



Cloning and endogenous expression of a *Eucalyptus grandis* UDP-glucose dehydrogenase cDNA

Mônica T. Veneziano Labate, Ana L. Ferreira Bertolo, Daniela Defávares do Nascimento, Gunta Gutmanis, Alexander de Andrade, Maria J. Calderan Rodrigues, Eduardo L.O. Camargo, Luis Felipe Boaretto, David H. Moon, Juliano Bragatto and Carlos A. Labate

Laboratório Max Feffer de Genética de Plantas, Departamento de Genética, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, SP, Brazil.

Abstract

UDP-glucose dehydrogenase (UGDH) catalyzes the oxidation of UDP-glucose (UDP-Glc) to UDP-glucuronate (UDP-GlcA), a key sugar nucleotide involved in the biosynthesis of plant cell wall polysaccharides. A full-length cDNA fragment coding for UGDH was cloned from the cambial region of 6-month-old *E. grandis* saplings by RT-PCR. The 1443-bp-ORF encodes a protein of 480 amino acids with a predicted molecular weight of 53 kDa. The recombinant protein expressed in *Escherichia coli* catalyzed the conversion of UDP-Glc to UDP-GlcA, confirming that the cloned cDNA encodes UGDH. The deduced amino acid sequence of the cDNA showed a high degree of identity with UGDH from several plant species. The Southern blot assay indicated that more than one copy of *UGDH* is present in *Eucalyptus*. These results were also confirmed by the proteomic analysis of the cambial region of 3- and 22-year-old *E. grandis* trees by 2-DE and LC-MS/MS, showing that at least two isoforms are present. The cloned gene is mainly expressed in roots, stem and bark of 6-month-old saplings, with a lower expression in leaves. High expression levels were also observed in the cambial region of 3- and 22-year-old trees. The results described in this paper provide a further view of the hemicellulose biosynthesis during wood formation in *E. grandis*.

Key words: UDP-glucose, UDP-glucuronate, hemicellulose, pectin, cell wall.

Received: July 7, 2009; Accepted: March 17, 2010.

Introduction

The biosynthesis of hemicelluloses and pectins in higher plants is mainly regulated by the cytosolic enzyme UDP-glucose dehydrogenase (UGDH) (EC 1.1.1.22), that catalyzes the sugar interconversion involving the four-electron, NAD⁺-linked, oxidation of UDP-glucose to UDP-glucuronate (Amino *et al.*, 1985; Gibeaut, 2000). The reaction is essentially irreversible, resulting in a unidirectional flow of UDP-glucuronate (UDP-GlcA) into a pool of sugar nucleotides, predominantly used for the biosynthesis of plant cell wall polysaccharides. UDP-GlcA is the precursor of many sugar nucleotides, including UDP-galacturonic acid (UDP-GalA), UDP-xylose (UDP-Xyl), UDP-arabinose (UDP-Ara) and UDP-apiose (UDP-Api), which are the substrates for the polymer synthases involved in the formation of pectins and hemicelluloses (Gibeaut and Carpita, 1994). UDP-glucuronate is the dominant sugar nucleotide precursor in the biosynthesis of hemicelluloses and pectins, providing half of the biomass of the primary cell wall

(Zabackis *et al.*, 1995; Reiter and Vanzin, 2001; Seifert, 2004).

The presence of isoforms for UGDHs was ignored in previous studies (Tenhaken and Thulke, 1996; Stewart and Copeland, 1998; Seitz *et al.*, 2000). In *Arabidopsis*, the *UGDH* gene family is represented by four highly similar *UGDH* isoforms: *UGDH1*, earlier described by Seitz *et al.* (2000), and three others (*UGDH2-4*), located on distinct chromosomes (Reiter and Vanzin, 2001). More recently, a fifth *UGDH* pseudogene (partial sequence) with a weaker similarity was detected (Klinghammer and Tenhaken, 2007). Moreover, in poplar, at least two isoforms were reported (Johansson *et al.*, 2002). A developmentally regulated and often transient expression pattern for *UGDH* transcripts was observed by Seitz *et al.* (2000), suggesting that the isoforms are expressed in cells only when sugar nucleotides derived from the UDP-GlcA are needed for the synthesis of new cell wall polymers. Developmental regulation was also suggested by Carvalho *et al.* (2008) in *E. grandis* wood-forming tissue, where one UGDH isoform is preferentially expressed.

The biosynthesis of UDP-GlcA in plants can also occur by the inositol oxygenation pathway. In young seed-

Send correspondence to Carlos A. Labate. Laboratório Max Feffer de Genética de Plantas, Departamento de Genética, Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, 13418-900 Piracicaba, SP, Brazil. E-mail: calabate@esalq.usp.

lings, this alternative pathway operates and oxidizes inositol directly to GlcA, as shown by isotope-labeling experiments (Loewus *et al.*, 1973; Roberts and Loewus, 1973; Sasaki and Taylor, 1984). The first irreversible step in the oxidation of myo-inositol is catalyzed by myo-inositol oxygenase (EC 1.13.99.1), generating D-glucuronate that is phosphorylated by D-glucuronokinase (EC 2.7.1.43) to D-glucuronate 1-P (Biswas *et al.*, 1984), and immediately transformed into UDP-glucuronate by UDP-glucuronate-1-phosphate uridylyltransferase (EC 2.7.7.44). Studies on enzymes participating in the sugar nucleotide interconversion showed that UGDH is often the least active enzyme of the pathway, and is also present in low amounts, leading to the conclusion that this enzyme might be rate-limiting for the synthesis of cell wall precursors, since there is a demand for UDP-glucuronate in all plant tissues and organs at all stages of development (Dalessandro and Northcote, 1977a,b; Robertson *et al.*, 1995; Gibeaut, 2000).

In wood-forming tissues, the role of UGDH in the regulation of hemicellulose biosynthesis is still poorly understood, particularly in fast growing trees such as *Eucalyptus spp.* In the present study, we cloned a UGDH-encoding cDNA from the cambial zone of 6-month-old *E. grandis* saplings and analyzed the expression pattern of transcripts from different developing tissues and plant ages by semi-quantitative RT-PCR. The cDNA sequences for UGDH from woody species have previously been documented for poplar (Johansson *et al.*, 2002) and cinnamon (NCBI, AY496079), and recently for *Eucalyptus gunnii*, although incomplete (EMBL, CT982226). Therefore, this is the first complete sequence of a UGDH cDNA reported for *E. grandis*. Moreover, this work led to the conclusion, based on Southern blot analysis, that there are at least 2 copies of the UGDH gene in *E. grandis*. We also isolated and identified three UGDH protein isoforms from the cambial region of 3- and 22-year-old trees, using 2-DE gels and LC-MS/MS.

Material and Methods

Plant material

Six-month-old saplings from a commercial clone of *E. grandis*, kindly provided by Suzano Papel e Celulose, were maintained in growth chambers under controlled conditions: 500 mol m⁻² s⁻¹ irradiance, 16/8 h light/dark photoperiod, and a 24 °C day/18 °C night temperature regime. The cambial region of 3- and 22-year-old-trees, containing the differentiating xylem and phloem tissues, was sampled by removing the bark and scrapping the inner parts of the stems with a razor blade, and immediately freezing it in liquid nitrogen, as described by Celedon *et al.* (2007).

Design of degenerate primers

The UGDH-encoding cDNA of *E. grandis* was cloned, using the RT-PCR technique employing degenerate

primers to highly conserved 5' and 3' ends (Kunihiro *et al.*, 2002). To accomplish this, a BLAST search (Altschul *et al.*, 1977) was performed for plant orthologous sequences encoding UGDH protein. A preliminary analysis of the highly conserved N and C terminal regions among amino acid sequences of UGDH was made by means of the ClustalW2 program (Thompson *et al.*, 1994), using soybean (AAB58398), *Arabidopsis* (BAB02581), rice (XP468764), cinnamon (AAR84297), poplar (AAF04455 and AAR32717), and taro (AA062313). The corresponding nucleotide sequences from soybean (U53418), *Arabidopsis* (AP001309), cinnamon (AY496079), poplar (AF0539973 and AY466400), taro (AY222335), rice (AK103919), wheat (BT009444) and maize (AY103689) were used to construct the degenerate primers flanking the *E. grandis* UGDH ORF. The forward (5'ATGGTGAAGATHHTGYTGYATY3') and reverse (5'TTADGCVAYVGCRRGGCATGTC3') primers were forced to flank the complete UGDH open reading frame. The alignment of the 21 nucleotides from the 5' and 3' ends of the nine sequences showed that the differences among them were mainly in the third base of the codon, not changing the amino acid, with few exceptions (Figure 1). These minor differences were overcome by the use of degenerate primers. Primer sequences are represented in standard IUB/IUPAC amino acid and nucleic acid codes.

Cloning and sequence analysis of the UGDH cDNA

Total RNA was isolated from the cambial region of 6-month-old saplings of *E. grandis* using the method of Salzman *et al.* (1999), and poly(A) mRNA was purified from 75 µg of total RNA, using the Dynabeads mRNA Purification kit (Dyna) as specified by the manufacturer, and eluted in 20 µL Tris-HCl 10 mM.

A UGDH encoding cDNA was obtained by RT-PCR, using the SuperScript One-step RT-PCR with Platinum Taq (Invitrogen), and the specific flanking primers (UGDH_ Forward: 5'CACCATGGTGAAGATHHTGYTGYATY3'; UGDH_Reverse: 5'TTADGCVAYVGCRRGGCATGTC3') for the corresponding open reading frame. The directional sequence CACC was added to the sense primer to facilitate the cloning orientation of the cDNA. For cDNA synthesis, a 50 µL RT-PCR reaction containing 25 µL reaction mix (0.4 mM of each dNTP, 2.4 mM of MgSO₄), 100 pg of mRNA, 0.2 µM of each primer, and 1 µL of RT/Platinum Taq mix was prepared. The cDNA was synthesized at 50 °C for 30 min and denatured at 94 °C for 2 min, and amplification was performed using 38 cycles of 15 s at 94 °C, 30 s at 60 °C, 1 min and 30 s at 72 °C, followed by 7 min at 72 °C. The 1443-bp blunt-end amplification product was cloned into the pENTR-Directional-TOPO[®] Cloning Vector, for entry into the Gateway System (Invitrogen), according to the manufacturer's instructions. A UGDH recombinant clone, selected in kanamycin

(50 $\mu\text{g mL}^{-1}$), was purified and screened by PCR for the presence of full-length *UGDH* cDNA, using degenerate primers and also by sequencing in both directions with the universal M13 primers and the BigDye Terminator Cycle sequencing kit (Perkin Elmer), to ensure the proper reading frame. The samples were loaded onto an ABI 3100 sequencer (Perkin Elmer). The complete nt sequence of the *E. grandis UGDH* was deposited in the NCBI GenBank under accession n° EF179384.

Expression of UGDH protein in *E. coli*

The RT-PCR product corresponding to the *E. grandis UGDH* cDNA was sub-cloned into the Gateway® pDEST17 vector (Invitrogen) in frame with a N-terminal 6x His tag, for expression of the recombinant protein, using the *E. coli* Expression System with Gateway® technology.

For the protein expression, the *UGDH* recombinant pDEST17 vector was introduced into chemically competent *E. coli* BL21-AI (Invitrogen) by heat shock. A recombinant clone, previously screened by PCR for the presence of the *UGDH* insert, was grown in LB medium containing ampicillin (100 $\mu\text{g mL}^{-1}$) at 37 °C and 200 rpm. This starter culture (1 mL, $\text{OD}_{600\text{nm}} = 0.7$) was used to inoculate 50 mL of LB medium supplemented with ampicillin (100 $\mu\text{g mL}^{-1}$) and left to grow at 37 °C (200 rpm) until an $\text{OD}_{600\text{nm}}$ of 0.4 was reached. After 3 h of induction with 0.2% (w/v) L-arabinose, the bacterial cells were harvested by centrifugation at 13,000 g for 1 min. The pellet was resuspended in 5 mL of lysis buffer (100 mM NaH_2PO_4 , pH 8.0; 10 mM Tris, containing 0.1% (w/v) lysozyme, 1 mM PMSF, 10 mM β -mercaptoethanol, 10 $\mu\text{g RNaseA mL}^{-1}$, 5 $\mu\text{g DNase mL}^{-1}$) and sonicated (three times of 10-second bursts at medium intensity). The resulting lysate was separated into soluble and insoluble fractions by centrifugation (3,000 g) for 15 min.

Purification of the recombinant protein

The histidine fusion protein was purified from the soluble fraction using a purification column prepared with ProBondNickel-Chelating Resin (Invitrogen), according to the manufacturer's instructions. The purity and M_r of the proteins in the eluted fractions collected were analyzed by 12.5% (w/v) SDS-PAGE, using Coomassie brilliant blue G-250 staining (Candiano *et al.*, 2004).

Protein concentration was determined using the Bradford assay (Bradford, 1976) and bovine serum albumin as a standard with protein dye reagent (Bio-Rad), following the supplier's instructions.

Enzyme assay

The activity of the purified Poly-His-Tagged eucalyptus UDP-GDH was measured at 30 °C (Stewart and Copeland, 1999) in a continuous assay, by monitoring, with a Hitachi U3300 spectrophotometer, the increase in absorbance at 340 nm due to the UDP-Glc-dependent formation

of NADH. Reaction mixtures for the standard assay contained, in a final volume of 1 mL, 20 mM Tris-HCL, pH = 8.0, and 2 mM MgCl_2 . The kinetic data for NAD^+ or UDP-Glc as substrates were determined in triplicate assays, keeping one of the substrates at saturating concentrations. The assay was initiated by the addition of UDP-Glc, and activity was calculated from linear initial reaction rates, based on the assumption that 2 mol of NADH were formed per mole of oxidized UDP-Glc.

Protein extraction for the two-dimensional gel electrophoresis

Total proteins from the cambial region of 3- and 22-year-old *E. grandis* trees were extracted as described by Hurkman and Tanaka (1986), with few modifications introduced by Celedon *et al.* (2007).

Two-dimensional gel electrophoresis

Protein samples (containing 2 μg of protein μL^{-1}) were applied onto a 4-7 linear immobilized pH gradient strip (18 cm, GE HealthCare) and focalized using the IPGphor apparatus (GE HealthCare). Strips were rehydrated for 12 h at 20 °C and 50 V. The proteins were prefocused at 100 V for 1 h and then at 200 V for 1 h, 400 V for 1 h, 700 V for 1 h, 1000 V for 1 h, and finally focused for a total of 70 KVh. After isoelectric focusing (IEF), the strips were kept at -80 °C until needed.

For the second-dimension analysis, the strips were kept at room temperature for 15 min in equilibration buffer [6 M urea, 2% (w/v) SDS, 50 mM Tris-HCl pH 6.8, 30% (v/v) glycerol] containing 1% (w/v) DTT, followed by incubation in the same buffer added with 2.5% (w/v) iodoacetamide and 0.001% bromophenol blue. The second-dimension electrophoresis was performed in 12% (w/v) polyacrylamide gels at 30 mA, until the dye reached the bottom of the gel. Three replicates were performed for each sample. Proteins were detected using Coomassie Brilliant Blue G-250 (Candiano *et al.*, 2004), with modifications. Gels were incubated twice for 1 h each in a solution containing 3% (v/v) phosphoric acid and 50% (v/v) ethanol, and once for 1 h in 2% (v/v) phosphoric acid. Protein detection was done after leaving the gels overnight in staining solution [17% (v/v) methanol, 15% (w/v) ammonium sulfate, 2% (v/v) phosphoric acid plus 0.1% (w/v) dye], followed by 3 washes in water (10 min each). The gels were then stored in 15% (w/v) ammonium sulfate for image analysis and spot selection.

In-gel protein digestion

Protein spots were excised from SDS-PAGE gels, cut into 1 mm cubes and washed with water for 15 min. For destaining, the gel pieces were washed several times with a 50% (v/v) acetonitrile (ACN) solution containing 50 mM ammonium bicarbonate, until complete removal of the Coomassie (G250) stain. The 2-DE gel spots were com-

```

1 atggtgaagatctgctgcatggctgctgctatgtoggggggctactatggccgtgatt
M V K I C C I G A G Y V G G P T M A V I 20
61 gctctcaagtgccctcagtagaagctgggctgcttgatattctctctcgcatacaa
A L K C F S V E V A V D I S V S R I Q 40
121 gctgggaacaggaacagctccctcctatgaaccaggccttgatcgpgtgggaagcaa
A W N S E Q L P I Y E F G L D A V V K Q 60
181 tgcggaggaagaacctctctcagcaccgatgggagaagcatgtcttggagctgac
C R G K N L F F S T V E K H V F E A D 80
241 attgtctctgctgctcaacccttaccagactcogagcttgggtgcaaggcaagct
I V F V S V N T [E] T K T R G L G A O K A 100
301 gcagatctgacctactgggaagtgacagccctatgatgctgctgcttccaaatctgat
A D L T Y W E S A A R N I A D V S K S D 120
361 aaaatctgttggagaatcgactctccggctaaagactcogcagggcaatagagaagatt
K I V V E K S T V F V K T A E A I E K I 140
421 ctgaccataaacagcaagggaatcaaatccagatctctgcaaccagagttccctgct
L T H N S K G I K F Q I L S N [E] E F L A 160
481 gagggaacggcaatcccaagatctcttctgctcagatctgctccttattggggcaggaa
E G T A I Q D L F A F D R V L I G O R E 180
541 actccagaagggcagaagccatccaaacctgaaggagctataatgctcattgggtccc
T F E G Q K A I Q T L K D V Y A H W F 200
601 gaagacccatcttaaacaccaccttggctgcagagctctcctaaatgggctgcaat
E A G R I L T N L W A E L S K L A A N 220
661 gcttcttggccacaaggaatctctctgctggaatgcccctgctccttctgcaaggcact
A F L A Q R I S S V N A M S L Q E R T 240
721 ggtgagatctgtcccaagctctctctgctgggtgaaggactcagcaattggctcccaag
G A D V A Q V S Y A V G K D S R I O P K 260
781 ttcctcaatgctgctgctgctgctggctcctgctcctcaaaagatattctcaactcg
F L N A S V G F G G S [C] F Q K D I L N L 280
841 gtgtacatctgtgagtgcaatggcctcctgaagtggctgactgctgcaaacagctgatt
V Y I C E C N G L P E V A E Y W K Q V I 300
901 aagatcaacgactatcagaaaggcccttggcaaccgctgctgctctctatgtctcaac
K I N D Y Q K A R F V N R V V S S M F N 320
961 acagcttcaaacagaagatcgcactctctggatcttgcccttaagaagatcacggggac
T V S N K K I A I L F A F K K D T G D 340
1021 acgctggagactcccgcaatctgactgcaaggctctcttgggagacaaggccagctca
T R E T P A I D V C K G L L G D K A R L 360
1081 agcatttatgctcctcaggtgactgaggagcagatccagagggacccttaccatgaaacag
S I Y D F Q V T E E Q I Q R V D L T M K F 380
1141 ttgatctgggaccaccagctccactcagccgatgagctcccaccctgtggaagcagctt
F D W H S F T H S T K W T G 400
1201 tctgactctgggagcctatgatgctgctcaaggatgcccctggcctctgcatcctcact
S V V W D A Y D A V K D A H G L C I L T 420
1261 gaggggatgagctcaagactctctgatcaccagagatatacagacaacatgcaagccog
E W D E F K T L D Y Q R I Y D N H Q K P 440
1321 gcttacatctttgatggcaggaagctctgcaagcttaacaagctgctgagatcggatcc
E Y I F D G B R N V V N K L R E I G F 460
1381 attgtctacttggcagccctgctgctcagggctcaaggacatgcccggctggca
I V Y S I G K C P E L D P W L K D M F A V A 480
1441 taa 1443

```

Figure 1 - Nucleotide (nt) and derived amino acid (aa) sequence of the cloned cDNA of the *E. grandis* *UGDH* gene. Underlined are the translation initiation nt motif (aa# 1 and 2), the NAD cofactor binding site (aa# 8-14) and the catalytic site (aa# 267-278, with a Cys residue, boxed). A putative glycosylation site (aa# 263-265) is underlined with a broken line. Pro residues at 89 and 156 are also boxed. The conserved N and C terminal regions (highlighted in dark grey) were used for the alignment of the 21 nucleotides from the 5' and 3'-ends (highlighted in light grey) of the cloned cDNA against poplar (AF053973 and AY466400), soybean (U53418), *Arabidopsis* (AP001309), taro (AY222335), cinnamon (AY496079), rice (AK103919), wheat (BT009444) and maize (AY103689), to produce the degenerate primers. The complete nt-sequence of *E. grandis* *UGDH* was deposited in the NCBI GenBank under accession n° EF179384.

pletely dehydrated with 100% (v/v) ACN for 10 min, rehydrated with 50 mM ammonium bicarbonate plus 20 mM dithiothreitol (DTT), and maintained for 40 min at 60 °C. This solution was discarded and replaced by 50 mM ammonium bicarbonate with 55 mM iodoacetamide, before keeping the tubes in darkness for 30 min. Then the gel pieces were dehydrated again with 100% (v/v) ACN and let to air-dry for complete removal of the solvent. Protein digestion was carried out with a solution of 10 ng µL⁻¹ trypsin (Promega) in 25 mM ammonium bicarbonate, for 12 h at 37 °C. Gel pieces were extracted twice with 50 µL of 60% (v/v) ACN containing 1% (v/v) formic acid (FA), and once with 50 µL of 100% ACN. All supernatants were combined and vacuum-dried. The peptides were then resuspended in 12 µL 1% (v/v) FA for MS analysis.

Protein identification by LC-MS/MS

The peptides were separated and identified by on-line chromatography, using a Cap-LC coupled to a Q-TOF Ultima API mass spectrometer (Waters, UK). 5 mL of sample were loaded onto a 0.18 mm x 23.5 mm NanoEase Trapping Column (Waters, UK) for pre-concentration and desalination, followed by peptide separation on a C18, 3.5 µm, 75 µm x 100 mm NanoEase Symmetry 300 LC column (Waters, UK). Peptides were eluted in a 60 min linear gradient of solvent B [95% (v/v) ACN containing 0.1% (v/v) formic acid in water] and solvent A [5% (v/v) ACN containing 0.1% (v/v) formic acid in water], at a flow rate of 250 nL min⁻¹.

MS spectra analysis

All analyses were performed using a positive ion mode at a 3 kV needle voltage. The mass range was set from 300 to 2000 m/z, and the MS/MS spectra acquired for the most intense peaks (≥ 15 counts). Multiple charged precursor ions were selected for fragmentation and peptide sequencing, using automated data-dependent-acquisition (DDA) MassLynx software (Waters), switching from the MS to the MS/MS mode and then returning to MS. The resulting fragmented spectra were processed using the ProteinLynx v. 4.0 software (Waters), and the MASCOT MS/MS Ion Search was used to blast the sequences against the SwissProt and NCBI databank. Combined MS-MS/MS searches were conducted with parent ion mass tolerance at 50 ppm, MS/MS mass tolerance of 0.1 Da, carbamidomethylation of cysteine (fixed modification) and methionine oxidation (variable modification). According to MASCOT probability analysis, only significant hits (p < 0.05) were accepted.

Southern blot analysis

Genomic DNA (15 µg) was isolated from *Eucalyptus grandis* leaves, as described by Doyle and Doyle (1987), and digested with *Hind*III, *Eco*RI, *Bam*HI, *Eco*RV, *Nco*I and *Sac*I restriction enzymes at 37 °C for 16 h. The DNA fragments were separated on 1% (w/v) agarose gel in 0.5 x TBE buffer (45 mM Tris-borate, 1 mM EDTA, pH 8.0), stained with ethidium bromide, visualized under ultraviolet light, and denatured before being transferred onto a Hybond N⁺ nylon membrane (GE HealthCare) by capillary transfer (Southern, 1975; Sambrook *et al.*, 1989). The membrane was probed using the AlkPhos Direct Labeling Module (GE HealthCare). The alkaline-phosphatase-labeled-PCR product (698-bp) resulting from the amplification of a conserved internal region (nt# 188-886, Figure 1) of the eucalyptus cDNA (1443-bp) cloned with the following primers (*ugdhdForward*: 5'-GAAGAACCTCTTCTTCAGCA-3' *ugdhdReverse*: 5'-AGTACTCAGCCACTTCAGGA-3') synthesized for the UGDH domain (pfam03721, pfam00984, pfam03720), used as probe, was

detected with the Gene Images CDP-Star Detection Module (GE HealthCare). Hybridization was carried out at 60 °C for 16 h, followed by stringent washes: twice in fresh primary wash buffer [2 M urea, 50 mM NaH₂PO₄, pH 7.0, 150 mM NaCl, 1 mM MgCl₂ containing 0.1% (w/v) SDS and 0.2% (w/v) blocking reagent] at 55 °C for 10 min, and once in secondary wash buffer [50 mM Tris, 100 mM NaCl containing 2 mM MgCl₂] at room temperature for 5 min. The hybridization signal was recorded on X-ray film (MXG/PLUS, Kodak) after 1 h exposure between intensifying screens (GE HealthCare), at room temperature.

The resulting exposed X-ray film was used as reference for localization of the bands.

Semi-quantitative real-time PCR (RT-PCR)

Tissue samples from 6-month-old saplings and from 3- and 22-year-old plants were used for total RNA isolation, as described by Salzman *et al.* (1999). Poly(A) mRNAs were purified from 75 µg of total RNA, using the Dynabeads mRNA Purification kit (Dyna) as specified by the manufacturer, and eluted in 20 µL Tris-HCl 10 mM. First-strand cDNAs were generated using the SuperScript III First-Strand Synthesis SuperMix in a 20 µL reaction mixture, containing 3/10 of the eluted mRNA, 50 ng random hexamers, and 1 µL annealing buffer. After 5 min of incubation at 65 °C, 10 µL of 2x First-Strand Reaction Mix plus 2 µL SuperScript III/RNase OUT Enzyme Mix, which includes the SuperScript III Reverse Transcriptase, were added, and the final reaction mixture was incubated for 10 min at 25 °C, followed by 50 min at 50 °C, and 5 min at 85 °C, following the manufacturer's instructions (Invitrogen). Semi-quantitative RT-PCR assays were performed, using 1/10 of the cDNA preparation per PCR, based on a preliminary semi-quantitative RT-PCR (data not shown). Primer sequences and amplicon sizes were the following: *UGDH* (247-bp; primers designed specifically for the cloned eucalyptus cDNA forward 5'-GCCCGTATGATTGCTGATGTC-3', reverse 5'-TCAAGGTTTGGATGGCTTTC-3'); *ubiquitin* (223-bp, primers to the *E. grandis* constitutive ubiquitin gene, forward 5'-CGATTGATTCTCAGCAAGC-3', reverse 5'-GGATGTTGTAGTCAGCCAAGG-3'). Amplification specificity was checked by melting-curve analysis, and PCR efficiency determined using standard curves for each primer pair constructed with serial dilutions (1:10, 1:100 and 1:1000) of the cDNA preparation.

Results and Discussion

PCR amplification of the full-length cDNA encoding *E. grandis* UGDH

In this study, the degenerate primers designed to anneal highly conserved 5' and 3' regions of already known genes, in combination with the RT-PCR technique, proved to be a potential strategy to overcome the difficulties in

cloning a particular novel gene, when cDNA or a genomic library are not available. The RT-PCR reaction produced a 1443-nt-long cDNA sequence, with an open reading frame of 480 amino acids encoding UGDH (Figure 1).

Nucleotide and deduced amino acid sequence analysis

The deduced amino acid sequence for eucalyptus shows conserved motifs (pfam03721, pfam00984, pfam03720). Figure 1 (underlined) presents the translation initiation nt motif (aa# 1 and 2), the NAD cofactor binding site (aa# 8-14), the catalytic site (aa# 267-278, with a Cys residue, boxed). A putative glycosylation site (aa# 263-265), also found in poplar (Johansson *et al.*, 2002), is indicated with a broken line, and Pro residues at aa# 89 and 156 are also boxed, representing the main chain bends in the protein structure (Hempel *et al.*, 1994). The derived amino acid sequence from the database search BLASTX analysis (Altschul *et al.*, 1997) showed a high degree of identity to the UGDH of soybean (92%), cinnamon (90%), *Arabidopsis* (90%), taro (89%), poplar (89%), and rice (85%), among several plant species, including some putatives UDP-glucose dehydrogenases, such as that of tobacco (91%). The high identity among UGDHs of higher plants suggests strict structural requirements for proper functioning of the protein. It is interesting to observe the high degree of amino acid sequence identity of the eucalyptus UGDH, even along the N and C terminal regions (Figure 1, highlighted in dark grey), where the differences in the first 25 aa, were less than 4% and 16%, respectively, compared to the leguminosa soybean and gramineae such as rice, maize and, particularly, wheat. The analysis of the phylogenetic tree, generated using the CLC Main Work Bench 5.5 (UPGMA algorithm with Bootstrap analysis, 100 replicates) for the eucalyptus UGDH protein sequence and eight other plant species, shows a high similarity of the *E. grandis* protein sequence to various plant sequences, particularly soybean and *Arabidopsis* (Figure S1, supplementary data).

Expression of recombinant eucalyptus UGDH

The recombinant UGDH has been found to be often present in protein inclusion bodies (Tenhaken and Thulke, 1996; Hinterberg *et al.*, 2002). It is known that eukaryotic proteins expressed in *E. coli* often form protein inclusion bodies, due to differences in the protein-folding systems between prokaryotes and eukaryotes (Oka and Jigami, 2006). Therefore, polyhistidine recombinant eucalyptus UGDH protein was produced in BL21-AI cells, and purified under a narrow range of optimized conditions, which included a low centrifugation of the lysate to recover, under native conditions, enough enzymatically active protein in the soluble fraction. The protein was purified at 4 °C, and immediately used for the enzyme assay.

The Coomassie-stained SDS-PAGE gel illustrates the L-arabinose induction of the UGDH recombinant protein in

E. coli (Figure 2). The identification of the protein was confirmed by LC-MS/MS, and the peptides sequenced covered 23.33% of the total protein, with an estimated molecular weight of 52.91 kDa, and a 100% probability that the peptides are related to the expected protein (Table S1, supplementary data).

The activity of the purified Poly-His-Tagged-UGDH measured at 30 °C (Stewart and Copeland, 1999), in a continuous assay, by spectrophotometrically monitoring the absorbance increase at 340 nm due to the UDP-Glc-dependent formation of NADH, showed that the recombinant protein expressed in *E. coli* catalyzed the conversion of UDP-Glc to UDP-GlcA (Figure 3), thus confirming that the cloned cDNA encodes the UGDH enzyme.

Enzyme kinetics

The recombinant protein expressed in *E. coli* catalyzed the conversion of UDP-Glc to UDP-GlcA, confirming that the cloned cDNA encodes the UGDH enzyme (Figure 3). The kinetic data were then fitted to the Michaelis-Menten equation to obtain the K_m and V_{max} values (Segel, 1976). For both substrates, UDP-Glc and NAD^+ , hyperbolic saturation curves were observed and the K_m values calculated, using the Origin-software[®], $60.7 \pm 8.5 \mu M$ for UDP-Glc ($V_{max} = 67.9 \pm 9.2 \mu mol \text{ min}^{-1} \text{ mg protein}^{-1}$) and $67.3 \pm 17.9 \mu M$ for NAD^+ ($V_{max} = 171.8 \pm 34.8 \mu mol \text{ min}^{-1} \text{ mg protein}^{-1}$) (Figures 3A and 3B, respectively). The estimated K_m of the eucalyptus enzyme for NAD^+ was similar to the value described for the recombinant protein from soybean ($70 \pm 5 \mu M$), whereas the K_m for UDP-Glc was around 3 times higher ($22 \pm 2 \mu M$) (Hinterberg *et al.*, 2002). A characterization of three isoforms of UGDH (UGDH2, UGDH3 and UGDH4) from *Arabidopsis* evidenced a similar affinity value for the cofactor NAD^+ as that observed for eucalyptus (40–45 μM), while major differences were observed in the K_m for UDP-Glc: $123 \pm 9 \mu M$, $335 \pm 16 \mu M$ and $171 \pm 9 \mu M$, respectively (Klinghammer and Tenhaken, 2007). The high affinity of the enzyme for NAD^+ was also observed in eucalyptus, confirming previous observations that the UGDHs are not limited by NAD^+ levels *in vivo*, while UDP-Glc affinity might be isoform-dependent (Hinterberg *et al.*, 2002; Klinghammer and Tenhaken, 2007).

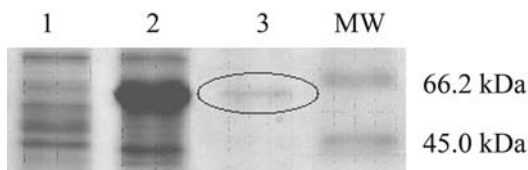


Figure 2 - Coomassie-stained SDS-PAGE of recombinant UGDH protein. The eucalyptus *UGDH* cDNA cloned into a His-tagged pDEST expression vector was used to transform the BL21 *E. coli* cells. (1) Extracts from uninduced cells; (2) L-arabinose-induced cells; and (3) the purified recombinant eucalyptus UGDH from the L-arabinose-induced *E. coli* culture. MW: Molecular weight marker (Sigma).

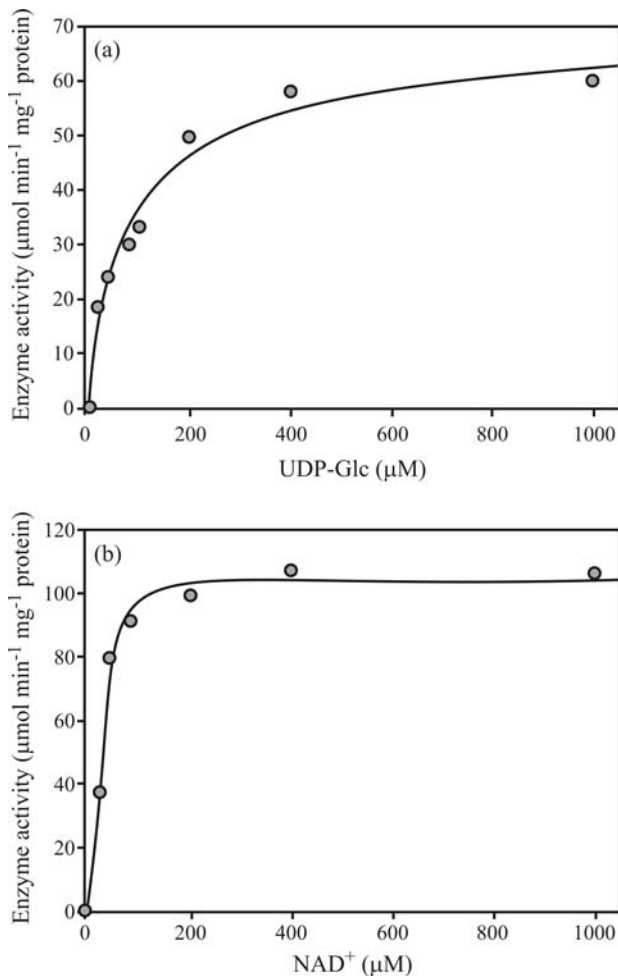


Figure 3 - Kinetic data of the eucalyptus UDP-GlucDH. The steady-state enzyme activity was measured as conversion of NAD^+ to NADH detected by the absorbance at 340 nm. Purified eucalyptus UDP-GlucDH was incubated for 1 min with increasing concentrations of NAD^+ (0–1 mM) in the presence of saturating substrate (A, UDP-Gluc, 1 mM), or with increasing concentrations of UDP-Gluc (0–1 mM) in the presence of saturating cofactor (B, NAD^+ , 1 mM).

Genomic copies of *UGDH*

Southern blot analysis indicated the existence of a least two copies of the *UGDH* gene within the eucalyptus genome (Figure 4). The existence of a single-copy *UGDH* gene has been described for soybean (Tenhaken and Thulke, 1996) and *Arabidopsis* (Tenhaken and Thulke, 1996; Seitz *et al.*, 2000). In poplar (Johansson *et al.*, 2002), two highly homologous genes were found, as also observed in maize (Kärkönen *et al.* 2005). However, a later evaluation of the *Arabidopsis* and poplar genome projects indicated multiple copies. The presence of four UGDH isoforms in *Arabidopsis*, highly similar to each other, reported by Reiter and Vanzin (2001), was confirmed by Klinghammer and Tenhaken (2007), who found a fifth *UGDH* pseudogene, located on chromosome III, with a weaker similarity to the others. In the *Populus trichocarpa* genome (*Populus* genome project), four *UGDH* genes were identified, one in

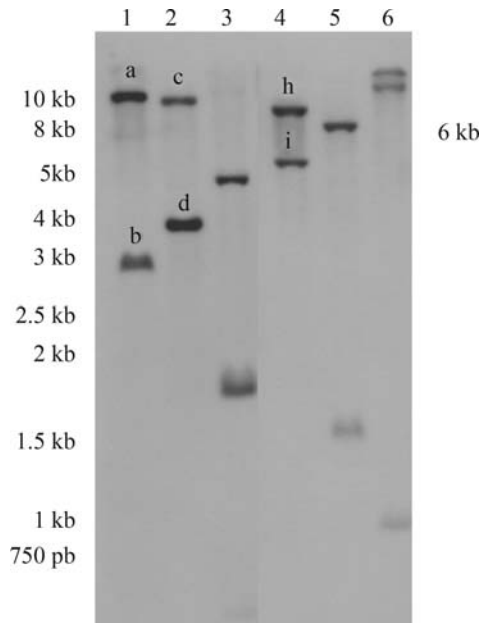


Figure 4 - Genomic Southern blot of *E. grandis* DNA, restricted with *Hind*III (1), *Eco*RI (2), *Bam*HI (3), *Eco*RV (4), *Nco*I (5), and *Sac*I (6). The sizes of the fragments are indicated in kilobase pairs (kb). The alkaline-phosphatase-labeled PCR product (698-bp) resulting from the amplification of a conserved internal region (nt# 188-886, Figure 1) of the cloned *E. grandis* cDNA (1443-bp) was used as probe.

linkage group IV (224080232), one in linkage group VIII (224098951), one in linkage group X (224112137), and one in linkage group XVII (224141486).

We further analyzed the possible number of copies for *UGDH* in the *Eucalyptus* genome by performing a *tblastx* analysis of the eucalyptus ESTs sequences available in the GenBank database. The results showed that the cloned *UGDH* sequence from *E. grandis* is highly similar (~99%) to the amino acid sequences of seven ESTs produced from the differentiating xylem of *Eucalyptus gunnii* (103476385, 103475842, 103479156, 103478292, 103476672, 103476952, 103480206) and five *Eucalyptus globulus* EST sequences, using mRNA isolated from leaf tissue under low temperature conditions (162324959, 162326130, 162329222, 162329166, 162326954). The phylogenetic tree generated using CLC Main Work Bench 5.5 (UPGMA algorithm with Bootstrap analysis, 100 replicates) for the complete nucleotide sequence of the cloned *E. grandis* *UGDH* mRNA (144926038) with these seven cambial *E. gunnii* EST and the five leaf *E. globulus* sequences (Figures 5A and 5B, respectively), suggests that there are at least two groups of nucleotide sequences, most probably indicating two isoforms of this gene in the RNA extracted from the cambial and leaf tissues.

The amplification product obtained from the eucalyptus genomic DNA against the primers flanking the *UGDH* open reading frame, corresponds exactly to the same size (1443-bp) of the eucalyptus *UGDH* cDNA (Figure 6), showing the lack of introns along the cloned *UGDH* gene

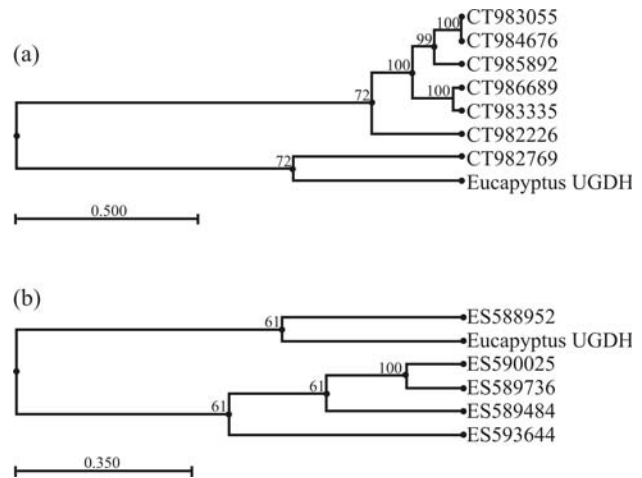


Figure 5 - Phylogenetic tree generated using CLC Main Work Bench 5.5 (UPGMA algorithm with Bootstrap analysis, 100 replicates) for the complete nucleotide sequence of the cloned *E. grandis* *UGDH* mRNA (144926038): A) using seven *E. gunnii* EST sequences (103476385, 103475842, 103479156, 103478292, 103476672, 103476952, 103480206) of mRNA isolated from differentiating xylem; B) using five *E. globulus* EST sequences (162324959, 162326130, 162329222, 162329166, 162326954) of mRNA isolated under low temperature conditions from leaf tissue.

sequence. The restriction map of the *UGDH* ORF (Figure S2, supplementary data) shows that the restriction enzymes *Eco*RI, *Hind*III and *Eco*RV do not cut the cloned fragment internally. This indicates that the bands observed in the Southern blot assay (Figure 4, lanes 1, 2 and 4), which are all larger than 1.6 kb, most probably contain the complete cloned *UGDH* ORF.

Moreover, in order to investigate the existence of *UGDH* protein isoforms, we produced 2DE-PAGE gels with proteins isolated from the cambial region of 3- and 22-year-old *E. grandis* trees. Figure 7 shows the identification of three distinct spots (1, 2 and 3) with the peptide data used to identify the proteins shown in Table S2 (supplementary data). Spot 3 showed a M_r of ~53.0 kDa and a pI of approximately 6.5, which is in accordance with the calculated values for the M_r and pI of the cloned sequence (53.09 kDa and 6.52, respectively). These values are similar to those obtained for soybean (52.94 kDa and 5.81) and poplar (52.99 kDa and 6.20). Thus, spots 1 and 2 represent proteins with different characteristics (lower pI and lower M_r), indicating the existence of at least one more copy of *UGDH* in the *E. grandis* genome. The differences between spots 1 and 2 might be explained by posttranslational modification, such as phosphorylation, although further analysis is needed to confirm this hypothesis.

***UGDH* expression in different tissues of 6-month-old saplings and in the cambial region of juvenile and mature *E. grandis* wood**

The expression of *UGDH* in different tissues of 6-month-old saplings and in the cambial region of 3- and 22-

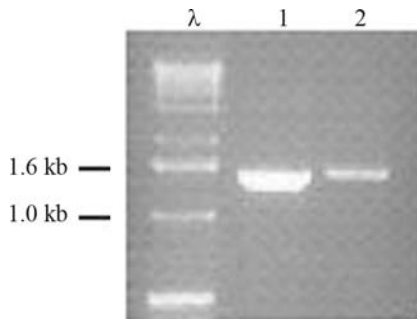


Figure 6 - Amplification product (1443-bp) of the eucalyptus genomic DNA (1), and the corresponding cDNA (2) for eucalyptus *UGDH*, against the degenerate primers flanking the *UGDH* open reading frame. λ : 1 kb DNA ladder (Invitrogen), kb: kilobase pair.

year-old trees was analyzed by semi-quantitative RT-PCR. In the 6-month-old saplings the maximum expression was observed in roots, whereas in stem and bark it was intermediate, and the leaves displayed the lowest expression (Figure 8). Similar results were observed in tobacco plants by Bindschedler *et al.* (2005), who compared the expression of *UGDH* and *NtADH2* (the dual specific *UGDH* similar to *ADH*), suggesting that the isoforms of *UGDH* might be preferentially involved in primary rather than in secondary growth, since they were strongly expressed in roots and in the youngest internodes of the stem, with the lowest level of expression observed in buds and in tobacco leaves, similar to the expression pattern observed by Seitz *et al.* (2000) in *Arabidopsis*.

We then checked the role of *UGDH* in the wood-forming tissues of juvenile and mature trees, by comparing the level of expression relative to the stem of 6-month-old saplings. Figure 9 shows that the expression of *UGDH* in the cambial region of 3- and 22-year-old trees was significantly higher than in the stem of 6-month-old saplings, indicating that the enzyme may have an important role in controlling hemicellulose biosynthesis during wood formation.

In this study, the strategy of using degenerate primers in combination with the RT-PCR technique, proved to be

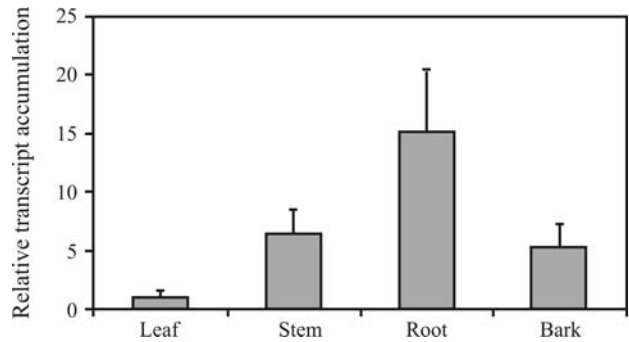


Figure 8 - Semi-quantitative RT-PCR of *UGDH* transcript accumulation in roots, stem, bark and leaves of 6-month-old *E. grandis* saplings. Transcript levels were normalized relative to the ubiquitin expression level as internal standard. Results are expressed as mean of three replicates and standard deviation relative to the leaf expression level, to which the value 1 on the linear scale was assigned.

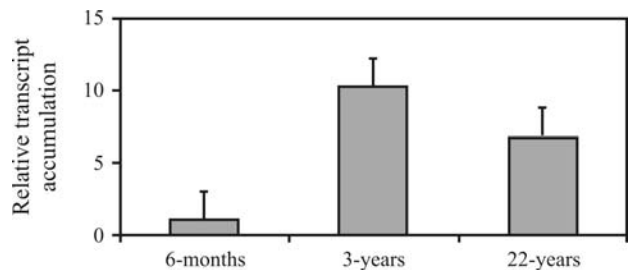


Figure 9 - Semi-quantitative RT-PCR of *UGDH* transcript accumulation of the cambial region from 3- and 22-year-old *E. grandis* trees. Transcript levels were normalized relative to the ubiquitin expression level as internal standard. Results are expressed as mean of three replicates and standard deviation relative to the stem expression level of the 6-month-old *E. grandis* saplings, to which the value 1 on the linear scale was assigned.

efficient to overcome the difficulties in cloning a particular novel gene, when cDNA or a genomic library are not available. We were able to clone and express the recombinant *UGDH* protein of *E. grandis* and to confirm its activity and sequence. The role of *UGDH* in wood-forming tissues is still poorly understood, and the isolation of *E. grandis* cDNA offers an opportunity to further understand the importance of this enzyme in trees.

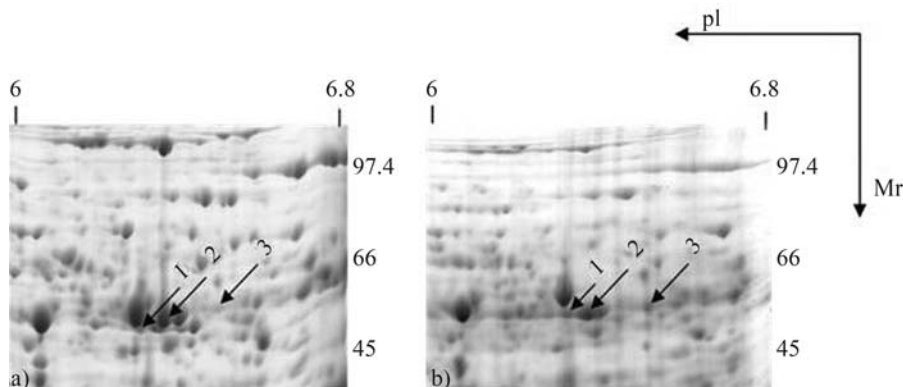


Figure 7 - 2-DE gels from the cambial region of 3-year-old (A) and 22-year-old (B) *E. grandis* trees. Arrows indicate the spot numbers of *UGDH* isoforms. Ip: Isoelectric-point, M_r : Molecular weight in kiloDaltons (kDa).

Acknowledgments

This work was financially supported by the Innovation Technology Program funded by FAPESP (01/11080-8) and Suzano Papel e Celulose S.A. Further support was provided by MEC-CAPEs with fellowships to Daniela Defávares do Nascimento and Gunta Gutmanis, CNPq with fellowship to Alexander de Andrade, and by FAPESP with fellowships to Ana Letícia Ferreira Bertolo, Luis Felipe Boaretto and Eduardo Leal de Oliveira Camargo. We would like to thank Dr. Siu Mui Tsai and Fabiana Cannavan from CENA/ESALQ/USP and Dr Maria Helena de Souza Goldman and Andrea Carla Quiapim from FFCLRP/USP, for their help with DNA sequencing. We would also like to thank Livia M. Franceschini for helping with the configuration of all figures.

References

- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller WE and Lipman DJ (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389-3402.
- Amino S, Takeuchi Y and Komamine A (1985) Changes in enzyme activities involved in formation and interconversion of UDP-sugars during the cell cycle in a synchronous culture of *Catharanthus roseus*. *Physiol Plant* 64:111-117.
- Bindschedler LV, Wheatley E, Gay E, Cole J, Cottage A and Bolwell GP (2005) Characterisation and expression of the pathway from UDP-glucose to UDP-xylose in differentiating tobacco tissue. *Plant Mol Biol* 57:285-301.
- Biswas BB, Ghosh B and Majumder AL (1984) myo-Inositol polyphosphates and their role in cellular metabolism: A proposed cycle involving glucose-6-phosphate and myo-inositol phosphates. In: Roodyn DB (ed) *Subcellular Biochemistry*. Plenum, London, pp 237-280.
- Bradford MM (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.
- Candiano G, Bruschi M, Musante L, Santucci L, Ghiggeri GM, Carnemolla B, Orecchia P, Zardi L and Righetti PG (2004) Blue silver: A very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis* 25:1327-1333.
- Carvalho MCG, Caldas DGG, Carneiro RT, Moon DH, Salvatierra GR, Franceschini LM, Andrade A, Celdon PAF, Oda S and Labate CA (2008) SAGE transcript profiling of the juvenile cambial region of *Eucalyptus grandis*. *Tree Physiol* 28:905-919.
- Celdon PAF, Andrade A, Meireles KGX, Carvalho MCG, Caldas DGG, Moon DH, Carneiro RT, Franceschini LM, Oda S and Labate CA (2007) Proteomic analysis of the cambial region in juvenile *Eucalyptus grandis* at three ages. *Proteomics* 7:2258-2274.
- Dalessandro G and Northcote DH (1977a) Possible control sites of polysaccharide synthesis during cell growth and wall expansion of pea seedlings (*Pisum sativum* L.). *Planta* 134:39-44.
- Dalessandro G and Northcote DH (1977b) Changes in enzymatic activities of nucleoside diphosphate sugar interconversions during differentiation of cambium to xylem in sycamore and poplar. *Biochem J* 162:267-279.
- Doyle JJT and Doyle JL (1987) Isolation of plant DNA from fresh tissue. *Focus* 12:13-15.
- Gibeaut DM and Carpita NC (1994) Biosynthesis of plant cell wall polysaccharides. *FASEB J* 8:904-915.
- Gibeaut DM (2000) Nucleotide sugars and glycosyltransferases for synthesis of cell wall matrix polysaccharides. *Plant Physiol Biochem* 38:69-80.
- Hempel J, Perozich J, Romovacek H, Hinich A, Kuo I and Feingold DS (1994) UDP-glucose dehydrogenase from bovine liver primary structure and relationship to other dehydrogenases. *Protein Sci* 3:1074-1080.
- Hinterberg B, Klos C and Tenhaken R (2002) Recombinant UDP-glucose dehydrogenase from soybean. *Plant Physiol Biochem* 40:1011-1017.
- Hurkman WJ and Tanaka CK (1986) Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. *Plant Physiol* 81:802-806.
- Johansson H, Sterky F, Amini B, Lundeberg J and Kleczkowski LA (2002) Molecular cloning and characterization of a cDNA encoding poplar UDP-glucose dehydrogenase, a key gene of hemicellulose/pectin formation. *Biochem Biophys Acta* 1576:53-58.
- Kärkönen A, Murigneux A, Martinant JP, Pepey E, Tatout C, Dudley BJ and Fry SC (2005) UDP-glucose dehydrogenases of maize: A role in cell wall pentose biosynthesis. *Biochem J* 391:409-415.
- Klinghammer M and Tenhaken R (2007) Genome-wide analysis of the UDP-glucose dehydrogenase gene family in Arabidopsis, a key enzyme for matrix polysaccharides in cell walls. *J Exp Bot* 58:3609-3621.
- Kunihiro S, Kawanishi Y, Sano M, Naito K, Matsuura Y, Tateno Y, Gojobori T, Yamagata Y, Abe K and Machida M (2002) A polymerase chain reaction-based method for cloning novel members of a gene family using a combination of degenerate and inhibitory primers. *Gene* 289:177-184.
- Loewus F, Chen MS and Loewus MW (1973) The myo-inositol oxidation pathway to cell wall polysaccharides. In: Lowes F (ed) *Biogenesis of Plant Cell Wall Polysaccharides*. Academic Press, New York, pp 1-27.
- Oka T and Jigami Y (2006) Reconstruction of *de novo* pathway for synthesis of UDP-glucuronic acid and UDP-xylose from intrinsic UDP-glucose in *Saccharomyces cerevisiae*. *FEBS J* 273:2645-2657.
- Reiter WD and Vanzin GF (2001) Molecular genetics of nucleotide sugar interconversion pathways in plants. *Plant Mol Biol* 47:95-113.
- Roberts RM and Loewus F (1973) The conversion of D-glucose-6-¹⁴C to cell wall polysaccharide material in *Zea mays* in presence of high endogenous levels of myoinositol. *Plant Physiol* 52:646-650.
- Robertson D, Beech I and Bolwell GP (1995) Regulation of the enzymes of UDP-sugar metabolism during differentiation of French bean. *Phytochemistry* 39:21-28.
- Sasaki K and Taylor IEP (1984) Specific labeling of cell wall polysaccharides with myo-[2-³H] inositol during germination and growth of *Phaseolus vulgaris* L. *Plant Cell Physiol* 25:989-997.
- Salzman RA, Fujita T, Zhu-Salzman K, Hasegawa PM and Versan RA (1999) An improved RNA isolation method for plant

- tissues containing high levels of phenolic compounds or carbohydrates. *Plant Mol Biol Rep* 17:11-17.
- Sambrook J, Fritsch EF and Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Segel IH (1976) *Biochemical Calculations*. 2nd edition. John Wiley & Sons, New York, 441 pp.
- Seifert GJ (2004) Nucleotide sugar interconversions and cell wall biosynthesis: How to bring the inside to the outside. *Curr Opin Plant Biol* 7:277-284.
- Seitz B, Klos C, Wurm M and Tenhaken R (2000) Matrix polysaccharide precursors in *Arabidopsis* cell walls are synthesized by alternative pathways with organ-specific expression patterns. *Plant J* 21:537-546.
- Stewart DC and Copeland L (1998) Uridine-5'-diphosphate dehydrogenase from soybean nodules. *Plant Physiol* 116:349-355.
- Stewart DC and Copeland L (1999) Kinetic properties of UDP-glucose dehydrogenase from soybean nodules. *Plant Sci* 147:119-125.
- Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* 98:503-517.
- Tenhaken R and Thulke O (1996) Cloning of an enzyme that synthesizes a key nucleotide-sugar precursor of hemicellulose biosynthesis from soybean: UDP-glucose dehydrogenase. *Plant Physiol* 112:1127-1134.
- Thompson JD, Higgins DG and Gibson TJ (1994) Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weigh matrix choice. *Nucleic Acids Res* 22:4673-4680.
- Zabackis E, Huang J, Muller B, Darvill AG and Albersheim P (1995) Characterization of the cell-wall polysaccharides of *Arabidopsis thaliana* leaves. *Plant Physiol* 107:1129-1138.

Internet Resources

- Mascot search engine for rapid protein identification using mass spectrometry data. MASCOT MS/MS Ion Search, <http://www.matrixscience.com/> (November 11, 2009).
- Multiple sequence alignment program ClustalW2, <http://www.ebi.ac.uk/Tools/clustalw2/index.html> (November 11, 2009).
- National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/> (November 11, 2009).
- Populus genome Project, http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html/ (November 11, 2009).

Supplementary Material

The following online material is available for this article:

Table S1 - Peptide sequencing data of the recombinant UGDH protein shown in Figure 2, obtained by LC-MS/MS.

Table S2 - Peptide sequencing data of UGDH proteins shown in Figure 7, identified by LC-MS/MS.

Figure S1 - Phylogenetic tree generated for the eucalyptus UGDH protein sequence and eight other plant species.

Figure S2 - Restriction map of the cloned *E. grandis* UGDH cDNA.

This material is available as part of the online article from <http://www.scielo.br/gmb>.

Associate Editor: Carlos F.M. Menck

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

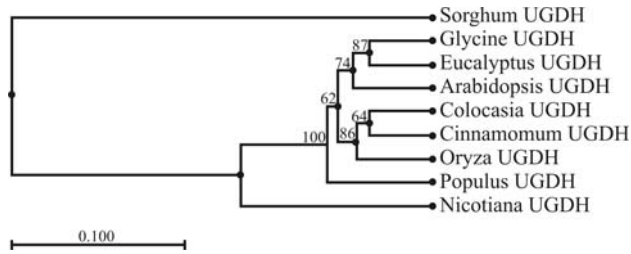


Figure S1 - Phylogenetic tree generated using CLC Main Work Bench 5.5 (UPGMA algorithm with Bootstrap analysis, 100 replicates) for the eucalyptus UGDH protein sequence and eight other plant species. The species and *gi* numbers are: *Eucalyptus grandis*, 144926039; *Cinnamomum osmophloeum*, 40317278; *Populus tomentosa*, 39939262; *Glycine Max*, 1518540; *Arabidopsis thaliana*, 11994517; *Sorghum bicolor*, 242047160; *Nicotiana tabacum*, 48093457; *Colocasia esculenta*, 29028306; *Oryza sativa Japonica*, Group 215737390.

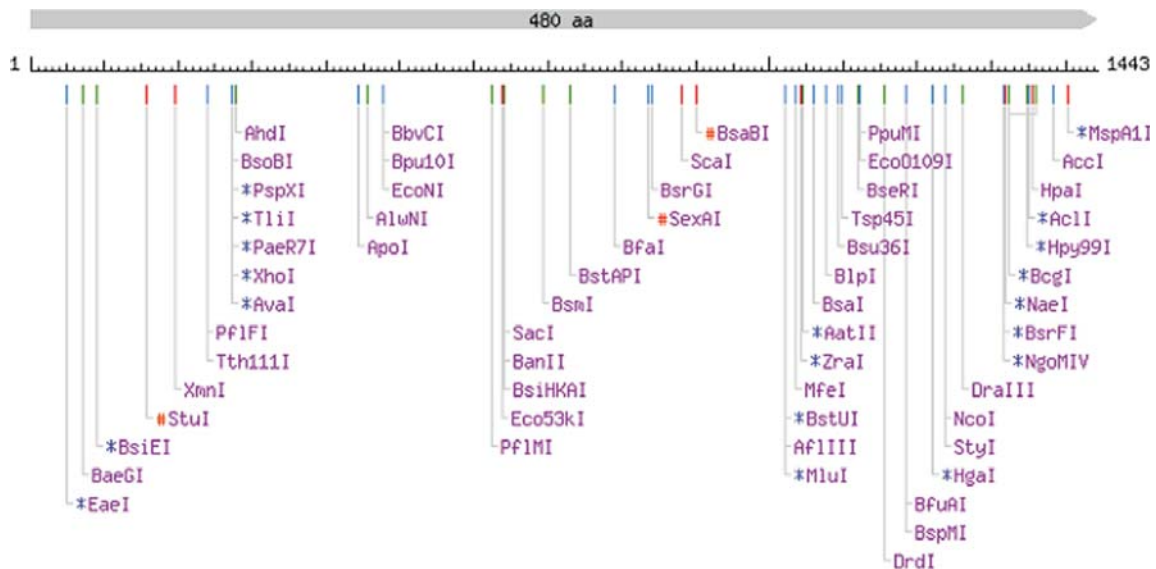


Figure S2 - Restriction map of the cloned *E. grandis* UGDH cDNA.

Table S1 - Peptide sequencing data of the recombinant UGDH protein of Figure 2, obtained by LC-MS/MS.

Name	% Probability	Peptide matches	Coverage (%)	Mr ¹⁾	pI ²⁾
EC 1.1.1.22					
AAB58398	100	12	23.33	52.90816	5.93

Submitted mass	Experimental mass	Peptide sequences
600.304	1198.592	(K)RAFFSTDVEK(H)
951.99	1901.964	(K)HVFEADIVFVSVSQPTK(T)
677.302	1352.588	(K)AADLTYWESAAR(M)
693.825	1385.634	(K)DVYAHWVPEDR(I)
738.401	1474.786	(R)ILTTNLWSAELSK(L)
752.393	1502.77	(R)ILTTNLWSAELSR(L)
537.79	1073.564	(K)LAANAFLAQR(I)
809.881	1617.746	(K)FLNASVGFGGSCFQK(D)
816.898	1631.78	(K)FLNASVGFGGSCFQK(D)
554.344	1106.672	(K)KIAILGFQK(K)
731.863	1461.71	(K)KVSVVWDAFEATK(D)
740.368	1478.72	(K)KVSVVWDAYDAVK(D)

¹⁾Molecular weight in kDa.

²⁾Isoelectric-point.

Table S2 - Peptide sequencing data of UGDH proteins of Figure 7, identified by LC-MS/MS.

Spot 1

Name EC 1.1.1.22	Probability (%)	Peptide matches	Coverage (%)	Theoretical ¹ Mr/pI ²	Experimental ¹ Mr/pI ²
<u>AAR32717</u>	100	4	9	52.963/5.92	57.907/6.39
Submitted mass	Charge	Experimental mass	Peptide sequences		
490.3026	2	490.3026	(K)IAILGFQFK(K)		
537.7942	2	1073.5738	(K)LAANAFLAQR(I)		
677.3171	2	1352.6196	(K)AADLTYWESAAR(T)		
738.3889	2	1474.7632	(R)ILTTNLWSAELSK(L)		

Spot 2

Name EC 1.1.1.22	Probability (%)	Peptide matches	Coverage (%)	Theoretical ¹ Mr/pI ²	Experimental ¹ Mr/pI ²
<u>AAO62313</u>	99.99	6	12	52.947/6.06	57.625/6.44
Submitted mass	Charge	Experimental mass	Peptide sequences		
382.2243	2	762.4340	(R)MIADVSK(S)		
398.2232	2	794.4318	(K)TLDYQR(I)		
537.8185	2	1073.6224	(K)LAANAFLAQR(I)		
677.3301	2	1352.6456	(K)AADLTYWESAAR(M)		
462.9116	3	1385.7130	(K)DVYAHWVPEDR(I)		
738.4020	2	1474.7894	(R)IITTNLWSAELSK(L)		

Spot 3

Name EC 1.1.1.22	Probability (%)	Peptide matches	Coverage (%)	Theoretical ¹ Mr/pI ²	Experimental ¹ Mr/pI ²
<u>AAR32717</u>	99.99	3	7	53.543/5.92	59.101/6.61
Submitted mass	Charge	Experimental mass	Peptide sequences		
490.3130	2	978.6114	(K)IAILGFQFK(K)		
677.3294	2	1352.6442	(K)AADLTYWESAAR(T)		
738.4080	2	1474.8014	(R)ILTTNLWSAELSK(L)		

¹Molecular weight in kDa.²Isoelectric-point.