



Cite this: DOI: 10.1039/c6gc01193d

Impact of engineered lignin composition on biomass recalcitrance and ionic liquid pretreatment efficiency†

Jian Shi,^{a,b,c} Sivakumar Pattathil,^{d,e} Ramakrishnan Parthasarathi,^{a,b} Nickolas A. Anderson,^f Jeong Im Kim,^f Sivasankari Venketachalam,^{d,e} Michael G. Hahn,^{d,e} Clint Chapple,^f Blake A. Simmons^{a,g} and Seema Singh^{*a,b}

Lignin plays important biological functions in plant cell walls, but also contributes to the recalcitrance of the walls to deconstruction. In recent years, genetic modification of lignin biosynthesis pathways has become one of the primary targets of plant cell wall engineering. In this study, we used a combination of approaches to characterize the structural and compositional features of wild-type *Arabidopsis* and mutants with distinct lignin monomer compositions: *fah1-2* (Guaiacyl, G-lignin dominant), *C4H-F5H* (Syringyl, S-lignin dominant), *COMT1* (G/5-hydroxy G-lignin dominant), and a newly developed *med5a med5b ref8* (*p*-hydroxyphenyl, H-lignin dominant) mutant. In order to understand how lignin modification affects biomass recalcitrance, substrate reactivity and lignin fractionation, we correlated these properties with saccharification efficiency after ionic liquid (IL) pretreatment. Results showed that the cleavage of β -O-4 linkages in the H- or S-lignin mutants was greater than that in G-lignin mutants. Furthermore, density functional theory (DFT) based calculations indicate higher chemical reactivity of the linkages between H- and S-lignin monomers, a possible cause of the reduced recalcitrance of H- or S-lignin mutants. Glycome profiling was conducted to study the impact of lignin modification on overall composition, extractability, integrity and lignin-associated features of most major non-cellulosic cell wall glycans in these mutants. This study provides insights into the role of lignin monomer composition on the enzymatic digestibility of biomass and the effect of lignin modification on overall wall structure and biomass pretreatment performance.

Received 27th April 2016,
Accepted 15th June 2016
DOI: 10.1039/c6gc01193d

www.rsc.org/greenchem

1. Introduction

Lignins are heterogeneous phenylpropanoid-derived polymers that are integrated with cellulose and hemicelluloses to form

plant secondary cell walls where they provide strength and rigidity of water-conducting and supportive tissues. Lignin is synthesized from coniferyl alcohol, sinapyl alcohol, and relatively small amounts of *p*-coumaryl alcohol, which upon polymerization form G, S and H units in lignin.^{1,2} It is believed that the presence of lignin hinders fermentable sugar extraction from the bulk of plant biomass by blocking enzyme access to cellulose and hemicelluloses.³ Therefore, genetic manipulation of lignin biosynthesis is being investigated as a strategy to reduce recalcitrance of bioenergy crops.^{4,5} Multiple strategies have been explored to modify the lignin biosynthesis pathways, including: (1) decreasing lignin content, which however, often poses a challenge in crops or trees, because of the potential for severe effects on biomass yield and integrity in vessels;⁶ (2) relocating lignin deposition so that lignin deposition is reduced in tissues other than vessels such that vessel integrity is maintained;^{7,8} (3) altering the relative proportions of lignin building blocks;^{4,9} and (4) introducing readily cleavable linkages into the backbone of the lignin polymer or altering linkages between lignin and polysaccharides.^{10,11}

^aDeconstruction Division, Joint BioEnergy Institute, 5885 Hollis St, Emeryville, CA 94608, USA. E-mail: seesing@sandia.gov; Fax: +1 510-486-4252; Tel: +1 925-294-4551

^bBiological and Materials Science Center, Sandia National Laboratories, 7011 East Avenue, Livermore, CA 94551, USA

^cDepartment of Biosystems and Agricultural Engineering, University of Kentucky, Lexington, KY 40546, USA

^dComplex Carbohydrate Research Center, University of Georgia, 315 Riverbend Road, Athens, GA 30602, USA

^eBioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

^fDepartment of Biochemistry, Purdue University, 175 South University Street, West Lafayette, IN 47907, USA

^gBiological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6gc01193d

The effects of lignin on enzymatic digestibility have received extensive attention. In general, it has been observed that the lower the lignin content of plant biomass, the higher the fermentable sugar release during enzymatic saccharification.^{12–15} Besides lignin content, there are a few other lignin-related factors that can influence digestibility, such as lignin composition (S/G ratio), the location of lignin deposition and its structure in biomass, and the linkages within lignin-carbohydrate complexes (LCCs). Lignin S/G ratios have been identified as one structural feature that impacts the recalcitrance of plant biomass, with higher S/G ratios leading to greater sugar release yields after pretreatment.^{16–18} Others have argued that factors beyond lignin content and S/G ratio may significantly influence biomass recalcitrance and sugar release. For example, a recent study on a collection of *Populus trichocarpa* samples has identified several unusual phenotypes that did not follow the dependency of sugar release performance on lignin content.¹⁹ These contrasting observations suggest that multiple factors influence biomass recalcitrance to conversion and digestibility. Thus, further investigations into the molecular features that define the impact of lignin on biomass recalcitrance are warranted.

Arabidopsis is a powerful model system for plant cell wall research. In this species, as well as in others, manipulations of lignin composition and deposition have, in some cases, resulted in stunted growth and developmental abnormalities *via* mechanism(s) that is/are currently poorly understood.⁸ For instance, disruption of the gene encoding the phenylpropanoid biosynthetic enzyme *p*-coumaroyl shikimate 3'-hydroxylase (C3'H), required for the production of G- and S- (but not H-) lignins, results in dwarfing and sterility leading to the suggestion that H-lignin may not be sufficient for normal plant growth and development. Surprisingly, a recent study showed that the growth and fertility of a C3'H-deficient *Arabidopsis* mutant could be rescued to near wild-type levels by disrupting two subunits of the Mediator complex without restoring C3'H function.⁹ These mutant plants contain lignin that is composed almost entirely of H subunits, indicating that H-lignin is sufficient for relatively normal growth and development, at least in *Arabidopsis*. This work demonstrated that manipulation of H-lignin levels may be a promising strategy for developing biofuel-suitable feedstocks.

Efficient utilization of lignocellulosic biomass usually requires pretreatment to loosen the cell wall structure by means of partially or fully dissolving/removing some components of the biomass. The ability of certain ionic liquids (ILs), *e.g.* 1-ethyl-3-methylimidazolium acetate [C₂C₁Im][OAc], to solubilize cellulose and/or lignin provides a highly effective pretreatment method and enables the possibility of lignin upgrading and valorization to improve overall biorefinery economics.^{20,21} Recent studies have demonstrated the preferential breakdown of G- and S-lignin during IL pretreatment using [C₂C₁Im][OAc] is dependent on both pretreatment conditions and the type of biomass feedstocks.^{17,22} The mechanism of lignin solubilization and depolymerization during IL pretreatment is not well understood, and there are several gaps in the

current understanding on how the composition and relative abundances of the phenylpropanoids in lignin affect the mode of depolymerization, cellulose reactivity and cellulase accessibility and their correlations with saccharification efficiency.

The objectives of this study are to: (1) characterize and compare the chemical composition and abundance of lignins present in untreated and IL-pretreated *Arabidopsis* genotypes that vary in their lignin structures using wet chemistry methods. These genotypes include *COL* (wild type), and the *fah1-2* (G-lignin dominant), *C4H-F5H* (S-lignin dominant), *COMT1* (G/5-hydroxy guaiacyl lignin dominant) and the newly developed H-lignin dominant *med5a med5b ref8* mutants; (2) monitor overall changes in wall structure and integrity using ELISA-based glycome profiling of non-cellulosic cell wall glycans; (3) monitor sugar yields during enzymatic hydrolysis of untreated and pretreated *Arabidopsis* samples and correlate these to lignin content and lignin monomer compositions; and (4) investigate changes in lignin inter-unit linkages, including β -aryl ethers (β -O-4), phenylcoumarans (β -5), resinols (β - β), and dibenzodioxocins (DBZO), lignin subunit compositions [guaiacyl (G), syringyl (S), and *p*-hydroxyphenyl (H)], and *p*-hydroxycinnamoyl ferulate (FA) and *p*-coumarate (*p*CA) contents of different *Arabidopsis* genotypes before and after IL pretreatment.

2. Results and discussion

2.1 Characterization of untreated *Arabidopsis* genotypes

Modification of the lignin biosynthetic pathway can be expected to not only change lignin content and composition, but also to affect other cell wall components in the plant biomass, such as cellulose and non-cellulosic matrix polysaccharides such as hemicelluloses and pectins. We determined the lignin content of five unpretreated *Arabidopsis* genotypes as measured by the two-step acid hydrolysis (Klason) lignin method and the acetyl bromide lignin method. We also analyzed the G-, S-, H-lignin subunit contents, using derivatization followed by reductive cleavage (DFRC) methods. Table 1 shows the composition of the unpretreated *Arabidopsis* samples. As the lignin biosynthesis pathways were genetically modified in the tested mutants, significant variations in Klason lignin content were observed across different genotypes. The *fah1-2* mutant has the greatest Klason lignin content (24.5%), followed by the *med5a med5b ref8* mutant, wild type and the *COMT1* mutant; while the *C4H-F5H* mutant contained the least amount of lignin (11.9%). It has been shown that *C4H-F5H* up-regulated (S-lignin dominant) transgenic poplar contains increased levels of acid-soluble lignin and reduced levels of acid-insoluble (Klason) lignin in comparison to the wild-type trees.¹⁸ Similarly, we observed a decrease in apparent lignin content using standard Klason lignin methods. It was also noticed that the sum of Klason lignin plus ethanol extractives²³ in the *C4H-F5H* line and the wild type are almost identical. The decrease in Klason lignin content in the *C4H-F5H* line is possibly due to the loss of extractable/soluble lignin during

Table 1 Chemical composition of Arabidopsis samples before and after IL pretreatment^a

	Sample ID Genotype	<i>COL</i> Wild type (WT)	<i>C4H-F5H</i> S	<i>med5a med5b ref8</i> H	<i>fah1-2</i> G	<i>COMT1</i> G/G'
Untreated	Glucan, %	24.7 ± 1.1	23.5 ± 0.5	14.7 ± 0.6	25.7 ± 1.2	25.4 ± 0.3
	Xylan, %	11.7 ± 0.5	12.7 ± 0.3	8.3 ± 0.7	12.6 ± 0.9	12.4 ± 0.7
	Klason lignin, %	17.4 ± 1.0	11.9 ± 1.1	19.0 ± 0.9	24.5 ± 1.3	14.7 ± 0.6
	AcBr lignin, %	28.5 ± 6.0	21.4 ± 5.6	20.9 ± 6.5	24.3 ± 4.8	28.9 ± 4.6
	Extractives, %	29.5 ± 2.3	35.5 ± 3.5	42.8 ± 2.1	26.4 ± 3.0	29.7 ± 1.9
Pretreated	Solid recovery, %	59.9 ± 1.6	56.1 ± 0.2	37.4 ± 1.8	62.3 ± 0.7	61.5 ± 2.8
	Glucan, %	38.5 ± 1.9 (7.2)	39.8 ± 2.1 (4.8)	35.1 ± 2.6 (5.2)	40.5 ± 3.2 (3.2)	40.7 ± 2.3 (1.2)
	Xylan, %	14.0 ± 0.4 (26.7)	18.6 ± 0.7 (15.6)	10.6 ± 0.5 (47.1)	15.6 ± 1.1 (22.8)	14.7 ± 0.8 (26.4)
	Klason lignin, %	20.0 ± 1.2 (24.0)	10.5 ± 0.9 (45.5)	13.7 ± 1.1 (70.4)	21.9 ± 2.0 (38.7)	17.5 ± 1.3 (19.3)

^a Compositions reported for untreated sample are based on the dry weight of untreated biomass; solid recoveries are based on the dry weight of untreated biomass, while the compositions for pretreated biomass are based on the dry weight of pretreated biomass; values in parentheses are percentage removal of each component (glucan, xylan or lignin) during IL pretreatment based on its original amount in untreated biomass; AcBr, acetyl bromide.

the two-step sulfuric acid hydrolysis. However, the determination of acid-soluble lignin using the two-step sulfuric acid hydrolysis method can be very challenging especially for genetically modified plant materials. For comparison, lignin content was also determined by an acetyl bromide (AcBr) method²⁴ and results indicate that the discrepancy between Klason lignin and AcBr lignin could be accounted for, at least in part, by the extracted lignin. Thus, the acetyl bromide method may provide more representative estimations on the total lignin contents of genetically modified Arabidopsis samples.

Compositional analyses of the lignins present in the different Arabidopsis genotypes were determined by DFRC method as described in previous studies^{25,26} (Table 2). *COL* (wild type, WT) contains predominantly G-lignin (68%) and S-lignin (31%), with a small amount of H-lignin (1.3%). In contrast, the *C4H-F5H* double mutant contained lignin that was almost exclusively composed of S-subunits (92%). The *fah1-2* and *COMT1* mutants both contain lignin composed primarily of G-lignin or 5-hydroxyl G-lignin. Although the *COMT1* mutants showed a similar G-lignin content as the *fah1-2* mutant, the ratio of eluted guaiacyl/5-hydroxyl guaiacyl monomers during pyrolysis was different as identified by the pyrolysis-GC/MS analysis, indicating the partial substitution of guaiacyl monomers with 5-hydroxy guaiacyl lignin monomers in the *COMT1* mutant (Fig. S1†). The newly developed *med5a med5b ref8* mutant contains ~89% H-lignin, and small amounts of S- and G-lignin types. Although lignin content and

extractability varied, the glucan and xylan contents, as determined following two-step acid hydrolysis protocol, in *COL* and the *fah1-2*, *C4H-F5H*, and *COMT1* mutants are similar. In contrast, the *med5a med5b ref8* mutant has much lower glucan and xylan contents than the other mutants.

2.2 Glycome profiling of lignin biosynthetic mutants

Changes in the overall non-cellulosic glycan composition and glycan extractabilities in the walls of various lignin biosynthetic mutants were examined by glycome profiling of cell walls isolated from each plant line. In these analyses, three independent biological replicates of each line were used and the data were represented as heatmaps generated from the average values of the raw data as shown in Fig. 1. Analyses of these averaged glycome profiling data delineated statistically significant changes in the extractabilities of cell wall glycans in the lignin biosynthetic mutant lines indicating an overall change in the cell wall architecture in these plants that result from changes in their lignin compositions. Compared to the wild-type lines, in general, all lignin biosynthetic mutant lines exhibited significant differences in their respective glycome profiles. The major variations observed are demarked in the figure as colored dotted blocks and arrows (Fig. 1). An enhanced abundance of xylan epitopes (comprised of both unsubstituted and substituted xylans) recognized by the xylan-3 through -7 groups of monoclonal antibodies (mAbs) (yellow dotted blocks) was observed in the carbonate extracts from all mutants lines. A similar enhancement in the abundance of

Table 2 Composition of H, G, S type lignin in untreated Arabidopsis genotypes

Sample ID		<i>COL</i> (WT)	<i>C4H-F5H</i> (S)	<i>med5a med5b ref8</i> (H)	<i>fah1</i> (G)	<i>COMT1</i> (G/G')
Amount (mmol g ⁻¹ cell wall)	H	0.9 ± 0.1	2.5 ± 0.3	33.5 ± 4.7	1.0 ± 0.6	0.8 ± 0.5
	G	45.9 ± 4.8	3.0 ± 0.2	2.0 ± 0.2	47.2 ± 2.4	31.9 ± 8.3
	S	20.7 ± 2.2	64.3 ± 10.1	1.2 ± 0.4	0.9 ± 0.3	0.5 ± 0.1
	Total	67.6 ± 7.0	69.8 ± 10.1	28.3 ± 3.7	49.0 ± 3.1	33.3 ± 8.5
mol%	H	1.3 ± 0.1%	3.6 ± 0.7%	88.9 ± 1.7%	1.9 ± 1.0%	2.6 ± 1.4%
	G	68.0 ± 0.2%	4.4 ± 0.8%	7.0 ± 0.8%	96.4 ± 1.5%	95.8 ± 1.6%
	S	30.7 ± 0.2%	92.0 ± 1.5%	4.2 ± 1.4%	1.7 ± 0.6%	1.6 ± 0.2%

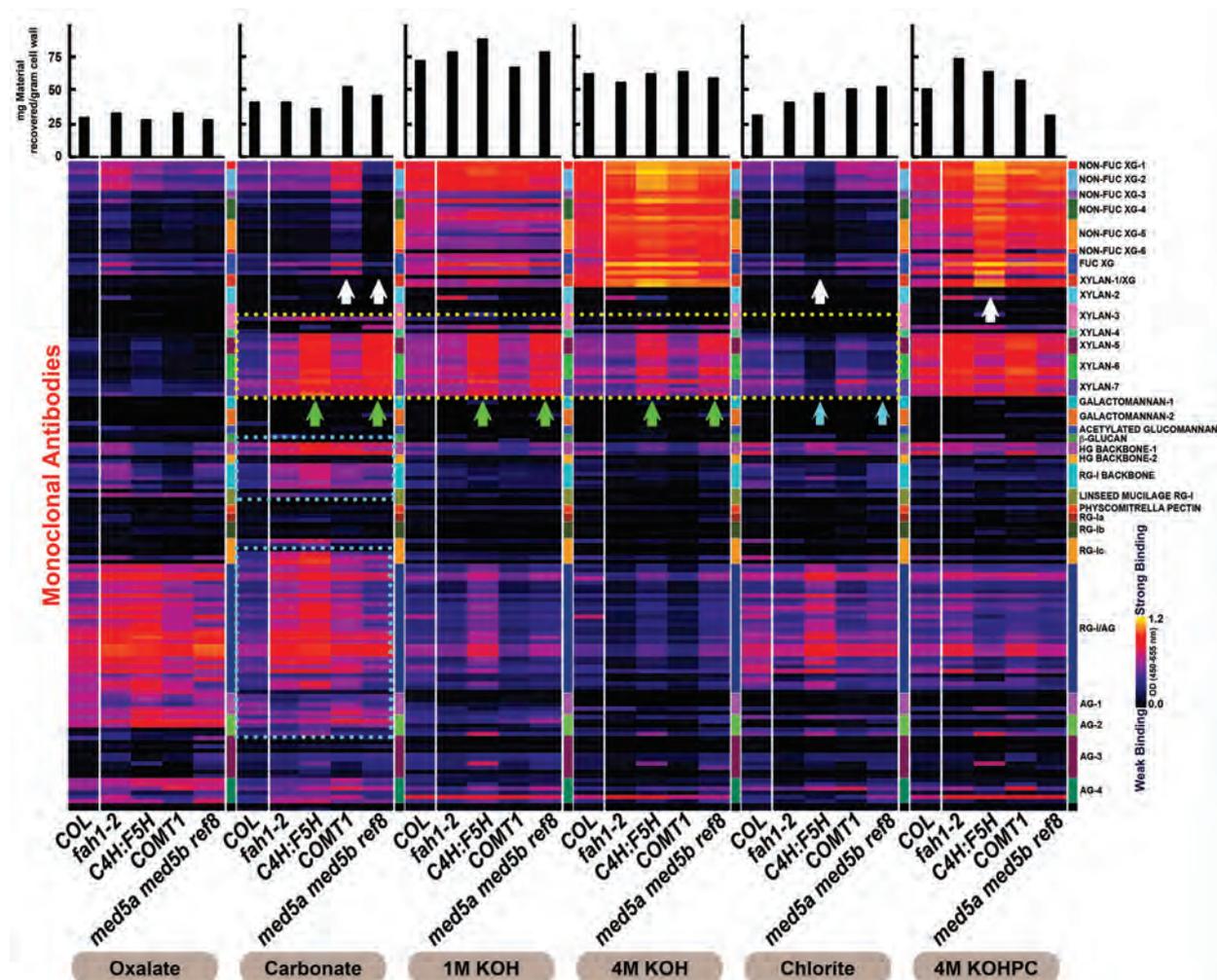


Fig. 1 Glycome profiling of cell walls isolated from stems of Arabidopsis wild type (*COL*), *fah1-2*, *C4H-F5H*, *COMT1* and *med5a med5b ref8* mutants. Sequential extracts were prepared from cell walls of each plant line using ammonium oxalate, sodium carbonate, potassium hydroxide (1 M KOH, 4 M KOH and 4 M KOHPC) and sodium chlorite as explained in the Methods section. These cell wall extracts were screened by ELISA using 155 plant glycan-directed mAbs (see ESI Table S1†). The resulting binding response data are represented as heatmaps with a bright yellow-red-black scale indicating the strength of the ELISA signal (bright yellow-red and black colors depict strong, medium, and no binding, respectively). The mAbs are grouped based on the cell wall glycans that they principally recognize as depicted in the panel at right-hand side of the figure. The gravimetric amounts of materials present in each extract are depicted as bar graphs at the top of the heatmaps.

pectic backbone epitopes [recognized by the homogalacturonan backbone-1 (HG backbone-1) and RG-I backbone groups of mAbs] and pectic arabinogalactan epitopes (recognized by the RG-I/AG group of mAbs) was observed in the carbonate extracts obtained from the mutant lines (light blue blocks).

Interestingly, in the case of *C4H-F5H* and *med5a med5b ref8* lines, more intense binding of most xylan directed mAbs was noted in the carbonate extract (green arrows) compared to other mutant lines. In *C4H-F5H* and *med5a med5b ref8* lines, such increases in the binding intensities of xylan-directed mAbs were also evident in the 1 M KOH and 4 M KOH extracts (green arrows). A concomitant reduction in the abundance of these xylan epitopes was noted in the chlorite extracts from these two lines, potentially due to reduced interactions

between xylan and lignin in these lines (light blue arrows). These results indicate more severe changes in the wall architecture in these two lines causing an enhanced extractability of xylans from the walls in concert with potentially reduced lignin-xylan associations. More subtle differences were also evident in the glycome profiles of the other lignin biosynthetic mutant lines. For instance, in the carbonate extract, a marginally enhanced extractability of xyloglucan epitopes was evident in *COMT1*, while *med5a med5b ref8* did not exhibit any changes in the extractability of xyloglucan epitopes (white arrows). In addition, *C4H-F5H* showed a reduced abundance of xyloglucan epitopes in the chlorite extract with a concomitant increase of these epitopes in the 4 M KOHPC extract. Statistical analysis of the averaged glycome profiling data validated the conclusions drawn above (see ESI Table S1†), and showed

significant variations on an extract-to-extract comparison basis. Principal components analyses substantiated that the variations noted in extract-to-extract comparisons among various lines are due to the differences in the plant genotypes. ANOVA delineated changes largely identical to the changes observed in the glycome profiles depicted in Fig. 1. In summary, the glycome profiling studies showed that altering the structure and composition of the lignin components results in overall changes in the structural architecture of cell walls. Changes in the wall architecture in *C4H-F5H* and *med5a med5b ref8* lines cause an enhanced extractability of xylans from the wall in concert with potentially reduced lignin–xylan associations.

2.3 Ionic liquid pretreatment of Arabidopsis samples

The significant compositional changes (Table 1) and altered extractabilities of specific classes of non-cellulosic glycans (Fig. 1) seen in the walls from the different untreated Arabidopsis genotypes examined in this study led us to examine the effects of pretreatment under mild conditions with an ionic liquid, $[C_2C_1Im][OAc]$, on biomass recalcitrance in these plant lines. Table 1 shows the composition of both untreated Arabidopsis and the pretreated materials recovered after mild IL pretreatment (120 °C for 1 h). After IL pretreatment, 56–62% of the solids were recovered for *COL*, *fah1-2*, *C4H-F5H*, *COMT1*, but solids recovery for the *med5a med5b ref8* triple mutant having an H genotype was much lower (37%). The acid-insoluble lignin contents of the recovered biomass varied, with pretreated *fah1-2* having the highest lignin content (22%), while pretreated *C4H-F5H* has the lowest lignin content (11%), indicating significant lignin removal during IL pretreatment of this mutant. Generally speaking, greater lignin removal was observed for the *C4H-F5H* (S-lignin dominant) and *med5a med5b ref8* (H-lignin dominant) mutants, compared with wild type and the G-lignin dominant mutants. Further, glycome profiling analyses (Fig. 1) had indicated that the most significant alterations in non-cellulosic glycan extractabilities occur in these same mutants hinting that overall cell wall architecture in these lines is more affected in comparison to other genotypes. These results indicate that H- and S-type lignins might be more labile to IL pretreatment, probably due to their relatively linear structure and the lower average dissociation energy of the inter-unit bonds in H- and S-lignin,^{27,28} and perhaps also due to reduced associations between lignin and non-cellulosic glycans in these plants.

2.4 Sugar yield from enzymatic hydrolysis of untreated and pretreated Arabidopsis samples

The effect of IL pretreatment on enzymatic saccharification of wild-type and mutant Arabidopsis biomass was examined under low and high enzyme loadings (Fig. 2). Saccharification of raw untreated Arabidopsis samples liberated approximately 20–35% of the glucose from biomass, except for the *med5a med5b ref8* (H-lignin dominant) mutant, from which 45% and 75% of the glucose was released at low and high enzyme loadings, respectively. Saccharification of IL pretreated Arabidopsis led to 70–80% glucose yield at low enzyme loading of 3 mg

enzyme protein per g pretreated biomass in 72 h. In contrast, at high enzyme loading of 10 mg enzyme protein per g pretreated biomass, nearly complete (>90% glucose yield) saccharification of IL-pretreated Arabidopsis was achieved within 24 h. The *fah1-2* samples with G-lignin-rich walls gave a lower glucose yield compared to other genotypes ($p < 0.05$), while the glucose yield from saccharification of IL-pretreated S-rich *C4H-F5H* were not significantly higher than those of other genotypes ($p > 0.05$). These results agree with a previous study showing enhanced saccharification of Arabidopsis that containing S-rich lignin in comparison with that of a G-lignin dominant genotype after liquid hot water pretreatment.²⁵ The surprisingly high sugar yield from saccharification of the untreated *med5a med5b ref8* (H-lignin dominant) mutant poses the question whether the lignin content, and/or the lignin type and lignin removal during IL pretreatment contribute dominantly to differences in sugar yields. Interestingly, the saccharification results obtained here do not correlate with the increased glycan extractability observed for both the *med5a med5b ref8* and *C4H-F5H* genotypes, which gave similar glycome profiles (Fig. 1), but showed different saccharification efficiencies after IL pretreatment (Fig. 2). Other studies have documented a correlation between enhanced cell wall extractability of non-cellulosic glycans and reduced recalcitrance, including in some lignin biosynthetic mutants.^{29,30} These data suggest that for the *med5a med5b ref8* mutant, the type of lignin had a greater influence on recalcitrance than did glycan extractability, perhaps due to variations in lignin-wall glycans crosslinks resulting from altered lignin compositions in these two genotypes.

In order to further elucidate the effect of lignin removal during IL pretreatment on sugar yield, we plotted the lignin content in untreated Arabidopsis genotypes and lignin removal during IL pretreatment as a function of the initial hydrolysis rate for untreated biomass and pretreated samples, respectively (as depicted in Fig. 3). For untreated Arabidopsis lines, no positive correlation was observed between lignin content and initial hydrolysis rate; however, it appeared that lignin genotype has significant effect on digestibility. In agreement with the overall high sugar yields shown in Fig. 2, the *med5a med5b ref8* (H-lignin dominant) mutant was more digestible than wild type and the other mutants regardless of lignin contents. Fig. 3b shows that IL pretreatment led to increases in initial hydrolysis rate for all tested genotypes, and minimized the differences between genotypes, although slightly higher initial hydrolysis rates were achieved for the *C4H-F5H* (S-lignin dominant) and *med5a med5b ref8* (H-lignin dominant) mutants. Thus, no positive correlation was observed between initial hydrolysis rate and lignin removal during pretreatment.

Results from this study do not appear to match some previous reports that the lower the lignin content of plant biomass, the higher the fermentable sugar release during enzymatic saccharification.^{12–14} However, a number of studies have shown that lignin content is only one of the contributing factors to biomass recalcitrance.^{19,31,32} Lignin-related factors, such as lignin composition (S/G ratio), the spatial distribution of lignins

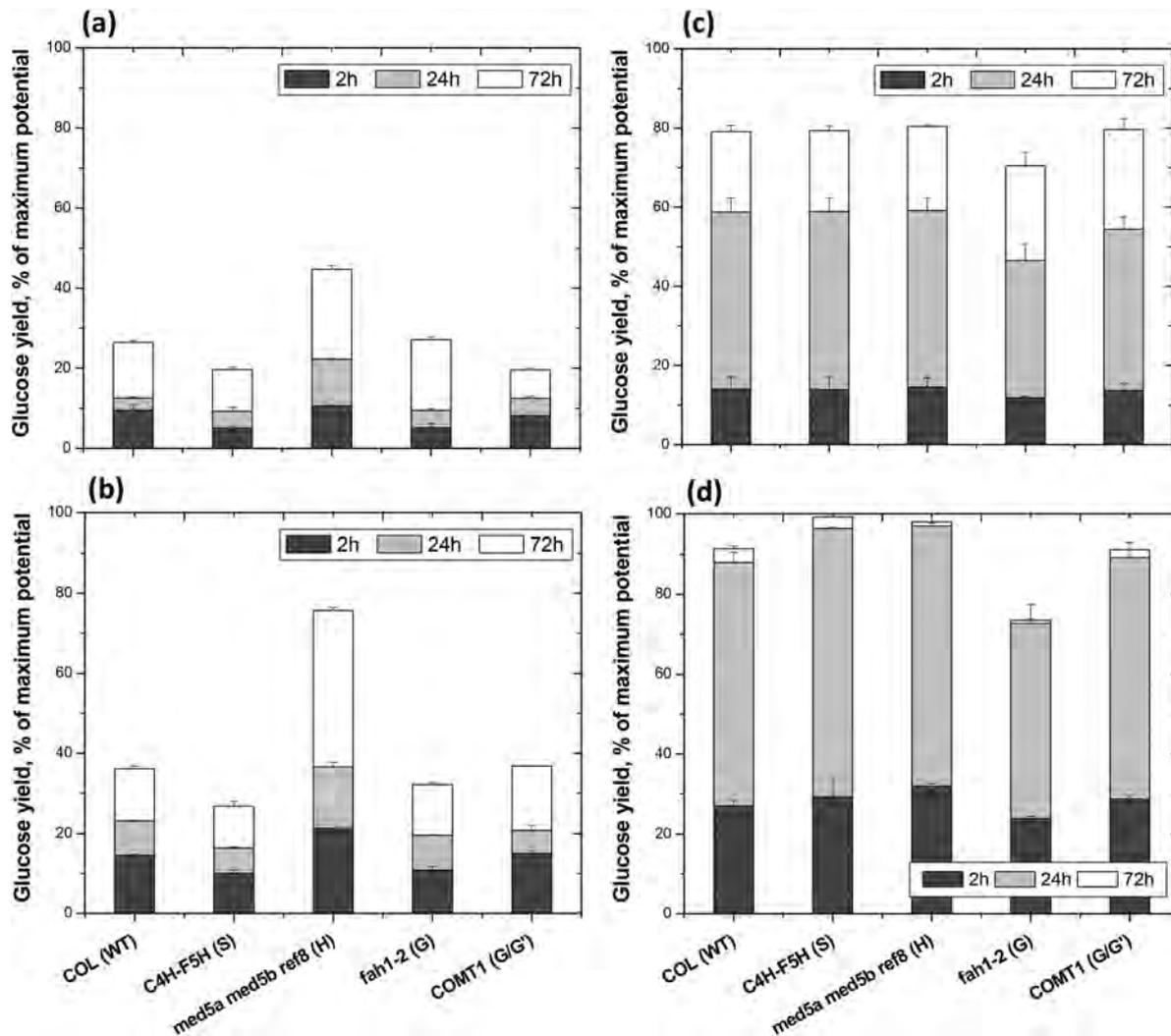


Fig. 2 Enzymatic glucan digestibility of untreated and IL pretreated Arabidopsis genotypes including *COL* (wild type), *fah1-2* (G-lignin dominant), *C4H-F5H* (S-lignin dominant), *COMT1* (G/5-hydroxy G-lignin dominant) and *med5a med5b ref8* (H-lignin dominant) mutants at (a) 3 mg, and (b) 10 mg enzyme protein per g biomass for untreated Arabidopsis samples, and (c) 3 mg, and (d) 10 mg enzyme protein per g biomass for IL pretreated Arabidopsis samples.

in biomass, and the linkages within lignin-carbohydrate complexes (LCCs) along with other non-lignin-related factors such as cellulose crystallinity, biomass porosity, and extractable compounds (*e.g.*, pectins) all influence biomass recalcitrance and affect sugar release during enzymatic saccharification.^{3,16–18,33}

The effects of lignin content differences may have been evened out by IL pretreatment due to the high effectiveness of certain ILs to deconstruct biomass.^{17,25,34} Altogether, the results presented here suggest that when comparing wild type and genetically engineered transgenic Arabidopsis, the lignin genotype outweighs lignin content in determining the recalcitrance/enzymatic digestibility of the biomass from this plant.

2.5 Lignin depolymerization during IL pretreatment

Pyrolysis-GC/MS analysis of untreated and pretreated Arabidopsis samples revealed distinct chromatographic patterns of lignin-derived compounds, but it is not clear from the

pyrolysis-GC/MS data alone how lignin was solubilized and depolymerized during IL pretreatment (Fig. S1†). In order to obtain additional information on this point, we monitored the lignin molecular weight distribution during IL pretreatment using gel permeation chromatography (GPC) of the extracted lignin samples (Fig. S2†). The relative areas of excluded and retained regions ($A_{E/R}$) in the chromatograms were used as an indicator of the size distribution of solubilized lignin after IL pretreatment, as reported in Table 3. Lignin was extracted from different streams during IL pretreatment and enzymatic hydrolysis, including, L_1 : lignin from untreated biomass, L_2 : solubilized lignin in $[C_2C_1Im][OAc]$, L_3 : lignin remaining in pretreated biomass. Lignin extracted from untreated Arabidopsis samples (L_1) showed a strong signal in the excluded region ($t < 13.4$ min) for all tested genotypes. The $A_{E/R}$ ranged from 2.8 to 5.5, suggesting that lignin of untreated Arabidopsis consisted mainly of large molecular mass materials. It should also

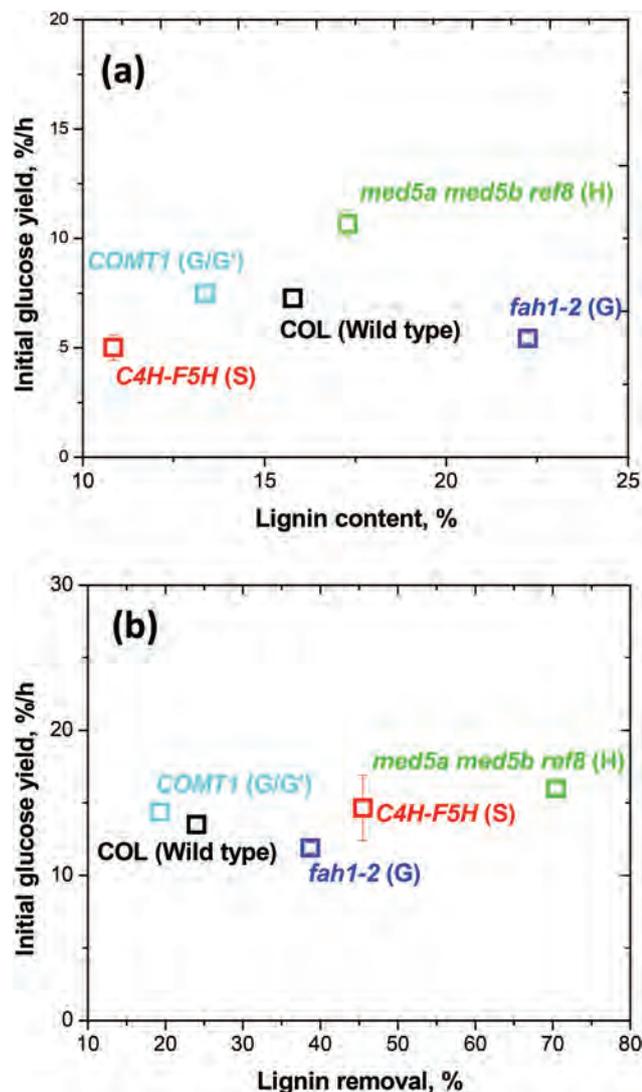


Fig. 3 (a) Lignin content in untreated Arabidopsis genotypes including COL (wild type), *fah1-2* (G-lignin dominant), *C4H-F5H* (S-lignin dominant), *COMT1* (G/5-hydroxy G-lignin dominant) and *med5a med5b ref8* (H-lignin dominant) mutants and (b) lignin removal during IL pretreatment depicted as a function of enzymatic glucan digestibility. The error bars represent the standard deviation of three replicates.

be noted that the S- and H-genotypes exhibited lower $A_{E/R}$ than those of wild type and G- or G/G'-genotypes ($p < 0.05$), indicating that lignins in the S- and H-genotypes tend to be small in terms of the average molecular mass materials.

As for the lignin solubilized during IL pretreatment (L_2), a distinct signal in the retained region ($t > 13.4$ min) was observed with $A_{E/R}$ in a range of 0.5 to 0.8 for different Arabidopsis genotypes, suggesting that lignin was solubilized and depolymerized in the liquid stream as previously observed.^{35,36} Interestingly, the $A_{E/R}$ for the S- or H-lignin dominant mutants was higher than wild type Arabidopsis and the G-lignin dominant mutants ($p < 0.05$), though the lignin removal for S- or H-lignin dominant mutants was greater than for the others. Due to the chemical nature of their corresponding building blocks, sinapyl alcohol and *p*-coumaryl alcohol, the S- and H-lignins show relatively linear structures and lower average bond dissociation energies between inter-unit linkages. This indicates that, under mild IL pretreatment conditions, the S- and H-lignin structures were readily cleaved into relatively large fractions before further depolymerization occurs. In contrast, for wild type or G-lignin dominant mutants, because the lignin structures are relatively rigid, only small fractions were cleaved off and become soluble, and, in turn, low $A_{E/R}$ were observed on IL-solubilized lignins in these plants. The lignins that remained in the pretreated solids (L_3) underwent severe depolymerization as indicated by the significantly smaller $A_{E/R}$ compared with that of untreated substrates across all Arabidopsis genotypes ($p < 0.05$).

2.6 2D HSQC NMR spectra of different regions for Arabidopsis cell wall biomass

2D HSQC NMR spectra of the aliphatic, anomeric and aromatic regions for untreated and pretreated Arabidopsis cell wall biomass are shown in Fig. S4.† The aliphatic regions of all spectra showed the presence of major inter-unit linkages: β -O-4 aryl ether (structure A), resinol (β - β , structure B), phenylcoumaran (β -5, structure C). Spirodienones were detected in all samples. The aromatic regions of the spectra confirm the presence of all S, G and H subunits in wild-type Arabidopsis, while

Table 3 Elution time and relative molecular mass of lignin extracted from different streams during IL pretreatment and enzymatic hydrolysis. L_1 : lignin from untreated biomass, L_2 : solubilized lignin in $[C_2C_{11}Im][OAc]$, L_3 : lignin remaining in pretreated biomass^a

Region		Elution time (min)	COL (WT)	<i>C4H-F5H</i> (S)	<i>med5a med5b ref8</i> (H)	<i>fah1-2</i> (G)	<i>COMT1</i> (G/G')
L_1	Excluded (%)	$t < 13.4$	84.7	73.7	78.9	81.7	82.2
	Retained (%)	$t > 13.4$	15.3	26.3	21.1	18.3	17.8
	$A_{Excluded/Retained}$ ($A_{E/R}$)	—	5.52 ^{AA}	2.81 ^{DA}	3.74 ^{CA}	4.47 ^{BA}	4.61 ^{BA}
L_2	Excluded (%)	$t < 13.4$	35.5	42.7	43.3	34.2	39.8
	Retained (%)	$t > 13.4$	64.5	57.3	56.7	65.8	60.2
	$A_{Excluded/Retained}$ ($A_{E/R}$)	—	0.55 ^{CC}	0.75 ^{AC}	0.76 ^{AC}	0.52 ^{CC}	0.66 ^{BC}
L_3	Excluded (%)	$t < 13.4$	63.8	52.3	56.6	61.6	58.8
	Retained (%)	$t > 13.4$	36.2	47.7	43.4	38.4	41.2
	$A_{Excluded/Retained}$ ($A_{E/R}$)	—	1.76 ^{AB}	1.10 ^{CB}	1.31 ^{BB}	1.60 ^{AB}	1.43 ^{AB}

^a $A_{E/R}$ stands for the ratio of peak areas in the excluded and retained regions; the first letters (A, B, C) on the right of the $A_{E/R}$ data represent the levels of significance within five Arabidopsis genotypes (among columns); the second letters (A, B, C) on the right of the $A_{E/R}$ data represent the levels of significance within three lignin streams: L_1 , L_2 , and L_3 , (among rows); significance is determined based on the p -values generated by Tukey's studentized range (HSD) test.

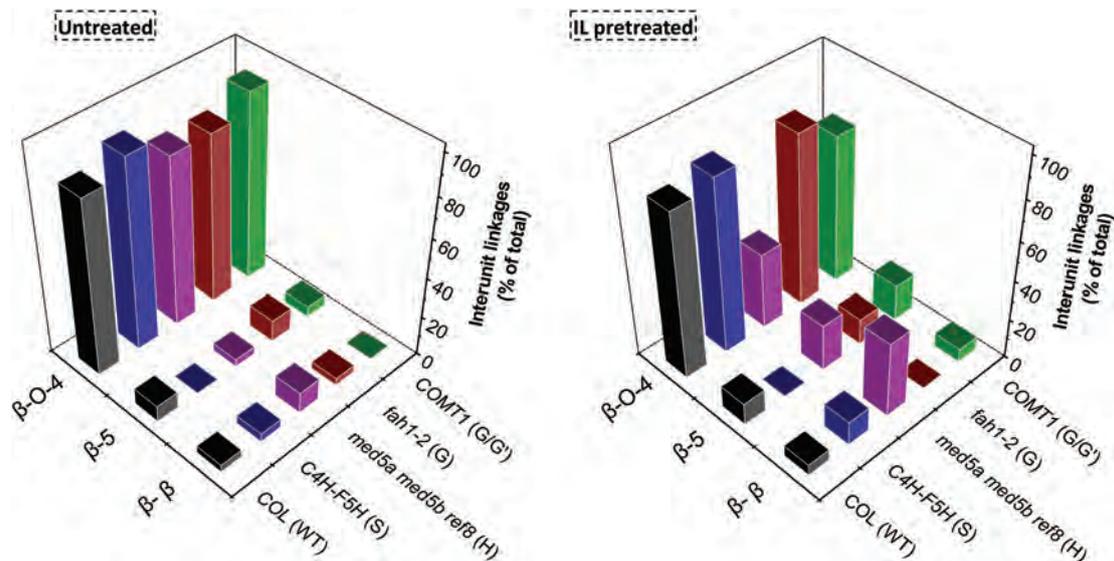


Fig. 4 Schematic diagram of changes in lignin interunit linkages [(β -O-4) β -aryl ethers, (β -5) phenylcoumarans, and (β - β) resinols] of different *Arabidopsis* genotypes including *COL* (wild type), *fah1-2* (G-lignin dominant), *C4H-F5H* (S-lignin dominant), *COMT1* (G/5-hydroxy G-lignin dominant) and *med5a med5b ref8* (H-lignin dominant) mutants before and after IL pretreatment.

H, S, or G subunits dominate in the various mutants and transgenic lines as expected.^{8,9,25} We further characterized and compared the lignin inter-unit linkages and the S, G, H ratios in both untreated and IL-pretreated *Arabidopsis* samples (Fig. 4). Results of these analyses show that β -aryl ethers (β -O-4, >85%) are the most abundant inter-unit linkages in untreated *Arabidopsis* samples followed by phenylcoumaran (β -5) or resinol (β - β) (<15% in sum) linkages depending on the lignin genotypes. The relative abundance of β -O-4 linkages in pretreated *Arabidopsis* samples decreased as a result of IL pretreatment, especially for the *med5a med5b ref8* (H-lignin dominant) mutant. Taken together, our results indicate that: (1) IL pretreatment cleaves β -O-4 linkages to a greater extent than β -5

or β - β linkages; and (2) β -O-4 linkages in *med5a med5b ref8* (H-lignin dominant) mutants are more vulnerable to IL pretreatment. No dibenzodioxocin (substructure D) was observed in any pretreated *Arabidopsis* samples, suggesting the removal of lignin branches. These results indicate that mixtures of smaller, more linear molecules are produced by IL treatment and this result is consistent with the GPC data (Table 3).

2.7 Chemical reactivity of inter-unit lignin linkages

Density functional theory (DFT) based calculations were carried out to understand the stability and reactivity of inter-unit lignin linkages upon manipulation of H, G and S units by replacing S and G units with H unit to form β -O-4, β -5, and

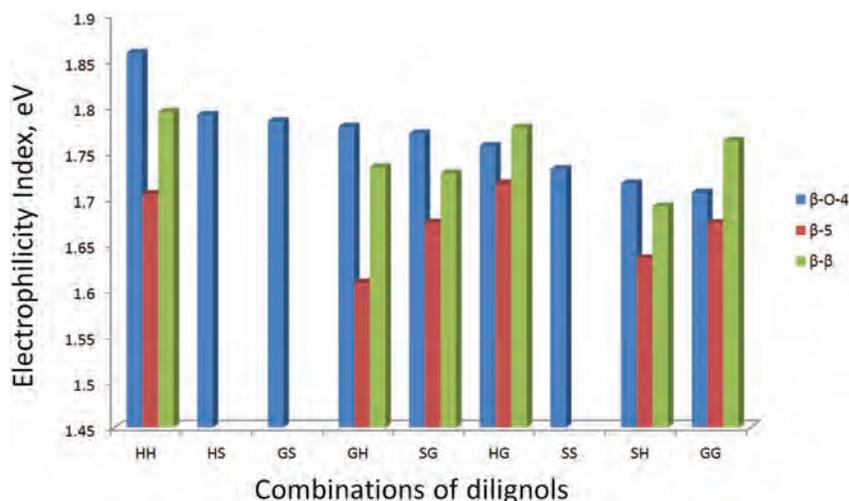


Fig. 5 Electrophilicity index as a descriptor of chemical reactivity of lignin dimers linked by (β -O-4) β -aryl ethers, (β -5) phenylcoumarans, and (β - β) resinols. See Fig. S4† for the optimized geometries of inter-unit lignin linkages.

β - β linkages. The optimized most stable geometries of β -O-4, β -5, and β - β linkages of lignin model compounds are shown in Fig. S4.† The electrophilicity index was calculated as a descriptor of reactivity that allows a quantitative classification of the global electrophilic nature of a molecule within a relative scale. As shown in Fig. 5, H-based lignin dimers have a higher electrophilicity index, indicating that they are chemically more reactive compared to other S- and G-lignin dimers. Replacing an S or G unit by an H unit increases the reactivity for GH (β -5 and β - β linkages) and SH (β -O-4 linkage) lignin dimers. Results show that H-based lignins are more chemically reactive (less stable) than other type of lignins irrespective of chemical pretreatment, a possible cause of the reduced recalcitrance of H-type dominant lignin mutants.

3. Conclusions

The impact of lignin composition on biomass recalcitrance is not fully understood, mainly due to the complications arising from the remarkable variances in plant type, the origin and maturity of plant materials, the cell-wall composition, and the selection of pretreatment method. In this study, we selected a diverse collection of Arabidopsis lines with distinct lignin monomer compositions. For the first time, we are able to characterize the structural and compositional features and study the impact of lignin modification on lignin-carbohydrate complex characteristics and the deconstruction of cell-wall compounds using a mild $[C_2C_1Im][OAc]$ ionic liquid (IL) pretreatment, decoupling many interfering factors. Glycome profiling revealed that altering the structure and composition of lignin components results in overall changes in the structural architecture of the cell walls, which in turn leads to changes in the extractability of wall polymers. Lignin composition significantly affects biomass recalcitrance, substrate reactivity and lignin removal during IL pretreatment, all of which are correlated with saccharification efficiency. Furthermore, as revealed by chromatography and spectrometry techniques, the variations in lignin monomer compositions led to different modes of lignin dissolution and depolymerization during $[C_2C_1Im][OAc]$ IL pretreatment. More interestingly, we found that the cleavage of β -O-4 linkages in the H-lignin dominant mutant was greater than those in the G-lignin dominant mutants, which can be explained by DFT-based chemical reactivity calculations. This study provides insights into the role of lignin monomer composition on the enzymatic digestibility of biomass and the effect of lignin modification on cell wall architecture and biomass pretreatment performance.

4. Experimental

4.1 Materials

The Arabidopsis *fah1-2*, *C4H-F5H*, *med5a med5b ref8*, and *COMT1* mutant lines were previously described.^{4,9} Plants were

grown in a greenhouse at Purdue University at a temperature range of 22–25 °C under ambient lighting during the winter of 2013 in standard potting mixture (Redi-Earth Plug and Seedling Mixture; Sun Gro Horticulture, Bellevue, WA, USA). For the glycome profiling experiments, dried stems were harvested from fully mature Arabidopsis plants and ground by a Wiley Mill through a 0.5 mm screen (Thomas Scientific, Swedesboro, NJ). For the other experiments, the ground stems were further separated by a vibratory sieve system (Endecotts, Ponte Vedra, FL). The Arabidopsis stem fragments falling between 20 and 80 mesh were collected for use in this study. 1-Ethyl-3-methylimidazolium acetate was purchased from BASF (Basionics™ BC-01, BASF, Florham Park, NJ) and used as the IL for all pretreatments. Commercial enzyme products, cellulase (Cellic® CTec2, Batch#VCN10007) and hemicellulase (Cellic® HTec2, Batch#VHN00002) were gifts from Novozymes, North America (Franklinton, NC). The protein contents of CTec2 and HTec2 were measured as 186.6 mg ml⁻¹ and 180.1 mg ml⁻¹, respectively, as determined with the Pierce™ bicinchoninic acid Protein Assay Kit (Thermo Scientific, Rockford, IL) using bovine serum albumin as protein standard.³⁷

4.2 IL pretreatment of Arabidopsis samples

A mild (120 °C for 1 h) IL pretreatment was used to process five Arabidopsis samples in sealed glass tubes each containing 0.3 g biomass and 3 g of $[C_2C_1Im][OAc]$ to give approximate 10% (w/w) biomass loading. After pretreatment, 8 mL of hot water were thoroughly mixed with the pretreatment slurry for cellulose regeneration. The mixtures of IL, water, and pretreated biomass were centrifuged to separate the solid and liquid phases. An aliquot of each supernatant was taken for lignin and sugar analysis. The remaining solids were washed 4 times with 10 mL of DI water and lyophilized in a FreeZone Freeze Dry System (Labconco, Kansas City, MO) for composition analyses and saccharification tests.

4.3 Enzymatic hydrolysis of pretreated Arabidopsis samples

Enzymatic saccharification of untreated and IL-pretreated Arabidopsis samples were run in duplicate by following the NREL LAP 9 “Enzymatic Saccharification of Lignocellulosic Biomass” protocol at NREL standard conditions (50 °C, 0.05 M citrate buffer, pH 4.8).³⁸ Citrate buffer (final molarity 50 mM), sodium azide (antimicrobial, final concentration of 0.02 g L⁻¹), enzymes, and DI water were mixed with pretreated solids to achieve a final solids loading of around 5% (w/w) [equivalent to 2.5% (w/w) glucan concentration]. CTec2 and HTec2 (Novozymes) were used at enzyme loadings of 3 and 10 mg CTec2 protein per g pretreated biomass supplemented with HTec2 at loadings of 0.07 and 0.26 mg enzyme protein per g glucan, respectively. Aliquots of each supernatant were taken at 2, 24, and 72 h of hydrolysis and were analyzed by HPLC for monosaccharides as described in the Analytical Methods section below. Enzymatic digestibility was defined as the glucose yield based on the maximum potential glucose from the glucans present in the biomass.

4.4 Analytical methods

Klason lignin, glucan, and xylan contents in the untreated and IL-pretreated *Arabidopsis* samples were determined using a two-step acid hydrolysis process according to the NREL Laboratory Analytical Procedure.³⁹ Monomeric sugars (glucose, xylose) and cellobiose released into the acid hydrolyzates were measured using an Agilent 1100 series HPLC equipped with a Biorad Aminex HPX-87H ion exchange column and refractive index detector, using 4 mM H₂SO₄ as the mobile phase at a flow rate of 0.6 mL min⁻¹ and a column temperature of 60 °C. Solids remaining after two-stage acid hydrolysis as held in Gooch crucible (medium porosity) were dried at 105 °C overnight, weighed, heated in a muffle furnace at 575 °C for 4 h, and weighed again. The mass of the oven-dried solids subtracted by the mass of ash in furnace-burned sample corresponds to the amount of Klason lignin (acid-insoluble lignin). Lignin contents were also determined by the acetyl bromide (AcBr) and DFRC method.^{24,26} The saccharification hydrolyzates were separated by centrifugation at 14 000g for 10 min followed by syringe filtration. The amounts of cellobiose, glucose, xylose, and arabinose released in the hydrolyzates were measured by HPLC following the same procedure as described above.

4.5 Sequential extraction and glycome profiling

Sequential cell wall extractions and glycome profiling of various biomass samples were carried out as described previously.⁴⁰ Plant glycan-directed mAbs⁴¹ were from laboratory stocks (CCRC, JIM and MAC series) at the Complex Carbohydrate Research Center (available through CarboSource Services; <http://www.carbosource.net>) or were obtained from BioSupplies (Australia) (BG1, LAMP). A description of the mAbs used in this study can be found in the ESI, Table S1,† which includes links to a web database, WallMabDB (<http://www.wallmabdb.net>) that provides detailed information about each antibody.

Statistical analysis of glycome profiling results was conducted using JMP Genomics 6.1 (SAS Institute). For these analyses, various plant lines (wild type and mutants) and cell wall fractions (oxalate, carbonate, 1 M KOH, 4 M KOH, chlorite and 4 M KOHPC) were considered to be fixed effects while biological replicates were considered random effects.

4.6 Characterization of lignin in liquid and residual solids

To understand changes in lignin molecular weight distributions during IL pretreatments, gel permeation chromatography (GPC) was performed on the lignin extracted from untreated wild-type *Arabidopsis* (L₁) by using an enzymatic mild acidolysis (EMAL) process,⁴² the soluble lignin in pretreatment hydrolyzates (L₂), and the insoluble lignin extracted from the residual solids after enzymatic hydrolysis using the EMAL process (L₃). An aliquot of the lignin stream (10 µL for L₁ and 2 mg for L₂ and L₃, respectively) was dissolved in 1 mL of *N*-methyl-2-pyrrolidone (NMP) and injected onto an Agilent 1200 series binary LC system (G1312B) equipped with a DA

(G1315D) detector, a PL-Gel™ 5 µm Mixed-D column (PL1110-6504, 300 mm L × 7.5 mm i.d., resolving range of 200 to 400 000 MW, Agilent Technologies, Santa Clara, CA) at 80 °C using NMP as the mobile phase at a flow rate of 0.5 mL per min. Absorbance of materials eluting from the column was detected at 290 nm (UV-A). Intensities were area normalized and molecular mass estimates were determined after calibration of the system with polystyrene standards.^{36,43}

Untreated and IL-pretreated *Arabidopsis* samples were analyzed by pyrolysis-GC/MS analysis to reveal the changes in lignin monomer compositions before and after IL pretreatment following a method described in Varanasi *et al.*¹⁷ In brief, roughly 1 mg of solids were loaded in a quartz tube and sealed with glass wool. The quartz tube with sample was pyrolyzed at 550 °C for 10 seconds using a Pyroprobe 5200 from CDS Analytical, Inc. (Chemical Data Systems, Oxford, PA, USA). The pyrolysis products were then injected into a GC/MS (Thermo Electron Corporation with Trace GC and Polaris-Q MS, now merged by Thermo Fisher Scientific, Austin, TX) equipped with a TR-SMS analytical column using helium as the carrier gas. The pyrolysis-GC/MS spectra of untreated and pretreated *Arabidopsis* samples are presented in Fig. S1.† The areas under the peaks of the chromatograms were calculated using the Qual browser software (1.4 SR1, Thermo Electron Corporation, 1998–2003). Percentages of H, G, and S type lignin were calculated based on the ratio of the summary peak areas of identified individual lignin compounds.

4.7 2D ¹³C-¹H HSQC NMR spectroscopy on untreated and IL pretreated biomass

Untreated *Arabidopsis* samples were extracted by water and 80% (v/v) ethanol to remove extractives and ball milled as previously described.²³ The milled samples (~50 mg) were then placed in NMR tubes with 600 µL DMSO-*d*₆ using a minute amount of [C₂C₁Im][OAc] as a co-solvent.^{44,45} The samples were sealed and sonicated until homogenous in a Branson 2510 table-top cleaner (Branson Ultrasonic Corporation, Danbury, CT) with the temperature of the bath closely monitored and maintained below 55 °C. For IL-pretreated *Arabidopsis*, 200 mg sample was enzymatically hydrolyzed with a combination of cellulase and hemicellulase (CTec2 and HTec2 in 10 : 1 v/v) at a dose of 20 mg enzyme protein per g pretreated biomass in shaker at 50 °C for 96 h. The residual lignin-rich samples were washed twice with 0.01 N hydrochloric acid water solution (pH = 2) and three times with water (10 mL each wash) and lyophilized in a FreeZone Freeze Dry System (Labconco, Kansas City, MO). The dried sample was milled with a bead beater and 15 mg of the milled sample was added to 600 µL DMSO-*d*₆ and sonicated until homogenous. The homogeneous lignin solutions were transferred to NMR tubes.

HSQC spectra were acquired at 25 °C using a Bruker Avance-600 MHz instrument equipped with a 5 mm inverse-gradient ¹H/¹³C cryoprobe using a q_hsqcetgp pulse program (ns = 200 for cell wall and 64 for lignin, ds = 16, number of increments = 256, *d*₁ = 1.0 s).⁴⁶ Chemical shifts were

referenced to the central DMSO peak (δ_C/δ_H 39.5/2.5 ppm). Assignment of the HSQC spectra was described elsewhere.⁴⁷ A semi-quantitative analysis of the volume integrals of the HSQC correlation peaks was performed using Bruker's Topspin 3.1 (Windows) processing software. A cosine squared function was applied to both F_2 (LB = -0.05, GB = 0.001) and F_1 (LB = -0.01, GB = 0.001) prior to 2D Fourier Transformation. Semi-quantitative evaluation of interunit linkages in lignins has been typically expressed as number of specific interunit linkages per 100 aromatic (lignin) monomers or C_9 units.^{43,48}

4.8 Density functional theory (DFT) calculation

The geometry optimizations of coniferyl, sinapyl, and *p*-coumaryl alcohol with inter-unit linkages, such as β -O-4, β -5, and β - β linkages of lignin model compounds were performed using density functional theory (DFT) with the M06-2X hybrid exchange–correlation functional and the 6-31+G(2d,2p) basis set.²⁸ Frequency calculations were carried out to verify that the computed structures corresponded to energy minima. In the present study, the DFT-based global reactivity descriptor-electrophilicity index was calculated to compare the chemical reactivity of all the lignin compounds in this study.⁴⁹

Acknowledgements

Work conducted at the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 and through the Center for Direct Catalytic Conversion of Biomass to Biofuels (C3Bio), an Energy Frontier Research Center funded by the U.S. DOE, Office of Science, Office of Basic Energy Sciences, Award number DE-SC0000997. The glycome profiling was supported by the BioEnergy Science Center administered by Oak Ridge National Laboratory, and funded by a grant (DE-AC05-00OR22725) from the Office of Biological and Environmental Research, Office of Science, United States, Department of Energy. The generation of the CCRC series of plant cell wall glycan-directed monoclonal antibodies used in this work was supported by the NSF Plant Genome Program (DBI-0421683 and IOS-0923992). This research used resources of the National Energy Research Scientific Computing Center and EMSL at Pacific Northwest National Laboratory. We thank Dr Noppadon Sathitsuksanoh and Dr Jeffrey Pelton for help on NMR and SEC analysis. We also acknowledge the National Science Foundation under Cooperative Agreement No. 1355438 for partially supporting the effort at University of Kentucky.

References

- C. L. Li, B. Knierim, C. Manisseri, R. Arora, H. V. Scheller, M. Auer, K. P. Vogel, B. A. Simmons and S. Singh, *Bioresour. Technol.*, 2010, **101**, 4900–4906.
- A. J. Ragauskas, G. T. Beckham, M. J. Biddy, R. Chandra, F. Chen, M. F. Davis, B. H. Davison, R. A. Dixon, P. Gilna, M. Keller, P. Langan, A. K. Naskar, J. N. Saddler, T. J. Tschaplinski, G. A. Tuskan and C. E. Wyman, *Science*, 2014, **344**, 1246843.
- M. E. Himmel, S.-Y. Ding, D. K. Johnson, W. S. Adney, M. R. Nimlos, J. W. Brady and T. D. Foust, *Science*, 2007, **315**, 804–807.
- N. D. Bonawitz, W. L. Soltau, M. R. Blatchley, B. L. Powers, A. K. Hurlock, L. A. Seals, J.-K. Weng, J. Stout and C. Chapple, *J. Biol. Chem.*, 2012, **287**, 5434–5445.
- Y. Mottiar, R. Vanholme, W. Boerjan, J. Ralph and S. D. Mansfield, *Curr. Opin. Biotechnol.*, 2016, **37**, 190–200.
- J. H. Jung, W. Vermerris, M. Gallo, J. R. Fedenko, J. E. Erickson and F. Altpeter, *Plant Biotechnol. J.*, 2013, **11**, 709–716.
- H. Yang, H. F. Sun, S. J. Zhang, B. D. Wu and B. C. Pan, *Environ. Sci. Pollut. Res.*, 2015, **22**, 10882–10889.
- N. D. Bonawitz and C. Chapple, *Curr. Opin. Biotechnol.*, 2013, **24**, 336–343.
- N. D. Bonawitz, J. Im Kim, Y. Tobimatsu, P. N. Ciesielski, N. A. Anderson, E. Ximenes, J. Maeda, J. Ralph, B. S. Donohoe and M. Ladisch, *Nature*, 2014, **509**, 376–380.
- C. Wilkerson, S. Mansfield, F. Lu, S. Withers, J.-Y. Park, S. Karlen, E. Gonzales-Vigil, D. Padmakshan, F. Unda and J. Rencoret, *Science*, 2014, **344**, 90–93.
- A. Ziebell, K. Gracom, R. Katahira, F. Chen, Y. Pu, A. Ragauskas, R. A. Dixon and M. Davis, *J. Biol. Chem.*, 2010, **285**, 38961–38968.
- S. L. Voelker, B. Lachenbruch, F. C. Meinzer and S. H. Strauss, *New Phytol.*, 2011, **189**, 1096–1109.
- F. Chen and R. A. Dixon, *Nat. Biotechnol.*, 2007, **25**, 759–761.
- H. Shen, X. He, C. R. Poovaiah, W. A. Wuddineh, J. Ma, D. G. Mann, H. Wang, L. Jackson, Y. Tang and C. Neal Stewart, *New Phytol.*, 2012, **193**, 121–136.
- J. Xu, X. Zhang, R. R. Sharma-Shivappa and M. W. Eubanks, *Bioresour. Technol.*, 2012, **116**, 540–544.
- B. H. Davison, S. R. Drescher, G. A. Tuskan, M. F. Davis and N. P. Nghiem, *Appl. Biochem. Biotechnol.*, 2006, **130**, 427–435.
- G. Papa, P. Varanasi, L. Sun, G. Cheng, V. Stavila, B. Holmes, B. A. Simmons, F. Adani and S. Singh, *Bioresour. Technol.*, 2012, **117**, 352–359.
- J. J. Stewart, T. Akiyama, C. Chapple, J. Ralph and S. D. Mansfield, *Plant Physiol.*, 2009, **150**, 621–635.
- M. H. Studer, J. D. DeMartini, M. F. Davis, R. W. Sykes, B. Davison, M. Keller, G. A. Tuskan and C. E. Wyman, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 6300–6305.
- D. Klein-Marcuschamer, B. A. Simmons and H. W. Blanch, *Biofuels, Bioprod. Biorefin.*, 2011, **5**, 562–569.
- N. M. Konda, J. Shi, S. Singh, H. W. Blanch, B. A. Simmons and D. Klein-Marcuschamer, *Biotechnol. Biofuels*, 2014, **7**, 1.
- P. Varanasi, P. Singh, M. Auer, P. D. Adams, B. A. Simmons and S. Singh, *Biotechnol. Biofuels*, 2013, **6**, 14.
- S. D. Mansfield, H. Kim, F. Lu and J. Ralph, *Nat. Protoc.*, 2012, **7**, 1579–1589.

- 24 R. S. Fukushima and R. D. Hatfield, *J. Agric. Food Chem.*, 2001, **49**, 3133–3139.
- 25 X. Li, E. Ximenes, Y. Kim, M. Slininger, R. Meilan, M. Ladisch and C. Chapple, *Biotechnol. Biofuels*, 2010, **3**, 1.
- 26 F. Lu and J. Ralph, *J. Agric. Food Chem.*, 1997, **45**, 2590–2592.
- 27 D. W. Cho, R. Parthasarathi, A. S. Pimentel, G. D. Maestas, H. J. Park, U. C. Yoon, D. Dunaway-Mariano, S. Gnanakaran, P. Langan and P. S. Mariano, *J. Org. Chem.*, 2010, **75**, 6549–6562.
- 28 R. Parthasarathi, R. A. Romero, A. Redondo and S. Gnanakaran, *J. Phys. Chem. Lett.*, 2011, **2**, 2660–2666.
- 29 Q. Zhao, Y. Tobimatsu, R. Zhou, S. Pattathil, L. Gallego-Giraldo, C. Fu, L. A. Jackson, M. G. Hahn, H. Kim, F. Chen, J. Ralph and R. A. Dixon, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 13660–13665.
- 30 S. Pattathil, T. Saffold, L. Gallego-Giraldo, M. O'Neill, W. S. York, R. A. Dixon and M. G. Hahn, *Ind. Biotechnol.*, 2012, **8**, 217–221.
- 31 R. Kumar and C. E. Wyman, *Biotechnol. Prog.*, 2009, **25**, 807–819.
- 32 X. Zhang, J. Xu and J. J. Cheng, *Energy Fuels*, 2011, **25**, 4796–4802.
- 33 S.-Y. Ding, Y.-S. Liu, Y. Zeng, M. E. Himmel, J. O. Baker and E. A. Bayer, *Science*, 2012, **338**, 1055–1060.
- 34 J. Shi, J. M. Gladden, N. Sathitsuksanoh, P. Kambam, L. Sandoval, D. Mitra, S. Zhang, A. George, S. W. Singer, B. A. Simmons and S. Singh, *Green Chem.*, 2013, **15**, 2579–2589.
- 35 A. George, K. Tran, T. J. Morgan, P. I. Benke, C. Berruenco, E. Lorente, B. C. Wu, J. D. Keasling, B. A. Simmons and B. M. Holmes, *Green Chem.*, 2011, **13**, 3375–3385.
- 36 J. Shi, K. Balamurugan, R. Parthasarathi, N. Sathitsuksanoh, S. Zhang, V. Stavila, V. Subramanian, B. A. Simmons and S. Singh, *Green Chem.*, 2014, **16**, 3830–3840.
- 37 J. Shi, K. W. George, N. Sun, W. He, C. Li, V. Stavila, J. D. Keasling, B. A. Simmons, T. S. Lee and S. Singh, *Bio-Energy Res.*, 2015, **8**, 1004–1013.
- 38 M. Selig, N. Weiss and Y. Ji, Enzymatic Saccharification of Lignocellulosic Biomass: Laboratory Analytical Procedure (LAP): Issue Date, 3/21/2008, National Renewable Energy Laboratory, 2008.
- 39 J. B. Sluiter, R. O. Ruiz, C. J. Scarlata, A. D. Sluiter and D. W. Templeton, *J. Agric. Food Chem.*, 2010, **58**, 9043–9053.
- 40 S. Pattathil, U. Avci, J. S. Miller and M. G. Hahn, in *Biomass Conversion: Methods and Protocols, Methods in Molecular Biology*, ed. M. Himmel, Humana Press, New York, NY, 2012, vol. 908, pp. 61–72.
- 41 S. Pattathil, U. Avci, D. Baldwin, A. G. Swennes, J. A. McGill, Z. Popper, T. Bootten, A. Albert, R. H. Davis and C. Chennareddy, *Plant Physiol.*, 2010, **153**, 514–525.
- 42 A. Guerra, I. Filpponen, L. A. Lucia and D. S. Argyropoulos, *J. Agric. Food Chem.*, 2006, **54**, 9696–9705.
- 43 N. Sathitsuksanoh, K. M. Holtman, D. J. Yelle, T. Morgan, V. Stavila, J. Pelton, H. Blanch, B. A. Simmons and A. George, *Green Chem.*, 2014, **16**, 1236–1247.
- 44 J.-L. Wen, S.-L. Sun, B.-L. Xue and R.-C. Sun, *Materials*, 2013, **6**, 359–391.
- 45 K. Cheng, H. Sorek, H. Zimmermann, D. E. Wemmer and M. Pauly, *Anal. Chem.*, 2013, **85**, 3213–3221.
- 46 S. Heikkinen, M. M. Toikka, P. T. Karhunen and I. A. Kilpeläinen, *J. Am. Chem. Soc.*, 2003, **125**, 4362–4367.
- 47 H. Kim and J. Ralph, *Org. Biomol. Chem.*, 2010, **8**, 576–591.
- 48 M. Sette, R. Wechselberger and C. Crestini, *Chem. – Eur. J.*, 2011, **17**, 9529–9535.
- 49 P. K. Chattaraj, U. Sarkar and D. R. Roy, *Chem. Rev.*, 2006, **106**, 2065–2091.