Received Date : 14-Apr-2016

Revised Date : 16-Jun-2016

Accepted Date : 28-Jun-2016

Article type : Original Article

A molecular dynamics study of the effect of glycosidic linkage type in the hemicellulose backbone on the molecular chain flexibility

Jennie Berglund[†], Thibault Angles d'Ortoli[‡], Francisco Vilaplana[†§], Göran Widmalm[‡], Malin Bergenstråhle-Wohlert[†], Martin Lawoko[†], Gunnar Henriksson[†], Mikael Lindström[†] and Jakob Wohlert[*†].

[†]*Wallenberg Wood Science Centre (WWSC), Department of Fiber and Polymer Technology, School of Chemical Engineering, Royal Institute of Technology KTH, SE-100 44 Stockholm, Sweden*

[‡]*Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, SE-106 91 Stockholm, Sweden*

[§]*Division of Glycoscience, School of Biotechnology, Royal Institute of Technology KTH, AlbaNova University Centre, SE-106 91 Stockholm, Sweden*

jebergl@kth.se; thibault.angles.dortoli@su.se; franvila@kth.se; goran.widmalm@su.se; malbe@kth.se; lawoko@kth.se; ghenrik@kth.se; mil@kth.se; jacke@kth.se*

*Corresponding author

Wallenberg Wood Science Centre, Teknikringen 56-58, SE-100 44 Stockholm, Sweden

Fax: +46 8790 8066

Running head: Molecular dynamics of glycosidic linkages

Keywords: Glucomannan, Xylan, Xyloglucan, Computer simulation, NMR spectroscopy.

SUMMARY

The macromolecular conformation of the constituent polysaccharides in lignocellulosic biomass influences their supramolecular interactions, and therefore their function in plants and their performance in technical products. The flexibility of glycosidic linkages from the backbone of hemicelluloses was studied by evaluating the conformational freedom of the φ and ψ dihedral angles using molecular dynamic simulations, additionally selected molecules were correlated with experimental data by NMR spectroscopy. Three types of β-(1→4) glycosidic linkages involving the monosaccharides (Glc*p*, Xyl*p* and Man*p*) present in the backbone of hemicelluloses were defined. Different di- and tetrasaccharides with combinations of such sugar monomers from hemicelluloses were simulated and free energy maps of the φ - ψ space and hydrogen bonding patterns were obtained. The glycosidic linkage between Glc-Glc or Glc-Man (C-type) was the stiffest with mainly one probable conformation; the linkage from Man-Man or Man-Glc (M-type) was similar but with an increased probability for an alternative conformation making it more flexible, and the linkage between two Xyl-units (X-type) was the most flexible with two almost equally populated conformations. Glycosidic linkages of the same type showed essentially the same conformational space in both disaccharides and in the central region of

tetrasaccharides. Different probabilities of glycosidic linkage conformations in the backbone of hemicelluloses can be directly estimated from the free energy maps, which to a large degree affect the overall macromolecular conformations of these polymers. The information gained contributes to an increased understanding of hemicelluloses' function both in the cell wall and in technical products.

INTRODUCTION

Plant cell walls are complex and dynamic polymeric networks, which confer not only structural stability to the cells, but also flexibility, permeability, signaling, and regulatory functions (Carpita and Gibeaut, 1993; Carpita and McCann, 2000). Primary plant cell walls are synthesized during cell growth and therefore have a dynamic, plastic and highly hydrated architecture. Furthermore, primary cell walls consist of non-oriented cellulose microfibrils, which are mainly interlocked by a matrix of hemicellulose and pectin polysaccharides and structural proteins. On the other hand, secondary cell walls have a fully differentiated structure in multiple lamellae, providing strength and rigidity to the plant woody tissues after cell growth. Secondary cell walls have a larger content of oriented cellulose microfibrils, embedded in a matrix of hemicelluloses and polyphenolic lignins, which contribute to the increased rigidity and hydrophobicity (Cosgrove and Jarvis, 2012; Burgert and Keplinger, 2013).

Hemicelluloses represent about one third of plant cell walls and have as such great potential to be used for new materials and applications replacing petroleum-based products. Hemicelluloses can be a source to biofuel (Pauly and Keegstra, 2008), a variety of platform chemicals (Bozell and Petersen, 2010), functional materials in oxygen barrier films and coatings (Hansen and Plackett, 2008), stabilizers, thickeners, and nutritional prebiotics (Zhang, 2008). In order to

technically use hemicelluloses we need extensive information about their properties which also is of great importance for their biological function. This is still not fully understood for hemicelluloses and in this study the goal is to increase the knowledge about how the structure of these polysaccharides affects their properties and function in the cell wall.

Hemicelluloses are heterogeneous polysaccharides with a backbone built up by neutral hexoses or pentoses connected through β-(1→4) glycosidic linkages in equatorial configuration, similarly to cellulose (Figure 1a). However, unlike cellulose, hemicelluloses are decorated with different chemical (e.g. acetyl groups) and glycosyl (neutral sugars and uronic acids) side-chain modifications. Xyloglucan is the main hemicellulose in the primary cell wall of flowering plants, with a polymeric backbone of D-glucopyranosyl units (as in cellulose) (Scheller and Ulvskov, 2010). On the other hand, xylans and glucomannans constitute the main hemicelluloses in woody tissues. Xylans consist of a backbone of D-xylopyranosyl units, whereas the backbone in glucomannans includes both D-mannopyranosyl and D-glucopyranosyl units (Timell, 1967; Scheller and Ulvskov, 2010). The nature, degree and pattern of substitutions differ not only amongst plant species, but also between tissues and developmental stages (Scheller and Ulvskov, 2010). When subjected to technical processing, native hemicelluloses in lignocellulosic biomass usually lose side chain saccharide groups and substituents, which results in more linear molecules with reduced molecular weight (Sjöström, 1993; Jacobs and Dahlman, 2001).

Apart from the presence of side-chain decorations, the differences in backbone conformation in hemicelluloses are of large importance for their biological and technical functionality. The most important factor for the macromolecular conformation of hemicelluloses is the flexibility of the backbone, which to a large extent is determined by the conformational freedom of the glycosidic linkages (Widmalm, 2013). The conformation of the glycosidic linkage depends on

the sugar residues present on each side of the linkage and is commonly expressed by two dihedral angles denoted φ and ψ, which refer to the rotations around the two covalent C-O bonds that form each glycosidic linkage (Figure 1c). There is typically an equilibrium between a few stable and meta-stable conformations, which are primarily stabilized by stereoelectronic and steric interactions, but could potentially also be influenced by hydrogen bonding between sugar residues.

The present study aims to evaluate how the conformational properties of the β-(1→4) glycosidic linkages between backbone sugar units vary in different hemicelluloses. We have identified three different types of linkages that are found in the backbone of hemicelluloses, each of them with a specific hydrogen bonding pattern (Figure 1b). The type C linkage involves a Glc*p* entity as the sugar containing the anomeric carbon in the glycosidic linkage and can be stabilized by hydrogen bonds on two sides of the glycosidic linkage simultaneously. Type M also has two possible hydrogen bonds; however, a change of conformation is expected to be necessary to change between hydrogen bonding possibilities (HB1 and HB2), due to the presence of a Man*p* unit as the sugar containing the anomeric carbon. Finally, the type X linkage can be found between Xyl*p* units and can only be stabilized by one hydrogen bond.

Molecular Dynamics (MD) simulations have been used previously to study conformational properties of glycosidic linkages in oligosaccharides (Campen et al., 2007; Peric-Hassler et al., 2010; Zaccheus et al., 2012; Wang et al., 2013) although hemicelluloses were not specifically targeted. Here, we use MD to study different disaccharides and tetrasaccharides mimicking the possible sugar combinations from the backbone of glucomannan, xylan and xyloglucan. The simulations are compared to effective proton-proton distances and transglycosidic *J* coupling constants determined by NMR experiments. From the simulations, the free energy space of the

glycosidic linkages is calculated to assess their conformational freedom, which is further discussed in terms of properties of plant cell wall polysaccharides from a technical and a biological perspective.

RESULTS AND DISCUSSION

As hemicelluloses are large polymeric macromolecules, they are not computationally effective to simulate. Therefore, small model systems (Table 1) are more appropriate for our interest in specific details of the glycosidic linkage conformation and hydrogen bonding.

The simulated structures were β-cellobiose (present in xyloglucan and glucomannan, but also cellulose), β-mannobiose (glucomannan), β-xylobiose (xylan) and β-(1→4)-linked D-glucose and D-mannose, one with glucose and one with mannose as the reducing end (glucomannan). Model tetrasaccharides were studied as well, with main focus on the central glycosidic linkage. These were β-mannotetraose, β-cellotetraose, β-xylotetraose and three combinations of tetraoses consisting of β-(1→4)-linked D-glucose and D-mannose, always with mannose as reducing and non-reducing end groups and the two middle sugar units were different combinations of glucose and mannose. Due to the interest in hydrogen bonding possibilities, allobiose and allotetraose, which have an axial hydroxyl group at the C3 position, were also simulated. Moreover, α-cellobiose was modelled for comparison with NMR data. In Table 1 information about all simulated saccharides is compiled.

Analysis of disaccharide conformations

The free energy maps of the dihedral angles φ and ψ for the three types of glycosidic linkages are presented in Figure 2. The two-letter codes refer to the model compounds listed in Table 1. Each

map is divided into four main regions: 1, 2, 3 and 4. Additionally, region 3 is divided into two parts where $3^{(+)}$ is the non-*exo*-anomeric region (Asensio et al., 2000) and $3^{(-)}$ corresponds to anti-φ conformation (Landersjö et al., 1997). Region 1 is the anti-ψ region (Dabrowski et al., 1995), and the minimum in region 4 is referred to as syn conformation (Dixon et al., 2003). For XX, region 4 is divided into two parts: $4^{(+)}$ and $4^{(-)}$, where the latter contains syn. A larger populated area in the energy map represents a higher flexibility of the glycosidic linkages. As shown in Table 1, GG and GM can be classified as being of glycosidic linkage type C, MM and MG of type M and XX of type X. Furthermore, data for the total probabilities of (φ,ψ) being found in each of the four regions are presented in Table 2. Standard errors were obtained from block averaging the data over 10 blocks of 5 ns each.

As expected, there was no significant difference between the free energy maps of structures with the same glycosidic linkage type. The free energy map for α-GG can be seen in Figure S1 (supporting information). The three different types of glycosidic linkages show similar topologies, albeit with some local differences. Region 4 is by far the most populated for all disaccharides investigated with a probability ranging between 93.1 to 97.4%, followed by region 3 (1.7 to 5.6%). The conformational space of both GG and MG are in agreement with earlier work using the GROMOS 45A4 force field (Peric-Hassler et al., 2010; Wang et al., 2013). The difference in probability between region 3 and 4 is large, which is also reflected in the difference between the corresponding free energies, which are 6.4 to 9.4 kJ mol^{-1} (the lowest value corresponds to type M and the highest to type X). Free energy maps of glycosidic linkages of type C and M are the most similar, but the minimum in $3^{(+)}$ is more populated for type M compared to type C. Type X shows two almost equally populated minima in region 4, and the

minimum in $3^{(+)}$ is not as well-defined and populated for XX as for the other disaccharides. This is not surprising since xylose, in contrast to the other studied sugar units, is a pentose.

The AA glycosidic linkage does not correspond to any of the hypothesized glycosidic linkage types and does not exist in hemicelluloses. This disaccharide was evaluated because of the different orientation of the hydroxyl group on the C3 position which is expected to hinder the formation of HB1 (as defined in Figure 1b). Despite this, the free energy map is quite similar to type C and type M, although a slightly larger flexibility in ψ is allowed compared to both C and M. Moreover, for AA the anti-φ region ($3^{(-)}$) is more populated than the anti-ψ (region 1), which is the opposite compared to the other disaccharides.

Assessment of the simulated systems by NMR experiments

Both *J* couplings and effective proton-proton distances can readily be calculated from an MD trajectory. The *J* couplings from the MD simulations were calculated according to corresponding equations in Säwén et al. (2010) for the disaccharides that were compared to NMR data. The *J* couplings from both literature and the NMR analysis are comparable to the ones obtained from the simulations; see the $^{3}J_{CH}$ comparison in Table 3, and $^{3}J_{CC}$ values in Table S1.

The distances between protons of the sugar residues of a glycosidic linkage are dependent on conformational preferences such as syn- and anti-conformations with respect to the dihedral angles φ and ψ. In GG severe spectral overlap of H3, H4 and H5 of the reducing end glycosyl residue significantly hampers determination of the H1'-H4 distance across the glycosidic linkage. It has, however, been possible to determine this distance in methyl β-cellobioside (Me GG) using a site-specifically deuterated isotopomer analogue of the compound resulting in $r_{H1'H4}$ ≈ 2.2 Å. For α-GG it was possible to obtain the corresponding distance by a 1D STEP-NOESY

NMR experiment, in conjunction with a regular 1D NOESY experiment to acquire a reference cross-relaxation rate (Figure 3).

Table 4 shows the proton-proton distances from the NMR experiments. In both the α- and β-anomeric forms of GG, the interglycosidic distance H1'-H4 is slightly longer as deduced from the MD simulation than determined with NMR, possibly due to overestimation of populations in non-*exo*-anomeric or anti-ψ conformations. The corresponding simulated interatomic distance in XX is longer than for GG, which is supported by the experimental NMR data, albeit in this case the presence of antiperiplanar conformations for XX may be underestimated in the simulations. Most importantly though, the transglycosidic H1'-H4 distances are similar between NMR experiments and MD simulations for the different compounds tested, and, furthermore, simulations capture the differences between the C and X type of linkages.

Hydrogen bonding in disaccharides

Three possible hydrogen-bonding positions in the glycosidic linkage (HB1 = O5'···H-O3, HB2a = O2'-H'···O6, HB2b = O2'···H-O6) have been considered (Figure 1b) and their respective probabilities in regions 3 and 4 are shown in Figure 4. Representative structures showing the conformations and hydrogen bonding possibilities in each free energy minima for GG, MM and XX is shown in Figure 5. In region 4, ($4^{(-)}$ in the case of XX) HB1 is common for all disaccharides, except for AA. On the other hand, HB1 did not occur in region 3 and instead HB2 was present for all structures in this region. Molecules with a glycosidic linkage of type C (GG and GM) attain both HB1 and HB2 in region 4. The probability of the two hydrogen bonds to be present at the same time was 11% for GG. Since the individual probabilities of HB1 and HB2 were 54% and 18% respectively, and their product becomes 10%, this indicates that, in region 4,

the two hydrogen bonds form independently of each other. Type C linkages rarely visit region 3, but when it does, the probability for HB2 is about 20-30%. Type M linkages (MM and MG) only attain HB1 in region 4 while they exhibit a significantly higher probability of HB2 in region 3 than the other disaccharides. Thus, these molecules may form hydrogen bonds of both type HB1 and HB2, but not at the same time as a consequence of having the hydroxyl group on C2' axially oriented. Finally, AA attains fewer hydrogen bonds across the glycosidic linkage compared to the other structures, as expected.

The hydrogen bonding follows the same pattern in all simulated glycosidic linkages of the same type. Data on hydrogen bonding across all glycosidic linkages in both di- and tetrasaccharides can be found in Table S2.

Ring conformations

An analysis of the pyranose ring conformations was conducted by calculating the pseudo-dihedral angles α_1 (C4-O5-C2-C1), α_2 (O5-C2-C4-C3) and α_3 (C2-C4-O5-C5) (Pickett and Strauss, 1970; Rao et al., 1998). This showed that the main conformation for all glucose and mannose units was the $^4C_1$ chair, which is as expected (Angyal, 1968). For the xylose units in XX the $^4C_1$ chair dominated although the presence of inverted chair conformations, $^1C_4$, was significant; 12.7% for the reducing sugar and 18.3% for the non-reducing sugar, which, assuming a two-state model, gives an energy difference between the two conformers of 4.7 kJ mol^{-1} and 3.7 kJ mol^{-1} respectively.

The energy difference between $^4C_1$ and $^1C_4$ in aqueous solution is significantly lower in xylose than in hexopyranoses, however, not as low as what was found here. Literature values end up between 9 to 10 kJ mol^{-1} (Angyal, 1968; Mayes et al., 2014), which indicate a population in $^1C_4$

of 2-3%. In fact, this correlates well with experimental NMR $^3J_{HH}$ coupling constants in the xylopyranose residues of XX (Figure S2, Table S4), which become nearly identical to the ones previously determined for 2-naphthyl β-D-xylopyranoside (Siegbahn et al., 2011), showing that the $^4C_1$ conformer is by far the dominating one (> 90%). The fact that the energy difference between $^4C_1$ and the inverted chair conformation $^1C_4$ is too low appears to be a shortcoming of the present force field. At present, there are no published studies that mention this, and further investigations into this matter are probably called for.

The ring conformation may, in some cases, affect the glycosidic conformation (Plazinski and Drach, 2015; Plazinski et al., 2016a; Plazinski et al., 2016b). To investigate this, the free energy map of xylobiose was decomposed into four parts, each of one corresponding to a different combination of ring conformations (Figure 6). Since the locations and relative depths of the minima were the same in all cases, we conclude that the ring conformation has only a negligible effect on the glycosidic linkage.

Analysis of tetrasaccharides

Figure 7 and Table 2 display the free energy maps and population statistics of the central glycosidic linkages in tetrasaccharides, which show essentially the same behavior as their corresponding disaccharide. Also the glycosidic linkages connected to the reducing and non-reducing end of tetrasaccharides show similar population statistics (Table S3). Just as for the disaccharides, region 4 is the most populated, for all bond types, with probabilities between 91.1 to 98.5%. The differences in populations that can be observed compared to the disaccharides are that $3^{(-)}$ is even less populated for the central glycosidic linkage in the tetrasaccharides, and for

MMMM and GGGG the population in region 1 increased by roughly a factor of two. However, based on the standard errors in the populations, these differences are not statistically significant.

The ring conformations of the sugar units in the tetrasaccharides were generally similar to the disaccharides, and for all mannose and glucose units the main conformation was $^4C_1$. For the two outer xylose units in β-xylotetraose $^4C_1$ dominated but also a significant fraction of $^1C_4$ was observed. Moreover, $^1C_4$ even dominated in the two central sugar units of xylotetraose. As previously discussed for xylobiose the high population of the inverted chair for xylose units is likely an artifact of the present force field parameters. Nevertheless, the observation that $^1C_4$ increases in sugar residues that were part of a larger molecule was observed in a previous study (Plazinski et al., 2016b) where it was suggested that this was caused by steric interactions between residues.

The hydrogen bonding over the glycosidic linkages in the tetrasaccharides (Figure 5) show essentially the same behavior as what is observed for the disaccharides of the same bond type. Still, one significant difference is that XX attained much more HB1 compared to XX-XX, but comparing with the outer glycosidic linkages in XXXX (Table S2) the amount of HB1 varies a great deal within the tetrasaccharides (23% in the reducing end, 7% in the middle, and 11% in the non-reducing end).

The role of environment for intramolecular hydrogen bonding

Wood in growing trees with a typical moisture content of 50% contains approximately one third of water, one third of air, and one third solid material. It is not possible to simulate the saccharides in an environment that accurately replicates the natural environment for hemicelluloses; still, given the large fraction of water in plant cell walls, it is not unreasonable to

instead consider aqueous solution. Another obvious advantage from using aqueous systems is that simulations can be readily validated using solution state NMR.

Previous simulations of saccharides in water show only negligible effect from hydrogen bonding on the conformation of glycosidic linkages (Peric-Hassler et al., 2010; Wang et al., 2013). This is due to that when an intramolecular hydrogen bond is broken, it can be immediately replaced by a hydrogen bond to a water molecule, which, from an energetic point of view, is just as favorable. This observation is supported by our data by the fact that allobiose and allotetraose, which has only limited possibility of forming hydrogen bonds of type HB1, show almost identical free energy maps for their glycosidic linkages as for linkages of both type C and M (Figure 2 and Figure S1). However, in another study where disaccharides were simulated in vacuum it was clear that hydrogen bonding caused a stabilization of the glycosidic linkages (Campen et al., 2007). In that study the free energy difference between region 4 and region 1 for α-cellobiose was calculated to 14.4 kJ mol^{-1}. Here, the corresponding free energy difference was 10.2 kJ mol^{-1}, i.e., significantly reduced. It is important however to point out that different force fields were used in those two cases.

One common plant polysaccharide that only contains glycosidic linkages of type C is cellulose, which in is predominantly found as crystalline or semi-crystalline aggregates. Although their assembly is driven by hydrophobic forces (Bergenstråhle et al., 2010) the crystal structure is determined by an intricate network of intra- and intermolecular hydrogen bonds (Nishiyama et al., 2002). Removal of these bonds results in complete rupture of the crystal structure, to a large extent accompanied by relaxation of the glycosidic linkages (Wohlert et al., 2012). This shows that in a water free environment, hydrogen bonds will indeed stabilize glycosidic linkage conformation, and even collectively force them into energetically strained conformations.

It seems obvious that hydrogen bonding is highly dependent on the environment and a hemicellulose molecularly dissolved in water is probably less restricted by intramolecular hydrogen bonds than a hemicellulose that is entrapped in cell wall assemblies with limited humidity.

Conformations of polysaccharide chains

To get an idea of what effect the glycosidic dihedral angles would have on the conformation of a longer hemicellulose chain, structures containing ten sugar residues with φ and ψ at their most probable values were generated (Figure 8, upper panel). The resulting xylan, mannan and glucan chains all formed elongated, twisted structures. The sum $\varphi + \psi$ can be used to characterize the conformation of a chain (Mazeau et al., 2005; French and Johnson, 2009), with, in the IUPAC definition, $\varphi + \psi = 300°/420°$ being indicative of a right/left-handed, 3_1-fold helical screw, and $\varphi + \psi = 360°$ giving a 2_1-fold helical screw.

The most favored conformation for type C and M is found in region 4 with $\varphi + \psi \approx 400°$, corresponding to an elongated screw, which can be seen in both the glucan and mannan chains (Figure 8).

Linkages of type X are found with the highest probability in region $4^{(+)}$, closely followed by region $4^{(-)}$ which have $\varphi + \psi \approx 440°$ and $\varphi + \psi \approx 380°$ respectively. The later results in an elongated screw, similar to a 2_1-fold helical screw, while the conformation from region $4^{(-)}$ forms a screw more like a left-handed 3_1-fold helix. The energy barrier between these two states ($4^{(+)}/4^{(-)}$) is only about 1.5 kJ mol^{-1} which indicates that the conformation of xylan molecules in solution is reversible and that these conformations can be interchangeable depending on surrounding environment.

Another interesting observation is that the glycosidic linkage conformation from region 1 for all glycosidic linkage types generates a bend of approximately 90°, as well as the uncommon region $3^{(-)}$ (Figure 8, lower panel). The probability for any of these bends is generally low and around 1% for studied structures (Table 2). However, hemicellulose chains usually have a DP > 100 and therefore these bends are of statistical relevance since only one bend creates a major change of direction for the polymer. An inverted chair conformation is also expected to have a similar effect on the direction of the molecules. Although extremely rare in both glucose and mannose, it occurs in xylose with a non-negligible probability as discussed previously, and may thus be of significant importance for the macromolecular conformation of xylan.

Significance of backbone flexibility for the ultrastructure of lignocellulosic biomass and properties of technical hemicelluloses

The previously described conformations of the glycosidic linkage (Figure 8) could explain why hemicelluloses can interact with cellulose crystals, i.e., by hydrogen bonding of polysaccharide hydroxyl groups and van der Waals interactions between the hemicellulose and cellulose, and at the same time act as a cross linker between polymers through covalent and non-covalent interactions in lignocellulosic biomass (Somerville et al., 2004; Cosgrove and Jarvis, 2012; Salmén, 2015). From our modeling results for the hemicellulose backbones with different glycosidic linkage types, preferred elongated chain configurations can be adopted that will favor interactions with cellulose microfibril surfaces. In addition, stiffer chains will also suffer smaller entropic penalties from the restrictions on chain mobility inferred by an interface, which also will favor adsorption to cellulose. Different statistically significant glycosidic linkage configurations can co-exist in the same hemicellulosic chain, which enable intramolecular regions with long

elongated screw conformations that could interact with the surfaces of cellulose microfibrils, together with specific local bends that would change the orientation of the hemicellulose chains. This observation is in agreement with the structural models established for xyloglucan in primary cell walls (Pauly et al., 1999; Cosgrove, 2005; Park and Cosgrove, 2012), where an individual xyloglucan molecule can act as a polymeric anchor to bind two adjacent cellulose microfibrils. A similar behavior was reported for xylan (Busse-Wicher et al., 2014), where interactions with both cellulose and other cell wall components were hypothesized to occur through combination of 2_1-fold and 3_1-fold helical screws, respectively. As previously stated the energy needed for transition between these two xylan conformations is rather low, so interchangeable configurations could be attained depending on the environment.

Although side chain substituents play a significant role in hemicelluloses, backbone configuration can be envisaged as the vertebral spine that modulates the main associative events between hemicelluloses and other lignocellulosic polymers. Xyloglucans have a large diversity of side chains, which may be involved in modulating the interactions with other cell wall components and with signaling events. It has been shown that both the backbone and side chain groups in xyloglucan are contributing to the interaction with cellulose microfibrils (Zhang et al., 2011). Side substituents can modulate xyloglucan backbone conformation from a putative twisted to planar form which influences the supramolecular interaction with cellulose microfibrils (Levy et al., 1991; Levy et al., 1997). Other similar MD studies pointed out that xyloglucans with smaller side chains, (Figure 1a) which have a larger contribution of the backbone conformation on the overall macromolecular conformation, can adopt flat local orientations interacting better with cellulose microfibrils (Hanus and Mazeau, 2006; Zhao et al., 2014).

The reason for the diversity in hemicellulose composition in different plants is not clear, although previous works suggested that the glucomannans might interact closer with cellulose in wood than xylans do (Åkerholm and Salmén, 2001). The present study has shown that glycosidic linkages of type M exhibit similar rigidity as type C, whereas type X exhibits more flexibility. This could in part explain why glucomannans should interact better with cellulose than xylan, since the chain of glucomannan should be stiffer and therefore easier to adsorb to cellulose. Furthermore, comparing linkage type C with type M the latter shows an increased population of non-*exo*-anomeric conformation and is therefore more flexible, which indicates that a higher content of glucose in glucomannans should give a higher stiffness of the chain. Here it is interesting to note that glucomannan in hardwoods generally has a higher content of glucose than in softwoods, but they occur in much lower amounts (Timell, 1967). The higher glucose content in hardwood glucomannans may give a stronger affinity to cellulose, and therefore it might be sufficient with lower amounts to provide the crosslinking support required in such tissues.

The observation that glucomannans interacts better with cellulose than xylan has also been proposed in the literature for technical wood (Clayton and Phelps, 1965). Wood processing (kraft/soda pulping, hot water extractions and sulphite pulping) causes changes to the structure of hemicelluloses, resulting in "technical hemicelluloses" with an increased linearity, a reduced presence of side chain saccharide groups and substituents, a decreased molar mass, and increased hydroxyl functionality (Enkvist et al., 1957; Sjöström, 1977; Johansson and Samuelson, 1977; Sjöström, 1993). The technically modified hemicelluloses (suggested structures are presented in Figure S3) thus resemble the models in this study from a view-point of hydrogen bonding possibilities and linearity, although the molar masses differ significantly. In fact, in earlier works, ivory nut mannan has been used as a model to study the recrystallization phenomena of

glucomannan onto cellulose (Chanzy et al., 1978). Such recrystallizations are well known to enhance pulp yields, and an increased fraction of glucomannan has also been shown to increase the pulp's strength properties (Salmen and Lindström, 2015). Furthermore, differences in flexibility between technical glucomannans and xylans may influence their potential in different types of applications of purified or partly purified hemicelluloses, as in barrier films or composites (Hartman et al., 2006; Escalante et al., 2012; Oinonen et al., 2013).

In conclusion

The present study uses MD computer simulations of selected di- and tetrasaccharides in aqueous solution to show that the conformational flexibility of the backbone in hemicelluloses to a large degree depend on the glycosidic linkages, which can be categorized as belonging to three different types depending on its monomeric neighbors. The most flexible one is type X (found in xylan) with two almost equally probable conformations. Type C (found in cellulose, glucomannan, and xyloglucan) is, as expected, the stiffest with primarily syn conformation while type M (found in glucomannan) is similar to type C but with an increased population of non-*exo*-anomeric conformation, which makes it more flexible. The difference between these three glycosidic linkage types can be connected to properties of the hemicellulose backbone. The results partially explains how the probable conformations of the backbone glycosidic linkages make it possible for hemicelluloses to interact with different types of wood polymers and act as a cross-linker in the cell wall. Our findings that glucomannan has a stiffer backbone than xylan can also explain why the interaction between glucomannan and cellulose are stronger than the interaction between xylan and cellulose (Clayton and Phelps, 1965; Åkerholm and Salmén, 2001) since the stiffer structure allows stronger binding to cellulose by hydrogen bonding and

van der Waals interactions, and from a smaller decrease in chain entropy upon adsorption. The NMR analysis corroborates the simulated systems, with one exception; xylose residues attain a larger fraction of the $^{1}C_4$ chair during the simulation compared to the real systems. This did not however affect the conformational space of the glycosidic linkage.

The effect of hydrogen bonding for stabilization of the glycosidic linkage conformation is limited in water solution, but may be of importance in the more complex environment of the cell wall.

The information gained about the flexibility of backbone glycosidic linkages in hemicelluloses contributes to an increased understanding of their function both in the cell wall and in technical products. In this context, disaccharides are judged to be good models for studying the behavior of a single glycosidic linkage also in polymers since no significant differences were observed between the conformational spaces of a glycosidic linkage in a disaccharide and the central glycosidic linkage in a tetrasaccharide. This means that the results presented here constitutes a good model for future simulations of polymer chains using e.g. coarse grain methods.

EXPERIMENTAL PROCEDURE

Molecular dynamics (MD) simulations

MD simulations of the molecules presented in Table 1 were performed using GROMACS 5.0.4 (Hess et al., 2008; Abraham et al., 2015) employing the GLYCAM06 (Kirschner et al., 2008) force field for carbohydrates and the TIP3P (Jorgensen et al., 1983) model for water. Molecules were built using the leap module in AMBER (Case et al., 2015); thereafter the topologies were converted to GROMACS format using ACPYPE (Sousa da Silva and Vranken, 2012). Disaccharides and tetrasaccharides were all simulated as a single molecule in aqueous solution

using a fully periodical box. For the disaccharides a 3x3x3 nm^3 box containing about 860 water molecules was used. For tetrasaccharides a dodecahedral box containing 1560 water molecules was employed.

Each run was carried out for 50 ns, at 1 atm pressure, and at a temperature of 300 K, using a *leap-frog* integrator with a 2 fs basic time step. The pressure was held constant by the Parrinello-Rahman barostat (Parrinello and Rahman, 1981) and the temperature was controlled by the Nosé-Hoover algorithm (Nosé, 1984; Hoover, 1985). Replica Exchange Molecular Dynamics (Sugita and Okamoto, 1999) was used at 12 different evenly spaced temperatures ranging from 300 K to 366 K (in 6 K increments) with attempts to exchange between neighboring replicas at every 10 steps for improved sampling.

Lennard-Jones interactions used a straight cutoff at 1.2 nm. For electrostatic forces the Particle Mesh Ewald (PME) summation (Darden et al., 1993; Essmann et al., 1995) with a 1.2 nm real space cutoff was applied. To keep bonds at their equilibrium values P-LINCS (Hess, 2008) was used for the saccharides and SETTLE (Miyamoto and Kollman, 1992) for water. According to the GLYCAM convention, no scaling was applied for 1-4 interactions.

Free energy maps of the (φ,ψ) dihedral space were constructed by recording the φ and ψ dihedral angles every ps in the 50 ns simulation and calculating their probabilities, which were converted to free energies by Boltzmann inversion. The contour plots were constructed with a grid spacing of 5° and with contour levels at every 1.5 kJ mol^{-1}. The maximum energy represents conformations that never are attained, and this maximum level was set to 19.5 kJ mol^{-1}, but represents energies above 19.5 kJ mol^{-1}. The hydrogen bonds were analyzed with the *gmx hbond* tool in GROMACS and the default geometrical criteria: the distance between the donor and acceptor $r \leq 0.35$ nm and the hydrogen-donor-acceptor angle $\leq 30°$.

Molecular graphics were made with VMD (Visual Molecular Dynamics) (Humphrey et al., 1996).

NMR spectroscopy

Disaccharide samples were xylobiose obtained from Carbosynth, Berkshire, UK; mannobiose from Megazyme, Wicklow, Ireland; and cellobiose from Aldrich, Beerse, Belgium. The samples were prepared in 5 mm NMR tubes in D_2O, pD ≈ 6, with a sample concentration of 84 – 92 mM. NMR experiments were carried out at 300 K on a 600 MHz Bruker AVANCE III spectrometer equipped with an inverse detection probe. The proton-proton cross-relaxation rate for the H1-H2 pair in α-cellobiose was measured using 1D ^{1}H,^{1}H-NOESY experiments (Stott et al., 1997) and that of H4-H1' employed a 1D ^{1}H,^{1}H-STEP-NOESY experiment (Hu et al., 2004; Pendrill et al., 2013). In all cases, selective excitation was achieved by single or double PFGSE modules utilizing 50 ms r-SNOB or i-SNOB-2 shaped pulses for the NOESY experiment and 50 – 65 ms r-SNOB shaped pulses for the STEP-NOESY experiment. The strengths of the first and second gradient pairs were 15% and 40%, respectively, of the maximum (55.7 G cm^{-1}) for the NOESY experiments. For the STEP-NOESY, the strengths of the gradients were set to 10% or 6.5% for the first and 45% or 15% for the second excitation, respectively. Due to the spectral overlap of the H4 resonance in α-cellobiose with those of H3 – H5 in β-cellobiose (Roslund et al., 2008), the resonance from H1 in the former was selectively excited and magnetization transferred to H4 using a 3.8 kHz DIPSI-2 spin lock with a duration of 140 ms prior to selective excitation of H4. In the STEP-NOESY experiments, zero-quantum coherences were suppressed using the scheme devised by (Thrippleton and Keeler, 2003) where a 30 – 50 ms adiabatic Chirp pulse with a bandwidth of 20 kHz was applied together with a gradient pulse with 3% of the maximum

power. In the NOESY experiment a 20 ms adiabatic Chirp pulse with a bandwidth of 40 kHz was used in combination with a gradient pulse at 6% of the maximum power. In the NOESY as well as the STEP-NOESY experiments, eight cross-relaxation delays between 80 – 400 ms were collected for each of the excited spins. A spectral width of 6 ppm was sampled using 16k data points and 512 transients were averaged. The repetition time was 8 – 10 s, i.e., in all cases longer than 5×T1. Prior to Fourier transformation, the FIDs of the 1D experiments were zero-filled to 262k points and multiplied by an exponential line-broadening function of 2 Hz. Baseline correction was performed prior to integration which used the same integration limits for all experiments within a series. The areas of relevant peaks were divided by the area of the inverted peak and least-square fitted to a first order function yielding the cross-relaxation rate constant.

The measurements of transglycosidic $^{3}J_{CH}$ coupling constants (Rundlöf et al., 1998) were carried out using one-dimensional long-range (1DLR) experiments essentially as devised by Nishida et al. (Nishida et al., 1995; Nishida et al., 1996) employing ^{13}C site-selective excitation with a Gaussian shaped pulse of 80 – 250 ms duration. The delay used for suppression of $^{1}J_{CH}$ was set to $(145\ Hz)^{-1}$ and the time of the delay between excitation and coherence transfer, for evolution of the long-range coupling, was set by using a nominal values of 8 – 16 Hz; an acquisition time of 3 s, 4096 – 6144 transients and 32 – 64k data points were used. Zero-filling was performed to 256k data points and an exponential line-broadening function, lb = 0.6 Hz, was employed. Subsequently, the $^{3}J_{CH}$ coupling constants were extracted by the *J* doubling methodology (del Río-Portilla et al., 1994) implemented in-house by a MATLAB script.

ACKNOWLEDGEMENT

This work was supported by the Knut and Alice Wallenberg's research foundation within the Wallenberg Wood Science Centre and by the Swedish Research Council. Computational resources were provided by the PDC Center for High Performance Computing through project PDC-2015-36.

Conflict of Interest

The authors do not have any conflicts of interest.

SUPPORTING INFORMATION

REFERENCES

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

Angyal, S.J. (1968) Conformational analysis in carbohydrate chemistry. I. Conformational free energies. The conformations and α : β ratios of aldopyranoses in aqueous solution. *Aust. J. Chem.*, 21, 2737-2746.

Asensio, J.L., Cañada, F.J., Cheng, X., Khan, N., Mootoo, D.R. and Jiménez-Barbero, J. (2000) Conformational Differences Between O- and C-Glycosides:The α-O-Man-(1→1)-β-Gal/α-C-Man-(1→1)-β-Gal Case- A Decisive Demonstration of the Importance of the *exo*-Anomeric Effect on the Conformation of Glycosides. *Chem. Eur. J.*, 6, 1035-1041.

Bell, N.G.A., Rigg, G., Masters, S., Bella, J. and Uhrín, D. (2013) Detecting low-level flexibility using residual dipolar couplings: a study of the conformation of cellobiose. *Phys. Chem. Chem. Phys.*, 15, 18223-18234.

Bergenstråhle, M., Wohlert, J., Himmel, M.E. and Brady, J.W. (2010) Simulation studies of the insolubility of cellulose. *Carbohydr. Res.*, 345, 2060-2066.

Bozell, J.J. and Petersen, G.R. (2010) Technology development for the production of biobased products from biorefinery carbohydrates - the US Department of Energy's "Top 10" revisited. *Green Chem.*, 12, 539-554.

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

Busse-Wicher, M., Gomes, T.C.F., Tryfona, T., Nikolovski, N., Stott, K., Grantham, N.J., Bolam, D.N., Skaf, M.S. and 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.

Campen, R.K., Verde, A.V. and Kubicki, J.D. (2007) Influence of glycosidic linkage neighbours on disaccharide conformation in vacuum. *J. Phys. Chem. B*, 111, 13775-13785.

Carpita, N.C. and Gibeaut, D.M. (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J*, 3, 1-30.

Carpita, N. and McCann, M. (2000) 'The cell wall', in *Biochemistry and Molecular Biology of Plants* (Buchanan, B.B., Gruissem, W. and Jones, R.L., eds) Rockville, MD: American Society of Plant Physiologists. pp. 52-108

Case, D.A., Berryman, J.T., Betz, R.M., Cerutti, D.S., Cheatham, T.E., Darden, T.A., Duke, R.E., Giese, T.J., Gohlke, H., Goetz, A.W., Homeyer, N., Izadi, S., Janowski, P., Kaus, J., Kovalenko, A., Lee, T.S., LeGrand, S., Li, P., Luchko, T., Luo, R. et al. (2015) *AMBER 2015*, University of California, San Francisco.

Chanzy, H., Dube, M. and Marchessault, R.H. (1978) Shish-kebab morphology Oriented recrystallization of mannan on cellulose. *Tappi J.*, 61, 81-82.

Clayton, D.W. and Phelps, G.R. (1965) The sorption of glucomannan and xylan on α-cellulose wood fibers. *J. Polym. Sci. Part C*, 11, 197-220.

Cosgrove, D.J. (2005) Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.*, 6, 850-861.

Cosgrove, D.J. and Jarvis, M.C. (2012) Comparative structure and biomechanics of plant primary and secondary cell walls. *Front. Plant Sci.*, 3, 206-211.

da Silva, A.W.S. and Vranken, W.F. (2012) ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Res. Notes*, 5, 367.

Dabrowski, J., Kozar, T., Grosskurth, H. and Nifant'ev, N.E. (1995) Conformational Mobility of Oligosaccharides: Experimental Evidence for the Existence of an "Anti" Conformer of the Gal.beta.1-3Glc.beta.1-OMe Disaccharide. *J. Am. Chem. Soc.*, 117, 5534-5539.

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

del Río-Portilla, F., Blechta, V. and Freeman, R. (1994) Measurement of Poorly Resolved Splittings by J Doubling in the Frequency Domain. *J. Magn. Reson. Ser. A*, 111, 132-135.

Dixon, A.M., Venable, R., Widmalm, G., Bull, T.E. and Pastor, R.W. (2003) Application of NMR, Molecular Simulation, and Hydrodynamics to Conformational Analysis of Trisaccharides. *Biopolymers*, 69, 448-460.

Ebringerová, A., Hromádková, Z. and Heinze, T. (2005) Hemicellulose. *Adv. Polym. Sci.*, 186, 1-67.

Enkvist, T., Alfredsson, B. and Martelin, J.-E. (1957) Determinations of the Consumption of Alkali and Sulfur at Various Stages of Sulfate, Soda and Alkaline and Neutral Sulfite Digestion of Spruce Wood. *Sven. Papperstidn.*, 60, 616-620.

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

Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H. and Pedersen, L.G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys.*, 103, 8577-8593.

French, A.D. and Johnson, G.P. (2009) Cellulose and the twofold screw axis: modeling and experimental arguments. *Cellulose*, 16, 959-973.

Hansen, N.M.L. and Plackett, D. (2008) Sustainable Films and Coatings from Hemicelluloses: A Review. *Biomacromolecules*, 9, 1493-1505.

Hanus, J. and Mazeau, K. (2006) The Xyloglucan-Cellulose Assembly at Atomic Scale. *Biopolymers*, 82, 59-73.

Hartman, J., Albertsson, A.-C. and Sjöberg, J. (2006) Surface- and Bulk-Modified Galactoglucomannan Hemicellulose Films and Film Laminates for Versatile Oxygen Barriers. *Biomacromolecules*, 7, 1983-1989.

Hatcher, E., Säwén, E., Widmalm, G. and MacKerell, Jr., A.D. (2011) Conformational Properties of Methyl β-Maltoside and Methyl α- and β-Cellobioside Disaccharides. *J. Phys. Chem. B*, 115, 597-608.

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. and Lindahl, E. (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, 4, 435-447.

Hoover, W.G. (1985) Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A*, 31, 1695-1697.

Hricovíni, M., Tvaroska, I. and Hirsch, J. (1990) Solution behaviour of methyl β-xylobioside: conformational flexibility revealed by n.m.r. measurements and theoretical calculations. *Carbohydr. Res.*, 198, 193-203.

Hu, H., Bradley, S.A. and Krishnamurthy, K. (2004) Extending the limits of the selective 1D NOESY experiment with an improved selective TOCSY edited preparation function. *J. Magn. Reson.*, 171, 201-206.

Hu, H. and Krishnamurthy, K. (2006) Revisiting the initial rate approximation in kinetic NOE measurements. *J. Magn. Reson.*, 182, 173-177.

Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, 14, 33-38.

Jacobs, A. and Dahlman, O. (2001) Characterization of the Molar Masses of Hemicelluloses from Wood and Pulps Employing Size Exclusion Chromatography and Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry. *Biomacromolecules*, 2, 894-905.

Johansson, M.H. and Samuelson, O. (1977) Alkaline destruction of birch xylan in the light of recent investigations of its structure. *Sven. Papperstidn.*, 80, 519-524.

Jorgensen, W.L., Chandrasekhar, J. and Madura, J.D. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79, 926-935.

Kirschner, K.N., Yongye, A.B., Tschampel, S.M., González-Outeiriño, J., Daniels, C.R., Foley, B.L. and Woods, R.J. (2008) GLYCAM06: A generalizable biomacromolecular force field. Carbohydrates. *J. Comput. Chem.* , 29, 622-655.

Landersjö, C., Stenutz, R. and Widmalm, G. (1997) Conformational Flexibility of Carbohydrates: A Folded Conformer at the Dihedral Angle of a Glycosidic Linkage. *J. Am. Chem. Soc.*, 119, 8695-8698.

Larsson, E.A., Staaf, M., Söderman, P., Höög, C. and Widmalm, G. (2004) Determination of the Conformational Flexibility of Methyl α-Cellobioside in Solution by NMR Spectroscopy and Molecular Simulations. *J. Phys. Chem. A*, 108, 3932-3937.

Levy, S., Maclachlan, G. and Staehelin, L.A. (1997) Xyloglucan sidechains modulate binding to cellulose during in vitro binding assays as predicted by conformational dynamics simulations. *Plant J.*, 11, 373-386.

Levy, S., York, W.S., Stuike-Prill, R., Meyer, B. and Staehelin, L.A. (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.

Macura, S., Farmer II, B.T. and Brown, L.R. (1986) An Improved Method for the Determination of Cross-Relaxation Rates from NOE Data. *J. Magn. Reson.*, 70, 493-499.

Mayes, H.B., Broadbelt, L.J. and Beckham, G.T. (2014) How sugars pucker: electronic structure calculations map the kinetic landscape of five biologically paramount monosaccharides and their implications for enzymatic catalysis. *J. Am. Chem. Soc.*, 136, 1008-1022.

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

Miyamoto, S. and Kollman, P.A. (1992) SETTLE: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.*, 13, 952-962.

Nishida, T., Widmalm, G. and Sándor, P. (1995) Hadamard long-range proton-carbon coupling constant measurements with band-selective proton decoupling. *Magn. Reson. Chem.*, 33, 596-599.

Nishida, T., Widmalm, G. and Sándor, P. (1996) Hadamard Long-Range Proton-Carbon Coupling Constant Measurements with Pulsed Field Gradients. *Magn. Reson. Chem.*, 34, 377-382.

Nishiyama, Y., Langan, P. and Chanzy, H. (2002) Crystal Structure and Hydrogen-Bonding System in Cellulose Iβ 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.

Oinonen, P., Areskogh, D. and Henriksson, G. (2013) Enzyme catalyzed cross-linking of spruce galactoglucomannan improves its applicability in barrier films. *Carbohydr. Polym.*, 95, 690-696.

Park, Y.B. and Cosgrove, D.J. (2012) A Revised Architecture of Primary Cell Walls Based on Biomechanical Changes Induced by Substrate-Specific Endoglucanases. *Plant Physiol.*, 158, 1933-1943.

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

Pauly, M., Albersheim, P., Darvill, A. and York, W.S. (1999) Molecular domains of the cellulose/xyloglucan network in the cell walls of higher plants. *Plant J.*, 20, 629-639.

Pauly, M. and Keegstra, K. (2008) Cell-wall carbohydrates and their modification as a resource for biofuels. *Plant J.*, 54, 559-568.

Pendrill, R., Säwén, E. and Widmalm, G. (2013) Conformation and Dynamics at a Flexible Glycosidic Linkage Revealed by NMR Spectroscopy and Molecular Dynamics Simulations: Analysis of β-L-Fucp-(1→6)-α-D-Glcp-OMe in Water Solution. *J. Phys. Chem. B*, 117, 14709-14722.

Peric-Hassler, L., Hansen, H.S., Baron, R. and Hünenberger, P.H. (2010) Conformational properties of glucose-based disaccharides investigated using molecular dynamics simulations with local elevation umbrella sampling. *Carbohydr. Res.*, 345, 1781-1801.

Pickett, H.M. and Strauss, H.L. (1970) Conformational Structure, Energy, and Inversion Rates of Cyclohexane and Some Related Oxanes. *J. Am. Chem. Soc.* , 92, 7281-7290.

Plazinski, W. and Drach, M. (2015) The influence of the hexopyranose ring geometry on the conformation of glycosidic linkages investigated using molecular dynamics simulations. *Carbohydr. Res.*, 415, 17-27.

Plazinski, W., Drach, M. and Plazinska, A. (2016a) Ring inversion properties of 1→2, 1→3 and 1→6-linked hexopyranoses and their correlation with the conformation of glycosidic linkages. *Carbohydr. Res.*, 423, 43-48.

Plazinski, W., Lonardi, A. and Hünenberger, P.H. (2016b) Revision of the GROMOS 56A6CARBO force field: Improving the description of ring-conformational equilibria in hexopyranose-based carbohydrate chains. *J.Comput. Chem.*, 37, 354-365.

Rao, V.S.R., Qasba, P.K., Balaji, P.V. and Chandrasekaran, R. (1998) *Conformation of Carbohydrates*, Amsterdam: Harwood Academic Publishers. pp. 49-90

Roslund, M.U., Tähtinen, P., Niemitz, M. and Sjöholm, R. (2008) Complete assignments of the ^{1}H and ^{13}C chemical shifts and $J_{H,H}$ coupling constants in NMR spectra of D-glucopyranose and all D-glucopyranosyl-D-glucopyranosides. *Carbohydr. Res.*, 343, 101-112.

Rundlöf, T., Kjellberg, A., Damberg, C., Nishida, T. and Widmalm, G. (1998) Long-range proton–carbon coupling constants in conformational analysis of oligosaccharides. *Magn. Reson. Chem.*, 36, 839-847.

Salmén, L. (2015) Wood morphology and properties from molecular perspectives. *Ann. For. Sci.*, 72, 679-684.

Salmén, L. and Lindström, M. (2015) Improved pulp yield and strength by retained glucomannans in kraft pulping of softwood. *Nord. Pulp Paper Res. J.*, 30, 584-590.

Scheller, H.V. and Ulvskov, P. (2010) Hemicelluloses. *Annu. Rev. Plant Biol.*, 61, 263-289.

Siegbahn , A., Aili, U., Ochocinska, A., Olofsson, M., Rönnols, J., Mani, K., Widmalm, G. and Ellervik, U. (2011) Synthesis, conformation and biology of naphthoxylosides. *Bioorg. Med. Chem.* , 19, 4114-4126.

Sjöström, E. (1977) The behaviour of wood polysaccharides during alkaline pulping processes. *Tappi J*, 60, 151-154.

Sjöström, E. (1993) *Wood Chemistry: Fundamentals and Applications*, 2nd edition, London: Academic Press. pp. 114-165

Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., Osborne, E., Paredez, A., Persson, S., Raab, T., Vorwerk, S. and Youngs, H. (2004) Toward a Systems Approach to Understanding Plant Cell Walls. *Science*, 306, 2206-2211.

Stott, K., Keeler, J., Van, Q.N. and Shaka, A.J. (1997) One-Dimensional NOE Experiments Using Pulsed Field Gradients. *J. Magn. Reson.*, 125, 302-324.

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

Säwén, E., Massad, T., Landersjö, C., Damberg, P. and Widmalm, G. (2010) Population distribution of flexible molecules from maximum entropy analysis using different priors as background information: application to the ϕ,ψ-conformational space of the α-(1→2)-linked mannose disaccharide present in N- and O-linked glycoproteins. *Org. Biomol. Chem.*, 8, 3684-3695.

Thrippleton, M.J. and Keeler, J. (2003) Elimination of Zero-Quantum Interference in Two-Dimensional NMR Spectra. *Angew. Chem. Int. Ed.*, 42, 3938-3941.

Timell, T.E. (1967) Recent Progress in the Chemistry of Wood Hemicelluloses. *Wood Sci. Technol.*, 1, 45-70.

Wang, D., Ámundadóttir, M.L., van Gunsteren, W.F. and Hünenberger, P.H. (2013) Intramolecular hydrogen-bonding in aqueous carbohydrates as a cause or consequence of conformational preferences: a molecular dynamics study of cellobiose stereoisomers. *Eur. Biophys. J.*, 42, 521-537.

Widmalm, G. (2013) A perspective on the primary and three-dimensional structures of carbohydrates. *Carbohydr. Res.*, 378, 123-132.

Wohlert, J., Bergenstråhle-Wohlert, M. and Berglund, L.A. (2012) Deformation of cellulose nanocrystals: entropy, internal energy and temperature dependence. *Cellulose*, 19, 1821-1836.

Zaccheus, M., Pendrill, R., Jackson, T.A., Wang , A., Auzanneau, F.-I. and Widmalm, G. (2012) Conformational Dynamics of a Central Trisaccharide Fragment of the LeaLex Tumor Associated Antigen Studied by NMR Spectroscopy and Molecular Dynamics Simulations. *Eur. J. Org. Chem.*, 4705-4715.

Zhang, Y.-H.P. (2008) Reviving the carbohydrate economy via multi-product lignocellulose biorefineries. *J. Ind. Microbiol. Biotechnol.*, 35, 367-375.

Zhang, Q., Brumer, H., Ågren, H. and Tu, Y. (2011) The adsorption of xyloglucan on cellulose: effects of explicit water and side chain variation. *Carbohydr. Res.*, 346, 2595-2602.

Zhao, Z., Crespi, V.H., Kubicki, J.D., Cosgrove, D.J. and 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.

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

TABLES

Table 1. All simulated oligosaccharide structures.

Abbrev.	Systematic notation	Common name	Linkage type
GG	β-D-Glc-(1→4)-β-D-Glc	β-Cellobiose	C
MM	β-D-Man-(1→4)-β-D-Man	β-Mannobiose	M
XX	β-D-Xyl-(1→4)-β-D-Xyl	β-Xylobiose	X
GM	β-D-Glc-(1→4)-β-D-Man		C
MG	β-D-Man-(1→4)-β-D-Glc		M
α-GG	β-D-Glc-(1→4)-α-D-Glc	α-Cellobiose	C
AA	β-D-All-(1→4)-β-D-All	β-Allobiose	
GGGG	β-D-Glc-(1→4)-β-D-Glc-(1→4)-β-D-Glc-(1→4)-β-D-Glc	β-Cellotetraose	CCC
MMMM	β-D-Man-(1→4)-β-D-Man-(1→4)-β-D-Man-(1→4)-β-D-Man	β-Mannotetraose	MMM
XXXX	β-D-Xyl-(1→4)-β-D-Xyl-(1→4)-β-D-Xyl-(1→4)-β-D-Xyl	β-Xylotetraose	XXX
MGGM	β-D-Man-(1→4)-β-D-Glc-(1→4)-β-D-Glc-(1→4)-β-D-Man		MCC
MGMM	β-D-Man-(1→4)-β-D-Glc-(1→4)-β-D-Man-(1→4)-β-D-Man		MCM
MMGM	β-D-Man-(1→4)-β-D-Man-(1→4)-β-D-Glc-(1→4)-β-D-Man		MMC
AAAA	β-D-All-(1→4)-β-D-All-(1→4)-β-D-All-(1→4)-β-D-All	β-Allotetraose	

Table 2. Probabilities [%] of different conformational states of the di- and tetrasaccharide glycosidic linkages.

Linkage type	Structure	1	2	3	$3^{(+)}$	$3^{(-)}$	4	$4^{(+)}$	$4^{(-)}$
C	GG	0.7 (0.3)	0.0	1.9 (0.2)	*1.9*	*0.1*	97.3 (0.4)		
C	GM	0.4 (0.1)	0.0	2.1 (0.1)	*2.1*	*0.0*	97.4 (0.2)		
M	MM	1.1 (0.4)	0.0	5.8 (0.4)	*5.6*	*0.1*	93.1 (0.7)		
M	MG	0.6 (0.3)	0.0	5.2 (0.4)	*5.1*	*0.1*	94.2 (0.4)		
X	XX	1.2 (0.3)	0.0	1.7 (0.2)	*1.4*	*0.3*	97.1 (0.3)	*53.4*	*43.7*
	AA	0.2 (0.1)	0.0	2.8 (0.3)	*2.4*	*0.5*	97.0 (0.3)		
C	GG-GG	1.5 (0.7)	0.0	2.0 (0.2)	*2.0*	*0.0*	96.5 (0.7)		
M	MM-MM	2.5 (1.3)	0.0	5.7 (0.2)	*5.7*	*0.1*	91.7 (1.2)		
X	XX-XX	0.4 (0.1)	0.0	1.2 (0.1)	*1.1*	*0.0*	98.5 (0.2)	*49.3*	*49.1*
C	MG-GM	1.6 (0.9)	0.0	2.5 (0.3)	*2.5*	*0.0*	95.9 (0.8)		
C	MG-MM	1.4 (0.9)	0.0	2.3 (0.1)	*2.3*	*0.0*	96.3 (0.9)		
M	MM-GM	3.7 (1.5)	0.0	4.7 (0.7)	*4.7*	*0.1*	91.6 (1.5)		

All values are rounded to one decimal. Numbers in parenthesis are the standard errors.

Table 3. $^{3}J_{CH}$ coupling constants for the four disaccharides that were studied by NMR spectroscopy.

		α-GG	GG	MM		XX
$^{3}J_{H1',C4}$	NMR	(3.95)[a] [4.0][b]	(3.97)[a] [4.1][c]	4.19 {4.32}[d]	NMR	4.25 {4.16}[e] [4.7][f]
	MD average	4.05	4.02	3.71	MD average	3.97
	MD reg. $3^{(+)}$	2.94	2.90	1.58	MD reg. $4^{(-)}$	4.18
	MD reg. 4	4.06	4.04	3.84	MD reg. $4^{(+)}$	3.82
$^{3}J_{C1',H4}$	NMR	(5.01)[a] [5.0][b]	n.d. [5.3][c]	n.d.	NMR	n.d. [5.1][f]
	MD average	5.58	5.65	5.75	MD average	4.69
	MD reg. $3^{(+)}$	5.02	5.06	5.13	MD reg. $4^{(-)}$	5.28
	MD reg. 4	5.58	5.65	5.77	MD reg. $4^{(+)}$	4.14

Heteronuclear scalar coupling constants in Hz. Values in () are from literature, [] are for literature values on the corresponding methyl cello- or xylobioside, and {} are NMR values for the reducing end α-anomeric form of the disaccharide. [a](Bell et al., 2013), [b](Hatcher et al., 2011), [c](Rundlöf et al., 1998), [d]α-MM, [e]α-XX, [f](Hricovíni et al., 1990)

Table 4. Comparison between averaged atom distances obtained with NMR and MD simulations.

Compound	$r_{H1',H4}$ MD[a]	$r_{H1',H4}$[b]	$\sigma_{H1',H4}$	r_{ref} MD[a]	σ_{ref}	ref pair[a]
α-GG	2.30	2.23	8.6[c]	2.37	6.0	H1,H2
Me α-GG[d]		[2.29]	[13.9]	[2.40]	[9.8]	[H1,H2]
Me GG[e]	2.29[f]	[2.19]	[8.9]	[2.54]	[3.7]	[H2,H4]
XX	2.39	2.61	6.6	2.64	6.1	H1,H5$_{ax}$
XX		2.64	6.6	2.85	4.2	H1,H3

Interatomic distances are in Å and cross-relaxation rates in x10^2s^{-1}. Literature values are in []. [a]Averaged over the MD simulation according to: $1/r_{ij}^{calc} = \langle r_{ij}^{-6} \rangle^{1/6}$; and [b] according to: $r_{ij}^{expt} = r_{ref}\left(\sigma_{ref}/\sigma_{ij}\right)^{1/6}$. [c]STEP-NOESY, [d] (Larsson et al., 2004), [e] (Hatcher et al., 2011), [f]MD simulation of β-D-Glc-(1→4)-β-D-Glc (GG).

FIGURE LEGENDS

Figure 1. a) General structures of relevant wood polysaccharides. a.1) cellulose, a.2) *O*-acetyl-galacto-glucomannan, a.3) xyloglucan, a.4) arabino-4-*O*-metylglucurono-xylan. Structures a.1), a.2) and a.4) are adapted from (Timell, 1967), and structure a.3) from (Ebringerová et al., 2005). b) The three studied glycosidic linkages defined as type C, M and X. c) Definition of the dihedral angles of the glycosidic linkage: φ = O5'-C1'-O4-C4; φ_H = H1'-C1'-O4-C4; $\varphi_{C2'}$ = C2'-C1'-O4-C4; ψ = C1'-O4-C4-C3; ψ_H = C1'-O4-C4-H4; ψ_{C5} = C1'-O4-C4-C5. The definition for φ and ψ are according to IUPAC and these are mainly used within this study, additionally φ_H, $\varphi_{C2'}$, ψ_H and $\psi_{C1'}$ are used for the *J* coupling calculations according to Karplus-type relationships (Säwén et al., 2010) where $\varphi_{O5'} = \varphi$, and $\psi_{C3} = \psi$.

Figure 2. Free energy maps of disaccharides. The global minimum of the energy in each map is anchored zero kJ mol^{-1} and each map is divided into four main regions.

Figure 3. Schematic of the STEP-NOESY experiment for measuring the cross-relaxation rate between H4 and H1' in α-GG; selective excitation of H1 (1), followed by isotropic mixing transfers magnetization to H4, which then can be selectively inverted (2), for subsequent cross-relaxation to H1' (bottom). ^{1}H,^{1}H-NOE buildup curves for α-GG obtained at a 600 MHz spectrometer frequency employing the PANIC approach (Macura et al., 1986; Hu and Krishnamurthy, 2006), in which $-I_f/I_i$ vs. τ_{mix} are plotted. The cross-relaxation rates are obtained from the slopes of the fitted data; reference distance H1-H2 (filled squares) and H4-H1' (filled triangles) (top).

Figure 4. Hydrogen bonding probability for the modelled disaccharides and tetrasaccharides. Normalized with the probability for each structure to be in the respective region.

Figure 5. Conformations of three disaccharides. Structures of GG, MM and XX corresponding to the three energy minima of each structure. The numbers in parenthesis represent the instantaneous (φ,ψ) coordinates in the corresponding free energy map. Hydrogen bonds are represented by the dashed lines.

Figure 6. Free energy maps of XX with different chair conformations. The percentage present on each map is corresponding to the probability for that combination of ring conformations.

Figure 7. Free energy maps of tetrasaccharides. The global minimum of the energy in each map is set to zero kJ mol^{-1} and each map is divided into four main regions in the same way as for the disaccharides.

Figure 8. Generated ten unit oligosaccharides with probable conformations at the glycosidic linkages. The structures on the lower row contain glycosidic linkages from two regions. The

notation X/Y means that 8 out of 9 linkages have a conformation from region X and one from region Y.