University of Reading

A solution NMR approach to determine the chemical structures of carbohydrates using the hydroxyl groups as starting points

Article

Published Version

Open Access ACS AuthorChoice License

Brown, G. D., Bauer, J., Osborn, H. M. I. ORCID: https://orcid.org/0000-0002-0683-0457 and Kuemmerle, R. (2018) A solution NMR approach to determine the chemical structures of carbohydrates using the hydroxyl groups as starting points. ACS Omega, 3 (12). pp. 17957-17975. ISSN 2470-1343 doi: 10.1021/acsomega.8b02136 Available at https://centaur.reading.ac.uk/82103/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1021/acsomega.8b02136

Publisher: ACS

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

University of Reading

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

This is an open access article published under an ACS AuthorChoice License, which permits copying and redistribution of the article or any adaptations for non-commercial purposes.

ACS OMEGA

Article

Cite This: ACS Omega 2018, 3, 17957−17975

http://pubs.acs.org/journal/acsodf

A Solution NMR Approach To Determine the Chemical Structures of Carbohydrates Using the Hydroxyl Groups as Starting Points

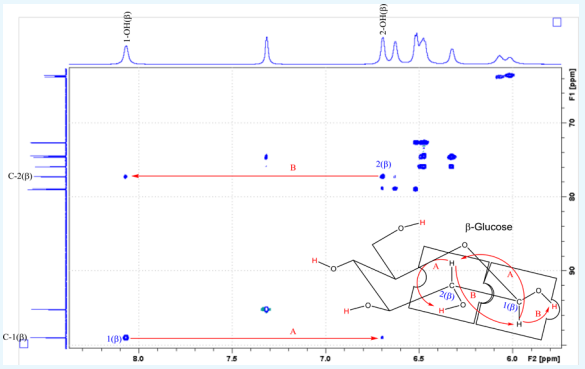
Geoffrey D. Brown,*[,†] Julia Bauer,[‡,§] Helen M. I. Osborn,[‡] and Rainer Kuemmerle[∥]

[†]Department of Chemistry and [‡]Reading School of Pharmacy, The University of Reading, Whiteknights, Reading RG6 6AP, United Kingdom

[∥]Bruker Biospin AG, NMR Division, Industriestrasse 26, CH-8117 Fallanden, Switzerland

Supporting Information

ABSTRACT: An efficient NMR approach is described for determining the chemical structures of the monosaccharide glucose and four disaccharides, namely, nigerose, gentiobiose, leucrose and isomaltulose. This approach uses the ^{1}H resonances of the −OH groups, which are observable in the NMR spectrum of a supercooled aqueous solution, as the starting point for further analysis. The 2D-NMR technique, HSQC-TOCSY, is then applied to fully define the covalent structure (i.e., the topological relationship between C−C, C−H, and O−H bonds) that must be established for a novel carbohydrate before proceeding to further conformational studies. This process also leads to complete assignment of all ^{1}H and ^{13}C resonances. The approach is exemplified by analyzing the monosaccharide glucose, which is treated as if it were an "unknown", and also by fully assigning all the NMR resonances for the four disaccharides that contain glucose. It is proposed that this technique should be equally applicable to the determination of chemical structures for larger carbohydrates of unknown composition, including those that are only available in limited quantities from biological studies. The advantages of commencing the structure elucidation of a carbohydrate at the −OH groups are discussed with reference to the now well-established 2D-/3D-NMR strategy for investigation of peptides/proteins, which employs the −NH resonances as the starting point.



INTRODUCTION

In recent years, there has been considerable interest in deciphering the various biological roles of carbohydrates, and as a result, it is now well established that carbohydrates are involved in a multitude of biological processes, including cell−cell recognition, differentiation, and adhesion processes.[1] Despite the importance of carbohydrates in biology, the determination of their covalent structure—the topological relationship between C−C, C−H, and O−H bonds that defines the ordering of these atoms within a molecule—remains a challenging analytical problem. NMR spectroscopy is the most powerful analytical technique currently available for determining the covalent structure of small organic molecules in solution;[2] however, the ^{1}H NMR spectroscopic analysis of a sugar is compromised because the hydrogen signals from the C−H units of the carbohydrate backbone normally occupy only a small portion of the full range of all possible ^{1}H chemical shifts (δ_H 3−5.5 ppm from the full range of 0−10 ppm). Chemical shifts in the ^{13}C NMR spectrum of a carbohydrate follow the same general trends as for ^{1}H NMR spectroscopy, although the probability of "accidental overlap" is less because the dispersion of ^{13}C shifts (δ_C 60−110 ppm) is approximately 20 times wider in the ^{13}C NMR spectrum (full range: 0−220 ppm).

There is a strong imperative to avoid accidental overlap because it leads to ambiguities in the logical process by which the interpretation of an NMR spectrum is made, which can quickly make structure elucidation/complete NMR assignment impossible. For the limited dispersion in a ^{1}H NMR spectrum, the probability of accidental overlap is normally reduced by spreading the chemical shift information out over a second dimension, for example, through application of a 2D experiment[3] in which ^{1}H is correlated with ^{13}C, such as heteronuclear single quantum coherence spectroscopy (HSQC). In the standard 2D-NMR approach to determining the covalent structure for a small organic molecule of unknown structure,[4] HSQC is used to create CH, CH_2 and CH_3 "jigsaw pieces", which are then connected to one another by "longer range" techniques, such as heteronuclear multiple bond coherence spectroscopy (HMBC) and proton correlation spectroscopy (^{1}H−^{1}H COSY) that depend on coupling through chemical bonds. In addition to determining the planar structure of the unknown, this strategy also results in a full set of NMR assignments. These assignments are an absolute

Received: August 22, 2018
Accepted: December 7, 2018
Published: December 21, 2018

Downloaded via UNIV OF READING on February 7, 2019 at 11:44:27 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

© 2018 American Chemical Society

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

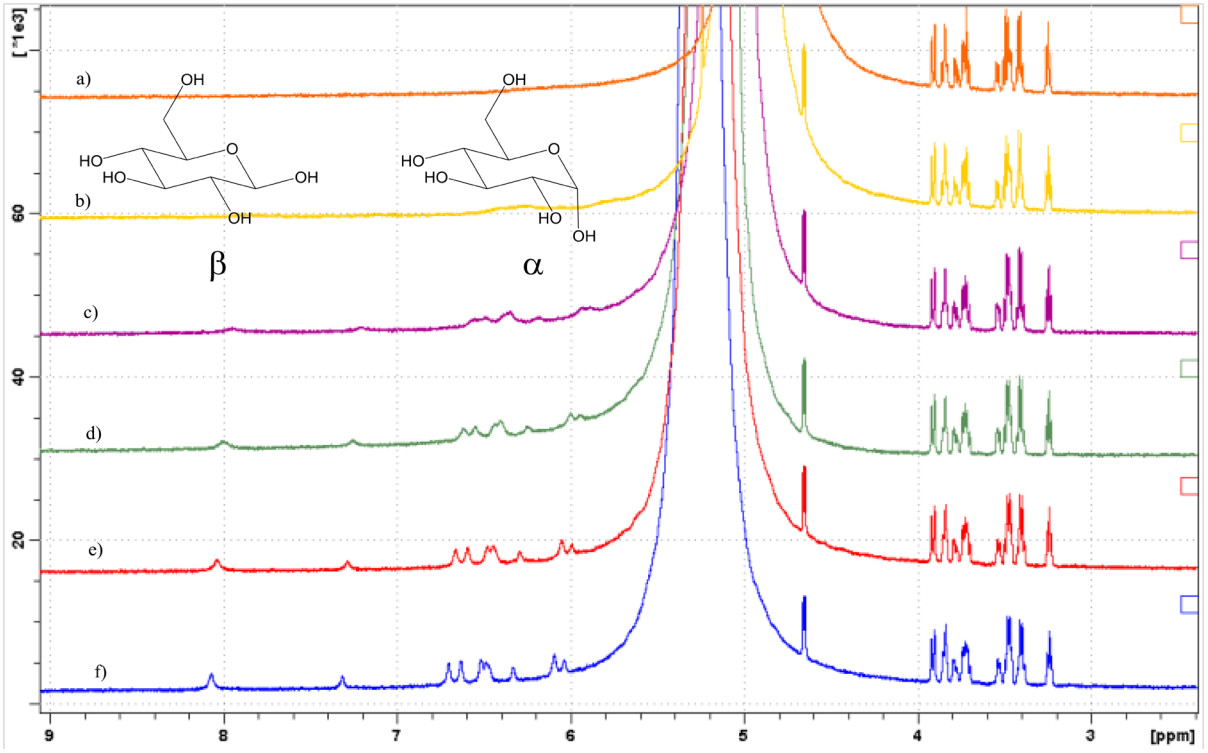


Figure 1. 700 MHz ^{1}H NMR spectra of a 100 mM solution of glucose, prepared in 1% D_2O/99% H_2O and contained in a single 0.7 mm i.d. capillary. Spectra were recorded at (a) 18 (291), (b) 8 (281), (c) −2 (271), (d) −7 (266), (e) −11 (262), and (f) −14 °C (259 K).

requirement for the final stage of the process of structure elucidation for a small molecule, in which through-space connections, often established by nuclear Overhauser spectroscopy (NOESY), are used to assign both relative configuration at stereocenters and to define geometrical isomerism at double bonds. For oligosaccharides, a complete 3D description can also include aspects of molecular conformation that are dictated by weaker intramolecular forces, such as hydrogen bonding between −OH groups, which may also be determined by NOESY[5] and other techniques.[6−9]

This strategy can be successful for the elucidation of the chemical structure of oligosaccharides of unknown constitution consisting of up to 3−8 monosaccharide units, when implemented with a high-field NMR spectrometer. This limit is a consequence of the accumulation of ambiguities for the through-bond connections made by HMBC and COSY in the standard strategy. This is, in turn, a result of the increasing probability of accidental overlap of both ^{13}C and ^{1}H resonances. One plausible strategy by which current limitations to the determination of the covalent structure of an oligosaccharide might be overcome would be to increase the dimensionality of the NMR experiment still further. Thus, if the resonances from an oligosaccharide could be spread out over a 3D volume instead of a 2D area, then the probability of accidental overlap would be even further diminished. This strategy has been applied with great success to the structure determination of proteins by NMR. The primary structure of a protein is usually known in advance, so a well-established suite of 3D-NMR techniques, such as HNCO, HNCA, and HN(CO)CA,[10] is used to "connect" three different nuclei (^{1}H, ^{13}C and ^{15}N) in the backbone of this ^{13}C/^{15}N doubly labeled protein to provide only an unambiguous assignment of all NMR resonances. This is then the cornerstone for the routine study of the secondary and tertiary structures of proteins that can consist of hundreds of amino acids. Although 3D-NMR has been described for oligosaccharides,[11] it has not found widespread application, partly because it is impractical to include a third nucleus for a carbohydrate, which consists only of the elements C, H, and O, where ^{17}O is quadrupolar[12] and therefore unsuited for high-resolution NMR studies. In addition, unlike proteins, it is difficult to completely enrich a carbohydrate with ^{13}C, and the non-availability of uniformly ^{13}C-labeled carbohydrates severely limits the possibilities of pulse sequence design.

An alternative strategy to improve ^{1}H dispersion would be to include the hydrogens from the −OH groups of a carbohydrate in the analysis.[13] In this approach, the ^{1}H chemical shifts of −OH groups would serve as an alternative starting point for structural assignment and the resolution between −OH resonances might be further increased by a 2D-NMR experiment that correlated their chemical shifts with the ^{13}C resonances of the carbons to which they are attached. Unfortunately, the hydroxyl resonances of a sugar are not normally seen when its ^{1}H NMR spectrum is acquired in D_2O solution at room temperature (as is standard practice) because the −OH groups undergo rapid chemical exchange with ^{2}H in the aqueous solvent. Two approaches are described in the literature to render these hydroxyl groups visible. The first is to acquire the NMR spectrum in an alternative solvent, such as dry d_6-DMSO.[2,13−16] The limitation here is that hydrogen bonds in organic solvents do not necessarily mirror the hydrogen bonds formed in aqueous solution, and the approach is therefore of questionable value for structural studies of biologically active oligosaccharides, which should be made under physiologically relevant conditions. The second option is to acquire the NMR spectrum from a supercooled aqueous solution,[17−20] in which the rate of −OH exchange has been

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

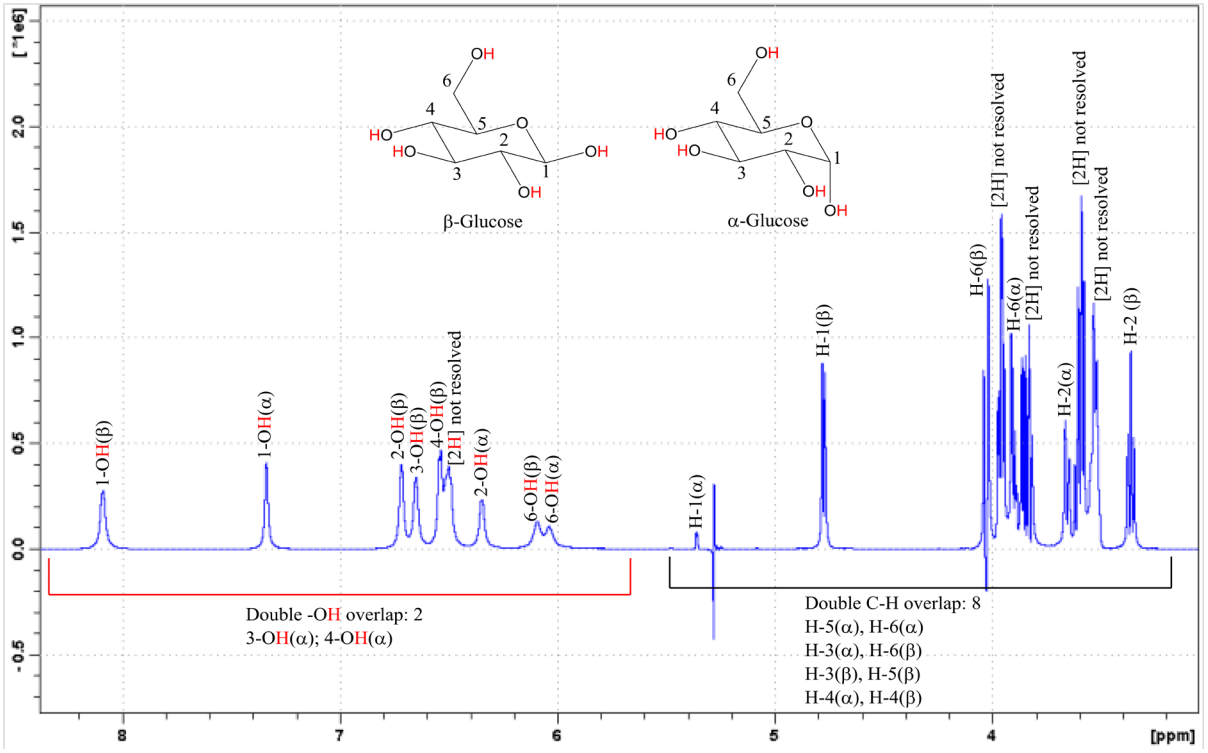


Figure 2. 700 MHz ^{1}H NMR spectrum with solvent suppression by excitation sculpting, acquired from three 0.7 mm i.d. capillaries, packed inside a 3 mm NMR tube at −14 °C (259 K). As previously, the sample is a 100 mM solution of glucose in 1% D_2O/99% H_2O.

slowed sufficiently that the hydroxyl groups can be resolved.[a] Most investigators consider this to be the better alternative because it favors the observation of more stable conformations already present in aqueous solution; however, both techniques have found quite considerable application in conformational studies of hydrogen bonding.[21−33] Because there are also some precedents for using supercooled aqueous solutions in NMR structural studies of biologically active proteins and RNA,[34−36] we have adopted this second approach, when exploring this alternative to the conventional NMR strategy for determining the covalent structure of carbohydrates in solution. We have probed the feasibility of using this alternative technique for elucidation of the chemical structure of the monosaccharide glucose, which is treated as if it were an unknown. We have also applied this approach to make full assignments of the NMR spectra of four disaccharides. It is proposed that this strategy could be extended to larger oligo- and polysaccharides of both known and unknown structure.

RESULTS AND DISCUSSION

Under standard ^{1}H NMR acquisition conditions (e.g., at room temperature for a sample contained in a 5 mm NMR tube), the −OH groups of a carbohydrate undergo rapid exchange with deuterium in the D_2O solvent, becoming converted to −OD groups, which are not visible in the spectrum. Therefore, to be able to use the −OH groups as a starting point for making NMR assignments of the covalent structure of a carbohydrate, the effects of such exchange have to be minimized. We have sought to achieve this in two ways: (i) by adopting the NMR spectrometer with the highest magnetic field strength available to us (this increases the frequency separation between exchanging species, thereby reducing the extent of associated line broadening) and (ii) by minimizing the deuterium content of the solvent. For the 700 MHz/TCI cryoprobe system employed in this study, it was possible to prepare samples in H_2O solutions that contained as little as 1% D_2O, which was the minimum quantity of deuterium compatible with locking and shimming for this system.

Reducing the Rate of −OH Exchange. The rate of −OH exchange was lowered by decreasing the temperature, while preventing the solution from freezing by maintaining it in a supercooled state in a capillary. Thus, for the above NMR system, we were routinely able to obtain NMR spectra from a 100 mM solution of glucose at −14 °C (259 K), when that solution was contained in a 0.7 mm i.d. capillary NMR tube. The reduction in the rate of exchange as the temperature is lowered toward −14 °C is illustrated by the gradual appearance of the −OH peaks from both α- and β-anomers of glucose in the ^{1}H NMR spectra that are displayed in Figure 1. Although there are a few examples in which −OH groups have proved useful for NMR studies at or above the normal freezing point of water,[5,9,37] most investigators have chosen to explore supercooled solutions at temperatures below 0 °C, which lead to a further sharpening of the −OH peaks. One consequence of this sharpening in the present study is that it was possible to acquire each of the HSQC-TOCSY spectra in a relatively short period of time (40 min). Hence, it is possible to collect all five of the 2D data sets that are required for the efficient approach described below in a reasonable time frame, even when the unknown carbohydrate is only available in limited quantities, for example, when isolated from a biological source.

The topmost ^{1}H NMR spectrum (Figure 1a), which was acquired at 18 °C (291 K), exhibits only those peaks that are normally observed for glucose, that is, several strongly overlapped multiplets for the methine (CH) and methylene (CH_2) protons between δ_H 3−5.5 ppm. When this NMR sample is cooled to 8 °C (281 K; Figure 1b), a new set of

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

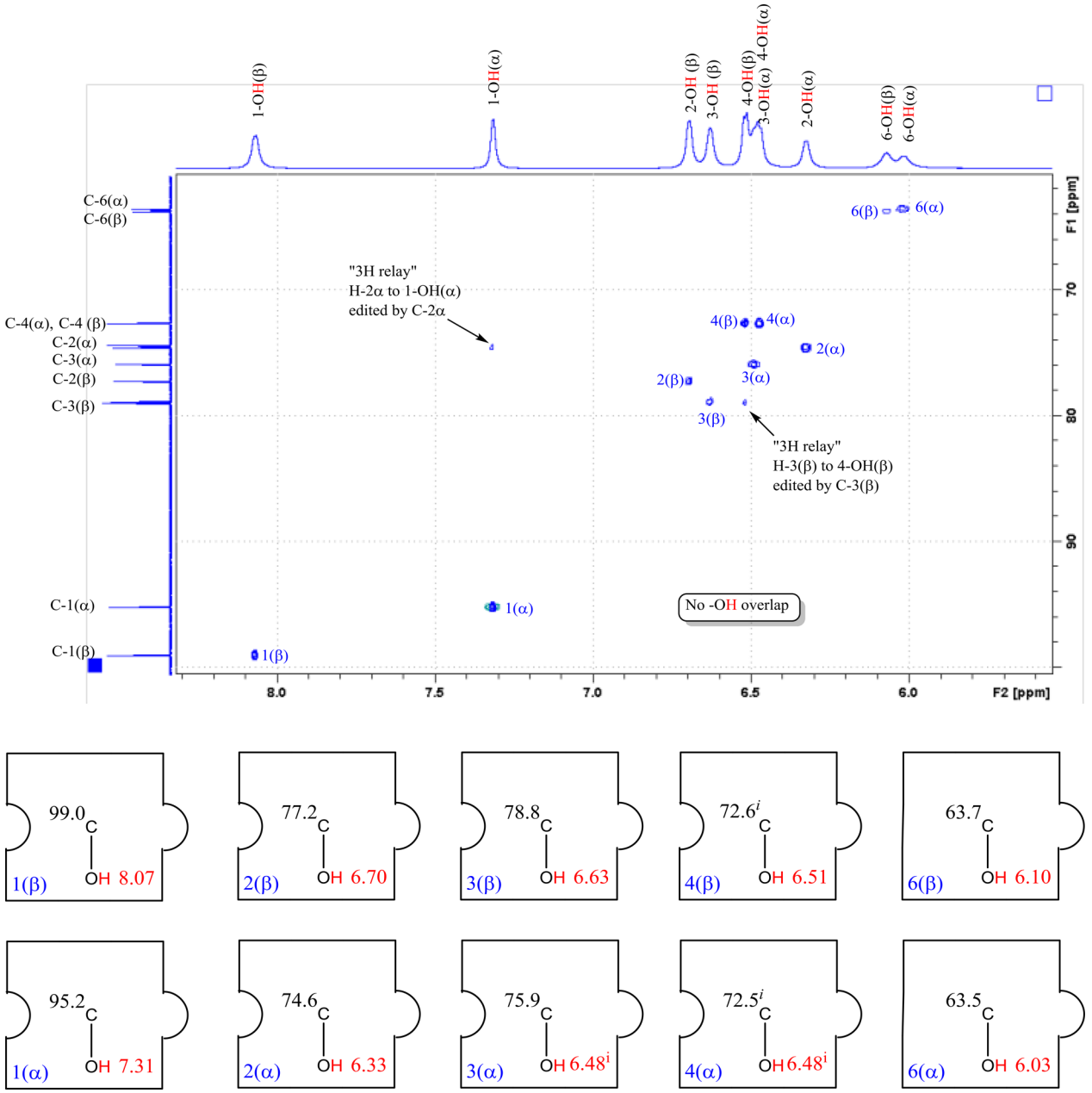


Figure 3. Expansion (δ_H 5.5−8.5 ppm in F2) of the 700 MHz HSQC-TOCSY spectrum of a solution of glucose in 1% D_2O/99% H_2O, supercooled to −14 °C. The corresponding (1D) ^{1}H NMR spectrum appears horizontally on top, and the (1D) ^{13}C NMR spectrum appears vertically on the left-hand side. The 10 2D peaks have also been represented within 10 jigsaw pieces. Ambiguities due to accidental chemical shift overlap in ^{13}C have been indicated by *i* in two of these jigsaw pieces and by i for ^{1}H overlap in a further two OH pieces. Two exceptional transfers over three hydrogen atoms of a coupled network have been indicated by arrows on the spectrum itself.

extremely broad peaks become barely discernible in the range δ_H 5.5−8.5 ppm. The ^{1}H NMR spectra in Figure 1c−f were recorded from supercooled aqueous solutions of glucose at −2 (271), −7 (266), −11 (262), and −14 °C (259 K), respectively. These spectra show a progressive sharpening of the "new" resonances between δ_H 5.5−8.5 ppm as the temperature is lowered to −14 °C (Figure 1f). In addition, there is a steady increase in chemical shift for all peaks in this range as the temperature is reduced.[26] Both observations are consistent with identification of the new resonances between δ_H 5.5−8.5 ppm with the −OH protons of glucose.

Improving the Quality of the NMR Spectrum from a Supercooled Capillary. Various solvent suppression techniques were investigated to reduce the ^{1}H NMR signal from the 99% H_2O solvent, which is off-scale for all spectra in Figure 1. Of these, excitation sculpting[38,39] had the least effect on the exchangeable −OH protons in glucose, as is shown in Figure 2.[b] Excitation sculpting has therefore been applied routinely to all NMR spectra reported herein. The quality of the spectrum in Figure 2 was further improved by packing three capillaries (1 mm o.d.) inside a standard 3 mm NMR tube to compensate for the relatively poor signal-to-noise ratio from the small amount of glucose that is contained within an individual NMR capillary tube.

Improving the Resolution between the −OH Groups. Nine broad singlets are resolved between δ_H 5.5 and 8.5 ppm

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

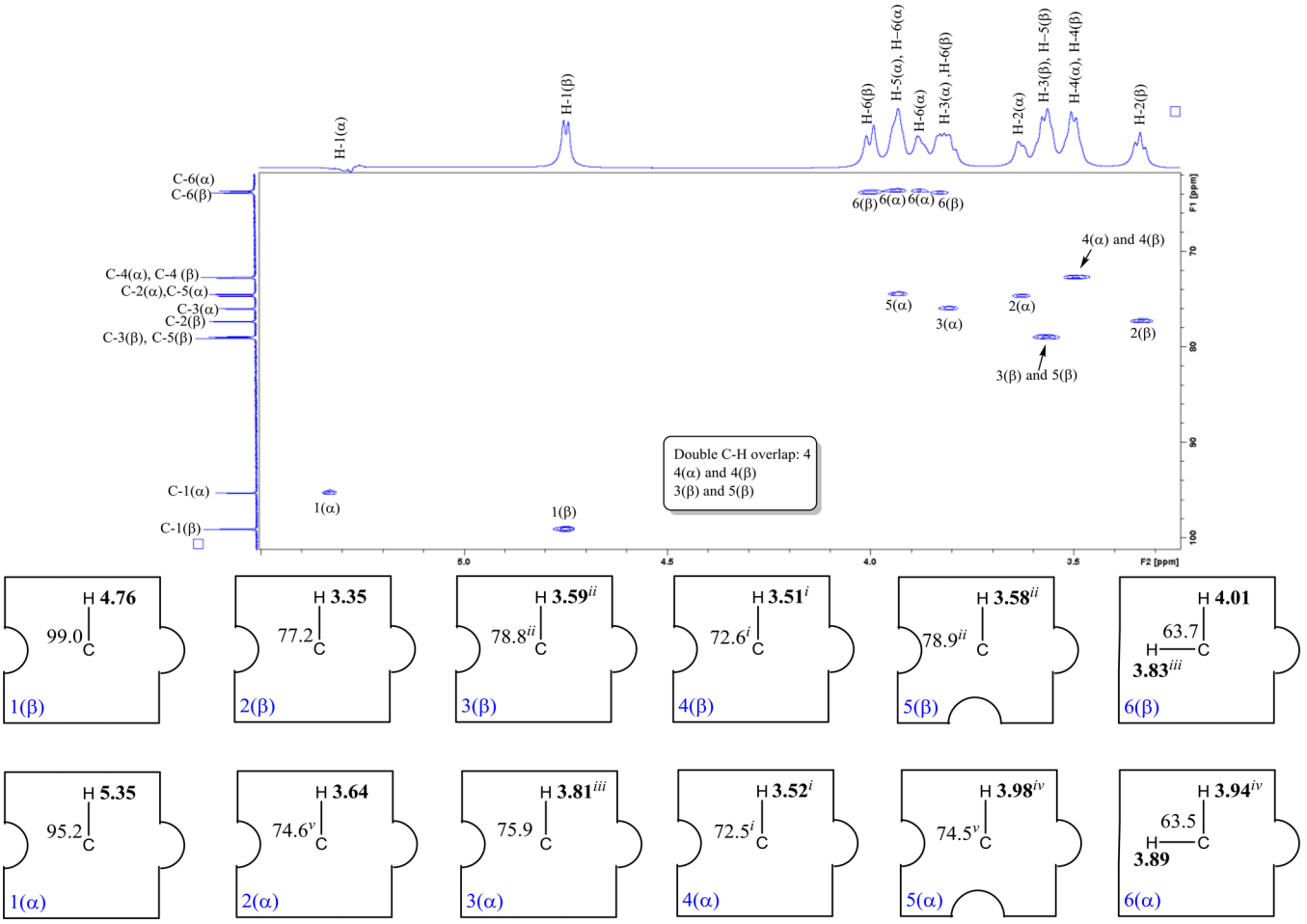


Figure 4. Expansion (δ_H 3.0−5.5 ppm in F2) of the 700 MHz HSQC spectrum of glucose in 1% D_2O/99% H_2O, supercooled to −14 °C. The 14 2D peaks (4 of which are not fully resolved) are represented as constituents of 12 jigsaw pieces, in which ambiguities due to accidental chemical shift overlap for both ^{1}H (C−H) and ^{13}C have been indicated by *i*, *ii*, *iii*, *iv* and *v*.

in the supercooled ^{1}H NMR spectrum of Figure 2, corresponding to the ten hydroxyl protons that are expected from the two anomers of glucose (i.e., 1-OH, 2-OH, 3-OH, 4-OH, and 6-OH for both α- and β-forms), with just one "double" overlap at δ_H 6.48 ppm for 3-OH(α) and 4-OH(α). The resolution for these −OH groups is significantly superior to the C−H protons between δ_H 3.0 and 5.5 ppm, for which only six resonances can be clearly identified,[c] with the remaining eight C−H protons appearing as four pairs of overlapped (sometimes strongly second-order) multiplets (i.e., H-5(α)/H-6(α), H-3(α)/H-6(β), H-3(β)/H-5(β), and H-4(α)/H-4(β) in Figure 2). Complete resolution of all 10 −OH groups in glucose can be achieved by introducing a second dimension, via the 2D-NMR experiment HSQC-TOCSY.[40,41] As shown in Figure 3, the ^{1}H resonance of every −OH group becomes correlated with the ^{13}C chemical shift of the carbon to which that −OH group is attached, resulting in 10 clearly distinguished 2D C−OH peaks. This 2D experiment confirms that the 3-OH(α) and 4-OH(α) resonances were accidentally overlapped at δ_H 6.48 ppm in the 1D ^{1}H NMR spectrum of Figure 2 and all 10 peaks have been represented within the 10 "jigsaw pieces" in Figure 3. It is these jigsaw pieces that will constitute the starting point for the enhanced approach to structural elucidation of carbohydrates, which is discussed in the next subsection.

The HSQC-TOCSY experiment is able to "edit" a ^{1}H NMR spectrum by sequentially relaying coherence from a C−H unit to all the other protons in a mutually coupled network. Glucose consists of two giant coupled networks (one for each anomer), comprising all C−H and −OH protons in the molecule.[42] The TOCSY mixing time[41] employed in Figure 3 has been set to a relatively short value (8 ms) to relay coherence only between hydrogens that are bonded to adjacent nuclei but no further along the network. This is the optimum condition for identifying the five H−C−O−H systems that reside within each anomer of glucose. In each case, coherence is being relayed from the C−H proton (δ_H 3−5.5 ppm, which is not shown in the expansion of Figure 3, but see Figure 4) to its associated −OH group (δ_H 5.5−8.5 ppm in the expansion of Figure 3) at the edited ^{13}C chemical shift of the carbon of the C−H unit.

Inclusion of −OH Groups in the Jigsaw Pieces Used for Chemical Structure Determination by NMR. As explained in the Introduction, the standard 2D-NMR approach to solving the structure of an unknown organic compound begins with an HSQC experiment,[2−4] which identifies all directly bonded ^{1}H and ^{13}C resonances via their large one-bond coupling constant ($^1J_{CH}$), and thereby defines all the CH, CH_2, and CH_3 pieces of a jigsaw (so called because, once connected to one another in the correct fashion, they define the planar covalent structure of the entire molecule). The

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

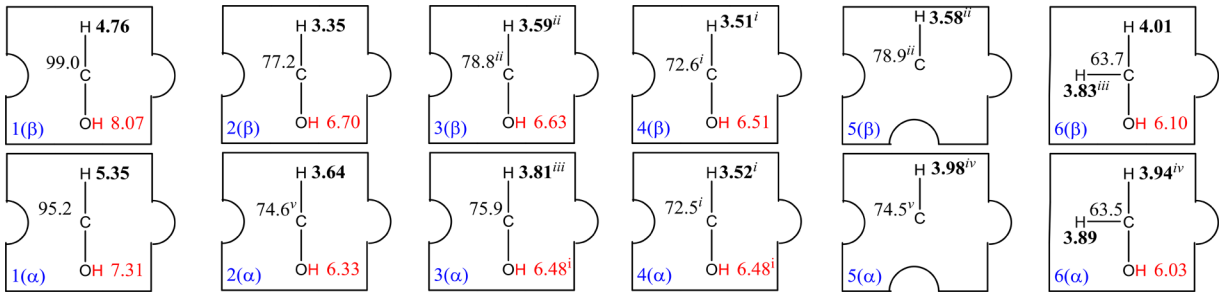


Figure 5. The 12 H−C−OH jigsaw pieces used in the enhanced NMR approach for determining the chemical structures of β- and α-glucose, which are treated as unknowns in the next section. Accidental chemical shift overlaps for ^{13}C and ^{1}H in a C−H unit are indicated by *i*, *ii*, *iii*, *iv* and *v*; accidental overlap within −OH is indicated by i.

HSQC spectrum of supercooled glucose (Figure 4) illustrates the creation of such pieces in the traditional strategy for which each C−H proton is correlated with the carbon to which it is directly attached.

The HSQC experiment can be thought of as editing the C−H protons by the ^{13}C chemical shift of their attached carbons, in much the same way that HSQC-TOCSY edits the −OH groups, as described above. Note, however, that the introduction of the second dimension is insufficient to overcome all instances of accidental overlap between the 14 C−H resonances, for which both the 3(β)/5(β) pieces and the 4(α)/4(β) pieces remain unresolved, as indicated by the annotation "Double C-H overlap: 4" in Figure 4. This stands in contrast to Figure 3, in which all 10 −OH protons have been resolved from one another.

Because the carbon chemical shifts for the C−H units in Figure 4 are the same as for the C−OH units in Figure 3, it is obviously possible to combine these two sets to create a new jigsaw unit, which is referred to as a H−C−OH piece in Figure 5 (six for each anomer). The advantage of doing this is that there are now two[d] ^{1}H chemical shifts available for each of the new pieces, and hence, whenever the proton spectrum suffers from poor dispersion, the shift that is better resolved can be selected for further analysis. For example, in the case of glucose, it is clear that the −OH resonances are better resolved than the C−H resonances for both anomers, as has been indicated directly in Figures 2−4. The more abundant β-anomer of glucose has been chosen in the next section to illustrate the enhanced approach for structure elucidation and accompanying NMR assignment, in which these H−C−OH jigsaw pieces are assembled by making connections between −OH chemical shifts in preference to C−H groups, as practiced in the standard strategy.

Using −OH Groups To Assemble the NMR Jigsaw for the β-Anomer. In the traditional NMR approach to structure elucidation, C−H jigsaw pieces must be assembled into the full molecular structure using other 2D-NMR techniques,[2−4] which make connections between C−H units over more than one bond. The two most important such long-range experiments are HMBC,[43] which relies on $^{2}J_{CH}$ and $^{3}J_{CH}$ couplings, and $^{1}H-^{1}H$ COSY,[44] which primarily identifies $^{3}J_{HH}$ couplings. In the alternative approach described herein, the mixing time of the HSQC-TOCSY experiment is made progressively longer to achieve the same purpose. Increasing the HSQC-TOCSY mixing time increases the maximum number of bonds over which coherence can be relayed between hydrogen atoms in a coupled network. This is illustrated in Figures 6, 7, and 8 for mixing times of 25, 45, and 60 ms, which are sufficient to transfer coherence between up to three, four, and five hydrogen atoms, respectively. This results in the appearance of several new 2D peaks, in addition to those already described in Figure 3, and the new correlations appearing in the spectral region δ_H 5.5−8.5 ppm connect each of the H−C−OH pieces to the other −OH groups.

In this section, we use these correlations to piece together the structure of glucose, as if it were a carbohydrate of unknown composition, to illustrate the potential of the enhanced approach. The anomeric C−H protons are generally selected as the starting point for commencing analysis of a carbohydrate in the standard strategy because they are few in number (one per monosaccharide) and their chemical shift range (δ_H 4.5−5.5 ppm) does not overlap with other methine and methylene protons (δ_H 3−4.5 and 3−4 ppm, respectively). The same is true for the anomeric −OH protons in the alternative approach, and the resonance at δ_H 8.07 ppm can be confidently identified with the anomeric position of the more abundant (β) anomer in Figures 2, 3, and 5. [Note that the carbon chemical shift at the anomeric position of a carbohydrate (δ_C 85−110 ppm) is also consistently higher than the rest of the molecule (δ_C 60−85 ppm).]

The horizontal arrow A that connects the 1(β) anomeric piece (Figure 3) with the new correlation for 2-OH(β)[e] at the edited ^{13}C chemical shift of C-1(β) in Figure 6 then defines a direct connection between the 1(β) and the 2(β) piece, which must therefore be located adjacent to the anomeric center, as shown by the topmost "jigsaw assembly" in Figure 6. New peak B, at the chemical shift of 1-OH(β) and C-2(β), confirms the connectivity between these same two −OH groups, with coherence being relayed in the opposite direction. Similarly, the two "reciprocal" correlations between the new peaks labeled C and D indicate that the 2(β) piece should, in turn, be connected with the 3(β) piece. A fifth correlation labeled E could represent a further ongoing connection from the 3(β) piece to 4(β), but this is a tentative conclusion (as indicated by the dotted lines for 4(β) in the bottom jigsaw assembly in Figure 6) because the carbon chemical shifts of C-3(β) and C-5(β) are accidentally overlapped (Figures 4 and 5). Hence, correlation E could equally well represent a connection between the 5(β) and 4(β) pieces. The decision would rest on the observation of a reciprocal correlation from the 4(β) piece to 3-OH(β), which is unfortunately absent[f] from this particular spectrum, although it does make an appearance in Figure 7 at the longer mixing time of 45 ms.

In conclusion then, the effect of increasing the HSQC-TOCSY mixing time from 8 to 25 ms is to link the three H−C−OH jigsaw pieces 1(β), 2(β), and 3(β) unambiguously into

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

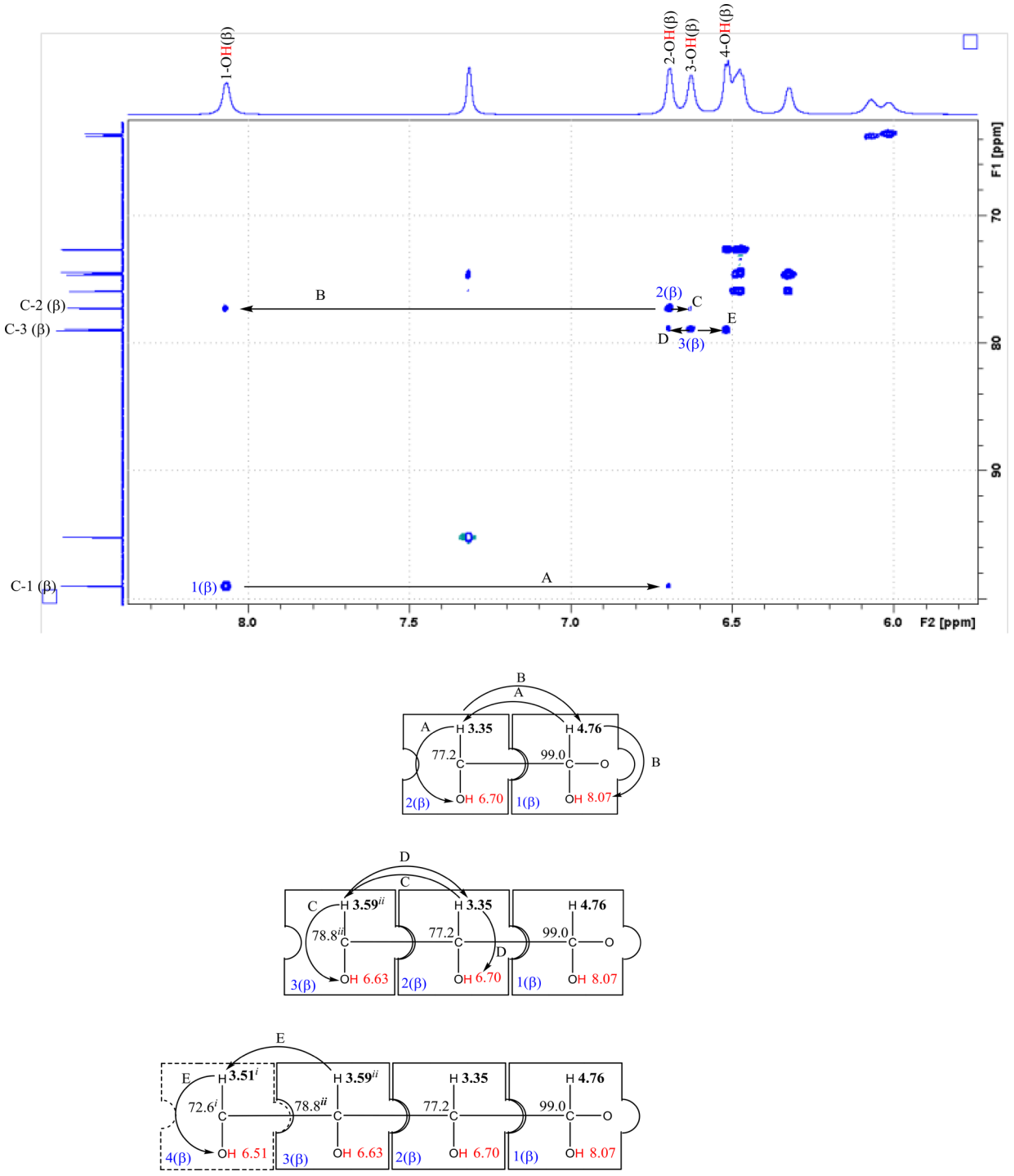


Figure 6. Expansion (δ_H 5.5−8.5 ppm in F2) of the HSQC-TOCSY spectrum of 100 mM glucose in 1% D_2O/99% H_2O supercooled to −14 °C. At the longer mixing time of 25 ms, coherence is relayed between up to three hydrogen atoms, forming a coupled network. This allows −OH chemical shifts to be linked by horizontal lines, thereby connecting together adjacent H−C−OH pieces of the NMR jigsaw.

a larger piece, based on the four correlations, labeled A−D in Figure 6. No further TOCSY correlations are expected for the right-hand side of this large piece, which represents the anomeric carbon, with a divalent oxygen atom for the unresolved linkage. The 4(β) piece has been tentatively attached to the left-hand side (dotted lines), in anticipation of the results of HSQC-TOCSY experiments with longer mixing times that produce additional unambiguous correlations, as described below.

Further extension of the mixing time to 45 ms allows coherence transfers to be made between as many as four hydrogen atoms in a coupled network. Thus, the three new peaks that appear in Figure 7, which have been connected by arrows labeled A′, C′, and D′, serve to confirm the full structure already proposed above for the larger piece in Figure 6. Arrows A′ and D′, which are the continuations of arrows A and D in Figure 6, form a pair of reciprocal correlations that represent connections between 1(β) and 3(β) separated by an intervening piece (topmost jigsaw in Figure 7). The arrow

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

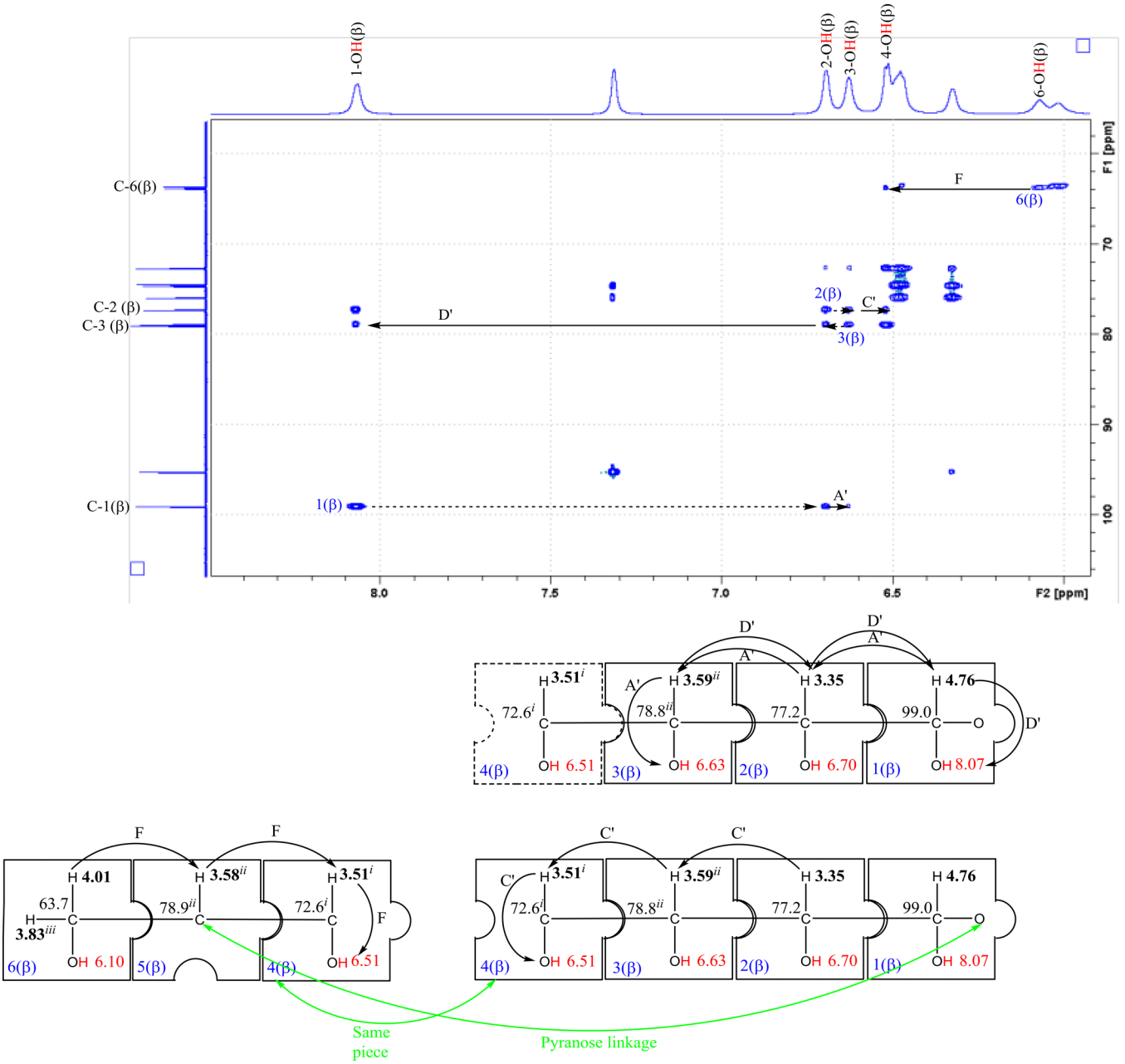


Figure 7. Expansion (δ_H 5.5−8.5 ppm in F2) of the HSQC-TOCSY spectrum of glucose in 1% D_2O/99% H_2O supercooled to −14 °C at an even longer mixing time of 45 ms. Coherence is now relayed between up to four hydrogens, forming a coupled network. This allows −OH chemical shifts to be linked by horizontal lines, thereby connecting together H−C−OH units separated by one intervening piece. It is not known which of the 6-CH_2 protons initiates correlation F, and the selection of δ_H 4.01 over δ_H 3.83 ppm is arbitrary.[45]

labeled as C′ represents a similar "skipped" connection from 2(β) to 4(β), which confirms that piece 4(β) should indeed have been included at the left-hand side of the large piece in Figure 6 (now represented by solid lines in Figure 7). The ambiguity for this connection has disappeared because neither the ^{13}C chemical shift for piece 2(β) nor the −OH chemical shift for 4-OH(β) suffers from accidental chemical shift overlap. Most importantly, however, this longer mixing time provides the means by which the (6β) CH_2−OH piece from Figure 5 can also be included in the NMR analysis via the correlation labeled F, which connects 4-OH(β) with 6-OH(β) at the ^{13}C chemical shift of C-6(β), as shown at the bottom left of Figure 7. Note that this final connection can be made even though there is no −OH group in the intervening 5(β) piece because coherence is still being relayed (over four hydrogen atoms) through the C−H proton at the 5-position.

The observation of this last connection completely defines the carbon backbone of the unknown because the two large pieces established in Figure 7 share the 4β unit in common. All that remains is to connect the two unsatisfied valencies between the 1(β) and 5(β) pieces to define the pyranose form of an aldohexose sugar, for which 5-OH has performed an intramolecular nucleophilic attack at C-1 to create a hemiacetal functional group at the anomeric position of this monosaccharide. The logical process that resulted in this covalent structure also generated the accompanying set of full NMR assignments that are shown in Figure 9. These could be used to fully identify the molecule as β-glucopyranose, based on

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

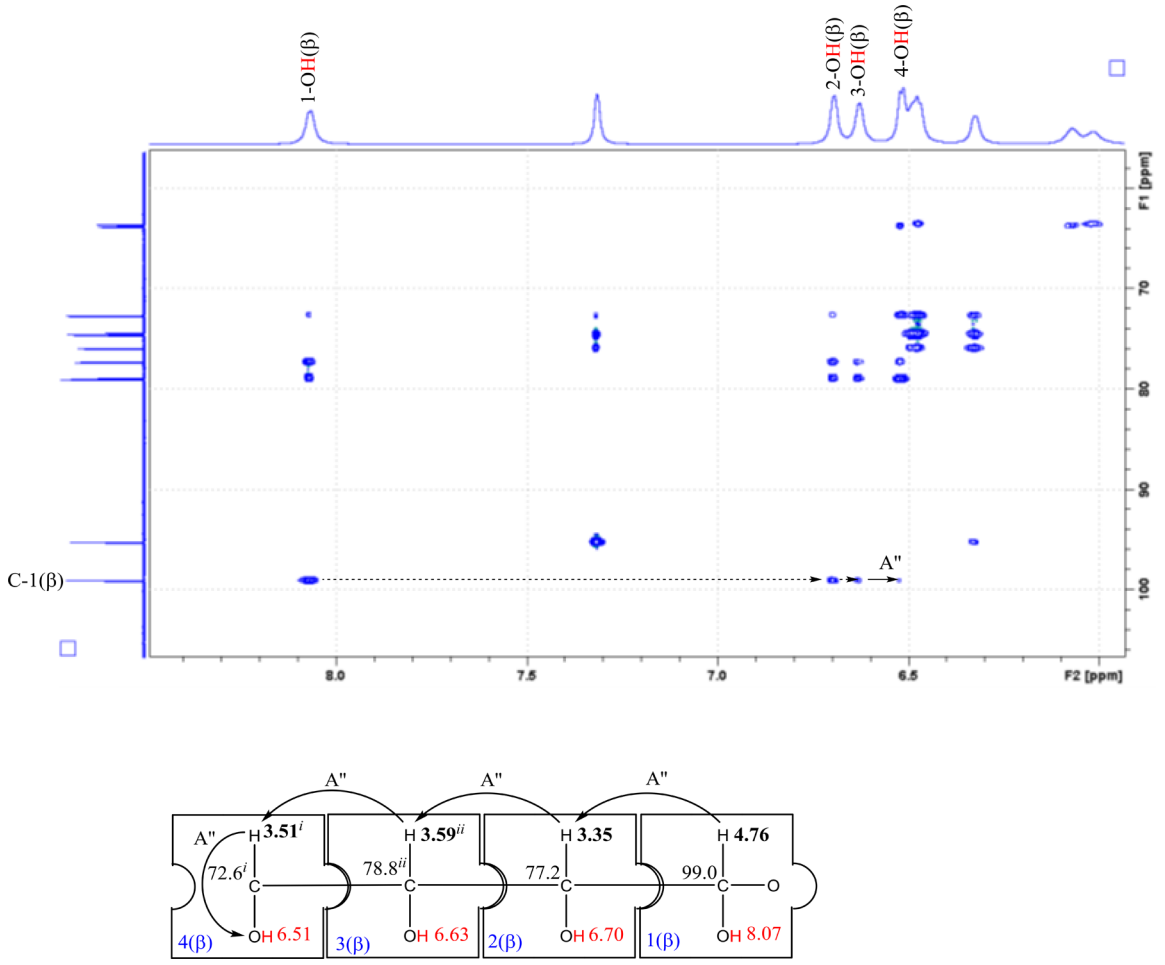


Figure 8. ^{1}H NMR assignments of the 10 −OH resonances (red) and 14 C−H resonances (bold) for the α- and β-anomers of glucose (100 mM) in a supercooled solution (−14 °C) in 1% D_2O/99% H_2O, as determined by the enhanced NMR approach. ^{13}C resonances common to these two sets of protons are shown below the corresponding ^{1}H assignments for the C−H groups. Accidental chemical shift overlaps for ^{13}C and ^{1}H resonances in C−H units are indicated by *i*, *ii*, *iii*, *iv*, and *v*; accidental chemical shift overlap for −OH resonances used in the alternative approach has been indicated by i.

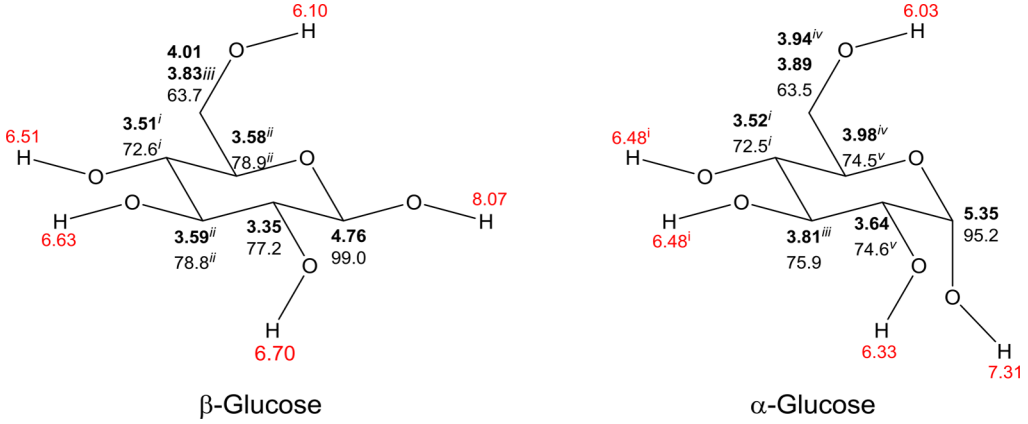


Figure 9. Expansion (δ_H 5.5−8.5 ppm in F2) of the HSQC-TOCSY spectrum of 100 mM glucose in 1% D_2O/99% H_2O supercooled to −14 °C at the very long mixing time of 60 ms. Coherence is now relayed between up to five hydrogen atoms in a coupled network.

additional NOESY data and/or an analysis of $^{3}J_{HH}$ coupling constants. For example, the coupling constants for H-1(β) (δ_H 4.76 ppm, d) and H-2(β) (δ_H 3.35 ppm, dd)—the only two ring protons to be resolved for the β-anomer in Figure 2—are all approximately 9 Hz. Hence, H-1(β), H-2(β), and H-3(β) must all be axial in the chair conformation shown in Figure 9.

Further extension of the HSQC-TOCSY mixing time results in the establishment of successively more distant linkages between H−C−OH pieces, such as that revealed by correlation A″ at a very long mixing time of 60 ms in Figure 8.

At these extremely long mixing times, nearly all[g] the −OH groups in a coupled network can be connected by a series of

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

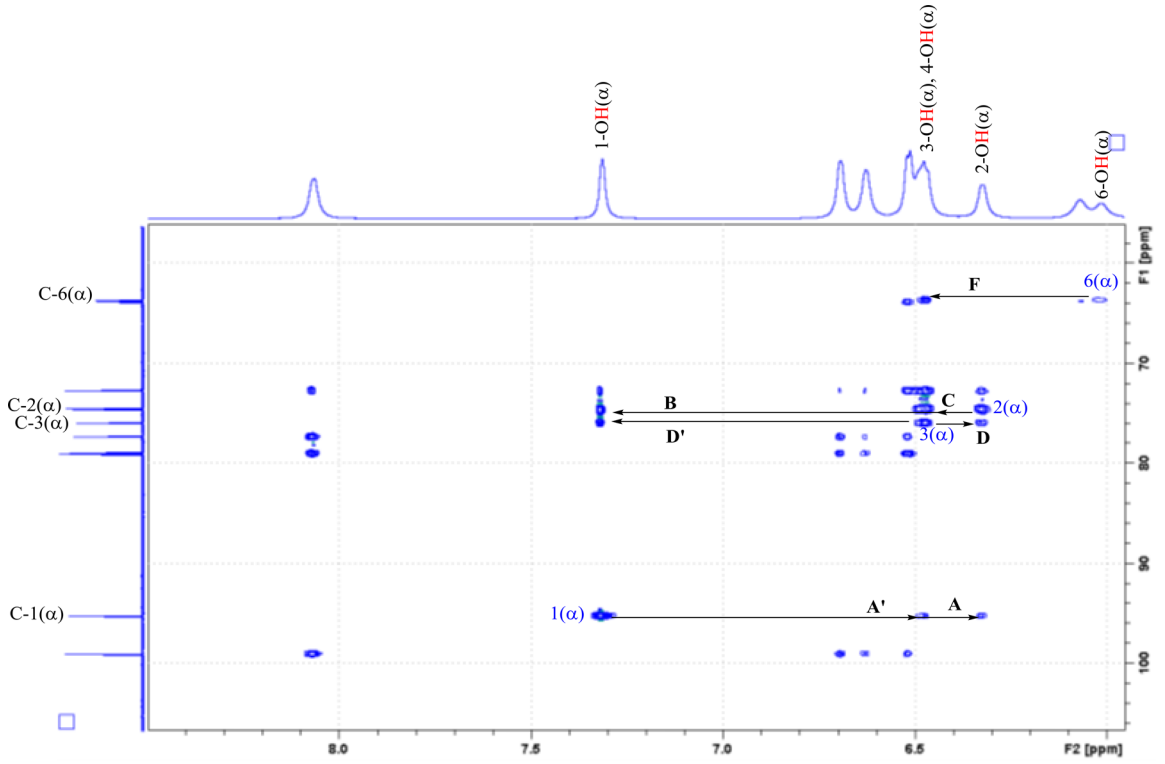


Figure 10. Expansion (δ_H 5.5−8.5 ppm in F2) of the HSQC-TOCSY spectrum of 100 mM glucose in 1% D_2O/99% H_2O supercooled to −14 °C at the extremely long mixing time of 90 ms.

horizontal arrows for the carbon chemical shift at which coherence transfer was initiated. This is illustrated in Figure 10 for the α-anomer of glucose at a mixing time of 90 ms, in which bold labels **B**, **C**, and **D** (which first appeared at a mixing time of 25 ms); **A**, **D′**, and **F** (45 ms); and **A′** (90 ms) are the equivalents of those already described for β-glucopyranose above. Such spectra provide a simple means for assessing whether or not more than one coupled network is present; in this particular case, it is obvious that there are two sets of ^{13}C chemical shifts, one for the β- and one for the α-anomer of glucose. Note that these two sets overlap at C-4, which has therefore not been included in any of the above analyses to avoid introducing an unnecessary further ambiguity to the structure elucidation/assignment process.

A Comparative Analysis of the Use of CH Resonances To Assemble the NMR Jigsaw for Glucose. It was anticipated in an earlier section that the advantage of having constructed H−C−OH pieces with two ^1H chemical shifts per piece would be fully realized when analyzing long-range HSCQ-TOCSY spectra, because it would be possible to select from one of two chemical shifts, whichever is better resolved, when attempting to connect two pieces together. The H−C−OH pieces allow connections to be made between −OH groups (via the alternative approach described above) or between C−H groups (traditional strategy), and we can readily alternate between the two strategies, as and when accidental overlap leads to ambiguities in either set of protons. This is illustrated explicitly in Figure 11 by the expansion of the C−H region of the HSQC-TOCSY spectrum of glucose at a mixing time of 8 ms, which has been taken from the same spectrum that is shown in Figure 3, as an expansion of the −OH region (see also Figure S2). It will be recalled that a mixing time of 8 ms is sufficient to relay coherence between two hydrogens that are bonded to adjacent nuclei; hence, the expansion of Figure 11 identifies relays within H−C−C−H units rather than within H−C−O−H units, as in Figure 3. The new correlations[h] that appear in the C−H expansion of the HSQC-TOCSY spectrum in Figure 11 therefore provide information about the linkages between adjacent C−H pieces. If glucose is, once again, treated as an unknown, then the interpretation of this half of the spectrum is more akin to the first stage of the jigsaw assembly process between −OH groups that is depicted in Figure 6.

Thus, the horizontal arrow labeled a in Figure 11, which connects the 1(β) piece for the anomeric center with the new correlation for H-2(β) at the edited ^{13}C chemical shift of C-1(β), mirrors correlation A in Figure 6. The reciprocal arrow b confirms this same connectivity with coherence being relayed in the opposite direction, just like connection B in Figure 6. However, ambiguity immediately sets in for arrows c and d, the reciprocal correlations that connect the 2(β) piece to the 3(β) piece (shown by dotted lines), because both the proton and carbon chemical shifts for H-3(β) and H-5(β) are accidentally overlapped (Figure 5). Note that this ambiguity was avoided in the −OH approach because there is no accidental overlap for 3-OH(β). The situation becomes even more challenging for correlation e, which is doubly ambiguous because the chemical shift of H-4(β) is also overlapped with H-4(α). Again, this was not an issue in the −OH approach, for which 4-OH(β) has a unique chemical shift. Most of the ambiguities that are caused by the accidental overlap of C−H proton chemical shifts are resolved as the mixing time of the HSQC-TOCSY experiment is further extended (see Figures S5−S9, right-hand side). However, on balance, it is the −OH strategy described in the previous section that offers the more straightforward approach to the structural elucidation of β-glucopyranose, when this molecule is treated as an unknown.

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

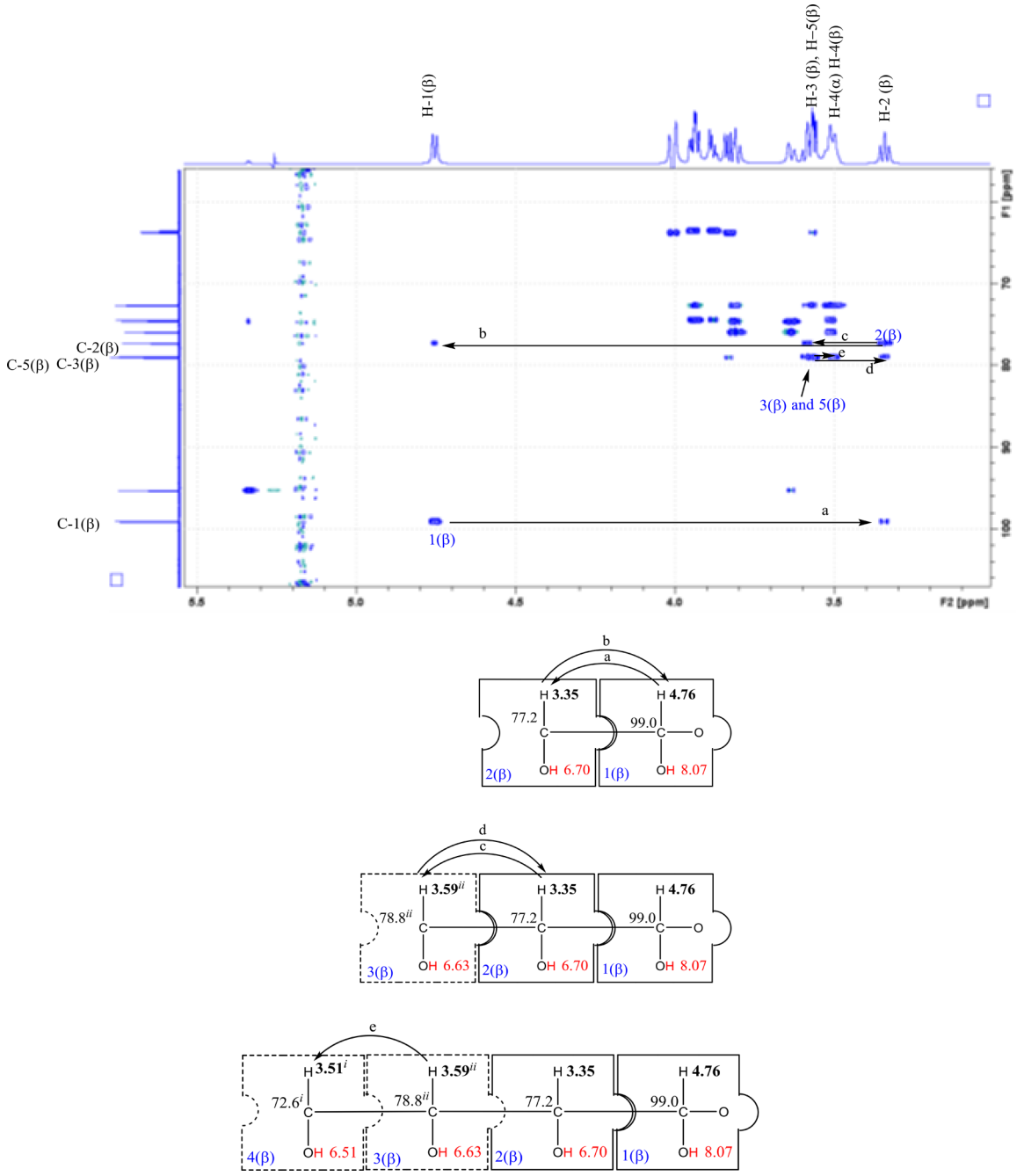


Figure 11. Expansion (δ_H 3−5.5 ppm in F2) of the HSQC-TOCSY spectrum of 100 mM glucose in 1% D_2O/99% H_2O supercooled to −14 °C. At the mixing time of 8 ms, coherence is relayed between just two hydrogen atoms in a coupled network, and C−H chemical shifts linked by horizontal lines connect together adjacent C−H pieces of the NMR jigsaw.

The more rapid onset of ambiguity for the conventional C−H strategy could have been predicted from the extent of accidental chemical shift overlap for C−H protons compared with −OH protons, as indicated directly on the 1D-NMR spectrum of Figure 2 and the 2D-NMR spectra of Figures 3 and 4. The reason that the alternative −OH approach is superior to the traditional strategy for glucose is that there is less overlap between −OH protons than between C−H protons in both 1D and 2D. The standard strategy remains essential for explicitly identifying pieces, such as 5(β), that do not contain an −OH group (Figure 5) which can only be inferred by the −OH approach (e.g., bottom left of Figure 7). It is also very useful whenever there is accidental overlap between −OH resonances. For example, 3-OH(α) and 4-OH(α) are overlapped in the α-anomer of glucose, whereas H-3(α) and H-4(α) are resolved from one another.

To conclude, the creation of H−C−OH pieces supplements the spectral region for C−H protons (δ_H 3−5.5 ppm) by an additional region for OH protons (δ_H 5.5−8.5 ppm) and thereby doubles the dispersion of the ^1H chemical shifts that is available for chemical structure elucidation. This increased dispersion is best utilized by alternating between the −OH and C−H strategies for determination of covalent structure described above to overcome challenges resulting from accidental overlap, whenever it arises in either O−H or C−H chemical shifts.

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

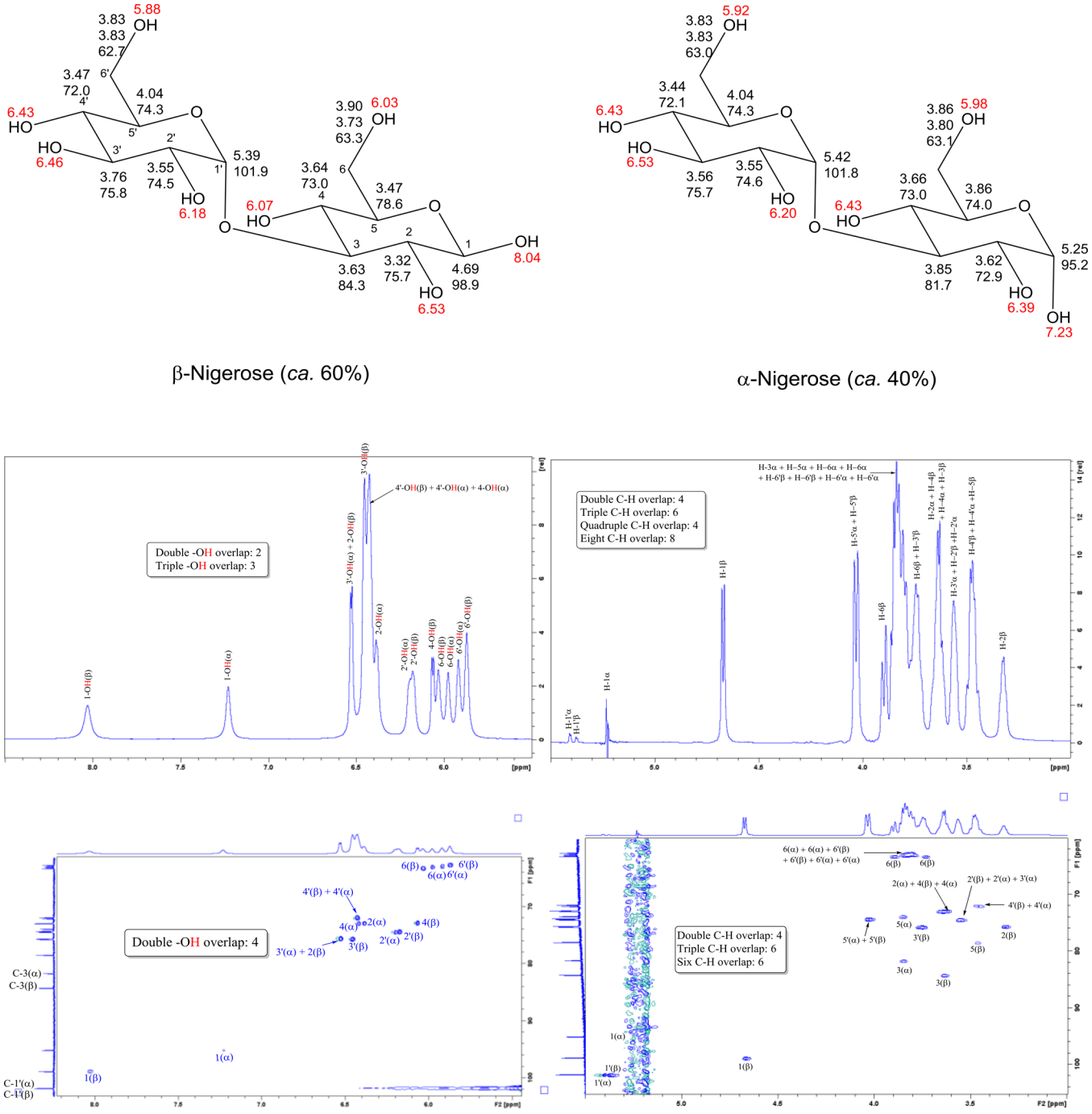


Figure 12. ^{1}H NMR assignments of −OH resonances (red) and C−H resonances (bold) for supercooled solutions of the α- and β-anomers for nigerose (Glcα1-3Glc). ^{13}C resonances that are shared by these two sets of protons are shown below the corresponding ^{1}H assignments for the C−H groups. Assignments were deduced from 1D NMR (top left is an expansion of the −OH region and top right is an expansion of the C−H region from the ^{1}H NMR spectrum) and 2D NMR [bottom left is an expansion of the −OH region from the HSQC-TOCSY spectrum (mixing time 8 ms) and bottom right is an expansion of the HSQC spectrum].

Nigerose and Gentiobiose. The usefulness of the alternative approach has been further exemplified below by its application to the complete ^{1}H and ^{13}C NMR assignments of more complex systems, such as the four commercially available disaccharides: nigerose (Glcα1-3Glc), β-gentiobiose (Glcβ1-6Glc), leucrose (Glcα1-5Fru), and isomaltulose (Glcα1-6Fru; palatinose). All these disaccharides incorporate glucose, and there is good agreement between the chemical shifts of −OH groups for glucose itself (Figure 9) and these various disaccharides which contain it (Figures 12−15).

Nigerose and β-gentiobiose (Sections S2.2 and S2.3) both exist as a mixture of anomers, just like the monosaccharide glucose. As for glucose, accidental overlap is found to be far less severe for the −OH groups than for the C−H groups in both 1D and 2D NMR, and this is illustrated directly on the spectra in Figures 12 and 13. Thus, there are 16 −OH groups in the two anomers of nigerose, 3 of which are triply overlapped at δ_H 6.43 ppm [4′-OH(β), 4′-OH(α), and 4-OH(α)] and 2 of which are doubly overlapped at δ_H 6.53 ppm [3′-OH(α) and 2-OH(β)], leaving 11 −OH groups that are

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

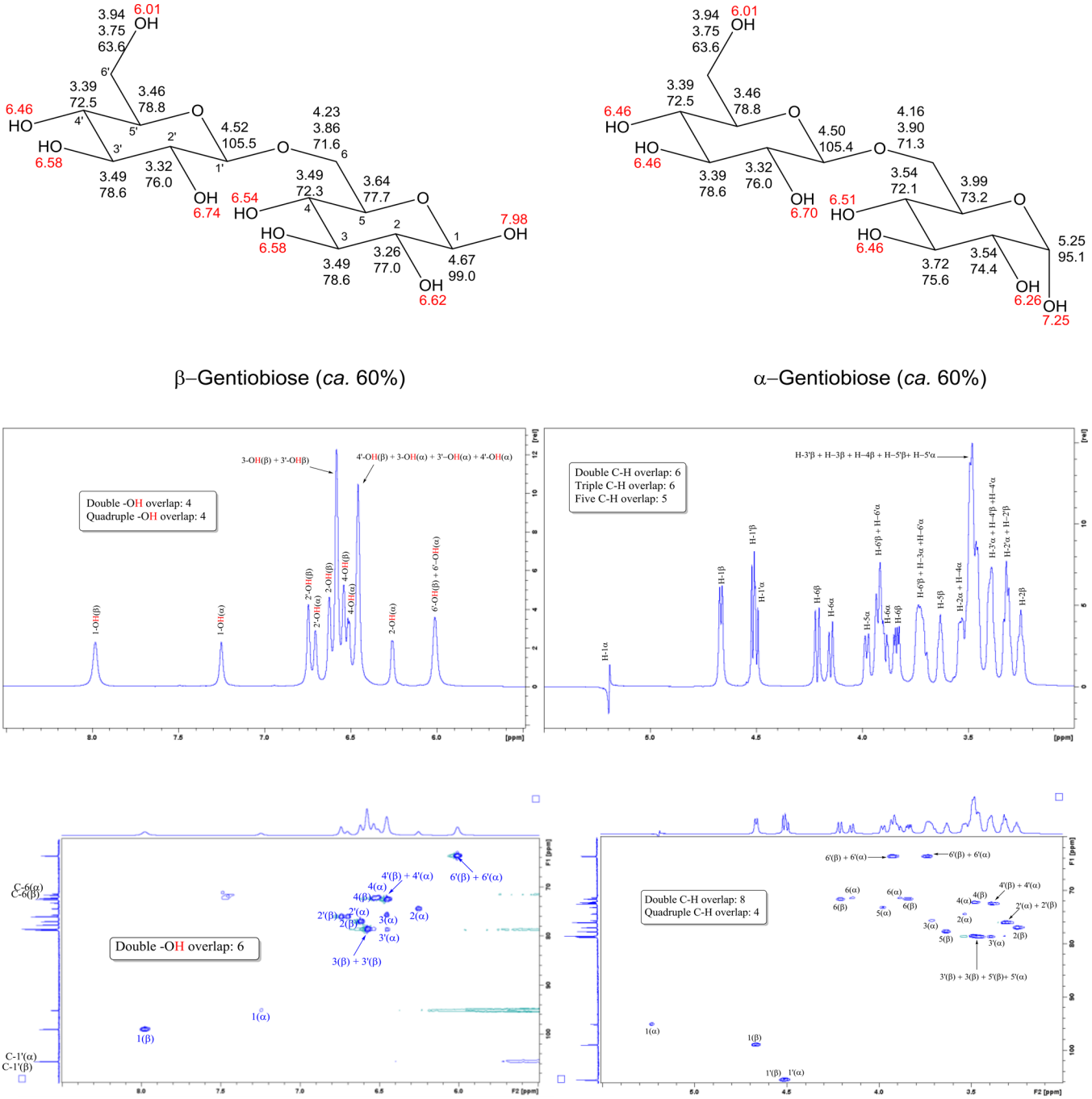


Figure 13. ^{1}H NMR assignments of −OH resonances (red) and C−H resonances (bold) for supercooled solutions of the α- and β-anomers for β-gentiobiose (Glcβ1-6Glc). ^{13}C resonances that are shared by these two sets of protons are shown below the corresponding ^{1}H assignments for the C−H groups. All assignments were deduced from 1D NMR (top left is an expansion of the −OH region and top right is an expansion of the C−H region from the ^{1}H NMR spectrum) and 2D NMR [bottom left is an expansion of the −OH region from the HSQC-TOCSY spectrum (mixing time 8 ms) and bottom right is an expansion of the C−H region of the HSQC spectrum].

resolved from one another. This compares with a complete resolution for just 6 out of 28 C−H resonances for the same two anomers, in which no less than 8 resonances are crammed into the region δ_H 3.80−3.86 ppm [H-3(α), H-5(α), H-6(α), H-6(α), H-6′(β), H-6′(β), H-6′(α), and H-6′(α)]. The situation does not improve greatly after introducing a second dimension, as shown by the HSQC spectrum of Figure 12, in which more than half of these C−H resonances remain overlapped. The C−H overlap for gentiobiose (Figure 13) is marginally less poor, but there can be little doubt that it is easier to elucidate the structures of both of these disaccharides using the alternative −OH approach in preference to the traditional C−H strategy.

Furthermore, it is immediately apparent that 8 out of 24 carbons are not involved in correlations in the −OH region of the HSQC-TOCSY spectra for either nigerose or gentiobiose (bottom left Figures 12 and 13). There are no quaternary carbons in either disaccharide, so there must be eight positions in the molecule that do not bear an −OH group.[46] In the case of nigerose, four of these carbon atoms are associated with the

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

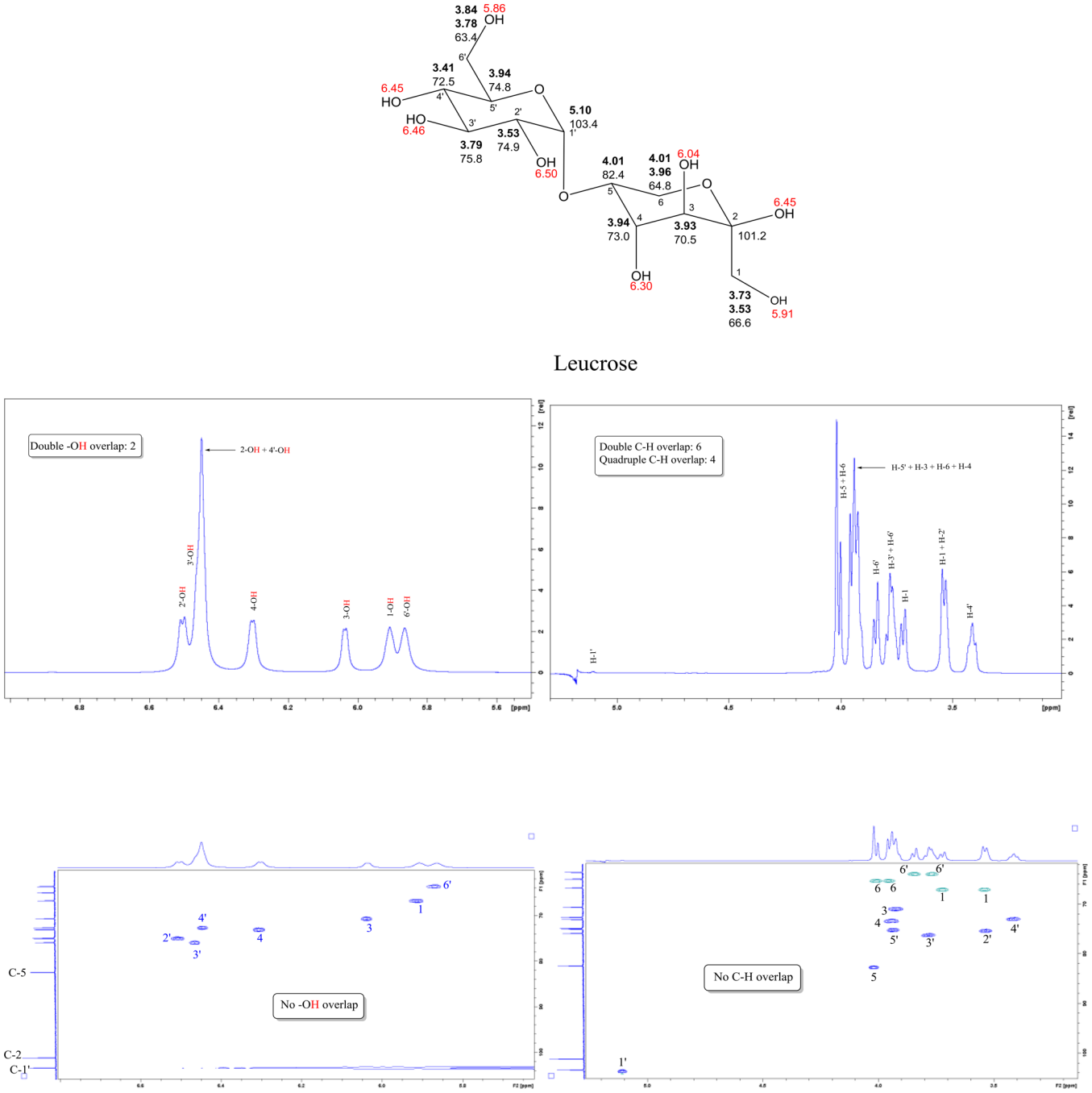


Figure 14. ^{1}H NMR assignments of −OH resonances (red) and C−H resonances (bold) for supercooled solutions of leucrose (Glcα1-5Fru). ^{13}C resonances that are shared by these two sets of protons are shown below the corresponding ^{1}H assignments for the C−H groups. All assignments were deduced by 1D NMR (top left is an expansion of the −OH region and top right is an expansion of the C−H region from the ^{1}H NMR spectrum) and 2D NMR [bottom left is an expansion of the −OH region from the HSQC-TOCSY spectrum (mixing time 8 ms) and bottom right is an expansion of the edited-HSQC spectrum].

internal oxygens of the two pyranose rings (δ_C 78.6 (C-5(β))/74.0 (C-5(α)) and δ_C 74.3 (C-5′(β))/74.3 (C-5′(α)), which have similar chemical shifts to those of glucopyranose in the previous section. The remaining four carbons, which appear as annotations on the ^{13}C projection of the HSQC-TOCSY spectrum in Figure 12, must be associated with the glycosidic linkage. Two of these carbons at δ_C 101.9 and 101.8 ppm are immediately identifiable as C-1′(β) and C-1′(α) by their distinctive chemical shifts (anomeric carbons appear in the range δ_C 85−115 ppm). The point of attachment for the other half of the disaccharide can be established once the remaining two non-hydroxyl-bearing carbons at δ_C 84.3 and 81.7 ppm are identified as C-3(β) and C-3(α) by application of the jigsaw assembly approach (it is the C−H strategy that is required for these units).

The 1,6-linkage in gentiobiose can be constructed by applying similar arguments to the HSQC-TOCSY spectrum shown in Figure 13, for which the absence of correlations to −OH protons at the four methine carbons at δ_C 71.6 (C-6(β))/71.3 (C-6(α)) and δ_C 105.5 (C-1′(β))/105.4 (C-1′(α)

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

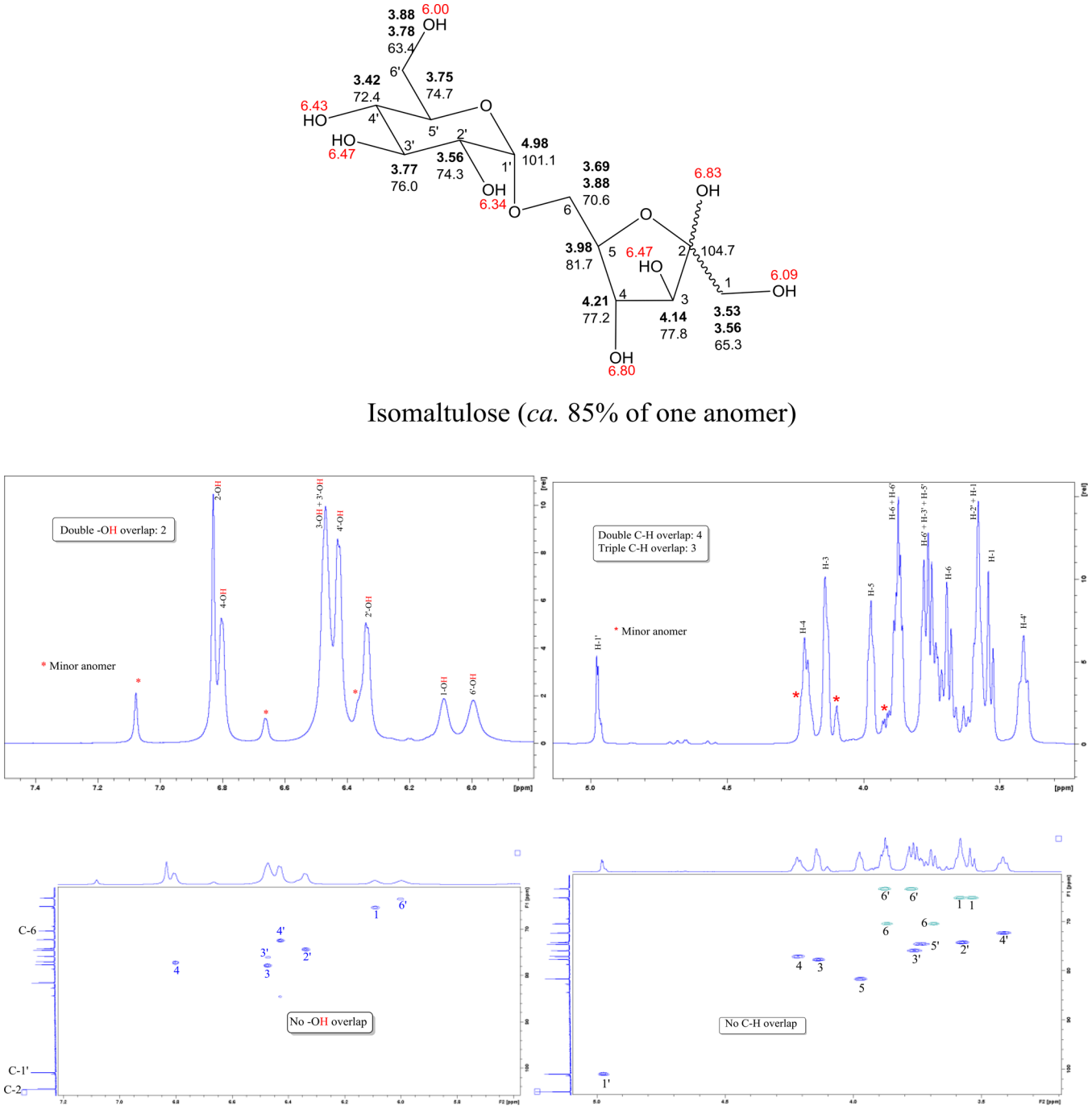


Figure 15. ^{1}H NMR assignments of −OH resonances (red) and C−H resonances (bold) for supercooled solutions of isomaltulose (Glcα1-6Fru). ^{13}C resonances that are shared by these two sets of protons are shown below the corresponding ^{1}H assignments for the C−H groups. All assignments were deduced from 1D NMR (top left is an expansion of the −OH region and top right is an expansion of the C−H region from the ^{1}H NMR spectrum) and 2D NMR [bottom left is an expansion of the −OH region from the HSQC-TOCSY spectrum (mixing time 8 ms) and bottom right is an expansion of the edited-HSQC spectrum].

are associated with the glycosidic linkage, with the other four (δ_C 77.7 (C-5(β))/73.2 (C-5(α)) and δ_C 78.8 (C-5′(β))/78.8 (C-5′(α)) being associated with the internal oxygen of the pyranose rings in each monosaccharide.

Leucrose and Isomaltulose. The disaccharides leucrose and isomaltulose (Sections S2.4 and S2.5) are both composed of glucose and fructose and have been included to illustrate a second kind of "missing" hydroxyl group that can be encountered in the −OH region of the HSQC-TOCSY spectrum. Once again, the extent of accidental overlap is less for the −OH groups than for the C−H groups in 1D NMR of both disaccharides (as noted directly on the ^{1}H NMR spectra in Figures 14 and 15) and is quite comparable to that of α- and β-glucose in glucose itself. The −OH and C−H regions are completely free from accidental overlap in all the 2D-NMR spectra of Figures 14 and 15—most probably because both disaccharides exist effectively in just one anomeric form.

The ^{1}H NMR spectrum of leucrose (Figure 14) presents eight −OH groups in the range δ_H 5.8−6.6 ppm. It might be inferred from this somewhat limited dispersion that the

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

glycosidic linkage must be at the C-1′ position of glucose because no −OH resonances are present with the high chemical shift that is characteristic of the anomeric hydroxyl group of an aldose sugar [e.g., δ_H 8.07 and 7.31 ppm for 1-OH(β) and 1-OH(α) in the glucose spectrum of Figure 2]. This inference is readily confirmed by the enhanced NMR approach, for which there are five carbons that do not show any correlations in the −OH region of the HSQC-TOCSY spectrum in Figure 14 (δ_C 103.4, 101.2, 82.4, 74.8, and 64.8 ppm), with the highest chemical shift corresponding to the glycosidic linkage at the C-1′ anomeric carbon of glucose.

More interestingly, it is observed that one of the eight −OH groups in leucrose (2-OH) also does not show any correlation in 2D HSCQ-TOCSY spectra (or, in fact, in any of the HSQC-TOCSY spectra in Figures S40−S44). The missing eighth −OH group must therefore be due to a hydroxyl proton that is not part of a coupled network; that is, it is attached to a carbon that lacks a C−H hydrogen in the H−C−OH piece from which coherence could have been relayed. The quaternary carbon of the hemiketal group in the ketose moiety of leucrose is thereby identified by the absence of correlations in HSQC-TOCSY for both the carbon at C-2 (δ_C 101.2 ppm) and its attached 2-OH group (δ_H 6.45 ppm).[i]

A similar argument can be made to identify the C-2 position in isomaltulose for which neither the 2-OH group (δ_H 6.83 ppm) nor its associated quaternary carbon (δ_C 104.7 ppm) appear at any HSQC-TOCSY mixing time (Figure 15 and Figures S50−S54).[i] Of the four remaining carbons that do not bear an −OH group (δ_C 101.1, 81.7, 74.7, and 70.6 ppm) in the HSQC-TOCSY spectrum of Figure 15, one (δ_C 70.6 ppm) is easily identified as a CH_2O piece by two negative cross-peaks (colored green rather than blue) in the corresponding edited-HSQC spectrum shown. The chemical shift of this CH_2O piece is 6−7 ppm higher than all other examples that appear in this manuscript with the exception of C-6 in gentiobiose and is therefore consistent with a glycosidic linkage between the 6-position of fructose and C-1′ of glucose (δ_C 101.1 ppm), which forces the fructose moiety in isomaltulose to adopt the furanose form.

CONCLUSIONS

With this study, it has proved possible to reduce the rate of exchange of the −OH protons of carbohydrates, which is fast on the NMR timescale at ambient temperature, to an appropriate level to determine the chemical structure of a carbohydrate in aqueous solution, in a reasonable time frame by the application of five 2D-NMR experiments: HSQC and HSQC-TOCSY, repeated at four different mixing times (8, 25, 45, and 60 ms). We have shown that, for a high-field 700 MHz NMR instrument equipped with a cryogenic probe, excellent results can be obtained from the pivotal HSQC-TOCSY experiment when studying a 100 mM solution in 1% D_2O/99% H_2O solvent, which is contained in a 0.7 mm i.d. capillary that is supercooled to −14 °C (259 K). Ideally, several such capillaries should be packed together inside a conventional NMR tube (3 or 5 mm) to improve the signal-to-noise ratio and/or reduce the acquisition time of NMR experiments. Under these conditions, the hydroxyl resonances in glucose and glucose-containing disaccharides appear as broad singlets (or occasionally as poorly resolved doublets) between the chemical shifts δ_H 5.5−8.5 ppm. Because this −OH range is distinct from that of the C−H resonances (δ_H 3−5.5 ppm), which are the only other hydrogens observable in a carbohydrate, the dispersion of the proton spectrum is effectively doubled. In addition, the −OH resonances suffer from less accidental overlap than the C−H resonances and therefore provide a better starting point for the structure elucidation/complete assignment of both mono- and di-saccahrides.

The inclusion of −OH groups in this expedited approach to determine the covalent structure of carbohydrates finds some parallels with the now well-exploited NMR structural determination of peptides and proteins that may contain hundreds of amino acids. Thus, the current protein NMR strategy also begins with the −NH resonances, which occupy an even wider range of distinctive chemical shifts (δ_H 6−10 ppm) for structured proteins. The analysis of oligosaccharides and polysaccharides is expected to be more complex than peptides and proteins because there are five OH's in each hexose carbohydrate monomer, whereas there is only one −NH amide linkage between each of the amino acids that constitute the linear structure of a peptide/protein. The number of OH protons requiring analysis decreases to four per monomer for a disaccharide connected by a single glycosidic linkage and to three for a (infinite) linear polysaccharide with two glycosidic linkages to each constituent monomer. A unique feature of our alternative approach is that it is also well suited to identify which hydroxyl group(s) in a carbohydrate monomer has become involved in the formation of these glycosidic linkages. It is this variability in glycosidic linkages that is responsible for the frequently branched and complex structures of oligosaccharides, and establishing the positions of these linkages is arguably the greatest challenge in structural carbohydrate chemistry today. By applying our alternative approach, this can now be achieved by identifying those carbons that do not bear −OH groups from an absence of correlations in the −OH region of an HQSC-TOCSY spectrum because such carbons must either be involved in the formation of a glycosidic linkage between two monosaccharides or in the formation of a pyranose/furanose ring within an aldose monosaccharide.[j] Previously, the only way to have identified the carbons involved in such glycosidic linkages would have been via the more laborious, and destructive, chemical derivatization of an oligosaccharide. This alternative −OH-based NMR approach should therefore be of value within many areas of chemistry and glycobiology, where a full understanding of the structure of a carbohydrate is essential for its subsequent application.

EXPERIMENTAL

All spectra were acquired using a Bruker AVIII 700 MHz instrument equipped with a TCI cryogenically cooled probe, with temperature regulation by a VT system able to achieve a minimum stable temperature of −14 °C (259 K). The temperature experienced by the sample was precalibrated using 99.8% d_4-MeOH as an NMR thermometer.[47,48]

Application of the Alternative NMR Approach to an Aqueous Solution of Glucose. Aliquots (15−20 μL) of a solution of D-glucose in 99% H_2O/1% D_2O [100 mM, incorporating 0.1 % (w/v) d_4-TSP as internal reference] were transferred to three capillary NMR tubes of 0.7 mm i.d. (Hilgenberg Part no. 86553), and the solutions were settled by centrifugation. These capillaries were then packed tightly together inside a 3 mm NMR tube.

^{1}H NMR spectra were recorded without solvent suppression (see Figure 1) using 8 repetitions of the Bruker pulse sequence

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

"zg", digitized over 65536 points, with a sweep width of 20 ppm centered at 4.7 ppm, an FID resolution of 0.44 Hz (acquisition time: 2.27 s), and a recycling delay of 1 s, leading to a total experiment time of 0.5 min. The time data were weighted using a negative exponential function (line broadening of 0.3 Hz), followed by 1D Fourier transformation to produce a spectrum consisting of 65536 points. ^{1}H NMR spectra were also recorded with water suppression by excitation sculpting (Figure 2) using the pulse sequence "zgesgp" (see the Supporting Information).

HSQC-TOCSY spectra were recorded using 16 repetitions of the Bruker pulse program "hsqcdietgpsisp.2" with 2048 points in F2 (sweep width 12 ppm, centered at 4.7 ppm) and 128 points in F1 (sweep width 60 ppm, centered at 78 ppm), resulting in an FID resolution of 8 Hz in F2 and 165 Hz in F1 (total acquisition time: 0.12 s in F2) and a recycling delay of 1 s, leading to a total experiment time of 40 min. The time data were weighted using a Qsine function (SSB = 2) in both dimensions, followed by 2D Fourier transformation with linear prediction (ncoeff = 16) in the second dimension, producing a spectrum with 2048 points in F2 and 1024 points in F1. The TOCSY mixing time of the experiment (d9) was varied as described in Figures 3, 6, 7, 8, and 10.

Application of the Alternative NMR Approach to Various Disaccharides. Aliquots (17 μL) of a solution of nigerose (Glcα1-3Glc) in 95% H_2O/5% D_2O (1 M, incorporating 0.1 % (w/v) d_4-TSP) were transferred to three 0.7 mm i.d. capillaries, which were packed together in a 3 mm NMR tube; then, NMR data were acquired as above. β-Gentiobiose (Glcβ1-6Glc), leucrose (Glcα1-5Fru), and isomaltulose (Glcα1-6Fru) were analyzed in the same way.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsomega.8b02136.

Experimental procedure and the NMR spectra of glucose, nigerose, gentibiose, leucrose, and isomaltulose (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: g.d.brown@reading.ac.uk.

ORCID

Helen M. I. Osborn: 0000-0002-0683-0457

Present Address

[§]Boehringer Ingelheim Pharma GmbH & Co. KG, Binger Straße 173, 55216 Ingelheim am Rhein, Germany

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The NMR instrumentation of the Chemical Analysis Facility (CAF) of the University of Reading was used to obtain the spectra described in this manuscript. J.B. and H.M.I.O. would like to thank the European Commission's Seventh Framework Programme for funding (FP7/2007-2013 under grant agreement number 215536 for funding the Marie Curie Initial Training Network EuroGlycoArrays).

■ ADDITIONAL NOTES

[a]A hybrid of the two techniques has also been described for aqueous solutions that incorporate an organic solvent (e.g., 15% d_6-acetone/85% H_2O), which are supercooled to between −8 and −12 °C.

[b]The chemical shift for the H-1(α) anomeric proton of α-glucose (δ_H 5.35 ppm) lies very close to the chemical shift for the water resonance at −14 °C. This signal was therefore unavoidably affected by all the water suppression techniques that were investigated.

[c]H-1(α) (δ_H 5.35 ppm, d, J = 4 Hz); H-1(β) (δ_H 4.76 ppm, d, J = 9 Hz); one of H-6(β) (δ_H 4.01 ppm, dd, J = 12, 2 Hz); one of H-6(α) (δ_H 3.89 ppm, dd, J = 12, 5 Hz); H-2(α) (δ_H 3.64 ppm, dd, J = 10, 4 Hz); H-2(β) (δ_H 3.35 ppm, dd, J = 9, 9 Hz).

[d]The two ^{1}H resonances in a H−C−OH jigsaw piece are correlated with one another in the ^{1}H−^{1}H COSY spectrum of a supercooled solution of glucose (Figure S4). This is a useful experiment if there is any difficulty in determining which ^{13}C chemical shift is common to the −OH groups in Figure 3 and the C−H groups in Figure 4.

[e]This new peak is in fact due to coherence being relayed over three ^{1}H atoms: H-1(β) → H-2(β) → 2-OH(β), at the edited ^{13}C chemical shift of the anomeric carbon C-1(β); while, the (previously identified) 1(β) piece was the result of a relay over just two ^{1}H atoms: H-1(β) → 1-OH(β), at this same edited ^{13}C chemical shift. The reason that 2-OH(β) and 1-OH(β) appear to be connected by arrow A is that both transfers start at H-1(β) (which resides in a region of the HSQC-TOCSY spectrum not included in the expansions of Figures 3 and 6), and H-1(β) is edited by the ^{13}C chemical shift of C-1(β). Similarly, new correlation B is due to a three-hydrogen coherence transfer: H-2(β) → H-1(β) → 1-OH(β), edited by the ^{13}C chemical shift of C-2(β).

[f]While the presence of a correlation in a 2D-NMR spectrum should always be considered as firm evidence for the corresponding through-bond connection in a molecule, the converse is not true: the absence of a correlation does not necessarily imply the absence of a bonding connection. This is a matter of common experience when interpreting an HMBC spectrum in the traditional strategy for small molecules, for which "expectable" $^2J_{CH}$ and $^3J_{CH}$ correlations are frequently discovered to be absent. The absence of such correlations is not important, so long as there are sufficient actual correlations for the extension of the NMR jigsaw to be logically justifiable.

[g]6-OH receives significantly weaker long-range correlations than the other −OH groups and lies below the threshold set for contour plots at all mixing times (see also Figures S11−S13). As the only primary hydroxyl group in glucose, 6-OH is likely to be participating in faster chemical exchange than the other more sterically hindered secondary hydroxyl groups. Also, when fully resolved, this −OH proton is expected to be a double doublet, rather than a doublet. These broadening mechanisms probably both contribute to the failure of 6-OH to appear as a long-range correlation in any of Figures 6, 7, 8 and 10.

[h]That is correlations that are in addition to those already seen in the HSQC spectrum of Figure 4.

[i]This conclusion is confirmed by the absence of any correlations for C-2 in the corresponding edited-HSQC spectra.

[j]A third possibility exists for ketose sugars, in which the quaternary carbon of the hemiketal functional group also

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

shows no correlations with its attached −OH group. However, this can be distinguished from the other two by the absence of correlations to both −OH and carbon in HSQC-TOCSY.

REFERENCES

(1) Varki, A. Biological roles of glycans. *Glycobiology* **2017**, *27*, 3.
(2) Claridge, T. W. D. *High-Resolution NMR Techniques in Organic Chemistry*, 3rd edition; Elsevier: Amsterdam, 2016.
(3) Willker, W.; Leibfritz, D.; Kerssebaum, R.; Bermel, W. Gradient selection in inverse heteronuclear correlation spectroscopy. *Magn. Reson. Chem.* **1993**, *31*, 287.
(4) Brown, G. D. The power of NMR: in two and three dimensions. *Educ. Chem.* **2008**, *45*, 152.
(5) Nestor, G.; Anderson, T.; Oscarson, S.; Gronenborn, A. M. Direct observation of carbohydrate hydroxyl protons in hydrogen bonds with a protein. *J. Am. Chem. Soc.* **2018**, *140*, 339.
(6) Battistel, M. D.; Azurmendi, H. F.; Yu, B.; Freedberg, D. I. NMR of glycans: shedding new light on old problems. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *79*, 48.
(7) Beecher, C. N.; Larive, C. K. Methods for Measuring Exchangeable Protons in Glycosaminoglycans. *Glycosaminoglycans: Chemistry and Biology*; Humana Press: New York, 2015; Vol. *1229*, p 173.
(8) Battistel, M. D.; Azurmendi, H. F.; Freedberg, D. I. Glycan OH exchange rate determination in aqueous solution: seeking evidence for transient hydrogen bonds. *J. Phys. Chem. B* **2017**, *121*, 683.
(9) Battistel, M. D.; Pendrill, R.; Widmalm, G.; Freedberg, D. I. Direct evidence for hydrogen bonding in glycans: A combined NMR and molecular dynamics study. *J. Phys. Chem. B* **2013**, *117*, 4860.
(10) Kay, L. E.; Ikura, M.; Tschudin, R.; Bax, A. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson.* **1990**, *89*, 496.
(11) Vuister, G. W.; de Waard, P.; Boelens, R.; Vliegenthart, J. F. G.; Kaptein, R. The use of three-dimensional NMR in structural studies of oligosaccharides. *J. Am. Chem. Soc.* **1989**, *111*, 772.
(12) Gerothanassis, I. P. Oxygen-17 NMR spectroscopy: basic principles and applications (Part I). *Prog. Nucl. Magn. Reson. Spectrosc.* **2010**, *56*, 95.
(13) Charisiadis, P.; Exarchou, V.; Troganis, A. N.; Gerothanassis, I. P. Exploring the "forgotten" -OH NMR spectral region in natural products. *Chem. Commun.* **2010**, *46*, 3589.
(14) Siebert, H.-C.; Andre, S.; Asensio, J. L.; Canada, F. J.; Dong, X.; Espinosa, J. F.; Frank, M.; Gilleron, M.; Kaltner, H.; Kozár, T.; Bovin, N. V.; von der Lieth, C.-W.; Vliegenthart, J. F. G.; Jimenez-Barbero, J.; Gabius, H. J. A new combined computational and NMR-spectroscopical strategy for the identification of additional conformational constraints of the bound ligand in an aprotic solvent. *ChemBioChem* **2000**, *1*, 185.
(15) Siebert, H.-C.; André, S.; Vliegenthart, J. F. G.; Gabius, H. J.; Minch, M. J. Suitability of binary mixtures of water with aprotic solvents to turn hydroxyl protons of carbohydrate ligands into conformational sensors in NOE and transferred NOE experiments. *J. Biomol. NMR* **2003**, *25*, 197.
(16) Xia, J.; Case, D. A. Sucrose in aqueous solution revisited, Part 2: Adaptively biased molecular dynamics simulations and computational analysis of NMR relaxation. *Biopolymers* **2012**, *97*, 289.
(17) Poppe, L.; van Halbeek, H. NMR spectroscopy of hydroxyl protons in supercooled carbohydrates. *Nat. Struct. Biol.* **1994**, *1*, 215.
(18) Sheng, S. Q.; van Halbeek, H. Evidence for a transient interresidue hydrogen bond in sucrose in aqueous solution obtained by rotating-frame exchange NMR spectroscopy under supercooled conditions. *Biochem. Biophys. Res. Commun.* **1995**, *215*, 504.
(19) Sandström, C.; Kenne, L. Hydroxy Protons in Structural Studies of Carbohydrates by NMR Spectroscopy. In *NMR Spectroscopy and Computer Modeling of Carbohydrates*; ACS Symposium Series 930; American Chemical Society: Washington, DC, 2006; p 114.
(20) Symons, M. C. R.; Benbow, J. A.; Harvey, J. M. Hydroxyl proton resonance shifts for a range of aqueous sugar solutions. *Carbohydr. Res.* **1980**, *83*, 9.
(21) Sandström, C.; Baumann, H.; Kenne, L. NMR spectroscopy of hydroxy protons of 3,4-disubstituted methyl α-D-galactopyranosides in aqueous solution. *J. Chem. Soc., Perkin Trans. 2* **1998**, *2*, 809.
(22) Sandström, C.; Magnusson, G.; Nilsson, U.; Kenne, L. Comparative ^{1}H NMR study of hydroxy protons in galabioside in its *S*-linked 4-thiodisaccharide analogue in aqueous solution. *Carbohydr. Res.* **1999**, *322*, 46.
(23) Bekiroglu, S.; Kenne, L.; Sandström, C. ^{1}H NMR studies of maltose, maltoheptaose, α-, β-, and γ-cyclodextrins, and complexes in aqueous solutions with hydroxy protons as structural probes. *J. Org. Chem.* **2003**, *68*, 1671.
(24) Besombes, S.; Utille, J.-P.; Mazeau, K.; Robert, D.; Taravel, F. R. Conformational study of a guaiacyl β-*O*-4 lignin model compound by NMR. Examination of intramolecular hydrogen bonding interactions and conformational flexibility in solution. *Magn. Reson. Chem.* **2004**, *42*, 337.
(25) Rohfritsch, P. F.; Frank, M.; Sandström, C.; Kenne, L.; Vliegenthart, J. F. G.; Kamerling, J. P. Comparative ^{1}H NMR and molecular modeling study of hydroxy protons of β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→2)-α-D-Man*p*-(1→O)$(CH_2)_7CH_3$ analogues in aqueous solution. *Carbohydr. Res.* **2007**, *342*, 597.
(26) Hakkarainen, B.; Kenne, L.; Lahmann, M.; Oscarson, S.; Sandström, C. NMR study of hydroxy protons of di- and trimannosides, substructures of Man-9. *Magn. Reson. Chem.* **2007**, *45*, 1076.
(27) Kindahl, L.; Sandström, C.; Norberg, T.; Kenne, L. ^{1}H NMR studies of hydroxy protons of ASN-and ser-linked disaccharides in aqueous solution. *J. Carbohydr. Chem.* **2000**, *19*, 1291.
(28) Battistel, M. D.; Shangold, M.; Trinh, L.; Shiloach, J.; Freedberg, D. I. Evidence for helical structure in a tetramer of α2-8 sialic acid: Unveiling a structural antigen. *J. Am. Chem. Soc.* **2012**, *134*, 10717.
(29) Battistel, M. D.; Azurmendi, H. F.; Frank, M.; Freedberg, D. I. Uncovering nonconventional and conventional hydrogen bonds in oligosaccharides through NMR experiments and molecular modeling: Application to sialyl Lewis-X. *J. Am. Chem. Soc.* **2015**, *137*, 13444.
(30) Norris, S. E.; Landström, J.; Weintraub, A.; Bull, T. E.; Widmalm, G.; Freedberg, D. I. Transient hydrogen bonding in uniformly ^{13}C,^{15}N-labeled carbohydrates in water. *Biopolymers* **2012**, *97*, 145.
(31) Craig, B. N.; Janssen, M. U.; Wickersham, B. M.; Rabb, D. M.; Chang, P. S.; O'Leary, D. J. Isotopic perturbation of intramolecular hydrogen bonds in rigid 1,3-diols: NMR studies reveal unusually large equilibrium isotope effects. *J. Org. Chem.* **1996**, *61*, 9610.
(32) Vasquez, T. E.; Bergset, J. M.; Fierman, M. B.; Nelson, A.; Roth, J.; Khan, S. I.; O'Leary, D. J. Using equilibrium isotope effects to detect intramolecular OH/OH hydrogen bonds: Structural and solvent effects. *J. Am. Chem. Soc.* **2002**, *124*, 2931.
(33) Loening, N. M.; Anderson, C. E.; Iskenderian, W. S.; Anderson, C. D.; Rychnovsky, S. D.; Barfield, M.; O'Leary, D. J. Qualitative and quantitative measurements of hydrogen bond mediated scalar couplings in acyclic 1,3-diols. *Org. Lett.* **2006**, *8*, 5321.
(34) Szyperski, T.; Mills, J. L. NMR-based structural biology of proteins in supercooled water. *J. Struct. Funct. Genomics* **2011**, *12*, 1.
(35) Skalicky, J. J.; Sukumaran, D. K.; Mills, J. L.; Szyperski, T. Toward structural biology in supercooled water. *J. Am. Chem. Soc.* **2000**, *122*, 3230.
(36) Schroeder, K. T.; Skalicky, J. J.; Greenbaum, N. L. NMR spectroscopy of RNA duplexes containing pseudouridine in supercooled water. *RNA* **2005**, *11*, 1012.
(37) Batta, G.; Kövér, K. E. Heteronuclear coupling constants of hydroxyl protons in a water solution of oligosaccharides: Trehalose and sucrose. *Carbohydr. Res.* **1999**, *320*, 267.
(38) Hwang, T.-L.; Shaka, A. J. Water suppression that works. Excitation sculpting using arbitrary waveforms and pulsed field gradients. *J. Magn. Reson., Ser. A* **1995**, *112*, 275.

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975

(39) Jerschow, A. Unwanted signal leakage in excitation sculpting with single axis gradients. *J. Magn. Reson.* **1999**, *137*, 206.
(40) Braunschweiler, L.; Ernst, R. R. Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. *J. Magn. Reson.* **1983**, *53*, 521.
(41) Bax, A. D.; Davis, D. G. MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J. Magn. Reson.* **1985**, *65*, 355.
(42) Zhao, H.; Pan, Q.; Zhang, W.; Carmichael, I.; Serianni, A. S. DFT and NMR studies of $^{2}J_{COH}$, $^{3}J_{HCOH}$, and $^{3}J_{CCOH}$ spin-couplings in saccharides: C−O torsional bias and H-bonding in aqueous solution. *J. Org. Chem.* **2007**, *72*, 7071−7082.
(43) Bax, A.; Summers, M. F. Proton and carbon-13 assignments from sensitivity-enhanced detection of heteronuclear multiple-bond connectivity by 2D multiple quantum NMR. *J. Am. Chem. Soc.* **1986**, *108*, 2093.
(44) Rance, M.; Sørensen, O.W.; Bodenhausen, G.; Wagner, G.; Ernst, R. R.; Wüthrich, K. Improved spectral resolution in COSY ^{1}H NMR spectra of proteins *via* double quantum filtering. *Biochem. Biophys. Res. Commun.* **1983**, *117*, 479.
(45) Olsson, U.; Säwén, E.; Stenutz, R.; Widmalm, G. Conformational flexibility and dynamics of two (1→6)-linked disaccharides related to an oligosaccharide epitope expressed on malignant tumour cells. *Chem. - Eur. J.* **2009**, *15*, 8886.
(46) Sheng, S.; Cherniak, R.; van Halbeek, H. A ^{1}H NMR spectroscopic approach to the unambiguous determination of glycosyl linkage positions in oligosaccharides. *Anal. Biochem.* **1998**, *256*, 63.
(47) Webb, A. G. Temperature measurements using nuclear magnetic resonance. *Ann. Rep. NMR Spectrosc.* **2002**, *45*, 1.
(48) Findeisen, M.; Brand, T.; Berger, S. A ^{1}H-NMR thermometer suitable for cryoprobes. *Magn. Reson. Chem.* **2007**, *45*, 175.

DOI: 10.1021/acsomega.8b02136
ACS Omega 2018, 3, 17957−17975