Identification of Carbohydrate Anomers Using Ion Mobility-Mass Spectrometry

J. Hofmann[1,2]†, H. S. Hahm[3,4]†, P. H. Seeberger[3,4]*, K. Pagel[1,2]*

[1] Fritz Haber Institute of the Max Planck Society, Faradayweg 4-6, 14195 Berlin, Germany.

[2] Institute for Chemistry and Biochemistry, Free University Berlin, Takustraße 3, 14195 Berlin, Germany.

[3] Max Planck Institute of Colloids and Interfaces, Department of Biomolecular Systems, Am Mühlenberg 1, 14476 Potsdam, Germany.

[4] Institute for Chemistry and Biochemistry, Free University Berlin, Arnimallee 22, 14195 Berlin, Germany.

† These authors contributed equally to this work

*Correspondence to: peter.seeberger@mpikg.mpg.de; kevin.pagel@fu-berlin.de

Carbohydrates are ubiquitous biological polymers that are important in a broad range of biological processes[1-3]. However, owing to their branched structures and the presence of stereogenic centres at each glycosidic linkage between monomers, carbohydrates are harder to characterize than are peptides and oligonucleotides[4]. Methods such as nuclear magnetic resonance spectroscopy can be used to characterize glycosidic linkages, but this technique requires milligram amounts of material and cannot detect small amounts of coexisting isomers[5]. Mass spectrometry, on the other hand, can provide information on carbohydrate composition and connectivity for even small amounts of sample, but it cannot be used to distinguish between stereoisomers[6]. Here, we demonstrate that ion mobility-mass spectrometry—a method that separates molecules according to their mass, charge, size, and shape—can unambiguously identify carbohydrate linkage-isomers and stereoisomers. We analysed six synthetic carbohydrate isomers that differ in composition, connectivity, or configuration. Our data show that coexisting carbohydrate isomers can be identified, and relative concentrations of the minor isomer as low as 0.1 per cent can be detected. In addition, the analysis is rapid, and requires no derivatization and only small amounts of sample. These results indicate that ion mobility-mass spectrometry is an effective tool for the analysis of complex carbohydrates. This method could have an impact on the field of carbohydrate synthesis similar to that of the advent of high-performance liquid chromatography on the field of peptide assembly in the late 1970s.

The inherent structural diversity of glycans (carbohydrates that comprise a large number of monosaccharides, linked by glycosidic bonds) poses a major analytical challenge to all aspects of the glycosciences[7,8], and is one reason why glycomics lags behind the advances that have been made in genomics[9] and proteomics[10]. The structure of a glycan is described by its composition, connectivity, and configuration (Fig. 1). The composition (Fig. 1a) is defined by its monosaccharides, the basic building blocks of oligosaccharides. These building blocks are often stereoisomers that differ only in their stereochemistry at one particular carbon atom, as in the case of glucose (Glc) and galactose (Gal). Each monosaccharide contains multiple hydroxyl groups that can be a point of attachment for a glycosidic bond with the next building block. Thus, unlike oligonucleotides and proteins, carbohydrates are not necessarily linear, but rather can be branched structures with diverse regiochemistry (that is, with linkages between different hydroxyl groups; Fig. 1b). In addition, a new stereocentre emerges

when a glycosidic bond is formed, because two monosaccharides can be connected in two different configurations (Fig. 1c). These α- and β-anomers are stereoisomers, even though the connectivity is identical. Anomers are diastereomers, not enantiomers, meaning that they differ in at least one, but not all, of their stereocentres; consequently, anomers may differ in their size and properties.

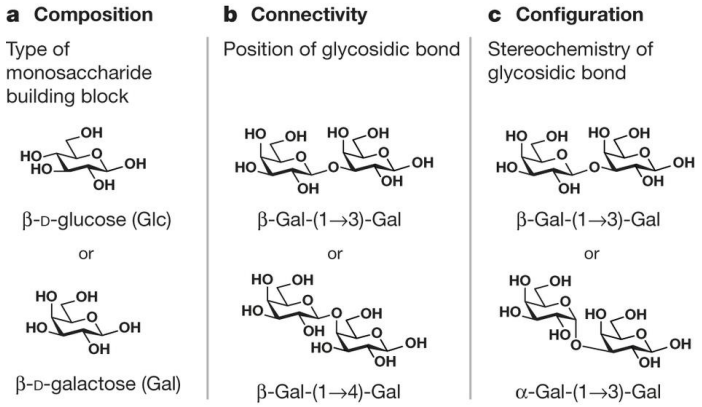


Figure 1. Structural features of complex carbohydrates. a, The composition of a carbohydrate is defined by its monosaccharide content. Monosaccharide building blocks are often isomers, as shown for glucose (Glc) and galactose (Gal), which differ only in their C4 stereochemistry. **b**, Because of the many possible functional groups, the formation of a new glycosidic bond can occur at several positions, resulting in different connectivities, such as those shown here. **c**, Each glycosidic linkage is a new stereocentre that can have either α- or β-configuration.

When synthesizing oligosaccharides, managing different compositions is straightforward, because specific building blocks are added stepwise to generate the desired structure[11]. Protective groups are used to define the connectivity by allowing the selective unveiling of specific hydroxyl groups[12], thus providing regiocontrol. In contrast, configurational control during glycosidic bond formation is the central challenge for chemical synthesis[11]. *Trans*-glycoside formation is aided by the use of participating protecting groups. *Cis*-glycoside formation, however, cannot rely on participation, and anomeric mixtures are frequently obtained[13].

The analysis of complex synthetic glycans is key to quality control but remains a major challenge. Glycan structure is typically ascertained by a combination of nuclear magnetic resonance spectroscopy (NMR) and mass spectrometry. Measuring a mass-to-charge ratio (*m/z*) with mass spectrometry is fast, requires very little sample and provides precise, high-resolution data about the sample composition. Detailed information regarding connectivity

can be obtained following derivatization (chemical modification) and/or elaborate tandem mass spectrometry analysis[6,14-16]. Nevertheless, with mass spectrometry it is not possible to analyse stereoisomers, since they generally cannot be distinguished from each other because of their identical atomic composition and mass. NMR experiments serve best to determine the configurational information of carbohydrates, but require large amounts of sample and are time-consuming; moreover, the resulting spectra are cumbersome to interpret when different stereoisomers need to be distinguished. In addition, the relative detection limit of 3% to 5% for larger oligosaccharides in NMR experiments is rather poor. Liquid chromatography can help to differentiate configurational isomers, but an unambiguous identification of one isomer in the presence of another is often not possible either[8].

A promising approach to overcoming these limitations is the combination of ion mobility spectrometry and mass spectrometry (IM–MS)[17,18]. Here, carbohydrate ions are separated not only according to their mass and charge, but also on the basis of their size and shape. IM–MS measures the time that ions require to drift through a cell that is filled with an inert neutral gas such as helium or nitrogen, under the influence of a weak electric field. While drifting, compact ions undergo fewer collisions with the gas than more extended ions, and therefore traverse the cell faster. This principle can allow for the separation of species with identical mass but different structure. The sample and time requirements of IM–MS are similar to those of a conventional mass spectrometry experiment, so the additional information is obtained at no extra cost. Moreover, the measured drift time can be converted into an instrument-independent, rotationally averaged collision cross-section (CCS). IM–MS has already been proven to be of value in the structural analysis of proteins and their assemblies[19,20], and also showed promise in the separation of carbohydrate and glycopeptide isomers[21-26]. Previous IM–MS studies of carbohydrates, however, were almost exclusively focused on pairs of regioisomers; data regarding the equally important compositional and configurational isomers is still lacking.

Here, we illustrate the utility of IM–MS for the in-depth structural analysis of carbohydrates by systematically investigating all types of isomerism simultaneously within one consistent and comparable set of compounds. Six trisaccharide isomers (Fig. 2a) that, owing to their similarity in structure, are difficult to distinguish using established techniques were prepared using automated glycan assembly (Extended Data Fig. 1)[11,27]. The six glycans share the reducing-end lactose motif Gal(β1→4)Glc, and an aminoalkyl linker placed during automated

synthesis for conjugation to carrier proteins or array surfaces. The non-reducing-end moiety was varied, to generate isomers that differ in composition, connectivity, or configuration. Each of the glycan pairs **1** + **2** and **4** + **5** share the same regiochemistry and stereochemistry of their glycosidic linkages, but differ in their composition. On the other hand, the trisaccharide pairs **1** + **4**, **2** + **5**, and **3** + **6** are connectivity isomers, where the terminal building block is connected through either a 1→3 or a 1→4 glycosidic linkage. Finally, the glycan pairs **2** + **3** and **5** + **6** are configurational isomers that differ in the stereochemistry of the terminal glycosidic linkage. We analysed all six carbohydrates separately as both positively and negatively charged ions, using a commercially available hybrid IM–MS instrument (see Methods)[28]. Although most previous studies focused on positive-ion adducts[21-26], we observed the most notable drift-time differences for deprotonated $[M\text{-}H]^-$

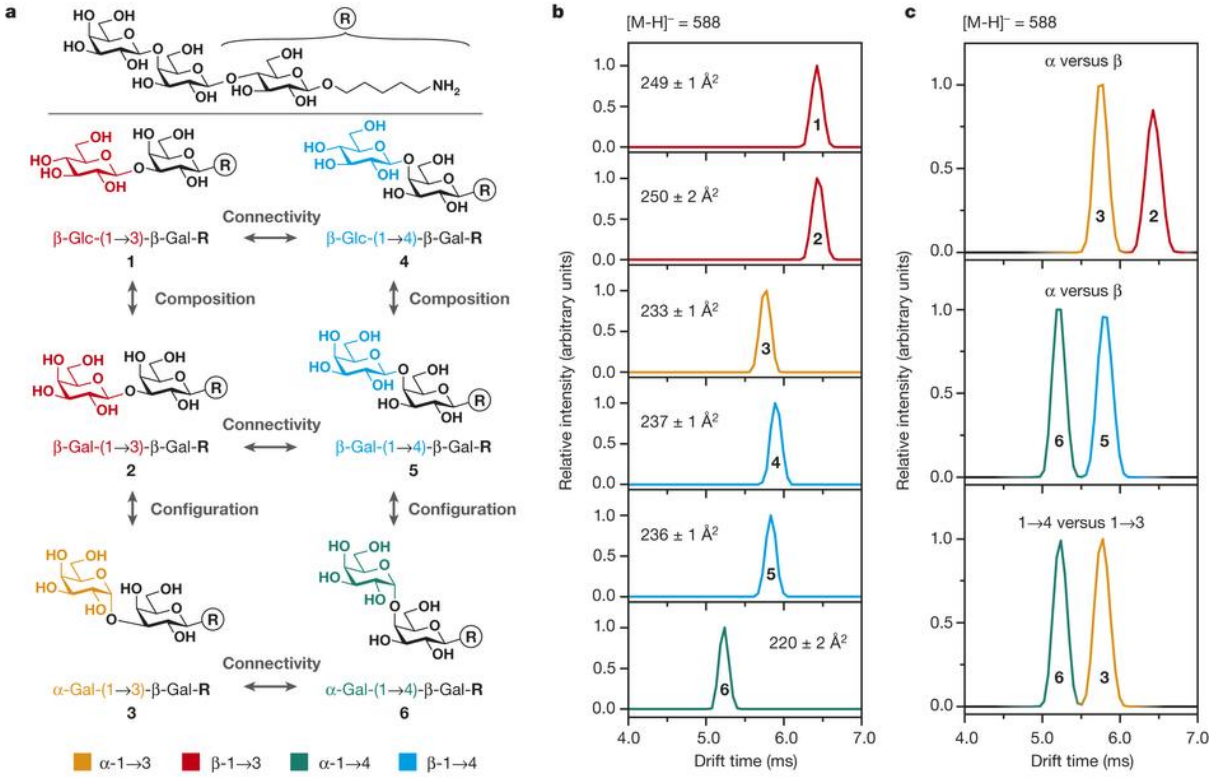


Figure 2. Structure and IM–MS data of trisaccharides 1–6. a, The synthetic trisaccharides **1**–**6** share the same disaccharide core, and differ merely in the composition, connectivity, or configuration of the last monosaccharide building block. **b**, IM–MS drift-time distributions (also known as arrival-time distributions) for trisaccharides **1**–**6** as $[M\text{-}H]^-$ ions. The values in $Å^2$ correspond to the estimated CCSs in the drift gas nitrogen and represent averages of three independent measurements. Although compositional isomers cannot be distinguished, connectivity and configurational isomers are clearly identified on basis of their CCSs. **c**, IM–MS drift-time distributions of isomeric mixtures show baseline separation between linkage- and stereoisomers.

ions (Fig. 2b and Extended Data Fig. 3). As a result, we achieve a higher separation capability, and therefore these results will be used for the analysis below. The drift times of the individual sugars were further converted into CCSs (Extended Data Table 2) using a previously reported calibration protocol to enable comparison[29,30].

A comparison of the drift times and CCSs of the six trisaccharides revealed both similarities and differences (Fig. 2b). The compositional isomers **1** and **2** exhibit drift times and CCSs that are almost identical to each other; the same is true of compositional isomers **4** and **5**. This observation is not surprising, given that the respective trisaccharide pairs differ only in the orientation of one hydroxyl group, at the C4 carbon atom. Such a minimal structural difference is not expected to result in a notable difference in CCS. However, the composition of carbohydrates is easily controlled during automated synthesis, because either glucose or galactose building blocks are being used. Thus, compositional differentiation is not essential for the quality control of glycan synthesis.

In contrast to compositional isomers, regioisomers (**1** + **4**, **2** + **5**, and **3** + **6**) can be distinguished readily from each other on the basis of their drift times and CCSs. Here, the trisaccharides containing 1→3 glycosidic linkages exhibit larger CCSs than their 1→4-linked counterparts. Analytically, however, the most striking observations resulted from comparison of the **2** + **3** and **5** + **6** configurational isomers. α-Linked trisaccharides **3** and **6** adopt a more compact structure than do the corresponding β-linked molecules, and both sets of anomers can be clearly differentiated. While trisaccharides that differ in both regiochemistry and stereochemistry (for example, **3** and **5**) exhibit similar CCSs in their respective deprotonated states, they can be distinguished in IM–MS as chloride adducts (Extended Data Fig. 3). In addition, deprotonated trisaccharide fragments generated from larger oligosaccharides exhibit highly diagnostic CCS values identical to those of their intact trisaccharide counterparts (Extended Data Fig. 4).

Although regiocontrol is well established during chemical oligosaccharide assembly, the formation of mixtures of stereoisomers is common when *cis*-glycosides are installed. The characterization and quality control of synthetic oligosaccharides would therefore benefit greatly from the ability to separate and identify different isoforms. We systematically analysed mixtures of connectivity and configuration isomers by IM–MS (Fig. 2c). Strikingly, the linkage isomers (1→3 versus 1→4), as well as the α- and β-anomers, are fully baseline-

separated (their peaks do not overlap). A similar quality of separation was also obtained in the quality control of a crude product mixture, where small amounts of an unintended by-product were clearly identifiable (Extended Data Fig. 6).

This encouraging result raises the question of whether an isomeric impurity can be not only identified qualitatively, but also determined quantitatively by IM–MS. We carried out experiments to address this question, where trisaccharide **2** was kept at a constant concentra-

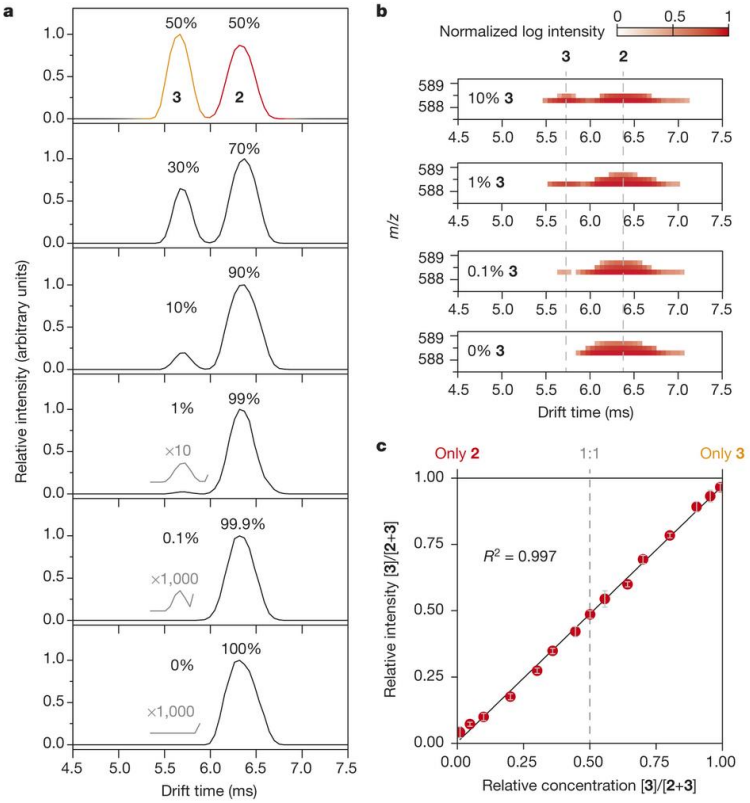


Figure 3. Relative quantification of configurational trisaccharide isomers. Mixtures of the configurational isomers **2** and **3** were measured using IM–MS. **a**, The amount of isomer **2** was kept constant, while isomer **3** was diluted to yield relative concentrations of 50%, 30%, 10%, 1%, 0.1%, and 0%. Minor components with relative concentrations as low as 0.1% can still be qualitatively detected. The grey traces are magnified by the values shown. **b**, Three-dimensional plot showing the separation of anomers **2** and **3**. The intensity is plotted using a logarithmic scale and impurities of 0.1% can be clearly identified without magnification. **c**, Plot of the relative IM-MS intensity of isomer **3** against the corresponding relative concentration, to illustrate the dynamic range of the method. A value of 0.50 represents equal amounts of isomer **2** and **3**, while values of 0 and 1 indicate the presence of only isomer **2** or isomer **3**, respectively. The grey error bars correspond to the double standard deviation observed for three independent replicates.

tion, while the content of the corresponding anomer **3** was gradually reduced to mimic different percentages of a typical synthetic impurity (Fig. 3a). As expected, the intensity of the IM–MS peak of trisaccharide **3** gradually declines with decreasing concentration. An impurity with a relative concentration as little as 1% is still clearly visible and exhibits a well resolved and well-shaped peak; however, although a relative content of 0.1% can still be identified qualitatively, it is close to the detection limit (see Methods and Extended Data Fig. 7b). To visualize better the large range of intensities, it is common to plot IM–MS data as a drift plot, with the drift time on the *x*-axis, *m/z* on the *y*-axis, and a logarithmic intensity scale (Fig. 3b). Here, a relative concentration of 0.1% is perfectly visible even without further magnification. We also tested the linearity of the IM–MS intensity for a broad range of concentrations with mixtures of anomers **2** and **3** (Fig. 3c). For that purpose, the relative peak area of the IM–MS signal of anomer **3** was plotted against the corresponding relative concentration. A value of 0.5 indicates an equal content of **2** and **3**, while values of 0 and 1 are expected for the pure oligosaccharides, respectively. Remarkably, the plot is strictly linear over the entire range of relative concentrations from 0.01 to 0.99, and very little deviation between different replicates is observed (for details see Methods and Extended Data Table 3). As a result, IM–MS can be used to estimate the relative content of a minor impurity when the investigated compounds are, like anomers, similar in structure and ionization efficiency.

In summary, we demonstrate that IM–MS is a powerful tool for the structural analysis and quality control of carbohydrates. Connectivity and configurational isomers can be separated efficiently with baseline resolution, especially when deprotonated ions are used, and the relative content of isomeric impurities can be determined quickly and easily (Extended Data Fig. 6). To our knowledge, no other experimental technique can provide the same structural information as quickly and with such minimal sample consumption. Larger glycans are known to separate less efficiently in IM–MS (Extended Data Fig. 5). However, their gas-phase fragments are similar to the oligosaccharides described here (Extended Data Fig. 4) and can therefore serve as diagnostic features for quality control and sequencing.

The full benefit of this method will become apparent once CCS data for carbohydrates and carbohydrate fragments, derived from synthetic and biological sources, are deposited in databases. These reference data will be essential for the quick and unambiguous identification of unknown compounds. The existence of commercially available mass spectrometers will enable IM–MS to become a routine technique for non-specialists or in automated analyses.

IM–MS has the potential to fundamentally change quality control and analysis in carbohydrate chemistry.

References

1. Dwek, R. A. Glycobiology: Toward understanding the function of sugars. *Chem. Rev.* **96**, 683-720 (1996).
2. Molinari, M. *N*-glycan structure dictates extension of protein folding or onset of disposal. *Nat. Chem. Biol.* **3**, 313-320 (2007).
3. Varki, A. Sialic acids in human health and disease. *Trends Mol. Med.* **14**, 351-360 (2008).
4. Bertozzi, C. R. & Rabuka, D. in *Essentials of Glycobiology* (eds Ajit Varki *et al.*) Ch. 2, (Cold Spring Harbor Laboratory Press, 2009).
5. Duus, J. Ø., Gotfredsen, C. H. & Bock, K. Carbohydrate Structural Determination by NMR Spectroscopy: Modern Methods and Limitations. *Chem. Rev.* **100**, 4589-4614 (2000).
6. Dell, A. & Morris, H. R. Glycoprotein structure determination by mass spectrometry. *Science* **291**, 2351-2356 (2001).
7. Service, R. F. Looking for a sugar rush. *Science* **338**, 321-323 (2012).
8. Marino, K., Bones, J., Kattla, J. J. & Rudd, P. M. A systematic approach to protein glycosylation analysis: a path through the maze. *Nat. Chem. Biol.* **6**, 713-723 (2010).
9. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304-1351 (2001).
10. Domon, B. & Aebersold, R. Mass spectrometry and protein analysis. *Science* **312**, 212-217 (2006).
11. Plante, O. J., Palmacci, E. R. & Seeberger, P. H. Automated solid-phase synthesis of oligosaccharides. *Science* **291**, 1523-1527 (2001).

12. Wang, Z. *et al.* A general strategy for the chemoenzymatic synthesis of asymmetrically branched *N*-glycans. *Science* **341**, 379-383 (2013).

13. Boltje, T. J., Buskas, T. & Boons, G.-J. Opportunities and challenges in synthetic oligosaccharide and glycoconjugate research. *Nature Chem.* **1**, 611-622 (2009).

14. Prien, J. M., Ashline, D. J., Lapadula, A. J., Zhang, H. & Reinhold, V. N. The high mannose glycans from bovine ribonuclease B isomer characterization by ion trap MS. *J. Am. Soc. Mass Spectrom.* **20**, 539-556 (2009).

15. Daikoku, S., Widmalm, G. & Kanie, O. Analysis of a series of isomeric oligosaccharides by energy-resolved mass spectrometry: a challenge on homobranched trisaccharides. *Rapid Commun. Mass Spectrom.* **23**, 3713-3719 (2009).

16. Harvey, D. J. Fragmentation of negative ions from carbohydrates: part 1. Use of nitrate and other anionic adducts for the production of negative ion electrospray spectra from *N*-linked carbohydrates. *J. Am. Soc. Mass Spectrom.* **16**, 622-630 (2005).

17. Bohrer, B. C., Merenbloom, S. I., Koeniger, S. L., Hilderbrand, A. E. & Clemmer, D. E. Biomolecule analysis by ion mobility spectrometry. *Annu. Rev. Anal. Chem.* **1**, 293-327 (2008).

18. Uetrecht, C., Rose, R. J., van Duijn, E., Lorenzen, K. & Heck, A. J. R. Ion mobility mass spectrometry of proteins and protein assemblies. *Chem. Soc. Rev.* **39**, 1633-1655 (2010).

19. Ruotolo, B. T. *et al.* Evidence for macromolecular protein rings in the absence of bulk water. *Science* **310**, 1658-1661 (2005).

20. Bleiholder, C., Dupuis, N. F., Wyttenbach, T. & Bowers, M. T. Ion mobility–mass spectrometry reveals a conformational conversion from random assembly to β-sheet in amyloid fibril formation. *Nature Chem.* **3**, 172-177 (2011).

21. Gabryelski, W. & Froese, K. L. Rapid and sensitive differentiation of anomers, linkage, and position isomers of disaccharides using High-Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS). *J. Am. Soc. Mass Spectrom.* **14**, 265-277 (2003).

22. Plasencia, M. D., Isailovic, D., Merenbloom, S. I., Mechref, Y. & Clemmer, D. E. Resolving and assigning *N*-linked glycan structural isomers from ovalbumin by IMS-MS. *J. Am. Soc. Mass Spectrom.* **19**, 1706-1715 (2008).

23. Zhu, M., Bendiak, B., Clowers, B. & Hill Jr., H. H. Ion mobility-mass spectrometry analysis of isomeric carbohydrate precursor ions. *Anal. Bioanal. Chem.* **394**, 1853-1867 (2009).

24. Williams, J. P. *et al.* Characterization of simple isomeric oligosaccharides and the rapid separation of glycan mixtures by ion mobility mass spectrometry. *Int. J. Mass Spectrom.* **298**, 119-127 (2010).

25. Fenn, L. S. & McLean, J. A. Structural resolution of carbohydrate positional and structural isomers based on gas-phase ion mobility-mass spectrometry. *Phys. Chem. Chem. Phys.* **13**, 2196-2205 (2011).

26. Both, P. *et al.* Discrimination of epimeric glycans and glycopeptides using IM-MS and its potential for carbohydrate sequencing. *Nature Chem.* **6**, 65-74 (2014).

27. Kröck, L. *et al.* Streamlined access to conjugation-ready glycans by automated synthesis. *Chem. Sci.* **3**, 1617-1622 (2012).

28. Pringle, S. D. *et al.* An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave IMS/oa-ToF instrument. *Int. J. Mass Spectrom.* **261**, 1-12 (2007).

29. Pagel, K. & Harvey, D. J. Ion mobility-mass spectrometry of complex carbohydrates – collision cross sections of sodiated *N*-linked glycans. *Anal. Chem.* **85**, 5138-5145 (2013).

30. Hofmann, J. *et al.* Estimating collision cross sections of negatively charged *N*-glycans using traveling wave ion mobility-mass spectrometry. *Anal. Chem.* **86**, 10789-10795 (2014).

Acknowledgments We thank the Free University Berlin and the Max Planck Society for generous financial support. J.H. and K.P. thank Gert von Helden, Justin L. P. Benesch and Weston B. Struwe for fruitful comments.

Author contributions P.H.S. and K.P. designed research; J.H. and H.S.H. performed research and contributed equally to the study. All authors analysed data and wrote the manuscript.

Author information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.H.S. (peter.seeberger@mpikg.mpg.de) or K.P. (kevin.pagel@fu-berlin.de).

METHODS

Materials. All chemicals were reagent grade and used as supplied. Before use, molecular sieves were activated by heating under high vacuum. All reactions were performed in oven-dried glassware under an argon atmosphere, unless noted otherwise. *N*,*N*-dimethylformamide (DMF), dichloromethane (DCM), toluene and tetrahydrofuran (THF) were purified in a cycle-Tainer solvent delivery system.

Pre-automation steps. All elements such as the synthesizer, modules, linker-bound resin (**10**), and a set of building blocks (**11**–**19**) were prepared (see Supplementary Information for details).

Automated glycan assembly. All automated glycosylations were performed on an automated oligosaccharide synthesizer demonstrator unit using anhydrous solvents of the cycle-Tainer solvent delivery system. Oligosaccharides **1**–**9** were synthesized following the schemes in Extended Data Figs 1 and 2 using sequences I to V (Extended Data Table 1)[11, 27, 31-33]. Detailed synthesis information and NMR data can be obtained from the Supplementary Information. To start the synthesis sequence, the resin was swollen in 2 ml DCM. The building blocks were coevaporated with toluene three times, dissolved in DCM under an argon atmosphere, and transferred into vials that were placed on the corresponding ports in the synthesizer. Reagents were dissolved in the corresponding solvents under an argon atmosphere in bottles that were placed on the corresponding ports in the synthesizer.

Module 1: acidic TMSOTf wash. The resin is washed three times with each of DMF, THF and DCM (each 2 ml for 25 s), and then once with 0.35 ml of a solution of trimethylsilyl trifluoromethanesulfonate (TMSOTf) in DCM at −20 °C. The resin is swollen in 2 ml DCM and the temperature of the reaction vessel is adjusted to the set temperature, T_a.

Module 2: glycosylation using thioglycoside. For glycosylation, the DCM is drained and a solution of thioglycoside building block (5 equivalents in 1.0 ml DCM) is delivered to the reaction vessel. After the set temperature (T_a) is reached, the reaction starts with the addition of 1 ml *N*-iodosuccinimide (NIS, 5 equivalents in 1.0 ml DCM) and trifluoromethanesulfonic acid (TfOH, 0.1 equivalents in 1.0 ml DCM) solution. The glycosylation is performed for time t_1 at temperature T_a and for time t_2 at temperature T_i. After the reaction, the solution is

drained and the resin is washed with DCM (six times with 2 ml for 15 s each). This procedure is repeated twice.

Module 3: Fmoc deprotection. The resin is washed with DMF (six times with 2 ml for 15 s) and swollen in 2 ml DMF; then the temperature of the reaction vessel is adjusted to 25 °C. For fluorenylmethyloxycarbonyl (Fmoc) deprotection, the DMF is drained and 2 ml of a solution of 20% triethylamine in DMF is delivered to the reaction vessel. After 5 minutes, the reaction solution is collected in the fraction collector of the oligosaccharide synthesizer and 2 ml of a solution of 20% triethylamine in DMF is delivered to the resin. This procedure is repeated three times.

Module 4: glycosylation using phosphate. For glycosylation the DCM is drained and a solution of phosphate building block (5 equivalents in 1.0 ml DCM) is delivered to the reaction vessel. After the set temperature (T_a) is reached, the reaction starts with the addition of 1 ml TMSOTf solution. The glycosylation is performed for t_1 at T_a, and for t_2 at T_i. After the reaction the solution is drained and the resin is washed with DCM (six times with 2 ml for 15 s). This procedure is repeated twice.

Module 5: levulinoyl (lev) deprotection. The resin is washed with DCM (six times with 2 ml for 25 s) and swollen in 1.3 ml DCM; the temperature of the reaction vessel is adjusted to 25 °C. For levulinoyl deprotection, 0.8 ml of the hydrazine hydrate solution is delivered into the reaction vessel. After 30 minutes the reaction solution is drained and the resin is washed with 0.2 M acetic acid in DCM and DCM (six times each with 2 ml for 25 s). The entire procedure is performed three times.

Purification of protected oligosaccharides. Following ultraviolet cleavage using the continuous-flow photoreactor[31], the crude molecules were confirmed with matrix-assisted laser desorption/ionization (MALDI) and crude NMR (^{1}H, and heteronuclear single quantum coherence (HSQC) spectra, recorded on a Varian Mercury 400 (400 MHz) or 600 (600 MHz) spectrometer). In addition, the crude material was analysed by high-performance liquid chromatography (HPLC; Agilent 1100 series spectrometer; column: Luna 5 μm, silica 100 Å (260 × 4.60 mm); flow rate 1 ml min^{-1}; eluents 5% DCM in hexane/5% DCM in ethyl acetate; gradient 20% (for 5 min), 60% (in 40 min), 100% (in 5 min); detection 280 nm, and evaporating light scattering detection (ELSD)). The samples were purified using preparative HPLC (Agilent 1200 series). The crude mixture was carefully dissolved in a minimum

volume of DCM and 0.9 ml of 20% hexane in ethyl acetate, and injected for purification using preparative HPLC (column: Luna 5 μm, silica 260 × 10 mm; flow rate 5 ml min^{-1}; eluents 5% DCM in hexane/5% DCM in ethyl acetate; gradient 20% (for 5 min), 60% (in 40 min), 100% (in 5 min); detection 280 nm, and ELSD) to afford the fully protected target oligosaccharide.

Oligosaccharide deprotection and final purification. To the solution of the fully protected oligosaccharide in methanol (0.2 ml μmol^{-1} of oligosaccharide), 58 μl of 0.5 M sodium methoxide (NaOMe) solution (0.25 equivalents per acetyl of benzoyl group) in methanol was added at 40 °C. The reaction was monitored by mass spectrometry until it was completed, then neutralized by 200 mg of Amberlite (400 mg per 100 μl NaOMe solution). After filtering off the suspension, the crude mixture was dissolved in methanol, ethyl acetate, and acetic acid (5:0.5:0.2, by volume), followed by the addition of 5% palladium on carbon (Pd/C) (50% by weight = Pd/oligosaccharide), and was then purged first with argon and then with hydrogen, and left to stir overnight at room temperature under balloon pressure. The reaction mixture was filtered through modified cellulose filter and washed with 20 ml water/methanol (9:1 in volume); the combined solution was evaporated under vacuum to provide the crude material. This material was analysed by reverse-phase HPLC (column Hypercarb (150 × 4.60 mm); flow rate 0.8 ml min^{-1}; eluents 0.1% formic acid in acetonitrile/0.1% formic acid in triple-distilled water; gradient 0% (for 10 min), 30% (in 30 min), 100% (in 5 min); detection ELSD). Subsequently the crude solution was purified by preparative reverse-phase HPLC (column Hypercarb, (150 × 10.00 mm); flow rate 3.6 ml min^{-1}; eluents 0.1% formic acid in acetonitrile/0.1% formic acid in triple-distilled water; gradient 0% (for 10 min), 30% (in 30 min), 100% (in 5 min); detection ELSD) to afford the unprotected oligosaccharide. All compounds were characterized by NMR, and high-resolution mass spectral analyses were performed using an Agilent 6210 electrospray ionization–time-of-flight spectrometer (Agilent Technologies). For details see Supplementary Information.

IM–MS. IM–MS experiments were performed on a travelling-wave quadrupole/ion mobility/orthogonal acceleration time-of-flight mass spectrometer, Synapt G2-S HDMS (Waters Corporation)[28], which was mass-calibrated before measurements using a solution of caesium iodide (100 mg ml^{-1}). IM–MS data analysis was performed using MassLynx 4.1, DriftScope 2.4 (Waters Corporation), and OriginPro 8.5 (OriginLab Corporation) software.

For IM–MS analysis, compounds **1**–**9** and the crude mixture **5**/**30** were each dissolved in water/methanol (1:1 by volume) to a concentration of 1–10 μmol l^{-1}. A nano-electrospray ionization source was used to ionize 3–5 μl of sample from platinum-palladium-coated borosilicate capillaries prepared in-house. Typical settings were: source temperature, 20 °C; needle voltage, 0.8 kV; sample cone voltage, 25 V; cone gas, off. The ion mobility parameters were optimized to achieve maximum resolution without excessive heating of the ions upon injection into the ion mobility cell. Values were: trap gas flow, 2 ml min^{-1}; helium cell gas flow, 180 ml min^{-1}; ion mobility gas flow, 90 ml min^{-1}; trap direct-current bias, 35 V; ion mobility wave velocity, 800 m s^{-1}; ion mobility wave height, 40 V. For MS/MS experiments the trap collision energy was increased to 30–60 V.

IM–MS spectra of each individual carbohydrate and three trisaccharide mixtures (**6** + **3**, **3** + **2** and **5** + **6**) were recorded separately in positive- and negative-ion mode. Drift-time distributions were extracted from raw data using MassLynx and drift times were determined manually via Gaussian fitting using Origin 8.5. For the measurement of the individual carbohydrates, the *m/z* signal intensity was kept at approximately 10^3 counts per second to avoid saturation and subsequent broadening of the corresponding drift peak (for an example see Extended Data Fig. 7). To avoid discrimination of a minor component, an average signal intensity of 10^4 counts per second was used for the semiquantitative assessment of mixtures (Extended Data Fig. 7b). Under these conditions, minor components with relative concentrations below 1% can still be detected qualitatively, but a semiquantitative assessment is no longer possible. For unknown mixtures, we therefore suggest acquiring data at both high- and low-intensity settings when possible. At high intensity, minor components with relative concentrations below 1% can be qualitatively detected, while the low-intensity case typically yields a better ion mobility resolution and enables a semiquantitative assessment (Extended Data Fig. 7). In addition, an acquisition at different intensity settings can help to evaluate mixtures in which the isomers cannot be fully resolved. For broad and inconclusive drift-time distributions, a comparison with neighbouring peaks of similar mass and charge can furthermore be used to distinguish between overlapping and saturated peaks[29].

CCS estimations were performed using an established protocol and dextran as calibrant (dextran1000, number average molecular weight 1,000; and dextran5000, number average molecular weight 5,000; Sigma Aldrich)[29,30]. The calibration solution consisted of 0.1 mg ml^{-1} dextran1000, 0.5 mg ml^{-1} dextran5000, and 1 mM NaH_2PO_4 in water/methanol (1:1 by

volume). The calibrant and each sample were measured on a travelling wave Synapt instrument at five wave velocities in both positive- and negative-ion mode. Drift times were extracted from raw data by fitting a Gaussian distribution to the drift-time distribution of each ion and corrected for their *m/z*-dependent flight time. CCS reference values[30] for dextran were corrected for charge and mass, and a logarithmic plot of corrected CCSs against corrected drift times was used as a calibration curve to estimate CCSs. One calibration curve was generated for every wave velocity and each ion polarity. The resulting five estimated CCSs for each sample ion were averaged. These measurements were repeated three times and the averaged values for different ions are presented in Extended Data Table 2. The reported error corresponds to the standard deviation obtained for three independent replicates.

Semiquantitative analysis of trisaccharide mixtures. For the semiquantitative analysis of anomeric trisaccharide mixtures, a quantification experiment was performed using isomers **2** and **3**. Stock solutions of **2** and **3** with identical concentration were prepared in water/methanol (1:1 by volume). Each stock solution was diluted individually to yield relative concentrations of 80%, 56%, 43%, 25%, 11%, 5%, 1%, 0.1%, and 0.01%. The serial dilutions were used to obtain isomer mixtures with concentration ratios [**3**]/[**2**] + [**3**] between 0 and 1 (see Extended Data Table 3). A value of 0.5 represents equal amounts of **2** and **3**, while 0 and 1 indicate the presence of only **2** or only **3**, respectively.

To achieve constant experimental conditions, we performed the semiquantitative analysis on a Synapt instrument equipped with an online nano-electrospray ionization source that was coupled to an ACQUITY ultraperformance liquid chromatography system (Waters). Settings were: eluents 0.1% formic acid in methanol/0.1% formic acid in water at a constant rate of 50%, flow rate 8 µl min^{-1}, sample injection 10 µl. Data were acquired in negative-ion mode with following settings: source temperature 80 °C, needle voltage 2.7 kV; sample cone voltage 25 V, desolvation temperature 150 °C, cone gas 0 l h^{-1}, nanoflow gas 1.3 bar, purge gas flow 500.0 ml h^{-1}. Ion mobility parameters were: trap gas flow, 0.4 ml min^{-1}, helium cell gas flow 180 ml min^{-1}, ion mobility gas flow 90 ml min^{-1}, trap direct-current bias 45 V, ion mobility wave velocity 800 m s^{-1}; ion mobility wave height 40 V.

Extraction of the drift-time distribution of the 588.4 *m/z* ion showed two separate drift times, each of which corresponded to one of the two isomers. The area under the drift-time distribution is related to the concentration of the sample. Therefore, the theoretical

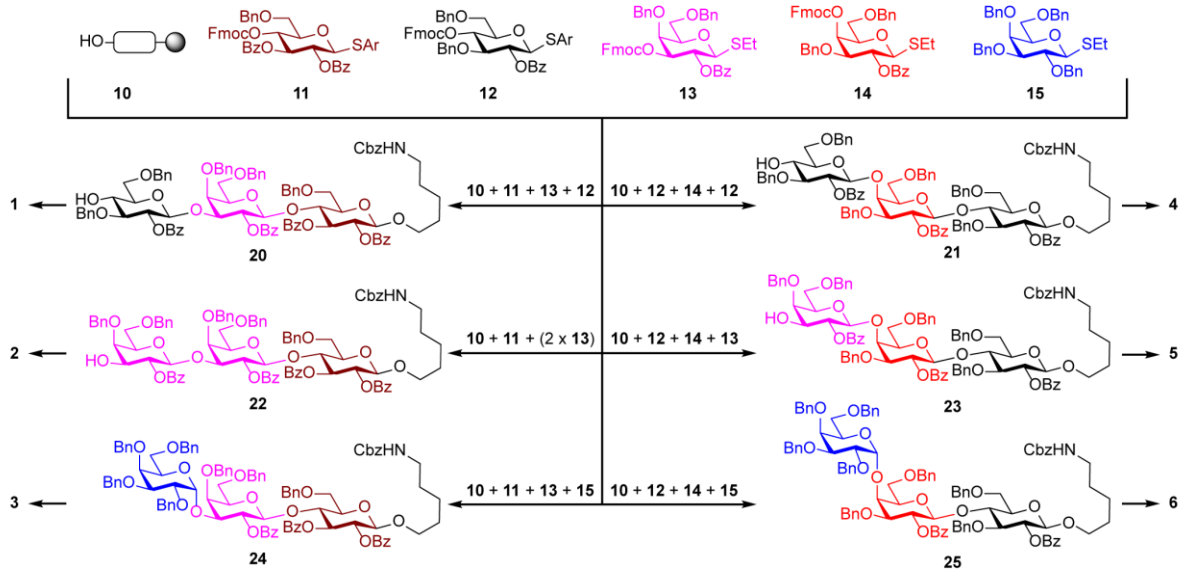
concentration ratio was compared to the ratio of the drift-time peak areas (*A*) such that the relative intensity is $A(\mathbf{3})/A(\mathbf{2}) + A(\mathbf{3})$ (Fig. 3c). A linear correlation was observed, demonstrating the semiquantification of one isomer in the presence of another, down to contents of 1% of the minor component. Relative concentrations between 1% and 0.1% were still qualitatively detectable, but a determination of the relative content was no longer possible owing to detector saturation caused by the major component.

31. Eller, S., Collot, M., Yin, J., Hahm, H. S. & Seeberger, P. H. Automated solid-phase synthesis of chondroitin sulfate glycosaminoglycans. *Angew. Chem. Int. Ed.* **52**, 5858-5861 (2013).

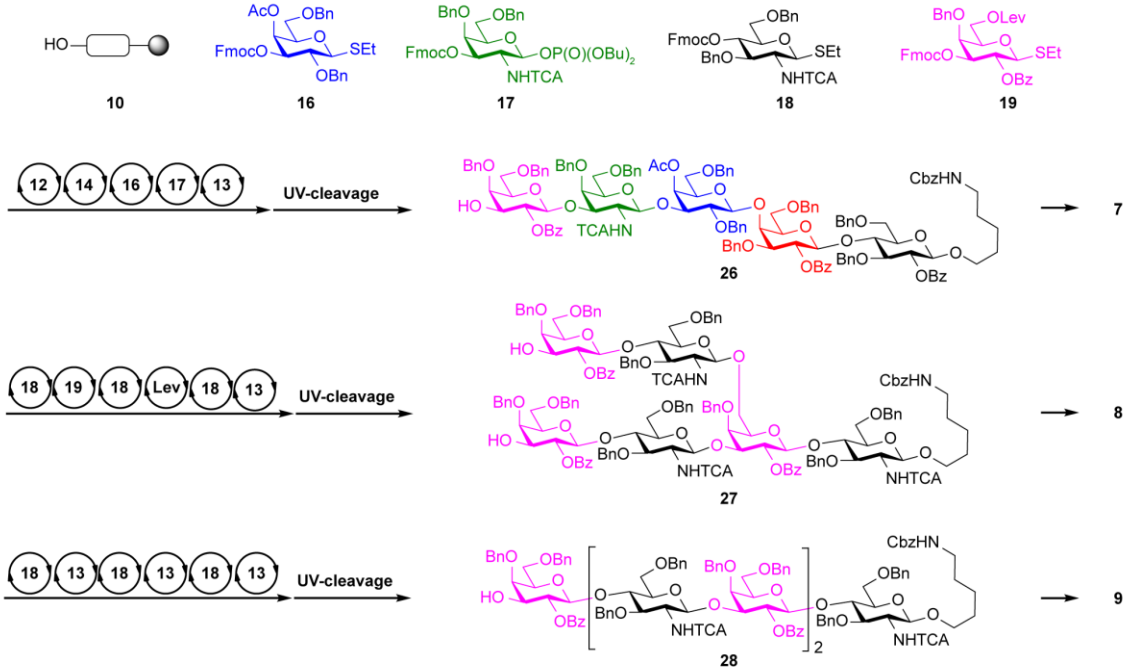
32. Martin, C. E., Weishaupt, M. W. & Seeberger, P. H. Progress toward developing a carbohydrate-conjugate vaccine against Clostridium difficile ribotype 027: synthesis of the cell-surface polysaccharide PS-I repeating unit. *Chem. Commun.* **47**, 10260-10262 (2011).

33. Werz, D. B., Carstagner, B. & Seeberger, P. H. Automated synthesis of the tumor-sssociated carbohydrate sntigens Gb-3 and Globo-H: Incorporation of α-galactosidic linkages. *J. Am. Chem. Soc.* **129**, 2770-2771 (2007).

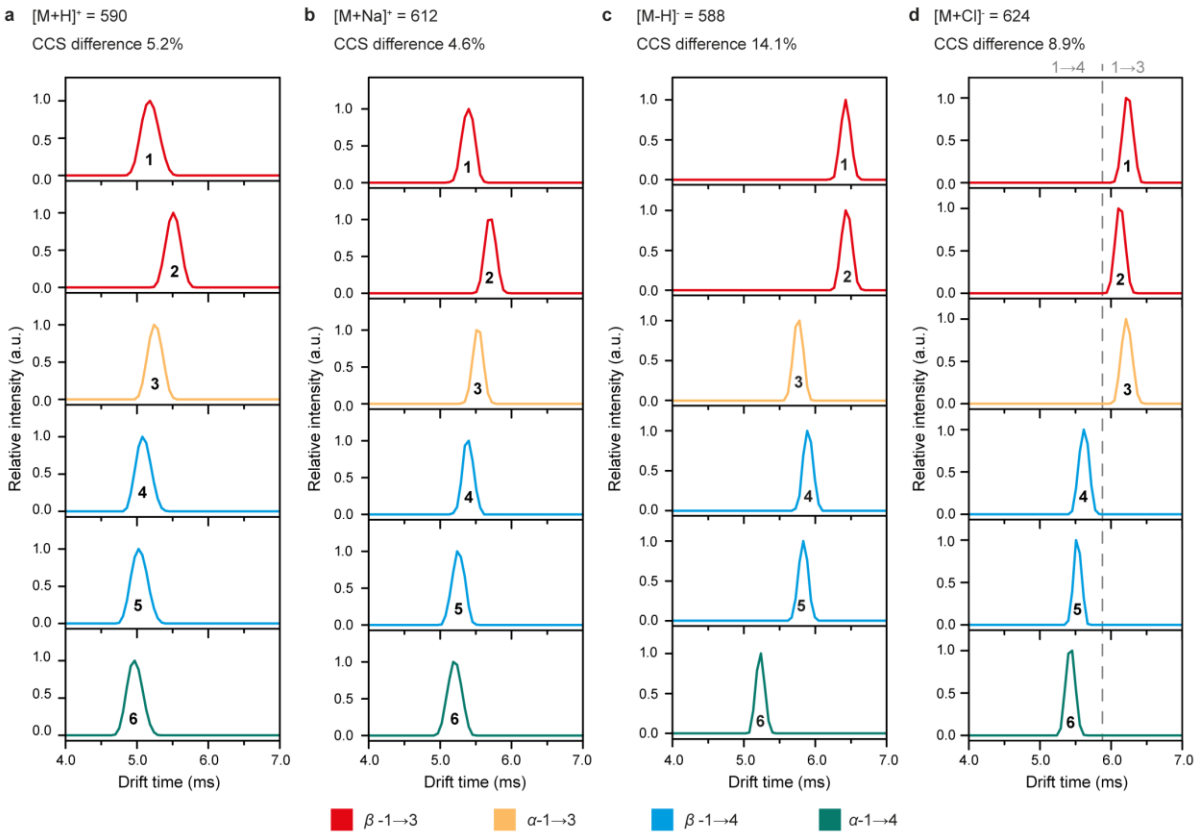
EXTENDED DATA



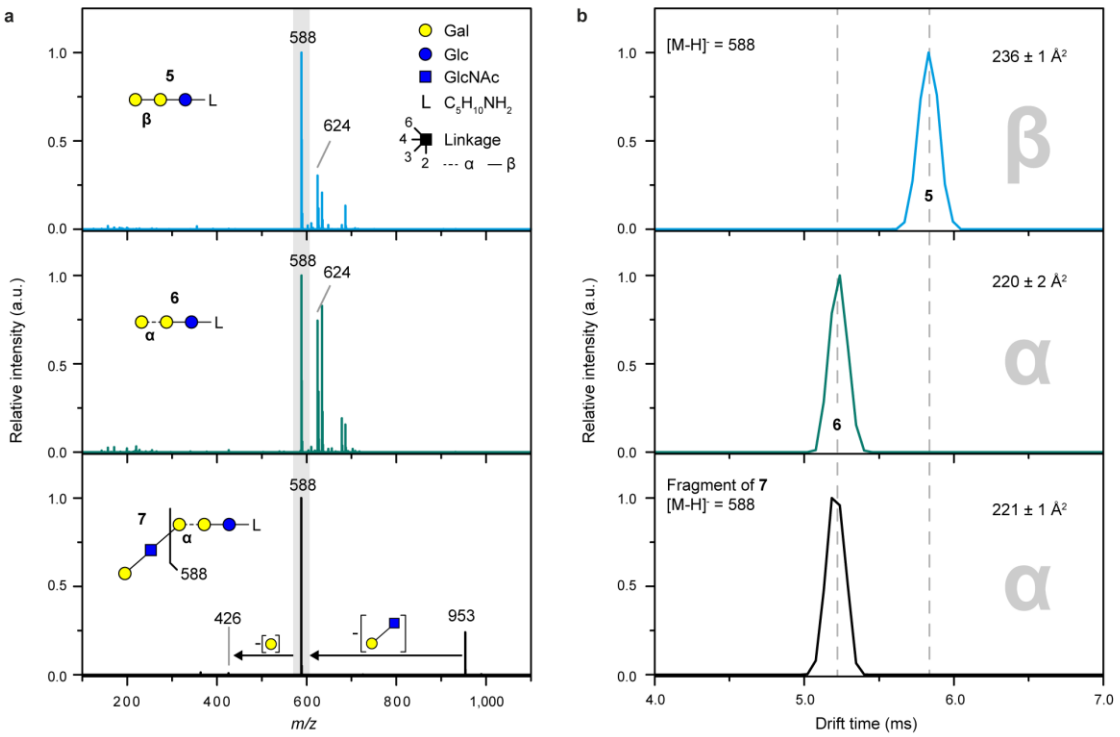
Extended Data Figure 1. Automated glycan synthesis of oligosaccharides 20-25. Ar, 2-methyl-5-tert-butylphenyl; Bn, benzyl; Bz, benzoyl; Cbz, carboxybenzyl; Et, ethyl; Fmoc, fluorenylmethyloxycarbonyl.



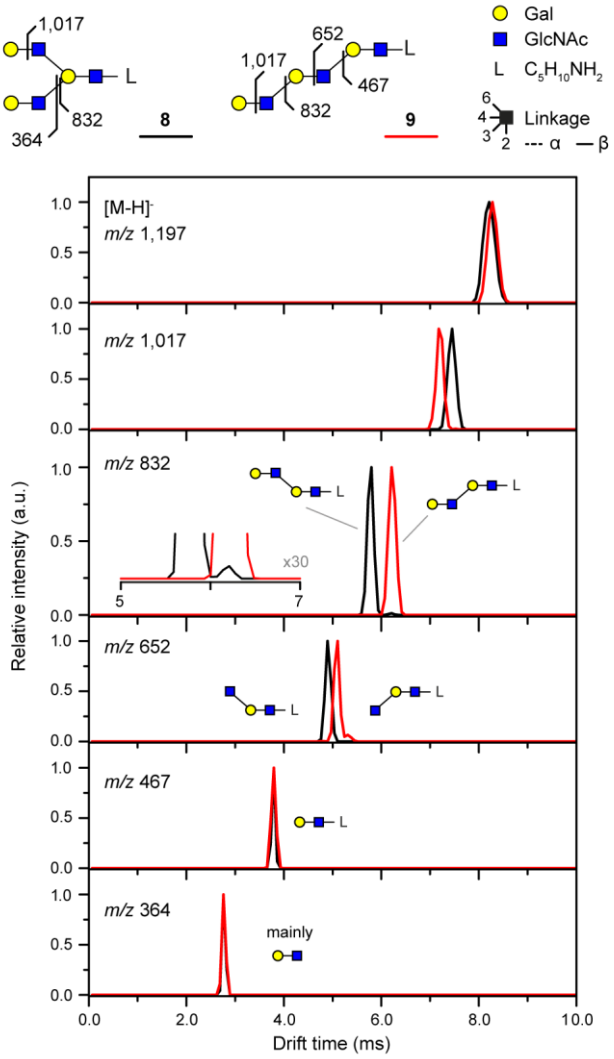
Extended Data Figure 2. Automated synthesis of oligosaccharides 26-28. Ac, acetyl; Bn, benzyl; Bu, butyl; Bz, benzoyl; Cbz, carboxybenzyl; Et, ethyl; Fmoc, fluorenylmethyloxy-carbonyl; lev, levulinoyl; TCA, trichloroacetimidate; UV, ultraviolet.



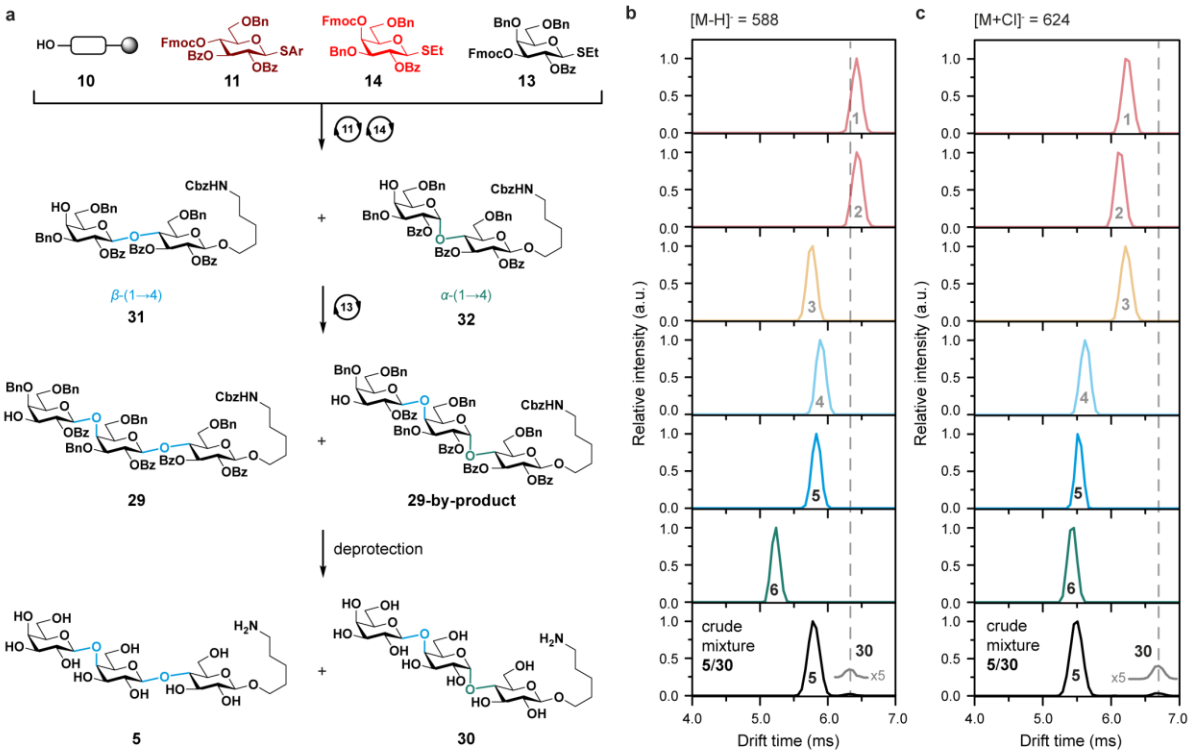
Extended Data Figure 3. Drift-time distributions of trisaccharides 1–6 as different species in positive- and negative-ion mode. The CCS difference between the most compact and the most extended isomer of each set is given as a percentage. Small CCS differences are observed in positive-ion mode (**a**, **b**), which makes an unambiguous identification of the trisaccharides difficult. The largest CCS differences are observed using deprotonated ions (**c**), allowing the identification of linkage isomers (for example, **3** + **6**) and stereoisomers (for example, **2** + **3**). A clear identification of regioisomers with a terminal 1→3 or 1→4 glycosidic bond can be obtained for chloride adducts (**d**).



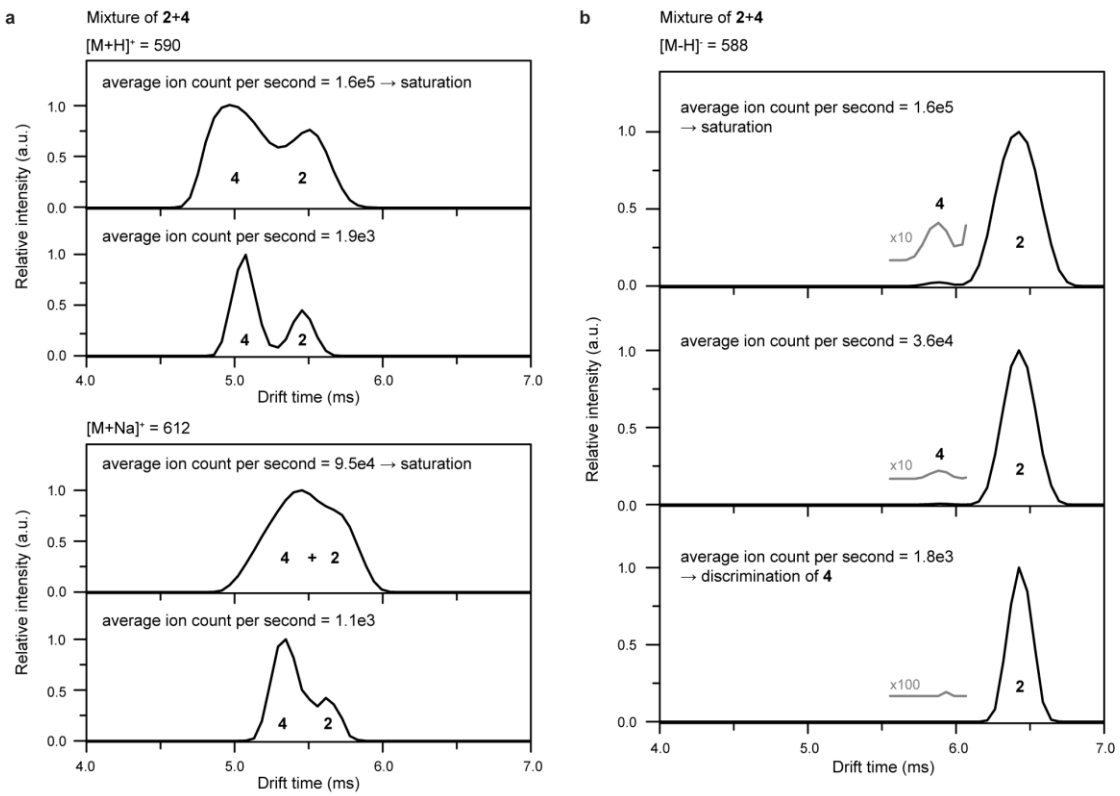
Extended Data Figure 4. Comparison of drift times and CCSs of structurally similar precursor ions and fragments. a, Mass spectra of trisaccharides **5** and **6**, as well as a tandem MS spectrum of **7** (β-Gal-(1→3)–β-GlcNAc-(1→3)–α-Gal-(1→4)–β-Gal-(1→4)–β-Glc-L; L = $C_5H_{10}NH_2$) in negative-ion mode. The pentasaccharide **7** has the same core structure as the trisaccharide **6**. Collision-induced dissociation of deprotonated **7** consequently produces a fragment with the same mass as the deprotonated precursor ion of **6**. **b**, Drift-time distributions of $[M\text{-}H]^-$ = 588 ions. The collision-induced dissociation fragment arising from deprotonated **7** exhibits a drift time and CCS identical to those of the intact deprotonated trisaccharide **6.** This indicates that glycans and glycan fragments with identical structures also exhibit identical CCSs. Seen from a broader perspective, this highlights the exceptional potential of negative-ion CCSs to be used as a diagnostic parameter for glycan sequencing.



Extended Data Figure 5. IM-MS differentiation and identification of the hexasaccharides 8 (black) and 9 (red). As deprotonated ions, **8** and **9** show almost identical drift times and therefore cannot be distinguished. However, smaller collision-induced dissociation fragments containing five, four, or three monosaccharide building blocks (*m/z* 1,017, 832, and 652, respectively) exhibit highly diagnostic drift times. At *m/z* 832, a double peak is observed for the branched oligosaccharide **8** (inset, black trace), because two isomeric fragments are formed. Both fragments can be detected simultaneously using IM–MS, with cleavage at the 3-antenna being clearly preferred. The disaccharide fragments at *m/z* 467 and 364 are identical for **8** and **9** and consequently exhibit identical drift times.



Extended Data Figure 6. Alternative synthesis of oligosaccharide 5 and corresponding IM-MS analysis. a, An alternative route for synthesizing **5** uses building block **11** instead of **12**, which results in a mixture of the disaccharides **31** and **32** and subsequently in a mixture of trisaccharides **5** and **30**. Neither the fully protected trisaccharides **29** and **29-by-product**, nor the deprotected sugars **5** and **30**, can be separated by HPLC. The formation of **29-by-product** can be detected using NMR analysis, but a clear structural assignment is not possible owing to the low relative concentration. **b**, $[M\text{-}H]^-$ = 588 and **c**, $[M+Cl]^-$ = 624 drift-time distributions of trisaccharides **1**–**6** compared to the drift time of the crude mixture consisting of **5** and **30** clearly reveal a content of about 5% by-product **30**. In particular, the drift time of the chloride adduct of **30** is very diagnostic, because it differs considerably from the drift times of all other trisaccharides investigated here.



Extended Data Figure 7. Correlation between signal intensity and ion mobility peak width in mixtures of 2 and 4. a, Drift-time distributions of $[M + H]^+$ and $[M + Na]^+$ ions at high (upper panels) and low (lower panels) signal intensity. The given average ion count per second corresponds to the signal detected for the major isotope peak. High signal intensities result in peak broadening and reduced ion mobility resolution, whereas a considerably improved separation is achieved at lower intensity. **b**, Drift-time distributions of $[M\text{-}H]^-$ ions from a mixture of <1% **4** and >99% **2**. Measurements at high signal intensity can be used to qualitatively detect **4**. At low intensity, however, **4** is discriminated and no signal can be detected.

Extended Data Table 1. Sequences and conditions for automated oligosaccharide synthesis.

sequence	module	details	condition
I	1	2.5 eq. of TMSOTf solution	-20 °C, for 1 min
	2	5 eq. building block for **11**, **12** and **18**, 5 eq. of NIS solution	T_a = -30 °C, t_1 = 5 min T_i = -30 °C, t_2 = 25 min
	3	Fmoc removal	r.t. for 5 min
II	1	2.5 eq. of TMSOTf solution	-20 °C, for 1 min
	2	5 eq. building block for **11**, **12** and **18**, 5 eq. of NIS solution	T_a = -40 °C, t_1 = 5 min T_i = -20 °C, t_2 = 25 min
	3	Fmoc removal	r.t. for 5 min
III	1	2.5 eq. of TMSOTf solution	-20 °C, for 1 min
	2	5 eq. of donor B, 5 eq. of NIS solution Fmoc removal	T_a = -40 °C, t_1 = 5 min T_i = -20 °C, t_2 = 25 min
IV	1	2.5 eq. of TMSOTf solution	-20 °C, for 1 min
	4	5 eq. building block **17**, 5 eq. of TMSOTf solution	T_a = -40 °C, t_1 = 5 min T_i = -20 °C, t_2 = 25 min
	3	Fmoc removal	r.t. for 5 min
V	1	2.5 eq. of TMSOTf solution	-20 °C, for 1 min
	5	5 eq. building block for **19**, 5 eq. of NIS solution	T_a = -40 °C, t_1 = 5 min T_i = -20 °C, t_2 = 25 min
	5	Lev removal	r.t. for 5 min

Eq., equivalent; Fmoc, fluorenylmethyloxycarbonyl; Lev, levulinoyl; NIS, N-iodosuccinimide; r.t., room temperature; TMSOTf, trimethylsilyl trifluoromethanesulfonate.

Extended Data Table 2. Estimated nitrogen CCSs ($^{TW}CCS_{N2}$) for trisaccharides **1-6** and by-product **30**.

substance	ion	$^{TW}CCS_{N2}$ in Å²	STD in Å²	ion	$^{TW}CCS_{N2}$ in Å²	STD in Å²
1	$[M+H]^+$	231.9	0.4	$[M+Na]^+$	236.2	0.6
2	$[M+H]^+$	238.7	0.8	$[M+Na]^+$	242.9	0.9
3	$[M+H]^+$	233.6	0.7	$[M+Na]^+$	239.6	0.6
4	$[M+H]^+$	229.8	0.8	$[M+Na]^+$	236.4	0.6
5	$[M+H]^+$	228.9	0.6	$[M+Na]^+$	233.6	0.7
6	$[M+H]^+$	227.0	0.8	$[M+Na]^+$	232.2	0.5
1	$[M-H]^-$	249.4	1.1	$[M+Cl]^-$	244.4	1.1
2	$[M-H]^-$	249.8	1.5	$[M+Cl]^-$	242.2	1.5
3	$[M-H]^-$	233.2	1.3	$[M+Cl]^-$	244.5	1.2
4	$[M-H]^-$	237.4	0.9	$[M+Cl]^-$	229.7	0.8
5	$[M-H]^-$	235.6	1.0	$[M+Cl]^-$	227.3	0.8
6	$[M-H]^-$	219.9	1.6	$[M+Cl]^-$	224.6	1.4
30	$[M-H]^-$	248.4	0.3	$[M+Cl]^-$	256.7	0.2

CCSs were estimated from travelling-wave (TW) measurements in nitrogen (N_2) using a previously described procedure[29,30]. Each $^{TW}CCS_{N2}$ is an average of three independent measurements with the corresponding standard deviation (STD).

Extended Data Table 3. Relative concentrations of **2** and **3** in the investigated mixtures and their corresponding relative concentration ratio x(**3**)=[**3**]/[**3**+**2**].

rel. conc. **3**	rel. conc. **2**	theoretical x(**3**)	measured Int_{rel}(**3**)	STD
1	100	0.01	0.04	0.011
5	100	0.05	0.07	0.004
11	100	0.10	0.10	0.007
25	100	0.20	0.18	0.005
43	100	0.30	0.27	0.005
56	100	0.36	0.35	0.005
80	100	0.44	0.42	0.010
100	100	0.50	0.49	0.007
100	80	0.56	0.55	0.016
100	56	0.64	0.60	0.005
100	43	0.70	0.69	0.008
100	25	0.80	0.78	0.005
100	11	0.90	0.89	0.010
100	5	0.95	0.93	0.012
100	1	0.99	0.97	0.007

Measured relative intensities $Int_{rel}(\mathbf{3}) = A(\mathbf{3})/A(\mathbf{2}) + A(\mathbf{3})$ were calculated from the drift peak areas (A) of the deprotonated species $[M\text{-}H]^-$ = 588.4. The standard deviation (STD) was obtained from three independent replicates.