A TEMPO-catalyzed oxidation-reduction method to probe surface and anhydrous crystalline-core domains of cellulose microfibril bundles

Tanã M. Shiga

University of São Paulo

Haibing Yang

Purdue University

Bryan W. Penning

USDA-ARS Wheat Quality Research Unit

Anna T. Olek

Purdue University

Maureen C. McCann

Purdue University

N.C. Carpita (✉ carpita@purdue.edu)

Purdue University

Research Article

Keywords: Arabidopsis (*Arabidopsis thaliana*), Catalytic delignification, Cellulose microfibrils, Cellulose bundling, Cotton (*Gossypium hirsutum*), Maize (*Zea mays*), Poplar (*Populus tremula* × *P. alba*)

Posted Date: March 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-188813/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Cellulose on April 15th, 2021. See the published version at <https://doi.org/10.1007/s10570-021-03815-9>.

Abstract

A modified TEMPO-catalyzed oxidation of the solvent-exposed glucosyl units of cellulose to uronic acids, followed by carboxyl reduction with NaBD 4 to 6-deutero- and 6,6-dideutero-glucosyl units, provided a robust method for determining relative proportions of disordered amorphous, ordered surface chains, and anhydrous core-crystalline residues of cellulose microfibrils inaccessible to TEMPO. Both glucosyl residues of cellobiose units, digested from amorphous chains of cellulose with a combination of cellulase and cellobiohydrolase, were deuterated, whereas those from anhydrous chains were undeuterated. By contrast, solvent-exposed and anhydrous residues alternate in surface chains, so only one of the two residues of cellobiosyl units was labeled. Although current estimates indicate that each cellulose microfibril comprises only 18 to 24 (1,4)- β -D-glucan chains, we show here that microfibrils of walls of Arabidopsis leaves and maize coleoptiles, and those of secondary wall cellulose of cotton fibers and poplar wood, bundle into much larger macrofibrils, with 67 to 86% of the glucan chains in the anhydrous domain. These results indicate extensive bundling of microfibrils into macrofibrils occurs during both primary and secondary wall formation. We discuss how, beyond lignin, the degree of bundling into macrofibrils contributes an additional recalcitrance factor to lignocellulosic biomass for enzymatic or chemical catalytic conversion to biofuel substrates.

Introduction

Cellulose microfibrils of flowering plants are commonly described as *para*-crystalline arrays containing up to 36 (1 $\oplus$ 4)- β -D-glucan chains with diameters of $\sim$ 3.2-3.5 nm (Ha et al. 1998; Kennedy et al. 2007a, b). Some microfibril diameters as low as 2 nm have been inferred on the basis of the ratio of surface to core chains evaluated by solid-state NMR (Ha et al. 1998), and more recent re-evaluation of solid-state NMR data indicate that diameters range from 2.3 to 3.0 nm, corresponding to a total of only 18 to 24 glucan chains in the crystalline core domain (Fernandes et al. 2011; Newman et al. 2013; Thomas et al. 2013; Jarvis 2013). These studies show that not all glucan chains are in crystalline conformations, that surface chains must adopt a different conformation than the crystalline core, and that additional amorphous chains can be accommodated (Vitor et al. 2002). Wide-angle X-ray scattering, which captures the contributions of surface chains in addition to those of the crystalline core, indicate larger diameters of 3.3 to 3.6 nm (Liu et al. 2013; Thomas et al. 2014), with adherent amorphous chains contributing to larger diameters than predicted for the crystalline cores alone. Zhao et al. (2007) show that amorphous cellulose in cotton must be removed by acid hydrolysis to reveal the microfibrillar structure, and imaging by atomic-force microscopy (AFM) suggests an 'elemental fibril' of 3nm x 5nm in maize stems (Ding and Himmel 2006). The consensus of these data suggests that the crystalline cores of an elemental microfibril are around 18 to 24 chains.

Regardless of the numbers of chains in microfibrils made by cellulose synthase complexes, microfibrils also aggregate into macrofibril bundles (Ding et al. 2014). Deuterium exchange experiments with plant native cellulose showed phases of rapid and much slower exchanged, interpreted as exchange with hydrogens in amorphous vs. crystalline domains (Frilette et al. 1948). Deuterium exchange followed by

rehydrogenation revealed an intermediate fraction of slow deuteration distinct from deuterium-resistant domains (Jeffries, 1963; Lindh and Salmén 2017). Fourier-transform infrared spectroscopy of these exchange properties indicated that cotton cellulose is a large aggregate of three bundles, each about 144 chains (Lindh and Salmén 2017). Based on electron microscopy, wood macrofibrils were variable in diameter, ranging from 14 to 23 nm, to as high as 60 nm (Donaldson 2007). Cotton fibers, for which secondary walls are nearly 98% cellulose (Meinert and Delmer 1977), showed bundling diameters of 20 to 30 nm after removal of amorphous cellulose by acid hydrolysis (Zhao et al. 2007). Imaging by AFM of macrofibrils in rehydrated state and without prior fixation indicated that elementary fibrils of maize stem cells coalesce into macrofibrils of over 20 nm in diameter (Ding et al. 2012). Thus, the recalcitrance of cellulose to biological or chemical conversion to biofuel substrates has at two levels of complexity: the variation in the number of chains in a single crystalline microfibril and the extent of bundling of microfibrils into macrofibrils that create larger zones of anhydrous material.

Here, we adapt a cellulose oxidation protocol to determine relative proportions of amorphous chains, surface chains of crystalline cellulose microfibrils, and the crystalline core chains. When catalyzed by 2,2,6,6-tetramethylpyperidine-1-oxyl (TEMPO), the *O*-6 primary alcohols of glucosyl residues exposed to water are oxidized by hypochlorite to yield glucuronic acid (de Nooy et al. 1995; Saito and Isogai 2004; Okita et al. 2010). In our modified method, we activated glucuronosyl oxidation products with a water-soluble diimide and then reduced them with NaBD₄, creating 6-deutero- and 6,6-dideuteroglucosyl residues that could be distinguished from undeuterated glucose by gas chromatography-mass spectrometry (GC-MS). Using this protocol, all glucosyl residues of amorphous cellulose and disordered surface chains would be expected to be oxidized, whereas none would be oxidized in the crystalline core domains that exclude TEMPO. However, in ordered surface chains, only one-half of the glucosyl residues have their *O*-6 primary alcohol exposed to the aqueous environment, as the other half are buried in the anhydrous domain of the crystalline core (Fig. 1). Thus, when partly oxidized and NaBD₄-reduced cellulose is enzymatically digested to cellobiose units and analyzed by electrospray ionization (ESI) MS, both glucosyl residues of formerly amorphous chains are 6- or 6,6-dideuterated, whereas units from the crystalline core remain undeuterated. However, the cellobiose units from chains that occupy the ordered surface domains will uniquely have only one 6-deutero or 6,6-dideuterated glucosyl residue (Fig. 1).

We show here that 86% of glucan chains of cellulose of cotton (*Gossypium hirsutum*) fibers and 67 to 72% of those of hybrid poplar (*Populus tremula* × *P. alba*) wood fibers were in an anhydrous domain unable to be oxidized by TEMPO catalysis. The bundling of microfibrils into larger anhydrous domains is not strictly a property of secondary wall cellulose, as the anhydrous domains of primary wall cellulose microfibrils of *Arabidopsis thaliana* leaf or etiolated maize (*Zea mays*) coleoptiles were 74 to 81% of the total residues, respectively. Although catalytic delignification of poplar wood particles more than doubles saccharification yield (Yang et al. 2019), this treatment increases the proportion of anhydrous domains from 67 to 72% to 74 to 79%. Swelling of the cotton fiber cellulose in trifluoroacetic acid (TFA) followed by gelatinization greatly increased the amorphous cellulose to almost 50% of the mass, but the remainder condensed into insoluble chains with ratios of anhydrous:surface chain of 2:1,

corresponding to fibrils of about 36 chains. The results of these experiments indicate that cellulose microfibrils aggregate into macrofibrils of over 100 chains and might represent a more significant factor of recalcitrance in the deconstruction of lignocellulosic biomass than previously reported.

Materials And Methods

Plant Materials

Cotton cellulose was obtained from Sigma Life Science (Sigmacell; Sigma-Aldrich Co., St Louis, MO; Cat no. S5504-500G). Barley mixed-linkage β -glucan, pectic 5-arabinan, and larch 4-xylan were obtained from Megazyme, Ireland. Arabidopsis (*Arabidopsis thaliana* Col-0) 21-d-old plantlets and maize (*Zea mays*) 3-d-old etiolated coleoptiles plus etiolated shoots were grown and isolated as described by Okekeogbu et al. (2019). Milled poplar (*Populus tremula* $\times$ *P. alba* cv. INRA 717-1B4) wood was obtained from regrowth of coppiced 6-year-old trees (Yang et al. 2019). Control poplar samples were assayed by Derivatization Followed by Reductive Cleavage (Lu and Ralph 1998) to contain 67% S-lignin; woody tissue from two lines of transgenic hybrid poplar with overexpression of the *FERULATE-5-HYDROXYLASE2* (*F5H*) gene under control of the *AtC4H* promoter represented high S-lignin (86 %), and woody tissue from two lines in which expression of *F5H* was attenuated by RNAi under control by an *AtC4H* promoter constituted the low S-lignin (52%) material (Yang et al. 2019). Woody material with exceptionally low S-lignin was obtained from trees in which expression of *CINNAMOYL O-METHYL TRANSFERASE* (*COMT*) was attenuated by RNAi under the *35S-CMV* promoter. COMT down-regulated plants are low in S-lignin (21%) but deposit atypical 5-OH-G lignin subunits. Some of the milled samples of hybrid poplar and each of these lignin variants were catalytically delignified using a Ni catalyst as described by Luo et al. (2016), washed with methanol and water, and freeze-dried.

Chemicals

Sodium hypochlorite solution, reagent grade, available chlorine 10-15% (Sigma-Aldrich Co.; Cat. No. **425044**); Sodium bromide, Sigma Life Science (Sigma-Aldrich, Cat. No. 71329-250G); 2,2,6,6-tetramethylpiperidine-1-oxyl free radical (TEMPO), CAS no. 15051-46-4 (Sigma-Aldrich; Cat No. 214000-1G); the cellulose solvent *N*-methylmorpholine-*N*-oxide (NMMO), (CAS no. 7529-22-8; Sigma-Aldrich; Cat. No. 67873-25G); Ionic liquid 1-butyl-3-methylimidazolium chloride (BMIM-Cl), (CAS no. 79917-90-1; Sigma-Aldrich Co. cat no. 77896); 1-cyclohexyl-3-(2-morpholinyl-4-ethyl) carbodiimide methyl-*p*-toluene sulfonate (CMC; Sigma-Aldrich; Cat. No. C106402); and sodium borodeuteride (CAS no. 15681-89-7; Sigma-Aldrich; Cat. No. 205591). Other chemicals and solvents were of laboratory grade from Sigma-Aldrich or Acros.

Cell Wall Isolation

Samples of Arabidopsis plantlets or maize coleoptiles were frozen in liquid nitrogen and pulverized. Cell walls from these powders were prepared by homogenization in 1% (w/v) SDS in 50 mM Tris-HCl, pH 7.2, at ambient temperature in a glass-glass motorized grinder (Kontes-Duall, Thomas Scientific). Wall

samples were warmed to 60°C and washed sequentially in additional 1% SDS, warm (50°C) water (3x), warm 50% (v/v) ethanol (3x), water (3x) at ambient temperature, with walls pelleted by centrifugation at $1200 \times g_{\text{max}}$ for 5 min after each step. Starch was extracted from the pellet by sonication in 90% (v/v) DMSO in water (Carpita and Kanabus, 1987). Cell walls were then washed four times with distilled water and then depectinated in hot (90°C) 5% (w/v) ammonium oxalate, followed by 0.1 M NaOH (supplemented with 3 mg mL⁻¹ NaBH₄) at ambient temperature as described in Okekeogbu et al. (2019); the pellets were acidified with glacial acetic acid and washed with several water washes. Milled poplar samples were suspended in water warmed to 50°C and washed sequentially in additional water (3x), 50% (v/v) ethanol (3x), and water (3x) at ambient temperature, and walls were pelleted after each step by centrifugation as described above.

TEMPO-oxidation

After initial experiments to determine the minimum alkalinity to drive oxidation, 10 to 50 mg of cellulose, polysaccharide or cell-wall material was suspended in 20 mL of 0.05 M of sodium phosphate [NaOH], pH 8.5, with 1.0 g of NaBr. Reactions were initiated with ~0.5 mg of TEMPO in 100 mL ethanol, and the suspensions were stirred continuously at ambient temperature. A solution of 15% sodium hypochlorite (5 mL) was added to initiate oxidation, and additional hypochlorite was added as needed to maintain pH between 8.5 and 9.0. The pH stabilized within 15 min but reactions were maintained for 2 h to ensure complete oxidation. Glacial acetic acid was then added to bring the pH to below 3.5. In some instances, the suspensions were dialyzed against running deionized water for 36 h, followed by exchanges with nanopure (18.2 MW) water, and then freeze-dried. Uronic acids were determined by a carbazole assay after reduction of neutral sugar interference with sulfamate (Fillisetti-Cozzi and Carpita 1991) and compared to total sugar determined by phenol-sulfuric assay (Dubois et al. 1956). From these sugar and uronic acid assays, yields of all samples were greater than 95%.

Carboxyl reduction

The TEMPO-generated carboxyl groups of cellulose and other uronosyl residues in the cell wall materials were activated with CMC and reduced with NaBD₄ to their respective 6,6-dideutero sugars according to Kim and Carpita (1992), as modified by Carpita and McCann (1997). Briefly, freeze-dried materials were suspended in about 2 mL of sodium acetate, pH 4.6, and 250 mg of CMC added with continuous stirring. The pH was monitored periodically, and, if necessary, additional sodium acetate, pH 4.6, was added periodically to maintain the pH below 5. After 2 h incubation, the reactions were chilled to ice temperature, and 300 mg of NaBD₄ added, with foaming reduced with drops of *n*-octanol. After 2 min, with 2 M imidazole [HCl] buffer, pH 7, to reduce alkalinity. After 1 h incubation, the reactions were acidified with glacial acetic acid and dialyzed against running deionized water for 48 h, followed by exchanges with nanopure water (18.2 MW) as described previously. From uronic acid assays, borodeuteride reduction was greater than 90%, and from phenol-sulfuric acid analysis, yields of all materials were greater than 95%.

Preparation of amorphous cellulose

Increases in amorphous cotton cellulose (Sigmacell) was generated by treatments of NaOH/urea, ionic liquids, or phosphoric acid, essentially as described (Kuo and Lee, 2009). For NaOH/urea, 100 mg of cellulose was mixed with a solution of 12% urea and 7% NaOH in water supplemented with 3 mg mL⁻¹ of NaBH₄ to prevent end peeling and stirred vigorously overnight. The mixture was chilled and transferred to a beaker using 10 volumes of 80% ethanol in water (v/v) and acidified with glacial acetic acid to below pH 4. The suspension was dialyzed for 48 h against running distilled water, then in nanopure water (18.2 MW) and freeze dried.

For samples treated with an ionic liquid and cellulose solvent, 100 mg of cellulose was mixed with 1.9 g of either 1-butyl-3-methylimidazolium (BMIM) chloride or *N*-methylmorpholine *N*-oxide (NMMO), respectively, and the mixtures were heated to 120°C with continuous stirring until liquefied and dissolved, and incubation continued at 120°C for 1 h. Mixtures were then chilled to ice temperature, and ten volumes of 80% ethanol in water (v/v) were added; the suspension was dialyzed as described above. For acid treatment, 100 mg of cellulose was mixed in 1.2 mL of 85% phosphoric acid in water (v/v) and stirred until dissolved. The mixture was incubated at 50°C with occasional stirring for 1 h. Mixtures were then chilled to ice temperature and neutralized with NaOH, and ten volumes of 80% ethanol in water (v/v) were added. The suspension was dialyzed as described above.

Generation of amorphous cellulose by TFA was performed according to Zhao et al. (2007) with modifications. Sigma cellulose (100 mg) was suspended in 2 mL of 99% TFA in a reaction tube, and the suspension was incubated at -20°C for 15 h, then heated at 55°C for 2.5 h. Ten volumes of absolute ethanol was added by vortex mixing, and the gelled suspension was stirred vigorously. The gelled cellulose was washed at ambient temperature four times with 80% ethanol in water (v/v), and four times in water.

Enzymatic conversion of amorphous cellulose to cellobiose units

Ammonium sulfate (3.2 M) suspensions of *Trichoderma longibrachiatum* cellulase (endo-1,4-b-D-glucanase EGII (EC 3.2.1.4; Megazyme, Ireland; Cat. No. E-CELTR; 1000U 0.8mL⁻¹) and microbial cellobiohydrolase II (EC 3.2.1.91; Megazyme, Ireland; Cat. No. E-CBHIIIM) were diluted 10-fold in 1 mM sodium acetate, pH 5.5, loaded on a 5-mL column of Sephadex G-25 (Pharmacia Biotech, Piscataway, NJ, USA), and eluted with additional 1 mM sodium acetate to exchange the ammonium sulfate. Water suspensions (0.5 mL in a 4-mL Teflon[®]-lined screw-cap vial) of the TFA-gelatinized amorphous celluloses were adjusted to 1 mM sodium acetate, pH 5.5. and 10 units each of the enzyme preparation added. The digestions were carried out to completion at 42°C for 24 h with gentle shaking. The reactions were stopped by addition of four volumes of ethanol and heating of the sealed vials to 105°C for 5 min. After cooling the suspensions were centrifuged at 1200 ´ g_{max} for 5 min to precipitate protein, and the clear supernatant liquid transferred to a fresh vial and dried under a stream of N₂ gas.

For digestion of the barley mixed-linkage glucan, a 3.2 M ammonium sulfate suspension of a lichenase preparation of *Bacillus subtilis* endo- α -D-glucanase (EC 3.2.1.73; Megazyme Ireland, Cat. No. E-LICHN, 1000 U mL⁻¹) was buffer-exchanged with 1 mM sodium acetate, pH 5.5, to remove the ammonium sulfate, and 10 mL of this preparation containing 10 units of the enzyme were added to 10 mg of the mixed-linkage glucan in 0.5 mL of water. Samples were incubated for 5 h at 42°C before the reactions were stopped by addition of four volumes of ethanol and heating of the sealed vials to 105°C for 5 min. After cooling the suspensions were centrifuged at 1200 $\times$ g_{max} for 5 min to precipitate protein, and the clear supernatant liquid transferred to a fresh vial and dried under a stream of N₂ gas.

Electrospray ionization MS

For electrospray ionization mass spectrometry, the dried products of enzymatic digestions were dissolved in 100 mL of 10 mM sodium acetate in 10% (v/v) methanol in water. Mass analysis was obtained in positive mode with an Agilent 6545 Q-TOF mass spectrometer with ESI capillary voltage of +3.5 kV, an N₂ temperature of 320°C, a drying gas flow rate of 8.0 mL min⁻¹, a nebulizer gas pressure of 35 psig, a fragmentor voltage of 135 V, a skimmer voltage of 65 V, and an octupole radio-frequency and voltage peak-to-peak (Vpp) of 750 V. Mass data (from m/z 80 to 1100) were collected using Agilent MassHunter Acquisition software (v. B.06). Mass spectral data analysis used Agilent MassHunter Qualitative Analysis (v. B.07) software. Values reported are the mean $\pm$ S.D. of three samples.

We determined empirically the ¹³C-spillover with cellobiose from native celluloses of cotton and poplar, which averaged 0.126. However, owing to the lower discrimination of ¹³C by C4 plants, the spillover in maize materials averaged 0.136. Thus, the algorithm determined the single and double spillover of the m/z 365 (undeuterated, anhydrous). The single spillover was subtracted from m/z 366, and the single and double spillover calculated from this remainder for m/z 367, and so forth. The single spillover values of m/z 366 and double spillover values from m/z 365 were subtracted from m/z 367 to get true m/z 367, and this process was iterated through the m/z 369. The m/z 365 was taken as anhydrous, the corrected m/z 366 and 367 summed to give the single and double deuteration (surface), and the corrected m/z 368 and 369 summed to give triple and quadruple deuteration (amorphous). An example of application of this algorithm is shown in Table S1.

Quantitation of cellulose and cell wall neutral monosaccharides

Samples of cell walls were hydrolyzed with 2 M TFA containing 400 nmol of *myo*-inositol (internal standard) for 90 min at 120°C in 1-mL conical Reacti-vials (Pierce Chemical). After hydrolysis, insoluble material was pelleted by centrifugation, and the supernatant TFA was collected and evaporated under a stream of filtered N₂. The insoluble cellulose was washed several times with water and collected by centrifugation.

Acid-soluble glucose determinations

To determine the extent of oxidation of surface chains and the depth of oxidation into the crystalline domains, TEMPO-catalyzed, NaBD₄-reduced cellulose materials (1-2 mg) were hydrolyzed in 1 mL of 2 M TFA containing 0.5 mmol of *myo*-inositol (internal standard) at 120°C for 90 min. The undigested crystalline cellulose was suspended in 0.8 mL of water, and 100 μ L was assayed for Glc equivalents using the phenol-sulfuric method (Dubois et al. 1956). The supernatant was evaporated at 40°C in a stream of air. The sugars were reduced with NaBH₄, and alditol acetates were prepared as described previously (Gibeaut and Carpita 1991). Derivatives were separated by gas-liquid chromatography on a 0.25-mm x 30-m column of SP-2330 (Supelco, Bellefonte, PA). After an initial loading hold at 100°C, the temperature was rapidly adjusted to 170°C at 25°C min⁻¹, then programmed to 240°C at 5°C min⁻¹ with 10-min hold at the upper temperature. Helium flow was 1 mL min⁻¹ with a splitless injection. The electron-impact mass spectrometry (EIMS) was with a Hewlett-Packard (Palo Alto, CA) MSD at 70 eV and a source temperature of 250°C. The proportion of 6,6-dideuteriogalactosyl was calculated using *m/z* 187/189, 217/219, and 289/291 according to the equation described by Kim and Carpita (1992), after correction of ¹³C spillover in undeuterated fragments as described above. Neutral sugar composition was verified by EIMS (Carpita and Shea 1989).

Results

Refinement of TEMPO-catalyzed oxidation and reduction

Saito and Isogai (2004) employed a titrimeter to hold the pH to 10.5 upon additions of hypochlorite over 24 h, but reported TEMPO-catalyzed oxidation of cellulose samples as complete in ~2 h. We tested the extent of TEMPO-catalyzed oxidation of cotton fiber cellulose held at pH between 6 and 10 with 0.1 M sodium phosphate and additions of hypochlorite. Preliminary experiments revealed oxidation of cellulose only above pH 8.0 (Fig. S1). When pH was held between pH 8.5 and 9.0, oxidation increased rapidly upon addition of hypochlorite with saturating oxidation after 2h (Fig. S2).

Cellulose oxidation was largely complete after 2 h, plateauing at less than 18% of the total cellulose. We then tested whether oxidation was truly complete for all solvent-exposed Glc residues or if oxidation might be restricted by oxidation of adjacent residues. By contrast, the degree of oxidation of a water-soluble barley mixed-linkage (1 $\rightarrow$ 3),(1 $\rightarrow$ 4)- β -D-glucan was greater than 90% based on yield by uronic acid assay (Fig. 2). As expected, because of the absence of available primary alcohols, 5-arabinans from rhamnogalacturonan-I pectin and hemicellulosic larch 4-xylans were resistant to oxidation. We then tested several protocols to solubilize cellulose and increase the proportion of amorphous chains, including NaOH/urea (Cai and Zhang 2006), 85% phosphoric acid, the cellulose solvent, *N*-methylmorpholine-*N*-oxide (NMMO) (Liu et al. 2009), the ionic liquid, 1-butyl-3-methylimidazolium chloride (BMIM-Cl) (Kuo and Lee 2009), and 99% TFA at -20°C (Zhao et al. 2007). We found that optimal oxidation was obtained by swelling the cellulose particles in TFA at -20°C followed by incubation at 55°C for 2.5 h (Fig. S3). This method generated a clear liquid that formed a translucent gel upon rapid injection of ethanol to 80% (v/v). This procedure resulted in large increases in the degree of oxidation, but was

limited to 53% of the total Glc residues because of annealing and precipitation of many of the (1 $\rightarrow$ 4)-b-d-glucan chains upon washing the gels with water and freeze-drying.

Analysis of cellobiose deuteration

To determine the relative proportions of glucan chains of cellulose in anhydrous, surface, and amorphous fractions, we used carbo-diimide activation of the uronosyl residues produced by TEMPO-catalyzed oxidation, followed by NaBD₄ reduction to produce 6,6-dideutero glucose residues. We digested these products with a mixture of cellulase and cellobiohydrolase to generate cellobiose residues, and determined by ESI-MS whether both residues, only one, or neither was reduced. To eliminate possible bias towards recovery of only the more soluble glucan chains, cellulose was gelatinized in TFA before cellulase/cellobiohydrolase digestion. Only primary alcohols of glucosyl residues of cellulose microfibrils in contact with water would have been susceptible to TEMPO-catalyzed oxidation (Fig. 1). All of the glucosyl residues of native amorphous or disordered (1 $\rightarrow$ 4)-b-d-glucan chains would be in contact with water, so these chains are converted completely to poly-(1 $\rightarrow$ 4)-b-d-glucuronan. Although the TEMPO-catalyzed oxidation should yield uronosyl residues, we noted that a portion of the products were aldehydes, and thus, were converted to their 6-deuterated ($M^+ + 1$) forms. In amorphous chains both glucosyl residues of cellobiose will be oxidized and converted to their 6-deuterated ($M^+ + 1$) or 6,6-dideuterated form ($M^+ + 2$), whereas the (1 $\rightarrow$ 4)-b-d-glucan chains of the anhydrous core crystalline domain will remain underivatized. However, for surface chains on the edge of a microfibril, only one of every two glucosyl residues is in contact with water, as the (1 $\rightarrow$ 4)-b-d-linkage turns each glucosyl residue nearly 180° with respect to each neighboring residue, such that for every glucosyl residue facing the solution, the primary alcohols of its neighboring residues are tucked within the crystalline core (Fig. 1). Upon oxidation of surface chains, a mixed polymer containing the repeating dimer, $\text{®4-b-d-glucuronosyl-(1®4)-b-d-glucosyl®}$, will be formed. Upon reduction, only one of the glucosyl residues of cellobiose from ordered surface chains would be 6-deuterated or 6,6-dideuterated (Fig. 1).

In electrospray ionization (ESI) MS, cellobiose adsorbs a $^{23}\text{Na}^+$ to form m/z 365 ($M + ^{23}\text{Na}^+$). Underivatized cotton cellulose or TEMPO-oxidized and then NaBH₄-reduced cellulose gave m/z 365 ($M + ^{23}\text{Na}^+$) additional +1 and +2 amu in ESI that were consistent with native ^{13}C -spillover (Fig. 3a). However, TEMPO-catalyzed cellulose oxidation of poplar samples, followed by diimide-activated reduction with NaBD₄, gave a complex series of cellobiose native m/z 365 (M^+) plus +1 through +5 amu consistent with the proportions of derivatization of one or both of the cellobiose glucosyl residues in excess of ^{13}C -spillover (Fig. 3b). The preliminary experiments had shown that oxidation yielded approximately equal amounts of aldehyde and carboxyl, for which doubling the amounts of ($M^+ + 1$) vs. the sum of ($M^+ + 3$) and ($M^+ + 4$) gave the relative amounts of singly and doubly deuterated glucosyl units, respectively. Digestion of the glucan with *B. subtilis* endoglucanase (lichenase), which cleaves the (1 $\rightarrow$ 4)-b-d-glucosyl residues adjacent to (1 $\rightarrow$ 3)-linkages (Anderson and Stone 1975), yields mostly cellobiosyl- and cellotriosyl-(1 $\rightarrow$ 3)-glucosyl oligomers that indicate almost complete oxidation and reduction to 6-deutero and 6,6-dideutero-glucose derivatives (Fig. 3c).

When NaBD₄-reduced cotton cellulose was hydrolyzed with 2 M TFA, which does not hydrolyze crystalline core domains, the TFA-soluble glucose recovered showed deuteration of 54.7% (Table S2), indicating oxidation of one-half of the surface chains and a small amount of amorphous cellulose. However, when TFA-solubilized cellulose was subjected to TEMPO-catalyzed oxidation and NaBD₄ reduction, considerably more amorphous chains were detected, representing about 40% of the total (Fig. 3d). Based on resistance to 2 M TFA hydrolysis after recovery of the TFA-soluble cellulose in water, extensive reannealing had occurred to give significant anhydrous and surface chains in a ratio of 2:1.

Based on analysis of the proportions of deuterated residues of cellobiose derivatives, cotton fiber cellulose yielded much higher proportions of anhydrous domains than would be predicted for microfibrils of 18 or 24 chains, or even 36 chains (Table 1; Fig. 3b). The anhydrous domain of cotton cellulose represented 86% of the glucose units in cellulose, with only 11% of the chains occupying an ordered surface domain and giving an anhydrous chain : surface chain ratio of 7.6. Only 2 to 3% of the chains were disordered or amorphous, where both glucosyl units were deuterated (Table 1). Iterating the 18- or 36-chain models of cellulose microfibrils to account for the higher proportion of anhydrous chains indicates an aggregation of well over 200 chains (Fig. 4). Cellulose chains in microfibrils from depectinated primary cell walls of *Arabidopsis* developing leaves and etiolated 3-day-old maize coleoptiles were 74% and 81% in anhydrous domains, with 17% and 14% in ordered surface chains, giving anhydrous chain : surface chain ratios of 4.3 and 5.9, respectively (Table 1). Depending on microfibril size and orientation, many more than 100 chains were bundled. By contrast to cotton cellulose, amorphous chains represented 10 and 12% of the cellulose in *Arabidopsis* and maize, respectively.

Lignin is considered a major source of recalcitrance in the deconstruction of cellulosic biomass by occluding cellulose to inhibit the action of cellulolytic enzymes. We determined the degrees of crystallinity of secondary wall cellulose in wood particles from transgenic fast-growing hybrid poplar with altered proportions of guaiacyl (G), 5-hydroxyguaiacyl (OH-G), and syringyl (S) monolignol subunits (Yang et al. 2019). The TEMPO-oxidation/NaBD₄-reduction reactions gave estimates of about 70 to 75% anhydrous domains depending on genotype, with surface chains comprising higher proportions than observed in cotton cellulose, giving anhydrous chain:surface chain ratios of 2.8 to 3.4 (Fig. S4, Table 2). To determine how lignin contributed to degree of crystallinity, we then evaluated the cellulose composition of these lignin variants after catalytic depolymerization of lignin (Yang et al. 2019). Crystallinity increased to 76 to 80% anhydrous as a result of delignification, regardless of the original lignin composition before removal, and increasing the anhydrous chain:surface chain ratios from 3.8 to 4.5 depending on genotype (Fig. S5, Table 2).

Discussion

The bundling of microfibrils into macrofibrils

The crystalline nature of cellulose is a key factor in recalcitrance to enzymatic digestion (Himmel et al. 2007). The degree of recalcitrance can be quantified by comparing the digestibility of native and

gelatinized cellulose. Over 80% of the low-temperature swollen cotton cellulose, and over 95% of the gelatinized cellulose, was digested by a standard cellulase Ctec2 cocktail in about 6 h at 37°C (Shiga et al. 2017). By contrast, only about 60% of control cotton fiber cellulose was digested to glucose using a 10-fold amount of Ctec2 after 3 days. Here, we employed TFA gelatinization after TEMPO-catalyzed oxidation and NaBD₄ reduction to completely digest the cellulose to cellobiose units with the combination of cellulase and cellobiohydrolase. In so doing, we could be confident that we were accounting for all of the anhydrous cellulose that would otherwise be resistant to enzymatic hydrolysis.

We generated cross-sectional models of cellulose microfibrils based on a rectangular packing arrangements of an 18-chain (6x3-chain) structure, as proposed by Fernandes et al. (2011), and a more compact arrangement (2x3-chain, 3x4-chain). We iterated additional layers vertically to make an 8x3 chain or laterally to make 6x4-chain configurations of 24-chain structures (Fig. 4). Surface chains are considered only those predicted to be solvent-exposed on the side; those on the top and bottom sandwiched by other chains are considered anhydrous because all *O*-6 glucosyl residues would be coordinated by their neighboring chains. Three of the four models based on 18 and 24 chains gave predicted anhydrous chain:surface chain ratios of 1:2, whereas the compact 18-chain model gave a 4:5 ratio. When two 18-chain microfibrils are bundled vertically, the ratio is maintained regardless of arrangement, but if bundled laterally, the ratio would switch to 2:1 in favor of anhydrous domains. A bundle of four 18-chain microfibrils retains the 2:1 ratio, but lateral extension of six 18-chain bundles increases anhydrous chains only to give ratios of 7:2, and eight 18-chain bundles give ratios of 5:1 (Fig. 4). A more compact arrangement of the 18-chain structure gives more complicated ratios based on aggregation into hexagonal symmetries, where four microfibrils give a 13:5, seven give a 7:2, and thirteen give approximately 3:1.

Extension of the 24-chain vertical model would maintain the 1:2 ratio; bundling two 24-chains laterally changes the ratio to 2:1, whereas lateral bundling of the 24-chain of the compact 6x4 chain model yields a 3:1 ratio. Iteration of the of 8x3-chain configurations into larger bundles produced ratios identical to the 6x3-chain configurations, but iteration of the 6x4-chain configurations resulted in larger increases in anhydrous domains, such that a 5:1 ratio was achieved in the six iterations (Fig. 4). Our data are consistent with macrofibril sizes within the range of the 20 to 60 nm diameter fibrils described by Donaldson (2007), and we propose that, to be consistent with our experimental ratios of about 4:1, would result from bundling of about five to six vertical 18- or 24-chain microfibrils, whereas this value could not be reached with even thirteen compact 18-chain microfibrils.

Microfibrils bundle into structures of 100-200 chains whose interior chains are unavailable for TEMPO-catalyzed oxidation and subsequent reduction. One cannot infer crystalline continuity during bundling of microfibrils. Although we term all undeuterated cellulose as 'anhydrous', the tight packing of microfibrils might produce junctions that are not strictly anhydrous. Solid-state NMR studies distinguish surface and anhydrous chains to suggest microfibril sizes of 18-24 chains, even in materials known to contain large bundles of microfibrils. Jeffries (1963) noted in deuterium-exchange experiments with cellulose, that three domains were evident: a domain of rapid exchange, representing about 20% of the mass, a second

domain of much slower exchange, and a bulk of the remainder that excluded deuterium even after several days of incubation. These data are consistent with those of Lindh and Salmén (2017), who describe how bundling of crystalline microfibrils but lack of crystalline continuity might create the domains of slow penetration of deuterated water distinct from the crystalline domains that exclude water completely. Our results are consistent with this conclusion, where we find that rapid oxidation of cotton cellulose represents 14 mole % surface and amorphous domains, those of Arabidopsis and maize primary wall present 26 and 19 mole %, respectively, and poplar secondary wall variants represents 28 to 33 mole % of the total mass (Tables 1 and 2). By oxidation of only the O-6 of the solvent-exposed glucosyl residues, we might avoid counting of deuteration of hydroxyls deeper into the crystalline domains of the microfibrils.

Although the crystallinity of cellulose microfibrils is considered a recalcitrance factor, physical behavior during conversion is complicated further by the degree that microfibrils bundle into larger aggregates, sometimes called macrofibrils (Ding et al. 2014), exhibiting even higher degrees of crystallinity. However, washing the gelatinized cellulose in water permits partial annealing to anhydrous states. Gelatinization of cellulose in TFA greatly increased rates of enzymatic digestion and degree of conversion to biofuel substrates, levulinic acid and furfural (Shiga et al. 2017). The proportion of anhydrous cellulose was greatly reduced but not eliminated (Table 1). About one-half of the cellulose was determined to be amorphous, but the proportion of the anhydrous chains decreased from 7.6:1 in native cellulose to 2:1 in TFA-gelatinized cellulose, consistent with models containing 36 to 96 chains depending on configuration (Fig. 4), a size that might render them susceptible to complete enzymatic digestion (Shiga et al. 2017).

Lignin and cellulose crystallinity

Lignin has been considered the major recalcitrance factor in biological or chemical conversion of biomass to biofuels or bio-based products (Wyman et al. 1999; Himmel et al. 2007). The prevailing hypothesis is that lignin interacts tightly with cellulose to block access of cellulases to its substrate (Chen and Dixon 2007). For this reason, a considerable effort has been devoted to lowering lignin content or altering composition to increase saccharification yield (Chanoca et al. 2019). Downregulation of several of the early genes of the monolignol synthesis pathway resulted in increased rates of saccharification, but often at the expense of biomass accumulation (Ralph et al. 2019). To improve saccharification yield without compromising growth and biomass quality, alternative strategies took advantage of the plasticity of lignin synthesis to retain normal lignin synthesis but adjust the pool of available monolignols. Overexpression of the *FERULATE 5-HYDROXYLASE (F5H)* gene in poplar resulted in more than 90% S-lignin content, which increased sugar yields from saccharification without growth penalty (Stewart et al. 2009). The conclusion was that conversion to mostly S-lignin resulted in shorter lignin chains with less branching, resulting in less occlusion of the cellulose and, consequently, more facile enzymatic digestion.

Several lines of evidence challenge the hypothesis that recalcitrance is based on the molecular association of lignin and cellulose. During steam-expansion pretreatments, maize stem lignin is melted and redistributed into small globules independent of the S-lignin composition (Donohoe et al. 2008). Variants with average lignin content exhibited high saccharification yields in poplar, indicating that

factors other than S-lignin content contributed to recalcitrance (Studer et al. 2011). In a maize recombinant-inbred population saccharification yield and lignin composition or abundance were uncorrelated, and QTL analysis determined that different genes contributed to these traits (Penning et al. 2014). Contrary to intuition, the lignified biomass from milled poplar wood had the least proportion of anhydrous glucan chains in cellulose, regardless of genotype (Table 2). Catalytic delignification enhanced saccharification yield, but despite removal of most of the lignin in all transgenic materials, high-S lignin materials showed enhanced cellulose digestibility (Yang et al. 2019). This enhanced saccharification yield occurs despite an increase in the proportion of anhydrous domains and, thus, apparent bundling (Table 2).

Determinants of recalcitrance comprise a combination of elements of molecular interactions within the cell wall, nanoscale composition of the cell wall, macroscale variations in cell-cell adhesion (McCann and Carpita 2015). The enhancement of saccharification potential observed after catalytic delignification is more likely a result of alterations of nanoscale and macroscale properties than in the molecular scale interactions of lignin and cellulose. As high S-lignin content results of more facile cell separation and comminution of poplar wood particles (Yang et al. 2020), enhanced saccharification in the CDL-treated materials might reflect enhanced access of cellulases into the biomass despite the increased cellulose bundling after delignification. Cellulose bundling is still a major factor of recalcitrance to digestibility. The question we ask now is if there is a genetic basis of bundling in the nanoscale architecture amenable to molecular design.

Declarations

Funding:

We thank our colleagues in the Center for Direct Catalytic Conversion of Biomass to Biofuels (C3Bio) for a decade of insights, interdisciplinary engagement and friendship. This work was supported by C3Bio, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award # DE-SC0000997. Additional support came from the from Coordenação para o Aperfeiçoamento de Pessoal de Nível Superior (CAPES - Process 10734-13-9) and Purdue University's International Programs in Agriculture (IPIA) in the College of Agriculture.

Conflicts of interest/Competing interests

Authors declare no conflicts or competing interests

Availability of data and material

Not applicable

Code availability

Not applicable

Authors' contributions: S.M.S., H.Y., and N.C.C. designed research; S.M.S., H.Y., A.T.O., and N.C.C. performed experiments; all authors analyzed data and contributed to the writing of the article.

References

Anderson MA, Stone BA (1975) A new substrate for investigating the specificity of b-glucan hydrolases. *FEBS Lett.* 52:202–207.

Cai J, Zhang L (2006) Unique gelation behavior of cellulose in NaOH/urea aqueous solution. *Biomacromolecules* 7:183–189

Carpita NC, Kanabus J (1987) Extraction of starch by dimethyl sulfoxide and quantitation by enzymatic assay. *Anal Biochem* 161:132–139

Carpita NC, Shea EM (1989) Linkage structure by gas chromatography-mass spectrometry of partially-methylated alditol acetates. In Biermann CJ, McGinnis GD, eds., *Analysis of Carbohydrates by GLC and MS*. CRC Press, Boca Raton, FL, pp. 155–216

Carpita NC, McCann MC (1997) Some new methods to study plant polyuronic acids and their esters. In: Townsend R, Hotchkiss A, eds. *Progress in Glycobiology*. New York: Marcell Dekker, pp. 595–611

Chanoca A, de Vries L, Boerjan W (2019) Lignin engineering in forest trees. *Front Plant Sci* 10:912

Chen F, Dixon RA (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nature Biotech* 25:759–761

de Nooy AEJ, Besemer AC, van Bekkum H (1995) Highly selective nitroxyl radical-mediated oxidation of primary alcohol groups in water-soluble glucans. *Carbohydr Res* 269:89–98

Ding S-Y, Himmel ME. (2006) The maize primary cell wall microfibril: A new model derived from direct visualization. *J Agric Food Chem* 54:597–606

Ding S-Y, Liu Y-S, Zeng Y, Himmel ME, Baker JO, Bayer EA (2012) How does plant cell wall nanoscale architecture correlated with enzymatic digestibility? *Science* 338:1055–1060

Ding S-Y, Zhao S, Zeng YN (2014) Size, shape, and arrangement of native cellulose fibrils in maize cell walls. *Cellulose* 21:863–871

Donaldson L (2007) Cellulose microfibril aggregates and their size variation with cell wall type. *Wood Sci Technol* 41: 443–460

Donohoe BS, Decker SR, Tucker MP, Himmel ME, Vinzant TB (2008) Visualizing lignin coalescence and migration through maize cell walls following thermochemical pretreatment. *Biotechnol Bioeng* 101:913–925

- Dubois M, Gilles KA, Hamilton JK, Rebers PA, Smith F (1956) Colorimetric method for determination of sugars and related substances. *Anal Chem* 28:350–356
- Fernandes AN, Thomas LH, Altaner CM, Callow P, Forsyth VT, Apperley DC, Kennedy CJ, Jarvis MC (2011) Nanostructure of cellulose microfibrils in spruce wood. *Proc Natl Acad Sci USA* 108:E1195–E1203
- Filisetti-Cozzi TMCC, Carpita NC (1991) Measurement of uronic acids without interference from neutral sugars. *Anal Biochem* 197:157-162
- Frilette VJ, Hanle J, Mark H (1948) Rate of exchange of cellulose with heavy water. *J Am Chem Soc* 70:1107–1113
- Gibeaut DM, Carpita NC (1991) Tracing cell wall biogenesis in intact cells and plants: Selective turnover and alteration of soluble and cell wall polysaccharides in grasses. *Plant Physiol* 97:551–561
- Ha M-A, Apperley DC, Evans BW, Huxham M, Jardine WG, Viëtor RJ, Reis D, Vian B, Jarvis MC (1998) Fine structure in cellulose microfibrils: NMR evidence from onion and quince. *Plant J* 16:183–190
- Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD (2007) Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* 315:804–807
- Jarvis ML (2013) Cellulose biosynthesis: Counting the chains. *Plant Physiol* 163:1485–1486
- Jeffries R (1963) An infra-red study of the deuteration of cellulose and cellulose derivatives. *Polymer* 4:375–289
- Kennedy CJ, Cameron GJ, Šturcová A, Apperley DC, Altaner C, Wess TJ, Jarvis MC (2007a) Microfibril diameter in celery collenchyma cellulose: X-ray scattering and NMR evidence. *Cellulose* 14:235–246
- Kennedy CJ, Šturcová A, Jarvis MC, Wess TJ (2007b) Hydration effects on spacing of primary-wall cellulose microfibrils: a small angle X-ray scattering study. *Cellulose* 14:401–408
- Kim J-B, Carpita NC (1992) Changes in esterification of the uronic acid groups of cell wall polysaccharides during elongation of maize coleoptiles. *Plant Physiol* 98:646–653
- Kuo CH, Lee K (2009) Enhancement of enzymatic saccharification of cellulose by cellulose dissolution pretreatments. *Carbohydr Polym* 77:41–46
- Lindh EL, Salmén L (2017) Surface accessibility of cellulose fibrils studied by hydrogen-deuterium exchange with water. *Cellulose* 24:21-33
- Liu CF, Zhang AP, Li WY, Sun RC (2009) Dissolution of cellulose in novel green solvent ionic liquids and its application. *Prog Chem* 21:1800–1806

- Liu J, Inouye H, Venugopalan N, Fischetti RF, Gleber SC, Vogt S, Cusumano JC, Kim JI, Chapple C, Makowski L (2013) Tissue specific specialization of the nanoscale architecture of Arabidopsis. *J Struct Biol* 184:103–114
- Lu F, Ralph J. (1998) The DFRC method for lignin analysis: 2. Monomers from isolated lignins. *J Agric Food Chem* 46:547–552
- Luo H, Klein IM, Jiang Y, Zhu H, Liu B, Kenttämä H, Abu-Omar MM (2016) Total utilization of Miscanthus biomass, lignin and carbohydrates, using earth abundant nickel catalyst. *ACS Sust Chem Eng* 4:2316–2322
- McCann MC, Carpita NC (2015) Biomass recalcitrance: A multi-scale, multi-factor and conversion-specific property. *J Exp Bot* 66:4109–4118
- Meinert MC, Delmer DP (1977) Changes in biochemical composition of cell-wall of cotton fiber during development. *Plant Physiol* 59:1088-1097
- Newman RH, Hill SJ, Harris PJ (2013) Wide-angle X-ray scattering and solid-state nuclear magnetic resonance data combined to test models for cellulose microfibrils in mung bean cell walls. *Plant Physiol* 163:1558–1567
- Okekeogbu IO, Pattathil S, Fernández-Niño SMG, Penning BW, Lao J, Heazlewood JL, Hahn MG, McCann MC, Carpita NC (2019) Glycome and proteomic components of Golgi membranes are common between two angiosperms with distinct cell wall structures. *Plant Cell* 31:1094-1112
- Okita Y, Saito T, Isogai A (2010) Entire surface oxidation of various cellulose microfibrils by TEMPO-mediate oxidation. *Biomacromolecules* 11:1696–1700
- Penning BW, Sykes RW, Babcock NC, Dugard CK, Held MA, Klimek JF, Shreve J, Fowler M, Gamblin D, Ziebell A, Davis M, Decker SR, Filley TR, Mosier NS, Springer NM, Thimmapuram J, Weil CF, McCann MC, Carpita NC (2014) Genetic determinants for enzymatic digestion of lignocellulosic biomass are independent of those for lignin abundance in a maize recombinant inbred population. *Plant Physiol* 165:1475–1487
- Ralph J, Lapierre C, Boerjan W (2019) Lignin structure and its engineering. *Curr Opin Biotechnol* 56:240–249
- Saito T, Isogai A (2004) TEMPO-mediated oxidation of native cellulose. The effect of oxidation conditions on chemical and crystal structures of the water-insoluble fractions. *Biomacromolecules* 5:1983–1989
- Shiga TM, Xiao W, Yang H, Zhang X, Olek AT, Donohoe BS, Makowski L, Hou T, Zhang Y, Zhao T, McCann MC, Carpita NC, Mosier NS (2017) Enhanced rates of enzymatic saccharification and catalytic synthesis of biofuel substrates in gelatinized cellulose generated by trifluoroacetic acid. *Biotechnol Biofuels* 10:310

Stewart JJ, Akiyama T, Chapple C, Ralph J, Mansfield SD (2009) The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiol* 150:621–635

Studer MH, DeMartini JD, Davis MF, Sykes RW, Davison B, Keller M, Tuskan GA, Wyman CE 2011. Lignin content in natural *Populus* variants affects sugar release. *Proc Natl Acad Sci USA* 108:6300–6305

Takaichi S, Saito T, Tanaka R, Isogai A (2014) Improvement of nanodispersibility of oven-dried TEMPO-oxidized celluloses in water. *Cellulose* 21:4093–4103

Thomas LH, Forsyth VT, Sturcova A, Kennedy CJ, May RP, Altaner CM, Apperley DC, Wess TJ, Jarvis MC (2013) *Structure of cellulose microfibrils in primary cell-walls from collenchyma*. *Plant Physiol* 161:465–476

Thomas LH, Forsyth VT, Martel A, Grillo I, Altaner CM, Jarvis MC (2014) Structure and spacing of cellulose microfibrils in woody cell walls of dicots. *Cellulose* 21:3887–3895

Viëtor RJ, Newman RH, Ha, MA, Apperley DC, Jarvis MC (2002) Conformational features of crystal-surface cellulose from higher plants. *The Plant Journal* 30:721–731

Wyman CE (1999) Biomass ethanol: technical progress, opportunities, and commercial challenges. *Annu Rev Energ Environ* 24:189–226

Yang H, Zhang X, Luo H, Liu B, Shiga TM, Li X, Kim JI, Rubinelli P, Overton JC, Subramanyam V, Abu-Omar M, Chapple C, Donohoe BS, Cooper BR, Mo H, Makowski L, Mosier NS, McCann MC, Meilan R, Carpita NC (2019) Overcoming cellulose recalcitrance in woody biomass for the lignin-first biorefinery. *Biotechnol Biofuels* 12:171

Yang H, Benatti MR, Karve RA, Fox A, Meilan R, Carpita NC, McCann MC (2020) Rhamnogalacturonan I is a determinant of cell-cell adhesion in poplar wood. *Plant Biotechnol* 18:1027-1040

Zhao H, Kwak JH, Zhang ZC, Brown HM, Arey BW, Holladay JE (2007) Studying cellulose fiber structure by SE, XRD, NMR, and acid hydrolysis. *Carbohydr Polym* 68:235–241

Tables

Table 1 Proportions of amorphous, surface and crystalline core chains from primary and secondary wall celluloses

Source	Anhydrous	Surface	Amorphous	Anhydrous: Surface
<i>mole%</i>				

Cotton cellulose	85.8±2.2	11.3±1.1	2.5±1.0	7.5
+TFA -20°C	69.9±1.0	19.7±0.1	11.1±1.0	3.5
+TFA -20°C, 55°C	48.4±0.4	23.7±0.8	27.9±1.2	2.0
Arabidopsis	74.0±2.7	17.4±2.2	8.6±0.6	4.3
Maize	80.7±3.9	13.6±3.9	5.7±0.1	5.9

Molar proportions were determined by the ratio of undeuterated cellobiose (anhydrous) (*m/z* 365), deuteration of one of the glucosyl residues (surface) (*m/z* 366 and *m/z* 367), and deuteration of both glucosyl residues (amorphous) (*m/z* 368 and *m/z* 369) by ESI-MS of products of cellulase and cellobiohydrolase hydrolysis. Abundance values from ESI-MS were corrected for ¹³C-spillover as described in Table S1. Values are the mean ± variance of two samples

Table 2 Proportions of amorphous, surface and crystalline core chains in cellulose from wood particles milled from hybrid poplar wood and its lignin variants before and after catalytic lignification (CDL-treatment)

	High-S	wild-type	High-G	High OH-G
<i>mole%</i>				
Control				
Anhydrous	69.5±2.9	74.8±2.6	70.5±1.2	69.8±3.0
Surface	25.5±1.7	22.1±1.8	25.2±0.3	25.2±2.0
Amorphous	5.0±1.2	3.1±0.7	4.3±0.9	5.0±1.0
Anhydrous:	2.7	3.4	2.8	2.8

surface

CDL-treated

Anhydrous	76.7±0.2	79.8±1.7	76.4±0.9	77.6±1.4
Surface	18.8±0.1	17.9±1.1	20.0±1.0	19.4±1.7
Amorphous	4.5±0.3	2.3±0.6	3.6±0.1	3.0±0.2
Anhydrous: surface	4.1	4.5	3.8	4.0

Molar proportions were determined by the ratio of undeuterated cellobiose (anhydrous) (m/z 365), deuteration of one of the glucosyl residues (surface) (m/z 366 and m/z 367), and deuteration of both glucosyl residues (amorphous) (m/z 368 and m/z 369) by ESI-MS of products of cellulase and cellobiohydrolase hydrolysis. Abundance values from ESI-MS were corrected for ^{13}C -spillover as described in Table S1. Values are the mean $\pm$ variance of two samples

Figures

Figure 1

Reaction scheme for surface chains of cellulose. Only solvent-exposed glucosyl residues of cellulose are accessible to TEMPO-catalyzed oxidation (red circles). All soluble amorphous cellulose residues (not shown) would be oxidized, and subsequently carbodiimide-activated carboxyl residues can be reduced by NaBD4 to 6,6-dideuteroglucosyl residues. By contrast, none of the residues within the cellulose crystalline matrix are solvent-exposed. Uniquely, alternate residues at the surface of the crystalline array are exposed, whereas the neighbors to either side are not (blue circles). Complete digestion of derivatized cellulose yields cellobiose units that possess deuteration of both glucosyl residues (amorphous cellulose), no deuteration (anhydrous cellulose), or only deuteration of one of the glucosyl residues (surface cellulose)

Figure 2

Uronic acid yields of the glucosyl residues of mixed-linkage (1→3),(1→4)-D-glucan, (1→4)-D-xylans and (1→5)-L-arabinans compared to those of cotton cellulose. Moles of uronic acid were determined by a method that reduces interference of neutral sugars (Filisetti-Cozzi and Carpita 1991), and compared to total moles of sugar determined by a phenol-sulfuric acid assay before reactions (Dubois et al. 1956)

Figure 3

Cellobiohydrolase digestions of TEMPO-oxidized and NaBD₄-reduced cellulose generates deuterated cellobiose representative of amorphous, surface and crystalline core chains. Electrospray ionization of partly deuterated cellulose resulted in cellobiose residues with M+1 and M+2 (m/z 366 and m/z 367), indicating one glucose residue, and M+3 and M+4 (m/z 368 and m/z 369 fragments, indicating both glucosyl residues of the disaccharide were oxidized and subsequently reduced. Undeuterated cellobiose arises only from anhydrous domains unexposed to solvent (m/z 365). Inset are the normalized m/z

abundances relative to m/z 365. A. Cellobiose from undeuterated cotton cellulose (Sigmacell) after digestion with cellulase and cellobiohydrolase. B. Cellobiose from cellulose in wood particles from a high S-lignin transgenic hybrid poplar cellulose. C. TEMPO-oxidized and NaBD4-reduced barley mixed-linkage (1→3),(1→4)-D-glucan yields, upon digestion with lichenase, a cellobiosyl-(1→3)-D-glucose trimer with all three glucosyl residues oxidized and reduced (m/z 533) and C-1 reduced (m/z 534). Inset: Trimer digestion product of untreated D-glucan (m/z 527). D. TEMPO-oxidized and NaBD4-reduced cotton cellulose after solubilization with TFA

Figure 4

Hypothetical anhydrous chain:surface chain ratios with bundles aggregated from 18- and 24-chain cellulose microfibrils and the number of glucan chains per bundle depending on configuration

This is a list of supplementary files associated with this preprint. Click to download.

- [Shigaetal.Supplementalmaterial.pdf](#)