Green Chemistry

PAPER

View Article Online
View Journal | View Issue

Cite this: *Green Chem.*, 2018, **20**, 2534

Differences in extractability under subcritical water reveal interconnected hemicellulose and lignin recalcitrance in birch hardwoods†

Antonio Martínez-Abad,[a] Nicola Giummarella,[b,c] Martin Lawoko *[b,c] and Francisco Vilaplana *[a,c]

Received 3rd February 2018,
Accepted 25th April 2018

DOI: 10.1039/c8gc00385h

rsc.li/greenchem

Hardwoods constitute an essential renewable resource for the production of platform chemicals and bio-based materials. A method for the sequential extraction of hemicelluloses and lignin from hardwoods is proposed using subcritical water in buffered conditions without prior delignification. This allows the cascade isolation of mannan, xylan and lignin-carbohydrate complexes based on their extractability and recalcitrance in birch lignocellulose. The time evolution of the extraction was monitored in terms of composition, oligomeric mass profiling and sequencing of the hemicelluloses, and molecular structure of the lignin and lignin-carbohydrate complexes (LCCs) by heteronuclear single quantum coherence nuclear magnetic resonance (2D HSQC NMR). The minor mannan and pectin populations are easily extractable at short times (<5 min), whereas the major glucuronoxylan (GX) becomes enriched at moderate extraction times. Longer extraction times result in major hydrolysis exhibiting GX fractions with tighter glucuronation spacing and lignin enrichment. The pattern of acetylation and glucuronation in GX is correlated with extractability and with connectivity with lignin through LCCs. This interconnected molecular heterogeneity of hemicelluloses and lignin has important implications for their supramolecular assembly and therefore determines the recalcitrance of hardwood lignocellulosic biomass.

Introduction

Lignocellulosic biomass is the most abundant renewable resource on Earth and constitutes the cornerstone in the production of sustainable materials, chemicals and energy with a balanced carbon cycle.[1,2] Wood lignocellulose consists of a complex network of oriented cellulose microfibrils embedded in a tightly intertwined matrix of hemicelluloses and lignin.[3] In hardwoods, glucuronoxylans (GX) are the most abundant hemicelluloses, although glucomannans (GM) are also observed in minor amounts[4–6] (Fig. 1A). Lignin is a disordered polymer with large structural diversity, which accounts for 20–30% in lignocellulosic biomass. Lignin consists of *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units linked by various interunit linkages (Fig. 1B), S and G units linked by β-O-4 bonds being the most abundant in hardwoods.[7] Biomass recalcitrance comprises the multiscale structural features that hinder the conversion of the biomass components into the final desired products.[2,8,9] Lignocellulose recalcitrance is determined by the composition and molecular structure of the individual components (cellulose, hemicellulose, and lignin), and the nature of their intra- and intermolecular interactions.[10,11] The molecular structure of hemicelluloses plays an important role in modulating the association with cellulose microfibrils[12–15] and the occurrence of covalent linkages with lignins forming so-called lignin carbohydrate complexes (LCCs) (Fig. 1C).[16,17] Although the occurrence of LCCs is still under debate due to questionable isolation procedures and low frequencies of detected bonds, recalcitrance in biomass fractionation has partly been attributed to these covalent linkages.

The economic viability of lignocellulosic biorefineries relies in the development of a multi-product portfolio. While cellulose has been widely exploited in the paper and pulp industry, the transformation of hemicellulose and lignin into high value products has been considered recently. Hemicelluloses in polymeric form display film-forming and oxygen barrier properties comparable to synthetic polymers together with

[a]*Division of Glycoscience, Department of Chemistry, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, AlbaNova University Centre, SE-106 91 Stockholm, Sweden. E-mail: franvila@kth.se*
[b]*Department of Fibre and Polymer Technology, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Teknikringen 56-58, SE-100 44 Stockholm, Sweden. E-mail: lawoko@kth.se*
[c]*Wallenberg Wood Science Centre, KTH Royal Institute of Technology, Teknikringen 56-58, SE-100 44 Stockholm, Sweden*

†Electronic supplementary information (ESI) available. See DOI: 10.1039/c8gc00385h

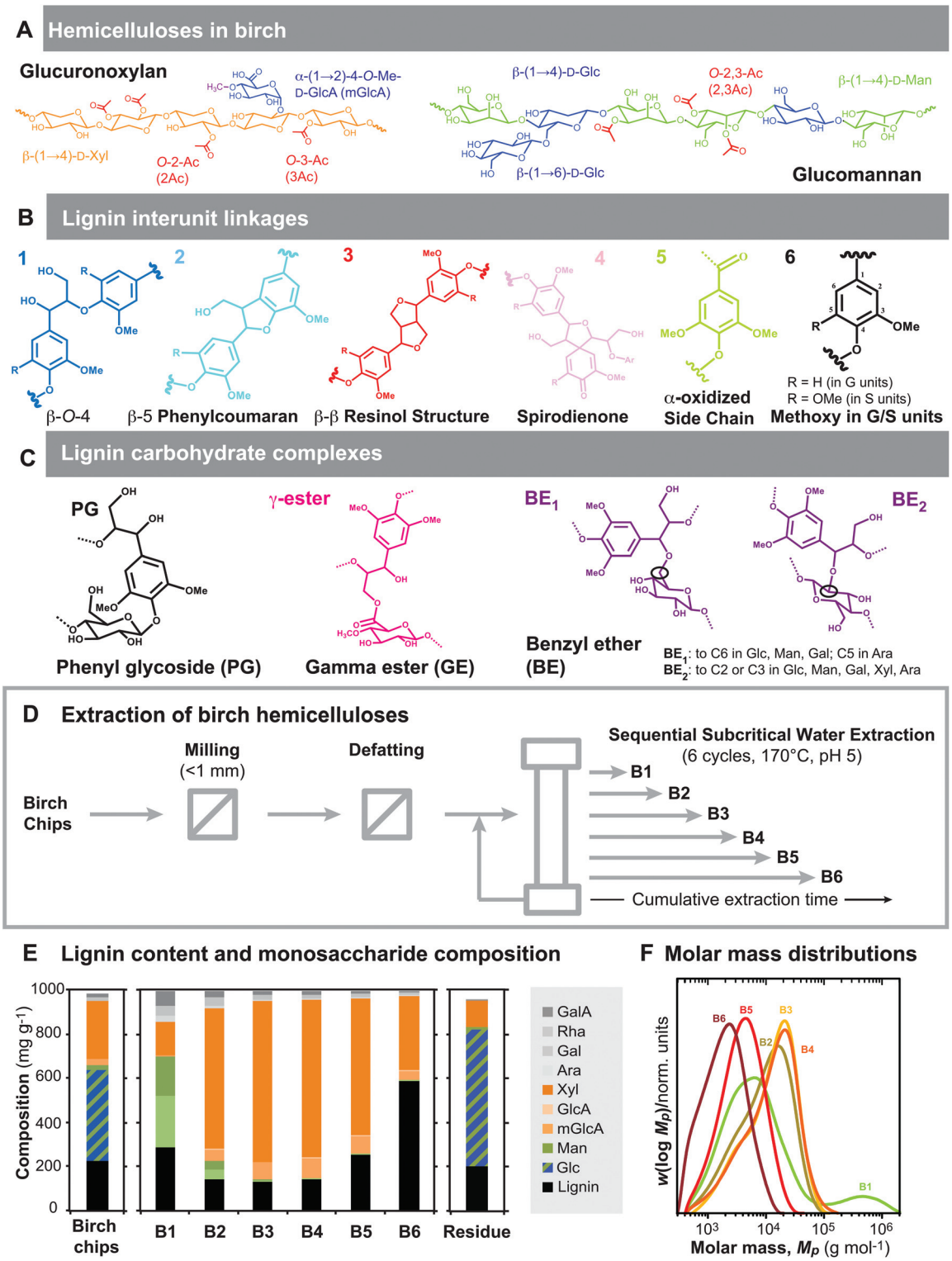

Fig. 1 Extraction, composition and molecular structure of the fractions from birch wood. (A) Hemicelluloses in birch. Glucuronoxylans (GX) consist of a backbone of β-(1 → 4)-linked xylopyranosyl (Xyl) units, decorated with 4-*O*-methyl glucuronic acid (mGlcA) units at the α-(1 → 2) position, and heavily acetylated (Ac) at the O-2 and/or O-3 positions.[6] Glucomannans (GM) consist of a backbone of β-(1 → 4)-linked mannopyranosyl (Man) or glucopyranosyl (Glc) units, with a minor occurrence β-(1 → 6)-Glc decorations and acetylated in the O-2 and/or O-3 positions of the Man units.[5] (B) Lignin interlinkage units (β-O-4′, β-5′, β-β′, 5-5′, 5-O-4′, and β-1′) connect the guaiacyl (G), and syringyl (S) units. (C) Lignin-carbohydrate complexes have been identified as phenyl glycosides, gamma esters, and benzyl ethers. (D) Experimental process for the sequential extraction of birch hemicelluloses using subcritical water. (E) Lignin content and monosaccharide composition of the birch chips, the consecutive extracts and the residue. (F) Molar mass distributions of the extracts.

excellent rheological properties,[18,19] with potential applications in the packaging sector and as texturizing agents. On the other hand, the use of lignin in material systems has not been very successful due to the molecular heterogeneity of available technical lignins[20] and molecular mass dispersity.[21–23] Large research efforts have been undertaken on unlocking the potential of lignin,[24] such as its use in polymer systems,[25,26] its thermal and catalytic valorisation,[27,28] and the preparation of lignin-based nanoparticles.[29] Most of the above-mentioned works have however been performed on technical lignins as Kraft and Organosolv lignins. However, studies of novel lignin streams originating from greener processes such as sub- or supercritical water extractions are still scarce. Recent research suggests that the dissolved lignin from such systems are less complex in structure than conventional ones, and are more similar to native lignins.[30]

Subcritical water extraction (SWE), also referred to in the literature as pressurized hot-water extraction, is a process based on the solubilisation of macromolecules in liquid water at increased pressure and temperature without any hazardous chemicals.[31] Under subcritical conditions, the density, surface tension, polarity, viscosity and diffusion of water change, and the mass transfer during extraction is enhanced. SWE is an important strategy for wood fractionation, as hemicelluloses can be separated for added value products, while the residual cellulose-rich solid can still be further processed.[32–34] In this study we have evaluated the capacity of SWE, performed in a sequential manner and without previous delignification, to selectively extract high yields of high molecular weight hemicelluloses and lignin from birchwoods. Structural changes in the extracted fractions have been monitored by mass-spectrometric (MS) approaches and nuclear magnetic resonance (NMR), in order to correlate molecular structure, connectivity of the lignocellulosic components, and extractability. The influence of the hemicellulose substitution pattern (acetylation, glucuronation) and the connectivity between hemicelluloses and lignins on the recalcitrance of hardwoods has been assessed. This fundamental understanding contributes to the development of sustainable processes for the fractionation of wood biopolymers in integral lignocellulosic biorefineries.

Results and discussion

Sequential subcritical water extraction of hemicelluloses and lignin: mass balances, composition, and molar mass distributions

We have proposed a sequential SWE process under buffered conditions from defatted birch chips without delignification to maximize the extraction of hemicellulose populations, minimize potential deacetylation and autohydrolysis, and evaluate the connectivity between lignin and hemicellulose (Fig. 1D). Preliminary assays were performed on the defatted birchwood chips in order to set the temperature conditions at 170 °C, and to optimize the duration of the consecutive cycles of SWE yielding extracts B1–B6 (data not shown). These conditions are comparable to similar procedures reported for the extraction of polymeric xylan from birchwood.[33,35] Table 1 shows the yields, composition and average molar mass of the starting birch chips (BC), the extracts (B1–B6) and the residue (R). After the SWE process, 68.5 wt% of the initial material is left in the residue, which indicates that the extraction yielded around 30 wt% of total dissolved solids, 23.6 wt% of them in B1–B6 as high molecular weight extracts (Table 1). From the mass balances, approximately 8% of the total material was lost in the extraction process; this corresponds with a low molar mass fraction that was essentially discarded during the dialysis of the extracts B5 and B6 (data not shown).

The lignin content and the monosaccharide composition profiles after acid hydrolysis from the sequential extracts show interesting differences with time (Fig. 1E). Hemicelluloses are primarily extracted throughout the sequential process, with an interesting evolution of the mannan and xylan compositions. Cellulose was not detected in the subcritical water extracts and was highly enriched in the residue. The absence of cellulose in the SWE extracts was verified by comparison of the monosaccharide composition after methanolysis and sulphuric acid hydrolysis (ESI Fig. S1†). Extraction of lignin, on the other hand, increased significantly at longer exposure times (Fig. 1E, Table 1). The molar mass profiles of the different extracts indicate that most of the fractions yield polymeric populations between 10^4 and 10^5 Da (Fig. 1F), although the occurrence of degradation is significant for longer extraction times. The comparison of the refractive index (DRI) and ultraviolet (UV) detection traces gives indirect evidence of the interconnectivity of the hemicellulose and lignin fractions in the molar mass distributions (ESI Fig. S2†).

Mannan was preferentially extracted at short times (<5 min), being the major component in extract B1. Pectic components are also enriched in extract B1 compared to the later extracts. Mannan is a minor hemicellulosic component in hardwoods (between 3–5 wt%),[36] in agreement with the monosaccharide composition of our birch chips (ESI Table S1†). Hardwood mannans are described as acetylated glucomannans, with reported mannose-to-glucose (Man : Glc) ratios of 2.1–2.4 : 1 for birch and degree of acetylation (DS_{ac}) of 0.2–0.3.[5] A bimodal molar mass distribution is observed for extract B1, indicating the presence of large populations of apparent molar masses between 10^5–10^6 Da, and smaller populations centered at 5000 Da. These populations have been ascribed to large charged polysaccharides (probably colloidal particles and or pectic substances) and neutral *O*-acetylated glucomannan with degree of polymerization of 15, respectively.[5] These mannan populations are easily extractable from the hardwood tissue, which may be related to their localization and structural role in the secondary cell wall of angiosperms.[37,38]

Xylan is the major hemicellulose in hardwoods, representing 30 wt% in the birchwood chips. Most part of this xylan (56.1%) is extracted in form of high molar mass polysaccharides throughout the entire SWE process (Table 1). Xylan is preferentially extracted in extracts B3 and B4 with high purity (around 95 wt% of the total carbohydrate content) and a stable

Table 1 Mass balances and composition of the different fractions

	BC	B1	B2	B3	B4	B5	B6	$\sum B_i$	R
Extraction times (min)	n.a.	5	15	20	20	60	120	n.a.	n.a.
Total yield (%)[a]	100	1.4	4.1	7.7	4.8	3.4	2.2	23.6	68.5
Xylan yield (%)[a]	100	0.8	9.7	21.4	13.4	8.1	2.8	56.1	27.2
Lignin yield (%)[a]	100	1.9	2.5	4.4	3.0	3.9	5.8	21.5	62.4
Carbohydrate content (mg g^{-1})[b]	766.5	715.1	863.1	873.5	862.1	747.1	410.2	n.a	752.6
Cellulose (%)[b]	56.7	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.		81.9
Xylan (%)[b]	39.6	22.5	81.1	93.5	94.8	94.8	93.5		15.6
Mannan (%)[b]	2.5	57.7	9.7	1.6	0.9	0.9	1.6		2.0
Pectin (%)[b]	4.4	19.8	9.2	4.9	4.4	4.3	4.9		0.6
mGlcA : Xyl[b]	0.11	0.04	0.08	0.10	0.13	0.13	0.11		0.01
Lignin content (mg g^{-1})[c]	220.5	284.9	151.0	148.6	159.9	272.0	597.6	n.a	201.0
Acetyl content (%)[d]	6.5	6.2	10.0	7.8	5.6	2.3	0.8	n.a	0.4
M_n (kDa)[e]	n.a.	421.1	5.1	6.9	6.3	2.7	1.4	n.a	n.a.
M_w (kDa)[e]		605.0	14.3	17.7	18.4	5.3	2.6		

[a] Yields determined gravimetrically and referred to the original birch chips. [b] Determined from the complete monosaccharide composition (Table S1). [c] Determined from the Klason lignin and the acid soluble lignin. [d] Determined after saponification and liquid chromatography (HPLC-UV). [e] Determined by size-exclusion chromatography (SEC). n.a: not applicable; n.d.: not determined.

and relatively high molecular weight (*ca.* 20 kDa; Fig. 1F). These values are, to the best of our knowledge, the highest reported in terms of yields and purity of polymeric xylan obtained by SWE, and similar to the overall hemicellulose extracted by a similar hydrothermal treatment.[30]

The molar mass of the extracted xylans are also close to values comparable to traditional alkaline extracted xylan.[39] This might be due to both efficient pH control due to the inclusion of a buffering system and the removal of saturated solutions in consecutive extraction cycles, which results in an increase of the dissolution rates and less exposure of extracted materials to the harsh extraction conditions. Although recovery of 99% of all xylan has been achieved in some studies, this only happens at the expense of extensive depolymerization to mono- or oligosaccharides.[40,41] Other studies reported high molar mass values and absence of lignin in the extracts at lower extraction temperatures, but at the expense of much lower xylan yields.[41,42] A compromise between molar mass, yield and purity is achieved at increased temperature and controlled mildly acidic pH. On the other hand, the similar composition and molar mass in the extracts B2–B4 indicates that the extraction at these time scales (0–60 min) is rather governed by the mass transfer and diffusion kinetics of the xylan polymers from the woody tissues to the subcritical water liquid phase. That would also explain why non-consecutive extraction under prolonged times does not improve yield or purity (data not shown). Hence, consecutive cycles are necessary in order to optimize the solid to liquid ratio while maintaining a high rate of extraction. In extracts B5 and B6, however, longer cycles are needed to keep up with the extraction yields. In these cases, the extracts show a drastic decrease in the molar mass, which indicates the occurrence of hydrolytic processes causing depolymerization (Fig. 1F). Even after 4 h extraction, 30% of all xylan remained in the residue. This evidences the recalcitrant structure of the secondary cell wall and suggests the existence of distinct crosslinked domains in xylan. These differences may be ascribed to the molecular structure of birch GX in terms of glucuronosyl (mGlcA) and/or acetyl (Ac) substitution patterns, which in turn govern the molecular and supramolecular interactions with cellulose and lignin. Interestingly, the glucuronic acid to xylose molar ratio (mGlcA : Xyl) in the extracts increases with extraction time (Table 1). The acetyl content, on the contrary, drastically decreases with extraction time in the xylan rich extracts (B2–B6). In addition to this, the contribution of lignin to the total yields is increasingly higher at longer extraction times in extracts B5 and B6 (Fig. 1E, Table 1). This suggests the connectivity of xylan to lignin as it has been suggested for SWE pre-treated woods[43] and their role in hemicellulose recalcitrance.[30]

Oligomeric mass profiling (OLIMP) of acetylation and glucuronation in birch GX during sequential SWE

In order to monitor the distribution of acetyl and mGlcA side groups in birch GX during the sequential SWE, intact and chemically deacetylated extracts were subjected to enzymatic digestion with a β-glucuronoxylanase (GH30). This enzyme specifically releases aldouronic acid oligosaccharides (UXOs) based on the recognition of the mGlcA substitution at the −2 position of the enzymatic active site, which enables the identification of the glucuronic acid spacing (Fig. 2A).[44,45] The oligomeric mass profiling (OLIMP) of the digested UXOs was analysed by electrospray mass spectrometry (ESI-MS). The chemically-deacetylated extracts were evaluated to investigate the potential steric hindrance of the acetyl groups on the enzymatic action of the glucuronoxylanase (Fig. 2B). A distribution of UXOs in the deacetylated xylan with mGlcA spacing every 3 to 8 xylose units can be observed, with a major relative intensity of X_5mU (a mGlcA unit every 5 Xyl units). The relative abundance of smaller UXOs (X_3U and X_4U) increases compared to the X_5U–X_8U signals in the more recalcitrant xylan extracts (B5 and B6; Fig. 2B), in agreement with the higher mGlcA : Xyl ratio with increasing extraction time (Table 1). This heterogeneous spacing of mGlcA substitutions contrasts with the consistent pattern of glucuronation and arabinosylation

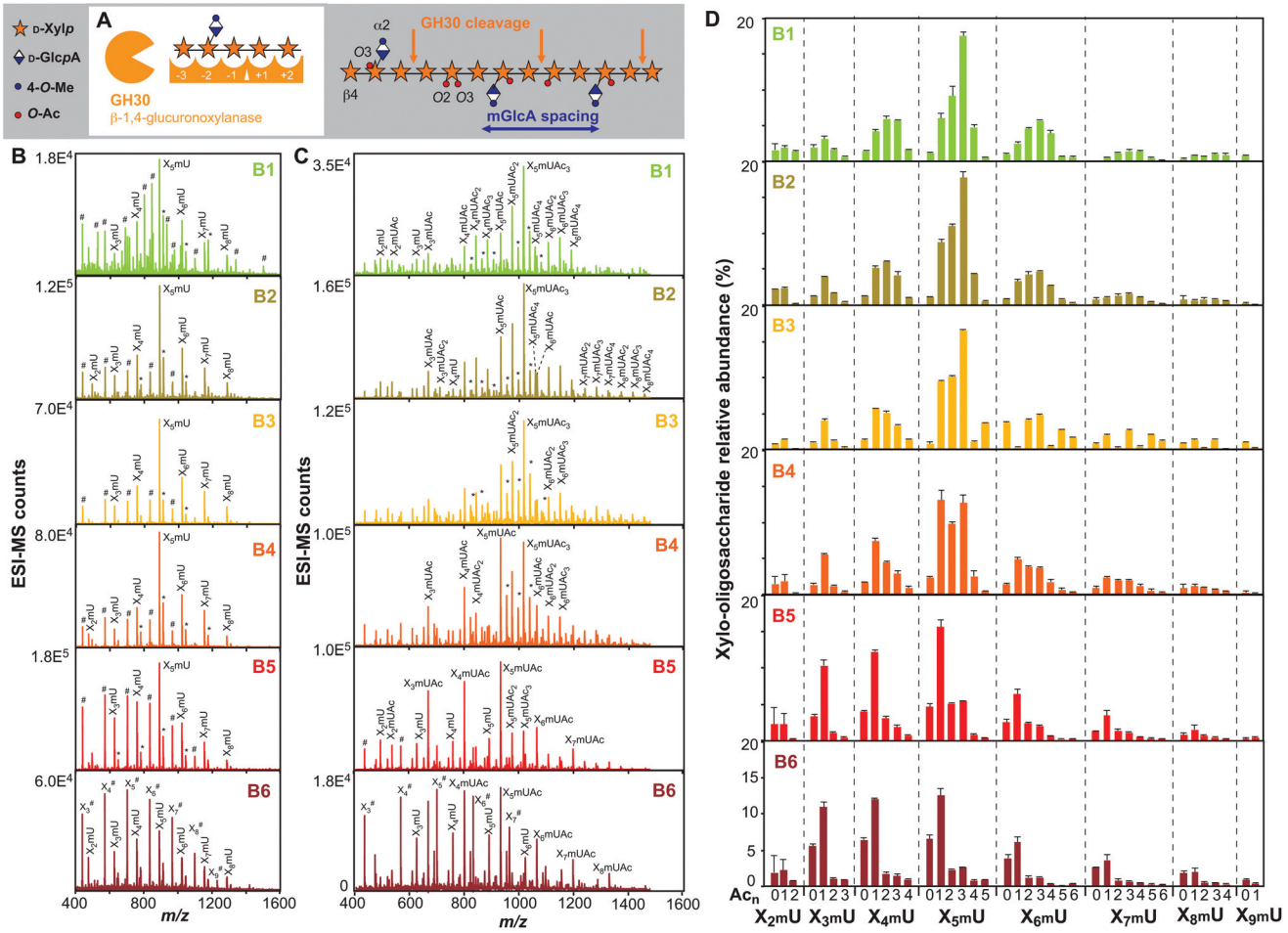

Fig. 2 Oligomeric mass profiling (OLIMP) of extracted glucuronoxylan (GX). (A) Substrate recognition by GH30 β-glucuronoxylanase and mGlcA spacing. (B) Electrospray mass spectrometry (ESI-MS) profiles from deacetylated GX extracts. (C) ESI-MS profiles from the intact GX extracts. (D) Relative abundance (%) of the alduronic acid oligosaccharides (UXOs), calculated from the total ESI-MS intensities from triplicate digestions. Peak assignment presented in ESI Table S2.† Note: X (xylose), mU (mGlcA), Ac (acetyl), # (linear XOs in the extracts prior to enzymatic action), * (sodiated adducts).

recently reported for gymnosperm xylans.[13,15] The mGlcA spacing differs as well from the one reported for GX from wild-type *Arabidopsis thaliana* (*A. thaliana*) stems. Instead, the mGlcA spacing is similar to the glucuronation pattern described by a specific glucuronosyl transferase (GUX2) from *A. thaliana*. GUX2 yields glucuronosyl side groups every 4–7 Xyl units with a higher abundance of the X_5mU spacing pattern, in contrast with the evenly-spaced glucuronation pattern generated by another glucuronosyl transferase GUX1.[45,46] No double glucuronated UXOs were detected in the GX extracts from birchwood; these unusual patterns were identified in our previous study in arabinoglucuronoxylan (AGX) from spruce wood.[15]

To study the evolution of the acetylation pattern with extraction time, OLIMP analysis was performed on the extracts without deacetylation (Fig. 2C). It is worth mentioning that non-acidic xylo-oligosaccharides (XOs) were detected for both deacetylated and acetylated GX extracts (marked with # in the ESI-MS profiles, Fig. 2B and C), especially for the recalcitrant extracts B5 and B6. These XOs were present prior to GH30 digestion and correspond with fragments released from the GX backbone during the SWE. Similar mGlcA spacing can be observed in the acetylated samples compared to the deacetylated ones, with abundance of acetylated X_5U isomers for the initial extracts (B1–B4), and a relative increase in shorter UXOs at longer times (B5 and B6). Previous studies on xylan acetylation in *A. thaliana* suggested that acetyl groups may hinder enzymatic access[47,48] and that the GH30 xylanase can recognize acetyl groups and cleave xylan even in the absence of mGlcA.[49] The absence of non-glucuronated XOs after enzymatic digestion and the similar distribution of UXOs in the original and deacetylated extracts suggest that acetylation did not affect enzymatic activity.

The relative abundances for each specific acetylated UXO isomer was evaluated as the sum of the ESI-MS peak intensities for all *m*/*z* adducts corresponding to each specific isomer (Fig. 2D). During the first stages of extraction (B1–B3), there is a distribution of differently acetylated isomers for each X_nU

species, corresponding to an average of 1 acetyl group per 2–3 xylose units. This is in agreement with findings in *A. thaliana*, where acetyl groups were found mostly, but not exclusively, every 2 xylose units.[49] In extracts coming from prolonged extraction time (B4–B6), acetylation drastically decreases, correlating with the results in Table 1. This decrease might indicate differences in the acetylation pattern in distinct xylan domains, or might arise from deacetylation with prolonged extraction times. Considering the very low amounts of acetyls in the residue, it is clear that a substantial fraction of acetyl groups was lost in the low molecular weight fractions B5 and B6. Interestingly, the acetylation profile of the extracts does not show a progressive linear tendency during sequential extraction. Indeed, X_5UAc_3 is the most abundant isomer in the initial extracts B1–B3; X_5UAc and X_5UAc_3 become the most abundant isomers in B4, and a major abundance of X_5UAc is observed for the latter extracts B5 and B6. This suggests that birch GX shows a controlled acetylation pattern, that deacetylation of GX during SWE is not random, or both.

Carbohydrate sequencing reveals that acetylation and glucuronation are interrelated in birchwood GX

OLIMP by ESI-MS enables the evaluation of the acetylation and glucuronation profiles in birch GX, but it does not provide the internal position of the acetyl and mGlcA groups in the UXOs, which could reveal the existence of any substitution patterning or correlation between acetylation and/or glucuronation. To achieve this, the acetylated aldouronic acid isomers were separated and sequenced using liquid chromatography with electrospray tandem mass spectrometry (LC-ESI-MS/MS). Based on the OLIMP profiles, samples B3 and B5 were selected as representative examples of two xylan populations with different extractability. The separation and oligosaccharide sequencing for the most representative isomers, namely X_5UAc_n and X_4UAc_n, are shown in Fig. 3 and Fig. S3,† respectively.

The single ion monitoring (SIM) chromatogram at *m*/*z* 1012 corresponding to the non-acetylated isomer X_5mU shows only one peak as expected, with higher intensity for the B5 extract (Fig. 3A and B). ESI-MS/MS fragmentation of the isomer (Fig. 3C) confirms the position of the mGlcA side group at the Xyl unit in the −2 position ($XXXU^4m^2X$), in agreement with the cleaving mechanism of the enzyme.[44] Surprisingly, the SIM chromatograms for the X_5UAc_n isomers exhibit much simpler profiles than expected, taking into account the statistical theoretical number of acetylated UXO isomers for each *m*/*z* ratio based on the multiple positions of the acetyl groups (Fig. 3A and B). The chromatogram at *m*/*z* 1054 (X_5mUAc) displays one significant peak for both B3 and B5 extracts. The fragmentation allows the univocal sequencing of the $XXXU^4m^{2+3a}X$ structure and reveals the specific placement of the acetyl group at the same Xyl unit carrying the mGlcA substitution in the −2 position (Fig. 3C). The same behaviour is observed for the X_4mUAc isomer at *m*/*z* 1054 ($XXU^4m^{2+3a}X$) (Fig. S3†). Interestingly, the triply acetylated aldouronic acid isomers (X_5mUAc_3) at *m*/*z* 1138 only display one isomeric peak in the SIM chromatograms (Fig. 3A and B). ESI-MS/MS fragmentation allows the complete sequencing of the oligosaccharide and reveals the evenly spaced position of the Ac groups in the Xyl units at both the −2 and −4 positions ($XX^{2a+3a}XU^4m^{2+3a}X$) in the backbone. Finally, two main isomers are distinguished for the double-acetylated aldouronic acid (X_5mUAc_2) at *m*/*z* 1096. The two isomers maintain one Ac group in the glucuronated Xyl unit at the −2 position, and one Ac group evenly spaced at the −4 Xyl unit in the backbone, either at the C-2 ($XX^{2a}XU^4m^{2+3a}X$) or the C-3 ($XX^{3a}XU^4m^{2+3a}X$) of the Xyl ring. Identical structures are sequenced for the X_4mUAc_2 isomers (Fig. S3†). Unfortunately, ESI-MS/MS fragmentation does not allow the assignation of the regioselective position of the Ac group within the Xyl ring. However, the sequencing reveals the specific and controlled position of the Ac groups in the acetylated UXOs and the correlation between acetylation and glucuronation in birch GX. This involves the consistent presence of an *O*-3-Ac group in the Xyl unit carrying mGlcA and potential additional Ac groups in even positions from the Xyl unit carrying mGlcA. This distinct pattern in birch GX agrees with previous findings suggesting that Ac groups are preferably evenly spaced on every other Xyl unit in GX from *A. thaliana*.[49–51]

The abundance of acetylation and glucuronation differs significantly in the extracts, suggesting the existence of heterogeneous xylan structures with distinct substituted domains. For the GX populations extracted during the initial stages (B2–B4), with stable composition and molecular weight, the decrease in acetylation might be due to inherent structural differences in the native GX or due to chemical deacetylation during SWE. The mGlcA seems to stabilize the acetyl group in the same xylose unit, while acetyl groups at position −4 are progressively lost (Fig. 3A and B). In B5 and B6, as the even spacing of acetylation is maintained, the higher abundance of X_3U and X_4U UXOS with tighter mGlcA spacing might be linked to domains with lower acetylation. The results evince the possibility of tuning the SWE process to extract fractions with targeted acetyl content, which influence solubility and other physico-chemical properties of the extracted hemicelluloses.

Insights on hemicellulose acetylation, lignin structure and LCC complexes by heteronuclear single quantum coherence nuclear magnetic resonance (2D HSQC NMR)

In order to gain more knowledge about the structural evolution during SWE of birch wood and the role of the glucuronation and acetylation patterns in birch GX on the interactions with lignin, 2D HSQC NMR analysis was performed in all native extracts (Fig. 4). The assignment of the main ^{13}C–^{1}H HSQC correlation signals is presented in the ESI Tables S3–S5† for the lignin, LCCs, and carbohydrate structural units, respectively. The main carbohydrate signals can be assigned to the hemicellulose populations, corresponding typically to partially acetylated GX and GM. The relative abundance of the mannan and xylan signals in the 2D HSQC NMR plots is in agreement with the compositional analysis, with acetylated GM being enriched in the extract B1, and acetylated GX in the

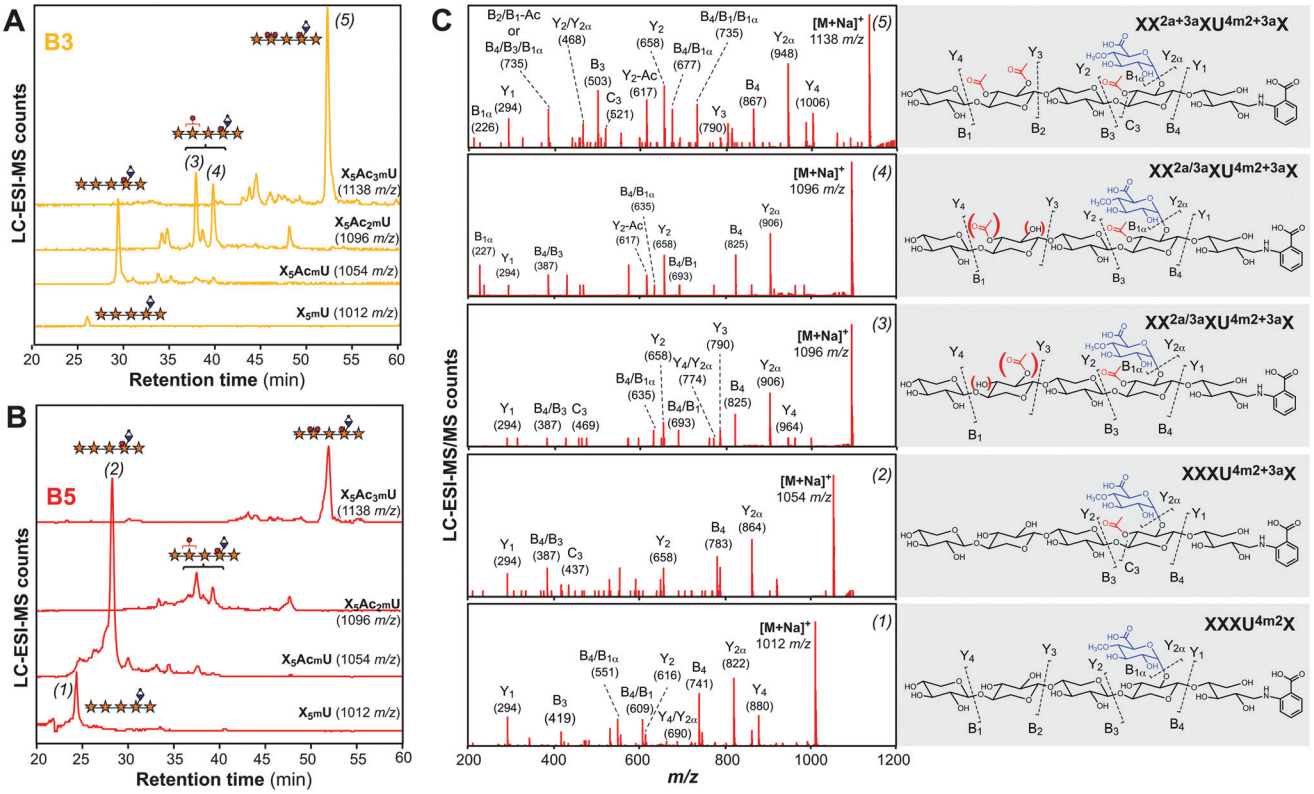

Fig. 3 Oligosaccharide sequencing of the birch GX extracts. Single ion monitoring (SIM) chromatograms for the X_5mUAc_n isomers in extract B3 (A) and extract B5 (B). (C) Fragmentation of the numbered peaks (1–5) and assignation of the ion fragments according to the nomenclature proposed by Domon and Costello.[52] The sequenced oligosaccharide structures are named following the systematic nomenclature proposed for xylo-oligosaccharides by Fauré *et al.*[53] Note: X (Xyl), mU (mGlcA), Ac (acetyl).

later extracts. 2D HSQC NMR enables the semi-quantitative determination of acetylation regioselectivity, providing the abundance of the acetyl groups in the *O*-2 (X_{2Ac}, M_{2Ac}), *O*-3 (X_{3Ac}), and *O*-2,3 ($X_{2,3Ac}$) units in the xylan and mannan backbone (Table 2). Acetyl quantification by NMR offers comparable results to the degree of acetylation obtained by HPLC. Sample B1 exhibits heavily acetylated mannan populations, in agreement with previously characterized acetylated GM from birch and aspen.[5] In addition, highly acetylated GX is also extracted in this fraction, in agreement with xylans extracted from ball-milled birch.[16] For the xylan-rich extracts B2 and B3, the decrease in acetylation in the *O*-3 or *O*-2,3 Xyl units is less noticeable than in *O*-2 acetylated Xyl units (Table 2). It is difficult to discern whether the differences are a result of chemical deacetylation of potentially more labile 2-*O*-Ac groups, or to real structural differences within xylan domains. The decrease in 2-*O*-Ac groups is not so evident in the LC-ESI-MS/MS profiles (Fig. 3A), which might be due to the fact that the GX domains bound to lignin are not hydrolysed by the GH30 xylanase and therefore might escape analysis.

The lignin structure in the extracts displays a substantial abundance of syringyl (S) compared to guaiacyl (G) units interconnected mostly by β-*O*-4 linkages (Table 2). This is in agreement with the typical lignin structure in hardwoods[7] and is specifically in good agreement with recent analysis on birch wood lignins.[16] The S/G ratio is normally used to gain insight on the monomeric composition of lignin. Here, we notice that most of the fractions have an S/G ratio between 4 and 5 with the exception of the B1 lignin, with a value of 2.4. The ratio of 5 has been reported to be an artefact resulting from incomplete solubility in d6-DMSO prior to NMR analysis, and was shown to decrease to a value of about 3 when solubility was complete upon acetylation of the fraction.[16] Thus, S/G ratios observed in B1 fraction are typical of native hardwood lignins. The aryl ether structures are mostly maintained throughout the extraction process until longer extraction times are reached (extracts B5 and B6), where a relative increase of the β–β-resinol structures is observed. Such inter-unit linkages involving carbon–carbon bonds are generally stable at several technical conditions. This increase is likely a result of an enrichment of carbon–carbon inter-unit linkages at the expense of the partial cleavage of aryl ether linkages.[30,54] The prevalence of high β-O-4 linkages is possibly attributed to the buffered system during SWE. The pH drop caused by deacetylation during SWE has been suggested to induce both acidolytic cleavage of aryl ethers leading to depolymerisation, and condensation reactions leading to increased levels of carbon–carbon inter-units in lignin.[30,54] The spirodienone structure (SD, Table 2) is known to be sensitive and would convert to β-1 type structures, which unfortunately may overlap with signals

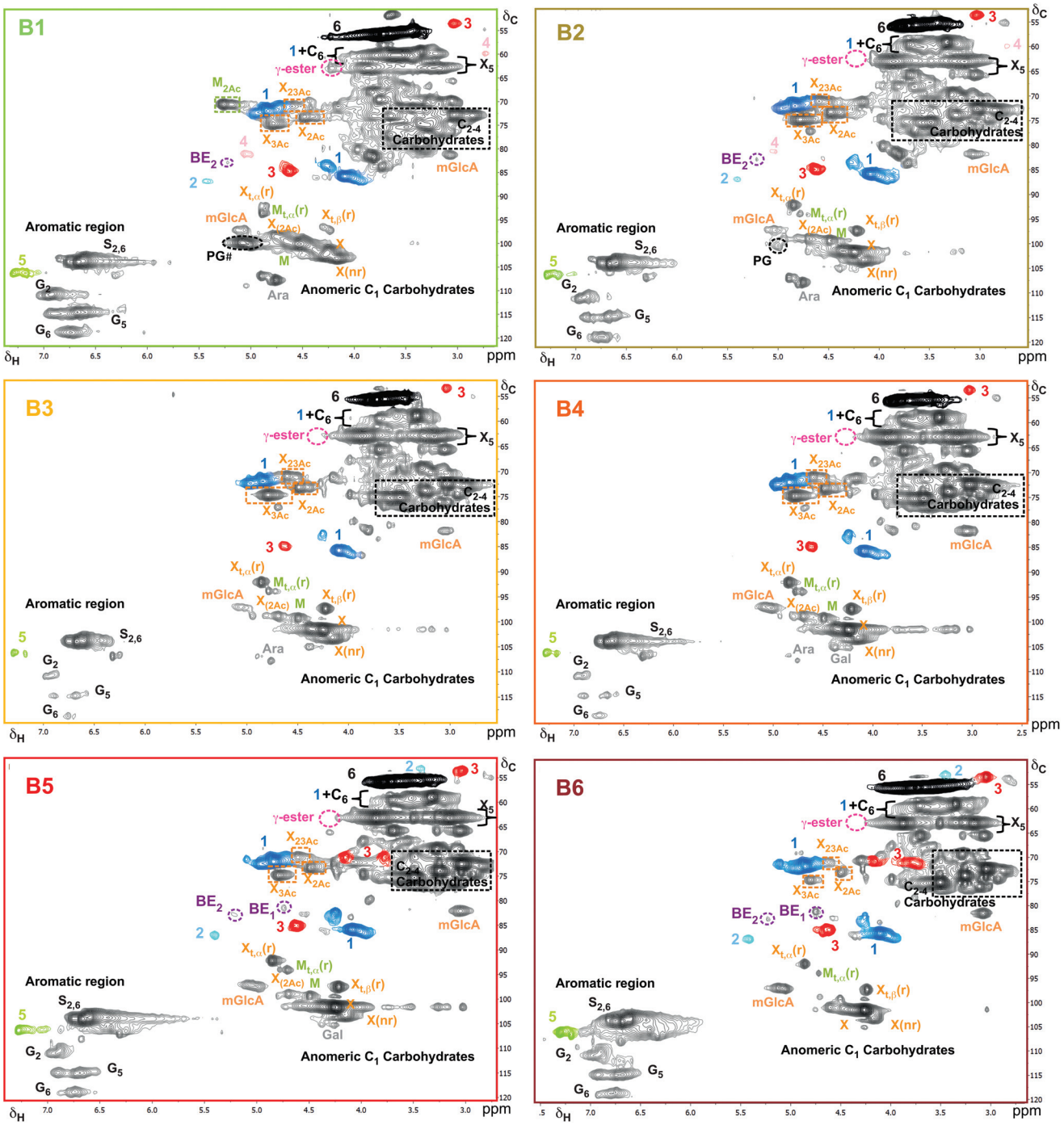

Fig. 4 Heteronuclear single quantum coherence nuclear magnetic resonance (2D HSQC NMR) spectra of birch fractions. The 2D HSQC NMR of all extracts is displayed in using the same nomenclature and color-coding as in Fig. 1A–C. Note: X (xylose); M (mannose); mGlcA (4-*O*-methyl glucuronic acid); t indicates carbon terminal in reducing end (r) or non-reducing end (nr) with anomeric conformation α or β; X_{2Ac} (xylose acetylated in position 2); X_{3Ac} (xylose acetylated in position 3); M_{2Ac} (mannose acetylated in position 2); C (carbon); S (syringyl unit); G (guaiacyl unit). The subscripted numbers indicate the carbon in either the aromatic or sugar ring.

from β-5 type structures. The presence of SD in extracts B1 and B2 may thus indicate that these lignins are probably more native in structure. The β-5 content in hardwoods is generally low and some fractions are devoid of such structures, while the last fraction (B6) had the highest levels. This suggests that β-5 structures constitute part of the recalcitrant network of structures, which are extracted later.

Three types of lignin-carbohydrate bonds are described in the literature, namely phenyl glycoside (PG), benzyl ether (BE) and gamma ester (GE) linkages (Fig. 1C).[55] PG can be clearly

Table 2 Semi-quantitative determination of acetylation regioselectivity and lignin structure by heteronuclear single quantum coherence nuclear magnetic resonance (2D HSQC NMR)

	Acetylation analysis by NMR					
	Acetylation on xylose (X) and mannose (M, mol%)				Degree of acetylation in xylan (DAc, mol%)	
	X_{2Ac}	X_{3Ac}	$X_{2,3Ac}$	M_{2Ac}	By NMR	By HPLC[a]
Error	±0.31	±0.35	±0.83		±1.50	
B1	25.04	26.43	11.13	32.95	62.60	n.c.
B2	17.53	15.03	6.51	–n.d.	39.07	45.3 ± 0.8
B3	8.71	14.06	7.59	n.d.	30.36	28.8 ± 0.8
B4	6.57	10.79	5.87	n.d.	23.23	21.0 ± 0.3
B5	5.72	4.71	2.02	n.d.	12.45	8.5 ± 0.2
B6	3.15	2.25	1.77	n.d.	7.17	6.9 ± 0.6
Lignin structure by NMR						
	S/G	βO4	β5	ββ	SD	Condensed/non condensed
Error	±0.63	±2.42	±0.08	±0.69		
		% C9-units				
B1	2.40	64.52	0.64	6.99	2.48	11.83%
B2	4.17	59.68	0.78	8.06	1.50	14.81%
B3	5.00	64.52		8.33		12.92%
B4	4.16	67.74		9.09		13.42%
B5	4.17	56.45		10.00		17.71%
B6	4.5	42.74	1.52	9.31		25.35%

[a] The DAc was calculated on the basis of the xylose content and assuming all acetyl groups in these extracts are part of the xylan. The nomenclature corresponds with Fig. 1 and 4. Note: n.c.: not calculated.

observed in the initial extracts (B1 and B2) and they are barely noticeable at longer extraction times. The glycosidic bond in this case is between C1-*O* of a Xyl and Man unit and C4 of lignin, assigned according to previous work.[56] This result is consistent with recent reports where PGs were extracted from birch at conditions as mild as 80 °C.[16] It can thus be concluded that these were not created during the SWE treatment. The early extraction of PG suggests that these LCCs are the most easily accessible and soluble. Similarly, γ-ester linkages were detected in the early extracts (B1). These esters are ascribed to linkages between mGlcA and γ-carbon of lignin. The presence of such γ-ester LCCs in the initial extracts has implications for the MS profiling discussed earlier. The mGlcA moieties bound to lignin are likely to be inaccessible to the enzymatic action and therefore will be undetectable by our MS profiling. Since γ-ester LCCs have been reported to arise from the linkage to mGlcA side groups, it seems that a population of them are easily extracted and are not detected at longer extraction times. On the other hand, benzyl ethers begin to appear in the extracts towards the end of the extraction in B5 and B6 fractions. In the B6 extract, where the BE signals are strongest, the β-O-4 content was lower. It is likely that cleavage of some β-O-4 in lignin and glycosidic bonds in hemicellulose, both within the LCC network, rendered the molecule soluble with retained BE linkage. This is also manifested in the SEC of the B6 sample (Fig. 1F). The BEs observed in our study are associated with both pentose and hexose units and could be linked to both Xyl and Man units. The BEs are mainly linked through C2, C3 or C6 hydroxyls of the sugar units. In the case of xylan, only C2 and C3 are available. However, if these are heavily acetylated as we observe, the probability of BE formation is limited and may explain the lower prevalence of BE to xylan when compared to softwoods, as discussed and experimentally evidenced in recent work.[16,30] These results suggest the evolution of extracted LCC with time, starting with phenyl glycosides and γ-esters and ending with benzyl ethers. However, this should be verified by the investigation of the residues, which is a subject for future studies.

Implications of the SWE process for birchwood recalcitrance and for the targeted isolation of hemicellulose and lignin populations

Our proposed sequential SWE protocol without previous delignification, together with the advanced characterization of the fractions by MS-based carbohydrate sequencing and 2D NMR offers new insights on the organization and connectivity of the hemicellulose and lignin populations, and their influence on lignocellulose recalcitrance in birchwood (Fig. 5). Glucomannan populations are easily extractable at short times, which could be related to their higher mobility in the lignocellulosic network and weaker connectivity to cellulose and lignin. On the other hand, glucuronoxylan is progressively enriched at longer extraction times (extracts B2–B4), with relatively stable purity, composition and molecular weight. Hemicellulose acetylation seems to exert an important role on the properties and connectivity of the hemicelluloses, and therefore on biomass recalcitrance. The highest degree of acetylation in extract B1 may be correlated with the mobility of the mannan and xylan populations, which may favour their diffusivity and extractability at shorter times. It is known that the solubility of GX is highly dependent on the acetylation degree. Non-acetylated xylan is not water soluble due to aggre-

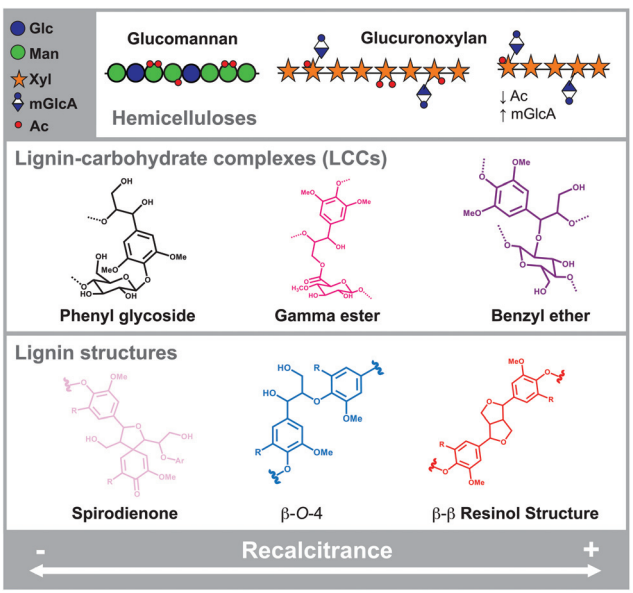

Fig. 5 Correlation between molecular features and recalcitrance.

gation, while high degrees of acetyl-substitution (DAc ≥ 2) again render xylan chains insoluble.[57] Oligosaccharide sequencing reveals the link between glucuronation and acetylation spacing in GX, with the acetyl groups predominantly evenly spaced with respect to the mGlcA. This may indicate an important role modulating the interactions with cellulose surfaces, as previously reported.[12,14] The decrease in acetylation may be caused by autohydrolysis despite the buffered pH used in the extraction. However, the acetyl unit adjacently placed to the mGlcA is relatively stable to degradation, which might indicate a role in regulating LCC formation, as previously suggested.[16] In a similar manner, the acetyl presence at C2 in the later extracts could be a result of protection from lignin attached at C3 as BE. For the more recalcitrant xylan populations (extracts B5 and B6), tighter glucuronation spacing is observed, together with a reduction of the acetylation. This tighter mGlcA spacing may have a role with the connectivity with lignin. However, a derived increase in the γ-ester LCCs is not observed. It is likely that the tighter-spaced uronic acids forming γ-ester and benzyl ethers are part of the same molecule, so when the ester linkage is cleaved during SWE, the mGlcA is retained in the non-dissolved residue due to the recalcitrant BE populations (present on the same molecule). These will only be released when the BE is dissolved through cleavage of aryl ethers and glycosidic bonds in hemicelluloses.

The lignin-carbohydrate complexes have a fundamental role in determining the recalcitrance of hardwoods towards hydrothermal treatment. In this study we can observe that phenyl glycosides are initially extracted with the hemicelluloses during the first stages, whereas γ-esters are progressively cleaved during the extraction, and finally benzyl ethers arise at longer extraction times. On a similar study using autohydrolysis at 180 °C of another hardwood, all three LCC types were detected in the solution phase, which could be due to the substantial cleavage of labile lignin–lignin bonds and glycosidic bonds in hemicelluloses at the higher treatment temperatures.[58] On the other hand, SWE performed at 160 °C on Norwegian spruce and birch showed that the phenyl glycosides were enriched in the extracts while benzyl ethers and γ-esters were enriched in the residue, suggesting the role of stable ester and ether bonds in retaining the hemicelluloses linked to hydrophobic lignin in the residue.[30] This again reinforces the role of LCCs in modulating the resilience of lignocellulosic biomass and the importance of tuning the temperature during SWE for the selective release of hemicellulose and LCC populations. Finally, lignin structure also plays a role in extractability during SWE. Stable content of β-O-4 linkages can be observed at shorter extraction times with a relative decrease at longer exposure, which correlates with a relative increase of the resilient β–β-resinol structures.

The implementation of the SWE process has the potential to provide well-defined hemicellulose fractions in terms of molecular structure and composition. Hemicellulose extraction by SWE has been already reported to be scalable, with the possibility of energy recovery and the use of ultrafiltration systems to concentrate extracts and re-use the effluents.[33] Of course, the technical and economic feasibility will be affected by the sequential nature of the extraction process, and in the end the cost-effectiveness will fully depend on the value of each of the extracts. Further life cycle cost analyses of the sequential process should be performed to cast light on the economic feasibility of the sequential process.

Conclusions

Sequential subcritical water extraction of birchwood using a buffered extraction solvent without previous delignification shows outstanding potential to extract specific hemicellulosic fractions with high yields, purity, and molecular weight. The different extractability of the hemicellulose populations and the control of the extraction times enable the targeted isolation of mannan and xylan populations with controlled glucuronation and acetylation. Mannans could be selectively extracted first, followed by high molar mass xylans with controlled substitution pattern, thus further tuning the solubility and physicochemical properties. These products might be of great potential interest in the food and packaging industry, as texturizing agents or high barrier packaging reinforcements.[18,19] The sequential extraction process using subcritical water is energy efficient, scalable and environmentally benign, and would add valuable fractions to the production chain of wood processing industries, increasing their competitiveness.

The combination of MS-based glycomic sequencing techniques with 2D HSQC NMR provided valuable information about the intramolecular substitution pattern of xylan (acetylation, glucuronation) and its potential role in the supramolecular assembly of birch hardwood, and lignocellulose recalcitrance. The correlation between acetylation and glucuronation in GX from birchwood is demonstrated, with consistent acetylation of xylose units carrying an mGlcA side group and further evenly spaced 2-*O* and 3-*O*-acetylations. This regularly-spaced distribution of acetylation may have an important role in modulating the association with cellulose surfaces, in agreement with similar results from model plant *A. thaliana*[12,14] The sequential SWE also reveals the existence of different xylan domains in birchwood, which could be separated by their recalcitrant nature. The more recalcitrant xylan populations show tighter glucuronation spacing and reduced acetylation, together with further additional lignin interactions by more stable benzyl ether linkages at the expense of depolymerization. This suggests that gamma esters and benzyl ethers maybe closely spaced in the molecule. This study offers new insights on the interconnected molecular features of hemicellulose and lignin structure that control recalcitrance during hydrothermal treatment, with fundamental implications for the development of 2nd generation biorefineries for the integral exploitation of the cellulose, hemicellulose and lignin components in hardwoods.

Experimental

Materials, extraction, and purification

Birchwoood chips (*Betula pendula*) were milled to sawdust (below 1 mm particle size) and washed consecutively with acetone and ethanol. The extractive-free sawdust was subject to SWE at pH 5 (formate buffer 0.2 M), 170 °C and 100 bar in a pressurized liquid extraction equipment (ASE-300, Dionex, USA). Extraction was carried out for 4 hours in total after consecutive cycles of 5, 15, 20, 20, 60 and 120 min, yielding extracts B1 to B6, respectively (Fig. 1D). The extracts were subjected to dialysis using Spectra/Por 3 membranes (Spectrum, USA) with a molecular weight cut-off of 3.5 kDa were used to remove salts and low molecular weight (LMW) molecules. The high molecular weight fractions were subsequently freeze-dried for further analysis.

Compositional analysis of the wood fractions

Monosaccharide composition. The sugar composition of the extracts was determined after acidic methanolysis[59,60] and sulfuric acid hydrolysis.[61] For methanolysis, freeze-dried samples (1 mg) were incubated with 1 mL of 2 M HCl in dry methanol for 5 h at 100 °C. Samples were then neutralized with pyridine, dried under a stream of air, and further hydrolyzed with 2 M TFA at 120 °C for 1 h. The samples were again dried under a stream of air and dissolved in water. A two-step sulfuric acid hydrolysis was performed by incubating 4 mg of sample with 250 μL of 72% H_2SO_4 at room temperature for 3 h; the solution was diluted with 1250 μL of deionized water and subsequently incubated at 100 °C for 3 hours. The monosaccharides were analyzed using high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) with a ICS-3000 system (Dionex) equipped with a CarboPac PA1 column (4 × 250 mm, Dionex), as previously reported.[62] The cellulose content was calculated by subtracting the glucose obtained from methanolysis from the total glucose as analyzed by sulfuric hydrolysis.[59] All experiments were carried out in triplicate.

Acetyl content. The acetyl content was determined after alkaline saponification with 0.8 M NaOH at 70 °C overnight using high performance liquid chromatography with UV detection at 210 nm (HPLC-UV; Dionex-Thermofisher, USA).[63] Separation was achieved with a ROA-Organic acid column (300 × 7.8 mm; Phenomenex, USA) at 50 °C in an isocratic flow of 2.5 mM H_2SO_4 at 0.5 mL min^{-1}.

Lignin content. The lignin content of the fractions was considered by Klason lignin, obtained as described by Effland,[64] together with acid soluble lignin.[65] The former was determined gravimetrically after acid hydrolysis of the extracts, whereas the latter was characterized by UV spectroscopy at 205 nm using an absorptivity of 113 g cm L^{-1}.

Molar mass distributions. The molar mass distributions of the extracted arabinoxylans were analysed by size exclusion chromatography coupled to a refractive index detector (SECcurity 1260, Polymer Standard Services, Mainz, Germany). The extracted fractions were dissolved in the SEC eluent consisting of dimethyl sulfoxide (DMSO, HPLC grade, Scharlab, Sweden) with 0.5% w/w LiBr (ReagentPlus) at 60 °C. The concentrations were adjusted between 0.5–2 g L^{-1} for optimized detection signal. SEC analyses were performed with a flow rate of 0.5 mL min^{-1} at 60 °C using GRAM PreColumn, 30 and 10 000 analytical columns (Polymer Standards Services, Mainz, Germany). Calibration was performed using pullulan standards provided by Polymer Standards Services (PSS, Mainz, Germany).

Profiling and sequencing of glucuronoxylan in birch extracts

Enzymatic digestion and chemical deacetylation. Chemical deacetylation was achieved by incubation in 0.1 M NaOH and 10 mM $NaBH_4$ for 2 h at 80 °C, subsequent neutralization with acetic acid, methanol wash, and air-drying. The xylan fractions were depolymerized using a GH30 *endo*-β-(1 → 4)-glucuronoxylanase (kindly provided by Prof. James F. Preston, University of Florida). End-point digestions were performed by incubating a 1 g L^{-1} solution of each extract with the enzyme (10 U mL^{-1}) for 16 h at 37 °C at the optimum pH of 5.5. After incubation, the solutions were boiled for 10 min, filtered, and kept at −20 °C for further analysis.

Oligomeric mass profiling (OLIMP). Oligomeric mass profiling (OLIMP) was performed by electrospray ionization mass spectrometry (ESI-MS) using a Q-TOF2 mass spectrometer in positive mode (Micromass, UK). The hydrolysates were desalted with HyperSep™ Hypercarb™ cartridges (Thermofischer, UK), dissolved in 50% acetonitrile, 0.1% formic acid and directly infused through a syringe pump at a rate of 5 μL min^{-1}. Capillary and cone voltage were set to 3.3 kV and 80 V, respectively. Mass spectra assignation has been performed using ChemDraw.

Oligosaccharide separation and sequencing by tandem LC-MS/MS. The digested oligosaccharides were chemically labelled and separated by liquid chromatography coupled with electrospray tandem mass spectrometry (LC-MS/MS, Micromass, UK). Derivatization was performed by reductive amination with anthranilic acid as previously described.[66] The labelled oligosaccharides were separated through a SB-C18 column (250 × 4.6 mm, Agilent Technologies) at a flow rate of 10 μL min^{-1} and a gradient of increasing acetonitrile content (10–40%) over 110 min. MS and MS/MS analysis was performed in positive mode at 3.3 kV and 45 V in the capillary and the cone, respectively. Argon was used as collision gas for MS/MS analysis of selected ions, at a voltage of 35–90 V, to analyze the fragmentation of the oligosaccharides. The assignation of the fragmentation spectra has been performed in-house using ChemDraw.

NMR analysis

An amount of 100 mg of sample was dissolved in 750 μl of deuterated DMSO-d_6. NMR spectra were recorded and processed as previously described.[67] All the C/H correlations of the studied inter-unit linkages are reported in ESI (Tables S3–S5†). Carbon 2 of aromatic groups, unsubstituted, was used as internal standard for the semi-quantification.[68] The central

DMSO (δ_C/δ_H = 39.5/2.5 ppm), signal was used as an internal reference. Acetyl content was determined according to previous work.[30] Shortly, 2D HSQC integral values of acetylated carbon cross peaks were normalized by the integral of the anomeric region of the carbohydrates, combined with the molar percentage of the acetylated sugar obtained from HPAEC analysis.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

FV and AMA thank the Swedish Research Council (Project 621-2014-5295) for their financial support. The Knut and Alice Wallenberg research foundation is acknowledged for financial support to NG within the Wallenberg Wood Science Centre.

References

1 R. A. Sheldon, *Green Chem.*, 2014, **16**, 950–963.
2 M. E. Himmel, S.-Y. Ding, D. K. Johnson, W. S. Adney, M. R. Nimlos, J. W. Brady and T. D. Foust, *Science*, 2007, **315**, 804–807.
3 D. Cosgrove and M. Jarvis, *Front. Plant Sci.*, 2012, **3**, 204.
4 A. Ebringerova, Z. Hromadkova and T. Heinze, in *Polysaccharides I*, ed. T. Heinze, 2005, vol. 186, pp. 1–67.
5 A. Teleman, M. Nordström, M. Tenkanen, A. Jacobs and O. Dahlman, *Carbohydr. Res.*, 2003, **338**, 525–534.
6 A. Teleman, M. Tenkanen, A. Jacobs and O. Dahlman, *Carbohydr. Res.*, 2002, **337**, 373–377.
7 O. Y. Abdelaziz, D. P. Brink, J. Prothmann, K. Ravi, M. Sun, J. García-Hidalgo, M. Sandahl, C. P. Hulteberg, C. Turner, G. Lidén and M. F. Gorwa-Grauslund, *Biotechnol. Adv.*, 2016, **34**, 1318–1346.
8 M. C. McCann and N. C. Carpita, *J. Exp. Bot.*, 2015, **66**, 4109–4118.
9 C. G. Yoo, Y. Yang, Y. Pu, X. Meng, W. Muchero, K. L. Yee, O. A. Thompson, M. Rodriguez, G. Bali, N. L. Engle, E. Lindquist, V. Singan, J. Schmutz, S. P. DiFazio, T. J. Tschaplinski, G. A. Tuskan, J.-G. Chen, B. Davison and A. J. Ragauskas, *Green Chem.*, 2017, **19**, 5467–5478.
10 S.-Y. Ding, Y.-S. Liu, Y. Zeng, M. E. Himmel, J. O. Baker and E. A. Bayer, *Science*, 2012, **338**, 1055–1060.
11 R. L. Silveira, S. R. Stoyanov, S. Gusarov, M. S. Skaf and A. Kovalenko, *J. Am. Chem. Soc.*, 2013, **135**, 19048–19051.
12 N. J. Grantham, J. Wurman-Rodrich, O. M. Terrett, J. J. Lyczakowski, K. Stott, D. Iuga, T. J. Simmons, M. Durand-Tardif, S. P. Brown, R. Dupree, M. Busse-Wicher and P. Dupree, *Nat. Plants*, 2017, **3**, 859–865.
13 M. Busse-Wicher, A. Li, R. L. Silveira, C. S. Pereira, T. Tryfona, T. C. F. Gomes, M. S. Skaf and P. Dupree, *Plant Physiol.*, 2016, **171**, 2418–2431.
14 M. Busse-Wicher, T. C. F. Gomes, T. Tryfona, N. Nikolovski, K. Stott, N. J. Grantham, D. N. Bolam, M. S. Skaf and P. Dupree, *Plant J.*, 2014, **79**, 492–506.
15 A. Martínez-Abad, J. Berglund, G. Toriz, P. Gatenholm, G. Henriksson, M. Lindström, J. Wohlert and F. Vilaplana, *Plant Physiol.*, 2017, **175**, 1579–1592.
16 N. Giummarella and M. Lawoko, *ACS Sustainable Chem. Eng.*, 2016, **4**, 5319–5326.
17 M. Lawoko, G. Henriksson and G. Gellerstedt, *Biomacromolecules*, 2005, **6**, 3467–3473.
18 A. Martínez-Abad, A. C. Ruthes and F. Vilaplana, *J. Appl. Polym. Sci.*, 2016, **133**(2), 709.
19 *Food Polysaccharides and Their Applications*, ed. A. M. Stephen, G. O. Phillips and P. A. Williams, CRC Press, Boca Raton, FL, USA, 2006.
20 C. Crestini, H. Lange, M. Sette and D. S. Argyropoulos, *Green Chem.*, 2017, **19**, 4104–4121.
21 O. Sevastyanova, M. Helander, S. Chowdhury, H. Lange, H. Wedin, L. Zhang, M. Ek, J. F. Kadla, C. Crestini and M. E. Lindström, *J. Appl. Polym. Sci.*, 2014, **131**, 40799.
22 A. Duval, F. Vilaplana, C. Crestini and M. Lawoko, *Holzforschung*, 2016, **70**, 11.
23 A. S. Jääskeläinen, T. Liitiä, A. Mikkelson and T. Tamminen, *Ind. Crops Prod.*, 2017, **103**, 51–58.
24 A. J. Ragauskas, G. T. Beckham, M. J. Biddy, R. Chandra, F. Chen, M. F. Davis, B. H. Davison, R. A. Dixon, P. Gilna, M. Keller, P. Langan, A. K. Naskar, J. N. Saddler, T. J. Tschaplinski, G. A. Tuskan and C. E. Wyman, *Science*, 2014, **344**(6185), 1246843.
25 A. Duval and M. Lawoko, *React. Funct. Polym.*, 2014, **85**, 78–96.
26 S. Sen, S. Patil and D. S. Argyropoulos, *Green Chem.*, 2015, **17**, 4862–4887.
27 W.-J. Liu, H. Jiang and H.-Q. Yu, *Green Chem.*, 2015, **17**, 4888–4907.
28 R. Rinaldi, R. Jastrzebski, M. T. Clough, J. Ralph, M. Kennema, P. C. A. Bruijnincx and B. M. Weckhuysen, *Angew. Chem., Int. Ed.*, 2016, **55**, 8164–8215.
29 W. Zhao, B. Simmons, S. Singh, A. Ragauskas and G. Cheng, *Green Chem.*, 2016, **18**, 5693–5700.
30 N. Giummarella and M. Lawoko, *ACS Sustainable Chem. Eng.*, 2017, **5**, 5156–5165.
31 L. Capolupo and V. Faraco, *Appl. Microbiol. Biotechnol.*, 2016, **100**, 9451–9467.
32 J. V. Rissanen, H. Grénaman, C. Xu, S. Willför, D. Y. Murzin and T. Salmi, *ChemSusChem*, 2014, **7**, 2947–2953.
33 P. O. Kilpeläinen, S. S. Hautala, O. O. Byman, L. J. Tanner, R. I. Korpinen, M. K. J. Lillandt, A. V. Pranovich, V. H. Kitunen, S. M. Willför and H. S. Ilvesniemi, *Green Chem.*, 2014, **16**, 3186–3194.
34 P. A. Penttilä, P. Kilpeläinen, L. Tolonen, J. P. Suuronen, H. Sixta, S. Willför and R. Serimaa, *Cellulose*, 2013, **20**, 2335–2347.
35 P. Kilpeläinen, K. Leppänen, P. Spetz, V. Kitunen, H. Ilvesniemi, A. Pranovich and S. Willför, *Nord. Pulp Pap. Res. J.*, 2012, **27**, 680–688.

36 T. E. Timell, *Wood Sci. Technol.*, 1967, **1**, 45–70.
37 M. G. Handford, T. C. Baldwin, F. Goubet, T. A. Prime, J. Miles, X. Yu and P. Dupree, *Planta*, 2003, **218**, 27–36.
38 J. S. Kim and G. Daniel, *Planta*, 2012, **236**, 1367–1379.
39 O. Grigoray, J. Järnström, E. Heikkilä, P. Fardim and T. Heinze, *Carbohydr. Polym.*, 2014, **112**, 308–315.
40 F. M. Yedro, D. A. Cantero, M. Pascual, J. García-Serna and M. J. Cocero, *Bioresour. Technol.*, 2015, **191**, 124–132.
41 K. Leppänen, P. Spetz, A. Pranovich, K. Hartonen, V. Kitunen and H. Ilvesniemi, *Wood Sci. Technol.*, 2011, **45**, 223–236.
42 S. Azhar, G. Henriksson, H. Theliander and M. E. Lindström, *Carbohydr. Polym.*, 2015, **117**, 169–176.
43 X. Chen, M. Lawoko and A. v. Heiningen, *Bioresour. Technol.*, 2010, **101**, 7812–7819.
44 F. J. St John, J. C. Hurlbert, J. D. Rice, J. F. Preston and E. Pozharski, *J. Mol. Biol.*, 2011, **407**, 92–109.
45 J. R. Bromley, M. Busse-Wicher, T. Tryfona, J. C. Mortimer, Z. Zhang, D. M. Brown and P. Dupree, *Plant J.*, 2013, **74**, 423–434.
46 E. A. Rennie, S. F. Hansen, E. E. K. Baidoo, M. Z. Hadi, J. D. Keasling and H. V. Scheller, *Plant Physiol.*, 2012, **159**, 1408–1417.
47 S. L. Chong, M. Derba-Maceluch, S. Koutaniemi, L. D. Gómez, S. J. McQueen-Mason, M. Tenkanen and E. J. Mellerowicz, *BMC Biotechnol.*, 2015, **15**, 56.
48 P. M. A. Pawar, M. Derba-Maceluch, S. L. Chong, L. D. Gómez, E. Miedes, A. Banasiak, C. Ratke, C. Gaertner, G. Mouille, S. J. McQueen-Mason, A. Molina, A. Sellstedt, M. Tenkanen and E. J. Mellerowicz, *Plant Biotechnol. J.*, 2016, **14**, 387–397.
49 M. Busse-Wicher, T. C. F. Gomes, T. Tryfona, N. Nikolovski, K. Stott, N. J. Grantham, D. N. Bolam, M. S. Skaf and P. Dupree, *Plant J.*, 2014, **79**, 492–506.
50 S. L. Chong, L. Virkki, H. Maaheimo, M. Juvonen, M. Derba-Maceluch, S. Koutaniemi, M. Roach, B. Sundberg, P. Tuomainen, E. J. Mellerowicz and M. Tenkanen, *Glycobiology*, 2014, **24**, 494–506.
51 P. M.-A. Pawar, M. Derba-Maceluch, S.-L. Chong, L. D. Gómez, E. Miedes, A. Banasiak, C. Ratke, C. Gaertner, G. Mouille, S. J. McQueen-Mason, A. Molina, A. Sellstedt, M. Tenkanen and E. J. Mellerowicz, *Plant Biotechnol. J.*, 2016, **14**, 387–397.
52 B. Domon and C. E. Costello, *Glycoconjugate J.*, 1988, **5**, 397–409.
53 R. Fauré, C. M. Courtin, J. A. Delcour, C. Dumon, C. B. Faulds, G. B. Fincher, S. Fort, S. C. Fry, S. Halila, M. A. Kabel, L. Pouvreau, B. Quemener, A. Rivet, L. Saulnier, H. A. Schols, H. Driguez and M. J. O'Donohue, *Aust. J. Chem.*, 2009, **62**, 533–537.
54 M. Leschinsky, G. Zuckerstätter, K. Weber Hedda, R. Patt and H. Sixta, *Holzforschung*, 2008, **62**, 653.
55 D. Fengel and G. Wegener, in *Wood chemistry, ultrastructure and reactions*, ed. W. De Gruyter, Berlin, 1984, pp. 167–174.
56 Y. Miyagawa, M. Takahito, H. Kamitakahara and T. Takano, *Holzforschung*, 2014, **68**, 747–760.
57 P. M. A. Pawar, S. Koutaniemi, M. Tenkanen and E. J. Mellerowicz, *Front. Plant Sci.*, 2013, **4**, 118.
58 R. H. Narron, H.-m. Chang, H. Jameel and S. Park, *ACS Sustainable Chem. Eng.*, 2017, **5**, 10763–10771.
59 M. M. Appeldoorn, M. A. Kabel, D. Van Eylen, H. Gruppen and H. A. Schols, *J. Agric. Food Chem.*, 2010, **58**, 11294–11301.
60 F. Bertaud, A. Sundberg and B. Holmbom, *Carbohydr. Polym.*, 2002, **48**, 319–324.
61 J. F. Saeman, W. E. Moore, R. L. Mitchell and M. A. Millett, *Tappi J.*, 1954, **37**, 336–343.
62 L. S. McKee, H. Sunner, G. E. Anasontzis, G. Toriz, P. Gatenholm, V. Bulone, F. Vilaplana and L. Olsson, *Biotechnol. Biofuels*, 2016, **9**, 1–13.
63 A. G. J. Voragen, H. A. Schols and W. Pilnik, *Food Hydrocolloids*, 1986, **1**, 65–70.
64 M. J. Effland, *Tappi J.*, 1977, **60**, 143–144.
65 TAPPI, Journal, 1991, TAPPI UM 250.
66 P. Mischnick, in *Mass Spectrometry of Polymers*, ed. M. Hakkarainen, Springer, Berlin Heidelberg, 2012, ch. 134, vol. 248, pp. 105–174.
67 N. Giummarella, L. Zhang, G. Henriksson and M. Lawoko, *RSC Adv.*, 2016, **6**, 42120–42131.
68 M. Sette, R. Wechselberger and C. Crestini, *Chem. – Eur. J.*, 2011, **17**, 9529–9535.

