Cellulose (2020) 27:7929–7953
https://doi.org/10.1007/s10570-020-03317-0

ORIGINAL RESEARCH

2D Assignment and quantitative analysis of cellulose and oxidized celluloses using solution-state NMR spectroscopy

Tetyana Koso · Daniel Rico del Cerro · Sami Heikkinen · Tiina Nypelö · Jean Buffiere · Jesus E. Perea-Buceta · Antje Potthast · Thomas Rosenau · Harri Heikkinen · Hannu Maaheimo · Akira Isogai · Ilkka Kilpeläinen · Alistair W. T. King

Received: 15 April 2020 / Accepted: 25 June 2020 / Published online: 27 July 2020

Abstract The limited access to fast and facile general analytical methods for cellulosic and/or biocomposite materials currently stands as one of the main barriers for the progress of these disciplines. To that end, a diverse set of narrow analytical techniques are typically employed that often are time-consuming, costly, and/or not necessarily available on a daily basis for practitioners. Herein, we rigorously demonstrate a general quantitative NMR spectroscopic method for structural determination of crystalline cellulose samples. Our method relies on the use of a readily accessible ionic liquid electrolyte, tetrabutylphosphonium acetate ([P_{4444}][OAc]):DMSO-d_6, for the direct dissolution of biopolymeric samples. We utilize a series of model compounds and apply now classical (nitroxyl-radical and periodate) oxidation reactions to cellulose samples, to allow for accurate resonance assignment, using 2D NMR. Quantitative heteronuclear single quantum correlation (HSQC) was applied in the analysis of key samples to assess its applicability as a high-resolution technique for following cellulose surface modification. Quantitation using HSQC was possible, but only after applying T_2 correction to integral values. The comprehensive signal assignment of the diverse set of cellulosic species in this study constitutes a blueprint for the direct quantitative structural elucidation of crystalline lignocellulosic, in general, readily available solution-state NMR spectroscopy.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10570-020-03317-0) contains supplementary material, which is available to authorized users.

T. Koso (✉) · D. Rico del Cerro · S. Heikkinen · J. E. Perea-Buceta · I. Kilpeläinen · A. W. T. King (✉)
Materials Chemistry Division, Department of Chemistry, Faculty of Science, University of Helsinki, Kumpula Campus, Helsinki, Finland
e-mail: tetyana.koso@helsinki.fi

A. W. T. King
e-mail: alistair.king@helsinki.fi

T. Nypelö
Division of Applied Chemistry, Department of Chemistry and Chemical Engineering, Chalmers University of Technology, Gothenburg, Sweden

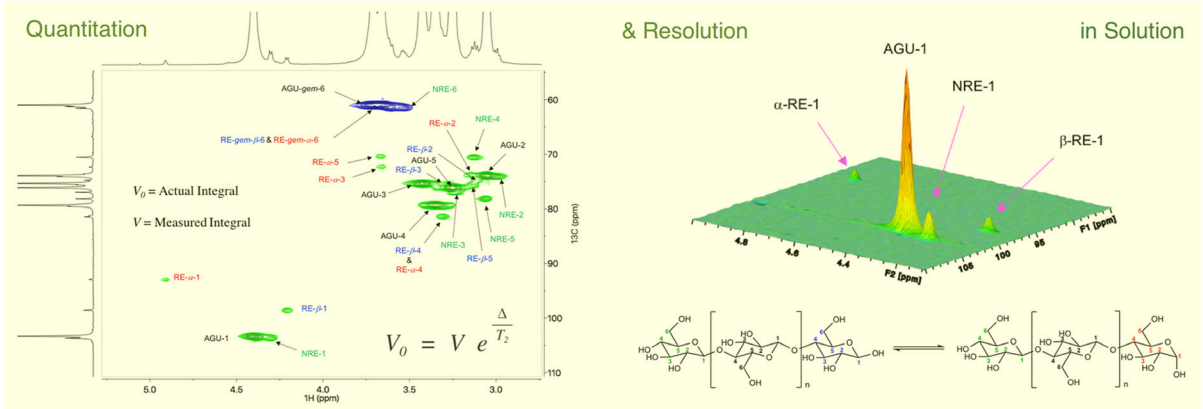
J. Buffiere
Department of Bioproducts and Biosystems, School of Chemical Engineering, Aalto University, Espoo, Finland

A. Potthast · T. Rosenau
Institute for Chemistry of Renewables, Department of Chemistry, University of Natural Resources and Life Sciences Vienna (BOKU), Wien, Austria

H. Heikkinen
The Finnish Institute for the Verification of the Chemical Weapons Convention (VERIFIN), Helsinki, Finland

H. Maaheimo
VTT Technical Research Centre of Finland Ltd., Espoo, Finland

Graphic abstract



Keywords TEMPO · Nitroxyl radical · Periodate · Ionic liquid · Cellulose dissolution · Quantitative HSQC

Introduction

Surface chemical modification of cellulosic materials is a logical approach to tune the properties and, thus, applicability of these bio-renewable polymers (Klemm et al. 2005). Unlike in small-molecule-based chemical disciplines, with cellulosics, currently there is no established general quantitative analytical technique to accurately assess chemical changes with sufficient resolution. This is in large part due to the poor solubility of cellulosic materials in common molecular solvents, preventing non-destructive solution-state analyses. This limitation has imposed researchers to rely on the poorer chemical resolution of solid-state techniques or indirect methods for characterization of samples, which contain a significant phase composition of crystalline cellulose. Typically, a succession of direct and indirect methods are applied for this task, affording partial insights. However, this process is often lengthy as a whole and its threads are difficult to bring together. Solid-state NMR, in particular, has found utility in the quantification of the different crystalline phases in celluloses (Newman 1999; Kono et al. 2002; Zuckerstätter et al. 2009). High resolution, using ultra-fast magic angle spinning (MAS), and multidimensional experiments are possible for solid-state NMR. However, spectral resolution is rather limited using typical MAS probes, preventing the accurate separation and quantitation of different chemical species. In addition, T_1 relaxation times in the solid state are typically very long, requiring labeling strategies to give sufficient signal-to-noise (S/N) for quantitative experiments. Chemical modification of nanocelluloses (Habibi et al. 2010), which by nature involves regioselective surface chemistry, represents a significant challenge due to the infancy of the field and complexity of the materials. This has been compounded by a flood of conceptual articles applying chemistries but which lack the analytical rigor of traditional chemistry disciplines. For accurate definition of feedstocks and reaction products, a multitude of complimentary methods are commonly used (Foster et al. 2018). However, until recently the one irreplaceable method to organic chemistry (solution-state NMR) has not seriously been considered.

The proposed solution-state NMR technique is practically applicable to all crystalline celluloses and even whole biomass samples, provided the molecular weights are not high enough to reduce spectral resolution and S/N (due to relaxation effects), as demonstrated by Holding et al. (2016). However, a specific requirement for a useable NMR method is to

A. Isogai
Department of Biomaterial Sciences, The University of Tokyo, Tokyo, Japan

have accurate assignments for the structural features common to the most studied samples. Another specific requirement is to allow for quantitation of chemical species, which is somewhat limited using NMR to analyze polymeric samples, even using basic 1D solution-state experiments. This is particularly difficult for the more complicated chemical modifications, or whole biomass samples, which would require resolution in several dimensions.

Previously, we published the use of a novel ionic liquid electrolyte, tetrabutylphosphonium acetate ([P_{4444}][OAc]):DMSO-d_6 for the solution-state NMR analysis of nanocelluloses (King et al. 2018). The choice of the ionic liquid electrolyte was discussed, in detail, in previous articles. (Deb et al. 2016; King et al. 2018) However, the choice is very much related to the high stability of tetraalkylphosphonium cations preventing reaction with solutes and, thus, artifact formation. In addition, the ability to dissolve cellulose efficiently at such low ratios of [P_{4444}][OAc] to DMSO allows for low viscosity solutions, thus, higher resolution spectra. Furthermore, as [P_{4444}][OAc] signals do not overlap with the cellulose resonances, in the ^{1}H and ^{13}C ppm domains, which makes [P_{4444}][OAc] ideal for this purpose. Direct-dissolution NMR solvents, based on the use of 1-ethyl-3-methylimidazolium acetate ([emim][OAc]) (Cheng et al. 2013) or tetrabutylammonium fluoride ([N_{4444}]F) (Heinze et al. 2000; Östlund et al. 2009), are problematic for fine chemical analysis of cellulosics. [emim][OAc] is known to react with cellulose (Liebert and Heinze 2008; Ebner et al. 2008; Clough et al. 2015) and high purity [N_{4444}]F is very unstable in non-protic solvents (Sun and DiMagno 2005). Unfortunately, both also have signals that overlap with the cellulose backbone resonances. Alternative low-cost and unreactive perdeuterated cellulose solvents have also not yet appeared.

In this work, we provide thorough characterization for a few different cellulose substrates, using the ([P_{4444}][OAc]):DMSO-d_6 electrolyte, before and after applying common oxidation schemes. Three cellulose substrates were used. The first is low degree of polymerization-cellulose nanocrystals (LDP-CNC), isolated by super-critical water (sc-H_2O) extraction of microcrystalline cellulose (MCC) (Buffiere et al. 2016). This was used as it is quite low molecular weight, offering good spectral resolution for signal assignment. The second was pristine cellulose nanocrystals (CNCs) derived from cotton. This is a representative CNC sample, also with relatively low molecular weight. The third substrate was MCC, a common cellulose model compound. The reaction products include cellulose which has been oxidized using either of two synthetically significant methods: 1) periodate oxidation (Kim et al. 2000; Nypelö et al. 2018) or 2) nitroxyl-radical (e.g., TEMPO)-oxidation (Isogai et al. 2011, 2018). The spin-systems are assigned (polymeric and terminal units) using a range of common NMR methods and with the help of the monomeric (glucose, gluconic acid and glucuronic acid) and dimeric model compounds (cellobiose and cellobionic acid). Standard heteronuclear single-quantum correlation (HSQC) NMR experiments are not quantitative. Therefore, a suitable quantitative HSQC sequence was tested, with and without T_2 correction, to demonstrate the accuracy of separation and quantitation of key chemical species, before and after oxidation. The results aim to illustrate the potential of this method, not only for analysis of cellulose and cellulose derivatives, but also as a method to improve quantitation in analysis of lignocellulosics in general.

Materials and methods

Raw materials and preparation of oxidized celluloses

MCC ($DP_{N\text{-}GPC}$ 153) was purchased from Sigma-Aldrich. The LDP-CNCs ($DP_{N\text{-}GPC}$ 37, 15 wt% dispersed in water) were the precipitated 'residue' from the sc-H_2O extraction of microcrystalline cellulose (MCC), as described by Buffiere et al. (2016). They were freeze-dried before use to remove as much free water as reasonably possible. Nitroxyl-radical oxidation of the LDP-CNCs was carried out in the $NaClO/NaClO_2$ system in the presence of 4-acetamido-2,2,6,6-tetramethylpiperidine 1-oxyl (4-AcNH-TEMPO), under acidic conditions (pH 5.8), according to Hirota et al. (2009). This yielded 4-AcNH-TEMPO-oxidized LDP-CNCs (TOx-LDP-CNCs). The prepared sodium polyglucuronic acid salt form of the TOx-LDP-CNCs was acidified to pH 1.0 and separated by centrifugation with subsequent water washing and freeze-drying. Pinnick oxidation of the reducing ends of LDP-CNCs was carried out under acidic conditions (pH 5.0) in the presence of one

weight equivalent of $NaClO_2$, to yield Pinnick-oxidized LDP-CNCs (POx-LDP-CNCs). The prepared salt form was acidified to pH 1.0 and further processed, as described above. Periodate-oxidized CNCs ($NaIO_4$-CNCs) were prepared from pristine CNCs (prepared from cotton by hydrolysis with H_2SO_4 followed by desulfation with HCl), as described in Nypelö et al. (2018). After oxidation with sodium periodate, a film from the oxidized CNCs was cast in a Petri dish, by initial sonication of cellulosic dispersion, casting and air-drying. A mixture of gluconic acid and the corresponding lactone was prepared by evaporation of an aqueous gluconic acid solution (49–53 wt%) in a rotary evaporator. Detailed procedures can be found in the Supporting Information. Cellobionic acid was purchased from Aldox, Dept. of Food Science and Technology, BOKU, Vienna, Austria. All other solvents and chemicals were commercially available from Sigma-Aldrich and VWR, except DMSO-d_6 (Eurisotop) and 4-AcNH-TEMPO (TCI Europe) and used as received, without further purification. More detailed information can be found in the Supporting Information.

Preparation of ionic liquid electrolyte and cellulosic samples

Tetrabutylphosphonium acetate ([P_{4444}][OAc]) was synthesized according to an optimized method by King et al. (2018). Briefly: Tetrabutylphosphonium chloride ([P_{4444}]Cl) was prepared from tri-*n*-butylphosphine by reaction with *n*-butyl chloride in a teflon-lined Parr acid-digestion reactor. [P_{4444}][OAc] was obtained by metathesis reaction of [P_{4444}]Cl with potassium acetate (KOAc) in isopropyl alcohol and purified by precipitation from chloroform to remove residual salts.

A stock solution of [P_{4444}][OAc]:DMSO-d_6 (20:80 wt%) for solution-state NMR analysis of the cellulosic materials was simply prepared by dissolution of crystalline [P_{4444}][OAc] into DMSO-d_6, in the w/w ratio of 1:4. Direct contact of [P_{4444}][OAc] with air was minimized, to avoid moisture uptake. It may be preferable to use a glove-box or argon flush especially in humid climates. In a typical sample dissolution procedure, 50 mg of dry cellulosic material were introduced into 950 mg of [P_{4444}][OAc]:DMSO-d_6 stock electrolyte, in 4 mL sealed vials, equipped with stirring bars. These were initially stirred at RT to see if the samples dissolved. If not, they were heated at 65 °C under inert atmosphere. Once the solutions were clear and visually isotropic, the samples were transferred into 5 mm NMR tubes (Wilmad-Labglass Co., USA) for analysis.

NMR experiments

Spectra were recorded using a Bruker Avance 600 MHz Avance III or NEO spectrometers. The majority of the experiments were recorded using a SmartProbe™ optimized for X-nucleus detection. For some samples, an inverse triple resonance probe-head ($^1H/^{19}F$, ^{13}C, ^{31}P) or a cryogenically-cooled quadruple resonance (^1H, ^{13}C, ^{31}P, ^{15}N) probe-head were used.

The key NMR experiments are as follows:

- Standard ^1H and ^{13}C 1D experiments were recorded for all samples. In some cases, instead of simple 1D ^{13}C experiments, ^{13}C (refocused) insensitive nuclei enhanced by polarization transfer (INEPT) experiments were recorded. They provided $> 2 \times$ improvement in S/N, at the expense of the loss of quaternary signals.
- Quantitative ^{13}C (inverse-gated ^1H-decoupling), was run for the 4-AcNH-TEMPO oxidized LDP-CNC sample, with a repetition delay of 8 s and a 30° pulse (King et al. 2018).
- Diffusion-edited 1D ^1H experiments were measured for all polymeric samples using a 1D bipolar-pulse pair stimulated echo (BPPSTE) pulse sequence (‘ledbpgp2s1d’ in the Bruker TopSpin 4.0 pulse program library).
- Multiplicity-edited HSQC (Willker et al. 1993), experiments (‘hsqcedetgp’, or ‘hsqcedetgpsisp2.2’ for increased sensitivity, in the Bruker TopSpin 4.0 pulse program library) were recorded for all samples.
- Quantitative Carr-Purcell-Meiboom-Gill (CPMG)-adjusted HSQC (Q-CAHSQC) experiments (Koskela et al. 2005) were recorded for the LDP-CNC, TOx-LDP-CNC and MCC samples. The sequence (‘qcahsqc’) was obtained directly from Bruker.
- Diffusion and multiplicity-edited HSQCs were measured for low LDP-CNC and TOx-LDP-CNC cellulose samples. 2D HSQC-total correlation spectroscopy (HSQC-TOCSY) (Schleucher et al. 1994) experiments (‘hsqcdietgpsisp.2’ in the

Bruker TopSpin 4.0 pulse program library), with short (15 ms) and long (120 ms) TOCSY mixing times, were recorded for all required samples to aid in resonance assignment.

- Heteronuclear multiple bond correlation (HMBC) (Bax and Summers 1986), experiments ('hmbcgplpndqf' in the Bruker TopSpin 4.0 pulse program library) were recorded for LDP-CNC and TOx-LDP-CNC cellulose samples.

All NMR measurements were conducted at a sample temperature of 65 °C. Typically, the time-domain size in the indirect ^{13}C-dimension (f1) for HSQC was 1024 and HMBC was 512, corresponding to 512 (td1/2) and 512 (td1) actual t_1-increments in the real data, for phase sensitive HSQC sequences and the magnitude mode HMBC sequence, respectively. High digital resolution was used as most samples were quite low molecular weight. Chemical shifts in ^{1}H and ^{13}C ppm scales were calibrated against the DMSO-d_6 signals (2.50 ppm for residual ^{1}H and 39.52 ppm for ^{13}C). All spectra were processed using Bruker TopSpin 4.0.6 (https://bruker.com/) and/or MestReNova 10.0.2 (https://mestrelab.com/) software. Further 1D data processing was completed using *Fityk* 1.3.1 (Wojdyr 2010) (https://fityk.nieto.pl/). Full NMR experimental and conditions are given in the Supporting Information.

Results and discussion

Cellulose model and methodology choice

For NMR analysis, the samples were dissolved in the [P_{4444}][OAc]:DMSO-d_6 (1:4 wt%) electrolyte at 5 wt%, at as low temperatures as possible (typically 25–80 °C). For the low DP samples this occurred rapidly at RT. This concentration of cellulosic materials allowed for detection of the low-intensity signals, such as the chain ends in the polymeric samples. All spectra were collected at the elevated temperature of 65 °C, as it offers further improvement in resolution and S/N, due to longer spin–lattice (T_1) and spin–spin (T_2) relaxation times. It is known that T_2 increases with elevated temperature, due to an approximate inverse-law relationship between T_2 and viscosity (Kim 2008). Significant improvements in resolution were also previously observed for the case of MCC dissolved in the homologous methyltrioctylphosphonium acetate ([P_{8881}][OAc]):DMSO-d_6 system (Holding et al. 2016).

Literature resonance assignment of the most basic monomeric units of oxidized celluloses are incomplete. Thus, a range of monomers and dimers were studied, specifically in the [P_{4444}][OAc]:DMSO-d_6 (1:4 wt%) electrolyte. Cellulose is also complicated by the fact that there are non-reducing end (NRE) and anomeric reducing-end (RE) units that differ in their chemical shifts from the corresponding species in the polymeric units. Separation of these species, using 2D correlation methods, is not guaranteed for high molecular weight samples. Hence, monomeric and dimeric models are described, in addition to the LDP-CNC sample ($DP_{N\text{-}GPC}$ 37). This sample is rather unique in the fact that there are not many sources of low DP cellulose accessible, in large enough quantities, for synthesis and assignment of the products.

The NMR spectra of native celluloses can provide information on the average chain length of the polymer, as the signals of the reducing end and non-reducing end are relatively well separated (King et al. 2018; Heise et al. 2019; Holding et al. 2016). However, the characterization of modified cellulose samples can be complicated, as both the location and substitution pattern may vary along the polymer chain, and the high molecular weight can preclude using more sophisticated NMR techniques. In addition, as the literature data on relevant monomeric units of oxidized cellulose is incomplete, a range of monomeric, dimeric and oligomeric models were chosen or prepared to aid the spectral interpretation of the oxidized samples. These include: glucose, cellobiose, LDP-CNC, glucuronic acid, gluconic acid and cellobionic acid.

Chemical shift assignment of cellulose and modified units

The full assignment for the dimers, polymeric units and terminal units in this study are shown in Fig. 1. These are in the [P_{4444}][OAc]:DMSO-d_6 (1:4 wt%) electrolyte at 65 °C, referenced against DMSO-d_6 (residual ^{1}H at 2.5 ppm and ^{13}C at 39.52 ppm). The following will be a description of how the assignments are made and further aspects of the study.

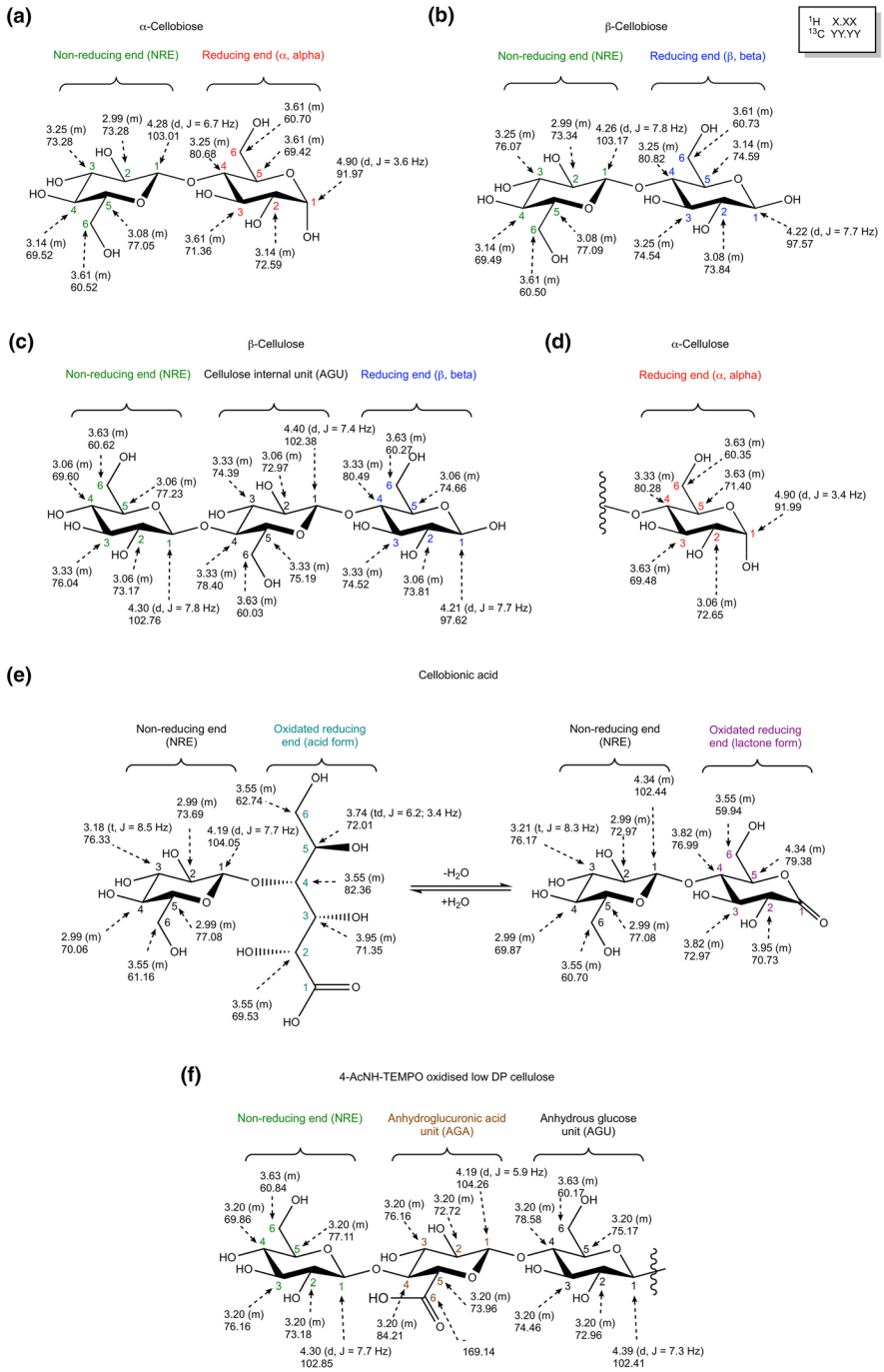


Fig. 1 Representative structures, atom numerations and resonance assignments for ^{1}H and ^{13}C NMR sets of: **a** α-anomer of cellobiose; **b** β-anomer of cellobiose; **c** LDP-CNC β-anomer of cellulose (DP_N 37); **d** LDP-CNC α-anomer of cellulose (DP_N 37); **e** equilibrium between cellobionic acid (turquoise) and cellobionolactone (purple); **f** 4-AcNH-TEMPO oxidized LDP-CNCs

Monomers

The HSQC, ^{13}C and HSQC-TOCSY spectra for the monomers are shown in Fig. S5-S13; glucose (Fig. S5-S8), gluconic acid, (Fig. S9-S11) and glucuronic acid (Fig. S12-S13). Tabulated chemical shift data, along with previous literature assignments, are shown in Tables S1 and S3: glucose (Table S1) (Roslund et al. 2008) and glucuronic acid (Table S3) (Agrawal 1992). Our assignments did not change significantly from those of the literature assignments.

Cellobiose

Cellobiose (Fig. 1a, b) consists of α an β anomers giving 24 identifiable correlations in HSQC and resonances the ^{13}C spectrum (Fig. 2a, b). The ratio of anomers is 34:66 (α: β) by ^{1}H NMR (by integration). HSQC-TOCSYs are shown in Fig. S14-S18, in the Supporting Information, allowing for complete assignment. The tabulated chemical shift data is also given in the Supporting Information (Table S4-S5), along with previous literature assignments in D_2O (Roslund et al. 2008) which do not show any major deviations from our data. With cellobiose, the spectra start to become quite a bit more complex than for the monomers, with many overlapping peaks in the C2–C5 region (65–85 ppm in the ^{13}C domain). HSQC-TOCSY spectra with short mixing times (15 ms, Fig. S14), which provides COSY-like correlations, was most useful for tracing the complete spin-systems for the anomers. In the case of cellobiose, which is not polymeric, we still term the glucopyranose with the hemiacetal anomeric carbon atom as the RE and the one with the glycosidic C1 as the NRE. HSQC-TOCSY with long mixing time (120 ms, Fig. S18), was used to easily visualize the corresponding TOCSY correlations for the separate RE and NRE spin-systems. Full assignments are given in Fig. 1a, b and Table S4-S5. At 2048 indirect (f1) increments in the HSQC (Fig. 2a) the resolution starts to approach that of the ^{13}C spectrum.

In the assignment of cross-peaks, we have tried to be consistent with the color labeling of the assigned correlations: (1) NREs are labeled in green; (2) nonmodified internal AGU correlations are labeled in black; (3) α and β REs are labeled in red and blue, respectively; (4) oxidized internal AGUs, anhydroglucopyranosiduronic acid (AGA) units are labeled in brown; (5) where appropriate, the open (acid) form of the unit was labeled turquoise and closed (lactone) form in purple. This applies to all the figures, except Fig. 8. Cross-peak coloration of the HSQC spectra, unless grayscale, depends on the multiplicity: for primary (CH) and tertiary (CH_3) substituted carbons, cross-peak correlations are shown in green; for secondary (CH_2) substituted carbons cross-peak correlations are shown in blue.

LOW-DP cellulose

LDP-CNCs ($DP_{N\text{-}GPC}$ 37) consist of chains of β-(1,4)-linked glucopyranose units terminated by RE and NRE groups (Fig. 1c, d). These are true nanocrystals (Fig. S35-S37), formed by partial depolymerization and recrystallisation of MCC using sc-H_2O. This 'residue' fraction (Buffiere et al. 2016) comprised of cellulose crystallite fragments consisting of both cellulose I and cellulose II allomorphs. This is illustrated by the microscopy combined with wide-angle X-ray scattering (WAXS) analysis of the material (Fig. S37), showing both phases of cellulose I and cellulose II. Some of the main distinguishable diffraction planes corresponding to cellulose I Miller indices (French 2014) are clearly visible at 15.6° for $(1\bar{1}0)$ and (110), and 22.3° for (200). For cellulose II there are also distinguishable peaks at 12.3° for the $(1\bar{1}0)$ and 20.1° for the (110) Miller indices (French 2014). Gel-permeation chromatography (GPC) showed a higher molecular weight residue, originating from MCC but the majority of material was as a lower molecular weight fraction, with a peak-maximum at a DP of $\sim$ 32 and overall $DP_{N\text{-}GPC}$ of 37 (Fig. S38).

The CH-1 region in the HSQC spectrum (Fig. 3a) was characterized by four clear signals and the signal with highest intensity was assigned as anhydroglucose unit (AGU)-1 that belongs to the bulk polymeric CH-1 (δ_H = 4.40 ppm (d); δ_C = 102.38 ppm), while the remaining signals correspond to NRE-1, RE-α-1 and RE-β-1. This region is characteristic of (hemi)acetals and such close grouping is caused by the rigid conformation adopted by the sugar unit, with the α-anomer showing a characteristic down-field shift (to $>$ 4.5 ppm) in the ^{1}H dimension and up-field shift in the ^{13}C domain. Detailed assignment of the remaining HSQC correlations was then completed using HSQC-TOCSY (Fig. S20-S22) and HMBC to

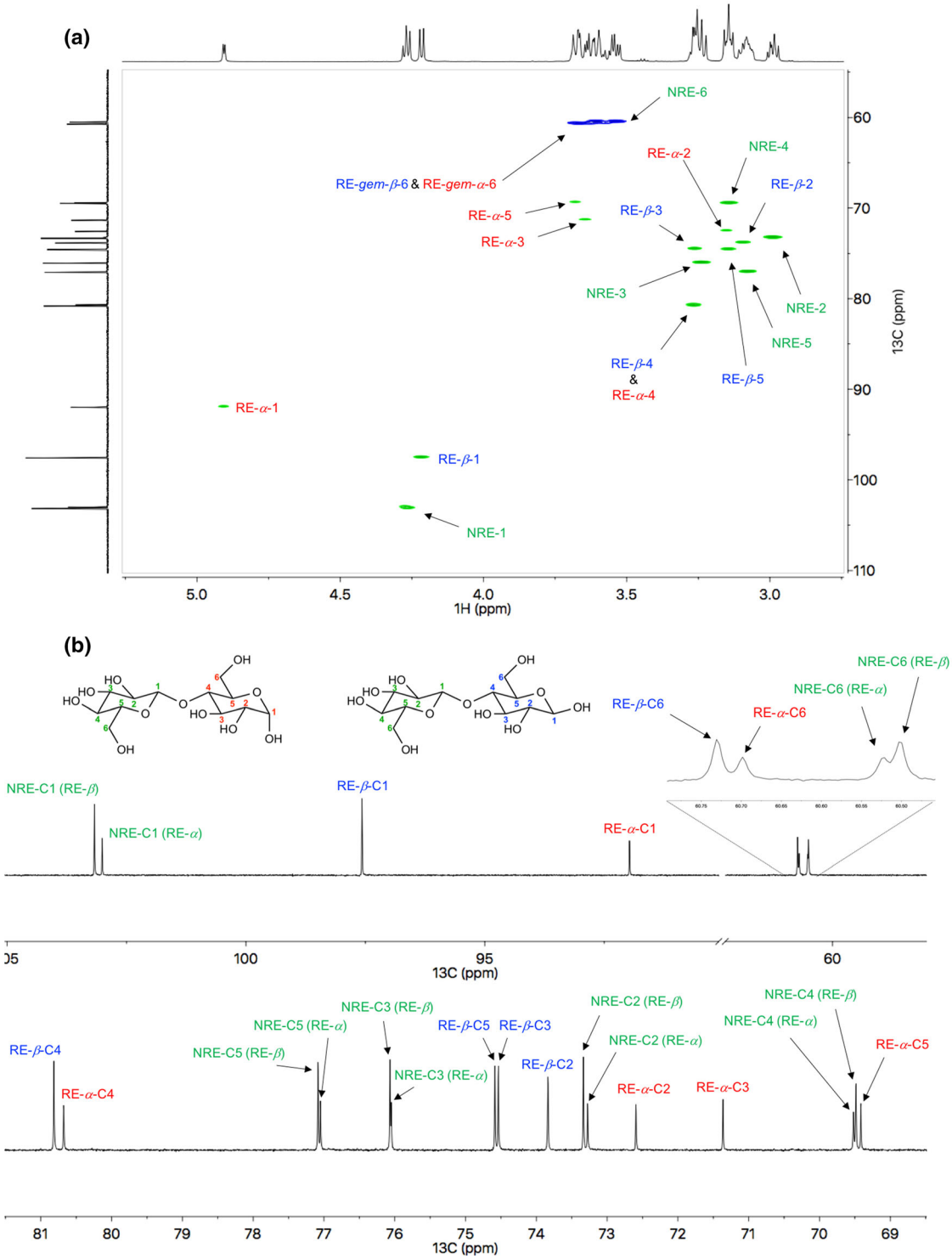


Fig. 2 Cellobiose spectra at 65 °C in $[P_{4444}][OAc]$/DMSO-d_6 (5 wt%): **a** Multiplicity-edited 2D HSQC (2048 time-domain data size in f1 corresponding to 1024 t_1-increments for the real spectrum); **b** ^{13}C spectrum. Non-reducing end (NRE) resonances shown in green, reducing end resonances α and β (RE- α and RE-β) shown in red and blue, respectively

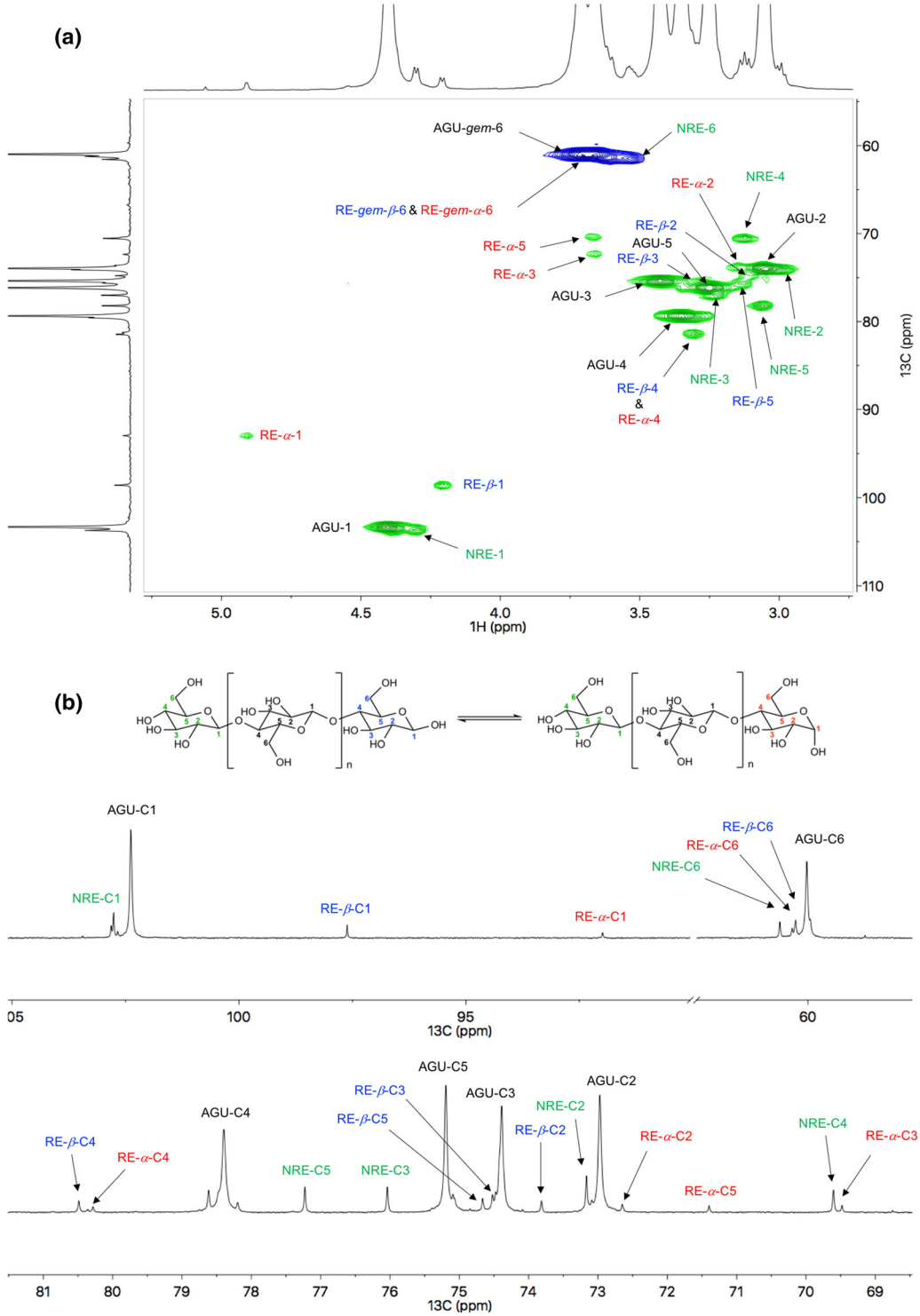


Fig. 3 LDP-CNCs spectra at 65 °C in $[P_{4444}][OAc]$:DMSO-d_6 (5 wt%): **a** Multiplicity-edited 2D HSQC (512 time-domain data size in f1, corresponding to 256 actual t_1-increments); **b** refocused ^{13}C INEPT. Non-reducing end (NRE) resonances shown in green, reducing end resonances α and β (RE-α and RE-β) shown in red and blue, respectively, internal (middle chain) anhydrous glucose unit resonances (AGU) shown in black

separate the CH-6 position correlations (Fig. S19). The ratio of α to β, of the ^{1}H spectrum is 38:62, by deconvolution (Supporting Information Fig. S3), using '*Fityk*' (Wojdyr 2010). The same method yielded a DPN-^{1}H of 15. While there is clearly error in this calculation, we favor the lower DP value, given by the NMR, as overestimation by GPC was also demonstrated in a previous publication (Heise et al. 2019), comparing 'CCOA labelling' (Röhrling et al. 2002) and 2 separate GPC systems; one calibrated using pullulan standards and one using multi-angle light scattering (MALLS) detection.

Throughout each of the experiments, the number of t_1-increments for the real spectrum (td1/2 for phase-sensitive HSQC) can be changed to improve resolution to the required level, to allow for separation of each signal. This is rather straightforward for low molecular weight compounds, such as glucose, cellobiose and to a lesser extent the LDP-CNCs. In this regard, the resolution in the indirectly detected ^{13}C-dimension in HSQC can start to approach that of the ^{13}C spectra, providing T_2 values are long enough to benefit from the further sampling. However, as molecular weights increase, the potential gain in resolution can often not be worth the additional collection times, with collection time proportional to the number of f1 increments. In addition, with shorter T_2 values the signals decay quickly and increased sampling will simply result in increased noise, with minimal increase in spectral resolution. Therefore, there is a trade-off between number of scans and number of increments, as molecular weight increases. To assess the resolution gain for a typical cellulose model, MCC ($DP_{N\text{-}GPC}$ 153), we measured the full-width at half maximum (FWHM) values (here in ppm units) in the ^{13}C dimension from HSQCs, gathered for different increment values (Fig. S2). The graph shows an inverse power function relating the resolution to the number of f1 increments. The optimum resolution, with little further trade-off in resolution vs collection time, can be achieved using 1024 increments, for the utilized spectral width of 24,883 Hz in the ^{13}C-dimension (corresponding to 24.3 Hz/pt digital resolution of the data in the f1-dimension). However, for most of the cases where good enough resolution is required for assignment of main peaks, 256 t_1-increments is sufficient and 512 still gives a reasonable improvement. This can be reduced further by using a smaller spectral width, as 24,883 Hz (165 ppm) is already rather wide, encompassing much more of the ^{13}C ppm domain that is necessary for unmodified cellulose. As molecular weight of the cellulose sample increases, the improvement in resolution with application of increasing increments is less apparent due to restricted motion, resulting in faster relaxation. However, 256–512 increments (i.e. 512–1024 time-domain size in f1) are perfectly reasonable values to achieve good S/N in an overnight run for assignment of NRE and RE signals for DP_N values of up to $\sim$ 200. It should also be considered that as the molecular weight increases, the relative abundance of NRE and RE resonances also decreases.

Nitroxyl-radical oxidized cellulose

Oxidation of cellulose with nitroxyl-radicals, such as TEMPO or AcNH-TEMPO, supposedly yield selective oxidation at the surface primary 6-hydroxyls (vs secondary 2- or 3-hydroxyls) to carboxylates but should also be capable of oxidizing the RE (hemi)acetal/aldehyde. Thus, the monomer unit in oxidized cellulose should be glucuronic acid (assuming each monomer is oxidized). If the terminal RE unit is oxidized, at the RE-1 position, gluconic acid should be the oxidized unit. Gluconic acid is available commercially as the sodium salt or as a solution in water, where it exists in equilibrium with the cyclic ester (lactone) form, dependent on water content. Indeed, drying mixtures of gluconic acid, even under ambient conditions, will induce lactonization (with loss of H_2O) to the δ-gluconolactone (Fig. S9). In this study, we dried a 49–53 wt% solution of gluconic acid in water using a rotary evaporator, at RT. The product was dissolved into the electrolyte and an HSQC spectrum was recorded (Fig. S9a). The HSQC spectrum clearly shows two separate spin-systems, i.e. open-chain and lactone forms. Both were assigned using 2D HSQC-TOCSY and compared with the HSQC spectrum for pure δ-gluconolactone (Fig. S10). Spectra of glucuronic acid (α- and β-anomers) were also taken for reference and can be found in the Supporting Information (Fig. S12-S13).

A water slurry of LDP-CNCs (15 wt%) was oxidized under mild acidic conditions (pH 5.8) with AcNH-TEMPO, in the presence of the $NaClO/NaClO_2$ oxidant system (see Supporting Information). Pinnick oxidation conditions were chosen to ensure complete conversion of aldehyde species to carboxylates. Water

soluble (high degree of oxidation (DO)) and a water insoluble (low DO) samples were recovered and separated by centrifugation. However, as the highly oxidized sodium carboxylate samples do not dissolve in the electrolyte, we were forced to acidify (Fujisawa et al. 2010) the fractions, for further NMR analysis. The principal structure of polyglucuronic acid is represented in Fig. 1f, with the AGA unit as the oxidized polymeric unit.

While the insoluble fraction expectedly consisted of minimally oxidized cellulose, the soluble TOx-LDP-CNC fraction had clearly identifiable correlations in the HSQC not corresponding to polymeric glucose resonances (Fig. 4). As with previous samples, the resonances for the AGA units were assigned using HSQC-TOCSY (Fig. S23-S26) but also using HMBC (Fig. S27), to further illustrate the linkage of the carboxylates C-6 to the H-5 position. The assignments for the AGA units were fully consistent with those of ^{13}C assignments for polyglucuronic acid in D_2O (Table S7 of the Supporting Information), from TEMPO oxidation of cellulose (Tahiri and Vignon 2000; Isogai et al. 2011). The RE and NRE peaks corresponding to glucose terminated chains are also assignable. One might assume that the NRE C6-OHs should be more accessible to oxidation than any other C6-OH. However, they are clearly present with the NRE more abundant than the RE signals, requiring scaling of the spectra close to the background to visualize the RE signals (Fig. 4-inset).

Cellobionic acid

Both the common nitroxyl-radical and Pinnick oxidation conditions should lead to oxidation of the reducing ends to carboxylates (Fig. 5). The Pinnick (acidic chlorite) oxidation at the reducing ends of CNCs is commonly used as the first step in reducing end functionalization, typically via amide formation and leading to nano-structures with self-assembly potential (Villares et al. 2018; Lin et al. 2019). To aid in the assignment of the terminal units in the oxidized products, we obtained a commercial sample of cellobionic acid and analyzed it in the $[P_{4444}][OAc]$:DMSO-d_6 electrolyte (Fig. 5). A doubling of the peaks was observed, consistent with partial conversion to the lactone form. The sample was also observed to be somewhat unstable at 65 °C, presumably decomposing by oligomerization. Therefore, the four spin-systems were assigned at 27 °C using HSQC-TOCSY (Fig. S29) and the two spin-systems corresponding to the acid form were identified by adding a drop of water into the NMR tube, allowing for almost complete conversion of the lactone form to the acid form (Fig. S31). The final assignments for the mixture of compounds at 65 °C showed little deviation from the sample at 27 °C. Thus, suitable model assignments for the oxidized reducing ends were afforded by the open-chain acid and closed-chain lactone spin-systems (Fig. 6).

Reducing end oxidation to carboxylate

Under nitroxyl-radical oxidation conditions (Hirota et al. 2009) we would have expected that the RE-1 position would have also been completely oxidized to carboxylates. However, expansion of the acetal region in the HSQC of the TOx-LDP-CNCs and increase in intensity shows presence of residual anomeric CH-1 resonances (Fig. 4a, inset). Clearly, complete oxidation of the reducing ends has not occurred. In addition to the cellobionic acid model compound, a further oxidation of the LDP-CNCs was performed under Pinnick oxidation conditions to allow for a more complete conversion of the reducing ends to gluconate moieties (or to the corresponding lactone). The HSQC spectrum for the oxidized POx-LDP-CNC product (Fig. 7a), where the spectral scale was increased to emphasize the baseline signals, shows signals corresponding to the open-chain acid spin-system, almost identical to that of cellobionic acid but signals corresponding to the closed-chain lactone form are absent. Reducing end signals are also present, again indicating incomplete oxidation. Similarly, if the nitroxyl-radical oxidized sample is also scaled to a similar scale, emphasizing baseline signals, the same peaks corresponding to the open-chain acid form are present (Fig. 7b). HSQC-TOCSY spectra of both these samples (Fig S29-S30) also allow for tracing of the spin-system, consistent with the cellobionic acid model (Fig S28). Therefore, there are now unequivocal solution-state NMR assignments for this functional moiety, which can be used for further understanding/optimization of associated chemistry.

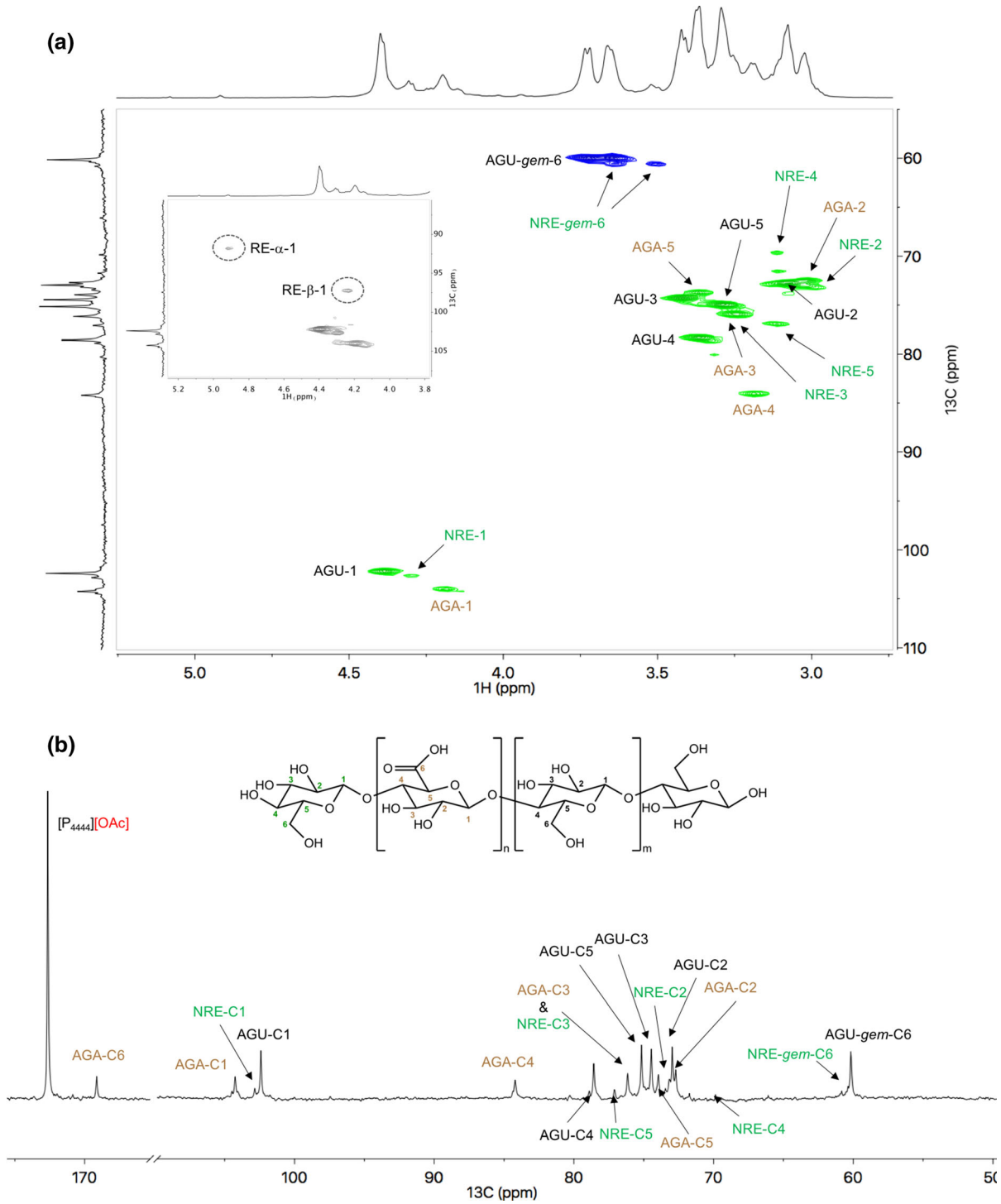


Fig. 4 Spectra of TOx-LDP-CNCs at 65 °C in $[P_{4444}][OAc]$:DMSO-d_6 (5 wt%): **a** Multiplicity-edited 2D HSQC (1024 time-domain data size in f1, corresponding to 512 actual t_1-increments); **b** refocused ^{13}C INEPT. Non-reducing end (NRE) of cellooligomeric resonances shown in green, anhydroglucopyranosiduronic acid (AGA) unit resonances are shown in brown, internal (middle chain) non-oxidized anhydroglucose unit resonances (AGU) shown in black

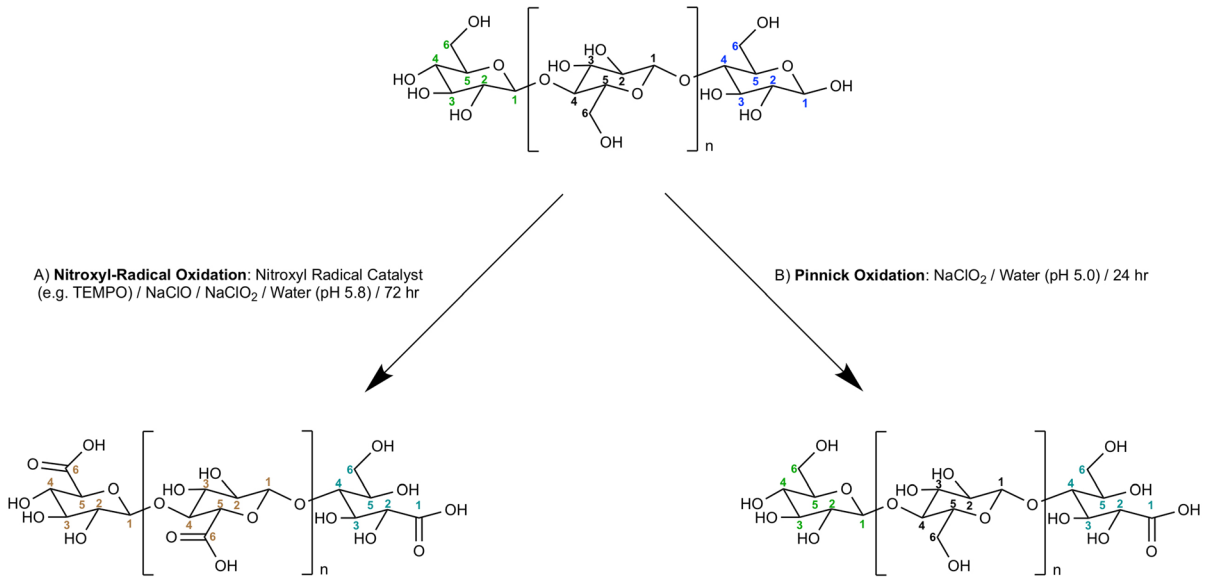


Fig. 5 Scheme and conditions for oxidation of cellulose under acidic nitroxyl-radical (e.g. TEMPO or AcNH-TEMPO) or Pinnick ($NaClO_2$) oxidation conditions

Assignment and stability of periodate-oxidized CNCs

Finally, a typical procedure for periodate oxidation ($NaIO_4$) of cellulose nanocrystals (CNCs) was performed (Nypelö et al. 2018). The resulting $NaIO_4$–CNCs were then dissolved in the electrolyte and analyzed. After analysis of the samples (overnight at 65 °C) the sample was brown, whereas other cellulose samples (nitroxyl-radical-oxidized and unoxidized) did not colorize significantly at all. This may indicate some kind of degradation or possibly iodine formation, from (per)iodate residues. The HSQC spectrum showed a forest of peaks (Fig. 8), clustered mainly around the CH2-5 region and from the high resolution of many of the signals in the ^{1}H spectrum (Fig. 8, top ^{1}H trace) it is quite clear that significant depolymerization had occurred. However, when the diffusion-edited ^{1}H spectrum (diffusion-editing filters out the slow-diffusing species) was collected (Fig. 8, bottom ^{1}H trace), broad signals corresponding to polymeric cellulose resonances were apparent. Based on our previous assignments for glucose (Fig. S5), cellobiose (Fig. 2) and cellulose (Fig. 3), almost all the correlations in the HSQC could be assigned, with only a few signals remaining unassigned. This indicated that a significant proportion of the cellulose was fragmented into glucose, cellobiose and likely other oligomers. Yet, polymeric cellulose also remained.

Hosoya et al. (2018) recently demonstrated that oxidation of cellulose at position 6 to carboxylate does not seem to introduce instability to cellulose, based on experimental kinetics and transition-state modelling. However, oxidation at positions 2 & 3 to ketones, and position 6 to aldehyde, does seem to introduce significant instability to cellulose. It is proposed that under alkaline conditions, β-elimination occurs leading to fragmentation of the sugar units. As the $[P_{4444}][OAc]$:DMSO-d_6 NMR electrolyte is rather a basic media, mainly attributable to the acetate anion and absence of any protic solvating species, it is apparent that the position 2 & 3 aldehydes that are formed during periodate oxidation also introduce significant instability to the cellulose polymer. Therefore, a mechanism can be proposed (Fig. 9) which accounts for the current NMR observations: periodate oxidation proceeds by oxidizing different points along cellulose chains, on the surface of the CNCs. After dissolution into the basic electrolyte, fragmentation at these oxidation sites occurs liberating the oligomeric, dimeric and monomeric sugars which linked the oxidation points on the surface chains. These are clearly resolvable using HSQC. Likewise, the untouched polymeric chains at the core of the CNCs are also resolvable and their presence is clearly

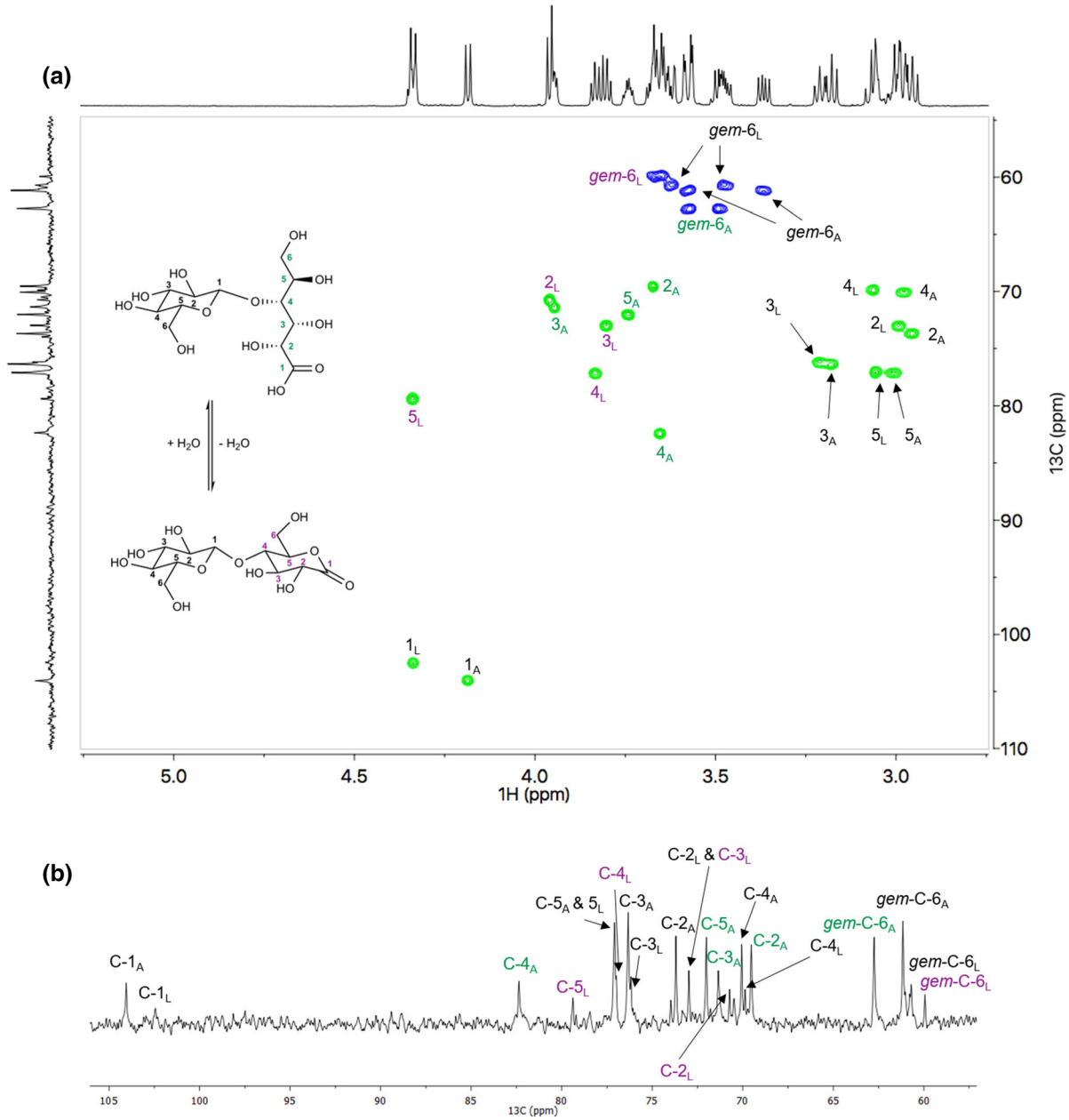


Fig. 6 Cellobionic acid spectra at 65 °C in [P_{4444}][OAc]:DMSO-d_6 (5 wt%): **a** Multiplicity-edited 2D HSQC (512 time-domain data size in f1, corresponding to 256 actual t_1-increments); **b** refocused ^{13}C INEPT. Glucopyranose unit resonances (black), open-chain acid unit (turquoise) and lactone unit (purple). 'A' and 'L' subscripts refer to 'acid' and 'lactone' forms

illustrated through the diffusion-edited ^1H spectrum, which filters out all low molecular-weight monomeric, dimeric and oligomeric species. This proposed mechanism is also consistent with previous mechanistic studies demonstrating that periodate oxidation on cellulose proceeds heterogeneously, by formation of oxidized domains on the crystallite surfaces (Kim et al. 2000). The 'unknown' low molecular weight residues, that remain unassigned in the HSQC (Fig. 8), may correspond to fragments not attached to the polymeric units, resulting from C2–C3 bond cleavage. Closer examination of the diffusion-edited ^1H and

HSQC spectra (Fig. S34) reveals some more complexity in the (hemi)acetal region, which may result from acetal formation with these fragments.

While it seems that periodate oxidation introduces instability through β-elimination under basic conditions, this method seems to allow for assessment of that stability and may offer a further method for validating the reported increase in stability of further modification schemes, e.g., through borohydride reduction of periodate-oxidized cellulose (Potthast et al. 2009).

Quantitation using HSQC

One drawback of solution-state ^{13}C NMR analysis is the low abundance of ^{13}C-nucleus leading to low sensitivity. Thus, high numbers of repetitions are required in order to obtain decent S/N ratios, for adequate quantitation accuracy. This is exacerbated by the requirement for longer relaxation delays. However, there is an increasing trend of deconvolution of ^{1}H spectra of polymers, as S/N is much better than for ^{13}C. Of course, not all ^{1}H resonance signals are easily identified and separated by deconvolution, due to the lower resolution of ^{1}H. Baseline correction can also be problematic and if one wishes to quantify the RE and NRE signals using this method, the errors very rapidly become large at a DP_N of > 100 (Holding et al. 2016). 2D HSQC provides vastly improved resolution of species over 1D experiments and significantly improved S/N over 1D ^{13}C-data, as it is a ^{1}H-detected experiment. In terms of analysis of celluloses and oxidized celluloses, separation of the main polymeric-1 cross-peak from the (nitroxyl-radical) oxidized polymeric-1 cross-peak and from the α-RE, and β-RE, cross-peaks is now very good using the $[P_{4444}][OAc]$:DMSO-d_6 electrolyte. This potentially allows for a rapid and reliable method for data extraction; requiring only phasing, baseline correction and 2D correlation peak integration of the raw data. NRE signals have volume overlapping with the main polymeric-1 correlation so this is not so clearly separable, except based on the assumption that there is the same amount of NRE species as there are RE species. The *geminal*-6 signals are also well separated from the rest of the cellulose backbone signals. Separation of other signals is possible but the error starts to increase the closer they are to each other, due to peak volume overlap. Samples with wider ranges of functionalities, such as lignocellulosic biomass samples or those that have resonances downfield from the cellulose polymeric-1 acetal correlations, are also easily separated.

The major drawbacks with 2D HSQC and polymeric samples are four-fold: 1) differences in $^{1}J(^{13}C–^{1}H)$ values for different ^{13}C–^{1}H pairs cause variations in intensities of those correlations. Typical HSQC experiments assume an average $^{1}J(^{13}C–^{1}H)$ value (typically 145 Hz) for the experiments, represented in a specific INEPT polarization transfer delay. 2) Coherence transfer periods, where sufficient time is given for ^{1}H magnetization to evolve, cause intensity variation of the correlation peaks as each resonance has different T_2 values, i.e., mainly during the INEPT delay periods, more or less signal is lost for different resonances, prior to acquisition. 3) Correlation-peak distortions arising from evolution of homonuclear $J(^{1}H–^{1}H)$ coupling during the INEPT steps can cause errors in integration. 4) Non-linear excitation bandwidth leads to variation of cross-peak signal intensity, especially across the ^{13}C frequency range at high field strengths. Several quantitative HSQC sequences or processing strategies have been developed that attempt to correct for these issues. Variations in $^{1}J(^{13}C–^{1}H)$ values have been corrected for by applying INEPT-delay modulation in the first 'quantitative HSQC' (Q-HSQC) experiment (Heikkinen et al. 2003). This corresponds to the application of a spread of INEPT delays covering the typical $^{1}J(^{13}C–^{1}H)$ coupling value range expected in organic materials. Signal losses due to variations in T_2 values as well as in $^{1}J(^{13}C–^{1}H)$ values have been accounted for in the 'time-zero HSQC' ($HSQC_0$) experiment (Hu et al. 2011), which records a loop of an increasing train of HSQC sequences ($HSQC_X$, X = 1–3), prior to actual acquisition. This has the effect of increasing coupling and relaxation effects for each loop, which can be extrapolated back to 'time-zero', where potentially all effects are removed. Obviously for $HSQC_0$, T_2 values have to allow recording of $HSQC_2$ and $HSQC_3$ data sets with intensity allowing reliable extrapolation. The 'quick-quantitative HSQC' (QQ-HSQC) experiment is a rather elegant method for potentially reducing the collection times by a factor of 4 (Peterson and Loening 2007). This encodes the INEPT-modulation into different vertical slices in the sample but only represents a possible doubling of S/N for the same collection times. However, this is only really

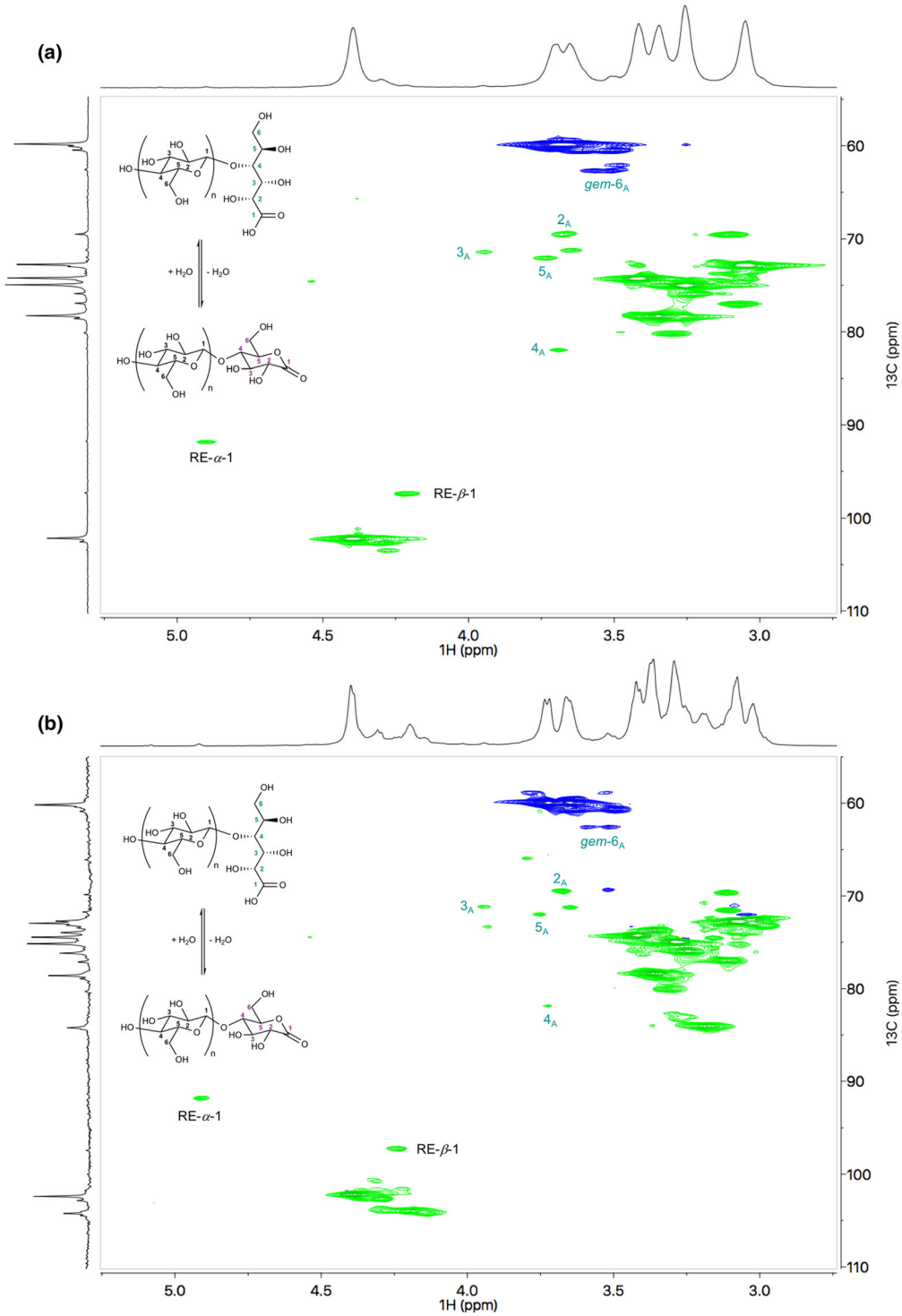


Fig. 7 Scaled (to emphasize baseline signals) multiplicity-edited HSQC spectra, at 65 °C in $[P_{4444}][OAc]$:DMSO-d_6 (5 wt%), showing presence of terminal gluconic acid moieties for: a) POx-LDP-CNCs (1024 time-domain data size in f1, corresponding to 512 actual t_1-increments) and b) TOx-LDP-CNCs (1024 time-domain data size in f1, corresponding to 512 actual t_1-increments)

applicable for small, slow-relaxing, molecules due to relaxation effects. The 'quantitative CPMG-adjusted HSQC' (Q-CAHSQC) experiment (Koskela et al. 2005) applies CPMG-INEPT steps to avoid cross-peak distortions, due to $J(^1H–^1H)$ coupling evolution. The 'quantitative, offset-compensated, CPMG-adjusted HSQC' (Q-OCCAHSQC) experiment (Koskela et al. 2010) applies novel broad-band pulses to reduce ^{13}C offset errors that are most prevalent on ultrahigh-field instruments, e.g. 1000 MHz, over wider frequency ranges (> 150 ppm). This is not really a concern for cellulose samples on a 600 MHz spectrometer, where the ^{13}C domain is rather narrow (< 50 ppm) but may start to become an issue for lignocellulose samples at ultrahigh field. The final sequence of interest is the 'quantitative, equal carbon HSQC' (QEC-HSQC) experiment (Mäkelä et al. 2016) where a refocusing period, after the first INEPT step, is used to discard the excess CH_2 and CH_3 magnetization. This yields the same signal intensity for each protonated carbon in the sample. This could potentially be of value in the analysis of lignocellulose samples but is not really necessary at this point for systems where the well separated resonances (and their multiplicities) are already reasonably well assigned. An overview of the different HSQC experiments is given in Table 1.

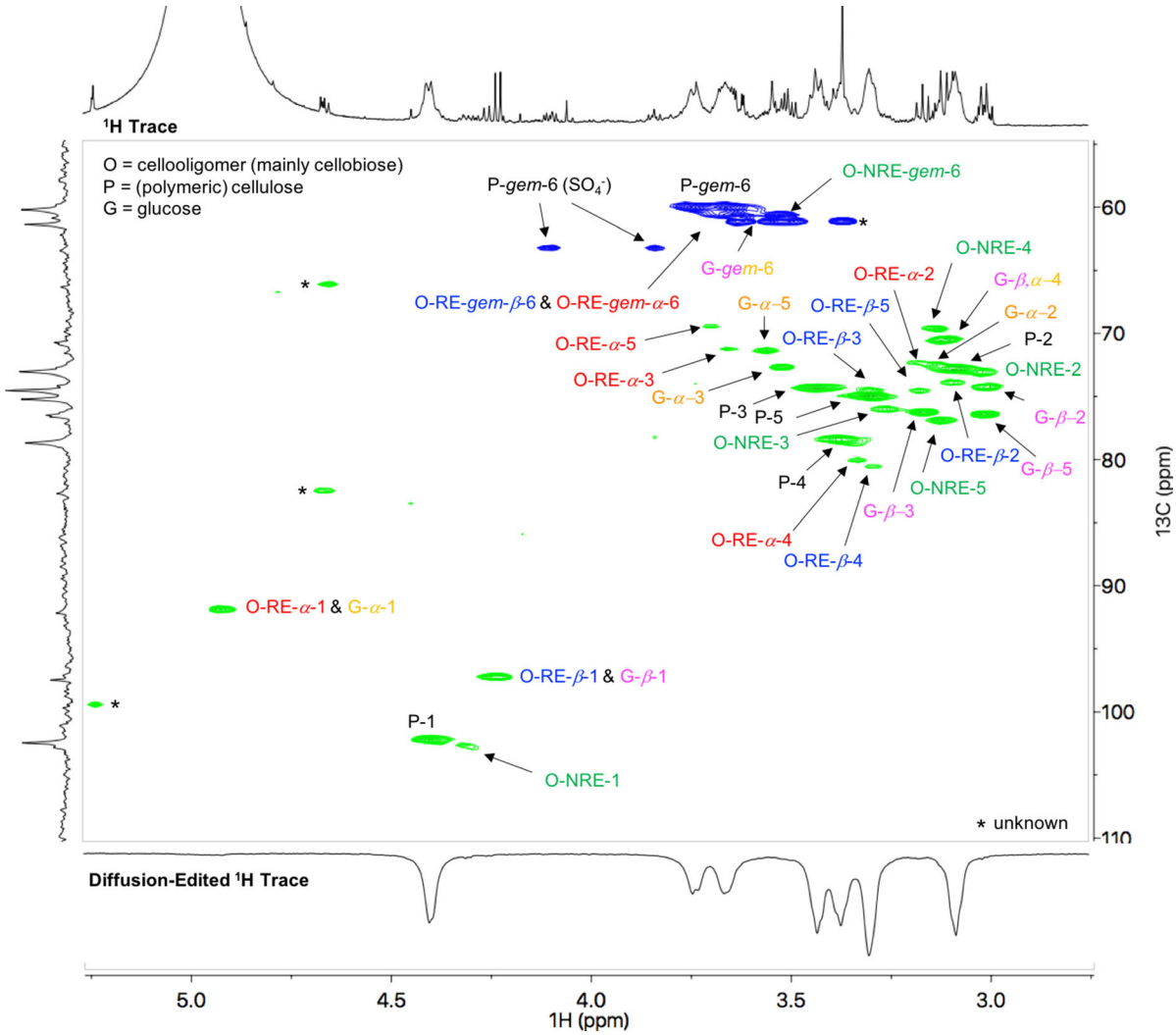


Fig. 8 Multiplicity-edited HSQC spectrum at 65 °C in $[P_{4444}][OAc]$:DMSO-d_6 (5 wt%) (1024 time-domain data size in f1, corresponding to 512 actual t_1-increments) of $NaIO_4$-CNCs

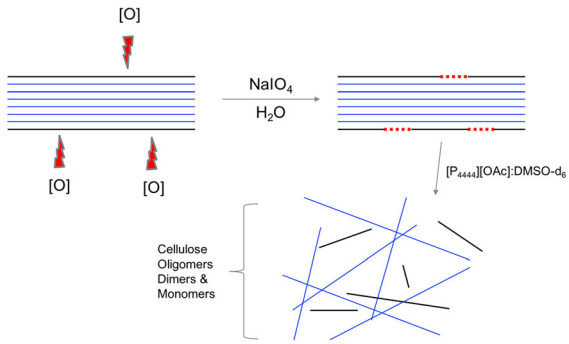


Fig. 9 Proposed mechanism for the observation of cellulose, cellooligomers, cellobiose and glucose in the HSQC spectrum of $NaIO_4$-CNCs, with the $[P_{4444}][OAc]$:DMSO-d_6 inducing β-elimination in oxidized units (Hosoya et al. 2018). '[O]' refers to 'oxidation' and the red dotted segments on the schematic surface refer to the oxidation sites, which are then cleaved upon dissolution into the basic $[P_{4444}][OAc]$:DMSO-d_6 electrolyte

For our purposes, the Q-CAHSQC sequence seems to be most suitable, if we can correct the processed integral data for T_2 losses. Issues with ^{13}C-offset, variation in $^{1}J(^{13}C–^{1}H)$ and relaxation can be accounted for mathematically, to some degree, in post-processing (Zhang and Gellerstedt 2007). The most concerning issue for cellulose is the very fast T_2 relaxation, that the $HSQC_0$ sequence accounts for, but the other approaches do not. If the T_2 values are known, it is possible to adjust for signal losses by application of Eq. 1, after integration of the 2D HSQC spectra; where Δ is the delay period in which T_2 losses occur, V is the measured correlation peak volume and V_0 is the theoretical correlation peak volume, with no losses due to relaxation:

$$V_0 = Ve^{\frac{\Delta}{T_2}} \quad (1)$$

In order to apply this correction, T_2 values for the resonances of interest must be measured. Zhang and Gellerstedt (2007) have shown that the 2D HSQC-CPMG sequence for determination of T_2 values of cellulose triacetate, gives inaccurate T_2 values. However, this is a chicken and egg scenario; losses due to the HSQC portion of the sequence obviously contribute to the inaccuracies, which becomes more of an issue where there is very high molecular weight material in the sample due to disproportionate loss of signal from those resonances. Therefore, the lower resolution 1D CPMG is the only real option for determining more representative 'average' T_2 values, for such samples. Nevertheless, loss of higher molecular weight signal during the INEPT delays is always going to be a problem. In terms of quantitation of chemical species; if the system is not complex, as is typically in cellulose samples, T_2 values can be relatively easily obtained. As such we determined T_1 and T_2 values for the LDP-CNC, MCC and TOx-LDP-CNC samples (Table 2).

After summing the appropriate delay times leading to T_2 losses, a Δ value of 13.9 ms was calculated for the Q-CAHSQC sequence, which is in a very similar range to some of the MCC T_2 values. Therefore, considerable signal loss is expected and needs to be corrected for. Equation 1 was applied in the correction of integral values. The main cross-peaks of interest were the polymeric CH-1 (AGU-1), NRE-1, α-RE-1, β-RE-1, oxidized polymeric CH-1 (AGA-1), AGU-*gem*-6 and oxidized RE-6 position (Ox-RE-*gem*-6). The MCC sample also clearly contained a little xylan, so the xylan 1 (AXU-1) and *geminal*-5 (AXU-*gem*-5) positions were also integrated in the MCC sample. The corrected results are given in Table 3 and were processed further to yield a few additional parameters: (1) the ratio of AGU-1 to AGU-*gem*-6 ($AGU_{1/6}$), (2) the DP_N from HSQC ($DP_{N\text{-}HSQC}$). (3) The % values of α-RE-1 by HSQC ($\%_{\alpha\text{-}HSQC}$) and β-RE-1 by HSQC ($\%_{\beta\text{-}HSQC}$), (4) The % values for AXU based on AXU-1 ($\%_{AXU\text{-}1}$) and AXU-*gem*-5 ($\%_{AXU\text{-}5}$). 5) The % values for oxidation of AGU to AGA ($\%_{AGA}$) and RE ($\%_{RE\text{-}Ox}$). In addition to the processed HSQC data, we have data from deconvolution of the ^1H spectra, for comparison: 1) the DP_N from ^1H ($DP_{N\text{-}^1H}$), 2) The % values of α-RE-1 by ^1H (%α-^1H) and β-RE-1 by ^1H (%β -^1H).

Clearly the T_2 correction has a big impact on the integral values, especially for those with short T_2 values, i.e. for the bulk polymeric AGU-1 and AGU-*gem*-6. As these positions are most likely to be used for quantitation, e.g. of DS values, it is clearly critical to do the T_2 correction. To compare how effective the quantitation is, the parameter $AGU_{1/6}$ shows how accurate integration of AGU-1 and AGU-*gem*-6 is, with the optimum value of 1. For the LDP-CNC sample, the value improves significantly after T_2 correction. For the MCC experiments, the lower-resolution but higher S/N experiments (ns = 40, td1 = 128) gave a value of 1.00 and 0.96, for the room temperature probe-head and He-cooled cryoprobe-head, respectively. This indicates that making all efforts to maximize S/N is critical for quantitation,

Table 1 Comparison of the different quantitative HSQC sequences available

Sequence	Development	Pros/Cons	References
Q-HSQC	INEPT modulation	Avoids J_{CH} errors Does not account for T_2 errors	Heikkinen et al. (2003)
Q-CAHSQC	CPMG-INEPT to avoid correlation distortions	Avoids J_{CH} errors Avoids correlation distortions leading to integration errors Does not account for T_2 errors	Koskela et al. (2005)
QQ-HSQC	4 different INEPT delays acquired at same time from 4 different slices corresponding to different J_{CH} values	Avoids J_{CH} errors Potentially doubles S/N (low molecular weight samples only) for the same collection times as other experiments Does not account for T_2 errors	Peterson and Loening (2007)
Q-OCCAHSQC	Shaped pulses	Avoids ^{13}C-offset effects in ultrahigh-fields Avoids J_{CH} errors Avoids correlation distortions leading to integration errors Does not account for T_2 errors	Koskela et al. (2010)
$HSQC_0$	Accounts for J_{CH} and T_2 variations by extrapolation back to 'time-zero'	Avoids J_{CH} errors Avoids T_2 errors Significantly increased collection time For short T_2 species the data points from the longest sequences might cause problems Requires more analysis i.e. extrapolation	Hu et al. (2011)
QEC-HSQC	Extra refocusing period after 1st INEPT step	Avoids J_{CH} errors Makes all carbons equal Does not account for T_2 errors	Mäkelä et al. (2016)

Table 2 T_1 and T_2 values for the LDP-CNC, MCC and TOx-LDP-CNC samples

Resonance	T_1, s			T_2, s		
	LDP-CNC	MCC	TOx-LDP-CNC	LDP-CNC	MCC	TOx-LDP-CNC
AGU-1	1.71	1.86	1.73	0.041	0.016	0.053
AGU-α-RE-1	1.71	2.25	1.70	0.173	0.152	0.188
AGU-β-RE-1	1.59	2.01	1.76	0.159	0.140[a]	0.184
AGU-NRE-1	1.47	2.00	1.58	0.096	0.146	0.066
AGU-2	1.82	1.90	–	0.067	0.033	–
AGU-3	1.74	1.86	–	0.063	0.031	–
AGU-4	1.73	1.83	–	0.059	0.030	–
AGU-5	1.63	1.77	–	0.063	0.030	–
AGU-*gem*-6a	1.19	1.61	1.12	0.036	0.016	0.046
AGU-*gem*-6b	1.21	1.67	1.14	0.039	0.018	0.050
AGA-1	–	–	2.06	–	–	0.033
AXU-1	–	–	–	–	0.091	–
AXU-*gem*-5a	–	–	–	–	0.124	–

[a]The AGU-β-RE-1 signal was overlapping with the larger AXU-1 signal so the value is an estimate, by comparing with the LDP-CNC values

Table 3 Uncorrected and T_2-corrected integral values and various other parameters calculated from the Q-CAHSQC peak volumes (by integration in MestreNova 10.0) and ^{1}H peak areas (by deconvolution using *Fityk* 1.3.1)

Sample Resonance	Uncorrected				T_2 Corrected			
	LDP-CNC [ns = 8, td1 = 512][a]	MCC [ns = 8, td1 = 640] (cryoprobe values)	MCC [ns = 40, td1 = 128] (cryoprobe values)	TOx-LDP-CNC [ns = 8, td1 = 512]	LDP-CNC [ns = 8, td1 = 512]	MCC [ns = 8, td1 = 512] (cryoprobe values)	MCC [ns = 32, td1 = 128] (cryoprobe values)	TOx-LDP-CNC [ns = 8, td1 = 512]
AGU-1	30.7	48.1 (109.4)	56.9 (55.8)	35.6	43.2	115.1 (261.8)	136.2 (133.5)	46.4
α-RE-1	1	1 (1)	1 (1)	1	1.08	1.10 (1.10)	1.10 (1.10)	1.08
β-RE-1	1.77	1.24 (3.14)	1.36 (1.62)	1.73	1.93	1.37 (3.47)	1.50 (1.79)	1.87
NRE	4.17[b]	5.64[b] (16.0)	7.72[b] (9.53)	3.91[b]	4.82[b]	6.21[b] (17.6)	8.50[b] (10.49)	4.83[b]
AGU-*gem*-6	72.7	94.9 (244.3)	130.7 (135.8)	89.5	105.0	213.3 (549.1)	294.0 (305.2)	119.7
AGA-1	–	–	–	13.8	–	–	–	20.9
Ox-RE-*gem*-6	–	–	–	1.86	–	–	–	2.04
AXU-*gem*-1	–	3.71[b] (8.49)	7.66[b] (6.36)	-	–	4.15[b] (9.50)	8.57[b] (7.12)	–
AXU-*gem*-5	–	5.0 (12.9)	5.65 (6.54)	-	–	5.83 (15.1)	6.59 (7.63)	–
$AGU_{1/6}$[c]	1.19	1.01 (0.90)	0.87 (0.82)	0.57	0.97	1.16 (1.03)	1.00 (0.96)	0.95
$DP_{N\text{-}HSQC}$[d]	12	22 (27)	25 (22)	14	17	50 (62)	57 (51)	–
$DP_{N\text{-}^1H}$	–	–	–	–	17[e]	58[f]	58[f]	–
$DP_{N\text{-}GPC}$	–	–	–	–	37	153	153	–
$\%_{\alpha\text{-}HSQC}$[g]	36.1	44.6 (24.2)	42.4 (38.2)	36.6	35.9	44.4 (24.0)	42.2 (38.0)	36.6
$\%_{\beta\text{-}HSQC}$[h]	65.9	55.4 (75.8)	57.6 (61.8)	63.4	64.1	55.6 (76.0)	57.8 (62.0)	63.4
%α-^{1}H[i]	–	–	–	–	38.1	–	–	–
%β-^{1}H[j]	–	–	–	–	61.9	–	–	–
$\%_{AXU\text{-}1}$[k]	–	6.9[b] (7.0)	11.5[b] (9.8)	–	–	3.2[b] (3.2)	5.5[b] (4.6)	–
$\%_{AXU\text{-}5}$[l]	–	9.9 (11.4)	9.5 (11.2)	–	–	2.3 (2.7)	2.2 (2.4)	–
$\%_{AGA}$[m]	–	–	–	22.9	–	–	–	27.9
$\%_{RE\text{-}Ox}$[n]	–	–	–	20.3	–	–	–	20.2
$S/N_{AGU\text{-}1}$[o]	486.6	110.5 (207.8)	179.3 (236.0)	418.8	-	–	–	–

[a]'ns' is the number of transients and 'td1′ is the number of f1 increments (real and imaginary)

[b]These values have most error due to overlap of AXU-1 with the AGU-1 peak

[c]Calculated as '2 × (AGU1 + α-RE-1 + β-RE-1 + NRE)/AGU-*gem*-6′ by HSQC peak volumes

[d]Calculated as'(AGU1 + α-RE-1 + β-RE-1 + NRE)/ (α-RE-1 + β-RE-1)' from the HSQC peak volumes

[e]Calculated as'(AGU1 + α-RE-1 + β-RE-1 + NRE)/ (α-RE-1 + β-RE-1)' from the ^{1}H peak areas

[f]Calculated as'(AGU-1 + α-RE-1 + β-RE-1 + NRE)/(α-RE-1 + β-RE-1)' from the ^{1}H peak areas, with the β-RE-1 peak area estimated to be'100 × β-RE-1/34′ (with 34:66 as the α:β-RE-1%ratio for cellobiose)

[g]Calculated as'100 × α-RE-1/(α-RE-1 + β-RE-1)' from the HSQC peak volumes

[h]Calculated as'100 × β-RE-1/(α-RE-1 + β-RE-1)' from the HSQC peak volumes

[i]Calculated as'100 × α-RE-1/(α-RE-1 + β-RE-1)' from the ^{1}H peak areas

[j]Calculated as'100 × β-RE-1/(α-RE-1 + β-RE-1)' from the ^{1}H peak areas

[k]Calculated as'100 × AXU-1/(AXU-1 + AGU-1 + α-RE-1 + β-RE-1 + NRE)' from the HSQC peak volumes

[l]Calculated as'100 × AXU-*gem*-5/(AXU-*gem*-5 + AGU-*gem*-6)' from the HSQC peak volumes

[m]Calculated as'100 × AGA-1/(AGA-1 + AGU-1 + α-RE-1 + β-RE-1 + NRE)' from the HSQC peak volumes

[n]Calculated as'50 × Ox-RE-*gem*-6/(α-RE-1 + β-RE-1)' from the HSQC peak volumes

[o]S/N ratio for the AGU-1 peak maximum. Values were determined using the MestreNova 10.0 'SNR Peak Calculator' script

even at the expense of resolution. The TOx-LDP-CNC sample also gave a significant improvement in $AGU_{1/6}$, from 0.57 to 0.95. $DP_{N\text{-}HSQC}$ values also changed significantly and the corrected values were more or less consistent with the $DP_{N\text{-}^1H}$ values. There is a noticeable difference between the $DP_{N\text{-}GPC}$ and those obtained from NMR. More accurate studies validating the use of NMR against both labelling and GPC studies are needed. Practically, the HSQC method is still limited in what samples can be studied for $DP_{N\text{-}HSQC}$ determination, as better S/N will be required with increases in molecular weight. For the same reason that HSQC-CPMG is not suitable for T_2 determination, for cellulose samples, HSQC on such samples is going to eliminate a significant proportion of the faster relaxing high molecular weight material, artificially decreasing the $DP_{N\text{-}HSQC}$ values somewhat. The ratios of α-RE-1 and β-RE-1 were relatively consistent between the corrected and uncorrected values, for HSQC and ^1H deconvolution. However, clearly the lower abundance of RE resonances for MCC causes significant error, although, this situation can be improved using a cryoprobe-head and possibly linear prediction. AXU contents for the two MCC experiments were relatively consistent, based on AXU-*gem*-5 integration. However, the higher resolution experiment gave more consistent values, based on integration of AXU-1 and AXU-*gem*-5, as the separation of these peaks from the cellulose resonances was much better in the higher resolution case. The degrees of oxidation for TOx-LDP-CNC, $\%_{AGA}$ and $\%_{RE\text{-}Ox}$, were also relatively consistent.

Overall, the CH1 peaks for the low molecular weight LDP-CNC and TOx-LDP-CNC samples are easily separable with 512 f1 increments (td1), or perhaps even less (of course depending on field strength), due to their slower T_2 relaxation (Fig. 10a, b). With the higher molecular weight MCC, separation of the CH1 resonances is definitely improved with the higher number of increments (Fig. 10c, d). While there is sufficient separation of the RE-1 and AGU-1 signals, so that f1 resolution can be lowered further (to allow for increased collection times), poor S/N is still an issue for the RE-1 signal in both spectra (Fig. 10c, d). This situation is improved somewhat with the use of the cryoprobe-head, where S/N is approximately doubled (Table 3). However, if quantitation of DS values is all that is required, lower resolutions are acceptable to reduce collection times to a few hr. If resolution eventually does become an issue in quantitation of the DS of some substituent, for higher molecular weight samples, then ball milling will likely have to be applied to reduce molecular weights (Ling et al. 2019), preventing disproportionate T_2 losses. However, this requires future work with well-defined samples over wide molecular weight ranges, both polydisperse and non-polydisperse.

Further applicability

This method is ideally suited to the analysis of nanocelluloses, due to the relatively low molecular weight that these samples show, in particular CNCs. However, higher molecular weight samples are also possible, which makes this method of significant wider value for following cellulose surface chemistry, where crystallinity is maintained. Indeed, it has been possible to dissolve and collect an HSQC spectrum for even bacterial nanocellulose, in a related solvent system (Holding et al. 2016). The main limitation here is the faster signal relaxation, which would have a significant effect on any HSQC quantitation, relative to the cellulose backbone signals. However, quantitation through 1D spectra, with the aid of signal deconvolution, would not be affected. Thus, a combination of 1D and 2D methods can be applied, optionally using the $[P_{4444}]^+$ signals themselves as internal standard.

Several solvent systems are known for analysis of whole biomass samples (Foston et al. 2016). Mansfield et al. (2012) have recommended the use of routine HSQC experiments for quantifying biopolymer species, in whole biomass samples. Their protocol demonstrates the swelling of planetary-milled wood samples in DMSO-d_6/pyridine-d_5 (4:1) to yield 'gelled' samples, yet with a solvent mixture that is unable to directly dissolve cellulose. This method was said to yield similar quantitation results to those samples which are fully processed into the solution-state, by peracetylation. However, wood samples are a difficult case, not only due to the insolubility of cellulose in common molecular solvents but also due to their recalcitrant nature, in general (Deb et al. 2016; Kyllönen et al. 2013; Kilpeläinen et al. 2007). Thus, there is still some way to go to establish quantitative conditions for whole biomass samples, even with non-derivatizing direct-dissolution cellulose solvents. The cellulose portion of these materials is always the most troublesome as it is such a rigid polymer, which suffers

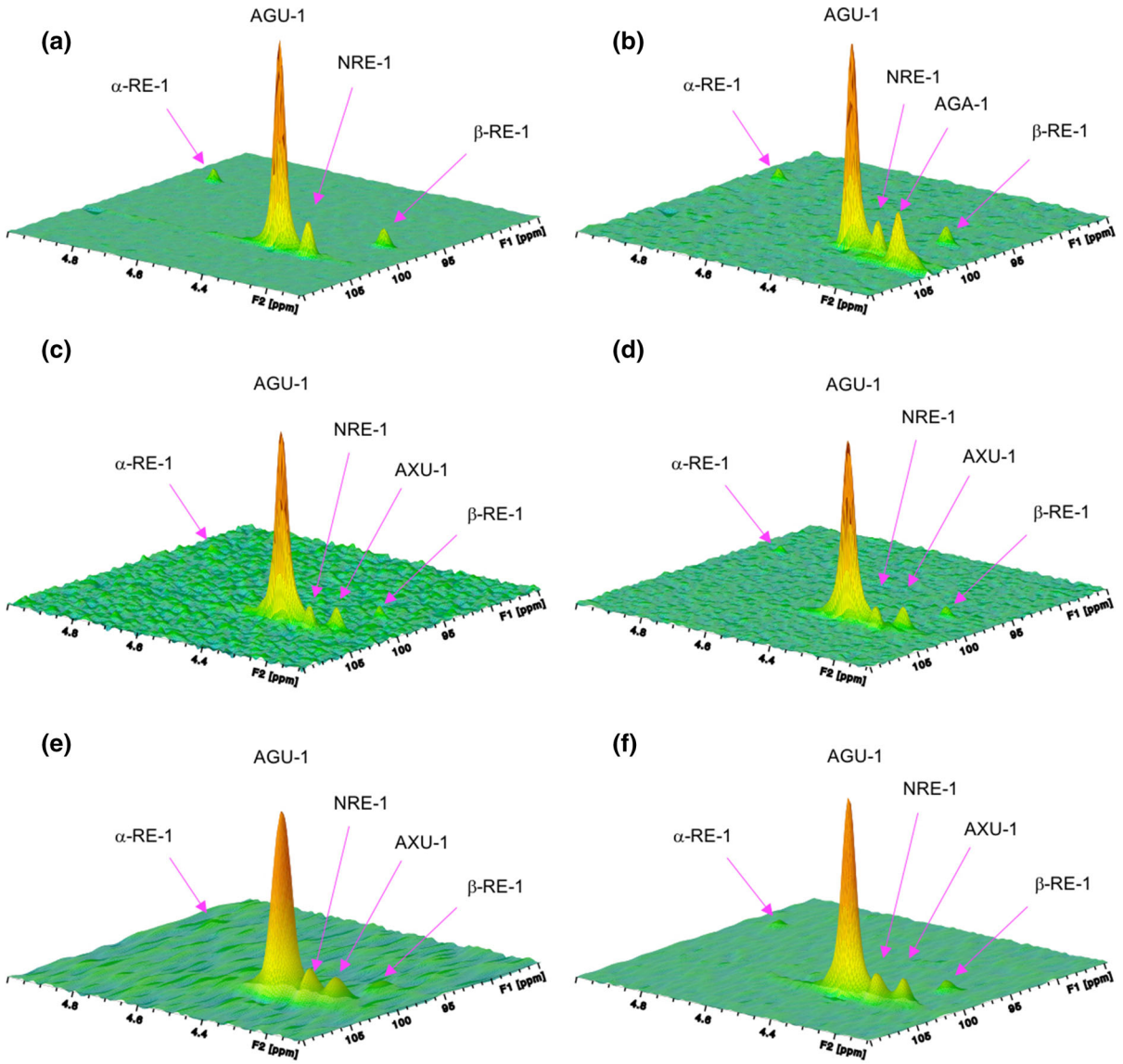


Fig. 10 2D Q-CAHSQC (quantitative HSQC) 3D projections of the CH1 region for: **a** LDP-CNCs, **b** TOx-LDP-CNCs, **c** MCC (ns = 8, td1 = 640) with a room temperature probe-head, **d** MCC (ns = 8, td1 = 640) with a cryoprobe-head, **e** MCC (ns = 40, td1 = 128) with a room temperature probe-head, and **f** MCC (ns = 40, td1 = 128) with a cryoprobe-head. F1 is the ^{13}C dimension and F2 is the ^{1}H dimension. No forward linear prediction was used to improve resolution

from fast T_2 relaxation. However, if extensive milling is applied to allow for full solubilization and low enough molecular weight (further increasing T_2 values), more accurate quantitation may be close. In this context, suitable stable direct-dissolution solvents for solution-state NMR analysis of whole biomass samples has been something that has also been lacking. Cheng et al. (2013) have already demonstrated this principal, by complete dissolution and analysis of ball-milled Miscanthus into a mixture of the cellulose-dissolving ionic liquid 1-ethyl-3-methylimidazolium acetate ([emim][OAc]) in DMSO-d_6. Perdeuterated [emim][OAc] ([emim][OAc]-d_{14}):DMSO-d_6 was then prepared and used for the application of a quantitative HSQC experiment on fully dissolved solutions. However, as mentioned previously, [emim][OAc] is known to react with cellulose (Liebert and Heinze 2008; Ebner et al. 2008, Clough et al. 2015). It also has

significant signal overlap with the polysaccharide spectral region and [emim][OAc]-d_{14} is much too expensive and laborious to prepare, for routine analyses. Nevertheless, the [emim][OAc]:DMSO-d_6 solutions were shown to be stable over a longer period (2 weeks), whereas the molecular solvent dispersions showed phase-separation. This is a good indication that the current solvent-system may open the window to a much wider range of samples.

Conclusions

The chemical shifts of polymeric units in cellulose, including NRE and RE units can be unambiguously assigned using solution-state NMR in a novel ionic liquid electrolyte, [P_{4444}][OAc]:DMSO-d_6. The main monomeric units in 4-AcNH-TEMPO oxidized cellulose (polyglucuronic acid) are also assigned, as are the terminal units for the unoxidized and oxidized materials. The latter has led to identification of the terminal open-chain gluconate moiety after both the used acidic 4-AcNH-TEMPO protocol and Pinnick oxidation conditions. However, in both instances RE groups remained in the oxidized products, indicating a further need for optimization of this reaction for different substrates or more defined structural characterization of substrates for surface oxidation sites, that may undergo β-elimination, yielding new reducing ends. Periodate oxidation of cellulose clearly introduces instability into cellulose, when the dry oxidation product was introduced to the basic electrolyte media. This degradation mechanism is thought to be similar to the β-elimination mechanisms, illustrated in previous publications, for aqueous alkaline media but obviously requires further study to elucidate the mechanism of degradation. Further investigations into how to stabilize the periodate oxidation products towards basic degradation, by further chemical modifications, are needed as periodate oxidation of cellulose is a widely utilized technique. However, NMR analysis in the electrolyte medium seems to be a useful probe into the stability of these compounds, in addition to providing the necessary chemical species resolution that other techniques cannot. Of course, this is also a direct method to follow the progress of oxidation reactions. Nitroxyl-radical-type oxidations (to 6 position carboxylates), under mild acidic conditions, seems to be quite robust, in terms of resulting product stability in the electrolyte and under aqueous alkaline conditions. Thus, avoiding aldehyde formation under alkaline oxidation conditions is clearly important in improving the quality of the oxidized products, by preventing losses and molecular weight reduction due to fragmentation of surface chains. Q-CAHSQC, with T_2 correction, seems to be a suitable experiment and processing combination to yield quantitative data from HSQC, without calibration against internal standards. While this is still not suitable for accurate determination of DP_N for higher molecular weight and low polydispersity samples, accurate DS and regioselectivity determination will be possible for certain chemical modifications, even at reducing ends in lower molecular weight samples, such as model CNCs. However, it should be stressed that this solvent system and processing strategy are not only applicable to nanocelluloses but offer the chance to significantly improve our opportunities for quantitative analysis of whole biomass samples, that contain a significant crystalline cellulose phase composition.

Acknowledgments Open access funding provided by University of Helsinki including Helsinki University Central Hospital. The authors would like to thank the Academy of Finland for funding under the project 'WTF-Click-Nano' (Project #: 311255). The authors would also like to thank Prof. Herbert Sixta for help in choosing the model cellulose materials.

References

Agrawal PK (1992) NMR spectroscopy in the structural elucidation of oligosaccharides and glycosides. Phytochemistry 31:3307–3330. https://doi.org/10.1016/0031-9422(92)83678-r

Bax A, Summers MF (1986) Proton and carbon-13 assignments from sensitivity-enhanced detection of heteronuclear multiple-bond connectivity by 2D multiple quantum NMR. J Am Chem Soc 108:2093–2094. https://doi.org/10.1021/ja00268a061

Buffiere J, Ahvenainen P, Borrega M et al (2016) Supercritical water hydrolysis: a green pathway for producing low-molecular-weight cellulose. Green Chem 18:6516–6525. https://doi.org/10.1039/C6GC02544G

Cheng K, Sorek H, Zimmermann H, Wemmer DE, Pauly M (2013) Solution-state 2D NMR spectroscopy of plant cell walls enabled by a dimethylsulfoxide-d_6/1-ethyl-3-methylimidazolium acetate solvent. Anal Chem 85:3213–3221. https://doi.org/10.1021/ac303529v

Clough MT, Geyer K, Hunt PA, Son S, Vagt U, Welton T (2015) Ionic liquids: not always innocent solvents for cellulose. Green Chem 17:231–243. https://doi.org/10.1039/C4GC01955E

Deb S, Labafzadeh SR, Liimatainen U et al (2016) Application of mild autohydrolysis to facilitate the dissolution of wood chips in direct-dissolution solvents. Green Chem 18:3286–3294. https://doi.org/10.1039/C6GC00183A

Ebner G, Schiehser S, Potthast A, Rosenau T (2008) Side reaction of cellulose with common 1-alkyl-3-methylimidazolium-based ionic liquids. Tetrahedron Lett 49:7322–7324. https://doi.org/10.1016/j.tetlet.2008.10.052

Fosten M, Samuel R, He J, Ragauskas AJ (2016) A review of whole cell wall NMR by the direct-dissolution of biomass. Green Chem 18:608–621. https://doi.org/10.1039/c5gc02828k

Foster EJ, Moon RJ, Agarwal UP, Bortner MJ, Bras J, Camarero-Espinosa S, Chan KJ, Clift MJD, Cranston ED, Eichhorn SJ, Fox DM, Hamad WY, Heux L, Jean B, Korey M, Nieh W, Ong KJ, Reid MS, Renneckar S, Roberts R, Shatkin JA, Simonsen J, Stinson-Bagby K, Wanasekara N, Youngblood J (2018) Current characterization methods for cellulose nanomaterials. Chem Soc Rev 47:2609–2679. https://doi.org/10.1039/C6CS00895J

French AD (2014) Idealized powder diffraction patterns for cellulose polymorphs. Cellulose 21:885–896. https://doi.org/10.1007/s10570-013-0030-4

Fujisawa S, Isogai T, Isogai A (2010) Temperature and pH stability of cellouronic acid. Cellulose 17:607–615. https://doi.org/10.1007/s10570-010-9407-9

Habibi Y, Lucia LA, Rojas OJ (2010) Cellulose nanocrystals: chemistry, self-assembly, and applications. Chem Rev 110:3479–3500. https://doi.org/10.1021/cr900339w

Heikkinen S, Toikka MM, Karhunen PT, Kilpeläinen I (2003) Quantitative 2D HSQC (Q-HSQC) via suppression of *J*-dependence of polarization transfer in NMR spectroscopy: application to wood lignin. J Am Chem Soc 125:4362–4367. https://doi.org/10.1021/ja029035k

Heinze T, Dicke R, Koschella A, Henning Kull A, Klohr E-A, Koch W (2000) Effective preparation of cellulose derivatives in a new simple cellulose solvent. Macromol Chem Phys 201:627–631. https://doi.org/10.1002/(SICI)1521-3935(20000301)201:6<627:AID-MACP627>3.0.CO;2-Y

Heise K, Koso T, Pitkänen L, Potthast A, King AWT, Kostiainen MA, Kontturi E (2019) Knoevenagle condensation for modifying the reducing end groups of cellulose nanocrystals. ACS Macro Lett 8:1642–1647. https://doi.org/10.1021/acsmacrolett.9b00838

Hirota M, Tamura N, Saito T, Isogai A (2009) Oxidation of regenerated cellulose with $NaClO_2$ catalyzed by TEMPO and NaClO under acid-neutral conditions. Carbohydr Polym 78:330–335. https://doi.org/10.1016/j.carbpol.2009.04.012

Holding AJ, Mäkelä V, Tolonen L et al (2016) Solution-state one- and two-dimensional nmr spectroscopy of high-molecular-weight cellulose. Chemsuschem 9:880–892. https://doi.org/10.1002/cssc.201501511

Hosoya T, Bacher M, Potthast A et al (2018) Insights into degradation pathways of oxidized anhydroglucose units in cellulose by β-alkoxy-elimination: a combined theoretical and experimental approach. Cellulose 25:3797–3814. https://doi.org/10.1007/s10570-018-1835-y

https://bruker.com/ Bruker TopSpin 4.0

https://fityk.nieto.pl/*Fityk* 1.3.1

https://mestrelab.com/ MestReNova 10.0

Hu K, Westler WM, Markley JL (2011) Simultaneous quantification and identification of individual chemicals in metabolite mixtures by two-dimensional extrapolated time-zero ^{1}H–^{13}C HSQC ($HSQC_0$). J Am Chem Soc 133:1662–1665. https://doi.org/10.1021/ja1095304

Isogai A, Saito T, Fukuzumi H (2011) TEMPO-oxidized cellulose nanofibers. Nanoscale 3:71–85. https://doi.org/10.1039/C0NR00583E

Isogai A, Hänninen T, Fujisawa S, Saito T (2018) Review: Catalytic oxidation of cellulose with nitroxyl radicals under aqueous conditions. Prog Polym Sci 86:122–148. https://doi.org/10.1016/j.progpolymsci.2018.07.007

Kim TH (2008) Pulsed NMR: Relaxation times as function of viscocity and impurities. 1–5

Kim U-J, Kuga S, Wada M, Okano T, Kondo T (2000) Periodate oxidation of crystalline cellulose. Biomacromol 1:488–492. https://doi.org/10.1021/bm0000337

Kilpeläinen I, Xie H, King AWT, Granström M, Heikkinen S, Argyropoulos DS (2007) Dissolution of Wood in Ionic Liquids. J Agric Food Chem 55:9142–9148. https://doi.org/10.1021/jf071692e

King AWT, Mäkelä V, Kedzior SA et al (2018) Liquid-state NMR analysis of nanocelluloses. Biomacromol 19:2708–2720. https://doi.org/10.1021/acs.biomac.8b00295

Klemm D, Heublein B, Fink H-P, Bohn A (2005) Cellulose: fascinating biopolymer and sustainable raw material. Angewandte Chemie Int Ed 44:3358–3393. https://doi.org/10.1002/anie.200460587

Kono H, Yunoki S, Shikano T, Fujiwara M, Erata T, Takai M (2002) CP/MAS ^{13}C NMR study of cellulose and cellulose derivatives. 1. complete assignment of the CP/MAS ^{13}C NMR spectrum of the native cellulose. J Am Chem Soc 124:7506–7511. https://doi.org/10.1021/ja010704o

Koskela H, Kilpeläinen I, Heikkinen S (2005) Some aspects of quantitative 2D NMR. J Magn Reson 174:237–244. https://doi.org/10.1016/j.jmr.2005.02.002

Koskela H, Heikkilä O, Kilpeläinen I, Heikkinen S (2010) Quantitative two-dimensional HSQC experiment for high magnetic field NMR spectrometers. J Magn Reson 202:24–33. https://doi.org/10.1016/j.jmr.2009.09.021

Kyllönen L, Parviainen A, Deb S, Lawoko M, Gorlov M, Kilpeläinen I, King AWT (2013) On the solubility of wood in non-derivatising ionic liquids. Green Chem 15:2374–2378. https://doi.org/10.1039/C3GC41273C

Liebert T, Heinze T (2008) Interaction of ionic liquids with polysaccharides 5. Solvents and reaction media for the modification of cellulose. BioResources 3:576–601

Lin F, Cousin F, Putaux J-L, Jean B (2019) Temperature-controlled star-shaped cellulose nanocrystal assemblies resulting from asymmetric polymer grafting. ACS Macro Lett 8:345–351. https://doi.org/10.1021/acsmacrolett.8b01005

Ling Z, Wang T, Makarem M, Santiago Cintrón M, Cheng HN, Kang X, Bacher M, Potthast A, Rosenau T, King H, Delhom CD, Nam S, Edwards JV, Kim SH, Xu F, French AD (2019) Effects of ball milling on the structure of cotton cellulose. Cellulose 26:305–328. https://doi.org/10.1007/s10570-018-02230-x

Mansfield SD, Kim H, Lu F, Ralph J (2012) Whole plant cell wall characterisation using solution-state 2D NMR. Nat Protoc 7:1579–1589. https://doi.org/10.1038/nprot.2012.064

Mäkelä V, Helminen JKJ, Kilpeläinen I, Heikkinen S (2016) Quantitative, equal carbon response HSQC experiment QEC HSQC. J Magn Reson 271:34–39. https://doi.org/10.1016/j.jmr.2016.08.003

Newman RH (1999) Estimation of the lateral dimensions of cellulose crystallites using ^{13}C NMR signal strengths. Solid State Nucl Magn Reson 15:21–29. https://doi.org/10.1016/S0926-2040(99)00043-0

Nypelö T, Amer H, Konnerth J et al (2018) Self-standing nanocellulose janus-type films with aldehyde and carboxyl functionalities. Biomacromol 19:973–979. https://doi.org/10.1021/acs.biomac.7b01751

Östlund Å, Lundberg D, Nordstierna L, Holmberg K, Nydén M (2009) Dissolution and gelation of cellulose in TBAF/DMSO solutions: the roles of fluoride ions and water. Biomacromol 10:2401–2407. https://doi.org/10.1021/bm900667q

Peterson DJ, Loening NM (2007) QQ-HSQC: a quick, quantitative heteronuclear correlation experiment for NMR spectroscopy. Magn Reson Chem 45:937–941. https://doi.org/10.1002/mrc.2073

Potthast A, Schiehser S, Rosenau T, Kostic M (2009) Oxidative modifications of cellulose in the periodate system—reduction and beta-elimination reactions. Holzforschung 63:12–17. https://doi.org/10.1515/HF.2009.108

Roslund MS, Tähtinen P, Niemitz M, 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. https://doi.org/10.1016/j.carres.2007.10.008

Röhrling J, Potthast A, Rosenau T, Lang T, Ebner G, Sixta H, Kosma P (2002) A novel method for the determination of carbonyl groups in cellulosics by fluorescence labeling. 1. Method development. Biomacromol 3:959–968. https://doi.org/10.1021/bm020029q

Schleucher J, Schwendinger M, Sattler M et al (1994) A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. J Biomol NMR 4:301–306. https://doi.org/10.1007/BF00175254

Sun H, DiMagno SG (2005) Anhydrous tetrabutylammonium fluoride. J Am Chem Soc 127:2050–2051. https://doi.org/10.1021/ja0440497

Villares A, Moreau C, Cathala B (2018) Star-like supramolecular complexes of reducing-end-functionalized cellulose nanocrystals. ACS Omega 3:16203–16211. https://doi.org/10.1021/acsomega.8b02559

Willker W, Leibfritz D, Kerssebaum R, Bermel W (1993) Gradient selection in inverse heteronuclear correlation spectroscopy. Magn Reson Chem 31:287–292. https://doi.org/10.1002/mrc.1260310315

Wojdyr M (2010) *Fityk*: a general-purpose peak fitting program. J Appl Cryst 43:1126–1128. https://doi.org/10.1107/S0021889810030499

Zhang L, Gellerstedt G (2007) Quantitative 2D HSQC NMR determination of polymer structures by selecting suitable internal standard references. Magn Reson Chem 45:37–45. https://doi.org/10.1002/mrc.1914

Zuckerstätter G, Schild G, Wollboldt RT, Weber HK, Sixta H (2009) The elucidation of cellulose supramolecular structure by ^{13}C CP-MAS NMR. Lenzinger Berichte 87:38–46