Syringyl Lignin Is Unaltered by Severe Sinapyl Alcohol Dehydrogenase Suppression in Tobacco

Abdellah Barakatea, Jennifer Stephensa,b, Alison Goldiea, William N. Hunterc, David Marshalla, Robert D. Hancockb, Catherine Lapierred, Kris Morreelef, Wout Boerjane,f, and Claire Halpina,2

a Division of Plant Sciences, College of Life Sciences, University of Dundee at the James Hutton Institute, Invergowrie, Dundee DD2 5DA, United Kingdom
b James Hutton Institute, Invergowrie, Dundee DD2 5DA, United Kingdom
c Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Dundee DD1 5EH, United Kingdom
d Institut Jean-Pierre Bourgin, Institut National de la Recherche Agronomique-AgroParisTech, Unité Mixte de Recherche 1318, 78026 Versailles, France
e Department of Plant Systems Biology, Flanders Institute for Biotechnology, B–9052 Ghent, Belgium
f Department of Plant Biotechnology, Ghent University, B–9052 Ghent, Belgium

The manipulation of lignin could, in principle, facilitate efficient biofuel production from plant biomass. Despite intensive study of the lignin pathway, uncertainty exists about the enzyme catalyzing the last step in syringyl (S) monolignol biosynthesis, the reduction of sinapaldehyde to sinapyl alcohol. Traditional schemes of the pathway suggested that both guaiacyl (G) and S monolignols are produced by a single substrate-versatile enzyme, cinnamyl alcohol dehydrogenase (CAD). This was challenged by the discovery of a novel sinapyl alcohol dehydrogenase (SAD) that preferentially uses sinapaldehyde as a substrate and that was claimed to regulate S lignin biosynthesis in angiosperms. Consequently, most pathway schemes now show SAD (or SAD and CAD) at the sinapaldehyde reduction step, although functional evidence is lacking. We cloned SAD from tobacco (Nicotiana tabacum) and suppressed it in transgenic plants using RNA interference–inducing vectors. Characterization of lignin in the woody stems shows no change to content, composition, or structure, and S lignin is normal. By contrast, plants additionally suppressed in CAD have changes to lignin structure and S:G ratio and have increased sinapaldehyde in lignin, similar to plants suppressed in CAD alone. These data demonstrate that CAD, not SAD, is the enzyme responsible for S lignin biosynthesis in woody angiosperm xylem.

INTRODUCTION

The manipulation of lignin could, in principle, facilitate efficient biofuel production from plant biomass. Despite intensive study of the lignin pathway, uncertainty exists about the enzyme catalyzing the last step in syringyl (S) monolignol biosynthesis, the reduction of sinapaldehyde to sinapyl alcohol. Traditional schemes of the pathway suggested that both guaiacyl (G) and S monolignols are produced by a single substrate-versatile enzyme, cinnamyl alcohol dehydrogenase (CAD). This was challenged by the discovery of a novel sinapyl alcohol dehydrogenase (SAD) that preferentially uses sinapaldehyde as a substrate and that was claimed to regulate S lignin biosynthesis in angiosperms. Consequently, most pathway schemes now show SAD (or SAD and CAD) at the sinapaldehyde reduction step, although functional evidence is lacking. We cloned SAD from tobacco (Nicotiana tabacum) and suppressed it in transgenic plants using RNA interference–inducing vectors. Characterization of lignin in the woody stems shows no change to content, composition, or structure, and S lignin is normal. By contrast, plants additionally suppressed in CAD have changes to lignin structure and S:G ratio and have increased sinapaldehyde in lignin, similar to plants suppressed in CAD alone. These data demonstrate that CAD, not SAD, is the enzyme responsible for S lignin biosynthesis in woody angiosperm xylem.
metabolic systems in plants, of which the lignin biosynthetic grid is one of the best studied examples. In this article, we describe the existence, expression, and cloning of SAD in tobacco (Nicotiana tabacum), a true wood-producing angiosperm that closely parallels poplar (Populus spp) and aspen in the effects of genetic manipulations to lignin biosynthetic genes (Franke et al., 2000; O’Connell et al., 2002; Ralph et al., 2008a). We have used RNA interference to severely suppress SAD expression in tobacco and have used these plants to directly evaluate the proposed role of SAD in S lignin biosynthesis.

RESULTS

Tobacco Contains a SAD Similar to Aspen SAD

Tobacco is an ideal model woody plant for rapid determination of the roles of various genes in lignification using reverse genetics strategies. To ultimately use such a strategy to illuminate the involvement of SAD in lignin biosynthesis in angiosperms, we first set out to investigate the existence of a SAD in lignifying tissue in tobacco using an antiserum raised against aspen SAD. This antiserum was previously shown to recognize SAD in a wide range of angiosperm species (Li et al., 2001). Preliminary studies demonstrated that the aspen SAD antiserum also recognized a specific protein band in tobacco stem extracts (see Supplemental Figure 1 online). To determine whether the tobacco SAD protein had a localization pattern similar to that of aspen SAD and consistent with a role in lignification, SAD and CAD localization studies were performed on tobacco stem sections. To ensure that only native SAD was being detected by the SAD antisera, SAD localization studies were also performed on CAD-deficient stems (CAD antisense plants) and yielded identical results to those reported here for the wild-type stems. Cross sections (70 μm) of stem were visualized by immunofluorescence microscopy after incubation with rabbit antiserum raised against either (a) tobacco CAD or (b) aspen SAD followed by Alexa Fluor 594-conjugated goat anti-rabbit IgG secondary antibody. Control sections incubated with nonimmunized rabbit serum gave no fluorescence signal. Sections incubated with anti-CAD or anti-SAD serum revealed that both proteins predominantly localize to the ray parenchyma cells in xylem, which can be identified by reference to a phloroglucinol-stained stem section (Figure 1A). In the cambial and phloem area of the sections, CAD (Figure 1B) is restricted to parenchyma cells extending outward from the rays and forming an outer border surrounding bundles of phloem cells that are not labeled. By contrast, SAD (Figure 1C) has a more uniform distribution in this area, suggesting that, at this stage of stem maturation, SAD is also expressed in phloem. The localization of tobacco CAD to parenchyma cells that provide lignin precursors to adjacent xylem vessels and xylem and phloem fibers corresponds with the previously determined expression pattern of the eucalyptus (Eucalyptus gunnii) CAD promoter (Feuillet et al., 1995) and aspen CAD protein (Li et al., 2001). The overlapping localization of tobacco SAD to the same cell types and additionally to phloem cells is consistent with a potential role in lignin biosynthesis and corresponds to the localization pattern reported for SAD in the sixth internode of aspen (Li et al., 2001).
Cloning and Phylogenetic Analysis of Tobacco SAD

To enable us to design primers to facilitate cloning of the tobacco SAD, a BLAST search (Altschul et al., 1997) was performed using the aspen SAD protein sequence to identify related sequences in protein databases. Many SAD-like sequences were identified, including the protein products of three Elicitor-inducible (Ei3) genes (two from Arabidopsis and one from tomato [Solanum lycopersicum]), which showed high sequence identity (~70%) to aspen SAD. The SAD and SAD-like nucleotide sequences were aligned with four CAD sequences (from tobacco, tomato, eucalyptus, and alfalfa [Medicago sativa]), and regions of sequence conserved in the SADs but not in the CADs were chosen for degenerate primer design (see Methods). The products amplified from tobacco genomic DNA with these primers were used to design gene-specific primers for Rapid Amplification of cDNA Ends of cDNA prepared from 8-week-old stem xylem RNA. Finally, specific primers were designed to the ends of the obtained cDNA-derived fragments and were used to amplify entire cDNAs. Five closely related cDNAs were isolated. In amino acid sequence, these cDNAs showed 91 to 96% identity to each other, 71 to 75% identity to aspen SAD, and only 54% amino acid identity to tobacco CAD. Overall, the extremely similar Nicotiana tabacum (Nt) SAD2 and Nt SAD4 protein sequences (>96% identical) are most closely related to that of aspen SAD. It is common for tobacco to contain two near-identical copies of a gene, one homeolog inherited from each of the two ancestral progenitor species of the tobacco allotetraploid genome. Subsequent BLAST searches of the recently released tobacco genome sequence revealed no additional genes with a closer relationship to aspen SAD than the Nt SAD genes described here. Thus, if all angiosperms require a specific SAD for lignification, as proposed by Li et al. (2001), the Nt SAD genes are the only obvious candidate orthologs of aspen SAD in the tobacco genome.

Phylogenetic analysis using the Tree-Puzzle software (Schmidt et al., 2002) reveals that all five tobacco SAD proteins (for clarity, illustrated by one label, Nt SAD) and aspen SAD [Populus tremuloides (Pt) SAD] cluster together on a branch that is distinct from that containing the “true” CADs involved in developmental lignification, denoted class I CADs by Raes et al. (2003) (Figure 2; see Supplemental Data Set 1 online). Both the tobacco and aspen SADs cluster with the class II CADs, a group of enzymes with diverse substrate preferences that includes several proteins associated with defense responses, such as the elicitor-inducible Ei3 proteins from Arabidopsis, celery (Apium graveolens), and parsley (Petroselinum crispum). The phylogeny supports the evolutionary relationship between aspen and tobacco SAD proteins and confirms that the genes we have isolated from tobacco are the true orthologs of the aspen gene.

Homology Modeling of Nt SAD2 and Nt SAD5

Homology models of SAD2 and SAD5 were generated using the Phyre Web server (Kelley and Sternberg, 2009) using the high-resolution crystal structure of Pt SAD as the template (Bomati and Noel, 2005). The model of SAD2 suggests that the active site is, in essence, identical to that of the aspen SAD enzyme. Four
essential residues, Cys-50, Ser-52, His-72, and Cys-166, identified by Bomati and Noel (2005) as necessary for tetrahedral coordination of the catalytic Zn\(^{2+}\) ion plus a coordinating water molecule (see Supplemental Figure 2 online), are strictly conserved in the SAD2 (and the homologous Nt SAD4) active site. The high degree of conservation also extends to residues involved in binding the cofactor and in forming a network of interactions in and around the active site. By contrast, the model of SAD5 reveals significant active site differences caused by the substitution of His-72 and Cys-166 (Cys-165 in SAD2) with Tyr-72 and Asn-165 of SAD5. The model of SAD5 predicts that the hydroxyl group of Tyr-72 would be positioned precisely where the catalytic Zn\(^{2+}\) binds in Pt SAD, and by implication also in SAD2 (see Supplemental Figure 2 online). These alterations therefore render it highly unlikely that SAD5 (and the similar Nt SAD1 and Nt SAD3 enzymes) can function in sinapaldehyde reduction. Thus, only SAD2 (and Nt SAD4) are likely to be functional orthologs of the aspen SAD enzyme, whereas Nt SAD1, Nt SAD3, and Nt SAD5 are merely “SAD-like” and were subsequently denoted as Nt SAD-L1, Nt SAD-L3, and Nt SAD-L5 when sequences were submitted to the Genbank database.

The homology models reveal active site residue substitutions in tobacco SADs similar to those that distinguish aspen SAD from classical CAD-like enzymes. According to the models of Bomati and Noel (2005), the two families of enzymes have complementary but opposite geometric substitution patterns for key active site aromatic residues. Aspen SAD has bulky aromatic residues (Trp-61 and Phe-289) on the left wall of the active site and small hydrophobic residues (Leu-122 and Gly-302) at the base and on the right wall of the active site. By contrast, aspen CAD has hydrophobic residues (Leu-61 and Pro-289) on the left wall and bulky aromatic residues (Trp-122 and Phe-302) on the right wall and at the base of the active site. These substitution patterns completely change the topology of the active site and may contribute to different substrate binding modes in classical CADs and SAD-like enzymes (Bomati and Noel, 2005). Tobacco SAD2 totally adheres to the configuration of residues expected in the active site of a SAD enzyme with bulky aromatic residues on the left wall (Trp-61 and Phe-289, exactly as in aspen SAD) and hydrophobic residues (Met-122 and Ala-301) at the base and right wall of the active site.

Substrate Preference of Tobacco CAD and SAD

The relative substrate preferences of tobacco CAD and SAD for coniferaldehyde and sinapaldehyde were evaluated by assaying the activity of the purified recombinant proteins after expression of CAD19 (encoding lignification-related tobacco CAD19), SAD2, and SAD-L5 coding sequences in Escherichia coli. Consistent with the predictions of the molecular modeling, SAD-L5 had no detectable activity toward coniferaldehyde or sinapaldehyde. By contrast, CAD19 and SAD2 exhibited activity toward both coniferaldehyde and sinapaldehyde. Although the \(K_m\) values were considerably lower for CAD19 than they were for SAD2, neither enzyme showed a clear preference for either substrate, as indicated by similar \(K_{cat}/K_m\) values (Table 1). Comparison of the tobacco SAD with the previously described aspen SAD (Li et al., 2001) revealed that the turnover number (\(K_{cat}\)) for sinapaldehyde substrate was virtually identical (4.3 s\(^{-1}\) for tobacco SAD and 200 min\(^{-1}\) or 3.3 s\(^{-1}\) for the aspen SAD). Thus, in addition to a close phylogenetic relationship to the aspen SAD gene and the high immunological and amino acid sequence identity between the SAD2 and aspen SAD proteins, the tobacco enzyme can also use both coniferaldehyde and sinapaldehyde as substrates, confirming that it is a true functional ortholog.
Table 1. Kinetic Properties of Recombinant Tobacco CAD and SAD

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrate</th>
<th>$K_m$ ($\mu$M)</th>
<th>$V_{max}$ ($\mu$mol/min/mg)</th>
<th>$K_{cat}$ (s$^{-1}$)</th>
<th>$V_{max}/K_m$</th>
<th>$K_{cat}/K_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAD</td>
<td>Coniferaldehyde</td>
<td>5.44 ± 1.32</td>
<td>25.3 ± 1.72</td>
<td>17.6</td>
<td>4.65</td>
<td>3.23</td>
</tr>
<tr>
<td>CAD</td>
<td>Sinapaldehyde</td>
<td>3.41 ± 1.11</td>
<td>18.1 ± 1.56</td>
<td>12.6</td>
<td>5.32</td>
<td>3.70</td>
</tr>
<tr>
<td>SAD</td>
<td>Coniferaldehyde</td>
<td>86.8 ± 18.8</td>
<td>3.12 ± 0.263</td>
<td>2.18</td>
<td>0.0359</td>
<td>0.0251</td>
</tr>
<tr>
<td>SAD</td>
<td>Sinapaldehyde</td>
<td>102 ± 19.8</td>
<td>6.19 ± 0.312</td>
<td>4.33</td>
<td>0.0607</td>
<td>0.0425</td>
</tr>
</tbody>
</table>

Data are means and SE corresponding to triplicate analyses. CAD19 and SAD2 were expressed in *E. coli*, and the purified proteins were assayed using coniferaldehyde and sinapaldehyde as substrates. Values are means ± SE for three independent reactions.

Production of Plants Suppressed in SAD Expression

To directly assess the role of SAD in lignin biosynthesis via reverse genetics, two hairpin RNA-producing gene-silencing vectors (Wesley et al., 2001) were designed to suppress expression of (i) SAD and (ii) SAD and CAD simultaneously. Both constructs should target all five tobacco SAD genes for suppression because of the high degree of homology between them. For the coordinate suppression of SAD and CAD, we used a strategy previously tested by us in conventional nonhairpin producing vectors where fusing partial sense sequences for different genes within an expression cassette was sufficient to effect suppression of multiple targets (Abbott et al., 2002). The simultaneous suppression of both SAD and CAD was performed to assess whether SAD and CAD might act redundantly to convert sinapaldehyde to sinapyl alcohol in lignin biosynthesis. For the preparation of the vectors, a 619-bp partial CAD cDNA was introduced alone, or fused to a 586-bp partial SAD cDNA, into a donor vector (pDONR201) to make an entry clone before being recombined into the gene-silencing cassette was sufficient to effect suppression of mul-

of each of the selected SAD and SAD CAD plants.

Given the potential complications inherent in enzyme assays of crude protein extracts, particularly the fact that assays performed in vitro may not truly differentiate between CAD and SAD, because both enzymes can use the same substrates, we performed two further tests. Both RT-PCRs and immunoblots were used to confirm the identity of the genes/proteins suppressed in a selection of the SAD and SAD CAD plants. Real-time RT-PCR performed with primers specific for the five SAD genes or for CAD confirmed that SAD transcripts were reduced in all SAD and SAD CAD lines evaluated (Figure 4A). In the SAD plant lines, SAD transcript levels were reduced by 93 to 98% (indicating, as expected, that all five closely related genes were suppressed), whereas in the SAD CAD lines, SAD transcripts were higher but nonetheless considerably reduced compared with the levels in the wild-type plants. CAD transcripts were also reduced in the SAD CAD lines compared with the wild type (see Supplemental Figure 4 online), confirming the enzyme assay data. Immunoblots (Figure 4B) confirmed that SAD and CAD protein levels were similarly altered. The SAD protein band evident in wild-type tobacco stems was hardly detectable in extracts from the SAD and SAD CAD plants, whereas levels of CAD protein were normal in SAD plants but greatly reduced in SAD CAD plants. These data confirm that we have successfully and severely suppressed expression of the gene that encodes the tobacco xylem protein that corresponds immunologically to aspen SAD.
SAD-Suppressed Plants Have No Detectable Changes in Lignin

The phenotypic consequences of SAD suppression were monitored throughout development. Some SAD and SAD CAD lines initially grew more slowly than wild-type plants, but average heights after 10 weeks of growth were normal except for lines SAD CAD7 and SAD CAD34, which were 10 to 20% shorter than the wild type. Leaves of some lines showed slight loss of symmetry, whereas roots seemed longer and less colored than wild-type roots. The basis of these phenotypes is under further study.

The effect of SAD suppression on lignin biosynthesis was investigated by several complementary techniques. Because of the complex and heterogeneous structure of lignin, no single technique yields complete, unbiased data, and several methods need to be used to get representative data on the content, structure, and composition of the polymer (Halpin, 2004).

Quantitative analysis of the lignin content of woody xylem tissues was determined by the gravimetric Klason method and...
SAD CAD lines (SC3, SC7, and SC34) had significantly increased phenolics that were released (Table 2). By comparison, all three the absorbance values of the lignin-derived low molecular weight wild type and both SAD lines tested (S4 and S18), according to could be degraded by such alkali treatment was similar in the under these conditions. The proportion of lignin structures that about the structure of the lignin polymer, because only a pro-

Relative SAD mRNA concentration

Relative SAD mRNA concentration in stems of the wild-type, SAD, and SAD CAD plants evaluated by real time RT-PCR. Means and sd for three separate experiments on independent plants of three SAD and three SAD CAD lines were calculated using the standard curve method (User Bulletin no. 2, ABI PRISM 7700 Sequence Detection System software).

Sinapaldehyde Accumulates in Lignin in SAD CAD but Not SAD Lines

The substrate for SAD’s proposed role in generating sinapyl alcohol for S lignin biosynthesis is sinapaldehyde, and one might therefore expect sinapaldehyde to accumulate when SAD activity is suppressed. Accumulating indene derivatives (peaks at mass-to-charge ratio [m/z] 384 and m/z 354) originating from sinapaldehyde and coniferaldehyde have previously been detected incorporated into lignin in the thioacidolysis products of CAD-antisense plants but are below the detection level in wild-type plants (Kim et al., 2002). Our thioacidolysis analysis of lignin in severely SAD-suppressed plants, whereas plants suppressed in both SAD and CAD displayed only the changes to lignin structure and composition that have been previously described for plants suppressed in CAD alone. Our data therefore did not detect any evidence for mutually redundant roles of CAD and SAD in lignin biosynthesis.
Yields of m/z 384 peak derived from sinapaldehyde incorporated into lignin. *significantly different to wild-type plants (P < 0.05).

Phenolic Profiling of SAD-Suppressed Tobacco Plants

Targeted metabolite profiling was performed on SAD plants to determine (a) whether intracellular sinapaldehyde pools were increased and (b) whether altered phenolic profiles would indicate which pathways were blocked by the deficiency in SAD. A comparative ultraperformance liquid chromatography-mass spectrometry (UPLC-MS) profiling of soluble phenolics was performed on methanol extracts from S4, S45, and wild-type plants. Analysis of extracts from stem xylem, cortex (i.e., soft tissues outside the woody xylem), leaves, and roots quantified 497, 867, 948, and 1471 m/z peaks, respectively. Data were subjected to rigorous statistical analysis using both principle component analysis and Student’s t tests. Only metabolites identified as differentially expressed by both statistical tests in all transgenic plants (i.e., five replicate plants of each of lines S4 and S45) compared with wild-type extracts were considered further. This joint statistical approach to evaluate the comparative profiling did not reveal any accumulation of sinapaldehyde or sinapic acid or of any other obvious aromatic that might be expected to accumulate if the lignin pathway is perturbed. No robust changes were pinpointed for xylem, bark, or root extracts, but one UPLC-MS peak (peak X) in leaf extracts was reduced by an average of 40% in all SAD plant samples compared with the wild type. Tandem mass spectrometry (MS/MS) fragmentation indicated that peak X is hexosylated dihydrodehydrodiconiferyl alcohol (DDDC) or isodihydrodehydrodiconiferyl alcohol (IDDDC), which may have arisen via reduction of the phenylcoumaran ring or the side-chain double bond of (hexosylated) DDDC (see Supplemental Figure 5 online). Because true differences may have been masked by the stringent significance threshold we initially used, we subsequently pursued a targeted search for hexosylated DDC in the transgenic lines, down to 55% of wild-type levels (see Supplemental Figure 5 online). None of these hexosylated phenylcoumarans were observed in xylem, bark, and root tissues, and no peaks associated with the aglycones, DDDC or...
IDDC, and DDC, were observed in the leaf profiles. We conclude that SAD could therefore possibly play a role in the production or modification of DDC phenylcoumarans in tobacco.

DISCUSSION

We have cloned two tobacco orthologs of the aspen SAD gene first isolated by Li et al. (2001). No closer relatives of the aspen SAD could be found in the tobacco genome. The tobacco cDNAs encode proteins that share 72% amino acid identity with aspen SAD and are more closely related to it than to tobacco CAD, with which they share only 56% amino acid identity. Like aspen SAD, the tobacco SAD cDNAs were cloned from lignifying xylem, and immunolocalization indicates that tobacco SAD is localized to the same cell types as the aspen protein, most obviously the ray parenchyma cells. Expression of recombinant tobacco SAD shows that, like aspen SAD, it is capable of reducing both sinapaldehyde and coniferaldehyde. Our data therefore confirm the data of Li et al. (2001), demonstrating that a second enzyme distinct from CAD, but with overlapping ability to reduce sinapaldehyde, is expressed in the lignifying xylem of at least some woody dicots.

Although the existence of SAD in lignifying tissues is clear, its function is not. Crude extracts from plants suppressed in expression of the tobacco SADs have decreased capacity to reduce sinapaldehyde when assayed in vitro. Despite this, the biosynthesis of S lignin units proceeds normally in tobacco stem xylem even though SAD transcripts are reduced to 2 to 7% of normal levels and SAD protein is almost undetectable. In addition, metabolite profiling did not reveal accumulation of soluble sinapaldehydes in the SAD plants. Only when CAD is additionally suppressed are significant changes to lignin detected, making the polymer more susceptible to alkaline extraction caused by an altered structure with increased incorporation of sinapaldehyde and coniferaldehyde units, reduced proportion of β-O-4 linked units, and reduced S:G ratio. These are well-documented characteristics of CAD suppression (Halpin et al., 1994; Baucher et al., 1996; Vailhé et al., 1998). No additional changes, or changes in the amplitude of these effects that could be assigned to concomitant SAD suppression, could be detected, despite the effectiveness with which both genes were suppressed by the chimeric pSADCAD construct.

To our knowledge, our results, from plants severely suppressed in SAD activity, provide the first direct and conclusive data to confirm other circumstantial evidence that SAD is not needed for the biosynthesis of S lignin units in dicots. Data from many different groups show that the accumulation in lignin of derivatives of sinapaldehyde, the precursor of S monomers, is a characteristic of severe CAD suppression (Halpin et al., 1994; Ralph et al., 1998, 2001; Kim et al., 2002; Sibout et al., 2003); indeed, the level of the indene derivative shows a close inverse correlation with CAD activity (Lapiere et al., 2004). Our data (see Supplemental Figure 1 online) demonstrate that SAD protein levels are not reduced in CAD-antisense tobacco lines that have previously been shown to accumulate sinapaldehyde (Halpin et al., 1994); therefore, reduction in expression of CAD alone is responsible for the accumulation of sinapaldehyde in lignin in these plants. A study of Arabidopsis cad mutants draws a similar conclusion. In double mutants in the lignification-related CAD genes showing large increases in sinapaldehyde in lignin, transcript levels of the genes most closely related to aspen SAD are not significantly altered (Sibout et al., 2005). Indeed, expression of a Populus tremula × tremuloides SAD (99.2% amino acid identity to aspen SAD) in these plants can restore only 18% of the S monomer yield (Sibout et al., 2005). Moreover, some plants may be able to bypass the requirement for a sinapaldehyde reduction step in lignin biosynthesis and alternatively produce sinapyl alcohol from coniferyl alcohol. Evidence for such a pathway comes from radiolabeled precursor feeding studies in several tree species (Chen et al., 1999, Matsui et al., 2000; Tsuji et al., 2004) and from investigation of the substrate specificities of Arabidopsis F5H and COMT (Humphreys et al., 1999). If such a pathway functions in vivo, even if only as a supplementary alternative route, it obviates the original arguments for the necessity of the existence of a specific SAD. Most significantly, a role for SAD in the S lignin synthesis is not consistent with a recent attempt (Shi et al., 2010) to provide a comprehensive description of lignin biosynthesis in Populus trichocarpa, a close relative of aspen for which a genome sequence is available (Tuskan et al., 2006). Quantitative criteria were applied to identify, from the 45,555 gene model set of P. trichocarpa genome annotation, the genes most likely to be monolignol biosynthetic genes based on transcript abundance and specificity for differentiating xylem. P. trichocarpa genes involved in S lignin synthesis, CAD1, CAD2, and COMT2 (encoding F5H and COMT, respectively), met these criteria, as did CAD1 (corresponding to aspen CAD), but CAD2 (the ortholog of aspen SAD) apparently did not, because it had low expression in all tissues tested. It was included in further analysis in any case, merely because of the belief from prior literature (Li et al., 2001) that it is important for the formation of sinapyl alcohol. However, collectively, the accumulating data from Arabidopsis, poplar, and tobacco argue forcefully that SAD does not play a major role in sinapaldehyde reduction during lignin biosynthesis in vivo; that role is reserved for CAD, and CAD cannot fully compensate if CAD is deficient.

This work highlights the caution that should be applied when interpreting the biological significance of enzyme assays performed in vitro. We have shown that whole stem extracts of our transgenic plants can have significant reductions in sinapaldehyde-reducing activity, but this does not necessarily translate (in the case of the SAD plants) into deficiencies in sinapyl alcohol supply for lignin biosynthesis in woody xylem tissues. There are several possible explanations for this. Whole stems at the age sampled here contain more epidermal, cortical, and pith tissue than developing xylem, and low levels of SAD expression can be detected in these tissues by immunoblot. Consequently, the enzyme assays are not reporting directly on sinapaldehyde-reducing activity in lignifying xylem alone. Indeed, neither sinapaldehyde nor coniferaldehyde can normally be detected in soluble phenolic extracts from xylem (Dauwe et al., 2007), suggesting that, in vivo, either (a) very low levels of substrate are sufficient to maintain sinapyl alcohol production and the slow accumulation of lignin over the lifetime of a plant or (b) a substrate-channeling mechanism exists to guide newly synthesized sinapaldehyde directly into lignin synthesis. Enzyme assays performed
in vitro in an excess of exogenously added substrate cannot inform on either of these possibilities. Furthermore, the apparent disappearance of most sinapaldehyde-reducing activity in the SAD plants despite normal CAD levels may simply be misleading. Both CAD and SAD have been reported to exhibit substantial substrate inhibition kinetics (Lauvergeat et al., 1995; Li et al., 2001; Bomati and Noel, 2005). Indeed, Lauvergeat et al. (1995), working with eucalyptus CAD, noted that lack of detection of CAD activity may merely reflect particularly strong substrate inhibition promoted by the high substrate concentrations used for assays in vitro. Bomati and Noel (2005) likewise urged caution in interpreting SAD activity data based on their discovery of unusual kinetic properties and substantial substrate inhibition that cause significant loss in turnover at higher substrate concentrations. These phenomena may help explain how our assays of the primary transformant plants could predominantly report only on the activity of SAD, if it was the case that CAD was inhibited to a higher degree than SAD by the high sinapaldehyde concentration used. All of these issues highlight the absolute necessity of focusing on measurable effects on the lignin polymer itself before assigning a role in its biosynthesis to specific genes or enzymes.

Similarly, assays performed in vitro on recombinant proteins do not necessarily identify an enzyme’s true substrate, they merely suggest its likely relative preference in vivo toward the few substrates tested in the assay. Consequently, predicting the function of an enzyme from its activity in vitro, even when associated with localization data (as in the case of aspen SAD) is fraught with potential error and likely to be misleading. Accumulating data suggest that the SAD gene family is not highly conserved across species in the way that the CAD family is and that many duplication and deletion events occurred during its evolutionary history (Guo et al., 2010). Genes within the SAD family can have different expression patterns and substrate specificities (Kim et al., 2004, 2007; Barakat et al., 2009), implying some divergence of function. Indeed, Bomati and Noel (2005) demonstrated that just two changes in the active site of aspen SAD were sufficient to completely alter specificity. A mutant aspen SAD where Leu-122 was replaced by Trp and Gly-302 was replaced by Phe was no longer highly specific for sinapaldehyde but was more specific for coniferaldehyde (Bomati and Noel, 2005). Overall, the SAD family does not seem to have the conservation of sequence, function, or substrate specificity expected of genes essential to lignin biosynthesis, and instead several family members have been suggested to be involved in stress or defense responses (Somssich et al., 1996; Brill et al., 1999; Montesano et al., 2003). It has been hypothesized that SAD might be a dehydrogenase that directs monolignol precursors toward plant defense and that experimentation in planta would be critical to the experimentally challenging objective of defining its true physiological function (Bomati and Noel, 2005). We present just such in planta data here and demonstrate that, although tobacco SAD maintains a functional relationship to aspen SAD, being able to use both coniferaldehyde and sinapaldehyde as substrates, it plays no detectable role in lignin biosynthesis.

If tobacco SAD is not needed for biosynthesis of S lignin units, what alternative function might it serve in lignifying tissues? Phenolic profiling of wild-type and SAD-suppressed plants by UPLC-MS suggest that SAD could possibly play a role in the production or modification of DCC phenylcoumarans, or neo-lignans. Although a wide spectrum of different lignans exist in plants, their functions are unclear, although roles in defense are indicated by the biological activity of many lignans, which may have antimicrobial, antifungal, and antifeedant properties, and some have antitumor and antimicrobial activity toward mammalian cells. In particular, DCC glucosides have been shown to have cytokinin-like activity in tobacco (Orr and Lynn, 1992). In this study, reductions in DDC derivatives were detected in leaf extracts in SAD-suppressed plants. Consistent with this observation, subsequent work using plants expressing a β-glucuroniidase gene fused to the SAD promoter demonstrated that SAD is expressed in both leaf veins and lamina.

Our results have relevance to contemporary theories regarding the evolution of S lignin biosynthesis in angiosperms. It has been proposed that SAD and S lignin units may have emerged together during angiosperm evolution (Li et al., 2001; Peter and Neale, 2004). This idea is consistent with the traditional supposition that the absence of S lignin units in conifers is caused by the low affinity of gymnosperm CADs for sinapaldehyde. However, several commentators have challenged such a generalization and have pointed out that some exceptional gymnosperms can make S lignin units and/or have CADs that display activity toward sinapaldehyde (Campbell and Sederoff, 1996; Whetten et al., 1998; Sibout et al., 2005; Uzal et al., 2009). Moreover, a couple of recent reports demonstrate the presence of low levels of S lignin in the red alga Calliarthron cheilosporioides (Martone et al., 2009) and in the lycophyte Selaginella moellendorfii (Weng et al., 2008), suggesting possible independent evolution of S lignin biosynthesis in different plant lineages (Weng and Chapple, 2010). A recent study of the evolutionary history and functional differentiation of the CAD/SAD gene family (Guo et al., 2010) confirmed that the clade containing aspen SAD was angiosperm-specific but suggested that the relaxed evolutionary selection within the clade, plus the high substrate versatility, pointed to relatively specialized functions for individual genes in processes such as stress resistance. These data and our evidence presented here indicate that SAD was not critical to the evolution of S lignin unit biosynthesis in angiosperms.

METHODS

Plant Growth Conditions

Tobacco (Nicotiana tabacum cv Samsun) was grown in tissue culture on solid Murashige and Skoog medium with 3% Suc at 25°C with a 16-h photoperiod (100 μmol m⁻² s⁻¹). At 4 weeks, plants were transferred to compost in the greenhouse for further growth and analysis.

SAD Homology Modeling

Homology models of Nt SAD2 and Nt SAD5 were generated using the Phyre Web server (Kelley and Sternberg, 2009). The most suitable template was the high-resolution crystal structure of Pt SAD, Protein Data Bank code 1YGD (Bomati and Noel, 2005). The high degree of amino acid sequence conservation (70%) provides confidence in the resulting models. Structures were examined using the graphics visualization software Crystallographic Object-Oriented Toolkit (Emsley and Cowtan, 2004).
Activity of Recombinant Tobacco CAD and SAD

Coding sequences of TCAD19, TSAD2, and TSAD5 were amplified by PCR using primers to incorporate convenient restriction sites (see Supplemental Table 2 online). After cloning and sequencing, the products were transferred into pET52b+ vector (Novagen) to fuse a His tag at the C terminus of the cloned sequence. The obtained clones, pET52b-TCAD19 and pET52b-TSAD2, were introduced into Escherichia coli strain B21 (DE3) (Novagen). After isopropyl-β-D-thiogalactopyranoside induction, cells were lysed, and the soluble recombinant proteins were purified using Ni-NTA columns (Novagen). The eluted fractions were concentrated, and the buffer was exchanged to 100 mM Tris-HCl (pH 8.8) using Amicon Ultra-15 filters (Millipore) before further purification of recombinant proteins by anion exchange chromatography on Q HyperD 20 resin prepacked into an optima 5/10 column (BioSepra). Elution was achieved using a linear gradient to 0.5 M KCl in 100 mM Tris (pH 8.8) followed by a step gradient to 1 M KCl. Fractions containing active recombinant protein were pooled and concentrated, and the buffer was exchanged to 100 mM Tris-HCl (pH 7.5) and 5 mM DTT.

The kinetic parameters of the purified proteins were determined using triplicate reaction mixtures at 30°C containing 33 mM sodium/potassium-phosphate buffer (pH 6.3), 0.5 mM NADPH, and increasing amounts of coniferaldehyde or sinapaldehyde and monitoring at A394. Initial reaction rates were fitted to the Michaelis-Menten equation to determine steady state kinetic parameters using the software Enzyme Kinetics module of SigmaPlot (SYSTAT Software).

Immunoblots

Protein extracts (30 to 50 μg) were separated by 12% SDS-PAGE along with prestained protein markers (New England Biolabs). Proteins were transferred to nitrocellulose membranes using a Trans-Blot SD semidyman blotter (Bio-Rad), treated with blocking buffer (2% dried milk in 200 mM Tris base, 75 mM NaCl, 0.1% Tween-20, pH 7.6), incubated with primary antibody (anti-CAD 1:10000 or anti-SAD 1:3000), washed three times, and then incubated with 1:10000 dilution of horseradish peroxidase-conjugated anti-rabbit IgG (New England Biolabs). Detection was performed using the Phototope-HRP Western Blot Detection Kit (New England Biolabs).

Fluorescence Microscopy

Cross sections (70 to 100 μm) of stem from 12-week-old plants cut on a Vibraslice machine were placed on Superfrost Plus slides and blocked in 1% BSA (in PBS, pH 7.2, 0.5% Tween-20) for 45 min. Sections were incubated for 1 h at room temperature in a 1:5 dilution of rabbit serum was used as the primary antibody for the controls. After washing, the sections were incubated with a 1:250 dilution Alexa Fluor 594-conjugated goat anti-rabbit IgG for 45 min, then washed again before mounting in Hydromount and viewing with a Leica SP2 confocal laser scanning system equipped with an inverted microscope. Normal rabbit serum was used as the primary antibody for the controls.

Cloning of Tobacco SAD

RNA was purified from dissected xylem taken from the base of an 8-week-old tobacco stem using the SV Total RNA Isolation System (Promega). cDNA ends were amplified using the GeneRacer Kit (Invitrogen). The SAD and SAD-like nucleotide sequences were aligned with four CAD sequences (from tobacco, tomato [Solanum lycopersicium], eucalyptus [Eucalyptus gunnii], and alfalfa [Medicago sativa]). To increase the chances of ultimately amplifying all potential SAD-like genes from tobacco xylem, regions of sequences that were conserved in SAD and SAD-like sequences but not in CAD sequences were chosen for degenerate primer design. Amplification of tobacco genomic DNA with these primers resulted in a single product of 354 bp that had high sequence identity with aspen SAD (70% identity at the amino acid level). Using the sequence of the tobacco fragment, two gene-specific primers and two nested primers were designed with which to amplify the ends of the gene from tobacco cDNA prepared from the RNA of 8-week-old stem xylem. Fragments of approximately 600 bp for the SAD 5’ end and approximately 900 bp for the SAD 3’ end were isolated, cloned into Zero Blunt TOPO vector (Invitrogen), sequenced, and aligned against aspen SAD. New specific primers were then designed to the ends of the tobacco SAD sequence, and the entire cDNA was amplified from total stem cDNA.

Phylogenetic Analysis

Protein sequences with homology to tobacco CAD and SAD genes were identified by BLAST searches (Altschul et al., 1997), and the top hits were used to construct a phylogenetic tree. The sequences were aligned using Clustal W (Thompson et al., 1994). The output file was run in Tree-Puzzle 5.0 using quartet puzzling (Schmidt et al., 2002) to estimate the phylogenetic tree. This is a maximum likelihood method, which provides good support values for the clusters within the tree. The amino acid substitution matrix used was JTT (Jones et al., 1992). Rate heterogeneity among sites was modeled using a gamma distribution with four categories. The tree was viewed in Tree-View v.1.6.6 (Page, 1996). Alignments are provided as Supplemental Data Set 1 online.

Preparation of pSAD and pSADCAD Constructs

Partial sequences of SAD cDNA and CAD cDNA were amplified by PCR and independently introduced into the donor vector (pDONR201) using Gateway technology (Invitrogen). Both entry clones were digested with HpI1 and MscI endonucleases and then religated to fuse together SAD and CAD partial sequences in one pDONR201 vector. These two entry clones were then recombined with pHellsgate 8, a modified version of pHellsgate (Wesley et al., 2001), to make pSAD and pSADCAD, respectively.

Production of Transgenic Plants

Transgenic tobacco plants were produced by Agrobacterium-mediated transformation as described in O’Connell et al. (2002). After regenerated plants had gone through two rounds of rooting on selective medium, they were transferred to nonselective medium for 4 weeks and then transferred to compost and grown in the greenhouse for 2 weeks before being screened for activity of SAD and CAD. Selected plants were cloned propagated to yield small populations of identical individuals for further analysis.

Determination of SAD Activity in Crude Plant Extracts

Stem samples were ground in liquid nitrogen, extracted with 100 mM Tris-HCl, pH 7.5, 20 mM β-mercaptoethanol, then centrifuged at 20,000g, 4°C for 15 min. The total protein concentration of the supernatant was determined using the Bradford assay. SAD activity was determined by monitoring the reduction of sinapaldehyde to sinapyl alcohol. Reactions (1 mL) contained 20 to 50 μL plant extract, 100 mM KNa phosphate buffer, pH 6.25, 0.01 mM sinapaldehyde, and 0.01 mM NADPH, and the decrease in absorbance at 340 nm was measured for 3 to 5 min at 30°C.

Real-Time PCR Analysis of SAD and SAD CAD Plants

Total RNA was extracted from 100 mg of stem tissue using the RiboPure RNA Kit (Ambion). The RNA was treated with DNaseI and then reverse
transcribed into first-strand cDNA using Superscript III reverse transcriptase (Invitrogen). Reactions (25 μL) were set up in triplicate with 1 μL cDNA (diluted 1:10), 12.5 μL2 × SYBR Green reporter dye, and 300 nM of SAD, CAD, or ubiquitin (UBI) UBI primers. Control reactions for each primer set contained no cDNA. The reactions were run on the ABI Prism 7700 Sequence Detector at 95°C for 15 min, followed by 40 cycles at 95°C for 15 s, 60°C for 30 s, and 72°C for 30 s. Primers used: NtUBi-F = GGA CCA GCA GAG GTT GAT CT; NtUBi-R = TCA GCC AAG GTC CTT CCA T; NtCAD-F = CAT TTT GGT TTT AAT CAG AGT GGA; NtCAD-R = CCC ATA TGT CCA ACT CCT CCT; NtSAD-F = GAC CCT GAC CAA ATG CAG; NtSAD-R = GTG CGC CAA C(CT)/A TTA CAA GC.

**Alkaline Extraction of EXR**

Xylem samples (1 g) from the base of 10-week-old tobacco stems were dried at room temperature, weighed, milled in a SPEX Certiprep freezer mill, and reweighed. Samples (300 mg) were extracted in a Soxhlet extractor with toluene: ethanol 1:1, ethanol, and H2O. The Klason line) and ground to pass a 0.5-mm sieve, and solvent was extracted in a Soxhlet extractor (toluene-ethanol 2:1, ethanol, and H2O). The method of Iiyama and Wallis (1990) was used. EXR (5 mg) was extracted in 1 mL methanol for the determination of UV absorbance at 280 nm.

**Acetyl Bromide Lignin Determination**

The method of Iiyama and Wallis (1990) was used. EXR (5 mg) was extracted in 1 mL methanol for the determination of UV absorbance at 280 nm.

**Klason Lignin Analysis**

The base of 10-week-old tobacco stems was collected (three plants per line) and ground to pass a 0.5-mm sieve, and solvent was extracted in a Soxhlet extractor (toluene-ethanol 1:1, ethanol, and H2O). The Klason lignin determination of the EXR was run in duplicate for each sample by the standard procedure (Dence, 1992). The SC between analytical duplicates was always less than 1% of the mean value.

**Thioacidolysis**

Lignin structure was investigated using thioacidolysis as previously described (Lapiere et al., 1999). All the experiments were run in duplicate, and 20 mg from the extract-free samples was prepared for Klason lignin analysis. The lignin-derived compounds were identified by gas chromatography–mass spectrometry of their trimethylsilyl derivatives. The quantitative evaluation of indenes originating from sinapaldehyde and coniferaldehyde that have undergone B-O coupling and lignification was performed from ion chromatograms reconstructed at m/z 384 and 354, respectively. The determination of the conventional G and S monomers was done from their prominent benzylic ions at m/z 269 or 299, respectively. All the calculations were done with the response factors of the main G and S monomers, relative to the C22 internal standard, and the yields were expressed on the basis of the extract-free lignin content of the samples.

**UPLC-MS Profiling**

The wild type and transgenic tobacco lines SAD4 and SAD45 were grown for 10 weeks in the greenhouse (16-h photoperiod at 22 to 25°C) and harvested just before flowering. After homogenization of leaf, xylem, bark, and root tissues in liquid nitrogen, liquid-liquid extraction was performed as previously described (Meyermans et al., 2000; Morreel et al., 2006).

Phenolics present in the aqueous phase were analyzed on a reversed phase Acquity BEH C18 (2.1 × 150 mm, 1.7 μm; Waters) column, heated at 40°C, using an Acquity Ultra-high Performance Liquid Chromatography (Waters) instrument coupled with an Acquity 2996 photodiode array detector (Waters) and a LCQ Classic ion trap mass spectrometer (ThermoQuest). A gradient from 0.1% triethylammonium acetate (pH 5, buffer A) to acetonitrile/methanol (75/25, v/v, to which 0.1% triethylammonium acetate was added, buffer B) was used for separation. At a flow rate of 0.2 mL/min, buffer A decreased from 95% to 50% in 25 min, and to 0% in 5 min. Ionization occurred in the negative mode using an Atmospheric Pressure Chemical Ionization source using the following conditions: capillary temperature 150°C, vaporizer temperature 350°C, sheath gas 25 (arb), aux gas 3 (arb), source current 5 μA. Quantification was based on full mass spectrometry (MS) (maximum ion time, 50 s; number of microscans, 1) data acquired in the m/z range from 100 to 1000 D. For structure elucidation, full MS scans were interrupted with data-dependent MS/MS scans (maximum ion time, 50 s; number of microscans, 2; collision energy, 35%).

Full MS data were integrated with the XCMS package (R version 2.4.1) using only the xcmsSet (fwhm 7, max 20, snthresh 5, step 0.1, steps 3, mzdiff 0.5) and group (bw 10, mzwid 0.5) functions. Statistical analysis was performed with Student’s t tests (α = 0.001), using a Welch correction in the case of heteroscedasticity, and principal component analysis with the prcomp function available in R version 2.4.1. All peaks were centered and scaled to unit variance. Peaks with loading factors that had an absolute value more than 2 s/d away from the mean loading value for a particular principal component were considered as major contributors to that principal component.

**MS Structure Elucidations**

Peak X, eluting at 12.8 min, showed a MS peak at m/z 581. MS/MS analysis yielded a base peak at m/z 521, resulting from the loss of the acetate adduct and three other major daughter ions at m/z 359, probably caused by the further loss of a hexose moiety, and at m/z 341 and 329, both latter ions likely derived from the m/z 359 ion by the detachment of H2O and CH3OH, respectively. The similarity of the MS/MS fragmentation behavior of the aglycone moiety to that of (8–5)-DDC (parent ion at m/z 357 and major daughter ions at m/z 339 and 327) (Morreel et al., 2004a, 2004b) suggests that peak X is hexosylated DDDD or IDDDD. These structures might arise from the reduction of the phenylcoumaran ring or the side chain double bond of hexosylated DDC.

A targeted search for hexosylated DDC (m/z 579) in the UPLC-MS chromatograms revealed a peak at 12.6 min. Because MS/MS fragmentation afforded the same neutral losses as observed for peak X, and because the m/z 357, 339, and 327 daughter ions referred to the presence of a DDC moiety (Morreel et al., 2004a, 2004b), this compound was identified as hexosylated DDC.

**Accession Numbers**

Sequence data from this article and for the phylogenetic analysis shown in Figure 2 can be found in the EMBL/GenBank data libraries under the following accession numbers: Ag (Alpinia graveolens), Q38707; At (Arabidopsis thaliana) CAD1, P42734; At4g39330; At CAD2, P48523; At3g19450; At CAD3, O65621; At1g43790; At CAD4, Q02971; At1g437980; At CAD5, Q02972; At2g37990; At CAD6, O49482; At4g34230; At CAD7, Q85275; At2g21730; At CAD8, Q95510; At2g21890; Ca (Camptotheca acuminata), Q7XAB2; Cr (Catharanthus roseus), O64940; Fxa (Fragaria × ananassa), AAD10227; Le CAD, X02655; Mc (Mesembryanthemum crystallinum),


