Beta galactosidases in Arabidopsis and tomato – a mini review

Balakumaran Chandrasekar† and Renier A.L. van der Hoorn*†1

†The Plant Chemetics Laboratory, Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, U.K.
‡The Plant Chemetics Laboratory, Max Planck Institute for Plant Breeding Research, Carl-von-Linne Weg 10, 50829 Cologne, Germany

Abstract
Beta galactosidases (BGALs) are glycosyl hydrolases that remove terminal β-D-galactosyl residues from β-D-galactosides. There are 17 predicted BGAL genes in the genomes of both Arabidopsis (BGAL1–17) and tomato (TBG1–17). All tested BGALs have BGAL activity but their distinct expression profiles and ancient phylogenetic separation indicates that these enzymes fulfil diverse, non-redundant roles in plant biology. The majority of these BGALs are predicted to have signal peptide and thought to act during cell wall-related biological processes. Interestingly, deletion of BGAL6 and BGAL10 in Arabidopsis causes reduced mucilage release during seed imbibition and shorter siliques respectively, whereas TBG4 depletion by RNAi decreases fruit softening in tomato. The majority of plant BGALs remain to be characterized.

Introduction
β-D-Galactose is found in many organisms and can be coupled to carbohydrates or non-carbohydrates via an O-glycosidic bond. Beta galactosidases (BGALs, EC 3.2.1.23) are glycosyl hydrolases (GHs) that remove the terminal β-D-galactosyl residues from the non-reducing end of these β-D-galactosides. BGALs performing this hydrolytic activity are found in GH families GH1, GH2, GH3, GH35, GH42, GH50 and GH59 [1]. BGALs from these GH families play major roles in different organisms. For instance, a GH35 enzyme GLB1 in humans is involved in removing terminal galactose residues from gangliosides in the lysosome [2]. Deficiency of GLB1 in humans causes Gangliosidosis due to the accumulation of toxic gangliosides. GALC is a GH59 BGAL which removes galactose from galactocerebrosides and its deficiency causes Krabbe disease in humans [3]. Finally, the frequently used bacterial LacZ gene encodes a GH2 family BGAL in Escherichia coli that is essential for lactose metabolism during glucose starvation [4,5].

Microbial BGALs are applied in dairy industry for the hydrolysis of lactose. These BGALs are known for their thermostability or activity at low temperatures. BGALs having optimal hydrolytic activity at low temperatures (e.g. 0 °C) have been identified in psychrophilic microbes like Arthrobacter sp., yeast, Pseudoalteromonas sp. and Paracoccus sp. [6–9]. These cold-adapted BGALs have applications in the food industry to remove lactose contaminations from heat-sensitive milk products. By contrast, BGALs having optimal hydrolytic activity at higher temperatures (e.g. 70 °C) have been identified in microbes like Bacillus stearothermophilus [10,11] and are used in industry for producing lactose-free milk [12]. Thermostable BGALs are also used in dairy industry to remove lactose contaminations from whey, a major by-product from cheese [13–16]. In addition to hydrolytic activities, some microbial BGALs also have transgalactosylating activity. Transgalactosylation is the process where BGAL transfers the released galactose to another carbohydrate instead of water. For example, microbial BGALs from different GH families have been used to synthesize β-galacto-oligosaccharides (GOS), an important human prebiotic diet. These BGALs transfer the hydrolysed galactose residue to acceptor lactose to build GOS [17–19].

Notably, all plant BGALs belong to family GH3. Typically, they follow the Koshland retaining mechanism, releasing galactose in their retained, β-anomeric conformation [20]. GH35 enzymes belong to clan GH-A in the CAZy database and fold as a (α/β)8 TIM barrel domain with the two catalytic glutamate residues [21]. One catalytic Glu residue acts as the proton donor and the other as a nucleophile during catalysis. In plants, β-D-linked galactosyl residues are found in glycolipids (e.g. monogalactosyldiacylglycerol, MGDG [22]), proteoglycans (e.g. arabinogalactan proteins [23]) and cell wall polysaccharides (e.g. xyloglucans and Rhamnogalacturonan I, RGI [24]). A biologically relevant substrate for BGALs during fruit ripening of tomato is galactan, a polymer of β-(1-4) D-galactose attached to RGI [25–27]. In this mini review, we discuss the BGALs of Arabidopsis thaliana and Solanum lycopersicum (tomato). We will discuss the phylogeny, domain architecture and expression patterns and summarize the biochemical and physiological functions.

Arabidopsis has 17 BGALs
The genome of A. thaliana contains 17 genes encoding putative BGALs, designated as BGAL1–17 (Figure 1A).
At2g04060 is not included because although it shares similarities to known BGAL sequences at its C-terminus, this protein does not have a GH35 domain and is probably a truncated duplicate of BGAL15 [28]. Phylogenetic analysis of the 17 BGAL Arabidopsis proteins has divided these proteins into seven different groups: group I (BGAL17), II (BGAL8-9), III (BGAL1–5, BGAL12), IV (BGAL10), V (BGAL7, BGAL13), VI (BGAL11, BGAL13, BGAL14) and VII (BGAL6, BGAL16) [29]. A second classification based on phylogenetic analysis of BGALs from various plant species, has divided the 17 BGALs into eight sub families: subfamily a1 (BGAL1–4, BGAL5, BGAL12), a2 (BGAL9), a4 (BGAL8), a5 (BGAL10), b (BGAL7, BGAL13), c1 (BGAL11, BGAL13, BGAL14), c2 (BGAL6, BGAL16) and d (BGAL17) [28].

**Arabidopsis BGAL proteins carry additional domains**

Of the 17 Arabidopsis BGAL proteins, 13 are predicted to have an N-terminal signal peptide that targets the protein to the endomembrane system. The four other BGAL proteins possibly locate in the cytoplasm or nucleus. The GH35 domain contains two active site glutamate residues. The active site consensus sequence G-G-P-[LIVM]-[x]-Q-[x]-E-[FY] is common to all GH35 BGALs, and contains the Glu residue (bold) that acts as a proton donor during hydrolysis. The motif P-N-K-x-x-K-P-KM-W-T-E-x-W is present in all BGALs except BGAL17, and carries the Glu residue (bold) that acts as a nucleophile [28]. Apart from the GH35 domain, ten BGALs also carry an additional gal_lectin domain in the C-terminus (Figure 1B). BGAL14 has an additional PRP1 domain C-terminal to the gal_lectin domain. BGAL11 and BGAL16 carry an additional BetaGal4.5 domain between GH35 and gal_lectin domain whereas BGAL13 carries a GH2N domain between GH35 and gal_lectin domain. The functional significance of these extra domains is yet unclear. It has been suggested that the gal_lectin domain might contribute to substrate specificity of BGAL [28,30].

**All tested Arabidopsis BGALs were detected in the cell wall**

Immunogold labelling followed by TEM of root sections has revealed that BGAL1 and BGAL12 reside in the thickened cell walls of xylem cells [30]. Cell wall localization of BGAL6 has been observed using a fusion with GFP, transiently expressed in *Nicotiana tabacum* leaves [31]. BGAL2 and BGAL5 have been detected in cell wall fractions of Arabidopsis leaves by dotblotting [33], whereas BGAL8 has been detected in the cell walls of Arabidopsis stems by proteomics [34]. In conclusion, all six characterized BGALs in Arabidopsis to date are localized in the cell wall, implicating their role in cell wall remodelling and expansion (Figure 1D).

**BGALs are differentially expressed in Arabidopsis organs**

Gene expression analysis using eFP browser [35] indicates that BGAL genes have distinct organ-specific expression patterns (Figure 1E). BGAL7, BGAL11 and BGAL13 are expressed mostly in flowers whereas BGAL17 is mostly expressed in the stem. Other BGALs are expressed in multiple organs, but still follow different expression patterns. BGAL12, BGAL14 and BGAL16 are poorly expressed in the selected tissues, consistent with RT-PCR analysis of these genes [28].

**Physiological roles of BGAL6 and BGAL10**

The physiological roles of only two Arabidopsis BGALs have been characterized. Mucilage mutant-2 (*mum2*) fails to extrude mucilage from the apoplast upon hydration and is caused by the *bgal6* mutant allele (Figure 2A), indicating that BGAL6 alters the hydration properties of mucilage by modifying carbohydrate structures [31]. By contrast, BGAL10 seems to be the only or main BGAL acting on xylglucan cell wall substrates because unusual xylglucan residues were observed in cell walls of *bgal10* mutant flowers [32]. This unusual xylglucan accumulation correlates with a reduced silique and sepal length of *bgal10* mutant plants (Figure 2B). Characterization of the physiological functions of the remaining Arabidopsis BGALs is an unexplored area in plant biology.
Figure 1 | BGALs of A. thaliana

(A) Protein name and accession numbers. (B) Domains in BGAL proteins. Represented domains are not on scale. (C) Biochemically characterized BGALs. (D) Experimentally validated subcellular protein localization. (E) Expression pattern of BGALs in various tissues of Arabidopsis. The expression data of Arabidopsis BGALs for these tissues were extracted using the eFP browser [35]. Expression data were converted into heat maps using R. Expression data for BGAL15 were not available (N).
Figure 2 | Phenotypes of bgal mutants of A. thaliana

(A) Failure to extrude the mucilage in bgal6 (mum2) mutant seeds. The wild-type and T-DNA insertion mutant seeds of bgal6 were stained with Ruthenium Red to visualize the mucilage layer in the seeds. These pictures are reproduced with permission from [31]. 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. and 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. (B) bgal10 T-DNA mutants have a reduced siliqua and sepal length. The pictures are reproduced with permission from [32]: Sampedro, J., Gianzo, C., Iglesias, N., Guitián, E., Revilla, G. and Zarra, I. (2012) AtBGAL10 is the main xyloglucan β-galactosidase in Arabidopsis, and its absence results in unusual xyloglucan subunits and growth defects. Plant Physiol. 158, 1146–1157.

Tomato also has 17 BGALs that are common to angiosperms

Also the tomato genome contains 17 genes encoding putative BGALs. All these proteins contain the GH35 domain with typical consensus sequences and both active site residues. Two additional proteins (Solyc07g038120 and Solyc07g038130), share some similarity with the GH35 domain but lack the active site consensus sequences. Seven genes are expressed during various stages of tomato fruit development [25]. These seven genes have been named TBG1–7 (tomato beta galactosidase). We named the remaining 10 tomato BGAL genes as TBG8–17, in chronological order of their accession number (Figure 3).

Phylogenetic analysis using protein sequences of both Arabidopsis and tomato revealed that tomato BGALs fall into the same seven groups as Arabidopsis BGALs (Figure 4). This indicates that BGAL diversification occurred early in plant evolution and that orthologues in Arabidopsis and tomato may have similar, distinct functions. To extend the evolutionary analysis of the BGAL family, we have included the 15 BGALs of the monocot Oryza sativa (Rice) [36] and six BGALs of the moss Physcomitrella patens [37] in our phylogenetic analysis (Figure 4). The grouping of the BGALs in the phylogenetic tree indicates that BGALs of groups I and VI existed since the evolution of land plants since they are also present in moss. By contrast, BGALs of groups II, III, IV and VII may have evolved later, but probably before the angiosperms evolved because they are present in both eudicot and monocot plant species. This includes groups III, IV and VII, of which individual BGALs were shown to act in fruit development (see below) and flower development [32] and seed mucilage release [31], consistent with the absence of these genes in moss. BGALs in groups II and III contain more than one orthologue, suggesting their distinct functions, whereas BGALs in group VI tend to have duplicated and diversified within each of the four plant species. The presence
of additional domains seems less conserved within plants. The C-terminal lectin domain is absent from group I and present in groups II, IV, V and VI, but irregularly present in groups III and VII. The GH2N domain is occasionally found in groups I, III, IV, VI and VIIIs and the remaining additional domains are only found in groups VI and VII. Thus, these data indicate that plant BGALs have ancient origins in the plant kingdom, though the additional domains are not consistently present.

Three biochemically characterized BGALs of tomato

Biochemical characterization has been performed for three tomato BGALs (Figure 3B). All these characterized TBGs are expressed during tomato fruit development. TBG4 is the first enzyme to be biochemically characterized. Earlier in 1980s, this enzyme (BGAL II) was found to be abundant in ripe tomato fruits and was purified from the tomato fruit extracts and shown to hydrolyse galactose residues from cell wall polysaccharides and artificial substrates [38,39]. In another study, the gene encoding TBG4 was cloned and the enzyme was expressed in yeast and purified. The purified TBG4 had β(1-4) galactosidase/exogalactanase activity, meaning it can hydrolyse galactose from galactan, lactose and synthetic substrates [40]. A similar strategy has been used to characterize TBG1 and TBG5, which both have β(1-4) galactosidase and β(1-4) exogalactanase activity [41,42]. The other 14 TBGs remain to be biochemically characterized.

Physiological role of tomato BGALs

The role of BGALs in tomato fruit ripening has been well studied. Fruit ripening is a complicated physiological process involving alterations in fruit texture and cell wall degradation. A key biochemical event during ripening is the loss of galactosyl residues from the cell wall fractions (mainly pectins) and the accumulation of soluble free galactose residues [43–46]. Galactose, when injected into tomato fruits, causes enhanced ethylene production and promotes early ripening [47]. Hence, galactose released from the cell walls...
during ripening might have the same effect as enhancing ethylene production.

Down-regulation of TBG4 transcript levels by antisense TBG4 resulted in transgenic tomato lines with reduced exogalactanase activity and low levels of galactose [40]. These TBG4-silenced plants also produced fruits with a 40% increased fruit firmness compared with the controls [40]. By contrast, down-regulating 90% of TBG1 transcript levels had no effect on exogalactanase activity or galactose levels and did not affect the firmness of the fruit [41].

Notably, antisense suppression of the TBG6 gene resulted in a tomato fruit with high BGAL activity at day 20 after pollination [48]. An interesting, unexpected observation is that although the mRNA levels of TBG6 were significantly down-regulated, the antisense lines had higher total BGAL activity than the wild-type plants. However, at 30 days after pollination or 3 days after the breaker stage, the total BGAL activity was comparable to that of the wild type. Unexpectedly, fruits from TBG6-silenced plants had reduced galactosyl residues in cell walls and enhanced fruit softening. In addition, TBG6 gene suppression also had some notable external and internal fruit morphological phenotypes. The fruits from these transgenic lines were elongated and had vertical ‘zipper like scars’ along their epidermis (Figure 5). Furthermore, the internal locular space was decreased or absent from these fruits.

Although the antisense approach had helped in these studies to understand the involvement of TBGs in tomato fruit development process, possible effects due to off-target effects cannot be neglected. Hence independent genetic knockouts by genome editing or complementation of the transgenic line with a synthetic TBG gene that is insensitive for silencing may be needed to confirm the role of TBGs during fruit development. To our knowledge, the exact
functions of other TBGs which are expressed during fruit development and elsewhere during development still remains to be elucidated.

Roles of BGALs in other plant species

The physiological roles of BGALs have also been studied in other plant species. BGAL activities are important during ripening of fruits like apple, mango, strawberry, banana and bell pepper [49–53]. BGALs acting on galactans are also involved in formation of secondary cell walls in flax fibres [54]. By contrast, a BGAL from radish seeds acts on β(1-3) and β(1-6) linked galactose residues on arabinoxylans (AGPs) [55,56]. The functional significance of BGALs in degrading this natural substrate is still unknown. In another study, up-regulation of a BGAL has been observed during abscission of mature orange fruits [57], suggesting that BGAL activity might play an important role during this abscission process. Hence it is evident from these studies that, similar to Arabidopsis and tomato, the majority of BGALs from other plant species also find their significance in cell wall associated biological processes.

Exciting directions for future BGAL research

Research into the biological and biochemical roles of plant BGALs has only just begun. Several major questions remain to be addressed. For example, what is the functional significance of additional domains in BGALs? One speculation is that the gal_lectin domain might assist in determining BGAL substrate specificity. Previously, the gal_lectin domain from a rice BGAL (LOC_Os03g06940) was shown to agglutinate or clump erythrocytes, which is a characteristic property of many lectins [58]. Hence this domain can function independently as lectin and might have a functional role in carbohydrate recognition.

Second, when and where are BGALs active? Being a large multigene family, characterizing their biochemical functions might be challenging because it requires the purification of each BGAL. In addition, TBG6 was found to be a difficult enzyme to overexpress and purify. Activity-based protein profiling (ABPP) might be helpful to solve these issues. ABPP involves chemical probes that label the active site residue of proteins in an activity-dependent manner. We have recently introduced cyclophellitol aziridine-based probes to monitor broad range of active glycosidases in plants [59]. To monitor the in vivo BGAL activity, Arabidopsis plants have also been treated with the artificial substrate X-Gal [60]. Performing assays with more natural substrates and locating the activity in tissues using microscopy would be relevant to associate BGAL activity with a biological process. For example, xyloglucan endotransglycosylase activity has been monitored in situ by infiltrating natural fluorogenic substrates for that enzyme [61]. Furthermore, cell-permeable fluorescent activity-based probes for glycosidases can be used to locate active enzymes in situ [62].

Third, what are the natural substrates of BGALs and what happens with the released products? Apart from galactan, β-D-galactosyl residues are also found in the glycolipids and many glycosylated proteins. Hence it would be interesting to determine if these are substrates for BGALs during various biological processes. There are different possible fates for the released galactose residues after hydrolysis. Galactose might be used: i) as energy source; ii) to build new glycoconjugates or iii) to initiate signalling cascades. In plants, sugars like glucose and sucrose can function as signalling molecules apart from being an energy source. Glucose plays a major role in the induction of senescence process and organ development [63, 64]. Likewise, sucrose signalling is important in regulating fructan and anthocyanin biosynthesis [65]. Hence experimental validation of galactose to behave as a signalling molecule can be interesting topic to investigate.

Funding

This work was supported by ERA-IB project "PRODuCE", the University of Oxford, and ERC Consolidator grant 'GreenProteases'.

References

21 Rojas, A.L., Nagem, R.A.P., Neustroev, K.N., Arand, M., Adamska, M.,
20 Zhang, S., McCarter, J.D., Okamura-Oho, Y., Yaghi, F., Hinek, A., Withers,
14 Panesar, P.S., Kumari, S., Panesar, R., Panesar, P.S., Kumari, S. and
13 Szczodrak, J. (2000) Hydrolysis of lactose in whey permeate by
7 Nakagawa, T., Ikehata, R., Uchino, M., Miyaji, T., Takano, K. and
35 Winter, D., Vinegar, B., Nahal, Y., Ammar, R., Wilson, G.V. and Provart,
20 Zhang, S., McCarter, J.D., Okamura-Oho, Y., Yaghi, F., Hinek, A., Withers,

Received 20 November 2015 doi:10.1042/BST20150217

©2016 Authors; published by Portland Press Limited