

Synthesis of an exhaustive library of naturally occurring Galf-Manp and Galp-Manp disaccharides. Towards fingerprinting according to the ring size by advanced mass spectrometry-based IM-MS and IRMPD.

Bénédicte Favreau,^a Ozgur Yeni,^{†b} Simon Ollivier,^{†c,d} Joël Boustie,^e Françoise Le Dévéhat,^e Jean-Paul Guégan,^a Mathieu Fanuel,^{c,d} Hélène Rogniaux,^{c,d} Richard Brédy,^b Isabelle Compagnon,^{*b} David Ropartz,^{*c,d} Laurent Legentil^a and Vincent Ferrières^{*a}

Nature offers a huge diversity of glycosidic derivatives. Amongst numerous structural modulations, the nature of the ring size of hexosides may induce significant differences on both biological and physicochemical properties of the glycoconjugate of interest. On this assumption, we expect that small disaccharides bearing either a furanosyl entity or a pyranosyl residue would give a specific signature, even in the gas phase. On the basis of the scope of mass spectrometry, two analytical techniques to register those signatures were considered, i.e. the ion-mobility (IM) and the infra-red multiple photon dissociation (IRMPD), in order to build up cross-linked databases. D-Galactose occurs in natural products in both tautomeric forms and presents all possible regioisomers when linked to D-mannose. Consequently, the four reducing Galf-Manp disaccharides as well as the four Galp-Manp counterparts were firstly synthesized according to a highly convergent approach, and IM-MS and IRMPD-MS data were secondly collected. Both techniques used afforded signatures, specific to the nature of the connectivity between the two glycosyl entities.

Introduction

A huge diversity of glycans is observed in Nature. This results from the variety of sequencing, branching, connectivity between monosaccharides, minor but very important functional group modulations, the traditional OH-group being opportunistically replaced by uronic acid, deoxy- or NHAc functions, conjugation to lipids, nucleosides or proteins through *O*- or *N*-, α - or β -glycosidic bonds, to name only the most widespread modulations. This biodiversity is also widened by considering the possible tautomerism between pyranosides and furanosides. As a consequence, glycoconjugates are involved in many biological processes such as photosynthesis, intercellular and host-pathogen recognition phenomena, maturation of proteins.¹ They can be also overexpressed by cancer cells² or used as markers of some infection diseases.³ To decipher the complex biological and physicochemical properties of this family of biomolecules, analytical strategies are still developed.⁴ Very recently, low-temperature scanning tunneling microscopy allowed direct imaging different conformers of some oligosaccharides with a sub-nanometer resolution.⁵ On another

side, neutron crystallography is able to study H-bond networks in protein binding sites.^{6, 7} Nevertheless, structural studies are mostly based on NMR⁸ and mass spectrometry (MS) techniques⁹ brilliantly complemented by modeling^{10, 11} and machine learning approaches.¹²⁻¹⁴ Despite these advances, no single analytical technique is able to solve all the structural elements in carbohydrates, and techniques still have to be improved in order to collect and cross-check data, to increase reliability of the analysis. They must also consider minor variations proposed by the living kingdom.

Mammals biosynthesize hexosyl-containing glycoconjugates exclusively in the pyranose form. However, some microorganisms are able to produce hexofuranosyl-containing polysaccharides and conjugates. For instance, galactose,¹⁵⁻¹⁷ *N*-acetyl-galactosamine^{18, 19} and fucose¹⁵ in their furanose form were identified in natural biomolecules with still unclear roles. Moreover, understanding why Nature sometimes prefers to biosynthesize hexofuranosides instead of the thermodynamically more stable pyranosidic counterparts, while the energy cost is higher, is not evident.

An analysis of the natural occurrence of galactofuranosides led us to focus on the galactofuranose (Galf)-mannopyranose (Manp) sequence, and so the four possible regioisomers (Fig. 1). The disaccharides with a (1→2)-, a (1→3)- or a (1→6)-linkage, but not a (1→4)-bond, were identified in microorganisms,¹⁵ some of them being pathogenic, for instance *Trypanosoma*, *Leishmania* and *Aspergillus* species. Surprisingly, the β -D-Galf-(1→4)-D-Manp sequence was exclusively found in lichens.²⁰⁻²⁴ On another side, the corresponding pyranosidic Galp-Manp isomers were identified in different microorganisms: *Klebsiella* for the (1→2)-disaccharide,²⁵ *Rahnella* for the (1→3)-derivative,²⁶ *Leishmania*, *Escherichia coli* for the non-digestible epilactose with the (1→4)-linkage,^{27, 28} and *Burkholderia*,²⁹ KLH

^a Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR – UMR 6226, F-35000 Rennes, France.

^b Université de Lyon, CNRS, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Lyon, France.

^c INRAE, UR BIA, F-44316 Nantes, France.

^d INRAE, BIBS Facility, F-44316 Nantes, France.

^e Univ Rennes, CNRS, ISCR – UMR 6226, F-35000 Rennes, France.

[†] These authors contributed equally to this work.

ORCID IDs: O. Yeni: 0000-0001-6171-3096 S. Ollivier: 0000-0002-7671-1736; J. Boustie: 0000-0002-3936-3859; M. Fanuel: 0000-0001-8384-8266; H. Rogniaux: 0000-0001-6083-2034; I. Compagnon: 0000-0003-2994-3961; D. Ropartz: 0000-0003-4767-6940; L. Legentil: 0000-0003-1402-150X; V. Ferrières: 0000-0002-2780-7774.

Electronic Supplementary Information available.

protein,³⁰ *Salmonella*³¹ for the β -D-Galp-(1 \rightarrow 6)-D-Manp. This resulting orthogonality in Nature strengthens the need for consolidating analytical data to differentiate Galf and Galp residues.

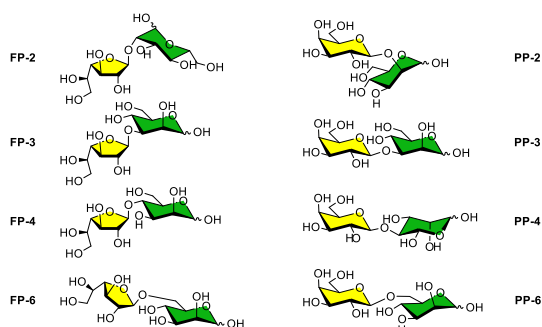


Fig. 1 Defined disaccharides for IM-MS and IRMPD fingerprinting.

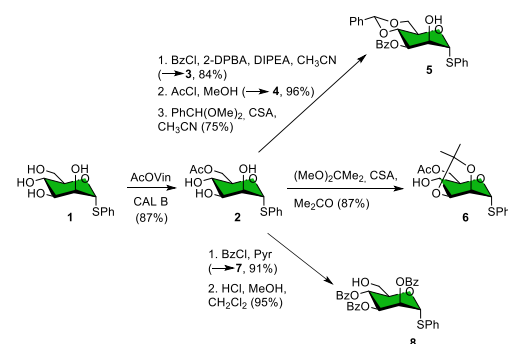
In this context, we have initiated a program dealing with new analytical approaches of hexofuranosides based on mass spectrometry since this technique can be augmented to provide further structural and conformational information. We thus anticipated that ion-mobility mass spectrometry (IM-MS)^{32, 33} and infra-red multiple photon dissociation (IRMPD)³⁴ spectroscopy are indeed suitable tools to assess the impact of the ring size and the branching pattern of small disaccharides on physicochemical properties. We thus expect that these techniques provide specific signatures for distinguishing furanosides from pyranosides. IM-MS separates ions, based on their gas-phase mobility in an electric field,³⁵ and can be used to determine specific structures.³⁶ IRMPD spectroscopy has been widely used, in combination with Quantum chemistry, to elucidate the conformation of biomolecular ions in the gas phase.³⁷ Previous results demonstrate that IRMPD spectroscopy has sufficient structural resolution to yield distinctive signatures and to unambiguously distinguish GalpNAc residues from GalfNAc ones in the case of monosaccharides.³⁸ Here we describe the synthesis of the eight isomers of naturally occurring reducing disaccharides Gal-Man and we demonstrate that individual signatures can be obtained using adequate MS-based strategies, namely IM-MS and IRMPD spectroscopy.

Results and discussion

Synthesis of Galf-containing disaccharides and of their Galp isomers.

To access the panel of the targeted unprotected and reducing disaccharides, four mannopyranosidic acceptors were first synthesized for further coupling with galactofuranosyl or galactopyranosyl donors. Considering the ease of preparation and structural modulation of thioglycosides, and the orthogonal properties towards other families of glycosyl donors, the phenyl thiomannopyranoside **1** was selected as the precursor for all acceptors (Sch 1). The primary position was first selectively acetylated with vinyl acetate thanks to the action of the *Candida antartica* lipase B (CAL B) to afford **2** isolated in 87% yield.³⁹ Interestingly, intermediate **2** can act as acceptor for direct O-3 glycosylation without further protections. The pioneering works of Taylor,⁴⁰ dealing with the activation of the equatorial

hydroxyl of 1,2-diols in carbohydrates with organoborinic catalysts, offers the opportunity to significantly increase the nucleophilicity of OH-3 in the *manno*-series. Consequently, this compound was indeed the key intermediate in our approach since it was either directly engaged in glycosylation reaction [(1 \rightarrow 3)-disaccharides] or converted into the other required acceptors for (1 \rightarrow 2)-, (1 \rightarrow 4)- and (1 \rightarrow 6)-linkages, thus limiting the number of intermediates and synthetic steps. The synthesis of acceptor **5** with free OH-2 started with the specific benzylation of **2** at position 3 in the presence of the borinate ester 2-aminoethylidiphenyl borinate (2-ADPB). The building block **3** thus obtained was then converted into the mannopyranoside **5** after selective deacetylation under acidic conditions and benzyldienation. This synthetic way was preferred to the sequence benzyldienation followed by 3-O-benzylation since it was anticipated the possible formation of the O-2/O-3 acetalization concomitantly with the desired O-4/O-6 protection. Nevertheless, it was taken advantage of the *cis* arrangement of the hydroxyl groups in compound **2** for obtaining the derivative **6** bearing free OH-4 by acido-catalyzed acetalization with dimethoxypropane. Finally, a two-step benzylation-deacetylation afforded **8** with a free primary position.

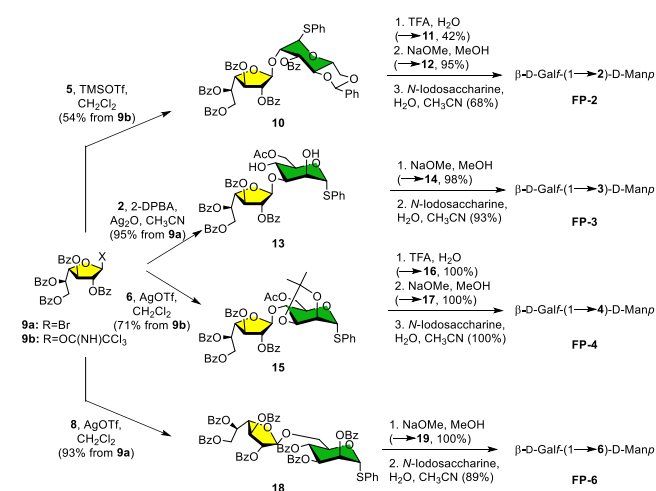


Sch. 1 Synthesis of required mannopyranosidic acceptors.

With the acceptors in hand, it was anticipated that glycosylation at positions 6 and 3 are easier than that at O-2 and O-4.⁴¹ Moreover, it is recognized that conditions should be tuned according to the nature of the leaving group on the glycosyl donors, but also on the basis of the cyclic size (furano vs. pyrano). Our results did not depart from these general observations. The synthesis of disaccharides **FP-2** to **FP-6** was firstly considered starting with the known galactofuranosyl bromide **9a** (Sch 2). As expected, the coupling reactions at O-2 and, in a lesser extent at O-4, were less efficient than the ones at O-3 under the Taylor's conditions and at the primary hydroxyl group. The preferred promoter was the soluble silver triflate in

all cases but the *O*-3 linkage, for which it was shown that silver oxide was more efficient.⁴² For this step, the yields ranged from 24% for the less reactive *O*-2 on the mannoside **5** to 93% and 95% for acceptors **2** and **8**, respectively. It was however observed that limitations to afford **10** and **15** resulted from low reactivity of the donor **9a** for **10** and formation of the corresponding orthoester for **15**. Nevertheless, glycosylation of **5** and **6** with the trichloroacetimidate **9b** in the presence of TMSOTf as the catalyst gave the desired disaccharides **10** and **15** in improved 54% and 71% yields, respectively.

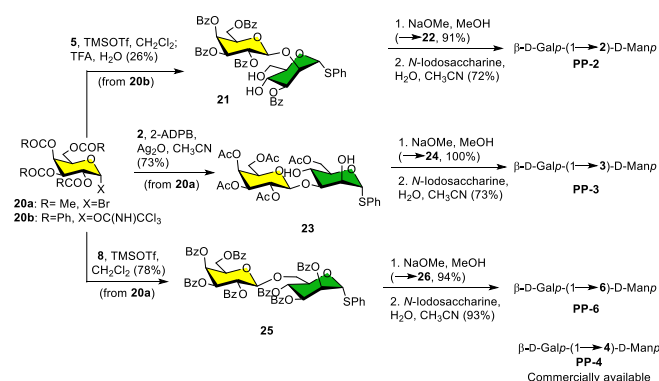
The disaccharides **10**, **13**, **15**, and **18** are characterized by two small J_{H1-H2} coupling constants close or less to 1 Hz for both residues near 5.7 and 5.4 ppm (see Supporting information). Moreover, 2D-NMR analysis revealed that the former signal is correlated with a peak around 85 ppm, while the latter is correlated with the signal at 104 ppm. These data corroborate both the α -mannopyranosidic and the expected β -galactofuranosidic linkages, respectively. Concerning the regioselectivity of the glycosylation of acceptor **2** catalyzed by 2-ADPB, it was also established on the ¹³C NMR data since the signal corresponding to C-3 was significantly up-shifted from 72.2 ppm for the monosaccharide **2** to 78.2 ppm for the disaccharide **13**. Finally, removal of protecting groups were performed under standard conditions: acidic hydrolysis for the acetals, Zemplen transesterification for the esters, and hydrolysis promoted by *N*-iodosaccharine⁴³ for the thioglycosides. It is noteworthy that the resulting reducing disaccharides were obtained in aqueous solution generally as mixtures of anomers as revealed by ¹H NMR and corroborates previous data for **FP-2**⁴⁴ and **FP-3**.⁴⁵ The targeted structures thus tightly approach those ensuing hydrolysis of polysaccharides extracted from biomass. The developed synthetic strategy therefore allowed the synthesis of all possible regioisomers *Galf*-*Manp* and complements well the one recently described for three *Galf*-*Manp* derivatives but bearing an amino spacer arm at the reducing end.⁴⁶ Both anomers were unambiguously characterized thanks to NMR spectra, with a precious help of TOCSY 1D-data, which allows isolation of the signals corresponding to the irradiated proton. Under these conditions, the spectrum of the minor β -anomer in water was fully assigned (See Supporting Information). This strategy was secondly extended to the preparation of the *Galp*-*Manp* isomers (Sch 3), except for the disaccharide **PP-4** which is commercially available. The desired intermediate **21** was obtained starting from the perbenzoylated galactopyranosyl trichloroacetimidate **20b** and the acceptor **5**, and isolated after removal of the benzylidene ring in 26% yield. Further deprotection steps gave the desired **PP-2**⁴⁷ in satisfactory yield. Considering the synthesis of **PP-3**, very efficient glycosidic coupling was obtained starting from the triol **2** and the bromide **20a** in the presence of the organoboron catalyst 2-ADPB and silver oxide. The standard deprotection steps afforded **PP-3** in 53% yield over the three steps. Finally, and surprisingly, the formation of the (1 \rightarrow 6)-linkage raised some difficulties. First attempts with the bromide **20a** activated by silver triflate transferred an acetyl group to the acceptor **8**. Unfortunately, adding diisopropylethylamine as a base (DIPEA)



Sch. 2 Synthesis of the *Galf*-containing disaccharides **FP**.

to avoid the acido-catalyzed transesterification also resulted in the formation of an orthoester in 73% yield (results not shown). This was circumvented by using the trichloroacetimidate **20b** and trimethylsilyl triflate (TMSOTf), so that the disaccharide **25** was isolated in a 78% yield. Action of methylate ion followed by hydrolysis of the thiomannoside gave the wanted disaccharide **PP-6**.⁴⁸

The configuration of the new glycosidic bonds was determined by 1D- and 2D-NMR analysis (see Supplementary Information). The main difference with the *Galf*-*Manp* disaccharides relied on higher J_{H1-H2} values measured for the *Galp* residue (near 8 Hz) compared to the low coupling constant (1 Hz) for the β -galactofuranosides.



Sch. 3 Synthesis of the *Galp*-containing disaccharides **PP**.

Ion mobility-mass spectrometry.

We analyzed the synthesized disaccharides using electrospray ionization (ESI) time-of-flight mass spectrometry, hyphenated with a high-resolution cyclic ion mobility (IM-MS) cell (Waters Select Series Cyclic IMS, Wilmslow, UK).⁴⁹ The set-up allows the gas phase separation of isomers and conformers prior to their mass measurement as shown in the literature.²⁻⁴ Samples were infused in the mass spectrometer at a flow rate of 5 μ L/min, and mass spectra were recorded in the positive ionization mode on a mass-to-charge ratio (m/z) range of 50-1200. We recorded the arrival time distributions (ATDs) of the $[M+Na]^+$ cations or

$[M+Li]^+$ cations after 4 passes around the cyclic mobility cell, *i.e.* approximately a path of 4 meters. The resolving power of ion mobility is proportional to the path length of the ions, which, for 4 passes, yields a theoretical resolution of about 200. Data were recorded with the Quartz v. 4 software and treated with the MassLynx v. 4.2 software (both from Waters, Wilmslow, UK).

First, we measured the ATDs of the eight synthesized products as $[M+Na]^+$ ions at m/z 365.1 (Fig. 2.a). The $[M+Na]^+$ cations are most commonly used for MS (and IM-MS) analyses of neutral oligosaccharides as they offer a good compromise between ionization, separation and MS/MS fragmentation. All the regioisomeric Galf-type **FP** samples individually showed a specific signature. As an example, the **FP-2** sample had an intense signal at 32.2 msec with another ATD of lower intensity at 33.3 msec. In contrast, the **FP-3** sample exhibited a triplet of ATDs respectively at 31.6, 32.5 and 33.5 msec. A less pronounced difference in ATD patterns was observed for the **PP**-type samples. However, by looking carefully at the ATDs, a specific signature could be established. For example, the **PP-3** showed a doublet of ATDs at 31.7 (minor species) and 33.2 msec, while the **PP-4** displayed a doublet at 31.7 (minor species) and 32.5 msec. Notably, the **FP**-type samples appeared to produce more complex ATD patterns than **PP**-type ones. This behaviour might be related to the greater flexibility of the furanose moiety,^{5,6} resulting in different conformations.

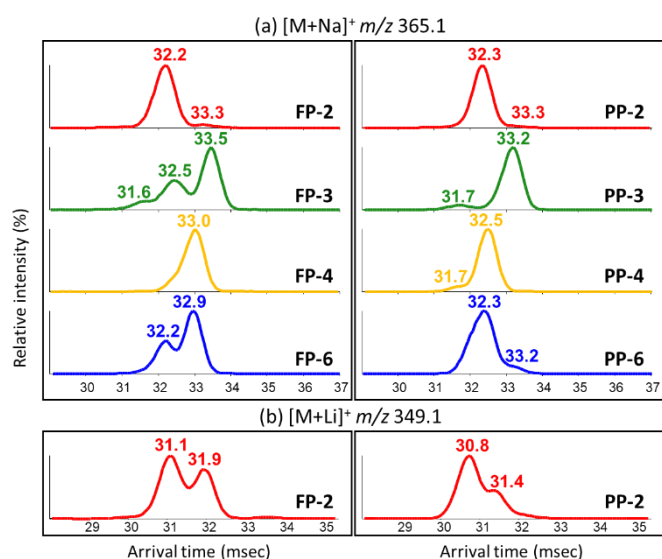


Fig. 2 Extracted monoisotopic ion mobility spectra for (a) the eight synthesized products as $[M+Na]^+$ (m/z 365.1) and (b) the 1→2-linked pair as $[M+Li]^+$ (m/z 349.1). The cyclisation of the galactose moiety is presented as columns, galactofuranose on the left and galactopyranose on the right. The regioisomerism of the glycosidic bond is presented as rows, from top to bottom in (a): 1→2 (red), 1→3 (green), 1→4 (orange), and 1→6 (blue); in (b): 1→2 (red).

Figure 2.a also shows that the related Galp/Galf pairs of disaccharides, analyzed as $[M+Na]^+$ species, have an obvious distinct signature, except for the disaccharides linked through (1→2)-bound. These two samples were thus analyzed as lithium adducts. As shown in Figure 2.b, the $[M+Li]^+$ ions led to an unambiguous IM-MS signature, which clearly differentiated **FP-2** from **PP-2**.

Exploration of the IRMPD sensitivity for the identification of the ring-size and regiochemistry

IRMPD (InfraRed Multiple Photon Dissociation) spectroscopy is an advanced MS-based scheme that adds a dimension of laser spectroscopy to mass spectrometry analysis. Its main advantage is to allow simultaneous measurements of the mass and the infrared fingerprint of an ion. Two ions of the same mass will generally feature distinctive IR fingerprints if they have different spatial arrangements (3D conformation or isomeric form).⁵⁰ Here we explore the sensitivity of this technique to resolve the two types of isomerism present in the synthesized model disaccharides, namely the regiochemistry of the Gal-Man bond and the ring-size of the galactose.

As mentioned in the previous section, Li^+ and Na^+ adducts are often used to increase ion signal and fragmentation in mass spectrometry of carbohydrates.⁵¹ In the context of IRMPD spectroscopy, different ion adducts will lead to different spectroscopic fingerprints. As previously reported, ion adducts can also be chosen to enhance the spectroscopic differentiation of carbohydrate isomers.⁵² Generally, protonated ions feature more distinguishing IR fingerprints than other charge states (*i.e.*, deprotonated ions or alkali adducts). In the present study, the disaccharides of interest are neutral and do not easily protonate, therefore we use ammonium adducts for ionization and enhanced spectroscopic diagnostic.

The eight IRMPD spectra presented in figure 3 were obtained using the set-up previously described.⁵³ In short, the standards were prepared in water/methanol (50/50) and 0.1% of ammonium sulfate was added to promote the formation of

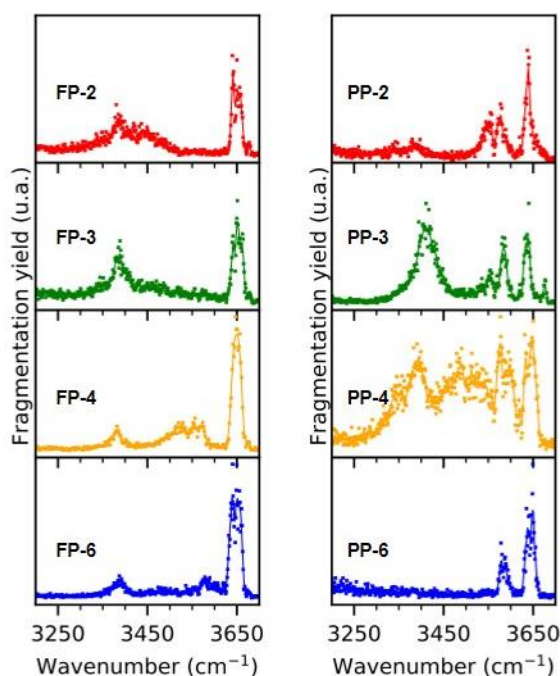


Fig. 3 Left panel: IRMPD spectra of galactofuranosyl-containing disaccharides. Right panel: galactopyranosyl-containing disaccharides. Color code for the regiochemistry of the glycosidic bond: Gal-(1→2)-Man (red), Gal-(1→3)-Man (green), Gal-(1→4)-Man (orange), Gal-(1→6)-Man (blue).

ionic complexes $M.NH_4^+$, which are easily detected by mass spectrometry. The ions of interest are detected at 360 m/z and isolated in the ion trap where they are irradiated by the laser to record their IR fingerprints. The IRMPD spectra are presented between 3200 and 3700 cm^{-1} , which corresponds to the active range of NH and OH vibrations (elongation modes). Firstly, all spectra display a common feature: an intense band around 3650 cm^{-1} . In addition, each spectrum consists of a unique combination of the following elements: (i) a feature at 3575 cm^{-1} which is observed for the pyranose form and absent for the furanose form. This constitutes a straightforward diagnostic of the ring size; (ii) a band around 3375 cm^{-1} and (iii) a region of vibrational activity in the 3450–3500 cm^{-1} range with distinctive widths and relative intensities, depending on the regiochemistry. Overall, the eight spectra are different, thus showing that IRMPD is sensitive to both the galactose ring-size and regiochemistry in the studied disaccharides.

Experimental

Synthesis

All detailed experimental synthesis, including NMR assignments of all intermediates and final disaccharides, are described in the Electronic Supplementary Material.

Ion mobility-mass spectrometry

The synthesised disaccharides were analysed on a Select Series Cyclic IMS™ (Waters, Wilmslow, UK), equipped with a cyclic Travelling Wave Ion Mobility Spectrometry (TWIMS) cell.⁴⁹ ATDs are reported according to Gabelica et al.⁵⁴

Samples were dissolved at 1 $\mu g/mL$ in 50/50 methanol/water (v/v), doped with 0.5 mM NaCl, and then infused at 5 $\mu L/min$ in the electrospray ion source. The β -D-Gal-(1 \rightarrow 2)-Manp furanoside/pyranoside pair was also analyzed at 1 $\mu g/mL$ doped with 0.1 mM LiCl. The disaccharides were analyzed in positive ion mode (in the m/z 50–1200 range), with MS/MS selection of m/z 365.1 in the quadrupole prior to ion mobility separation. The detailed instrumental parameters are given in Suppl. Text.

The following sequence of events was used to perform IMS analyses at a TW height of 22V and TW velocity of 375 m/s: (i) injection in the IMS cell (10 msec); (ii) separation (see Table S1 in Supplementary Information); and (iii) ejection to TOF analyser (13.2 msec). As long as the ions remain in the 4th pass, their arrival times do not change: the slight differences in separation times serve the purpose of setting a window to observe the full ATD.

Data were processed using Driftscope 2.9 and MassLynx 4.2 (Waters, Wilmslow, UK) to extract the monoisotopic ion mobility spectra.

IRMPD

IRMPD spectra are obtained using a commercial mass spectrometer equipped with an electrospray ion source and a Paul trap (LCQ classic ThermoFisher) with custom made modifications. The trap was drilled to allow irradiation of the ion cloud by the beam of a YAG-pumped tunable infrared OPO/OPA system (laserVision), which delivers 8 ns,

10Hz, 13 mJ pulses in the 2700–3700 cm^{-1} . A mechanical shutter ensures the synchronisation of the laser injection with the desired stage of the MS sequence.

The samples were prepared at 30 μM in water/methanol (50/50) and 0.1% of ammonium sulfate was added to promote the formation of ionic complexes $[M+NH_4]^+$. The ions are produced by electrospray ionization and directed to the ion trap through a set of ion multipoles and isolated based on their mass-to-charge ratio at m/z 360. The isolated ions are further held in the trap for a period of 800 ms, during which they are irradiated by the laser. If the IR wavelength is resonant with one of the vibrational frequencies of the ions, multiple photons can be absorbed and their energy redistributed through internal vibrational coupling (IVR). This increases the internal energy of the ion, which results in fragmentation, in a very similar way to traditional collisional activation. The trapped ion (remaining parent and photofragments) are then ejected from the trap and the resulting mass spectrum is recorded. This sequence is repeated thorough the spectral range of interest and the photofragmentation yield is plotted as a function of the wavenumber.

Conclusions

The results presented in this work, combining different fields in glycosciences, allowed to highlight the impact of the ring size of the non-reducing galactosyl residue for each possible connectivity in the Galf-Manp disaccharides. On one hand, this study showed that the metal adducts of the synthesized compounds have different gas-phase conformations, resulting in specific signatures in IM-MS for all Gal-Man disaccharides, according to the ring size for the galactosyl residue and to the connectivity between both entities. The same structural differences were strengthened on another hand, registering IR signals of the disaccharides in the gas phase.

As a result, we expect that the IM-MS and IRMPD data can be used as cross-libraries of reference fingerprints for further analysis of unknown disaccharides, and that they constitute the fundamentals of bringing out Galf for pattern recognition in longer oligosaccharides.³ Therefore, we expect that this study paves the way to the screening of galactofuranose in natural polysaccharides.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors thank the Agence Nationale de la Recherche (ANR) to financially support the ALGAIM-MS project (ANR-18-CE29-0006-02, <https://algaims-35.webself.net/accueil>). The NMR platform PRISM (Université de Rennes 1, <https://biosit.univ-rennes1.fr/prism-plate-forme-rennaise-dimagerie-et-spectroscopie-multimodales>) has also to be acknowledged for

having recording ^1H and ^{13}C spectra at 500 and 125 MHz, respectively.

Notes and references

1. B. B. Allan and W. E. Balch, *Science*, 1999, **285**, 63-66.
2. J. Chen, T. Liu, J. Gao, L. Gao, L. Zhou, M. Cai, Y. Shi, W. Xiong, J. Jiang, T. Tong and H. Wang, *Adv. Sci.*, 2016, **3**, 1600270.
3. D. A. Osibe, S. Lei, B. Wang, C. Jin and W. Fang, *Mycoses*, 2020, **63**, 644-652.
4. C. J. Gray, L. G. Migas, P. E. Barran, K. Pagel, P. H. Seeberger, C. E. Eysers, G.-J. Boons, N. L. B. Pohl, I. Compagnon, G. r. Widmalm and S. L. Flitsch, *J. Am. Chem. Soc.*, 2019, **2019**, 14463-14479.
5. X. Wu, M. Delbianco, K. Anggara, T. Michnowicz, A. Pardo-Vargas, P. Bharate, S. Sen, M. Pristl, S. Rauschenbach, U. Schlickum, S. Abb, P. H. Seeberger and K. Kern, *Nature*, 2020, **582**, 375-378.
6. O. O. Gerlits, L. Coates, R. J. Woods and A. Kovalevsky, *Biochemistry*, 2017, **56**, 4747-4750.
7. F. Manzoni, J. Wallerstein, T. E. Schrader, A. Ostermann, L. Coates, M. Akke, M. P. Blakeley, E. Oksanen and D. T. Logan, *J. Med. Chem.*, 2018, **61**, 4412-4420.
8. A. Gimeno, P. Valverde, A. Arda and J. Jimenez-Barbero, *Curr. opin. Stru. Biol.*, 2020, **62**, 22-30.
9. O. Aizpurua-Olaizola, J. S. Torano, J. M. Falcon-Perez, C. Williams, N. Reichardt and G.-J. Boons, *Trends Anal. Chem.*, 2018, **100**, 7-14.
10. R. Marchetti, S. Perez, A. Arda, A. Imberty, J. Jimenez-Barbero, A. Silipo and A. Molinaro, *ChemistryOpen*, 2016, **5**, 274-296.
11. S.-J. Park, J. Lee, Y. Qi, N. R. Kern, H. S. Lee, S. Jo, I. Joung, K. Joo, J. Lee and W. Im, *Glycobiology*, 2019, **29**, 320-331.
12. S. Kumozaki, K. Sato and Y. Sakakibara, *IEEE/ACM Trans. Comput. Biol. Bioinform.*, 2015, **2015**, 1267-1274.
13. F. Li, C. Li, M. Wang, G. I. Webb, Y. Zhang, J. C. Whisstock and J. Song, *Bioinformatics*, 2015, **31**, 1411-1419.
14. D. Bojar, D. M. Camacho and J. J. Collins, *bioRxiv*, 2020, <https://doi.org/10.1101/2020.01.10.902114doi>.
15. P. Peltier, R. Euzen, R. Daniellou, C. Nugier-Chauvin and V. Ferrières, *Carbohydr. Res.*, 2008, **343**, 1897-1923.
16. M. R. Richards and T. L. Lowary, *ChemBioChem*, 2009, **10**, 1920-1938.
17. C. Marino, A. Rinflerch and R. M. d. Lederkremer, *Future Sci.*, 2017, **3**, FSO199.
18. N. P. Arbatsky, A. S. Shashkov, S. S. Mamyan, Y. A. Knirel, K. Zych and Z. Sidorczyk, *Carbohydr. Res.*, 1998, **310**, 85-90.
19. F. S. Michael, C. M. Szymanski, J. Li, K. H. Chan, N. H. Khieu, S. Larocque, W. W. Wakarchuk, J. R. Brisson and M. A. Monteiro, *Eur. J. Biochem.*, 2002, **269**, 5119-5136.
20. P. A. J. Gorin, *Carbohydr. Res.*, 1985, **142**, 253-267.
21. M. Iacomini, C. L. Schneider and P. A. J. Gorin, *Carbohydr. Res.*, 1985, **142**, 237-251.
22. S. Omarsdottir, B. O. Petersen, B. S. Paulsen, A. Togola, J. O. Duus and E. S. Olafsdottir, *Carbohydr. Res.*, 2006, **341**, 2449-2455.
23. E. R. Carbonero, F. R. Smiderle, A. H. P. Gracher, C. G. Mellinger, G. Torri, T. Ahti, P. A. J. Gorin and M. Iacomini, *Carbohydr. Polym.*, 2006, **63**, 13-18.
24. S. Omarsdottir, J. Freysdottir and E. S. Olafsdottir, *Phytomed.*, 2007, **14**, 179-184.
25. S. K. Das and N. Roy, *Carbohydr. Res.*, 1995, **271**, 177-183.
26. E. L. Zdorovenko, L. D. Varbanets, G. V. Zatonsky, V. V. Kachala, G. M. Zdorovenko, A. S. Shashkov and Y. A. Knirel, *Carbohydr. Res.*, 2008, **343**, 2494-2497.
27. W. Mu, Q. Li, C. Fan, C. Zhou and B. Jiang, *Appl. Microbiol. Biotechnol.*, 2013, **97**, 1821-1827.
28. O. V. Sizova, A. J. Ross, I. A. Ivanova, V. S. Borodkin, M. A. J. Ferguson and A. V. Nikolaev, *ACS Chem. Biol.*, 2011, **6**, 648-657.
29. L. Pol-Fachin, R. V. Serrato and H. Verli, *Carbohydr. Res.*, 2010, **345**, 1922-1931.
30. H. Geyer, M. Wuhrer, A. Resemann and R. Geyer, *J. Biol. Chem.*, 2005, **280**, 40731-40748.
31. P. J. Garegg, C. Henrichson, T. Norberg and P. Ossowski, *Carbohydr. Res.*, 1983, **119**, 95-100.
32. J. Hofmann and K. Pagel, *Angew. Chem. Int. Ed.*, 2017, **56**, 8342-8349.
33. S. Warnke, A. B. Faleh, R. P. Pellegrinelli, N. Yalovenko and T. R. Rizzo, *Faraday Discuss.*, 2019, **217**, 114-125.
34. C. J. Gray, I. Compagnon and S. L. Flitsch, *Cur. Opin. Stru. Biol.*, 2020, **62**.
35. V. Gabelica and E. Marklund, *Cur. Opin. Chem. Biol.*, 2018, **42**, 51-59.
36. J. Hofmann, H. S. Hahm, P. H. Seeberger and K. Pagel, *Nature* 2015, **526**, 241-244.
37. N. C. Polfer and J. Oomens, *Mass Spec. Rev.*, 2009, **28**, 468-494.
38. B. Schindler, L. Legentil, A.-R. Allouche, V. Ferrières and I. Compagnon, *Phys. Chem. Chem. Phys.*, 2019, **21**, 12460-12467.
39. H. K. Weber, H. Weber and R. J. Kazlauskas, *Tetrahedron Asym.*, 1999, **10**, 2635-2638.
40. M. S. Taylor, *Acc. Chem. Res.*, 2015, **48**, 295-305.
41. C. M. Pedersen, J. Olsen, A. B. Brka and M. Bols, *Chem. Eur. J.*, 2011, **17**, 7080-7086.
42. L. Legentil, Y. Cabezas, O. Tasseau, C. Tellier, F. Daligault and V. Ferrières, *J. Org. Chem.*, 2017, **82**, 7114-7122.
43. P. K. Mandal and A. K. Misra, *Synlett*, 2007, 1207-1210.
44. P. A. J. Gorin, E. M. Barreto-Bergter and F. S. Cruz, *Carbohydr. Res.*, 1981, **88**, 177-188.
45. C. Marino, A. Chioconci, O. Varela and R. M. d. Lederkremer, *Carbohydr. Res.*, 1998, **311**, 183-189.
46. V. B. Krylov, D. A. Argunov, A. S. Solovev, M. I. Petruk, A. G. Gerbst, A. S. Dmitrenok, A. S. Shashkov, J.-P. Latgé and N. E. Nifantiev, *Org. Biomol. Chem.*, 2018, **16**, 1188-1199.
47. S. Huang, H. Yu and X. Chen, *Angew. Chem. Int. Ed.*, 2007, **46**, 2249-2253.
48. H. Y. a. X. Chen*, *Org. Lett.*, 2006, **8**, 2393-2396.
49. K. Giles, J. Ujma, J. Wildgoose, S. Pringle, K. Richardson, D. Langridge and M. Green, *Anal. Chem.*, 2019, **91**, 8564-8573.
50. B. Schindler, L. Barnes, G. Renois, C. Gray, S. Chambert, S. Fort, S. Flitsch, C. Loison, A. R. Allouche and I. Compagnon, *Nat. Commun.*, 2017, **8**, 973.
51. M. J. Kailemia, L. R. Ruhaak, C. B. Lebrilla and I. J. Amster, *Anal. Chem.*, 2014, **86**, 196-212.
52. B. Schindler, L. Barnes, C. J. Gray, S. Chambert, S. L. Flitsch, J. Oomens, R. Daniel, A. R. Allouche and I. Compagnon, *J. Phys. Chem. A.*, 2017, **121**, 2114-2120.

53. B. Schindler, J. Joshi, A.-R. Allouche, D. Simon, S. Chambert, V. Brites, M.-P. Gaigeot and I. Compagnon, *Phys. Chem. Chem. Phys.*, 2014, **16**, 22131-22138.
54. V. Gabelica, A. A. Shvartsburg, C. Afonso, P. Barran, J. L. P. Benesch, C. Bleiholder, M. T. Bowers, A. Bilbao, M. F. Bush, J. L. Campbell, I. D. G. Campuzano, T. Causon, B. H. Clowers, C. S. Creaser, E. D. Pauw, J. Far, F. Fernandez-Lima, J. C. Fjeldsted, K. Giles, M. Groessl, C. J. H. Jr, S. Hann, H. I. Kim, R. T. Kurulugama, J. C. May, J. A. McLean, K. Pagel, K. Richardson, M. E. Ridgeway, F. Rosu, F. Sobott, K. Thalassinou, S. J. Valentine and T. Wyttenbach, *Mass Spec. Rev.*, 2019, **38**, 291-320.