J Biomol NMR (2014) 59:95–110
DOI 10.1007/s10858-014-9830-6

ARTICLE

NMR structure analysis of uniformly ^{13}C-labeled carbohydrates

Carolina Fontana · Helena Kovacs · Göran Widmalm

Received: 19 March 2014 / Accepted: 16 April 2014 / Published online: 26 April 2014

Abstract In this study, a set of nuclear magnetic resonance experiments, some of them commonly used in the study of ^{13}C-labeled proteins and/or nucleic acids, is applied for the structure determination of uniformly ^{13}C-enriched carbohydrates. Two model substances were employed: one compound of low molecular weight [(UL-^{13}C)-sucrose, 342 Da] and one compound of medium molecular weight (^{13}C-enriched O-antigenic polysaccharide isolated from *Escherichia coli* O142, ~10 kDa). The first step in this approach involves the assignment of the carbon resonances in each monosaccharide spin system using the anomeric carbon signal as the starting point. The ^{13}C resonances are traced using ^{13}C–^{13}C correlations from homonuclear experiments, such as (H)CC–CT–COSY, (H)CC–NOESY, CC–CT–TOCSY and/or virtually decoupled (H)CC–TOCSY. Based on the assignment of the ^{13}C resonances, the ^{1}H chemical shifts are derived in a straightforward manner using one-bond ^{1}H–^{13}C correlations from heteronuclear experiments (HC–CT–HSQC). In order to avoid the $^{1}J_{CC}$ splitting of the ^{13}C resonances and to improve the resolution, either constant-time (CT) in the indirect dimension or virtual decoupling in the direct dimension were used. The monosaccharide sequence and linkage positions in oligosaccharides were determined using either ^{13}C or ^{1}H detected experiments, namely CC–CT–COSY, band-selective (H)CC–TOCSY, HC–CT–HSQC–NOESY or long-range HC–CT–HSQC. However, due to the short T_2 relaxation time associated with larger polysaccharides, the sequential information in the O-antigen polysaccharide from *E. coli* O142 could only be elucidated using the ^{1}H-detected experiments. Exchanging protons of hydroxyl groups and *N*-acetyl amides in the ^{13}C-enriched polysaccharide were assigned by using HC–H2BC spectra. The assignment of the *N*-acetyl groups with ^{15}N at natural abundance was completed by using HN–SOFAST–HMQC, HNCA, HNCO and ^{13}C-detected (H)CACO spectra.

Keywords Carbohydrates · ^{13}C-uniform labeling · NMR · Structure determination

Introduction

Carbohydrates, also known as glycans, are one of the major classes of biopolymers found in nature. They play an essential role in a wide range of biological processes, for instance, in bacterial recognition and initiation of the host immune response (Aich and Yarema 2009; Ghazarian et al. 2011; Varki et al. 2009). The number of different structures that can be generated with just a few monosaccharides is enormous when compared to other biopolymers, making a detailed structural analysis crucial for the understanding of the role of these molecules in biological systems. Nuclear magnetic resonance (NMR) spectroscopy is one of the most powerful techniques to study these molecules in solution. The classical approach in NMR structural elucidation of carbohydrates takes advantage of the higher sensitivity and abundance of the ^{1}H spin nuclei: the different spin systems are characterized using proton–proton correlations from homonuclear experiments, and the protons are then

Electronic supplementary material The online version of this article (doi:10.1007/s10858-014-9830-6) contains supplementary material, which is available to authorized users.

C. Fontana · G. Widmalm (✉)
Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, 106 91 Stockholm, Sweden
e-mail: gw@organ.su.se

H. Kovacs
Bruker BioSpin AG, Industriestrasse 26, 8117 Fällanden, Switzerland

connected to their respective ^{13}C resonances through one-bond heteronuclear correlations. However, the structural characterization of glycans by this approach is seriously hindered by severe spectral overlap of the ^{1}H resonances.

The use of ^{13}C-detected experiments is well established in the NMR spectroscopy of ^{13}C-enriched proteins (Bermel et al. 2006, 2008) and nucleic acids (Farès et al. 2007; Fiala and Sklenár 2007; Richter et al. 2010), but only limited use has been made of ^{13}C-enriched carbohydrates (Battistel et al. 2012; Harris et al. 1997; Kiddle and Homans 1998; Kjellberg et al. 1998; Martin-Pastor et al. 2003; Martin-Pastor and Bush 2000; Norris et al. 2012; Wang et al. 2008; Xu and Bush 1998; Yu et al. 1993). Glycans do not necessarily assume a particular fold free in solution, instead, they may appear in extended, flexible conformations with mere transient structural elements (Martin-Pastor and Bush 2000; Sarkar et al. 2013). This dynamic behavior renders sharper NMR signals albeit with a reduced chemical shift dispersion. Thus, there is a certain analogy to intrinsically disordered proteins (IDP) (Felli and Pierattelli 2012). Recently, ^{13}C-detected NMR methods have become instrumental in the investigations of IDPs (Felli and Pierattelli 2012; Sibille and Bernadó 2012), and we here explore their applicability for glycans. ^{13}C-labeling avoids the problem of the reduced ^{1}H chemical shifts dispersion often found in carbohydrates and facilitates the complete NMR analysis through the large chemical shift dispersion of the ^{13}C spins. Consequently, given the access to ^{13}C-labeled carbohydrates (Fairweather et al. 2004; Kamiya et al. 2011; Kato et al. 2010), a set of optimal pulse sequences for the structural elucidation needs to be defined.

Even though the spectral dispersion of ^{13}C spins is by far larger than that of the ^{1}H spins, the resolution of the NMR spectra of uniformly ^{13}C-labeled compounds is reduced by the large homonuclear one-bond ^{13}C–^{13}C couplings, as well as a variety of smaller ^{13}C–^{13}C long-range couplings. Since carbohydrate spectra are often very crowded, the removal of the splitting of the resonances in the direct and/or indirect dimensions is a critical point to be addressed. In this regard, the main strategies discussed in this work include the use of constant-time (CT) experiments for removal of the large homonuclear splitting in the indirect dimension, or the use of in-phase anti-phase (IPAP) or double in-phase anti-phase (DIPAP) schemes (Bermel et al. 2006) for virtual decoupling in the direct dimension.

The two ^{13}C-enriched compounds used herein are uniformly ^{13}C-labelled [UL-^{13}C]-sucrose and uniformly ^{13}C-enriched O-antigen polysaccharide of *Escherichia coli* (*E. coli*) O142 (Landersjö et al. 1997), whereas the [1-^{13}C]-enriched O-antigen polysaccharide of *E. coli* O91 (Lycknert and Widmalm 2004) ($\sim$10 kDa) is used for the determination of the polymer dynamics (cf. Fig. 1). As one of the major targets of the host immune response, the O-antigen polysaccharide plays a critical role in host-pathogen interactions. This is the most variable part of the lipopolysaccharide and its serological specificity is used as one of the major bases for serotyping schemes in gram-negative bacteria (DebRoy et al. 2011; Stenutz et al. 2006). In the case of *E. coli*, 174 serogroups are currently described and some of them are considered pathogenic, such is the case of the aforementioned *E. coli* O142 and O91 serogroups which have been classified as enteropathogenic *E. coli* (EPEC) (Bugarel et al. 2011) and Shiga toxing-producing *E. coli* (STEC) (Son et al. 2014), respectively. Analogously to what is found in proteins and nucleic acids, the NH groups of amino sugars can also act as hydrogen donors and influence the three-dimensional structure of carbohydrates through inter-residue hydrogen bonding interactions (Norris et al. 2012). Furthermore, the many hydroxyl groups available in these molecules can act either as hydrogen donors or acceptors but, since the H_2O molecules can also compete for the same hydrogen bond, the use of aprotic co-solvents and/or low temperatures may be required to detect such kind of interactions (Battistel et al. 2013; Norris et al. 2012). Whether hydrogen bonds involving hydroxyl groups may or may not influence the three-dimensional fold of carbohydrates under physiological conditions is a question that still remains unanswered; they have, however, proven to play a critical role in carbohydrate–protein interactions such as in the recognition by antibodies (Villeneuve et al. 2000). As a consequence, the possibility to detect NH and OH protons by NMR spectroscopy, and unambiguously assign these resonances, is of paramount importance in understanding the function of these molecules in biological systems.

Materials and methods

Sample preparation

Commercially available [UL-^{13}C]-sucrose (99 % ^{13}C-enrichment) was purchased from ISOTEC. The ^{13}C-enriched O-antigen polysaccharide of *E. coli* O142 was obtained as previously reported by supplementing the LB growth medium with [UL-^{13}C]-D-glucose and subsequently isolating the lipopolysaccharide from the outer membrane of the bacterium followed by pertinent purification (Landersjö et al. 1997; Norris et al. 2012). The [1-^{13}C]-enriched O-antigen polysaccharide of *E. coli* O91, available from a previous study (Kjellberg et al. 1999), had been prepared in a similar way. The concentration of sucrose was 8–22 mg mL^{-1} (23–64 mM) and the concentration of

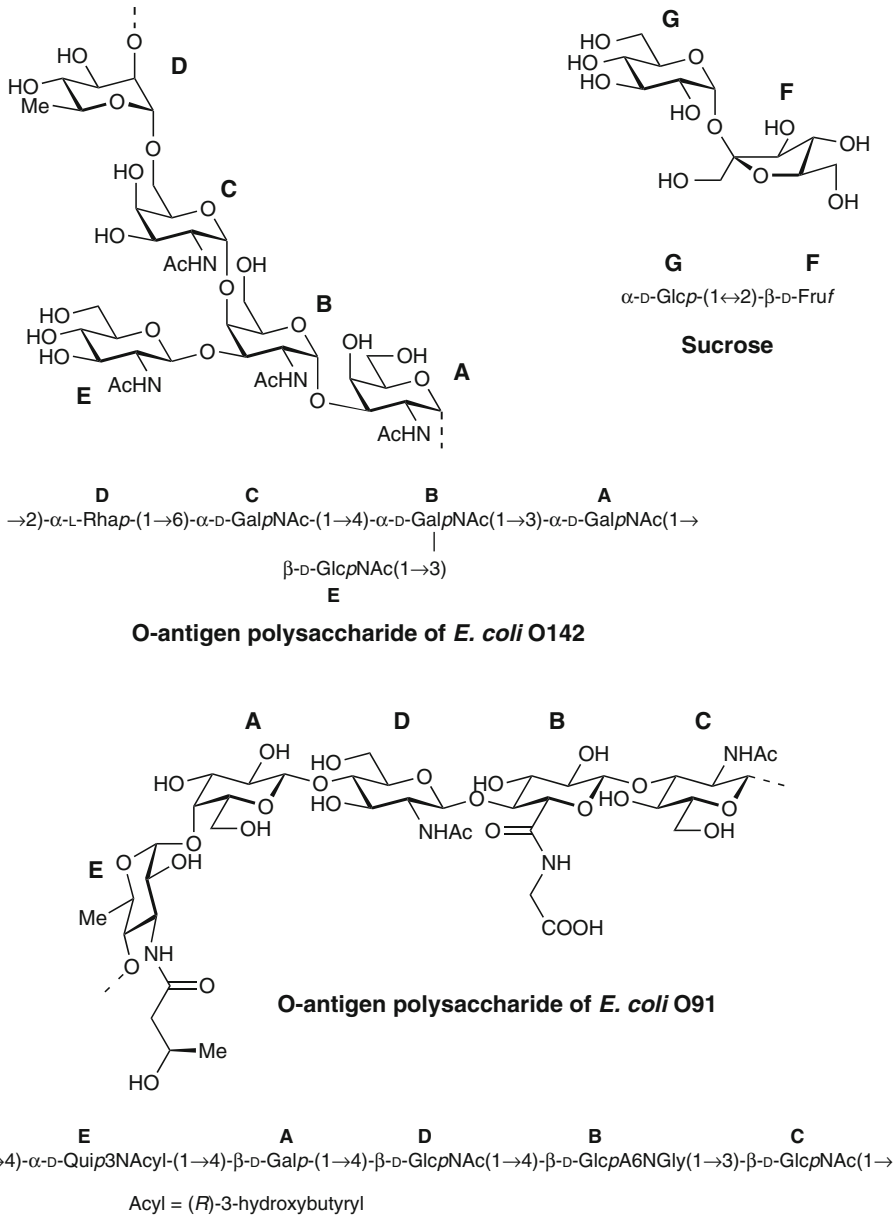

Fig. 1 Structures of the ^{13}C-enriched compounds used in this study. Representation of the structures of the repeating units of the O-antigen polysaccharides of *E. coli* O142 and O91 (*top left* and *bottom*, respectively), and sucrose (*top right*), in schematic and standard nomenclature. Sugar residues are denoted by *capital letters*

O142 was typically 3.0–6.0 mg mL^{-1} (corresponding to an effective repeating unit concentration of 3–6 mM). The O142 O-antigen polysaccharide chain consists of a basic unit of five monosaccharides that is repeated ~10 times in a single molecule. Furthermore, the O142 polysaccharide contains four *N*-acetyl groups per repeating unit, with ^{15}N at natural abundance in the present case.

NMR spectroscopy

Unless otherwise specified the experiments were carried out on Bruker Avance III 600 MHz or 700 MHz spectrometers equipped with a 5 mm TCI CryoProbe^{TM} (with the nuclei ^{1}H–^{13}C/^{15}N).

NMR spectra of [UL-^{13}C]-sucrose

Unless otherwise specified the experiments were recorded at 25 °C in D_2O solution (4 mg in 0.5 mL) at magnetic field strength of 16.4 T.

The CC–CT–COSY (correlated spectroscopy) (Bermel et al. 2003) spectrum of Fig. 2b was acquired over a spectral region of 60 ppm in both dimensions, using acquisition times of 97 and 10 ms in *F2* and *F1*, respectively, and 16 scans per increment. A CT delay (2*T*) of 12 ms was employed. The spectrum of Fig. 7a was acquired with the same parameters as above but using acquisition times of 97 and 20 ms in *F2* and *F1*, respectively, and a 2*T* delay of 22 ms. The CC–CT–TOCSY (total correlated spectroscopy)

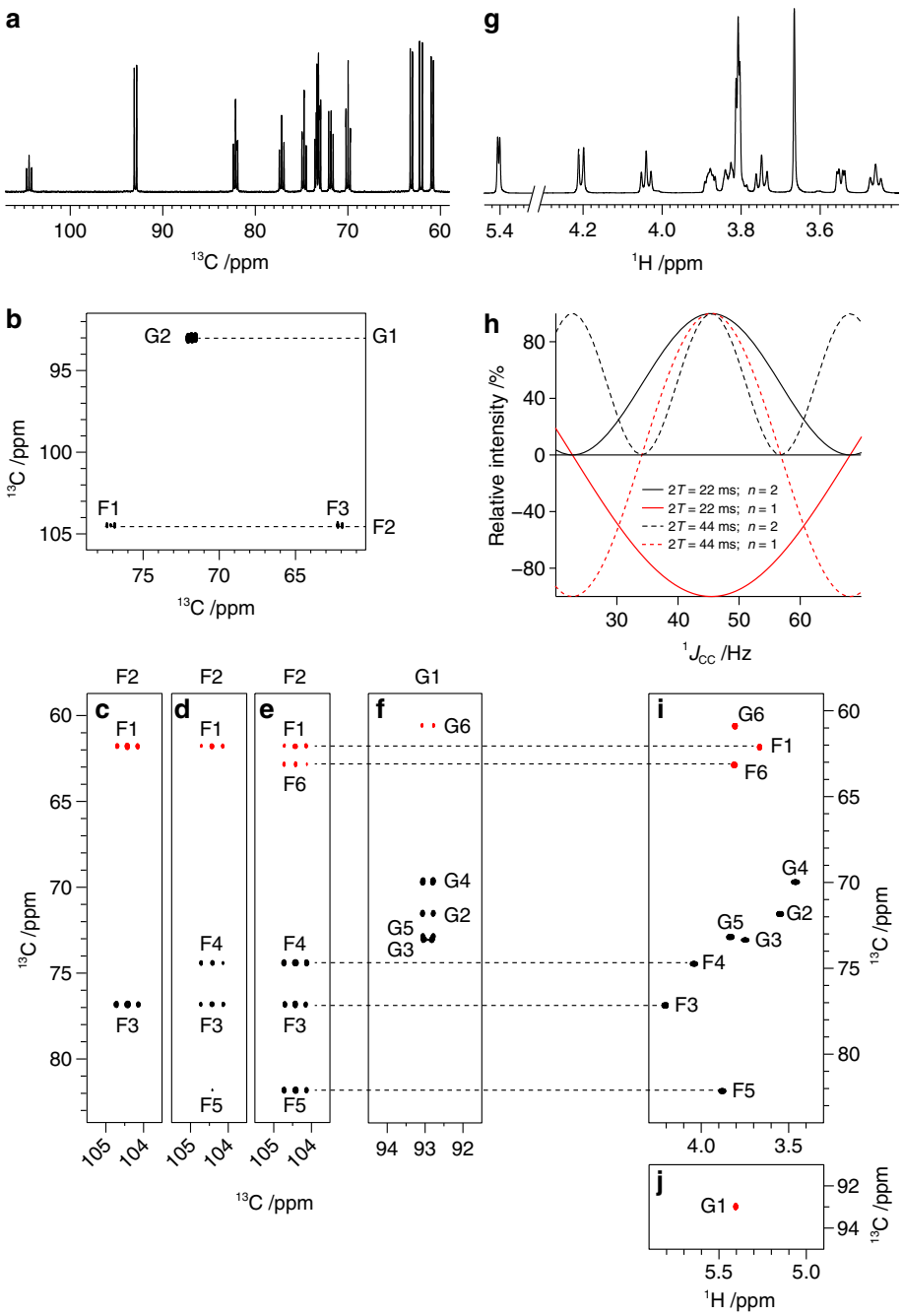

Fig. 2 ^{13}C and ^{1}H chemical shift assignments of oligosaccharides. **a** ^{1}H-decoupled ^{13}C spectrum of [UL-^{13}C]-sucrose, **b** selected region of the CC–CT–COSY spectrum ($2T = 12$ ms) showing correlations from anomeric carbons. **c–f** Selected regions of the CC–CT–TOCSY spectra of [UL-^{13}C]-sucrose ($2T = 22$ ms) recorded with different mixing times ($\tau_m = 4.7$, 9.4, 14.1 and 18.8 ms, from panel **c** to **f**, respectively) showing correlations from the anomeric carbon of the fructose (**c–e**) and glucose (**f**) residues. **g** ^{13}C-decoupled ^{1}H NMR spectrum of [UL-^{13}C]-sucrose. **h** Plot of the expected peaks intensities in the HC–CT–HSQC spectrum as a function of the one-bond carbon–carbon couplings and the number of neighboring aliphatic carbons (n) using two different CT values.(Vuister and Bax 1992) **i, j)** The HC–CT–HSQC spectrum of [UL-^{13}C]-sucrose ($2T = 22$ ms) showing the region for the ring atoms and those from hydroxymethyl groups (**i**), as well as the anomeric region (**j**). The sign of the ^{13}C magnetization is opposite for carbons directly attached to an odd versus an even number of neighboring aliphatic carbons (shown in *red* and *black color*, respectively)

(Eletsky et al. 2003) spectra of Fig. 2c–f were recorded over a spectral region of 60 ppm in both dimensions using acquisition times of 97 and 22 ms in *F2* and *F1*, respectively, 8 scans per increment, and a CT delay ($2T$) of 22 ms. The total length of the FLOPSY-16 mixing sequence (Kadkhodaie et al. 1991) is given by the following equation: mixing time (τ_m) = 188.448 × (length 90° pulse) × n, where n is the number of times the cycle is repeated. Since the length of the 90° pulse was 25 µs, spectra with four different mixing times (4.7, 9.4, 14.1 and 18.8 ms) were recorded.

The HC–CT–HSQC (heteronuclear single quantum coherence) (Vuister and Bax 1992) spectrum (Fig. 2i, j) was recorded over a spectral width of 6 × 80 ppm, using acquisition times of 122 and 9 ms in *F2* and *F1*, respectively, 8 scans per increment and a $2T$ value of 22 ms.

The band-selective (H)CC–TOCSY spectrum (Fig. 7b) was recorded over a spectral width of 70 ppm in both dimensions, using acquisition times of 83 and 8 ms in *F2* and *F1*, respectively, and 2 scans per increment. For the anomeric selective adiabatic ^{13}C–^{13}C-spinlock a constant adiabaticity (ca) WURST–2 shape (Kupče et al. 1998) with

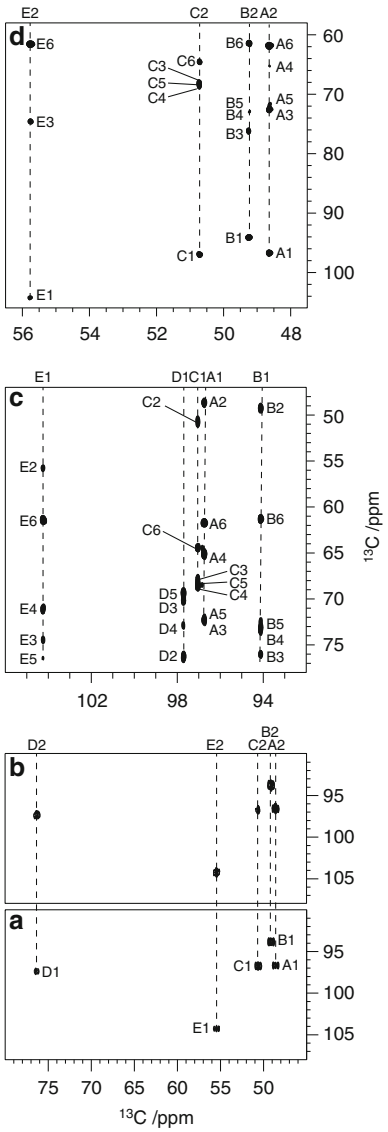

Fig. 3 ^{13}C chemical shift assignment in polysaccharides. Selected region of the **a** (H)CC–CT–COSY ($2T = 10$ ms) and **b** (H)CC–NOESY ($\tau_m = 500$ ms) spectra of the ^{13}C-enriched O-antigen polysaccharide of *E. coli* O142 showing correlations from the C2 carbons to the anomeric resonances. Selected regions of (H)CC–TOCSY spectra ($\tau_m = 20$ ms) recorded with: **c** virtual decoupling of $^1J_{C1,C2}$ couplings in the direct *dimension* using the IPAP scheme (Bermel et al. 2006) and **d** simultaneous decoupling of $^1J_{C1,C2}$ and $^1J_{C2,C3}$ couplings in the direct *dimension* using the DIPAP scheme (Bermel et al. 2006)

5.3 kHz nominal sweep, 800 μs duration and amplitude power index of 2 was chosen. The shape was then expanded with a p5p9 phase cycle (Kupče et al. 1998) which gave a total duration of 36 ms and a total rotation of zero in a 4.3 kHz wide region at the peak power of 2.6 kHz. The average power of the shape was, however, only 37 %. The selective Pc9_4_90.1000 excitation pulse of 1 ms (Kupče and Freeman 1994) and the adiabatic mixing pulse (caWURST, 5.3 kHz, 36 ms) were centered at the middle of the region for the anomeric carbons, and a total spinlock time of 144 ms was employed.

The long-range HC–CT–HSQC spectrum (Fig. 7c) was recorded using the same conditions as for the HC–CT–HSQC spectrum described above, but the delay for evolution of the proton–carbon couplings was optimized for $^nJ_{CH} = 12$ Hz instead of 145 Hz. The HC–CT–HSQC–NOESY spectrum (Fig. 7d) was acquired at a magnetic field strength of 14.1 T, using a room temperature TXI probe. The pulse sequence used to carry out this experiment was derived from the same HC–CT–HSQC sequence described above (Vuister and Bax 1992) but inserting a NOESY (nuclear Overhauser effect spectroscopy) block (Parella et al. 1997) just prior to the acquisition. The spectrum was recorded over a spectral region of 7 × 60 ppm using acquisition times of 122 and 14 ms in *F2* and *F1*, respectively, 4 scans per increment, a $2T$ value of 22 ms and a mixing time of 500 ms.

NMR spectra of the ^{13}C-enriched O-antigen polysaccharide of *E. coli* O142

The experiments were recorded at different temperatures ranging from 2 to 70 °C. The (H)CC–CT–COSY, (H)CC–NOESY, virtual-decoupled (H)CC–TOCSY, HC–H2BC (heteronuclear 2-bond correlation), HC(C)H–COSY and HN–SOFAST–HMQC (band-selective optimized flip angle short transient HMQC) experiments were acquired in H_2O/D_2O 95:5 solution (2–3 mg in 0.5 mL).

The (H)CC–CT–COSY spectrum (Fig. 3a) was recorded at 3 °C and at a magnetic field strength of 14.1 T using a pulse sequence similar to the one described above, but with a proton starting block implemented for better sensitivity. The spectrum was recorded over a spectral region of 120 × 100 ppm using acquisition times of 113 and 7 ms in *F2* and *F1*, respectively, 80 scans per increment and a $2T$ value of 10 ms. The (H)CC–NOESY spectrum (Fig. 3b) was recorded under the same conditions as for the (H)CC-CT-COSY, using a standard pulse sequence (Bertini et al. 2003, 2004) but with a proton starting block implemented for better sensitivity. The spectrum was recorded over a spectral region of 180 × 180 ppm using acquisition times of 75 and 4 ms in *F2* and *F1*, respectively, 128 scans per increment and $\tau_m = 500$ ms.

The (H)CC–TOCSY spectrum ($\tau_m = 20$ ms) with virtual decoupling of $^1J_{C1,C2}$ in the direct dimension (Fig. 3c) was recorded at 40 °C at a magnetic field strength of 14.1 T using a 2D version of the IPAP (Bermel et al. 2006) pulse sequence described by Richter et al. (Richter et al. 2010) for investigations of ribose in RNA. For the virtual decoupling scheme a band-selective 180° refocusing Reburp.1000 pulse of 1.35 ms and a 90° excitation Eburp2.1000 pulse of 1.2 ms were applied at the center of

the anomeric carbon resonances (~100 ppm), and a 180° refocusing Reburp.1000 pulse of 1.00 ms was applied off-resonance at the center of the ring carbon resonances (~62 ppm). The spectrum was recorded over a spectral region of 70 ppm using acquisition times of 97 and 12 ms in *F2* and *F1*, respectively, and 40 scans per increment. The (H)CC-TOCSY spectrum ($\tau_m = 20$ ms) with simultaneous virtual decoupling of the $^1J_{C1,C2}$ and $^1J_{C2,C3}$ couplings in the direct dimension (Fig. 3d) was recorded at 40 °C at a magnetic field strength of 14.1 T using a 2D version of the DIPAP (Bermel et al. 2006) pulse sequence described by Richter et al. (Richter et al. 2010). For the virtual decoupling scheme the following 180° selective refocusing pulses were applied: a 2.2 ms Reburp.1000 pulse centered at the middle of the C2 carbon resonances (~51 ppm), a 2.2 ms double selective Reburp.1000 pulse centered at two positions (middle of the C1 resonances at ~100 ppm and middle of the C2 resonances at ~51 ppm) and a 1.0 ms Reburp.1000 pulse centered at the middle of the C2 and C3 resonances (~62 ppm). The spectrum was recorded over a spectral region of 70 ppm using acquisition times of 97 and 24 ms in *F2* and *F1*, respectively, and 40 scans per increment.

The HC–H2BC (Nyberg et al. 2005) spectrum for hydroxyl–^{1}H assignments (Fig. 5b) was recorded at 2 °C at a magnetic field strength of 16.4 T. It was acquired over a spectral region of 14 × 94 ppm using acquisition times of 209 and 4 ms in *F2* and *F1*, respectively, and 64 scans per increment.

The experiments for amide-^{1}H and *N*-acetyl groups assignments (HC–H2BC, HC(C)H–COSY, HN–SOFAST–HMQC, BEST–HNCA, BEST–HNCO and (H)CACO spectra) were recorded at 40 °C at a magnetic field strength of 16.4 T. The HC–H2BC (Nyberg et al. 2005) spectrum (Fig. 6b) was recorded over a spectral region of 14 × 110 ppm using acquisition times of 209 and 3 ms in *F2* and *F1*, respectively, and 64 scans per increment. The 2D HC(C)H–COSY spectrum (Fig. 6d) was acquired as a ^{1}H–^{13}C plane of the 3D HC(C)H–COSY experiment. The spectrum was recorded over a spectral region of 14 × 110 ppm, using acquisition times of 122 and 6 ms in *F3* and *F2*, respectively, and 8 scans per increment. The HN–SOFAST–HMQC (Schanda and Brutscher 2005) spectrum (Fig. 6e) was recorded over a spectral region of 6 × 40 ppm using acquisition times of 46 and 23 ms in *F2* and *F1*, respectively, and 8 scans per increment. A recycle delay of 0.1 s and the following ^{1}H shaped pulses were employed: a 120° excitation Pc9_4_120.1000 pulse of 2.57 ms (centered at 8.0 ppm) and a 180° refocusing Q3_surbob.1 pulse of 1.2 ms (centered at 8.0 ppm). The BEST–HNCA (Lescop et al. 2007; Schanda et al. 2006; Schanda and Brutscher 2005) spectrum (Fig. 6c) was recorded over a spectral region of 12 × 80 ppm, using

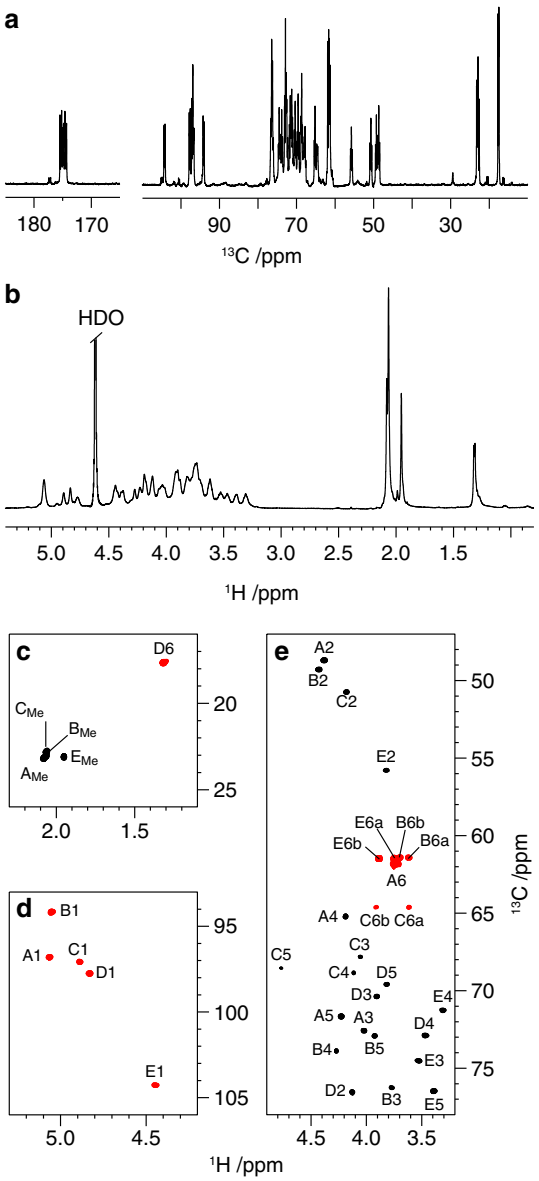

Fig. 4 ^{1}H chemical shift assignment in polysaccharides. **a** ^{13}C and **b** ^{13}C-decoupled ^{1}H NMR spectra of the ^{13}C-enriched O-antigen polysaccharide of *E. coli* O142 (1.5 mg in 0.5 mL of D_2O). **c**–**e** The HC–CT–HSQC spectrum ($2T = 22$ ms) of the same polysaccharide, showing the region for methyl groups (**c**) and anomeric resonances (**d**), as well as the region for the ring atoms and those from hydroxymethyl groups (**e**). The signs of the cross-peaks are opposite for carbons directly attached to an odd versus an even number of neighboring aliphatic carbons (shown in *red* and *black color*, respectively)

acquisition times of 122 and 3 ms in *F3* and *F1*, respectively, and 256 scans per increment. The BEST–HNCO (Lescop et al. 2007; Schanda et al. 2006; Schanda and Brutscher 2005) spectrum (Fig. 6f) was recorded over a spectral region of 12 × 6 ppm, using acquisition times of 122 and 19 ms in *F3* and *F1*, respectively, and 128 scans per increment. In both cases (BEST–HNCA and BEST–HCNO experiments) a recycle delay of 0.2 s and the

following shaped pulses were employed: selective 90° ^{1}H excitation Pc9_4_90.1000 pulse of 2.5 ms, 180° ^{1}H refocusing Reburp.1000 pulse of 1.714 ms, 90° ^{1}H excitation Eburp2.1000 pulse of 1.645 ms, 180° ^{1}H Bip720,50,20.1 pulse of 0.171 ms), 90° ^{13}C excitation Q5.1000 pulse of 274 μs and 180° ^{13}C refocusing Q3.1000 pulse of 219 μs; all ^{1}H selective pulses were centered at 8.0 ppm, the offset for the carbonyl carbons was 180/173 ppm (BEST–HNCA/BEST–HNCO experiments, respectively), and the offset of the $C_{aliphatic}$ was 62/18 ppm (BEST-HNCA/BEST–HNCO experiments, respectively). The relaxation-optimized (H)CACO (Bermel et al. 2009a, b) spectrum (Fig. 6g) was recorded over a spectral region of 20 × 1 ppm, using acquisition times of 145 and 57 ms in *F2* and *F1*, respectively, and 8 scans per increment. A recycle delay of 0.2 s and the following ^{13}C shaped pulses were employed: 90° excitation Q5.1000 pulse of 274 μs, 180° refocusing Q3.1000 pulse of 219 μs and a highly CH_3-selective 180° refocusing Q3.1000 pulse of 600 μs; the offsets for the CH_3 and CO carbon resonances were 22 and 174 ppm, respectively. The $1/(4J_{CA,CO})$ delay was set to 2.5 ms.

The ^{1}H-decoupled ^{13}C (Fig. 4a) and the HC–CT–HSQC (Fig. 4c–e) spectra were acquired at 40 °C in D_2O solution (1.5 mg in 0.5 mL) and a magnetic field strength of 16.4 T. The HC–CT–HSQC spectrum (Fig. 4c-e) was recorded over a spectral region of 6 × 110 ppm using acquisition times of 122 and 10 ms in *F2* and *F1*, respectively; (Vuister and Bax 1992) 4 scans per increment were used, and a 2*T* value of 22 ms. The offset for the selective pulse used for the refocusing of the carbonyl resonances was set at 174 ppm.

The ^{13}C-decoupled ^{1}H (Fig. 4b) and the HC–CT–HSQC–NOESY (Fig. 8a, b) spectra were acquired at 40 °C and at a magnetic field strength of 14.1 T, using a room temperature TXI (with the nuclei ^{1}H–^{13}C/^{31}P) probe. The 2D spectrum was recorded over a spectral region of 7 × 110 ppm using acquisition times of 122 and 15 ms in *F2* and *F1*, respectively, 16 scans per increment, a CT value of 44 ms and a mixing time of 100 ms. The same pulse sequence was employed as for [UL-^{13}C]-sucrose.

The long-range HC–CT–HSQC experiment was recorded at 70 °C at a magnetic field strength of 16.4 T. The spectrum (Fig. 8c, d) was recorded over a spectral region of 6 × 110 ppm using acquisition times of 122 and 7 ms in *F2* and *F1*, respectively, 32 scans per increment and a delay for evolution of the proton–carbon couplings optimized for $^{n}J_{CH} = 20$ Hz. The offset for the selective pulse used for the refocusing of the carbonyl resonances was set at 174 ppm.

Relaxation measurements on the [1-^{13}C]-enriched O-antigen polysaccharide of *E. coli* O91

^{13}C T_2 and T_1 relaxation times were measured at 59 °C and at a magnetic field strength of 14.1 and 16.4 T, using 5 mm room temperature BBO (X–^{1}H) probes. In the case of T_2 relaxation, the CPMG (Car-Purcel-Meiboom-Gill) pulse sequence for heteronuclei was employed, with ^{1}H-decoupling during acquisition and ^{1}H-refocussing pulses (32 and 34 μs at 14.1 and 16.4 T, respectively) placed at even echos. The ^{13}C-refocussing pulses were 48 and 39.6 μs at a magnetic field strength of 14.1 and 16.4 T, respectively. Twelve different relaxation delays (between 4.4 and 132 ms at 14.1 T; and between 4.6 and 139 ms at 16.4 T) were used; the CPMG delay was set to 0.25 ms and the recovery delay was 4 s. T_1 relaxation measurements were carried out using the inversion recovery pulse sequence with ^{1}H-decoupling applied during the T_1 relaxation delay. Ten different relaxation delay times (10–2,500 ms) were used, and the recovery delay was 4 s.

Results

The two-dimensional (2D) NMR experiments discussed below are summarized in Table 1. The experiments useful

Table 1 Summary of the 2D NMR experiments used in the structural elucidation of ^{13}C-enriched carbohydrates

Oligosaccharides	Polysaccharides
^{13}C chemical shifts assignments	
CC–CT–COSY	(H)CC–CT–COSY
CC–CT–TOCSY	(H)CC–TOCSY
	(H)CC–TOCSY–IPAP
	(H)CC–TOCSY–DIPAP
	(H)CC–NOESY
^{1}H chemical shifts assignments	
HC–CT–HSQC	HC–CT–HSQC
HC(C)H–TOCSY	HC(C)H–TOCSY
	HC(C)H–COSY
	HC–HSQC–TOCSY
HC–CT–H2BC (for hydroxyl protons)	HC–CT–H2BC (for hydroxyl protons)
Sequence determination	
Long-range HC–CT–HSQC	long-range HC–CT–HSQC
HC–CT–HSQC–NOESY	HC–CT–HSQC–NOESY
Band-selective (H)CC–TOCSY	
CC—CT–COSY	
***N*-acetyl groups assignments**	
	HC–CT–H2BC (for amide protons)
	HN–SOFAST–HMQC
	HNCO
	HNCA
	(H)CACO

for oligosaccharides are reported in the left column whereas the ones suitable for polysaccharide investigations are given in the right column. Further, in Table 1 the 2D experiments are grouped according to their purpose of either ^{13}C or ^{1}H signal assignment, sequential assignment of the monosaccharide units and determination of the glycosidic linkage positions, or assignment of *N*-acetyl groups and their locations.

^{13}C chemical shift assignments

Three experiments were evaluated for the assignment of ^{13}C resonances in uniformly labeled ^{13}C carbohydrates. In this approach the distinctive ^{13}C anomeric signals are identified in the 1D ^{1}H-decoupled ^{13}C spectra and used as starting point for the assignments in the respective spin systems. For instance, the resonances of the carbons directly attached to these atoms can be readily identified using CC–CT–COSY experiments. In order to remove the one-bond ^{13}C–^{13}C splitting in the indirect dimension, a CT version of this experiment was employed (Bermel et al. 2003; Machonkin et al. 2002; Rance et al. 1984). Considering that the average $^{1}J_{CC}$ coupling in carbohydrates is usually $\sim$45 Hz, the optimal CT value ($2T$) to be used in this case is $\sim$11.1 ms [corresponding to $1/(2 \times {}^{1}J_{CC})$]. In this kind of experiments, the CT length restricts the maximum number of increments that can be used and, consequently, this limits the maximum possible resolution that can be achieved in the indirect dimension. Therefore, the $2T$ value can be increased in cases where improved resolution is required in the indirect dimension; however, one should consider that in such a case long-range correlations may also appear in the spectrum. Furthermore, the 2D CC–TOCSY and 3D HC(C)H–TOCSY experiment have previously proved to be useful in the assignments of ^{13}C resonance signals of polysaccharides (Kjellberg et al. 1998; Linnerborg et al. 1999). Unlike the HH–TOCSY experiment, the ^{13}C–^{13}C correlations observed in the CC–TOCSY spectrum are not sensitive to the configuration of the sugar residue (i.e., they are not dependent on whether the sugar residue is for example glucose, galactose or mannose), and the large magnitude of the $^{1}J_{CC}$ couplings facilitates rapid coherence transfer from the anomeric carbon to the most distant carbons of the monosaccharide backbone. In the case of well resolved anomeric resonances, the CT version of this experiment (Eletsky et al. 2003) is the most efficient way to trace all the ^{13}C resonances in each monosaccharide starting from the respective anomeric signals.

The ^{1}H-decoupled ^{13}C spectrum of [UL-^{13}C]-sucrose is shown in Fig. 2a. The two distinctive anomeric signals found at 93.0 and 104.5 ppm were used as a starting point for the assignments in the respective spin systems, and a CC–CT–COSY experiment was used to reveal correlations from the anomeric carbons to the directly attached carbons (Fig. 2b). Furthermore, in the CC–CT–TOCSY spectrum recorded with the shortest mixing time two correlations were observed from the anomeric carbon of the fructofuranosyl residue (denoted **F2** in Fig. 2c) to the C1 and C3 carbons in the same residue (denoted **F1** and **F3**). Subsequent correlations to carbon C4, C5 and C6 (denoted **F4**, **F5** and **F6** in Fig. 2d and 2e) were observed in the spectra recorded with increasing mixing times (9.4 and 14.1 ms). The resonances in the glucopyranosyl residue (**G**) were identified using the same strategy, and a spectrum recorded with a mixing time of 18.8 ms was required to observe the correlation from the C1–C6 carbon in the same residue (denoted **G1**–**G6** in Fig. 2f, respectively). If the CT delay is set to $1/{}^{1}J_{CC}$ (which is $\sim$22.2 ms) the cross-peaks from ^{13}C spins attached to an odd and even number of neighboring aliphatic carbons will appear with different signs in the NMR spectrum (in red and black in Fig. 2c–f, respectively). As a consequence, the resonances of the carbon atoms located at the terminal ends of the monosaccharide backbone (C1 and C6 in the case of hexoses) will appear with a different sign than the resonances of the internal carbons. In the spectrum shown in Fig. 2c, this editing capacity is used as an argument to assign the red colored resonance to C1 and the black colored resonance to C3, without further analysis of other regions of the NMR spectrum. Furthermore, the appearance of a second 'red colored' cross-peak in the spectra of Fig. 2e, f indicates that the complete spin system has been revealed for the respective monosaccharides. This editing capacity is similar to that of the CT–HSQC experiment and will be discussed in more detail below.

The same approach as described above for oligosaccharides was used for the assignment of the ^{13}C and ^{1}H resonances of the ^{13}C-enriched O-antigen polysaccharide of *E. coli* O142 (Fig. 1 *top left*). Initially, the (H)CC-CT-COSY experiment was used to reveal correlations from the anomeric carbons to the directly attached C2 carbons (Fig. 3a). For large polysaccharides the (H)CC–NOESY experiment (Fig. 3b) (Bertini et al. 2003, 2004) is also an alternative to the (H)CC–CT–COSY experiment. An array of (H)CC–TOCSY experiments with different mixing times can be employed to trace all the carbon resonances in the different spins-systems, starting from the anomeric carbon signals. In the case of poorly dispersed ^{13}C anomeric resonances (such as the residues **A**, **C** and **D** in the O-antigen polysaccharide of *E. coli* O142), the virtually decoupled version of this experiment allows improved resolution through the removal of the large homonuclear ^{13}C–^{13}C splitting in the direct dimension and collapsing the ^{13}C-multiplets into a single line. Thus, the assignments of the ^{13}C resonances were achieved using (H)CC–TOCSY correlations from the anomeric carbons of each

monosaccharide residue, and the IPAP scheme (Bermel et al. 2006) for virtual decoupling of $^{1}J_{C1,C2}$ couplings was used to improve the resolution of the anomeric carbon resonances in the direct dimension (Fig. 3c). Furthermore, a (H)CC–TOCSY experiment recorded using the DIPAP scheme (Bermel et al. 2006) (Fig. 3d) was useful to observe correlations from the nitrogen bearing carbons (C2 carbons in residues **A**, **B**, **C** and **E**) with virtual decoupling of both $^{1}J_{C1,C2}$ and $^{1}J_{C2,C3}$ couplings in the direct dimension. The ^{1}H-decoupled ^{13}C NMR spectrum of the ^{13}C-enriched O-antigen polysaccharide of *E. coli* O142 is shown in Fig. 4a.

^{1}H chemical shift assignments and exchangeable protons

The ^{13}C-decoupled ^{1}H NMR spectra of [UL-^{13}C]-sucrose and the ^{13}C-enriched O-antigen polysaccharide of *E. coli* O142 are shown in Fig. 2g and 4b, respectively. Once the ^{13}C chemical shifts have been assigned, an HC–CT–HSQC spectrum can be used to correlate the carbon resonances to their respective protons. Originally developed for proteins, the CT version of the HSQC experiment (Santoro and King 1992; Vuister and Bax 1992) allows the removal of the homonuclear ^{13}C–^{13}C splittings in the indirect dimension, improving the resolution in the crowded areas of carbohydrates spectra. The editing capability of this experiment depends on the CT period and differs from that of the regular multiplicity-edited HSQC. In this experiment, the sign and relative intensity of the cross-peaks are directly proportional to $\cos^{n}[2\pi(^{1}J_{CC})T]$, where $2T$ is the duration of the CT period and n is the number of directly attached carbons (Vuister and Bax 1992). Consequently, in a spectrum recorded with a CT period of 22 ms, the maximum intensity will be observed for carbons with $^{1}J_{CC}$ ~45 Hz, which is the average value usually observed in carbohydrates. In addition, the sign of the respective cross-peaks will depend on the number of neighboring aliphatic carbons (n); thus, in the case of ^{13}C spins with $^{1}J_{CC}$ couplings in the range between ~23 and 68 Hz, the sign of the magnetization will be negative for atoms located at the end of the monosaccharide backbone ($n = 1$, red solid line in Fig. 2h) and positive for the remaining carbons ($n = 2$, black solid line in Fig. 2h). If the maximum resolution that can be achieved with a shorter CT value is not enough to resolve the resonances in the NMR spectrum, the $2T$ value can be doubled to 44 ms. In such a case (dashed lines in Fig. 2h), the resonances from carbon atoms with $^{1}J_{CC}$ couplings ~45 Hz will also display the maximum possible intensity, but the aforementioned editing capacity will be lost. Furthermore, the intensity of the cross-peaks from atoms with $^{1}J_{CC}$ couplings ~34 and 57 Hz will be considerably reduced or not observed at all.

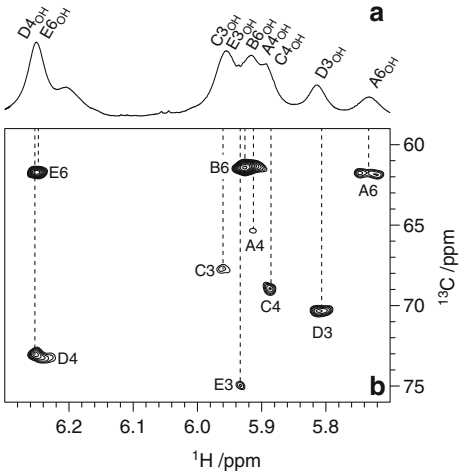

Fig. 5 ^{1}H chemical shift assignment of OH-groups. **a** The region for the hydroxyl protons of the 1D ^{1}H spectrum of the natural abundance O-antigen polysaccharide of *E. coli* O142. **b** Selected region of the HC–H2BC spectrum of the ^{13}C-enriched O-antigen polysaccharide of *E. coli* O142 showing correlations from the hydroxyl protons. The experiments were run at a magnetic field strength of 16.4 T and a temperature of 2 °C

The HC–CT–HSQC spectra of [UL-^{13}C]-sucrose and the ^{13}C-enriched O-antigen polysaccharide of *E. coli* O142, recorded with a $2T$ delay of 22 ms, are shown in Fig. 2i, j and Fig. 4c–e, respectively. The assignment of the ^{1}H resonances of each monosaccharide residue were carried out in a straightforward manner using the ^{13}C assignments from the CC–CT–TOCSY, or virtually decoupled CC–TOCSY spectra, and correlating them to their respective protons resonances in the HC–CT–HSQC spectrum.

In the case of exchanging protons not directly attached to carbon atoms (e.g. hydroxyl (Battistel et al. 2013) or amide (Norris et al. 2012) protons) the assignments were carried out in a H_2O/D_2O mixture using ^{1}H detected experiments such as HC-H2BC (Nyberg et al. 2005), BEST–HNCA (Lescop et al. 2007), HC(C)H–COSY, HC(C)H–TOCSY (Kay et al. 1993) and/or HC–HSQC–TOCSY (Kövér et al. 1997). In the case of [UL-^{13}C]–sucrose both HC–H2BC (Nyberg et al. 2005) and HC(C)H–TOCSY (Kay et al. 1993) experiments proved to be useful for the assignment of hydroxyl protons (Fig. S1 in Supplementary Material). In the case of the O-antigen polysaccharide of *E. coli* O142, the HC–H2BC spectrum proved to be useful not only for the assignment of the proton resonances in hydroxyl groups at low temperatures (2 °C), where exchange is sufficiently slowed down (Fig. 5), but also for the assignment of the amide protons of *N*-acetylated aminosugars (such as those in the residues **A**, **B**, **C** and **E** in the O-antigen polysaccharide of *E. coli* O142, cf. Figs. 1 *top left* and 6 a, b). Alternatively, the

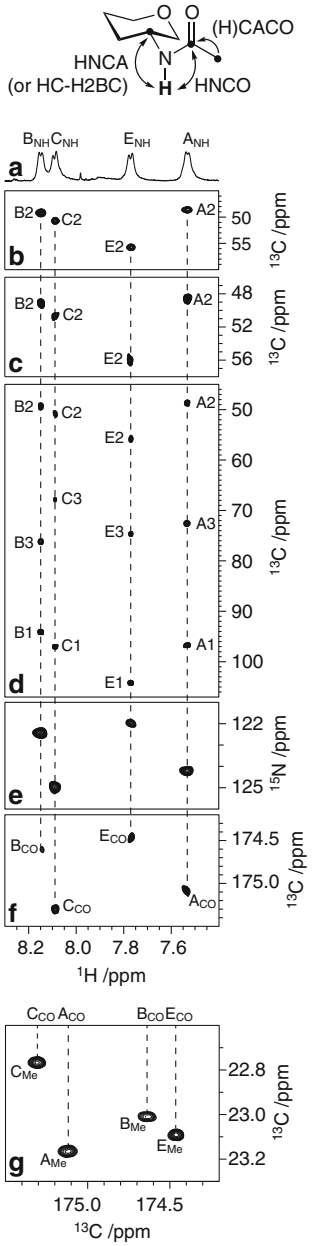

Fig. 6 Assignment of *N*-acetyl groups. **a** Selected region of the ^{1}H NMR spectrum of the ^{13}C-enriched O-antigen polysaccharide of *E. coli* O142 showing the amide protons of residues **A**, **B**, **C** and **E**. Selected regions of the: **b** HC–H2BC, **c** ^{1}H–^{13}C plane of the 3D BEST–HNCA, **d** ^{1}H–^{13}C plane of the 3D HC(C)H–COSY, **e)** HN–SOFAST–HMQC and **f)** ^{1}H–^{13}C plane of the 3D BEST–HNCO spectra of the ^{13}C-enriched O-antigen polysaccharide from *E. coli* O142 (^{15}N at natural abundance) showing correlations to the amide protons. **g** Selected region of the (H)CACO spectrum showing correlations from methyl to carbonyl carbons in the respective *N*-acetyl groups. The experiments were recorded in a H_2O/D_2O 95:5 solution (2 mg of the polysaccharide in 0.5 ml of the solvent) at a magnetic field strength of 16.4 T and at 40 °C. The correlations used for the assignment of resonances from the *N*-acetyl moieties, are highlighted in the schematic representation located on the *top* of the spectra

same correlations from amide protons to C2 carbons observed in the HC–H2BC spectrum of Fig. 6b can also be observed in the BEST–HNCA spectrum of Fig. 6c. Another set of three experiments that make up an alternative approach to the assignment of the amide protons resonances are: HC(C)H–COSY (Fig. 6d), HC(C)H–TOCSY and HC–HSQC–TOCSY (Fig. S2a-c in Supplementary Material, respectively). The latter experiments allow the assignment of the ^{1}H-amide resonances through additional correlations to more distant carbons in the respective spin systems. Once the amide protons have been assigned, they can be connected to their respective ^{15}N resonances using a HN-SOFAST–HMQC experiment (Fig. 6e) and to the carbonyl carbons of the respective *N*-acetyl moieties using a BEST–HNCO spectrum (Fig. 6f). The carbonyl carbons can subsequently be linked to their respective methyl carbons using a (H)CACO spectrum (Fig. 6g).

Sequence determination

Two ^{13}C-detected experiments were evaluated for elucidation of inter-residue correlations in [UL-^{13}C]-sucrose. First, the CC–CT–COSY experiment described above was optimized for detection of long-range carbon–carbon correlations using 2*T* values between 20 and 50 ms. In the spectrum recorded with a 2*T* value of 22 ms (Fig. 7a), five different correlations were observed from the anomeric carbon of residue **F** (C2): two corresponding to one-bond correlations (C1 and C3), one corresponding to an intra-residue two-bond correlation (C4) and, finally, two inter-residue correlations to the C1 and C2 carbons in residue **G** (two- and three-bonds correlations, respectively, highlighted by a green and an orange oval in Fig. 7a, respectively). Most importantly, a ^{13}C–^{13}C inter-residue correlation was observed in the band-selective (H)CC–TOCSY experiment between the anomeric carbon (C2) of residue **F** and the anomeric carbon (C1) of residue **G** (highlighted by a green oval in Fig. 7b). It should be noted that the $^{2}J_{CC}$ coupling constant of the carbon atoms at the glycosidic linkage is only 2.4 Hz (Duker and Serianni 1993), thus requiring the long mixing time of 144 ms in the experiment. The main limitation in the use of ^{13}C-detected experiments such as (H)CC–CT–COSY and band-selective (H)CC–TOCSY, for detection of long-range correlations in large polysaccharides, lies in their intrinsic lower sensitivity. The intensity of the cross-peaks from large molecules can be considerably diminished due to magnetization losses during the long delays required for long-range coupling evolution, often rendering these correlations undetectable.

Long-range through-bond proton-carbon correlations can also be used to determine the sequence of monosaccharide

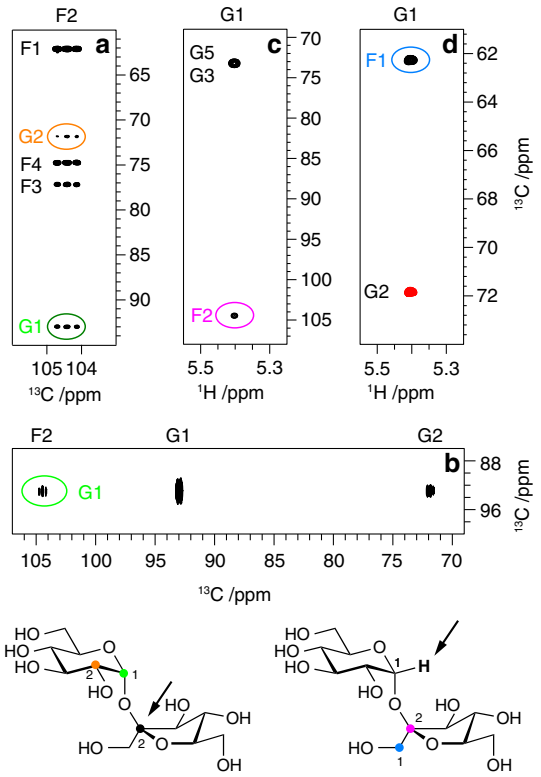

Fig. 7 Determination of monosaccharide sequence in oligosaccharides. Selected regions of the **a** CC–CT–COSY ($2T = 22$ ms), **b** band-selective CC–TOCSY ($\tau_m = 144$ ms), **c** long-range HC–CT–HSQC ($2T = 22$ ms and optimized for $^{n}J_{CH} = 12$ Hz) and **d** HC–CT–HSQC–NOESY ($2T = 22$ ms, $\tau_m = 500$ ms) spectra of [UL-^{13}C]-sucrose showing correlations from the anomeric atoms. The inter-residue correlations observed in the different spectra are indicated with *colored ovals* and also highlighted in the schematic representation of the sucrose structure using the *same color*. The correlations are observed from the atom indicated with an arrow to those at the colored positions. In the HC–CT–HSQC–NOESY spectrum the sign of the ^{13}C magnetization is opposite for carbons directly attached to an odd versus an even number of neighboring aliphatic carbons (shown in *black* and *red* color, respectively), and opposite to the respective autopeaks

residues in oligo- and polysaccharides using ^{1}H-detected experiments. The HC–CT–HSQC experiment described above was optimized for detection of heteronuclear long-range correlations (long-range HC–CT–HSQC) and allowed identification of an inter-residue correlation in [UL-^{13}C]-sucrose from H1 in residue **G** to C2 in residue **F** (highlighted by a magenta oval in Fig. 7c). A CT version of the HC-HSQC-NOESY experiment was also implemented, and allowed identification of a through-space correlation in [UL-^{13}C]-sucrose from the anomeric proton (H1) of residue **G** and the proton(s) directly attached to the C1 carbon of residue **F** (highlighted by a blue oval in Fig. 7d). The standard version of this experiment was recently used in a study of ^{13}C,^{15}N-labeled sialic acid oligomers (Battistel et al. 2012). All the inter-residue correlations aforementioned are illustrated in the schematic chemical representations of sucrose located on the bottom of Fig. 7, using the same color-coding as in the spectra.

Since ^{1}H-detected experiments offer better sensitivity than the ^{13}C-detected experiments, the long-range HC–CT–HSQC and HC–CT–HSQC–NOESY experiments were successfully employed to determine the monosaccharide sequence and linkage positions of the repeating unit of a large polysaccharide (the O-antigen polysaccharide of *E. coli* O142). Particularly, the HC–CT–HSQC–NOESY experiment is more sensitive for larger polysaccharides than small oligosaccharides, allowing the acquisition of high signal-to-noise spectra in a relatively short time (for example, an acceptable signal-to-noise spectrum of 1.5 mg of the O-antigen polysaccharide of *E. coli* O142 was obtained in ~30 min at a magnetic field strength of 14.1 T using a room-temperature inverse-detection probe). The inter-residue correlations observed from the anomeric carbons in the ^{13}C-enriched O-antigen polysaccharide of *E. coli* O142 are shown in Fig. 8a, b, and they are highlighted by colored ovals. These correlations are also illustrated in the schematic chemical representation of the polysaccharide (located to the right of the spectrum) using the same color-coding as in the spectrum. On the other hand, due to the long delay required for long-range proton-coupling evolution, the long-range HC–CT–HSQC experiment showed lower sensitivity. This problem could be overcome to some extent by recording the experiment at a higher temperature and using a reduced long-range evolution delay (optimized for $^{n}J_{CH}$ ~20 Hz instead of a smaller magnitude that is theoretically required). The long-range ^{1}H,^{13}C correlations observed in the long-range HC–HSQC spectrum of the O-antigen polysaccharide of *E. coli* O142 (recorded at 70 °C and a magnetic field strength of 16.4 T) are shown in Fig. 8c, d, where they are highlighted by colored ovals. These correlations are also represented by the same color-coding in the schematic chemical representation of the polysaccharide shown to the left of the spectrum.

Contrary to proteins, polysaccharides are not globular but extended or random structures in which different segments may display different flexibility (Martin-Pastor and Bush 2000; Soltesova et al. 2013). To evaluate the effect of the temperature and the spectrometer magnetic field strength on the relaxation parameters of this kind of structures, a polysaccharide model of approximately the same molecular weight as the O-antigen polysaccharide of *E. coli* O142, but with a simplified labeling pattern, was selected: the [1-^{13}C]-labeled O-antigen polysaccharide of *E. coli* O91 (Fig. 1 *bottom*) (Lycknert and Widmalm 2004). The T_2 and T_1 relaxation times were measured for the anomeric carbons and the data are compiled in Table 2. The results revealed that increasing the temperature by 22° (from 37 to 59 °C) significantly lengthened the T_2

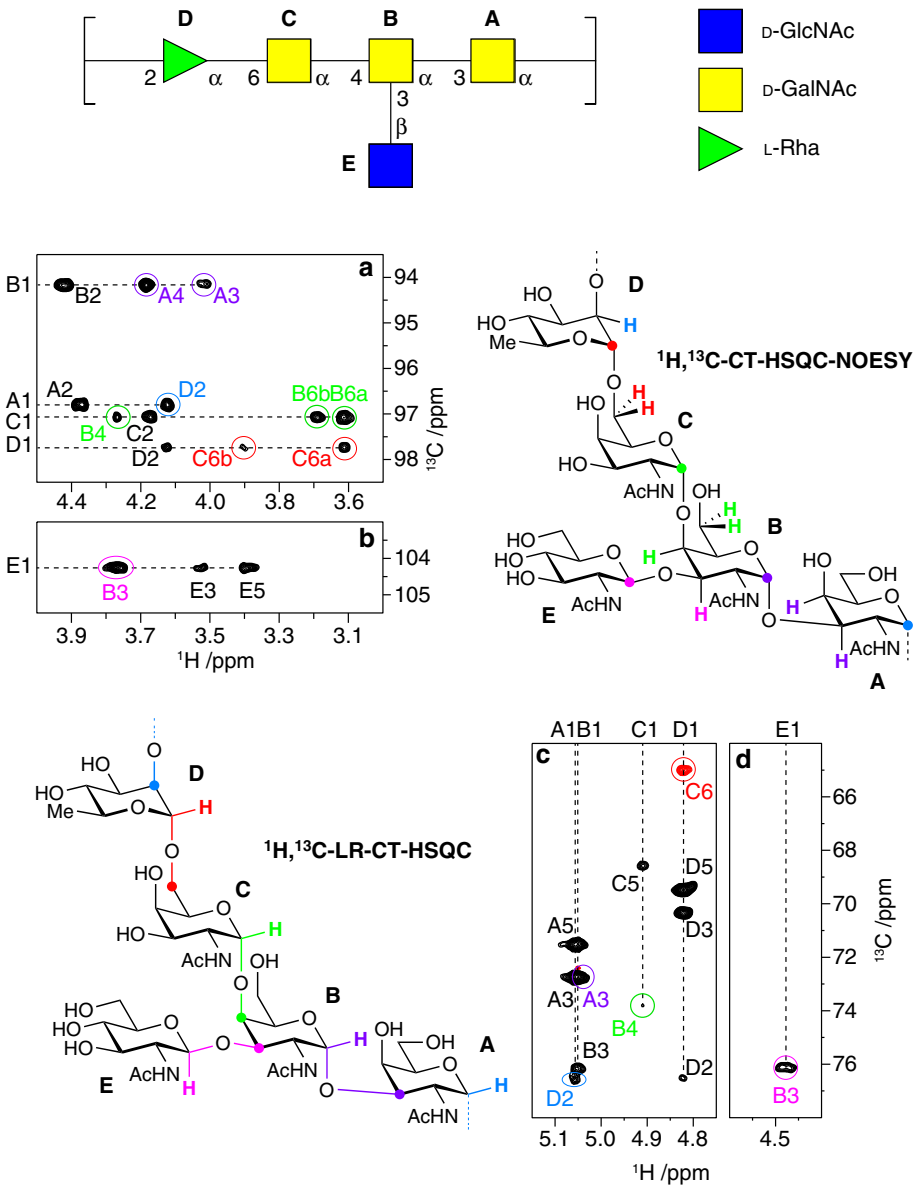

Fig. 8 Sequence determination in polysaccharides. The structure of the repeating unit of the O-antigen polysaccharide of *E. coli* O142 is shown in CFG (consortium for functional glycomics) notation on the top of the figure. Selected regions of the HC–CT–HSQC–NOESY ($2T = 44$ ms, $\tau_m = 100$ ms) (**a, b**) and long-range HC–CT–HSQC ($2T = 22$ ms and optimized for $^{n}J_{CH} = 20$ Hz) (**c, d**) spectra of the ^{13}C-enriched O-antigen polysaccharide of *E. coli* O142, showing intra- and inter-residue correlations from anomeric atoms (the latter are highlighted by *colored ovals*). The *same color* nomenclature is used to illustrate the respective inter-residue correlations in the structures located on the *right* or *left side* of the respective spectra

Table 2 Comparison of the ^{13}C relaxation data of the anomeric resonances of the [1-^{13}C]-labeled O-antigen polysaccharide of *E. coli* O91 at two different magnetic field strengths and temperatures

Residue	14.1 T				16.4 T	
	37 °C		59 °C		59 °C	
	T_2 (ms)[a]	T_1 (ms)[a]	T_2 (ms)	T_1 (ms)	T_2 (ms)	T_1 (ms)
A	68	484	90	462	101	527
B	72	503	117	479	126	545
C	48	467	77	433	76	515
D	78	507	118	489	130	541
E	62	461	89	414	103	483

[a] Data from literature (Lycknert and Widmalm 2004)

relaxation times (by as much as 32–63 %) whereas changing the magnetic field strength from 14.1 to 16.4 T lead only to minor improvements (16 % in the best of the cases). These results are in agreement with the observation that in the long-range HC–HSQC spectra the intensities of the inter-residue correlations were considerably reduced if longer delays than the ones described above were employed, or if the experiments were recorded at 40 °C instead of 70 °C.

Disscusion and conclusions

In this study a selection of experiments were evaluated for the structural analysis of ^{13}C-enriched carbohydrates. The

strategy consists of three steps: (1) assignment of the ^{13}C resonances within each monosaccharide spin system, (2) assignment of the ^{1}H resonances and (3) determination of the monosaccharide sequence and linkage positions.

In the first step, the ^{13}C chemical shift assignments are obtained (in both oligo- and polysaccharides) by employing ^{13}C-detected experiments such as (H)CC–COSY, (H)CC–NOESY and (H)CC–TOCSY, while using the anomeric resonances of each monosaccharide as the starting point for the assignments. In the case of well resolved ^{13}C anomeric resonances the use of CC–CT–TOCSY experiments with different mixing times (e.g. ~5, 10, 15 and 20 ms) is the method of choice, allowing the assignment of the resonances within each monosaccharide spin system in a straightforward manner. However, the use of (H)CC–COSY and (H)CC–NOESY experiments may also play an important role in the discrimination between one-bond and long-range correlations in cases where this distinction cannot be achieved using CC–CT–TOCSY experiments. For example, in the case of [UL-^{13}C]-sucrose two new correlations were observed in the spectrum recorded with a mixing time of 9.4 ms (Fig. 2d) that were not present in the spectrum recorded with a mixing time of 4.7 ms (Fig. 2c). In cases of significant spectral overlap in the carbon anomeric region, the use of virtually decoupled (H)CC–TOCSY experiments may help to alleviate the overlap of signals in that region through the removal of the $^{1}J_{CC}$ splitting in the direct dimension. Due to its selective nature and more demanding setup, the latter experiment is only recommended if the regular CC–CT–TOCSY experiment fails to provide sufficient resolution for the assignment of the resonances.

Subsequently, an HC–CT–HSQC spectrum allows the assignment of the proton resonances of ^{1}H nuclei directly attached to ^{13}C atoms, based on the assignments obtained in the previous step. Furthermore, if a CT value of 22 ms is used, the editing capacity of both CC–CT–TOCSY and HC–CT–HSQC experiments can be used to identify the start and end points of the spin systems, since the cross-peaks of the carbons located at the terminal ends of the monosaccharide backbone (C1 and C6 in the case of hexoses) appear with opposite signs with respect to the cross-peaks originating from the intermediate carbons (C2-C5) in the sequence of atoms. An additional HC–CT–HSQC experiment with a CT value of 44 ms may also be recorded if higher resolution is required in certain regions of the spectrum.

In the case of exchangeable protons (Battistel et al. 2014) (hydroxyl or amide protons), the use of HC–H2BC proved highly informative. This experiment is remarkably more sensitive and easier to interpret than the HC(C)H–COSY experiment, that can also be used to achieve the same purpose in the case of strong overlap in the HC–H2BC spectrum. For instance, if the C2 resonances are not well resolved, unambiguous assignment of the amide protons of 2-amino sugars may not be possible using an HC–H2BC spectrum, but the HC(C)H–COSY experiment will allow correlation of the respective amide protons to other carbons within the same spin system. Furthermore, the HC–H2BC experiment only works if the nitrogen bearing carbon (Cx) is protonated as it relies on homonuclear COSY-type transfer between the amide proton and the Cx-proton, followed by an HMQC-transfer to the Cx-carbon. In the HC(C)H–COSY experiment, however, the amide proton connectivities appear because of a magnetization transfer over long-range couplings (Hu et al. 2010) between the amide protons and the Cx (two-bond correlation), C(x − 1) and C(x + 1) (three-bond correlation) carbons. This being the case, the experiment can be made more favorable for the amide proton connectivities by extending the HC-transfer delay. As a comparison, the two-bond and three-bond carbon-proton couplings of the hydroxyl protons in sucrose are smaller, and known to be ~1.5–3.9 Hz (Batta and Kövér 1999). Alternatively, HC(C)H–TOCSY and HC–HSQC–TOCSY experiments can also be used for assignment of exchangeable protons (Fig. S2b and S2c in Supplementary Material, respectively). Both HC–H2BC and HC–HSQC–TOCSY experiments were originally intended for unlabeled samples but, in fact, they provide higher sensitivity on [UL-^{13}C]-materials than the HC(C)H–COSY and HC(C)H–TOCSY experiments (cf. Fig. S2d-f in Supplementary Material) that are dedicated to ^{13}C-labeled samples.

The exchange rate of the amide protons in the *N*-acetyl groups renders them observable even at 40 °C although, contrary to proteins, the NH groups in polysaccharides are only rarely hydrogen bonded in a stable structure. For maximal sensitivity in the *N*-acetyl assignment we chose the so called SOFAST–HMQC and the relaxation optimized ^{1}H-detected versions of the 3D HNCA and HNCO experiments published by Brutscher and co-workers (Lescop et al. 2007; Schanda et al. 2006; Schanda and Brutscher 2005). These experiments use band-selective amide proton pulses and water flip-back pulses for minimal perturbation of aliphatic and water ^{1}H spins. Since water is neither excited nor dephased, even signals from amide protons in fast exchange with water are retained. The large amount of aliphatic and water ^{1}H spin polarization along the Z-axis by the end of these pulse sequences then enhances longitudinal (spin–lattice) relaxation of amide hydrogen spins via dipole–dipole interactions (NOE effects) and hydrogen exchange (Lescop et al. 2007; Schanda et al. 2006; Schanda and Brutscher 2005). For the ^{13}C-detection we used the ^{1}H-start relaxation-optimized experiment called (H)CACO (Bermel et al. 2009a) that also allows rapid pulsing through realigning ^{1}H magnetization

along Z-axis, developed particularly for studies of inherently disordered proteins. The four 2D spectra for the assignment of *N*-acetyl groups are thus HN–SOFAST–HMQC (Fig. 6e), the ^{1}H–^{13}C-planes of 3D BEST–HNCA (Fig. 6c) and 3D BEST–HNCO (Fig. 6f) and the ^{13}C-detected (H)CACO (Fig. 6g). It is worth noting that all four experiments showed high sensitivity although, in contrast to previously reported triple resonance spectroscopy on glycans (Norris et al. 2012; Wang et al. 2008), the O-antigen polysaccharide sample used in our study was only ^{13}C-labeled, not ^{15}N-labeled.

For the sequence determination in oligosaccharides four experiments were found useful: (H)CC–CT–COSY, band-selective (H)CC–TOCSY, HC–CT–HSQC–NOESY and long-range HC–CT–HSQC; but only the two latter ones showed enough sensitivity to be used for large polysaccharides. In the case of large molecules, the main limitation is the loss of the magnetization during the long delays required for long-range couplings to evolve, due to fast T_2 relaxation. In the case of the long-range HC–CT–HSQC experiment, this problem could be overcome to certain extent by increasing the temperature and shortening the long-range evolution delay. On the other hand, the band-selective (H)CC–TOCSY experiment proved valuable in the case of [UL-^{13}C]-sucrose, but its use is limited to cases in which the resonances of the carbons at the linkage positions are not overlapping with other resonances (in other words, the adiabatic mixing pulse should be selective to the carbon resonances at the linkage positions).

In conclusion, by means of the experiments discussed above, we have been able to unambiguously assign not only the ^{1}H and ^{13}C resonances of a small ^{13}C-labeled carbohydrate (sucrose) and a ^{13}C-enriched O-antigen polysaccharide (from *E. coli* O142), including its ^{15}N resonances, but also to determine the monosaccharide sequence and linkage positions. Uniform ^{13}C-labeling of the carbohydrate samples allowed us to extend the selection of NMR experiments to constitute a comprehensive toolbox for future studies of this type of biomolecules.

Acknowledgments This work was supported by grants from the Swedish Research Council and the Knut and Alice Wallenberg Foundation. The research that has led to these results has received funding from the European Commission's Seventh Framework Programme FP7/2007–2013 under grant agreement no. 215536.

References

Aich U, Yarema KJ (2009) Glycobiology and immunology. carbohydrate-based vaccines and immunotherapies. Wiley, Hoboken, pp 1–53

Batta G, Kövér KE (1999) Heteronuclear coupling constants of hydroxyl protons in a water solution of oligosaccharides: trehalose and sucrose. Carbohydr Res 320:267–272. doi:10.1016/S0008-6215(99)00183-4

Battistel MD, Shangold M, Trinh L, Shiloach J, Freedberg DI (2012) Evidence for helical structure in a tetramer of α2-8 sialic acid: unveiling a structural antigen. J Am Chem Soc 134:10717–10720. doi:10.1021/ja300624j

Battistel MD, Pendrill R, Widmalm G, Freedberg DI (2013) Direct evidence for hydrogen bonding in glycans: a combined NMR and molecular dynamics study. J Phys Chem B 117:4860–4869. doi:10.1021/jp400402b

Battistel MD, Azurmendi HF, Yu B, Freedberg DI (2014) NMR of glycans: shedding new light on old problems. Prog Nucl Magn Reson Spectrosc. doi:10.1016/j.pnmrs.2014.01.001

Bermel W, Bertini I, Felli IC, Kümmerle R, Pierattelli R (2003) ^{13}C direct detection experiments on the paramagnetic oxidized monomeric copper, zinc superoxide dismutase. J Am Chem Soc 125:16423–16429. doi:10.1021/ja037676p

Bermel W, Bertini I, Felli IC, Piccioli M, Pierattelli R (2006) ^{13}C-detected protonless NMR spectroscopy of proteins in solution. Prog Nucl Magn Reson Spectrosc 48:25–45. doi:10.1016/j.pnmrs.2005.09.002

Bermel W, Felli IC, Kümmerle R, Pierattelli R (2008) ^{13}C direct-detection biomolecular NMR. Concepts Magn Reson Part A 32A:183–200. doi:10.1002/cmr.a.20109

Bermel W, Bertini I, Csizmok V, Felli IC, Pierattelli R, Tompa P (2009a) H-start for exclusively heteronuclear NMR spectroscopy: the case of intrinsically disordered proteins. J Magn Reson 198:275–281. doi:10.1016/j.jmr.2009.02.012

Bermel W, Bertini I, Felli IC, Pierattelli R (2009b) Speeding up ^{13}C direct detection biomolecular NMR spectroscopy. J Am Chem Soc 131:15339–15345. doi:10.1021/ja9058525

Bertini I, Felli IC, Kümmerle R, Moskau D, Pierattelli R (2003) ^{13}C–^{13}C NOESY: an attractive alternative for studying large macromolecules. J Am Chem Soc 126:464–465. doi:10.1021/ja0357036

Bertini I, Felli I, Kümmerle R, Luchinat C, Pierattelli R (2004) ^{13}C-^{13}C NOESY: a constructive use of ^{13}C–^{13}C spin-diffusion. J Biomol NMR 30:245–251. doi:10.1007/s10858-005-1679-2

Bugarel M, Martin A, Fach P, Beutin L (2011) Virulence gene profiling of enterohemorrhagic (EHEC) and enteropathogenic (EPEC) *Escherichia coli* strains: a basis for molecular risk assessment of typical and atypical EPEC strains. BMC Microbiol 11:142. doi:10.1186/1471-2180-11-142

DebRoy C, Roberts E, Fratamico PM (2011) Detection of O antigens in *Escherichia coli*. Anim Heal Res Rev 12:169–185. doi:10.1017/S1466252311000193

Duker JM, Serianni AS (1993) (^{13}C)-substituted sucrose: ^{13}C–^{1}H and ^{13}C–^{13}C spin coupling constants to assess furanose ring and glycosidic bond conformations in aqueous solution. Carbohydr Res 249:281–303. doi:10.1016/0008-6215(93)84096-O

Eletsky A, Moreira O, Kovacs H, Pervushin K (2003) A novel strategy for the assignment of side-chain resonances in completely deuterated large proteins using ^{13}C spectroscopy. J Biomol NMR 26:167–179. doi:10.1023/A:1023572320699

Fairweather JK, Him JLK, Heux L, Driguez H, Bulone V (2004) Structural characterization by ^{13}C-NMR spectroscopy of products synthesized in vitro by polysaccharide synthases using ^{13}C-enriched glycosyl donors: application to a UDP-glucose:(1 → 3)-β-D-glucan synthase from blackberry (*Rubus fructicosus*). Glycobiology 14:775–781. doi:10.1093/glycob/cwh097

Farès C, Amata I, Carlomagno T (2007) ^{13}C-detection in RNA bases: revealing structure—chemical shift relationships. J Am Chem Soc 129:15814–15823. doi:10.1021/ja0727417

Felli IC, Pierattelli R (2012) Recent progress in NMR spectroscopy: toward the study of intrinsically disordered proteins of increasing

size and complexity. IUBMB Life 64:473–481. doi:10.1002/iub.1045

Fiala R, Sklenár V (2007) ^{13}C-detected NMR experiments for measuring chemical shifts and coupling constants in nucleic acid bases. J Biomol NMR 39:153–163. doi:10.1007/s10858-007-9184-4

Ghazarian H, Idoni B, Oppenheimer SB (2011) A glycobiology review: carbohydrates, lectins and implications in cancer therapeutics. Acta Histochem 113:236–247. doi:10.1016/j.acthis.2010.02.004

Harris R, Rutherford TJ, Milton MJ, Homans SW (1997) Three-dimensional heteronuclear NMR techniques for assignment and conformational analysis using exchangeable protons in uniformly ^{13}C-enriched oligosaccharides. J Biomol NMR 9:47–54. doi:10.1023/A:1018671517876

Hu X, Carmichael I, Serianni AS (2010) *N*-acetyl side-chains in saccharides: NMR *J*-coupling equations sensitive to CH-NH and NH-CO bond conformations in 2-acetamido-2-deoxy-aldohexopyranosyl rings. J Org Chem 75:4899–4910. doi:10.1021/jo100521g

Kadkhodaie M, Rivas O, Tan M, Mohebbi A, Shaka AJ (1991) Broadband homonuclear cross polarization using flip–flop spectroscopy. J Magn Reson 91:437–443. doi:10.1016/0022-2364(91)90210-K

Kamiya Y, Yamamoto S, Chiba Y, Jigami Y, Kato K (2011) Overexpression of a homogeneous oligosaccharide with ^{13}C labeling by genetically engineered yeast strain. J Biomol NMR 50:397–401. doi:10.1007/s10858-011-9525-1

Kato K, Yamaguchi Y, Arata Y (2010) Stable-isotope-assisted NMR approaches to glycoproteins using immunoglobulin G as a model system. Prog Nucl Magn Reson Spectrosc 56:346–359. doi:10.1016/j.pnmrs.2010.03.001

Kay LE, Xu GY, Singer AU, Muhandiram DR, Forman-Kay JD (1993) A gradient-enhanced HCCH–TOCSY experiment for recording side-chain ^{1}H and ^{13}C correlations in H_2O samples of proteins. J Magn Reson, Ser B 101:333–337. doi:10.1006/jmrb.1993.1053

Kiddle GR, Homans SW (1998) Residual dipolar couplings as new conformational restraints in isotropically ^{13}C-enriched oligosaccharides. FEBS Lett 436:128–130. doi:10.1016/S0014-5793(98)01112-0

Kjellberg A, Nishida T, Weintraub A, Widmalm G (1998) NMR spectroscopy of ^{13}C-enriched polysaccharides: application of ^{13}C–^{13}C TOCSY to sugars of different configuration. Magn Reson Chem 36:128–131. doi:10.1002/(SICI)1097-458X(199802)36:2<128:AID-OMR226>3.0.CO;2-L

Kjellberg A, Weintraub A, Widmalm G (1999) Structural determination and biosynthetic studies of the O-antigenic polysaccharide from the enterohemorrhagic *Escherichia coli* O91 using ^{13}C-enrichment and NMR spectroscopy. Biochemistry 38:12205–12211. doi:10.1021/bi9910629

Kövér KE, Hruby VJ, Uhrín D (1997) Sensitivity- and gradient-enhanced heteronuclear coupled/decoupled HSQC–TOCSY experiments for measuring long-range heteronuclear coupling constants. J Magn Reson 129:125–129. doi:10.1006/jmre.1997.1265

Kupče Ē, Freeman R (1994) Wideband excitation with polychromatic pulses. J Magn Reson Ser A 108:268–273. doi:10.1006/jmra.1994.1123

Kupče Ē, Schmidt P, Rance M, Wagner G (1998) Adiabatic mixing in the liquid state. J Magn Reson 135:361–367. doi:10.1006/jmre.1998.1607

Landersjö C, Weintraub A, Widmalm G (1997) Structural Analysis of the O-Antigenic polysaccharide from the enteropathogenic *Escherichia coli* O142. Eur J Biochem 244:449–453. doi:10.1111/j.1432-1033.1997.t01-1-00449.x

Lescop E, Schanda P, Brutscher B (2007) A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. J Magn Reson 187:163–169. doi:10.1016/j.jmr.2007.04.002

Linnerborg M, Weintraub A, Widmalm G (1999) Structural studies utilizing ^{13}C-enrichment of the O-antigen polysaccharide from the enterotoxigenic *Escherichia coli* O159 cross-reacting with *Shigella dysenteriae* type 4. Eur J Biochem 266:246–251. doi:10.1046/j.1432-1327.1999.00851.x

Lycknert K, Widmalm G (2004) Dynamics of the *Escherichia coli* O91 O-antigen polysaccharide in solution as studied by carbon-13 NMR relaxation. Biomacromolecules 5:1015–1020. doi:10.1021/bm0345108

Machonkin TE, Westler WM, Markley JL (2002) ^{13}C{^{13}C} 2D NMR: a novel strategy for the study of paramagnetic proteins with slow electronic relaxation rates. J Am Chem Soc 124:3204–3205. doi:10.1021/ja017733j

Martin-Pastor M, Bush CA (2000) Comparison of the conformation and dynamics of a polysaccharide and of its isolated heptasaccharide repeating unit on the basis of nuclear Overhauser effect, long-range C–C and C–H coupling constants, and NMR relaxation data. Biopolymers 54:235–248. doi:10.1002/1097-0282(20001005)54:4<235:AID-BIP10>3.0.CO;2-V

Martin-Pastor M, Canales-Mayordomo A, Jiménez-Barbero J (2003) NMR experiments for the measurement of proton–proton and carbon–carbon residual dipolar couplings in uniformly labelled oligosaccharides. J Biomol NMR 26:345–353. doi:10.1023/A:1024096807537

Norris SE, Landström J, Weintraub A, Bull TE, Widmalm G, Freedberg DI (2012) Transient hydrogen bonding in uniformly ^{13}C,^{15}N-labeled carbohydrates in water. Biopolymers 97:145–154. doi:10.1002/bip.21710

Nyberg NT, Duus JØ, Sørensen OW (2005) Heteronuclear two-bond correlation: suppressing heteronuclear three-bond or higher NMR correlations while enhancing two-bond correlations even for vanishing $^{2}J_{CH}$. J Am Chem Soc 127:6154–6155. doi:10.1021/ja050878w

Parella T, Sánchez-Ferrando F, Virgili A (1997) Quick recording of pure absorption 2D TOCSY, ROESY, and NOESY spectra using pulsed field gradients. J Magn Reson 125:145–148. doi:10.1006/jmre.1996.1069

Rance M, Wagner G, Sørensen OW, Wüthrich K, Ernst RR (1984) Application of ω_1-decoupled 2D correlation spectra to the study of proteins. J Magn Reson 59:250–261. doi:10.1016/0022-2364(84)90169-0

Richter C, Kovacs H, Buck J, Wacker A, Fürtig B, Bermel W, Schwalbe H (2010) ^{13}C-direct detected NMR experiments for the sequential J-based resonance assignment of RNA oligonucleotides. J Biomol NMR 47:259–269. doi:10.1007/s10858-010-9429-5

Santoro J, King GC (1992) A constant-time 2D Overbodenhausen experiment for inverse correlation of isotopically enriched species. J Magn Reson 97:202–207. doi:10.1016/0022-2364(92)90250-B

Sarkar A, Fontana C, Imberty A, Pérez S, Widmalm G (2013) Conformational preferences of the O-antigen polysaccharides of *Escherichia coli* O5ac and O5ab using NMR spectroscopy and molecular modeling. Biomacromolecules 14:2215–2224. doi:10.1021/bm400354y

Schanda P, Brutscher B (2005) Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. J Am Chem Soc 127:8014–8015. doi:10.1021/ja051306e

Schanda P, Van Melckebeke H, Brutscher B (2006) Speeding up three-dimensional protein NMR experiments to a few minutes. J Am Chem Soc 128:9042–9043. doi:10.1021/ja062025p

Sibille N, Bernadó P (2012) Structural characterization of intrinsically disordered proteins by the combined use of NMR and SAXS. Biochem Soc Trans 40:955–962. doi:10.1042/BST20120149

Soltesova M, Kowalewski J, Widmalm G (2013) Dynamics of exocyclic groups in the *Escherichia coli* O91 O-antigen polysaccharide in solution studied by carbon-13 NMR relaxation. J Biomol NMR 57:37–45. doi:10.1007/s10858-013-9763-5

Son I, Binet R, Maounounen-Laasri A, Lin A, Hammack TS, Kase JA (2014) Detection of five Shiga toxin-producing *Escherichia* coli genes with multiplex PCR. Food Microbiol 40:31–40. doi:10.1016/j.fm.2013.11.016

Stenutz RR, Weintraub A, Widmalm G (2006) The structures of *Escherichia coli* O-polysaccharide antigens. FEMS Microbiol Rev 30:382–403. doi:10.1111/j.1574-6976.2006.00016.x

Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME (2009) Essentials of glycobiology, vol 2. Cold Spring Harbor, New York

Villeneuve S, Souchon H, Riottot M–M, Mazié J-C, Lei P-s, Glaudemans CPJ, Kováč P, Fournier J-M, Alzari PM (2000) Crystal structure of an anti-carbohydrate antibody directed against *Vibrio cholerae* O1 in complex with antigen: molecular basis for serotype specificity. Proc Natl Acad Sci U S A 97:8433–8438. doi:10.1073/pnas.060022997

Vuister GW, Bax A (1992) Resolution enhancement and spectral editing of uniformly ^{13}C-enriched proteins by homonuclear broadband ^{13}C decoupling. J Magn Reson 98:428–435. doi:10.1016/0022-2364(92)90144-V

Wang W, Sass HJ, Zähringer U, Grzesiek S (2008) Structure and dynamics of ^{13}C,^{15}N-labeled lipopolysaccharides in a membrane mimetic. Angew Chemie Int Ed 47:9870–9874. doi:10.1002/anie.200803474

Xu Q, Bush CA (1998) Measurement of long-range carbon–carbon coupling constants in a uniformly enriched complex polysaccharide. Carbohydr Res 306:335–339. doi:10.1016/S0008-6215(97)10099-4

Yu L, Goldman R, Sullivan P, Walker GF, Fesik SW (1993) Heteronuclear NMR studies of ^{13}C-labeled yeast cell wall β-glucan oligosaccharides. J Biomol NMR 3:429–441. doi:10.1007/BF00176009