Investigating the Role of Extensin Proteins in Poplar Biomass Recalcitrance

Margaret B. Fleming,[a,b] Stephen R. Decker,[c,d] and Patricia A. Bedinger [a,*]

The biological conversion of cellulosic biomass to biofuel is hindered by cell wall recalcitrance, which can limit the ability of cellulases to access and break down cellulose. The purpose of this study was to investigate whether hydroxyproline-rich cell wall proteins (extensins) are present in poplar stem biomass, and whether these proteins may contribute to recalcitrance. Three classical extensin genes were identified in *Populus trichocarpa* through bioinformatic analysis of poplar genome sequences, with the following proposed names: *PtEXTENSIN1* (Potri.001G019700); *PtEXTENSIN2* (Potri.001G020100); *PtEXTENSIN3* (Potri.018G050100). Tissue print immunoblots localized the extensin proteins in poplar stems to regions near the vascular cambium. Different thermochemical pretreatments reduced but did not eliminate hydroxyproline (Hyp, a proxy for extensins) from the biomass. Protease treatment of liquid hot water-pretreated poplar biomass reduced Hyp content by a further 16% and increased subsequent glucose yield by 20%. These data suggest that extensins may contribute to recalcitrance in pretreated poplar biomass, and that incorporating protease treatment into pretreatment protocols could result in a small but significant increase in the yield of fermentable glucose.

Keywords: Extensin; Hydroxyproline-rich glycoprotein; Cellulosic biofuel; Biomass recalcitrance; Acid fungal protease; Poplar; Pretreatment

Contact information: a: Biology Department, Colorado State University, Campus Delivery 1878, Fort Collins, CO 80523-1878 USA; b: USDA Agricultural Research Service, National Laboratory for Genetic Resources Preservation, 1111 South Mason St, Fort Collins, CO 80524, USA; c: Biosciences Center, National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO, 80401 USA; d: BioEnergy Science Center (BESC), National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO, 80401 USA;
** Corresponding author: Patricia.Bedinger@colostate.edu*

INTRODUCTION

Plant cell walls are complex structures composed of polysaccharides (cellulose, hemicelluloses, and pectin), lignin, and structural proteins (Carpita and McCann 2000). Cellulose, a polymer of glucose, is of high interest as a renewable energy source for biofuel production. Cellulose can be broken down by cellulases into glucose, which can then be converted into fuel (Béguin and Aubert 1994). However, cell wall recalcitrance, a result of the crystalline structure of cellulose and other aspects of cell wall architecture, can limit how well cellulases can access and digest cellulose (Himmel 2008).

One way to address cell wall recalcitrance is to pretreat biomass before attempting the enzymatic digestion of polysaccharides. Commonly used thermochemical pretreatments disrupt the cell wall structure both physically and chemically. Typical pretreatments include dilute acid, alkaline peroxide, liquid hot water (LHW), ammonia steam explosion, ammonia fiber expansion, and ionic liquid (Chaturvedi and Verma 2013).

Pretreatments target lignin and hemicelluloses, which are abundant components of the cell wall and are known contributors to recalcitrance (Kumar *et al*. 2009). Pretreatment increases glucose yield, but yields must increase further for biofuels to be cost-competitive with conventional fossil fuels (Yang and Wyman 2008).

Structural cell wall proteins are a minor component of cell walls, comprising no more than 1% of wood (extrapolated from Dill *et al*. 1984). Nevertheless, these proteins could contribute to recalcitrance; the role of structural cell wall proteins in biomass digestibility has not yet been investigated. Extensins are structural cell wall proteins forming a sub-group of the hydroxyproline-rich glycoproteins (HRGPs). Extensins contain at least two repeats of a diagnostic SPPP or SPPPP primary sequence motif (Showalter 1993). The primary amino acid sequence directs substantial post-translational changes, including hydroxylation of select prolines to hydroxyproline (Hyp) (Lamport 1965; Kieliszewski *et al*. 2011), the addition of an arabinosyl side-chain (containing 3-5 arabinoses) to Hyp (Lamport 1967; Lamport and Miller 1971), and the addition of galactose to serine (Lamport *et al*. 1973). The mature protein has a rigid, outstretched polyproline-2 helix because of the series of proline and Hyp residues, as well as the protective and stabilizing β-L-arabinofuranoside "sheath" around the polypeptide chain (van Holst and Varner 1984; Stafstrom and Staehelin 1986). Extensins secreted into the cell wall are initially soluble and salt-elutable, but they eventually become cross-linked within the cell wall (Heath and Northcote 1971; Fry 1982; Cooper and Varner 1984).

Extensin cross-linking has been proposed to occur between several different molecule pairs. First, extensin-pectin cross-links have been found in cotton (Qi *et al*. 1995), sugarbeet pectin extracts (Masuda *et al*. 1989; Nuñez *et al*. 2009), and soybean suspension-cultured cells (Moore 1973). The linkage has not yet been characterized. Extensin-extensin cross-links have been formed *in vitro* (Held *et al*. 2004), and recent evidence indicates extensin-extensin crosslinks occur *in vivo* (Cegelski *et al*. 2010), most likely by condensation of tyrosine residues into isodityrosine or di-isodityrosine. Finally, extensin-lignin cross-links have been identified in soybean suspension-cultured cells (Moore 1973), callus tissue of *Pinus elliottii* (Whitmore 1982), and in elicitor-challenged suspension-cultured pine cells (Lange *et al*. 1995). Recent *in vitro* experiments provide evidence for the mechanism of protein-lignin interactions, and demonstrate cross-linking of the amino acids C, Y, and T with coniferyl alcohol (the subunit of G-lignin) (Cong *et al*. 2013).

Once extensins are covalently cross-linked within cell walls, they are extremely difficult to remove (for example, Heath and Northcote 1971; Mort and Lamport 1977; Lamport 1980). Although the protein content of wood is low (experimentally determined N contents of wood range from 0.03 to 0.1% (Cowling and Merrill 1966)), intractably cross-linked proteins within wood may have a small but significant role in biomass recalcitrance. For example, LHW pretreatment mainly affects pectin and hemicellulose, with some lignin extraction as well. The extent of these modifications is dependent on both time and temperature. As the pectin-extensin cross-linking chemistry is not yet understood, it is difficult to postulate whether LHW pretreatment cleaves this bond. In the case of lignin-extensin cross-links, although partial removal of lignin through extraction by LHW occurs, the lignin-extensin crosslink may also limit the extent of this extraction. Cross-linking may also impede the complete enzymatic breakdown of cellulose by physically blocking the advancement of processive cellulases.

Extensin proteins in poplar stem biomass were investigated in this study. Three classical extensin genes were found using bioinformatic analysis of poplar genome sequences, and extensin proteins were localized in poplar stems using tissue print

immunoblots. The removal of extensins from poplar biomass by thermochemical pretreatments was assessed by measuring Hyp content in the biomass after pretreatment. Finally, the effect of protease treatment on the release of glucose from pretreated poplar biomass was analyzed.

EXPERIMENTAL

Bioinformatic Analysis of Poplar Extensins

The 37 extensins found by Guo *et al.* (2014) in *Populus trichocarpa* were originally identified because they contained two or more repeats of SP_3 or SP_4, as well as a signature extensin protein domain in Interpro (IPR006706, IPR006041, IPR003882, IPR003883, PR01217, or PTHR23201). In this study, these 37 genes were further analyzed to identify "classical" extensin genes, as defined in Showalter *et al*. (2010) and Kieliszewski *et al*. (2011). SignalP 4.1 (www.cbs.dtu.dk/services/SignalP/) was used with its default settings to predict the presence of a signal peptide (Petersen *et al*. 2011). In genes with predicted signal peptides, protein sequence homology to annotated extensin proteins was identified with HMMER3, using a phmmer search of the SwissProt database (hmmer.janelia.org/) (Finn *et al*. 2011). Genes with additional domains (chimeric extensins) or pollen Ole e 1 allergen/extensin domains were excluded, as were those with similarities to non-extensin HRGPs (such as arabinogalactan proteins or proline-rich proteins). Genes with a signal peptide and a predicted repetitive extensin domain but no additional functional domains were classified as classical extensins.

Introns were identified in classical extensin genes using PopGenIE 3.0 (popgenie.org) (Sjödin *et al.* 2009). The organ with the highest expression for each gene was identified using the PopGenIE 3.0 exPlot tool, based on expression values in the *Populus balsamifera* developmental tissue series (Wilkins *et al*. 2009) and the *P. trichocarpa* tissue series (Yang *et al*. 2008). *P. trichocarpa* transcriptome data from Hefer *et al.* (2015) were also considered.

Tissue Print Immunoblots

Tissue prints were made as described in Cassab and Varner (1989). Pieces of nitrocellulose paper were soaked for 30 min in 0.2 M $CaCl_2$ and dried. Freehand, 1- to 3-mm cross-sections of stems of *Populus alba* x *Populus tremuloides* grown in hydroponic media were made with a double-edged razor blade. The sections were rinsed in distilled water for 3 s, dried with a KimWipe, and firmly pressed for 20 s onto the prepared nitrocellulose.

Printed nitrocellulose was air-dried, blocked in 5% non-fat milk in Tris-buffered saline with Tween (TBST: 10 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.05% Tween® 20) overnight at 4 °C, and washed five times for 10 min in TBST. Prints were treated for 3 h at room temperature with one of three different antibodies diluted in TBST according to the recommendations of the providers: a 1:10 dilution of mAb JIM20 (CarboSource, Athens, GA, USA), a 1:10 dilution of mAb LM1 (PlantProbes, Leeds, UK), or a 1:500 dilution of polyclonal carrot extensin-1 antibody, gE-1 (a kind gift from Dr. L. A. Staehelin). Tissue prints were washed four times for 15 min in TBST at room temperature, then treated overnight at 4 °C with a 1:10,000 dilution in TBST of goat anti-rat alkaline-phosphatase conjugated secondary antibody (Sigma-Aldritch, USA) for detecting mAbs JIM20 and LM1, or goat anti-rabbit alkaline-phosphatase conjugated secondary antibody

(Sigma) for detecting gE-1. Tissue prints were then washed with TBST and developed with 5-bromo-4-chloroindoxyl-phosphate and nitroblue-tetrazolium as previously described (Cassab and Varner 1989). For controls, several tissue prints were treated with the secondary antibody only or with the alkaline phosphatase substrates directly after blocking; no signal was detected in either control treatment. The developed tissue prints were imaged with an Epson Perfection V700 Photo flat-bed scanner at 12,800 dpi (Epson, USA).

The cross-sections that were used for the tissue prints were then dipped for 5 seconds in 0.25% Toluidine Blue O (Electron Microscopy Sciences, USA, #22050), mounted in glycerol, and imaged with a Leica 5500 microscope using a Leica DFC 450 color camera with Leica Application Suite V4.1 software (Leica Microsystems, Germany). Images were composited using Image Composite Editor (http://research.microsoft.com/en-us/um/redmond/groups/ivm/ice/).

Pretreatment of Poplar Biomass

Dried, milled poplar biomass (described in Selig *et al*. 2010) was pretreated with LHW or dilute acid at the National Renewable Energy Laboratory (NREL, Golden, CO) in a 4-L ZC® vertically stirred reactor (ZipperClave,®Autoclave Engineers, Erie, PA, USA). Pretreatments were performed separately, as previous work indicates that combinations have no additive effect (Cara *et al.* 2007). Steam was directly injected into the bottom of the reactor through ports in a rotary-plow type of agitator, and constant temperature was achieved by controlling the steam pressure in the reactor. The ZC reactor has an electrical heating blanket set at reaction temperature to lessen steam condensation from heat losses through the reactor wall. The contents of the ZC reactor typically reached reaction temperature within 5 to 10 s of starting the steam flow as measured by two thermocouples, one in the bottom and one near the middle of the reactor. After pretreatment, the steam pressure was slowly released through a condenser over a period of 15 to 30 s to lessen boil-over. In LHW pretreatment, poplar was mixed with water, loaded into the ZipperClave®, and treated at 180 °C for 40 min. In acid pretreatment, poplar was mixed with 0.5% sulfuric acid, impregnated into the biomass under vacuum with mixing, loaded into the ZipperClave® pretreatment reactor, and treated at 150 °C for 20 min. In alkaline peroxide pretreatment, 4 g of poplar biomass was mixed with 100 mL of 1% hydrogen peroxide solution (pH 11.5, adjusted with 12 M NaOH) and incubated with vigorous shaking at 65 °C for 3 h (Selig *et al.* 2009). After all pretreatments, the biomass was extensively washed with deionized water and dried before use in further experiments.

Compositional Analysis of Biomass

Each type of pretreated biomass (pretreatment: none, LHW, dilute acid, or alkaline peroxide) was analyzed using a scaled-down version of the NREL standard Laboratory Analytical Procedures for the compositional analysis of structural carbohydrates, lignin, protein, and ash (http://www.nrel.gov/biomass/pdfs/42618.pdf). Briefly, 100 mg of biomass was hydrolyzed in 72% (w/w) H_2SO_4 for 2 h at 30 °C. After dilution with distilled water to 4% (w/w) H_2SO_4, the samples were autoclaved for 60 min at 121 °C to hydrolyze oligosaccharides formed in the first stage. After neutralization with $CaCO_3$, monomeric sugars were quantified by liquid chromatography. Structural polysaccharide content was back-calculated after adjustment for losses caused by degradation against sugar recovery standards. Each sample was analyzed in duplicate using 100 mg of biomass per sample. The percentage of glucose content in each sample was used to determine the efficiency of glucose release after the biomass digestibility assay (described below).

Hyp Assay

Samples of untreated and LHW, dilute acid, and alkaline peroxide-pretreated poplar weighing 15 mg were rehydrated in 2.1 mL of 30 mM sodium citrate (pH 4.5) for 3 h at room temperature in 2-mL screwtop tubes (Sarstedt, Germany). Buffer (as a control) or 4.5 µL of Fermgen protease (Genencor, Palo Alto, CA, USA) at 70 mg/mL (determined with the Pierce bicinchoninic acid assay, ThermoFisher Scientific, USA) was added to the rehydrated biomass, and the samples were incubated in a 50 °C rotisserie oven with end-over-end rotation for 24 h. Multiple samples for each treatment and biomass type were pooled to achieve a final mass of ~100 mg. The pooled samples were boiled overnight in distilled water, washed three times with distilled water, and dried overnight at 50 °C to remove any soluble proteins (particularly monomeric, non-covalently-bound extensins, or other soluble HRGPs) (Lamport 1969). Samples were re-weighed, and amino acid hydrolysis was performed by adding 1.6 mL of 6 M HCl and incubating in closed Sarstedt tubes at 100 °C for 18 h. Then, 500 µL of the supernatant were transferred to a new tube and pH-adjusted by the addition of 12 M NaOH to pH 3.0 (+/− 0.1). A higher pH produced a dark brown color in the samples that interfered with subsequent colorimetric analysis. Samples were centrifuged at 21,130 g for 1 min to clarify the supernatant. A colorimetric assay for Hyp content was then performed on the clarified supernatant (Lamport 2013). The absorbance at 560 nm of each sample was measured in triplicate (200 µL each) in a 96-well plate using a BioTek Synergy HT plate reader (BioTek, USA). Data were analyzed using Microsoft Excel 2010 (Microsoft, USA). Hyp standards with the same salt concentrations (4.2 M NaCl) and pH (3) as the samples were used to construct a standard curve.

Protease Treatment

Dried poplar biomass samples (untreated and LHW, dilute acid, and alkaline peroxide-pretreated) weighing 5 mg were rehydrated in 700 µL of 30 mM sodium citrate (pH 4.5) for 3 h at room temperature in 2-mL screwtop tubes (Sarstedt). Buffer (as a control) or 1.5 µL of Fermgen was added to the rehydrated biomass, and samples were incubated in a 50 °C rotisserie oven with end-over-end rotation for 24 h. Treatment was halted by boiling the tubes in a water bath for 10 min. To demonstrate that these conditions would result in extensin degradation, the same amount of Fermgen was added to 18.5 µL of crude extensin protein in 30 mM sodium citrate buffer (extracted from carrot root disks with $CaCl_2$ and concentrated with an Amicon PM-10 membrane (Stuart and Varner 1980)), and incubated under identical conditions. Two µL of protease-treated and untreated samples were each spotted onto nitrocellulose membranes that were then probed with antibodies gE-1 and JIM20 (as described for tissue prints); no extensin signal was detected after protease treatment (data not shown). The biomass was then centrifuged at 21,300 g for 30 s and washed to remove any residual protease activity, using two 500-µL washes with 1 M NaCl followed by four 500-µL washes with deionized water, and finally adding 500 µL of 30 mM sodium citrate buffer, pH 4.5 (final concentration 21.4 mM) (Kumar and Wyman 2009). The biomass was vortexed and then centrifuged for each wash. Only 500 µL of the total 700 µL was removed after each centrifugation to avoid disturbing the biomass pellet. Control experiments using bovine serum albumin (BSA) as a substrate verified that no Fermgen activity remained in the samples after boiling and washing.

For a time-course experiment, 5 mg of dried poplar biomass samples of all four types was rehydrated as described above. Every two hours for ten hours, 1.5 µL of Fermgen were added to the rehydrated biomass (9 µL total), while the samples incubated in a 50 °C

rotisserie oven with end-over-end rotation for 24 h. The treatment was stopped as described above before proceeding with a digestibility assay.

Biomass Digestibility Assay

The high-throughput protocols of Santoro *et al*. (2010) and Selig *et al*. (2010) for the enzymatic digestion of biomass into simple sugars were adapted to laboratory scale. To 5 mg of protease-treated poplar biomass in 700 µL of 21.4 mM sodium citrate (pH 4.5) (prepared as described above), 50 µL of a solution containing Cellic CTEC2 (70 mg/g biomass, Novozymes, Denmark), Cellic HTEC2 (2.5 mg/g biomass, Novozymes), 0.007% NaN_3 (final concentration), and 30 mM sodium citrate (pH 4.5) were added. The excess enzyme loading ensured enzyme activity would not be a limiting factor. Samples were incubated in a 50 °C rotisserie oven with end-over-end rotation for one week.

The tubes were centrifuged at 21,130 g for 1 minute. The supernatant was removed to a new tube, diluted 1:10 in water, and tested for glucose content using a glucose oxidase/peroxidase (GOPOD) assay (Megazyme International Ireland, Ireland). Triplicate 20-µL aliquots of sample, glucose standard, a Fermgen-only mixture, or a CTEC2/HTEC2 mixture were mixed with 200 µL GOPOD reagent and incubated at 40 °C for 45 minutes in a 96-well plate. The absorbance at 510 nm was read in a BioTek Synergy HT plate reader. Data were analyzed using Microsoft Excel 2010. Glucose standards were used to construct a standard curve. The absorbance from the enzyme-only mixtures was subtracted from the absorbance of the samples to correct for glucose present in the enzymes.

RESULTS

Three Classical Extensin Genes in Poplar

Previous work by Guo *et al.* (2014) identified 37 genes in the *P. trichocarpa* genome (v. 3.0) with signature SPPP or SPPPP extensin motifs. Further diagnostic criteria were used to identify those genes encoding classical extensins. First, the encoded protein must include an N-terminal signal peptide for it to serve as a structural cell wall protein. According to SignalP analysis, 22 of the 37 genes encode proteins with N-terminal signal peptides. Thirteen genes exhibited both signal peptides and significant homology to annotated extensins, found by querying the SwissProt database with HMMER3, where the sequence returned with greatest homology to the queried gene was an extensin (E-value < $3x10^{-9}$). Finally, classical extensins consist primarily of a single domain made up of repetitions containing $SP_{3–6}$ and YXY, unlike other subgroups of HRGPs or extensin chimera proteins. Three poplar genes met all of our criteria: Potri.001G019700, Potri.001G020100, and Potri.018G050100 (Fig. 1, Table 1). Proposed names for these three genes are: *PtEXTENSIN1* (Potri.001G019700), *PtEXTENSIN2* (Potri.001G020100), and *PtEXTENSIN3* (Potri.018G050100). *PtEXT3* had highest similarity to carrot (*Daucus carota*) extensin (Chen and Varner 1985a, 1985b) based on the results of the HMMER3 search, while the other two genes had highest similarity to *Arabidopsis Extensin-2* (Yoshiba *et al.* 2001). Gene structure was also examined, as the 20 classical extensins in *Arabidopsis* have either no introns or a single intron (Showalter *et al*. 2010). *PtEXT1* and *PtEXT3* lack introns, while *PtEXT2* has a single intron (Table 1).

> Potri.001G019700 PtEXT1
MIYALAFCVVATSVVAKEPYYYKSPPPPLKSSPPPSPSPPPPYHYSSPPPPKKSPPPPYIYKSPPPPLKSPPPPYHY
SSPPPPKKSPPPPYHYSSPPPPKKSPPPPYVYKSPPPPSPSPPPPYHYSSPPPPKKSPPPPYVYKSPPPPSPSPPPP
YHYSSPPPPKKSPPPPYIYKSPPPPSPSPPPPYHYSSPPPPKK**SPHPP**YVYKSPPPPHY

> Potri.001G020100 PtEXT2
MENRGRMGHLSPMIHAIAICLVATSVVAYEPYYYKSPPPPSQSPPPPYHYSSPPPPKKSPPPPYHYTSPPPPKKSPP
PPYHYS**SPPQP**KKSPPPPYHYSSPPPPKKSPPPPYHYSSPPPPKK**SPLPP**YHYSSPPPPKKSPPPPYHYSSPPPPKK
SPPPQYHYTSPPPPKKSPPPPYHYTSPPPPKKSPPPPYHYTSPPPPKKSPPPPYHYTSPPPPKKSPPPPYHYSSPPP
PKKSPPPPYHYTSPPPPKKIEIVDPW

> Potri.018G050100 PtEXT3
MDLLHSVMLYFSLALLLSSSSEATDTTFSRNSALWLTYKSPPPPFHNKHKSPPPPHKYKSPPPPHHKCKYSPPPPV
YTYRSPPPPTPMHHKSPPPSPHMFKSPPPPPYRYISPPPPPPHPPCHAYKYLSPPPPFSYKYASPPPPKHHHHHKHK
WSPYPFITYMSPPPPHHNYPDYHYSSPPPPPPIVAY

Fig. 1. Complete amino acid sequences of the three poplar classical extensin genes, with signal peptide sequences highlighted in yellow, $SP_{3–6}$ repeats in green, and H or K in blue. YXY motifs are underlined.

The three poplar extensin proteins were dominated by repeated motifs. The mature PtEXT1 protein consists of 12 repeats of (S/K/L)(P/K)SPPPPY(Y/H/I/V)Y(K/S)SPPPP, with only a few non-repeat amino acids at the N- and C-termini (Fig. 1). PtEXT2 consists of 14 repeats of YHY(K/S/T)SPPPPKKSPPPP, again with only a few non-repeat amino acids at the N- and C-termini. These repeats are similar to the tomato P3 extensin repeat sequence SPPPPSPSPPPPYYYK (Smith *et al*. 1986), a motif found in extensins across the plant kingdom (Kieliszewski *et al*. 2011). PtEXT3 has 12 $SP_{3–6}$ motifs that make up 40% of the 167-amino-acid protein, including five repeats of the sequence $YXYXSP_{4–6}$.

Table 1. Classical Extensin Genes in *P. trichocarpa*

Locus ID	Homologous Protein (Organism)	E-value (HMMER)	Repeated Motif	#R	#AA	$SP_3/SP_4/SP_5/SP_6$ Repeats	#YXY/H/K	Max Hyp (%)	Number of Introns
Potri.001G019700 (PtEXT1)	Extensin-2 (*A. thaliana*)	$1.10x10^{-70}$	(S/K/L)(P/K)SPPPPY(Y/H/I/V)Y(K/S)SPPPP	12	197	0/22/0/0	12/8/21	46	0
Potri.001G020100 (PtEXT2)	Extensin-2 (*A. thaliana*)	$6.70x10^{-75}$	YHY(K/S/T)SPPPPKKSPPPP	14	229	1/24/0/0	14/13/27	44	1
Potri.018G050100 (PtEXT3)	Extensin (*D. carota*)	$2.70x10^{-9}$	$YXYXSP_{4–6}$	5	167	1/6/3/2	5/19/14	32	0

#R=number of times repeated motif occurs in the protein sequence; #AA=number of amino acids in mature protein after signal peptide cleavage; Max Hyp (%) = theoretical maximum mol % Hyp.

Expression of the three poplar extensin genes was examined using poplar microarray data available at popgenie.org. The organ or tissue with the highest identified expression of each gene depends somewhat on the specific microarray experiment. Based on the *P. balsamifera* developmental tissue dataset, all three genes are most highly expressed in young leaves (Wilkins *et al*. 2009). Based on the *P. trichocarpa* tissues dataset, *PtEXT1* is most highly expressed in mature leaves and roots, *PtEXT2* is most highly expressed in roots and young leaves, and *PtEXT3* is most highly expressed in young and mature leaves (Yang *et al*. 2008). Based on the recent work of Hefer *et al*. (2015), at least one poplar extensin gene (*PtEXT1*) is expressed in young woody tissue.

Localization of Extensin Proteins in Poplar Stems

To investigate whether extensin proteins are present in poplar wood, tissue prints were made with poplar stem sections to detect extensin proteins using three different anti-extensin antibodies (Fig. 2).

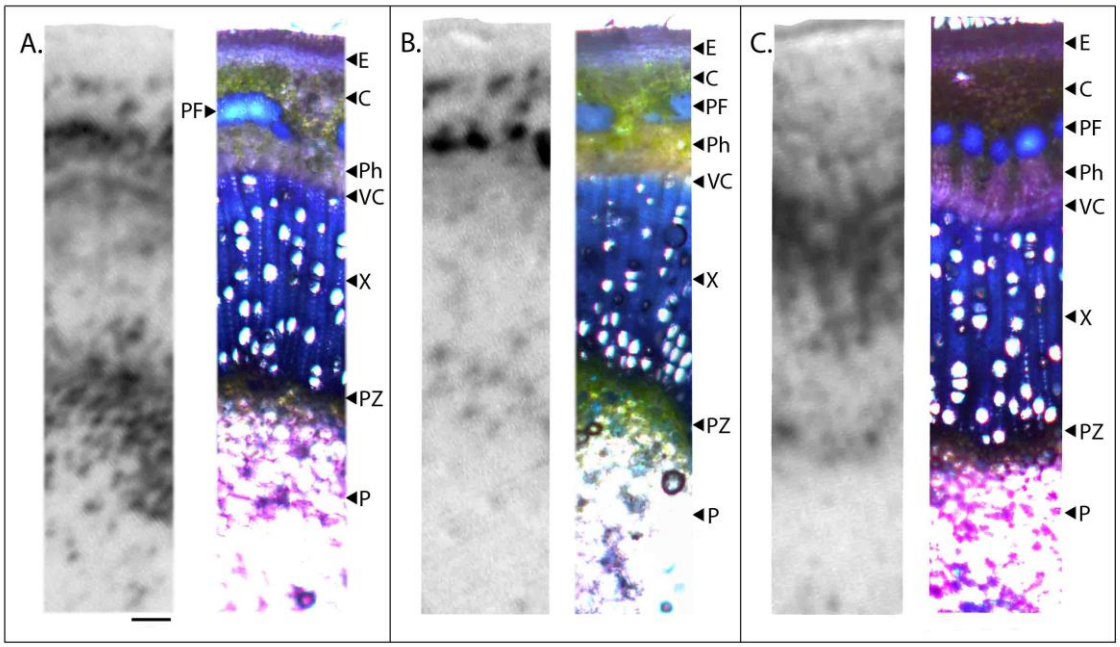

Fig. 2. Tissue print immunoblots showing extensin localization in poplar stems. A) Print probed with JIM20 (left) and its corresponding stained section (right); B) Print probed with LM1 (left) and its corresponding section (right); C) Print probed with gE-1 (left) and its corresponding section (right). P=pith; PZ=perimedullary zone; X=xylem; VC=vascular cambium; Ph=phloem; PF=phloem fiber; C=cortex; E=epidermis. Bar=100 µm.

JIM20 is an anti-extensin monoclonal antibody (mAb) raised to pea guard cell protoplasts, shown to be specific for extensins (Smallwood *et al*. 1994). LM1 is a mAb raised to rice cell wall material that has been shown to recognize monocot threonine-rich HRGPs and dicot extensins (Smallwood *et al*. 1995). The anti-carrot extensin antibody (gE-1) is a polyclonal antibody raised to purified glycosylated carrot extensin-1 (Stafstrom and Staehelin 1988). Using poplar stem tissue prints, extensins were detected adjacent to the vascular cambium with all three antibodies. Extensin localization was detected predominantly in the phloem with both mAbs, JIM20 and LM1 (Figs. 2(a) and (b)). However, extensin localization was detected most prominently in xylem with gE-1, with a very low signal (if any) in phloem (Fig. 2(c)). All three antibodies also produced a detectable signal in the perimedullary zone (the outermost layer of the pith).

Effect of Pretreatments and Protease Treatment on Hyp Content

Having verified that extensin proteins are present in the woody tissues of poplar, the next question is whether standard pretreatments remove extensins from poplar stem biomass. Extensin is the most abundant source of Hyp in plant cell walls; in some cases a single extensin accounts for 90% of the Hyp, making Hyp content an accepted proxy for extensin content (Lamport 1969; Pope 1977; Tierney and Varner 1987). In untreated poplar biomass, 193 µg of Hyp per gram dry weight was measured (Fig. 3, black bars). In poplar biomass pretreated with LHW, 156 µg Hyp g^{-1} dry weight was measured (a 20% decrease,

$p < 0.01$, Student's T-test). Lower Hyp content was found in poplar biomass pretreated with alkaline peroxide (126 μg Hyp g^{-1} dry weight, a 34% decrease, $p < 0.01$) and in poplar biomass pretreated with dilute sulfuric acid (112 μg Hyp g^{-1} dry weight, a 42% decrease, $p < 0.01$). Thus, all pretreatments removed some Hyp from pretreated poplar biomass, but in all cases, over 50% of the original Hyp content remained.

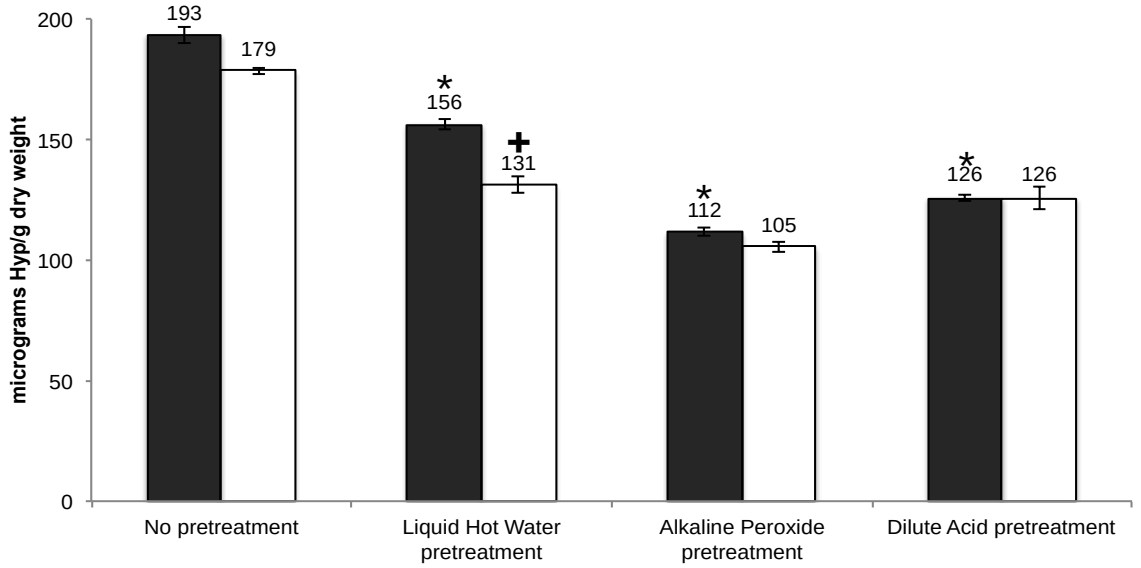

Fig. 3. Hyp remaining in poplar biomass after different pretreatments and buffer-only (black bars) or Fermgen (white bars) treatment. Bars show standard error, n=3. *, $p < 0.05$ compared with no pretreatment; **+**, $p < 0.05$ compared with buffer-only treatment of the same biomass.

To test whether protease treatment could remove extensins from poplar biomass, the Hyp content of the four types of poplar biomass (untreated or pretreated with LHW, alkaline peroxide, or dilute acid) was measured after incubation with Fermgen, an acid protease that has been used to digest proteins in starchy biomass to promote fermentation. Protease treatment decreased Hyp content in untreated biomass from 193 to 179 μg Hyp g^{-1} dry weight, but this difference was not significant (a 7% decrease, $p=0.10$). Protease treatment significantly decreased Hyp content in LHW-pretreated biomass, from 156 to 131 μg Hyp g^{-1} dry weight (a 16% decrease, $p = 0.04$ (Fig. 3)). Protease treatment did not change the Hyp content of either alkaline peroxide or dilute acid-pretreated poplar biomass.

Protease Treatment Can Increase Glucose Release from Poplar Biomass

Protease treatment of the four types of poplar biomass (untreated or pretreated with LHW, alkaline peroxide, or dilute acid) was also tested for its effect on glucose yield from the digestion of biomass with cell-wall-degrading enzymes. Protease treatment with Fermgen caused a significant increase in glucose released from LHW-pretreated poplar after digestion with cell-wall-degrading enzymes, from 30% of the total glucan content before protease treatment to 36% after protease treatment (a 20% increase (Fig. 4)). Protease treatment with Fermgen caused no significant change in the amount of glucose released from the three other types of biomass (untreated, alkaline peroxide-pretreated, or acid-pretreated). Treatment of LHW-pretreated poplar biomass with chymotrypsin did not affect the release of glucose (data not shown).

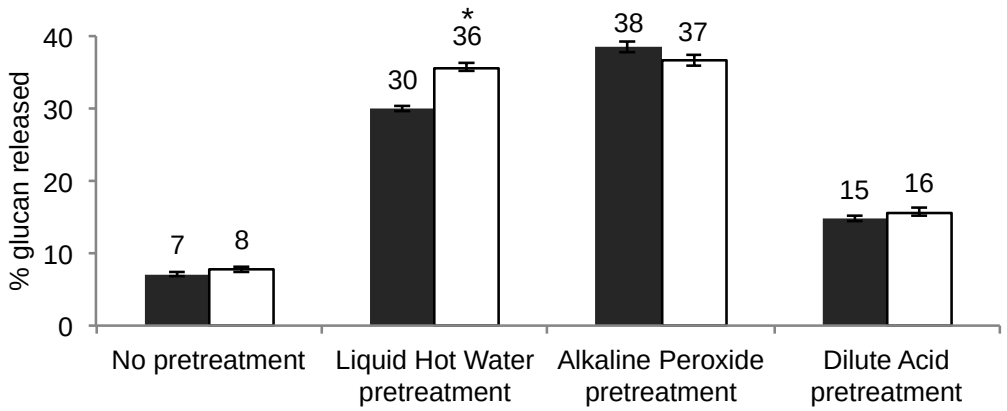

Fig. 4. Effect of Fermgen treatment on glucose release from pretreated poplar biomass. Glucose released after week-long cell wall digestion with Cellic CTEC2/HTEC2 without (black) or with (white) prior treatment with Fermgen protease. Bars show standard error, n=6. *, $p < 0.005$ in comparison with no Fermgen treatment (Student's T-test).

To test whether the concentration of Fermgen was limiting in protease treatments of poplar biomass, glucose release was compared after protease treatment with single and multiple (six) doses of Fermgen administered over a 24-h period, followed by cell wall digestion. Poplar biomass treated with multiple doses of Fermgen showed no significant increase in glucose release compared with poplar biomass treated with a single dose of Fermgen (p-values > 0.05 (Fig. 5)). An ANOVA of all Hyp and glucose release data indicates a significant negative correlation between Hyp content and glucose release (p-value = 0.04).

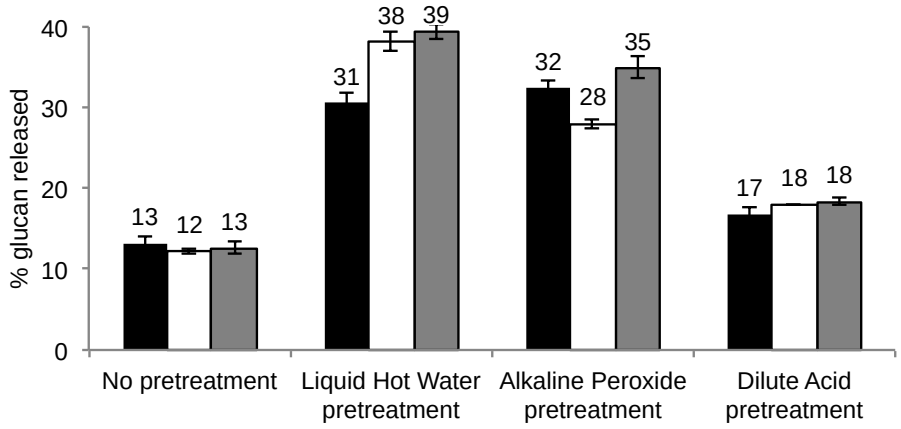

Fig. 5. Effect of multiple doses of Fermgen on glucose release. Glucose released after week-long cell wall digestion with Cellic CTEC2/HTEC2 after prior treatment with different doses of Fermgen protease. Black=no Fermgen; white=single dose of Fermgen; grey=six doses of Fermgen over 24 hours. Bars show standard error, n=3.

DISCUSSION

Of the 37 previously identified poplar extensin genes (Guo *et al*. 2014), which include extensin chimeras and hybrids such as leucine-rich repeat extensin proteins (Rubinstein *et al.* 1995), proline-rich extensin-like receptor kinases (Nakhamchik *et al*.

2004), and extensin-arabinogalactan proteins (Lind *et al.* 1994), three poplar genes show characteristics that are shared by the classical extensins, in that the encoded proteins 1) include an N-terminal signal peptide, 2) consist largely of repeats containing SP_n (12 in *PtEXT1*, 14 in *PtEXT2*, and 12 in *PtEXT3*), 3) do not belong to other subgroups of HRGPs or contain any other domains indicative of extensin chimeric proteins, and 4) have no introns (*PtEXT1* and *PtEXT3*) or a single intron (*PtEXT2*) (Fig. 1, Table 1). The three poplar extensin proteins also exhibit several features thought to be important in forming a covalently linked extensin matrix in the cell wall: strong periodicity, to allow self-assembly; repetition of the YXY cross-linking motif, to allow isodityrosine formation; and regular placement of positively charged residues (H and K), to allow electrostatic interactions with negatively charged cell wall components, such as pectins (Cannon *et al*. 2008; Valentin *et al*. 2010). Since the three poplar classical extensins have an abundance of YXY motifs and lack SPSP and tri-C motifs, they can be classified as Group IIb extensins (Kieliszewski *et al*. 2011; Saha *et al*. 2013).

The presence of only three classical extensin genes in poplar is somewhat surprising, since the smaller *Arabidopsis* genome encodes 20 classical extensins (Showalter *et al*. 2010). However, the relatively large number of extensins in *Arabidopsis* may have resulted from multiple genome duplication events that occurred after the lineage diverged from poplar and other dicots (Guo *et al*. 2014).

Post-translational modifications, specifically proline hydroxylation and glycosylation, are thought to be critical to extensin function (Velasquez *et al*. 2011). Generally, however, it is thought that the "rule" for proline hydroxylation is that all prolines within a block of two or more prolines will be hydroxylated, although whether a particular proline is hydroxylated depends on tissue identity, growth conditions, and context within neighboring amino acid sequences (Shpak *et al.* 2001; Shimizu *et al*. 2005; Estévez *et al.* 2006). Following this rule, the potential molar percentage of Hyp in the three poplar extensins, assuming maximal hydroxylation, is 51% (PtEXT1), 49% (PtEXT2), and 35% (PtEXT3) (Table 1). These values for Hyp composition are comparable to experimentally determined values for other extensins, including those from runner bean (35.3%) (O'Neill and Selvendran 1980), tomato P1 (31.5%) and P2 (41.8%) (Smith *et al*. 1984), and Douglas fir PHRGP (28.1%) and P2 (29.2%) (Kieliszewski *et al*. 1992).

In general, the methods for detecting the presence of extensin proteins in woody tissue are indirect and rely on the detection of high concentrations of Hyp, for example, in the wood of European beech (*Fagus sylvatica*) (Dill *et al*. 1984) and spruce (*Picea abies*) (Westermark *et al*. 1986), in the callus tissue of *Pinus elliottii* (Whitmore 1982), and in the suspension-cultured cells of sycamore (*Acer pseudoplatanus*) (Lamport 1967) and Douglas fir (*Pseudotsuga menziesii* (*Mirbel*) *Franco*) (Kieliszewski *et al*. 1992). A Hyp/Pro-rich extensin-like protein of unknown sequence has been purified from xylem in the gymnosperm loblolly pine (*Pinus taeda* L.) and detected *in situ* in wood using antibodies raised to the protein (Bao *et al*. 1992).

In this paper, direct evidence for extensin localization in woody tissues in poplar, a woody angiosperm, was obtained using tissue print immunoblots probed with three different antibodies specific for extensins (Fig. 2). Much of the extensin signal was found near the vascular cambium, a result congruent with the model that soluble extensins are secreted in expanding cell walls and then cross-linked as cells mature. The recent findings of Hefer *et al*. (2015), who found high expression of the *PtEXT1* gene in developing xylem tissue, are consistent with these results. It should be noted that tissue prints can only detect extensins that are transferred to the nitrocellulose paper. The distribution of extensins may

be greater than what we detected, as cross-linked extensins cannot be transferred, while the transfer of soluble extensins may be limited by the interference of other cell wall components, such as lignin. Nevertheless, extensins clearly are present in woody tissue.

The detection of extensins in woody poplar tissue suggests that extensins could play a role in recalcitrance, since even a low level of cross-linking within the cell wall could impede the passage of processive cellulases along cellulose microfibrils, or stabilize the architectural arrangements of other cell wall components. Recent work in rice demonstrates a negative correlation between Hyp content and glucose release following either LHW or dilute base pretreatments (Tanger *et al*. 2015). We also found a significant negative correlation between Hyp content and glucose release. Although in theory biomass pretreatment could reduce extensin content, earlier work showing that extensins are not released from cell walls after either strong acid or base digestions (Heath and Northcote 1971; Lamport 1980) suggests that the relatively mild pretreatments currently favored by the industry would not remove extensins. Indeed, 81% of the initial Hyp content remained in poplar biomass after LHW pretreatment (Fig. 3). More robust pretreatments (alkaline peroxide and dilute acid) removed more Hyp, but even so, more than half remained after pretreatment.

It has been proposed that the protease removal of extensins from woody biomass is responsible for improved paper pulp or bast fiber production (Pokora and Johnson 1994; Sung *et al*. 2011). Using Hyp content as a proxy for extensins, Lamport (1965) found that proteases released up to 30% of Hyp from the cell walls of suspension-cultured sycamore cells. We measured a 20% reduction in Hyp content in LHW-pretreated woody poplar biomass after treatment with Fermgen, an acid fungal protease (Fig. 3). The resistance of alkaline peroxide or dilute acid-pretreated poplar biomass to protease treatment suggests that the extensins remaining in this material after pretreatment may be inaccessible to the protease. As six doses of Fermgen had no greater effect than a single dose of Fermgen, the protease itself was not limiting. This result is consistent with other researchers' efforts to extract extensins – most treatments are able to remove a portion of the extensin pool, but an intractable portion remains (for example, Biggs and Fry 1990; Lamport 1969; Qi *et al*. 1995).

The effectiveness of protease treatment of LHW-pretreated biomass in reducing recalcitrance, as shown by increased glucose yields, was directly tested, and a statistically significant increase of 20% over the baseline was observed (Fig. 4). Whether it is feasible or desirable to include a protease treatment step during the conversion of biomass to biofuel depends on the pretreatment protocol utilized. Protease treatment of biomass that was pretreated with dilute acid or alkaline peroxide did not release any additional Hyp, and did not increase glucose yields. In contrast, LHW-pretreated biomass exhibited both lower Hyp content and higher glucose yields following protease treatment. Independently of the protease treatment, alkaline peroxide pretreatment was the most effective, releasing 32-38% of the glucose (Fig. 4, 5). This result is in line with work in other systems demonstrating alkaline peroxide pretreatment is more effective than LHW or dilute acid (Cao *et al*. 2012; Saha and Cotta 2010). The differences in glucose release with pretreatment type is beyond the scope of this research, but provisionally may be attributed to the known effects of alkaline peroxide on lignin (Selig *et al*. 2009), which is well-established as a significant contributor to biomass recalcitrance.

We have demonstrated that extensin removal from LHW-pretreated biomass using protease treatment has the potential to increase access of cell wall-degrading enzymes to cellulose to improve biofuel production. Mild biomass pretreatment with LHW has a

number of advantages in biofuel production, since it is generally less expensive in terms of reagents, waste cleanup, and equipment, and fewer toxic and inhibitory byproducts are formed. Fermgen protease has already been successfully utilized in large-scale ethanol production, and even an incremental yield increase can be significant given the scale and fixed costs of industrial biofuel production. Further work will be necessary to determine the technical and economic feasibility of incorporating protease treatment at industrial scales.

CONCLUSIONS

1. Three poplar genes (*PtEXT1, PtEXT2,* and *PtEXT3*) encoded classical extensins, and extensin proteins were detected in woody poplar tissue.
2. Extensin proteins were present in woody tissue even after standard biomass pretreatments, as measured by Hyp content.
3. Glucose yield from LHW-pretreated poplar biomass was improved by a protease treatment that also reduces Hyp content. Such improvement was not seen for other pretreatments.

ACKNOWLEDGMENTS

This work was submitted by M.B.F. in partial fulfillment of the requirements of Colorado State University for the degree of Doctor of Philosophy in Botany.

The authors thank the Colorado Center for Biorefining and Biofuels (Grant #013-7), the Colorado State University Libraries Open Access Research and Scholarship Fund, CarboSource (Athens, GA) for JIM20, PlantProbes (Leeds, UK) for LM1, Dr. L. A. Staehelin (University of Colorado, Boulder) for gE-1, Dr. Suzanne Royer and Dr. Lorrie Anderson (CSU) for imaging assistance, Bronwyn Phillips and Olivia Todd for research assistance, Dr. Melvin P. Tucker (NREL) for providing the liquid hot water and dilute-acid pretreated poplar, and Dr. Emily Fox (University of Maryland) for a helpful and critical review of the manuscript.

The pretreatment and analytical work was partially supported by the U.S. Department of Energy under Contract No. DE-AC36-08GO28308 with the National Renewable Energy Laboratory. Funding was provided by the DOE Office of Energy Efficiency and Renewable Energy, Bioenergy Technologies Office.

REFERENCES CITED

Bao, W., O'Malley, D. M., and Sederoff, R. R. (1992). "Wood contains a cell-wall structural protein," *Proc. Natl. Acad. Sci. U. S. A.* 89(14), 6604-6608. DOI: 10.1073/pnas.89.14.6604

Béguin, P., and Aubert, J. P. (1994). "The biological degradation of cellulose," *FEMS Microbiol. Rev.* 13(1), 25-58. DOI: 10.1111/j.1574-6976.1994.tb00033.x

Biggs, K. J., and Fry, S. C. (1990). "Solubilization of covalently bound extensin from capsicum cell walls," *Plant Physiol.* 92, 197-204. DOI: 10.1104/pp.92.1.197

Cannon, M. C., Terneus, K., Hall, Q., Tan, L., Wang, Y., Wegenhart, B. L., Chen, L., Lamport, D. T. A., Chen, Y., and Kieliszewski, M. J. (2008). "Self-assembly of the plant cell wall requires an extensin scaffold," *Proc. Natl. Acad. Sci. U. S. A.* 105(6), 2226-2231. DOI: 10.1073/pnas.0711980105

Cassab, G. I., and Varner, J. E. (1989). "Tissue printing on nitrocellulose paper: A new method for immunolocalization of proteins, localization of enzyme activities and anatomical analysis," *Cell Biol. Int. Rep.* 13(1), 147-152. DOI: 10.1016/S0309-1651(89)80016-5

Cao, W., Sun, C., Liu, R., Yin, R., and Wu, X. (2012). "Comparison of the effects of five pretreatment methods on enhancing the enzymatic digestibility and ethanol production from sweet sorghum bagasse," *Biores. Technol.* 111, 215-221. DOI: 10.1016/j.biortech.2012.02.034

Carpita, N., and McCann, M. C. (2000). "The cell wall," in: *Biochemistry and Molecular Biology of Plants*, Buchanan, B., Gruissem, W., and Jones, R. (eds.), American Society of Plant Physiologists, Rockville, MD, pp.45-110.

Cegelski, L., O'Connor, R. D., Stueber, D., Singh, M., Poliks, B., and Schaefer, J. (2010). "Plant cell-wall cross-links by REDOR NMR spectroscopy," *J. Am. Chem. Soc.* 132(45), 16052-16057. DOI: 10.1021/ja104827k

Chaturvedi, V., and Verma, P. (2013). "An overview of key pretreatment processes employed for bioconversion of lignocellulosic biomass into biofuels and value added products," *3 Biotech* 3(5), 415-431. DOI: 10.1007/s13205-013-0167-8

Chen, J., and Varner, J. E. (1985a). "An extracellular matrix protein in plants: Characterization of a genomic clone for carrot extensin," *EMBO J.* 4(9), 2145-2151.

Chen, J., and Varner, J. E. (1985b). "Isolation and characterization of cDNA clones for carrot extensin and a proline-rich 33-kDa protein," *Proc. Natl. Acad. Sci. U. S. A.* 82(13), 4399-4403. DOI: 10.1073/pnas.82.13.4399

Cong, F., Diehl, B. G., Hill, J. L., Brown, N. R., and Tien, M. (2013). "Covalent bond formation between amino acids and lignin: Cross-coupling between proteins and lignin," *Phytochemistry* 96, 449-56. DOI: 10.1016/j.phytochem.2013.09.012

Cooper, J. B., and Varner, J. E. (1984). "Cross-linking of soluble extensin in isolated cell walls," *Plant Physiol.* 76(2), 414-417. DOI: 10.1104/pp.76.2.414

Cowling, E. B., and Merrill, W. (1966). "Nitrogen in wood and its role in wood deterioration," *Can. J. Bot.* 44(11), 1539-1554. DOI: 10.1139/b66-167

Dill, I., Salnikow, J., and Kraepelin, G. (1984). "Hydroxyproline-rich protein material in wood and lignin of *Fagus sylvatica*," *Appl. Environ. Microbiol.* 48(6), 1259-1261.

Estévez, J. M., Kieliszewski, M. J., Khitrov, N., and Somerville, C. (2006). "Characterization of synthetic hydroxyproline-rich proteoglycans with arabinogalactan protein and extensin motifs in *Arabidopsis*," *Plant Physiol.* 142(2), 458-470. DOI: 10.1104/pp.106.084244

Finn, R. D., Clements, J., and Eddy, S. R. (2011). "HMMER web server: Interactive sequence similarity searching," *Nucleic Acids Res.* 39(S2), W29-W37. DOI: 10.1093/nar/gkr367

Fry, S. C. (1982). "Isodityrosine, a new cross-linking amino acid from plant cell-wall glycoprotein," *Biochem. J.* 204(2), 449-455. DOI: 10.1042/bj2040449

Guo, L., Chen, Y., Ye, N., Dai, X., Yang, W., and Yin, T. (2014). "Differential retention and expansion of the ancestral genes associated with the paleopolyploidies in modern rosid plants, as revealed by analysis of the extensins super-gene family," *BMC Genomics* 15, 612. DOI: 10.1186/1471-2164-15-612

Heath, M., and Northcote, D. H. (1971). "Glycoprotein of the wall of sycamore tissue-culture cells," *Biochem. J.* 125(4), 953-961. DOI: 10.1042/bj1250953

Held, M. A., Tan, L., Kamyab, A., Hare, M., Shpak, E., and Kieliszewski, M. J. (2004). "Di-isodityrosine is the intermolecular cross-link of isodityrosine-rich extensin analogs cross-linked *in vitro*," *J. Biol. Chem.* 279(53), 55474-55482. DOI: 10.1074/jbc.M408396200

Hefer, C., Mizrachi, E., Myburg, A., Douglas, C. J., and Mansfield, S. D. (2015). "Comparative interrogation of the developing xylem transcriptomes of two wood-forming species: *Populus trichocarpa* and *Eucalyptus grandis*," *New Phytologist* 206(4), 1391-1405. DOI: 10.1111/nph.13277

Himmel, M. E. (2008). *Biomass Recalcitrance*, Blackwell Publishing Ltd., Oxford, UK. DOI: 10.1002/9781444305418

Kieliszewski, M., De Zacks, R., Leykam, J. F., and Lamport, D. T. A. (1992). "A repetitive proline-rich protein from the gymnosperm Douglas fir is a hydroxyproline-rich glycoprotein.," *Plant Physiol.* 98(3), 919-926. DOI: 10.1104/pp.98.3.919

Kieliszewski, M. J., Lamport, D. T. A., Tan, L., and Cannon, M. C. (2011). "Hydroxyproline-rich glycoproteins: Form and function," in: *Annual Plant Reviews. Plant Polysaccharides: Biosynthesis and Bioengineering*, P. Ulvskov (ed.), Wiley Blackwell, Oxford, UK, pp. 321-342. DOI: 10.1002/9781444391015.ch13

Kumar, R., and Wyman, C. (2009). "Does change in accessibility with conversion depend on both the substrate and pretreatment technology?" *Biores. Technol.* 100(18), 4193-4202. DOI: 10.1016/j.biortech.2008.11.058

Kumar, P., Barrett, D. M., Delwiche, M. J., and Stroeve, P. (2009). "Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production," *Ind. Eng. Chem. Res.* 48(8), 3713-3729. DOI: 10.1021/ie801542g

Lamport, D. T. A. (1965). "The protein component of primary cell walls," *Adv. Bot. Res.* 2, 151-218. DOI: 10.1016/S0065-2296(08)60251-7

Lamport, D. T. A. (1967). "Hydroxyproline-O-glycosidic linkage of the plant cell wall glycoprotein extensin," *Nature* 216(5122), 1322-1324. DOI: 10.1038/2161322a0

Lamport, D. T. A. (1969). "The isolation and partial characterization of hydroxyproline-rich glycopeptides obtained by enzymic degradation of primary cell walls," *Biochemistry* 8(3), 1155-1163. DOI: 10.1021/bi00831a049

Lamport, D. T. A. (1980). "Structure and function of plant glycoproteins," in: *The Biochemistry of Plants*, Vol. 3, Preiss, J. (ed.), Academic Press, New York, NY, pp. 501-541. DOI: 10.1016/b978-0-12-675403-2.50019-3

Lamport, D. T. A. (2013). "Hydroxyproline assay using NaBr/NaOCl," *Bio-Protocol* 3(19), e917.

Lamport, D. T. A., Katona, L., and Roerig, S. (1973). "Galactosylserine in extensin," *Biochem. J.* 133(1), 125-132. DOI: 10.1042/bj1330125

Lamport, D. T. A., and Miller, D. H. (1971). "Hydroxyproline arabinosides in the plant kingdom," *Plant Physiol.* 48(4), 454-456. DOI: 10.1104/pp.48.4.454

Lange, B. M., Lapierre, C., and Sandermann, H. (1995). "Elicitor-induced spruce stress lignin: Structural similarity to early developmental lignins," *Plant Physiol.* 108(3), 1277-1287. DOI: 10.1104/pp.108.3.1277

Lind, J. L., Bacic, A., Clarke, A. E., and Anderson, M. A. (1994). "A style-specific hydroxyproline-rich glycoprotein with properties of both extensins and arabinogalactan proteins," *Plant J.* 6(4), 491-502. DOI: 10.1046/j.1365-313X.1994.6040491.x

Masuda, H., Komiyama, S., and Sugawara, S. (1989). "Cell wall proteins from sugar beet cells in suspension culture," *Plant Physiol.* 89(2), 712-716. DOI: 10.1104/pp.89.2.712

Moore, T. S. (1973). "An extracellular macromolecular complex from the surface of soybean suspension cultures," *Plant Physiol.* 51(3), 529-536. DOI: 10.1104/pp.51.3.529

Mort, A. J., and Lamport, D. T. A. (1977). "Anhydrous hydrogen fluoride deglycosylates glycoproteins," *Anal. Biochem.* 82(2), 289-309. DOI: 10.1016/0003-2697(77)90165-8

Nakhamchik, A., Zhao, Z., Provart, N. J., Shiu, S.-H., Keatley, S. K., Cameron, R. K., and Goring, D. R. (2004). "A comprehensive expression analysis of the *Arabidopsis* proline-rich extensin-like receptor kinase gene family using bioinformatic and experimental approaches," *Plant Cell Physiol.* 45(12), 1875-1881. DOI: 10.1093/pcp/pch206

Nuñez, A., Fishman, M. L., Fortis, L. L., Cooke, P. H., and Hotchkiss, A. T. (2009). "Identification of extensin protein associated with sugar beet pectin," *J. Agric. Food Chem.* 57(22), 10951-10958. DOI: 10.1021/jf902162t

O'Neill, M. A., and Selvendran, R. R. (1980). "Glycoproteins from the cell wall of *Phaseolus coccineus*," *Biochem. J.* 187(1), 53-63. DOI: 10.1042/bj1870053

Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). "SignalP 4.0: Discriminating signal peptides from transmembrane regions," *Nat. Methods* 8(10), 785-786. DOI: 10.1038/nmeth.1701

Pokora, A. R., and Johnson, M. A. (1994). "Protease catalyzed treatments of lignocellulose materials," US Patent 5374555.

Pope, D. G. (1977). "Relationships between hydroxyproline-containing proteins secreted into the cell wall and medium by suspension-cultured *Acer pseudoplatanus* cells," *Plant Physiol.* 59(5), 894-900. DOI: 10.1104/pp.59.5.894

Qi, X., Behrens, B. X., West, P. R., and Mort, A. J. (1995). "Solubilization and partial characterization of extensin fragments from cell walls of cotton suspension cultures. Evidence for a covalent cross-link between extensin and pectin," *Plant Physiol.* 108(4), 1691-1701. DOI: 10.1104/pp.108.4.1691

Rubinstein, A. L., Broadwater, A. H., Lowrey, K. B., and Bedinger, P. A. (1995). "Pex1, a pollen-specific gene with an extensin-like domain," *Proc. Natl. Acad. Sci. U. S. A.* 92(8), 3086-3090. DOI: 10.1073/pnas.92.8.3086

Saha, B. C., and Cotta, M. A. (2010). "Comparison of pretreatment strategies for enzymatic saccharification and fermentation of barley straw to ethanol," *New Biotechnol.* 27(1), 10-16. DOI: 10.1016/j.nbt.2009.10.005

Saha, P., Ray, T., Tang, Y., Dutta, I., Evangelous, N. R., Kieliszewski, M. J., Chen, Y., and Cannon, M. C. (2013). "Self-rescue of an EXTENSIN mutant reveals alternative gene expression programs and candidate proteins for new cell wall assembly in *Arabidopsis*," *Plant J.* 75(1), 104-116. DOI: 10.1111/tpj.12204

Santoro, N., Cantu, S. L., Tornqvist, C.-E., Falbel, T. G., Bolivar, J. L., Patterson, S. E., Pauly, M., and Walton, J. D. (2010). "A high-throughput platform for screening milligram quantities of plant biomass for lignocellulose digestibility," *BioEnergy Res.* 3(1), 93-102. DOI: 10.1007/s12155-009-9074-6

Selig, M., Vinzant, T., Himmel, M., and Decker, S. (2009). "The effect of lignin removal by alkaline peroxide pretreatment on the susceptibility of corn stover to purified cellulolytic and xylanolytic enzymes," *Appl. Biochem. Biotechnol.* 155(1-3), 397-406. DOI: 10.1007/s12010-008-8511-x

Selig, M. J., Tucker, M. P., Sykes, R. W., Reichel, K. L., Brunecky, R., Himmel, M. E.,

Davis, M. F., and Decker, S. R. (2010). "Lignocellulose recalcitrance screening by integrated high-throughput hydrothermal pretreatment and enzymatic saccharification," *Ind. Biotechnol.* 6(2), 104-112. DOI: 10.1089/ind.2010.0009

Shimizu, M., Igasaki, T., Yamada, M., Yuasa, K., Hasegawa, J., Kato, T., Tsukagoshi, H., Nakamura, K., Fukuda, H., and Matsuoka, K. (2005). "Experimental determination of proline hydroxylation and hydroxyproline arabinogalactosylation motifs in secretory proteins," *Plant J.* 42(6), 877-889. DOI: 10.1111/j.1365-313X.2005.02419.x

Showalter, A. M. (1993). "Structure and function of plant cell wall proteins," *Plant Cell* 5(1), 9-23. DOI: 10.1105/tpc.5.1.9

Showalter, A. M., Keppler, B., Lichtenberg, J., Gu, D., and Welch, L. R. (2010). "A bioinformatics approach to the identification, classification, and analysis of hydroxyproline-rich glycoproteins," *Plant Physiol.* 153(2), 485-513. DOI: 10.1104/pp.110.156554

Shpak, E., Barbar, E., Leykam, J. F., and Kieliszewski, M. J. (2001). "Contiguous hydroxyproline residues direct hydroxyproline arabinosylation in *Nicotiana tabacum*," *J. Biol. Chem.* 276(14), 11272-11278. DOI: 10.1074/jbc.M011323200

Sjödin, A., Street, N. R., Sandberg, G., Gustafsson, P., and Jansson, S. (2009). "The *Populus* Genome Integrative Explorer (PopGenIE): A new resource for exploring the *Populus* genome," *New Phytologist* 182(4), 1013-1025. DOI: 10.1111/j.1469-8137.2009.02807.x

Smallwood, M., Beven, A., Donovan, N., Neill, S. J., Peart, J., Roberts, K., and Knox, J. P. (1994). "Localization of cell wall proteins in relation to the developmental anatomy of the carrot root apex," *Plant J.* 5(2), 237-246. DOI: 10.1046/j.1365-313X.1994.05020237.x

Smallwood, M., Martin, H., and Knox, J. P. (1995). "An epitope of rice threonine- and hydroxyproline-rich glycoprotein is common to cell wall and hydrophobic plasma-membrane glycoproteins," *Planta* 196(3), 510-522. DOI: 10.1007/BF00203651

Smith, J. J., Muldoon, E. P., and Lamport, D. T. A. (1984). "Isolation of extensin precursors by direct elution of intact tomato cell suspension cultures," *Phytochemistry* 23(6), 1233-1239. DOI: 10.1016/S0031-9422(00)80432-X

Smith, J. J., Muldoon, E. P., Willard, J. J., and Lamport, D. T. A. (1986). "Tomato extensin precursors P1 and P2 are highly periodic structures," *Phytochemistry* 25(5), 1021-1030. DOI: 10.1016/S0031-9422(00)81547-2

Stafstrom, J. P., and Staehelin, L. A. (1986). "The role of carbohydrate in maintaining extensin in an extended conformation," *Plant Physiol.* 81(1), 242-246. DOI: 10.1104/pp.81.1.242

Stafstrom, J. P., and Staehelin, L. A. (1988). "Antibody localization of extensin in cell walls of carrot storage roots," *Planta* 174(3), 321-332. DOI: 10.1007/BF00959517

Sung, W. L., Wood, M., and Huang, F. (2011). "Enzymatic preparation of plant fibers," US Patent #0312066.

Tanger, P., Vega-Sánchez, M. E., Fleming, M., Tran, K., Singh, S., Abrahamson, J. B., Jahn, C. E., Santoro, N., Naredo, E. B., Baraoidan, M., Danku, J. M. C., Salt, D. E., McNally, K. L., Simmons, B. A., Ronald, P. C., Leung, H., Bush, D. R., McKay, J. K., and Leach, J. E. (2015). "Cell wall composition and bioenergy potential of rice straw tissues are influenced by environment, tissue type, and genotype," *BioEnergy Res.* 8(3), 1165-1182. DOI: 10.1007/s12155-014-9573-y

Tierney, M. L., and Varner, J. E. (1987). "The extensins," *Plant Physiol.* 84(1), 1-2. DOI:

10.1104/pp.84.1.1
Valentin, R., Cerclier, C., Geneix, N., Aguié-Béghin, V., Gaillard, C., Ralet, M.-C., and Cathala, B. (2010). "Elaboration of extensin-pectin thin film model of primary plant cell wall," *Langmuir* 26(12), 9891-9898. DOI: 10.1021/la100265d
van Holst, G.-J., and Varner, J. E. (1984). "Reinforced polyproline II conformation in a hydroxyproline-rich cell wall glycoprotein from carrot root," *Plant Physiol.* 74(2), 247-251. DOI: 10.1104/pp.74.2.247
Velasquez, S. M., Ricardi, M. M., Dorosz, J. G., Fernandez, P. V, Nadra, A. D., Pol-Fachin, L., Egelund, J., Gille, S., Harholt, J., Ciancia, M., *et al.* (2011). "O-glycosylated cell wall proteins are essential in root hair growth," *Science* 332(6036), 1401-1403. DOI: 10.1126/science.1206657
Westermark, U., Hardell, H.-L., and Iversen, T. (1986). "The content of protein and pectin in the lignified middle lamella/primary wall from spruce fibers," *Holzforschung* 40(2), 65-68. DOI: 10.1515/hfsg.1986.40.2.65
Whitmore, F. W. (1982). "Lignin-protein complex in cell walls of *Pinus elliottii*: Amino acid constituents," *Phytochemistry* 21(2), 315-318. DOI: 10.1016/S0031-9422(00)95258-0
Wilkins, O., Nahal, H., Foong, J., Provart, N. J., and Campbell, M. M. (2009). "Expansion and diversification of the *Populus* R2R3-MYB family of transcription factors," *Plant Physiol.* 149(2), 981-993. DOI: 10.1104/pp.108.132795
Yang, B., and Wyman, C. E. (2008). "Pretreatment: The key to unlocking low-cost cellulosic ethanol," *Biofuels, Bioprod. Biorefin.* 2(1), 26-40. DOI: 10.1002/bbb.49
Yang, X., Kalluri, U. C., Jawdy, S., Gunter, L. E., Yin, T., Tschaplinski, T. J., Weston, D. J., Ranjan, P., and Tuskan, G. A. (2008). "The F-box gene family is expanded in herbaceous annual plants relative to woody perennial plants," *Plant Physiol.* 148(3), 1189-1200. DOI: 10.1104/pp.108.121921
Yoshiba, Y., Aoki, C., Iuchi, S., Nanjo, T., Seki, M., Sekiguchi, F., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2001). "Characterization of four extensin genes in *Arabidopsis thaliana* by differential gene expression under stress and non-stress conditions," *DNA Res.* 8(3), 115-122. DOI: 10.1093/dnares/8.3.115

Article submitted: November 19, 2015; Peer review completed: February 6, 2016; Revised version received: March 25, 2016; Accepted: March 31, 2016; Published: April 13, 2016.
DOI: 10.15376/biores.11.2.4727-4744