Regular Motifs in Xylan Modulate Molecular Flexibility and Interactions with Cellulose Surfaces[1][CC-BY]

Antonio Martínez-Abad,[a,2] Jennie Berglund,[b,2] Guillermo Toriz,[c,d] Paul Gatenholm,[c] Gunnar Henriksson,[b] Mikael Lindström,[b] Jakob Wohlert,[b,3] and Francisco Vilaplana[a,b,3]

[a]Division of Glycoscience, School of Biotechnology, KTH Royal Institute of Technology, AlbaNova University Centre, 106 91 Stockholm, Sweden
[b]Wallenberg Wood Science Centre, Department of Fiber and Polymer Technology, School of Chemical Engineering, KTH Royal Institute of Technology, 100 44 Stockholm, Sweden
[c]Wallenberg Wood Science Centre, Department of Chemistry and Chemical Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden
[d]Wood Cellulose and Paper Research Department, University of Guadalajara, 44100 Guadalajara, Mexico

ORCID IDs: 0000-0003-0277-2269 (J.B.); 0000-0001-6732-2571 (J.W.); 0000-0003-3572-7798 (F.V.).

Xylan is tightly associated with cellulose and lignin in secondary plant cell walls, contributing to its rigidity and structural integrity in vascular plants. However, the molecular features and the nanoscale forces that control the interactions among cellulose microfibrils, hemicelluloses, and lignin are still not well understood. Here, we combine comprehensive mass spectrometric glycan sequencing and molecular dynamics simulations to elucidate the substitution pattern in softwood xylans and to investigate the effect of distinct intramolecular motifs on xylan conformation and on the interaction with cellulose surfaces in Norway spruce (*Picea abies*). We confirm the presence of motifs with evenly spaced glycosyl decorations on the xylan backbone, together with minor motifs with consecutive glucuronation. These domains are differently enriched in xylan fractions extracted by alkali and subcritical water, which indicates their preferential positioning in the secondary plant cell wall ultrastructure. The flexibility of the 3-fold screw conformation of xylan in solution is enhanced by the presence of arabinofuranosyl decorations. Additionally, molecular dynamic simulations suggest that the glycosyl substitutions in xylan are not only sterically tolerated by the cellulose surfaces but that they increase the affinity for cellulose and favor the stabilization of the 2-fold screw conformation. This effect is more significant for the hydrophobic surface compared with the hydrophilic ones, which demonstrates the importance of nonpolar driving forces on the structural integrity of secondary plant cell walls. These novel molecular insights contribute to an improved understanding of the supramolecular architecture of plant secondary cell walls and have fundamental implications for overcoming lignocellulose recalcitrance and for the design of advanced wood-based materials.

Secondary plant cell walls are formed after primary cell wall deposition once the cells have ceased to expand, resulting in woody tissues with remarkable material properties. While primary cell walls have a plastic, dynamic, and highly hydrated architecture enabling cell growth, secondary cell walls consist of a less hydrated, dense and thick network of biomacromolecules with fully differentiated structure in multiple lamellae, providing mechanical strength and rigidity to vascular plants. Secondary cell walls are composed of oriented cellulose microfibrils that are embedded in a tightly connected matrix of hemicelluloses and polyphenolic lignins, with well-defined ordering from the nanoscale to the macroscale (Cosgrove and Jarvis, 2012; Burgert and Keplinger, 2013). The cellulose microfibrils in secondary cell walls can associate in bundles that limit the accessibility of the original microfibril surfaces and exhibit increased disorder levels toward the surface, which may indicate a certain microfibril twisting (Fernandes et al., 2011; Cosgrove and Jarvis, 2012). The hemicelluloses are distributed between the cellulose microfibrils and also are connected to lignin, but the nature of these associations is still not well understood. Supramolecular interactions between cellulose microfibrils, hemicelluloses, and lignins in secondary plant cell walls are fundamental for its biomechanical integrity and functionality in different environments (Whitney et al., 1999, 2006; Åkerholm and Salmen, 2001; Ryden et al., 2003; Salmén,

[1] The research was supported by the Swedish Research Council (project no. 621-2014-5295) and the Knut and Alice Wallenberg Foundation (Wallenberg Wood Science Centre).

[2] These authors contributed equally to the article.

[3] Address correspondence to jacke@kth.se or franvila@kth.se.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Francisco Vilaplana (franvila@kth.se).

A.M.-A. and F.V. conceived and performed the mass spectrometric sequencing of xylan and made structural figures; J.B., J.W., F.V., G.H., and M.L. conceived the molecular dynamics experiments, while J.B. and J.W. performed them; G.T., P.G., A.M.-A., and F.V. extracted and purified the different xylan fractions and contributed to fundamental compositional analysis; A.M.-A., J.B., J.W., and F.V. analyzed the experimental data and wrote the article, with critical contributions from all authors.

www.plantphysiol.org/cgi/doi/10.1104/pp.17.01184

2004; Cosgrove and Jarvis, 2012; Burgert and Keplinger, 2013). However, the molecular features that modulate the interactions among cell wall biopolymers, their 3D macromolecular conformations in planta, and the nature of the nanoscale forces that control cell wall integrity are still controversial (Silveira et al., 2013; Cosgrove, 2014; Mikkelsen et al., 2015). Spatial and temporal heterogeneity in the composition and molecular structure of plant cell wall polysaccharides may be directed by biosynthetic processes in order to modulate their molecular diversity, adaptability, and biological function (Burton et al., 2010); however, the biological causes and effects of these fine structural variations are unknown. We are starting to understand how the molecular structure of hemicelluloses regulates the association with cellulose microfibrils (Bromley et al., 2013; Busse-Wicher et al., 2014, 2016a, 2016b) and the occurrence of covalent linkages with lignins in lignin-carbohydrate complexes (Lawoko et al., 2005). The presence of intermolecular forces and the interconnected supramolecular architecture in lignocellulosic biomass are responsible for its recalcitrance against conversion into valuable biofuels and platform chemicals (Mortimer et al., 2010; Ding et al., 2012; Silveira et al., 2013).

Xylans represent the main hemicellulose component in the secondary plant cell walls of flowering plants (hardwoods), and together with galactoglucomannans, they are present as well in the secondary cell walls of conifers (also referred to as softwoods). In conifers, a close association between glucomannan and cellulose has been reported, whereas for xylan that association is less clear (Åkerholm and Salmen, 2001; Salmén, 2004). The general molecular structure of xylans consists of a backbone of β-(1→4) D-xylopyranose (Xyl) units, which can be decorated by glycosyl substitutions and chemically modified (by acetylation) depending on the plant type, tissue, and developmental stage. In particular, conifer xylans are of the arabinoglucuronoxylan (AGX) type, with α-(1→2) D-GlcA in its 4-*O*-methylated form (mGlcA) and α-(1→3) L-arabinofuranose (Ara) as the backbone main decorations and with no reported acetylations (Fig. 1; Escalante et al., 2012; Busse-Wicher et al., 2016b; McKee et al., 2016). A regular distribution of glucuronic acids was reported previously in the xylans extracted from softwoods, whereas the pattern in hardwoods seemed to be irregular (Jacobs et al., 2001). Recent studies, however, have shown that glycosidic and acetyl decorations are preferably evenly spaced in specific domains of the xylan backbone (Bromley et al., 2013; Busse-Wicher et al., 2014; Chong et al., 2014), which could sterically allow favorable interactions with the hydrophilic surfaces of cellulose microfibrils (Busse-Wicher et al., 2016a, 2016b). Moreover, backbone substitution influences the solubility of hemicelluloses and their macroscopic properties (Höije et al., 2008; Sternemalm et al., 2008; Pitkänen et al., 2009; Escalante et al., 2012; Bosmans et al., 2014; Littunen et al., 2015).

In this study, we investigate the substitution pattern of softwood xylans and the effect of distinct intramolecular motifs on the conformation in solution and on the interaction with cellulose surfaces, combining comprehensive mass spectrometry (MS)-based glycan sequencing and in silico molecular dynamic (MD) simulations. The presence of regular motifs with both evenly spaced and consecutive glycosyl decorations on the backbone is confirmed for xylan extracted from Norway spruce (*Picea abies*) secondary cell walls. MD simulations suggest that these substitution motifs in xylan modulate the affinity for cellulose hydrophilic and hydrophobic surfaces and stabilize the conformational transition of xylan from a 3-fold screw in solution to a 2-fold screw onto the cellulose microfibrils. This offers new molecular insights on the structural integrity and the supramolecular architecture of plant secondary cell walls, with fundamental implications to overcome plant biomass recalcitrance and an optimized utilization of wood-derived materials.

RESULTS

Enzymatic Profiling Reveals Regular Ara and mGlcA Substitution Patterns in Spruce Xylan

Two different xylan fractions were extracted from spruce wood using an alkaline process (AGX-A) and a hydrothermal process by subcritical water (AGX-H), as presented in Supplemental Figure S1. The carbohydrate composition of the two purified xylans is presented in Supplemental Table S1. AGX-A shows lower relative glucuronation and higher arabinosylation compared with AGX-H. These differences in substitution may arise from the different extractability of distinct xylan populations in softwoods or may be induced by the extraction process. Detailed fingerprinting of the substitution pattern of AGX was performed using xylanolytic enzymes with known substrate recognition toward the presence of decorated glycosyl units (Fig. 1; Supplemental Fig. S2; Pell et al., 2004; Vardakou et al., 2008; Pollet et al., 2010; St John et al., 2011). End-point incubation of AGX with a GH10 β-xylanase releases xylobiose (X_2) and Xyl as main products, together with two main limiting decorated oligosaccharides (P_3U^m; P_3 and P_4). Sequencing of these decorated oligosaccharides by liquid chromatography (LC) coupled with electrospray ionization (ESI)-tandem mass spectrometry (MS/MS) confirms that these structures correspond with U^mXX and A^3X with the substituted sugar unit in the nonreducing end and with XA^3X (Supplemental Fig. S3), in agreement with the cleavage mechanism of GH10 β-xylanases (Pell et al., 2004; Pollet et al., 2010). Hydrolysis by a GH11 β-xylanase provides X_2 and X_3 as the main hydrolytic products, together with XU^mXX and XA^3X as the main limiting decorated oligosaccharides (Supplemental Fig. S3; Vardakou et al., 2008; Pollet et al., 2010). Minor amounts of longer oligosaccharides can be observed in the digestion products of GH10 and GH11. This confirms the combined presence of domains in AGX with loose distribution of substitutions that favor enzymatic saccharification together with clustered substitution

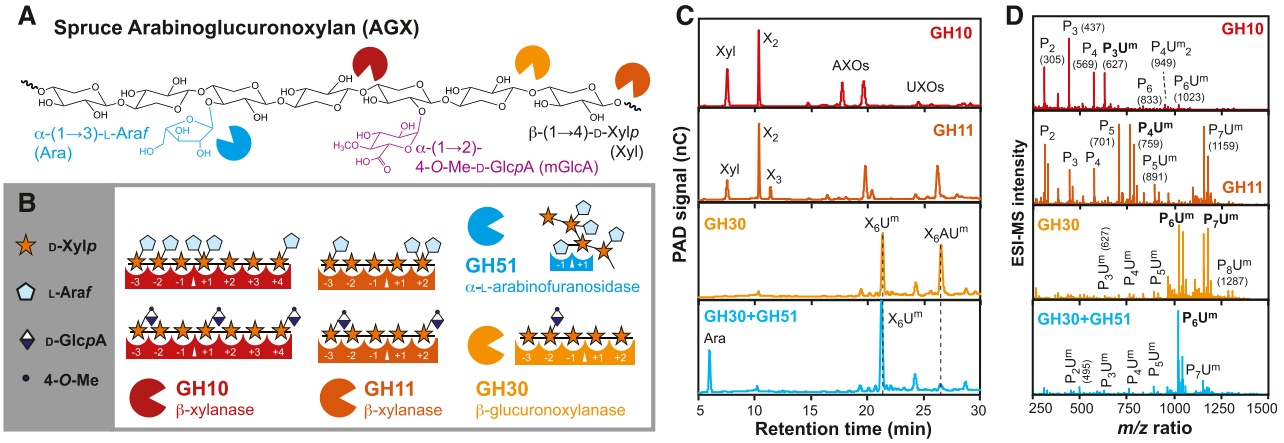

Figure 1. Enzymatic profiling of spruce xylan. A, General structure of AGX. B, Substrate recognition pattern of the xylanolytic enzymes used for AGX digestion, based on previous literature (Pell et al., 2004; Vardakou et al., 2008; Pollet et al., 2010; St John et al., 2011). The position of the hydrolytic activity is marked with a wedge at the glycosidic linkage between the −1 and +1 subsites. C, Oligosaccharide fingerprinting with high-pH anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) for the spruce AGX-A. D, Oligomeric mass profiling with ESI-MS of the AGX-A. Note that P refers to a pentose (Xyl or Ara) and U^m refers to mGlcA.

regions that prevent the hydrolytic action of the xylanases by steric hindrance, as we observed in our previous study (McKee et al., 2016).

The use of a specific β-glucuronoxylanase (GH30) releases acidic oligosaccharides based on the recognition of the (m)GlcA substitution at the xylosyl residue in position −2, which enables the identification of the GlcA spacing (Fig. 1; St John et al., 2011; Bromley et al., 2013). Interestingly, two main aldouronic acid oligosaccharides (P_6U^m and P_7U^m) can be observed in AGX-A (Fig. 1). Smaller oligosaccharides (P_2U^m–P_5U^m) can be observed as well with minor relative abundance. Further incubation with an additional *exo*-α-arabinofuranosidase (GH51) removes the potential Ara substitutions and provides singly substituted aldouronic acids with increasing lengths (Fig. 1). Interestingly, the peak corresponding to P_7U^m disappears from the HPAEC-PAD and ESI-MS profiles, whereas the relative abundance of P_6U^m increases. Therefore, these oligosaccharides can be ascribed to a xylohexaose containing a single Ara and mGlcA substitutions and a xylohexaose containing one mGlcA substitution, respectively, indicating the existence of a predominant repetitive spacing of mGlcA on every six Xyl units in the backbone. Further sequencing of the oligosaccharide motifs in AGX-A was performed by LC-ESI-MS/MS. Single ion monitoring (SIM) reveals the presence of isobaric oligosaccharides in the different series from P_3U^m and P_7U^m (Fig. 2). Remarkably, few isomers can be observed for each specific mass-to-charge ratio (*m/z*), indicating the presence of nonrandom structures for each isobaric series. Two main peaks can be observed in the SIM chromatograms for AGX-A at *m/z* 1,263 and 1,423. ESI-MS/MS fragmentation of the main oligosaccharide after derivatization at *m/z* 1,263 confirms the placement of the mGlcA substitution in the xylosyl residue at position −2 from the reducing end; this corresponds with the structure $XXXXU^mX$ (Fig. 2), following the systematic nomenclature proposed for xylooligosaccharides (Fauré et al., 2009), in agreement with the cleavage pattern for GH30 glucuronoxylanases. The identical position of the mGlcA substitution was verified for other acidic oligosaccharides in the series (XU^mX, XXU^mX, $XXXU^mX$, and $XXXXXU^mX$), as presented in Supplemental Figure S4.

The SIM chromatogram for *m/z* 1,423 revealed the presence of two oligosaccharides with Ara substitutions that disappear with the addition of the GH51 arabinofuranosidase. The structure of the main oligosaccharide could be fully sequenced (Fig. 2), and the position of the Ara substituent was located two Xyl units away from the mGlcA substitution point (xylosyl residue at position −4 from the reducing end), corresponding with the structure XXA^3XU^mX. A similar relative placement of the Ara unit two residues away from the mGlcA was determined for the A^3XU^mX fragment sequenced for *m/z* 1,103 (Supplemental Fig. S4). The minor oligosaccharide motif A^3XXXU^mX also was sequenced (Supplemental Fig. S4), with the Ara substitution located in the nonreducing end, four Xyl units away from the U^m substitution point (xylosyl residue at position −6 from the reducing end). It is significant that the placement of the Ara and mGlcA substitutions in these motifs always occurs in even Xyl positions, confirming a controlled level of regularity in spruce AGX. These two main repeating oligosaccharide motifs ($XXXXU^mX$ and XXA^3XU^mX) also were found in the digestion profile of the alcohol-insoluble residue of the debarked stem of three other coniferous species, namely Douglas fir (*Pseudotsuga menziesii*), black pine (*Pinus nigra*), and juniper (*Juniperus communis*; Busse-Wicher

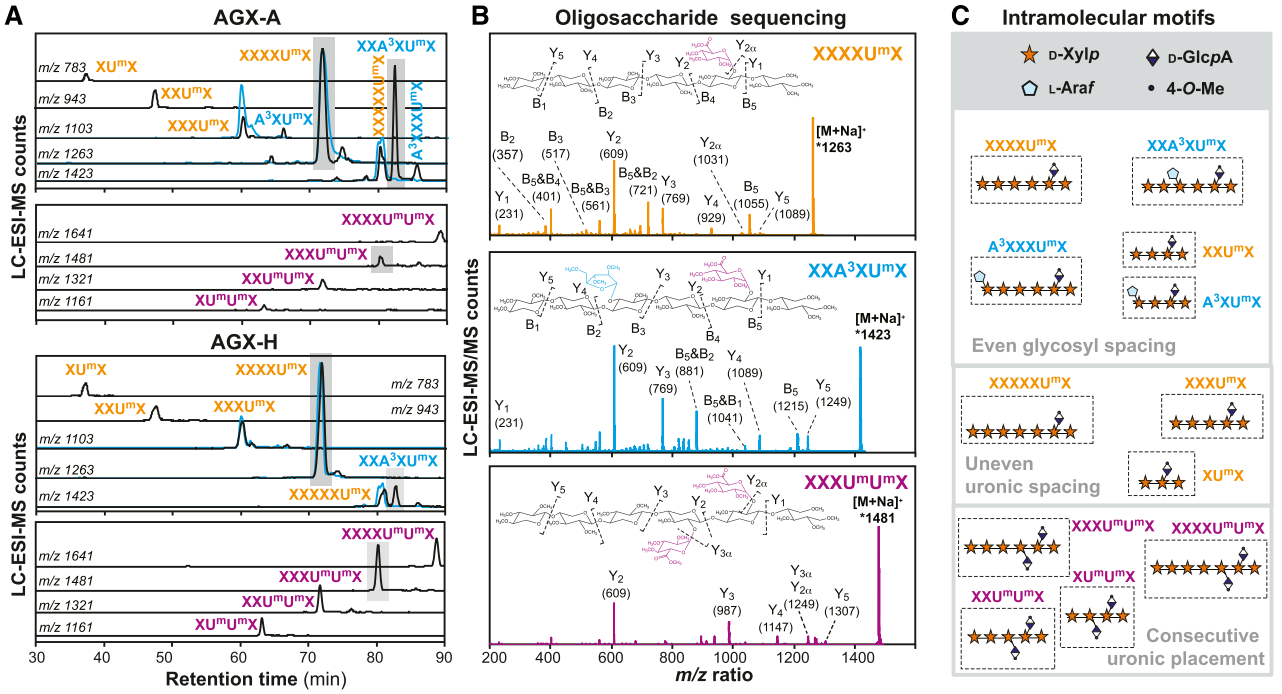

Figure 2. Comprehensive MS sequencing of selected oligosaccharides from spruce xylan. A, SIM chromatograms of the oligosaccharides released by the GH30 glucuronoxylanase alone (in black) or in combination of a GH51 arabinofuranosidase (in blue). B, Oligosaccharide fragmentation by MS/MS (fragment assignation according to Domon and Costello, 1988). C, Classification of the oligosaccharide motifs in spruce xylan based on the substitution pattern.

et al., 2016b). We can hereby confirm that these structures also are found in lignified softwoods from the genus *Picea*, and we here extend the regular and even placement of Ara and mGlcA substitutions to other minor motifs (A^3XXXU^mX and A^3XU^mX) also present in AGX.

Differences in Substitution Pattern between Xylan Populations Extracted by Alkali and Subcritical Water

The digestion of AGX-H with GH10 and GH11 releases similar oligosaccharide profiles as for AGX-A, with a marked presence of longer recalcitrant oligosaccharides (Supplemental Fig. S2). The SIM chromatograms for AGX-H after digestion with the GH30 glucuronoxylanase (Fig. 2) verify the occurrence of $XXXXU^mX$ as the main repetitive motif, with a lower abundance of XXA^3XU^mX. The loss of Ara substitutions in AGX-H could be ascribed to hydrolytic cleavage during subcritical water, since Ara moieties are more labile compared with the resistant mGlcA substituents (Willför et al., 2009). The use of a buffered pH mitigates backbone degradation by autohydrolysis (Ruthes et al., 2017); however, the occurrence of backbone hydrolysis during subcritical water extraction cannot be completely excluded. Higher relative abundance of shorter alduronic acids (XU^mX, XXU^mX, and $XXXU^mX$) can be observed, corresponding with intramolecular domains with tighter mGlcA spacing, in agreement with the higher mGlcA-Xyl ratio (Supplemental Table S1). AGX-H exhibits a series of acidic oligosaccharides with multiple mGlcA units ($P_4U^m{}_2$, $P_5U^m{}_2$, $P_6U^m{}_2$, and $P_7U^m{}_2$) barely detected in AGX-A (Fig. 2). Remarkably, separation and sequencing of the oligosaccharides corresponding to *m/z* 1,481 ($P_6U^m{}_2$) showed that only one isomer is present and reveals the univocal and adjacent position of the mGlcA substitution in neighboring Xyl units (Fig. 2), corresponding to structure $XXXU^mU^mX$. An identical structure with consecutive placement of mGlcA units was confirmed for the other oligosaccharides in the $X_nU^mU^mX$ series (Supplemental Fig. S5). This pattern of glucuronation in xylans, where two mGlcAs are attached consecutively to neighboring Xyl*p* backbone units, has been reported previously for larch (*Larix* spp.; Shimizu et al., 1978), sugi (*Cryptomeria japonica*) and hinoki (*Chamaecyparis obtuse*; Yamasaki et al., 2011), and suggested for spruce (Jacobs et al., 2001). The presence of major domains with even distribution of the mGlcA units along the backbone and minor domains with uneven glucuronation was already reported in glucuronoxylan from Arabidopsis (*Arabidopsis thaliana*), which was ascribed to the action of two distinct glucuronyltransferases (Bromley et al., 2013). This structural information reveals that the intramolecular glycosylation pattern in AGX is more complex than reported previously (Busse-Wicher et al., 2016b). We here confirm (1) the relative abundance of regular motifs with even placement of Ara and mGlcA

units along the Xyl backbone, with a predominant spacing of mGlcA every six units, (2) the presence of minor motifs with tighter and odd glucuronation spacing, and (3) the existence of domains with adjacent mGlcA substitutions (Fig. 2).

Influence of Regular Substitution Motifs on the Backbone Conformation in Solution

Four oligosaccharides with a common xylohexaose backbone were considered to study the effect of the substitution pattern on the xylan conformation in solution, in agreement with the sequenced structures in Figure 2B: (1) unsubstituted xylohexaose (X_6); (2) xylohexaose with one mGlcA α-(1→2)-linked to a xylosyl residue in position −2 ($XXXXU^{m}X$); (3) xylohexaose with an mGlcA α-(1→2)-linked to a xylosyl residue in position −2 and an Ara α-(1→3)-linked to a xylosyl residue in position −4 ($XXA^{3}XU^{m}X$); and (4) a xylohexaose with two mGlcA substituents α-(1→2)-linked to consecutive xylosyl residues in positions −2 and −3 ($XXXU^{m}U^{m}X$). The nomenclature used for the specific glycosidic linkages in the xylo-oligosaccharides (XOs) is presented in Figure 3. Further details about the molecular dynamic simulations of the sequenced spruce XOs are presented in Supplemental Text S1 and Supplemental Figures S6 to S8. The conformation of the xylan backbone can be modeled by the two dihedral angles (φ and ψ) related to the rotations around the two covalent C-O bonds that form each glycosidic linkage (Fig. 3). Previous modeling studies show that the preferred conformation for xylan in solution is a twisted 3_1-fold helical screw conformation (Mazeau et al., 2005; Busse-Wicher et al., 2014), the same as was reported previously for the experimentally determined crystal structure of xylan hydrate (Nieduszynski and Marchessault, 1971). Free energy surfaces for the backbone glycosidic linkage conformations are presented in Figure 4 (for X_6 and $XXA^{3}XU^{m}X$) and Supplemental Figure S9 (for all other glycosidic linkages). They are divided into four different regions, where region 4 is by far the most populated one. Region 4 is further divided into two parts, $4^{(-)}$ and $4^{(+)}$. The sum $\varphi + \psi$ is indicative of the saccharide chain conformation (French and Johnson, 2009), where a value of 420° represents a left-handed 3_1-fold helical screw and 360° indicates a flat 2_1-fold conformation. Here, $4^{(+)}$ corresponds to a mean value of $\varphi + \psi$ of approximately 440°, and $4^{(-)}$ gives $\varphi + \psi \approx 390°$. Since these subregions are nearly equally populated, the average value of $\varphi + \psi$ becomes ~415°, meaning that the conformation can be characterized on average as a flexible 3_1-fold helical screw. This is in agreement with our previous observations for X_2 and X_4, where the same simulation methodology was used as in this study (Berglund et al., 2016).

The introduction of a single α-(1→2)-linked mGlcA, as in $XXXXU^{m}X$ and $XXA^{3}XU^{m}X$, did not have a significant impact on the conformational space of the glycosidic linkages in aqueous solution. However, even if the main conformation of $XXXU^{m}U^{m}X$ still is the same, the probability of finding GL4 in region 1 increased 4-fold compared with X_6 (Supplemental Fig. S9), which causes a backbone twist resulting in the two mGlcA substituents pointing out in the same direction (Supplemental Fig. S10). Nevertheless, this effect is not present for the control simulations on the deprotonated oligosaccharides $XXXU^{m(-)}U^{m(-)}X$ (Supplemental Fig. S11). This is not surprising, since the side groups then would repel each other, and the conformational behavior seems to be pH dependent. Otherwise, deprotonation showed a small effect on the solution structures on the acidic oligosaccharides (Supplemental Fig. S11).

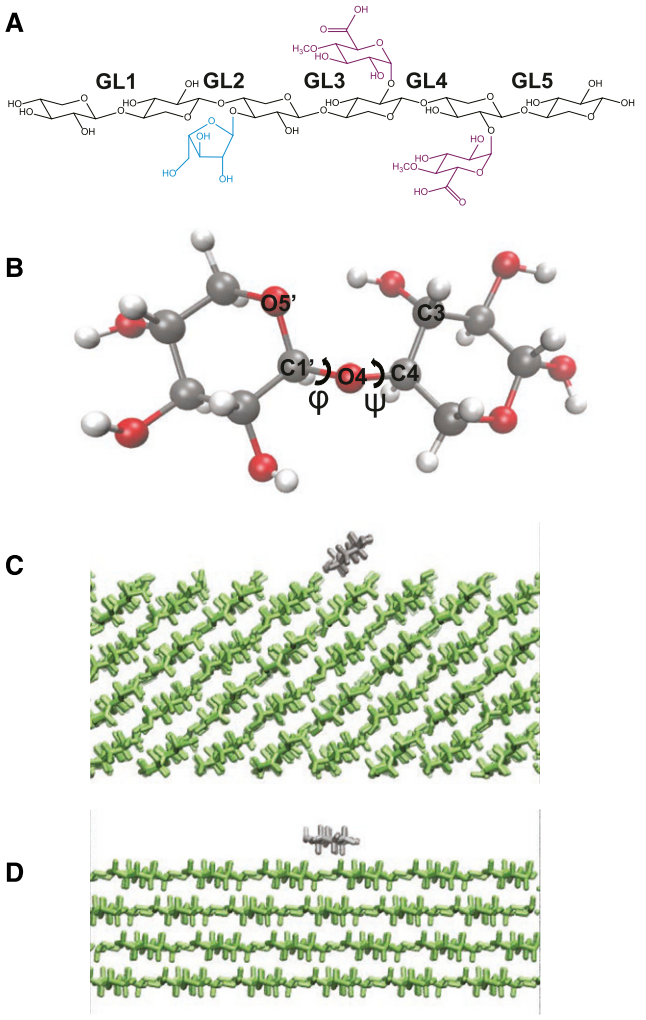

Figure 3. A, Nomenclature of the glycosidic linkages (GL) of the simulated XOs. B, Definition of dihedral angles in the backbone GLs: φ = O5′-C1′-O4-C4 and ψ = C1′-O4-C4-C3. For the GL connecting Ara to the xylopyranose backbone, the dihedral angles are defined as φ = O4′-C1′-O3-C3 and ψ = C1′-O3-C3-C2, and for the mGlcA GLs, φ = O5′-C1′-O2-C2 and ψ = C1′-O2-C2-C1. C, A xylooligomer (gray) on a (1-10) cellulose surface (green). D, A xylooligomer (gray) on a (200) cellulose surface. The views are directed along the cellulose chain axis (from the nonreducing end to the reducing end).

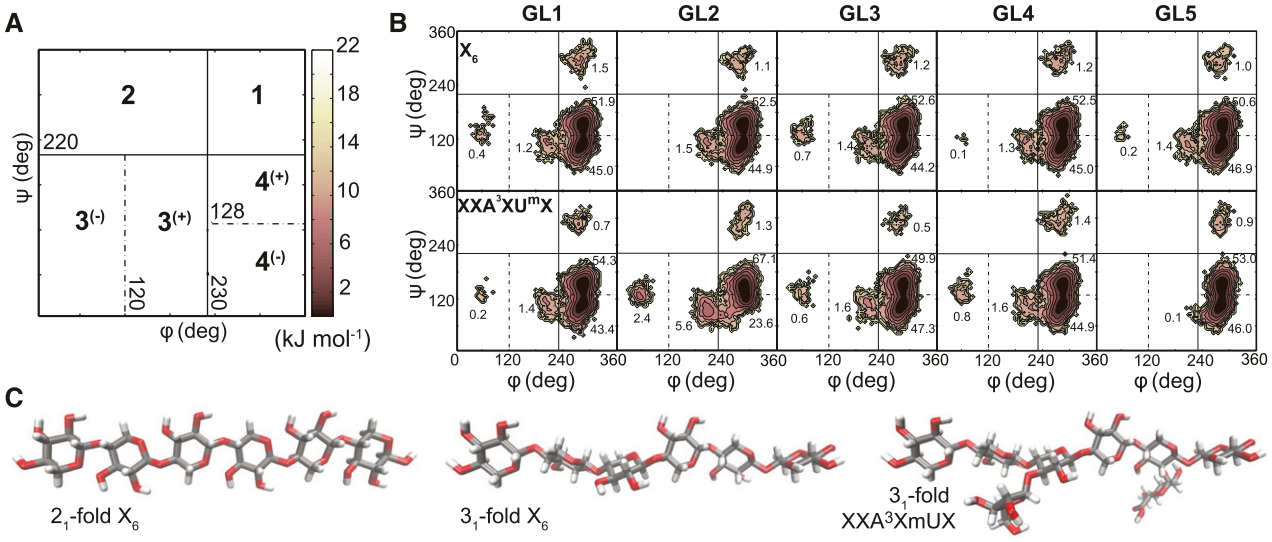

Figure 4. Conformation of XOs in water solution. A, Division of the free energy surfaces into regions [1, 2, $3^{(+)}$, $3^{(-)}$, $4^{(+)}$, and $4^{(-)}$]. B, Free energy surfaces (φ, ψ) for the backbone glycosidic linkages in X_6 and XXA^3XU^mX. The probabilities (%) for each region within each free energy surface are presented, and the SE for the probabilities is 0.0 to 1.6. C, Structure and conformation of X_6 and XXA^3XU^mX. For the 2_1-fold conformation, the glycosidic linkages are (280,° 80°), and for the 3_1-fold conformation, they are (290°, 130°). The linkage connecting the Ara side group is (295°, 250°), and that for mGlcA is (80°, 90°), which are common minima for the respective side groups (Supplemental Fig. S15).

Deprotonation of acidic groups may have significance on pH and ionic strength regulation in plant cell walls (White et al., 2014). Furthermore, the presence of an α-(1→3)-linked Ara disturbed the conformational space of GL2 in XXA^3XU^mX and became more prone toward twisted conformations [$4^{(+)}$ and $3^{(+)}$], with a decrease in the probability for $4^{(-)}$ (Supplemental Fig. S10). This change in backbone flexibility may be caused by steric hindrance from the Ara decoration, which removes the possibility for hydrogen bonding over the glycosidic linkage (O5′…H-O3; Supplemental Table S2). On the other hand, both Ara and mGlcA are highly hydrated. Therefore, side groups can be envisioned as highly hydrated moieties of the xylan macromolecule.

Xylan Binding onto Cellulose Hydrophilic and Hydrophobic Surfaces

Solid-state NMR on never-dried Arabidopsis stems (Simmons et al., 2016) and technical pulp model systems (Larsson et al., 1999; Teleman et al., 2001) indicate that xylan undergoes a conformational change when interacting with cellulose. Based on in silico predictions of NMR chemical shifts by density functional theory and MD simulations, this change has been identified as the transition from a 3_1-fold helical structure to a flat 2_1-fold screw (Busse-Wicher et al., 2016b; Simmons et al., 2016), which intuitively would facilitate close interaction between the xylan polymer and the flat cellulose surface, and even cocrystallization. Therefore, we initially placed the XOs in conformations that corresponded to the cellulose Iβ structure (which constitute a 2_1-fold screw) in an extra, hypothetical surface layer.

The cross section of cellulose microfibrils has been a topic of debate over the last decade. The crystal structure of the cellulose Iβ allomorph (Nishiyama et al., 2002) offers several possible crystallographic planes to be exposed at the surfaces, where the (1-10), (110), and (010) surfaces, on the one hand, and the (200) surface, on the other hand, are generally referred to as hydrophilic and hydrophobic cellulose surfaces, respectively. Depending on which model is assumed for the cross section, different crystallographic planes will be exposed, and in various proportions (see discussion in Cosgrove, 2014). Traditionally, the choice has been to assume a square cross section, primarily exposing the (110) and (1-10) surfaces, with the hydrophobic (200) face only present at two corner chains. This model has the benefit of being essentially hydrophilic, which makes sense in an aqueous environment. Ding and Himmel (2006) have advocated a model that builds on the rectangular model but in which several corner chains are removed, producing a diamond-shaped cross section. This model thus has a larger proportion of hydrophobic surface exposed to the surroundings. Finally, the rectangular model (Fernandes et al., 2011) was recently proposed as the model that best fitted a combination of small-angle neutron scattering and wide-angle x-ray scattering on spruce samples. This model exposes the (010) and (200) planes and, thus, has equal proportions of hydrophilic and hydrophobic characteristics. At present, it is not known which model is closest to the truth, although it is clear that it can have

great implications on the interactions between microfibrils and with other constituents of the plant cell wall. In this work, we chose to limit the investigation to one hydrophilc (1-10) surface and the hydrophobic (200) surface. The implications of this choice are elaborated further in "Discussion."

Figure 5 shows the free energy space of φ and ψ from XOs interacting with both hydrophilic and hydrophobic cellulose surfaces. Since a sum of $\varphi + \psi = 360°$, which corresponds to φ/ψ values along the diagonal, indicates a rigid 2_1-fold screw (French and Johnson, 2009), it is clear that this conformation is maintained for the three central glycosidic linkages of X_6 throughout the simulations. Interestingly, just as in solution, the presence of an Ara side group impacts the backbone conformation also when interacting with a surface. Common conformations of XXA^3XU^mX systems are shown in Figure 5. Although the 2_1-fold conformation is still common, the area in the free energy space becomes larger and more twisted glycosidic linkages become frequent. The XXA^3XU^mX (both protonated and deprotonated) docked on the (200) surface undergoes dynamic changes during MD simulations, from the Xyl backbone parallel to the cellulose chains to a backbone twist in the Xyl at position −4 (Supplemental Video S1), and eventually it returns to the parallel conformation again. This flexibility could favor the affinity of these XXA^3XU^mX motifs toward cellulose microfibrils that exhibit twisted conformations in the axial dimension (Fernandes et al., 2011). On the other hand, mGlcA substitutions do not affect significantly the conformational space of the xylan backbone, either in their protonated or deprotonated state (Supplemental Figs. S12–S14). The conformational space for mGlcA both in solution and on the hydrophilic surface was similar, with the main minima at $\varphi = 80°$ and $\psi = 90°$. However, when interacting with the hydrophobic surface, the minima became more localized (Supplemental Fig. S15), indicating restricted motion. Interestingly, the mGlcA side group, especially at −3 in $XXXU^mU^mX$ interacting with (200) cellulose, shows minima around $\varphi = 100°$ and $\psi = 300°$. This conformation of the mGlcA unit at −3 seems to interact better with the hydrophobic surface, as depicted in Figure 5. Additional snapshots from the simulations of the different XOs on the (1-10) and the (200) surfaces are depicted in Supplemental Figures S16 to S18.

Quantification of Molecular Adsorption Interactions between Xylan and Cellulose Surfaces by Pulling Out Experiments

After extensive equilibration of the oligosaccharides adsorbed on cellulose surfaces, the XOs were slowly pulled off the cellulose surface by applying a harmonic potential on the surface normal component of the center of mass distance between the cellulose and the XO, with a reference value that increased with a constant velocity of 0.1 nm ns^{-1}, until the oligosaccharide was completely separated from the surface (see description in "Materials and Methods," and Supplemental Text S1). These simulations were used to calculate the reversible work of adhesion (Gibbs' free energy) of the oligosaccharides to cellulose (Table I; Fig. 6). The calculated free energies indicate that side groups generally improve the interaction between XOs and cellulose surfaces, with the structure $XXXU^mU^mX$ overall exhibiting the strongest ones. The weakest interactions with the (1-10) surface were those of the unsubstituted X_6 structure, together with $XXXXU^mX$ having a single substitution, followed closely by XXA^3XmUX. It is noteworthy that the latter two were positioned with their side groups pointing away from the surface, whereas $XXXU^mU^mX$, which exhibited the strongest interaction, has one side group pointing directly to the cellulose. This is in agreement with a recent study that showed that α-(1→2)-linked substitutions improve the interaction between xylan and cellulose (Pereira et al., 2017). These results also clearly show that the interaction with the hydrophobic (200) surface is stronger compared with the hydrophilic (1-10) one. This was also suggested previously (Pereira et al., 2017), where it was seen that xylan diffused from hydrophilic to hydrophobic surfaces at elevated temperatures. Therefore, a substituted structure is a clear advantage for the interaction with the (200) hydrophobic surface, where the weakest interaction was observed for the unsubstituted structure. In this case, for the substituted oligomers, no side groups are pointing away from the surface, as all residues lie in the same plane. The effect of deprotonation on the interactions with the different surfaces shows no clear trend on either hydrophobic or hydrophilic surfaces. Hydrogen bonding of the XOs was analyzed with both cellulose and water (Supplemental Fig. 19). As expected, hydrogen bonds rarely occur between the XOs and the hydrophobic surface (≤1), while four to five hydrogen bonds were formed on average with the hydrophilic one.

The free energy is a state variable; thus, the calculated desorption energies depend only on the end states and not on the actual path between them. But the reaction path can still convey useful information about the system. Since the pull force was applied to the center of mass, the XOs could adopt to the pulling by following a relatively low-energy trajectory. This means that parts of the XOs that were weakly bound were more likely to detach first. The simulation trajectories showed that detachment always started at a chain end, followed by a process during which the rest of the XO was slowly peeled off. For the unsubstituted XOs, detachment was initiated randomly at either chain end, but for the substituted ones, detachment was much more likely to start at the end farthest away from the substitution, which, in this case, was the nonreducing end. This supports the notion of the side groups acting as anchors of the xylooligosaccharides to the cellulose surface. A series of snapshots showing the conformation of a disubstituted XO during a pull-off simulation is given in Supplemental Figure S20.

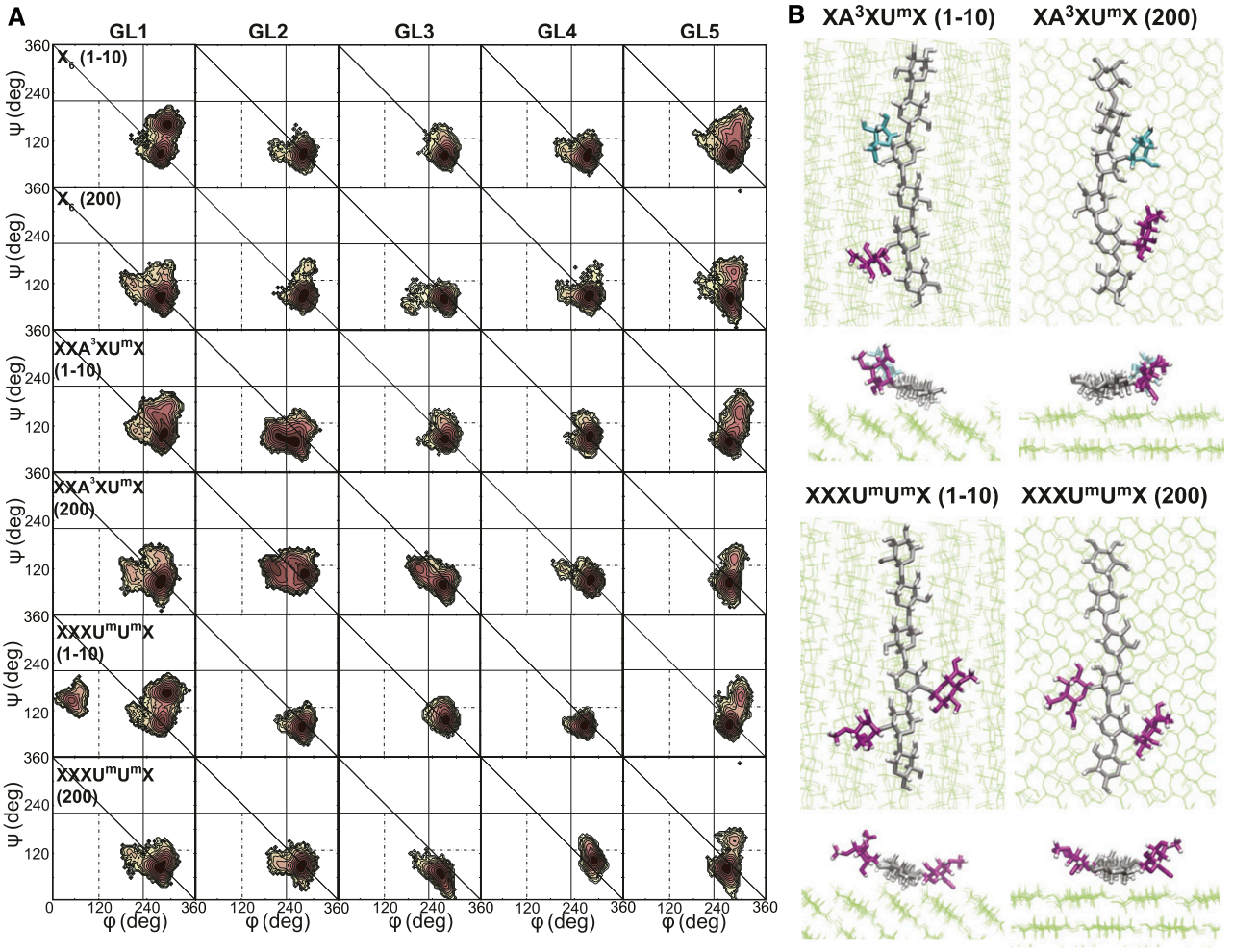

Figure 5. Conformations of XOs located on cellulose surfaces. A, Free energy surfaces of X_6, XXA^3XU^mX, and $XXXU^mU^mX$. The energy bar and notation of the glycosidic linkages are the same as in Figures 3 and 4. B, Snapshots of XXA^3XU^mX and $XXXU^mU^mX$ on (1-10) and (200) surfaces. Additional snapshots from the simulations of XOs on cellulose surfaces are presented in Supplemental Figures S17 to S19.

DISCUSSION

In this study, we report the existence of intramolecular motifs with distinct placement of glycosyl substitutions along the backbone in xylan extracted from spruce secondary cell walls. We have identified a major motif with a regular substitution pattern consisting of $XXXXU^mX$ and XXA^3XU^mX, with one mGlcA unit regularly placed every six Xyl units and an Ara moiety locked −2 units from the mGlcA, respectively. These oligosaccharide motifs also have been found in xylan extracted from the stem of other coniferous species (Busse-Wicher et al., 2016b), and their occurrence here is independently verified in softwoods from the genus *Picea*. This prevalent even spacing of substitutions in xylan had been proposed as well for acetylation in Arabidopsis glucuronoxylan (Busse-Wicher et al., 2014; Chong et al., 2014). Additionally, we have identified the presence of minor xylan domains with a tighter and uneven placement of mGlcA substitutions and a novel repetitive motif with two mGlcAs placed in adjacent positions. This demonstrates the occurrence of a more complex and controlled molecular regularity in softwood xylans than hitherto reported. Bromley et al. (2013) already proposed the presence of molecular domains with different uronic acid spacing for glucuronoxylan in Arabidopsis, which was attributed to the action of two different glucuronyltransferases. The presence of distinct domains with even and clustered placement of glycosyl decorations may be directed by the 3D spatial arrangement of the glucuronyltransferase and arabinofuranosyltransferase catalytic units with respect to the nascent xylan chain. Indeed, the biosynthesis of AGX is still not well known (Rennie and Scheller, 2014), and future studies on the catalytic machinery and enzyme complexes in the Golgi apparatus should cast further light.

The hydration properties and the conformation of cell wall polysaccharides in aqueous solutions have

Table I. *Equilibrium free energies needed to pull the XOs off the cellulose surfaces*
Estimated errors are shown in parentheses.

Oligosaccharide Structure	ΔG	
	Hydrophilic Surface (1-10)	Hydrophobic Surface (200)
	kJ mol^{-1}	
X_6	71.6 (2.7)	91.2 (2.3)
XXXXU^{m}X	72.4 (3.2)	113.2 (2.0)
XXXU^{m}U^{m}X	104.4 (7.7)	114.3 (5.1)
XXA^{3}XU^{m}X	80.0 (3.8)	114.1 (3.9)
XXXXU$^{m(-)}$X	68.3 (2.2)	98.4 (4.4)
XXXU$^{m(-)}$U$^{m(-)}$X	92.9 (2.5)	123.3 (2.7)
XXA^{3}XU$^{m(-)}$X	92.0 (3.5)	108.8 (4.6)

been matters of intense discussion in recent times. The different β-(1→4) glycosidic linkage types present in the backbone of hemicelluloses show distinct conformation flexibility in solution prone to a 3_1-fold helix, with the xylan backbone being the most flexible compared with the (gluco)mannan and glucan (cellulose) backbone types (Berglund et al., 2016). Here, the effect of the repetitive glycosidic substituents in stabilizing the solution conformation of the xylan backbone has been evaluated. The presence and the protonation state of α-(1→2)-mGlcA substitutions do not seem to exert a strong influence on the conformation of the neighboring β-(1→4)-Xyl units. On the other hand, the presence of α-(1→3)-Ara increases chain flexibility, which can be attributed to the hindered intramolecular hydrogen bonding possibilities of the xylan backbone caused specifically by the presence of a substituent at the *O*-3 position. Importantly, the presence of Ara and mGlcA side groups affects the probability of interaction with surrounding water by hydrogen bonding compared with the xylan backbone, where intermolecular hydrogen bonds also play a role (Berglund et al., 2016). Therefore, these glycosyl substitutions enhance the hydration state of the xylan macromolecule locally in these positions and prevent the possibility of intermolecular and intramolecular interactions between xylan backbone segments from the same or different macromolecules. This has a strong influence on the aggregation of xylan in dispersions and on the hydrodynamic properties of xylan-based materials. Indeed, higher Ara and mGlcA contents have been reported to hinder the supramolecular aggregation of xylans in aqueous systems (Sternemalm et al., 2008; Pitkänen et al., 2009; Bosmans et al., 2014) and to influence the reactivity and macroscopic properties of xylan-based materials (Höije et al., 2008; Escalante et al., 2012; Littunen et al., 2015).

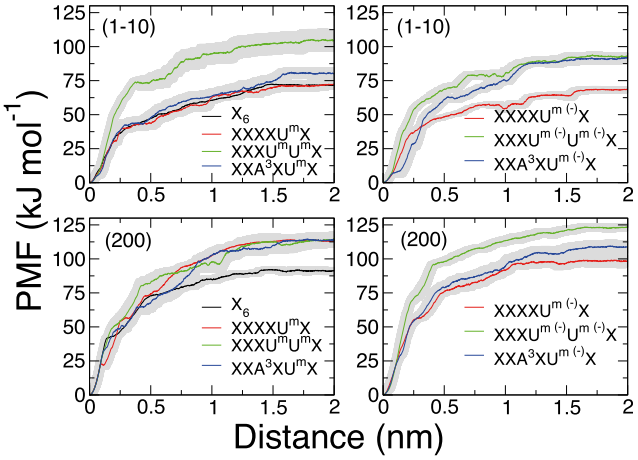

Figure 6. Pulling-out experiments of the different XOs from the hydrophilic (1-10) and hydrophobic (200) cellulose microfibril surfaces. The diagrams exhibit the pulling mechanical free energies (PMF; in kJ mol^{-1}) for X_6, XXXXU^{m}X, XXA^{3}XU^{m}X, and XXXU^{m}U^{m}X (in protonated and deprotonated forms) as a function of the distance (in nm) from the celluloses surfaces.

Our MD data show the differences in the xylan conformation depending on the surrounding environment in the plant cell wall. The conformation of XOs in solution is that of a relatively flexible 3_1-fold screw, whereas XOs that are docked onto cellulose surfaces retain a flat, relatively rigid, 2_1-fold screw conformation. This is in agreement with both modeling studies and NMR data showing a distinct change in conformation in xylan when there is cellulose present (Teleman et al., 2001; Simmons et al., 2016). Here, we examined these observations for XOs bound to the hydrophilic (1-10) surface and the hydrophobic (200) one, and our simulations show that the 2_1-fold screw backbone conformation does not differ significantly between these surfaces. Interactions between xylan and cellulose on the hydrophilic surface have been analyzed previously in terms of hydrogen bonding (Zhang et al., 2015; Busse-Wicher et al., 2016b). Interestingly, we find that interactions between XOs and cellulose are stronger on the hydrophobic surface than on the hydrophilic one. As our analysis shows that there is virtually no hydrogen bonding taking place between the xylan and the cellulose on these surfaces, it is clear that the interaction is of a different nature. It is instructive to make an analogy to the structure of cellulose, which consists of stacked sheets with an intricate network of in-plane hydrogen bonds. Although the recalcitrance of this structure often has been ascribed to hydrogen bonds alone, computer simulations show that the in-plane hydrogen bonds are 1 order of magnitude weaker than

the intraplane interactions that exist between the sheets (Bergenstråhle et al., 2010). This interaction arises as a hydrophobic effect from the relatively large penalty in free energy of hydrating the pseudoflat hydrophobic surfaces of the pyranose rings. For the case of xylan, the flat 2_1-fold screw conformation mimics that of cellulose chains, and a similar effect arises here. On the hydrophilic surfaces, the interaction is a combination of both hydrogen bonds and hydrophobic effects, but on the (200) surface, hydrophobic effects dominate, and the interaction becomes stronger. Although the conformation of the backbone is the same on both surfaces considered here, the hydrophobic surfaces seem to stabilize the conformation of the side groups docked on them compared with the hydrophilic surfaces. This conclusion is in agreement with previous observations reporting the migration of an AGX chain from the hydrophilic surface to the hydrophobic surface of cellulose upon thermal treatment at 160°C using MD simulations (Pereira et al., 2017).

The MD docking experiments of different XOs with distinct Ara and mGlcA patterns show that the placement of the substitutions is well tolerated with both hydrophilic (1-10) and hydrophobic (200) surfaces. In line with our findings, MD simulations of xylan adsorption onto a cellulose surface in gas phase showed that side groups did not affect the ability of one xylan molecule to adsorb onto a surface (Mazeau and Charlier, 2012). In addition to this, our free energy calculations showed that side groups, and especially the $XXXU^mU^mX$ structure, where mGlcA are pointing toward the cellulose surface, had a positive impact on the interaction between xylan and cellulose. These observations challenge a previous study, where it was suggested that the common structure XXA^3XU^mX, with side groups spaced at every other xylan unit, is favored, since the side groups can point out to the same side, which makes xylan fit well onto the cellulose (Busse-Wicher et al., 2016b). Orienting the side groups away from the surface could potentially allow for a closer packing of multiple xylan chains, possibly allowing the formation of a layer of single xylan molecules fully coating the cellulose surface, as is proposed for the hydrophilic (110) surface (Busse-Wicher et al., 2016b; Simmons et al., 2016). On the other hand, however, it is possible that the gain in interaction energy by having the side groups in contact with the cellulose would compensate for the loss of near-crystalline order and the stabilizing effect of neighboring chains. The effect of consecutive side chains on the subsequent binding of other xylan macromolecules onto cellulose surfaces, and the energetic balance between the anchoring effect of the side groups of a single xylan chain in contact with the cellulose surfaces and the stabilization by the binding of two adjacent xylan chains, should be the subject of further studies. In addition, these observations likely depend on which cellulose crystal model is used. On the hydrophilic (010) and (020) surface models used previously (Busse-Wicher et al., 2016b), xylan adsorbs into distinct grooves present at the surfaces. Here, xylan having the side groups pointing out on the same side does indeed seem favored. But if the hydrophobic (200) surface is considered instead, having side groups pointing out on opposite sides is not an obstacle for close xylan-cellulose interaction. On the contrary, having side groups on alternating sides improves the interaction, as in the case of $XXXU^mU^mX$. An interesting parallelism can be made with xyloglucan, another highly substituted hemicellulose that exhibits strong interaction with cellulose, where the side group residues are believed to anchor the xyloglucan onto the cellulose (Levy et al., 1991; Hanus and Mazeau, 2006; Cosgrove, 2014; Zhao et al., 2014). The fact that the part of the XOs without side groups detached before substituted residues in the pull-off simulations further strengthens this idea (Supplemental Fig. S20).

In our study, the XOs were oriented parallel to the cellulose chains, as suggested earlier (Busse-Wicher et al., 2016b), although theoretically other directions also are possible. Considering that the antiparallel structure of cellulose II is thermodynamically more stable than parallel cellulose I (Langan et al., 2001; Goldberg et al., 2015), it is not unlikely that hemicelluloses deposited next to the cellulose fibrils orient in an antiparallel fashion. Moreover, other orientations (e.g. perpendicular or random) cannot be completely ruled out. There also could be a difference between the (1-10) surface investigated here and the (110) surface, which closely resembles (1-10) but has a slightly different tilt and spacing of the cellulose chains. Another complicating factor is that the Iβ crystallographic unit cell contains two chains, implying that only every other surface chain is identical. Therefore, the binding free energy also could depend on where on the surface the xylan is deposited. Finally, the role of hydrophobic surfaces in wood microfibrils is a question that will depend on which cross section is assumed. In the rectangular model, the hydrophobic surfaces are abundant and quite extended, similar to the computational model used here. However, considering a diamond-shaped cross section, the hydrophobic surfaces are abundant but much narrower, only two to three chains wide, whereas in a square cross section, they are present only as single chains at the corners. The above-mentioned limitations do limit, to some extent, the general applicability of our results and should be a subject for future studies.

Our results indicate that the presence of glycosyl substitutions in xylan with controlled and well-defined placement along the backbone is not only sterically tolerated on both hydrophilic and hydrophobic surfaces but also favored in terms of xylan conformation and adsorption onto cellulose surfaces. This observation is contradictory to previous studies of in vitro adsorption of xylans from solution onto cellulose surfaces, in which a lower degree of substitution was correlated to a higher level of xylan adsorption (Köhnke et al., 2008, 2011; Bosmans et al., 2014). Additionally, technical xylans with lower glycosyl substituents formed during chemical pulping and readsorbed onto cellulose

microfibrils improve pulp strength (Danielsson and Lindström, 2005). However, the nature of the intermolecular interfaces between xylans adsorbed in vitro onto cellulose are different from those of xylans docked onto the surface of cellulose microfibrils during plant cell wall biosynthesis and assembly. In vitro adsorption processes may involve not only the quantification of xylan directly in contact with the cellulose surfaces but also xylan that accumulates in multiple layers due to crowding effects. In this latter case, the presence of side groups in xylans that specifically are beneficial for cellulose interactions might, at the same time, impose hindrances for additional xylan molecules to adsorb next to it, resulting in an overall decreased in vitro-quantified adsorption. This self-aggregation behavior of xylan with reduced substitution has been observed in plant cell wall analog models based on bacterial cellulose (Mikkelsen et al., 2015). This indicates that the results of in vitro binding experiments must be taken with caution (Cosgrove, 2014; Wang et al., 2015).

Finally, our integrated results by MS-based sequencing of spruce xylan and MD simulations impose further questions regarding the association of xylan with cellulose microfibrils and lignin in plant secondary cell walls. The major and minor domains with even and clustered backbone glycosylation in xylan are differently enriched in the AGX extracted from an alkaline procedure with previous delignification (AGX-A) and in the AGX populations extracted with subcritical water without delignification (AGX-H), respectively. The relative abundance of the minor domains with tighter and consecutive substitution spacing is small compared with the major domains, but they become enriched in the AGX populations extracted with subcritical water. We here hypothesize that the minor domains represent AGX populations with closer connection with lignin. The function of tighter mGlcA positioning might be to provide local environments with acidic pH in the plant cell wall, which could catalyze the formation of lignin-carbohydrate complexes through covalent cross-links with lignol precursors (β-*O*-4 bonds and γ-ester bonds) during the complex radical polymerization processes that occur during lignification. We have observed that the presence of $XXXU^mU^mX$ moieties induces conformational changes that could end with the two mGlcA moieties pointing out in the same direction, which would sterically favor the random formation of such carbohydrate-lignin bonds (Supplemental Fig. S10B). The presence of major and minor xylan domains with distinct glucuronation patterns already has been reported (Bromley et al., 2013). These domains were hypothesized to be present in the same xylan molecules based on the impossibility of separating such domains based on charge and size. Unfortunately, our study cannot ascertain whether the two extracted AGX-A and AGX-H fractions represent two different populations in planta or whether the differences in the relative amounts of the major and minor domains in our AGX-A and AGX-H populations arise from different original xylan molecules. We hypothesize that the alkaline treatment with previous delignification may alter the minor domains with consecutive glucuronation, which may be in closer contact with lignin. On the other hand, the subcritical water extraction may preserve intact these minor domains, becoming enriched in the AGX-H fractions. Indeed, subcritical water extraction is able to retain the feruloylation in arabinoxylan extracted from wheat (*Triticum aestivum*) bran (Ruthes et al., 2017). The presence of the major and minor domains with even and clustered glucuronation in the same or different xylan molecules is fundamental to their ability to cross-link between cellulose microfibrils and/or with lignin. Further studies on intact secondary cell walls by nondestructive techniques such as solid-state NMR may cast some light on the specific placement of xylan domains with respect to cellulose microfibrils and lignin components.

CONCLUSION

Our MS sequencing data on the substitution pattern of softwood xylan, together with the molecular modeling of xylan conformation in solution and interaction with cellulose surfaces, offer fundamental molecular insights into the assembly of secondary plant cell walls. The presence of predominant regular motifs with even placement of glycosyl (mGlcA and Ara) substitutions along the backbone is here confirmed for spruce AGX, together with minor motifs with odd and consecutive glucuronation spacing. The presence of glycosyl substitutions has no strong influence on the flexible 3_1-fold screw of the xylan backbone in solution; however, they may act as local moieties prone to hydration. Xylan oligosaccharides adopt a flat, relatively rigid, 2_1-fold screw conformation when they are docked onto hydrophilic and hydrophobic cellulose surfaces. This conformation is stabilized by the presence of glycosyl substituents and is not hindered by the consecutive placement of mGlcA substitutions, as demonstrated by the simulations. These results highlight the importance of nonpolar driving forces for the structural integrity of secondary plant cell walls, and they indicate that both hydrophobic and hydrophilic cellulose surfaces both should be strongly considered in future molecular models of plant cell walls. It is worth mentioning here that our modeling experiments consider a simplified model of conifer secondary cell walls without the interference of other cell wall components, such as *O*-acetylated galactoglucomannan and lignin. Indeed, previous studies have reported the close association of glucomannan and cellulose fibers in spruce wood (Åkerholm and Salmen, 2001), which may alter the interactions between xylan and cellulose. The confirmation of the specific placement of xylan domains with distinct substitution motifs with respect to cellulose microfibrils (including the different hydrophilic and hydrophobic surfaces), galactoglucomannan, and lignin components in lignified cell walls by nondestructive techniques remains an exciting challenge.

Until then, these new molecular insights on the presence of intramolecular repetitive motifs in xylans and their potential effects on the association with cellulose microfibrils contribute to an improved understanding of the supramolecular architecture of secondary cell walls, with fundamental implications for the development of processes to overcome lignocellulosic recalcitrance and for the design of advanced wood-based materials and products.

MATERIALS AND METHODS

Purification of Softwood Xylan and Compositional Analysis

AGX was extracted and purified from Norway spruce (*Picea abies*) secondary cell walls using two complementary processes: (1) alkali extraction after delignification (AGX-A) following the procedure reported by Escalante et al. (2012) and (2) subcritical water extraction and enzymatic purification (AGX-H; for details, see Supplemental Fig. S1). The monosaccharide composition of the xylans was analyzed after acid methanolysis (1 M trifluoroacetic acid, 3 h, 100°C) and HPAEC-PAD (ICS-3000; Dionex; McKee et al., 2016).

Digestion of Spruce AGX Using Xylanolytic Enzymes

The purified spruce xylans were depolymerized using different xylanolytic enzyme combinations (Fig. 1A). A GH10 *endo-β*-(1→4)-xylanase from *Cellvibrio mixtus* (Megazyme), a GH11 *endo-β*-(1→4)-xylanase from *Cellvibrio japonicus* (Nzytech), and a GH30 *endo-β*-(1→4)-glucuronoxylanase (kindly provided by James F. Preston) were used for the specific cleavage of the xylan backbone. A GH51 *exo-α*-arabinofuranosidase (Megazyme) was used to additionally remove Ara*f* substitutions. End-point digestions were performed by incubating 200 μL of AGX (1 g L^{-1}) with the enzymes (10 units mL^{-1}) for 16 h at 37°C in acetate or citrate buffers at the optimum pH for each enzyme. After incubation, the solutions were boiled for 10 min, filtered, and kept at −20°C for further analysis.

Oligosaccharide Profiling and Sequencing

The oligosaccharide profiles after enzymatic digestion were analyzed by HPAEC-PAD as reported previously (McKee et al., 2016). Linear xylooligosaccarides (X_2-X_6; Megazyme) were used as external standards. Oligomeric mass profiling was performed by ESI-MS using a Q-TOF2 mass spectrometer (Micromass). The hydrolysates were desalted with HyperSep Hypercarb cartridges (Thermo Fisher), dissolved in 50% acetonitrile and 0.1% formic acid, and infused directly into the positive mode-operated Q-TOF2 mass spectrometer through a syringe pump at a rate of 5 μL min^{-1}. Capillary and cone voltages were set to 3.3 kV and 80 V, respectively.

Oligosaccharide sequencing was achieved after the separation of labeled oligosaccharides by LC-ESI-MS/MS. Derivatization was performed by reduction in 2% borohydride (30 min) and permethylated in dimethyl sulfoxide with CH_3I as described previously (Ciucanu and Kerek, 1984). The organic phase was recovered after partition in CH_2Cl_2:water, dried, and resuspended in 50% acetonitrile. The labeled oligosaccharides were separated through an SB-C18 column (250 × 4.6 mm; Agilent Technologies) in a Capillary LC (Micromass) at a flow rate of 10 μL min^{-1} and a gradient of increasing acetonitrile content (30%–60%) over 110 min. MS and MS/MS analyses were performed with a quadrupole time-of-flight (Q-TOF) system (Waters) in positive mode at 3.3 kV and 60 V in the capillary and the cone, respectively. Argon was used as the collision gas for MS/MS analysis of selected ions, at a voltage of 35 to 90 V, to analyze the diagnostic fragmentation patterns of the oligosaccharides.

MD Simulations

MD simulations were performed on xylan oligomers both free in solution and docked to cellulose surfaces. The simulations lasted for 50 ns and were run with GROMACS 5.1.2 (Hess et al., 2008; Abraham et al., 2015), employing the GLYCAM06 force field for carbohydrates (Kirschner et al., 2008) and the TIP3P water model (Jorgensen et al., 1983). Lennard-Jones interactions used a cutoff of 1.2 nm, and electrostatic forces were calculated using particle mesh Ewald summation (Darden et al., 1993; Essmann et al., 1995) with a 1.2-nm real space cutoff. Bonds were kept at their equilibrium values by P-LINCS (Hess 2008) for the saccharides and by SETTLE (Miyamoto and Kollman, 1992) for water. In simulations of solution structures, constraints were applied to all covalent bonds, whereas for simulations with cellulose, constraints were used on bonds involving hydrogens only. In addition, harmonic restraints were used on xylan residues to keep them in the $^4C_1$ conformation. No scaling was used for one to four interactions, according to the GLYCAM convention.

During the simulations, constant pressure was kept at 1 atm using a Parrinello-Rahman barostat (Parrinello and Rahman, 1981), and temperature was controlled using a Nosé-Hoover thermostat (Nosé, 1984; Hoover, 1985). In simulations of xylan docked to cellulose, the temperature was maintained at 300 K, whereas for simulations in solution, a replicate-exchange scheme was applied (Sugita and Okamoto, 1999), with 12 replicates at temperatures ranging from 300 K to 366 K (evenly spaced by 6 K) and with attempts to exchange between neighboring replicates at every 10 steps.

Free energies for the glycosidic bond dihedral angles (φ and ψ) were calculated by Boltzmann inversion of the probabilities. Grid spacing in the plots is 5°, and contour levels are drawn at every 2 kJ mol^{-1}. Energy is anchored to 0 kJ mol^{-1}, and the highest energy level, for conformations that are never attained, is set to 22 kJ mol^{-1}. Hydrogen bonds were analyzed by the gmx hbond tool in GROMACS with the default geometrical criteria (donor-acceptor distance < 0.35 nm and hydrogen-donor-acceptor angle $< 30°$). The free energy of adsorption of docked xylan structures was calculated from the exponential average over 25 nonequilibrium pull-off simulations of 20 ns each, using Jarzynski's equality (Jarzynski, 1997). Additional information on the construction of the model systems and the free energy calculations is given in Supplemental Text S1.

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Scheme of the extraction and purification processes of spruce AGX.

Supplemental Figure S2. Enzymatic profiling of AGX-H.

Supplemental Figure S3. LC-ESI-MS/MS fragmentation and sequencing of relevant oligosaccharides released by digestion with GH10 and GH11 *β*-xylanases.

Supplemental Figure S4. LC-ESI-MS/MS fragmentation and sequencing of the X_nU^m oligosaccharide series released by digestion with a GH30 *β*-glucuronoxylanase.

Supplemental Figure S5. LC-ESI-MS/MS fragmentation and sequencing of the oligosaccharide series with consecutive glucuronation released by digestion with a GH30 *β*-glucuronoxylanase.

Supplemental Figure S6. Charge distribution of the protonated GlcA unit that was used in this work.

Supplemental Figure S7. Pucker space for Ara*f*.

Supplemental Figure S8. Glycosidic linkage conformational space in X_6, comparison between the force fields GLYCAM06 and CHARMM.

Supplemental Figure S9. All free energy surfaces from the different xylooligomers in water solution.

Supplemental Figure S10. Conformations of substituted xylooligomers in solution.

Supplemental Figure S11. Free energy surfaces for $XXXXU^{m(-)}X$ and $XXXU^{m(-)}U^{m(-)}X$ in water solution.

Supplemental Figure S12. Free energy surfaces of backbone glycosidic linkages of xylooligomers on a hydrophilic (1-10) cellulose surface.

Supplemental Figure S13. Free energy surfaces of backbone glycosidic linkages of xylooligomers on a hydrophobic (200) cellulose surface.

Supplemental Figure S14. Free energy surfaces of backbone glycosidic linkages of deprotonated xylooligomers on hydrophilic (1-10) and hydrophobic (200) cellulose surfaces.

Supplemental Figure S15. Free energy surfaces from the linkages connecting the side groups with the Xyl backbone.

Supplemental Figure S16. Snapshots from the simulations of xylooligomers on a (1-10) cellulose surface.

Supplemental Figure S17. Snapshots from the simulations of xylooligomers on a (200) cellulose surface.

Supplemental Figure S18. Snapshots from the simulations of XXA^3XU^mX on (1-10) and (200) cellulose surfaces.

Supplemental Figure S19. Average number of hydrogen bonds between xylooligomers and the cellulose surface and xylooligomer and the surrounding water.

Supplemental Figure S20. Snapshots from one of the pulling simulations of $XXXU^mU^mX$ from the (1-10) surface.

Supplemental Table S1. Composition of the AGX (as wt % of dry weight) by monosaccharide analysis.

Supplemental Table S2. Average number of hydrogen bonds for the X_6 and XXA^3XU^mX xylooligomers in water solution.

Supplemental Video S1. XXA^3XU^mX on (200) showing how the XO move on the cellulose surface.

Supplemental Text S1. Additional information from the MD simulations.

ACKNOWLEDGMENTS

We thank James F. Preston (University of Florida) for providing the GH30 glucuronoxylanase and the PDC Center for High Performance Computing for providing the computational resources (project PDC-2016-58).

Received September 26, 2017; accepted October 23, 2017; published October 25, 2017.

LITERATURE CITED

Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindahl E (2015) GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX **1-2:** 19–25

Åkerholm M, Salmen L (2001) Interactions between wood polymers studied by dynamic FT-IR spectroscopy. Polymer (Guildf) **42:** 963–969

Bergenstråhle M, Wohlert J, Himmel ME, Brady JW (2010) Simulation studies of the insolubility of cellulose. Carbohydr Res **345:** 2060–2066

Berglund J, Angles d'Ortoli T, Vilaplana F, Widmalm G, Bergenstråhle-Wohlert M, Lawoko M, Henriksson G, Lindström M, Wohlert J (2016) A molecular dynamics study of the effect of glycosidic linkage type in the hemicellulose backbone on the molecular chain flexibility. Plant J **88:** 56–70

Bosmans TJ, Stépán AM, Toriz G, Renneckar S, Karabulut E, Wågberg L, Gatenholm P (2014) Assembly of debranched xylan from solution and on nanocellulosic surfaces. Biomacromolecules **15:** 924–930

Bromley JR, Busse-Wicher M, Tryfona T, Mortimer JC, Zhang Z, Brown DM, Dupree P (2013) GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. Plant J **74:** 423–434

Burgert I, Keplinger T (2013) Plant micro- and nanomechanics: experimental techniques for plant cell-wall analysis. J Exp Bot **64:** 4635–4649

Burton RA, Gidley MJ, Fincher GB (2010) Heterogeneity in the chemistry, structure and function of plant cell walls. Nat Chem Biol **6:** 724–732

Busse-Wicher M, Gomes TCF, Tryfona T, Nikolovski N, Stott K, Grantham NJ, Bolam DN, Skaf MS, Dupree P (2014) The pattern of xylan acetylation suggests xylan may interact with cellulose microfibrils as a twofold helical screw in the secondary plant cell wall of *Arabidopsis thaliana*. Plant J **79:** 492–506

Busse-Wicher M, Grantham NJ, Lyczakowski JJ, Nikolovski N, Dupree P (2016a) Xylan decoration patterns and the plant secondary cell wall molecular architecture. Biochem Soc Trans **44:** 74–78

Busse-Wicher M, Li A, Silveira RL, Pereira CS, Tryfona T, Gomes TCF, Skaf MS, Dupree P (2016b) Evolution of xylan substitution patterns in gymnosperms and angiosperms: implications for xylan interaction with cellulose. Plant Physiol **171:** 2418–2431

Chong SL, Virkki L, Maaheimo H, Juvonen M, Derba-Maceluch M, Koutaniemi S, Roach M, Sundberg B, Tuomainen P, Mellerowicz EJ, et al (2014) O-Acetylation of glucuronoxylan in *Arabidopsis thaliana* wild type and its change in xylan biosynthesis mutants. Glycobiology **24:** 494–506

Ciucanu I, Kerek F (1984) A simple and rapid method for the permethylation of carbohydrates. Carbohydr Res **131:** 209–217

Cosgrove DJ (2014) Re-constructing our models of cellulose and primary cell wall assembly. Curr Opin Plant Biol **22:** 122–131

Cosgrove DJ, Jarvis MC (2012) Comparative structure and biomechanics of plant primary and secondary cell walls. Front Plant Sci **3:** 204

Danielsson S, Lindström ME (2005) Influence of birch xylan adsorption during kraft cooking on softwood pulp strength. Nordic Pulp & Paper Research Journal **20:** 436–441

Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. J Chem Phys **98:** 10089–10092

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

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

Domon B, Costello CE (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconj J **5:** 397–409

Escalante A, Gonçalves A, Bodin A, Stephan A, Sandström C, Toriz G, Gatenholm P (2012) Flexible oxygen barrier films from spruce xylan. Carbohydr Polym **87:** 2381–2387

Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. J Chem Phys **103:** 8577–8593

Fauré R, Courtin CM, Delcour JA, Dumon C, Faulds CB, Fincher GB, Fort S, Fry SC, Halila S, Kabel MA, et al (2009) A brief and informationally rich naming system for oligosaccharide motifs of heteroxylans found in plant cell walls. Aust J Chem **62:** 533–537

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

French AD, Johnson GP (2009) Cellulose and the twofold screw axis: modeling and experimental arguments. Cellulose **16:** 959–973

Goldberg RN, Schliesser J, Mittal A, Decker SR, Ana Filipa LOMS, Freitas VLS, Urbas A, Lang BE, Heiss C, Ribeiro da Silva MDMC, et al (2015) A thermodynamic investigation of the cellulose allomorphs: cellulose(am), cellulose Ib(cr), cellulose II(cr), and cellulose III(cr). J Chem Thermodyn **81:** 184–226

Hanus J, Mazeau K (2006) The xyloglucan-cellulose assembly at the atomic scale. Biopolymers **82:** 59–73

Hess B (2008) P-LINCS: a parallel linear constraint solver for molecular simulation. J Chem Theory Comput **4:** 116–122

Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput **4:** 435–447

Höije A, Sternemalm E, Heikkinen S, Tenkanen M, Gatenholm P (2008) Material properties of films from enzymatically tailored arabinoxylans. Biomacromolecules **9:** 2042–2047

Hoover WG (1985) Canonical dynamics: equilibrium phase-space distributions. Phys Rev A Gen Phys **31:** 1695–1697

Jacobs A, Larsson PT, Dahlman O (2001) Distribution of uronic acids in xylans from various species of soft- and hardwood as determined by MALDI mass spectrometry. Biomacromolecules **2:** 979–990

Jarzynski C (1997) Nonequilibrium equality for free energy differences. Phys Rev Lett **78:** 2690–2693

Jorgensen WL, Chandrasekhar J, Madura JD (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys **79:** 926–935

Kirschner KN, Yongye AB, Tschampel SM, González-Outeiriño J, Daniels CR, Foley BL, Woods RJ (2008) GLYCAM06: a generalizable biomolecular force field. Carbohydrates. J Comput Chem **29:** 622–655

Köhnke T, Östlund A, Brelid H (2011) Adsorption of arabinoxylan on cellulosic surfaces: influence of degree of substitution and substitution pattern on adsorption characteristics. Biomacromolecules **12:** 2633–2641

Köhnke T, Pujolras C, Roubroeks JP, Gatenholm P (2008) The effect of barley husk arabinoxylan adsorption on the properties of cellulose fibres. Cellulose **15:** 537–546

Langan P, Nishiyama Y, Chanzy H (2001) X-ray structure of mercerized cellulose II at 1 A resolution. Biomacromolecules **2:** 410–416

Larsson PT, Hult EL, Wickholm K, Pettersson E, Iversen T (1999) CP/MAS ^{13}C-NMR spectroscopy applied to structure and interaction studies on cellulose I. Solid State Nucl Magn Reson **15:** 31–40

Lawoko M, Henriksson G, Gellerstedt G (2005) Structural differences between the lignin-carbohydrate complexes present in wood and in chemical pulps. Biomacromolecules **6:** 3467–3473

Levy S, York WS, Stuike-Prill R, Meyer B, Staehelin LA (1991) Simulations of the static and dynamic molecular conformations of xyloglucan: the role of the fucosylated sidechain in surface-specific sidechain folding. Plant J **1:** 195–215

Littunen K, Kilpeläinen P, Junka K, Sipponen M, Master ER, Seppälä J (2015) Effect of xylan structure on reactivity in graft copolymerization and subsequent binding to cellulose. Biomacromolecules **16:** 1102–1111

Mazeau K, Charlier L (2012) The molecular basis of the adsorption of xylans on cellulose surface. Cellulose **19:** 337–349

Mazeau K, Moine C, Krausz P, Gloaguen V (2005) Conformational analysis of xylan chains. Carbohydr Res **340:** 2752–2760

McKee LS, Sunner H, Anasontzis GE, Toriz G, Gatenholm P, Bulone V, Vilaplana F, Olsson L (2016) A GH115 α-glucuronidase from Schizophyllum commune contributes to the synergistic enzymatic deconstruction of softwood glucuronoarabinoxylan. Biotechnol Biofuels **9:** 2

Mikkelsen D, Flanagan BM, Wilson SM, Bacic A, Gidley MJ (2015) Interactions of arabinoxylan and (1,3)(1,4)-β-glucan with cellulose networks. Biomacromolecules **16:** 1232–1239

Miyamoto S, Kollman PA (1992) SETTLE: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. J Comput Chem **13:** 952–962

Mortimer JC, Miles GP, Brown DM, Zhang Z, Segura MP, Weimar T, Yu X, Seffen KA, Stephens E, Turner SR, et al (2010) Absence of branches from xylan in Arabidopsis gux mutants reveals potential for simplification of lignocellulosic biomass. Proc Natl Acad Sci USA **107:** 17409–17414

Nieduszynski I, Marchessault RH (1971) Structure of beta-D-(1→4′)xylan hydrate. Nature **232:** 46–47

Nishiyama Y, Langan P, Chanzy H (2002) Crystal structure and hydrogen-bonding system in cellulose Ibeta from synchrotron x-ray and neutron fiber diffraction. J Am Chem Soc **124:** 9074–9082

Nosé S (1984) A molecular dynamics method for simulations in the canonical ensemble. Mol Phys **52:** 255–268

Parrinello M, Rahman A (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. J Appl Phys **52:** 7182–7190

Pell G, Taylor EJ, Gloster TM, Turkenburg JP, Fontes CMGA, Ferreira LMA, Nagy T, Clark SJ, Davies GJ, Gilbert HJ (2004) The mechanisms by which family 10 glycoside hydrolases bind decorated substrates. J Biol Chem **279:** 9597–9605

Pereira CS, Silveira RL, Dupree P, Skaf MS (2017) Effects of xylan side-chain substitutions on xylan-cellulose interactions and implications for thermal pretreatment of cellulosic biomass. Biomacromolecules **18:** 1311–1321

Pitkänen L, Virkki L, Tenkanen M, Tuomainen P (2009) Comprehensive multidetector HPSEC study on solution properties of cereal arabinoxylans in aqueous and DMSO solutions. Biomacromolecules **10:** 1962–1969

Pollet A, Delcour JA, Courtin CM (2010) Structural determinants of the substrate specificities of xylanases from different glycoside hydrolase families. Crit Rev Biotechnol **30:** 176–191

Rennie EA, Scheller HV (2014) Xylan biosynthesis. Curr Opin Biotechnol **26:** 100–107

Ruthes AC, Martínez-Abad A, Tan HT, Bulone V, Vilaplana F (2017) Sequential fractionation of feruloylated hemicelluloses and oligosaccharides from wheat bran using subcritical water and xylanolytic enzymes. Green Chem **19:** 1919–1931

Ryden P, Sugimoto-Shirasu K, Smith AC, Findlay K, Reiter WD, McCann MC (2003) Tensile properties of Arabidopsis cell walls depend on both a xyloglucan cross-linked microfibrillar network and rhamnogalacturonan II-borate complexes. Plant Physiol **132:** 1033–1040

Salmén L (2004) Micromechanical understanding of the cell-wall structure. C R Biol **327:** 873–880

Shimizu H, Hashi M, Sakurai K (1978) Isolation from a softwood xylan of oligosaccharides containing two 4-O-methyl-D-glucuronic acid residues. Carbohydr Res **62:** 117–126

Silveira JA, Fort KL, Kim D, Servage KA, Pierson NA, Clemmer DE, Russell DH (2013) From solution to the gas phase: stepwise dehydration and kinetic trapping of substance P reveals the origin of peptide conformations. J Am Chem Soc **135:** 19147–19153

Simmons TJ, Mortimer JC, Bernardinelli OD, Pöppler AC, Brown SP, deAzevedo ER, Dupree R, Dupree P (2016) Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. Nat Commun **7:** 13902

Sternemalm E, Höije A, Gatenholm P (2008) Effect of arabinose substitution on the material properties of arabinoxylan films. Carbohydr Res **343:** 753–757

St John FJ, Hurlbert JC, Rice JD, Preston JF, Pozharski E (2011) Ligand bound structures of a glycosyl hydrolase family 30 glucuronoxylan xylanohydrolase. J Mol Biol **407:** 92–109

Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. Chem Phys Lett **314:** 141–151

Teleman A, Larsson PT, Iversen T (2001) On the accessibility and structure of xylan in birch kraft pulp. Cellulose **8:** 209–215

Vardakou M, Dumon C, Murray JW, Christakopoulos P, Weiner DP, Juge N, Lewis RJ, Gilbert HJ, Flint JE (2008) Understanding the structural basis for substrate and inhibitor recognition in eukaryotic GH11 xylanases. J Mol Biol **375:** 1293–1305

Wang T, Park YB, Cosgrove DJ, Hong M (2015) Cellulose-pectin spatial contacts are inherent to never-dried *Arabidopsis thaliana* primary cell walls: evidence from solid-state nuclear magnetic resonance. Plant Physiol **168:** 871–884

White PB, Wang T, Park YB, Cosgrove DJ, Hong M (2014) Water-polysaccharide interactions in the primary cell wall of Arabidopsis thaliana from polarization transfer solid-state NMR. J Am Chem Soc **136:** 10399–10409

Whitney SE, Gothard MGE, Mitchell JT, Gidley MJ (1999) Roles of cellulose and xyloglucan in determining the mechanical properties of primary plant cell walls. Plant Physiol **121:** 657–664

Whitney SE, Wilson E, Webster J, Bacic A, Reid JS, Gidley MJ (2006) Effects of structural variation in xyloglucan polymers on interactions with bacterial cellulose. Am J Bot **93:** 1402–1414

Willför S, Pranovich A, Tamminen T, Puls J, Laine C, Suurnäkki A, Saake B, Uotila K, Simolin H, Hemming J, et al (2009) Carbohydrate analysis of plant materials with uronic acid-containing polysaccharides: a comparison between different hydrolysis and subsequent chromatographic analytical techniques. Ind Crops Prod **29:** 571–580

Yamasaki T, Enomoto A, Kato A, Ishii T, Shimizu K (2011) Structural unit of xylans from sugi (*Cryptomeria japonica*) and hinoki (*Chamaecyparis obtuse*). J Wood Sci **57:** 76–84

Zhang N, Li S, Xiong L, Hong Y, Chen Y (2015) Cellulose-hemicellulose interaction in wood secondary cell-wall. Model Simul Mater Sci Eng **23:** 085010

Zhao Z, Crespi VH, Kubicki JD, Cosgrove DJ, Zhong L (2014) Molecular dynamics simulation study of xyloglucan adsorption on cellulose surfaces: effects of surface hydrophobicity and side-chain variation. Cellulose **21:** 1025–1039