Identification and characterisation of *Arabidopsis* glycosyltransferases capable of glucosylating coniferyl aldehyde and sinapyl aldehyde☆

Eng-Kiat Lim, Rosamond G. Jackson, Dianna J. Bowles*

CNAP, Department of Biology, University of York, York YO10 5DD, UK

Received 31 January 2005; revised 3 April 2005; accepted 4 April 2005

Available online 25 April 2005

Edited by Ulf-Ingo Flügge

Abstract This study describes the substrate recognition profile of UGT72E1, an UDP–glucose:glycosyltransferase of *Arabidopsis thaliana* that is the third member of a branch of glycosyltransferases, capable of conjugating lignin monomers and related metabolites. The data show that UGT72E1, in contrast to the two closely related UGTs 72E2 and 72E3, is specific for sinapyl and coniferyl aldehydes. The biochemical properties of UGT72E1 are characterised, and are compared with that of UGT72E2, which is capable of glycosylating the aldehydes as well as coniferyl and sinapyl alcohols.
© 2005 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: Glycosyltransferase; Lignin; Monolignol; Coniferyl aldehyde; Sinapyl aldehyde

1. Introduction

Lignins are naturally occurring polymers that provide strength, protection, and water impermeability to the polysaccharide matrix of the plant cell wall. Recent studies on lignin composition reveal that, in addition to the three classical monolignols (*p*-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol), the polymers consist of many other monomers such as coniferyl aldehyde and sinapyl aldehyde [1–3]. Whilst the biosynthetic routes of these monomers are still under debate [4–6], it is generally thought that the cellular homeostasis of these molecules is regulated through glucosylation [1,7]. This modification step increases the solubility and stability of the monomers and provides access to the membrane transport systems for transportation and storage [8]. The glucosides of lignin monomers have been reported in many gymnosperms and angiosperms [9], and recently have been found to accumulate in light-grown *Arabidopsis* roots [7]. As yet, it is unclear whether these glucosides are transported out of the cell for polymerisation into lignin, although a coniferyl alcohol glucoside β-glucosidase has been immunolocalised within xylem secondary walls [10,11].

Glucosylation of lignin monomers usually occurs at the hydroxyl group on the phenolic ring, leading to the formation of 4-*O*-glucosides. The reaction is catalysed by UDP–glucose:coniferyl alcohol glucosyltransferase (UGT) (EC 2.4.1.111). This enzyme was first partially purified from cell suspension cultures of Paul's scarlet rose thirty years ago [12], but the corresponding gene sequence has not been identified. In a recent study, when 36 recombinant *Arabidopsis* glycosyltransferases (UGTs) containing a consensus sequence of 44 amino acids involved in nucleotide sugar-binding were screened in vitro for activity towards a range of phenylpropanoid derivatives, two UGTs, 72E2 and 72E3, were found to form the 4-*O*-glucosides of two monolignols coniferyl alcohol and sinapyl alcohol [13].

A total number of 107 UGTs carrying the same consensus has now been identified from the complete *Arabidopsis* genome, and a comprehensive phylogenetic tree of these UGTs has been constructed [14–16]. The two UGTs 72E2 and 72E3 conjugating monolignols were classified into the phylogenetic Group E. This study describes the characterisation of the most closely related enzyme to UGTs 72E2 and 72E3 in the same phylogenetic branch (Fig. 1). The results show that the enzyme, UGT72E1, displays similar but distinct catalytic activity from UGTs 72E2 and 72E3. UGT72E1 is highly specific to coniferyl aldehyde and sinapyl aldehyde, and shows only negligible activity towards coniferyl alcohol and sinapyl alcohol.

2. Materials and methods

2.1. Recombinant UGT purification

Recombinant UGTs 72E1 and 72E2 were expressed as fusion proteins with glutathione-*S*-transferase (GST) attached to the N-terminus of the UGTs. The GST gene fusion vector pGEX-2T (Amersham Pharmacia Biotech) containing the cDNA of UGT72E1 or UGT72E2 was transformed into *E. coli* XL-1 Blue for recombinant protein expression. The bacterial cells were grown in 1 L of 2× YT medium containing 50 μg/ml ampicillin at 20 °C until A_{600} reading reached 1.0. The culture was then incubated with 1 mM isopropyl-1-thio-β-D-galactopyranoside for 24 h at 20 °C. Cells were harvested (5000 × *g* for 5 min), resuspended (5 ml of ice-cold phosphate-buffered saline), disrupted by French Press (Thermo Electron) with 1400 psi, and centrifuged again (40 000 × *g* for 5 min). The supernatant was mixed with 100 μl of 50% glutathione-coupled Sepharose (Amersham Pharmacia Biotech) at room temperature for 30 min. The beads were washed with phosphate-buffered saline,

☆ Nucleotide sequence data reported are available in the DDBJ/EMBL/GenBank databases under the accession numbers AL049862 (UGT72E1), AB018119 (UGT72E2) and AF077407 (UGT72E3).

*Corresponding author. Fax: +44 190 432 8772.
E-mail address: djb32@york.ac.uk (D.J. Bowles).

Abbreviations: UGT, UDP–glucose:glycosyltransferase; GST, glutathione-*S*-transferase; HPLC, high pressure liquid chromatography; NMR, nuclear magnetic resonance

0014-5793/$30.00 © 2005 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.febslet.2005.04.016

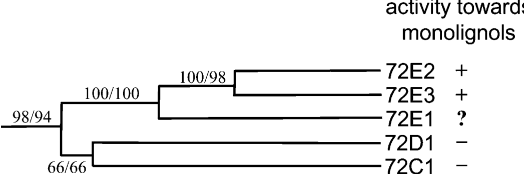

Fig. 1. Phylogenetic relationship of the *Arabidopsis* UGTs with activity towards monolignols. A branch in Group E on the phylogenetic tree of *Arabidopsis* UGTs [14,15] contains the UGTs previously assayed for activity towards monolignols [13]. Whereas 72E2 and 72E3 showed significant activity (+), and 72C1 and 72D1 showed no activity (−), 72E1 was not assayed [13].

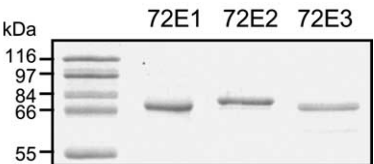

Fig. 2. SDS–PAGE analysis of UGTs 72E1. 72E2 and 72E3. The UGTs were analysed on a 10% (w/v) polyacrylamide gel and visualised with Coomassie Brilliant Blue staining.

and the absorbed proteins were eluted with 20 mM reduced-form glutathione according to the manufacturer's instructions. The concentration of the protein was determined with Bio-Rad Protein Assay Dye using bovine serum albumin as reference.

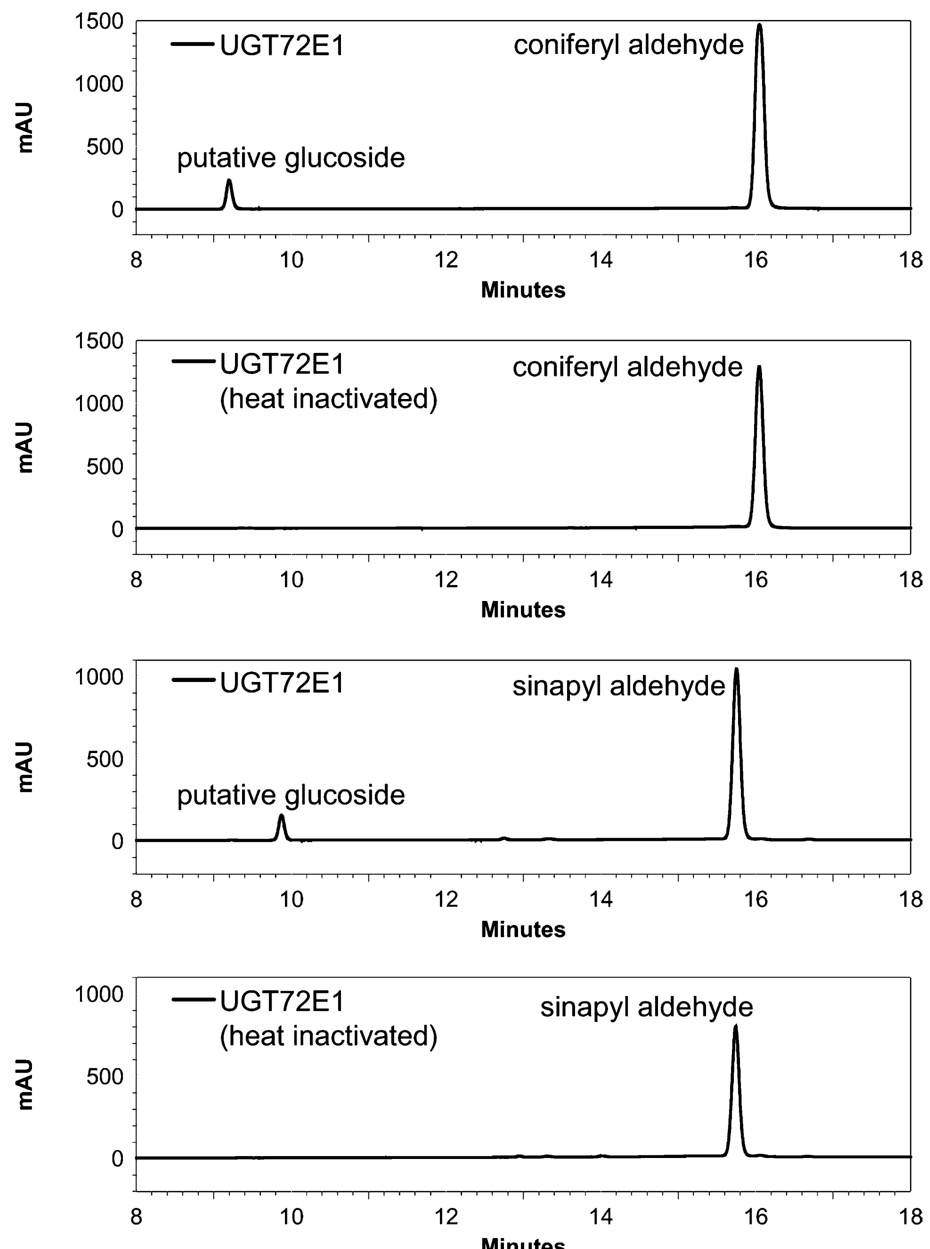

Fig. 3. HPLC analysis of the reaction mixes containing UGT72E1, UDP–glucose and the substrate coniferyl aldehyde or sinapyl aldehyde. A linear gradient of acetonitrile from 10% to 47% was used in the HPLC analysis. UGT72E1 heat-inactivated at 100 °C for 5 min was used as the negative control in the assays. The chromatograms were monitored at 340 nm.

2.2. Glycosyltransferase activity assay

The glycosyltransferase activity assay was carried out following the conditions described in our previous study [13] with modification. The assay mix (200 µl) contained 0.2 µg of recombinant protein, 5 mM UDP–glucose and 1 mM phenylpropanoid substrate. The reaction was carried out at pH 7.0 (100 mM Tris–HCl) and 30 °C for 30 min due to the enzyme pH optima and linearity of the reaction, and was stopped by the addition of 20 µl of trichloroacetic acid (240 mg/ml), quick-frozen and stored at −20 °C prior to the reverse-phase HPLC analysis. Each recombinant UGT activity assay containing a single substrate was analysed using one of the methods described in the following section.

2.3. HPLC analysis of the in vitro reaction mixtures

Reverse-phase HPLC (SpectraSYSTEM HPLC systems and UV6000LP Photodiode Array Detector, ThermoQuest) analysis was carried out using a Columbus 5 µ C_{18} column (250 × 4.60 mm, Phenomenex) maintained at 30 °C. A linear gradient of acetonitrile in H_2O (all solutions contained 0.1% trifluoroacetic acid) at 1 ml/min over 20 min, was used to separate the glucose conjugates from their aglycone. The HPLC methods were described as the following: cinnamic acid, $\lambda_{288\ nm}$, 10–55% acetonitrile; *p*-coumaric acid, $\lambda_{311\ nm}$, 10–25% acetonitrile; caffeic acid, $\lambda_{311\ nm}$, 10–16% acetonitrile; ferulic acid, $\lambda_{311\ nm}$, 10–35% acetonitrile; sinapic acid, $\lambda_{306\ nm}$, 10–40% acetonitrile; *p*-coumaryl aldehyde, $\lambda_{315\ nm}$, 10–46% acetonitrile; coniferyl aldehyde and sinapyl aldehyde, $\lambda_{340\ nm}$, 10–47% acetonitrile; *p*-coumaryl alcohol, $\lambda_{254\ nm}$, 10–27% acetonitrile; coniferyl alcohol, $\lambda_{260\ nm}$, 10–30% acetonitrile; sinapyl alcohol, $\lambda_{270\ nm}$, 10–25% acetonitrile. The retention time of the glucose conjugates analysed was as follows: feruloyl-4-*O*-glucoside, 7.9 min; sinapoyl-4-*O*-glucoside, 8.5 min; coniferyl aldehyde-4-*O*-glucoside, 9.4 min; sinapyl aldehyde-4-*O*-glucoside, 10.1 min; coniferyl alcohol-4-*O*-glucoside, 8.1 min; sinapyl alcohol-4-*O*-glucoside, 8.9 min. The data were acquired and analysed using the software ChromQuest version 2.51.

2.4. ^1H NMR analysis

The aldehyde glucosides synthesised in the enzymatic reactions were purified by HPLC and were collected with a Gilson FC 240 fraction collector. The samples were freeze-dried and resuspended in deuterated methanol. The NMR spectra of the glucosides were acquired on a Bruker AMX 500-MHz NMR spectrometer at 22 °C. The data were processed and analysed using Bruker XWIN-NMR software version 2.6.

2.5. Steady-state enzyme kinetic measurements

Michaelis–Menten kinetics of the enzymes were measured over a range of substrate concentrations of 0–0.5 mM in the presence of 0.2 µg of recombinant protein, 100 mM Tris–HCl, pH 7.0, and 5 mM UDP–glucose. The reactions (200 µl) were carried out at 30 °C for 30 min and were stopped by the addition of 20 µl of trichloroacetic acid (240 mg/ml), quick-frozen and stored at −20 °C prior to the reverse-phase HPLC analysis. The kinetic parameters were derived using Hyperbolic Regression Analysis of Hyper32 programme available from www.homepage.ntworld.com/john.easterby (Copyright J.S. Easterby).

3. Results

3.1. Stability of coniferyl aldehyde glucoside and sinapyl aldehyde glucoside

In our previous study of UGTs 72E2 and 72E3, no significant activity towards coniferyl aldehyde and sinapyl aldehyde was detected [13]. When the reaction conditions were further investigated, the results revealed that 2-mercaptoethanol and trichloroacetic acid, which were included in the initial study, together destabilise the aldehyde glucosides formed in the reactions, thereby decreasing the levels of products detected (data not shown). Thus, in this study 2-mercaptoethanol was omitted from the in vitro analyses.

Three UGTs 72E1, 72E2, and 72E3 were expressed in *E. coli* and were purified (Fig. 2) for in vitro biochemical characterisation towards coniferyl aldehyde and sinapyl aldehyde. Both UGTs 72E1 and 72E2 were found to form putative products which were absent from the reaction mixes containing UGT72E3. Fig. 3 shows the reverse-phase HPLC analyses of the reaction mixes of UGT72E1. The putative products were absent in the reactions mixes containing heat-inactivated enzyme. These products were further purified and analysed by NMR. In comparison to that of the aglycone, the NMR spectrum of the product of coniferyl aldehyde showed a significant up-field shift at the proton attached to the C5 position (Table 1), indicating that the neighbouring C4-OH was conjugated. A similar chemical shift was not observed with the product of sinapyl aldehyde since there is no proton adjacent to the C4–OH (Table 1). The NMR spectra shown in Table 1 are near-identical to those reported for coniferyl aldehyde-4-*O*-glucoside and sinapyl aldehyde-4-*O*-glucoside by other research groups [3,9], and thus confirm the identities of the products.

Table 1
^1H NMR spectral data of coniferyl aldehyde, sinapyl aldehyde and the corresponding glucosides produced by the UGTs described in this study

Position	Coniferyl aldehyde	Coniferyl aldehyde 4-*O*-glucoside	Sinapyl aldehyde	Sinapyl aldehyde 4-*O*-glucoside
	MeO, 2, 8, O, 9, H, HO, 7, 5, 6	MeO, 2, 8, O, 9, H, Glc-O, 7, 5, 6	MeO, 2, 8, O, 9, H, HO, 7, MeO, 6	MeO, 2, 8, O, 9, H, Glc-O, 7, MeO, 6
C2	7.24 (*d*, 1.5)	7.32 (*d*, 1.5)	6.97 (*s*)	6.99 (*s*)
C5	6.84 (*d*, 8.2)	7.21 (*d*, 8.2)	–	–
C6	7.16 (*dd*, 8.2, 1.5)	7.26 (*dd*, 8.2, 1.5)	6.97 (*s*)	6.99 (*s*)
C7	7.70 (*d*, 16.0)	7.62 (*d*, 16.0)	7.56 (*d*, 16.0)	7.59 (*d*, 16.0)
C8	6.63 (*dd*, 16.0, 8.0)	6.71 (*dd*, 16.0, 8.0)	6.67 (*dd*, 16.0, 8.0)	6.68 (*dd*, 16.0, 8.0)
C9	9.56 (*d*, 8.0)	9.61 (*d*, 8.0)	9.57 (*d*, 8.0)	9.58 (*d*, 8.0)
OCH_3	3.89 (*s*)	3.90 (*s*)	3.88 (*s*)	3.89 (*s*)
Glc–C1	–	5.06 (*d*, 7.5)	–	5.04 (*d*, 7.5)
Glc–C2–C5	–	3.38-3.54 (*m*)	–	3.33–3.47 (*m*)
Glc–C6	–	3.88 (*dd*, 12.0, 2.0)	–	~3.84 (interrupted)
		3.68 (*dd*, 12.0, 5.5)		~3.66 (interrupted)

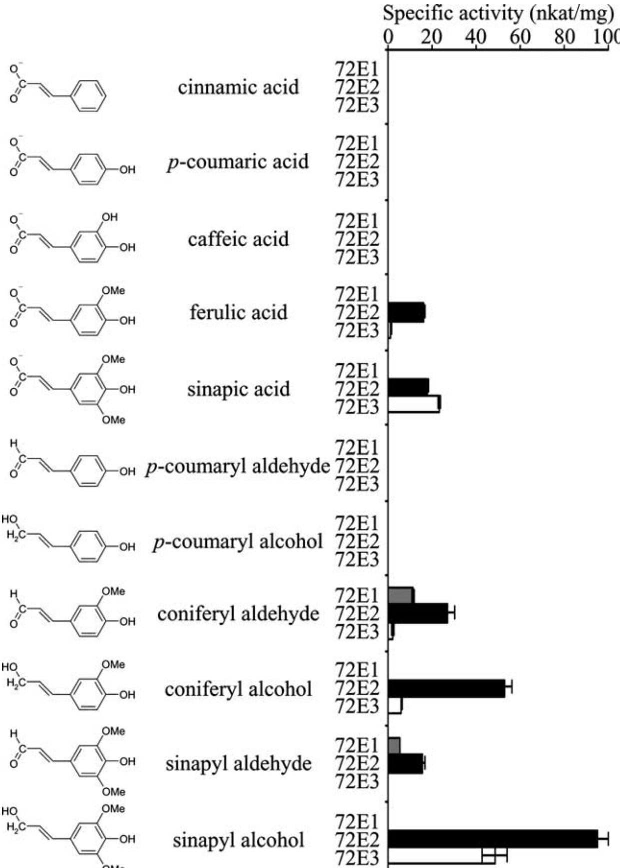

Fig. 4. Specific activity of UGTs 72E1, 72E2 and 72E3 towards various phenylpropanoid derivatives. The specific activity was defined as nmol of substrates converted into glucose conjugates/s (nanokat, nkat) by 1 mg of protein. The glucosides formed in the reactions were quantified using the extinction coefficient of the aglycones after hydrolysis of the purified glucosides.

3.2. Activity of UGTs 72E1, 72E2 and 72E3 towards various phenylpropanoid derivatives

When the three recombinant UGTs were analysed in vitro against 11 phenylpropanoid derivatives, different substrate recognition profiles were obtained. As summarised in Fig. 4, in addition to coniferyl aldehyde and sinapyl aldehyde, UGT72E2 conjugates a range of substrates including ferulic acid, sinapic acid, coniferyl alcohol and sinapyl alcohol in the presence of UDP–glucose. In contrast, UGT72E1 shows high substrate specificity towards coniferyl aldehyde and sinapyl aldehyde whereas UGT72E3 displays significant activity towards sinapic acid, coniferyl alcohol and sinapyl alcohol. Since this study focuses on the enzymes capable of glucosylating coniferyl aldehyde and sinapyl aldehyde, only UGTs 72E1 and 72E2 were characterised further in the steady-state kinetic analyses.

3.3. Steady-state kinetics of UGT72E1 towards coniferyl aldehyde and sinapyl aldehyde

The steady-state kinetics of UGT72E1 were determined using two acceptors coniferyl aldehyde and sinapyl aldehyde. The kinetic constants are summarised in Table 2. The k_{cat}/K_m values of UGT72E1 towards coniferyl aldehyde and sinapyl aldehyde are 4.51 and 3.10 $mM^{-1}\ s^{-1}$, respectively. These values are much lower than that of UGT72E2. Taken together with the lower K_m values of UGT72E2, the results suggest that UGT72E2 has higher substrate binding affinity than UGT72E1 towards the aldehydes. Among the four substrates analysed in this study, coniferyl aldehyde is the preferred substrate for UGT72E2 (Table 2).

4. Discussion

In this study, we report an *Arabidopsis* enzyme UGT72E1 that is highly specific to coniferyl aldehyde and sinapyl aldehyde in vitro. In contrast to the broad substrate recognition of UGT72E2, UGT72E1 did not conjugate cinnamic acids, and showed only negligible activity towards coniferyl alcohol and sinapyl alcohol, irrespective to the similar phenolic ring structures of these compounds (Fig. 4). When another *Arabidopsis* enzyme UGT72E3, that is capable of conjugating sinapyl alcohol in vitro, was analysed towards coniferyl aldehyde and sinapyl aldehyde, no significant activity was observed (Fig. 4). Thus, despite their amino acid sequence similarity of over 65%, these three UGTs have different substrate recognition profiles in vitro. Whilst UGT72E2 recognises both monolignols and the related aldehydes, UGT72E1 conjugates only coniferyl aldehyde and sinapyl aldehyde. Although UGT72E1 is highly specific to the aldehydes, it has a much lower substrate affinity than UGT72E2 (Table 2). UGT72E2 clearly shows higher activities for in vitro conversion of coniferyl aldehyde and sinapyl aldehyde into glucosides (Fig. 4). These studies provide some indication of substrate recognition and affinity in vitro, further studies will be required to investigate the in vivo substrates of these enzymes.

In planta monolignols are considered to be mainly used in lignin biosynthesis. These monomers may be exported out of the cell through Golgi-mediated secretion or directly by membrane-bound transporters [1,10]. In contrast to the destination of monolignols, coniferyl aldehyde, and sinapyl aldehyde can act as the precursors of ferulic acid and sinapic acid as well as lignin polymers [4]. In this context, glucosylation of coniferyl aldehyde and sinapyl aldehyde may regulate both lignin biosynthesis and the metabolism of other phenylpropanoids

Table 2
Summary of the steady-state kinetic parameters of UGTs 72E1 and 72E2

	72E1			72E2		
	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1}\ s^{-1}$)	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1}\ s^{-1}$)
Coniferyl aldehyde	0.27	1.22	4.52	0.02	1.37	68.50
Sinapyl aldehyde	0.46	1.43	3.11	0.02	1.08	52.50
Coniferyl alcohol	–	–	–	0.06	3.17	52.83
Sinapyl alcohol	–	–	–	0.15	4.52	30.13

such as ferulic acid, 5-hydroxyferulic acid, sinapic acid and their derivatives. Given that three *Arabidopsis* UGTs have been identified with the capability of glucosylating lignin monomers in vitro and their substrate recognition profile characterised, analysis of transgenic plants with upregulation or downregulation of these UGTs may provide further understanding in the cellular homeostasis of lignin monomers in plant cells, flux through the phenylpropanoid metabolic pathway, and the biosynthesis of lignin polymers.

Acknowledgements: We thank Professor John Ralph (U.S. Dairy Forage Research Center) for providing *p*-coumaryl aldehyde and *p*-coumaryl alcohol. Heather Fish (Department of Chemistry, York) is thanked for help with NMR studies. This study was supported by The Garfield Weston Foundation and Biotechnology and Bioscience Research Council (BBSRC) (Grant No.: 87/P12844).

References

[1] Boerjan, W., Ralph, J. and Baucher, M. (2003) Lignin biosynthesis. Annu. Rev. Plant Biol. 54, 519–546.
[2] Whetten, R.W., MacKay, J.J. and Sederoff, R.R. (1998) Recent advances in understanding lignin biosynthesis. Annu. Rev. Plant Physiol. Plant Mol. Biol. 49, 585–609.
[3] Kim, H., Ralph, J., Lu, F., Ralph, S.A., Boudet, A.-M., MacKay, J.J., Sederoff, R.R., Ito, T., Kawai, S., Ohashi, H. and Higuchi, T. (2003) NMR analysis of lignins in CAD-deficient plants. Part 1. Incorporation of hydroxycinnamaldehydes and hydroxybenzaldehydes into lignins. Org. Biomol. Chem. 1, 268–281.
[4] Nair, R.B., Bastress, K.L., Ruegger, M.O., Denault, J.W. and Chapple, C. (2004) The *Arabidopsis thaliana* reduced epidermal fluorescence1 gene encodes an aldehyde dehydrogenase involved in ferulic acid and sinapic acid biosynthesis. Plant Cell 16, 544–554.
[5] Humphreys, J.M. and Chapple, C. (2002) Rewriting the lignin roadmap. Curr. Opin. Plant Biol. 5, 224–229.
[6] Dixon, R.A., Chen, F., Guo, D. and Parvathi, K. (2001) The biosynthesis of monolignols: a "metabolic grid", or independent pathways to guaiacyl and syringyl units?. Phytochemistry 57, 859–873.
[7] Hemm, M.R., Rider, S.D., Ogas, J., Murry, D.J. and Chapple, C. (2004) Light induces phenylpropanoid metabolism in *Arabidopsis* roots. Plant J. 38, 765–778.
[8] Hösel, W. (1981) Glycosylation and glycosidases in: The Biochemistry of Plant (Stumpf, P.K. and Conn, E.E., Eds.), Secondary Plant Products, vol. 7, pp. 725–753, Academic Press, New York.
[9] Steeves, V., Förster, H., Pommer, U. and Savidge, R. (2001) Coniferyl alcohol metabolism in conifers – I. Glucosidic turnover of cinnamyl aldehydes by UDPG:coniferyl alcohol glucosyltransferase from pine cambium. Phytochemistry 57, 1085–1093.
[10] Samuel, A.L., Rensing, K.H., Douglas, C.J., Mansfield, S.D., Dharmawardhana, D.P. and Ellis, B.E. (2002) Cellular machinery of wood production: differentiation of secondary xylem in *Pinus contorta* var. *latifolia*. Planta 216, 72–82.
[11] Dharmawardhana, D.P., Ellis, B.E. and Carlson, J.E. (1995) A β-glucosidase from lodgepole pine xylem specific for the lignin precursor coniferin. Plant Physiol. 107, 331–339.
[12] Ibraham, R.K. and Grisebach, H. (1976) Purification and properties of UDP–glucose:coniferyl alcohol glucosyltransferase from suspension cultures of Paul's Scarlet rose. Arch. Biochem. Biophys. 176, 700–708.
[13] Lim, E.-K., Li, Y., Parr, A., Jackson, R., Ashford, D.A. and Bowles, D.J. (2001) Identification of glucosyltransferase genes involved in sinapate metabolism and lignin synthesis in *Arabidopsis*. J. Biol. Chem. 276, 4344–4349.
[14] Li, Y., Sandie, B., Lim, E.-K. and Bowles, D.J. (2001) Phylogenetic analysis of the UDP-glycosyltransferase multigene family of *Arabidopsis thaliana*. J. Biol. Chem. 276, 4338–4343.
[15] Ross, J., Li, Y., Lim, E.-K. and Bowles, D.J. (2001) Higher plant glycosyltransferases. Genome Biol. 2, 3004.1–3004.6.
[16] Lim, E.-K. and Bowles, D.J. (2004) A class of plant glycosyltransferases involved in cellular homeostasis. EMBO J. 23, 2915–2922.