DOI: 10.1002/chem.200801081

Rapid Identification of Common Hexapyranose Monosaccharide Units by a Simple TOCSY Matching Approach

Katelijne Gheysen,[a] Camelia Mihai,[a, b] Karel Conrath,[c] and José C. Martins*[a]

Dedicated to Professor R. Willem on the occasion of his 60th birthday

Abstract: Solution NMR spectroscopy is a well established technique for non-destructive characterization of the structures and conformations of complex oligo- and polysaccharides. One of the key experiments involves the use of 2D TOCSY to collect the ^{1}H spins into groups that can be associated with the individual saccharide units that are present in the molecule under study. It is well known that the magnetization transfer rate through the ^{1}H spin system during the TOCSY spin lock period is sensitive to the intervening ^3J-(H,H) scalar couplings, and therefore also to the saccharide stereochemistry. Here, we have investigated the potential to extract information on the stereochemistry of hexapyranose monosaccharide units directly from TOCSY spectra. Through a systematic experimental investigation of the magnetization transfer initiated from the anomeric ^{1}H resonance in D-glucose, D-galactose and D-mannose it is shown that a 100 ms spin lock time provides optimal spectroscopic discrimination between these three commonly occurring building blocks. A simple matching scheme is proposed as a new tool for rapid attribution of the TOCSY traces originating from the anomeric ^{1}H resonances towards the underlying monosaccharide type. The scheme appears robust with regard to structural variations and fairly tolerant to incidental overlap. Its application provides useful guidance during the subsequent NMR assignment process, as demonstrated with the PS7F polysaccharide from *Streptococcus pneumonia*. In addition, we show that our scheme affords a clear-cut distinction between the α- and β-epimers of D-mannose-type units, which can be difficult to discriminate by NMR analysis. Application to the N-glycan 100.2 demonstrates the potential and wide applicability of this new discrimination approach.

Keywords: conformation analysis · NMR spectroscopy · oligosaccharides · polysaccharides · stereochemistry

Introduction

Solution NMR spectroscopy is a well established technique for the structural and conformational characterization of complex oligo- and polysaccharides and their conjugates with proteins and lipids.[1–4] Although the particular approach used for the NMR investigation depends on the sample of interest and the user, it will generally always involve 2D TOCSY spectra recorded at one or several mixing times.[5–7] Indeed, each monosaccharide unit will generally constitute a single ^{1}H spin system that is isolated from its neighbouring monosaccharide unit(s) due to the intervening glycosidic link(s). As a result, analysis of the various cross-peak networks in the TOCSY spectrum allows all resonances belonging to the same monosaccharide unit to be grouped together. Most often this analysis is performed by use of the well resolved anomeric ^{1}H as starting point. Once all the ^{1}H resonances have been grouped into monosaccharide

[a] K. Gheysen, Dr. C. Mihai, Prof. Dr. J. C. Martins
NMR and Structure Analysis Unit
Department of Organic Chemistry
Universiteit Gent, Krijgslaan 281, S4, 9000 Gent (Belgium)
Fax: (+32) 9-264-4972
E-mail: Jose.Martins@UGent.be

[b] Dr. C. Mihai
High Resolution NMR Centre (HNMR)
Department of Polymer Science and Structural Chemistry (POSC)
Vrije Universiteit Brussel, Pleinlaan 2
1050 Brussel (Belgium)

[c] Dr. K. Conrath
GlaxoSmithKline, Biologicals
Rue de l'Institut 89, 1330 Rixensart (Belgium)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/chem.200801081.

units, their precise location in the unit can subsequently be obtained from a combination of COSY- and NOESY-type experiments, supplemented whenever possible by ^{1}H–^{13}C correlation techniques.[1–5]

As the architecture of the monosaccharide spin system is mostly linear, the transfer of magnetization during the TOCSY spin lock from the anomeric ^{1}H to the opposite end of the furanose or pyranose ring will depend on the magnitude of the intervening $^{3}J(H,H)$ scalar coupling constants.[6] These are strongly dependent on the associated torsion angle. In hexapyranoses these are narrowly constrained, due to their cyclic structures. A large coupling constant (8–10 Hz), involving two neighbouring axial C–H bonds, is expected to allow a fast transfer of magnetization, whereas a small coupling constant (< 4 Hz) typical of an axial–equatorial or equatorial–equatorial C–H bond pair will considerably reduce transfer efficiency. The differences in hexapyranose stereochemistry lead to a different sequence of the axial/equatorial disposition of the neighbouring C–H bond pairs and associated scalar coupling constants along the ring. This will in turn translate into different transfer efficiencies during the TOCSY spin lock period, such that differences will occur in the intensity build-up of the TOCSY cross-peaks linking the anomeric ^{1}H with the other ^{1}H spins in the hexapyranose ring. In principle, therefore, it should be possible to derive the hexapyranose stereochemistry from its TOCSY pattern. In order to explore this, we have investigated the intensity build-up in the TOCSY pattern for a number of hexapyranose monosaccharides that are quite commonly encountered in natural oligo- and polysaccharides. Both MLEV17[7a] and DIPSI-2[8] mixing schemes are investigated. We show that D-glucose, D-mannose and D-galactose monosaccharide units can easily be distinguished from a qualitative inspection of the TOCSY cross-peak intensities in a single 100 ms TOCSY spectrum. A simplified representation of the number and minimal intensities of the TOCSY cross-peaks emanating from the anomeric ^{1}H resonance that holds in the presence both of glycosidic linkages and of N- and O-acetylation is presented. Through matching of experimentally observed TOCSY cross-peak pattern intensities and patterns with those of our scheme, information relating to the basic stereochemistry of the hexapyranose structure can easily be obtained prior to the NMR assignment. This knowledge is very useful for guiding the NMR assignment process, especially when limited quantities prevent involvement of ^{13}C nuclei, either by 1D or 2D NMR techniques, in the analysis.

The general applicability is demonstrated with PS7F (Figure 1, **1**), a capsular bacterial polysaccharide of interest for vaccine formulations targeted against *Streptococcus pneumoniae*. In addition, we show that our approach can also be used to discriminate between α- and β-D-mannose, a feature experimentally verified on the commercially available N-glycan 100.2 (Figure 1, **2**).[9] The latter observation is especially relevant, since such discrimination from the anomeric ^{1}H and/or ^{13}C chemical shift is not without ambiguity in the case of mannose.[1]

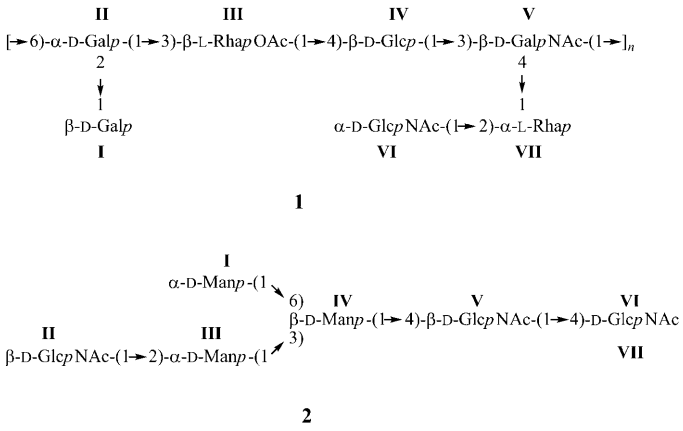

Figure 1. Structure of the capsular bacterial polysaccharide PS7F (**1**) repeat unit, and the N-linked asialo, galactosylated monoantennary N-glycan 100.2 (**2**). For both structures the Roman numerals identify the individual monosaccharides as used throughout the text.

Results and Discussion

Many different monosaccharide building blocks can contribute to the structures of biologically relevant oligosaccharides or be involved in the monomer repeating units in capsular polysaccharides. However, the basic structure and stereochemistry often boils down to that of the α and β isomers of D-glucose, D-mannose and D-galactose (Figure 2). These were therefore selected for the initial investigation. Neuraminic acid, another important oligosaccharide building block, was not considered, because it does not feature an anomeric resonance and is readily discernable from the other types. Since the anomeric proton resonances are generally well separated from the bulk of the saccharide signal, they provide the most convenient starting point for analysis of 2D TOCSY spectra. Their TOCSY traces should allow the other members of the same monosaccharide ^{1}H spin system to be identified through a chain of cross-peaks connecting the anomeric H1 all the way up to the exocyclic methylene H6 protons. Here, selective 1D TOCSY spectra obtained by use of the MLEV17 spin lock scheme[10–12] were recorded as a time-efficient means to monitor the build-up of relayed magnetization intensity. Except for the simplest of spin systems,[7b,13–15] the description of the magnetization transfer during isotropic mixing is a complicated matter and relies on numerical simulations.[16,17] Fortunately, the kinetics of the transfer in linear spin systems such as those of monosaccharides (Figure 2) are sufficiently simplified to afford a qualitative picture. Magnetization originating from an anomeric H1 will initially transfer to H2 at a rate determined by the $^{3}J(H1,H2)$ scalar coupling. As the spin lock time is increased, the original H1 magnetization will progressively transfer beyond H2 into the linear chain, with each subsequent transfer step occurring at a rate governed by the intervening $^{3}J(H,H)$ scalar couplings.

This process is nicely illustrated in the case of β-D-glucose (Figure 3a), in which all endocyclic $^{3}J(H,H)$ values are in the

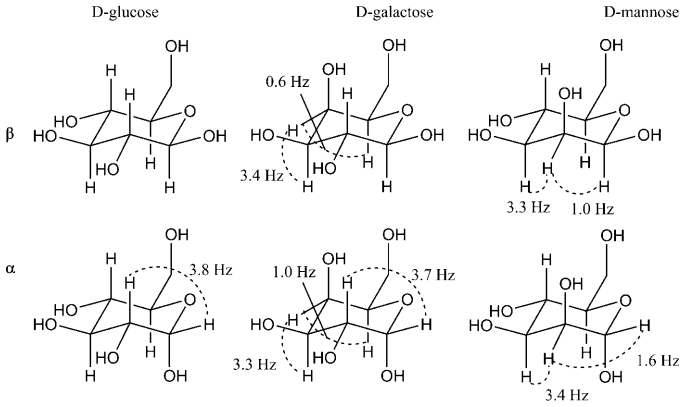

Figure 2. Structures of the studied monosaccharides: D-glucose, D-galactose and D-mannose. The small ^3J(H,H) coupling constants (<4 Hz) that are responsible for a bottleneck in the transfer of magnetization are explicitly indicated.

8.0–9.5 Hz range, due to the axial–axial relationships of neighbouring C–H bond vectors. After selective excitation of the anomeric H1 resonance, its intensity gradually decreases with increasing spin lock time, more than 10% of the excited magnetization being transferred to H2 after 10 ms. As the spin lock increases, the other ^{1}H resonances appear one by one, in a sequence and with a time lag consistent with the increasing number of transfer steps necessary to traverse the linear spin system. Transfer of up to nearly 10% of the magnetization of H1 to H5 is observed at 100 ms spin lock time for both the α- and β-D-glucose isomers (Figure 3a,b), and significant intensity (>1.5%) is seen for one or both methylene H6 protons, despite the smaller exocyclic ^3J(H5,H6) coupling constants (5.4 Hz and 1.6 Hz). If the build-up in relayed magnetization for both isomers is compared, the smaller initial ^3J(H1,H2) coupling (3.8 Hz) in α-D-glucose mostly reduces the absolute intensity of the transfer, as is evident from the different intensity scales in Figure 3a and b. It also increases the time lag preceding the appearance of a certain spin. It should be noted that the detailed evolution of the intensity build-up profile at each proton as a function of the spin lock time also illustrates the limitations of this qualitative interpretation. At the longer mixing times needed to allow transfer over the complete spin system, the relative intensities will not necessarily reflect the distance from the originally excited spin. The intensity at H3, for instance, becomes equal to or even larger than that of H2 beyond 50 ms. This underscores the well known fact that a simplistic interpretation of relative cross-peak intensities in the TOCSY pattern in terms of distance from the originating proton should be avoided, or applied with great care, even in such linear systems.

D-Mannose and D-galactose are the C2 and C4 epimers, respectively, of D-glucose. As a result of the equatorial C–H bonds at these corresponding positions, the ^3J(H,H) coupling with the preceding and following proton should be considerably reduced (Figure 2). In β-D-galactose the magnetization transfer from H3 to H4 slows down considerably, ^3J(H3,H4) being 3.4 Hz, while transfer from H4 to H5 effectively grinds to a halt because ^3J(H4,H5) amounts to only 0.6 Hz (Figure 3c). Essentially the same analysis applies for α-D-galactose, in which ^3J(H3,H4) and ^3J(H4,H5) amount to 3.3 and only 1.0 Hz, respectively. The additional small ^3J(H1,H2) coupling (3.7 Hz) in relation to α-D-galactose further reduces the absolute intensity of the relayed signal resonances (Figure 3d). The lack of any measurable build-up in intensity beyond H4 up to a spin lock time of 100 ms is therefore a characteristic signature for D-galactose.

In α- and β-D-mannose, the initial ^3J(H1,H2) scalar couplings are reduced to 1.6 and 1.0 Hz, respectively (Figure 2), such that the first magnetization transfer step is already severely attenuated. This is especially the case for β-D-mannose, in which the relative intensity of the H2 resonance builds up to barely 10% after a 100 ms spin lock (20% for α-D-mannose). In addition, the ^3J(H2,H3) couplings are only 3.4 and 3.3 Hz respectively. As a result a significant (> 1.5%) build-up in intensity is limited to H3 for α-D-mannose (Figure 3e, f) and even H2 for β-D-mannose (Figure 3g).

The data collected from the intensity profiles for the various hexapyranoses at 30, 60 and 100 ms spin lock times are presented in a qualitative fashion in Figure 4. This provides a handier representation, similar to the appearance of a cross-peak trace in a 2D TOCSY spectrum. Filled and open ovals indicate that 1.5 and 0.5%, respectively, of the anomeric resonance intensity at that specific mixing time is relayed to the corresponding ^{1}H in the monosaccharide spin system. These thresholds are not arbitrary but are chosen such that maximum discrimination between the monosaccharide types is achieved. This simplified representation allows for a more direct comparison of the magnetization transfer in these three hexapyranose core structures, and clearly reveals the positions of the magnetization transfer bottlenecks resulting from small ^3J(H,H) scalar couplings in the linear spin system. D-mannose can already be distinguished from D-galactose and D-glucose at 30 ms spin lock time. Indeed, in the latter case significant intensity is already apparent at H3 and even H4, while in that of D-mannose the very small ^3J(H1,H2) and ^3J(H2,H3) scalar coupling constants limit the transfer to H2. A clear differentiation between D-glucose and D-galactose becomes apparent at a 60 ms spin lock time, as transfer is halted at H4 in both D-galactose isomers, while the D-glucose isomers show significant intensity at H5. At 100 ms spin lock time, the three monosaccharide types can still be differentiated, but now with the added bonus that the α- and β-D-mannose isomers can clearly be distinguished from one another. Indeed, only H2 meets a threshold in β-D-mannose (1.5%), while in α-D-mannose the 0.5% threshold is achieved for most other resonances in the spin system as well. Possible confusion with D-glucose can easily be avoided by noting that in the latter case intensities above 1.5% are observed for all resonances, those of H3 and H4 being well above this threshold. Such α/β differentiation cannot be achieved, however, for D-glucose or D-galactose at any mixing time investigated.

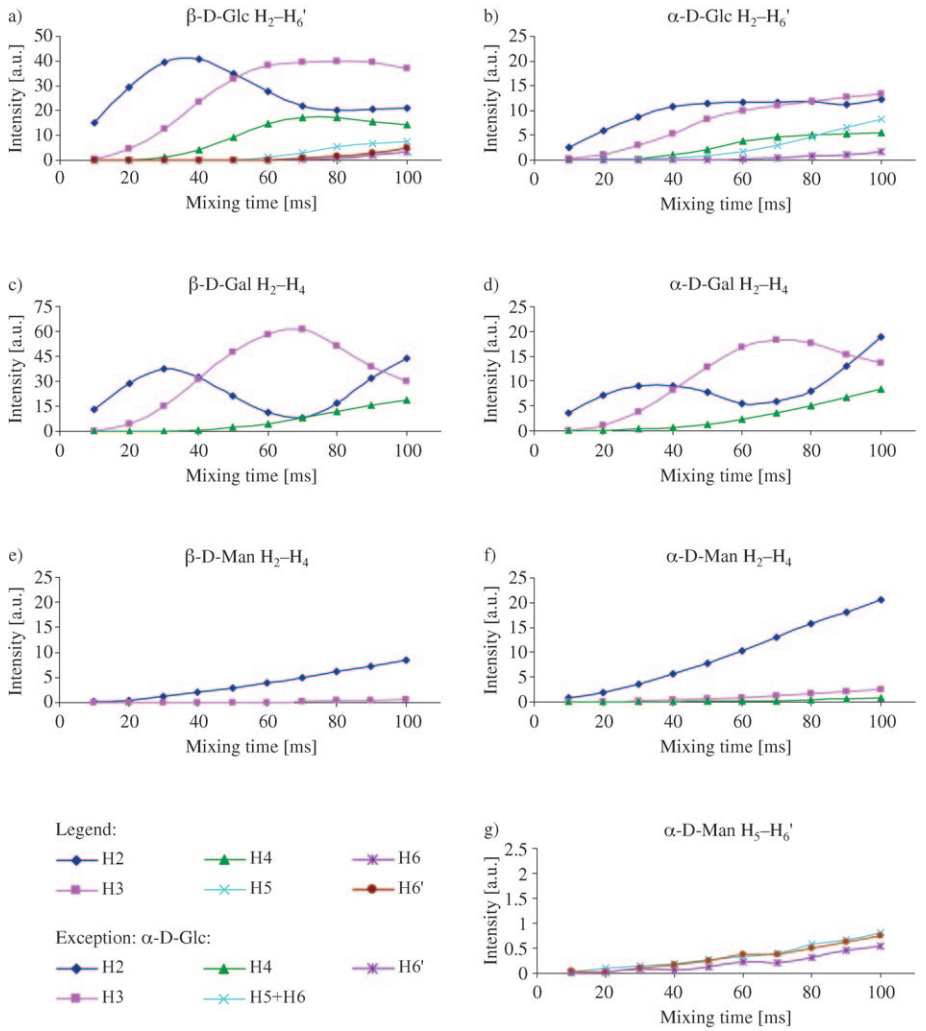

Figure 3. Intensity build-up profiles of hexapyranose resonances as a function of MLEV17 spin lock time after selective excitation of the anomeric proton of: a) β-D-glucose, b) α-D-glucose, c) β-D-galactose, d) α-D-galactose, e) β-D-mannose, and f),g) α-D-mannose. For each monosaccharide investigated, the intensities of the proton resonances are expressed relative to that of the anomeric proton in the 10 ms 1D TOCSY spectrum. Profiles are only shown for those resonances that obtained at least 0.5 % of the start intensity of the anomeric proton. Note that for β-D-glucose (a) and β-D-galactose (b) the scales of the ordinates are modified by factors of 2 and 3, respectively, whereas for (g) this factor is 10 with respect to (f). The relayed intensities are significantly larger in these cases. For α-D-glucose (b) the resonances belonging to H5 and H6 partially overlap and are shown as one build-up pattern.

The analysis performed with MLEV17 spin lock mixing was repeated with DIPSI-2,[8] another widely adopted spin lock scheme. While there are differences in the variation of the relayed magnetization intensities with mixing time along the various monosaccharide spin systems, scoring of these intensities at 1.5 and 0.5 % of the anomeric signal yields practically identical results at 30, 60 and 100 ms spin lock times (Supporting Information). We thus conclude that it is possible to differentiate D-glucose, D-mannose and D-galactose units with the aid of a single 100 ms TOCSY spectrum. Conversely, these monosaccharides can also be identified by mapping the number of the relayed resonances (1D TOCSY) or cross-peaks (2D TOCSY) originating from their anomeric resonance that reach the corresponding intensity thresholds and matching these to the schematic representation in Figure 4. In addition, for D-mannose the α and β stereochemistry can also be differentiated. This could potentially allow early identification of the various hexapyranose types and thus be useful as an independent source of information that can subsequently be used to guide the NMR-based assignment process.

Can this matching procedure really be of practical use in the structure analysis process of oligo- and polysaccharides? Two major concerns can be anticipated.

Firstly, the matching process described above assumes that all relayed resonances or cross-peaks are resolved from one another. In realistic cases however, it can be expected that incidental overlap between resonances of the same or different monosaccharide units will complicate the matching process. This is adequately addressed by using a 100 ms spin lock time to perform the analysis. At this longest mixing time, up to six relayed resonances with significant intensity can be expected for a D-glucose unit, as opposed to three for D-galactose and only two or one for α- and β-D-mannose. Recognition of three or more relayed resonances (1D) or cross-peaks (2D TOCSY) thus eliminates mannose-type units, whereas recognition of at least four cross-peaks further eliminates galactose units, demonstrating that full resolution of all signals in a TOCSY trace is not mandatory for successful typing.

Secondly, the differentiation in monosaccharide type described above must hold irrespective of the presence of one or more glycosidic linkages and or commonly occurring derivatizations such as N- and O-acetylation. These will modulate the ^3J(H,H) coupling constants as a result of changes in dihedral angles or substituent effects. Consequently, the intensities of the relayed magnetization are affected. In spite of these concerns, we find that our approach holds admirably well, as illustrated by the following examples.

In *N*-acetylglucosamine (GlcNAc), the scalar couplings involving H2 are influenced by the N-acetylation. For the α

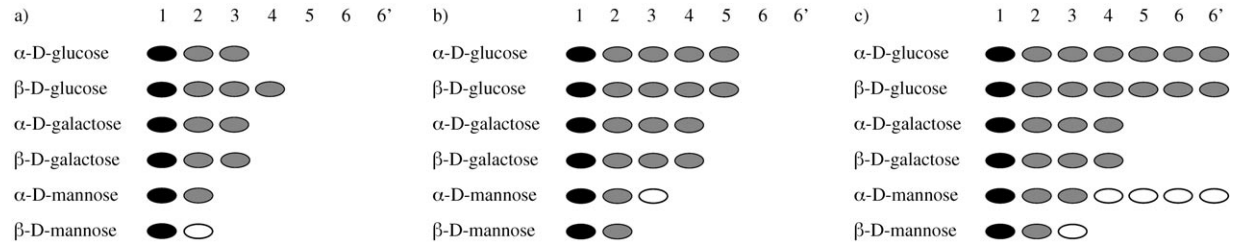

Figure 4. TOCSY pattern matching scheme. TOCSY patterns starting from the anomeric resonance (black oval) for each studied type of monosaccharide, at applied MLEV17 mixing times of a) 30 ms, b) 60 ms, and c) 100 ms. Resonances with intensities of at least 1.5% of the anomeric resonance at each of the mixing times used are represented as grey ovals. Resonances with more than 0.5% but less than 1.5% are represented as open ovals. The latter suggest that the resonances will appear in the 2D TOCSY spectrum, but only with weak intensity. Intensities lower than 0.5% are in most cases not traceable. The Scheme in (c) is applied on **1** and **2**.

anomer, ^{3}J(H1,H2) and ^{3}J(H2,H3) change from 3.8 and 9.6 Hz to 3.6 and 10.6 Hz, respectively, while for the β anomer these change from 8.0 and 9.1 Hz to 8.5 and 10.4 Hz. Despite changes up to 1.3 Hz, the couplings remain large, and magnetization transfer beyond H2 remains efficient, such that intensity is relayed up to H6 (Figure 5a,b), yielding the same magnetization transfer pattern as observed for the D-glucose core structure (Figure 4c). Despite some overlap, over four cross-peaks are easily recognized, as would be expected for a D-glucose-type unit. In the case of lactose, the presence of the β(1–4) glycosidic link between β-D-galactose and D-glucose does not impair the differentiation of the two types of hexapyranose (Figure 5c,e). Indeed, prior to any assignment, it is immediately clear that the TOCSY trace originating from the anomeric ^{1}H at δ 4.44 ppm is limited to three intense cross-peaks, a clear signature for a D-galactose, while a quick count in the traces at 4.65 and 5.21 ppm reveals five or more cross-peaks. Indeed, complete magnetization transfer from the anomeric resonance up to H6 is expected in the α- and β-D-glucose unit, while this should be halted at H4 in the β-D-galactose unit (Figure 4). Here, assignment of the α/β configuration is firmly established from the characteristic anomeric H1 chemical shift.

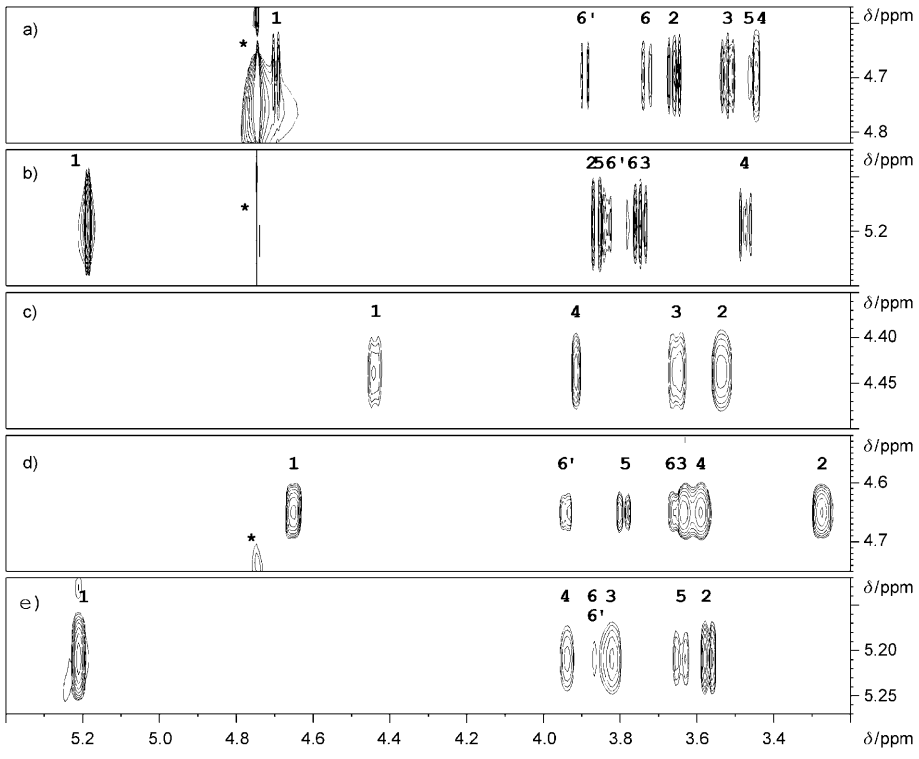

Figure 5. TOCSY patterns of a), b) *N*-acetylglucosamine, and c), d), e) lactose. Labels indicate the position of the ^{1}Hs in the corresponding hexapyranose units. On application of the TOCSY matching scheme, the identities of the monosaccharides were easily determined as a) β-D-glucose, b) α-D-glucose, c) β-D-galactose, d) β-D-glucose, and e) α-D-glucose. The individual resonances are labelled as shown on the spectrum. The * represents the residual water peak.

The PS7F capsular polysaccharide (**1**) from *Streptococcus pneumoniae* provided us with a realistic case with which to apply our matching procedure. Upon production of this polysaccharide for vaccine formulation, independent validation of the structure of the oligosaccharide repeat unit is mandatory. NMR provides the only nondestructive method to this end. Although a complete ^{1}H and ^{13}C assignment has been reported previously,[18] the resonance assignment was performed independently: that is, without use of any prior NMR knowledge. The monomeric repeat unit features seven monosaccharide units (Figure 1): two β-D-galactoses (I, V), one α-D-galactose (II), one α- (VI) and one β-D-glucose (IV) and one α- (VII) and one β-L-rhamnose (III). Several are involved in two glycosidic linkages, and the α-D-galactose (II) and one of

the β-D-galactose units (V) even in three. In view of the presence of L-rhamnose in PS7F (**1**), the magnetization transfer in L-rhamnose monosaccharides was also investigated by 1D TOCSY. In spite of its L configuration and the 6-deoxy derivatization, L-rhamnose can be regarded as having a mannose-type topology when it comes to the pattern of ^{3}J-(H,H) coupling constants. As a result, α-/β-L-rhamnose displays the same magnetization transfer pattern as found in α-/β-D-mannose when a 100 ms TOCSY spectrum is analysed (Supporting Information). Ultimately, the two can be differentiated by the presence or absence, respectively, of a methyl group in the spin system. As observed in the case of D-mannose, the TOCSY pattern also allows the α/β stereochemistry in L-rhamnose to be differentiated. From an initial analysis, seven traces associated with anomeric ^{1}H resonances can be identified in the 100 ms TOCSY spectrum (Figure 6). Classification can readily be made for most of these, on the basis of the number and intensity (strong/weak) of the cross-peaks as schematized in Figure 4. Traces E and F each have a clear galactose signature, three intense correlation peaks being easily identified in either case. The anomeric ^{1}H chemical shifts for both are characteristic for a β configuration, indicating these must belong to I and V.

Further differentiation, however, is only available upon complete resonance assignment (see below). Trace B shows four correlation peaks. With one exception, these are all of low intensity. Through reference to our matching scheme (Figure 4), these point to an α-D-mannose-type monosaccharide. Since only L-rhamnose is present, this is attributed to the α-L-rhamnose (VII). The highest-field trace, G, shows four intense correlations and a fifth weaker one. This agrees well with a glucose type, in which H6 has been reached. In view of the anomeric ^{1}H chemical shift, this is assigned to β-D-glucose (IV).

The attribution of the three remaining traces is less straightforward, mainly due to overlap issues, but should be consistent with the remaining galactose (II), rhamnose (III) and glucose (VI) units. Trace D, for instance, shows only two clear cross-peaks, together with a medium-intensity one. This does not immediately fit the Scheme in Figure 4. However, if the composition of PS7F and the possibility of overlap is taken into account, this pattern can only be accommodated by a D-glucose. Indeed, a mannose-type component (such as L-rhamnose) would show only a single strong correlation, whereas no additional weaker ones would be expected for a galactose type. Together with the anomeric ^{1}H chemical shift, the trace is assigned to an α-D-glucose and therefore unit VI in PS7F. Trace A is attributed to a galactose type, as two intense correlations but no additional weak ones are observed. Although a glucose type cannot totally be excluded, this would

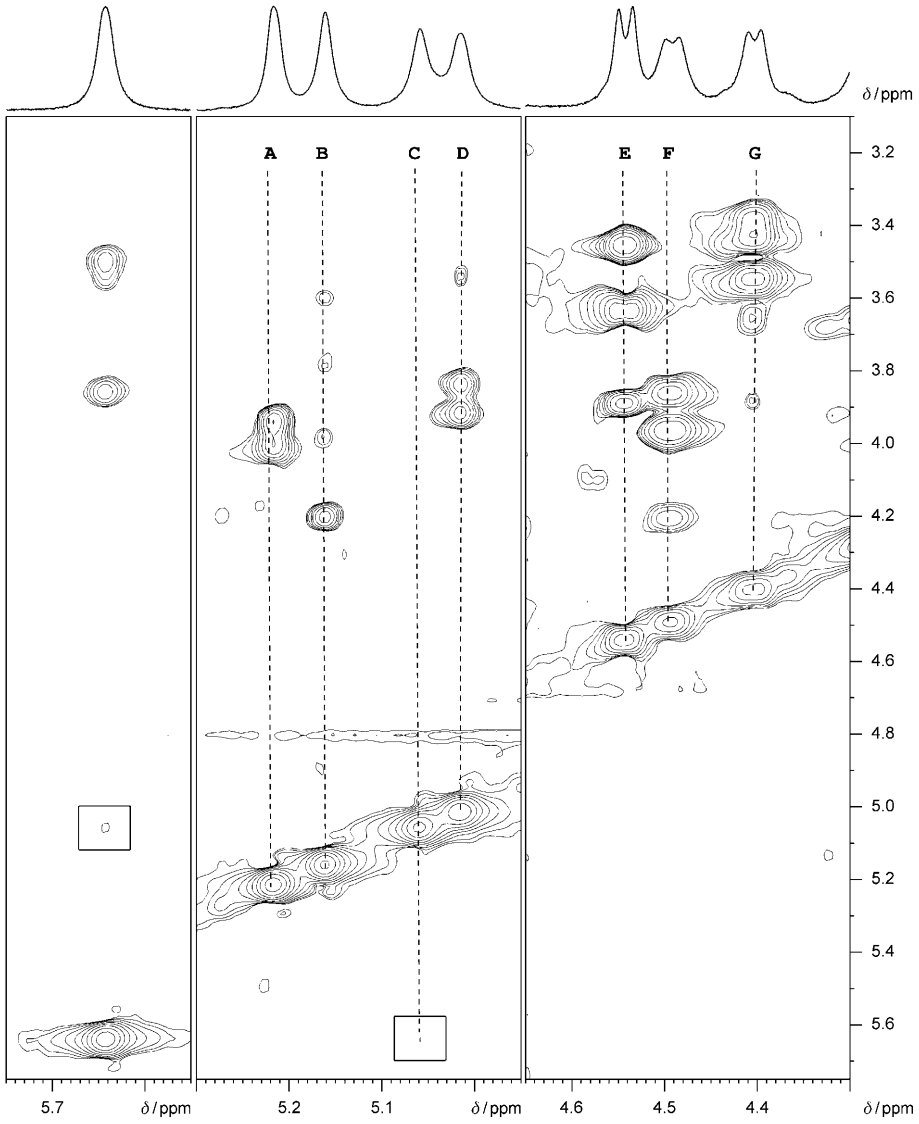

Figure 6. Anomeric part of the 100 ms 2D TOCSY spectrum of the PS7F capsular polysaccharide from *Streptococcus pneumoniae*. The 1D spectrum is shown on top and clearly shows the line-broadening typical for polysaccharides. The TOCSY traces contributed by each monosaccharide are labelled A to G, starting downfield. In order to provide an unbiaised impression of the matching scheme, no labels identifying the individual resonances are used.

imply extensive overlap, and would require another attribution for trace D, as no glucose units are left to be assigned. It is tentatively assigned to α-D-galactose (II).

This leaves trace C, which initially shows no obvious correlations to other ^{1}H spins in the TOCSY spectrum. However, when the intensity levels are scaled close to the noise level, a single correlation is found to the lowest-field, non-anomeric ^{1}H resonance at 5.64 ppm (square boxes in Figure 6). This is consistent with a D-mannose type, or in this case the remaining β-L-rhamnose (III). This is supported by noting that the unusually low-field signal can then be attributed to the only 2-O-acetylation occurring in **1**. The low intensity of the H1–H2 cross-peak is attributed to a further reduction of the ^{3}J(H1,H2) scalar coupling due to the effects of 2-O-acetylation.

The knowledge obtained from this initial analysis provided an excellent basis to guide the more extensive analysis required for full resonance assignment. The usual strategy—consisting of the combined analysis of 2D gCOSY, TOCSY, ^{1}H{^{13}C}-gHSQC and ^{1}H{^{13}C}-gHMBC experiments—was used and is described in short hereafter. Firstly, the precise topology of each ^{1}H-spin system identified from the TOCSY spectrum was established by mapping the direct connectivities in a gCOSY spectrum. The ^{1}H assignments obtained were extended to the ^{13}C nuclei with the aid of the ^{1}H{^{13}C}-gHSQC spectrum. From these data, the complete monosaccharide spin systems were assembled and validated by ^{1}H-{^{13}C}-gHMBC. Indeed, for the β-rhamnose and the galactose units, the small couplings involving H2 and H4, respectively, sever the ^{1}H spin system into two individual and unlinked parts that cannot be connected by TOCSY, but can be "sewn" back together through correlations afforded by ^{n}J-(C,H) scalar couplings. Finally, the ^{1}H{^{13}C}-gHMBC also provides connectivities across the glycosidic bond, thereby establishing sequence-specific resonance assignment. When these are analysed for PS7F, the position of each monosaccharide unit and its identity is finally established. (Supporting Information)

The type and, where available, α/β stereochemistry originally obtained from our matching scheme is fully consistent with the results from the complete assignment process. The chemical shifts attributed to the newly produced PS7F assigned here are identical to those reported before in the literature, further validating our assignment. It also shows that the α/β configurations of the two L-rhamnose units, which cannot be unambiguously discriminated because of their insufficient anomeric ^{1}H chemical shift separation (here 0.10 ppm), were correctly established from our matching scheme.

To demonstrate further the potential for α/β discrimination of mannose-type monosaccharides and their derivatives, we applied our matching approach to the N-glycan 100.2 **2**. This commercially available biantennary structure forms the core of many N-glycan oligosaccharides. It features two α-D-mannose units (I, III), one β-D-mannose (IV), two β-D-2-*N*-acetylglucosamines (II, V) and one D-2-*N*-acetylglucosamine at the reducing end. The last of these would therefore be expected to feature both α (VI) and β (VII) stereochemistry. To the best of our knowledge, the complete NMR characterization of this system has not so far been reported, although limited ^{1}H assignments for this segment as a fragment in larger oligosaccharides are available.[19] Given its considerable cost, the amount of material purchased was limited to 100 μg, representing 90 nmol of material. To optimize sensitivity, the compound was analysed by use of a 1 mm probe at 700 MHz. Again, the initial analysis focussed on the anomeric region of the 2D TOCSY spectrum, shown in Figure 7. Despite the great similarity in monosaccharide type contents, sufficient dispersion is available for seven unique traces clearly to be distinguished. These were labelled A to G and analysed with the aid of our matching scheme (Figure 4) while the anomeric ^{1}H chemical shift was taken into account.

From a first inspection the traces can be divided in two categories. Traces E, F and G all contain multiple intense cross-peaks and can therefore immediately be assigned to three D-GlcNAc units. Given their H1 chemical shifts, these all have the β-stereochemistry; together they represent units II, V and VII. The three mannose units and the remaining α-D-GlcNAc are therefore to be found in traces A–D. Trace D is easily recognized as a β-D-mannose, in view of the fact that one strong (to H2) and only one weak correlation (to H3) are visible. Traces B and C each show the typical signature for a β-D-mannose, consisting of two intense cross-peaks and three or two, respectively, weaker cross-peaks. More quantitatively, when the intensities of the traces are appropriately scaled to the diagonal peak (set to 100 %), the expected pattern emerges. This leaves trace A, which is tentatively assigned as the one corresponding to the reducing α-D-GlcNAc unit.

After these preliminary assignments of the anomeric region, a complete resonance assignment by the procedure described above was attempted. Unfortunately, the limited amount of compound impaired the signal-to-noise ratio of the ^{1}H{^{13}C}-gHMBC such that we were only able to establish clear long-range ^{3}J(C,H) correlations over the three glycosidic links involving the branching β-D-mannose IV as well as a ^{3}J(C,H) in support of the β(1–2) link between II and III (Supporting Information). This allows traces B and C to be assigned to the α-D-mannose units III and I, respectively, and traces F and G to β-D-GlcNAc V and II, respectively, in **2**. Through elimination, traces A and E correspond to the reducing D-GlcNAc (VI, VII), assigned to the α and β isomers, respectively, on the basis of their H1 chemical shift values. Despite the incomplete assignment, the successful attribution of the mannose core structure supports the α/β stereochemistry originally proposed through our matching approach. This is not possible from, for example, the anomeric H1 chemical shift, as clearly shown by the fact that the H1 of the only β-D-mannose, at 4.77 ppm, is located relatively close to that of α-D-mannose IV at 4.91 ppm (Supporting Information).

The presence of bottlenecks in the TOCSY transfer pathways of monosaccharides was recognized early on by various

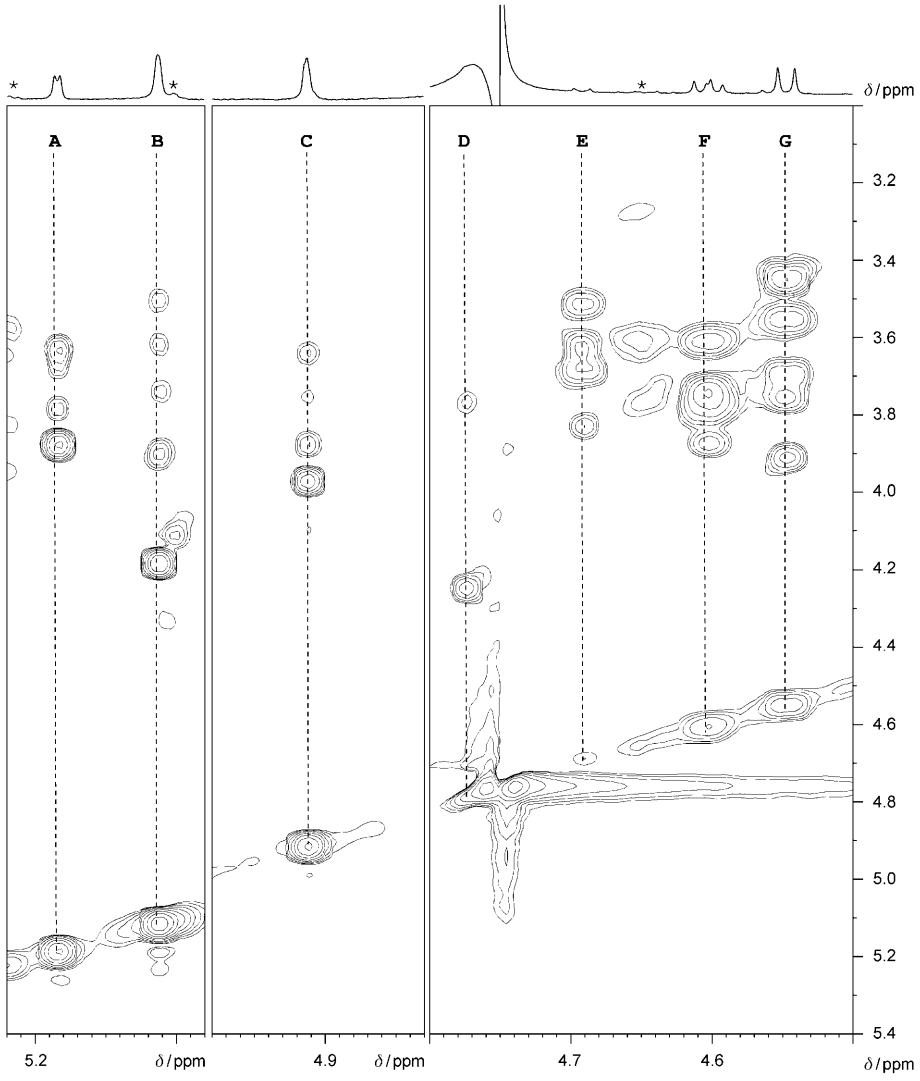

Figure 7. Anomeric part of the 100 ms 2D TOCSY spectrum of the N-glycan 100.2. The TOCSY traces contributed by each monosaccharide are labelled A to G, starting downfield. In order to provide an unbiaised impression of the matching scheme, no labels identifying the individual resonances are used. The * in the 1D spectrum shown on top marks contributions from impurities present in the sample, and determined by integration to contribute ≈15%.

groups. Bax and co-workers demonstrated the use of a relayed TOCSY method to overcome the bottlenecks associated with small scalar couplings.[6] In their work, it was shown that even at spin lock times exceeding 200 ms, no relayed magnetization is seen beyond H4 in D-galactose. Systematic investigations of the intensity build-up profile as a function of mixing time in saccharides have been reported previously for D-ribose and D-deoxyribose moieties in nucleotides.[18,21] While a link between TOCSY patterns and stereochemistry has undoubtedly been surmised, a systematic investigation to provide solid grounds for this assumption has, to the best of our knowledge, not been presented so far for common hexapyranose units.

The interest of the work presented here lies in the effective demonstration that the dependency of the bottlenecks' locations in hexapyranose units can indeed be exploited and used as an additional source of information, available at the outset of the subsequent assignment procedure. The 100 ms mixing time required is sufficiently short to be applicable to polysaccharides, where relaxation can lead to considerable line broadening and complicate analysis.[4,21] Moreover, the proposed scheme appears to be robust with regard to derivatizations of the hexapyranose units and tolerates some resonance overlap. Furthermore, it does not rely on chemical shift information, because this can be considerably influenced by variation in the structure (e.g., acetylation) and conformation. The matching scheme is not very sensitive to the spin lock scheme used, with similar results being obtained for the MLEV17 and DIPSI-2 schemes that are most often used. While other schemes, such as WALT17, have not been tested explicitly they will most probably yield similar results. Interestingly, our investigation also revealed the possibility of discriminating α from β stereoisomers in the case of mannose-type hexapyranose units. This is especially welcome in view of the difficulties that can be encountered in discriminating these through their H1 chemical shifts. So far, unambiguous α/β discrimination in D-mannose units and their derivatives has only been possible through measurement of the $^{1}J(^{13}C,^{1}H)$ scalar coupling constant of the anomeric C–H unit,[22] typically available from F_2-coupled ^{1}H–^{13}C HMQC or HSQC experiments. This requires one more measurement, otherwise of little use, than our ^{1}H TOCSY-based matching approach. More importantly, however, our 1D and 2D TOCSY matching-based approach is much more sensitive, because it only relies on the ^{1}H nucleus. This should allow the range of NMR-based α/β discrimination of D-mannose-type stereochemistry to be extended to considerably more limited quantities of material.

Conclusion

Detailed analysis of the transfer of magnetization from the anomeric ^{1}H resonance to the other ^{1}H resonances in the most common hexapyranose monosaccharides has been presented. The numbers and intensities of relayed resonances (1D) or cross-peaks (2D) in 100 ms TOCSY spectra, conveniently summarized in a qualitative scheme, are shown to afford clear-cut discrimination between glucose-, galactose- and mannose-type stereochemistries in hexapyranoses. This is highly useful in guiding the subsequent assignment process. In addition, a novel, sensitive TOCSY-based approach that allows simple discrimination between the α and β configurations in D-mannose and L-rhamnose hexapyranose units has been presented. When no information relating to the saccharide content is available, our approach could possibly supplement data on saccharide content and stereochemistry obtained by the common analysis methods based on chemical derivatization, in a non-destructive way. Given that both 1D and 2D TOCSY techniques are nowadays available in a highly standardized, even automated, format on all commercial NMR platforms, our method should be easily and widely applicable.

Experimental Section

Sample preparation for NMR: All solutions were prepared in D_2O (99.9%) with 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) as internal chemical shift reference standard and a small amount of NaN_3 to prevent the growth of micro-organisms. NE-HP5–7" NMR tubes (New Era, Inc., 5 mm) were used throughout. Monosaccharides were purchased from Sigma–Aldrich and were dissolved in D_2O to afford 30 mM concentrations. PS7F (**1**) was obtained by fermentation of the serotype specific strain, extraction and purification of the polysaccharide by the standard procedures used at GSK Biologicals. N-glycan 100.2 (**2**) was purchased from Dextra Laboratories and used without further purification. A total of 100 μg was dissolved in D_2O (10 μL) and transferred into a 1 mm capillary (Cortec) for measurement by use of the 1 mm probe. The NMR spectra showed extra resonances of minor intensity indicating that the concentration of 8.9 mM should be considered an upper limit. These signals did not interfere significantly with the analysis of the main compound (or compound of interest). The PS7F sample was prepared by dissolving solid material (20 mg), previously lyophilized from D_2O, in D_2O (2 mL). The lyophilized material was reconstituted in a total volume of two millilitres with D_2O (99.9%) as solvent.

NMR data acquisition and processing: Unless mentioned otherwise, spectra were recorded on a Bruker Avance II NMR spectrometer running TopSpin 1.3 software, fitted with a ^{1}H/^{119}Sn/BB *z*-gradient probe (TBI, 5 mm) and operating at 700.13 MHz for ^{1}H and 176.05 MHz for ^{13}C. Except when otherwise mentioned, the temperature used was 298 K. The 1D selective TOCSY spectra were acquired by use of a gradient-enhanced selective inversion scheme with either a MLEV17 (selmlgp) or a DIPSI-2 (seldigp) spinlock sequence. On each monosaccharide sample a series of ten 1D selective TOCSY spectra was recorded for mixing times of 10 ms to 100 ms with 10 ms steps and for each spin lock mixing sequence, a 10 s relaxation delay was used throughout. For D-mannose, overlap with the residual water peak was avoided by measuring at 302 K. The selective 180° pulse used was a standard 160 ms Gaussian pulse with a 6.7 Hz B_1 field strength. The 1D raw data, accumulated from 16 scans, 8k data points each, were multiplied with an exponential function with 1.5 Hz line broadening, prior to zero-filling to 32k followed by Fourier transformation and baseline correction. Integration of resonances in each recorded set was performed relative to the intensity of the excited anomeric resonance in the 10 ms 1D TOCSY spectrum and expressed as a percentage. For the polysaccharide **1** and asialo N-glycan **2**, gCOSY, ^{1}H-{^{13}C}gHSQC and ^{1}H{^{13}C}-gHMBC spectra were acquired by use of the standard sequences in the Bruker pulse programme library.[23] 2D TOCSY spectra were recorded with 100 ms of either MLEV or DIPSI-2 spin lock. The assignment of the PS7F **1** was entirely performed by use of 500 MHz data recorded on a Bruker DRX 500 instrument fitted with a ^{1}H/^{13}C/^{15}N TXI-*z*-gradient probehead. Because of the limited amount of the asialo N-glycan **2**, a ^{1}H/^{13}C/^{15}N TXI-Z probe (1 mm) was used. For homonuclear 2D experiments typically 512 increments, consisting of 1 to 32 scans of 2k datapoints each, were recorded with a 1.27 s relaxation delay and a ^{1}H spectral width of 12.02 ppm. The residual water signal was removed by selective presaturation during the relaxation delay when appropriate. For heteronuclear 2D experiments, a ^{13}C spectral width of 140 ppm (gHSQC) or 180 ppm (gHMBC) and 16–128 scans were used. Processing consisted of apodization with a squared cosine window function, followed by zero-filling before Fourier transformation to a 2k×2k complex dataset. A second-order polynomial baseline correction was applied when necessary. All ^3J(H,H) scalar couplings for the various monosaccharides were measured from the 1D spectra after mild Gaussian apodization.

Acknowledgements

K.G. is grateful to the Research Council of UGent for a PhD grant (GOA project 01.G015.07). C.M. thanks the SBO program of the IWT for a postdoctoral research grant. J.C.M. wishes to acknowledge the Fund for Scientific Research—Flanders (FWO-Vlaanderen) for various equipment grants (G.0365.03, G.0064.07). The 700 MHz equipment of the Interuniversitary NMR Facility was financed by Ghent University, the Free University of Brussels (VUB) and the University of Antwerp through the "Zware Apparatuur" Incentive of the Flemish Government.

[1] P. K. Agrawal, *Phytochemistry* **1992**, *31*, 3307–3330.
[2] J. Duus, C. H. Gotfredsen, K. Bock, *Chem. Rev.* **2000**, *100*, 4589–4614.
[3] J. Jiménez-Barbero, T. Peters, *NMR Spectroscopy of Glycoconjugates*, Wiley-VCH, **2002**.
[4] C. Jones, *J. Pharm. Biomed. Anal.* **2005**, *38*, 840–850.
[5] S. W. Homans, R. A. Dwek, J. Boyd, N. Soffe, T. W. Rademacher, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 1202–1205.
[6] F. Inagaki, I. Shimada, D. Kohda, A. Suzuki, A. Bax, *J. Magn. Reson.* **1989**, *81*, 186–190.
[7] a) A. Bax, D. G. Davis, *J. Magn. Reson.* **1985**, *65*, 355–360; b) L. Braunschweiler, R. R. Ernst, *J. Magn. Reson.* **1983**, *53*, 521–528.
[8] S. P. Rucker, A. J. Shaka, *Mol. Phys.* **1989**, *68*, 509–517.
[9] N. Tomiya, J. Awaya, M. Kurono, S. Endo, Y. Arata, N. Takahashi, *Anal. Biochem.* **1988**, *171*, 73–90.
[10] H. Kessler, H. Oschkinat, C. Griesinger, *J. Magn. Reson.* **1986**, *70*, 106–133.
[11] J. Stonehouse, P. Adell, J. Keeler, A. J. Shaka, *J. Am. Chem. Soc.* **1994**, *116*, 6037–6038.
[12] K. Stott, J. Stonehouse, J. Keeler, T. L. Hwang, A. J. Shaka, *J. Am. Chem. Soc.* **1995**, *117*, 4199–4200.
[13] N. Chandrakumar, *J. Magn. Reson.* **1986**, *67*, 457–465.
[14] B. Luy, O. Schedletzky, S. J. Glaser, *J. Magn. Reson.* **1999**, *138*, 19–27.
[15] M. Rance, *Phys. Lett.* **1989**, *154*, 242–247.
[16] J. Cavanagh, M. Rance, *J. Magn. Reson.* **1990**, *87*, 408–414.
[17] J. P. M. Vanduynhoven, J. Goudriaan, C. W. Hilbers, S. S. Wijmenga, *J. Am. Chem. Soc.* **1992**, *114*, 10055–10056.
[18] M. Moreau, J. C. Richards, M. B. Perry, P. J. Kniskern, *Carbohydr. Res.* **1988**, *182*, 79–99.

[19] H. Ogawa, A. Yoneda, N. Seno, M. Hayashi, I. Ishizuka, S. Hase, I. Matsumoto, *Eur. J. Biochem.* **1995**, *230*, 994–1000.
[20] S. S. Wijmenga, M. W. Mooren, C. W. Hilbers, *NMR of nucleic acids: from spectrum to structure*, in *NMR of Macromolecules*, *A Practical Approach* (Ed.: G. C. K. Roberts), **1993**, pp. 217–288.
[21] S. J. F. Vincent, *NMR Experiments for Large Carbohydrates, in NMR Spectroscopy of Glycoconjugates* (Eds.: J. Jiménez-Barbero, T. Peters), Wiley-VCH, **2003**, pp. 95–108.
[22] K. Bock, C. Pedersen, *J. Chem. Soc. Perkin Trans. 2* **1974**, 293–299.
[23] S. Berger, S. Braun, *200 and More NMR Experiments*, Wiley-VCH, **2004**, and references therein.

Received: June 4, 2008
Published online: August 26, 2008

