Glycan Fingerprinting using Cold-Ion Infrared Spectroscopy

Eike Mucha, Ana Isabel González Flórez, Mateusz Marianski, Daniel A. Thomas, Waldemar Hoffmann, Weston B. Struwe, Heung S. Hahm, Sandy Gewinner, Wieland Schöllkopf, Peter H. Seeberger*, Gert von Helden*, and Kevin Pagel*

E. Mucha, Dr. A. I. González Flórez, Dr. M. Marianski, Dr. D. A. Thomas, W. Hoffmann, S. Gewinner, Dr. W. Schöllkopf, Dr. G. von Helden, Prof. Dr. K. Pagel
Fritz-Haber-Institut der Max-Planck-Gesellschaft
Faradayweg 4–6, 14195 Berlin (Germany)
E-mail: helden@fhi-berlin.mpg.de

E. Mucha, W. Hoffmann, Prof. Dr. K. Pagel
Institut für Chemie und Biochemie der Freien Universität Berlin
Takustraße 3, 14195 Berlin (Germany)
E-mail: kevin.pagel@fu-berlin.de

Dr. H. S. Hahm, Prof. Dr. P. H. Seeberger
Max-Planck-Institut für Kolloid- und Grenzflächenforschung
Am Mühlenberg 1, 14476 Potsdam (Germany)
E-mail: peter.seeberger@mpikg.mpg.de

Prof. Dr. P. H. Seeberger
Institut für Chemie und Biochemie der Freien Universität Berlin
Arnimallee 22, 14195 Berlin (Germany)

Dr. W. B. Struwe
Oxford Glycobiology Institute, Department of Biochemistry, University of Oxford,
OX1 3QU Oxford (United Kingdom)

Abstract

The diversity of stereochemical isomers present in glycans and glycoconjugates poses a formidable challenge for comprehensive structural analysis. Typically, sophisticated mass spectrometry (MS)-based techniques are used in combination with chromatography or ion mobility separation. However, coexisting structurally similar isomers often render an unambiguous identification impossible. Other powerful techniques such as gas-phase infrared (IR) spectroscopy have been limited to smaller glycans, as conformational flexibility and thermal activation during the measurement result in poor spectral resolution. Here, we show that this limitation can be overcome using cold-ion spectroscopy. The vibrational fingerprints of cold oligosaccharide ions exhibit a wealth of well-resolved absorption features that are diagnostic for minute structural variations. The unprecedented resolution of cold-ion spectroscopy coupled with tandem MS may render this the key technology to unravel complex glycomes.

Main Text

Carbohydrates are ubiquitous in nature, and are historically associated with their prominent roles as structural scaffolds and energy sources within the cell. However, shorter chains, often referred to as oligosaccharides or glycans, are also essential in numerous biological signaling processes.[1] The field of glycomics, which aims to comprehensively elucidate the structure and functions of glycans, is currently profiting from technical breakthroughs in automated chemical synthesis and analysis but remains largely under-explored when compared to genomics and proteomics. The inherent structural diversity of oligosaccharides creates major challenges for progress in glycobiology. Whereas DNA and proteins are exclusively assembled in a linear and template-driven fashion, glycans are non-template derived, branched and exhibit a complex stereo- and regiochemistry. Exploring the structure, shape and the resulting functions of oligosaccharides is difficult.

Detailed structural analyses of glycans typically involve liquid chromatography (LC) and/or mass spectrometry-based techniques (MS) that require only small amounts of sample.[2] In addition to the molecular composition, sophisticated MS techniques such as sequential mass spectrometry (MS^n) or chemical derivatization can yield information about the sequence and connectivity of the constituting monosaccharide building blocks.[3] Ion mobility-mass spectrometry (IM-MS) has proven a powerful tool to rapidly separate and identify connectivity and configurational isomers of carbohydrates.[4]

Combining MS and infrared (IR) spectroscopy provides a very sensitive alternative by interrogating the vibrational modes of isolated molecules in the controlled environment of the gas phase. Exploring the characteristic absorption bands of molecules is a routine means to deduce structural information concerning functional groups, hydrogen bonding patterns as well as preferred

molecular conformations. As such, gas-phase IR spectroscopy is a well-established tool for the structural analysis of peptides, proteins and small molecules.

Over the past two decades, the use of gas-phase IR spectroscopy for glycan analysis has been explored. Especially for larger species, spectral congestion due to peak broadening was a major issue limiting the informational content of the resulting signatures. Resolution of carbohydrate ion IR spectra obtained at around 300 K by IR multiple photon dissociation (IRMPD) was very limited already at the mono-[5] and disaccharide[6] level (see Figure S1). Two major factors are likely to cause this peak broadening. First, the conformational flexibility of oligosaccharide ions at room temperature generally appears to be high. More importantly, the sequential absorption of multiple photons in IRMPD leads to considerable thermal activation of the ions before dissociation, such that spectra of hot molecules are obtained.[7] To overcome this limitation experimentally, small neutral carbohydrates have been investigated in the low temperature environment of a free jet expansion. In combination with UV-IR double resonance spectroscopy, enhanced resolution IR-spectra provided information about the structural preferences of isolated mono- and disaccharides and their tendency to form intramolecular hydrogen bond networks.[8] However, mostly oven evaporation was used to transfer oligosaccharides into the gas phase, which, in addition to other technical problems, limits the method to smaller molecules. We overcome this challenge by combining nanoelectrospray ionization (nESI), mass spectrometry and ultra-cold gas-phase IR spectroscopy which leads to diagnostic fingerprints with an exceptionally high resolution and informational content.

The experimental setup (Figure 1a) consists of a mass spectrometer that is connected to the Fritz Haber Institute IR free-electron laser (FHI FEL[9]) and equipped with a source for superfluid helium droplets (for details see SI). Isolated molecular ions are generated using nESI, to allow for the ionization of large molecules. These ions are mass-to-charge selected and accumulated in a linear ion trap. Superfluid helium droplets with an average size of 10^5 helium atoms are subsequently guided through the trap where they can pick up ions, which are rapidly cooled to the internal temperature of the droplet (0.37 K). While the droplet itself is optically transparent, irradiation with intense, narrow-bandwidth IR radiation can lead to the ejection of ions from the droplet if the wavelength is in resonance with a vibrational excitation of the ion. Monitoring the ejection efficiency as a function of the wavelength yields a highly reproducible (Figure S2) IR spectrum of ultra-cold molecular ions.

In order to systematically assess the potential of cold-ion spectroscopy to disentangle the fine structural details in complex oligosaccharides, a series of well-defined aminoalkyl-linked synthetic standards were prepared using automated glycan assembly (for details see SI). Monosaccharides are the simplest sugars and represent the fundamental building blocks of oligosaccharides. Different monosaccharides often share the same mass and only differ in the stereochemistry at single carbon atoms (such as galactose and glucose in Figure 1b). Using the cold-ion spectroscopy technique, IR spectra of protonated, aminopentyl-linked monosaccharides α- and β-galactose **1** and **2**, as well as

β-glucose **3** were recorded (Figure 1c). Each spectrum exhibits a unique absorption pattern with a variety of well-resolved bands, characteristic for each monosaccharide. In some cases, the line-width is equal to the corresponding bandwidth of the laser radiation (FWHM ≈ 4 cm^{-1}). This unprecedented resolution reveals remarkable differences in the overall IR spectra and allows a simple, fingerprint-based discrimination between the isomeric monosaccharides **1**, **2**, and **3**.

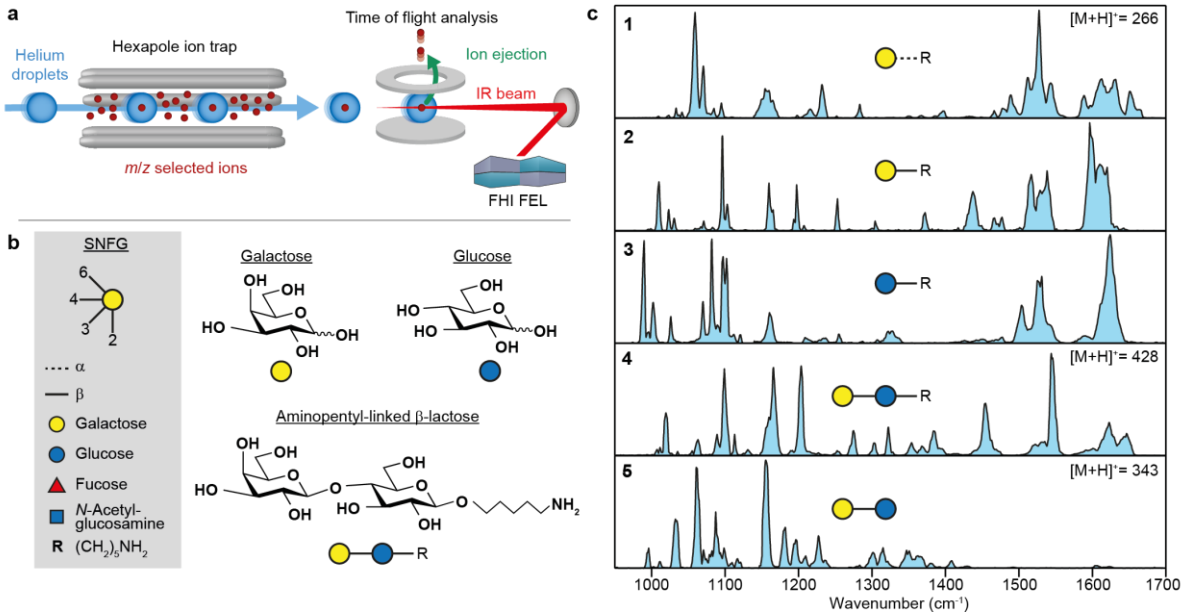


Figure 1. **a**, Schematic diagram of the experimental setup. Ions in the gas phase are mass-to-charge selected and accumulated inside an ion trap. Traversing helium droplets can pick up trapped ions, which immediately cools them to 0.37 K. Subsequently, the doped droplets are irradiated with IR radiation of a defined wavelength using the free-electron laser. **b**, representative structures of carbohydrates investigated in this work and their corresponding symbol nomenclature for glycans (SNFG) that is used throughout this article.[11] **c**, IR spectra of aminoalkyl-linked mono- and disaccharides α-galactose **1**, β-galactose **2**, β-glucose **3**, β-lactose **4** and free lactose **5**, investigated as $[M+H]^+$-ions.

Moving to disaccharides that consist of two building blocks that are joined via a glycosidic bond, the structural complexity increases significantly. Individual monosaccharides can be linked together at different hydroxyl groups (connectivity isomers) and with a certain stereochemical orientation (configurational isomers) leading to a number of isomeric disaccharides. The increasing size of the system and the torsional angles of the glycosidic bond create an additional dimension to the conformational complexity, i.e. the three-dimensional orientation of single isomers. The IR spectra obtained for two lactose variants (Figure 1c) reveal that, although an additional building block is present, β-lactose **4** shows a similar spectral appearance as the monosaccharides described above with a diagnostic pattern of highly-resolved absorption bands. The IR spectrum of free lactose **5**,

where the anomeric center can adopt both an α- as well as a β-configuration and where the location of the charge is uncertain due to the absence of the usually protonated aminoalkyl linker, was recorded for comparison. Nonetheless, the IR spectrum of **5** exhibits a well-resolved pattern between 950 cm^{-1} and 1450 cm^{-1} in which some, but not all bands resemble those of the linker-containing analogue. The observed subtle differences may result from interactions between the charged linker and hydroxyl groups of the disaccharide in **4** or the distinct charge distribution in **5**. The absence of characteristic absorption bands above 1450 cm^{-1} for **5** underscores that this region is mainly governed by vibrational modes of the aminoalkyl linker.

Next, six trisaccharides (**6**-**11**) that share the same aminopentyl-linked β-lactose core and only differ in the connectivity (regiochemistry of the glycosidic bond), configuration (stereochemistry of the glycosidic bond) or composition (type of monosaccharide) of the last building block were analyzed.[4a] These molecules resemble an extreme set of structural isomers, which cannot be distinguished using established LC-MS techniques and therefore serve as a benchmark to systematically assess the utility of this method. The structures of trisaccharides **6–11**, as well as the corresponding IR spectra, are shown in Figure 2. In general, each spectrum exhibits a large number of highly-resolved absorption bands and no significant spectral congestion. The glycan pairs **7**/**8** and **9**/**10** share the same composition and connectivity, but differ in the stereochemistry of the terminal glycosidic linkage (α vs. β). Due to distinct differences in their absorption pattern these configurational isomers can be readily distinguished. Similar results are obtained for connectivity isomers **6**/**11**, **7**/**10** and **8**/**9** for which the last building block is either connected via a 1→3 or a 1→4 glycosidic bond. Again, the fine details revealed in their IR spectra lead to an unambiguous identification of these connectivity isomers. Most striking are the observations for the compositional isomers **6**/**7** and **10**/**11**, which share the same connectivity and configuration, but either exhibit a glucose or galactose unit as the terminal building block. Each trisaccharide pair only differs in the orientation of a single hydroxyl group. This minute structural variation is very difficult to disentangle using established techniques. Surprisingly, distinct differences in the vibrational signatures, especially above 1300 cm^{-1}, are observed in cold-ion spectroscopy. As an example, a characteristic high intensity band at 1450 cm^{-1} is found in the spectrum of trisaccharide **7**, but is absent in the spectrum of the corresponding compositional isomer **6**. Taken together, each of the six trisaccharide isomers **6–11** exhibits a unique IR signature, which allows for their unambiguous identification. Considering the marginal differences between these isomers and the intrinsic conformational flexibility of carbohydrates, this observation highlights the exceptional resolving power of the presented experimental approach.

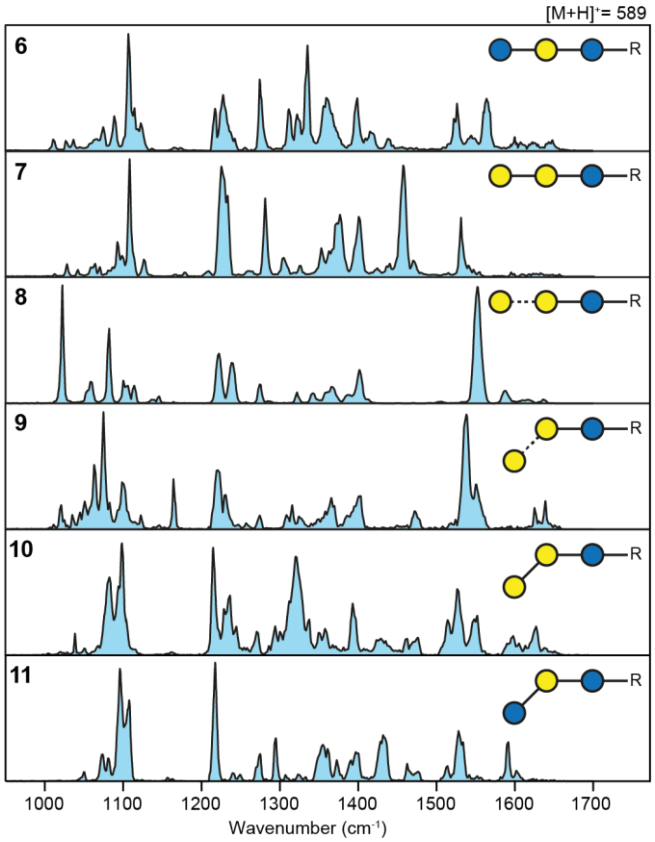


Figure 2. The synthetic trisaccharides **6–11** only differ in the connectivity (1→3 or 1→4), configuration (α or β) or composition (Glc or Gal) of the terminal building block. Despite the marginal structural differences between these species, each trisaccharide can be readily distinguished by its unique absorption pattern that contains a variety of resolved bands. Structures are depicted using the SNFG nomenclature[11] described in Figure 1b.

To extend the scope of the method from synthetic standards to naturally occurring glycans, the blood group related antigens Lewis b (Le^b, **12**) and Lewis y (Le^y, **13**) were investigated as sodium adducts. These isomeric tetrasaccharides each consist of an *N*-acetylglucosamine, a galactose and two fucose units and differ in their glycosidic linkages. In the absence of a linker, the anomeric center is not defined and can adopt an α- as well as a β-configuration. Although the number of expected absorption bands for molecules of this size is large, the corresponding infrared spectra (Figure 3) exhibit a remarkably small number of well-resolved features that allow for an unambiguous discrimination between the two isomeric species. Likewise, the two well-characterized milk sugar tetrasaccharides lacto *N*-neotetraose (**14**) and lacto *N*-tetraose (**15**), which only differ in the connectivity of the terminal galactose unit, exhibit distinct and well-resolved absorption features that allow for their discrimination.

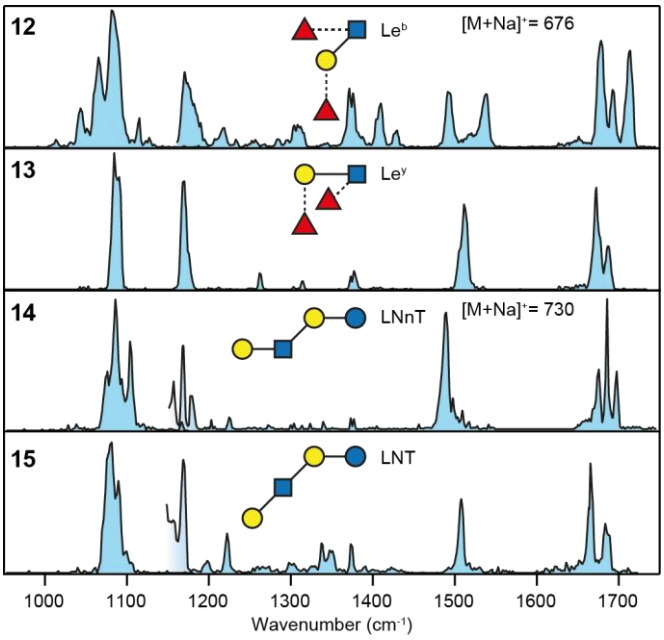


Figure 3. As representatives for biologically glycans, the blood group-related antigens Le^b **12** and Le^y **13**, as well as milk sugar tetrasaccharides lacto-*N*-neotetraose **14** and lacto-*N*-tetraose **15** were investigated as $[M+Na]^+$-ions. The characteristic absorption patterns allow for an unambiguous distinction between the corresponding isomers. Structures are depicted using the SNFG nomenclature[11] described in Figure 1b. The leap around 1150 cm^{-1} resides from two slightly adjusted experimental conditions (for details see SI).

In summary, cold-ion spectroscopy has the potential to overcome all the challenges previously encountered in oligosaccharide analysis. Its unprecedented resolving power provides highly diagnostic absorption patterns—a spectral fingerprint that is unique for each oligosaccharide. For larger glycans, the variety of fragments obtained by sequential MS also enables fragment-based fingerprinting (Fig. S3). These spectra can be stored in databases to provide reference data for quality control or serve as increments for the identification of unknown samples. The sophisticated experimental setup presented here will probably not find any commercial application in the glycosciences, but similar results can likely be obtained on more widely used instruments where cold-ion spectroscopy is performed using cold-ion traps and commercially available bench top laser systems.[10] Although challenging, theoretical spectra can also be calculated (for details see SI). As such, cold-ion spectroscopy has the potential to serve as the ultimate analytical tool to fully disentangle the detailed structure of oligosaccharides and might prove an enabling technology for the glycosciences.

Bibliography

[1] a) A. Varki, *Glycobiology* **1993**, *3*, 97-130; b) R. A. Dwek, *Chem. Rev.* **1996**, *96*, 683-720.

[2] a) K. Mariño, J. Bones, J. J. Kattla, P. M. Rudd, *Nat. Chem. Biol.* **2010**, *6*, 713-723; b) L. Veillon, Y. Huang, W. Peng, X. Dong, B. G. Cho, Y. Mechref, *Electrophoresis* **2017**, Accepted, doi:10.1002/elps.201700042

[3] a) N. Viseux, E. de Hoffmann, B. Domon, *Anal. Chem.* **1997**, *69*, 3193-3198; b) A. Dell, H. R. Morris, *Science* **2001**, *291*, 2351-2356; c) L. R. Ruhaak, A. M. Deelder, M. Wuhrer, *Anal. Bioanal. Chem.* **2009**, *394*, 163-174.

[4] a) J. Hofmann, H. S. Hahm, P. H. Seeberger, K. Pagel, *Nature* **2015**, *526*, 241-244; b) P. Both, A. P. Green, C. J. Gray, R. Sardzik, J. Voglmeir, C. Fontana, M. Austeri, M. Rejzek, D. Richardson, R. A. Field, G. Widmalm, S. L. Flitsch, C. E. Eyers, *Nat. Chem.* **2014**, *6*, 65-74; c) H. Hinneburg, J. Hofmann, W. B. Struwe, A. Thader, F. Altmann, D. Varon Silva, P. H. Seeberger, K. Pagel, D. Kolarich, *Chem. Commun.* **2016**, *52*, 4381-4384; d) C. J. Gray, B. Thomas, R. Upton, L. G. Migas, C. E. Eyers, P. E. Barran, S. L. Flitsch, *Biochim. Biophys. Acta* **2016**, *1860*, 1688-1709.

[5] a) S. Rudić, H.-b. Xie, R. B. Gerber, J. P. Simons, *Mol. Phys.* **2012**, *110*, 1609-1615; b) R. Sagar, S. Rudic, D. P. Gamblin, E. M. Scanlan, T. D. Vaden, B. Odell, T. D. W. Claridge, J. P. Simons, B. G. Davis, *Chem. Sci.* **2012**, *3*, 2307-2313.

[6] N. C. Polfer, J. J. Valle, D. T. Moore, J. Oomens, J. R. Eyler, B. Bendiak, *Anal. Chem.* **2006**, *78*, 670-679.

[7] J. Oomens, B. G. Sartakov, G. Meijer, G. von Helden, *Int. J. Mass Spectrom.* **2006**, *254*, 1-19.

[8] a) R. A. Jockusch, R. T. Kroemer, F. O. Talbot, L. C. Snoek, P. Carcabal, J. P. Simons, M. Havenith, J. M. Bakker, I. Compagnon, G. Meijer, G. von Helden, *J. Am. Chem. Soc.* **2004**, *126*, 5709-5714; b) F. O. Talbot, J. P. Simons, *Phys. Chem. Chem. Phys.* **2002**, *4*, 3562-3565; c) B. Brauer, M. Pincu, V. Buch, I. Bar, J. P. Simons, R. B. Gerber, *J. Phys. Chem. A* **2011**, *115*, 5859-5872.

[9] W. Schöllkopf, S. Gewinner, H. Junkes, A. Paarmann, G. von Helden, H. Bluem, A. M. M. Todd, in *Proc. SPIE Int. Soc. Opt. Eng., Vol. 9512*, **2015**, p. 95121L.

[10] a) N. Heine, K. R. Asmis, *Int. Rev. Phys. Chem.* **2015**, *34*, 1-34; b) J. Roithova, A. Gray, E. Andris, J. Jasik, D. Gerlich, *Acc. Chem. Res.* **2016**, *49*, 223-230; c) L. Voronina, A. Masson, M. Kamrath, F. Schubert, D. Clemmer, C. Baldauf, T. Rizzo, *J. Am. Chem. Soc.* **2016**, *138*, 9224-9233; d) C. T. Wolke, J. A. Fournier, E. Miliordos, S. M. Kathmann, S. S. Xantheas, M. A. Johnson, *J. Chem. Phys.* **2016**, *144*, 074305.

[11] A. Varki, R. D. Cummings, M. Aebi, N. H. Packer, P. H. Seeberger, J. D. Esko, P. Stanley, G. Hart, A. Darvill, T. Kinoshita, J. J. Prestegard, R. L. Schnaar, H. H. Freeze, J. D. Marth, C. R. Bertozzi, M. E. Etzler, M. Frank, J. F. Vliegenthart, T. Lutteke, S. Perez, E. Bolton, P. Rudd, J. Paulson, M. Kanehisa, P. Toukach, K. F. Aoki-Kinoshita, A. Dell, H. Narimatsu, W. York, N. Taniguchi, S. Kornfeld, *Glycobiology* **2015**, *25*, 1323-1324.

Author Contributions

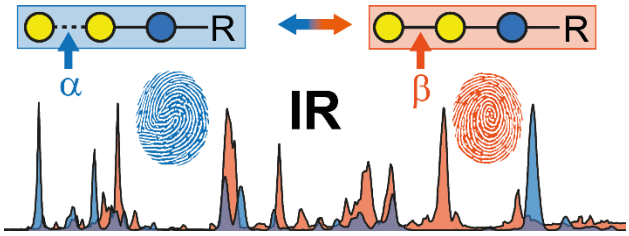
P.H.S., G.v.H. and K.P. designed the research.

E.M., A.I.G.F., M.M., D.A.T., W.H., W.B.S. and H.S.H. performed the research.

S.G. and W.S. operated the IR free-electron laser.

All authors analyzed data and wrote the manuscript.

Table of Content



Fingerprints, please! The diversity of stereochemical isomers present in glycans poses a formidable challenge for comprehensive structural analysis. Here we show that cold-ion spectroscopy has the potential to serve as the ultimate analytical tool to fully disentangle the detailed structure of oligosaccharides. Even for minute structural variations of isomeric glycans, the remarkable resolving power provides unique and highly diagnostic absorption patterns—a true spectral fingerprint.

Keywords:

Carbohydrates, Oligosaccharides, IR Spectroscopy, Mass Spectrometry, Glycans