Broad-range Glycosidase Activity Profiling

Balakumaran Chandrasekar‡§, Thomas Colby¶, Asif Emran Khan Emon§, Jianbing Jiang‖, Tram Ngoc Hong‡§, Joji Grace Villamor§, Anne Harzen‖, Herman S. Overkleeft∥∥, and Renier A. L. van der Hoorn‡**

Plants produce hundreds of glycosidases. Despite their importance in cell wall (re)modeling, protein and lipid modification, and metabolite conversion, very little is known of this large class of glycolytic enzymes, partly because of their post-translational regulation and their elusive substrates. Here, we applied activity-based glycosidase profiling using cell-permeable small molecular probes that react covalently with the active site nucleophile of retaining glycosidases in an activity-dependent manner. Using mass spectrometry we detected the active state of dozens of myrosinases, glucosidases, xylosidases, and galactosidases representing seven different retaining glycosidase families. The method is simple and applicable for different organs and different plant species, in living cells and in subproteomes. We display the active state of previously uncharacterized glycosidases, one of which was encoded by a previously declared pseudogene. Interestingly, glycosidase activity profiling also revealed the active state of a diverse range of putative xylosidases, galactosidases, glucanases, and heparanase in the cell wall of Nicotiana benthamiana. Our data illustrate that this powerful approach displays a new and important layer of functional proteomic information on the active state of glycosidases. Molecular & Cellular Proteomics 13: 10.1074/mcp.O114.041616, 2787–2800, 2014.

Carbohydrates are present in all kingdoms of life and are particularly prominent in plants (1). Plants produce carbohydrates as one of their major constituents through their photosynthetic activity. The simplest synthesized forms of carbohydrates are monosaccharide sugars such as glucose, which provides energy for various cellular activities. Carbohydrates also exist in very complex forms. Monosaccharide sugars are attached to one another through covalent glycosidic linkage, which generates di-, oligo-, and polysaccharides. Carbohydrates also attach to non-carbohydrate species (lipids, proteins, hormones) through a glycosidic linkage to form glyco-conjugates (2).

Glycosidic bonds are hydrolyzed by a group of enzymes termed glycosyl hydrolases (GHs)1 or glycosidases (3). Because of the tremendous carbohydrate diversity, there are a vast variety of glycosidases, including glucosidases, xylosidases, and galactosidases, that preferentially hydrolyze their respective glycoside substrates. In general, the number of glycosidase-related genes in plants (for instance, Arabidopsis) is relatively high when compared with that in other sequenced organisms (for instance, human) (4). This signifies the unique importance of glycosidases in plants as opposed to other organisms. Based on protein sequence similarities, glycosidases are classified into different GH families. Members of the same GH family share a common mechanism of glycosidic bond cleavage (5).

Mechanistically, glycosidases are classified as retaining or inverting enzymes (6). To hydrolyze the glycosidic bond, both retaining and inverting enzymes carry two catalytic glutamate or aspartate residues (or both) (7). Of these two catalytic residues, one acts as a proton donor and the other as a nucleophile/base. The distance between these catalytic residues in the active site of the glycosidases determines the mechanism of hydrolysis. Retaining enzymes have two catalytic residues separated by a distance of ~5.5 Å, and their hydrolysis mechanism retains the net anomeric configuration of the C1 atom in the sugar molecule. In contrast, inverting enzymes have catalytic residues that are ~10 Å apart, and these enzymes invert the overall anomeric configuration of the C1 carbon atom in the released sugar (8).

Both retaining and inverting glycosidases are present abundantly in plants. The genome of Arabidopsis thaliana encodes for 400 glycosidases, of which 260 are retaining enzymes and 140 are inverting enzymes. Genetic, molecular, and biochemical approaches revealed that glycosidases are localized in...
Glycosidase Activity Profiling

different cellular compartments and are important for various biological processes. The majority of plant glycosidases reside in the cell wall, and these enzymes can play major roles in cell wall restructuring (9). Other characterized glycosidases reside in other compartments to regulate glycosylation of proteins and hormones. Despite the importance of GH enzymes, physiological and biochemical functions are assigned to only a few glycosidases (9).

Activity-based protein profiling (ABPP) is a powerful tool for monitoring the active state of multiple enzymes without knowledge of their natural substrates (10, 11). ABPP involves chemical probes that react with active site residues in an activity-dependent manner. Thus ABPP displays the availability and reactivity of active site residues in proteins, which are hallmarks for enzyme activity (12). ABPP is particularly attractive because the profiling can be done without purifying the enzymes and can be performed in cell extracts or in living cells. Another key advantage of ABPP is that the activities of large multigene enzyme families can be monitored using broad-range probes. ABPP has had a significant impact on plant science. After the introduction of probes for papain-like cysteine proteases (13, 14), these probes revealed increased protease activities in the tomato and maize apoplasts during immune responses (15, 16) and that these immune proteases are targeted by unrelated inhibitors secreted by fungi, oomycetes, and nematodes (17–24). Likewise, probes for the proteasome displayed unexpected increased proteasome activity during immune responses (25) and revealed that the bacterial effector molecule syringolin A targets the nuclear proteasome (26). We anticipate that more regulatory mechanisms will be discovered through the use of probes introduced for serine hydrolases, metalloproteases, vacuolar proteases, ATP binding proteins, and glutathione transferases (27–32).

Cyclophellitol-azidine-based probes were previously used in animal proteomes to target retaining glucosidas (33). Here we established and applied glycosidase profiling in plants. We discovered that cyclophellitol-azidine-based probes targeted an unexpectedly broad range of glycosidases representing members of at least seven different GH families. We used these probes to study the active state of glycosidases present in living cells, in different organs and plant species, and in the apoplast of Nicotiana benthamiana.

**EXPERIMENTAL PROCEDURES**

**Probe Synthesis**—The synthesis of JJB70 and JJB111 is described in the supplemental material. The synthesis of KY371 and JJB111 has been described previously (33). Aliquots of these compounds are available upon request.

**Biological Materials**—A. thaliana plants (ecotype Columbia, Col-0) were grown at 24 °C (day)/20 °C (night) in a growth chamber under a 12-h light regime. Leaves from rosettes of 4-week-old Arabidopsis plants were used for protein extraction. Cell cultures of A. thaliana (ecotype Landsberg erecta) were grown and weekly subcultured in L9 liquid medium (4.41 g/l MS-Basal medium including nitsch vitamins (Duchefa M0256), 30 g/l sucrose, 5.4 µM α-naphthyleacetic acid, 0.23 µM kinetin, adjusted to pH 5.6). N. benthamiana plants were grown at 22 °C and 60% relative humidity under a 12-h light regime. 4-week-old N. benthamiana plants were used for apoplastic fluid isolation. Commercially available xylosidase was purchased from PROZOMIX (Nothumberland, UK).

**Sample Preparation for Labeling Assays**—

**Extracts for Small-scale Labeling**—Leaves of various plant species or different organs of Arabidopsis plants were homogenized with 500 µl of 50 mM MES buffer at pH 6.0 or 50 mM MOPS buffer at pH 7.5. For the pH course experiment, the leaves were homogenized in sterile water. Two leaf discs (1.5-cm diameter) were taken from the leaves of different plant species and homogenized with 500 µl of 50 mM MOPS buffer at pH 7.5. After the tissues had been ground in a 1.5-ml tube, the samples were centrifuged at 10,000g and 4 °C for 10 min to remove cell debris, and the supernatant containing the soluble proteins was used for labeling.

**Extracts for Large-scale Labeling**—2 g of Arabidopsis leaf rosettes were frozen with liquid nitrogen and ground with a mortar and pestle. 5 ml of extraction buffer (50 mM MES buffer at pH 6.0, 0.1% Triton X-100, and 10% glycerol) were added to the leaf powder, vortexed, and centrifuged at 4000 rpm and 4 °C for 30 min to remove cell debris. The supernatant containing the soluble proteins was used for large-scale labeling.

**Apoplastic Fluid Isolation**—The apoplastic fluids were isolated by means of vacuum infiltration and centrifugation (34, 35). In brief, the leaves of N. benthamiana were submerged in ice-cold sterile water, and vacuum was applied. The water infiltrated the intercellular spaces of leaves as the vacuum was slowly released. The leaves were then dried briefly and centrifuged at 3000 rpm and 4 °C for 15 min. Protease inhibitor mixture solution (Complete Protease Inhibitor Tablets, Roche Applied Sciences) was added to the isolated apoplastic fluids at 4X final concentration.

**Labeling Plant Extracts**—Leaf or organ extracts of Arabidopsis containing 1.0 to 1.5 mg/ml total soluble proteins were pre-incubated with 50 µM KY371 or DMSO for 30 min at pH 6.0 or 7.5. These extracts were incubated with 2 µM JJB70 or JJB111 for 1h at room temperature in the dark at 50 °C total reaction volume. Equal volumes of DMSO were added for no probe control. For the pH course experiment, Arabidopsis leaf extracts were incubated with suitable pH buffer (50 mM Sodium acetate buffer for pH 4.0, 50 mM MES buffer for pH 5.0–6.5, 50 mM MOPS buffer for pH 7.0–7.5, Tris buffer for pH 8–10) and 2 µM JJB70 in the dark for 1h at room temperature at 50 °C total reaction volume.

**Peptide-N-Glycosidase F (PNGaseF) Treatment of Labeled Proteins**—10 µl of JJB70 or DMSO incubated leaf extracts of Arabidopsis were treated with 1.5 µl of 10X glycoprotein denaturing buffer (New England BioLabs) and heated at 100 °C for 10 min. The denatured proteins were treated with 3 µl of H2O, 2 µl of 10% Nonidet P-40, 2 µl of 10X G7 buffer (New England BioLabs), 1.5 µl PNGaseF (New England BioLabs) or H2O and incubated at 37 °C for 1h.

**Labeling Apoplastic Fluids**—Apoplastic fluids of N. benthamiana were pre-incubated with 50 µM KY371 or DMSO for 30 min at pH 5.0 with 50 mM MES buffer. These apoplastic fluids were incubated with 2 µM JJB70 or DMSO in the dark for 1h at room temperature for labeling at 50 µl total reaction volume. For pH course experiment, apoplastic fluids were incubated with suitable pH buffer (50 mM Sodium acetate buffer for pH 4.0, 50 mM MES buffer for pH 5.0–6.5, 50 mM MOPS buffer for pH 7.0–7.5, Tris buffer for pH 8–10) and 2 µM JJB70 in the dark for 1h at room temperature at 50 µl total reaction volume.

**Analysis of Labeled Proteins**—The labeling reactions were stopped by adding gel loading buffer containing β-mercaptoethanol at 1X final concentration and heating at 95 °C for 10 min, unless indicated.
otherwise. The labeled proteins were separated on 12% protein gels at 200 volts for 1 h. The JJB70-labeled proteins were detected in the protein gels with the Typhoon FLA 9000 scanner (GE Healthcare Life Sciences) using excitation wavelength at 472 nm and BBP1 filter (GE Healthcare Life Sciences). The fluorescence of the labeled proteins was quantified using ImageQuant (GE Healthcare Life Sciences). JJB111-labeled proteomes were transferred from the protein gels onto PVDF membranes, incubated with streptavidin HRP (ultra sensitive, Sigma) and detected using chemiluminescent substrates (SuperSignal West Chemiluminescent substrates, Thermo scientific).

In Vivo Labeling of Arabidopsis Cell Cultures—500 μl of an Arabidopsis cell cultures were transferred to a 12-well tissue culture plate. The medium was removed from the cell culture, replaced by 500 μl fresh L9 medium and kept shaking gently for 30 min before labeling. The cell cultures were preincubated with 50 μl of KY371 or DMSO for 30 min under gentle shaking. The cell culture was labeled by the solution was incubated with 100 μl of the JJB70 solution for 30 min. The precipitated proteins were dissolved in 2 ml of 1.2% SDS/PBS buffer. The final wash was with 10 ml of water. The labeled proteins were transferred from the protein gels with the Typhoon FLA 9000 scanner (GE Healthcare Life Sciences). Peptides were separated on 12% polyacrylamide gels at 200 V for 1 h, and the protein gels were stained between 400 and 1600 Da using a Top20 method with active exclusion for 30 s in a window from 0.2 Da below to 1.5 Da above the precursor mass. The resulting RAW files were de-convoluted and converted to MGF format using MS Convert from the Proteowizard 2.1 package, accepting only the six strongest peaks per 100-Da window.

Database Searching—The N. benthamiana database (v. 0.4.4, 76,379 sequences) was downloaded from the SOL genomics network. The Arabidopsis database was based on TAIR10 (35,384 sequences). 1095 common artifact sequences were added to both databases, and both were then reversed and concatenated to provide decoys for false discovery rate calculation. MS/MS spectra were searched against the described databases using MASCOT 2.4, permitting JJB111 (735,386 Da) as an additional variable modification of Asp and Glu residues in the searches. For both searches, a Mascot score of 38 represented a 95% certainty cutoff.

Other Mascot parameters were as follows: precursor mass tolerance, 0.4 Da; fragment mass tolerance, 0.5 Da; one C13 permitted; one tryptic miss-cleavage allowed; fixed modifications of cysteine (57.02 Da, carbamidomethyl); variable modifications of methionine (15.99 Da, oxidation).

Bioinformatics—

Glycosyl Hydrolase Family Classification—GH family classification of identified glycosidases in the Arabidopsis leaf proteome was performed based on the annotations done for the genome of A. thaliana in the CAZY database (38). GH family classification of identified glycosidases in the secreted leaf proteome of N. benthamiana was done by submitting protein sequences to the Pfam database (39).

Phylogenetic Tree Construction—Protein sequences of 258 retaining glycosidases in the genome of Arabidopsis were retrieved from the TAIR database (40). The protein sequences of AT4G33850 and A. thaliana IG002N01.5 were not available in the TAIR database, so they were retrieved from the National Center for Biotechnology Information (NCBI) (www.ncbi.nlm.nih.gov/). The sequences were aligned with ClustalX2 (41). The pairwise alignment gap opening penalty 0.75 and gap extension penalty 0.75 were used, whereas for multiple alignment gap opening penalty 15 and gap extension penalty 0.3 were used. The alignment files were saved in phylip format, and R script was implemented to construct the tree (42, 43). A neighbor-joining algorithm was used for tree construction from the calculated distance matrix in the program.

A. thaliana Leaf Transcriptome/Proteome Analysis—The transcript levels (average of three replicate experiments) of all 260 retaining glycosidases from the vegetative rosette (age 1.09) were retrieved from the AtGenExpress database (44). Spectral counts (Qty spectra) of all 260 retaining glycosidases detected in the juvenile leaves of Arabidopsis were retrieved from the pep2pro database (45).
three activity-based probes that carry an aziridine analog of cyclophellitol as their warhead. JJB70 carries a bodipy as its reporter tag and is slightly different from the previously described ABP3 (33) in that it has the triazole linker in reverse orientation (Fig. 1A). JJB111 (previously called ABP4) and KY371 carry biotin and alkyne minitags, respectively (Fig. 1A) (33). The electrophilic acyl-aziridine group in the probes reacts with the nucleophilic active site glutamate/aspartate residue of retaining glycosidases, resulting in a reversible covalent ester bond (Fig. 1B).

To test plant extracts for labeling, Arabidopsis leaf extracts were incubated with and without 2 $\mu$M JJB70 or JJB111. After 1 h of labeling, proteins were separated on protein gels, and labeled proteins were detected via fluorescent scanning (JJB70) or protein blot analysis using streptavidin-linked horseradish peroxidase (for JJB111). Two major signals were detected at 68 kDa (gray circle) and 75 kDa (black circle) (Fig. 1C). These signals were not detected in the no-probe control and were suppressed in the samples preincubated with KY371 (Fig. 1C). These data indicate that JJB70 and JJB111 target similar proteins in Arabidopsis leaf extracts and that labeling with these probes is specific. Further characterization of labeling parameters revealed that a one-hour labeling time is sufficient to reach maximum labeling, that labeling is saturated at 1 $\mu$M JJB70, and that labeling is optimal at slightly acidic pH (supplemental Fig. S1). These experiments show that labeling strongly depends on labeling conditions.

**Active Site Labeling of Myrosinases TGG1 and TGG2**—To identify the labeled proteins, large-scale labeling was performed from JJB111-labeled leaf extracts. The labeled proteins were enriched on streptavidin beads, separated on a protein gel, stained with SYPRO Ruby, and detected via fluorescent scanning. The detected protein bands were excised and treated with trypsin, and the peptides were identified via ion-trap mass spectrometry. Similar to small-scale labeling, two strong signals at 68 and 75 kDa were enriched, and these signals were absent in the no-probe control (Fig. 2A). MS analysis showed that the majority of the peptide spectra from the 68-kDa signal originated from the thioglucoside glucohydrolase TGG2, whereas the 75-kDa signal contained peptides mostly from TGG1 (Fig. 2B). TGG1 and TGG2 are myrosinases that mediate the conversion of glucosinolates during herbivore attack (46). Myrosinases have one catalytic glutamate residue; the second catalytic residue has been replaced by glutamine residue (47).

To identify the labeling site, the expected modification by JJB111 (735.386 Da) on potentially reactive residues (E and D)
was included in the searches. One labeled peptide of TGG1 was identified (black in Fig. 2B). Further analysis of the MS2 fragmentation spectrum revealed that b-ions up to T419 and y-ions down to E420 of this peptide were unmodified. In contrast, b-ions beyond E420 or y-ions starting at T419 or before carried the modification (Fig. 2C). These data demonstrate that the modification was at either E420 or T419. This is consistent with the fact that E420 has been annotated as the nucleophilic active site residue (48). Notably, in addition to the labeled active site peptide, the spectra of the unmodified active site peptide were found among the peptide spectra. This can be explained by the fact that the ester bond between the probe and the active site (Fig. 1B) can hydrolyze in water.

This is the first time that a modification of the active site with cyclophellitol-aziridine has been identified through MS.

**TGG1 and TGG2 Cause the Major Signals in Arabidopsis Leaf Extracts**—The TGG1 and TGG2 signals migrate at a higher apparent molecular weight (MW) than calculated from the plain amino acid sequence of the mature enzyme (59 and 60.4 kDa, respectively). The increased apparent MW could be due to N-glycosylation, as TGG1 and TGG2 carry nine and four putative N-glycosylation sites, respectively (Fig. 2E). To examine whether TGG1 and TGG2 are N-glycosylated, we incubated the labeled leaf proteins with and without PNGaseF, separated on protein gel and detected by in-gel fluorescence scanning. Labeling of mutant Arabidopsis lines confirms that signals are caused by TGG1 and TGG2. Leaf proteins of wild type (WT) and tgg1-3, tgg2-1 and tgg1-3/tgg2-1 mutant plants were extracted with and without 500 mM NaCl and labeled with 2 μM JJB70. Mixtures of these extracts (mix) were used as no-probe-control.
E), demonstrating that TGG1 and TGG2 are N-glycosylated. The presence of N-glycosylation is also consistent with the absence of tryptic peptides carrying putative glycosylation sites from the MS data (Fig. 2B).

To confirm that the major signals were caused by myrosinases TGG1 and TGG2, leaf extracts of wild-type, tgg1-3, tgg2-1, and tgg1-3/tgg2-1 knockout plants (46) were incubated with and without JJB70. Surprisingly, both the major signals were absent in the tgg1-3 and tgg1-3/tgg2-1 mutants, and only the 68-kDa signal was absent in the tgg2-1 mutant (Fig. 2F). It has previously been noted that the solubility of TGG2 depends on the presence of TGG1, unless the proteins are extracted in the presence of 500 mM NaCl (48). Indeed, the 68-kDa signal was recovered by extraction in 500 mM NaCl only in the tgg1-3 mutant and remained absent in tgg2-1 and tgg1-3/tgg2-1 mutants. A third signal at 60 kDa was detected in all plants upon extraction in 500 mM NaCl. Taken together, these data demonstrate that TGG1 and TGG2 cause the two major signals in Arabidopsis leaf extracts labeled with JJB70 and JJB111.

**JJB Probes Label a Second Layer of Glycosidases Representing New GH Families**—Further examination of the labeling profiles of the tgg1-3/tgg2-1 double mutant revealed another 13 weak signals at longer exposure times (Fig. 3A). These signals were also present in wild-type plants but were covered by the strong TGG1 and TGG2 signals (Fig. 3A). Importantly, pre-incubation with KY371 blocked the labeling of all these proteins in both wild-type and tgg1-3/tgg2-1 plants (Fig. 3B). These data demonstrate that JJB70 labeled proteins in addition to myrosinases TGG1 and TGG2.

To identify these additional labeled proteins, we extended the MS analyses of JJB111-labeled proteins by analyzing the weaker signals. This analysis revealed an additional 18 proteins that were annotated as glycosidases in the CAZY database.
The genome of *A. thaliana* encodes for 260 retaining glycosidases subdivided into 24 GH families. Mapping the identified glycosidases onto a genome-wide phylogenetic tree of *Arabidopsis* retaining glycosidases showed that the detected glycosidases were not clustered in the tree and represented members from families GH1, GH3, GH35, and GH79 (Fig. 4A). This broad range of GH families extends the target list of these probes beyond the previously described GH1, GH3, and GH30 families (33). Of the 20 identified glycosidases, only eight enzymes had been previously characterized for their biochemical and molecular functions. This included three enzymes from the GH1 family (TGG1, TGG2, and BGLU44), three enzymes from the GH3 family (BXL1, BXL4, and BXL6), one enzyme from the GH35 family (BGAL6), and one enzyme from the GH79 family (GUS2) (48, 50–55). These data show that the probes displays a broad range of active glycosidases.

**Selectivity of Glycosidase Labeling**—To investigate possible preferential labeling of certain glycosidases (e.g. TGG1 and TGG2) in *Arabidopsis* leaves, we compared our data with the abundance of transcript and peptide spectra present in the AtGenExpress database and the Arabidopsis proteome. The abundance of transcript and peptide spectra present in organs of *Arabidopsis*, we labeled extracts from seeds, seedlings, leaves, senescing leaves, and flowers with JJB70. Profiles were very different between organs (Fig. 5A). Notably, many signals were detected in flower extracts. Labeling of these proteins was blocked upon pre-incubation with KY371, and most signals remained in the *tgg1-3/tgg2-1* double mutant extracted with and without NaCl (Fig. 5B).

To extend glycosidase profiling to different plant species, leaf extracts of other dicot and monocot plants were incubated with and without JJB70. Labeling was observed in all tested plant extracts with different profiles and different intensities (Fig. 5C). Labeling of these proteins was suppressed by pre-incubation with KY371. Taken together, these data demonstrate that glycosidase activity profiling is broadly applicable in plant science.

To test whether JJB70 can label proteins in living cells, *Arabidopsis* cell cultures were pre-incubated with and without KY371 and labeled with JJB70 in duplicate. Proteins were extracted by grinding the cell cultures with SDS-containing gel loading buffer to stop the labeling reaction. JJB70 labeling caused two major signals at 56 and 72 kDa, and these signals were suppressed by KY371 pre-incubation (Fig. 5D). Protein
extraction in the presence of JJB70 also caused labeling (Control 1), but this ex vivo labeling was completely abolished if the extraction was performed in SDS-containing gel loading buffer (Control 2; Fig. 5D). This shows that JJB70 does not label proteins under denaturing conditions and demonstrates that the probe can label proteins in living cells. Interestingly, the labeling profile of the cell culture medium was distinct from that of the cells, as it displayed two faint signals at 55 kDa that were suppressed upon pre-incubation with KY371, indicating that the probe also displayed activities of secreted glycosidases.

**Glycosidase Profiling of Secreted Proteomes**—To investigate the secreted glycosidases further, we identified active glycosidases in the extracellular space (apoplast) of *N. benthamiana*.
thamiana, an important model plant having large leaves that are ideal for apoplast extraction. The apoplast of plants contains many glycosidases involved in cell wall remodeling and defense (57, 58). To monitor the active state of apoplastic glycosidases, we isolated apoplastic proteomes from N. benthamiana leaves and incubated them with and without 2 μM JJB70 for 1 h. The labeled proteins were separated on protein gel and detected via fluorescent scanning. Three major signals at 45, 55, and 72 kDa; a weak signal at 43 kDa; and a series of weaker signals below 40 kDa were detected in the probe-labeled sample (Fig. 6A). Similar but much weaker signals were detected in the total extracts (Fig. 6A). These signals were absent in the no-probe control and suppressed in the samples pre-incubated with KY371. Furthermore, labeling was optimal at acidic pH (Fig. 6B), consistent with the pH of the apoplast.

To identify the labeled proteins, we purified JJB111-labeled proteins, separated them on protein gels, and detected them through SYPRO Ruby staining. Labeled proteins were enriched in the JJB111-labeled sample, and these signals were not detected in the no-probe control (Fig. 6C). Seven protein bands from the JJB111-labeled sample and the corresponding regions from no-probe control were excised and treated with trypsin. MS analysis of peptides identified 19 different glycosidases from six different GH families: GH1, GH3, GH5, GH35, GH51, and GH79 (supplemental Table S2). Members of the GH5 and GH51 families had not been detected before in labeling experiments, which signifies the potential for JJB probes to label retaining glycosidases from additional GH families. Of the 19 identified glycosidases, BGAL (NbS00024332g0007) and α-L-arabinofuranosidase/β-D-xylosidase (NbS00011746g0004.1) mi-
grated at apparent MWs of 45 and 56 kDa, respectively, which are much lower than their theoretical MWs (89.7 and 83.9, respectively; Fig. 6C). The MS analysis identified 12 unique peptides from BGAL and four unique peptides of /H9251-L-arabinofuranosidase/H9252-D-xylosidase. All the identified peptides of BGAL originated from the GH35 catalytic domain, and not from the C-terminal half of this protein (Fig. 6D). This indicates that the active state of BGAL is a 45-kDa truncated protein consisting of only the catalytic domain.

DISCUSSION

We introduced broad-range activity profiling of glycosidases and demonstrated its potential to uncover unexpected...
post-translational regulation of these enzymes. Using cyclophellitol-aziridine probes, we detected the active state of nearly 40 different retaining glycosidases of Arabidopsis and N. benthamiana. These identified glycosidases belong to seven different GH families and are likely to have different substrate specificities. The majority of the detected glycosidases are uncharacterized enzymes. Glycosidase activity profiling revealed the active state of apoplastic glycosidases in N. benthamiana.

Activity-based Glycosidase Profiling—Several experiments have demonstrated that glycosidase labeling is activity dependent. First, glycosidase labeling depends on pH. For example, myrosinases TGG1 and TGG2 and apoplastic glycosidases were labeled intensively at slightly acidic or acidic pH, consistent with in vitro studies of these enzymes (48, 59) and the location of these enzymes in the acidic apoplast, lysosome, or vacuole (60). In general, biomolecules have evolved to perform their functions in the conditions of the compartments where they are located (61). Second, the identified glycosidases are all retaining enzymes. This selectivity is in agreement with the proposed mechanism of these probes to label the active site of retaining glycosidases with a covalent glycosyl-enzyme intermediate. Third, spectral analysis of the JJB111-labeled peptide of TGG1 indicated that the nucleophilic glutamate active site residue was indeed labeled by the probe. This confirms that the warhead of the probe enters the substrate-binding pocket of glycosidases to label the active site. Identification of both labeled and unlabeled versions of active site peptides from TGG1 is explained by the reversible nature of the ester bond formed between the probe and the enzymes, sensitive for hydrolysis during the chemical proteomics workflow. This also explains why we did not identify labeled peptides from the other detected glycosidases. Fourth, labeling of proteins in cell cultures was abolished under denaturing conditions, implying that a native protein structure is essential for labeling. Fifth, competition assays by pre-incubation with glycosidase inhibitor KY371 blocked labeling. Taken together, these results show that the probes labeled glycosidases based on their active state rather than their abundance.

Broad-range Glycosidase Profiling—We discovered the potential of JJB probes to label a greater number of glycosidases than originally thought. Cyclophellitol-aziridine-based probes were initially designed to label retaining glycosidases (33). In addition to glucosidases, we detected glycosidases with additional specificities including xyllosidases, galactosidases, glucuronidase, and glucanases. Interestingly, we also identified an α-1-arabinofuranosidase and heparanase, whose substrates are α-1-arabinofuranoside and heparan sulfate, respectively. Furthermore, labeling of a well-characterized xyllosidase confirmed that the JJB probes do not target only glycosidases. Detection of a large number of different glycosidases in plant proteomes is consistent with the fact that plant genomes have more glycosidase-related genes than animals (4). We detected the active state of 39 different glycosidases representing six different retaining GH families of Arabidopsis and N. benthamiana. The labeled purified xylosidase is a representative of the seventh GH family. In addition to labeling members of GH1, GH3, and GH30 (33), we have now detected labeling of members of the GH5, GH35, GH79, GH51, and GH52 families, which constitute 100 genes in Arabidopsis. Thus JJB probes offer great potential for studying the active state of a large number of glycosidases in plants.

Opportunities Offered by Glycosidase Profiling—Being a large superfamily, glycosidases are challenging to characterize. Glycosidase profiling can be used to study when and which of these enzymes are in their active state. Some glycosidases do not show activity when heterologously expressed and purified. For example, recombinant TGG2 did not have myrosinase activity upon heterologous expression and in vitro enzyme assays (62). This highlights the importance of profiling activities of enzymes under native conditions. To our knowledge, activities contributed by TGG1 or TGG2 have not been studied in wild-type Arabidopsis plants before. In order to study myrosinase activities contributed by TGG2, a tedious extraction and purification procedure had to be followed using tgg1 mutant plants (48). We displayed the active states of TGG1 and TGG2 in wild-type plants without using knockout lines and without complex extraction procedures. Thus glycosidase activity profiling can significantly simplify studies of myrosinases and other glycosidases.

The synthesis of putative substrates via organic chemistry to characterize glycoside specificities is a bottleneck in glycobiology (58). Furthermore, some glycosidases, such as plant α-D-xyllosidases, hydrolyze only their natural substrates and not commercially available synthetic substrates (58, 63). Glycosidase activity profiling bypasses these bottlenecks by displaying active site availability and reactivity, which are hallmarks of enzyme activity (12). Thus active glycosidases can be monitored with ABPP without knowledge of the natural substrates. We also detected active glycosidases that were assumed to be inactive. Detection of TGG3, for example, showed that this enzyme was active, even though it was originally classified as a pseudogene (49). Glycosidase profiling is not restricted to the proteomes from which we have identified labeled proteins. We also profiled for active glycosidases in extracts of different organs and leaves of different dicot and monocot plant species. Thus glycosidase profiling can be used to unravel the roles of glycosidases in other plant species. We have also demonstrated that both the fluorescent probe JJB70 and the inhibitor KY371 enter living cells and label glycosidases in vivo. These experiments demonstrate that KY371 is an excellent broad-range glycosidase inhibitor in living cells. KY371 can now be used in chemical knockout studies to display phenotypes associated with a loss of broad-range glycosidase function at a time point and dose of choice.
Active Glycosidases in the Apoplast—The extracellular space and the cell wall jointly constitute an important part of a plant called the apoplast. In this study, we chose N. benthamiana as a model system to study the active glycosidases in the leaf apoplast. Apoplastic glycosidases can selectively alter cell wall polysaccharides to remodel the cell wall architecture during plant growth and development. By labeling apoplastic proteomes, we detected the active state of four putative bi-functional α-L-arabinofuranosidase/β-D-xylosidases, two putative β-D-xylosidases, five putative β-D-glucosidases, two putative β-D-galactosidases, one putative heparanase, one putative β-D-glucanase, and one putative α-L-arabinofuranosidase. Glycosidases with these substrate-hydroryzizing activities process the major cell wall polysaccharides, including glucan, xylan, arabinoxylan, galactan, and arabinan, during cell wall assembly and reorganization (9). Thus glycosidase profiling reveals the active state of glycosidases that likely play a role in cell wall remodeling. Furthermore, we also detected two GH3 family enzymes with unknown glycosidic activities. Thus glycosidase profiling allows for the detection of the active state of these enzymes without knowledge of their substrates.

The activity of glycosidases is tightly regulated at the post-translational level. For example, cell wall proteome analysis suggests that glycosidases can be regulated by proteases (64). Consistent with this, we found that the labeled BGAL (NbS00024332g0007) was a truncated protein consisting of only the GH35 catalytic domain. Such truncation has also been observed in an animal lysosomal β-galactosidase where the C-terminal half is removed to release a mature enzyme in an acidic environment (65). Interestingly, a C-terminal truncation of a bacterial β-galactosidase resulted in an efficient transgalactosylase (66). The truncated BGAL in the plant apoplast might have acquired transglycosylase activity and might alter cell wall polysaccharides to remodel the cell wall architecture.

REFERENCES

1. Sturgeon, R. J. (2003) Carbohydrates. In eLS, John Wiley & Sons, online article available at: 10.1002/9780470015902.a0020548.pub2
2. Cheeke, P. R. (2001) Glycosides: naturally occurring. In eLS, John Wiley & Sons, online article available at: 10.1002/9780470015902.a0000689
3. Hancock, S. M., and Withers, S. G. (2014) Glycosidases: functions, families and folds. In eLS, John Wiley & Sons, online article available at: 10.1002/9780470015902.a0020548.pub2
4. Coutinho, P. M., Stamm, M., Bianc, E., and Henrisat, B. (2003) Why are there so many carbohydrate-active enzyme-related genes in plants? Trends Plant Sci. 8, 563–565
5. Henrisat, B. (1991) A classification of glycosyl hydrolases based on amino-acid-sequence similarities. Biochem. J. 280, 309–316
6. Sinnott, M. L. (1990) Catalytic mechanism of enzymic glycosyl transfer. Chem. Rev. 90, 1171–1202
7. Mocjar, C. D., and Withers, S. G. (1994) Mechanisms of enzymatic glycoside hydrolysis. Curr. Opin. Struct. Biol. 4, 885–892
8. Davies, G., and Henrisat, B. (1993) Structures and mechanisms of glycosyl hydrolases. Structure 3, 853–859
9. Minici, Z. (2008) Physiological roles of plant glycoside hydrolases. Planta 227, 723–740
10. Cravatt, B. F., Wright, A. T., and Kozarich, J. W. (2008) Activity-based protein profiling: from enzyme chemistry to proteomic analysis. Annu. Rev. Biochem. 77, 383–414
11. Kolodziejek, I., and Van der Hoorn, R. A. L. (2010) Mining the active proteome in plant science and biotechnology. Curr. Opin. Biotechnol. 21, 225–233
12. Kobe, B., and Kemp, B. E. (1999) Active site-directed protein regulation. Nature 402, 373–376
13. Richau, K. H., Kaschani, F., Verdoes, M., Panstruya, T. C., Niessen, S., Stuber, K., Colby, T., Overkleeft, H. S., Bogoy, M., and Van der Hoorn, R. A. L. (2012) Subclassification and biochemical analysis of plant papain-like cysteine proteases displays subfamily-specific characteristics. Plant Physiol. 158, 1583–1599
14. Van der Hoorn, R. A. L., Leeuwenburgh, M. A., Bogoy, M., Joosthen, M. H., and Peck, S. C. (2004) Activity profiling of papain-like cysteine proteases in plants. Plant Physiol. 135, 1170–1178
15. Shabab, M., Shindo, T., Gu, C., Kaschani, F., Panstruya, T., Chintna, R., Harzen, A., Colby, T., Kamoun, S., and Van der Hoorn, R. A. L. (2008) Fungal effector protein AVR2 targets diversifying defense-related Cys proteases of tomato. Plant Cell 20, 1169–1183
16. Van der Linde, K., Mueller, A. N., Hemetsberger, C., Kashani, F., Van der Hoorn, R. A. L., and Doehlemann, G. (2012) The maize cystatin CC9 interacts with apoplastic cysteine proteases. Plant Signal. Behav. 7, 1397–1401
17. Dong, S. M., Starm, R., Cano, L. M., Song, J., Sklenar, J., Yoshida, K., Bozkurt, T., Oto, T., Li, Z. Y., Tian, M. Y., Win, J., Banfield, M. J., Jones, A. M. E., Van der Hoorn, R. A. L., and Kamoun, S. (2014) Effector specialization in a lineage of the Irish potato famine pathogen. Science 343, 552–555
18. Hörger, A. C., Ilyas, M., Stephan, W., Tellier, A., Van der Hoorn, R. A. L., and rose, L. E. (2012) Balancing selection at the tomato RCR3 gueradee gene family maintains variation in strength of pathogen defense. PLoS Genet. 8, e1002813
19. Kaschani, F., Shabab, M., Bozkurt, T., Shindo, T., Schornack, S., Gu, C., Ilyas, M., Win, J., Kamoun, S., and Van der Hoorn, R. A. L. (2010) An effector-targeted protease contributes to defense against Phytophthora infestans and is under diversifying selection in natural hosts. Plant Physiol. 154, 1784–1804
20. Lozano-Torres, J. L., Wilbers, R. H. P., Gawronski, P., Boshoven, J. C., Finkers-Tomczak, A., Cordewener, J. H. G., America, A. H. P., Overmars, H. A., Van ‘t Klooster, J. W., Baranowski, L., Sobczak, M., Ilyas, M., Van der Hoorn, R. A. L., Schots, A., de Wit, P. J. G. M., Bakker, J., Gouverse, A., and Smant, G. (2012) Dual disease resistance mediated by the immune receptor Cf-2 in tomato requires a common virulence target of a fungus and a nematode. Proc. Natl. Acad. Sci. U.S.A. 109, 10119–10124
21. Mueller, A. N., Ziemann, S., Treitschke, S., Assmann, D., and Doehlemann, G. (2013) Compatibility in the Ustilago maydis-maize interaction requires inhibition of host cysteine proteases by the fungal effector Pit2. PLoS Pathog. 9, e1003177
22. Rooney, H. C., Van ’t Klooster, J. W., Van der Hoorn, R. A. L., Joosthen,
64. Minic, Z., Jamet, E., Negroni, L., Arsene der Garabedian, P., Zivy, M., and Jouanin, L. (2007) A sub-proteome of Arabidopsis thaliana mature stems trapped on Concanavalin A is enriched in cell wall glycoside hydrolases. J. Exp. Bot. 58, 2503–2512
65. Van der Spoel, A., Bonten, E., and d’Azzo, A. (2000) Processing of lysosomal beta-galactosidase—the C-terminal precursor fragment is an essential domain of the mature enzyme. J. Biol. Chem. 275, 10035–10040
66. Jorgensen, F., Hansen, O. C., and Stougaard, P. (2001) High-efficiency synthesis of oligosaccharides with a truncated beta-galactosidase from Bifidobacterium bifidum. Appl. Microbiol. Biot. 57, 647–652