2011

Switchgrass Contains Two Cinnamyl Alcohol Dehydrogenases Involved in Lignin Formation

Aaron J. Saathoff
USDA-ARS, asaathoff2@unl.edu

Christian M. Tobias
USDA-ARS, christian.tobias@ars.usda.gov

Scott E. Sattler
USDA-ARS, Scott.Sattler@ars.usda.gov

Eric J. Haas
Creighton University, EricHaas@creighton.edu

Paul Twigg
Creighton University, twiggp@unk.edu

See next page for additional authors

Bioenerg. Res. (2011) 4:120–133
DOI 10.1007/s12155-010-9106-2

Switchgrass Contains Two Cinnamyl Alcohol Dehydrogenases Involved in Lignin Formation

Aaron J. Saathoff · Christian M. Tobias · Scott E. Sattler · Eric J. Haas · Paul Twigg · Gautam Sarath

Published online: 28 September 2010

Abstract Lignin content of switchgrass (*Panicum virgatum* L.), a bioenergy species, is a critical determinant of biomass quality since it can negatively impact conversion of biomass into liquid fuels via biochemical platforms. Cinnamyl alcohol dehydrogenase (CAD) is a key enzyme in lignin biosynthesis. Here, we have shown that cv. Kanlow switchgrass contains at least two closely related CAD genes (*PviCAD1* and *PviCAD2*) that code for proteins containing highly conserved domains and residues that identify them as bona fide CADs. Both recombinant proteins displayed substrate kinetics consistent with their presumed role in cell wall lignification. Proteomic and immunoblotting detected CAD containing spots in internode protein extracts, and proteomic analyses demonstrated that both CADs were expressed. In planta CAD activity, CAD protein levels were observed at all stages of tiller development. A real-time qPCR analysis of the two CADs and one CAD-like sequence indicated that transcripts coding for PviCAD1 were present in greater abundance than those coding for PviCAD2. Transcripts for a third CAD-like sequence (*PviAroADH*) were present at intermediate levels as compared to *PviCAD1* and *CAD2*. The predicted protein sequence of PviAroADH indicated that it was an enzyme unrelated to lignification based on phylogenetic and protein modeling data.

Keywords Cinnamyl alcohol dehydrogenase · Internodes · *Panicum virgatum* L. · Proteomic identification · Recombinant enzyme characterization · Tillers · Switchgrass

A. J. Saathoff · S. E. Sattler · G. Sarath (✉)
Grain, Forage, and Bioenergy Research Unit, USDA-ARS,
University of Nebraska,
137 Keim Hall,
Lincoln, NE 68583-0937, USA
e-mail: Gautam.Sarath@ars.usda.gov

A. J. Saathoff · S. E. Sattler · G. Sarath
Department of Agronomy and Horticulture, University of Nebraska,
137 Keim Hall,
Lincoln, NE 68583, USA

C. M. Tobias
Genomics and Gene Discovery Unit, USDA-ARS,
800 Buchanan St,
Albany, CA 94710, USA

E. J. Haas
Department of Chemistry, Creighton University,
2500 California Plaza,
Omaha, NE 68178, USA

P. Twigg
Department of Biology, University of Nebraska-Kearney,
Kearney, NE 68849, USA

Introduction

Second generation biofuels, if properly developed, are likely to play a part in alleviating carbon emissions while at the same time improving energy security [3]. In the US, switchgrass (*Panicum virgatum* L.) is being developed as a potential feedstock for these purposes, and intensive efforts are underway towards developing cultivars to serve as dedicated feedstocks. Perennial plants, such as switchgrass, have the advantage of requiring fewer chemical inputs than row crops, enhance soil organic carbon, and can be grown on marginal land. On-farm field research has also reported a highly positive overall energy balance for switchgrass grown for biomass applications [44].

Understanding the mechanisms of cell wall synthesis is important since plants bred or engineered with easier to digest cell walls may enhance fuel yields or ease down-

stream processing steps resulting in reduced energy requirements for bioconversion [12, 40]. Although broadly similar in composition to dicots, grass secondary cell walls have significant amounts of lignin, ferulic, and *p*-coumaric acids, which can exist freely or form ester- or ether-linkages between lignin and cell wall polysaccharides [35, 51]. As a major component of grass secondary cell walls, lignin has been implicated as a key factor in the recalcitrance of biomass towards hydrolytic enzymes used in cell wall deconstruction [10, 17, 28].

Lignin is a complex heteropolymer whose monomers are derived from the deamination of phenylalanine followed by numerous side chain modifications, hydroxylations, and *O*-methylations of cinnamic acid [7]. The heteropolymer has generally been characterized as being mainly comprised of three different phenylpropanoid subunits: *p*-hydroxyphenyl (H-lignin), guaiacyl (G-lignin), and syringyl (S-lignin) [28], although other derivatives such as acylated lignin units and hydroxycinnamyl aldehydes were also found capable of being incorporated [7]. Monocot and dicot lignin generally have contained similar amounts of G and S units, but grass lignin was also found to contain a small, but significant amount of H units [51].

Lignin biosynthesis has been suggested to require a suite of at least ten enzymes [28]. A key enzyme involved in the synthesis of lignin subunits is cinnamyl alcohol dehydrogenase (CAD). Essentially, CAD catalyzes the conversion of cinnamyl aldehydes to their corresponding alcohols (H, G, and S lignin monomers) which are subsequently incorporated into the lignin polymer. CAD is a member of the alcohol dehydrogenase superfamily and has been characterized from a number of different species [19, 22, 29–32, 42, 56]. The crystal structure of an *Arabidopsis* CAD, AtCAD5, has been solved [55]. Disruption of CAD through natural or engineered mutations can lead to plants with altered lignin levels and differential incorporation of monolignals into lignin [2, 21, 34]. For instance, a sorghum brown midrib phenotype, *bmr6*, was recently shown to be caused by a nonsense mutation in CAD [38, 43] and plants with this mutation had an altered lignin composition [34, 37].

Reduction in lignin amounts or changes in lignin composition can have beneficial effects. In forage grasses, decreasing lignin resulted in enhanced dry matter digestability for grazing animals [6, 11]. These same types of changes have also been shown to enhance sugar release or actual ethanol yields from cellulosic biomass of both dicots and monocots [14, 15]. CAD occupies a central role in lignin biogenesis and alteration in its activity often results in changing lignin content, leading to plants with potentially improved quality. However, plants contain a number of CAD and CAD-like proteins which do not all participate in lignin biosynthesis [23], and it is important to biochemically characterize those CAD or CADs that can participate in lignin biosynthesis. Such studies will clearly be an important part of selective breeding of switchgrass biomass cultivars with desirable traits for biofuel production.

Materials and Methods

Identification of PviCAD1, PviCAD2, and PviAroADH

PviCADs were identified during an initial analysis of a switchgrass EST resource [49]. Several switchgrass cDNA clones were identified during single pass library sequencing as being closely related to the *ZmCAD*2 gene of *Zea mays* encoding cinnamyl alcohol dehydrogenase. However, none contained the entire coding region. Therefore, RT-PCR was performed to isolate full length coding sequences using 1 μg total RNA isolated from root tissue of the cv. ‘Alamo’ or stem tissue from the cv. ‘Kanlow’. These were reverse transcribed with Thermoscript II reverse transcriptase (Invitrogen Corp., Carlsbad, CA, USA) according to manufacturer’s instructions and amplified with the primers: 5′-CATATGGGCAGCCTGGCGTCG-3′ (PviCAD1 start) and 5′-GAATTCAGTTGGCCGGCGCCC-3′ (PviCAD1 stop). Products from both reactions were then cloned into pCR2.1 and sequenced. One isolate from cv. Alamo root RNA (*PviCAD*1, GenBank accession no. GU045611) and one from cv. Kanlow stem RNA (*PviCAD*2, GenBank accession no. GU045612), which fell into two separate clusters of sequences after alignment with existing switchgrass EST sequences, were chosen for further analysis. These isolates were subcloned by digestion with *Eco*RI and *Nde*I into the corresponding sites of pET28a (EMD Chemicals, Inc., Madison, WI, USA). Recombinant proteins were produced in *Escherichia coli*, and purified as described earlier [43]. A survey of the available EST sequences and the number of clones identified for each CAD-like sequence are shown in Table 1. We have not yet found other unique CAD-like sequences in the available EST collections.

The putative PviAroADH coding region was assembled from available switchgrass EST sequences and corresponds to dbEST accession number GD038270 and NCBI Unigene assembly Pvr.3084. These sequences were identified as potentially coding for CAD-like proteins through a TBLASTN search, followed by Clustal and manual assembly of overlapping sequences. Primers based on the assembled sequence were used to amplify a product of the expected size by RT-PCR using internode RNA as a template. Identity of the product was confirmed by DNA sequencing. Despite numerous attempts, we were unable to obtain soluble, active

Table 1 Representation of individual cDNA clones by library

	Tissue									
	ST	AP	CR	FB	ET	LF/SD	RT	SL	CA	Total
PviCAD1	0 (0.000)[a]	2 (0.007)	6 (0.015)	6 (0.022)	3 (0.011)	4 (0.014)	10 (0.036)	0 (0.000)	0 (0.000)	36 (0.016)
PviCAD2	0 (0.000)	5 (0.017)	6 (0.015)	7 (0.026)	5 (0.018)	3 (0.011)	12 (0.043)	2 (0.016)	0 (0.000)	46 (0.020)
PviAroADH	6	1	3 (0.007)	2 (0.007)	0 (0.000)	2 (0.007)	5 (0.018)	1 (0.008)	0 (0.000)	22 (0.010)

ST abiotic and biotic stress (normalized cDNA library), *AP* apex and stem library, *CR* crown library, *FB* early floral bud library, *ET* etiolated seedling library, *LF/SD* flowering/seed development library, *RT* root library, *SL* seedling library, *CA* callus (normalized cDNA library)

[a] Numbers in parenthesis are % of total number of clones with high-quality sequence data

recombinant protein for PviAroADH and therefore could not document its enzymatic characteristics.

Enzyme Assays

Enzyme activity on each substrate was measured using reaction conditions similar to those previously published [19, 53] with adaptations for use with a microplate reader (BioTek Synergy HT, BioTek Instruments, Winooski, VT, USA). However, substrate levels were varied while cofactor concentration was kept constant at 200 μM in order to kinetically characterize the enzymes. Enzyme dilutions were prepared immediately before their use from 100 μL aliquots that were stored at −80°C, and fresh dilutions were prepared approximately every 30 min in order to minimize the effects of enzyme activity loss. The amount of protein used for each assay was 3.51 and 4.73 ng for PviCAD1 and PviCAD2, respectively. Reaction rates were calculated from the first 90 s of absorbance data using PROC GLM in SAS 9.1 (SAS Institute Inc., Cary, NC, USA) and Michaelis–Menten kinetic parameters were estimated using nonlinear least squares whereby multiple initial starting values were chosen for each curve in order to ensure reliable convergence.

Plant Materials

P. virgatum cv. Kanlow N1 [52] was grown in fields at the University of Nebraska Agricultural Research Experiment Station near Mead, NE. Tillers were harvested at anthesis about 4 cm above soil surface, stripped of leaves and sheaths, separated into individual internodes, and immediately placed on dry ice for transport to the laboratory where it was stored at −80°C until used. The internode subtending the peduncle was labeled as internode 1 and the lowest (most mature internode) was labeled as internode 6 [39]. Most tillers generally possessed six well-distinguishable internodes. Tillers containing fewer than six internodes were not used.

Preparation of Internode Extracts for CAD Activity Assays

Internodes were ground using dry ice and a coffee grinder and the powdered material was used for further analyses. Centrifuge tubes (2.0 mL) were filled with approximately 300 mg of plant material, 1 mL of buffer (100 mM Tris–HCl, pH 7.5; 5 mM DTT; 5% (*v*/*v*) ethylene glycol) and 10 μL of a protease inhibitor mixture (Sigma–Aldrich #P9599). Samples were placed on ice and sonicated using a Branson Digital Sonifier 450 (Branson Ultrasonic Corp., Danbury, CT, USA) at 20 W three times with a 15 s pulse, with samples placed for 30 s into an ice-ethanol bath between pulses to allow for sufficient cooling. Sonicated extracts were then centrifuged for 15 min at 14,000 RPM and 4°C. Supernatant was decanted into new tubes and kept on ice until assayed for CAD activity. Protein content was determined using a colorimetric assay (Pierce 660 nm Protein Assay, Pierce Biotechnology, Rockford, IL, USA) and lysozyme was used for generating standard curves.

One- and Two-Dimensional Gel Electrophoresis

For one-dimensional SDS-PAGE, approximately 300 mg of ground internode materials were extracted as described earlier [43] and proteins were separated on 12% polyacrylamide gels [26]. Approximately 25 μg of protein was loaded into each well. Separated proteins were transferred to nitrocellulose membranes for immunoblotting as described below.

For two-dimensional gel electrophoresis, proteins were first extracted from 1 g samples of switchgrass internodes essentially according to a previously published method [54]. Final protein pellets were taken up in 180 μL of sample buffer (8 M urea, 50 mM DTT, 4% CHAPS, and 0.2% ampholytes) and used for isoelectric focusing (IEF). IEF was performed using the ZOOM IPGRunner System (Invitrogen Corp., Carlsbad, CA, USA) and 7.0 cm immobilized pH gradient (IPG) strips with a pH range of

5.0–8.0. IPG strips were rehydrated overnight with 155 μL of protein extract dissolved in the sample buffer described previously. After rehydration, IPG strips were prepared for IEF using the manufacturer's protocol. IEF was conducted using a power supply (Thermo Electron Corp. PRO-6000) with the following steps: 175 V for 30 min, 175–2,000 V linear increase over 45 min, and 2,000 V for 3 h. After IEF, IPG strips were equilibrated in a buffer containing 6 M urea, 375 mM Tris–Cl, pH 8.0, 2.5% SDS, and 20% glycerol. Strips were then reduced in equilibration buffer with 65 mM DTT followed by alkylation in equilibration buffer containing 135 mM iodoacetamide. The second dimension was run on 12% polyacrylamide gels (C.B.S. Lite Slab Gel Kit, C.B.S. Scientific Company, Del Mar, CA, USA) and stained using Coomassie brilliant blue. We had conducted a larger proteomic analysis (data not shown) of several spots obtained from the two-dimensional gel shown in Fig. 5a; however, for this study we focused on the spots labeled 1–7 since this was the region expected to contain CAD proteins based on the predicted pI and molecular weights. We used the 2-D immunoblots as a guide to narrow these analyses. Protein spots numbered 1–7 in Fig. 5b were analyzed by mass spectrometry as described earlier [20].

Immunoblots

Proteins separated by SDS-PAGE were blotted to nitrocellulose membranes and probed with polyclonal antibodies raised to CAD as described earlier [43]. Polyclonal antibodies raised against soybean root nodule ascorbate peroxidase (AscPx) were used as loading controls. Preliminary data indicated that the signal for AscPx was relatively consistent across many different switchgrass tissues at similar protein loads (not shown) and could therefore serve as a control.

Real-Time qPCR

RNA was isolated from *P. virgatum* internodes using a previously published method [48]. After treating total RNA with RQ1-RNase-free DNase (Promega, Madison WI USA), 1 μg from each preparation was reverse transcribed using an anchored oligo dT primer mixture along with random hexamers and the Transcriptor First Strand cDNA kit (Roche Diagnostics, Indianapolis, IN, USA). Real-time PCR used an ABI Prism 7000 Sequence Detection System (Applied Biosystems, Foster City, CA, USA) and the SYBR Green PCR Master Mix. Primers used for these reactions were designed using the included Primer Express software and are given in Table 2. Reactions were performed in quadruplicate. Delta Ct values for the target genes were generated relative to PviCAD1 transcript level in internode 1. Three control genes were used for normalization: UCE1, eIF4α, and GAPDH. Normalization factors were generated by geometric averaging of the control genes using geNorm version 3.4 [50], and a suitable pairwise variation (*V*) value indicated that the selected combination of control genes were stable. Normalization factors were then applied to the delta Ct values to generate relative expression levels.

Table 2 Primer sequences

PviCAD1	FWD	5′-AGGCCAACGTTGAGCAGTA
	RVS	5′-CGATCCCTGCTGGTCTGG
PviCAD2	FWD	5′-GCGAGGTGGTGGAGGTC
	RVS	5′-CGATCCCGGCAGGCCT
PviAroADH	FWD	5′-CAGCCCCATGAAGTACCA
	RVS	5′-TTCATTGTCAGCAAGAGCG

Protein Modeling

Models of PviCAD1 and PviAroADH were created with SWISS-MODEL [4]. Both models used the published *Arabidopsis* CAD5 structure [55] as a template for model building (PDB ID 2CF6).

Results

Phylogenetic Analyses of Three Switchgrass CAD-Like Protein Sequences

ClustalW2 was used to align PviCAD1, PviCAD2, and PviAroADH protein sequences (Fig. 1). The results indicated that PviCAD1 and PviCAD2 share a very high degree of identity (95%) to each other, but only 45% identity to PviAroADH. In order to compare the protein sequences with other plant CADs, sequence data was obtained from publicly available databases and used to construct a phylogenetic tree (Fig. 2). The tree included a number of CAD and CAD-like proteins identified in different species including CADs from *Arabidopsis* (AtCAD1-9) as well as CADs from sorghum (*Sorghum bicolor*, SbCAD1-7). Other CADs utilized in this analysis have either been shown or suggested to directly impact lignification in sorghum (*S. bicolor*, bmr6 [43]), rice (*Oryza sativa*, OsCAD2 [56]), perennial ryegrass (*Lolium perenne* L., LpeCad1 [32]), corn (*Z. mays*, ZmCAD2 [18]), and sugarcane (*Saccharum officinarum*, SoCAD [45]). All of the CADs that have been shown to be involved in lignin biosynthesis clustered into a single clade. In turn, this clade further subdivided into monocot

```
PviCAD1         MG--SLASERTVVG----WAARDAAGHLSPYTYTVRNTGPEDVVVKVLYCGICHTDIHQA 54
PviCAD2         MG--SLASERTVVG----WAARDATGHLSPYTYTVRKTGPEDVVVKVLYCGICHTDIHQA 54
PviAroDH        MAPVAAAAEQQHTGKAAALAARDASGHLAPLTITRRSTGDDDVAIKILYCGMCHSDLHFI 60
                *.  : *:*:  .*     *****:***:* * * *.** :**.:*:****:**:*:*
PviCAD1         KNHLGASKYPMVPGHEVVGEVVEVGPEVSKHRVGDVVGVGVIVGCCRECRPCKANVEQYC 114
PviCAD2         KNHLGASKYPMVPRHEVVGEVVEVGPEVSKHRVGDVVGVGVIVGCCRECRPCKANVEQYC 114
PviAroDH        KNEWNNAMYPMVPGHEIAGVVTEVGKNVTKFKAGDRVGVGCMVNSCQSCDRCDEGFENHC 120
                **. . : ***** **:.* *.*** :*:*.:.** **** :*..*:.*  *. ..*::*

PviCAD1         NKRIWSYNDVYTDGRPTQGGFASTMVVDQKFVVPIPAGLAPEQAAPLLCAGVTVYSPLKH 174
PviCAD2         NKRIWSYNDVYTDGRPTQGGFASTMVVDQKFVVPIPAGLAPEQAAPLLCAGVTVYSPLKH 174
PviAroDH        RGIIFTYNSVDRDGTVTYGGYSSNVVVHERFVVRFPDAMPLDQGAPLLCAGITVYSPMKY 180
                .  *::**.*  **  * **::*.:**.::*** :* .:. :*.*******:*****:*:

PviCAD1         FGLTAPGLRGGILGLGGVGHMGVKVAKALGHHVTVISSSSRKRAEAMDDLGADAYLVSSD 234
PviCAD2         FGLTAPGLRGGILGLGGVGHMGVKVAKAMGHHVTVISSSSRKRAEAMDELGADAYLVSSD 234
PviAroDH        HGLNVPGKHVGVLGLGGLGHVAVKFAKAFGMKVTVISTSPAKKQEALERLGADAFIVSKS 240
                .**..** : *:*****:**:.**.***:* :*****:*. *: **:: *****::**..

PviCAD1         AGAMAAAADSLDYIIDTVPVHHPLEPYLALLRLDGKHVLLGVVGEPLSFVSPMVMLGRKS 294
PviCAD2         AEAMAAAADSLDYIIDTVPVHHPLEPYLALLRLDGKHVLLGVVGEPLSFVAPMVMLGRKA 294
PviAroDH        ADEMKAAVATMDGIINTVSANMSIAPYMGLLKPNGKMIMVGLPVKPLEIPPFDLIMGNKT 300
                *  * **. ::* **:**..: .: **:.**: :** :::*:  :**.: .  :::*.*:

PviCAD1         VTGSFIGSVDETAELLRFCVDKGLTSQIEVVKMGYVNEALERLERNDVRYRFVVDVAGSN 354
PviCAD2         VTGSFIGSIDETAELLRFCVDKGLTSQIEVVKMGYVNEALERLERNDVRYRFVVDVAGSN 354
PviAroDH        LAGSCIGGMRDTQEMIDVAAKHGVTADIEVVGAEYVNTAMERLAKADVRYRFVIDIGNTL 360
                ::** **.: :* *:: ....:*:*::****   *** *:*** : *******:*:..:

PviCAD1         IEEQAAAAGAPAN 367
PviCAD2         IEE---ATGAPAN 364
PviAroDH        KSSE--------- 364
```

Fig. 1 Alignment of PviCAD1, PviCAD2, and PviAroDH protein sequences using ClustalW2

and dicot CADs, with PviCAD1 and PviCAD2 being most closely related to the other grass CADs, namely, ZmCAD2, SoCAD, and bmr6. In contrast, PviAroADH showed significant homology to a putative corn aromatic alcohol dehydrogenase and clustered with CAD-like sequences including AtCAD8 (ELI3) that have been associated with the reduction of aromatic aldehydes and not with cell wall lignification [47].

Biochemical Characterization of Recombinant PviCAD1 and PviCAD2

Recombinant enzyme activity of PviCAD1 and PviCAD2 were assayed on a variety of substrates in order to examine substrate preferences and enzyme kinetics. Both forward (aldehyde-to-alcohol) and reverse (alcohol-to-aldehyde) reaction rates were characterized using coniferaldehyde, sinapaldehyde, coniferyl alcohol, sinapyl alcohol, and coumaryl alcohol as substrates. Results from the enzyme assays followed Michaelis–Menten kinetics for all of the tested substrates. Estimates for kinetic constants K_m, V_{max}, and k_{cat}, and k_{cat}/K_m are provided in Table 3. For the alcohol substrates, PviCAD1 had the lowest K_m at 0.98 μM for coniferyl alcohol followed by coumaryl (4.2 μM) and sinapyl alcohols (15.6 μM), respectively; V_{max} values were 170 nkat mg^{-1} for coniferyl alcohol, 320 nkat mg^{-1} for coumaryl alcohol, and 600 nkat mg^{-1} for sinapyl alcohol, respectively. On aldehyde substrates, PviCAD1 displayed the lowest K_m for coniferaldehyde at 10.9 μM while the K_m for sinapaldehyde was 14.5 μM. Also, PviCAD1 had significantly higher V_{max} values on both aldehyde substrates when compared to alcohol substrates. The V_{max} of 2,530 nkat mg^{-1} for sinapaldehyde was significantly higher than a V_{max} of 1,280 nkat mg^{-1} when coniferaldehyde was used as a substrate. Because PviCAD2 shared a highly similar sequence to PviCAD1, it was only tested on the aldehyde substrates. The results were similar to PviCAD1, although the calculated K_m values for PviCAD2 were slightly lower. Both enzymes consistently exhibited higher activity on sinapaldehyde than on coniferaldehyde. A table of CAD kinetic parameters was collected from the published literature and is presented in Table 5. The results for PviCAD1 and PviCAD2 kinetics fall within the range of CAD kinetic data that were reported. PviAroADH was not analyzed because we could not obtain active soluble protein.

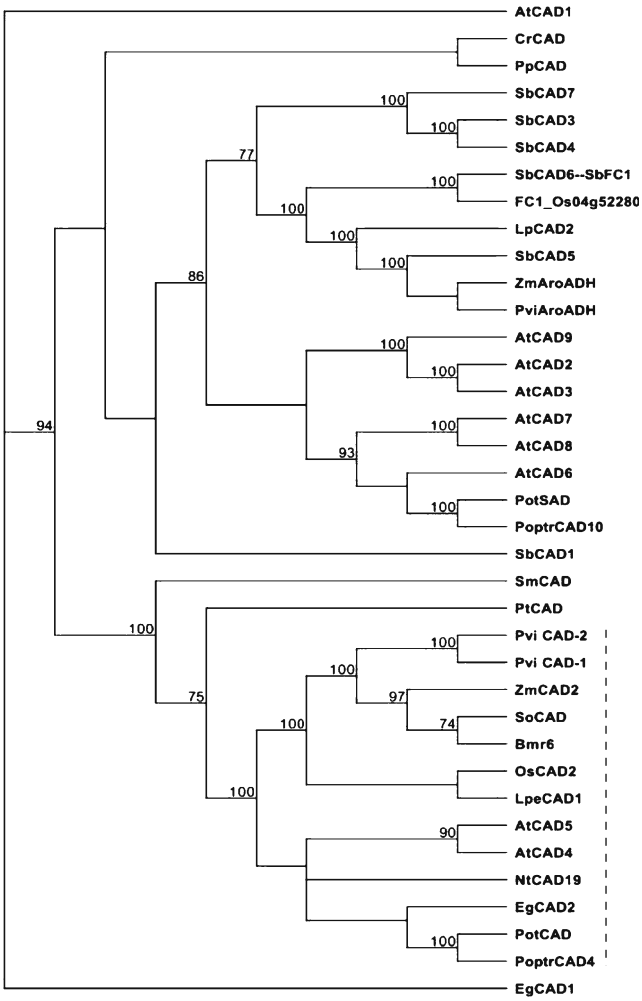

Fig. 2 Phylogenetic analysis of CAD sequences. This tree contains both putative and biochemically characterized CADs. The tree was constructed as indicated previously [43]. GenBank accession numbers are provided except for *Arabidopsis*, *Chlamydomonas*, *Physcomitrella*, *Selaginella*, and *Sorghum* sequences; in these cases, numbers from their corresponding genome projects were used. *Arabidopsis* AtCAD1 (At1g72680), AtCAD2 (At2g21730), AtCAD3 (At2g21890), AtCAD4 (At3g19450), AtCAD5 (At4g34230), AtCAD6 (At4g37970), AtCAD7 (At4g37980), AtCAD8 (At4g37990), and AtCAD9 (At4g39330); *C. reinhardtii* CrCAD (CHLREDRAFT_190510); *E. gunnii* EgCAD1 (CAA61275) and EgCAD2 (CAA46585); *L. perenne* LpeCAD1 (AF010290.1) and LpCAD2 (AF472592.1); *N. tabacum* NtCAD19 (CAA44217); *O. sativa* FC1 (Os04g52280) and OsCAD2 (Os02g0187800); *P. patens* PpCAD (87951 scaffold_163:497997..49922); *P. taeda* PtCAD (CAA86073); *P. tremuloides* PotCAD (AAF43140) and PotSAD (AAK58693); *P. trichocarpa* PoptrCAD4 (estExt_Genewise1_v1.C_LG_IX2359) and PoptrCAD10 (grail3.0004034803); *S. officinarum* (CAA13177); *S. moellendorffii* SmCAD (estExt_fgenesh2_pg.C_390191); *S. bicolor* SbCAD1 (Sb06g001430.1), SbCAD3 (Sb02g024210.1), SbCAD4 (Sb02g024190.1), SbCAD5 (Sb07g006090.1), SbCAD6 (Sb06g028240.1), SbCAD7 (Sb02g024220.1), and Bmr6 (Sb04g005950.1); and *Z. mays* ZmCAD2 (BM1 locus; NM_001112184). The *dashed line* indicates CADs that are likely to be involved in monolignol biosynthesis and corresponds to the Class I CADs identified by Barakat et al. [5]

CAD Activity in Switchgrass Internodes

Native CAD activity in switchgrass internodes and leaves was analyzed in clarified tissue homogenates. Individual internodes were analyzed because they give insight into CAD activity across different phases of stem development. In all internodes there was significantly greater activity when sinapaldehyde was used as a substrate (Fig. 3), suggesting that the CAD or CADs present in these extracts shared similar substrate preferences. Leaf and sheath extracts showed little or no activity (data not shown). CAD activity was highest in extracts from internode 1 (3.43 nmol sinapaldehyde reduced per milligram of protein per minute) and lower in all other internodes.

CAD Protein Levels in Internodes

Immunoblot results of protein extracts from Kanlow N1 internodes are shown in Fig. 4 and revealed that immunoreactive bands at the expected size of CAD were present in all internode extracts at approximately similar levels.

Proteomic Identification of PviCAD1 and PviCAD2 and Other Proteins

Two-dimensional gel electrophoresis followed by immunoblotting and mass spectrometry of selected spots was performed using protein extracts from internode 3. The

Table 3 Estimated kinetic constants

Enzyme	Substrate	K_m (μM)	V_{max} (nkat mg^{-1})	k_{cat} (s^{-1})	k_{cat}/K_m (μM^{-1} s^{-1})
PviCAD1	Sinapyl alcohol	15.6±2.5	600±30	46.4	2.98
	Coumaryl alcohol	4.2±1.1	320±18	24.7	5.88
	Coniferaldehyde	10.9±1.8	1,280±60	99.6	10.2
	Sinapaldehyde	14.5±1.8	2,530±100	196	13.6
	Coniferyl alcohol	0.98±0.1	170±2.6	13.2	13.4
PviCAD2	Coniferaldehyde	3.83±0.8	1,760±70	138	35.8
	Sinapaldehyde	9.5±2.0	2,500±140	195	20.5

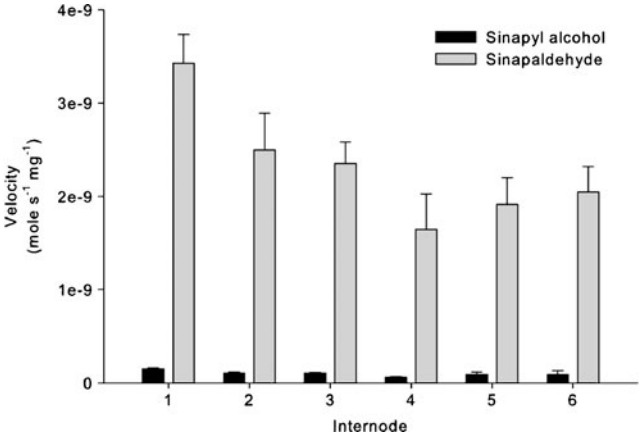

Fig. 3 CAD activity in switchgrass internodes. Internodes were numbered from least mature (top of plant; *1*) to most mature (bottom of plant; *6*). Results are from duplicate experiments. *Error bars* represent the standard deviation (n=5 or 6)

immunoreactive region of the 2-D blots were used a guide to identify putative protein spots that could contain CAD proteins. A representative gel stained with Coomassie brilliant blue, a magnified region showing the numbered spots cut out for proteomic analysis, and an immunoblots are shown (Fig. 5a–c). Spots labeled 1 through 7 shown in Fig. 5b were excised and subjected to proteomic analysis. Proteomic analyses (Table 4) showed that peptides common to both PviCAD1 and PviCAD2 proteins were present in spots 1–4 and 7, although the number of peptides attributable to PviCAD1 and 2 were found in greatest abundance in spots 1 and 2, with a lesser number of peptides in spots 3, 4, and 7. Peptides unique to PviCAD2 were only found in spots 1 and 2. The predicted C-terminal residue for PviCAD1 (FVVDVAGSNIEEQAAAAGAPAN) was identified in spots 1–3. Taken together these data suggest that the abundance of PviCAD1-related peptides was probably greater than those attributable solely to

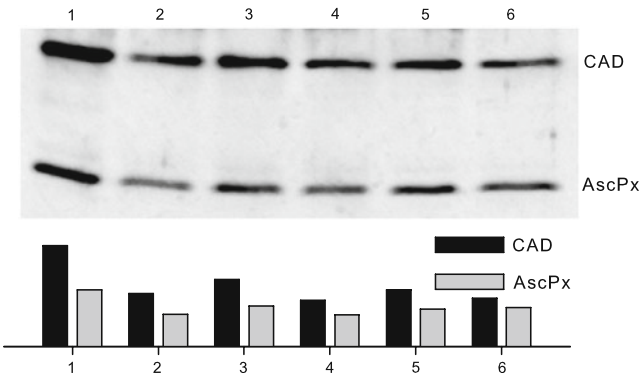

Fig. 4 Immunoblots of switchgrass internode extracts probed with CAD-specific polycolonal antibodies. Antibodies generated to ascorbate peroxidase (*AscPx*) were used a loading control. Pixel intensities for the signal arising from the two probed proteins are shown in the bottom bar graph

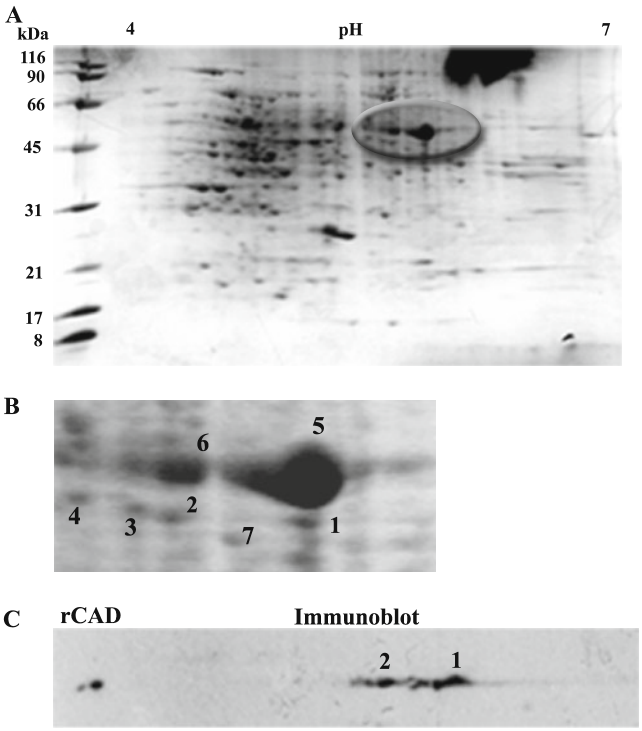

Fig. 5 **a** Two-dimensional gel of internode protein extract stained with Coomassie brilliant blue. The region of the gel selected for proteomic analyses is highlighted; **b** zoomed image of the 2-D gel with numbered protein spots that were analyzed by mass spectrometry; **c** An immunostained image from a comparable 2-D gel blotted onto nitrocellulose and developed with polyclonal antibodies to CAD is shown in the *bottom panel*. The signal arising from recombinant PviCAD1 is indicated as rCAD and was loaded in the marker lane to permit identification of potential spots that could contain native switchgrass CAD (see image above)

PviCAD2. Spots 5 and 6 were dominated by peptides arising from the large subunit of Rubisco (Table 4). Immunoblots of 2-D gels showed the presence of at least three or four immunoreactive spots at the expected size of CAD (~42–45 kDa, Fig. 5c).

Real-Time qPCR Analysis of PviCAD1, PviCAD2, and PviAroADH Transcripts

Analysis of candidate reference genes using geNorm [50] indicated that *UCE1*, *eIF4α*, and *GAPDH* would comprise a suitable set of reference genes. Using this approach, delta Ct values were normalized and relative expression levels of *PviCAD1*, *PviCAD2*, and *PviAroADH* in switchgrass internodes indicated that *PviCAD1* transcripts were considerably more abundant in internodes as compared to the transcript levels for *PviCAD2* (Fig. 6). Average *PviCAD2* transcript levels were 44-fold lower than the *PviCAD1* transcript levels in internode 1. Additionally, *PviAroADH* transcript was also consistently more abundant than *PviCAD2*. Compared to *PviCAD1*, relative

Table 4 Proteomic results

Protein	Sequence	Molecular weight	Ion score	Spot number
PviCAD1 and PviCAD2	TVVGWAAR	858.48	41	1,2
	HFGLTAPGLR	1,067.58	61	1,2,3,4
	GLTSQIEVVK	1,072.59	66	1,2,3
	ANVEQYCNK	1,124.5	42	1,2,3
	MGVYNEALER	1,180.55	83	1,2,3
	ANVEQYCNKR	1,280.59	38	1,2
	GGILGLGGVGHMGVK	1,350.75	66	1,2,3
	VGDVVGVGVIVGCCR	1,544.8	68	1,2,3,4
	FCVDKGLTSQIEVVK	1,721.9	1	1,2,3
	VLYCGICHTDIHQAK	1,813.86	52	1,2,3
	IWSYNDVYTDGRPTQGGFASTMVVDQK	3,050.46	76	1,2
	FVVPIPAGLAPEQAAPLLCAGVTVYSPLK	2,977.6	91	2,3
PviCAD1	NTGPEDVVVK	1,056.53	50	1,2,3,4
	ALGHHVTVISSSSR	1,449.77	71	1,2
	DAAGHLSPYTYTVR	1,549.75	74	1,2
	KSVTGSFIGSVDETAELLR	2,008.07	47	1,2,3,4,7
	FVVDVAGSNIEEQAAAAGAPAN	2,100.01	157	1,2,3
	HVLLGVVGEPLSFVSPMVMLGR	2,368.28	72	1,2,3
	YPMVPGHEVVGEVVEVGPEVSK	2,336.17	77	2,3,4
PviCAD2	KTGPEDVVVK	1,067.58	61	1,2
	AVTGSFIGSIDETAELLR	1,878.0	143	1,2
	HVLLGVVGEPLSFVAPMVMLGR	2,352.27	58	1,2
Ribulose 1,5,-bisphosphate	EITLGFVDLLR	1,274.72	78	5,6
	DDENVNSQPFMR	1,466.61	78	5,6
	DDFIEKDR	1,036.48	67	5,6
	VTPQPGVPPEEAGAAVAAESSTGTWTTVWTDGLTSLDR	3,583.86	101	5,6

expression values for *PviAroADH* ranged from 1.8- to 5-fold lower. Expression patterns differed among the three transcripts. *PviCAD1* was most abundant in internode 1 and displayed lower transcript abundance in all other internodes, with the lowest abundance in internode 4. *PviCAD2* transcript levels were low throughout the length of the tiller, with the highest abundance in internode 6. *PviAroADH* transcript levels were relatively constant across internodes 1–5 with the highest abundance in internode 6.

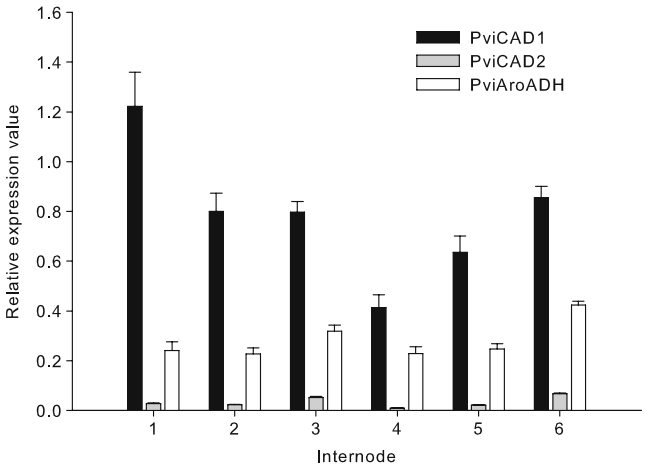

Fig. 6 Relative expression levels of *PviCAD1*, *PviCAD2*, and *PviAroDH* in switchgrass internodes. Internode numbering as described for Fig. 3. *Error bars* represent the standard deviation

Detection of Motifs Present in CAD Enzymes with Good Catalytic Activity Against Hydroxycinnamylaldehydes

Based on the phylogenetic analyses, the AtCAD5 crystal structure [55], models of sorghum Bmr6 and a related sorghum CAD-like protein [43], and the enzyme activity data presented in Fig. 3 and Table 3, we looked for specific sequence motifs that could be used to distinguish CADs that are most likely to function in lignification versus CAD-like sequences that are unlikely to have significant role in secondary cell wall formation. Alignment of several grass CADs along with AtCAD8, SbCAD5, PviAroADH, and the putative *Z. mays* AroADH (enzymes not associated with

lignification, nor displaying significant in vitro activity against monolignals) indicated that there are several key conserved residues present in CADs with documented catalytic activity against coniferaldehyde or sinapaldehyde as compared to CAD-like proteins that display poor activity against these substrates (Fig. 7). The proposed mechanism for aldehyde reduction in AtCAD5 appears to involve Thr49, His52, and Asp57 [55]. Sattler et al. [43] had shown that the equivalent residue for Asp57 in monocot CADs was a histidine. As shown in Fig. 7, monocot and dicot CADs known to be associated with cell wall lignification and displaying significant catalytic activity against monolignals have the 57HL58 or 57DL58 motif. This motif is present in PviCAD1 and PviCAD2 but not in PviAroADH. Also, key residues identified in the substrate binding domain in AtCAD5 [55] and sorghum bmr6 [43] are apparently absolutely conserved in all other CADs with a demonstrated role in cell wall lignification; these include the 118IWSY121 and 274LGV276 or 274MGV276 motifs (Fig. 7). The IWSY and LGV motifs are present in PviCAD1 and 2 but absent in PviAroADH and related proteins. In these enzymes, two critical amino acid substitutions are evident, first, the leucine next to the catalytically important His57 or Asp57 has been replaced with tryptophan, and second, the tryptophan in the IWSY motif has been replaced with phenylalanine.

PviCAD1 and PviAroADH Protein Structures

PviCAD1 and PviAroADH structures were modeled using the AtCAD5 [55] structure in order to minimize template bias in the modeling. Amino acids 6 to 356 of PviCAD1 show 71.5% sequence identity with the AtCAD5 sequence, whereas amino acids 15 to 359 of PviAroADH align with only 48.1% sequence identity to AtCAD5. The residues thought to be catalytically important appeared to have the same general conformation between PviCAD1 and AtCAD5 (Fig. 8a and b). The overlay (Fig. 8c) further shows the basic similarities and differences between these two CADs; most residues have a similar orientation. However, the orientation of the catalytically important His57 is apparently different than the Asp57 of AtCAD5. Similarly, Trp119 and Tyr121 in PviCAD1 are oriented in a manner that could improve substrate binding. Results for PviAroADH (Fig. 8d) showed that, when AtCAD5 coordinates were used as a template, the alignment produced a structure that was markedly different from the active site of either AtCAD5 or PviCAD1. Significantly, the Trp64 next to the predicted catalytically important Glu63 residue appears to protrude into the putative substrate binding pocket and the replacement of Trp119 in the CADs with high catalysis against cinnamylaldehydes with Phe125 in PviAroADH appears to also change the putative substrate binding pocket relative to PviCAD1, suggesting that PviAroADH is likely to prefer substrates other than cinnamylaldehydes.

Discussion

PviCAD1 and PviCAD2 encode for the dominant CADs involved in switchgrass lignification

CAD is involved in the phenylpropanoid pathway in plants and catalyzes the final step in monolignol biosynthesis by converting cinnamyl aldehydes to the corresponding cinnamyl alcohols [28]. Here, two CADs and one CAD-like gene were cloned from switchgrass EST resources. Although *PviCAD1* and *PviCAD2* were cloned from different cultivars, analysis of existing Kanlow CAD clones and

```
                                   * *
PviCAD1    49 T D I H Q A K N H L G A S K Y P M V P G H E---I W S Y---L G V 276
PviCAD2    49 T D I H Q A K N H L G A S K Y P M V P G H E---I W S Y---L G V 276
BMR6       49 T D I H Q A K N H L G A S K Y P M V P G H E---I W S Y---L G V 276
ZmCAD2     49 T D I H Q A K N H L G A S K Y P M V P G H E---I W S Y---L G V 276
LpeCad1    49 T D L H Q T K N H L G A S K Y P M V P G H E---I W S Y---M G V 277
AtCAD5     49 T D L H Q T K N D L G M S N Y P M V P G H E---I W S Y---M G V 276
PoptrCAD4  49 T D I H Q I K N D L G M S H Y P M V P G H E---I W S Y---M G V 276
ZmMDH      52 S D L H S I K N E W H N A T Y P M V P G H E---I F T Y---V G L 279
PviAroDH   55 S D L H F I K N E W N N A M Y P M V P G H E---I F T Y---V G L 282
SbCAD5     61 S D L H S I K N E W G N A M Y P M V P G H E---I F T Y---V G L 290
AtCAD8     48 S D L H M V K N E W G M S T Y P L V P G H E---I Q T Y---V G A 275
PtSAD      52 S D L H S I K N D W G F S M Y P L V P G H E---I L T Y---V G A 279
Poptr10    52 S D L H S I K N D W G F S M Y P L V P G H E---I L T Y---V G A 279
```

Fig. 7 Alignment of protein residues for the two switchgrass CADs as well as the putative switchgrass aromatic dehydrogenase. The alignment also includes sequences from sorghum (BMR6 and SbCAD5), maize (ZmCAD2 and ZmMDH), ryegrass (LpeCAD1), and *Arabidopsis* (AtCAD5 and AtCAD8). Alignment was done using ClustalW2. The amino acids shown include those in the active site as proposed by Youn et al. [55]

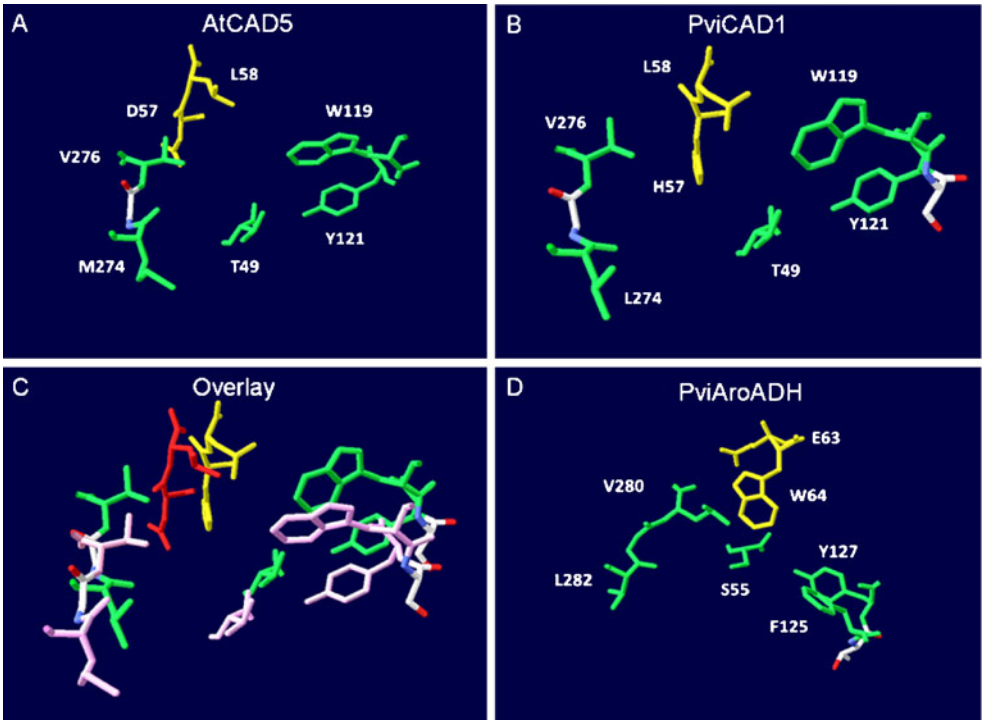

Fig. 8 Comparison of key residues in the predicted active site of AtCAD5 (**a**), PviCAD1 (**b**), and PviAroDH (**d**). Overlay of AtCAD5 (*pink/yellow residues*) and PviCAD1 (*green/red residues*) is shown in **c**. PviCAD1 and PviAroDH structures were based off of the AtCAD5 crystal structure. Only residues that are likely to be involved in substrate docking or the catalytic reaction are shown

proteomic data, especially identification of the PviCAD1 C-terminal peptide indicated that PviCAD1 sequence was identical in both cultivars. Despite the fact that many genes coding for CADs or CAD-like proteins are found in all currently known plant genomes (for example: [5, 22]), genetic and enzyme activity data from several studies support the notion that one or at most a few CADs actually function in cell wall lignification [14, 38, 43, 46, 56]. The data presented here indicate that a similar situation exists in switchgrass, where *PviCAD1*and *PviCAD2* encode functional CAD proteins that appear to be involved in switchgrass lignin biosynthesis.

CAD activity and protein was detected in all internodes, which indicated that CAD is active in lignifying tissues at different stages of maturity. These data were consistent with earlier studies that had shown lignification is greater in more mature internodes suggesting that cell wall lignification continues to take place across all stages of internode maturity. Our studies focused on the stems since these tissues comprise a significant proportion of biomass at harvest, and represent tissues with the most lignin content. Ultimately, improving quality parameters in switchgrass biomass for liquid fuels will probably result in significant reduction in stem lignin content [39].

Immunoblot results from 2-D gel electrophoresis indicated that different forms of CAD may exist in switchgrass internodes, although these differences were not fully identified. Multiple spots can be artifacts due to processes such as deamidation which are frequently encountered during sample preparation for 2-D gel analyses [41] and it is unclear if differences included CAD isozymes, post-translational modifications, or were simply artifacts. However, our polyclonal antibodies will detect other related CAD-like proteins in tissue extracts [43]. Proteomic analyses revealed a greater abundance of peptides matching PviCAD1 suggesting that PviCAD1 was present at higher levels or peptides derived with PviCAD1 were preferentially enriched during proteomic analyses. Enrichment of PviCAD1 relative to PviCAD2 was supported by qPCR data that showed greater transcript abundance for mRNA coding for *PviCAD1* relative to mRNA coding for *PviCAD2*. However, these results do not preclude the possibility that PviCAD2 may be present in specific tissues or have other temporally or spatially defined developmental roles as was suggested for CADs in aspen [27] and other species [5, 24]. Kanlow switchgrass is an allotetraploid cultivar [1, 33] and it is possible that PviCAD1 and PviCAD2 are encoded by the two different genomes. In many polyploid species, translation of alleles is not equal, and dominance by one or the other genome is observed [1, 13]. Our data suggests that *PviCAD1* and *PviCAD2* transcripts are present in all internodes. Furthermore, these tissues appear to contain active CAD proteins which could participate in the continued lignification of tissues such as the stem sclerenchyma and parenchyma. Previously published results show that these tissues are progressively more lignified in the basal and more mature internodes as compared to the topmost internode [39]. Future functional genomic studies in switchgrass should allow greater insights into these phenomena.

Relationships of PviCAD1, PviCAD2, and PviAroADH to Other CAD and CAD-Like Proteins

PviCAD1 and PviCAD2 displayed a preference for aldehyde substrates, which was unsurprising because these enzymes are expected to catalyze the final step in monolignol biosynthesis where cinnamaldehydes are converted to their corresponding alcohols. Although PviCAD1 velocity and turnover (k_{cat}) values were higher for the aldehyde substrates, the two lowest K_m values were for coniferyl and coumaryl alcohol, respectively. The low PviCAD1 K_m observed with coniferyl alcohol results in a higher k_{cat}/K_m value when compared to the other alcohols. Otherwise, PviCAD1 efficiency with the aldehydes was higher than with alcohols. PviCAD2 was even more efficient with aldehyde substrates than PviCAD1, although the exact localization and temporal control of these two enzymes are not currently known. A similar preference for aldehyde substrates, particularly sinapaldehyde has been reported for sorghum and rice CADs [43, 56]. However, reported kinetic data for other CADs did not necessarily show this type of strong preference (Table 5). In *Arabidopsis*, AtCAD5 had only a 12.4% higher V_{max} for sinapaldehyde [22]; in *Eucalyptus gunnii* the velocity increase was 31–33%, depending upon the CAD subunit composition [19]. In contrast, CAD velocity on sinapaldehyde was markedly lower in *Populus euramericana* [42], *Picea abies* [16, 29], *Populus tremoloides* [27], and *Pinus thunbergii* [25]. Outside of monocots, the only other kinetically

Table 5 Reported CAD kinetic parameters

Species	Substrate	K_m (μM)	V_{max} (nkat mg^{-1})	Reference
Pinus thunbergii	Coniferaldehyde	9.1	3.3	[25]
	Sinapaldehyde	–	0.073	
	p-coumaraldehyde	30	5.7	
	Cinnamaldehyde	14	7.6	
Picea abies	Coniferaldehyde	3.6	1,724	[29]
	Sinapaldehyde	83	167	
	4-coumaraldehyde	12.5	2,857	
Forsythia suspensa	Cinnamyl alcohol	156	39%[e]	[31]
	p-coumaryl alcohol	132	112%	
	Coniferyl alcohol	32	100%	
Eucalyptus gunnii	Coniferaldehyde	4.5[a]/5.2[b]	2,104[a]/1,580[b]	[19]
	Sinapaldehyde	6.8/2.5	2,758/2,107	
	p-coumaraldehyde	5.1/16	1,244/2,107	
	Coniferyl alcohol	2.3/23	539/1,817	
	Sinapyl alcohol	6.6/4.5	717/562	
	p-coumaryl alcohol	35/64	812/1,150	
Populus euramericana	Coniferaldehyde	0.77	52.2	[42]
	p-coumaraldehyde	1.2	17	
	Sinapaldehyde	4.8	19.5	
Populus tremoloides	Coniferaldehyde	2.3[c]	3.5[c]	[27]
	Sinapaldehyde	9.1	1.7	
	p-coumaraldehyde	6.2	2.8	
	Caffealdehyde	37	2.5	
	5-OH coniferaldehyde	17.5	2.8	
Arabidopsis thaliana (AtCAD5)[d]	Coniferaldehyde	35	157.4	[22]
	Sinapaldehyde	20	177	
	p-coumaraldehyde	13	187.3	
	Caffeylaldehyde	68	94.1	
	5-OH coniferaldehyde	22	106.9	
Medicago sativa[d]	Coniferaldehyde	1.3	47.8%[e]	[9]
	Sinapaldehyde	6.9	67.2%	
	Cinnamaldehyde	9.2	100%	
Oryza sativa[d]	Coniferaldehyde	4.4	32.9	[56]
	Sinapaldehyde	20.8	87.0	
Lolium perenne L.	Coniferaldehyde	1.9	57.8%[e]	[32]
	Sinapaldehyde	6.3	75.7%	
	Cinnamaldehyde	5.8	100%	

[a] HeteroCAD
[b] HMW HomoCAD
[c] PtCAD
[d] Recombinant protein
[e] Only relative values were reported

characterized enzyme for which a relatively similar increase in activity was observed is PtSAD [27], although questions have been raised about its proposed specific physiological role in angiosperms [2].

A recent analysis of *Populus trichocarpa* CAD genes suggested that the CAD protein sequences present in the *Populus* genome could be separated into three classes [5]. Based on these classes, *PviCAD1* and *PviCAD2* fall into Class I, which includes *PoptrCAD4* and *AtCAD5* and further indicates the broad conservation of the lignin-pathway related CADs in plants. The phylogenetic analysis reported here places PviCAD1 and PviCAD2 into a distinct clade that contains all of the other known well-characterized CADs that have a role in catalyzing the last step in monolignol biosynthesis and are thus bona fide CAD genes. Also, this group further divides into two major sub-groups encompassing the monocots and dicots; *PviCAD1* and *PviCAD2* fall into a group that contains all of the other known monocot CADs, including the recently reported sorghum BMR6 [43]. However, *PviAroADH*, *PtSAD*, and its apparent *Populus* ortholog *Poptr10* are not included within this clade which is in agreement with Barakat et al. [5], and suggests these genes may not be involved in lignin monomer synthesis or may have other, more primary, roles such as plant defense or catalytic activity under specific conditions [8, 46]. We have been unable to produce soluble, active recombinant PviAroADH in *E. coli* and could not directly characterize this protein. A sorghum ortholog (Sb_02g024190; 61% identity, 5e-116) of PviAroADH displayed poor activity against monolignol and monolignal substrates [43], suggesting that PviAroADH could exhibit similar properties. Future biochemical evaluation of recombinant PviAroADH should clarify these points and indicate if it could participate in the lignification process.

Sequence Motifs of Lignifying CADs and Predicted Model of PviCAD1

Sequence alignment of binding pocket and active site residues based on AtCAD5 [55], showed that among lignifying CADs (i.e., CADs shown by activity and/or genetics to function in the biosynthesis of monolignols), most amino acids were highly conserved, with the exception that in monocots His57 is present instead of the dicot Asp57 [43].

When the CAD-like sequences including PviAroADH, PtSAD, and AtCAD8 were compared to bona fide CADs, some differences were evident. First, PviCAD1 and other Class I (lignin) CADs have an invariant Gln53, which is more variable at the equivalent position in other CAD-like sequences. Second, the lignin CADs have either a His57/Leu58 or Asp57/Leu58 motif, while the CAD-like sequences contain either Glu/Trp or Asp/Trp at equivalent positions. Finally, lignin CADs contain an invariant Trp119 which is not conserved in the other CADs. In examining the binding pocket models for PviCAD1, AtCAD5, and PviAroADH, the differences are more apparent (Fig. 8a–d). Trp119, which resides towards the back of the binding pocket, likely stabilizes the aromatic ring of cinnamaldehydes through pi-bonding; this residue has a similar conformation in both PviCAD1 and AtCAD5 (Fig. 8a and b, respectively). Additionally, Youn et al. [55] noted that Trp119 and Phe299 shrunk the size of the substrate binding pocket compared to PtSAD (Leu122/Gly302), which likely resulted in AtCAD5 having increased substrate specificity. These residues (Trp119/Phe299) are replaced by Phe125 and Cys305 in PviAroADH. Also, His57 and Leu58 in PviCAD1 are replaced by Glu63 and Trp64 in PviAroADH. Taken together, these amino acid changes could dramatically alter the binding pocket conformation of PviAroADH (Fig. 8d) and, although a member of the alcohol dehydrogenase superfamily, this protein is likely to have a biological role that is different than PviCAD1 in switchgrass.

Given that switchgrass is a good candidate feedstock for biorefineries [36, 40] and that CAD mutants have been shown to impact lignin composition and conversion to ethanol in sorghum [15, 34, 43], the data presented here will be useful for understanding and manipulating lignin synthesis for future development of switchgrass bioenergy cultivars.

Acknowledgments This work was supported by the USDA-ARS CRIS project 5440-21000-028-00D and in part by the Office of Science (BER), US Department of Energy grant number DE-AI02-09ER64829. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture.

References

1. Adams KL, Wendel JF (2005) Allele-specific, bidirectional silencing of an alcohol dehydrogenase gene in different organs of interspecific diploid cotton hybrids. Genetics 171:2139–2142
2. Anterola AM, Lewis NG (2002) Trends in lignin modification: a comprehensive analysis of the effects of genetic manipulations/mutations on lignification and vascular integrity. Phytochemistry 61:221–294
3. Antizar-Ladislao B, Turrion-Gomez JL (2008) Second-generation biofuels and local bioenergy systems. Biofuel Bioprod Bior 2:455–469
4. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22:195–201
5. Barakat A, Bagniewska-Zadworna A, Choi A, Plakkat U, DiLoreto DS, Yellanki P et al (2009) The cinnamyl alcohol dehydrogenase gene family in *Populus:* phylogeny, organization, and expression. BMC Plant Biol 9:26

6. Baucher M, Bernard-Vailhe MA, Chabbert B, Besle JM, Opsomer C, Van Montagu M et al (1999) Down-regulation of cinnamyl alcohol dehydrogenase in transgenic alfalfa (*Medicago sativa* L.) and the effect on lignin composition and digestibility. Plant Mol Biol 39:437–447
7. Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol 54:519–546
8. Boudet AM, Hawkins S, Rochange S (2004) The polymorphism of the genes/enzymes involved in the last two reductive steps of monolignol synthesis: what is the functional significance? CR Biol 327:837–845
9. Brill EM, Abrahams S, Hayes CM, Jenkins CLD, Watson JM (1999) Molecular characterisation and expression of a wound-inducible cDNA encoding a novel cinnamyl-alcohol dehydrogenase enzyme in lucerne (*Medicago sativa* L.). Plant Mol Biol 41:279–291
10. Carroll A, Somerville C (2009) Cellulosic biofuels. Annu Rev Plant Biol 60:165–182
11. Casler MD, Buxton DR, Vogel KP (2002) Genetic modification of lignin concentration affects fitness of perennial herbaceous plants. Theor Appl Genet 104:127–131
12. Chapple C, Ladisch M, Meilan R (2007) Loosening lignin's grip on biofuel production. Nat Biotechnol 25:746–748
13. Chaudhary B, Flagel L, Stupar RM, Udall JA, Verma N, Springer NM et al (2009) Reciprocal silencing, transcriptional bias and functional divergence of homeologs in polyploid cotton (*Gossypium*). Genetics 182:503–517
14. Chen F, Dixon RA (2007) Lignin modification improves fermentable sugar yields for biofuel production. Nat Biotechnol 25:759–761
15. Dien B, Sarath G, Pedersen J, Sattler S, Chen H, Funnell-Harris D et al (2009) Improved sugar conversion and ethanol yield for forage sorghum (*Sorghum bicolor* L. Moench) lines with reduced lignin contents. BioEnergy Res 2:153–164
16. Galliano H, Heller W, Sandermann H (1993) Ozone induction and purification of spruce cinnamyl alcohol dehydrogenase. Phytochemistry 32:557–563
17. Grabber JH, Mertens DR, Kim H, Funk C, Lu FC, Ralph J (2009) Cell wall fermentation kinetics are impacted more by lignin content and ferulate cross-linking than by lignin composition. J Sci Food Agric 89:122–129
18. Halpin C, Holt K, Chojecki J, Oliver D, Chabbert B, Monties B et al (1998) Brown-midrib maize (bm1)—a mutation affecting the cinnamyl alcohol dehydrogenase gene. Plant J 14:545–553
19. Hawkins SW, Boudet AM (1994) Purification and characterization of cinnamyl alcohol dehydrogenase isoforms from the periderm of *Eucalyptus gunnii* Hook. Plant Physiol 104:75–84
20. Kayser JPR, Kim JG, Cerny RL, Vallet JL (2006) Global characterization of porcine intrauterine proteins during early pregnancy. Reproduction 131:379–388
21. Kim H, Ralph J, Lu FC, Ralph SA, Boudet AM, MacKay JJ et al (2003) NMR analysis of lignins in CAD-deficient plants. Part 1. Incorporation of hydroxycinnamaldehydes and hydroxybenzaldehydes into lignins. Org Biomol Chem 1:268–281
22. Kim S-J, Kim M-R, Bedgar DL, Moinuddin SGA, Cardenas CL, Davin LB et al (2004) Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in *Arabidopsis*. Proc Natl Acad Sci U S A 101:1455–1460
23. Kim SJ, Kim KW, Cho MH, Franceschi VR, Davin LB, Lewis NG (2007) Expression of cinnamyl alcohol dehydrogenases and their putative homologues during *Arabidopsis thaliana* growth and development: lessons for database annotations? Phytochemistry 68:1957–1974
24. Koutaniemi S, Warinowski T, Kärkönen A, Alatalo E, Fossdal C, Saranpää P et al (2007) Expression profiling of the lignin biosynthetic pathway in Norway spruce using EST sequencing and real-time RT-PCR. Plant Mol Biol 65:311–328
25. Kutsuki H, Shimada M, Higuchi T (1982) Regulatory role of cinnamyl alcohol dehydrogenase in the formation of guaiacyl and syringyl lignins. Phytochemistry 21:19–23
26. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685
27. Li LG, Cheng XF, Leshkevich J, Umezawa T, Harding SA, Chiang VL (2001) The last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase. Plant Cell 13:1567–1585
28. Li X, Weng JK, Chapple C (2008) Improvement of biomass through lignin modification. Plant J 54:569–581
29. Luderitz T, Grisebach H (1981) Enzymic synthesis of lignin precursors comparison of cinnamoyl-CoA reductase and cinnamyl alcohol: NADP+ dehydrogenase from spruce (*Picea abies* L.) and soybean (*Glycine max* L.). Eur J Biochem 119:115–124
30. Ma QH (2010) Functional analysis of a cinnamyl alcohol dehydrogenase involved in lignin biosynthesis in wheat. J Exp Bot 61:2735–2744
31. Mansell RL, Gross GG, Stockigt J, Franke H, Zenk MH (1974) Purification and properties of cinnamyl alcohol dehydrogenase from higher plants involved in lignin biosynthesis. Phytochemistry 13:2427–2435
32. McAlister FM, Lewis-Henderson WR, Jenkins CLD, Watson JM (2001) Isolation and expression of a cinnamyl alcohol dehydrogenase cDNA from perennial ryegrass (*Lolium perenne*). Aust J Plant Physiol 28:1085–1094
33. Okada M, Lanzatella C, Saha MC, Bouton J, Wu R, Tobias CM (2010) Complete switchgrass genetic maps reveal subgenome collinearity, preferential pairing, and multilocus interactions. Genetics 185:745–760
34. Palmer NA, Sattler SE, Saathoff AJ, Funnell D, Pedersen JF, Sarath G (2008) Genetic background impacts soluble and cell wall-bound aromatics in brown midrib mutants of sorghum. Planta 229:115–127
35. Pauly M, Keegstra K (2008) Cell-wall carbohydrates and their modification as a resource for biofuels. Plant J 54:559–568
36. Perlack R, Wright L, Turhollow A, Graham R, Stokes B, Erbach D (2005) Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply. In: Energy USDoE (ed), Oak Ridge, TN, p 76
37. Pillonel C, Mulder MM, Boon JJ, Forster B, Binder A (1991) Involvement of cinnamyl alcohol dehydrogenase in the control of lignin formation in *Sorghum bicolor* L. Moench Planta 185:538–544
38. Saballos A, Ejeta G, Sanchez E, Kang C, Vermerris W (2009) A genomewide analysis of the cinnamyl alcohol dehydrogenase family in sorghum [*Sorghum bicolor* (L.) Moench] ddentifies SbCAD2 as the brown midrib6 gene. Genetics 181:783–795
39. Sarath G, Baird LM, Vogel KP, Mitchell RB (2007) Internode structure and cell wall composition in maturing tillers of switchgrass (*Panicum virgatum* L.). Bioresour Technol 98:2985–2992
40. Sarath G, Mitchell RB, Sattler SE, Funnell D, Pedersen JF, Graybosch RA et al (2008) Opportunities and roadblocks in utilizing forages and small grains for liquid fuels. J Ind Microbiol Biotech 35:343–354
41. Sarioglu H, Lottspeich F, Walk T, Jung G, Eckerskorn C (2000) Deamidation as a widespread phenomenon in two-dimensional polyacrylamide gel electrophoresis of human blood plasma proteins. Electrophoresis 21:2209–2218
42. Sarni F, Grand C, Boudet AM (1984) Purification and properties of cinnamoyl-CoA reductase and cinnamyl alcohol dehydrogenase from poplar stems (*Populus X euramericana*). Eur J Biochem 139:259–265
43. Sattler SE, Saathoff AJ, Haas EJ, Palmer NA, Funnell-Harris DL, Sarath G et al (2009) A nonsense mutation in a cinnamyl alcohol dehydrogenase gene is responsible for the sorghum brown midrib6 phenotype. Plant Physiol 150:584–595

44. Schmer MR, Vogel KP, Mitchell RB, Perrin RK (2008) Net energy of cellulosic ethanol from switchgrass. Proc Natl Acad Sci U S A 105:464–469
45. Selman-Housein G, Lopez MA, Hernandez D, Civardi L, Miranda F, Rigau J et al (1999) Molecular cloning of cDNAs coding for three sugarcane enzymes involved in lignification. Plant Sci 143:163–171
46. Sibout R, Eudes A, Mouille G, Pollet B, Lapierre C, Jouanin L et al (2005) CINNAMYL ALCOHOL DEHYDROGENASE-C and -D are the primary genes involved in lignin biosynthesis in the floral stem of *Arabidopsis*. Plant Cell 17:2059–2076
47. Somssich IE, Wernert P, Kiedrowski S, Hahlbrock K (1996) *Arabidopsis thaliana* defense-related protein ELI3 is an aromatic alcohol:NADP(+) oxidoreductase. Proc Natl Acad Sci U S A 93:14199–14203
48. Suzuki Y, Kawazu T, Koyama H (2004) RNA isolation from siliques, dry seeds, and other tissues of *Arabidiopsis thaliana*. Biotechniques 37:542–544
49. Tobias CM, Twigg P, Hayden DM, Vogel KP, Mitchell RM, Lazo GR et al (2005) Analysis of expressed sequence tags and the identification of associated short tandem repeats in switchgrass. Theor Appl Genet 111:956–964
50. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3 http://genomebiology.com/2002/3/7/research/0034.1
51. Vogel J (2008) Unique aspects of the grass cell wall. Curr Opin Plant Biol 11:301–307
52. Vogel KP, Mitchell KB (2008) Heterosis in switchgrass: biomass yield in swards. Crop Sci 48:2159–2164
53. Wyrambik D, Grisebach H (1975) Purification and properties of isoenzymes of cinnamyl-alcohol dehydrogenase from soybean-cell-suspension cultures. Eur J Biochem 59:9–15
54. Xu CP, Xu Y, Huang BR (2008) Protein extraction for two-dimensional gel electrophoresis of proteomic profiling in turfgrass. Crop Sci 48:1608–1614
55. Youn B, Camacho R, Moinuddin SGA, Lee C, Davin LB, Lewis NG et al (2006) Crystal structures and catalytic mechanism of the *Arabidopsis* cinnamyl alcohol dehydrogenases AtCAD5 and AtCAD4. Org Biomol Chem 4:1687–1697
56. Zhang KW, Qian Q, Huang ZJ, Wang YQ, Li M, Hong LL et al (2006) GOLD HULL AND INTERNODE2 encodes a primarily multifunctional cinnamyl-alcohol dehydrogenase in rice. Plant Physiol 140:972–983