that the galactosyl content of the cell wall increases and total β-galactosidase activity decreases in *Gal-1* knockout mutant. Iglesias et al. (2006) reported β-galactosidase activity on xyloglucan oligosaccharides in the apoplastic fluid. To corroborate further the role of β-galactosidases in cell wall modification, it is essential to determine their substrate specificity, by isolating and characterizing the proteins. To this end, Ahn et al. (2007) expressed Gal-4 (At5g56870), one of the seventeen β-galactosidases, in *E. coli* and insect cells. It was shown that

in addition to synthetic substrates, Gal-4 hydrolyzes β-(1, 3) and β-(1, 4) linked galacto-oligosaccharides (Ahn et al., 2007).

It is likely that the 17 BGALs differ in temporal and spatial expression and in their ability to hydrolyze various cell wall polysaccharides. For annotating them and understanding their function better, it is desirable to determine the expression profile and substrate specificity of all 17 BGALs. Currently, we are focused on the largest subfamily (Gal-1 through Gal-5 and Gal-12), which consists of six genes, amongst the 17 BGALs. Of these, we studied Gal-5 and Gal-2 for this paper. First, we conducted database analyses using microarray data available for these genes. Second, we studied spatial expression and subcellular localization by immunoblotting using peptide-specific antibodies. Third, we purified the recombinant proteins expressed in *P. pastoris* and investigated their substrate specificity using synthetic galactosides and natural polysaccharides derived from *Arabidopsis* cell wall.

2. Results and Discussion

2.1. In silico analysis of Gal-5 and Gal-2

BLAST searches of the NCBI database using Gal-5 and Gal-2 sequences as queries show that the sequences of 20 other plant β-galactosidases share 38-81% identity with all *Arabidopsis* β-galactosidases except Gal-17. Gal-5 and Gal-2 proteins share 60% sequence identity with each other. The sequence differences among β-galactosidases suggest differences in their enzymatic properties (e.g., substrate specificity) and biological functions.

In-silico characterization of predicted protein sequences of Gal-5 and Gal-2 using SignalP and TargetP shows that these proteins contain a signal peptide that targets them to the endoplasmic reticulum from which they can be secreted to the cell wall via vesicular pathways (Bendtsen et al., 2004; Emanuelsson et al., 2000). Phylogenetic analysis of Gal-5 and Gal-2 proteins shows that they cluster with other plant β-galactosidases that are known to be cell wall modifying enzymes (Fig. 1). In tomato fruit, the galactosyl content of cell wall and fruit firmness depend on two β-galactosidases, TBG6 and TBG4 (Smith et al., 2002), which are closely related to Gal-5 and Gal-2, respectively (Fig. 1).

2.2. Microarray data for expression of At1g45130 (*Gal-5*) and At3g52840 (*Gal-2*) transcripts

To study the expression profiles of the *Gal-5* and *Gal-2* genes, we used microarray data from the integrated database Genevestigator (Zimmermann et al., 2004). We analyzed the relative abundance of At1g45130 (*Gal-5*) and At3g52840 (*Gal-2*) transcripts for different tissues, developmental stages, and various stress conditions. Table 1 shows mean gene expression data across many microarray experiments carried out in different laboratories. Relatively small standard errors for the mean expression values indicate the consistency and reliability of the data. The mean expression of *Gal-5* is high in stem, root and silique, while the mean expression of *Gal-2* is high in the petiole of rosette leaves and stem. The transcript levels do not change notably throughout developmental stages from seedling to adult plants. *Gal-5* and *Gal-2* transcripts were absent in the pollen (data not shown). It is possible that other members of BGAL

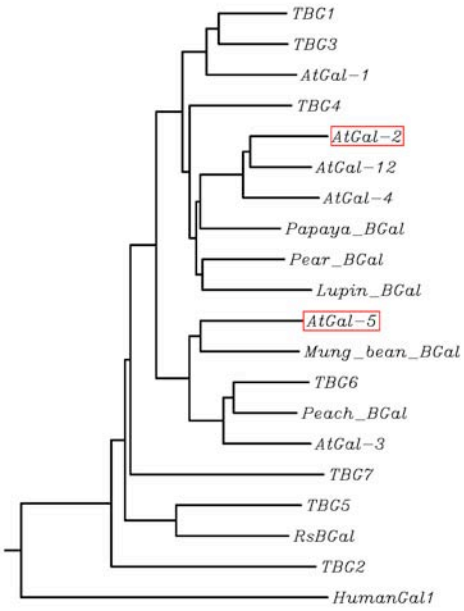


Fig. 1. Phylogenetic tree of selected plant β-galactosidases. The tree was constructed from the alignment of the amino acid sequences using Clustal W software in http://align.genome.jp/. **AtGal-1**, *A. thaliana* β-galactosidase-1/AtBGAL1 (**NP187988**); **AtGal-2**, *A. thaliana* β-galactosidase-2/AtBGAL2 (**NP190852**); **AtGal-3**, *A. thaliana* β-galactosidase-3/AtBGAL3 (**NP849506**); **AtGal-4**, *A. thaliana* β-galactosidase-4/AtBGAL4 (**CAB64740**); **AtGal-5**, *A. thaliana* β-galactosidase-5/AtBGAL5 (**NP175127**); **AtGal-12**, *A. thaliana* β-galactosidase-12/AtBGAL12 (**NP849553**); **Pear BGal**, *P. pyrifolia* (**BAB21492**); **Papaya BGal**, *C. papaya* (**AAC77377**); **TBG1**, *L. esculentum* (**AAF21626**); **TBG2**, *L. esculentum* (**AAF70821**); **TBG3**, *L. esculentum* (**AAF70822**); **TBG4**, *L. esculentum* (**AAC25984**); **TBG5**, *L. esculentum* (**AAF70824**); **TBG6**, *L. esculentum* (**AAF70825**); **TBG7**, *L. esculentum* (**AAF70823**); **Mung bean BGal**, *V. radiata* (**AAF67341**); **Peach BGal**, *P. persica* (**AAW47739**); **Lupin BGal**, *L. angustifolius* (**CAA09467**); **RsbGal**, *R. sativus* β-galactosidase 1 (**BAD20774**); **Human BGal**, *H. sapiens* (**P16278**).

Table 1. Expression profile of Gal-5 and Gal-2 using microarray data

Organ/tissue	*# of chips*	*At1g45130 Gal-5*	*At3g52840 Gal-2*	*At4g05320 UBQ10*
Root	200	11.67±0.04	8.55±0.06	14.7±0.02
Petiole	12	10.62±0.11	12.82±0.2	14.8±0.07
Rosette	604	8.93±0.05	11.14±0.04	14.7±0.01
Stem	7	12±0.04	11.26±0.22	14.6±0.24
Cauline leaf	3	7.3±0.33	10.69±0.11	14.8±0.04
Flower	58	9.94±0.25	9.12±0.11	14.5±0.07
Silique	11	11.83±0.33	10.12±0.23	14.8±0.04
Seedling	386	10.57±0.05	10.9±0.03	14.8±0.01

Second column shows the number of microarray chips available for a given tissue. The organ/tissue specific expression levels for Gal-5 (At1g45130) and Gal-2 (At3g52840) as $\log_2(n)$ are shown along with the standard errors in columns three and four. Ubiquitin10 (*UBQ10*/Atg05320) used as controls for comparison are shown in column five. The Gene Atlas Tool of Genevestigator (https://www.genevestigator.ethz.ch/at/) was used for analysis.

family are expressed in pollen to rescue Gal-5 and Gal-2. Indeed, Hruba et al. (2005) showed that *Gal-7*, *Gal-11* and *Gal-13* are specifically expressed in the pollen. It is worth mentioning that microarray results are in good agreement with recent RT-PCR results for selected organs (Ahn et al., 2007).

Comprehensive microarray data are available for *Arabidopsis* genes expressed in shoots and roots under various stress conditions (http://www.weigelworld.org/resources/microarray/AtGenExpress). According to the microarray data, expression levels of *Gal-5* and *Gal-2* transcripts are not significantly different under most of the stress conditions, except UV-B light and osmotic (mannitol) stresses. In comparison to the control, maximal change (~7-fold increase) of *Gal-2* expression is in roots (harvested after 3 hours) treated with UV-B light. For *Gal-5*, maximal change (~3-fold decrease) is in root (harvested after 24 hours) treated by osmotic stress. An RT-PCR study of β-galactosidase gene expression in plant organs under various stress conditions and hormone treatments may further corroborate the role of BGALs in cell wall modification, by demonstrating coordinated expression of β-galactosidase genes and the genes whose products are known to be cell wall-specific (e.g. expansin, polygalacturonase).

2.3. Expression of Gal-5 and Gal-2 in *Arabidopsis*: western blotting studies using peptide-specific antibodies

Before the injection of immunogens, rabbit preimmune sera were tested by ELISA for immunoreactivity with total protein extracts of *Arabidopsis* and synthetic peptides (data not shown). Two rabbits with the lowest preimmune serum reactivity were chosen for the injection of Gal-5 and Gal-2 peptide conjugates, respectively. The peptide-specific antisera were tested for

their specificity by dot blotting. Anti- Gal-2 peptide antibody recognized Gal-2 synthetic peptide and the recombinant Gal-2 protein (*E. coli* and *P. pastoris* expressed), and anti- Gal-5 peptide antibody recognized Gal-5 peptide and recombinant Gal-5. Preimmune sera did not show any detectable immunoreactivity with Gal-5 and Gal-2 on the blots.

The expression pattern of the two β-galactosidases in different *Arabidopsis* tissues was studied by western blotting using antibodies raised against Gal-5 and Gal-2 specific peptides. Immunoreactive bands of ~75 kD were detected in 4 to 5-week old plants (Fig. 2A), indicating Gal-5 and Gal-2 proteins do not undergo proteolytic processing that has been reported to produce 35- to 50-kD fragments for β-galactosidases in apple (Ross et al., 1995), lupin (Buckeridge et al., 1994) and radish (Kotake et al., 2005). Such fragments were not detected on immunoblots even after using high protein loads (Fig. 2B). Under our experimental conditions, levels of Gal-2 protein were similar in all organs (root, leaf, stem, flower, and silique), whereas Gal-5 levels were different. Gal-5 levels were higher in stem and leaves, but lower in roots and silique. We compared our results with the previous studies on expression of β-galactosidase genes by Perez (2004) and Ahn et al. (2007). These authors found Gal-2 transcripts with moderate levels of expression in all organs, which is in agreement with our result for Gal-2. In the case of Gal-5, they found higher level of expression in roots than we did. We propose that this difference in Gal-5 expression is due to the difference in transcriptional and translational stages of regulation, though we do not exclude other factors, such as plant age, growth conditions, and extraction methods used in the experiments. It is worth mentioning that there is agreement between our data and those of Perez (2004) in that Gal-5 is not detectable in mature roots, although Perez (2004) found Gal-5 expression in root elongation and root hair zones of juvenile plants.

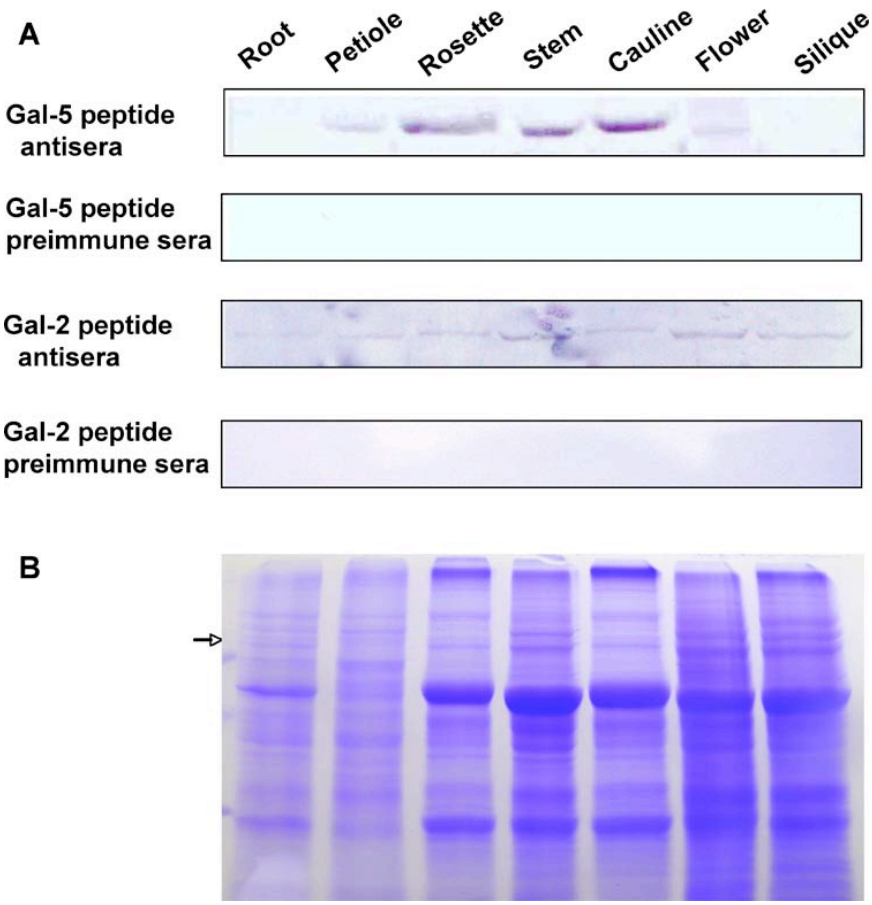


Fig. 2. Organ-specific expression of Gal-5 and Gal-2. A. Western blot and B. Coomassie Blue stained SDS-PAGE. Total proteins (30 µg) from *Arabidopsis* (4 weeks old) root, petiole of rosette leaves, rosette leaves, stems, cauline leaves, flowers, and siliques were separated by 10% SDS-PAGE and transferred to nitrocellulose membranes. Identical membranes were incubated with rabbit preimmune sera and immune antisera against Gal-5 and Gal-2 peptides. The arrow marks the position of immunoreactive bands.

2.4. Cell wall localization of Gal-5 and Gal-2 proteins: dot blotting

We isolated cell wall from rosette leaves of *Arabidopsis* to confirm the presence of Gal-5 and Gal-2 proteins in the cell wall. Five different fractions (S1, soluble 1; S2, soluble 2; S3, soluble 3; CW4, extractable with $CaCl_2$; and CW5, extractable with LiCl) were obtained. These fractions were assayed for β-galactosidase activity using *p*NPGal as a substrate. S2 and S3 fractions did not have detectable activity. Specific activities of S1, CW4, and CW5 were 0.06, 0.12, and 20.1 nmole *p*NP/min/mg, respectively, indicating that fraction CW5 (LiCl-soluble) had the highest specific β-galactosidase activity. Fraction CW5 contained the lowest amount of protein (Fig. 3, bottom row) among the three fractions. The immunoblotting data showed that antiserum to intact Gal-2 protein had high immunoreactivity (Fig. 3, top blot) with fraction CW4 ($CaCl_2$-soluble) and weak immunoreactivity with fractions S1 and CW5. However, this antiserum is not specific for Gal-2; it recognizes also other cell wall-bound β-galactosidases. In contrast, the reactivity of the Gal-2 peptide-specific antiserum was strongest with fraction CW5 (Fig. 3, second blot), indicating that Gal-2 is enriched in CW5 and it requires LiCl for complete release from the cell wall. In the case of Gal-5, the preimmune serum from the rabbit immunized with Gal-5 peptide had considerable background activity with cell wall components (Fig. 3, fourth blot). Although the immune serum from the same rabbit reacted more strongly with fractions CW4 and CW5 than the preimmune serum, the difference between the specific and the background reactions was not as striking as for Gal-2. Taken together, our enzyme activity and dot immunoblotting data indicate both Gal-5 and Gal-2 are present in, and tightly associated with, the cell wall in *Arabidopsis*. ELISA data (not shown) also confirmed that Gal-5 and Gal-2

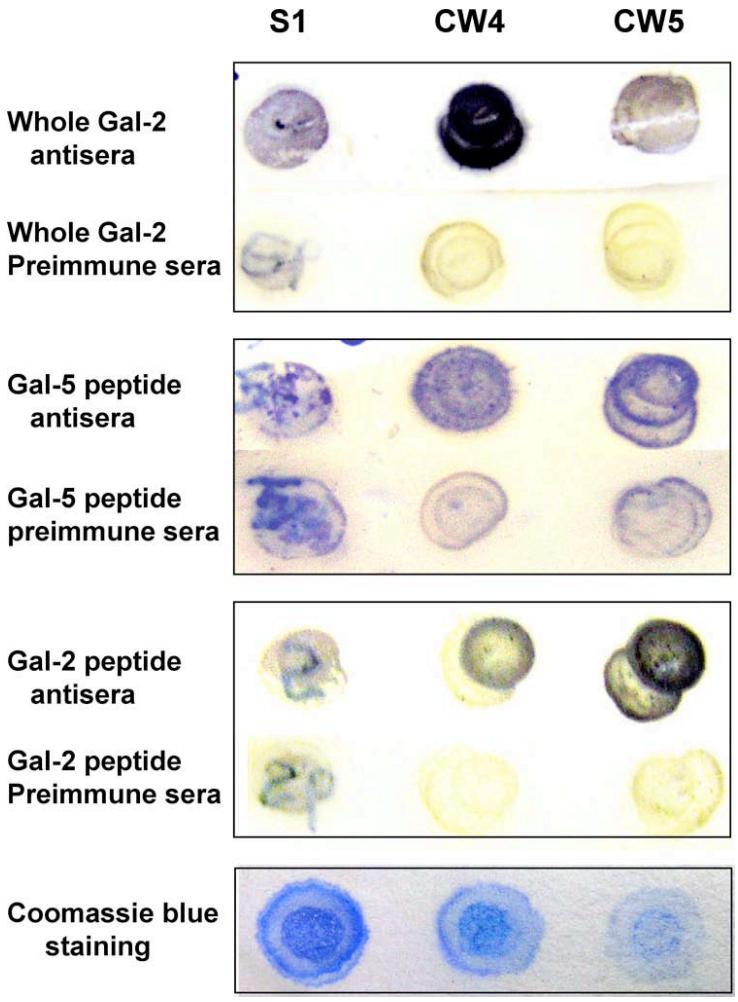


Fig. 3. Localization of Gal-5 and Gal-2 in cell wall fractions by dot blotting. Cell walls were isolated from *Arabidopsis* rosette leaves. Three fractions (S1, soluble; CW4, $CaCl_2$-soluble; CW5, LiCl-soluble) with β-galactosidase activity were spotted on nitrocellulose strips at the same place multiple times. Identical membranes with protein spots were incubated with anti-whole Gal-2, anti- Gal-5 peptide-, and anti- Gal-2 peptide- specific antisera. Total protein spots were stained with Coomassie Blue R-250. These dot immunoblotting data show the presence of Gal-5 and Gal-2 in the cell wall.

proteins were present in cell wall fractions, supporting our hypothesis that Gal-5 and Gal-2 proteins are bound to the cell wall.

2.5. Expression of Gal-5 and Gal-2 in *P. pastoris* and purification

The recombinant proteins were expressed under the control of the AOX (alcohol oxidase) promoter in *P. pastoris*. Gal-5, 700 amino acids long, (79-kD protein, calculated) and Gal-2, 719 amino acids long, (81-kD protein, calculated), were expressed and secreted into the culture medium. Optimization of induction and time-course studies of expression were done to obtain the best expression level for recombinant Gal-5 and Gal-2. Results from the induction time course (data not shown) showed that β-galactosidase activity was secreted into the culture medium and was detectable after 24 hrs of induction on 1% methanol, and it peaked after 72-96 hours. While β-galactosidase activity of Gal-5 and Gal-2 transformants increased during the course of induction, no detectable activity was observed in the control *P. pastoris* transformed with an empty vector.

Recombinant Gal-5 and Gal-2 were purified from the culture medium by ion exchange chromatography. Cation exchange chromatography using Sulphoxyethyl (SE) cellulose was efficient in purification of these enzymes because *P. pastoris* culture medium contains low levels of secreted endogenous protein, and both enzymes bind SE due to their high positive net charge at around pH 6. It allowed 2.7-fold purification of Gal-5 and 2-fold purification of Gal-2 in a single step. Purified Gal-5 and Gal-2 appeared on SDS –PAGE gel as single bands with an estimated monomeric molecular weight of ~75 kD (Fig. 4A, B), indicating their purification to near homogeneity. This result was supported by acidic native polyacrylamide gel electrophoresis

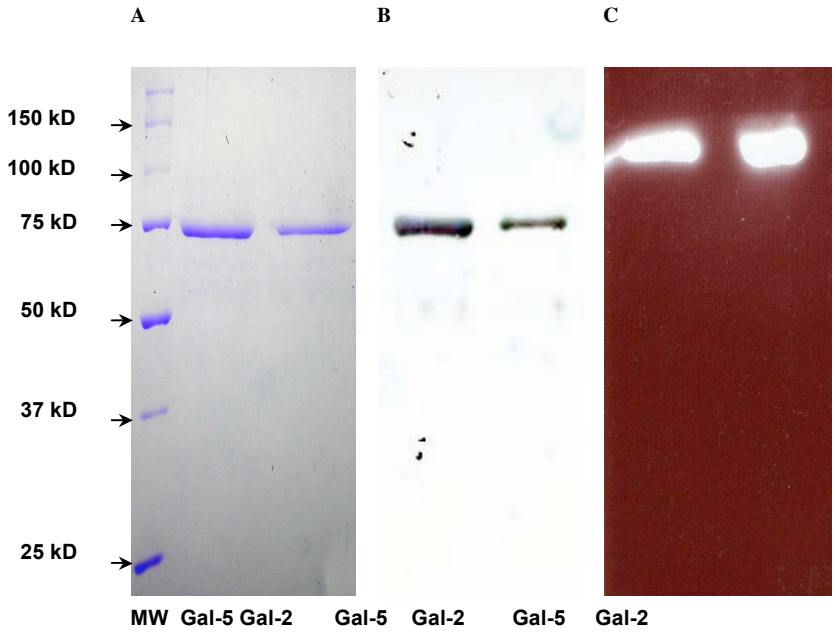


Fig. 4. SDS-PAGE (A), western blot (B), and zymogram (C) of purified Gal-5 and Gal-2. **A**-10% SDS gel was stained with Coomassie Blue R-250. B-Purified proteins separated on 10% SDS-PAGE were transferred to a nitrocellulose membrane, which was incubated with rabbit anti- whole Gal-2 antiserum. C-Proteins were subjected to 8% acidic native gel and stained with 4-methylumbelliferyl galactoside. **Lane MW**, **molecular weight standard**; **Lane 1, Gal-5**; **Lane 2, Gal-2**.

that showed a single activity band on the zymogram for each enzyme (Fig. 4C). It should be noted that the monomeric molecular weight of Gal-5 and Gal-2 obtained by SDS-PAGE is slightly lower than the estimation based on amino acid sequence (79- kD for Gal-5 and 81-kD for Gal-2). Peptide mass fingerprinting (MALDI-TOF) analysis showed that both ends of the Gal-5 polypeptide sequence were present in the peptide mixture, which ruled out proteolytic modification of Gal-5 during purification. Thus, the lower experimentally estimated monomeric molecular weight of Gal-5 and Gal-2 is likely due to their high content of hydrophobic amino acids (Gal-5 40%, Gal-2 42% compared to BSA (bovine serum albumin) 34%). Hydrophobic amino acids bind SDS at a higher ratio compared to hydrophilic ones, thereby causing slightly faster mobility of hydrophobic proteins on SDS-PAGE (Bayreuther et al., 1980).

The fact that Gal-5 and Gal-2 bind to a cation exchanger at pH~6 and migrate through acidic native gels indicates that these proteins have high p*I* values, consistent with the predicted p*I*s of 8.1 and 8.6 for Gal-5 and Gal-2, respectively. It is interesting to note that basic p*I* values were also observed for kiwifruit, tomato, and carambola β-galactosidases (Ross et al., 1993; Carey et al., 1995; Balasubramaniam et al., 2005). These results support the hypothesis that the cell wall-associated proteins interact ionically with the acidic cell wall matrix. Pectic polygalacturonic acid provides negative charges and contributes to the interaction with high p*I* proteins (Jamet et al., 2006).

2.6. Properties of recombinant Gal-5 and Gal-2

Recombinant Gal-5 and Gal-2 displayed optimal activity on *p*NP-β-D-galactopyranoside between pH 4.0 and 4.5. The pH-activity profile had a narrow range in that less than 50% of

maximal activity was observed at pH values below 3.5 and above 6.0. These data are consistent with the pH optima of most of the well-characterized plant β-galactosidases, which range from 3.0 to 5.0 (Carey and et al, 1995; Cicek, 1998; Edwards Mary, 1988; Gross, 2001; Kotake et al, 2005). It is generally believed that the pH of the cell wall is acidic, around 5.5. During growth and expansion of the cell, the pH of the cell wall falls below 5.5 (Grebe, 2005), which is optimal for the activity of most cell wall modifying enzymes. When β-galactosidase activity was measured as a function of temperature, both Gal-5 and Gal-2 showed temperature optima at 40 °C and were stable up to 55 °C (data not shown). At temperatures above 55 °C, both enzymes were inactivated irreversibly.

2.7. Aglycone, glycone and natural substrate specificity

To determine the specificity of the enzymes for the aglycone moiety, β-galactosides with a variety of aglycones such as *ortho*-nitrophenyl (*o*NP), *para*-nitrophenyl (*p*NP), 4-methylumbelliferyl (4MU), 5-bromo-4-chloro-3-indolyl (X) and 6-bromo-2-naphthyl (6BN) were tested. All of them were hydrolyzed by both Gal-5 and Gal-2, albeit with different efficiencies (Table 2). These results indicate that they have a broad specificity with respect to the aglycone moiety. To investigate the specificity for the glycone moiety, *p*NP-β-D-galactopyranoside (*p*NPGal), *p*NP−β-D-mannopyranoside, *p*NP−β-D-fucopyranoside, *p*NP−β-D-xylofuranoside, *p*NP−β-D-arabinopyranoside, and *p*NP−α-L-arabinopyranoside (*p*NPAra) were tested. The results are summarized in Table 3, which show that only *p*NP-β-D-galactopyranoside and its 6-deoxy analogue *p*NP−β-D-fucopyranoside were hydrolyzed. Gal-5 and Gal-2 failed to hydrolyze *p*NPAra, showing that the glycone specificity of these two enzymes is strict. Thus,

Table 2. Aglycone specificities of Gal-5 and Gal-2

Aglycone	Relative activity[a], %	
	Gal-5	**Gal-2**
***para*- nitrophenyl-** (*p*NPGal)	100	100
***ortho*-nitrophenyl-** (*o*NPGal)	61	72
4-methylumbelliferyl- (4MUGal)	18	20
5-bromo-4-chloro-3-indolyl- (X-Gal)	43	56
6-bromonaphthyl- (6BNGal)	22	30

[a]Activities of Gal-5 and Gal-2 were assayed in reaction mixtures containing 2.5 mM substrate in NaOAc buffer pH 4.6. Aglycone specificity is expressed as a percentage of activity against *p*NPGal (100% ~ 0.03 units (nkat)). For insoluble aglycones, amounts of galactose produced as a result of hydrolysis were measured by the galactose dehydrogenase assay.

Table 3. Sugar specificities of Gal-5 and Gal-2

Glycone	Relative activity[a], %	
	Gal-5	**Gal-2**
*p*NP-β-D-**galactopyranoside**	100	100
*p*NP-β-D-**fucopyranoside**	25	21
*p*NP-β-D-**glucopyranoside**	<1	<1
*p*NP-β-D-**mannopyranoside**	0	0
*p*NP-β-D-**xylofuranoside**	0	0
*p*NP-β-D-**arabinopyranoside**	0	0
*p*NP-α-L-**arabinopyranoside**	0	0

[a]Activities of Gal-5 and Gal-2 were assayed in reaction mixtures containing 2.5 mM substrate in NaOAc buffer pH 4.6. Sugar specificity is expressed as a percentage of activity against *p*NPGal (100% ~ 0.03 unit (nkat)).

Gal-5 and Gal-2 are highly specific for β-galactopyranoside and discriminate sugars based on the configuration of the hydroxyl group at C4 and C3 positions.

Kinetic parameters of Gal-5 and Gal-2 were determined with *p*NP-β-D-galactopyranoside. K_m values for *p*NP-β-D-galactopyranoside for the two enzymes were similar (0.28±0.06 mM for Gal-5 and 0.40±0.02 mM for Gal-2), but k_{cat} values were different (1.55 s^{-1} for Gal-5 and 6.03 s^{-1} for Gal-2). Their catalytic efficiencies (k_{cat}/K_m) differed to some extent (5.54 $s^{-1}mM^{-1}$ for Gal-5 and 15.21 $s^{-1}mM^{-1}$ for Gal-2). K_m values for *o*NPGal and *p*NPFuc were 0.83±0.015 and 3.75±0.55 mM for Gal-5, respectively, and 0.72±.016 and 6.4±2.8 mM for Gal-2, respectively.

Inhibitory effects of several sugars and sugar derivatives were tested using *p*NPGal as a substrate. γ-Galactonelactone and D-galactose were the most effective inhibitors for Gal-5 and Gal-2 activity. Their K_i values were 44 μM and 7.4 mM, respectively for Gal-5 and 98 μM and 4.5 mM for Gal-2. Also D-fucose, methyl-α-D-galactoside and raffinose were weaker inhibitors for both enzymes while *p*NPGlc, *p*NPAra, lactose, IPTG, galacturonic acid, L-arabinose, and D-mannose did not show any inhibitory effects. Ag^{+}, Hg^{2+} and SDS strongly inhibited activity of both enzymes when *p*NPGal was used as a substrate.

β-galactosidases from different plants or within the same plant are known to differ considerably in their linkage specificity (Kotake et al., 2004; Ishimura et al., 2005; Buckeridge et al., 2005). Using β-(1, 4), β-(1, 3) and β-(1, 6) linked galacto-oligosaccharides, we investigated the linkage specificity of Gal-5 and Gal-2. As can be seen from Fig. 5, both Gal-5 and Gal-2 hydrolyze β-(1, 4) (lanes 3-4 and 5-6) and β-(1, 3) (lanes 8-9 and 10-11) linkages, whereas the β-(1,6) linkages in galacto-oligosaccharides were less susceptible to hydrolysis (lanes 13-16).

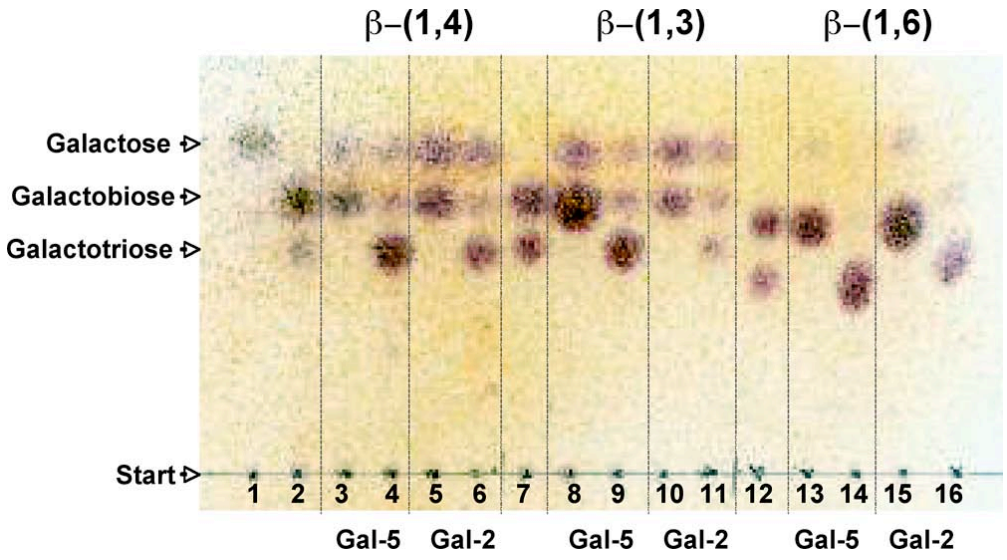


Fig. 5. Linkage specificity of Gal-5 and Gal-2. After hydrolysis of galactobioses and galactotrioses by Gal-5 and Gal-2, products were separated by TLC and developed with naphthoresorcinol (see methods). **Lane 1**, monogalactose, Lane **2**- β–(1, 4)- linked galactobiose and galactotriose; **Lane 3 and 4**, hydrolysis product of β–(1, 4)- linked galactobiose and galactotriose by **Gal-5**; **Lane 5 and 6**, hydrolysis product of β–(1, 4)- linked galactobiose and galactotriose by **Gal-2**; **Lane 7**, β–(1, 3)- linked galactobiose and galactotriose; **Lane 8 and 9**, hydrolysis product of β–(1, 3)- linked galactobiose and galactotriose by **Gal-5**; **Lane 10 and 11**, hydrolysis product of β–(1, 3)-linked galactobiose and galactotriose by **Gal-2**; **Lane 12**, β–(1, 6)- linked galactobiose and galactotriose; **Lane 13 and 14**, hydrolysis product of β–(1, 6)- linked galactobiose and galactotriose by **Gal-5**; **Lane 15 and 16**, hydrolysis product of β–(1, 6)-linked galactobiose and galactotriose by **Gal-2**.

Ahn et al. (2007) showed that a member of the family, Gal-4, preferentially cleaves β-(1, 4) and β-(1, 3) linkages. Thus, the three *Arabidopsis* paralogs, Gal-5, Gal-2, and Gal-4, might act on the same natural substrates with β-(1, 4) and β-(1, 3) linkages. To probe the natural substrate specificity of Gal-5 and Gal-2, more complex oligo-/polysaccharides were tested and the results are shown in Table 4. L-Arafase (α-L-arabinofuranosidase) pretreated (to remove arabinose) lupin galactan, a polymer of β-(1, 4) linked galactose, was hydrolyzed to some extent, whereas gum arabic and gum guar were not hydrolyzed. Commercially prepared apple pectin (Sigma, P8471) was the best substrate among the polysaccharides tested. In the case of larchwood arabinogalactan and gum arabic galactan, we were unable to measure hydrolysis, since the negative controls had high background due to the presence of components that interfere with the galactose dehydrogenase assay (#2570-050, Interscientific, Hollywood FL). The fact that Gal-5 and Gal-2 hydrolyzed lupin galactan suggests that these enzymes have exo-galactanase activity. Exo-galactanase activity has been reported for lupin β-galactosidase (Buckeridge et al., 1994 and Buckeridge et al., 2005), a tomato β-galactosidase (TBG4) (Carey et al. and 1995; Ishimaru et al. 2005;), and apple β-galactosidase (Ross et al., 1994).

The strict specificity of Gal-5 and Gal-2 for galactose and their ability to hydrolyze β-(1, 4) and β-(1, 3) linkages in galactooligosaccharides and β-(1, 4) linkages in lupin galactan suggest that cell wall polysaccharides rich in galactan are more likely to be natural substrates for these enzymes. Rhamnogalacturonan-I (RG-I) and rhamnogalacturonan-II (RG-II), the pectic polysaccharides of *Arabidopsis* cell wall, contain side chains rich in terminal galactose residues with β-(1, 4) and β-(1, 3) linkages (Zablackis et al, 1995). Besides pectins, xyloglucans from hemicellulose are also known to have terminal β-(1, 2) linked galactose residues that are susceptible to cleavage by β-galactosidases. Interestingly, xyloglucan oligosaccharides were

resistant to the action of Gal-5 and Gal-2. Both enzymes released galactose only from the hot water-soluble and ammonium oxalate-soluble pectic fractions of *Arabidopsis* cell wall, though at a slow rate (Table 4). Gal-5 and Gal-2 (~0.03nkat) released ~1 μg of galactose from 2 mg of *Arabidopsis* cell wall. The crude extract from rosette leaves released even less galactose under the same conditions. This observation is likely due to the complex structure of cell wall polysaccharides (e.g. RG-I) that contain side chains with linear and branched α-L-arabinose and β-D-galactose residues that sometimes can be substituted by α-L-fucose, β-D-glucuronic acid, and 4-O-methyl-β-D-glucuronic acid residues. Such substitutions make galactose residues inaccessible to the action of β-galactosidases and hence limit hydrolysis (Zablackis et al., 1995, Kotake et al. 2005, Iglesias et al., 2006). It should be noted that the limited action of Gal-5 and Gal-2 on the side chains of the pectic backbone can increase porosity of the matrix that creates microenvironments in cell walls in vivo, which in turn may control accessibility of other cell wall-degrading enzymes to their substrates (Smith et al., 2002, Verbelen and Vissenberg, 2007).

Conclusions

Most of the enzymes and structural proteins that are directly involved in construction and functioning of *Arabidopsis* cell wall are encoded by multigene families (Farrokhi et al., 2006). These families consist of members sharing structural similarity, but differing in their temporal and spatial expression profiles and physiological functions. GH family 35 enzymes, consisting of 17 putative β-galactosidases in *Arabidopsis,* are believed to be involved in cell wall dynamics (Imoto et al., 2005, Ahn et al., 2007, Verbelen and Vissenberg, 2007). We studied two members

Table 4. Natural substrate specificities of Gal-5 and Gal-2.

Substrate	Linkage of terminal residue	Activity[a]		
		Gal-5	**Gal-2**	**Crude**
Arabidopsis cell wall, hot water soluble pectin	Unknown	**0.7**	**0.7**	**0.3**
Arabidopsis cell wall, oxalate soluble pectin	Unknown	**0.6**	**1.3**	**0.6**
Arabidopsis cell wall, alkali soluble hemicellulose I	Unknown	**tr**	**0.7**	**0.6**
Arabidopsis cell wall, alkali soluble hemicellulose II	Unknown	**tr**	**tr**	**tr**
Lupin galactan	Galβ–(1, 4)Gal	**1.7**	**2.5**	**N/A**
Apple pectin	unknown	**0.8**	**4.5**	**1.5**
Gum Arabic	Galβ–(1, 3), (1, 6) Gal	**0**	**0**	**N/A**
Gum Guar	Manβ-(1, 4)Man, Galα–(1, 6)Gal	**0**	**0**	**N/A**

[a]- μg of Gal released from 2 mg of polysaccharide by 0.03 nkat (units) of enzyme at 25ºC in 24 h,
tr- trace (poor hydrolysis), N/A-not tested

of the *Arabidopsis* β-galactosidase family, Gal-5 and Gal-2. Western blot analysis using peptide-specific antibodies revealed organ-specific expression of the two genes encoding these enzymes. We showed that Gal-5 and Gal-2 are present in and tightly associated with the cell wall in *Arabidopsis* by using peptide-specific antisera and dot blotting. Recombinant Gal-5 and Gal-2 expressed in *P. pastoris* hydrolyzed various synthetic galactosidases, galacto-oligosaccharides and cell wall-derived polysaccharides. Both enzymes preferentially cleaved galactosides containing β-(1, 4) and β-(1, 3) linkages. The properties of the enzymes and their natural substrate specificities suggest that they may have the potential to be involved in modification of pectic polysaccharides of cell wall matrices. Further studies are needed to understand their biological roles as pectin-modifying enzymes.

3. Experimental

3.1. Materials

cDNAs for Gal-5 (pda05881 or pda06378) and Gal-2 (pda01770) in pBluescript vector were obtained from RIKEN, Institute of Physical and Chemical Research, Japan. Enzymes for the cloning were from Stratagene (La Jolla, CA) and NEB (Ipswich, MA). Easy select *Pichia* expression kit was from Invitrogen (Carlsbad, CA). Synthetic substrates and other chemicals were from Sigma (St. Louis, Mo). The imject maleimide activated BSA conjugation kit was from Pierce (Rockford, IL). Total Galactose Neonatal Screening Test Kit was from Interscientific (Hollywood, FL). Galactose dehydrogenase was from Roche (Indianapolis, IN). Galacto-oligosaccharides were a gift from Dr. Yoichi Tsumuraya and Dr. Toshihisa Kotake of Saitama University, Japan. Lupin galactan was a gift from Dr. David Smith of USDA, Beltsville, MD.

3.2. Methods

Database analysis

Signal sequence predictions and subcellular targeting of predicted proteins were done using SignalP (Bendtsen et al., 2004) and TargetP (Emanuelsson et al., 2000). Amino acid sequences of β-galactosidases were aligned using software at http://align.genome.jp. A phylogenetic tree was constructed from the alignment using PAUP 4.0.

Microarray expression analysis

Expression profile analysis was done using *Arabidopsis* gene expression datasets from the Genevestigator website (http://www.genevestigator.ethz.ch). Using Gene Atlas Tool, the organ/tissue-specific expression levels for At1g45130 (*Gal-5*) and At3g52840 (*Gal-2*) were estimated by Genevestigator software along with Atg05320 (Ubiquitin10, *UBQ10*) as the control for comparison. A given gene was scored as "expressed" if data from the Digital Northern Tool gave signal values higher than 200 with $p<0.06$ (Zimmermann, 2004). The p value for *Gal-5* and *Gal-2* was $p=0.00164$. Response Viewer Tool was used to verify up- and down- regulated genes under different abiotic and biotic stresses. Reliability and reproducibility of analyses were evaluated by the number of chips and replicates in individual experiments.

Plant materials

Arabidopsis seeds (Col-O) were obtained from the Arabidopsis Biological Resource Center (ABCR), Ohio State University Seed Stock Center (Columbus, OH). For germination, seeds were surface-sterilized with 3% hypochlorite for 10 minutes followed by washes with

dH_2O (three times) and finally suspended in 0.1% agarose. Sterilized seeds were kept at 4ºC for 3-4 days and seedlings were germinated on half strength Murashige-Skoog salt-agar plates for 10-14 days with 16 h day and 8 h night cycles. Seedlings were transferred to soil and grown at 16/8 h day/night cycle. Plants were harvested when 4-5 weeks old, and immediately frozen in liquid nitrogen and kept at -80ºC until use. For western blot analysis, *Arabidopsis* tissue was ground with sand (0.3 g/1 g tissue). Total proteins were solubilized in 6M urea (1g tissue: 2 ml solvent). Cell wall polysaccharide isolation was done as described in Li et al. (2001).

Expression of Gal-2 in E. coli and preparation of rabbit antisera

The Gal-2 mature protein coding sequence was cloned into pET21a vector and expressed in *E. coli* BL21 codon plus cells. Cells were suspended in lysis buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.02% SDS, 1 mM PMSF) and broken up using a French press. After extensive washing of soluble fractions with lysis buffer, insoluble proteins (the inclusion body fraction) were solubilized in 6 M urea and separated on a 10% SDS-PAGE preparative gel. The gel was stained (30 min) with Coomassie brilliant blue R-250 and the band corresponding to the Gal-2 polypeptide was excised. The excised band was destained in 50% methanol with several changes of solution and rehydrated in a minimum amount of 1X PBS at 4ºC overnight. After rehydration, the band was ground in a pre-chilled mortar. Ground powder was suspended in 1X PBS containing 0.2% SDS and 0.5% 2-mercaptoethanol (v/v) and heated at 75ºC for 15 min. After cooling, the suspension was mixed with 1 volume of Freund's Complete Adjuvant (Sigma) and used for immunization. Rabbit anti-Gal-2 sera were raised by repeated injection of antigen mixed with Freund's Incomplete Adjuvant at two-week intervals. Synthetic peptides (Gal-2: CSGKIRAPTILMKMIPTS and Gal-5: CSGVAFLTNYHMNAPAKVV) were conjugated to the

BSA using an imject maleimide activated BSA kit (Pierce, Rockford, IL) according to the vendor's protocol. Peptide-specific antisera were raised by injecting BSA-conjugated synthetic peptides with Freund's adjuvants. A small volume of trial bleeding was taken at two-week intervals to monitor the change in antisera titer during the course of immunization. All antisera were diluted twice with glycerol and stored at -20°C until usage.

Expression of recombinant Gal-5 and Gal-2 in P. pastoris and purification

The mature protein coding sequences of Gal-5 (S24 through N724) and Gal-2 (V28 through K727) cDNAs were amplified by the primer pair 5'-CAC CGT GGT CAC TTA TGA TCA CAA AGC-3' and 5'-CCA ATG AAA GAG GGT AAC AAA GGGC-3' for Gal-2 and 5'-AGG TGA ATT CCA GTG TAG TAG TGT AAC CTA CG-3' and 5'-TTT GCG GCC GCA AGT TAG TTT ACT GAT CTC TTC ACA AC-3' for Gal-5 from cDNA inserts of plasmids obtained from RIKEN, using the high-fidelity *Pfu* Turbo DNA polymerase. The inserts were cloned into pPICZα *Pichia* expression vector to express Gal-5 and Gal-2 as the yeast α factor secretion signal fusion protein to facilitate secretion of recombinant proteins into culture medium. After confirming the accuracy of the sequence and the correct reading frame, linearized plasmids were transformed into *P. pastoris* by electroporation. Recombinant enzyme production was under the control of the alcohol oxidase (AOX) promoter induced by methanol. Production of recombinant proteins was monitored by assaying β-galactosidase activity toward *p*NPGal in culture supernatant samples taken every 24 h. After 72 hours of induction, cells were pelleted and culture supernatant was used for further purification of the recombinant enzymes.

Gal-5 and Gal-2 were purified from culture supernatants by ion exchange chromatography using Sulphoxyethyl (SE) cellulose. In a typical experiment, 100 ml of culture

supernatant was filtered, diluted five times with degassed dH_2O to reduce ionic strength of the medium and loaded onto the column (1.5 cm x 3 cm) pre-equilibrated with buffer A (20 mM potassium phosphate, pH 6.0). After washing with 10 column volumes of buffer A, bound proteins were eluted in one step with 150 mM NaCl in buffer A (flow rate 1ml/min). All fractions were checked for β−galactosidase activity using *p*NPGal as substrate. Fractions with highest β-galactosidase activity were used for further experiments. Protein concentration was determined by the Bradford method (Bio-Rad Protein Assay Reagent kit) using BSA as a standard.

Cell wall isolation and extraction of cell wall-bound proteins

We isolated cell wall from rosette leaves of *Arabidopsis* using the procedure described by Feiz et al. (2006). Rosette leaves were ground in a blender whose cup was dipped at intervals into liquid nitrogen to maintain low temperature during grinding. The cell wall fraction was washed extensively with 3 L of wash buffer on a metal net (75 μm pore size). After washing, the cell wall fraction was lyophilized. The lyophilized cell wall material was ground to a fine powder by grinding in a blender, and was then used to extract wall-bound proteins. Five different (S1, soluble 1; S2, soluble 2; S3, soluble 3; CW4, extractable with $CaCl_2$; and CW5, extractable with LiCl) fractions were obtained. These fractions were assayed for β-galactosidase activity. Of these, three fractions with β-galactosidase activity were further analyzed for immunoreactivity. They were spotted multiple times on nitrocellulose strips to increase antigen (Gal-5 and Gal-2) concentration and incubated with preimmune (control) and immune sera from rabbits immunized with whole Gal-2 polypeptide and unique peptides derived from Gal-5 and Gal-2 sequences.

SDS-PAGE, native PAGE and western blotting

SDS-PAGE was performed as described by Laemmli (1970). Native PAGE was performed in acidic gels using the protocol on the website (http://wolfson.huji.ac.il/purification/Protocols/PAGE_Acidic.html). After electrophoresis, the gel was rinsed in a wash buffer (100 mM acetate buffer pH 4.6) for 2x15 min. The zymogram was developed by incubating the gel in 0.5 mM 4-MUGal in wash buffer at 37ºC for 20 min and photographed under UV light. For immunoblotting, the gel was soaked in a blotting buffer (10 mM CAPS, pH 11 with 10 % (v/v) methanol) for 2x15 min. Proteins were transferred onto a nitrocellulose (0.45 μm, Protran) membrane using a Bio-Rad Mini trans blot cell at 50V, at 4ºC overnight following the vendor's protocol. For immunodetection, 2000-times dilution of anti-Gal-2 antiserum or 1000-times dilution of peptide-specific antiserum was used as primary antibody and 2000-times dilution of goat anti-rabbit antibody conjugated with peroxidase (A0545, Sigma, Saint Louis, MO) as secondary antibody. Immunoreactive bands were visualized by the deposition of 4-chloronaphthol after oxidation by HRP (horse radish peroxidase) using the substrate solution (21 ml of PBS pH 7.4, mixed with 5.5 ml of 3.3 mg/ml 4-chloronaphthol in 100% MeOH and 10 μl of 30% H_2O_2).

β-galactosidase activity assay

One hundred μl of 5 mM *p*NPGal in 100 mM NaOAc buffer pH 4.6, 80 μl H_2O and 20 μl of the diluted enzyme solution were mixed and incubated at 37ºC for up to 30 minutes. The reaction was stopped by adding 100 μl of 1M Na_2CO_3. The absorbance was measured at 405 nm to quantify the amount of *p*NP released after hydrolysis. This standard protocol was used for all activity assays with *p*NPGal, if not otherwise stated. Boiled enzyme or buffer solution was used

as a control. One unit (nkat) of enzyme activity is defined as an amount of enzyme that is able to produce one nmole of *p*NP per second at 37ºC.

For the determination of natural substrate specificity, galacto-oligosaccharides (20 mM), xyloglucan oligosaccharides (1 μg/μl) and polysaccharides (1% (w/v)) were prepared in water. Final concentration of substrates was 4 mM for oligosaccharides and 0.5% for polysaccharides in 100 mM acetate buffer pH 4.6. Reaction mixture was incubated with 0.03 units/nkats enzymes at room temperature for 24 h. Reaction mixtures containing no enzyme and no substrate were used as controls. Reactions for polysaccharides were stopped by adding 1 ml of 100% EtOH to 0.4 ml of reaction mix to precipitate proteins and polysaccharides. After centrifugation, the supernatant was transferred into a new microfuge tube and vacuum dried. Dried mixtures were dissolved in 100 μl of dH_2O, and total galactose produced as a result of hydrolysis was quantified using galactose dehydrogenase assay kit (Interscientific, Hollywood, FL). Products of hydrolysis of oligosaccharides were analyzed by thin layer chromatography (TLC) on silica gel $60F_{254}$ (EM Science, Germany) using 3:2:1 (v/v/v) butanol: acetic acid: water as the solvent and detected by heating TLC plates after spraying with 0.2% (w/v) naphthoresorcinol in 1:19 H_2SO_4: ethanol (v/v) (Ahn Young Ock, 2004)

Acknowledgements

We are grateful to Drs. Yoichi Tsumuraya and Dr. Toshihisa Kotake of Saitama University, Japan for providing galacto-oligosaccharides, and Dr. David Smith of USDA, Beltsville, MD for providing lupin galactan. The authors also wish to thank Dr. Farooqahmed Kittur of Virginia Tech for much help and critical reading of the manuscript. This research is funded by the *Arabidopsis* 2010 Project of the National Science Foundation (MCB-0115937) and a research grant award from the Virginia Academy of Science to D. Gantulga.

REFERENCES

Ahn Y.O., Zheng M., Winkel B., Bevan D. R., Esen A., Shin-Han S., Benson J., Peng H., Miller J.T., Cheng C., Poulton J.E., and Shih M., 2007. Functional genomic analysis of *Arabidopsis thaliana* Glycoside Hydrolase Family 35. Phytochemistry 68, 1510-1520.

Ahn Y.O, Mizutani M., Hiromichi S., and Kanzo S., 2004. Furcatin Hydrolase from *Viburnum furcatum* Blume is a Novel Disaccharide-specific Acuminosidase in Glycosyl Hydrolase Family 1. J. Biol. Chem. 279, 23405-23414.

Balasubramaniam S., Lee H., Lazan H., Othman R., and Ali Z.M., 2005. Purification and properties of a β-galactosidase from carambola fruit with significant activity toward cell wall polysaccharides. Phytochemistry 66, 153-163.

Bayreuther K., Bieseler D.J., Ehring R., Griesser H.W., Mieschendahl M., Muller-Hill B., Triesch I., 1980. Investigation of Structure and Function of Lactose Permease of *Escherichia coli*. Biochem. Soc. Trans. 8, 675-676.

Bendtsen J., Nielsen H., von Heijne G., and Brunak S., 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. 340, 783-795.

Bradford M.M., 1976 A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248-254.

Buckeridge M. S. and Reid J. S., 1994. Purification and properties of a novel beta-galactosidase or exo-(1-->4)-beta-D galactanase from the cotyledons of germinated *Lupinus angustifolius L.* seeds. Planta 192(4), 502-511.

Buckeridge M.S., Huntcheon I.S., and Grant Reid J.S., 2005. The role of exo-(1-4)-β-galactanase in the mobilization of polysaccharides from the cotyledon cell wall of *Lupinus angustifolus* following germination. Ann. Botany. 96, 435-444.

Carey A.T., Holt K., Picard S., Wilde R., Tucker G.A., Bird C.R., Schuch W., Seymour G.B., 1995. Tomato exo-(1-4)-D-galactanase. Isolation, changes during ripening in normal and mutant tomato fruit, and characterization of related cDNA clone. Plant Physiol. 108, 1099-1107.

Chantarangsee M., Fujimura T., Fry S.C., Ketudat-Cairns J.R., 2007. Molecular characterization of β-galactosidases from germinating rice (Oryza sativa). Plant Sci. 173, 118-134

Coutinho, P.M. and Henrissat, B., 1999. Carbohydrate-active enzymes: an integrated database approach. In "Recent Advances in Carbohydrate Bioengineering", H.J. Gilbert, G. Davies, B. Henrissat and B. Svensson eds. The Royal Society of Chemistry, Cambridge, pp. 3-12.

Davies G. and Henrissat B., 1995. Structures and mechanism of glycosyl hydrolases. Structure. 3, 853-859.

de Alcantara P.H.N., Martim L., Silva C.O., Dietrich S.M.C., Buckeridge M.S., 2006. Purification of a β-galactosidase from cotyledons of *Hymenaea courbaril L.* (Leguminosae). Enzyme properties and biological function. Plant Physiol. Biochem. 44, 619-627.

Edwards M., Bowman Y., Dea I., and Reid J., 1988. A β-D-galactosidase from Nasturtium (*Tropaeolum majus L.*) cotyledons. J. Biol. Chem. 9, 4333-4337.

Emanuelsson O., Nielsen H., Brunak S. and von Heijne G., 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J. Mol. Biol. 300, 1005-1016.

Farrokhi N., Burton R. A., Brownfield L., Hrmova M., Wilson S. M., Bacic A., and Fincher G. B., 2006. Plant cell wall biogenesis: genetic, biochemical and functional genomics approaches to the identification of key genes. Plant Biotech. 4, 145-167

Feiz L., Irshad M., Pont-Lezica R. F., Canut H., and Jamet E., 2006. Evaluation of cell wall preparations for proteomics: a new procedure for purifying cell walls from *Arabidopsis* hypocotyls. Plant Methods, 2:10

Grebe M., 2005, Growth by Auxin: when a weed needs acid. Science 310, 60-61

Jamet E., Canut H., Boudart G. and Pont-Lezica R.F., 2006. Cell wall proteins: a new insight through proteomics. Trends Plant Sci. 11, 33-39.

Kotake T., Dina S., Konichi T., Kaneko S., Igarashi K., Samejima M., Watanbe Y., Kimura K., Tsumuraya Y., 2005. Molecular cloning of a β-galactosidase from radish that specifically hydrolyzes β-(1-3) and β-(1-6)-galactosyl residues of arabinogalactan protein. Plant Physiol. 138, 1563-1576.

Iglesias N., Abelenda J. A., Rodino M., Samredro J., Revilla G., and Zarra I., 2006. Apoplastic glycosidase active against xyloglucan oligosaccharides of *Arabidopsis thaliana*. Plant Cell Physiol. 47(1), 55-63.

Imoto K, Yokoyama R, Nishitani K., 2005. Comprehensive approach to genes involved in cell wall modifications in *Arabidopsis thaliana*. Plant Mol Biol. 58(2), 177-92.

Ishimaru M., Smith D.L., and Gross K.C., 2005. Yeast expressed tomato β-galactosidase 1, 4, and 5 have activity against synthetic and plant-derived cell wall substrates. HortScience 40, 1092.

Hruba P., Honys D., Twell D., Capkova V., Tupy J., 2005. Expression of β-galactosidase and β-xylosidase genes during microspore and pollen development. Planta 220, 931-940

Laemmli U.K., 1970. Cleavage of structural proteins during assembly of the head of bacteriophage T4. Nature 277, 680-685.

Lazan H. Ng S.Y., Goh L.Y, and Ali Z.M., 2004. Papaya β-galactosidase/galactanase isoforms in differential cell wall hydrolysis and fruit softening during ripening. Plant Physiol. Biochem. 42, 847-853.

Li S.C., Han J.W., Chen K.C., and Chen C.S., 2001. Purification and characterization of isoforms of β-galactosidases in mung bean seedling. Phytochemistry 57, 349-359

Perez Almeida I. B., 2004. *Arabidopsis* cell wall beta-galactosidase gene family: Expression, catalytic activities and biological function in galactose dynamics. Dissertation. Purdue University ETD.

Ross, G.S., Wegrzyn T., MacRae E.A., Redgwell R.J. 1994. Apple β-galactosidase activity against cell wall polysaccharides and characterization of related cDNA clone. Plant Physiol. 106, 521-528.

Smith D.L., and Gross K.C., 2000. A family of at least seven β-galactosidase genes is expressed during tomato fruit development. Plant Physiol. 123, 1173-1183.

Smith D.L., Abbott J.A., and Gross K.C., 2002. Down-regulation of tomato β-galactosidase 4 results in decreased fruit softening. Plant Physiol. 129, 1755-1762

Vervelen J. P. and Vissenberg K. The expanding cell. Plant Cell monographs, Springer. Vol. 5, 2007

Zablackis E., Huang J., Muller B., Darvill A. G., and Albersheim P., 1995. Characterization of the cell wall polysaccharides of *Arabidopsis thaliana* leaves. Plant Physiol. 107, 1129-1138.

Zimmermann, P., Hirsch-Hoffmann Matthias, Hennig Lars, Gruissem Wilhelm 2004. GENEVESTIGATOR. Arabidopsis microarray database and analysis tool. Plant Physiol. 136, 2621-2632.

Chapter 3

Comparative characterization of the *Arabidopsis* Subfamily III β-galactosidases

Keywords: *Arabidopsis*; Cell wall; Glycosyl hydrolase family 35; Subfamily III, β-Galactosidase; cell wall modification;

Abbreviations: *Arabidopsis thaliana, Arabidopsis*; Gal-1, (AtBGal1)–β−galactosidase-1, Gal-2, (AtBGal2)–β−galactosidase-2, Gal-3, (AtBGal3)–β−galactosidase-3, Gal-4, (AtBGal4)–β−galactosidase-4, Gal-4, (AtBGal5)–β−galactosidase-5; Gal-12, (AtBGal12)–β-galactosidase-12; *p*NP, para-nitrophenyl; GH, glycosyl hydrolase.

Abstract

The *Arabidopsis* genome encodes 17 predicted β−galactosidases that have been classified into seven subfamilies based on sequence similarity. The largest of these, Subfamily III, consists of six genes, *Gal-1* (At3g13750), *Gal-2* (At3g52840), *Gal-3* (At4g36360), *Gal-4* (At5g56870), *Gal-5* (At1g45130), and *Gal-12* (At4g26140), several of which were characterized in previous studies. We report here the purification and biochemical characterization of recombinant Gal-1, Gal-3, Gal-4 and Gal-12 from *Pichia pastoris*, completing the analysis of all six recombinant proteins, as well as the isolation and characterization of the native Gal-2 protein from *Arabidopsis* leaves. Comparison of the relative expression levels of the six *Gal* genes uncovered evidence of differential regulation. In addition, this study provides further support for the proposed function of the Subfamily III β−galactosidases in cell wall modification based on organ-specific expression by immunoblotting and subcellular localization by immunoEM of Gal-1 and Gal-12. Our study suggests that despite small differences in individual biochemical characteristics and expression patterns, each member of the family has the potential to contribute to cell wall dynamics.

1. Introduction

β-galactosidases (EC 3.2.1.23) are enzymes that cleave substrates containing galactosyl moieties, such as lactose, glycolipids, proteoglycans, oligosaccharides, and polysaccharides. β-galactosidases are of interest because of their use in the dairy and food industry, agriculture, and biotechnology (Callahan, 1999; Smith and Gross, 2000; Tanthanuch et al., 2008). Based on amino acid sequence similarities, β-galactosidases have been found to fall into four of the 113 current glycosyl hydrolase (GH) families, GH-1, GH-2, GH-35 and GH-42 (Coutinho and Henrissat, 1999; http://www.cazy.org/). β-galactosidases belonging to GH-1, GH-2, and GH-42 are found predominantly in microorganisms, whereas the GH-35 enzymes are found in both prokaryotes and eukaryotes. In plants the β-galactosidases belonging to GH-35 family constitute gene subfamilies with high sequence similarities, where in other eukaryotes the total number of β-galactosidase genes is much smaller.

Plant β-galactosidases play important roles in the metabolism of galactosyl conjugates during carbohydrate reserve mobilization, cell wall expansion and degradation, and turnover of signaling molecules (de Alcantara et al., 2006; Esteban et al., 2003; McDougall and Fry, 1990). A correlation between β-galactosidase expression and cell wall disassembly has been reported in fruit ripening of tomato (Carey et al., 1995), apple (Ross et al., 1994), kiwi (Ross et al., 1993), pear (Tateishi et al., 2001), and papaya (Lazan et al., 2004). Besides fruit ripening, β-galactosidases have been shown to participate in seed germination in *Arabidopsis* (Dean et al., 2007), radish (Kotake et al., 2005; Sekimata et al., 1989), and rice (Chantarangsee et al., 2007). β-galactosidase activities associated with various physiological processes have been observed in different plant tissues,

including the cotyledon of lupin (Buckeridge et al., 2005), floral tissue of Sandersonia flower (O'Donoghue et al., 2002), seedling of mung bean (Li et al., 2001), and pollen of tobacco (Hruba et al., 2005). The analysis of individual β-galactosidases purified from different plants has led to better understanding of their functions and natural substrate specificities. Further genetic and bioinformatics studies are necessary to understand the interrelationships among β-galactosidase isoforms and the mechanisms by which they participate in the regulation of growth and development.

For a systematic characterization of β-galactosidases, the most attractive organism is a model plant with a completely sequenced genome. In *Arabidopsis thaliana*, a gene family of 17 *β-galactosidases* has been identified (Ahn et al., 2007). Among these, the members of Subfamily III, which consists of Gal-1, Gal-2, Gal-3, Gal-4, Gal-5 and Gal-12, is the largest subfamily. The presence of such a large subfamily raises the question of functional redundancy of these isoforms in plants. Many homologs of Subfamily III genes have been studied in other plant species (Buckeridge et al., 2005; Carey et al., 1995; Chantarangsee et al., 2007; Lazan et al., 2004; Tateishi et al., 2001; Trainotti et al., 2001), but a characterization of the entire subfamily from a single species has not yet been carried out. We recently reported the biochemical characterization of Gal-4 (Ahn et al., 2007) and Gal-2 and Gal-5 (Gantulga et al., 2008), showing that these enzymes have high hydrolase activity with specific cell wall polysaccharides. We also demonstrated the presence of Gal-2 and Gal-5 proteins in the cell wall of rosette leaves by immuno dot-blotting. Our results suggested that the sequence similarities of the genes in the subfamily translate into functional similarities of the proteins in cell wall modification.

In this work we extend the previous analysis of Gal-2, Gal-4, and Gal-5 to complete the biochemical characterization of the Subfamily III β-galactosidases. We also describe the purification and characterization of a native β-galactosidase from this subfamily from *Arabidopsis* leaves. Detailed organ-specific expression of all six genes in the subfamily was studied by semi-quantative RT-PCR and compared with the publicly-available microarray data. We show that methyl jasmonate treatment, known to arrest growth and control defense responses in plants, down-regulates the expression of the *Gal* genes. Immunolocalization of Gal proteins using transmission electron microscopy (TEM) of roots confirmed the predicted localization of the Subfamily III galactosidases in the cell wall in vascular and epidermal tissues of mature root.

2. Results and discussion

2.1 In silico *characterization of the* Arabidopsis *Subfamily III β-galactosidases*

Previous analysis of the evolutionary relationships among the GH Family 35 *β-galactosidase* genes in 23 eukaryotic genomes showed that the 17 *Arabidopsis* genes fall into two groups and seven subfamilies (Ahn et al., 2007). Fig. 1 shows an unrooted tree based on the amino acid sequences predicted for these 17 genes. Gal-1, Gal-2, Gal-3, Gal-4, Gal-5, and Gal-12 comprise the largest subfamily, designated as Subfamily III (Gantulga et al., 2008). These six proteins share 60-81% sequence identity with each other. The other subfamilies each consist of one to three members. All six members of Subfamily III are predicted to have signal peptides and basic pIs (7.2-8.6), consistent with a cellular destination in the cell wall (Ahn et al., 2007). This localization was recently

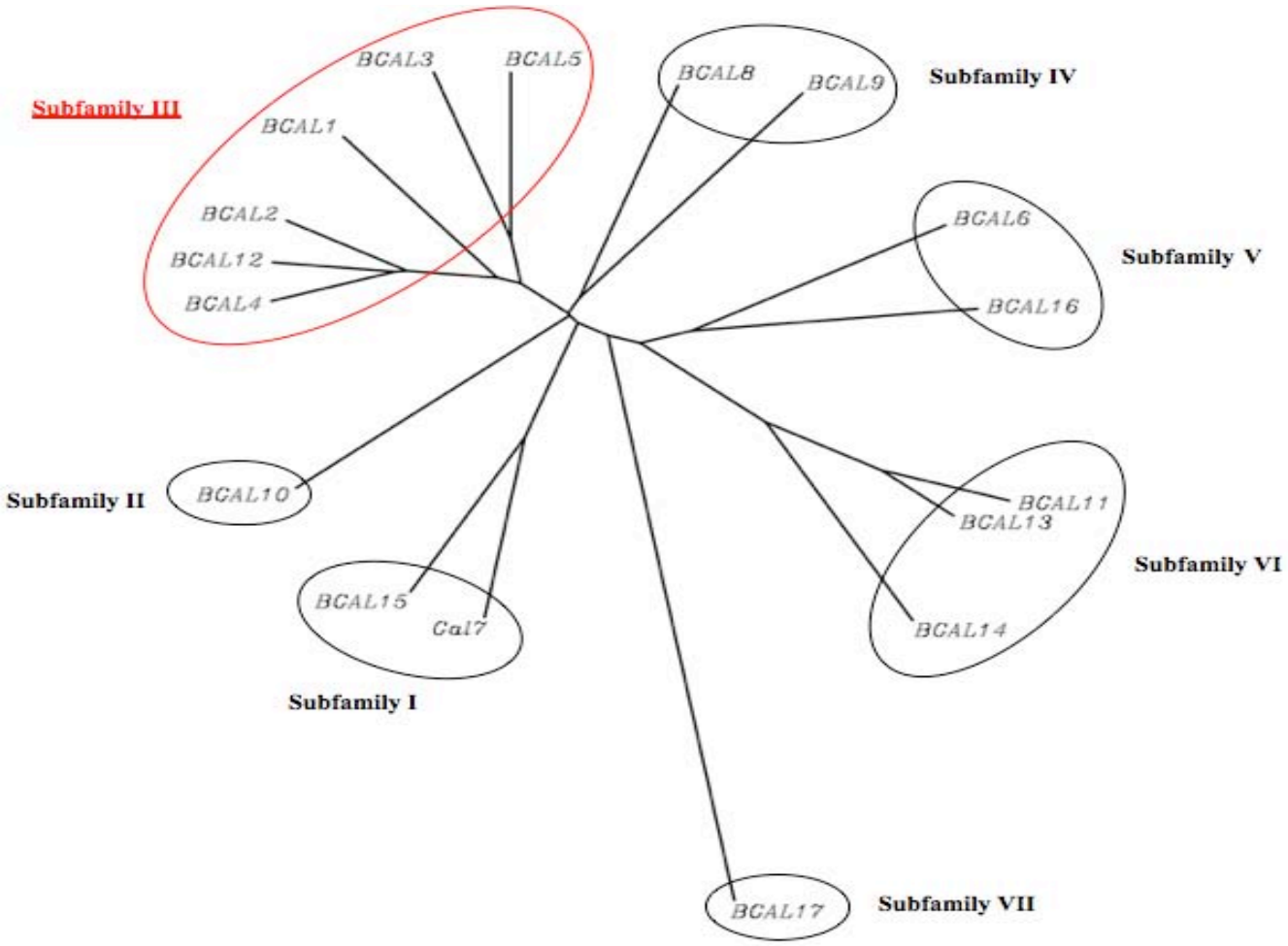


Fig. 1. Phylogenetic relationships among the *Arabidopsis thaliana* β-galactosidases.

The tree of the β−galactosidase family is derived from an alignment of 17 members that are divided into 7 subfamilies based on the amino acid sequence.

confirmed for Gal-2 and Gal-5 (Gantulga et al., 2008). The six enzymes differ with regard to the presence of a C-terminal lectin-like domain in Gal-1 and Gal-3, which places these two isoforms with the higher molecular mass Group 2 β-galactosidases (Ahn et al., 2007). It has been suggested that the lectin-like domains may enhance catalytic efficacy by anchoring the enzymes to their polymeric substrates.

2.2 Expression of β-galactosidases in Arabidopsis

As a first approach to characterizing the relative expression patterns of the Subfamily III genes, the publicly available microarray data were examined using Genevestigator (Zimmermann et al., 2004). We previously used this database to determine that *Gal-2* and *Gal-5* exhibit differential, tissue-specific expression patterns (Gantulga et al., 2008). Extending this analysis to include all six members of Subfamily III showed that the genes are expressed in all tissues examined, but to different degrees (Fig. 2). *Gal-1* was the most abundantly expressed of the six genes in most tissues, with the highest levels in the petiole, stem, and cauline leaf and moderate levels in all other organs. *Gal-3* had the highest expression in the stem and flowers, while *Gal-4* transcripts were most abundant in silique, stem, root and cauline leaves. The relative levels of the *Gal-12* transcripts were low in almost all organs except root, specifically root hairs.

To confirm these findings and to extend the analysis of expression patterns to include such organs as mature root, cauline leaf, petiole of the rosette leaf, and silique, semi-quantitative RT-PCR was performed. These experiments provided further evidence for temporal and spatial differences in the expression of the six Subfamily III genes (Fig. 3A). Most of the genes are expressed in all organs, although the relative abundance of the

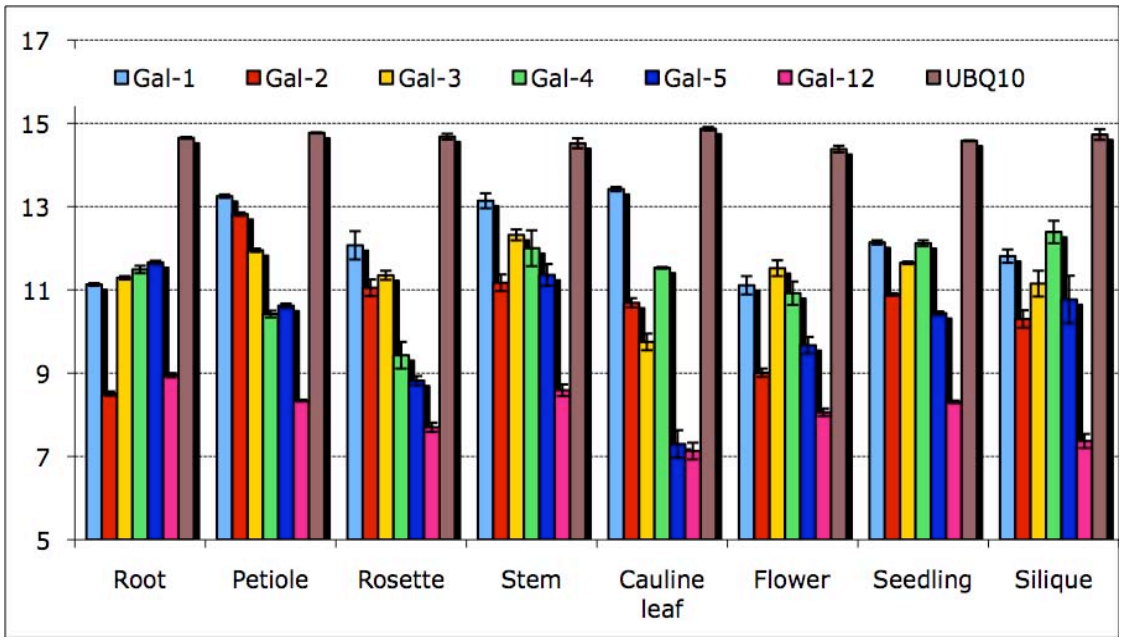


Fig. 2. Expression profile of the Subfamily III β-galactosidases from the public microarray data. The organ/tissue-specific expression levels for *Gal-1*, *Gal-2*, *Gal-3*, *Gal-4*, *Gal-5*, and *Gal-12* are shown as log2(n) based on data obtained using Genevestigator {Zimmermann, 2004 #4}. Data for *ubiquitin-10* (UBQ10) was included for comparison. The Gene Atlas Tool of Genevestigator (https://www.genevestigator.ethz.ch/at/) was used for the analysis. Numbers on the Y axis represent the average signal intensity.

transcripts varied. Only three of the genes, *Gal-1, Gal-3, and Gal-4,* were expressed in siliques. *Gal-1* transcripts were detected in all organs: root, leaf, stem, flower, siliques, and young seedling, Fig. 3A. The *Gal-1* transcript level was the highest in root followed by leaves and flower. The *Gal-2* transcript was the most abundant in young seedling and was lowest in root and flower. The *Gal-3* transcript was highest in stem and silique, but was low in all other organs examined. The *Gal-4* transcript was abundant in root, stem and flower, but was low in leaves. The *Gal-5* and *Gal-12* transcripts were detected at the highest levels in green tissues and were lowest in roots. We detected two bands for the *Gal-12* RT-PCR product. This is consistent with the gene model for *Gal-12* in TAIR, which reflects alternative splicing resulting in products that differ in size by 100 bp. In summary, the *Gal-2, Gal-5, Gal-12* genes are expressed predominantly early in development of the *Arabidopsis* plant, while *Gal-3* and *Gal-4* are expressed later in development. In contrast, *Gal-1* gene expression was similar in all stages and tissues examined (Fig. 2).

We investigated the effects of a variety of stress conditions and hormone treatments known to regulate plant growth and development (Nemhauser et al., 2006) on expression of the Subfamily III genes in seedlings. Little or no difference was observed in the transcript levels of the six genes in response to a variety of treatments, including salt, cold, osmotic shock, or treatment with salicylic acid, 2,4-D, ACC, and ABA (data not shown). This observation is consistent with publicly available microarray data (Genevestigator) where the change in expression of these genes in response to similar treatments was not significant compared to the cases of top-ranked, differentially-regulated genes. However, notably reduced transcript levels were found for all six genes

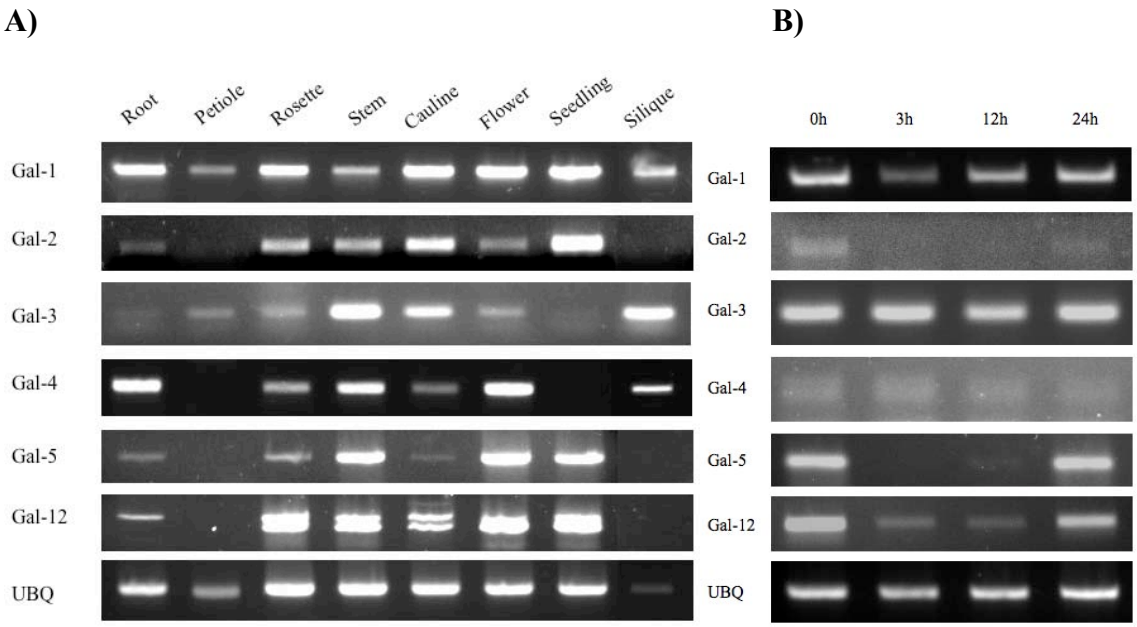


Fig. 3. Semiquantitative RT-PCR analysis of Subfamily III gene expression. A, organ specific expression in five-week-old plant tissues; B, time course of the response to jasmonic acid in ten-day-old seedlings. Relative expression levels of the six *β–galactosidase* genes are shown relative to the *ubiquitin-10* (UBQ) gene.

in seedlings treated with methyl jasmonate. Analysis of the time course of the response to methyl jasmonate showed that expression of *Gal-1, Gal-2, Gal-5* and *Gal-12* decreased notably only after 3h of treatment, indicating that these are late-response genes (Fig. 3B). The transcript levels returned to pre-treatment level within 24 h. Under our experimental conditions, expression of *Gal-3* and *Gal-4* was not affected as dramatically as the other four genes. As jasmonate is known to modulate pathways controlling cell growth of *Arabidopsis* (Liu and Mehdy, 2007; Staswick et al., 1992), these data provide further support to the hypothesis that the Subfamily III β-galactosidases are involved in growth control.

2.3 Expression and purification of recombinant Gal-1, Gal-3, Gal-4 and Gal-12 in P. pastoris

We recently reported the biochemical characterization of recombinant Gal-2 and Gal-5 expressed in *Pichia pastoris* (Gantulga et al., 2008) and of Gal-4 expressed in *E. coli* and insect cells (Ahn et al., 2007) . In order to facilitate direct comparison of the biochemical characteristics of the six Subfamily III enzymes, Gal-1, Gal-3, Gal-4, and Gal-12 were also expressed in *P. pastoris.* The mature proteins without signal peptides were targeted for secretion into the culture medium, as had been done previously for Gal-2 and Gal-5 (Gantulga et al., 2008). Gal-1 and Gal-4 were purified by cation exchange chromatography using sulphoxyethyl cellulose (SE). Gal-3 and Gal-12 were purified by hydrophobic interaction chromatography followed by cation exchange chromatography. The purified proteins appeared as single bands by SDS –PAGE analysis, with estimated monomeric molecular weights in the range of 80~90 kD (data not shown), consistent with

sizes predicted for these proteins, 90 kD, 92 kD, 78 kD and 79 kD, respectively for Gal-1, Gal-3, Gal-4, and Gal-12. It is worth mentioning that substantial differences were observed in the efficiency with which the recombinant proteins were secreted into the culture medium. Although Gal-2, Gal-4, Gal-5, and Gal-12 were present at relatively high levels in the medium, the levels of Gal-1 and Gal-3 were extremely low, despite the moderate overall expression levels based on total β-galactosidase activity. The majority of the activity remained in the cell pellet according to this assay. A modified method was developed for purification of Gal-1 and Gal-3 from *P. pastoris,* starting with larger initial culture volumes and longer induction times that substantially enhanced recovery (see Experimental, section 3.2). It is interesting to note that Gal-1 and Gal-3 are the two members of the subfamily with a C-terminal lectin-like domain. It is thus possible that these enzymes are retained within the yeast cell due to interaction of this domain with the cell wall.

2.4 Properties of recombinant β-galactosidases

The biochemical properties of the recombinant Gal-1, Gal-3, Gal-4, and Gal-12 proteins produced in *P. pastoris* were characterized as described previously for Gal-2 and Gal-5 (Gantulga et al., 2008). All four enzymes showed maximum activity with *p*NP-β–D-galactopyranoside at pH ~4.0 and in the temperature range of ~40-50°C (data not shown). The enzymes were stable at temperatures up to 50^{o}C and lost their activity rapidly above 55^{o}C, presumably due to heat denaturation (data not shown). The biochemical properties and substrate specificity of Gal-4 were similar, whether the protein was expressed in *P. pastoris, E.coli,* or insect cells (Ahn et al., 2007).

To determine the substrate preferences of the β-galactosidases for different sugar and aglycone moieties, the ability of the recombinant Gal-1, Gal-3, Gal-4 and Gal-12 enzymes to cleave different glycoside substrates was tested using (i) *p*NP-β-D-galactopyranoside (*p*NPGal), (ii) *p*NP-β-D-mannopyranoside, (iii) *p*NP-β-D-fucopyranoside, (iv) *p*NP-β-D-xylofuranoside, (v) *p*NP-β-D-arabinopyranoside, (vi) *p*NP-α-L-arabinopyranoside (*p*NPAra), (vii) *o*NP-β-D-galactopyranoside (*o*NPGal), (viii) 4-methylumbelliferyl-β-D-galactopyranoside (4-MUGal), (ix) 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal), and (x) 6-bromo-2-naphthyl-β-D-galactopyranoside (6-BNGal). All four enzymes showed a strict specificity for galactose and its 6-deoxy analogue, fucose (Tables I and II), indicating that only sugars with hydroxyl groups at the C3 and C4 positions are bound and hydrolyzed by these enzymes. None of the other glycosides that were tested were hydrolyzed by any of the family III isoforms (Table II). In contrast to the strict sugar specificity, the enzymes showed a broad aglycone specificity. The four enzymes hydrolyzed *o*NPGal, 4-MUGal, X-Gal, and 6-BNGal, although to different extents (data not shown).

Kinetic parameters for Gal-1, Gal-3, Gal-4, and Gal-12 with several different substrates were estimated from Lineweaver-Burk plots. The apparent K_m and V_{max} values for *p*NPGal, *p*NPFuc, and *o*NPGal are shown in Table I. These experiments show that the different isozymes have different K_m and V_{max} values with these substrates, indicating that the enzymes have different active site requirements for substrate binding and hydrolysis. Although no three-dimensional structures are yet available for plant β-galactosidases, the predicted active site residues are almost identical among the six isoforms of subfamily III.

Table I. Kinetic parameters[a]

Enzyme	*p*NPGal		*o*NPGal		*p*NPFuc	
	K_m mM	V_{max} nkat s^{-1}	K_m mM	V_{max} nkat s^{-1}	K_m mM	V_{max} nkat s^{-1}
Gal-1	3.48 (±0.11)	1324 (±67)	2.5 (±0.19)	1099 (±35)	5.34 (±1.05)	285 (±3.5)
Gal-3	0.32 (±0.06)	977 (±22)	1.9 (±0.25)	644 (±22)	5.69 (±1.9)	347 (±71)
Gal-4	1.71 (±0.13)	560 (±38)	1.28 (±0.18)	831 (±200)	3.6 (±0.8)	53.4 (±3)
Gal-12	0.71 (±0.02)	314 (±20)	4.2 (±0.4)	481 (±73)	15.8 (±0.1)	250 (±76)
nGal-2	0.7 (±0.09)	265 (±21)	n/a	n/a	n/a	n/a

[a]Enzyme activities were assayed in reaction mixture containing 0.1, 0.5, 1, 2.5, and 5 mM substrates in 50 mM sodium acetate buffer pH 4.6. K_m and V_{max} values were calculated from Lineweaver-Burk plots. Numbers represent mean ± SE (standard error) n=3 repicates.
n/a- not applicable.

Table II. Sugar specificities of Gal-1, Gal-3, Gal-4 and Gal-12

Substrate	Relative activity[a], %				
	Gal-1	**Gal-3**	**Gal-4**	**Gal-12**	**nGal-2**
*p*NP-β-D-**galactopyranoside**	100	100	100	100	100
*p*NP-β-D-**glucopyranoside**	<1	<1	<1	<1	<1
*p*NP-β-D-**mannopyranoside**	0	0	0	0	n/a
*p*NP-β-D-**xylofuranoside**	0	0	0	0	n/a
*p*NP-β-D-**arabinopyranoside**	0	0	0	0	n/a
*p*NP-α-L-**arabinopyranoside**	0	0	0	0	n/a

[a]Enzyme activities were assayed in reaction mixtures containing 2.5 mM substrate in 50 mM NaOAc buffer pH 4.6. Sugar specificity is expressed as a percentage of activity against *p*NPGal (100% ~ 0.04 nkat/unit). Numbers represent mean ± SE (standard error) for N=3 replicates.
n/a- not applicable.

A number of sugars and sugar derivatives were tested for their ability to inhibit the hydrolysis of *p*NPGal by the four enzymes. The strongest inhibitors were γ-galactone-lactone and D-galactose; the K_i values are presented in Table III. D-fucose, methyl-α-D-galactoside, and raffinose, also inhibited hydrolysis of *p*NPGal by the enzymes (data not shown). All were competitive inhibitors, leading to increased Km values for *p*NPGal. Enzyme activities were not affected by the presence of *p*NPGlc, *p*NPAra, lactose, IPTG, galacturonic acid, L-arabinose, and D-mannose in the reaction mixtures. Common non-competitive inhibitors and denaturants were also tested for effects on the activities of these proteins. Ag^{+} (silver), Hg^{2+} (mercury), and SDS completely abolished the activities of all four enzymes at concentrations of 15 μM, 15 μM and 0.05%, respectively. This indicates that these enzymes are very sensitive to these types of inhibitors.

In order to explore the action of the enzymes on natural polysaccharide substrates, the linkage specificities of the four β-galactosidases were examined. All four enzymes hydrolyzed β-(1, 4)- and β-(1, 3)-linked galacto-oligosaccharides, but only Gal-12 was able to cleave the β-(1, 6) linkage in galactobiose and galactotriose (data not shown). Together with the previous findings for Gal-2 and Gal-5 (Gantulga et al., 2008), this indicates that Subfamily III β-galactosidases prefer β-(1, 4)- and β-(1, 3)- linkages. These are the types of linkages found in pectic polysaccharides of the cell wall (Zablackis et al., 1995), indicating that pectic galactans may be the natural substrates for these enzymes. Indeed, all four enzymes hydrolyzed hot water- and oxalate-soluble pectins from *Arabidopsis* cell wall and commercially prepared apple pectin, releasing sufficient quantities of free galactose to be detectable by galactose dehydrogenase assay (Table IV),

Table III. Inhibition coefficients

Enzyme	K_i[a]	
	γ-galactonelactone μM	D-galactose mM
Gal-1	267 (±32)	2.83 (±0.5)
Gal-3	40 (±21)	1.0 (±0.2)
Gal-4	132 (±29)	9 (±1.3)
Gal-12	453 (±67)	6.3 (±0.9)

[a] Enzyme activities were assayed in reaction mixtures containing 0.1, 0.5, 1, 2.5, and 5 mM *p*NPGal in 50 mM NaOAc buffer pH 4.6 in the presence of 0, 2.5, 5, and 10 mM D-galactose or 0, 25, 50, 100 μM of γ-galactone-lactone. K_i values were calculated from a Lineweaver-Burk plot. Numbers represent mean ± SE for N=3.

Table IV. Natural substrate specificities

Substrate[a]	Activity[b]				
	Gal-1	**Gal-3**	**Gal-4**	**Gal-12**	**nGal-2**
Hot water soluble pectin	**1.7 (±0.8)**	**3.3 (±0.3)**	**tr[b]**	**2.3 (±0.14)**	**tr**
Oxalate soluble pectin	**1.9 (±0.3)**	**tr**	**1 (±0.15)**	**1.8 (±0.19)**	**1.14 (±0.2)**
Alkali soluble hemicellulose I	**3.4 (±0.5)**	**tr**	**tr**	**tr**	**0.393 (±0.05)**
Alkali soluble hemicellulose II	**1.7 (±0.7)**	**1.9 (±0.9)**	**tr**	**tr**	**tr**
Apple pectin	**13 (±2.2)**	**2.3 (±0.1)**	**4.9 (±0.8)**	**6.4 (±1.7)**	**4.5 (±0.9)**

[a]All polysaccharides were extracted from *Arabidopsis* cell wall except apple pectin, which was obtained from Sigma (St.Louis, MO).
[b]µg of D-galactose released from 2.5 mg of polysaccharide by 0.04 nkat/units enzymes at 25°C for 30 h,.
[c] tr - trace (poor hydrolysis)
Numbers represent mean ± SE for N=3.

similar to what was previously shown for Gal-2 and Gal-5 (Gantulga et al., 2008) . Gal-1 and Gal-3 also had weak but detectable activities with the alkali-soluble cell wall fractions. None of the enzymes appeared to release galactose from xyloglucan oligosaccharides, larchwood arabinogalactan, or oat xylan (data not shown). This observation indicates that the Subfamily III enzymes are unlikely to be involved in modification of these types of polysaccharides in *Arabidopsis*.

Four of the Subfamily III enzymes, Gal-1, Gal-2, Gal-4, and Gal-12, are predicted to have at least one N-glycosylation site. However, only two of the recombinant proteins, Gal-1 and Gal-12, appeared to be glycosylated when expressed in yeast, based both on the size of the recovered protein and the ability to bind to a ConA column (data not shown and Gantulga et al., 2008). The result for Gal-4 is consistent with previous findings that this protein is also not glycosylated when expressed in insect cells (Ahn et al., 2007). Gal-2, Gal-3, Gal-4, Gal-5 and Gal-12 migrated through acidic native gels, consistent the ability of these proteins to bind to cation exchange resin and the prediction that these proteins have a basic pI based on amino acid sequence. While Gal-1 was also predicted to have a basic pI, the recombinant Gal-1 protein did not migrate into the acidic native gels, suggesting that it forms large aggregates in solution.

Although both Gal-1 and Gal-3 contain a C-terminal lectin-like domain, we were not able to detect any lectin activity associated with these proteins. Neither protein bound to galactosyl- or mannosyl-agarose columns, nor did the proteins show hemagglutinating activity in a rabbit red blood-cell assay. The biochemical functions of these lectin-like domains therefore remain to be elucidated.

Taken together, these results indicate that the *Arabidopsis* Subfamily III β-galactosidases are exo-galactanases that primarily act on pectic polysaccharides of the cell wall. Therefore, these enzymes may play important roles in cell wall remodeling during plant growth and development.

2.5 Isolation and characterization of native Gal-2 protein

Although previous microarray (Hruba et al., 2005), proteomic (Jamet et al., 2006), and genetic (Ahn et al., 2007; Iglesias et al., 2006; Perez, 2004) studies, as well as the RT-PCR experiments described above, all indicate that the Subfamily III *β-galactosidase* genes are expressed in *Arabidopsis,* to the best of our knowledge the presence of functional β-galactosidase enzyme has not yet been verified. Although it is desirable to carry out detailed biochemical characterizations using native enzymes, the isolation and purification of enzymes from native sources can be challenging. We therefore focused on achieving this goal for a representative member of Subfamily III. A combination of affinity, hydrophobic interaction, and ion-exchange chromatography was used to isolate and biochemically characterize a native β-galactosidase protein from mature rosette leaves, the tissue that can be most easily collected in large quantities from *Arabidopsis*. Purification was monitored using a β-galactosidase activity assay, SDS-PAGE, and immunoblotting until a single protein band was obtained (Fig. 4). MALDI-TOF analysis after in-gel trypsin digestion gave a peptide mass fingerprint consistent with that of Gal-2 (Fig. 4C). We therefore designated the preparation as native Gal-2 (nGal-2). This is the first successful isolation of a native β-galactosidase protein from *Arabidopsis*.

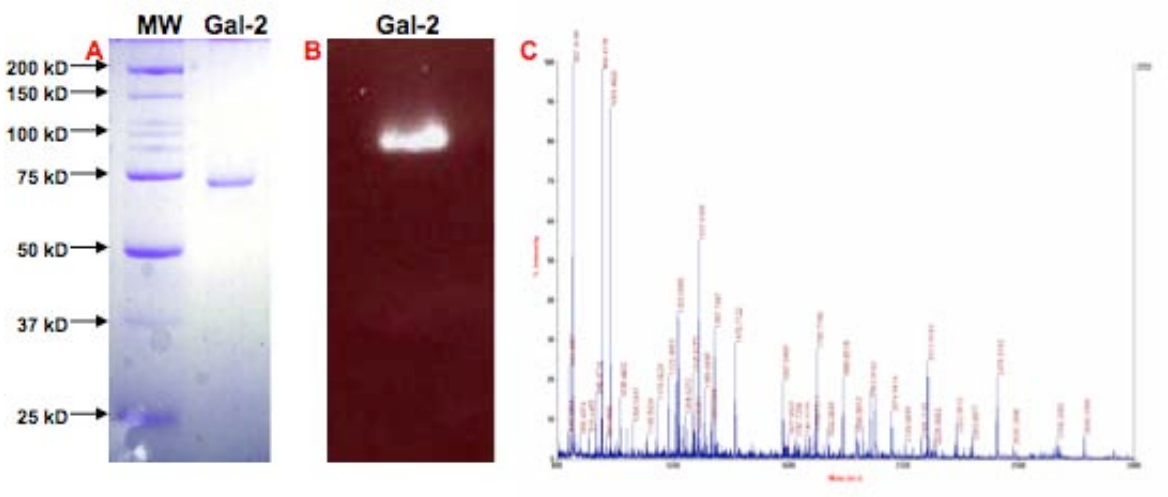


Fig. 4. SDS-PAGE (A), native PAGE (B), and MALDI-TOF spectrum (C) of purified native Gal-2. (A) 10% SDS-PAGE gel stained with Coomassie Blue R-250 (Lane 1, molecular weight standard; Lane 2, purified native Gal-2); (B) 8% acidic native gel

The biochemical properties and substrate specificities of nGal-2 were studied and compared with those of the recombinant enzyme. nGal-2 showed a monomeric molecular weight of ~75kD by SDS-PAGE. This is consistent with the size predicted for this protein, indicating that the nGal-2 polypeptide is intact and that it does not undergo proteolytic processing that has been reported for other plant β-galactosidases (Kotake et al., 2005; Ross et al., 1994; Triantafillidou and Georgatsos, 2001). The size of the protein, together with the fact that it did not bind to ConA (not shown), indicated that it was not glycosylated. This is also consistent with our observations that recombinant Gal-2 expressed in *P. pastoris* is also not glycosylated (Gantulga et al., 2008). nGal-2 has a basic pI that allows the enzyme to migrate into an acidic native gel, showed a strict sugar specificity toward galactose, with some activity towards fucose, and hydrolyzed the artificial substrates, *p*NPGal, *o*NPGal, X-Gal, and *p*NFuc (Table I, Table II, and Fig. 4). The pH optimum of the enzyme was ~4.0, which is typical for plant β-galactosidases (Buckeridge and Reid, 1994; Carey et al., 1995; Kotake et al., 2005; Li et al., 2001; Ross et al., 1994). Hydrolysis of *p*NPGal by nGal-2 showed typical Michaelis-Menten kinetics (Table II). The linkage specificity of nGal-2 was the same as that of the recombinant Gal-2, in that it preferred to cleave β-(1, 4) and β-(1, 3) linkages in galacto-oligosaccharides (data not shown). nGal-2 also hydrolyzed apple pectin and ammonium oxalate-soluble pectic fractions of *Arabidopsis* cell walls, releasing monomeric galactose, similar to the recombinant enzyme (Table IV).

2.6 Localization of β-galactosidase activity in whole tissues

To further explore the potential roles of the Subfamily III enzymes, an effort was made to examine the tissue-specific distribution and subcellular localization of the isozymes. As an initial approach, X-Gal was used as a substrate at pH4.6 to visualize β-galactosidase activity in 14-day-old *Arabidopsis* seedlings, as well as other tissues from 5-week-old plants. Control tissues incubated with X-Gal in the presence of galactone-lactone (a galactosidase inhibitor) or glucono-lactone (a glucosidase inhibitor) or with X-Glu (a glucosidase substrate) at pH 4.6 did not show any staining (not shown). This indicates that any observable staining resulted only from β-galactosidases with acidic pH optima. β-galactosidase activity was present primarily in vascular tissues of young leaves (Fig 5A) and flowers (Fig. 5B and C), as well as in the silique and in seeds (Fig. 5D and E). The intense staining observed in the anther and stigma (Fig. 5B and C) is consistent with a role for β-galactosidases in pollen development and germination. In root, β-galactosidase activity was restricted to the elongation zones, and no staining was detected in the root tip (not shown).

2.7 Organ specific expression of Gal-1 and Gal-12: Immunoblotting using anti-peptide antibodies

Isoform-specific antipeptide antibodies were prepared for the two enzymes with the highest (Gal-1) and lowest (Gal-12) levels of expression from microarray data among the members of Subfamily III. The antibody preparations were affinity-purified and then used to detect the relative abundance and size of the native Gal-1 and Gal-12 proteins in *Arabidopsis* root, petiole, rosette leaf, stem, cauline leaf, flower, and silique (Fig. 6). Gal-

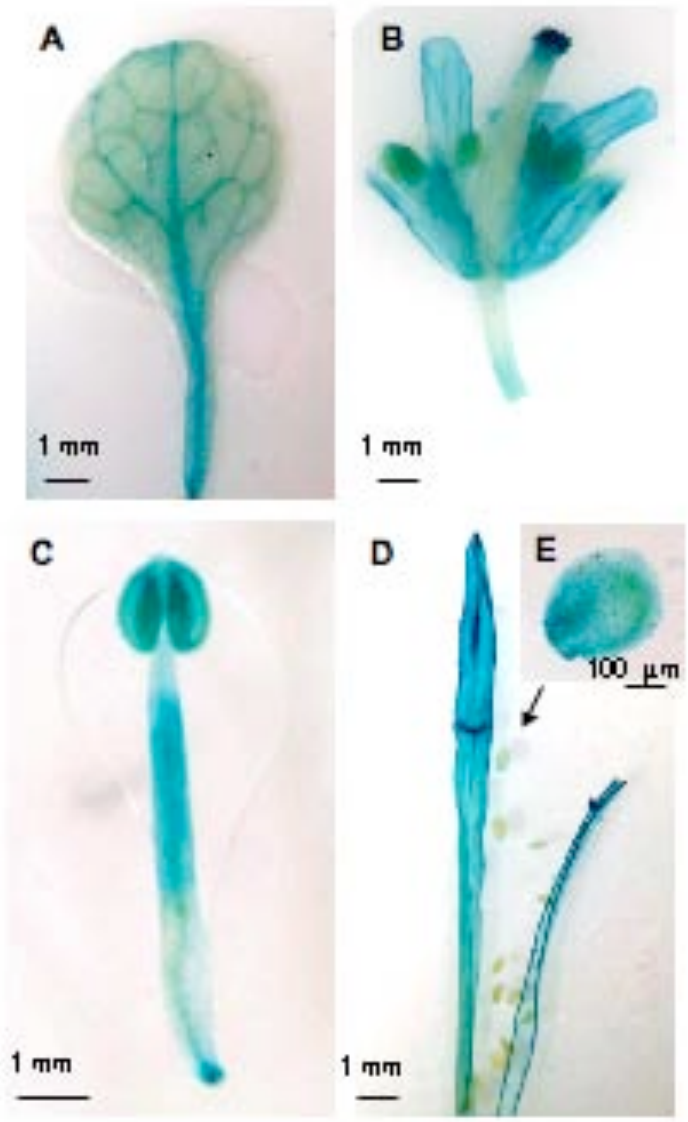


Fig. 5. Localization of β-galactosidase activity in *Arabidopsis* whole tissues. Tissues from four-week-old plants were stained with X-Gal. (A) cotyledon; (B) flower; (C) stamen; (D) silique; (E) seed. The arrow in D indicates the location of the seed shown in E.

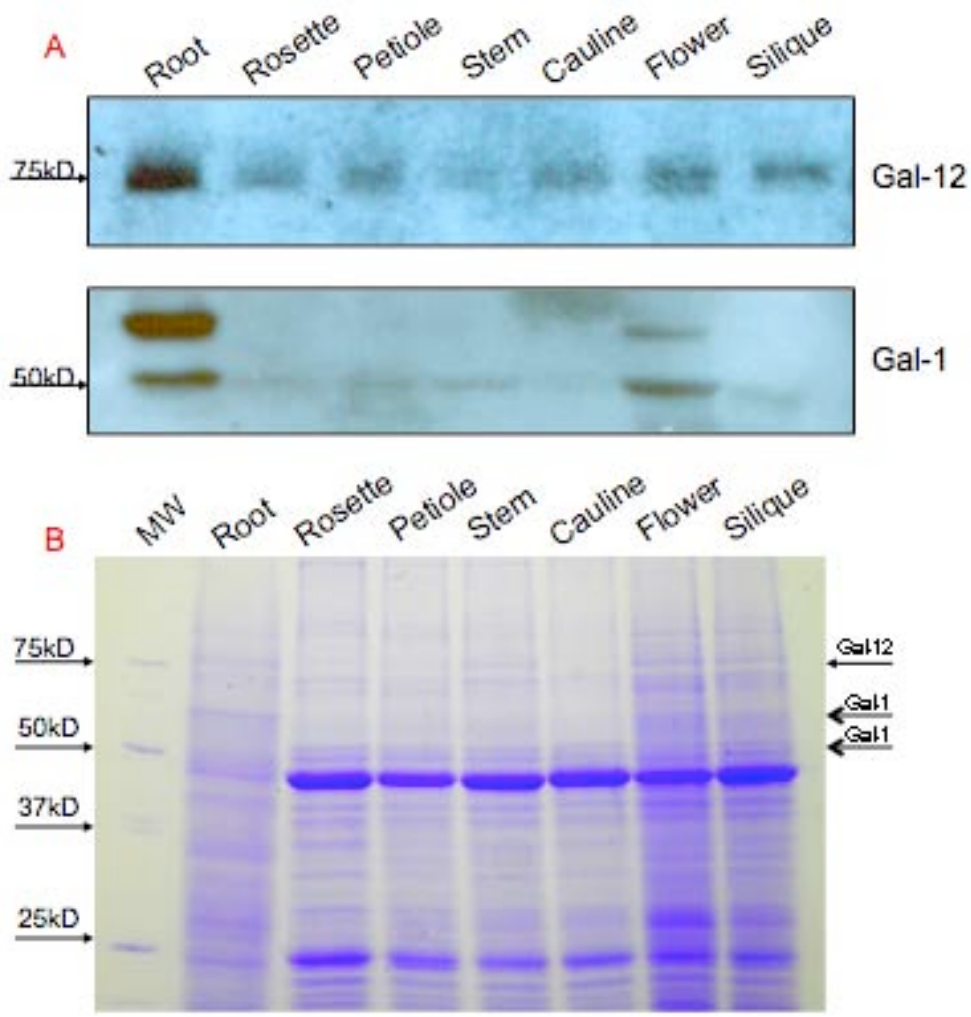


Fig. 6 Immunoblot analysis of organ-specific expression of Gal-12 and Gal-1.

Arabidopsis total protein extracts (10 μg) were separated on a 10% SDS–PAGE gel. The proteins were either transferred to PVDF membranes and incubated with affinity purified rabbit antibodies against Gal-12 and Gal-1 peptides (A) or stained with Coomassie blue R-250 (B). The arrows in panel B mark the approximate size and positions of the immunoreactive bands in panel A.

12 protein was detected as a single band of approximately 75 kD in all organs, consistent with the predicted molecular mass for this enzyme, indicating that native Gal-12 is not glycosylated, as was also observed for the protein produced in *P. pastoris,* and that it is not proteolytically processed. The relative abundance of Gal-12 protein was highest in the root, moderate in flowers, and low in all other organs. We observed discrepancies in the the relative abundance of the transcript and protein in some cases. This could be due to the differences in transcriptional and translational regulation, as proposed previously for Gal-5 (Gantulga et al., 2008).

The anti-peptide antibody for Gal-1 reacted with two bands of ~50 and ~60 kD in all tissues (Fig. 6). Unlike Gal-2 and Gal-5 (Gantulga et al., 2008) and the Gal-12 protein, Gal-1 protein appeared to be post-translationally processed into fragments much smaller than the predicted polypeptide size of 90 kD. It should be noted that proteolytic processing has also been observed for β-galactosidases in apple (Ross et al., 1994), rice (Chantarangsee et al., 2007; Kaneko and Kobayashi, 2003), barley (Triantafillidou and Georgatsos, 2001), and radish (Kotake et al., 2005; Sekimata et al., 1989). Gal-1 protein was most abundant in root and flower, which is consistent with the microarray and RT-PCR data (Figs. 2 and 3), but as with Gal-12 the mRNA and protein expression patterns are not exactly the same, pointing to possible posttranscriptional or posttranslational regulation.

The overall expression patterns of Gal-1 and Gal-12 proteins are similar to that of of Gal-2, with accumulation of the protein in most organs, while Gal-5 protein is found only in green tissues (Gantulga et al., 2008). Both Gal-1 and Gal-12 are most abundant in the root, suggesting possible cooperation between the enzymes in growth control in this

organ. Specific antibodies will have to be developed against Gal-3 and Gal-4 in order to extend this analysis to the entire subfamily.

2.8 Immunolocalization of Gal-1 and Gal-12

Although there is good evidence for a role for β-galactosidases in cell wall modification, the subcellular localization of the enzymes in plant has not yet been determined at high resolution. The anti-peptide antibodies developed for Gal-1 and Gal-12 were therefore used to localize these proteins by immuno-electron microscopy. According to the immunoblotting and total β-galactosidase activity-staining experiments (Figs. 5 and 6), Gal-1 and Gal-12 are likely to be most strongly expressed in the elongation zone of the mature root. The localization experiments were therefore performed using root tissue. A light microscopy image of a transverse section of an unlabelled root tissue is shown in Fig. 7A. Fig. 7B shows an enlargement of a xylem cell shown in Fig. 7A by transmission electron microscopy (TEM). The sections from the same tissue, with similar morphology as in Fig. 7B, were then labeled either with purified anti-Gal-12 or with anti-Gal-1 antibody, and then with a gold-conjugated secondary antibody.

Transmission electron microscopy images revealed that Gal-12 and Gal-1 proteins are localized in thickened walls of xylem cells. For Gal-12, the areas near the pit of xylem vessels were labeled more intensely, Fig. 7C. Less intense labeling of Gal-12 was observed in the cell walls of the epidermal cells and root hairs. In the case of Gal-1, only xylem cell wall was labeled; no gold particles were observed in any other areas of the root section. Unlike Gal-12, which often appeared as clusters of particles inside and around

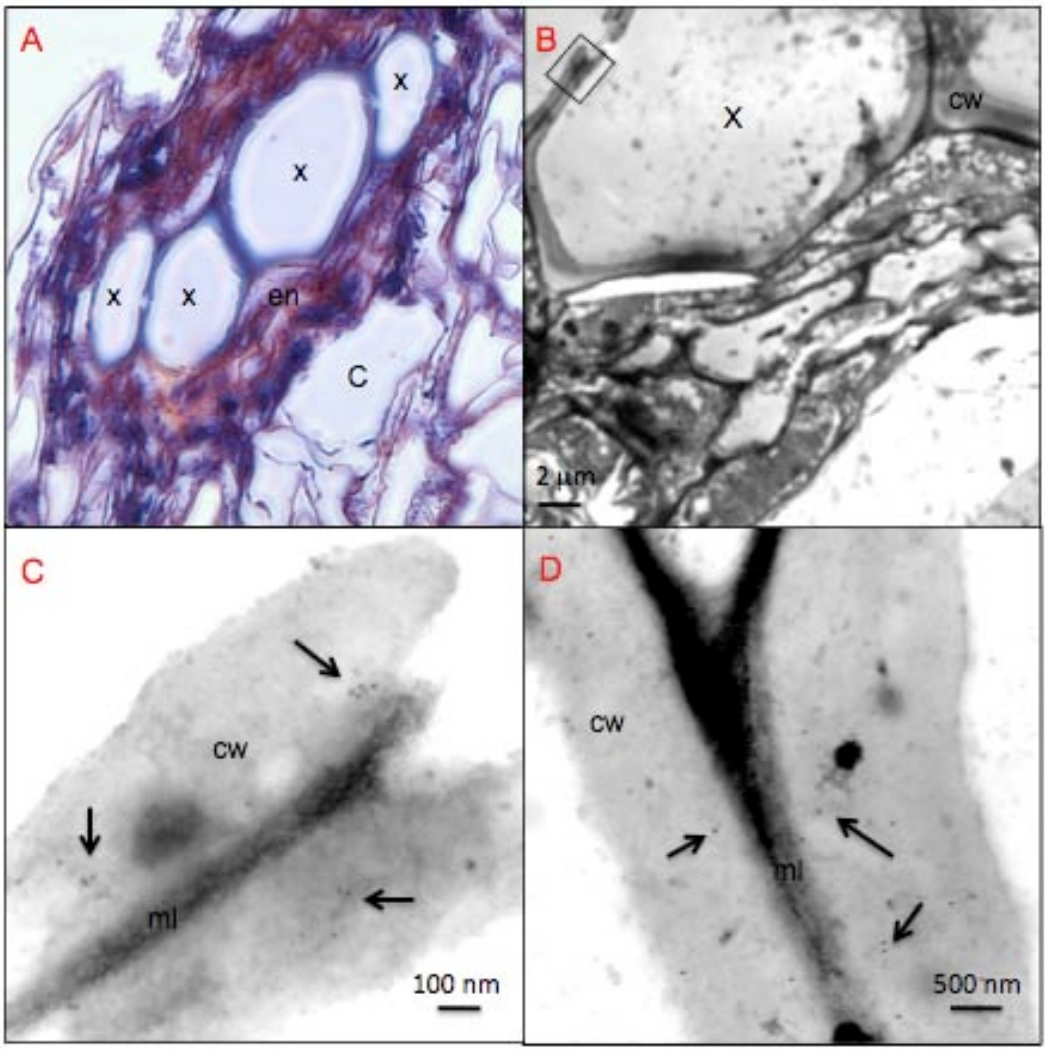


Fig. 7 Subcellular localization of Gal-1 and Gal-12. (A) Light microscopy image of an unlabeled toluidine-blue stained transverse section of the root. (B) TEM image of the xylem cell. (C) Gal-12 labeled xylem cell wall. (D) Gal-1 labeled xylem cell wall. "x"-xylem cell; "c"-cortex, "en"-endodermis; and "cw"-cell wall. Arrows indicate locations of the gold particles.

small, electron dense areas, Gal-1 labels appeared as single particles scattered and evenly distributed along the cell wall. Control sections incubated with a buffer in place of the primary antibodies showed no labeling, indicating the labeling of Gal-1 and Gal-12 in Fig. 7C-D was specific. Localization of β-galactosidase proteins in the thickened areas of the root cell wall provides strong evidence that the enzymes encoding the Subfamily III genes are involved in the degradation of cell wall polysaccharides.

2.9 Conclusions

Here we have described the characterization of four recombinant proteins expressed in *P. pastoris*, completing the systematic analysis of the Subfamily III β-galactosidases of *Arabidopsis*. A native protein from the subfamily was also purified and shown to have biochemical properties similar to those of the recombinant protein. All of the enzymes were found to have a strict galactose specificity and a broad aglycone specificity. The enzymes all cleave β-(1, 4) and β-(1, 3) linked galacto-oligosaccharides, while only Gal-12 hydrolyzed β-(1, 6) linked galacto-oligosaccharides. All of the enzymes also have activity toward pectic polysaccharides, releasing monomeric galactose as a product. This indicates that Subfamily III β-galactosidases are exo-galactanases that primarily act on galactan side chains of cell wall polysaccharides. It will be interesting to examine these differences at the structural level. Unfortunately, although structures have been solved for three GH35 enzymes, these are from microbial species that do not have sufficient similarity to the plant enzymes to allow for reliable homology modeling.

Spatial expression of six *Gal* genes was studied by RT-PCR, which revealed that the *Gal* genes are differentially regulated. Methyl jasmonate known to regulate growth and defense response in plants down-regulates expression of at least three out of six *Gal*

genes. Isoform-specific antibodies showed that levels of the Gal-1 and Gal-12 proteins were high in root, moderate in flower, and low in all other organs examined. While the Gal-12 polypeptide in plant extracts was similar to the predicted size, Gal-1 appears to undergo posttranslational cleavage to two smaller fragments. We also observed numerous discrepancies between mRNA and protein levels which suggests possible transcriptional and/or post-transcriptional regulation of *Gal* gene expression.

To the best of our knowledge, we show for the first time the cell wall localization of β-galactosidases by TEM. Mature root sections labeled with Gal-12 and Gal-1 antibodies showed specific labeling of cell wall of vascular and epidermal tissues. More detailed study is necessary to establish whether Gal-1 and Gal-12 labeling corresponds to regions where β-(1, 4)-galactan accumulates, suggested to be associated with elongation and differentiation of root cells (Martin et al., 2008; McCartney et al., 2003). It will also be interesting to study the localization of Gal-1 and Gal-12, as well as other Subfamily III Gals, in tissues of other organs. Our preliminary study revealed labeling in the cell wall of rosette leaf and flower for Gal-1 and Gal-2 (data not shown). It may be possible to derive more direct insights into the physiological functions of these enzymes through genetic analysis. However, this may require construction of double or triple mutants or RNAi lines because preliminary characterization of T-DNA insertion lines for Gal-1, Gal-3, and Gal-12 uncovered no obvious phenotypes (data not shown; Nickolas Carpita, Perdue University, personal communication)..

One of the central goals of the NSF Arabidopsis-2010 project is to characterize all putative genes in this model plant. As for seventeen β-galactosidases in GH Family 35, seven have been functionally characterized to date. In addition to the six enzymes that

belong to Subfamily III (this report and Ahn et al., 2007; Gantulga et al., 2008), Gal-6 from Subfamily V has been characterized by Dean et al. (2007). Clearly, much remains to be learned about the functions of the remaining members of the large family and whether the subfamilies reflect functional redundancy of the isozymes.

3. Experimental

3.1 Plant material

Arabidopsis Col-O seeds were obtained from the Arabidopsis Biological Resource Center (Columbus, OH). Plants were grown as described previously (Gantulga et al., 2008). For stress and hormone treatments, 14-day old seedlings on plates were sprayed directly. At different time points after treatment, seedlings were harvested, snap frozen in liquid nitrogen and kept at -80^{o}C until RNA isolation. Cell wall isolation and extraction of cell wall polysaccharides were carried out as described in Li et al (2001). For immunoblot analysis, *Arabidopsis* tissue was ground in 1X SDS Laemmli sample buffer containing 6M urea at a ratio of 1:4 (0.5 g tissue: 2 ml solvent).

3.2 Expression and purification of recombinant Gal-1, Gal-3, Gal-4, and Gal-12 in **P. pastoris**

The mature protein coding sequences of Gal-1, amino acids Ser33 to Trp815, Gal-3, residues Gly32 to Pro825, Gal-4, residues Ser28 to Thr697, and Gal-12, residues Ile28 to Lys701, were cloned into the pPICZα expression vector (Invitrogen, Carlsbad, CA) for expression fused to the C-terminus of the yeast α factor secretion signal to facilitate

secretion of recombinant proteins into the culture medium. Expression of recombinant Gal-4 and Gal-12 proteins was carried out following the manufacturer's protocol. A modified protocol was used for expression of Gal-1 and Gal-3, as follows. Initial biomass was obtained by growing 1L-1.5L culture for 2-3 days, then transferring the cells into 100 ml of induction medium to initiate expression of the recombinant proteins. Supernatant fractions from seven-day old cultures were used for purification of the enzymes.

Gal-1 and Gal-4 were purified from culture supernatants by ion exchange chromatography using Sulfoxyethyl (SE 53) cellulose (Whatman, UK). One hundred ml of culture supernatant was filtered, diluted five times with degassed dH_2O to reduce the ionic strength of the medium and loaded onto the column pre-equilibrated with buffer A (20 mM potassium phosphate, pH 6.0). After washing with 10 column volumes of buffer A, bound proteins were eluted in one step with 200 mM NaCl in buffer A. All fractions were checked for β-galactosidase activity using pNPGal as substrate. Fractions with the highest β-galactosidase activity were used for further experiments.

Gal-3 and Gal-12 proteins were purified using hydrophobic interaction chromatography and ion exchange chromatography. Solid ammonium sulfate was added to the culture supernatant to 30% final concentration, and filtered supernatant was loaded onto a ToyoPearl Butyl column (Tosoh Bioscience, Montgomeryville, PA) pre-equilibrated with 25 mM potassium phosphate buffer, pH 6.0 with 30% ammonium sulfate. After washing the column with 5 volumes of equilibration buffer, bound proteins were eluted with a linear 30-0% ammonium sulfate gradient in 25 mM potassium phosphate buffer, pH 6.0. Active fractions were pooled and desalted using a 30 kD cutoff centrifugal concentrator (Vivascience, UK) with three changes of 10 mM sodium acetate

buffer pH 5.0. The sample was then loaded onto a cation-exchange Accell-CM column (Waters, Milford, MA) pre-equilibrated with desalting buffer and eluted with a 0-300 mM NaCl gradient. Active fractions were pooled and concentrated using a centrifugal filter and stored on ice at 4^{o}C for further use.

3.3 Purification of native Gal-2 from **Arabidopsis** *rosette leaves*

Fresh rosette leaves (100 g) from 4-week-old *Arabidopsis* plants were homogenized in 500 ml extraction buffer (EB buffer, 10 mM potassium phosphate, pH 7.0, 10 mM EDTA, 1 mM $MgCl_2$, 5% glycerol, 1 mM PMSF). The resulting homogenate was incubated on ice for 2 h and centrifuged at 12000 x g for 20 min at 4^{o}C. The pellets were re-extracted with EB and centrifuged. The two supernatants were combined and filtered. The crude extract was loaded onto a Concanavalin-A (ConA) agarose column (Vector Laboratories, Burlingame, CA) equilibrated with ConA buffer (10 mM TRIS-HCl, pH 7.0, 1 mM $CaCl_2$, 500 mM NaCl). Pass-through fractions were collected, solid ammonium sulfate (AS) was added to 25% final concentration, and the solution was incubated at 4^{o}C for 1 h. Precipitated proteins were removed by centrifugation and the supernatant was loaded onto a ToyoPearl Butyl (TPB) column, equilibrated with TPB buffer (20 mM potassium phosphate, pH 6.2, 25% AS). After washing the column with 50 ml of TPB buffer (5 bed volumes), bound proteins were eluted with a linear gradient of AS from 25-0% in 20 mM potassium phosphate, pH 6.2. All fractions were checked for β-galactosidase activity using *p*NPGal as a substrate. Two peaks were detected. The fractions from the later-eluting peak were pooled and loaded onto the Sulfoxyethyl cellulose (SE) (Whatman, UK) equilibrated with SE buffer (20 mM potassium phosphate,

pH 6.2) after a 10-fold dilution with SE buffer. All β-galactosidase activity was bound to the column. Bound fractions were eluted with a linear 0-500 mM NaCl gradient in SE buffer. A single activity peak was detected using *p*NPGal as a substrate. Active fractions were pooled and concentrated on a small SE column. Purified protein was subjected to MALDI-TOF/MS after trypsin digestion. Peptide fragment masses were analyzed using MASCOT at http://www.matrixscience.com.

3.4 β-galactosidase activity assay

The standard protocol described in Gantulga et al. (2008) was used for all activity assays with *p*NPGal, unless stated otherwise. The hemagglutination assay for lectin activity was carried out as described in Kittur et al. (2007).

In order to determine natural substrate specificities, galacto-oligosaccharides (20 mM), xyloglucan olisaccharides (XGOs) (1 mg/ml), and polysaccharides (1% (w/v)) were prepared in water. Final concentrations were 4 mM for oligosaccharides, 0.5 mg/ml for XGOs, and 0.5% for polysaccharides in 100 mM acetate buffer pH 4.6. Reaction mixtures were incubated with 0.04 units/nkat of enzyme at room temperature for 30 h. Reaction mixtures containing no enzyme and no substrate were used as controls. Quantification of galactose produced after hydrolysis of polysaccharides by β-galactosidases was carried out as described in Gantulga et al. (2008). Hydrolysis of XGOs was analyzed by MALDI-TOF as described in Iglesias et al. (2006).

3.5 Localization of β-galactosidase activity in Arabidopsis whole tissues

Ten-day-old seedlings and tissues from 4-week-old plants were washed twice with dH_2O and incubated with staining solution (50 mM sodium acetate buffer pH 4.6, 10 mM EDTA, 0.1% sarkosyl, 20% methanol, 1 mM X-Gal (5-bromo-3-indolyl-β-D-galactopyranoside)) at 37°C for 2h. Tissues were then de-stained in methanol 3 times for 1h for each until the indigo staining was visible (Chantarangsee et al., 2007).

3.6 RT-PCR analysis

Total RNA was extracted from 100 mg of seedling and from mature plant tissues using Trizol reagent (Invitrogen, Carlsbad, CA) and treated with Turbo RNase free DNase (Ambion, Foster City, CA) to remove contaminating genomic DNA. cDNA was prepared from these RNA samples (2-5 μg) using the SuperScript II reverse transcriptase kit (Invitrogen, Carlsbad, CA) and oligo(dT) primers. The reverse transcription reaction was diluted 100-300 times with nuclease-free water and used for PCR analysis of gene expression using GoTaq Green mix (Promega, Madison, WI). *Ubiquitin-10* (At4g05320) expression was analyzed as a control.

3.7 Generation of peptide-specific antibodies

For generating isozyme specific antibodies, unique peptides CSQTSRMKMVRVPVHGGL and CSPDEKFLTGEDPLLTIG derived from Gal-1 and Gal-12 polypeptide sequences, respectively, were designed. Each peptide sequence was compared against the *Arabidopsis* genome sequence using tblastn (blast.ncbi.nlm.nih.gov) to confirm that these represented unique epitopes. Peptide synthesis and generation of antibodies were carried out using the Genscript antibody service (Piscataway, NJ). The

isozyme-specific antibodies were purified using antigen-affinity columns, prepared by coupling Gal-1 synthetic peptide or purified recombinant Gal-12 protein to AffiGel-10 (Bio-Rad, Hercules, CA) according to the manufacturer's protocol. The specificity of the antibodies was verified by probing dot blots containing recombinant protein for all six Subfamily III members under the conditions used for the immunoblot analysis of *Arabidopsis* tissues (data not shown).

3.8 SDS-PAGE, native PAGE and immunoblotting

SDS-PAGE was performed as described by Laemmli (1970). Immunoblots were carried out as described in Gantulga et al. (2008) except that detection was by chemiluminescence using the ELC-Plus kit (GE Healthcare, Piscataway, NJ) and CL–Exposure film (Thermo, Rockford, IL).

3.9 Immunolocalization

For immunolocalization studies, *Arabidopsis* root samples were prepared as described in Saslowsky et al. (2001). Thin sections were collected on Formvar-coated nickel grids. For TEM analysis, ultra-thin LR White sections were first incubated with 100 mM glycine pH 7.0 to block residual aldehyde groups. Then grids were transferred to 10% Triton-X and incubated for 15 min. Nonspecific binding sites were blocked with 10% (w/v) normal goat serum (Sigma, St. Louis, MO) in TBS for 1 h at room temperature. Grids were incubated at room temperature with affinity purified anti-Gal-1 antibody diluted 1:2 and anti-Gal-12 diluted 1:5 in TBS containing 10% (w/v) normal goat serum overnight at 4^{o}C. After three 10 min washes in TBST, grids were incubated with 10 nm

gold-conjugated goat anti-rabbit antibody (Sigma, St. Louis, MO) diluted 1:100 for Gal-12 and 1:50 for Gal-1 in 10% (w/v) normal goat serum in TBS for 1 h at room temperature. Grids were washed three times in TBST for 10 min and three times in water for 10 min. Sections were counterstained with 2% (w/v) uranyl acetate for 15 min and washed in distilled water. As a control, primary antibodies were replaced with buffer. Immunogold-labeled sections were examined with a JEOL (Tokyo, Japan) JEM 100CX-II transmission electron microscope. Thick sections of the same samples were collected on glass slides, stained with toluidine blue, mounted on permount, and examined by light microscopy.

Acknowledgements

We are grateful to Drs. Yoichi Tsumuraya and Dr. Toshihisa Kotake of Saitama University, Japan for providing galacto-oligosaccharides. The authors also wish to thank Drs. R.A Walker, D. Cimini and J. Sible for the light microscopy analysis and K. Lowe for the Tem analysis. This research is funded by the *Arabidopsis* 2010 Project of the National Science Foundation (MCB-0115937) to Esen et al., a research grant award from the Virginia Academy of Science to D. Gantulga and Department of Biological Sciences, VT.

REFERENCES

Ahn, Y. O., Zheng, M., Bevan, D. R., Esen, A., Shiu, S. H., Benson, J., Peng, H. P., Miller, J. T., Cheng, C. L., Poulton, J. E., Shih, M. C., 2007. Functional genomic analysis of *Arabidopsis thaliana* glycoside hydrolase family 35. Phytochemistry 68, 1510-1520.

Buckeridge, M. S., Hutcheon, I. S., Reid, J. S. G., 2005. The role of exo-(1 -> 4)-β-galactanase in the mobilization of polysaccharides from the cotyledon cell walls of *Lupinus angustifolius* following germination. Ann. Bot. 96, 435-444.

Buckeridge, M. S., Reid, J. S. G., 1994. Purification and properties of a novel beta-galactosidase or exo-(1-]4)-β-D-galactanase from the cotyledons of germinated *Lupinus-angustifolius* L. seeds. Planta 192, 502-511.

Callahan, J. W., 1999. Molecular basis of GM1 gangliosidosis and Morquio disease, type B. Structure-function studies of lysosomal beta-galactosidase and the non-lysosomal beta-galactosidase-like protein. Biochimica Et Biophysica Acta-Molecular Basis of Disease 1455, 85-103.

Carey, A. T., Holt, K., Picard, S., Wilde, R., Tucker, G. A., Bird, C. R., Schuch, W., Seymour, G. B., 1995. Tomato exo-(1-]4)-β-D-galactanase - isolation, changes during ripening in normal and mutant tomato fruit, and characterization of a related cDNA clone. Plant Physiology 108, 1099-1107.

Chantarangsee, M., Tanthanuch, W., Fujimura, T., Fry, S. C., Cairns, J. K., 2007. Molecular characterization of β-galactosidases from germinating rice (*Oryza sativa*). Plant Science 173, 118-134.

Coutinho, P. M., Henrissat, B., 1999. Carbohydrate active enzymes: integrated database approach In: H. J. Gilbert g. Davies, B. H. a. B. S. (Ed.), Recent advances in carbohydrtae bioengineering. Rolyal Society of Chemistry, pp. 2-12.

de Alcantara, P. H. N., Martim, L., Silva, C. O., Dietrich, S. M. C., Buckeridge, M. S., 2006. Purification of a β-galactosidase from cotyledons of *Hymenaea courbaril* L. (Leguminosae). Enzyme properties and biological function. Plant Physiol. Biochem. 44, 619-627.

Dean, G. H., Zheng, H., Tewari, J., Huang, J., Young, D. S., Hwang, Y. T., Western, T. L., Carpita, N. C., McCann, M. C., Mansfield, S. D., Haughn, G. W., 2007. The Arabidopsis *MUM2* gene encodes a β-galactosidase required for the production of seed coat mucilage with correct hydration properties. Plant Cell 19, 4007-4021.

Esteban, R., Dopico, B., Munoz, F. J., Romo, S., Martin, I., Labrador, E., 2003. Cloning of a *Cicer arietinum* β-galactosidase with pectin-degrading function. Plant Cell Physiol. 44, 718-725.

Gantulga, D., Turan, Y., Bevan, D. R., Esen, A., 2008. The *Arabidopsis* At1g45130 and At3g52840 genes encode beta-galactosidases with activity toward cell wall polysaccharides. Phytochemistry 69, 1661-1670.

Hruba, P., Honys, D., Twell, D., Capkova, V., Tupy, J., 2005. Expression of β-galactosidase and β-xylosidase genes during microspore and pollen development. Planta 220, 931-940.

Iglesias, N., Abelenda, J. A., Rodino, M., Sampedro, J., Revilla, G., Zarra, I., 2006. Apoplastic glycosidases active against xyloglucan oligosaccharides of *Arabidopsis thaliana*. Plant Cell Physiol. 47, 55-63.

Jamet, E., Canut, H., Boudart, G., Pont-Lezica, R. F., 2006. Cell wall proteins: a new insight through proteomics. Trends Plant Sci. 11, 33-39.

Kaneko, S., Kobayashi, H., 2003. Purification and characterization of extracellular β-galactosidase secreted by supension cultured rice (*Oryza sativa* L.) cells. Biosci Biotechnol Biochem 67, 627-630.

Kittur, F. S., Lalgondar, M., Yu, H. Y., Bevan, D. R., Esen, A., 2007. Maize β-glucosidase-aggregating factor is a polyspecific jacalin-related chimeric lectin, and its lectin domain is responsible for β-glucosidase aggregation. J Biol Chem 282, 7299-7311.

Kotake, T., Dina, S., Konishi, T., Kaneko, S., Igarashi, K., Samejima, M., Watanabe, Y., Kimura, K., Tsumuraya, Y., 2005. Molecular cloning of a β-galactosidase from radish that specifically hydrolyzes β-(1 -> 3)- and β-(1 -> 6)-galactosyl residues of arabinogalactan protein. Plant Physiology 138, 1563-1576.

Lazan, H., Ng, S. Y., Goh, L. Y., Ali, Z. M., 2004. Papaya β-galactosidase/galactanase isoforms in differential cell wall hydrolysis and fruit softening during ripening. Plant Physiol. Biochem. 42, 847-853.

Li, S. C., Han, J. W., Chen, K. C., Chen, C. S., 2001. Purification and characterization of isoforms of β-galactosidases in mung bean seedlings. Phytochemistry 57, 349-359.

Liu, C., Mehdy, M. C., 2007. A nonclassical arabinogalactan protein gene highly expressed in vascular tissues, AGP31, is transcriptionally repressed by methyl jasmonic acid in Arabidopsis. Plant Physiol 145, 863-874.

Martin, I., Jimenez, T., Esteban, R., Dopico, B., Labrador, E., 2008. Immunolocalization of a cell wall β-galactosidase reveals its developmentally regulated expression in *Cicer arietinum* and its relationship to vascular tissue. J. Plant Growth Regul. 27, 181-191.

McCartney, L., Steele-King, C. G., Jordan, E., Knox, J. P., 2003. Cell wall pectic (1->4)-β-d-galactan marks the acceleration of cell elongation in the Arabidopsis seedling root meristem. Plant J 33, 447-454.

McDougall, G. J., Fry, S. C., 1990. Xyloglucan oligosaccharides promote growth and activate cellulase: evidence for a role of cellulase in cell expansion. Plant Physiol 93, 1042-1048.

Nemhauser, J. L., Hong, F., Chory, J., 2006. Different plant hormones regulate similar processes through largely nonoverlapping transcriptional responses. Cell 126, 467-475.

O'Donoghue, E. M., Somerfield, S. D., Heyes, J. A., 2002. Organization of cell walls in *Sandersonia aurantiaca* floral tissue. Journal of Experimental Botany 53, 513-523.

Perez, A. I. B., 2004. Arabidopsis cell wall β-galactosidase gene family: expression, catalytic activities and biological function in galactose dynamics., Botany and Plant Pathology, vol. Ph.D. Purdue, west Lafayette, p. 141.

Ross, G. S., Redgwell, R. J., Macrae, E. A., 1993. Kiwifruit β-galactosidase - isolation and activity against specific fruit cell-wall polysaccharides. Planta 189, 499-506.

Ross, G. S., Wegrzyn, T., Macrae, E. A., Redgwell, R. J., 1994. Apple β-galactosidase - activity against cell-wall polysaccharides and characterization of a related cDNA clone. Plant Physiology 106, 521-528.

Saslowsky, D., Winkel-Shirley, B., 2001. Localization of flavonoid enzymes in Arabidopsis roots. Plant J. 27, 37-48.

Sekimata, M., Ogura, K., Tsumuraya, Y., Hashimoto, Y., Yamamoto, S., 1989. A β-balactosidase from radish (*Raphanus sativus* L.) seeds. Plant Physiol 90, 567-574.

Smith, D. L., Gross, K. C., 2000. A family of at least seven β-galactosidase genes is expressed during tomato fruit development. Plant Physiology 123, 1173-1183.

Staswick, P. E., Su, W. P., Howell, S. H., 1992. Methyl jasmonate inhibition of root-growth and induction of a leaf protein are decreased in an *Arabidopsis thaliana* mutant. Proc. Natl. Acad. Sci. U. S. A. 89, 6837-6840.

Tanthanuch, W., Chantarangsee, M., Maneesan, J., Ketudat-Cairns, J., 2008. Genomic and expression analysis of glycosyl hydrolase family 35 genes from rice (*Oryza sativa* L.). BMC Plant Biol 8, 84.

Tateishi, A., Inoue, H., Shiba, H., Yamaki, S., 2001. Molecular cloning of β-galactosidase from Japanese pear (*Pyrus pyrifolia*) and its gene expression with fruit ripening. Plant Cell Physiol 42, 492-498.

Trainotti, L., Spinello, R., Piovan, A., Spolaore, S., Casadoro, G., 2001. β-Galactosidases with a lectin-like domain are expressed in strawberry. J Exp Bot 52, 1635-1645.

Triantafillidou, D., Georgatsos, J. G., 2001. Barley β-galactosidase: Structure, function, heterogeneity, and gene origin. Journal of Protein Chemistry 20, 551-562.

Zablackis, E., Huang, J., Muller, B., Darvill, A. G., Albersheim, P., 1995. Structure of plant-cell walls .34. Characterization of the cell-wall polysaccharides of *Arabidopsis thaliana* leaves. Plant Physiology 107, 1129-1138.

Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L., Gruissem, W., 2004. GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. Plant Physiol 136, 2621-2632.

Chapter 4

Conclusion

Conclusion

The *Arabidopsis* β-galactosidase GH35 Subfamily III, consisting of six members, was systematically characterized with respect to enzyme activity, substrate specificity, organ- and tissue-specific expression constituting a first complete characterization of the entire subfamily from a single species (Ahn et al., 2007; Gantulga et al., 2008a; Gantulga et al., 2008b) . Based on identification of natural substrates and subcellular localization, we conclude that these β-galactosidases are most likely involved in cell wall modification during plant growth and development.

All six members of Subfamily III were expressed in *Pichia pastoris* as recombinant proteins and characterized with a special emphasis on substrate specificities. A native β-galactosidase purified from *Arabidopsis* rosette leaves displayed biochemical characteristics similar to those of the recombinant enzyme, suggesting that the properties of the recombinant proteins accurately reflect those of the native counterparts. The natural substrate specificities of these six isoforms suggest that Subfamily III β-galactosidases are exo-galactanases that primarily act on galactan side chains of cell wall polysaccharides.

Spatial expression of the six *Gal* genes was studied by analysis of publicly available microarray data and by RT-PCR experiments, revealing different expression patterns in different organs. Methyl jasmonate, which is known to regulate plant growth and defense, was found to down regulate the expression of at least three *Gal* genes. Peptide-specific antibodies were generated against four isoforms, Gal-1, Gal-2, Gal-5, and Gal-12, and purified using antigen-affinity columns. These antibodies were used to

study the expression of these four isoforms at the protein level by immunoblotting. The organ-specific expression patterns were generally consistent with microarray and RT-PCR data, although some discrepancies were observed between relative abundances of the transcripts and proteins. This suggests that there are distinct mechanisms of regulation for transcription and translation. The sizes of the Gal-2, Gal-5 and Gal-12 polypeptides in plant extracts were similar to the predicted sizes for these enzymes, whereas Gal-1 appeared to undergo a posttranslational cleavage to two smaller fragments, similar to what has been observed for β-galactosidases in apple (Ross et al., 1994), radish (Kotake et al., 2005), and barley (Triantafillidou and Georgatsos, 2001).

To explore the localization of the β-galactosidase proteins, enzyme activity was first visualized *in vivo* using the substrate, X-Gal. β-galactosidase activity was present mostly in vascular tissues of young leaves, flowers, and roots, as well as in seed. Cell wall localization of Gal-1, Gal-2, Gal-5, and Gal-12 was shown by immuno-dot blotting and western blotting of apoplastic and cell wall bound protein extracts. Subcellular localization of β-galactosidase proteins was further studied by immunofluorescence and immuno-electron microscopy in organs where β-galactosidase activity was abundant based on the staining experiment. In particular, immunogold labeled particles for Gal-1 and Gal-12 proteins were detected in the cell wall of vascular and epidermal tissues in mature root. To the best of our knowledge, this is the first time that the cell wall localization of β-galactosidases has been shown at high resolution by transmission electron microscopy. Localization of the β-galactosidase proteins in the cell wall and the identified natural substrate suggests that these proteins have a potential to modify cell wall structure.

Determination of functions and characteristics of the Subfamily III genes is a valuable contribution to the NSF *Arabidopsis* 2010 project, whose central goal is to determine the functions of all 24,000 predicted genes in this important model plant. The systematic approach taken in this work, from biochemical characterization to subcellular localization through expression patterns, can be successfully applied in determining the functions of the remaining *Arabidopsis* β-galactosidases as well as many other genes. The results presented in this dissertation can be helpful in further studies aimed at understanding the metabolic and gene regulatory networks in which β-galactosidases participate to modify cell wall structures.

Finally, as for further research on Subfamily III β-galactosidases, interesting goals are to increase catalytic efficiency and alter substrate specificity of the enzymes by site-directed mutagenesis. Mutant genes encoding highly efficient and controllable β-galactosidases can be inserted into the genome of different plants for specific modification of cell wall structures. Such transgenic plants may have improved agronomic qualities, enhanced defenses against herbivores, drought tolerance, and long term post-harvest shelf life.

References

Ahn, Y. O., Zheng, M., Bevan, D. R., Esen, A., Shiu, S. H., Benson, J., Peng, H. P., Miller, J. T., Cheng, C. L., Poulton, J. E., Shih, M. C., 2007. Functional genomic analysis of *Arabidopsis thaliana* glycoside hydrolase family 35. Phytochemistry 68, 1510-1520.

Gantulga, D., Turan, Y., Bevan, D. R., Esen, A., 2008a. The *Arabidopsis* At1g45130 and At3g52840 genes encode beta-galactosidases with activity toward cell wall polysaccharides. Phytochemistry 69, 1661-1670.

Gantulga, D., Winkel, B., Bevan, D., Esen, A., 2008b. Comparative characterization of the *Arabidopsis* Subfamily III β-galactosidases. in preparation.

Kotake, T., Dina, S., Konishi, T., Kaneko, S., Igarashi, K., Samejima, M., Watanabe, Y., Kimura, K., Tsumuraya, Y., 2005. Molecular cloning of a β-galactosidase from radish that specifically hydrolyzes β-(1 -> 3)- and β-(1 -> 6)-galactosyl residues of arabinogalactan protein. Plant Physiol. 138, 1563-1576.

Ross, G. S., Wegrzyn, T., Macrae, E. A., Redgwell, R. J., 1994. Apple β-galactosidase - activity against cell-wall polysaccharides and characterization of a related cDNA clone. Plant Physiol. 106, 521-528.

Triantafillidou, D., Georgatsos, J. G., 2001. Barley β-galactosidase: Structure, function, heterogeneity, and gene origin. J. Protein Chem. 20, 551-562.