Exploration of high-resolution ion mobility-mass spectrometry for structural glycosciences

Simon Ollivier

▶ **To cite this version:**

Simon Ollivier. Exploration of high-resolution ion mobility-mass spectrometry for structural glycosciences. Other. Nantes Université, 2022. English. ⟨NNT : 2022NANU4069⟩. ⟨tel-04175261⟩

HAL Id: tel-04175261

https://theses.hal.science/tel-04175261v1

Submitted on 2 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HAL Authorization

DOCTORAT ECOLOGIE BRETAGNE GEOSCIENCES LOIRE AGRONOMIE ALIMENTATION

Nantes Université

THESE DE DOCTORAT DE

NANTES UNIVERSITE

ECOLE DOCTORALE N° 600
Ecole doctorale Ecologie, Géosciences, Agronomie et Alimentation
Spécialité : « *Sciences agronomiques* »

Par
Simon OLLIVIER

Exploration of high-resolution ion mobility-mass spectrometry for structural glycosciences

Volume 1

Thèse présentée et soutenue à Nantes, le 29 novembre 2022
Unité de recherche : INRAE, UR Biopolymères Interactions et Assemblages

Rapporteurs avant soutenance :

Dr. Valérie GABELICA	Directrice de Recherche, INSERM, Institut Européen de Chimie et Biologie, Pessac
Pr. Guillaume VAN DER REST	Professeur, Université Paris-Saclay

Composition du Jury :

Président :	Pr. Patrick GIRAUDEAU	Professeur, Nantes Université
Examinateurs :	Dr. Sarah CIANFERANI	Directrice de Recherche, CNRS, Université de Strasbourg
	Dr. Philippe DUGOURD	Directeur de Recherche, CNRS, Université Lyon 1
	Dr. Laurent LEGENTIL	Chargé de Recherche, HDR, CNRS, Ecole Nationale Supérieure de Chimie de Rennes
Dir. de thèse :	Dr. Hélène ROGNIAUX	Ingénieure de Recherche, HDR, INRAE Nantes
Co-enc. de thèse :	Dr. David ROPARTZ	Ingénieur d'Etudes, INRAE Nantes

Invité(s)

Dr. Kevin GILES	Senior Research Fellow, Waters Corp., Wilmslow, UK

A mes parents,

“All that is gold does not glitter,

Not all those who wander are lost;

The old that is strong does not wither,

Deep roots are not reached by the frost”

J.R.R. Tolkien

The Lord of the Rings: The Fellowship of the Ring

Remerciements

Ce travail de thèse a été réalisé au sein du laboratoire de spectrométrie de masse de la plateforme BIBS (Bioressources : Imagerie, Biochimie & Structure) de l'unité Biopolymères Interactions Assemblages (INRAE, UR 1268).

Je remercie tout d'abord ma directrice de thèse Hélène Rogniaux et mon co-encadrant David Ropartz pour l'accueil chaleureux qu'ils m'ont réservé au sein de ce laboratoire. Merci pour votre implication, pour vos conseils, pour avoir été chaque jour à l'écoute, et pour m'avoir suivi pendant ces trois ans. Ce fut (et restera) un plaisir de travailler avec vous !

Je tiens à remercier Valérie Gabelica et Guillaume van der Rest pour avoir accepté d'évaluer ce travail en qualité de rapporteurs, ainsi que Sarah Cianférani, Philippe Dugourd, Patrick Giraudeau et Laurent Legentil pour avoir accepté de faire partie de mon jury.
I would also like to thank Kevin Giles for accepting our invitation to be part of my PhD committee. Also, many thanks to you and the rest of the Waters team for always being available to answer any question we could have regarding Cyclic IMS.

Je remercie également tous les partenaires avec lesquels j'ai eu la chance de collaborer tout au long de ma thèse, et dont les noms apparaîtront au fil des Chapitres.
Je souhaite notamment remercier tous les partenaires du projet ALGAIMS, Vincent, Laurent, Bénédicte, Joël, Françoise, Camille, Mathis, Isabelle et Oznur, pour toutes ces discussions enrichissantes au fil de ces trois ans. Tout ne s'est pas passé comme prévu, mais je pense qu'au final nos résultats en valaient la peine.
Un grand merci également aux collègues du LSMBO avec qui j'ai eu l'occasion de travailler, et tout particulièrement à Sarah et Evolène. Le pharmacien qui sommeille en moi a pris beaucoup de plaisir à travailler sur des mAbs avec vous ! Evolène, merci beaucoup pour ton enthousiasme durant les quelques semaines que tu as passées chez nous, pour tes conseils sur la SEC, et pour m'avoir fait découvrir des outils de traitement des données IMS qui m'ont changé la vie !
Enfin, merci à Gabrielle et toute l'équipe du TBI, pour les heures passées à comprendre ce qui pouvait bien se passer dans ce fameux milieu réactionnel.

Je remercie aussi tous les membres de la plateforme BIBS – trop nombreux pour que je les cite tous ici – pour la bonne ambiance qui règne au sein de l'équipe, que ce soit pendant les séminaires ou pendant la vie de tous les jours.
Un immense merci notamment au reste des membres du laboratoire de spectrométrie de masse pour tous ces moments partagés ensemble – y compris aux anciens, Bastien, Léry et Jessica, qui sont partis vers de nouveaux horizons. Mathieu, Mehdi, merci pour votre soutien pendant ma rédaction, et aussi pour notre petite escale à New York sur le chemin de l'ASMS. Je ne suis pas près d'oublier les quelques 25 km par jour (et mes pieds non plus !) mais la Statue de la Liberté, le Vanderbilt et le T-Rex en valaient largement la peine !! ;-) Adrien, je te souhaite un bon rétablissement, ainsi que bon courage et beaucoup de réussite pour la fin de ta thèse. J'espère que tu pourras vite revenir, ton bureau fait un peu vide en face du mien… Et enfin, merci à Alice, qui m'a bien aidée au cours de son stage, pour avoir bien géré les manips de SEC et donc pour m'avoir libéré du temps pour la rédaction.

Pour mes années pré-thèse, je souhaite également remercier à nouveau Joël Boustie et Françoise Lohézic-Le Dévéhat, mais aussi Béatrice Legouin-Gargadennec et Marylène Chollet-Krugler pour m'avoir fait découvrir le monde de la Recherche pendant mes années d'études à la fac' de Pharmacie de Rennes. Je souhaite également remercier les personnes qui m'ont initié à la spectrométrie de masse : merci à Nicolas Le Yondre et Philippe Jéhan du CRMPO à Rennes de m'avoir fait découvrir cette technique un peu magique, et merci aussi à Damien Olivier-Jiminez et Pierre-Marie Allard de m'avoir enseigné les réseaux moléculaires lors de mon séjour à Genève dans l'équipe de Jean-Luc Wolfender – je les utilise encore, même si j'ai arrêté la métabo'.

Evidemment, je n'oublie pas tous les proches qui m'ont accompagné pendant toutes ces années, et tout particulièrement ma famille et mes parents.
Maman, Papa, merci pour votre soutien sans faille depuis toujours. Merci de m'avoir encouragé à suivre la voie qui me plaisait, de m'avoir supporté quand j'étais d'humeur irascible, et de m'avoir encouragé dans les moments difficiles.

Table of Contents

TABLE OF ILLUSTRATIONS 9

A. Figures 9

B. Tables 13

ABBREVIATIONS 15

SYNTHESIS OF SCIENTIFIC PRODUCTIONS 17

A. Publications 17

B. Oral communications 17

C. Poster presentations 18

D. Teaching activity 18

E. Internship supervision 18

GENERAL INTRODUCTION 19

PART I – INTRODUCTION 23

CHAPTER 1 – CHARACTERIZING CARBOHYDRATES: AN ALL-TIME ANALYTICAL CHALLENGE 25

A. The "basic" structure of carbohydrates 25

B. How seemingly minor structural changes highly impact the properties of carbohydrates 29

C. Overview of analytical methods for carbohydrate characterization: advantages and limitations 32

CHAPTER 2 – THE SPECIAL CONTRIBUTION OF ION MOBILITY TO GLYCOMICS: WHAT ADVANTAGES DOES IT BRING COMPARED TO OTHER ANALYTICAL TECHNIQUES? 43

A. History of Ion Mobility Spectrometry (IMS) 43

B. Expected benefits of high-resolution IMS (HR-IMS) for glycosciences 52

C. State of the art of IMS applied to carbohydrate analysis at the beginning of this thesis 54

D. Current developments in IMS-MS for glycosciences 55

CHAPTER 3 – THE SELECT SERIES CYCLIC IMS INSTRUMENT 57

A. Instrument geometry 57

B. Operating modes 65

PART II – CAN CYCLIC IMS RETAIN AND RESOLVE ALL TYPES OF ISOMERISM IN OLIGOSACCHARIDES? 73

CHAPTER 4 – DISTINGUISHING LINKAGE REGIOISOMERS AND FURANOSE/PYRANOSE TAUTOMERS IN GALACTOMANNAN DISACCHARIDES 75

A. Context....75

B. Objectives 76

C. Summary of the results 76

D. Conclusions....78

E. Scientific Communication 79

CHAPTER 5 – INVESTIGATING THE BEHAVIOR OF FURANOSIDES UPON COLLISION-INDUCED DISSOCIATION 99

A. Context....99

B. Objectives 100

C. Summary of the results 100

D. Conclusions....103

E. Scientific Communication 103

CHAPTER 6 – ANOMERIC MEMORY OF THE GLYCOSIDIC BOND IN IMS^N EXPERIMENTS 119

A. Context....119

B. Objectives 119

C. Summary of the results 120

D. Conclusions....122

E. Scientific Communication 123

CHAPTER 7 – SEPARATION OF UNDERIVATIZED MONOSACCHARIDE STEREOISOMERS 135

A. Context....135

B. Objectives 136

C. Methods 136

D. Results and discussion....137

E. Conclusions....145

F. Scientific Communication 146

PART II – CONCLUSIONS 147

PART III – CAN WE ACHIEVE A COMPLETE STRUCTURAL DETERMINATION OF COMPLEX OLIGOSACCHARIDES USING CIMS? 149

CHAPTER 8 – FIRST INSIGHTS TOWARDS COMPLETE OLIGOSACCHARIDE SEQUENCING WITH CIMS 151

A. Context....151

B. Objectives 151

C. Methods *153*
D. Results and Discussion *154*
E. Conclusions and Perspectives *165*
F. Scientific Communication *166*
CHAPTER 9 – MOLECULAR NETWORKING OF TANDEM IMS DATA TO ORGANIZE GLYCOMICS DATASETS 167
A. Context *167*
B. Objectives *168*
C. Summary of the results *169*
D. Conclusions and Perspectives *171*
E. Scientific Communication *172*
PART III – CONCLUSIONS AND PERSPECTIVES 213
PART IV – CAN CIMS BE USED TO ADDRESS THE POLYSACCHARIDE CHALLENGE? 215
CHAPTER 10 –....COUPLING SIZE EXCLUSION CHROMATOGRAPHY WITH CIMS FOR THE ANALYSIS OF INTACT POLYSACCHARIDES 217
A. Context *217*
B. Objectives *218*
C. Methods *218*
D. Results and Discussion *221*
E. Conclusions and Perspectives *239*
F. Scientific Communication *239*
PART IV – CONCLUSIONS AND PERSPECTIVES 241
GENERAL CONCLUSIONS AND PERSPECTIVES 243
REFERENCES 249
RÉSUMÉ EN FRANÇAIS 263
A. La mobilité ionique cyclique peut-elle préserver et résoudre tous les types d'isomérie dans les oligosaccharides ? *266*
B. Peut-on réaliser une détermination structurale complète d'oligosaccharides complexes en utilisant la cIMS ? *267*
C. La cIMS peut-elle être utilisée pour faire face au défi de la caractérisation structurale à l'échelle des polysaccharides ? *269*
D. Perspectives *271*

Table of Illustrations

A. Figures

Figure 1. Example of D- and L- oses, illustrated using glucose. 25
Figure 2. Linear structures of the monosaccharide building blocks with 3, 4, 5 and 6 carbons (Fischer representation) for the (A) D-aldoses and (B) D-ketoses series. Glucose, used to illustrate the different types of isomerisms in this Chapter, is highlighted in blue. 26
Figure 3. Illustration of the four possible heterocycles that can be formed by a monosaccharide, exemplified on D-glucose 27
Figure 4. Mutarotation and tautomeric equilibrium, illustrated on glucose 27
Figure 5. A. Numbering of the carbons/OH groups illustrated on D-glucose. B. Existing glucobiose species. 28
Figure 6. Examples of xyloglucan oligosaccharides (to retain readability, the linkage regioisomerism is not considered in the representations). 29
Figure 7. Most common chemical modifications found in polysaccharides 30
Figure 8. Schematic model of the structure of pectin[12,13] (to retain readability, the linkage regioisomerism is not considered in the representations). 31
Figure 9. Principle of anion exchange chromatography 34
Figure 10. Principle of hydrophilic interaction liquid chromatograhy 35
Figure 11. Principle of ion pairing chromatography 35
Figure 12. Schematic of the dipole-dipole interaction occurring between the analytes and the PGC phase 36
Figure 13. Principle of SEC separation 36
Figure 14. Principle of MALDI ionization 38
Figure 15. Domon and Costello nomenclature for the fragmentation of oligosaccharides. 39
Figure 16. Timeline indicating some key milestones of the history of IMS, and the number of publications indexed in the web of science (WOS) for IMS and IMS-MS 43
Figure 17. Top 15 research categories associated with the topic “ion mobility spectrometry”, by number of publications indexed in the Web Of Science for the period 2006–present[63] 44
Figure 18. Principle of DTIMS separation 46
Figure 19. Principle of TWIMS separation 46
Figure 20. Principle of TIMS separation 47
Figure 21. Principle of DMA separation 48
Figure 22. Principle of FAIMS separation 48
Figure 23. Ion cyclotron mobility spectrometer introduced by Merenbloom *et al.*[101] 51
Figure 24. The cyclic TWIMS cell of the SELECT SERIES Cyclic IMS (courtesy of Waters Corp.) 51
Figure 25. Serpentine SLIM built using PCBs (© Pacific Northwest National Laboratory) 52
Figure 26. Resolving power required to separate the different types of isomers, adapted from Dodds *et al.*[98] 53
Figure 27. Scheme of the cIMS instrument 57
Figure 28. A. Schematic of a basic electrospray ion source. B. Schematic of the ZSpray ion source from Waters. 58

Figure 29. The three models of ion formation from ESI droplets. 59
Figure 30. Schematic of a quadrupole mass filter 60
Figure 31. Stability diagram of ions in a linear quadrupole 60
Figure 32. A. The ExD cell module for the Select Series Cyclic IMS, B. schematic of the ExD cell. (© e-MSion, Inc.) 62
Figure 33. A-B. Schematic of the cIMS cell and C. schematic of the multifunction array region (courtesy of Waters Corp.) 63
Figure 34. A. Principle of a reflectron TOF analyzer; B. The triple-reflection (W-mode) TOF analyzer available on the cIMS setup (courtesy of Waters Corp.) 64
Figure 35. Profile of the resolving power of a TOF analyzer (for a multiple-reflection TOF, as a function of the time of flight and of the number of turns). Reprinted from Ayet San Andrés *et al*.[137] 64
Figure 36. Sequence of events used for single/multipass separations 65
Figure 37. A. Site of fragmentation in the mass spectrometer during CID-IMS experiments; B. Graph plotting the IMS arrival time vs *m/z* of the fragments in a CID-IMS experiment 66
Figure 38. A. Site of fragmentation in the mass spectrometer during IMS-CID experiments; B. IMS arrival time vs *m/z* of the fragments of IMS-CID data 67
Figure 39. Sites of fragmentation in the mass spectrometer during CID-IMS-CID experiments 67
Figure 40. Flowchart describing the different kinds of cIMS-based selection of ions that can be performed using the SELECT SERIES Cyclic IMS 68
Figure 41. Sequence of events used for head-and-tail IMS^n experiments, illustrated for a single stage of selection 70
Figure 42. Sequence of events used for slicing IMS^n experiments, illustrated for a single stage of selection and fragmentation 71
Figure 43. A. Schematic of the multifunction array region used to set the direction of the ion path. B. Potential energy diagrams for the selection to pre-store and reinjection stages, without and with activation upon reinjection. Color code for the direction of the array's traveling wave: forward (yellow), sideways (orange), backward (red). 72
Figure 44. Overview of the isomerisms that can be found in carbohydrates 73
Figure 45. Structures of the Gal-Man disaccharides in the reference library of IM-MS and IRMPD spectra 76
Figure 46. Extracted monoisotopic ion mobility spectra for the eight synthesized products as $[M + Na]^+$ (*m/z* 365.1). Standard deviations are given for three replicates. For both parts, the left panel corresponds to the galactofuranosyl-containing disaccharides and right panel to the pyranosyl-containing ones. Color for the regiochemistry of the glycosidic bond: Gal-(1→2)-Man (red), Gal-(1→3)-Man (green), Gal-(1→4)-Man (orange), and Gal-(1→6)-Man (blue) 77
Figure 47. Extracted monoisotopic ion mobility spectra for the eight synthesized products as $[M + Li]^+$ (*m/z* 349.1). Standard deviations are given for three replicates. For both parts, the left panel corresponds to the galactofuranosyl-containing disaccharides and right panel to the pyranosyl-containing ones. Color for the regiochemistry of the glycosidic bond: Gal-(1→2)-Man (red), Gal-(1→3)-Man (green), Gal-(1→4)-Man (orange), and Gal-(1→6)-Man (blue) 78
Figure 48. Structures of the studied trisaccharides and reference disaccharides 99
Figure 49. Reference profile of Man-(α1,6)-Man (top) and IM profile of the fragment at *m/z* 351 from Gal*f*-(β1,4)-Man-(α1,6)-Man (bottom) 100
Figure 50. Reference profile of Gal*f*-(β1,4)-Man (top) and IM profile of the fragment at *m/z* 351 from Gal*f*-(β1,4)-Man-(α1,6)-Man (bottom) 101

Figure 51. Interconversion between the two conformational populations of Man-(α1,6)-Man, as observed using multistage IMS 102
Figure 52. Comparison of the experimental CCS profile of the Y_2 fragment from Gal*f*-(β1,4)-Man-(α1,6)-Man to the conformers calculated by MD analysis of Man-(α1,6)-Man 102
Figure 53. Reaction catalyzed by Uhgb_MS with unknown reaction product(s) 120
Figure 54. Drift time distributions of the IMS^2 fragments from pure β1,2- and α1,2-linked mannooligosaccharide standards (blue) and of the mannotriose from the reaction medium (red), analyzed as $[M + Li]^+$ adducts. The non-reducing end only yielded B_2 fragments of interest (left column), while the reducing end only yielded Y_2 fragments (right column). 121
Figure 55. Drift time distributions of the IMS^2 fragments from pure β1,2- and α1,2-linked mannotriose (blue) of the IMS^2 fragments from the mannotetraose produced by Uhgb_MS (red) and of the IMS^3 fragments produced after refragmentation of the Y_3 mannotetraose fragment (violet), analyzed as $[M + Li]^+$ adducts. 122
Figure 56. Static traveling wave separation. The WH remains constant over time. 137
Figure 57. Separation with a ramping of the traveling wave height. Different WH are applied over time. 138
Figure 58. Schematic of the strategy combining static and ramping wave height separations to resolve additional populations 139
Figure 59. Fischer representation of the 13 studied monosaccharides, sorted by isomeric family.....140
Figure 60. Arrival time distributions for the Glc, Gal and Man monosaccharides (colored), and for an equimolar mixture of the three hexoses at 0.1 µg/mL (solid black line). Inset: result of the Drifscope peak detection for the mixture.....141
Figure 61. Arrival time distributions for the Ara, Xyl and Rib monosaccharides (colored), and for mixtures of Ara 10 µg/mL, Xyl 1 µg/mL and Rib 0.5 µg/mL (left panel) or only Ara 10 µg/mL and Xyl 1 µg/mL (right panel) (solid black lines). Inset: results of the Drifscope peak detections for the mixtures.....142
Figure 62. Arrival time distributions for the Fuc and Rha monosaccharides (colored), and for an equimolar mixture of the two deoxyhexoses at 5 µg/mL (solid black line). Inset: result of the Drifscope peak detection for the mixture.....143
Figure 63. Arrival time distributions for the GlcA and GalA monosaccharides (colored), and for an equimolar mixture of the two HexA at 5 µg/mL (solid black line). Inset: result of the Drifscope peak detection for the mixture. 144
Figure 64. Arrival time distributions for the GlcNAc, ManNAc and GalNAc monosaccharides (colored), and for an equimolar mixture of the three HexNAc at 1 µg/mL (solid black line). Inset: result of the Drifscope peak detection for the mixture. 145
Figure 65. Cyclic IMS can retain and resolve all types of isomerism in oligosaccharides ...147
Figure 66. Conceptual scheme of the sequencing of oligosaccharides by IMS^n.....152
Figure 67. Examples of CID fragmentation spectra produced by activation upon reinjection from the pre-store (+190 V) for four different Glc-OS DP5, displaying different linkage anomerism and regioisomerism. Parent ion ($(Glc)_5$) was isolated at *m/z* 835.3 as $[M + Li]^+$. The monosaccharidic fragments of interest (C_1 or Y_1, corresponding to $(Glc)_1$) are highlighted in purple at *m/z* 187.1.....155
Figure 68. Monosaccharidic fragments released by fragmenting the $[M + Li]^+$ species isolated at *m/z* 835.3 of maltopentaose (α1,4-glucan) and cellopentaose (β1,4-glucan) in red, are compared to the reference profile of Glc in blue. 155

Figure 69. Monosaccharidic fragments released by fragmenting the $[M + Li]^+$ species isolated at *m/z* 835.3 of branched and linear mannopentaose in red, are compared to the reference profile of Man in blue. 156
Figure 70. The monosaccharidic fragment released by fragmenting the $[M + Li]^+$ species isolated at *m/z* 835.3 of galactopentaose (β1,4-galactan) in red, is compared to the reference profile of Gal in blue. 157
Figure 71. The monosaccharidic fragment released by fragmenting the $[M + Li]^+$ species isolated at *m/z* 817.3 of xylohexaose (β1,4-xylan) in red, is compared to the reference profile of Xyl in blue.....157
Figure 72. Comparison of the monosaccharidic fragments released from different gluco-oligosaccharides isolated at *m/z* 835.3 as $[M + Li]^+$ with α (DP5 dextran in (α1,6) and maltopentaose in (α1,4)) or β linkages (cellopentaose in (β1,4) and laminaripentaose in (β1,3)). In blue, the IM profile of pure glucose is given for reference. 158
Figure 73. Comparison of the dehydrated monosaccharidic fragments (*m/z* 169.1) released from different Glc-OS of DP5 with α or β linkages. Parent ion was isolated at *m/z* 835.3 as $[M + Li]^+$. IM profiles of the dehydrated monosaccharidic fragment are presented (B_1 fragment, detected at *m/z* 169.1 as $[M + Li]^+$ ion). The presented ATDs were acquired after three passes around the IMS cell (static WH conditions). 160
Figure 74. Comparison of the dehydrated monosaccharidic fragments (B_1, *m/z* 169.1 as $[M + Li]^+$ species) released from Glc-OS DP5, Man-OS DP5 and Gal-OS DP5. Parent ion was isolated at *m/z* 835.3 ($(Hex)_5$ as $[M + Li]^+$). The presented ATDs were acquired after three passes around the IMS cell (static WH conditions). 161
Figure 75. Fragmenting a Glc-(β1,4)-Glc or a Man-(β1,4)-Man bond gives the same B_1 fragment. The involved dehydration mechanism is presented for both Glc and Man subunits (R = oligosaccharidic chain). 161
Figure 76. IMS^n study of maltopentaose. The selection steps and separation times are given for each stage of the sequences.....163
Figure 77. CID fragmentation spectrum of ^{18}O-labeled XXXG xyloglucan (+210 V upon reinjection in the mobility cell), for the $[M + Li]^+$ species at *m/z* 1071.4. Color code for the annotation of the fragments: peak highlighted in green = unambiguous fragment, in red = ambiguous fragment, in blue = loss of $H_2{}^{18}O$ from another fragment. 164
Figure 78. IMS/IMS fragmentation spectra of β1,4-glucans (A: cellotetraose, and B: cellopentaose) *vs*. α1,4-glucans (C: maltotetraose, and D: maltopentaose). The precursor ion is identified by an arrow in each spectrum. 168
Figure 79. Families of oligosaccharides included in the set of compounds used to evaluate the interest of molecular networking (DP4 are shown as example).....169
Figure 80. MS-based molecular network for the set of standard oligosaccharides 169
Figure 81. IM-based molecular network for the set of standard oligosaccharides 170
Figure 82. Extended IM-MN in which the initial IM-MN was challenged with structures containing moieties (subunits or linkages) from multiple standards in the initial network. These structures are highlighted in yellow.....171
Figure 83. Illustration of the LC-CID-IMS acquisitions, and IM-MN for the *Lasallia pustulata* digest sample. 172
Figure 84. Examples of mass spectra when infusing A. proteins and B. polysaccharides in electrospray. C. Separation of charge states in spectrum B using IMS. 221
Figure 85. SEC separation of three native protein standards (β-lactoglobuline, yeast alcohol dehydrogenase and pyruvate kinase) on two columns of 200 Å (top) and 450 Å porosity (bottom).....224

Figure 86. A. Isopropanol disrupts hydrophobic interactions between the analyte and the SEC material; B. Total ion chromatogram for the P-100 pullulan standard (*ca.* 100 kDa) at varying concentrations of isopropanol. 225
Figure 87. A. Salt disrupts electrostatic interactions between the analyte and the SEC material; B. Total ion chromatogram for the P-100 pullulan standard (*ca.* 100 kDa) at varying concentrations of $AcONH_4$. 226
Figure 88. Total ion chromatograms for the pullulan standards P-200, P-100, P-50 and P-20. 227
Figure 89. A. Regions integrated in the TIC of the P-50 fraction; B. Mass spectrum for the peak at 19 min (blue); C. Mass spectrum for the peak at 22 min (pink); D. Extracted ion chromatograms for the base peaks of spectra B and C (respectively, *m/z* 4398 and 3063). 228
Figure 90. Backbone pattern of pullulans. 229
Figure 91. A–C. P-200; D–F. P-100; G–I. P-50; J–L. P-20 standards. A,D,G,J. Slices done in the TIC to perform deconvolution; B,E, H, K. *m/z* vs retention time plots; C, F, I, L. Deconvolved Zero Charge Mass Spectra vs retention time plots. 230
Figure 92. Total ion chromatograms and 2D retention time *vs* arrival time for fractions A. F30.8, B. F40.6, C. F50.7, and D. F60.8. 233
Figure 93. A. Selected SEC-IM regions for the F50.7 AX fraction, B. Mass spectra for the selected regions. 234
Figure 94. A. TIC of the F50.7 AX fraction, with the selected region highlighted in orange; B. Raw mass spectrum for the orange region (in the *m/z* 3500–6000 range); C. Zoom-in on the *m/z* 4050–4750 range of the raw spectrum; D. Smoothed mass spectrum for the *m/z* 4050–4750 range. 236
Figure 95. A. Smoothed mass spectrum for the F50.7 AX fraction (see Figure 94B); B. Extracted ion chromatograms for DP/z 28–38 (smoothed); C. Examples of 2D retention time vs arrival time for DP/z 28, 33, and 38. D. Mass spectrum of DP/z 28 in population ① using only the SEC dimension as a filter. E. Example mass spectra for DP112 and DP140, isolated at DP/z 28 in population ① using SEC-IM filtering (the theoretical isotopic patterns for these species are given in gray for reference). 238

B. Tables

Table 1. Summary of the cIMS sequences used to separate the monosaccharides 145
Table 2. Voltages used to set the direction of the ions in the multifunction array region of the Cyclic IMS spectrometer for each type of cIM event, with the default settings (top) and the settings adapted for polysaccharides (bottom). The voltages changed for analysis of high molecular weight biomolecules are highlighted in red. 220
Table 3. Specific deconvolution parameters changed in UniDec to deconvolve the Zero Charge Mass Spectrum of each standard. 231
Table 4. Structural characteristics of the selected arabinoxylan fractions. 232

Abbreviations

3C5NBA: 3-carboxy-5-nitrophenyl boronic acid

ACN: acetonitrile

ADHy: yeast Alcohol DeHydrogenase

ALGAIMS (ANR Project): Analysis of Lichen Galf-containing polysaccharides using IRMPD and ion-mobility MS

APCI: Atmospheric Pressure Chemical Ionization

Ara: Arabinose

ATD: Arrival Time Distribution

AX: ArabinoXylan

CCS: Collision Cross Section

CCSD: Collision Cross Section Distribution

CDMS: Charge Detection MS

CEM: Chain-Ejection Model

CID: Collision Induced Dissociation

cIMS: cyclic IMS

CRM: Charge Residue Model

DDA: Data Dependent Acquisition

DESI: Desorption ESI

DMA: Differential Mobility Analyzer

DMS: Differential Mobility Spectrometry

DP: Degree of Polymerization

DT: Drift Tube

DTD: Drift Time Distribution

DTIMS: Drift Tube IMS

ECD: Electron Capture Dissociation

EDD: Electron Detachment Dissociation

EED: Electronic Excitation Dissociation

EI: Electron Impact

EID: Electron-Induced Dissociation

EIM: Extracted Ion Mobility spectrum

ESI: ElectroSpray Ionization

ETD: Electron Transfer Dissociation

ExD: Electron-based Dissociation methods (also the name of the fragmentation cell commercialized by e-MSion, Inc.)

FAIMS: Field Asymmetric IMS

FID: Flame Ionization Detector

FITDOG: Fenton's Initiation Towards Defined Oligosaccharide Groups

FTICR: Fourier-Transform Ion Cyclotron Resonance

FTMS: Fourier-Transform MS

Fuc: Fucose

FWHM: Full Width at Half Maximum

Gal: Galactose

GalA: Galacturonic Acid

Gal*f*: Galactofuranose

GalNAc: *N*-Acetyl-Galactosamine

Gal-OS: Galacto-OligoSaccharide

Gal*p*: Galactopyranose

GC: Gas Chromatography

GH: Glycoside Hydrolase

Glc: Glucose

GlcA: Glucuronic acid

GlcNAc: *N*-Acetyl-Glucosamine

Glc-OS: Gluco-OligoSaccharide

GNPS: Global Natural Products Social MN platform

He-CTD: Helium-Charge Transfer Dissociation

Hex: Hexose

HexA: Hexuronic Acid

HexNAc: *N*-Acetyl-Hexosamine

HILIC: Hydrophilic Interaction LIquid Chromatography

HMW: High Molecular Weight

HPAEC: High-Performance Anion Exchange Chromatography

HR-IMS: High-Resolution IMS

IEM: Ion Evaporation Model

IgG: Immunoglobulin G

IM: Ion Mobility

IM-MN: Ion Mobility-Molecular Networking

IMS/IMS: tandem IMS

IMS: Ion Mobility Spectrometry

IMS^n: multistage IMS

IP(R): Ion Pairing (Reagent)

IPA: IsoPropyl Alcohol

IRL: Internal Residue Loss

IRMPD: InfraRed Multiple Photon Dissociation

LAESI: Laser Ablation ESI

LC: Liquid Chromatography

LMW: Low Molecular Weight

mAb: monoclonal Antibody

MALDI: Matrix-Assisted Laser Desorption/Ionization

MALS: Multi-Angle Light Scattering

Man: Mannose

Man-OS: Manno-OligoSaccharide

Man*p*: Mannopyranose

MD: Molecular Dynamics

MeOH: Methanol

MN: Molecular Network

Mn: Number-average molecular mass

MS/MS: Tandem MS

MS: Mass Spectrometry

MSI: Mass Spectrometry Imaging

MW: Molecular Weight

Mw: Weight-average molecular mass

nIM-MS: native IM-MS

NMR: Nuclear Magnetic Resonance

nMS: native MS

oa-TOF: orthogonal acceleration TOF

OS: OligoSaccharide

PAD: Pulsed Amperometry Detector

PCB: Printed Circuit Board

Pent: Pentose

PGC: Porous Graphitic Carbon

PK: Pyruvate Kinase

PS: PolySaccharide

RF: RadioFrequency

RG: RhamnoGalacturonan

Rha: Rhamnose

Rib: Ribose

R_p: Resolving power

RP-LC: Reverse Phase Liquid Chromatography

SEC: Size Exclusion Chromatography

SIMMS²: Shotgun IMS-MS Sequencing

SLIM: Structures for Lossless Ion Manipulation

SNFG: Symbol Nomenclature For Glycans

SUPER (SLIM): Serpentine Ultralong Path with Extended Routing (SLIM)

t_A: Arrival Time

t_d: drift Time

TIC: Total Ion Chromatogram

TIMS: Trapped IMS

TOF: Time-Of-Flight

TW: Traveling Wave

TWIMS: Traveling Wave IMS

Uhgb_MS: Unknown human gut bacterium_Mannoside Synthase

Vpp: Volt peak-to-peak

VUV: Vacuum UltraViolet

WH: Wave Height

WOS: Web Of Science

WV: Wave Velocity

XIC: eXtracted Ion Chromatogram

XUV: eXtreme UltraViolet

XUV-DPI: XUV-Dissociative PhotoIonization

XUVPD: eXtreme UltraViolet PhotoDissociation

Xyl: Xylose

Xyl-OS: Xylo-OligoSaccharide

αCD: α-CycloDextrin

β-LG: β-LactoGlobuline

Synthesis of scientific productions

A. Publications

- S. Ollivier, L. Tarquis, M. Fanuel, A. Li, J. Durand, E. Laville, G. Potocki-Veronese,* D. Ropartz* and H. Rogniaux (2021) **Anomeric retention of carbohydrates in multistage cyclic ion mobility (IMS^n): *de novo* structural elucidation of enzymatically produced mannosides**. *Analytical Chemistry* 93 (15), 6254-6261 (https://doi.org/10.1021/acs.analchem.1c00673)
- B. Favreau, O. Yeni,[§] S. Ollivier,[§] J. Boustie, F. Le Dévéhat, J.-P. Guégan, M. Fanuel, H. Rogniaux, R. Brédy, I. Compagnon,* D. Ropartz,* L. Legentil and V. Ferrières* (2021) **Synthesis of an exhaustive library of naturally occurring Gal*f*-Man*p* and Gal*p*-Man*p* disaccharides. Toward fingerprinting according to the ring size by advanced mass spectrometry-based IM-MS and IRMPD.** *The Journal of Organic Chemistry* 86, 6390-6405 (https://doi.org/10.1021/acs.joc.1c00250)
- S. Ollivier, M. Fanuel, H. Rogniaux* and D. Ropartz (2021) **Molecular networking of high-resolution tandem ion mobility spectra: a structurally relevant way of organizing data in glycomics?** *Analytical Chemistry* 93 (31), 10871-10878 (https://doi.org/10.1021/acs.analchem.1c01244)
- S. Ollivier, M. Fanuel, H. Rogniaux* and D. Ropartz (2022) **Using a cyclic ion mobility spectrometer from tandem ion mobility experiments.** *JoVE (Journal of Visualized Experiments)* 179, e63451 (https://doi.org/10.3791/63451)
- D. Ropartz, M. Fanuel, S. Ollivier, A. Lissarrague, M. Benkoulouche, L. A. Mulard, I. André, D. Guieysse and H. Rogniaux* (2022) **Combination of high-resolution multistage ion mobility and tandem MS with high energy of activation to resolve the structure of complex chemoenzymatically synthesized glycans.** *Analytical Chemistry* 94 (4), 2279-2287 (https://doi.org/10.1021/acs.analchem.1c04982)
- E. Deslignière, S. Ollivier, A. Ehkirch, A. Martelet, D. Ropartz, N. Lechat, O. Hernandez-Alba, J.-M. Menet, S. Clavier, H. Rogniaux, B. Genet, and S. Cianférani* (2022) **Combination of IM-Based Approaches to Unravel the Coexistence of Two Conformers on a Therapeutic Multispecific mAb.** *Analytical Chemistry*, *94* (22), 7981–7989 (https://doi.org/10.1021/acs.analchem.2c00928)

B. Oral communications

- S. Ollivier*, M. Fanuel, A. Li, E. Laville, D. Guieysse, L. Tarquis, G. Potocki-Veronese, H. Rogniaux and D. Ropartz ; **Can high-resolution IMS/IMS solve the intra-chain anomerism of carbohydrates?** ; RCJSM, 10/12/2020, online
- S. Ollivier*, M. Fanuel, D. Ropartz and H. Rogniaux ; **Molecular networking of tandem ion mobility data and its implications for glycomics** ; JFSM 2021, 22/06/2021, online

- S. Ollivier*, M. Fanuel, D. Ropartz and H. Rogniaux ; **Molecular networking of tandem ion mobility data and its implications for glycomics** ; Journées scientifiques de l'ED EGAAL, 30/06 & 01/07/2021, online
- S. Ollivier*, M. Fanuel, D. Ropartz and H. Rogniaux ; **New prospects for the structural characterization of oligosaccharides: an overview of novel analytical methods relying on cyclic ion mobility mass spectrometry** ; EPNOE 7th International Polysaccharide Conference, 11-15/10/2021, Nantes, FR
- S. Ollivier* ; **Aperçu de nouvelles méthodes analytiques basées sur la spectrométrie de masse à mobilité ionique cyclique : nouvelles perspectives pour la caractérisation structurale des oligosaccharides.** Animation scientifique de l'axe Analyses Structurales et Métabolomique de BioGenOuest, 09/12/2021, online
- S. Ollivier* ; **Ion mobility-mass spectrometry : new prospects for the structural characterization of complex polysaccharides** ; Young Researchers Day of the BIA research unit, 07/07/2022, Nantes, FR

C. Poster presentations

- S. Ollivier*, M. Fanuel, C. Guitteny, F. Le Dévéhat, D. Ropartz and H. Rogniaux ; **From tandem ion mobility on a Cyclic IMS platform to ion mobility-molecular networking of glycans** ; 70th ASMS Conference on Mass Spectrometry and Allied topics, 08/06/2022, Minneapolis, MN, USA
- S. Ollivier*, L. Legentil, M. Fanuel, L.-P. David, O. Yeni, I. Compagnon, V. Ferrières, D. Ropartz and H. Rogniaux ; **On the behavior of Gal*f*-containing oligosaccharides upon fragmentation – a multistage high-resolution Ion Mobility study** ; Analytics2022, 06/09/2022, Nantes, FR

D. Teaching activity

- Workshop Analytics2022 (Nantes, FR) : **Mobilité ionique cyclique de haute résolution pour l'analyse de carbohydrates complexes et d'anticorps monoclonaux thérapeutiques.** Duration: 2h. S. Ollivier et E. Deslignère (LSMBO Strasbourg). 05/09/2022
- Master 2 CQPS (Nantes Université). Lecture on Ion Mobility Spectrometry. Duration: 3h (yearly). S. Ollivier. 2021 and 2022

E. Internship supervision

- Alice Pétrau, 2 months, 2022, L3 "Sciences pour la santé", La Rochelle Université. **Evaluation de la chromatographie d'exclusion stérique couplée à la spectrométrie de masse pour l'analyse de polysaccharides d'origine végétale.**

General introduction

Carbohydrates are the most abundant biopolymers on Earth: they are ubiquitous among living organisms and are found among animals, plants, algae, fungi and microorganisms alike. From a phylogenetic standpoint, they are the least conserved biomolecules in Life. Indeed, depending on the producing organism and even on its life stage, the organ or the tissue, their structures and compositions vary strongly. Their structures can even be impacted by environmental or abiotic conditions, for instance tides impact the sulfation pattern of algal polysaccharides. Carbohydrates serve a variety of biological functions ranging from cell signaling to key structural roles (e.g., in the cell wall). They also find numerous *ex vivo* applications including in food, cosmetic, pharmaceutical industries and in biorefinery. However, the biological properties of carbohydrates—as well as the resulting industrial applications—largely depend on their structure. Many subtle features in carbohydrates' structure are critical to their properties but remain complicated to decipher by current analytical methods. The consequence is that many carbohydrates remain structurally unknown, limiting the study of their structure-function relationship and making carbohydrates a probably underexploited resource.

For other types of biopolymers such as proteins, mass spectrometry (MS) has established as a core tool affording the characterization of up to thousands of proteins per hour, including a fine determination of their potential modifications. In structural glycosciences however, MS lags far behind. A major setback in the use of MS for the characterization of oligo- and polysaccharides is the preponderance of isomerism in these compounds. First, numerous types of isomeric subunits exist among the main families of carbohydrates (*i.e.*, among hexoses, pentoses, uronic acids, etc.). Each of these families in fact comprises different stereoisomers, which can even sometimes be the orientation of a single hydroxyl function (epimers). A further complexity comes from the coexistence of cyclization tautomers, a single type of monosaccharide being able to form 6-membered rings (pyranose form) and/or 5-membered rings (furanose form). The formation of the glycosidic bond between subunits increases the complexity by adding two additional types of isomerism: (i) the regioisomerism of the linkage, which can happen on any hydroxyl function, and (ii) the α/β anomerism of the bond, resulting from the α and β anomers existing for every branched monosaccharide.

All of the above factors induce an extreme structural diversity of carbohydrates. It has been computed for instance that there are over 12 million possible structures for an oligosaccharide built from six hexose subunits.[1] Finally, various chemical modifications (acetylation, *N*-

acetylation, sulfation, anhydro-bridges, methyl-etherification, etc.) can occur at different available positions of the subunits, further amplifying the isomerism possibilities.

With respect to these cases of isomerism, ion mobility spectrometry (IMS) is an interesting complement to MS. IMS is a gas-phase analysis technique, which separates ion populations through an interaction with a buffer gas and an electric field. The resulting separation depends in part on the three-dimensional structure of the ions, making it sensitive to all kinds of isomers and gas-phase conformers (provided the apparatus has sufficient IMS resolving power). IMS can readily be integrated in mass spectrometers and is now commercially available with high resolving powers. A notable setup is the Cyclic IMS instrument (cIMS) from Waters Corp. (Wilmslow, UK) that was launched in 2019 under the name SELECT SERIES Cyclic IMS. This instrument offers ultrahigh IMS resolution but also a high flexibility regarding the manipulation of the ions, allowing collision induced dissociation (CID)-IMS, IMS-CID as well as multistage IMS (IMS^n) experiments. **The objective of my thesis was to explore the potentialities of the cIMS instrument for the structural characterization of carbohydrates.**

The major part of the work conducted during this PhD thesis aimed at answering the question: **can we achieve a complete structural determination of complex oligosaccharides using cIMS?** We addressed this point through two intermediary questions: (i) can cIMS retain and resolve all types of isomerism found in carbohydrates? (ii) can we design a cIMS-based approach to solve the complete structure of an oligosaccharide?

We addressed the first question through multiple studies, in which the carbohydrates exhibited different types of isomerisms, *i.e.*, linkage anomerism, linkage regioisomerism, furanose/pyranose cyclization tautomerism, and finally isomeric and epimeric subunits. The results were the basis for the next step, *i.e.*, the second issue, which we approached in two distinct ways. The first way is based on the construction of a library of IMS profiles of monosaccharides. The oligosaccharide to be sequenced is then fragmented by CID in several steps of IMS^n, to be reduced gradually to monosaccharide fragments that can be identified in the library. The oligosaccharide is finally reconstructed from the fragments in a “bottom-up” strategy. The second way consisted in a structure-linked data organization strategy—inspired by the molecular networks used in metabolomics—that compared the ion mobility dimension of oligosaccharides’ fragmentation spectra. This strategy achieved a classification that was of high structural relevance.

The last part of my work consisted in questioning how far we could scale-up the previous results, which were obtained on rather small oligosaccharides, to address higher degrees of polymerization or even full-length polysaccharides. To answer this question, we implemented a method hailing from native IM-MS of protein complexes: we combined the high resolving

power of cyclic IMS-MS with upstream size-exclusion chromatography (SEC) to study the different populations of polysaccharides unavoidably present in any biological fraction.

The manuscript is thus divided as described thereafter:

- **Part I (Introduction**) gives an overview of the literature regarding carbohydrates and the corresponding analytical challenges related to their structure determination. It also presents the state of the art in IMS, and details the geometry and operation modes of the SELECT SERIES Cyclic IMS spectrometer used in my thesis.

The results are then presented in three parts: **II, III and IV**.

- **Part II** describes studies that established that cIMS was able to discriminate unequivocally between the different types of isomerism found in oligosaccharides. The presented studies include the determination of the intrachain anomerism of enzymatically produced mannosides using multiple stages of ion mobility and fragmentation, the building of an IMS database of Gal*f*/Gal*p*-containing galactomannan disaccharides, and further studies probing into the gas-phase behavior of galactomannan oligosaccharides upon fragmentation.
- **Part III** follows up on how to resolve the complete structure of an oligosaccharide using cIMS coupled with fragmentation approaches. It presents the results of a shotgun sequencing approach starting with only the monosaccharides as a reference. Because this approach requires extensive manipulations and is time-consuming, I also propose a more generic approach of ion mobility-molecular networking that we designed as a quick and efficient way to gain some insight on the structures present in a set of oligosaccharides, by establishing structural relationships between the compounds.
- **Part IV** focuses on preliminary studies concerning the analysis of intact polysaccharides by cIMS-MS. It brings new prospects regarding the potentialities of the SEC-cIMS-MS coupling for identification of polysaccharides. It also highlights how far along we still are—in terms of processing methods—before we can fully exploit the highly multiplexed data generated on entire polysaccharides.
- Finally, **Part V** (**Conclusion and perspectives**) presents general conclusions about the potential of cIMS for the study of oligo- and polysaccharides, and outlooks on how to further expand and apply the methods we have developed.

Part I – Introduction

The first Chapter of this section introduces the analytical challenges in characterizing oligo- and polysaccharides. We then provide an overview of the state of the art in IMS and IMS-MS and the positioning of the technique in carbohydrate analysis. Finally, the last Chapter describes in detail the instrument on which the experiments of this thesis were conducted, *i.e.*, the SELECT SERIES Cyclic IMS.

Chapter 1. **Characterizing carbohydrates: an all-time analytical challenge**

Chapter 2. **The special contribution of Ion Mobility to glycomics: What advantages does it bring compared to other analytical techniques?**

Chapter 3. **The SELECT SERIES Cyclic IMS instrument**

Chapter 1 – Characterizing carbohydrates: an all-time analytical challenge

Carbohydrates are ubiquitous molecules found in all life forms, including algae, fungi, plants, bacteria, etc. They are an abundant and renewable resource, which makes them highly interesting for green chemistry, their industrial use having grown up about 17% between 2017 and 2021.[2] The major sources of polysaccharides are plants (e.g., for cellulose and pectin), marine animals (e.g., for chitin/chitosan) and algae (e.g., for carrageenans).[2,3]

Carbohydrates have a wide range of functions and potential applications depending on their structures, hence the importance of their characterization. *In vivo*, they are major components of the plant cell wall, but also serve as energy storage and for cell signaling.[4] *Ex vivo*, they also find applications in numerous industrial fields: (i) in the biorefinery industry, lignocellulose obtained from the plant cell wall can be used to make biofuels,[5] (ii) in the food industry, carrageenans and agars can be used as gelling or thickening agent depending on their sulfation pattern,[6] (iii) in the cosmetic and pharmaceutical industries, polysaccharides (notably starch and cellulose) are used as excipients.[2] Because carbohydrates are biocompatible and biodegradable components, they have raised special interest from the pharmaceutical industry to make nanoformulations[7] or carriers for osteogenic components.[2] Sulfated polysaccharides also offer an interesting range of bioactivities including immunomodulatory, antitumor, antiviral and antioxydant activities.[3]

A. The "basic" structure of carbohydrates

Polysaccharides are heterogeneous biopolymers constituted of monosaccharide subunits, linked through glycosidic bonds.

1. Monosaccharides: the building blocks

Monosaccharides are poly-hydroxy-aldehydes (aldoses) or poly-hydroxy-ketones (ketoses), made of three or more carbon atoms. Depending on its chirality, each monosaccharide can be dextrorotatory (D) or levorotatory (L), as illustrated for glucose enantiomers (**Figure 1**).

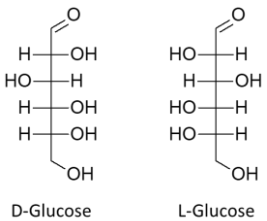


Figure 1. Example of D- and L- oses, illustrated using glucose.

To illustrate the heterogeneity of the isomeric building blocks, **Figure 2** represents the different structures of monosaccharides (Fischer representation) for trioses, tetroses, pentoses and hexoses of the D series. This shows that, by adding one carbon in the structure (with the exception of Erythrulose), it multiplies the number of possible structures by two. These characteristics result in numerous isomeric subunits: for instance, there exist 24 isomers of hexoses (8 aldoses and 4 ketoses for both D- and L- series).

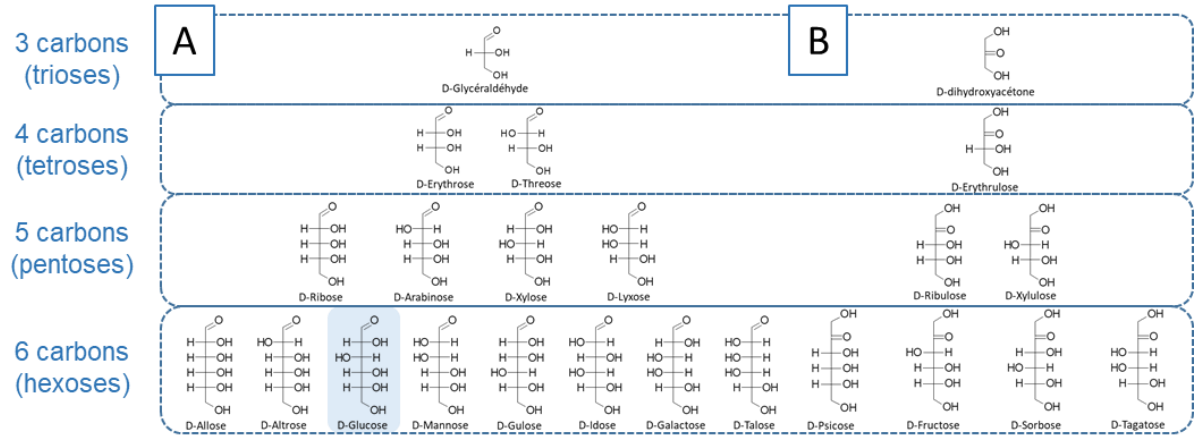


Figure 2. Linear structures of the monosaccharide building blocks with 3, 4, 5 and 6 carbons (Fischer representation) for the (**A**) D-aldoses and (**B**) D-ketoses series. Glucose, used to illustrate the different types of isomerisms in this Chapter, is highlighted in blue.

2. Cyclization of monosaccharides

In natural oligo- and polysaccharides, the monosaccharides are not found in their linear form but rather in a cyclized form. Indeed, monosaccharides undergo an intramolecular hemiacetylation that can form two types of heterocycles: a pyrane (6-membered ring; 5 carbons and 1 oxygen) or a furane (5-membered ring; 4 carbons and 1 oxygen). This is the *furanose/pyranose tautomerism*.

This hemiacetylation additionally introduces a chiral carbon in the monosaccharide compared to its linear form—called anomeric carbon because it is subject to anomerism.[a] The cyclized monosaccharides thus form α or β anomers depending on the orientation of the hydroxyl group carried by the anomeric carbon. Including the pyrane/furane forms and the α/β anomerism, each monosaccharide can thus form four different heterocycles (**Figure 3**). This brings for instance the number of potential isomeric monosaccharidic units to 96 (*i.e.*, 24 x 4), for hexoses.

[a] Anomerism is a type of isomerism relatively specific of the chemistry of cyclic hemiacetals/hemiketals. In biochemistry, it is thus rarely found outside of carbohydrate chemistry.

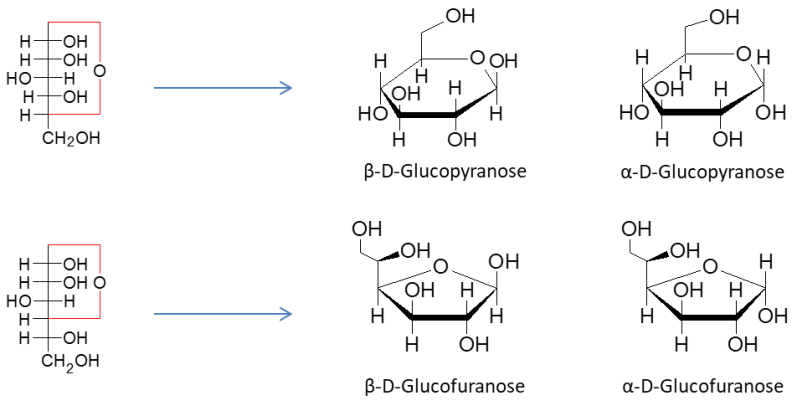


Figure 3. Illustration of the four possible heterocycles that can be formed by a monosaccharide, exemplified on D-glucose

Despite all these potential combinations, not all of these forms actually occur in Nature, and those who exist do so in varying proportions. Indeed, a mutarotation phenomenon (**Figure 4**) naturally occurs at the anomeric carbon, through a series of openings/closings of the heterocycle. This gives different ratios of each forms. For instance, D-glucose exists at ~ 65% in the form of β-pyranose, ~ 35% of α-pyranose, and less than 1% of the furanose forms.[8] The occurring monosaccharide subunits are also restricted by the biosynthetic capabilities of the producing organisms.

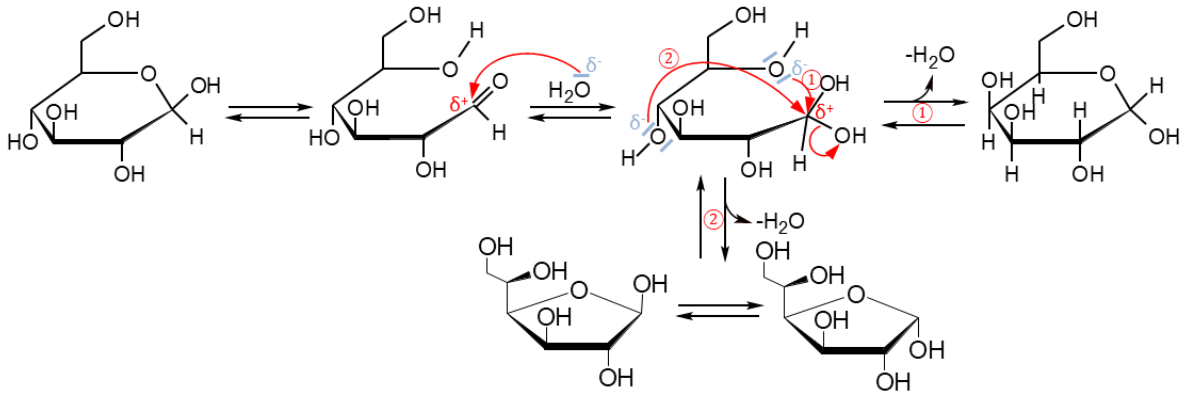


Figure 4. Mutarotation and tautomeric equilibrium, illustrated on glucose

Because of the wide diversity and high structural similarity of all of these monosaccharide units, a graphical representation is usually used to symbolize each monosaccharide when drawing oligo- and polysaccharides. Currently, there is no stable nomenclature for the representation of carbohydrates, but the symbol nomenclature for glycans[9,10] (SNFG) is becoming widely adopted. This nomenclature is presented in **Appendix 1**, and the SNFG representation of all oligosaccharides studied in this thesis is presented in **Appendix 2**. Nevertheless, it is not exhaustive and some monosaccharides are absent from this nomenclature (e.g., aceric acid).

3. The glycosidic bond

In polysaccharides, the monosaccharidic units are linked together through glycosidic bonds. The term glycosidic bond most often refers to a covalent bond formed between the hemiacetal group of a saccharide with a free hydroxyl group of a second saccharide.[b] This reaction is an intramolecular dehydration, “freezing” the mutarotation phenomenon as well as the reducing activity of the saccharide(s) whose C-1 carbon is involved in the linkage. This results in a *linkage anomerism*, or *intrachain anomerism*.

Additionally, any free OH group can in theory react with the hemiacetal (up to 5 potential positions for an hexose) resulting in a *linkage regioisomerism*. Taking the example of a D-glucose homodimer (glucobiose), this gives 9 possible combinations (**Figure 5**).

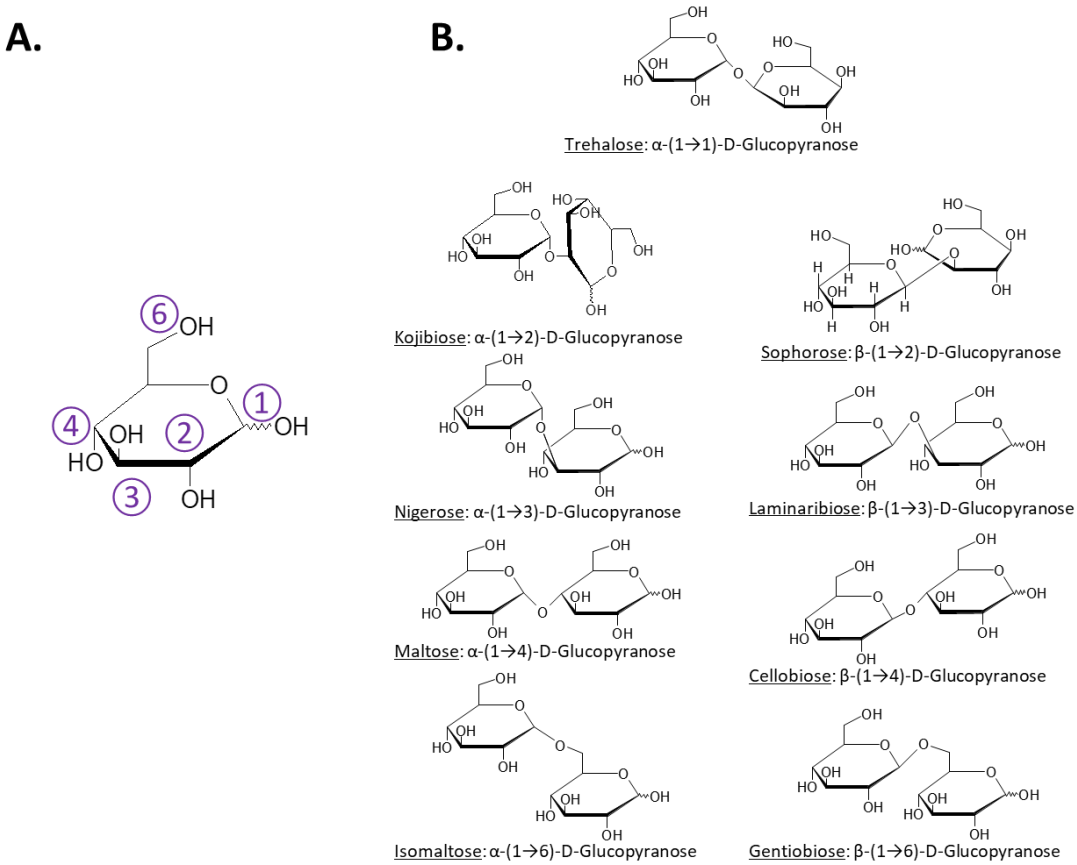


Figure 5. A. Numbering of the carbons/OH groups illustrated on D-glucose.
B. Existing glucobiose species.

The diversity in the primary structure of polysaccharides, considering both the different types of monosaccharides and all the linkage isomers, pauses a major analytical challenge. For

[b] This type of glycosidic bond formed through an oxygen atom is more specifically called an *O*-glycosidic bond. Other types of glycosidic bonds exist where the hemiacetal carbon is linked to another heteroatom—possibly not coming from a carbohydrate—including *N*-glycosidic, *S*-glycosidic or *C*-glycosidic bonds.

instance, for a reducing hexasaccharide, there can theoretically be over 10^{12} possible isomeric structures.[1]

B. How seemingly minor structural changes highly impact the properties of carbohydrates

4. The backbone of polysaccharides and their ramifications

When considering the complete structure of polysaccharides, a key point is that they are made up of an assembly of subunits, including a main stretch called the *backbone*. The end of the polysaccharide backbone where the terminal subunit has a free anomeric carbon is called the *reducing end*. The other end of the backbone, where the anomeric carbon of the terminal subunit is involved in a glycosidic bond, is called the *non-reducing end*.

However, an important particularity of polysaccharides—when compared to other types of biopolymers like proteins or nucleic acids—is that they are rarely linear. Indeed, every free OH group in the backbone can be branched by lateral subunits. This can end with a single lateral subunit at a given position, or the lateral subunit can be elongated to give rise to a lateral chain. This can result in highly ramified structures, an example of which is xyloglucan. Xyloglucan is part of the hemicellulose found in the (primary) cell wall of plants, and is especially abundant in dicots. The backbone of xyloglucan is made of β1,4-linked D-glucose subunits that can either remain unsubstituted (G residue), or substituted with different sidechains.[11] The major types of sidechains include substitution by a single α-D-xylose (X residue), that can be substituted by a β-D-galactose (L residue), that can itself be substituted by an α-L-fucose (F residue), as illustrated in **Figure 6**.

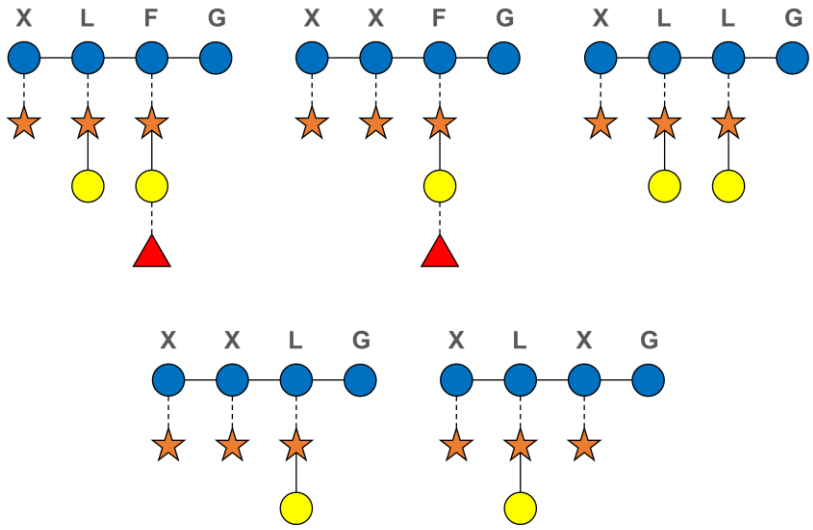


Figure 6. Examples of xyloglucan oligosaccharides (to retain readability, the linkage regioisomerism is not considered in the representations).

5. Possible modifications

On top of the huge diversity of monosaccharides and glycosidic linkages, and of the ramified structure of polysaccharides, the monosaccharidic subunits can further be functionalized by modifications on the OH groups that are not involved in glycosidic bonds. The most common modifications are depicted in **Figure 7**.

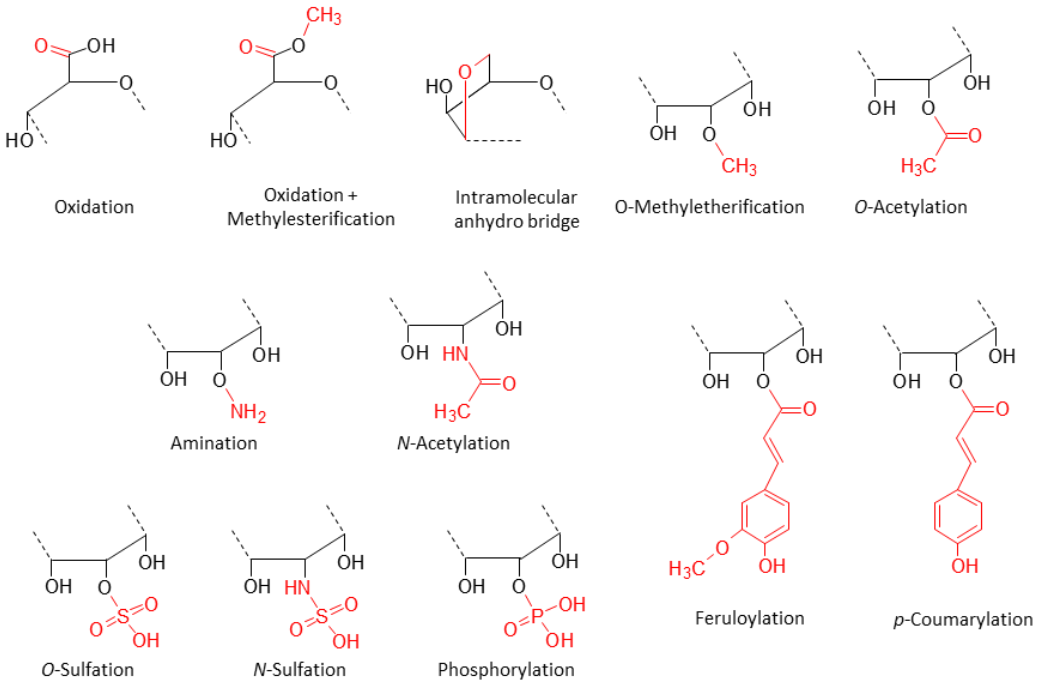


Figure 7. Most common chemical modifications found in polysaccharides

One notable modification is the oxidation of the C-6 carbon of a hexose into a carboxylic acid, giving rise to a very common type of monosaccharidic unit called uronic acid. For instance, in plants, α-D-galacturonopyranose acid is a major constituent of pectins and can represent up to 35% of the dry weight of the cell wall of dicots.[12]

Pectin (found in vegetables, notably citrus and apples) is a prominent example of the diversity that can be found in polysaccharides as a result of the above-mentioned structural features. It is built by over 15 different types of monomers, including both ramifications and chemical modifications (**Figure 8**). Pectin is a supramolecule composed of three domains. The first is a unique polysaccharide, named homogalacturonan, which consist of an α-D-galacturonopyranose linear chain. The two other domains are much more complex and heterogeneous. They are called rhamnogalacturonans I (RG-I) and II (RG-II). RG-I consists of a backbone composed of a strict alternation of α-L-rhamnose and α-D-galacturonopyranose, and bear different side chains including two different types of arabinogalactans, arabinans and xylogalacturonans. RG-II are even more complex, they are composed of a backbone of α-D-galacturonopyranose presenting five types of heterogeneous lateral chains labeled A–E.

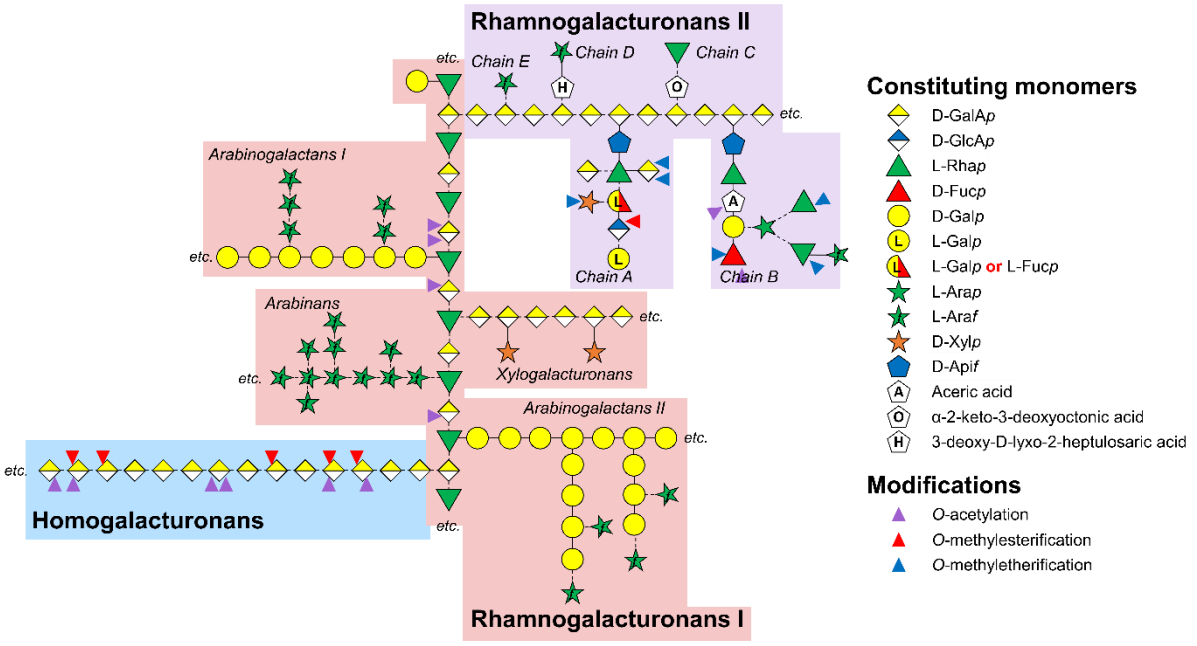


Figure 8. Schematic model of the structure of pectin[12,13] (to retain readability, the linkage regioisomerism is not considered in the representations).

6. Structure-function relationship of carbohydrates

All of the above-mentioned chemical features have an important impact on the properties and functions of the polysaccharides. The chemical properties of polysaccharides can have an impact on their *in vivo* structural properties, their physical properties, or their biological activities.

For instance, in the case of pectin, the acidity of the polysaccharide (due to the high content in uronic acids) and the chemical modifications (notably the *O*-methylesterification) confer particular properties to pectin as a plant cell wall component. These are referred to as 'matrix properties'.[12] In the plant cell wall, the pectic matrix is a more soluble component that embeds the cellulosic network (that itself bears the load of the physical constraints applied to the plant) and provides an environment that allows microscale deformations of this network. This matrix is also involved in controlling the cell wall porosity, and is the main adhesive material between cells.

The substitution of the backbone can also impact the interactions between two types of polysaccharides. This is notably the case for the interactions between heteroxylans and cellulose in the plant cell wall.[14] The nature of the substitution and modifications of the xylans (acetyl, arabinosyl, or glucuronosyl residues), as well as their pattern and proportions, will impact the interactions with cellulose. In particular, they will affect the solubility and the hydrophilicity of the xylan components.

Finally, the branching degree of the polysaccharide and modifications will have a major impact on their three-dimensional conformations, as pointed out by recent reviews.[15,16] For instance, more linear polysaccharides can adopt helical conformations, ranging from single-helix to triple-helix.[16] However, the relationship between the tertiary structure of polysaccharides and their functions is not completely clear as of yet. For example, a recent review on the immunomodulatory activities of polysaccharides pointed out that sometimes, triple-helix conformations conferred higher activity to different types of polysaccharides but also lowered their activity in other cases.[16]

Overall, the literature shows that the many different features found in carbohydrates can impact their physicochemical, functional and biological properties. However, in many cases, the structural data remains insufficient to fully understand the structure-function relationships of carbohydrates. Notably, major limitations come from the difficulty of obtaining highly pure fractions, as well as of the lack of fast and sensitive analytical methods able to accurately determine the chemical composition and purity of polysaccharides.

C. Overview of analytical methods for carbohydrate characterization: advantages and limitations

Because of the high structural complexity of carbohydrates, their structural characterization remains highly challenging with current analytical approaches, as recently discussed by Gray *et al*.[17] Each method has its own strengths and weaknesses, affording incomplete structural information. It is thus necessary to cross the data from multiple analytical techniques to completely resolve the structure of carbohydrates. In this section, we will describe the main methods used to characterize oligo- and polysaccharides, distinguishing non-MS-based approaches from MS-based approaches.

7. Non-MS-based methods

a) Nuclear Magnetic Resonance

Nuclear Magnetic Resonance (NMR) has been a method of choice for characterizing carbohydrates for decades.[18] NMR gives complete structural information, affording the identification of constitutive subunits including their stereoisomerism (anomerism, epimerism, etc.). A recent review highlights that the use of NMR to characterize polysaccharides structures, as well as for conformational or interaction studies, is still expanding.[19] The method also has the notable advantages of giving quantitative results and being non-destructive. Still, it suffers from a number of drawbacks.

Notably, NMR is limited by the size of polysaccharides. High molecular weight polysaccharides (up to megadaltons) are often poorly soluble, which limits the use of liquid state NMR. Other

limiting factors are the polydispersity and the important structural variability behind every degree of polymerization (DP). Among others, the position of the lateral chains on the backbone or of the chemical modifications of the subunits, are subject to variation and induce an important heterogeneity of the samples. As a result, the prism of NMR spectroscopy provides an average of all structures in presence. It is therefore important to produce highly pure polysaccharides for NMR characterization, often requiring multiple purification steps. However, even with multiple purification steps, it is not always possible to obtain completely pure polysaccharides (*i.e.*, multiple structures often coexist in fractions). Considering the sensitivity of NMR, the minimal amount of sample required for said fractions would be in the range of the hundreds of µg, which is challenging for natural compounds and especially for minor components. However, recent approaches promise to increase the sensitivity of NMR using Dynamic Nuclear Polarization, also affording much faster acquisitions.[20]

b) Chromatographic methods[c]

Many chromatographic methods can be used to analyse oligo- and polysaccharides, either intact or after degradation and/or derivatization.

Gas chromatography (GC), coupled to flame ionization detectors (FID) or electron impact (EI) mass spectrometers is often used to determine the osidic composition and/or the subunit linkages. Osidic composition is determined by total hydrolysis of the polysaccharides followed by a chemical derivatization—required to make the oligosaccharides volatile for GC analysis—and dosing of the alditol acetates, by comparison to standards at a specific retention time. The linkage regioisomerism can also be determined in this way but requires additional derivatization steps (polysaccharide permethylation, followed by hydrolysis, and then peracetylation). This double derivatization allows to distinguish the OH groups involved in glycosidic bond from the others: the free OH groups of the polysaccharide are methylated and the OH groups involved in glycosidic bonds are acetylated. Overall, these approaches are powerful and robust methods, but they only give a global view of the types of monosaccharides and linkages present in the sample. The connectivity within the polysaccharide itself is lost. Worse, the information about different polysaccharides present in the same sample can be mixed.

Liquid chromatography (LC) is an interesting alternative to GC when dealing with oligosaccharides, which cannot be made volatile enough to be separated in GC. The oligosaccharides may be produced by a mild/controlled hydrolysis and/or by an enzymatic depolymerisation of the polysaccharide. The advantage of analysing oligosaccharides instead

[c] Most chromatographic methods can be coupled to mass spectrometry, but this remains optional.

of monosaccharides is that the connectivity of the building blocks as it occurs in the polymer is in part maintained.

However, oligosaccharides being hydrophilic, not all LC stationary phases are adapted to their separation.

- *Ion exchange chromatography* is often used to separate oligosaccharides, and in particular *high-performance anion exchange chromatography* (HPAEC). The stationary phase is made of a resin coated with positively-charged counter ions (**Figure 9**). For neutral oligosaccharides, it notably makes use of the fact that these species are negatively charged at high pH levels. However, this operation at high pH makes it incompatible to MS, and HPAEC is most often found coupled to pulsed amperometry detectors (HPAEC-PAD).[d]

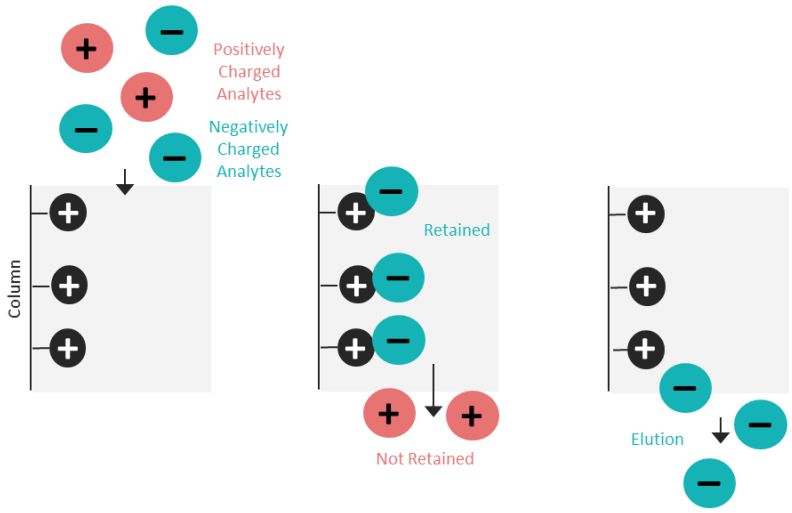


Figure 9. Principle of anion exchange chromatography

Because HPAEC cannot easily be coupled to MS, other chromatographic methods with more compatible mobile phases will be preferred:

- *Hydrophilic interaction LC* (HILIC) is an interesting method to separate oligosaccharides, including neutral ones. This variant of normal-phase chromatography separates analytes according to differences in their polarity. A layer of water is formed around a normal-phase stationary phase, and elution is then performed using RP-type solvents (**Figure 10**). Depending on their polarity, the analytes will exchange at different rates between the water layer and the organic solvent.

[d] Pulsed amperometry detectors are used when analyzing carbohydrates because on classical amperometry detectors, these compounds forms an insulating layer on the electrode resulting in a loss of signal over time. Electric pulses, on the other hand, detach the compounds from the electrode.

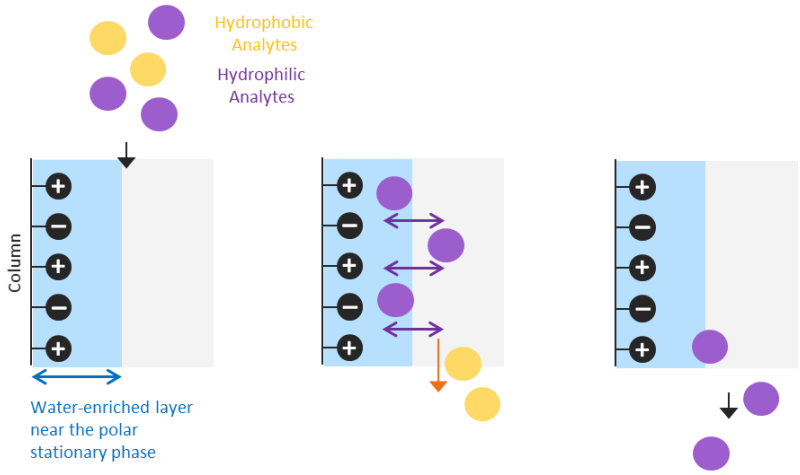


Figure 10. Principle of hydrophilic interaction liquid chromatograhy

- *Ion pairing reversed phase liquid chromatography* (IP-RP-LC) affords a better separation compared to classical RP-LC, which is not efficient for separating carbohydrates because of their high hydrophilicity. By adding an amphiphilic ion pairing reagent (IPR) such as hexyl- or heptylamine,[21,22] an interaction is created with the stationary phase. The IPR agent forms an ionic bond with the analyte, and its lipophilic tail interacts with the stationary phase, causing retention (**Figure 11**). However, this technique requires the analyte to be charged and is thus mostly adapted to acidic oligosaccharides (uronic acid or sulfated).

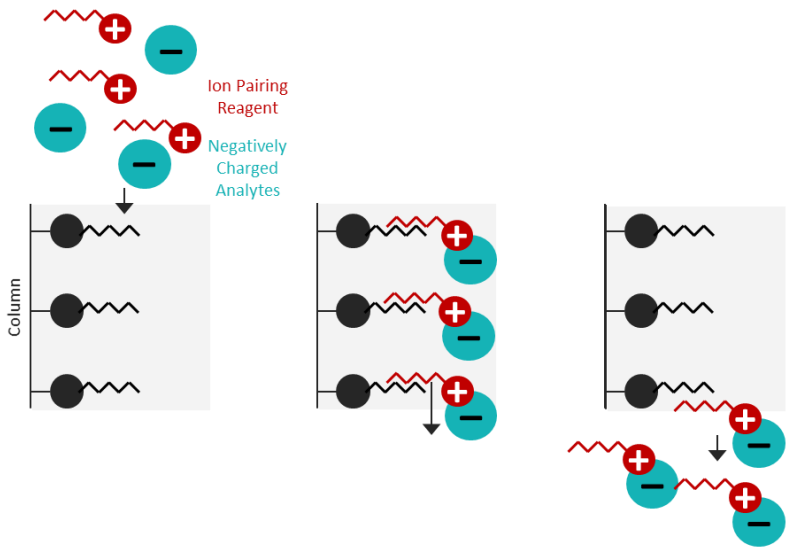


Figure 11. Principle of ion pairing chromatography

- *Porous Graphitic Carbon* (PGC) is a chromatographic separation more adapted to the separation of neutral oligosaccharides. It operates using charge-induced dipole–dipole interaction between the analyte and layers of graphitic carbon.[23] The separation will thus be function of the surface of interaction between the oligosaccharide and the

carbon layer (**Figure 12**), meaning that linear oligosaccharides will theoretically be eluted later than branched oligosaccharides of similar degree of polymerization.

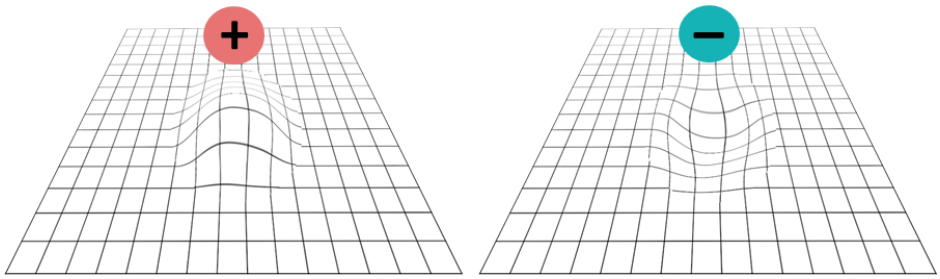


Figure 12. Schematic of the dipole-dipole interaction occurring between the analytes and the PGC phase

Size-exclusion chromatography (SEC) is more adapted to the analysis of intact polymers. Contrary to other types of stationary phases, it does not involve any chemical interaction. This method is based exclusively on the porosity of the stationary phase, which contains a range of small to bigger pores (**Figure 13**). The smaller analytes (*i.e.*, with a less important hydrodynamic volume) will enter almost all of the pores while larger analytes will only go through the larger ones. As such, the path that smaller analytes take to exit the column is longer than that of large analytes, which results in later elution.

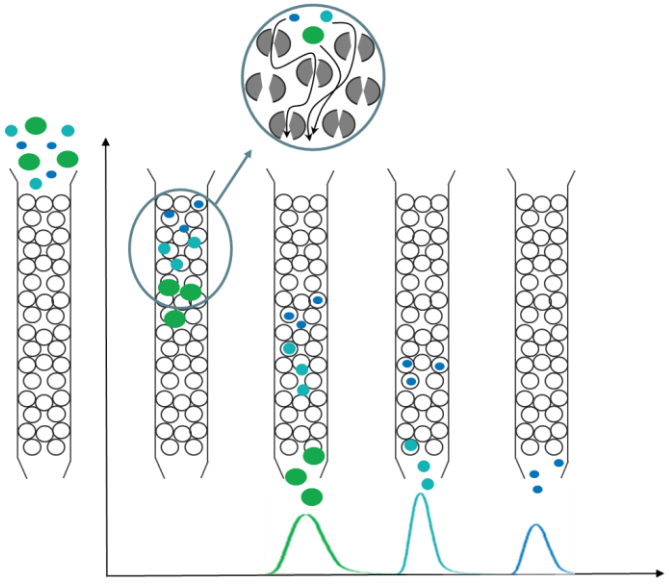


Figure 13. Principle of SEC separation

SEC is a chromatographic method designed to work on analytes ranging from a few kDa up to MDa levels. In the context of carbohydrate analysis, SEC is thus used to separate intact polysaccharides. Since polysaccharides do not absorb ultraviolet light (unlike proteins), SEC is mostly coupled with refractive index detectors or multi-angle light scattering detectors (MALS) for carbohydrates. MALS affords the measurement of the weight-average (Mw) and

number-average molecular masses (Mn), and thus the polydispersity (Mw/Mn) of a polysaccharide.

SEC can also be coupled to MS, as exemplified for proteins.[24] However, polysaccharides pose additional challenges with regard to SEC-MS analysis, which will be further discussed in **Part IV** of this manuscript.

8. MS-based methods

MS is a method of high interest for the structural characterization of carbohydrates, notably because of its fast acquisition rate, high sensitivity and relatively high tolerance to mixtures (especially when coupled to chromatographic approaches).

a) MS and ionization methods in the context of carbohydrates

A key point in the analysis of biomolecules by MS is the generation of ions from the neutral species in the samples, since only ionized species can fly through the mass spectrometer to the detector. Ions must also be "desorbed" from the condensed phase (e.g., liquid) into the gas phase. Ideally, ionization/desorption methods should occur in "soft" conditions to preserve the integrity of the species. As of today, there are two "soft" methods of ionization/desorption that are predominantly used for the analysis of biomolecules by MS.

The first is electrospray ionization (ESI), described in 1989 by Fenn and collaborators.[25] This ionization method was the primary method used in this thesis and will be further described alongside the instrument in **Chapter 3**. Briefly, the analytes are ionized directly from a solution: the solution is nebulized at atmospheric pressure, through a metallic capillary submitted to an electric field (in the order of 3000 V/cm). This generates charged droplets, and the charge is then transferred to the analytes.

The second one is matrix-assisted laser desorption/ionization (MALDI) introduced in 1988 by Karas and Hillenkamp,[26] and by Tanaka.[27] MALDI is based on the irradiation of the sample by a laser, at the same time as a molecule absorbing at the appropriate wavelength (depending on the laser type), which is called "matrix". The sample is deposited on a conductive surface, and during the irradiation, there will be an energy uptake by the matrix. The matrix subsequently releases energy, generating desorption from the surface (**Figure 14**) in the form of a plasma containing the analytes. In the plasma, collisions happen between the matrix ions and the analytes, resulting in a charge transfer from the matrix to the analyte and inducing ionization. MALDI has some advantages, notably including a tolerance to salts higher than ESI, the ability to analyze samples with low solubility, or the fact that it can be naturally implemented for MS imaging (MSI).[28] In the context of carbohydrates, MSI can notably be used to study the *in planta* distribution of cell wall polysaccharides according to their structure.[29,30]

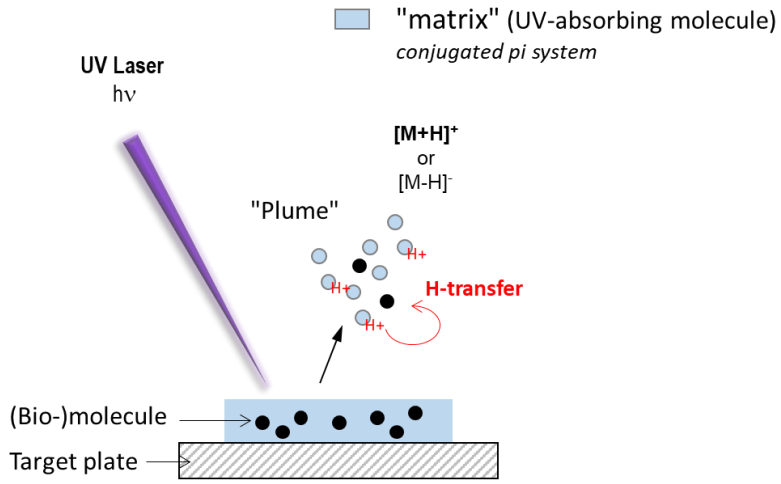


Figure 14. Principle of MALDI ionization

Regarding the species observed when ionizing oligosaccharides in positive mode, they tend to form mostly metal-cations. This is in contrast to the protonated species that can be observed for other types of biomolecules (peptides, proteins). Metal-cations are often formed by adduction of alkali metals already present in the analysis media (Na, K, resulting in $[M + Na]^+$ or $[M + K]^+$ species for instance), but it is also possible to introduce on purpose some metal species like lithium or alkali-earth metals (Mg, Ca). Metal adduction is an ionization process that is less favorable/efficient than protonation or deprotonation of analytes, leading to a sensitivity in the order of the µg/mL for carbohydrates (in both ESI and MALDI).

In negative mode, deprotonated molecules ($[M – H]^–$ species) are mostly observed, but the deprotonation is relatively inefficient for neutral oligosaccharides compared to acidic oligosaccharides.

Overall, the generation of “intact” ions by soft ionization methods such as ESI or MALDI affords further fine structural analysis, which can be achieved by selecting the ion and fragmenting it in so-called MS/MS approaches.

b) Fragmentation: the Domon and Costello nomenclature

In the case of MS/MS characterization of oligosaccharides, a specific nomenclature for the fragments is used. This nomenclature has been proposed by Domon and Costello in 1988 (**Figure 15**). [31] It describes:

- Glycosidic bond cleavages yielding fragments corresponding to *intact* oligosaccharides. The fragments containing the reducing end are labeled Y, the fragments containing the non-reducing end are labeled C.

- Glycosidic bond cleavages yielding fragments corresponding to *dehydrated* oligosaccharides. The fragments containing the reducing end are labeled Z, the fragments containing the non-reducing end are labeled B.
- Intracyclic (or cross-ring) fragments correspond to the concomitant fragmentation of two bonds of the ring. The fragments containing the reducing end are labeled X, the fragments containing the non-reducing end are labeled A. Superscript numbers placed before the A/Z letter indicated between which bonds of the cycle the cleavage occurs.

For all of the fragment types, a subscript number placed after the letter indicates the position of the fragment in the oligosaccharidic chain (in terms of number of subunits).

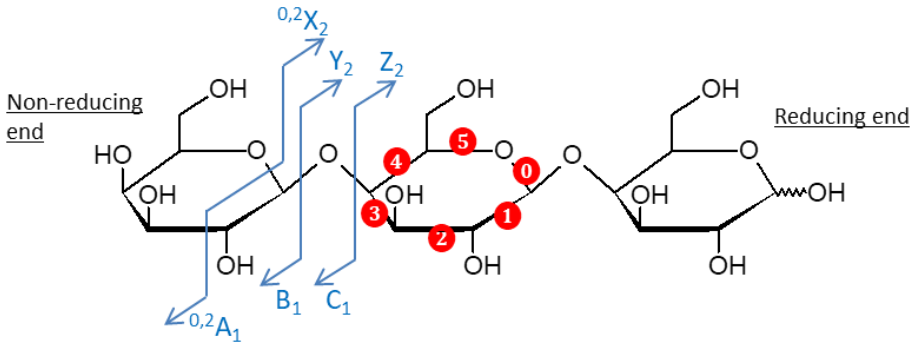


Figure 15. Domon and Costello nomenclature for the fragmentation of oligosaccharides.

c) Fragmentation methods

(1) Collisional activation: the gold-standard in mass spectrometry

There exist a number of different fragmentation techniques used in mass spectrometry,[32] but the most common one is collision-induced dissociation (CID), sometimes also called collision-activation dissociation (CAD).

The principle of CID is to accelerate the ions by applying a voltage, increasing their kinetic energy. They will then collide with neutral gas molecules (usually Ar, He or N_2), which induces a conversion of the kinetic energy into internal energy. This step is called the *activation*.

The internal energy will then be distributed among the vibrational modes of the precursor ion. When the energy surpasses a threshold on a given bond, said bond will break and generate fragments. This step is called *unimolecular dissociation*. Because different types of bonds have different energy thresholds, more fragile bonds such as ether bonds—glycosidic bonds in the case of carbohydrates —will be the main locations of the fragmentation process induced by CID.

CID is a very efficient technique that generates fragments with a high intensity relative to the precursor ion, contributing to its popularity in the MS community. Decent MS/MS spectra can thus be acquired within seconds in many situations.

(2) The need for alternative fragmentation methods in the case of carbohydrates

The major problem of using CID for the fragmentation of oligosaccharides is that the most fragile bonds will fragment first. This implies that labile modifications, (such as sulfation, which is prevalent in many functionally important carbohydrates) will be lost and that the vast majority of the fragments formed will be glycosidic bond cleavages.

On the other hand, intracyclic fragments are the most informative to determine the linkage regioisomerism, the localization of the modifications and of the lateral branching. It was shown that these fragments can be promoted by the use of higher energy fragmentation techniques that can induce radical-driven fragmentation pathways and generate these informative cross-ring fragments. Electron-based methods (ExD) have shown to be useful, notably electron detachment dissociation (EDD),[33] electronic excitation dissociation (EED)[34] or electron transfer dissociation (ETD).[35] Until recently, these methods were restricted to instruments equipped with a Fourier-transform ion cyclotron resonance (FTICR) cell or they required the insertion of an electron transfer reagent in the mass spectrometer (in the case of ETD).

Photo-dissociation using VUV[36,37] or XUV[38,39] photons produced by synchrotron radiation is also able to trigger radical-driven fragmentation pathways through a dissociative photoionization process, and generates informative fragments required for carbohydrate characterization. Still, the method requiring access to a synchrotron is a major limitation.

On the other hand, a method was recently developed based on ion-ion interactions between the analyte and highly accelerated helium cations, called helium-charge transfer dissociation (He-CTD). He-CTD has been shown to give similar results to XUVPD on carbohydrates,[40] and can be settled on benchtop instruments.[41]

d) Benefit of adding an additional dimension to MS to solve isomerism problems

Despite the advances in structural characterization with MS and MS/MS approaches, MS is—as its name implies—only sensitive to the mass of the ions. This makes it blind to a number of structural features found in carbohydrates, and especially those related to stereoisomerism phenomenon.

Previous studies have shown the interest of adding another gas-phase spectroscopic or spectrometric dimension to MS, helping solve the isomerism problems. For instance, Schindler *et al*. have used infrared multiple photon dissociation (IRMPD) spectroscopy to help solve problems related to the intrachain anomerism[42] or the furanose/pyranose tautomerism.[43] IRMPD spectroscopy is based on a photo-dissociation approach: the parent ion is irradiated with infrared photons at different wavelengths, inducing slow thermal dissociation. Then, the photofragmentation yield at each wavenumber is used as a signature of the compound.

In this thesis, we opted for the use of ion mobility spectrometry (IMS), which we will describe in detail in **Chapter 2**.

Chapter 2 – The special contribution of Ion Mobility to glycomics: What advantages does it bring compared to other analytical techniques?

In the previous Chapter, I presented the challenge of characterizing carbohydrates and the limitations of existing techniques. I mentioned ion mobility spectrometry (IMS) as a possible way to circumvent the problem of isomerism in mass spectrometry. In this second Chapter, I will thus start by introducing the principles of IMS, the different technologies available and the latest innovations developed to improve the potential of IMS in different fields.

In a second part, I will focus on recent applications developed in the context of carbohydrate characterization that demonstrate, through the efforts implemented by several groups throughout the world, the potential of IMS for glycosciences.

A. History of Ion Mobility Spectrometry (IMS)

1. Origins of IMS

Of particular interest is the history of IMS, a summary of which is presented in **Figure 16**.

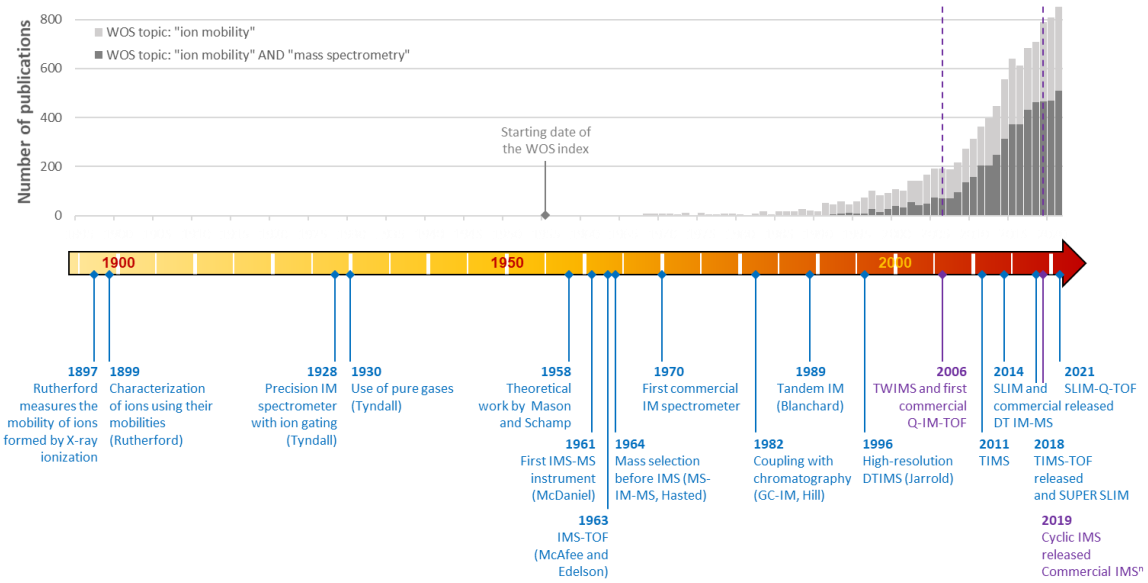


Figure 16. Timeline indicating some key milestones of the history of IMS, and the number of publications indexed in the web of science (WOS) for IMS and IMS-MS

Indeed, in the timeline of analytical chemistry, IMS is an old technique that was first described in 1897[44]—15 years before MS[45]—when Ernest Rutherford first measured the mobility of ions in the gas-phase. During this early period of IMS, foundational studies already established that

there was an effect of collisions, temperature, pressure and voltages on the gas phase mobility of ions.[46,47] But in the 30's and 40's, MS arose at the expense of IMS, notably because MS data were easier to interpret (at the time, the ion sources and operating pressure of IMS resulted in complicated reactions). Only during the late 50's did IMS regain attention with theoretical work by Mason and Schamp[48] and McDaniel,[49,50] which described the gas-phase behavior of ions according to their shape and size, and to collision phenomena. Notably, McDaniel was also the pioneer in coupling IMS to MS (with a magnetic sector mass spectrometer) back in 1961.[51] Two years later, McAfee and Edelson introduced the first IMS-TOF instrument.[52]

The first commercialization of an IMS-MS instrument dates back to the 70's. It was an ambient pressure ion mobility spectrometer interfaced with a quadrupole mass spectrometer, and was commercialized under the name "Plasma Chromatograph".[53] Yet, it was only in the 80's and 90's that the field of IMS really expended, notably through developments from the Smith,[54] Jarrold,[55,56] Bowers[57,58] and Hill[59,60] groups. These groundbreaking works notably afforded improvements in the use of electric fields to manipulate and direct ion beams under vacuum conditions, and laid the foundation of modern IMS.

As shown by the major increase in the use of IMS in the literature, the field finally took off in 2006, when Waters Corporation commercialized the first Q-IM-TOF instrument—the Synapt HDMS System—which gave access to IMS to a broader community.[61,62] Since 2006, IMS has found many applications, as illustrated by the research categories associated with IMS in the Web Of Science (**Figure 17**). The most represented categories cover wide and various fields including physics and physical chemistry, biochemistry and molecular biology, food and environmental science, or pharmaceutical applications. Most represented categories also include, as expected, topics directly related to developments in analytical chemistry and spectrometry, notably engineering development for IMS instrumentation.

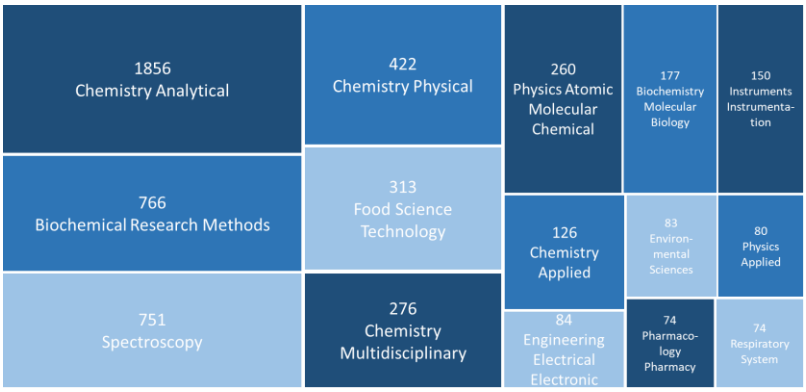


Figure 17. Top 15 research categories associated with the topic "ion mobility spectrometry", by number of publications indexed in the Web Of Science for the period 2006–present[63]

2. Principles of modern IMS

Modern IMS instrumentation can operate according to different principles. All of these have in common that ions will be separated in the gas-phase under the influence of electric fields, which is the fundamental principle of IMS.[64] The commercially available set ups are presented thereafter. Here, we will distinguish the setups operating using linear methods (*i.e.*, the mobility is assumed to be independent from the electric field) and nonlinear methods (*i.e.*, the method exploits the dependence of the mobility on the electric field).[65]

a) Linear methods

Drift Tube IMS (DTIMS) is the historical form of ion mobility spectrometers. DTIMS is a time-dispersive IMS technique. A so-called "drift tube" is filled with a static buffer gas, typically He or N_2 (**Figure 18**).[66] In this static buffer gas, ions are subjected to a linearly decreasing electric field (E) to guide them through the IMS cell. The velocity at which the ions go through the cell is a function of their size and 3D conformation: ions are slowed down by collisions with the buffer gas—the more important the surface interacting with the gas, the slower the ions.[67] DTIMS allows to directly measure the mobility parameter (K), directly linked to the shape of the ions. From this direct measurement of K, the so-called collision cross section (CCS, expressed in Å² or in nm²) of the ions can be derived from the Mason-Schamp equation[68] (**Equation 1**).

$$CCS = \Omega = \frac{3}{16}\sqrt{\frac{2\pi}{\mu\, k_B\, T}} \times \frac{z\, e}{N_0\, K_0} \qquad (1)$$

Where μ is the reduced mass of the ion-gas pair, k_B is the Boltzmann constant, T is the temperature, z is the ion absolute integer charge, e is the charge of the electron, N_0 is Loschmidt's number, and K_0 is the reduced mobility of the ion[e].

To this day, DTIMS is the most widespread method from which the CCS can be directly computed, and it is thus the main primary method[f] for CCS measurement.[65] Note that CCS is in fact more of a momentum transfer cross section that represents the surface (rotationally averaged) of an ion that is interacting with the buffer gas during the separation. It is a property

[e] The term "reduced" mobility refers to mobilities corrected to standard temperature (T_0 = 273 K) and pressure (p_0 = 1013 hPa) values: $K_0 = K\frac{T_0}{T}\frac{p}{p_0}$

[f] Definition of a primary method of measurement: "a method having the highest metrological qualities, whose operation can be completely described and understood, for which a complete uncertainty statement can be written in terms of the SI, and whose results are, therefore, accepted without reference to a standard of the quantity being measured" (BIPM, 1995).[69]

of the ion-gas pair, meaning that the same ion will have different CCSs when studied using two drift tubes operating with different buffer gases.

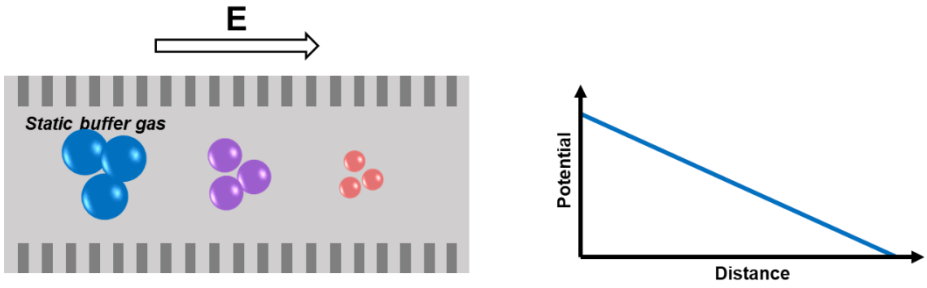


Figure 18. Principle of DTIMS separation

Traveling Wave IMS (TWIMS) is also a time dispersive IMS method. However, unlike DTIMS that operates using a uniform E field, TWIMS uses a dynamic electric field in the form of waves of potential, constantly propagating through a drift cell (**Figure 19**), hence the name "traveling wave".[61] The drift cell is filled with a static buffer gas, with a pressure in the order of a few millibars (most commonly N_2, possibly He or Ar).[70]

Like in DTIMS, the ions will travel through the cell under the influence of the field. They are carried by the waves and will be slowed down by collisions with the buffer gas. Precisely, the ions "surf" on the potential waves and experience friction with the buffer gas, which makes them undergo a phenomenon called "roll-over" that results in making them slip backwards in a next wave.[61] The bigger ions experiencing more friction, they will roll-over more frequently than compact ions. The more compact ions will thus come out of the TWIMS cell first, and the more extended ones come out last.

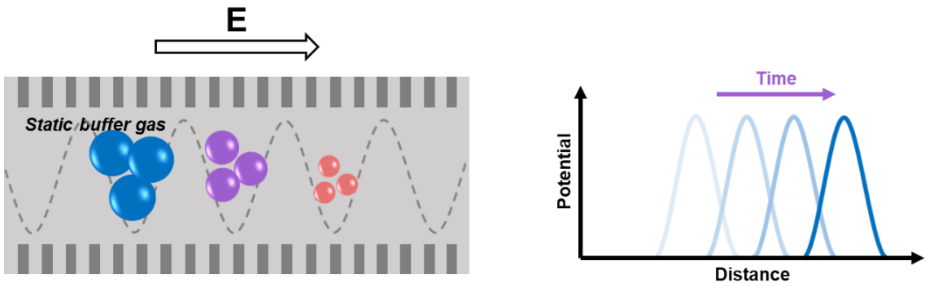


Figure 19. Principle of TWIMS separation

Because the electric field used in TWIMS is not static, it is not possible to directly measure K—and thus deduce the CCS—using TWIMS. However, it is possible to perform a CCS calibration using ^{DT}CCS values of reference ions.[65,71]

Trapped IMS (TIMS) is not a time dispersive IMS method, unlike DTIMS and TWIMS. This more recent concept functions as a "trap-and-release" system, or field-dispersive.

In TIMS, a high flow of buffer gas (N_2) pushes the ions into the drift cell while RF voltages confine the ions radially. A non-uniform electric field is applied as an axial field gradient in a direction that goes against the elution by the gas flow.[72,73] This results in trapping the ions in the tunnel (**Figure 20**). To be precise, the ions will be trapped at the point where the strength of the electric field is such that the velocity of the ions equals the velocity of the buffer gas.[74,75] Ions with lower mobilities (bigger) will be trapped closer to the exit compared to ions with higher mobilities (more compact), for which a higher electric field is needed to compensate the action of the buffer gas. Then, the electric field in the mobility cell is gradually lowered, inducing a sequential elution of the ions.

Opposite to DTIMS and TWIMS, the more extended ions (or low-mobility ions) are eluted first from the mobility cell.

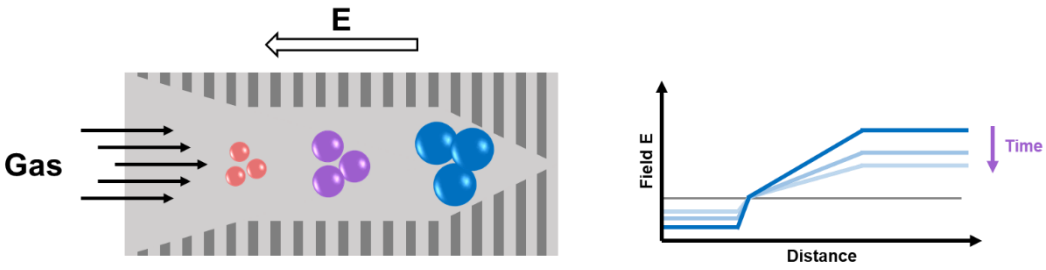


Figure 20. Principle of TIMS separation

As for TWIMS, TIMS can not be used as a direct method for CCS measurement because of the complexity of the electric fields involved. However, because it remains in the low-field limit, it is possible to perform CCS (or K_0) calibration using reference ^{DT}CCS values.[65] The TIMS technology has been commercialized by Bruker.[76]

Differential Mobility Analyzers (DMA) are a type of spatially dispersive IM spectrometers, generally used for the study of nanoparticles.[77] In DMA, a uniform electric field is generated between two electrodes[g] (**Figure 21**, a voltage is applied on one and the other is grounded). Ions enter the mobility cell through a slit in the first electrode and migrate towards the second electrode under the influence of the electric field. Meanwhile, the ions are also transported by a stream of buffer gas introduced in a direction orthogonal to the electric field lines. At a given voltage, only ions with a specific mobility reach the exit slit on the second electrode.[78,79] During DMA analyses, voltages are scanned to sequentially sample the ions of different mobilities.

DMA can act as a primary method of CCS measurement, as K_0 can be determined with precise knowledge of the voltages, temperature, pressure and sheath gas flowrate, however it is often

[g] In DMA the electrodes can be found in different geometries—parallel-plate, concentric cylinders, or radial

more practical to perform a CCS or K_0 calibration.[65] DMA can notably be used on very large analytes that are not accessible with DTIMS (e.g. viruses).[80,81]

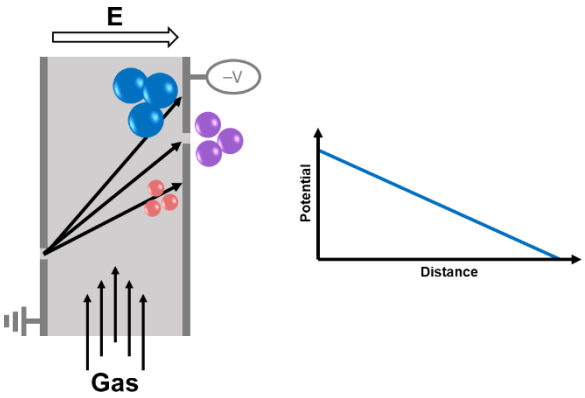


Figure 21. Principle of DMA separation

b) Nonlinear methods

Field Asymmetric IMS (FAIMS) or Differential Mobility Spectrometry (DMS) is a spatially dispersive IMS technique that separates ions according to their differential mobilities between high and low fields. This non-linear method uses the fact that the gas-phase mobility of any ion depends on the intensity of the electric field E, and uses two (or more) E levels to separate mixtures of ions. E is varied temporally in a periodic asymmetric waveform, generating a spatial dispersion of the different ion populations (**Figure 22**). A compensation voltage is applied at the exit of the IMS cell to redirect the ions and select which population makes it to the MS detector at a given time.[65,82] Two geometries exist with cyclindrical electrodes[83] or with parallel plates.[84] The concentric cylindrical design is commercialized by Thermo Fisher Scientific.

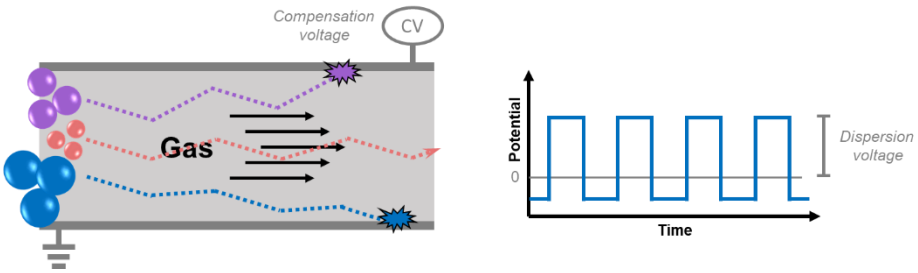


Figure 22. Principle of FAIMS separation

3. Coupling IMS with mass spectrometry (MS)

A strong asset of today's IMS is the fact that it can be integrated in mass spectrometers (IMS-MS), anywhere between the ion source and the mass analyzer. Notably, different kinds of ion sources and analyzers can be coupled to IMS resulting in a variety of instruments and applications. However, each technology has its specificities and not every mass analyzer, for instance, can be coupled to any IMS technique.[85] This will be detailed in the following section.

a) Compatible ion sources with IMS

Many types of ionization methods have been successfully coupled to IMS for decades, highlighting the potential of IMS-MS as an additional separation dimension in various fields of application of MS.

Electrospray ionization (ESI) is the ion source that is most often found on IMS-MS instruments. ESI is notably found on the Cyclic IMS (cIMS) setup and will be described later, alongside the cIMS instrument (**Chapter 3**). ESI can be used through direct infusion or coupled to liquid chromatography (LC), in a timescale that stays compatible with IMS-MS. One of its variants, desorption ESI (DESI), is also available for Waters instruments and can be used for MS imaging (MSI) in conjunction with IMS-MS.

Matrix-assisted laser desorption ionization (MALDI) is another ion source that is compatible with IMS-MS. Like DESI, it allows to perform MS imaging. IMS-MS setups with MALDI ion sources notably include the timsTOF fleX instrument (Bruker, Bremen, Germany) and the previous generation of linear TWIMS from Waters (Synapt series).

Besides ESI and MALDI, other types of ionization methods have been used in IMS-MS analyses. A non-exhaustive list is provided hereafter: laser ablation ESI (LAESI),[86] paper spray ionization,[87] atmospheric pressure chemical ionization (APCI),[88] corona discharge ionization,[89] and the historical ^{63}Ni beta ionization.[90]

b) Interfacing IMS to mass analyzers

One of the main difficulties in interfacing IMS to MS lies with the compatibility of the two techniques in terms of scanning rate. IMS is indeed a relatively fast scanning technology, on a millisecond timescale, making time-of-flight (TOF) the easiest mass analyzer to couple to IMS because TOF operates on a microsecond timescale.[85] [h] Such IMS-TOF setups afford the acquisition of several mass spectra for every mobility scan, with a high mass resolving power. The IMS dimension also helps separate the ions of interest from background ions (noise). By summing a few mobility scans it becomes possible to observe ions at very low signal-to-noise ratios.

However, some other types of mass analyzers are more difficult to couple to IMS, specifically Fourier transform MS (FTMS) analyzers. These include the orbitrap and Fourier transform ion cyclotron resonance (FTICR) analyzers. Despite their high MS performance and especially their high MS resolving power, FTMS instruments have a scan rate on the order of a hundred

[h] This statement stands for classical TOF analyzers, the multi-reflecting TOF (MRT, Waters, Wilmslow, UK) analyzers that were recently released operate on a timescale nearing the millisecond—making them harder to couple to IMS.

millisecond or a second, which means on the same order if not slower than the IMS step. This is a major limitation because this implies that MS data cannot be properly aligned with the IM dimension.

Nevertheless, spatially dispersive IM analyzers such as FAIMS can be coupled to MS analyzers under these constraints because they involve scanning voltages instead of time. FAIMS has been successfully combined to orbitrap[91] and FTICR[92] instruments.

The community has endeavored to couple other IMS methods with FTMS analyzers by slowing down the ion mobility dimension to make it compatible with the MS scan rate.[85] For instance, ion gating combined with multiplexing approaches were developed to trap ions before and after DTIMS separation and thus "divide" the mobility dimension to make it compatible with orbitrap analysis.[93] Some people have also worked on coupling TIMS with FTICR analyzers using gating approaches,[94] selective accumulation,[95] or nonlinear mobility scan functions.[96] Others have also developed a FT-IMS multiplexing approach to overcome the duty cycle mismatch between DTIMS and orbitraps.[97]

In summary, there are a number of MS analyzers that can be considered in combination with IMS, the most straightforward being the TOF analyzer that has a compatible duty cycle. The TOF analyzer is the analyzer implemented on the cIMS instrument and is described later in the manuscript (**Chapter 3**).

4. Towards high-resolution IMS

The resolving power (R_p) in IMS is the capacity to separate ions of close CCSs and is measured by the ratio Ω/δΩ. The resolving power of the first generation of commercial IMS-MS instruments in the 2010s was on the order of 40–60, which was limiting for a number of applications.[98] Years of collective efforts have been aimed at raising the IM resolving power and thereby increasing the level of information brought by IMS. A first step towards improved R_p on commercial setups was the introduction of the TIMS in 2018 by the Bruker company,[76] which can reach R_p up to 400.[99]

However, the most promising approach towards (ultra)high resolving power in IMS currently seems to be to increase the length of the separation path. In DTIMS and TWIMS, R_p scales with the square root of the charge of ions (z), the ion path length (L), the electric field (E), and buffer gas temperature (T) (**Equation 2**).[100]

$$R_p \propto \sqrt{\frac{zLE}{T}} \qquad (2)$$

a) Cyclic IMS

In 2009, D. E. Clemmer and collaborators introduced the concept of a cyclic IM device to increase the separation path length under the name “ion cyclotron mobility spectrometry”.[101] The idea is that by making ion undergo multiple passes in a single device, the separation path length—and thus R_p—could theoretically be increased almost infinitely. This cyclotron spectrometer consisted of four drift tubes put in series (**Figure 23**).

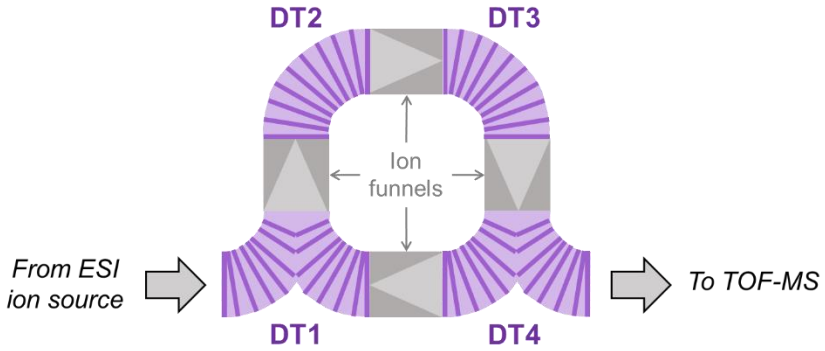


Figure 23. Ion cyclotron mobility spectrometer introduced by Merenbloom *et al.*[101]

The concept of a cyclic multipass device was developed by Giles *et al.* in combination with the TWIMS technology instead of the serial drift tubes.[102] The main advantage of switching to TWIMS is that TWIMS uses a dynamic electric field that can be maintained in a single racetrack over extended path lengths without needing to increase voltages (contrary to DTIMS). In their prototype, the mobility ring was placed orthogonally to the main ion optical axis to allow for finer ion manipulations (**Figure 24**). This setup has been launched commercially by Waters in 2019 under the name “SELECT Series Cyclic IMS” (Wilmslow, UK) and was the main instrument used in the context of this thesis. It will thus be integrally described in **Chapter 3**. As of the writing of this thesis, the maximum R_p recorded on the SELECT Series Cyclic IMS is 920.[103]

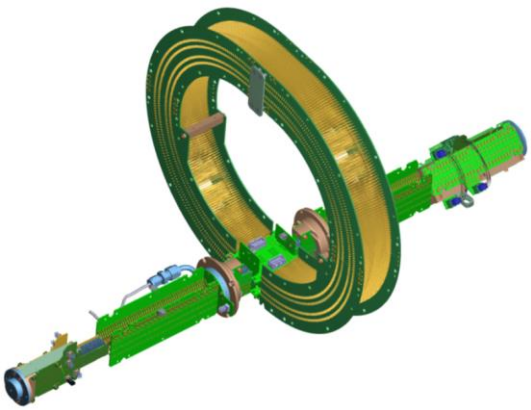


Figure 24. The cyclic TWIMS cell of the SELECT SERIES Cyclic IMS (courtesy of Waters Corp.)

b) Structures for Lossless Ion Manipulation (SLIM)

The SLIM technology, introduced by the group of R. D. Smith,[104] is another TWIMS-based IMS technique aiming at increasing the path length of ions. A particularly ingenious serpentine design on printed circuit boards (PCBs, **Figure 25**) allowed to optimize the space and path length and was shown to achieve maximal ion transmission.[105] Later on, this initial design was followed by a multipass SLIM design (SUPER SLIM).[106] These developments resulted in R_p in excess of 1800.[106] Since then, a number of developments were conducted, notably to enhance the duty cycle[107] or to achieve fine control of ions on PCBs, for instance, in a multi-level SLIM with ion elevators.[108]

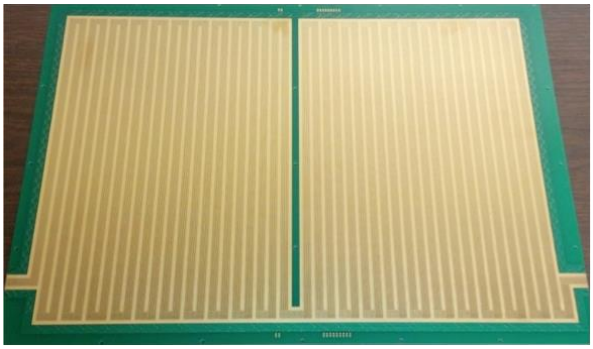


Figure 25. Serpentine SLIM built using PCBs (© Pacific Northwest National Laboratory)

Since SLIM are built using PCBs, they can be build in-house although some specific components are not easily accessible to all laboratories such as trapping ion funnels or traveling wave generators. Recently, Kinlein *et al*.[109] discussed these issues and attempted to provide solutions to make prototyping of SLIM devices easier. SLIM has been commercialized in 2020 under the name MOBIE (Mobilion Systems, Chadds Ford, PA, USA) in partnership with Agilent.

B. Expected benefits of high-resolution IMS (HR-IMS) for glycosciences

In **Chapter 1**, I described the inherent difficulties that analysts are faced with when attempting to characterize carbohydrates. These difficulties have much to do with the structural diversity of carbohydrates and their inherent isomerism. The following paragraph thus emphasizes on two points specifically: (i) the characterization of carbohydrates isomers by MS and tandem MS is complicated by the occurrence of numerous levels of isomerism, and (ii) the chromatographic separation of carbohydrates remains challenging, even when using specific stationary phases such as PGC.

The greatest expected benefit of high-resolution approaches in IMS is the separation of isomeric compounds that would not be separated in chromatography, which is a major concern for carbohydrates. Dodds *et al.*[98] have previously described the IMS resolving power required

to separate isomers expressed as percent difference in CCS for protonated molecules (**Figure 26**). To separate constitutional isomers (in the case of carbohydrates, regioisomers and cyclization tautomers), a R_p of around 100 is needed to achieve baseline separation. This is already above the abilities of classical DTIMS and linear TWIMS instruments ($R_p \approx 50$), but can readily be achieved with HR-IMS.

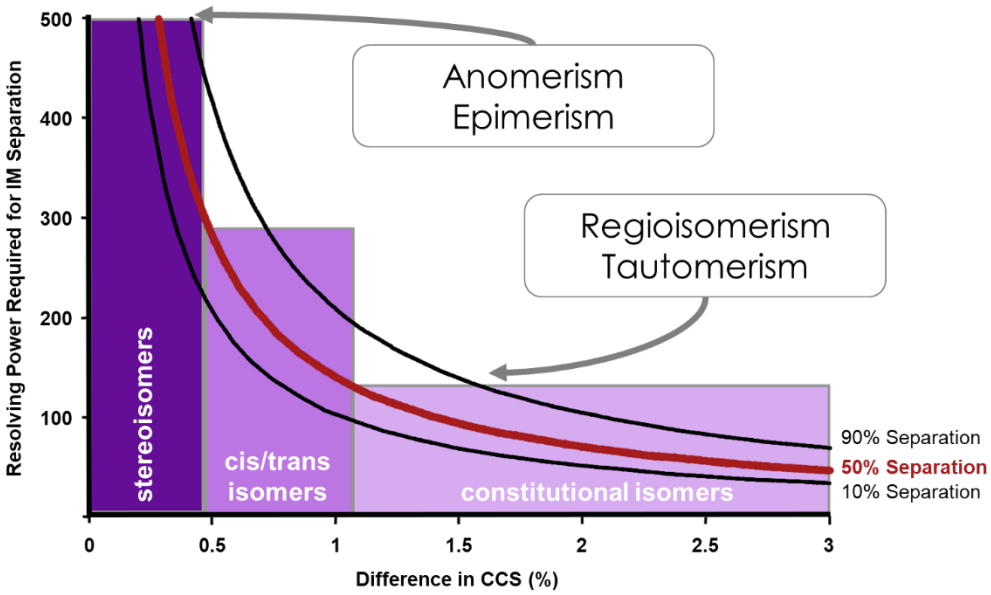


Figure 26. Resolving power required to separate the different types of isomers, adapted from Dodds *et al*.[98]

To separate closer structures such as stereoisomers (in the case of carbohydrates, anomers and epimers), a higher R_p is required, in excess of 500. This resolving power is now reachable with HR-IMS, the cIMS setup having notably reached $R_p > 900$ on oligosaccharides[103] and the SLIM device $R_p > 1800$.[106]

Another benefit that can be expected of HR-IMS is the separation of species that could not be (fully) resolved using chromatographic approaches. Incompletely resolved chromatographic separation happens frequently in situations of extreme molecular diversity, with thousands or tens of thousands of analytes in the sample (as may be encountered in “omics” analyses). Because elution of an analyte in liquid chromatography (LC) yields peaks with a second/minute timescale, it can be coupled to IMS-MS to generate a three-dimensional LC-IMS-MS data. This hopefully allows to completely resolve all species present in a mixture, even more if IMS operates with high resolving power. Unresolved separation in LC is also directly linked to the problem of isomerism: the closer the structures, the more difficult it is to find a stationary phase and/or solvent system able to separate them in chromatography.

C. State of the art of IMS applied to carbohydrate analysis at the beginning of this thesis

Prior to the start of my thesis, several interesting experiments suggested that IMS, notably with an increase in resolving power, would prove extremely useful in the study of carbohydrates. This section presents an overview of previous work conducted on IMS analysis for carbohydrates before I started my PhD.

1. Pioneering experiments at low IM resolution

Even before the introduction of HR-IMS setups, IMS showed promising results supporting its potential for glycomics. Notably, Gray *et al.*[110] demonstrated that IMS-MS combined to CID could identify monomers within a carbohydrate sequence, with fragments of different glucobioses having the same mobility as glucose. Furthermore, in this study, α-linked glucobioses exhibited an additional IM peak compared to the β-linked ones. IMS was also shown to be able to separate linkage anomers,[111] and used to build CCS databases that could serve to identify glycans (notably of mammalian *N*-glycans).[112] A notable drawback of such database approaches is the availability of standards that can be limited for certain types of oligosaccharides (including plant glycans), making it difficult to build an exhaustive database.

In the early 2010s, Huang and Dodds showed the interest of varying the ion type to improve the IMS separation, by studying group I and group II metal adducts.[113,114] They notably showed that the CCS increased with group I metal ion radius, in an isomer specific manner. Different metal ions binding differently to carbohydrates, changing the cation could also result in distinctive IM profiles.

2. Insight brought by the first high-resolution IMS experiments

On the one hand, a number of studies were conducted on oligosaccharide standards, providing a first step towards resolving the isomerism problems in carbohydrates. Deng *et al*.[115] were the first to demonstrate the use of a HR-IMS system to analyze carbohydrates, interestingly resolving multiple components of pentasaccharides (*i.e.*, multiple IM peaks for a single pentasaccharide). Nagy *et al*.[116] further showed that non-reducing oligosaccharides exhibited a single IM peak while reducing oligosaccharides exhibited multiple peaks. Ujma *et al*.[117] then demonstrated that these multiple IM peaks actually resulted from the co-existence of reducing-end anomers of the pentasaccharides (α-anomer, β-anomer and open-chain intermediary).

On another hand, Ropartz *et al*.[103] also showed the interest of HR-IMS coupled to fragmentation approaches for the structural characterization of oligosaccharides in biological samples. In this study, post-IMS fragmentation was performed on a mixture of oligoporphyrans extracted from red algae. Precise time alignment of fragments to their cognate parent ion after

IMS allowed unambiguous assignment of fragments from a given isomer and thereby, led to the unambiguous structural determination of nearly all of the expected oligoporphyrans studied.

The experiments described above, both in low and high resolution IMS, built a solid foundation to support the use of IMS-MS in structural glycosciences. Even more decisive results could be expected thanks to the arrival of commercial instruments with high resolution IMS and specific functionalities like the SELECT SERIES Cyclic IMS. This is what motivated my thesis work.

D. Current developments in IMS-MS for glycosciences

To conclude this Chapter, other groups conducted several interesting pieces of work concomitantly to this thesis. They are summarized thereafter.

1. SLIM- MS coupled to infrared spectroscopy

One very interesting approach, developed by the group of T. R. Rizzo, combines the power of HR-IMS (on a home-built SLIM device) to gas-phase infrared spectroscopy in a cryogenic ion trap.[118] They used this setup to separate oligosaccharides using SLIM IMS, and then identify the species in a mixture using IR data and comparison to reference spectra. This strategy was applied to a number of glycan types, including glycans cleaved from antibodies[119] or human milk oligosaccharides.[120]

This setup notably permitted the identification of the reducing-end anomeric peaks described above by comparison with α-methylated and β-methylated standards.[121] The group further conducted work that consolidated some previous evidence that there was an “anomeric memory” of glycans upon collisional dissociation (Schindler *et al*.[42]), in other words, that anomerism of the bond was maintained in CID experiments.[122]

The implementation of on-board CID in the SLIM part of their device allowed Rizzo’s group to improve the structural information obtained through the use of tandem IMS experiments.[123] Notably, this has allowed the identification of N-glycan positional isomers.[124] The group is now developing strategies aiming to move towards high-throughput analysis of glycans.[125,126]

2. Other works using the cIMS setup on glycans

The group of G. Nagy recently studied the impact of group I metal adduction using HR-IMS on a cIMS instrument (as a function of degree of polymerization, on α1,4- and β1,4-glucans).[127] Notably, they showed that varying the counter cation could yield more distinctive HR-IMS profiles between the α- and β-glucans (but there was no general trend regarding which cation was better). They later developed a temporal compression strategy designed to refocus HR-IMS peaks on the cIMS setup, as well as to improve peak intensity. The proof of concept of

this method allowed the separation of monosaccharides adducted to α-cyclodextrin, and of lactose isotopologues.[128]

Some work has also been conducted on the coupling of HILIC chromatography with a multipass cIMS approach to resolve heparan sulfate oligosaccharide positional isomers.[129]

3. Shotgun sequencing approach (Pagel's group)

Finally, Miller *et al*.[130] recently published a noteworthy approach on heparan sulfate oligosaccharides, called shotgun IMS-MS sequencing (SIMMS^2). The strategy consists of fragmenting saccharides prior to IM analysis, and then compare the IM spectrum of the generated fragments to the IM spectra of a database built from smaller oligosaccharides. By assigning several fragments, the user can gain good sequence coverage and assign the sequence of the parent oligosaccharide.

This approach appears to be very promising, especially as this study was conducted using low-resolution DTIMS. It can thus be expected that the results could be further improved by relying on HR-IMS.

Chapter 3 – The SELECT SERIES Cyclic IMS instrument

All experiments presented in this manuscript were conducted on the SELECT Series Cyclic IMS instrument (Waters, Wilmslow, UK), abbreviated hereafter as “cIMS”. This Chapter therefore provides a complete overview of the instrument geometry, as well as of the different sequences of analyses that are available on the instrument. Indeed, the intricate geometry offers a wide range of possibilities for ion manipulation, and some of the experiments performed amidst our work require a detailed description to be properly understood.

A. Instrument geometry

The cIMS setup is a q-IM-TOF instrument (**Figure 27**). Ions are first generated in an electrospray ion source (ESI) with a Z-spray™ geometry. The ion beam is introduced into the mass spectrometer through a stepwave ion guide (ensuring that the remaining neutral molecules do not go in). Ions can then be selected according to their *m/z* using a quadrupole before any IMS and/or fragmentation experiment. The quadrupole is directly followed by a linear “Trap” cell used to perform pre-IMS CID. After this, the cyclic mobility cell is found, preceded by a helium cell used to cool down ions before their injection into the IMS cell. The IMS cell is followed by another CID cell, this time a segmented quadrupole called “Transfer cell”. Finally, high-resolution MS analysis is performed with a high performance TOF analyzer with a dual V or W geometry (for enhanced sensitivity or resolution, respectively).

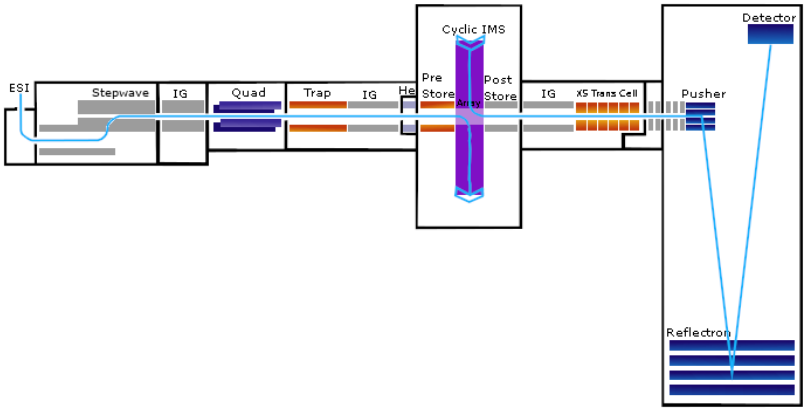


Figure 27. Scheme of the cIMS instrument

Additionally, the setup can be augmented with an additional fragmentation cell—the ExD cell[i]—used to perform electron-based fragmentation. An ExD WK-150 cell commercialized by e-MSion, Inc. (Portland, Oregon) in partnership with Waters has been implemented in our

[i] “ExD” is a term found in the literature to describe to electron-based dissociation methods in general, but is also the name chosen by e-MSion Inc. for the commercialization of their fragmentation cell.

laboratory. The ExD cell can be positioned either: (i) between the trap cell and the IMS cell to perform pre-IMS fragmentation, or (ii) between the IMS cell and the transfer cell to perform post-IMS fragmentation.

A brief mechanistic overview of the main parts of the spectrometer is given thereafter.

1. Electrospray ion source

The cIMS instrument comes with an electrospray ion source for atmospheric pressure ionization. The electrospray method was mainly developed by John B. Fenn and coworkers in the 1980s—owing him the 2002 Nobel Prize in Chemistry—and relies on the generation of charged droplets of an analyte solution.

In the basic electrospray design described by Fenn et al.[131] the liquid sample is introduced in the ion source through a needle, which is enclosed in a cylindrical electrode where high voltages are applied to generate the spray of charged droplets. A drying gas flow is applied to help desolvate the molecules and generate gas-phase ions (**Figure 28A**). In the Waters ZSpray ion source, an additional nebulizer gas is added. To limit the introduction of neutrals, contaminants, or residual solvent droplets into the mass spectrometer, the spray is positioned orthogonally to the sampling cone and a counter-gas at low flow comes through the sampling cone (**Figure 28B**).

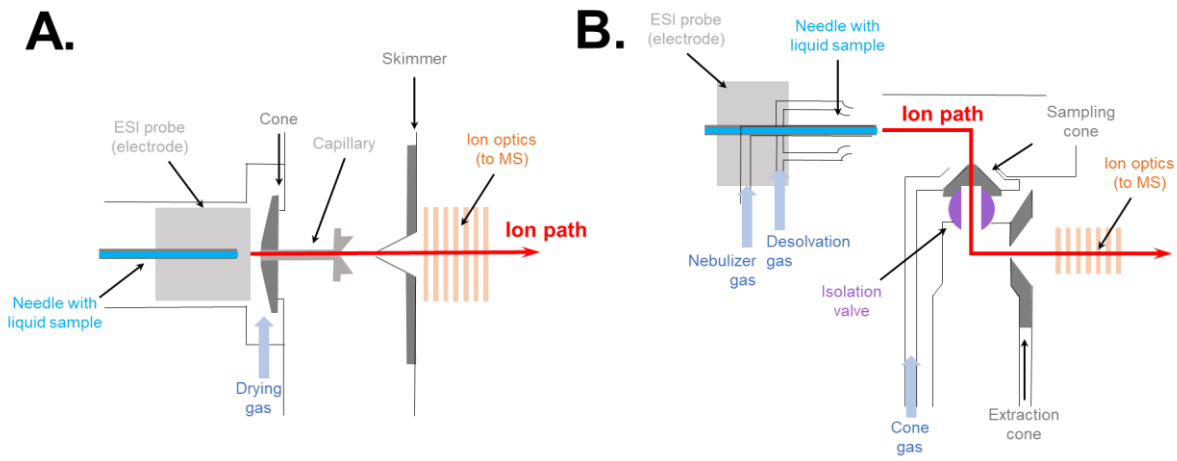


Figure 28. A. Schematic of a basic electrospray ion source. **B.** Schematic of the ZSpray ion source from Waters.

One of the main advantages of ESI is its ability to perform soft ionization of small and large molecules alike, *i.e.*, to ionize the molecules while inducing minimal activation and fragmentation. This property is inherently linked to the way in which the ions are generated from charged droplets of the sample solution. Different models have been proposed to explain how gas-phase ions are then generated (**Figure 29**). The base idea is that evaporation would

increase the surface-charge density of the droplets, eventually reaching the Rayleigh limit and thus inducing Coulombic explosion. The ion evaporation model (IEM) proposes that the decrease in droplet size and build-up of an electrostatic field at the droplet surface would lift the solute ions into the gas phase. The charge residue model (CRM) proposes that the series of Coulombic explosions would end up giving droplets so small that they would only contain ultimately a single charged molecule.[132] Finally, a third model has been proposed for larger disordered polymers (possibly including polysaccharides), combining properties of both previous models, the chain ejection model.[133]

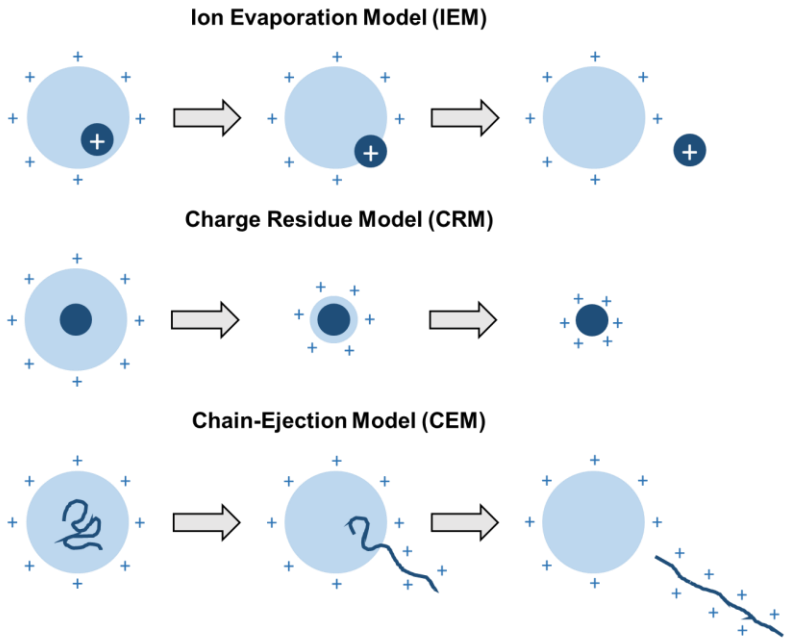


Figure 29. The three models of ion formation from ESI droplets.

Overall, ESI is a soft ionization method affording the generation of ions without much denaturation of the parent compound, mainly generating (de)protonated molecules or metal adducts. It has become arguably the most popular ionization method in mass spectrometry applied to biomolecules, especially because it can be easily coupled to liquid chromatography.

2. Quadrupole

The cIMS setup includes an extended mass range quadrupole (up to *m/z* 8000) for mass selection of ions prior to fragmentation and IMS analyses. Quadrupoles can serve both as ions guides, as mass filters, as mass analyzers, or even as linear ion traps. They are also very intricate pieces of technology and their operating principle relies on several complex equations that define the ion motions within the electric fields of the quadrupole.[134]

The quadrupole, originally invented by Nobel laureate Wolfgang Paul, consists of four parallel metal rods. The opposing rods are connected in pairs, and a radio frequency voltage (RF) is applied between the two pairs (**Figure 30**).

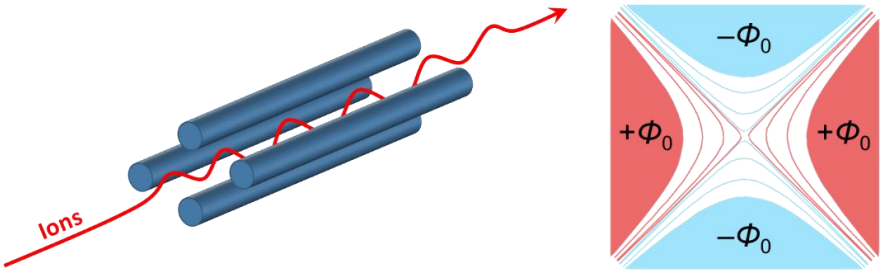


Figure 30. Schematic of a quadrupole mass filter

For a given ratio of voltages, only ions within a given *m/z* range will have a stable trajectory within the quadrupole, and the others will not make it through the filter, being ejected or colliding with the rods—as expressed by the Mathieu equation (**Equation 3**) and stability diagram (**Figure 31**):

$$\frac{d^2x}{d\xi^2} + (a_x - 2q_x cos2\xi)x = 0 \qquad (3)$$

with a_x and q_x being the key parameters defining the stability of the ion motion, and ξ being proportional to the time-varying radio-frequency. The term a_x is a function of the DC voltage applied pole to ground (U), whereas the term q_x is a function of the zero-to-peak alternating voltage applied pole to ground (V_{RF}).

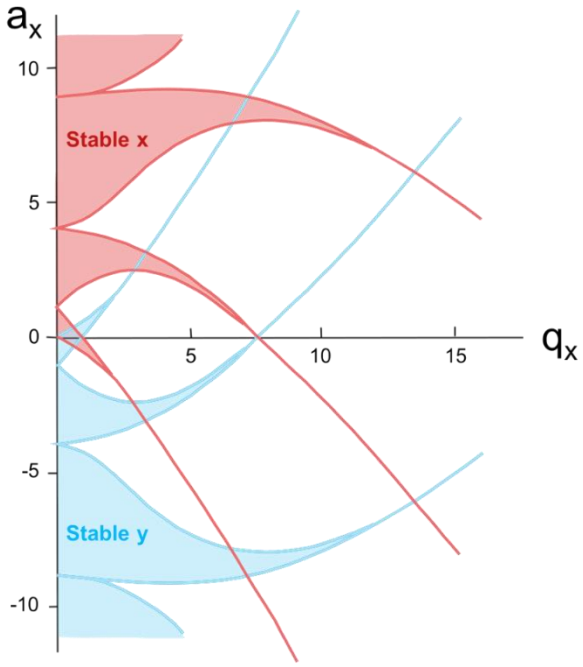


Figure 31. Stability diagram of ions in a linear quadrupole

This differential stability of ions is most useful, as it allows easily selecting a given *m/z* that will pass through the rods and be considered for further analysis.

3. Collision cells for CID

The Cyclic IMS setup includes two collisions cells for CID fragmentation, operating using N_2 as a collision gas.

- The first collision cell, located before the ion mobility cell, is a linear ion trap. This cell ("Trap") operates at a pressure of 2.5e-2 mbar.
- The second collision cell, located after the ion mobility cell, is a segmented quadrupole ("Transfer cell"). It operates at a pressure of 1.8e-2 mbar. Note that CID fragmentation is slightly more efficient in segmented quadrupoles compared to linear ion traps, *i.e.*, the same fragmentation pattern will be obtained while applying a lower activation voltage in the transfer cell compared to the trap cell.

Additionally, the cIMS affords a third potential location to perform CID: the pre-array store cell that precedes the IMS cell and can be used to store ions between different rounds of IMS analysis. The voltages in this pre-store cell can be lifted while reinjecting ions in the IMS cell, producing an acceleration of the ions upon reinjection. This approach has been widely used during my PhD work and will thus be further described below.

However, this approach uses the drift gas of the IMS cell as a proxy for the collision gas. Since CID fragmentation efficiency is highly dependent on the probability of collision (and thus on the gas pressure), CID performed under these conditions is less efficient compared to other collision cells, as the gas pressure is more adapted to IMS analysis (1.7 mbar).

4. ExD cell

The ExD cell (commercialized by e-MSion, Inc.) is a recent innovation that allows users to perform electron-based dissociation on many MS platforms (including Waters, Thermo-Fisher Scientific, and Agilent instruments). It is a great innovation as such methods were previously limited to FTICR instruments.[32]

The ExD cell is composed of a series of guiding electrodes used to direct the ion beam to a rhenium filament where a current is applied, generating a cloud of electrons. These electrons are confined near the filament using two lens magnets. An additional annular lens is used to confine the ions in the cloud of electrons. A scheme of the ExD cell is provided in **Figure 32**. The user can then fine-tune the voltages applied on the lenses and the filament bias to adjust the exposure of the ions of interest to the electrons, and perform electron capture dissociation (ECD) or electron-induced dissociation (EID).

The basic principles, advantages and limits of ECD and EID will be presented thereafter. For more details about electron-mediated fragmentation methods, reviews are available elsewhere.[32,135]

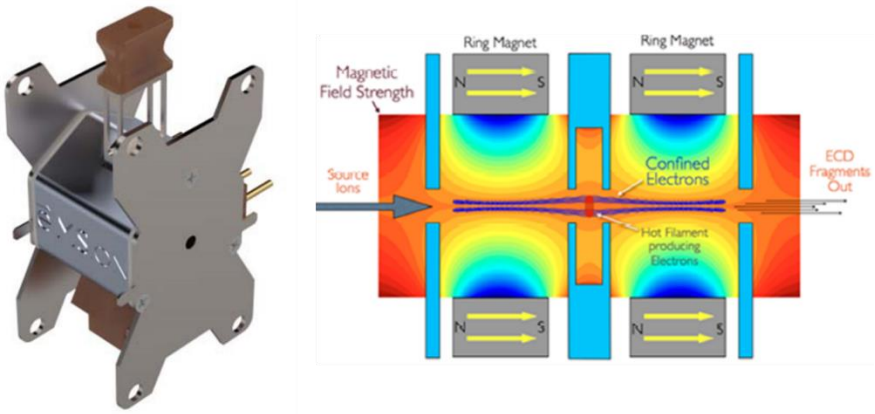


Figure 32. A. The ExD cell module for the Select Series Cyclic IMS, **B.** schematic of the ExD cell. (© e-MSion, Inc.)

- ECD results from the capture of low energy electrons (0.2–1 eV) by multiply charged cations. An exothermic positive-negative charge recombination occurs, inducing the formation of radical cations. Fragmentation is then triggered depending on the site of radical formation, which explains why the cleaved bonds are not necessarily the most labile ones. This explains why ECD fragmentation spectra are generally more structurally informative compared to CID, but the fragmentation efficiency is lower and longer acquisition times are required. Additionally, because ECD involves the capture of an electron by a cation, it can only be used on multiply charged species (singly charged cations would become neutral after capturing the electron).
- EID results from the irradiation of ions by high-energy electrons (6–20 eV). The fragmentation is triggered by an electron ionization-like vibrational excitation. The fragments formed by EID are relatively similar to that formed by ECD, but EID has the notable advantage that it can be used on both anions and cations, including on singly charged species. The fragmentation efficiency is in the same range compared to ECD, and it thus requires longer acquisition times than CID.

5. Cyclic TWIMS

The core part of the cIMS instrument is the cyclic TWIMS cell. It is a closed-loop IMS cell with a 98-cm single pass length. It consists of over 600 electrodes supported by PCBs, placed orthogonally to the ion path direction of the mass spectrometer (**Figure 33A**). The PCBs are spaced by 5 cm, and the electrodes by 0.5 cm, with additional repellers where voltages are applied to confine the ions in the mobility cell (**Figure 33B**). The racetrack is filled with N_2 as the IM gas.[102]

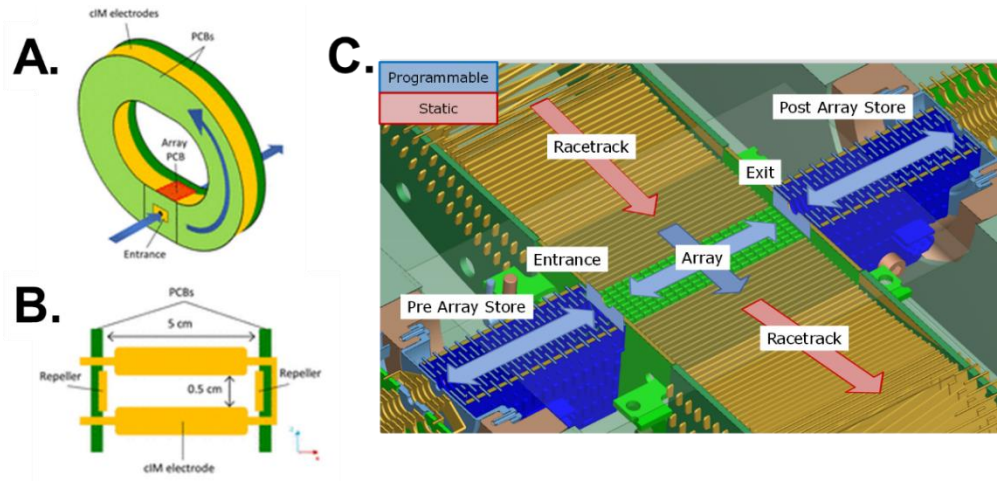


Figure 33. A-B. Schematic of the cIMS cell and **C.** schematic of the multifunction array region (courtesy of Waters Corp.)

The junction between the IMS cell and the rest of the spectrometer is done through a multifunction array region (**Figure 33C**). This array uses a traveling wave to control the ions' path and can send them sideways (for separation), forward (for acquisition and/or ejection), or even backwards to the pre-array store cell.

This peculiar ion optics afford an extended range of gas-phase ion manipulations, notably giving access to multi-stage IMS (that I will detail in the later parts of this Chapter).

As mentioned above, the IMS resolving power of traveling wave separations is proportional to the square root of the ion path length. The single pass IMS resolving power (expressed as FWHM CCS/δCCS) of the cyclic IMS is described to be around 70 for a single pass, and is thus $70.\sqrt{(n)}$ for n passes.

6. High-resolution TOF analyzer

Finally, the cIMS device is equipped with a high-resolution orthogonal acceleration time-of-flight (oa-TOF) analyzer. TOF analyzers measure the time it takes an ion (*in vacuo*) to fly from the entrance of the analyzer to the detector. To increase the mass resolving power of TOF analyzers, the ion path in the flight tube typically makes a U-turn to compensate for initial distribution of kinetic energy. While making the U-turn, the ions with higher kinetic energy will dive deeper through the reflective lenses (**Figure 34A**), which has the effect of refocusing the ion packets.[136]

In an oa-TOF like the one found on the cIMS instrument, a pusher will send packets of ions using impulsions on a nanosecond timescale, and said packets will then fly through the analyzer on a microsecond timescale. This rapid analysis is one of the main strength of TOF analyzers, making it perfectly compatible with TWIMS separation, which operates on a millisecond timescale.

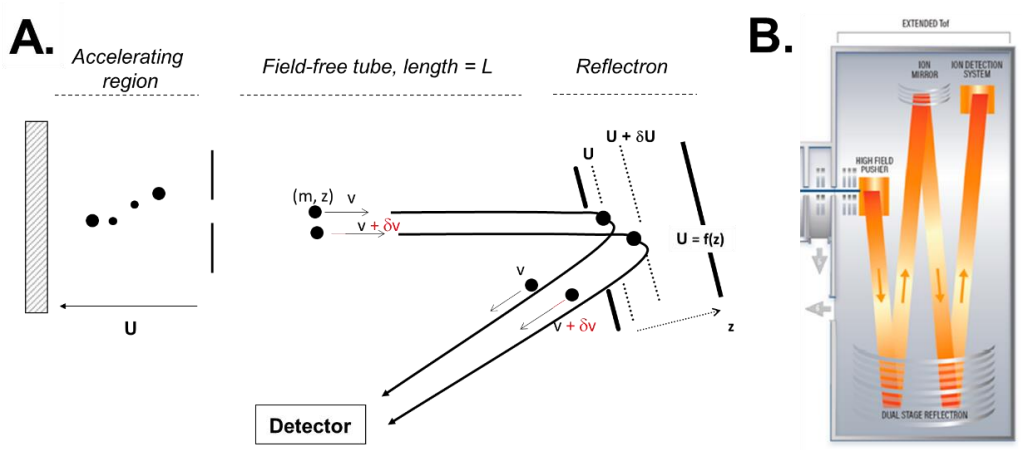


Figure 34. A. Principle of a reflectron TOF analyzer; **B.** The triple-reflection (W-mode) TOF analyzer available on the cIMS setup (courtesy of Waters Corp.)

Another strong point of modern TOFs is their mass resolving power. For instance, the cyclic IMS has two TOF operation modes:

- A sensitivity mode, where a single reflection of the ion beam is performed, reaches a theoretical mass resolving power (FWHM) R > 60,000
- An enhanced resolution mode, where three reflections of the ion beam are performed (**Figure 34B**), reaches R > 80,000

It is notable that the resolving power of TOF analyzers increases with the flight time, *i.e.*, with the *m/z* of the populations under study (**Figure 35**). This is of special interest as it allows analyzing high molecular weight species with high MS resolution, whereas other high-resolution FT-MS analyzers (FTICR, Orbitraps) suffer a loss in resolution with increasing *m/z*.

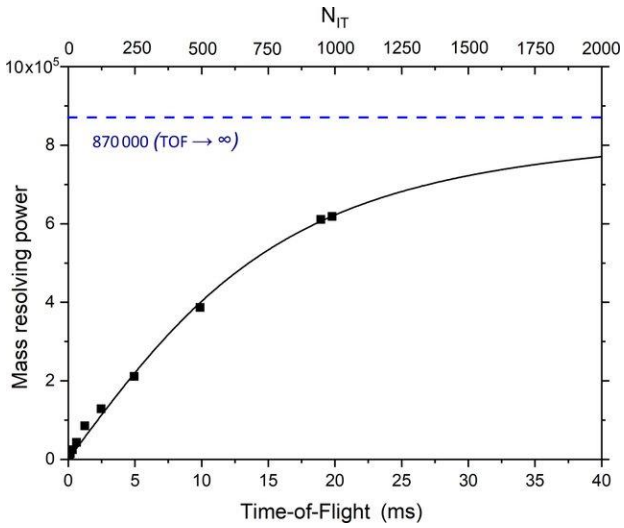


Figure 35. Profile of the resolving power of a TOF analyzer (for a multiple-reflection TOF, as a function of the time of flight and of the number of turns).
Reprinted from Ayet San Andrés *et al*.[137]

B. Operating modes

This section describes the different types of acquisition that can be performed on the Cyclic IMS instrument and that I have performed in the span of this thesis. In the following description, I will consider that an *m/z* isolation has been performed using the quadrupole, unless stated otherwise.

1. Single and multipass separations

The most basic operation that can be performed using cIMS is to simply separate ions around the cyclic mobility cell. The cIMS is controlled using a sequence of events applied to the multifunction array. The ions are first injected into the ion mobility cell by applying a forward traveling wave in the multifunction array, then the direction of the array switches to a sideways traveling wave to separate ions over one or multiple passes (until the resolution is satisfactory), and finally, the direction switches back to a forward wave to send the ions to the TOF analyzer (**Figure 36**).

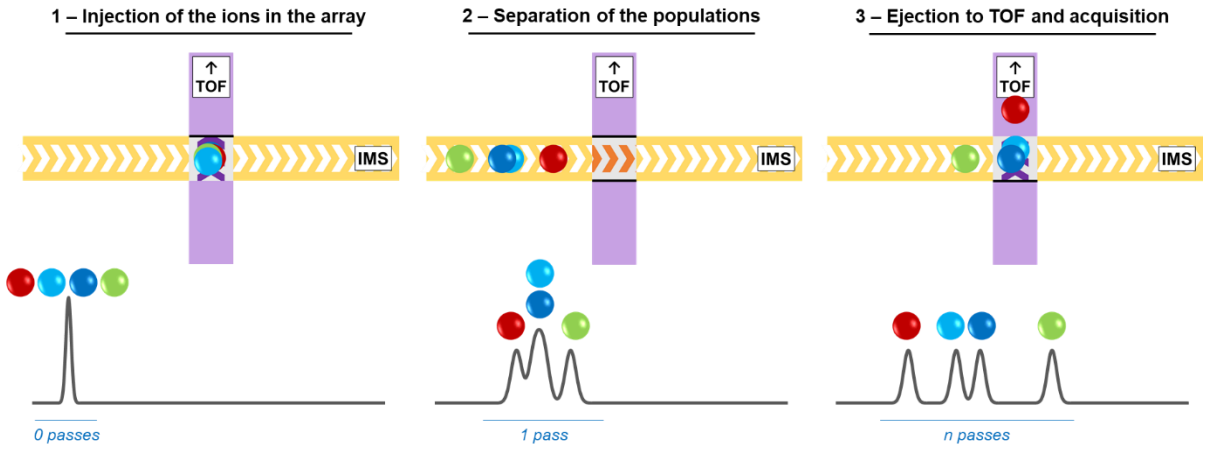


Figure 36. Sequence of events used for single/multipass separations

A major limitation of the cyclic design is that, when performing multipass separation, the faster populations can catch up to the slower ones. It then becomes difficult to identify the different peaks. This phenomenon is called “wrap-around”, and can be solved using IMS^n strategies that require more complex cIMS sequences. These strategies will be described later in this Chapter.

2. CID-IMS[j]

The first type of fragmentation-based analysis that can be performed using the cIMS setup is CID-IMS. In CID-IMS, the ion of interest is first fragmented in the trap collision cell to generate the MS/MS fragmentation spectrum (**Figure 37A**). Because all of the fragments are generated prior to IMS separation, it is then possible to acquire the ion mobility profile of all of the fragments. When analyzing the IM profile of fragments in a graph plotting the arrival time vs *m/z*, the fragments form a diagonal starting from the precursor ion (**Figure 37B**).

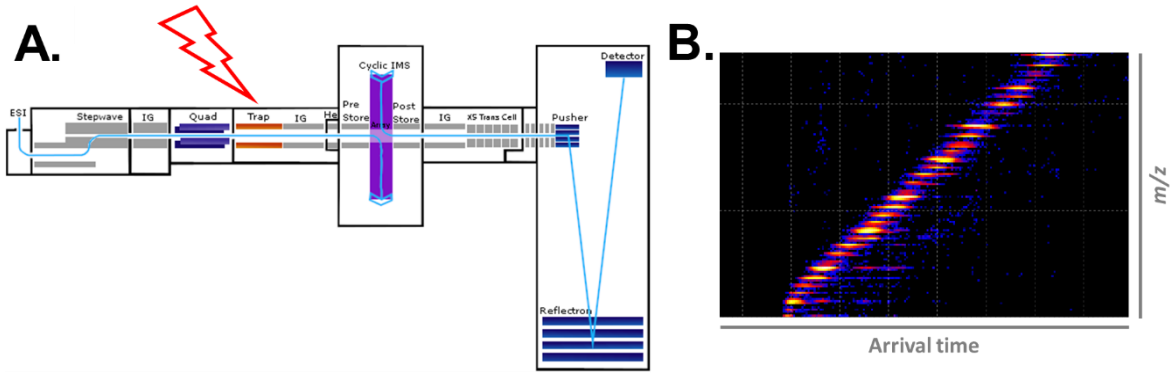


Figure 37. A. Site of fragmentation in the mass spectrometer during CID-IMS experiments; **B.** Graph plotting the IMS arrival time vs *m/z* of the fragments in a CID-IMS experiment

Note, however, that since the fragments may expand through a wide range of mobilities, the smaller fragments may rapidly catch up to the larger ones when performing multi-pass (high-resolution) IMS separations, resulting in the wrap-around phenomenon described above.

3. IMS-CID[k]

A second type of analysis that can be performed on the instrument is IMS-CID. During this type of analysis, the parent ions are first separated using high-resolution ion mobility. When the desired separation is attained, the ions are fragmented in the transfer collision cell (**Figure 38A**). The fragmentation occurs on a timescale that is compatible with the ion mobility separation. In a graph plotting the IMS arrival time vs *m/z* of the fragments, the fragments are thus aligned at the arrival time of their related precursor. This allows to unequivocally assign the fragments arising from a given precursor, and to reconstruct the MS/MS spectra of all the species separated upstream using high resolution IMS.

[j] On an instrument equipped with the ExD system, ExD-IMS analyses can be performed instead of CID-IMS analyses.
[k] On an instrument equipped with the ExD system, IMS-ExD analyses can be performed instead of IMS-CID analyses.

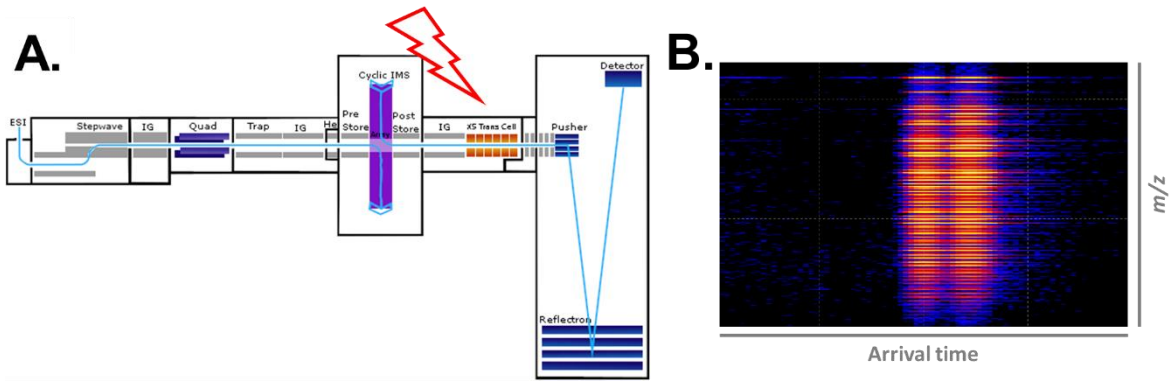


Figure 38. A. Site of fragmentation in the mass spectrometer during IMS-CID experiments; **B.** IMS arrival time vs *m/z* of the fragments of IMS-CID data

4. CID-IMS-CID[l]

If desired, the instrument can also be operated to perform CID-IMS-CID, *i.e.*, the ions will be fragmented twice to perform MS^3 (**Figure 39**), but with a step of ion mobility in-between the fragmentation steps. This helps distinguishing MS^2 fragments. For instance, this can be useful in structural studies where two isomeric fragments are generated during the first stage of CID: the two isomeric fragments can be separated by IMS and characterized in the second stage of CID.

In most cases, this will not be the preferred strategy, because the various fragments generated during the first stage of CID will overlap during high-resolution IMS separations, complicating the interpretation of the spectra generated in the second stage of CID. The approach of choice is instead to perform multistage IMS experiments (IMS^n), as described below.

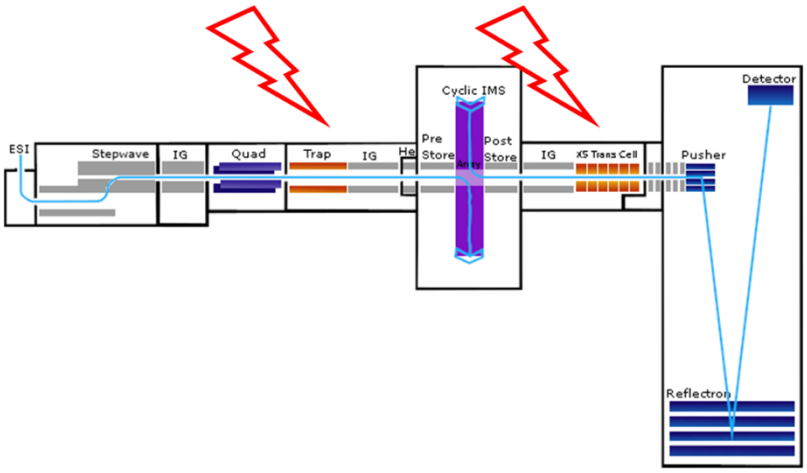


Figure 39. Sites of fragmentation in the mass spectrometer during CID-IMS-CID experiments

[l] It is not possible to use two ExD cells at once on the instrument; it is thus possible to perform ExD-IMS-CID or CID-IMS-ExD but not ExD-IMS-ExD.

5. IMS^n

I finally present the different types of experiments that are afforded by the multifunction array region. This region can be tuned finely to perform two main types of IMS^n experiments, by selecting a specific ion population: the “head-and-tail” and the “slicing” experiments (**Figure** 40).

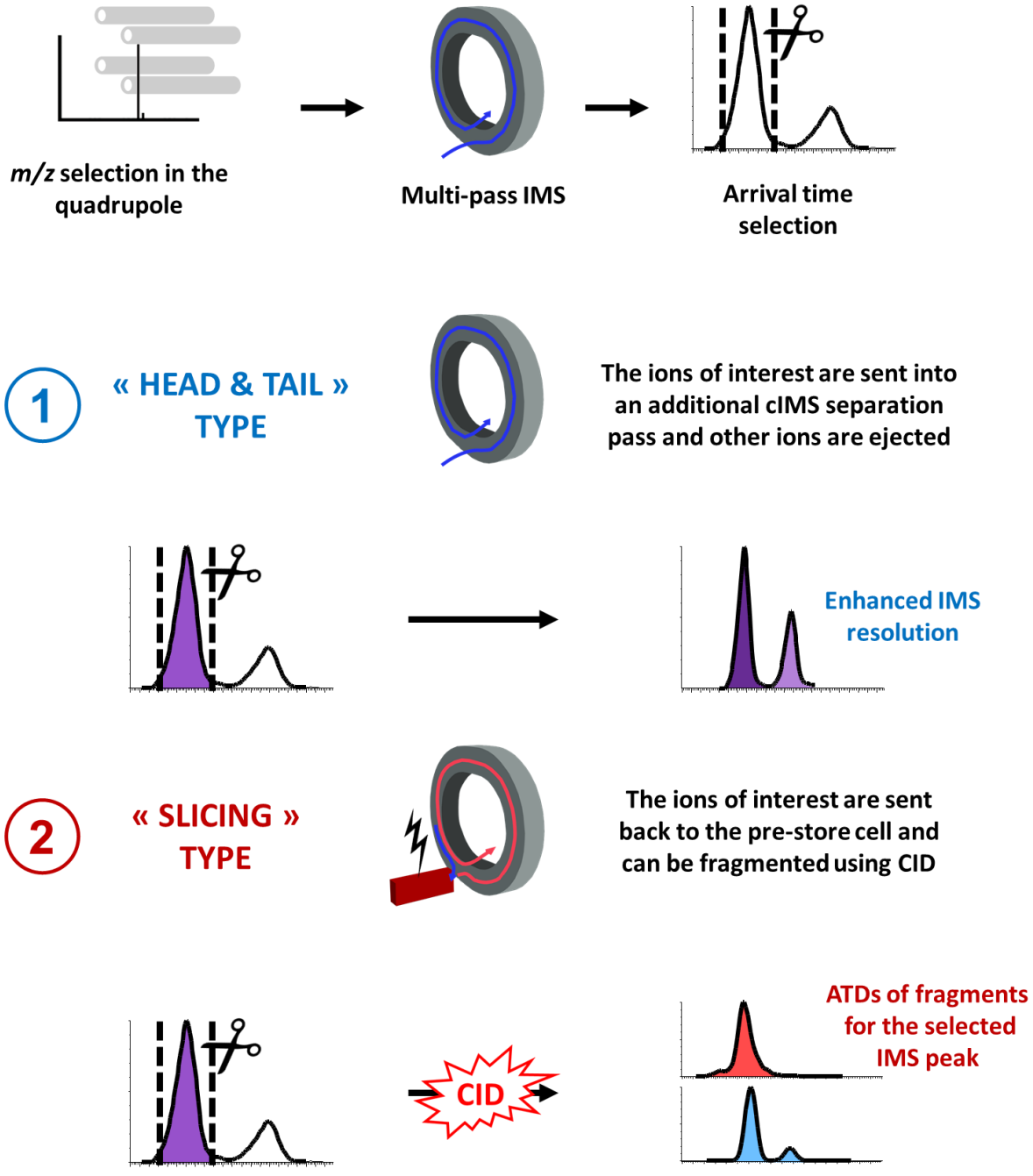


Figure 40. Flowchart describing the different kinds of cIMS-based selection of ions that can be performed using the SELECT SERIES Cyclic IMS

a) Head-and-tail

The first type of IMS^n experiment that can be performed is the so-called “head-and-tail” or “top-and-tail” experiment. This experiment mainly serves the purpose of avoiding the wrap-around phenomenon (see above). It is designed to specifically eject undesired ion populations—*i.e.*, ions with higher or lower mobilities than the population of interest—from the ion mobility cell, before they have enough time to catch up to the desired population (for the faster) or before the population of interest catches up to them (for the slower).

By switching the multifunction array region in the “eject” direction, the undesired ions are ejected from the mobility cell, thus clearing the IM range (**Figure 41**). This allows to keep separating the populations of interest without being confronted to superimposition with the undesired population(s).

The main advantage of the “head-and-tail” experiment is that the ion population of interest stays in the racetrack while ejecting the undesired populations. Therefore, the separation that was already obtained on the species of interest is preserved and can be enhanced by additional passes.

b) Slicing

The second type of IMS^n experiment—most useful for structural characterization studies—is the “slicing” experiment (**Figure 42**). Like the “head-and-tail” experiment, it empties the mobility cell of the undesired ion populations, but opposite to the “head-and-tail” experiment, the ions of interest are sent backwards in the pre-array store cell (*i.e.*, they do not remain in the cyclic racetrack). This is achieved by controlling the direction of the traveling wave in the multifunction array region (**Figure 43A**). Then, the ion population of interest is stored in the pre-array store cell for an amount of time defined by the user, after which it can be reinjected in the cyclic IMS racetrack (**Figure 43B**). It is notably possible to activate/fragment the selected population upon reinjection, by increasing the reinjection voltages and making the ions collide with the IM gas.

The main advantages include the possibility to perform CID between two rounds of IMS (IMS-CID-IMS) affording “true” IMS^n studies. Another notable advantage is the possibility to trap ions in the pre-store cell for a defined amount of time, which can for instance be useful to align two mobility experiments where the fragments of two precursors with different arrival times are to be compared.

The slicing experiment, however, also has the disadvantage that the ions of interest leave the racetrack and thus completely lose all the separation performed during the early stages of IMS. This phenomenon induces a slight loss in duty cycle.

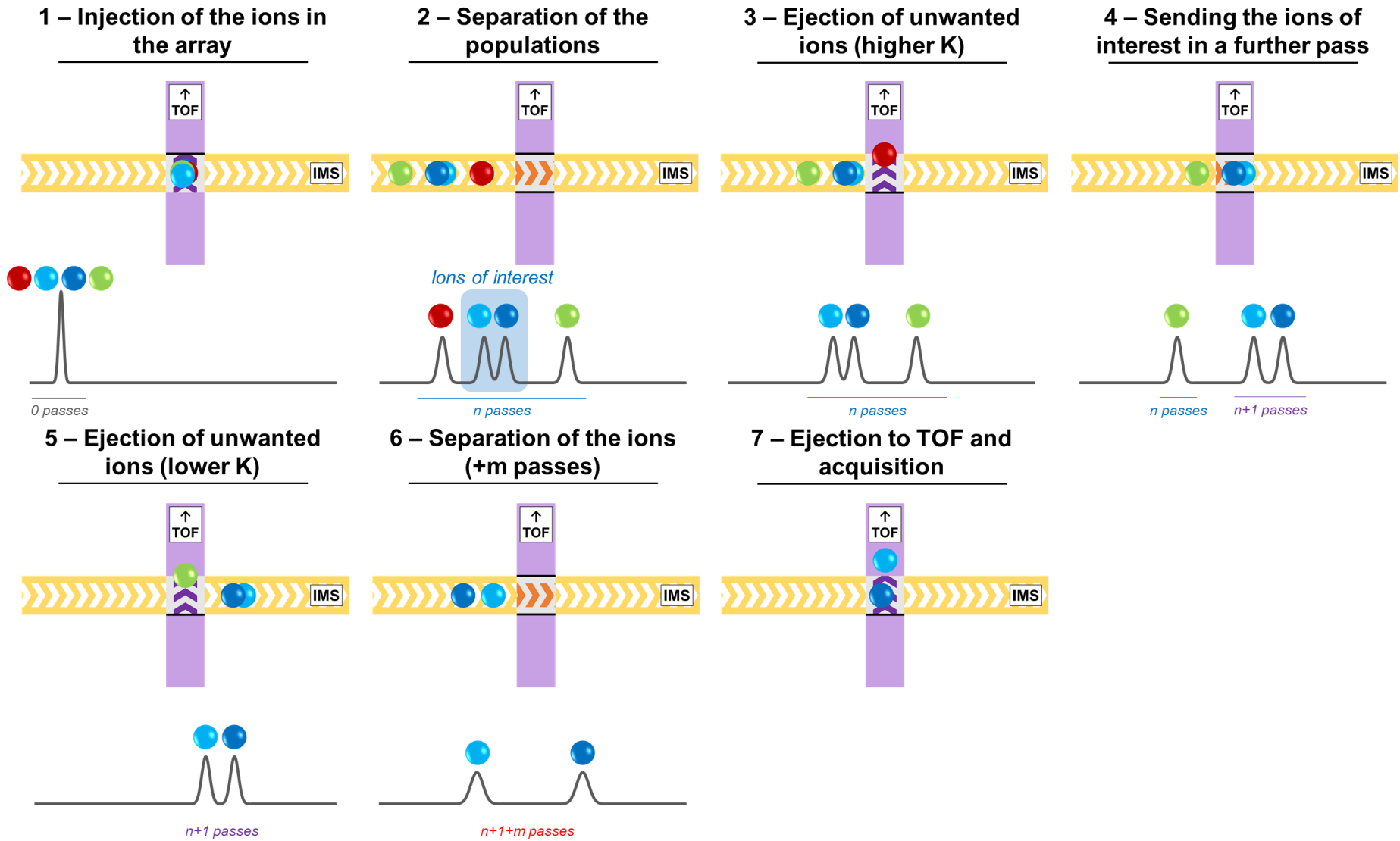


Figure 41. Sequence of events used for head-and-tail IMS^n experiments, illustrated for a single stage of selection

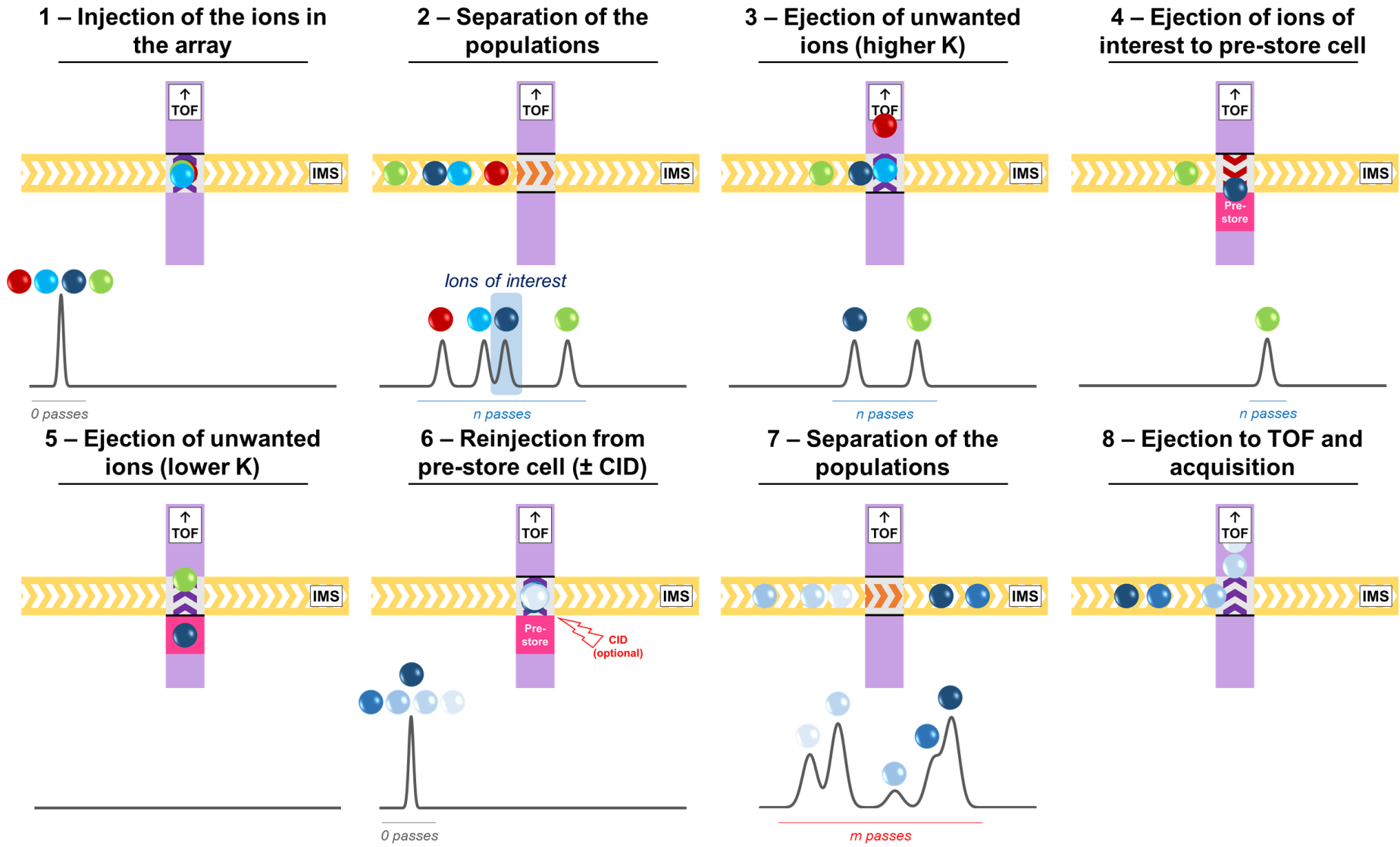


Figure 42. Sequence of events used for slicing IMS^n experiments, illustrated for a single stage of selection and fragmentation

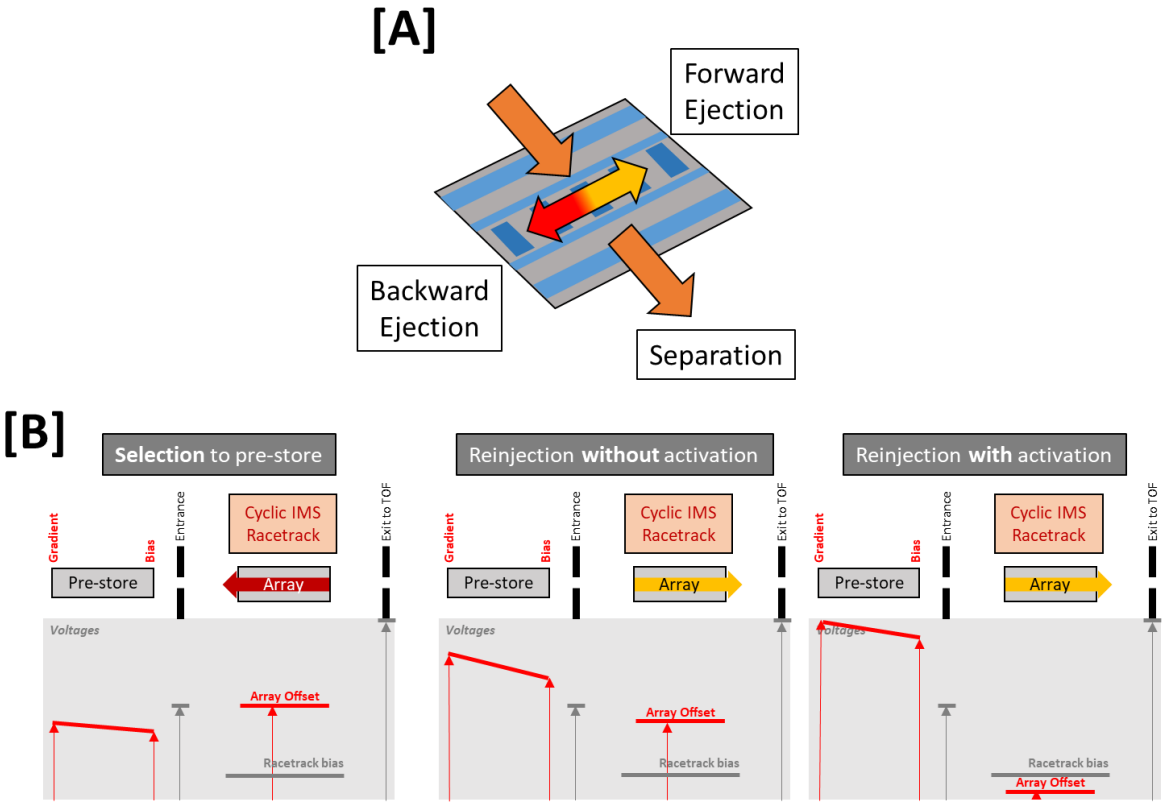


Figure 43. A. Schematic of the multifunction array region used to set the direction of the ion path. **B.** Potential energy diagrams for the selection to pre-store and reinjection stages, without and with activation upon reinjection. Color code for the direction of the array's traveling wave: forward (yellow), sideways (orange), backward (red).

Part II – Can Cyclic IMS retain and resolve all types of isomerism in oligosaccharides?

During my PhD work, the first main question that we attempted to answer was: ***Can we achieve a complete structural determination of complex oligosaccharides using cIMS?*** (with the implicit assumption that IMS is combined with MS).

To address this question, we have defined an intermediate question which is a prerequisite to the first one: ***Is cIMS—and the different experiments that can be associated with it, such as MS/MS—capable of retaining and resolving all types of isomerism that can be found in oligosaccharides?*** The different kinds of isomerisms have been detailed in the Introduction section and are summarized in **Figure 44**.

The second Part of this manuscript therefore focuses on establishing that cIMS is able to unequivocally discriminate between the different types of isomerisms.

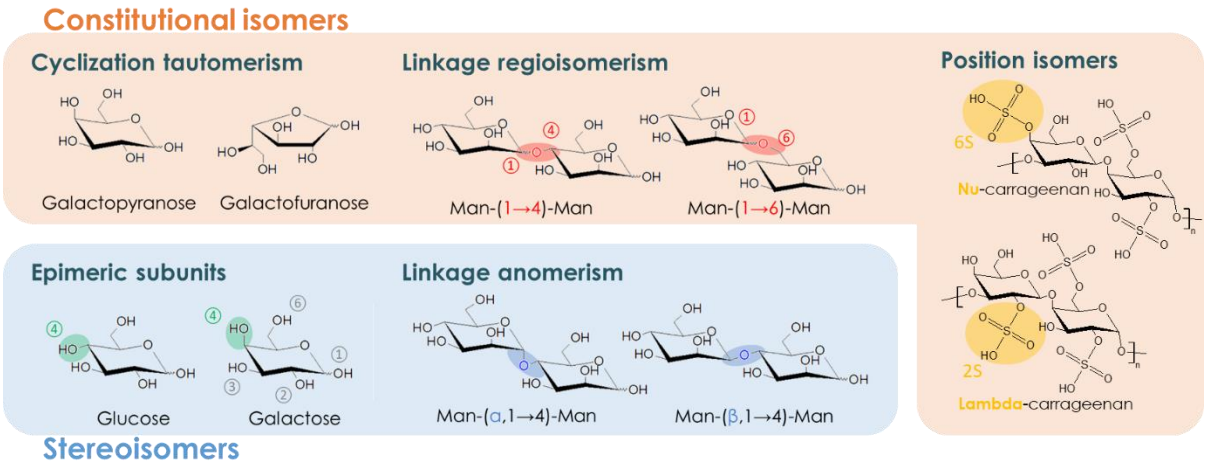


Figure 44. Overview of the isomerisms that can be found in carbohydrates

A previous study of our group has already demonstrated that the position isomerism of functional groups could be solved by post-IMS fragmentation approaches.[103]

My work consisted in going more in depth by exploring the resolution of constitutional isomerisms (the linkage regioisomerism and the furanose/pyranose tautomerism), as well as of stereoisomerisms (the linkage anomerism and the distinction of epimeric subunits).

Part II is divided into four Chapters that describe the results I obtained on these different analytical challenges. The carbohydrates studied for each isomerism originated from different studies, with different scientific backgrounds that I will briefly describe. A final Chapter provides a synthesis and conclusion of all these studies.

Chapter 4. **Distinguishing linkage regioisomers and furanose/pyranose tautomers in galactomannan disaccharides**

Chapter 5. **Investigating the behavior of furanosides upon CID**

Chapter 6. **Anomeric memory of the glycosidic bond in IMS^n experiments**

Chapter 7. **Separation of underivatized monosaccharides with cIMS**

Synthesis and Conclusions of Part II

Chapter 4 – Distinguishing linkage regioisomers and furanose/pyranose tautomers in galactomannan disaccharides

A. Context

In Nature, hexoses and their derivatives most often adopt a pyranose configuration because it is more thermodynamically stable than the furanose form.[138] Interestingly, there are some exceptions, notably with the occurrence of galactofuranose(Gal*f*)-containing polysaccharides. From a biological standpoint, it would be interesting to understand why some organisms (including fungi, bacteria, mycobacteria) biosynthesize hexoses in the form of five-membered rings. In addition, there is increasing interest in finding new applications for carbohydrates as discussed in **Chapter 1**. Gal*f*-containing carbohydrates are no exception to this general trend, although knowledge about Gal*f*-containing carbohydrates is limited and few related applications have been reported to date. For instance, Gal*f* has been reported across the tree of Life in many harmless species/genera such as *Penicillium*, *Agelas* (sponge) or *Rocella* (lichen) but also in highly infectious pathogens like *Mycobacterium tuberculosis, Trypanosoma cruzi* or *Shigella dysenteriae*.[139] For such pathogenic organisms, Gal*f*-containing components (including the biosynthetic pathways and involved enzymes) have even been targeted to develop new therapeutic strategies.[140] A major obstacle lies in the ability to accurately screen and identify such structures in biological matrices.

In an attempt to develop new methods to identify Gal*f*-containing oligosaccharides, a project was built focusing on MS coupled to gas-phase analysis techniques—IMS and IRMPD spectroscopy. This ANR-funded project, called ALGAIMS[m] for Analysis of Lichen Gal*f*-containing polysaccharides using IRMPD and ion-mobility MS, involves our group and partners from the universities of Rennes 1[n] (in charge of organic synthesis of standards, as well as extraction and hydrolysis of polysaccharides), and Lyon 1[o] (in charge of IRMPD spectroscopy, as well as theoretical calculations).

[m] Details about the ALGAIMS project are available from the website of the French National Research Agency (https://anr.fr/Projet-ANR-18-CE29-0006) and the website of the project (https://algaims-35.webself.net)

[n] Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR – UMR 6226, F-35000 Rennes, France, *and* Univ Rennes, CNRS, ISCR (Institut des Sciences Chimiques de Rennes) - UMR 6226, F-35000 Rennes, France

[o] Univ Lyon, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Villeurbanne, France

The first stage of the project consisted in the construction of a library of reference Gal*f*-containing disaccharides—since reducing Gal*f* is unstable—to acquire reference IM-MS and IRMPD spectra. This library is intended to be used in combination with fragmentation approaches to search for the corresponding Gal*f*-Man*p* motifs in larger oligosaccharides. The ability to identify Gal*f*-containing oligosaccharides would then be useful to characterize the related polysaccharides through a bottom-up approach, in lichen or other organisms.

B. Objectives

The library of Galf-containing disaccharides includes all possible regioisomers of Gal*f*-Man*p* disaccharides and their Gal*p* counterparts (**Figure 45**). From an analytical point of view and in line with the objectives of my PhD work, my goal was to establish, from this library of compounds, whether IM-MS with high resolving power was able to discriminate between all the cases of regioisomerism and furanose/pyranose tautomerism. A parallel study was performed by other partners using IRMPD. **Publication 1** describes the chemical synthesis of the library and the results obtained by both methods.

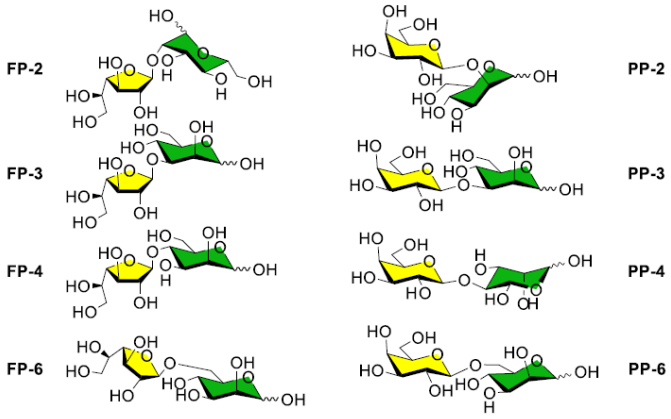


Figure 45. Structures of the Gal-Man disaccharides in the reference library of IM-MS and IRMPD spectra

C. Summary of the results

We recorded the ion mobility after four passes around the cyclic mobility cell, *i.e.*, approximately a path of four meters. The resolving power of ion mobility is proportional to the path length of the ions, which, for four passes, yields a theoretical CCS/δCCS resolving power of about 200, meaning that cyclic IM-MS can resolve conformers that differ by less than about 1 Å^2 in the case of the disaccharides studied here (*i.e.*, having CCSs of around 170 Å^2). As shown in the literature, this level of IM resolution allows the gas-phase separation of isomers and conformers prior to their mass measurement.[103,117,141]

After mass selection of each disaccharide as $[M + Na]^+$ ions at *m/z* 365.1 with the quadrupole filter, we recorded the set of gas-phase conformations translated into CCS distributions (CCSDs). The CCS values were evaluated for all observed peaks. As shown in **Figure 46** for the eight FP and PP disaccharides, a single disaccharide can give a single conformer (e.g., FP-2) or multiple ones with different CCSs, intensities, and peak shapes (e.g., FP-3). These features of the CCSD act as a fingerprint of each disaccharide.

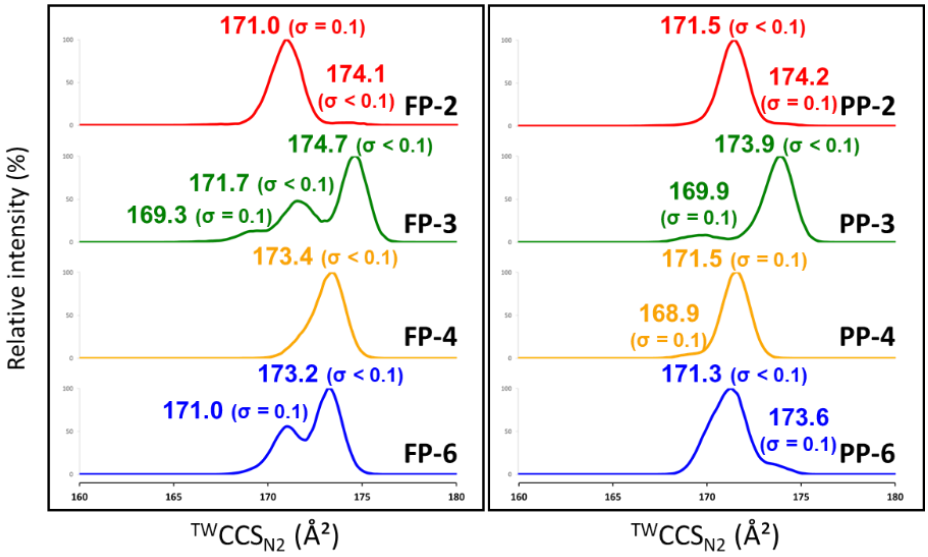


Figure 46. Extracted monoisotopic ion mobility spectra for the eight synthesized products as $[M + Na]^+$ (*m/z* 365.1). Standard deviations are given for three replicates. For both parts, the left panel corresponds to the galactofuranosyl-containing disaccharides and right panel to the pyranosyl-containing ones. Color for the regiochemistry of the glycosidic bond: Gal-(1→2)-Man (red), Gal-(1→3)-Man (green), Gal-(1→4)-Man (orange), and Gal-(1→6)-Man (blue).

Figure 46 can be read "horizontally" to differentiate isomers differing by the Gal*f* or Gal*p* residue. Here, two kinds of signatures could be observed: (i) the isomers that each display a single peak of different CCS, like FP-4 and PP-4, or (ii) the isomers that share CCSs but differ through the global distribution of the peaks, like FP-6 and PP-6. Notably, the FP type samples appeared to produce more complex ATD patterns than PP-type ones. This behavior might be related to the greater flexibility of the furanose moiety,[138,142] resulting in different conformations. The reading can also be made "vertically", this time to differentiate isomers differing in their regiochemistry.

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 and ion transmission through IM drift cells (and also because Na is generally the main alkali metal found in the samples). Alkali metal cations also favor the separation of conformers compared to protonated adducts. However, although the $[M + Na]^+$ adducts already provided distinctive signatures in most cases, some ambiguities remained regarding the samples FP-2 and PP-2 for instance.

They had similar profiles with a CCS difference below 1 Å^2. Moreover, regarding the regiochemistry, samples PP-2 and PP-4 have the same CCS and only differed through additional minor conformations (168.9 Å^2 for PP-4). The samples were thus also analyzed as lithium adducts. As shown in **Figure 47**, we achieved a better distinction in the IM-MS profiles for FP-2 and PP-2, both in terms of the CCS values of the peaks and in terms of the number and intensities of the conformers. The measurement of the IM-MS of the $[M + Li]^+$ adduct also yielded a completely different pattern for PP-2 compared to PP-6. Although samples PP-4 and PP-6 have strikingly similar CCSDs when analyzed as $[M + Li]^+$ adducts, the combination of the $[M + Na]^+$ and $[M + Li]^+$ fingerprints is unambiguous and allows identification of both the galactosyl configuration and the regiochemistry of the glycosidic bond.

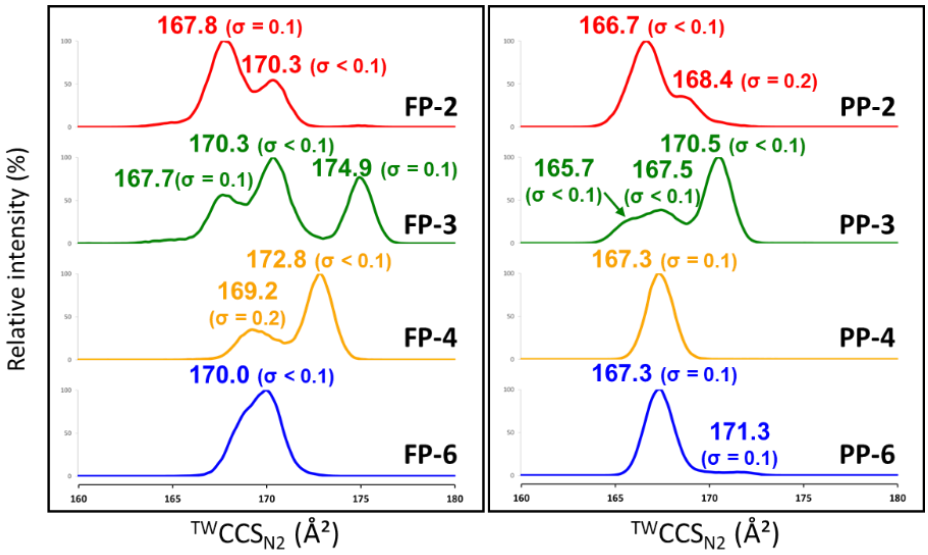


Figure 47. Extracted monoisotopic ion mobility spectra for the eight synthesized products as $[M + Li]^+$ (*m/z* 349.1). Standard deviations are given for three replicates. For both parts, the left panel corresponds to the galactofuranosyl-containing disaccharides and right panel to the pyranosyl-containing ones. Color for the regiochemistry of the glycosidic bond: Gal-(1→2)-Man (red), Gal-(1→3)-Man (green), Gal-(1→4)-Man (orange), and Gal-(1→6)-Man (blue).

D. Conclusions

In this study, we have shown that cIMS is able to discriminate between all regioisomers and tautomers in the library of eight Gal-Man disaccharides, provided that two adduct species (*i.e.*, $[M + Na]^+$ and $[M + Li]^+$) are analyzed.

Furthermore, we have also notably observed more complex IM patterns for furanosides compared to pyranosides, which raises the question of the gas-phase stability of such oligosaccharides. Indeed, the preservation of the Gal*f*-Man*p* motif upon fragmentation is a major concern—the identification of Gal-Man fragments coming from larger oligosaccharides with a library search may be compromised by, e.g., gas-phase rearrangements.

E. Scientific Communication

These results have been published in the following article presented as **Publication 1**. The supporting data for this publication are provided as **Appendix 3**.

Peer-reviewed article[p]

Favreau, B.; Yeni, O.[#]; Ollivier, S.[#]; Boustie, J.; Le Dévéhat, F.; Guégan, J.-P.; Fanuel, M.; Rogniaux, H.; Brédy, R.; Compagnon, I.; Ropartz, D.; Legentil, L.; Ferrières, V. Synthesis of an Exhaustive Library of Naturally Occurring Gal*f*-Man*p* and Gal*p*-Man*p* Disaccharides. Toward Fingerprinting According to Ring Size by Advanced Mass Spectrometry-Based IM-MS and IRMPD. *J. Org. Chem.* **2021**, 86, 6390–6405. https://doi.org/10.1021/acs.joc.1c00250.

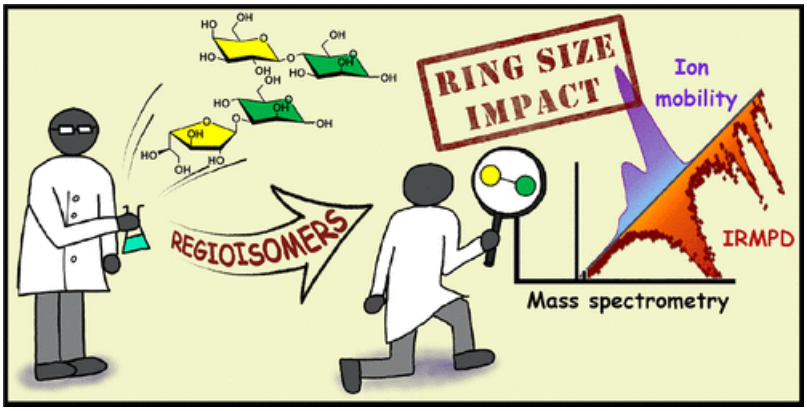
[p] # : Equal contribution

Publication 1 – Synthesis of an Exhaustive Library of Naturally Occurring Gal*f*-Man*p* and Gal*p*-Man*p* Disaccharides. Toward Fingerprinting According to Ring Size by Advanced Mass Spectrometry-Based IM-MS and IRMPD

Bénédicte Favreau, Oznur Yeni,[#] Simon Ollivier,[#] Joël Boustie, Françoise Le Dévéhat, Jean-Paul Guégan, Mathieu Fanuel, Hélène Rogniaux, Richard Brédy, Isabelle Compagnon, David Ropartz, Laurent Legentil, and Vincent Ferrières*

[#] Equal contribution

* Corresponding author



Published in: The Journal of Organic Chemistry, 2021, 86, 6390-6405

JOC The Journal of Organic Chemistry

pubs.acs.org/joc Article

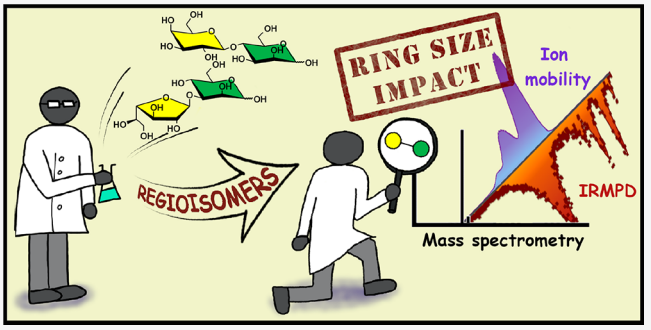
Synthesis of an Exhaustive Library of Naturally Occurring Gal*f*-Man*p* and Gal*p*-Man*p* Disaccharides. Toward Fingerprinting According to Ring Size by Advanced Mass Spectrometry-Based IM-MS and IRMPD

Bénédicte Favreau, Oznur Yeni,[§] Simon Ollivier,[§] Joël Boustie, Françoise Le Dévéhat, Jean-Paul Guégan, Mathieu Fanuel, Hélène Rogniaux, Richard Brédy, Isabelle Compagnon, David Ropartz, Laurent Legentil, and Vincent Ferrières*

Cite This: *J. Org. Chem.* 2021, 86, 6390–6405 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Nature offers a huge diversity of glycosidic derivatives. Among 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 infrared 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 Gal*f*-Man*p* disaccharides as well as the four Gal*p*-Man*p* counterparts were first synthesized according to a highly convergent approach, and IM-MS and IRMPD-MS data were second 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, and 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, and maturation of proteins.[1] They can be also overexpressed by cancer cells[2] 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 being developed.[4] Very recently, low-temperature scanning tunneling microscopy allowed direct imaging of different conformers of some oligosaccharides with a subnanometer resolution.[5] On the contrary, neutron crystallography is able to study H-bond networks in protein binding sites.[6,7] Nevertheless, structural studies are mostly based on NMR[8] and mass spectrometry (MS) techniques[9] brilliantly complemented by modeling[10,11] and machine learning approaches.[12−14] Despite these advances, no single analytical technique is able to solve all of 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*-acetylgalactosamine,[18,19] and fucose[15] 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 thermodynami-

Received: February 1, 2021
Published: April 20, 2021

Downloaded via INRAE on May 7, 2021 at 08:23:23 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ACS Publications © 2021 American Chemical Society

cally 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 (Gal*f*)-mannopyranose (Man*p*) sequence and thus the four possible regioisomers (Figure 1). The disaccharides with a (1→2)-, (1→3)-, or (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-Gal*f*-(1→4)-D-Man*p* sequence was exclusively found in lichens.[20−24] On another side, the corresponding pyranosidic Gal*p*-Man*p* 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 nondigestable epilactose with the (1→4)-linkage,[27,28] and *Burkholderia*,[29] KLH protein,[30] *Salmonella*[31] for the *β*-D-Gal*p*-(1→6)-D-Man*p*. This resulting orthogonality in Nature strengthens the need for consolidating analytical data to differentiate Gal*f* and Gal*p* residues.[32]

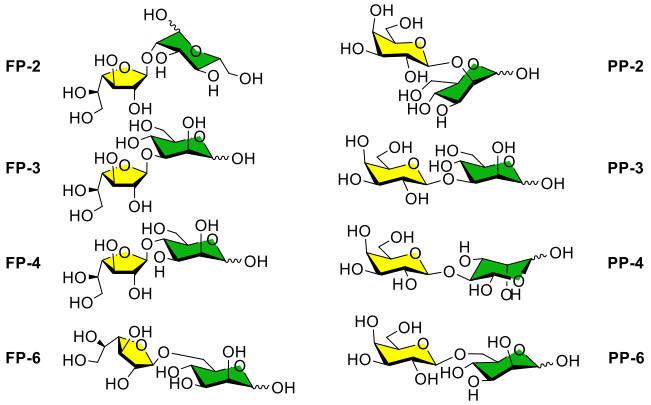


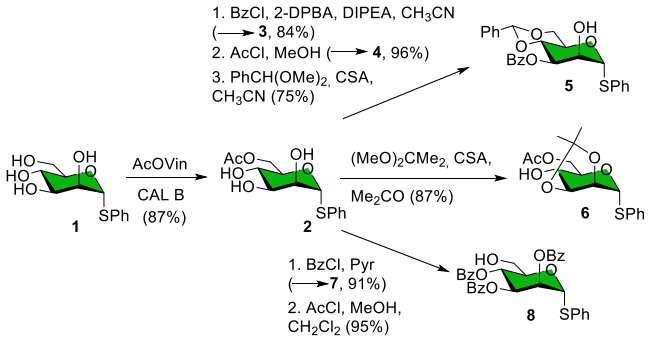
Figure 1. Defined disaccharides for IM-MS and IRMPD fingerprinting

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)[33,34] and infrared multiple photon dissociation (IRMPD)[35] 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,[36] and can be used to determine specific structures.[37,38] IRMPD spectroscopy has been widely used, in combination with quantum chemistry, to elucidate the conformation of biomolecular ions in the gas phase.[39] Previous results demonstrate that IRMPD spectroscopy has sufficient structural resolution to yield distinctive signatures and to unambiguously distinguish Gal*p*NAc residues from Gal*f*NAc ones in the case of monosaccharides.[40] 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. We expect that these signatures will be useful for the identification of the corresponding Gal-Man patterns in larger polysaccharides. Upon fragmentation, galactomannans will release specific disaccharidic fragments, whose fingerprints will match that of the synthetic library. The MS/MS strategy has been applied to the identification of the monosaccharide content (analyses were done on the monosaccharide fragments released upon fragmentation),[41−43] and IRMPD signatures of methyl furanoside and pyranoside derived from *N*-acetylgalactosamine have been previously reported.[40] We anticipate, however, that monosaccharidic fragments may isomerize upon fragmentation, especially the thermodynamically less stable hexofuranosyl residues, and this would adversely affect the analysis of monosaccharide entities in larger oligosaccharides. Therefore, we present the fingerprints of Gal*f*-Man*p* and Gal*p*-Man*p* disaccharides, for future analysis of oligosaccharides. The strategy offers the advantage that no additional efforts will be further required in terms of synthesis, as the eight disaccharides synthesized in this work will provide the library necessary for the structural study of larger polymers.

RESULTS AND DISCUSSION

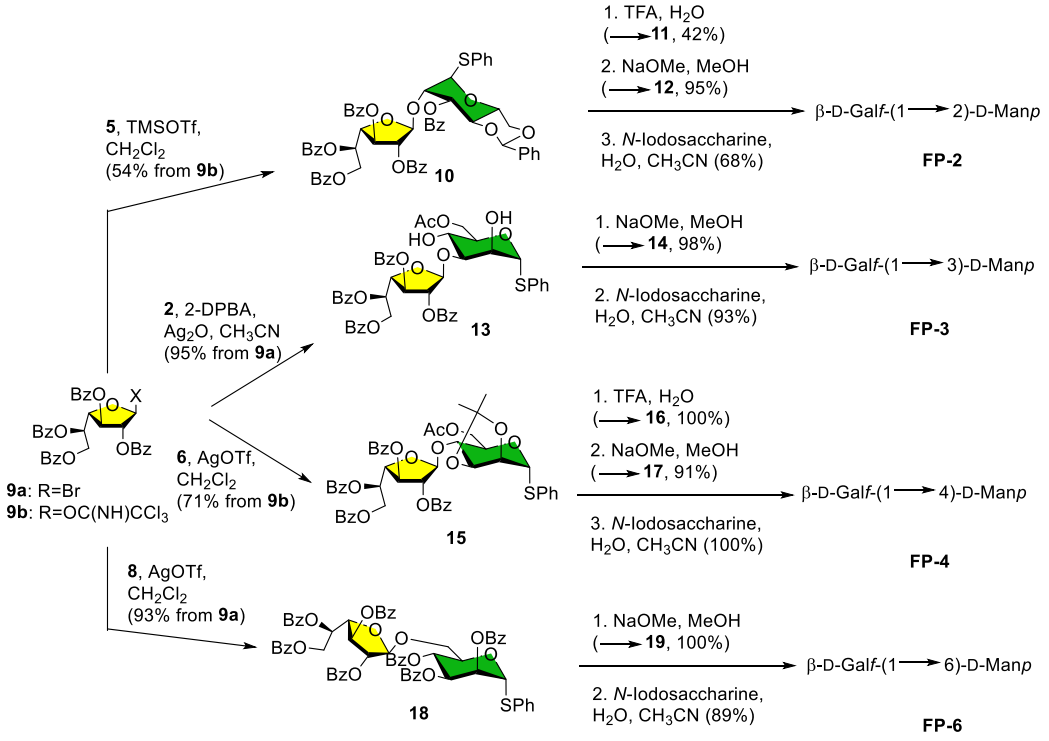
Synthesis of Gal*f*-Containing Disaccharides and of Their Gal*p* Isomers. To access the panel of the targeted unprotected and reducing disaccharides, four mannopyranosidic acceptors were first synthesized for further coupling with galactofuranosyl or galactopyranosyl donors. Considering the ease of preparation and structural modulation of thioglycosides, and the orthogonal properties toward other families of glycosyl donors, the phenyl thiomannopyranoside **1** was selected as the precursor for all acceptors (Scheme 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.[44] This biocatalyzed acylation was interestingly performed on a large 20 g scale, and this consolidates our choice toward this compound as a key intermediate in our synthesis strategy. Moreover, intermediate **2** can act as acceptor for direct O-3 glycosylation without further protections. The pioneering works of Taylor,[45] 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→3)-disaccharides] or converted into the other required acceptors for (1→2)-, (1→4)-, and (1→6)-linkages, thus limiting the number of intermediates and synthetic steps. The synthesis of acceptor **5**

Scheme 1. Synthesis of the Required Mannopyranosidic Acceptors



https://doi.org/10.1021/acs.joc.1c00250
J. Org. Chem. 2021, 86, 6390−6405

Scheme 2. Synthesis of the Gal*f*-Containing Disaccharides FP



with free OH-2 started with the specific benzoylation of **2** at position 3 in the presence of the borinate ester 2-aminoethyldiphenyl borinate (2-ADPB). The building block **3** thus obtained was then converted into the mannopyranoside **5** after selective deacetylation under acidic conditions and benzylidenation. This synthetic way was preferred to the sequence benzylidenation followed by 3-*O*-benzoylation 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 took advantage of the *cis* arrangement of the hydroxyl groups in compound **2** for obtaining the derivative **6** bearing free OH-4 by acid-catalyzed acetalization with dimethoxypropane. Finally, a two-step benzoylation−deacetylation of **2** afforded **8** with a free primary position.

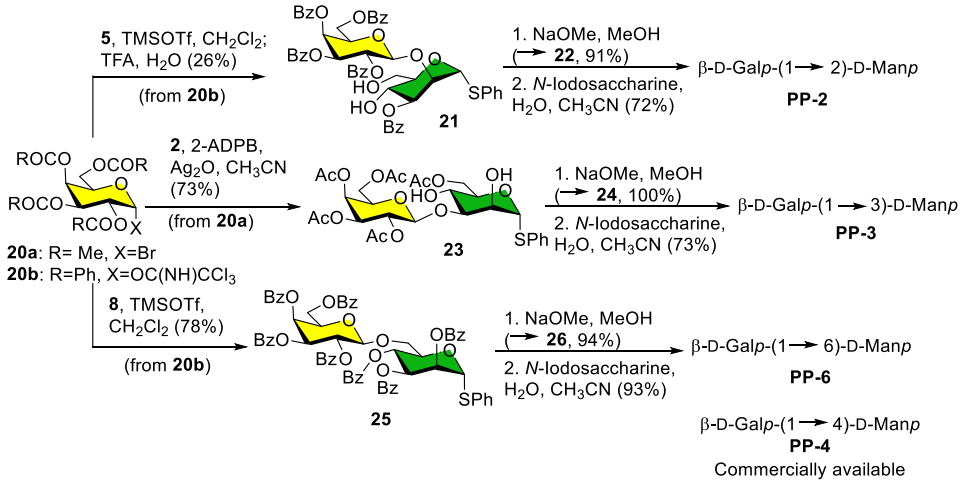
With the acceptors in hand, it was anticipated that glycosylation at positions 6 and 3 is easier than that at O-2 and O-4.[46] Moreover, it is recognized that conditions should be tuned according to the nature of the leaving group on the glycosyl donors but also on the basis of its furanose or the pyranose form. Our results did not depart from these general observations. The synthesis of disaccharides **FP-2**−**FP-6** was first considered starting with the known galactofuranosyl bromide **9a** (Scheme 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 except the O-3 linkage, for which it was shown that silver oxide was more efficient.[47] For these steps, 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 the 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 ^{13}C 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 was performed under standard conditions: acidic hydrolysis for the acetals, Zemplen transesterification for the esters, and hydrolysis promoted by *N*-iodosaccharine[48] for the thioglycosides. It is noteworthy that the resulting reducing disaccharides were obtained in aqueous solution generally as mixtures of anomers as revealed by ^{1}H NMR and corroborates previous data for **FP-2**[49] and **FP-3**.[50] 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 Gal*f*-Man*p* and complements well the one recently described for three Gal*f*-Man*p* derivatives but bearing an amino spacer arm at the reducing end.[51] Both anomers were unambiguously characterized thanks to the NMR spectra, with the 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 the Supporting Information).

This strategy was next extended to the preparation of the Gal*p*-Man*p* isomers (Scheme 3), except for the disaccharide

https://doi.org/10.1021/acs.joc.1c00250
J. Org. Chem. 2021, 86, 6390−6405

Scheme 3. Synthesis of the Gal*p*-Containing Disaccharides PP



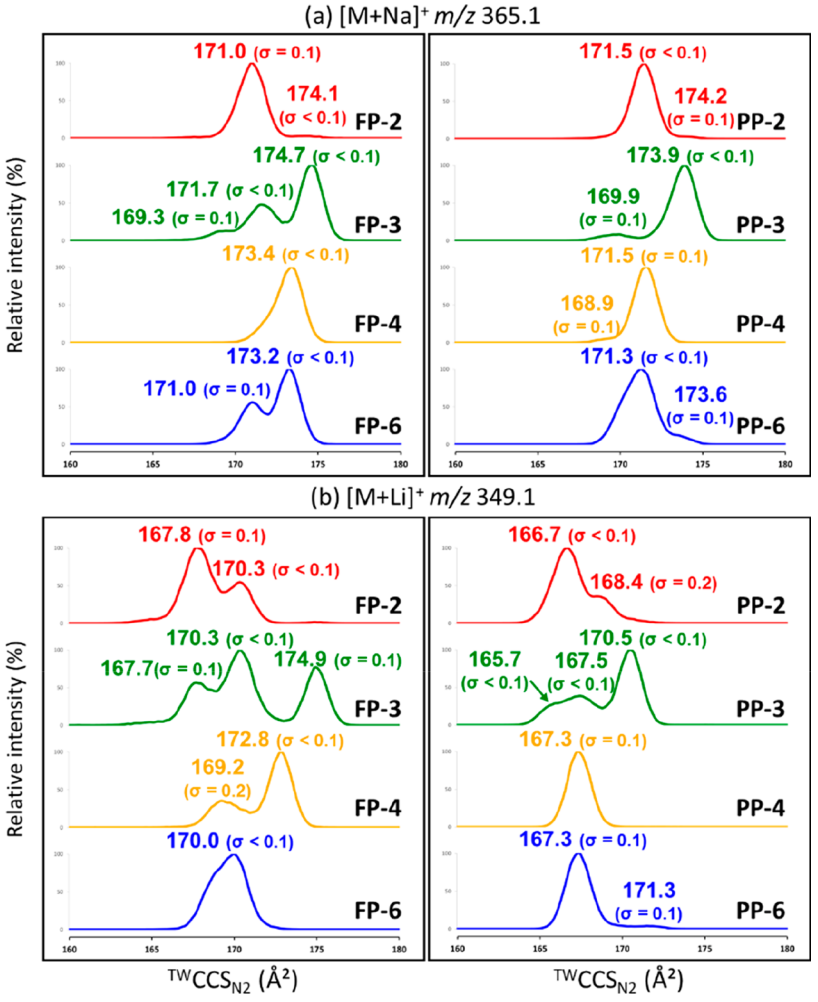


Figure 2. Extracted monoisotopic ion mobility spectra for the eight synthesized products as (a) $[M + Na]^+$ (m/z 365.1) and (b) as $[M + Li]^+$ (m/z 349.1). Standard deviations are given for three replicates. For both parts, the left panel corresponds to the galactofuranosyl-containing disaccharides and right panel to the pyranosyl-containaing ones. Color for the regiochemistry of the glycosidic bond: Gal-(1→2)-Man (red), Gal-(1→3)-Man (green), Gal-(1→4)-Man (orange), and Gal-(1→6)-Man (blue).

PP-4, which is commercially available. 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. Unfortunately, a poor glycosylation yield was obtained using the same donor **20a** and the acceptor **5**. The targeted intermediate **21** was nevertheless prepared from the perbenzoylated galactopyranosyl trichloroacetimidate **20b**[52] and the same acceptor and isolated after removal of the

https://doi.org/10.1021/acs.joc.1c00250
J. Org. Chem. 2021, 86, 6390−6405

benzylidene ring in 26% yield. Further deprotection steps gave the desired product **PP-2**[53] in satisfactory yield. Finally, and surprisingly, the formation of the (1→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 diisopropylethyl amine as a base (DIPEA) to avoid the acid-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 sodium methoxide followed by hydrolysis of the thiomannoside gave the wanted disaccharide **PP-6**.[54]

Mass Spectrometry Analysis. As a preliminary mass spectrometry analysis, the spectra of all disaccharides were recorded using traditional MS/MS approaches. It was evidenced that the Gal*f* moiety cannot be differentiated from its Gal*p* isomer, as the products do not generate specific fragments for a defined Gal*f*-Man*p*/Gal*p*-Man*p* pair (see Figure S1). We thus turned to gas phase analysis techniques that are sensitive to both the spatial arrangement and the regiochemistry of the ions: ion mobility (IM-MS) and IRMPD spectroscopy.

Ion Mobility–Mass Spectrometry. We first analyzed the synthesized disaccharides using positive electrospray ionization (ESI) time-of-flight mass spectrometry, hyphenated with a high-resolution cyclic ion mobility cell (Waters Select Series Cyclic IMS, Wilmslow, UK).[55] Arrival time distributions (ATDs) were then calibrated to give collision cross sections (CCSs), which represent the effective surface of the ions that interacts with the buffer gas during ion mobility experiments. We recorded the ion mobility of the $[M + Na]^+$ cations or $[M + Li]^+$ cations after four passes around the cyclic mobility cell, i.e., approximately a path of four meters. The resolving power of ion mobility is proportional to the path length of the ions, which, for four passes, yields a theoretical CCS/δCCS resolution of about 200, meaning that cyclic IM-MS can resolve conformers that differ by less than about 1 Å^2 in the case of the disaccharides studied here (i.e., having CCSs of around 170 Å^2). As shown in the literature, this level of IM resolution allows the gas-phase separation of isomers and conformers prior to their mass measurement.[56−58] After mass selection of each disaccharide using a quadrupole filter, we recorded the set of gas-phase conformations for each ion, translated into CCS distributions (CCSDs). As shown in Figure 2 for the eight **FP** and **PP** disaccharides, a single disaccharide can give a single conformer or multiple ones with different CCSs, intensities, and peak shapes. These features of the CCSD act as a fingerprint of each disaccharide. Notably, fitting gaussians under the curve can be used in order to estimate the number of conformers in CCSDs (full, deconvolved CCSDs are given in Figure S2).

Second, we measured the CCSDs of the eight synthesized products as $[M + Na]^+$ ions at m/z 365.1 (Figure 2a). 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 and ion transmission through IM drift cells, more than ammonium adducts. Alkali metal cations also favor the separation of conformers compared to protonated adducts. Figure 2.a can first be read "horizontally" to differentiate isomers differing by the Gal*f* or Gal*p* residue. Here, considering the maximum standard deviations of CCSs observed during replicates (differences are significant if above 0.5 Å^2), two kinds of signatures could be observed: (i) the isomers each display a single peak of different CCS or (ii) the isomers have shared CCSs but differ through the global distribution of the peaks. For instance, in (i), the case of **FP-4** and **PP-4**, their pattern is rather similar with one major peak (note, however, a slight broadening of the peak for **FP-4**, which suggests the presence of two conformers) but these single peaks have different CCSs (173.4 Å^2 for **FP-4** and 171.5 Å^2 for **PP-4**). In (ii), the case of **FP-3** and **PP-3**, a difference of less than 1 Å^2 is measured on the most extended conformation (174.7 Å^2 versus 173.9 Å^2). However, the overall pattern is different, with three conformations separated for **FP-3** compared to two conformations for **PP-3**. Notably, the **FP**-type samples appeared to produce more complex ATD patterns than **PP**-type ones. This behavior might be related to the greater flexibility of the furanose moiety,[59,60] resulting in different conformations.

The same reasoning can be made "vertically" to differentiate isomers differing in their regiochemistry. As an example, the **FP-2** sample had an intense signal at 171.0 Å^2 with another peak of lower intensity at 174.1 Å^2. In contrast, the **FP-3** sample exhibited a triplet of CCSDs, respectively, at 169.3, 171.7, and 174.7 Å^2. Although less pronounced differences in CCSD patterns were observed for the **PP**-type samples, most signatures remained distinctive. For example, the **PP-3** showed a doublet of CCSDs at 169.9 (minor species) and 173.9 Å^2, while the **PP-4** displayed a doublet at 168.9 (minor species) and 171.5 Å^2.

However, although the $[M + Na]^+$ adducts already provided distinctive signatures in most cases, some ambiguities remained regarding the samples **FP-2** and **PP-2**. They had similar profiles with a CCS difference below 1 Å^2. Moreover, regarding the regiochemistry, samples **PP-2** and **PP-4** have the same CCS and only differed through minor conformations. The samples were thus also analyzed as lithium adducts. As shown in Figure 2b, we achieved a better distinction in the IM-MS profiles for **FP-2** and **PP-2**, both in terms of the CCS values of the peaks and in terms of the number and intensities of the conformers (both harbor two peaks at, respectively, 167.8 and 170.3 Å^2 and 166.7 and 168.4 Å^2). The measurement of the IM-MS of the $[M + Li]^+$ adduct also yielded a completely different pattern for **PP-2** compared to **PP-6**. Although samples **PP-4** and **PP-6** have strikingly similar CCSDs when analyzed as $[M + Li]^+$ adducts (a main peak at 167.3 Å^2, with similar peak shapes), the combination of the $[M + Na]^+$ and $[M+Li]^+$ fingerprints is unambiguous and allows identification of both the galactosyl configuration and the regiochemistry of the glycosidic bond.

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).[43] 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 form of the galactosyl entity as a five- or a six-membered ring.

https://doi.org/10.1021/acs.joc.1c00250
J. Org. Chem. 2021, 86, 6390−6405

As mentioned in the previous section, Li^+ and Na^+ adducts are often used to increase ion signal and fragmentation in mass spectrometry of carbohydrates.[61] In the context of IRMPD spectroscopy, different ion adducts will lead to different spectroscopic fingerprints.[39,62] As previously reported, ion adducts can also be chosen to enhance the spectroscopic differentiation of carbohydrate isomers.[62] 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 setup previously described.[63] In short, the

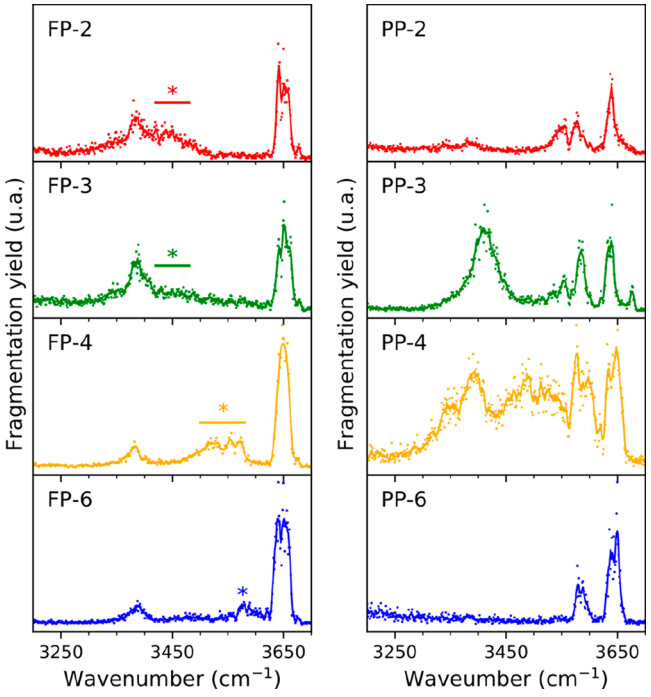


Figure 3. Left pannel: IRMPD spectra of galactofuranosyl-containing disaccharides. Right panel: galactopyranosyl-containing disaccharides. Color code for the regiochemistry of the glycosidic bond: Gal-(1→2)-Man (red), Gal-(1→3)-Man (green), Gal-(1→4)-Man (orange), Gal-(1→6)-Man (blue).

standards were prepared in water/methanol (50/50), and 0.1% of ammonium sulfate was added to promote the formation of ionic complexes $M{\cdot}NH_4^+$, 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).

First, all spectra display a common feature: an intense band around 3650 cm^{-1}. In addition, another intense feature is observed at 3575 cm^{-1} only for the pyranose form, which makes a straightforward distinction between pyranose and furanose fingerprints. The 3200−3550 cm^{-1} range is unique for each pyranose: **PP-6** hardly presents any vibrational activity besides the pyranose diagnostic doublet. **PP-2** presents a distinctive feature at 3547 cm^{-1} and a weak NH band at 3380 cm^{-1}. **PP-3** and **PP-4** show an intense and broad NH feature in the same area. Additionally, **PP-4** shows a broad and intense feature around 3500 cm^{-1}.

Spectra are less distinctive of the regiochemistry of furanosyl-containing disaccharides, but some differences can be noticed. Besides the main band at 3650 cm^{-1} (furanose diagnostic), all four spectra show a NH band in a close range (from 3380 to 3390 cm^{-1}). In addition, **FP-6** shows a weak band at 3580 cm^{-1} (marked with a blue star); **FP-4** features a group of partly resolved bands in the 3500−3570 cm^{-1} region (marked with an orange star); both **FP-2** and **FP-3** show a broadening of their NH band toward the high wavenumber, ranging to 3490 cm^{-1} (marked with red and green stars). **FP-2** and **FP-3** only differ by the relative intensity of this shoulder (**FP-2**: high, **FP-3**: low).

Overall, the IRMPD spectra bear a strong diagnostic of the ring size: either one dominant band at 3650 cm^{-1} for furanose or a doublet of bands at 3650/3575 cm^{-1} for pyranose. In the case of Gal*p*-containing disaccharides, the IR fingerprint is also diagnostic of the regiochemistry because all **PP** spectra are distinguishable. For Gal*f*-containing disaccharides, the IRMPD approach proves less distinguising of the regiochemistry: **FP-2** and **FP-3** have close fingerprints.

CONCLUSION

The results presented in this work, combining different fields in glycosciences, allowed us to highlight the impact of the ring size of the nonreducing galactosyl residue for each possible connectivity in the Gal*f*-Man*p* 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 libraries of reference fingerprints for further analysis of unknown disaccharides and that they constitute the fundamentals of bringing out Gal*f* for pattern recognition in longer oligosaccharides. Therefore, we expect that this study paves the way to the screening of galactofuranose in natural polysaccharides.

EXPERIMENTAL SECTION

Synthesis. *General Experimental Details.* All reagents were purchased from commercial sources and were used without further purification unless noted. 2,3,5,6-Benzoylgalactofuranosyl bromide **9a**[64] and trichloroacetimidate **9b**[65] as well as 2,3,4,6-benzoylgalactopyranosyl trichloroacetimidate **20b**[52] were synthesized according to literature procedures. Unless otherwise stated, all reactions were monitored by TLC on silica gel 60 F_{254}. TLC spots were detected under 254 nm UV light or by staining with cerium ammonium molybdate solution. For reactions that proceeded by heating, a metallic heating mantle was used. Column chromatography was performed on silica gel (50 μm). High-resolution masses were recorded in positive mode using direct electrospray ionization on a Waters Q-Tof 2 spectrometer.

NMR spectra were recorded on a Bruker Avance III 400 spectrometer operating at 400.13 MHz for ^{1}H, equipped with a BBFO probe with a Z-gradient coil and a GREAT 1/10 gradient unit or on a Bruker Avance III operating at 500.13 MHz for ^{1}H equipped with ^{1}H and ^{13}C cryo-probe. All experiments were carried out at 25 °C, and references for chemical shifts are external. Coupling constants J were calculated in hertz (Hz). Proton and carbon NMR peaks were unambiguously assigned by COSY (double quantum filtered with

https://doi.org/10.1021/acs.joc.1c00250
J. Org. Chem. 2021, 86, 6390−6405

gradient pulse for selection), HSQC (gradient echo-anti echo selection and shape pulse), and HMBC (echo-anti echo gradient selection, magnitude mode) correlation experiments as well as TOCSY. Mutiplicity are annotated as s (singlet), d (doublet), t (triplet), m (multiplet), or o (overlapping signals).

The zg30 Bruker pulse program was used for 1D ^{1}H NMR, with a TD of 64k, a relaxation delay d1 = 2s and 32 scans or more. The spectrum width was set to 18 ppm. 2D COSY experiments were acquired using the cosygpdqf pulse program. Matrices consisting of 256−512 (t1) × 2048 (t2) complex data points were recorded. An AQ time of 0.25−0.5 s was used. Processing was performed with a QSINE function in both dimensions (SSB = 0). 2D TOCSY experiments were acquired using the dipsi2sisp pulse program with a mixing time of 150 ms in major cases. Matrices consisting of 256−400 (t1) × 4k−8k (t2) complex data points were recorded; 8−16 scans are carried out per t1 increment with 1.5 s recovery delay (d1) and AQ time of 0.3 to 0.7 s. Processing was performed with a QSINE function in both dimensions (SSB = 2). ^{13}C NMR spectra were recorded at 100.61 MHz or at 125.13 MHz on cryo-probe. Several sequences as jmod, dept135, or zgpg30 were used with n*1024 scans depending on the concentration of the sample. TD was set to 64k and a relaxation delay of 2 s for a spectral width of 220 ppm was used.

2D HSQC (^{1}H−^{13}C) experiments were acquired using the hsqcetgpsisp2.2 pulse program for high sensitivity with an AQ = 0.25 to 0.5 s, d1 = 1.5 s, ns = 2 to 24 depending on the concentration. Generally, 256 experiences were acquired (t1). Fourier transform was performed in both dimensions with a SINE function (SSB = 3). 2D HMBC (^{1}H−^{13}C) experiments were performed with either the hmbcplrndqf or the impact-hmbc pulse program. D1 was set to 0.3 to 1.5 s, delay d6 for the evolution of long-range coupling was set to 65 ms, NS was depending on the concentration and the number of series was set at 256 to 512(t1). All of the data were processed with a SINE function in both dimensions (SSB = 3). To identify the ^{1}H frequency of the mixture of anomers, the 1D-Tocsy experiment (selcssfdizs.2) from the Bruker library was used with 150 ms for mixing time and 16 for TD0 when the highest selectivity was required.

Further details, including NMR data, are described in the Supporting Information.

*Synthesis of Glycosyl Acceptors. Phenyl 6-O-Acetyl-1-thio-α-D-mannopyranoside (**2**).* To a solution of phenyl 1-thio-α-D-mannopyranoside **1** (20 g, 73.5 mmol) in dry THF (600 mL) were added vinyl acetate (17.6 mL, 0.19 mol) and Novozym 435 (3 g). The reaction mixture was allowed to stire at 45 °C for 65 h. The reaction mixture was filtered and concentrated under reduced pressure. The resulting residue was purified by column chromatography on silica gel (DCM/MeOH gradient from 100:0 up to 90:10) to afford **2** (19.97 g, 87%) as a beige solid. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.50−7.44 (m, 2H, H_{arom}), 7.31−7.22 (m, 3H, H_{arom}), 5.58 (d, 1H, J = 1.3 Hz, H-1), 4.49 (dd, 1H, J = 12.6, 6.9 Hz, H-6), 4.37−4.31 (m, 2H, H-5, H6′), 4.27 (dd, 1H, J = 3.3, 1.3 Hz, H-2), 3.91 (dd, 1H, J = 9.4, 3.3 Hz, H-2), 3.78 (t, 1H, J = 9.4 Hz, H-3), 2.03 (s, 3H, CH_3). ^{13}C{1H} NMR ($CDCl_3$, 100 MHz): δ 172.2 (CO), 133.7, 131.8, 129.1, 127.7 (C_6H_5), 87.9 (C-1), 72.2, 72.2 (C-2, C-3), 71.4 (C-5), 68.2(C-4), 63.9 (C-6), 21.0 (CH_3). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{14}H_{18}O_6SNa$ 337.0716; Found 337.0718.

*Phenyl 6-O-Acetyl-3-O-benzoyl-1-thio-α-D-mannopyranoside (**3**).* To a solution of **2** (1.0 g, 3.2 mmol) in dry acetonitrile (20 mL) were added 2-aminoethyldiphenyl borinate (788 mg, 3.5 mmol), benzoyl chloride (1.2 mL, 10.0 mmol), and *N*,*N*-diisopropylethylamine (1.7 mL, 9.8 mmol) under anhydrous conditions. The reaction mixture was stirred at 40 °C for 17 h. The reaction mixture was then concentrated under reduced pressure, and the resulting residue was partitioned between DCM and aqueous saturated $NaHCO_3$ solution. The organic layer was washed with aqueous saturated $NaHCO_3$ solution (2×) and brine, dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography on silica gel (DCM/MeOH gradient from 100:0 to 90:10) to afford **3** (1.17 g, 84%) as a white solid. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 8.12−8.08 (m, 2H, H_{arom}), 7.62−7.28 (m, 8H, H_{arom}), 5.58 (d, 1H, J = 1.6 Hz, H-1), 5.34 (dd, 1H, J = 9.8, 3.1 Hz, H-3), 4.56 (dd, 1H, J = 12.1, 5.1 Hz, H-6), 4.46 (dd, 1H, J = 3.1, 1.6 Hz, H-2), 4.44 (ddd, 1H, J = 9.8, 5.1, 2.2 Hz, H-5), 4.31 (dd, 1H, J = 12.1, 2.2 Hz, H-6′), 4.11 (t, 1H, J = 9.8 Hz, H-4), 2.09 (s, 3H, CH_3). ^{13}C{1H} NMR ($CDCl_3$, 100 MHz): δ 171.7, 166.6 (CO), 133.8, 131.8, 130.0, 129.3, 128.7, 127.9 (C_6H_5), 88.0 (C-1), 75.1 (C-3), 72.1, 70.9 (C-2, C-5), 66.2 (C-4), 63.4 (C-6), 21.0 (CH_3). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{21}H_{22}O_7SNa$ 441.0978; Found 441.0982.

*Phenyl 3-O-Benzoyl-1-thio-α-D-mannopyranoside (**4**).* To a solution of **3** (1.42 g, 2.41 mmol) in dry MeOH (30 mL) was added acetyl chloride (360 μL, 5.09 mmol) under anhydrous conditions. The reaction mixture was stirred at room temperature for 17 h. The reaction was then stopped by the addition of saturated aqueous $NaHCO_3$. The aqueous layer was extracted with DCM. The combined organic layers were dried over $MgSO_4$, filtered, and concentrated under reduced pressure to afford **4** (1.22 g, 96%) as a white solid. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 8.13−8.00 (m, 2H, H_{arom}), 7.63−7.27 (m, 8H, H_{arom}), 5.55 (d, 1H, J = 1.6 Hz, H-1), 5.35 (dd, 1H, J = 9.7, 3.1 Hz, H-3), 4.44 (dd, 1H, J = 3.1, 1.6 Hz, H-2), 4.35 (t, 1H, J = 9.7 Hz, H-4), 4.26 (dt, 1H, J = 9.7, 3.3 Hz, H-5), 3.96 (dd, 1H, J = 12.1, 3.7 Hz, H-6), 3.86 (dd, 1H, J = 12.1, 2.8, H-6′). ^{13}C{1H} NMR ($CDCl_3$, 100 MHz): δ 167.9, 167.0 (CO), 134.3, 133.7, 132.0, 130.0, 129.8, 129.3, 128.8, 127.9 (C_6H_5), 88.2 (C-1), 75.5 (C-3), 73.7 (C-5), 71.0 (C-2), 65.9 (C-4), 62.0 (C-6). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{19}H_{20}O_6SNa$ 399.0873; Found 399.0875.

*Phenyl 3-O-Benzoyl-4,6-O-benzylidene-1-thio-α-D-mannopyranoside (**5**).* To a solution of **1** (620 mg, 1.65 mmol) in dry acetonitrile (10 mL) were added benzaldehyde dimethyl acetal (372 μL, 2.48 mmol) and camphorsulfonic acid (77 mg, 0.33 mmol) under anhydrous conditions. The reaction mixture was allowed to stir at 55 °C for 8 h. Then the medium was concentrated under reduced pressure and partitioned between DCM and aqueous saturated $NaHCO_3$ solution. The organic layer was washed with aqueous saturated $NaHCO_3$ solution (2×) and water. The combined organic phases were dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography on silica gel (cyclohexane/EtOAc gradient from 100:0 to 70:30) to afford **5** (573 mg, 75%) as a white solid. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 8.08 (dd, 2H, J = 8.4, 1.4 Hz, H_{arom}), 7.61−7.43 (m, 7H, H_{arom}), 7.37−7.29 (m, 6H, H_{arom}), 5.63 (s, 1H, H-7), 5.61 (d, 1H, J = 1.4 Hz, H-1), 5.60 (dd, 1H, J = 10.2, 3.2 Hz, H-3), 4.59 (dd, 1H, J = 3.2, 1.4 Hz, H-2), 4.54 (td, 1H, J = 10.2, 4.9 Hz, H-5), 4.37 (t, 1H, J = 10.2 Hz, H-4), 4.28 (dd, 1H, J = 10.2, 4.9 Hz, H-6), 3.92 (t, 1H, J = 10.2 Hz, H-6′). ^{13}C{1H} NMR ($CDCl_3$, 100 MHz): δ 165.6 (CO), 137.2, 133.6, 133.4, 132.0, 130.0, 129.7, 129.3, 129.2, 128.6, 128.4, 128.0, 126.3 (C_6H_5), 102.0 (C-7), 88.6 (C-1), 76.4 (C-4), 71.6, 71.4 (C-3, C-2), 68.7 (C-6), 65.3 (C-5). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{26}H_{24}O_6SNa$ 487.1186; Found 487.1187.

*Phenyl 6-O-Acetyl-2,3-O-isopropylydene-1-thio-α-D-mannopyranoside (**6**).* To a solution of **2** (2.0 g, 3.37 mmol) in anhydrous acetone (30 mL) were added 2,2-dimethoxypropane (1.6 mL, 12.74 mmol) and camphorsulfonic acid (147 mg, 0.637 mmol) under anhydrous conditions. The mixture was stirred at room temperature for 17 h. Then the reaction was stopped by the addition of triethylamine and concentrated under reduced pressure. The resulting residue was dissolved in DCM, and the organic layer was washed with saturated aqueous $NaHCO_3$ (3×) followed by water. The organic layers were dried over $MgSO_4$, filtered, and evaporated under reduced pressure. The resulting residue was purified by column chromatography on silica gel (cyclohexane/EtOAc gradient from 100:0 up to 80:20) to afford **6** (1.95 g, 87%) as a colorless oil. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 7.51−7.47 (m, 2H, H_{arom}), 7.34−7.27 (m, 3H, H_{arom}), 5.82 (s, 1H, H-1), 4.47−4.41 (m, 1H, H-6), 4.35 (dd, 1H, J = 5.6, 1.0 Hz, H-2), 4.22−4.15 (m, 3H, H-3, H-5, H-6′), 3.60 (dd, 1H, J = 10.2, 7.5 Hz, H-4), 2.02 (s, 3H, CH_3), 1.52 (s, 3H, CH_3), 1.37 (s, 3H, CH_3). ^{13}C{1H} NMR ($CDCl_3$, 100 MHz): δ 171.7 (CO), 133.0, 132.0, 129.2, 127.9 (C_6H_5), 110.0 ($C(CH_3)_2$), 84.0 (C-1), 78.1 (C-3), 76.3 (C-2), 69.9 (C-4), 69.4 (C-5), 63.4 (C-6), 28.3, 26.4, 20.9

https://doi.org/10.1021/acs.joc.1c00250
J. Org. Chem. 2021, 86, 6390−6405

(CH_3). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{17}H_{22}O_6SNa$ 377.1029; Found 377.1031.

*Phenyl 6-O-Acetyl-2,3,4-tri-O-benzoyl-1-thio-α-D-mannopyranoside (**7**).* To a solution of **2** (10 g, 31.8 mmol) in dry pyridine (300 mL) was added dropwise benzoyl chloride (14.8 mL, 0.13 mol) at 0 °C under anhydrous conditions. The reaction mixture was stirred at room temperature for 17 h. The reaction mixture was then concentrated under reduced pressure. The residue was dissolved in EtOAc, and the organic layer was washed with aqueous 1 M HCl, followed by water, saturated aqueous $NaHCO_3$, and brine. The resulting aqueous layers were extracted with EtOAc, and the combined organic layers were dried over $MgSO_4$, filtered, and evaporated under reduced pressure. The resulting residue was purified by column chromatography on silica gel (cyclohexane/EtOAc gradient from 100:0 up to 80:20) to afford **7** (18.12 g, 91%) as a white solid. ^1H NMR ($CDCl_3$, 400 MHz): δ 8.13−8.08 (m, 2H, H_{arom}), 8.03−7.98 (m, 2H, H_{arom}), 7.90−7.83 (m, 2H, H_{arom}), 7.71−7.23 (m, 14H, H_{arom}), 6.02 (t, 1H, J = 10.1 Hz, H-4), 5.97 (dd, 1H, J = 3.2, 1.6 Hz, H-2), 5.86 (dd, 1H, J = 10.1, 3.2 Hz, H-3), 5.79 (d, 1H, J = 1.6 Hz, H-1), 4.88 (ddd, 1H, J = 10.1, 5.5, 2.7 Hz, H-5), 4.43 (dd, 1H, J = 12.2, 5.5 Hz, H-6), 4.30 (dd, 1H, J = 12.2, 2.7 Hz, H-6′). $^{13}C\{1H\}$ NMR ($CDCl_3$, 100 MHz): δ 170.6, 165.7, 165.5, 165.4 (CO), 133.8, 133.7, 133.5, 132.8, 132.3, 130.0, 130.0, 129.9, 129.4, 129.4, 129.0, 128.9, 128.7, 128.6, 128.5, 128.3 (C_6H_5), 86.0 (C-1), 72.0 (C-2), 70.4 (C-3), 69.8 (C-5), 67.2 (C-4), 62.9 (C-6). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{35}H_{30}O_9SNa$ 649.1503; Found 649.1501.

*Phenyl 2,3,4-Tri-O-benzoyl-1-thio-α-D-mannopyranoside (**8**).* To a solution of **7** (4.89 g, 7.82 mmol) in a mixture of dry DCM/dry MeOH 1:1 (80 mL) was added acetyl chloride (830 μL, 11.73 mmol) under anhydrous conditions. The mixture was stirred at room temperature for 17 h. The reaction was then stopped by the addition of saturated aqueous $NaHCO_3$. The resulting aqueous layer was extracted with DCM (2×). The combined organic layers were dried over $MgSO_4$, and solvent was removed under reduced pressure to afford **8** (4.34 g, 95%) as a white solid. The measured spectroscopic data are in concordance with the literature.[66] HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{33}H_{28}O_8SNa$ 607.1397; Found 607.1405.

*Synthesis of Galactofuranosyl-Containing Disaccharides **FP**. Phenyl 2,3,5,6-Tetra-O-benzoyl-β-D-galactofuranosyl-(1→2)-3-O-benzoyl-4,6-O-benzylidene-1-thio-α-D-mannopyranoside (**10**).* To a solution of trichloroacetimidate galactofuranosyl **9b** (72 mg, 0.1 mmol) and **5** (35 mg, 0.075 mmol) in anhydrous DCM (5 mL) were added activated 4 Å molecular sieves (100 mg). The mixture was cooled to 0 °C and stirred for 15 min before dropwise addition of TMSOTf (0.02 M in DCM, 1.1 mL, 0.022 mmol). The temperature was maintained at 0 °C, and the reaction was followed by TLC (cyclohexane/EtOAc 8/2). After 2 h, no starting material remained so triethylamine was added and the solution was filtered on a pad of Celite. The resulting filtrate was evaporated under vacuo, and the residue was purified by column chromatography on silica gel (cyclohexane/EtOAc gradient from 90:10 up to 70:30) to give **10** (42 mg, 54%) as an oil. ^1H NMR ($CDCl_3$, 400 MHz): δ 8.20 (ddd, 2H, J = 7.1, 2, 1.5 Hz, H_{arom}), 8.00 (ddd, 2H, J = 8.3, 2.8, 1.7 Hz, H_{arom}), 7.92 (ddd, 2H, J = 8.1, 2.9, 1.6 Hz, H_{arom}), 7.87 (ddd, 2H, J = 8.0, 2.9, 1.7 Hz, H_{arom}), 7.80 (ddd, 2H, J = 8.1, 2.9, 1.7 Hz, H_{arom}), 7.60 (tt, 1H, J = 7.5, 1.6 Hz, H_{arom}), 7.53 (ddd, 2H, J = 6.5, 2.9, 1.7 Hz, H_{arom}), 7.50−8.44 (m, 7H, H_{arom}), 7.40−7.27 (m, 12H, H_{arom}), 7.24−7.16 (m, 3H, H_{arom}), 5.68 (ddd, 1H, J = 7.5, 4.1, 3.0 Hz, H-5b), 5.67 (d, 1H, J = 1.6 Hz, H-1a), 5.63 (s, 1H, H-1b), 5.63 (dd, 1H, J = 4.1, 0.9 Hz, H-3b), 5.62 (dd, 1H, J = 10.3, 3.4 Hz, H-3a), 5.56 (d, 1H, J = 0.9 Hz, H-2b), 5.43 (s, 1H, H-7), 4.82 (dd, 1H, J = 3.4, 1.5 Hz, H-2a), 4.45 (dd, 1H, J = 10.3, 9.4 Hz, H-4a), 4.38 (t, 1H, J = 4.1 Hz, H-4b), 4.30 (dd, 1H, J = 10.3, 5.0 Hz, H-6a), 4.27 (dd, 1H, J = 12.2, 7.5 Hz, H-6b), 4.55 (td, 1H, J = 10.3, 5.0 Hz, H-5a), 4.12 (dd, 1H, J = 12.2, 3.0 Hz, H-6′b), 3.96 (t, 1H, J = 10.3 Hz, H-6′a). $^{13}C\{1H\}$ NMR ($CDCl_3$, 100 MHz): δ 165.96, 165.70, 165.68, 165.48, 165.44 (CO), 137.28, 133.91, 133.61, 133.57, 133.44, 133.30, 133.10, 132.11, 130.06, 129.93, 129.80, 129.66, 129.55, 129.54, 129.37, 129.36, 129.26, 128.75, 128.69, 128.64, 128.59, 128.52, 128.46, 128.45, 128.05, 126,28, 125,74 (C_6H_5), 103.97 (C-7), 102.00 (C-1b), 85.96 (C-1a), 83.05 (C-4b), 81.53 (C-2b), 77.68 (C-3b), 76.43 (C-4a), 74.81 (C-2a), 70.56 (C-3a), 70.35 (C-5b), 68.67 (C-6a), 65.45 (C-5a), 63.65 (C-6b). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{60}H_{50}O_{15}SNa$ 1065.2763; Found 1065.2761.

*Phenyl 2,3,5,6-Tetra-O-benzoyl-β-D-galactofuranosyl-(1→2)-3-O-benzoyl-1-thio-α-D-mannopyranoside (**11**).* To a solution of disaccharide **10** (440 mg, 0.42 mmol) in DCM (5 mL) were added 1 drop of water and TFA (180 μL, 1.33 mmol). The mixture was stirred during 1 h 30 at room temperature. The reaction was then quenched by addition of triethylamine and concentrated under reduced pressure. The residue was purified by column chromatography on silica gel (DCM/MeOH gradient from 100:0 to 95:5) to afford **11** (171 mg, 42%) as a white solid. ^1H NMR ($CDCl_3$, 400 MHz): δ 8.16−8.10 (m, 2H, H_{arom}), 8.00−7.93 (m, 4H, H_{arom}), 7.90−7.86 (m, 2H, H_{arom}), 7.82−7.78 (m, 2H, H_{arom}), 7.64−7.27 (m, 16H, H_{arom}), 7.25−7.15 (m, 4H, H_{arom}), 5.79 (dt, 1H, J = 7.5, 3.4 Hz, H-5b), 5.68 (d, 1H, J = 1.7 Hz, H-1a), 5.62 (d, 1H, J = 4.5 Hz, H-3b), 5.52 (d, 1H, J = 0.8 Hz, H-2b), 5.44 (s, 1H, H-1b), 5.40 (dd, 1H, J = 9.8, 3.2 Hz, H-3a), 4.69 (dd, 1H, J = 3.2, 1.7 Hz, H-2a), 4.47 (t, 1H, J = 9.8 Hz, H-4a), 4.43 (t, 1H, J = 4.5 Hz, H-4b), 4.37 (dd, 1H, J = 12.1, 7.5 Hz, H-6b), 4.31 (dt, 1H, J = 9.8, 3.6 Hz, H-5a), 4.24 (dd, 1H, J = 12.1, 3.4 Hz, H-6′b), 3.96 (d, 2H, J = 3.6 Hz, H-6a). $^{13}C\{1H\}$ NMR ($CDCl_3$, 100 MHz): δ 166.3, 166.0, 165.7, 165.6, 165.6 (CO), 133.9, 133.8, 133.6, 133.5, 133.3, 133.1, 132.3, 130.1, 129.9, 129.8, 129.7, 129.6, 129.5, 129.4, 129.1, 128.8, 128.7, 128.6, 128.5, 128.5, 128.4, 128.1 (C_6H_5), 103.8 (C-1b), 85.1 (C-1a), 82.9 (C-4b), 82.1 (C-2b), 77.7 (C-3b), 74.2 (C-2a), 74.1 (C-3a), 73.7 (C-5a), 70.3 (C-5b), 66.4 (C-4a), 63.6 (C-6b), 62.3 (C-6a). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{53}H_{46}O_{15}SNa$ 977.2450; Found 977.2452.

*Phenyl β-D-Galactofuranosyl-(1→2)-1-thio-α-D-mannopyranoside (**12**).* To a solution of disaccharide **11** (122 mg, 0.013 mmol) in anhydrous MeOH (2 mL) was added a solution of MeONa in MeOH (0.054M, 473 μL, 25 μmol). The mixture was then stirred overnight at room temperature. Then the solution was neutralized by the addition of IR 120-H^+ form resin. After filtration, the filtrate was evaporated under reduced pressure. The resulting residue was suspended in water and the biphasic solution was washed with DCM (3×). The aqueous phase was finally freeze-dried to afford **12** (53 mg, 95%) as a pulverulent white powder. ^1H NMR (D_2O, 400 MHz): δ 7.64−7.57 (m, 2H, H_{arom}), 7.47−7.40 (m, 3H, H_{arom}), 5.69 (s, 1H, H-1a), 5.15 (d, 1H, J = 1.0 Hz, H-1b), 4.32 (dd, 1H, J = 3.5, 1.3 Hz, H-2a), 4.21−4.16 (m, 1H, H-5a), 4.16−4.14 (m, 1H, H-2b), 4.11−4.05 (m, 2H, H-4b, H-3b), 3.93 (dd, 1H, J = 9.8, 3.5 Hz, H-3a), 3.88−3.76 (m, 3H, H-5b, H-6a), 3.75 (t, 1H, J = 9.8 Hz, H-4a), 3.70 (dd, 1H, J = 11.8, 4.4 Hz, H-6b), 3.65 (dd, 1H, J = 11.6, 7.3 Hz, H-6′b). $^{13}C\{1H\}$ NMR (D_2O, 100 MHz): δ 132.6, 132.3, 129.4, 128.4 (C_6H_5), 104.8 (C-1b), 84.8 (C-1a), 82.9 (C-4b), 81.1 (C-2b), 76.7 (C-3b), 75.5 (C-2a), 73.7 (C-5a), 70.6 (C-5b), 70.3 (C-3a), 67.2 (C-4a), 62.7 (C-6b), 60.6 (C-6a). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{18}H_{26}O_{10}SNa$ 457.1139; Found 457.1142.

*β-D-Galactofuranosyl-(1→2)-D-mannopyranose (**FP-2**).* To a solution of the disaccharide **12** (35 mg, 0.08 mmol) in a mixture of acetonitrile/H_2O 1:1 (1 mL) was added *N*-iodosaccharine (87 mg, 0.28 mmol). The resulting mixture was stirred for 2 h 30 at room temperature. Then the solvent was evaporated under reduced pressure. The resulting residue was partitioned between DCM and water. The resulting organic layer was further extracted with water (2×). The combined aqueous layers were finally freeze-dried. The resulting solid was purified by column chromatography on silica gel (EtOAc/AcOH/H_2O/MeOH gradient from 9:0.5:0.5:0 up to 5:2:2:1) to afford **FP-2** (19 mg, 68%) as a solid. HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{12}H_{22}O_{11}Na$ 365.1054; Found 365.1057.

FP-2*α* (major). ^1H NMR (D_2O, 400 MHz): δ 5.32 (d, 1H, J = 1.6 Hz, H-1a), 5.12 (d, 1H, J = 2 Hz, H-1b), 4.15 (t, 1H, J = 2.2 Hz, H-2b), 4.01 (dd, 1H, J = 3.4, 1.6 Hz, H-2a), 4.09−4.07 (m, 2H, H-3b, H-4b), 3.89 (dd, 1H, J = 9.7, 3.4 Hz, H-3a), 3.87 (dd, 1H, J = 11.7, 2.0 Hz, H-6a), 3.85−3.81 (m, 1H, H-5a, H-5b), 3.65 (t, 1H, J = 9.7 Hz, H-4a), 3.76 (dd, 1H, J = 11.7, 6.0 Hz, H-6′a), 3.72 (dd, 1H, J = 11.8, 4.3 Hz, H-6b), 3.65 (dd, 1H, J = 11.8, 7.3 Hz, H-6′b). $^{13}C\{1H\}$

https://doi.org/10.1021/acs.joc.1c00250
J. Org. Chem. 2021, 86, 6390−6405

NMR (D_2O, 100 MHz): δ 105.7 (C-1b), 91.3 (C-1a), 82−7 (C-4b), 81.0 (C-2b), 76.5 (C-3b), 69.4 (C-3a),72.5 (C-5a), 70.6 (C-5b), 75.5 (C-2a), 67.1 (C-4a), 62.7 (C-6b), 60.8 (C-6a).

FP-2*β* (minor). ^{1}H NMR (D_2O, 400 MHz): δ 4.98 (s, 1H, H-1a), 5.27 (d, 1H, *J* = 1.7 Hz, H-1b), 4.23 (dd, 1H, *J* = 3.8, 1.7 Hz, H-2b), 4.11 (d, 1H, *J* = 3.6 Hz, H-2a), 4.16−4.12 (m, 1H, H-4b), 4.13 (dd, 1H, *J* = 6.2, 3.8 Hz, H-3b), 3.67 (dd, 1H, *J* = 9.8, 3.6 Hz, H-3a), 3.89−3.85 (o, 1H, H-6a), 3.85−3.81 (m, 1H, H-5b), 3.69−3.65 (o, 1H, H-5a), 3.90−3.87 (o, 1H, H-4a), 3.75−3.68 (o, 1H, H-6′a), 3.75−3.70 (o, 1H, H-6b), 3.67−3.63 (o, 1H, H-6′b). ^{13}C{1H} NMR (D_2O, 100 MHz): δ 108.2 (C-1b), 94.0 (C-1a), 83.0 (C-4b), 80.7 (C-2b), 76.4 (C-3b), 69.4 (C-3a),72.2 (C-5a), 70.6 (C-5b), 76.7 (C-2a), 67.0 (C-4a), 62.7 (C-6b), 61.0 (C-6a).

*Phenyl 2,3,5,6-Tetra-O-benzoyl-β-D-galactofuranosyl-(1→3)-6-O-acetyl-1-thio-α-D-mannopyranoside (**13**).* Bromide donor **9a** (200 mg, 0.3 mmol), acceptor **2** (47 mg, 0.15 mmol), and 2-aminoethyl diphenylborinate (34 mg, 0.15 mmol) were suspended in dry acetonitrile (4 mL) in the presence of activated 4 Å molecular sieves under a nitrogen atmosphere. The mixture was stirred 30 min at room temperature before addition of Ag_2O (35 mg, 0.15 mmol). The mixture was then allowed to stir at room temperature in the dark, and the reaction was followed by TLC (cylohexane/EtOAc 7:3 and DCM/MeOH 9:1). After 2 h of stirring, no acceptor remained. Then, the solution was filtered on a plug of Celite, and the filtrate was evaporated under vacuum. The resulting crude was purified by column chromatography on silica gel (cyclohexane/EtOAc gradient from 9:1 up to 7:3) to give **13** (127 mg, 95%) as a white solid. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 8.14−7.90 (m, 8H, H_{arom}), 7.61−7.27 (m, 12H, H_{arom}), 5.95 (td, 1H, *J* = 5.3, 3.5 Hz, H-5b), 5.73 (dd, 1H, *J* = 5.9, 2.1 Hz, H-3b), 5.62 (d, 1H, *J* = 1.3 Hz, H-1a), 5.49 (s, 1H, H-1b), 5.48 (dd, 1H, *J* = 2.2, 0.8 Hz, H-2b), 4.93 (dd, 1H, *J* = 5.9, 3.5 Hz, H-4b), 4.81 (d, 2H, *J* = 5.3 Hz, H-6b), 4.42 (dd, 1H, *J* = 12.1, 5.8 Hz, H-6a), 4.34 (dd, 2H, *J* = 12.2, 2.1 Hz, H-6′a), 4.34−4.32 (m, 1H,H-2a), 4.32−4.28 (m, 1H, H-5a), 3.98−3.90 (m, 2H, H-3a, H-4a), 2.06 (s, 3H, CH_3). ^{13}C{1H} NMR ($CDCl_3$, 100 MHz): δ 171.4, 166.5, 166.3, 165.8, 165.7 (CO), 133.9, 133.6, 133.5, 133.4, 131.8, 130.1, 130.1, 130.0, 130.0, 129.6, 129.5, 129.2, 128.8, 128.7, 128.7, 128.6, 128.6, 127.8 (C_6H_5), 104.1 (C-1b), 87.4 (C-1a), 83.2 (C-2b), 81.4 (C-4b), 78.2 (C-3a), 77.0 (C-3b), 71.5 (C-5a), 70.2 (C-5b), 69.6 (C-2a), 66.3 (C-4a), 63.7 (C-6a), 63.2 (C-6b), 21.2 (CH_3). HRMS (ESI) *m*/*z*: $[M + Na]^+$ Calcd for $C_{48}H_{44}O_{15}SNa$ 915.2293; Found 915.2289.

*Phenyl β-D-Galactofuranosyl-(1→3)-1-thio-α-D-mannopyranoside (**14**).* To a solution of **13** (400 mg, 0.448 mmol) in anhydrous MeOH (5 mL) was added MeONa (0.54 M in MeOH, 170 μL, 0.090 mmol). The mixture was stirred for 17 h at room temperature. Then the solution was neutralized by the addition of IR 120-H^+ form resin. After filtration, the filtrate was evaporated under reduced pressure. The resulting residue was suspended in water, and the biphasic solution was washed with DCM (3×). The aqueous phase was finally freeze-dried to afford **14** (190 mg, 98%) as a white powder. ^{1}H NMR (D_2O, 400 MHz): δ 7.63−7.58 (m, 2H, H_{arom}), 7.47−7.40 (m, 3H, H_{arom}), 5.57 (d, 1H, *J* = 1.7 Hz, H-1a), 5.19 (d, 1H, *J* = 1.6 Hz, H-1b), 4.44 (dd, 1H, *J* = 3.2, 1.7 Hz, H-2a), 4.24−4.20 (m, 1H, H-5a), 4.18 (dd, 1H, *J* = 3.2, 1.6 Hz, H-2b), 4.14−4.07 (m, 2H, H-3b, H-4b), 3.96 (dd, 1H, *J* = 9.7, 3.2 Hz, H-3a), 3.90−3.85 (m, 1H, H-5b), 3.82 (t, 1H, *J* = 9.7 Hz, H-4a), 3.85−3.76 (m, 2H, H-6a), 3.74 (dd, 1H, *J* = 11.8, 4.5 Hz, H-6b), 3.69 (dd, 1H, *J* = 11.8, 7.4 Hz, H-6′b). ^{13}C{1H} NMR (D_2O, 100 MHz): δ 132.6, 132.2, 129.4, 128.4 (C_6H_5), 104.3 (C-1b), 87.8 (C-1a), 82.9 (C-4b), 81.3 (C-2b), 77.0 (C-3b), 75.6 (C-3a), 73.5 (C-5a), 70.7 (C-5b), 68.1 (C-2a), 65.3 (C-4a), 62.8 (C-6b), 60.7 (C-6a). HRMS (ESI) *m*/*z*: $[M + Na]^+$ Calcd for $C_{18}H_{26}O_{10}SNa$ 457.1139; Found 457.1138.

*β-D-Galactofuranosyl-(1→3)-D-mannopyranose (**FP-3**).* To a solution of **14** (14 mg, 0.032 mmol) in a mixture of acetonitrile/H_2O 1:1 (4 mL) was added *N*-iodosaccharine (35 mg, 0.113 mmol). The mixture was stirred for 2 h at room temperature. Then Et_3N was added, and the resulting reaction mixture was concentrated under reduced pressure. The residue was partitioned between DCM and water. The resulting organic layer was further extracted with water (2×). The combined aqueous layers were finally freeze-dried. The resulting solid was purified by column chromatography on silica gel (EtOAc/AcOH/H_2O/MeOH gradient from 9:0.5:0.5:0 up to 5:2:2:1) to afford **FP-3** (10 mg, 93%) as a solid. HRMS (ESI) *m*/*z*: $[M + Na]^+$ Calcd for $C_{12}H_{22}O_{11}Na$ 365.1054; Found 365.1058.

FP-3*α* (major). ^{1}H NMR (D_2O, 400 MHz): δ 5.24 (d, 1H, *J* = 1.8 Hz, H-1a), 5.18 (d, 1H, *J* = 1.5 Hz, H-1b), 4.18 (dd, 1H, *J* = 3.0, 1.5 Hz, H-2b), 4.15 (dd, 1H, *J* = 3.2, 1.8 Hz, H-2a), 4.10 (dd, 1H, *J* = 6.6, 3.0 Hz, H-3b), 4.09 (dd, 1H, *J* = 6.6, 3.8 Hz, H-4b), 3.95 (dd, 1H, *J* = 9.5, 3.2 Hz, H-3a), 3.92 (dd, 1H, *J* = 12.3, 2.3 Hz, H-6a), 3.90−3.85 (m, 2H, H-5a, H-5b), 3.77 (t, 1H, *J* = 9.5 Hz, H-4a), 3.77 (dd, 1H, *J* = 12.3, 6.1 Hz, H-6′a), 3.74 (dd, 1H, *J* = 11.6, 4.4 Hz, H-6b), 3.69 (dd, 1H, *J* = 11.6, 7.0 Hz, H-6′b). ^{13}C{1H} NMR (D_2O, 100 MHz): δ 104.4 (C-1b), 93.9 (C-1a), 83.0 (C-4b), 81.4 (C-2b), 77.0 (C-3b), 75.1 (C-3a), 72.6 (C-5a), 70.7 (C-5b), 67.4 (C-2a), 65.1 (C-4a), 62.7 (C-6b), 61.1 (C-6a).

FP-3*β* (minor). ^{1}H NMR (D_2O, 400 MHz): δ 5.20 (d, 1H, *J* = 1.4 Hz, H-1b), 4.92 (s, 1H, H-1a), 4.18 (dd, 1H, *J* = 3.0, 1.5 Hz, H-2b), 4.10 (d, 1H, *J* = 3.1 Hz, H-2a), 4.11 (dd, 1H, *J* = 6.6, 3.0 Hz, H-3b), 4.09 (dd, 1H, *J* = 6.6, 3.8 Hz, H-4b), 3.83−3.79 (m, 1H, H-3a), 3.94 (dd, 1H, *J* = 11.9, 2.2 Hz, H-6a), 3.90−3.85 (m, 1H, H-5b), 3.83−3.74 (o, 1H, H-6′a), 3.74 (dd, 1H, *J* = 11.6, 4.4 Hz, H-6b), 3.69 (dd, 1H, *J* = 11.6, 7.0 Hz, H-6′b), 3.68 (dd, 1H, *J* = 9.9, 3.1 Hz, H-4a), 3.45 (ddd, 1H, *J* = 9.9, 6.3, 2.2 Hz, H-5a). ^{13}C{1H} NMR (D_2O, 100 MHz): δ 104.1 (C-1b), 93.6 (C-1a), 82.9 (C-4b), 81.4 (C-2b), 77.4 (C-3a), 77.1 (C-3b), 76.0 (C-5a), 70.7 (C-5b), 67.7 (C-2a), 64.9 (C-4a), 62.7 (C-6b), 61.0 (C-6a).

*Phenyl 2,3,5,6-Tetra-O-benzoyl-β-D-galactofuranosyl-(1→4)-6-O-acetyl-3,4-O-isopropylidene-1-thio-α-D-mannopyranoside (**15**).* To a solution of galactofuranosyl trichloroacetimidate **9b** (50 mg, 70 μmol) and **6** (18 mg, 50 μmol) in anhydrous DCM (5 mL) were added activated 4 Å molecular sieves (100 mg). The mixture was cooled to 0 °C and stirred for 15 min before dropwise addition of TMSOTf (0.02 M in DCM, 750 μL, 15 μmol). The temperature was maintained at 0 °C, and the reaction was followed by TLC (cyclohexane/EtOAc 7:3). After 2 h, no starting material remained so triethylamine was added and the solution was filtered on a pad of Celite. The resulting filtrate was evaporated under vacuo and the residue was purified by column chromatography on silica gel (cyclohexane/EtOAc gradient from 9:1 up to 7:3) to give **15** (34 mg, 71%) as an oil. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 8.14−8.07 (m, 4H, H_{arom}), 8.00−7.96 (m, 2H, H_{arom}), 7.87−7.83 (m, 2H, H_{arom}), 7.63−7.27 (m, 17 H, H_{arom}), 6.18 (dt, 1H, *J* = 8.0, 3.5 Hz, H-5b), 5.84 (s, 1H, H-1a), 5.61 (dt, 1H, *J* = 5.4, 0.9 Hz, H-3b), 5.39−5.36 (m, 2H, H-2b, H-1b), 4.84 (dd, 1H, *J* = 5.4, 2.9 Hz, H-4b), 4.79 (dd, 1H, *J* = 12.1, 3.5 Hz, H-6b), 4.74 (dd, 1H, *J* = 12.1, 8.0 Hz, H-6′b), 4.42−4.31 (m, 3H, H-5a, H-6a), 4.35 (dd, 1H, *J* = 5.5, 0.9 Hz, H-2a), 4.27 (dd, 1H, *J* = 7.3, 5.5 Hz, H-3a), 3.92 (dd, 1H, *J* = 9.8, 7.3 Hz, H-4a), 1.96 (s, 3H, $COCH_3$), 1.65 (s, 3H, CH_3), 1.37 (s, 3H, CH_3). ^{13}C{1H} NMR ($CDCl_3$, 100 MHz): δ 170.7, 166.4, 165.9, 165.9, 165.8 (CO), 133.7, 133.5, 133.4, 133.2, 133.0, 131.9, 130.2, 130.1, 130.0, 129.9, 129.7, 129.3, 129.2, 128.9, 128.7, 128.6, 128.5, 128.5, 127.9 (C_6H_5), 110.1 ($C(CH_3)_2$), 105.7 (C-1b), 84.1 (C-1a), 82.4 (C-2b), 81.6 (C-4b), 77.4, 77.3 (C-3a, C-3b), 76.5 (C-2a), 75.1 (C-4a), 70.3 (C-5b), 68.3 (C-5a), 64.8 (C-6b), 62.8 (C-6a), 28.0, 26.6 (CH_3), 21.0 ($COCH_3$). HRMS (ESI) *m*/*z*: $[M + Na]^+$ Calcd for $C_{51}H_{48}O_{15}SNa$ 955.2606; Found 955.2605.

*Phenyl 2,3,5,6-Tetra-O-benzoyl-β-D-galactofuranosyl-(1→4)-6-O-acetyl-1-thio-α-D-mannopyranoside (**16**).* To a solution of **15** (108 mg, 0.116 mmol) in DCM (5 mL) were added 1 drop of water and TFA (44 μL, 0.579 mmol). The mixture was stirred during 15 h at room temperature. Then the reaction was quenched by addition of triethylamine, and the solvent was removed under reduced pressure. The resulting residue was purified by column chromatography on silica gel (cyclohexane/EtOAc gradient from 1:0 to 6:4) to afford **16** (103 mg, 100%) as a white solid. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 8.12−7.89 (m, 7H, H_{arom}), 7.62−7.27 (m, 18H, H_{arom}), 6.00 (dt, 1H, *J* = 6.2, 4.7 Hz, H-5b), 5.73 (dd, 1H, *J* = 5.2, 1.8 Hz, H-3b), 5.60 (d, 1H, *J* = 1.4 Hz, H-1a), 5.47 (dd, 1H, *J* = 1.9, 0.8 Hz, H-2b), 5.39 (s, 1H, H-1b), 4.91 (dd, 1H, *J* = 5.1, 4.6 Hz, H-4b), 4.81 (dd, 1H, *J* =

https://doi.org/10.1021/acs.joc.1c00250
J. Org. Chem. 2021, 86, 6390−6405

11.9, 4.7 Hz, H-6b), 4.72 (dd, 1H, J = 11.9, 6.2 Hz, H-6′b), 4.49−4.38 (m, 3H, H-5a, H-6a), 4.21 (dd, 1H, J = 2.6, 1.4 Hz, H-2a), 3.96−3.90 (m, 2H, H-3a, H-4a), 2.00 (s, 3H, CH_3). $^{13}C\{1H\}$ NMR ($CDCl_3$, 100 MHz): δ 170.9, 166.3, 165.8, 165.7, 165.7 (CO), 133.9, 133.8, 133.6, 133.6, 133.4, 131.7, 130.1, 130.1, 130.0, 129.9, 129.5, 129.4, 129.3, 128.8, 128.7, 128.7, 128.7, 128.6, 127.8 (C_6H_5), 108.0 (C1-b), 87.2 (C-1a), 82.6 (C-2b), 82.0 (C-4b), 78.7 (C-4a), 77.2 (C-3b), 71.9 (C-2a), 70.8 (C-3a), 70.4 (C-5b), 69.4 (C-5a), 63.1, 63.0 (C-6a, C-6b), 21.2 (CH_3). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{48}H_{44}O_{15}SNa$ 915.2293; Found 915.2290.

*Phenyl β-D-Galactofuranosyl-(1→4)-1-thio-α-D-mannopyranoside (**17**).* To a solution of **16** (275 mg, 0.31 mmol) in anhydrous MeOH (3 mL) was added MeONa (0.54 M in MeOH, 120 μL, 0.062 mmol). The mixture was stirred for 17 h at room temperature. Then the solution was neutralized by the addition of IR 120-H^+ form resin. After filtration, the filtrate was evaporated under reduced pressure. The resulting residue was suspended in water, and the biphasic solution was washed with DCM (3×). The aqueous phase was finally freeze-dried to afford **17** (122 mg, 91%) as a white powder. ^1H NMR (D_2O, 400 MHz): δ 7.58−7.54 (m, 2H, H_{arom}), 7.43−7.37 (m, 3H, H_{arom}), 5.50 (d, 1H, J = 1.8 Hz, H-1a), 5.08 (d, 1H, J = 2.2 Hz, H-1b), 4.24 (dd, 1H, J = 3.3, 1.8 Hz, H-2a), 4.24−4.20 (m, 1H, H-5a), 4.13−4.05 (m, 3H, H-2b, H-3b, H-4b), 3.95 (dd, 1H, J = 9.5, 3.3 Hz, H-3a), 3.85 (t, 1H, J = 9.6 Hz, H-4a), 3.83−3.78 (m, 3H, H-6a, H-5b), 3.69 (dd, 1H, J = 11.7, 4.6 Hz, H-6b), 3.64 (dd, 1H, J = 11.7, 7.3 Hz, H-6′b). $^{13}C\{1H\}$ NMR (D_2O, 100 MHz): δ 132.6, 132.3, 129.4, 128.3 (C_6H_5), 107.9 (C-1b), 87.9 (C-1a), 82.6 (C-4b), 81.0 (C-2b), 75.9 (C-3b), 75.2 (C-4a), 72.4 (C-5a), 71.3 (C-2a), 70.5 (C-5b), 69.7 (C-3a), 62.7 (C-6b), 60.1 (C-6a). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{18}H_{26}O_{10}SNa$ 457.1139; Found 457.1135.

*β-D-Galactofuranosyl-(1→4)-D-mannopyranose (**FP-4**).* To a solution of **17** (100 mg, 0.23 mmol) in a mixture of acetonitrile/H_2O 1:1 (10 mL) was added *N*-iodosaccharine (249 mg, 0.81 mmol). The mixture was stirred for 1 h at room temperature. Then the reaction mixture was concentrated under reduced pressure, and the residue was partitioned between DCM and water. The resulting organic layer was further extracted with water (2×). The combined aqueous layers were finally freeze-dried. The resulting solid was purified by column chromatography on silica gel (EtOAc/AcOH/H_2O/MeOH gradient from 9:0.5:0.5:0 up to 5:2:2:1) to afford **FP-4** (87 mg, 100%) as a solid. HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{12}H_{22}O_{11}Na$ 365.1054; Found 365.1058.

FP-4*α* (major). ^1H NMR ($(CD_3)_2CO$, 500 MHz): δ 5.07 (d, 1H, J = 1.7 Hz, H-1a), 5.01 (d, 1H, J = 2.5 Hz, H-1b), 3.99 (dd, 1H, J = 4.5, 2.5 Hz, H-2b), 4.091 (dd, 1H, J = 6.6, 3.3 Hz, H-4b), 4.04 (dd, 1H, J = 6.6, 4.5 Hz, H-3b), 3.86−3.81 (m, 1H, H-4a), 3.84−3.82 (m, 1H, H-2a), 3.83−3.79 (m, 1H, H-3a), 3.78−3.69 (m, 3H, H-6a, H-5a), 3.73 (dt, 1H, J = 6.4, 3.3 Hz, H-5b), 3.57−3.61 (m, 2H, H-6b). $^{13}C\{1H\}$ NMR (D_2O, 125 MHz): δ 109.07 (C-1b), 94.81 (C-1a), 83.80 (C-4b), 81.92 (C-2b), 77.27 (C-3b), 75.62 (C-4a), 72.00 (C-5a), 71.74 (C-2a), 71.32 (C-5b), 70.22 (C-3a), 63.46 (C-6b), 61.27 (C-6a).

FP-4*β* (minor). ^1H NMR ($(CD_3)_2CO$, 500 MHz): δ 5.01 (d, 1H, J = 2.5 Hz, H-1b), 4.75 (d, 1H, J = 1.1 Hz, H-1a), 4.093 (dd, 1H, J = 6.6, 3.3 Hz, H-4b), 4.04 (dd, 1H, J = 6.6, 4.5 Hz, H-3b), 3.98 (dd, 1H, J = 4.5, 2.5 Hz, H-2b), 3.86 (dd, 1H, J = 3.3, 1.1 Hz, H-2a), 3.78 (dd, 1H, J = 12.1, 2.2 Hz, H-6a), 3.75 (t, 1H, J = 9.6 Hz, H-4a), 3.726 (dt, 1H, J = 6.4, 3.3 Hz, H-5b), 3.70 (dd, 1H, J = 12.1, 4.6 Hz, H-6′a), 3.60 (dd, 1H, J = 9.6, 3.3 Hz, H-3a), 3.57−3.61 (m, 2H, H-6b), 3.31 (ddd, 1H, J = 9.6, 4.6, 2.2 Hz, H-5a). $^{13}C\{1H\}$ NMR (D_2O, 125 MHz): δ 109.07 (C-1b), 94.70 (C-1a), 83.93 (C-4b), 81.88 (C-2b), 77.27 (C-3b), 76.09 (C-5a), 75.04 (C-4a), 73.08 (C-3a), 71.99 (C-2a), 71.32 (C-5b), 63.41 (C-6b), 61.23 (C-6a).

*Phenyl 2,3,5,6-Tetra-O-benzoyl-β-D-galactofuranosyl-(1→6)-2,3,4-tri-O-benzoyl-1-thio-α-D-mannopyranoside (**18**).* To a solution of donor **9a** (850 mg, 1.29 mmol) and acceptor **8** (502 mg, 0.86 mmol) in anhydrous dichloromethane (10 mL) were added activated 4 Å molecular sieves. The media was stirred during 15 min at room temperature under nitrogen atmosphere. Then AgOTf (244 mg, 0.95 mmol) was added, and the reaction was stirred in the dark for 6 h. Then the reaction was quenched by addition of triethylamine, and the resulting suspension was filtered on Celite. The filtrate was concentrated under reduced pressure, and the residue was purified by column chromatography on silica gel (cyclohexane/EtOAc gradient from 1:0 up to 7:3) to afford 18 (930 mg, 93%) as a white powder. ^1H NMR ($CDCl_3$, 400 MHz): δ 8.11−7.81 (m, 16H, H_{arom}), 7.65−7.40 (m, 13H, H_{arom}), 7.38−7.21 (m, 10H, H_{arom}), 7.17−7.11 (m, 1H, H_{arom}), 5.98 (t, 1H, J = 10.0 Hz, H-4a), 5.98−5.95 (m, 1H, H-2a), 5.95−5.94 (m, 1H, H-5b), 5.86 (dd, 1H, J = 10.0, 3.2 Hz, H-3a), 5.76 (d, 1H, J = 1.5 Hz, H-1a), 5.61 (ddd, 1H, J = 5.1, 1.4, 0.6 Hz, H-3b), 5.56 (d, 1H, J = 1.4 Hz, H-2b), 5.39 (s, 1H, H-1b), 4.91 (ddd, 1H, J = 8.9, 6.5, 2.1 Hz, H-5a), 4.67−4.63 (m, 2H, H-6b), 4.61 (dd, 1H, J = 5.1, 3.7 Hz, H-4b), 4.05 (dd, 1H, J = 11.5, 2.3 Hz, H-6a), 3.92 (dd, 1H, J = 11.5, 6.5 Hz, H-6′a). $^{13}C\{1H\}$ NMR ($CDCl_3$, 100 MHz): δ 166.2, 165.8, 165.7, 165.7, 165.6, 165.6, 165.3 (CO),133.7, 133.6, 133.4, 133.4, 133.3, 133.1, 132.3, 132.2, 130.1, 130.1, 130.0, 130.0, 129.9, 129.9, 129.8, 129.7, 129.6, 129.4, 129.4, 129.3, 129.2, 129.0, 129.0, 128.8, 128.8, 128.7, 128.6, 128.5, 128.5, 128.4, 128.4, 128.3 (C_6H_5), 106.6 (C-1b), 86.2 (C-1a), 82.0 (C-4b), 81.9 (C-2b), 77.7 (C-3b), 72.2 (C-5b), 71.4 (C-5a), 70.5, 70.5 (C-2a, C-3a), 67.6 (C-4a), 66.9 (C-6a), 63.7 (C-6b). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{67}H_{54}O_{17}SNa$ 1185.2974; Found 1185.2972.

*Phenyl β-D-Galactofuranosyl-(1→6)-1-thio-α-D-mannopyranoside (**19**).* To a solution of **18** (688 mg, 0.59 mmol) in anhydrous MeOH (10 mL) was added MeONa (5.4 M in MeOH, 27 μL, 0.14 mmol). The mixture was stirred for 17 h at room temperature. Then the solution was neutralized by the addition of IR 120-H^+ form resin. After filtration, the filtrate was evaporated under reduced pressure. The resulting residue was suspended in water and the biphasic solution was washed with DCM (3×). The aqueous phase was finally freeze-dried to afford **19** (257 mg, 100%) as a powder. ^1H NMR (CD_3OD, 400 MHz): δ 7.57−7.53 (m, 2H, H_{arom}), 7.37−7.26 (m, 3H, H_{arom}), 5.37 (d, 1H, J = 1.2 Hz, H-1a), 4.93 (s, 1H, H-1b), 4.21 (ddd, 1H, J = 8.7, 6.3, 2.3 Hz, H-5a), 4.08 (dd, 1H, J = 2.9, 1.6 Hz, H-2a), 4.02 (dd, 1H, J = 11.0, 2.3 Hz, H-6a), 4.00−3.95 (m, 3H, H-2b, H-3b, H-4b), 3.74−3.64 (m, 4H, H-3a, H-4a, H-6′a, H-5b), 3.61 (d, 2H, J = 6.2 Hz, H-6b). $^{13}C\{1H\}$ NMR (CD_3OD, 100 MHz): δ 135.7, 133.2, 132.9, 130.1, 128.6 (C_6H_5), 109.9 (C-1b), 90.5 (C-1a), 84.8 (C-4b), 83.0 (C-2b), 79.1 (C-3b), 74.4 (C-5a), 73.6 (C-2a), 73.1 (C-3a), 72.5 (C-5b), 69.1 (C-4a), 68.2 (C-6a), 64.6 (C-6b). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{18}H_{26}O_{10}SNa$ 457.1139; Found 457.1140.

*β-D-Galactofuranosyl-(1→6)-D-mannopyranose (**FP-6**).* To a solution of **19** (270 mg, 0.62 mmol) in a mixture of acetonitrile/H_2O 1:1 (10 mL) was added *N*-iodosaccharine (731 mg, 2.36 mmol). The mixture was stirred for 1 h at room temperature. Then the reaction mixture was concentrated under reduced pressure, and the residue was partitioned between DCM and water. The resulting organic layer was further extracted with water (2×). The combined aqueous layers were finally freeze-dried. The resulting solid was purified by column chromatography on silica gel (EtOAc/AcOH/H_2O/MeOH gradient from 9:0.5:0.5:0 up to 5:2:2:1) to afford **FP-6** (193 mg, 89%) as a solid. HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{12}H_{22}O_{11}Na$ 365.1054; Found 365.1058.

FP-6*α*. ^1H NMR (CD_3OD, 500 MHz): δ 5.06 (d, 1H, J = 1.8 Hz, H-1a), 4.95 (s, 1H, H-1b), 4.02 (dd, 1H, J = 4.7, 3.2 Hz, H-4b), 3.99−3.96 (m, 2H, H-2b, H-3b), 3.96 (dd, 1H, J = 10.9, 2.5 Hz, H-6a), 3.89 (ddd, 1H, J = 9.6, 5.6, 2.5 Hz, H-5a), 3.79 (dd, 1H, J = 3.3, 1.8 Hz, H-2a), 3.75−3.71 (m, 1H, H-5b), 3.75 (dd, 1H, J = 9.2, 3.3 Hz, H-3a), 3.65 (dd, 1H, J = 10.9, 5.6 Hz, H-6′a), 3.64−3.59 (m, 2H, H-6b), 3.61 (t, 1H, J = 9.6 Hz, H-4a). $^{13}C\{1H\}$ NMR (CD_3OD, 125 MHz): δ 109.45 (C-1b), 95.60 (C-1a), 84.97 (C-4b), 82.12 (C-2b), 78.82 (C-3b), 72.41 (C-2a), 72.31, 72.28 (C-5a, C-5b), 71.92 (C-3a), 68.96 (C-4a), 68.17 (C-6a), 64.18 (C-6b).

FP-6*β*. ^1H NMR (CD_3OD, 500 MHz): δ 4.95 (s, 1H, H-1b), 4.73 (d, 1H, J = 0.9 Hz, H-1a), 4.02−4.01 (o, 1H, H-3b), 4.02−3.96 (o, 1H, H-6a), 4.00−3.98 (o, 1H, H-4b), 3.98−3.96 (o, 1H, H-2b), 3.91−3.88 (o, 1H, H-5b), 3.81 (dd, 1H, J = 3.3, 0.9 Hz, H-2a), 3.66−3.60 (m, 3H, H-6′a, H-6b), 3.56 (t, 1H, J = 9.6 Hz, H-4a), 3.44 (dd, 1H, J = 9.6, 3.3 Hz, H-3a), 3.35 (ddd, 1H, J = 9.6, 5.9, 2.6 Hz, H-5a).

https://doi.org/10.1021/acs.joc.1c00250
J. Org. Chem. 2021, 86, 6390−6405

$^{13}C\{1H\}$ NMR (CD_3OD, 100 MHz): δ 109.56 (C-1b), 95.35 (C-1a), 85.00 (C-4b), 82.16 (C-2b), 78.80 (C-3b), 76.37 (C-5a), 75.01 (C-3a), 72.72 (C-2a), 72.31 or 72.28 (C-5b), 68.48 (C-4a), 68.06 (C-6a), 64.18 (C-6b).

Synthesis of Galactopyranosyl-Containing Disaccharides ***PP****. Phenyl 2,3,4,6-Tetra-O-benzoyl-β-D-galactopyranosyl-(1→2)-3-O-benzoyl-1-thio-α-D-mannopyranoside (**21**).* To a solution of trichloroacetimidate galactopyranosyl **20b** (207 mg, 0.28 mmol) and **5** (100 mg, 0.21 mmol) in anhydrous DCM (10 mL) were added activated 4 Å molecular sieves (100 mg). The mixture was cooled to 0 °C and stirred for 15 min before dropwise addition of TMSOTf (0.02 M in DCM, 3.2 mL, 0.06 mmol). The temperature was maintained at 0 °C and reaction was followed by TLC (cyclohexane/EtOAc 8/2). After 2 h, no starting material remained, so triethylamine was added and the solution was filtered on a pad of Celite. The filtrate was evaporated in vacuo. The resulting residue was dissolved in DCM (10 mL), and water (10 μL) and TFA (124 μL, 1.62 mmol) were added successively. The reaction mixture was stirred at room temperature for 24 h. Then triethylamine was added, and the reaction media was washed with water (3×), dried over $MgSO_4$, and evaporated. The residue was purified by column chromatography on silica gel (cyclohexane/EtOAc gradient from 90:10 up to 60:40) to give **21** (53 mg, 26% over two steps) as an oil. ^1H NMR ($CDCl_3$, 400 MHz): δ 8.12 (dd, 2H, J = 8.5, 1.3 Hz, H_{arom}), 8.09 (dd, 2H, J = 8.6, 1.3 Hz, H_{arom}), 7.94 (dd, 2H, J = 8.5, 1.2 Hz, H_{arom}), 7.88 (dd, 2H, J = 8.4, 1.3 Hz, H_{arom}), 7.76 (dd, 2H, J = 8.4, 1.2 Hz, H_{arom}), 7.68 (tt, 1H, J = 7.5, 1.3 Hz, H_{arom}), 7.59−7.51 (m, 3H, H_{arom}), 7.49−7.18 (m, 12H, H_{arom}), 7.14 (td, 2H, J = 7.3, 1.3 Hz, H_{arom}), 7.07 (dd, 2H, J = 7.3, 1.3 Hz, H_{arom}), 5.89 (dd, 1H, J = 3.5, 0.9 Hz, H-4b), 5.81 (dd, 1H, J = 10.5, 7.8 Hz, H-2b), 5.61 (dd, 1H, J = 10.5, 3.5 Hz, H-3b), 5.24 (d, 1H, J = 2.1 Hz, H-1a), 5.21 (dd, 1H, J = 9.7, 3.1 Hz, H-3a), 4.75 (d, 1H, J = 7.8 Hz, H-1b), 4.56 (dd, 1H, J = 3.1, 2.1 Hz, H-2a), 4.53 (t, 1H, J = 9.7 Hz, H-4a), 4.14 (dd, 1H, J = 6.5, 0.9 Hz, H-5b), 4.04 (apparent td, 1H, J = 9.7, 3.0 Hz, H-5a), 4.00 (dd, 1H, J = 11.4, 6.5 Hz, H-6b), 3.93 (dd, 1H, J = 11.4, 6.5 Hz, H-6′b), 3.79 (dd, 1H, J = 12.1, 3.6 Hz, H-6a), 3.68 (dd, 1H, J = 12.1, 2.4 Hz, H-6′a). $^{13}C\{1H\}$ NMR ($CDCl_3$, 100 MHz): δ 166.77, 165.89, 165.57, 165.48 (COPh), 134.03, 133.89, 133.56, 133.53, 133.38, 130.31, 130.13, 130.07, 129.95, 129.92, 129.90, 129.64, 129.45, 129.33, 129.22, 129.14, 128.88, 128.80, 128.72, 128.54, 128.46, 128.35, 127.25 (C_6H_5), 101.44 (C-1b), 85.28 (C-1a), 77.90 (C-2a), 73.80 (C-3a), 73.55 (C-5a), 71.60 (C-5b), 71.01 (C-3b), 70.24 (C-2b), 68.00 (C-4b), 65.11 (C-4a), 61.97 (C-6a), 61.70 (C-6b). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{53}H_{46}O_{15}SNa$ 977,2455; Found 977,2536.

*Phenyl β-D-Galactopyranosyl-(1→2)-1-thio-α-D-mannopyranoside (**22**).* To a solution of **21** (36 mg, 0.05 mmol) in anhydrous MeOH (1 mL) was added MeONa (0.054 M in MeOH, 200 μL, 11 μmol). The mixture was stirred for 17 h at room temperature. Then the solution was neutralized by the addition of IR 120-H^+ form resin. After filtration, the filtrate was evaporated under reduced pressure. The resulting residue was suspended in water and the biphasic solution was washed with DCM (3×). The aqueous phase was finally freeze-dried to afford **22** (20 mg, 91%) as a powder. ^1H NMR (D_2O, 400 MHz): δ 7.64−7.57 (m, 2H, H_{arom}), 7.47−7.39 (m, 3H, H_{arom}), 5.67 (d, 1H, J = 1.4 Hz, H-1a), 4.91 (dd, 1H, J = 3.4, 1.1 Hz, H-4b), 4.46 (d, 1H, J = 7.5 Hz, H-1b), 4.44 (dd, 1H, J = 3.5, 1.4 Hz, H-2a), 4.17 (ddd, 1H, J = 9.9, 4.5, 2.9 Hz, H-5a), 3.94 (dd, 1H, J = 9.9, 3.5 Hz, H-3a), 3.85−3.81 (m, 2H, H-6a), 3.81 (dd, 1H, J = 12.0, 7.9 Hz, H-6b), 3.80 (t, 1H, J = 9.9 Hz, H-4a), 3.74 (dd, 1H, J = 12.0, 3.9 Hz, H-6′b), 3.66 (ddd, 1H, J = 7.9, 3.9, 1.1 Hz, H-5b), 3.64 (dd, 1H, J = 11.1, 3.3 Hz, H-3b), 3.57 (dd, 1H, J = 11.1, 7.5 Hz, H-2b). $^{13}C\{1H\}$ NMR (D_2O, 100 MHz): δ 132.71, 132.14, 129.45, 128.47 (C_6H_5), 101.45 (C-1b), 85.83 (C-1a), 78.17 (C-2a), 75.25 (C-5b), 73.56 (C-5a), 72.47 (C-3b), 70.40 (C-2b), 70.32 (C-3a), 68.60 (C-4b), 66.97 (C-4a), 61.11 (C-6b), 60.19 (C-6a). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{18}H_{26}O_{10}SNa$ 457.1139; Found 457.1141.

*β-D-Galactopyranosyl-(1→2)-D-mannopyranoside (**PP-2**).* To a solution of **22** (18 mg, 4 μmol) in a mixture of acetonitrile/H_2O 1:1 (5 mL) was added *N*-iodosaccharine (45 mg, 15 μmol). The mixture was stirred for 2 h at room temperature. Then the reaction mixture was concentrated under reduced pressure, and the residue was partitioned between DCM and water. The resulting organic layer was further extracted with water (2×). The combined aqueous layers were finally freeze-dried. The resulting solid was purified by column chromatography on silica gel (EtOAc/AcOH/H_2O/MeOH gradient from 9:0.5:0.5:0 up to 5:2:2:1) to afford **PP-2** (10 mg, 72%) as a solid. HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{12}H_{22}O_{11}Na$ 365.1054; Found 365.1056.

PP-2*α* (major). ^1H NMR ($(CD_3)_2CO$, 500 MHz): δ 5.22 (d, 1H, J = 1.8 Hz, H-1a),4.38−4.33 (m, 1H, H-1b), 3.98 (dd, 1H, J = 3.4, 1.8 Hz, H-2a), 3.85−3.84 (m, 1H, H-3b), 3.77 (dd, 1H, J = 9.5, 3.4 Hz, H-3a), 3.75−3.72 (m, 2H, H-5a, H-6b), 3.74−3.72 (m, 2H, H-6a), 3.65 (dd, 1H, J = 11.7, 4.0 Hz, H-6′b), 3.65 (t, 1H, J = 9.5 Hz, H-4a), 3.60 (ddd, 1H, J = 7.8, 4.2, 0.9 Hz, H-5b), 3.59−3.56 (m, 1H, H-4b), 3.57−3.55 (m, 1H, H-2b). $^{13}C\{1H\}$ NMR ($(CD_3)_2CO$, 125 MHz): δ 103.3 (C-1b), 92.92 (C-1a), 79.16 (C-2a), 75.99 (C-5b), 73.3 (C-4b), 73.09 (C-5a), 71.10 (C-2b), 70.32 (C-3a), 69.25 (C-3b), 67.87 (C-4a), 61.66 (C-6b), 60.86 (C-6a).

PP-2*β* (minor). ^1H NMR ($(CD_3)_2CO$, 500 MHz): δ 4.84 (d, 1H, J = 0.9 Hz, H-1a), 4.44 (d, 1H, J = 7.7 Hz, H-1b), 4.05 (dd, 1H, J = 2.9, 0.9 Hz, H-2a), 3.85−3.84 (o, 1H, H-3b), 3.81 (o, 2H, H-6a), 3.73 (o, 1H, H-6b), 3.67 (o, 1H, H-6′b), 3.64 (o, 1H, H-2b), 3.62 (o, 1H, H-5b), 3.58−3.54 (o, 1H, H-3a), 3.56 (o, 1H, H-4b), 3.54 (o, 1H, H-4a), 3.25 (ddd, 1H, J = 8.3, 5.9, 2.2 Hz, H-5a). $^{13}C\{1H\}$ NMR ($(CD_3)_2CO$, 125 MHz): δ 105.22 (C-1b), 94.37 (C-1a), 81.50 (C-2a), 77.14 (C-5a), 76.21 (C-5b), 73.44 (C-4b), 73.14 (C-3a), 71.81 (C-2b), 69.28 (C-3b), 67.70 (C-4a), 61.52 (C-6a), 61.61 (C-6b).

*Phenyl 2,3,4,6-Tetra-O-acetyl-β-D-galactopyranosyl-(1→3)-6-O-acetyl-1-thio-α-D-mannopyranoside (**23**).* Bromide donor **20a** (956 mg, 2.3 mmol), acceptor **2** (487 mg, 1.55 mmol), and 2-aminoethyl diphenylborinate (349 mg, 1.55 mmol) were suspended in anhydrous CH_3CN (20 mL) in the presence of activated 4 Å molecular sieves under a nitrogen atmosphere. The mixture was stirred for 15 min at room temperature before addition of Ag_2O (359 mg, 1.55 mmol). The mixture was then allowed to stir at room temperature in the dark for 3 h. Then the solution was filtered on a plug of Celite, and the filtrate was evaporated under vacuum. The resulting crude was purified by column chromatography on silica gel (cyclohexane/EtOAc gradient from 10:1 up to 6:4) to give **23** (1.46 g, 73%) as a white solid. ^1H NMR ($CDCl_3$, 400 MHz): δ 7.50−7.46 (m, 2H, H_{arom}), 7.34−7.27 (m, 3H, H_{arom}), 5.59 (d, 1H, J = 1.3 Hz, H-1a), 5.43 (dd, 1H, J = 3.4, 1.0 Hz, H-4b), 5.25 (dd, 1H, J = 10.6, 7.9 Hz, H-2b), 5.07 (dd, 1H, J = 10.6, 3.4 Hz, H-3b), 4.62 (d, 1H, J = 7.9 Hz, H-1b), 4.40−4.30 (m, 2H, H-6a), 4.34−4.28 (m, 1H, H-5a), 4.19 (dd, 1H, J = 11.4, 5.3 Hz, H-6b), 4.13 (dd, 1H, J = 11.4, 7.6 Hz, H-6′b), 4.09 (dd, 1H, J = 3.4, 1.3 Hz, H-2a), 4.05 (ddd, 1H, J = 7.6, 5.3, 1.0 Hz, H-5b), 3.91 (t, 1H, J = 9.0 Hz, H-4a), 3.74 (dd, 1H, J = 9.0, 3.4 Hz, H-3a), 2.17 (s, 3H, CH_3), 2.11 (s, 3H, CH_3), 2.08 (s, 3H, CH_3), 2.02 (s, 3H, CH_3), 2.01 (s, 3H, CH_3). $^{13}C\{1H\}$ NMR ($CDCl_3$, 100 MHz): δ 171.0, 170.6, 170.2, 170.1, 170.1 (CO), 133.0, 131.9, 129.3, 128.0 (C_6H_5), 102.1 (C-1b), 87.1 (C-1a), 84.5 (C-3a), 71.6 (C-2a), 71.5 (C-5b), 70.9 (C-5a), 70.5 (C-3b), 69.2 (C-2b), 66.9 (C-4a), 66.3 (C-4b), 63.5 (C-6a), 61.8 (C-6b), 21.0, 20.8, 20.7, 20.7 (CH_3). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{28}H_{36}O_{15}SNa$ 667.1667; Found 667.1665.

*Phenyl β-D-Galactopyranosyl-(1→3)-1-thio-α-D-mannopyranoside (**24**).* To a solution of **23** (350 mg, 0.54 mmol) in anhydrous MeOH (5 mL) was added MeONa (0.054 M in MeOH, 2 mL, 0.11 mmol). The mixture was stirred for 17 h at room temperature. Then the solution was neutralized by the addition of IR 120-H^+ form resin. After filtration, the filtrate was evaporated under reduced pressure. The resulting residue was suspended in water and the biphasic solution was washed with DCM (3×). The aqueous phase was finally freeze-dried to afford **24** (235 mg, 100%) as of a white powder. ^1H NMR (D_2O, 400 MHz): δ 7.62−7.57 (m, 2H, H_{arom}), 7.46−7.40 (m, 3H, H_{arom}), 5.58 (d, 1H, J = 1.6 Hz, H-1a), 4.54 (d, 1H, J = 7.6 Hz, H-1b), 4.42 (dd, 1H, J = 3.3, 1.7 Hz, H-2a), 4.20 (ddd, 1H, J = 9.7, 5.5, 2.5 Hz, H-5a), 4.07 (dd, 1H, J = 9.7, 3.1 Hz, H-3a), 3.94 (d, 1H, J = 3.3 Hz, H-4b), 3.86 (t, 1H, J = 9.6 Hz, H-4a), 3.88−3.75 (m, 4H, H-6a, H-6b), 3.74−3.70 (m, 1H, H-5b), 3.68 (ddd, 1H, J = 9.9, 3.3,

https://doi.org/10.1021/acs.joc.1c00250
J. Org. Chem. 2021, 86, 6390−6405

0.5 Hz, H-3b), 3.62 (dd, 1H, J = 9.9, 7.6 Hz, H-2b). ^{13}C{1H} NMR (D_2O, 100 MHz): δ 132.7, 132.1, 129.4, 128.4 (C_6H_5), 101.0 (C-1b), 87.7 (C-1a), 78.6 (C-3a), 75.3 (C-5b), 73.6 (C-5a), 72.6 (C-3b), 70.7 (C-2b), 69.1 (C-2a), 68.6 (C-4b), 65.4 (C-4a), 61.0, 60.7 (C-6a, C-6b). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{18}H_{26}O_{10}SNa$ 457.1139; Found 457.1143.

*β-D-Galactopyranosyl-(1→3)-D-mannopyranose (**PP-3**).* To a solution of **24** (231 mg, 0.53 mmol) in a mixture of acetonitrile/H_2O 1:1 (5 mL) was added *N*-iodosaccharine (576 mg, 1.86 mmol). The mixture was stirred for 2 h at room temperature. Then the reaction mixture was concentrated under reduced pressure, and the residue was partitioned between DCM and water. The resulting organic layer was further extracted with water (2×). The combined aqueous layers were finally freeze-dried. The resulting solid was purified by column chromatography on silica gel (EtOAc/AcOH/H_2O/MeOH gradient from 9:0.5:0.5:0 up to 5:2:2:1) to afford **PP-3** (132 mg, 73%) as a solid. HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{12}H_{22}O_{11}Na$ 365.1054; Found 365.1055.

PP-3*α* (major). ^{1}H NMR (D_2O, 400 MHz): δ 5.24 (d, 1H, J = 1.9 Hz, H-1a), 4.55 (d, 1H, J = 7.6 Hz, H-1b), 4.14 (dd, 1H, J = 3.2, 1.9 Hz, H-2a), 4.08 (dd, 1H, J = 9.4, 3.2 Hz, H-3a), 3.96−3.94 (m, 1H, H-4b), 3.94−3.88 (m, 2H, H-6a), 3.90−3.85 (m, 1H, H-5a), 3.82−3.74 (m, 3H, H-4a, H-6b), 3.76−3.71 (m, 1H, H-5b), 3.71 (dd, 1H, J = 9.8, 3.3 Hz, H-3b), 3.626 (dd, 1H, J = 9.8, 7.6 Hz, H-2b). ^{13}C{1H} NMR (D_2O, 100 MHz): δ 101.00 (C-1b), 93.74 (C-1a), 78.11 (C-3a), 75.30 (C-5b), 72.63 (C-3b), 72.32 (C-5a), 70.75 (C-2b), 68.39 (C-2a), 68.66 (C-4b), 65.27 (C-4a), 61.1 (C-6a), 60.96 (C-6b).

PP-3*β* (minor). ^{1}H NMR (D_2O, 400 MHz): δ 4.92 (d, 1H, J = 0.7 Hz, H-1a), 4.56 (d, 1H, J = 7.7 Hz, H-1b), 4.17 (dd, 1H, J = 3.1, 0.7 Hz, H-2a), 3.96−3.94 (o, 1H, H-4b), 3.92 (dd, 1H, J = 9.7, 3.1 Hz, H-3a), 3.92 (dd, 1H, J = 12.5, 2.5 Hz, H-6a), 3.82−3.77 (o, 2H, H-6b), 3.76−3.71 (o, 1H, H-5b), 3.75 (dd, 1H, J = 12.5, 6.4 Hz, H-6′a), 3.72−3.69 (o, 1H, H-3b), 3.71 (t, 1H, J = 9.7 Hz, H-4a), 3.631 (dd, 1H, J = 10.0, 7.7 Hz, H-2b), 3.44 (ddd, 1H, J = 9.7, 6.4, 2.5 Hz, H-5a). ^{13}C{1H} NMR (D_2O, 100 MHz): δ 100.75 (C-1b), 93.53 (C-1a), 80.18 (C-3a), 72.62 (C-5b), 75.90 (C-5a), 75.32 (C-3b), 70.75 (C-2b), 68.79 (C-2a), 68.65 (C-4b), 65.07 (C-4a), 61.09 (C-6b), 60.99 (C-6a).

*Phenyl 2,3,4,6-Tetra-O-benzoyl-β-D-galactopyranosyl-(1→6)-2,3,4-O-benzoyl-1-thio-α-D-mannopyranoside (**25**).* To a solution of galactopyranosyl trichloroacetimidate **20b** (329 mg, 0.44 mmol) and **8** (200 mg, 0.34 mmol) in anhydrous DCM (15 mL) were added activated 4 Å molecular sieves (100 mg). The mixture was stirred for 30 min at room temperature before the dropwise addition of TMSOTf (0.02 M in DCM, 5.1 mL, 0.10 mmol). After 3 h, triethylamine was added, and the solution was filtered on a pad of Celite. The resulting filtrate was evaporated under vacuo, and the residue was purified by column chromatography on silica gel (cyclohexane/EtOAc gradient from 9:1 up to 8:2) to give **25** (309 mg, 78%) as an oil. ^{1}H NMR ($CDCl_3$, 400 MHz): δ 8.07−7.98 (m, 6H, H_{arom}), 7.95−7.86 (m, 4H, H_{arom}), 7.81−7.76 (m, 4H, H_{arom}), 7.63−7.55 (m, 4H, H_{arom}), 7.55−7.33 (m, 16H, H_{arom}), 7.31−7.21 (m, 6H, H_{arom}), 5.95 (dd, 1H, J = 3.4, 1.0 Hz, H-4b), 5.90 (dd, 1H, J = 2.9, 1.6 Hz, H-2a), 5.86 (dd, 1H, J = 10.4, 8.0 Hz, H-2b), 5.82 (dd, 1H, J = 9.7, 2.7 Hz, H-4a), 5.79 (dd, 1H, J = 9.7, 2.9 Hz, H-3a), 5.66 (d, 1H, J = 1.6 Hz, H-1a), 5.54 (dd, 1H, J = 10.4, 3.4 Hz, H-3b), 4.91 (d, 1H, J = 8.0 Hz, H-1b), 4.94−4.88 (m, 1H, H-5a), 4.56 (dd, 1H, J = 11.3, 6.7 Hz, H-6b), 4.34 (dd, 1H, J = 11.3, 6.7 Hz, H-6′b), 4.22 (dd, 1H, J = 6.7, 1.0 Hz, H-5b), 4.17 (dd, 1H, J = 11.8, 1.9 Hz, H-6a), 4.04 (dd, 1H, J = 11.8, 6.1 Hz, H-6′a). ^{13}C{1H} NMR ($CDCl_3$, 100 MHz): δ 166.10, 165.72, 165.68, 165.57, 165.54, 165.49, 165.47 (COPh), 133.66, 133.63, 133.42, 133.38, 133.36, 133.25, 133.05, 132.15, 130.15, 130.10, 129.97, 129.93, 129.90, 129.88, 129.72, 129.49, 129.31, 129.27, 129.08, 129.01, 128.94, 128.77, 128.71, 128.63, 128.61, 128.42, 128.41, 128.27, 128.24 (C_6H_5), 101.95 (C-1b), 86.12 (C-1a), 71.85 (C-5a), 71.52 (C-5b), 72.00 (C-2a, C-3b), 70.45 (C-3a), 69.66 (C-2b), 69.02 (C-6a), 68.21 (C-4b), 67.45 (C-4a), 61.92 (C-6b). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{67}H_{54}O_{17}SNa$ 1185.2979; Found 1185.2961.

*Phenyl β-D-Galactopyranosyl-(1→6)-1-thio-α-D-mannopyranoside (**26**).* To a solution of **25** (77 mg, 0.084 mmol) in anhydrous MeOH (1 mL) was added MeONa (0.054 M in MeOH, 312 μL, 0.017 mmol). The mixture was stirred for 17 h at room temperature. Then the solution was neutralized by the addition of IR 120-H^+ form resin. After filtration, the filtrate was evaporated under reduced pressure. The resulting residue was suspended in water and the biphasic solution was washed with DCM (3×). The aqueous phase was finally freeze-dried to afford **26** (34 mg, 94%) as a powder. ^{1}H NMR (CD_3OD, 400 MHz): δ 7.58−7.50 (m, 2H, H_{arom}), 7.37−7.25 (m, 3H, H_{arom}), 5.39 (d, 1H, J = 1.5 Hz, H-1a), 4.28 (d, 1H, J = 7.7 Hz, H-1b), 4.22 (ddd, 1H, J = 9.7, 5.3, 2.2 Hz, H-5a), 4.13 (dd, 1H, J = 11.2, 2.3 Hz, H-6b), 4.09 (dd, 1H, J = 3.3, 1.5 Hz, H-2a), 3.86 (dd, 1H, J = 11.2, 5.4 Hz, H-6′b), 3.85−3.67 (m, 5H, H-3a, H-4a, H-6a, H-4b), 3.57 (dd, 1H, J = 9.7, 7.6 Hz, H-2b), 3.51−3.47 (m, 1H, H-5b), 3.46 (dd, 1H, J = 9.7, 3.4 Hz, H-3b). ^{13}C{1H} NMR (CD_3OD, 100 MHz): δ 135.5, 133.4, 130.2, 128.7 (C_6H_5), 105.3 (C-1b), 90.7 (C-1a), 76.8 (C-5b), 74.8 (C-3b), 74.6 (C-5a), 73.6 (C-2a), 73.0 (C-3a), 72.6 (C-2b), 70.3 (C-4a or C-4b), 69.8 (C-6a), 68.7 (C-4a or C-4b), 62.5 (C-6b). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{18}H_{26}O_{10}SNa$ 457.1139; Found 457.1143.

*β-D-Galactopyranosyl-(1→6)-α-D-mannopyranose (**PP-6**).* To a solution of **26** (53 mg, 0.122 mmol) in a mixture of ACN/H_2O 1:1 (10 mL) was added *N*-iodosaccharine (130 mg, 0.427 mmol). The mixture was stirred for 1 h at room temperature. Then the reaction mixture was concentrated under reduced pressure, and the residue was partitioned between DCM and water. The resulting organic layer was further extracted with water (2×). The combined aqueous layers were finally freeze-dried. The resulting solid was purified by column chromatography on silica gel (EtOAc/AcOH/H_2O/MeOH gradient from 9:0.5:0.5:0 up to 5:2:2:1) to afford **PP-6** (39 mg, 93%) as a solid.

PP-6*α* (major). ^{1}H NMR (D_2O, 400 MHz): δ 5.18 (d, 1H, J = 1.6 Hz, H-1a), 4.45 (d, 1H, J = 7.8 Hz, H-1b), 4.18 (dd, 1H, J = 11.4, 1.8 Hz, H-6a), 4.07−3.96 (m, 1H, H-5a), 3.94 (dd, 1H, J = 3.3, 1.6 Hz, H-2a), 3.93 (d, 1H, J = 3.3 Hz, H-4b), 3.91 (dd, 1H, J = 11.4, 5.9 Hz, H-6′a), 3.86 (dd, 1H, J = 9.7, 2.2 Hz, H-3a), 3.81 (dd, 1H, J = 11.6, 7.9 Hz, H-6b), 3.77 (o, 1H, H-6′b), 3.75 (t, 1H, J = 9.7 Hz, H-4a), 3.75−3.70 (m, 1H, H-5b), 3.66 (dd, 1H, J = 9.8, 3.3 Hz, H-3b), 3.57 (dd, 1H, J = 9.8, 7.8 Hz, H-2b). ^{13}C{1H} NMR (D_2O, 100 MHz): δ 103.3 (C-1b), 94.1 (C-1a), 75.1 (C-5b), 72.6 (C-3b), 71.3 (C-5a), 70.7 (C-2b), 70.6 (C-2a), 70.1 (C-3a), 68.8 (C-6a), 68.6 (C-4b), 66.5 (C-4a), 61.0 (C-6b). HRMS (ESI) m/z: $[M + Na]^+$ Calcd for $C_{12}H_{22}O_{11}Na$ 365.1054; Found 365.1058.

PP-6*β* (minor). ^{1}H NMR (D_2O, 400 MHz): δ 4.91 (t, 1H, J = 1.0 Hz, H-1a), 4.46 (d, 1H, J = 7.8 Hz, H-1b), 4.23 (dd, 1H, J = 11.2, 1.9 Hz, H-6a), 3.95 (dd, 1H, J = 3.3, 1.6 Hz, H-2a), 3.94 (o, 1H, H-4b), 3.85 (o, 1H, H-6′a), 3.77 (o, 2H, H-6b), 3.71 (o, 1H, H-5b), 3.71−3.66 (m, 2H, H-3a, H-4a), 3.66 (dd, 1H, J = 9.8, 3.3 Hz, H-3b), 3.57 (dd, 1H, J = 9.8, 7.8 Hz, H-2b), 3.55 (o, 1H, H-5a). ^{13}C{1H} NMR (D_2O, 100 MHz): δ 103.3 (C-1b), 93.7 (C-1a), 75.1 (C-5b), 75.0 (C-5a), 72.9 (C-3a), 72.6 (C-3b), 71.2 (C-2a), 70.7 (C-2b), 68.8 (C-6a), 68.6 (C-4b), 66.4 (C-4a), 61.0 (C-6b).

Ion Mobility−Mass Spectrometry. The synthesized disaccharides were analyzed on a Select Series Cyclic IMS (Waters, Wilmslow, UK), equipped with a cyclic Traveling Wave Ion Mobility Spectrometry (TWIMS) cell.[55] Arrival time distributions (ATDs) were recorded with the Quartz v. 6 software and processed with the MassLynx v. 4.2 software (both from Waters, Wilmslow, UK). ATDs are reported according to Gabelica et al.[67]

Samples were dissolved at 1 μg/mL in 50/50 methanol/water (v/v), doped with 0.5 mM NaCl, and then infused at 5 μL/min in the electrospray ion source. Samples were also analyzed at 1 μg/mL doped with 0.1 mM LiCl. All IM-MS analyses were performed in triplicates. 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 the Supporting Information.

The following sequence of events was used to perform IMS analyses at a TW height of 22 V and TW velocity of 375 m/s: (i)

https://doi.org/10.1021/acs.joc.1c00250
J. Org. Chem. 2021, 86, 6390−6405

injection in the IMS cell (10 ms); (ii) separation (see Table S1); and (iii) ejection to TOF analyzer (13.2 ms). As long as the ions remain in the fourth 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. ATDs were then calibrated to give CCSDs using MajorMix calibration solution (Waters, Wilmslow, UK)[68,69] (R^2 = 0.999).

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, 10 Hz, 13 mJ pulses in the 2700−3700 cm^{-1}. A mechanical shutter ensures the synchronization 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. The data shown are averaged three times.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.joc.1c00250.

^1H and ^{13}C NMR spectra of acceptors as well as galactofuranosyl- and galactopyranosyl-containing disaccharides, IM-MS and IRMPD technical descriptions, and MS/MS and IRMPD spectra of the eight disaccharides (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Vincent Ferrières − *Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France; Université de Lyon, CNRS, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Lyon, France; INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France; Univ Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France*; Email: vincent.ferrieres@ensc-rennes.fr

Authors

Bénédicte Favreau − *Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France; Université de Lyon, CNRS, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Lyon, France; INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France; Univ Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France*

Oznur Yeni − *Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France; Université de Lyon, CNRS, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Lyon, France; INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France; Univ Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France*

Simon Ollivier − *Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France; Université de Lyon, CNRS, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Lyon, France; INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France; Univ Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France*; orcid.org/0000-0002-7671-1736

Joël Boustie − *Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France; Université de Lyon, CNRS, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Lyon, France; INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France; Univ Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France*

Françoise Le Dévéhat − *Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France; Université de Lyon, CNRS, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Lyon, France; INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France; Univ Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France*

Jean-Paul Guégan − *Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France; Université de Lyon, CNRS, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Lyon, France; INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France; Univ Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France*

Mathieu Fanuel − *Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France; Université de Lyon, CNRS, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Lyon, France; INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France; Univ Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France*; orcid.org/0000-0001-8384-8266

Hélène Rogniaux − *Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France; Université de Lyon, CNRS, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Lyon, France; INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France; Univ Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France*; orcid.org/0000-0001-6083-2034

Richard Brédy − *Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France; Université de Lyon, CNRS, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Lyon, France; INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France;*

Univ Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France

Isabelle Compagnon − *Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France; Université de Lyon, CNRS, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Lyon, France; INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France; Univ Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France;* orcid.org/0000-0003-2994-3961

David Ropartz − *Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France; Université de Lyon, CNRS, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Lyon, France; INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France; Univ Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France;* orcid.org/0000-0003-4767-6940

Laurent Legentil − *Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France; Université de Lyon, CNRS, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Lyon, France; INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France; Univ Rennes, CNRS, ISCR − UMR 6226, F-35000 Rennes, France;* orcid.org/0000-0003-1402-150X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.joc.1c00250

Author Contributions

[§]O.Y. and S.O. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank the Agence Nationale de la Recherche (ANR) for financially supporting 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) is also acknowledged for recording ^{1}H and ^{13}C spectra at 500 and 125 MHz, respectively.

■ REFERENCES

(1) Allan, B. B.; Balch, W. E. Protein Sorting by Directed Maturation of Golgi Compartments. *Science* **1999**, *285*, 63−66.

(2) Chen, J.; Liu, T.; Gao, J.; Gao, L.; Zhou, L.; Cai, M.; Shi, Y.; Xiong, W.; Jiang, J.; Tong, T.; Wang, H. Variation in Carbohydrates between Cancer and Normal Cell Membranes Revealed by Super-Resolution Fluorescence Imaging. *Adv. Sci.* **2016**, *3*, 1600270.

(3) Ahamefula Osibe, D.; Lei, S.; Wang, B.; Jin, C.; Fang, W. Cell wall polysaccharides from pathogenic fungi for diagnosis of fungal infectious disease. *Mycoses* **2020**, *63*, 644−652.

(4) Gray, C. J.; Migas, L. G.; Barran, P. E.; Pagel, K.; Seeberger, P. H.; Eyers, C. E.; Boons, G.-J.; Pohl, N. L. B.; Compagnon, I.; Widmalm, G. r.; Flitsch, S. L. Advancing Solutions to the Carbohydrate Sequencing Challenge. *J. Am. Chem. Soc.* **2019**, *141*, 14463−14479.

(5) Wu, X.; Delbianco, M.; Anggara, K.; Michnowicz, T.; Pardo-Vargas, A.; Bharate, P.; Sen, S.; Pristl, M.; Rauschenbach, S.; Schlickum, U.; Abb, S.; Seeberger, P. H.; Kern, K. Imaging single glycans. *Nature* **2020**, *582*, 375−378.

(6) Gerlits, O. O.; Coates, L.; Woods, R. J.; Kovalevsky, A. Mannobiose Binding Induces Changes in Hydrogen Bonding and Protonation States of Acidic Residues in Concanavalin A as Revealed by Neutron Crystallography. *Biochemistry* **2017**, *56*, 4747−4750.

(7) Manzoni, F.; Wallerstein, J.; Schrader, T. E.; Ostermann, A.; Coates, L.; Akke, M.; Blakeley, M. P.; Oksanen, E.; T. Logan, D. Elucidation of Hydrogen Bonding Patterns in Ligand-Free, Lactose- and Glycerol-Bound Galectin-3C by Neutron Crystallography to Guide Drug Design. *J. Med. Chem.* **2018**, *61*, 4412−4420.

(8) Gimeno, A.; Valverde, P.; Arda, A.; Jimenez-Barbero, J. Glycan structures and their interactions with proteins. A NMR view. *Curr. Opin. Struct. Biol.* **2020**, *62*, 22−30.

(9) Aizpurua-Olaizola, O.; Torano, J. S.; Falcon-Perez, J. M.; Williams, C.; Reichardt, N.; Boons, G.-J. Mass spectrometry for glycan biomarker discovery. *TrAC, Trends Anal. Chem.* **2018**, *100*, 7−14.

(10) Marchetti, R.; Perez, S.; Arda, A.; Imberty, A.; Jimenez-Barbero, J.; Silipo, A.; Molinaro, A. Rules of Engagement" of Protein-Glycoconjugate Interactions:AMolecular View Achievable by using NMR Spectroscopy and Molecular Modeling. *ChemistryOpen* **2016**, *5*, 274−296.

(11) Park, S.-J.; Lee, J.; Qi, Y.; Kern, N. R.; Lee, H. S.; Jo, S.; Joung, I.; Joo, K.; Lee, J.; Im, W. CHARMM-GUI Glycan Modeler for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology* **2019**, *29*, 320−331.

(12) Kumozaki, S.; Sato, K.; Sakakibara, Y. A Machine Learning Based Approach to de novo Sequencing of Glycans from Tandem Mass Spectrometry Spectrum. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **2015**, *12*, 1267−1274.

(13) Li, F.; Li, C.; Wang, M.; Webb, G. I.; Zhang, Y.; Whisstock, J. C.; Song, J. GlycoMine: a machine learning-based approachfor predicting N-, C- and O-linked glycosylationin the human proteome. *Bioinformatics* **2015**, *31*, 1411−1419.

(14) Bojar, D.; Camacho, D. M.; Collins, J. J. Using Natural Language Processing to Learn the Grammarof Glycans. *bioRxiv* **2020**, DOI: 10.1101/2020.01.10.902114doi.

(15) Peltier, P.; Euzen, R.; Daniellou, R.; Nugier-Chauvin, C.; Ferrières, V. Recent knowledge and innovations related to hexofuranosides: structure, synthesis and applications. *Carbohydr. Res.* **2008**, *343*, 1897−1923.

(16) Richards, M. R.; Lowary, T. L. Chemistry and Biology of Galactofuranose-Containing Polysaccharides. *ChemBioChem* **2009**, *10*, 1920−1938.

(17) Marino, C.; Rinflerch, A.; de Lederkremer, R. M. Galactofuranose antigens, a target for diagnosis of fungal infections in humans. *Future Sci.* **2017**, *3*, FSO199.

(18) Arbatsky, N. P.; Shashkov, A. S.; Mamyan, S. S.; Knirel, Y. A.; Zych, K.; Sidorczyk, Z. Structure of the *O*-specific polysaccharide of a serologically separate Proteus penneri strain 22. *Carbohydr. Res.* **1998**, *310*, 85−90.

(19) Michael, F. S.; Szymanski, C. M.; Li, J.; Chan, K. H.; Khieu, N. H.; Larocque, S.; Wakarchuk, W. W.; Brisson, J. R.; Monteiro, M. A. The structures of the lipooligosaccharide and capsule polysaccharide of Campylobacter jejuni genome sequenced strain NCTC 11168. *Eur. J. Biochem.* **2002**, *269*, 5119−5136.

(20) Gorin, P. A. J. Structural diversity of D-galacto-D-mannan components isolated from lichens having ascomycetous mycosymbionts. *Carbohydr. Res.* **1985**, *142*, 253−267.

(21) Iacomini, M.; Schneider, C. L.; Gorin, P. A. J. Comparative studies on the polysaccharides of Cladonia alpestris (reindeer moss), Cladonia confusa, and Cladonia amaurocraea. *Carbohydr. Res.* **1985**, *142*, 237−251.

(22) Omarsdottir, S.; Petersen, B. O.; Paulsen, B. S.; Togola, A.; Duus, J. O.; Olafsdottir, E. S. Structural characterisation of novel lichen heteroglycans by NMR spectroscopy and methylation analysis. *Carbohydr. Res.* **2006**, *341*, 2449−2455.

(23) Carbonero, E. R.; Smiderle, F. R.; Gracher, A. H. P.; Mellinger, C. G.; Torri, G.; Ahti, T.; Gorin, P. A. J.; Iacomini, M. Structure of two glucans and a galactofuranomannan from the lichen Umbilicaria mammulata. *Carbohydr. Polym.* **2006**, *63*, 13−18.

(24) Omarsdottir, S.; Freysdottir, J.; Olafsdottir, E. S. Immunomodulating polysaccharides from the lichen Thamnolia vermicularis var. subuliformis. *Phytomedicine* **2007**, *14*, 179−184.
(25) Das, S. K.; Roy, N. Synthesis of the tetrasaccharide repeating unit of the antigen from Klebsiella type 20. *Carbohydr. Res.* **1995**, *271*, 177−183.
(26) Zdorovenko, E. L.; Varbanets, L. D.; Zatonsky, G. V.; Kachala, V. V.; Zdorovenko, G. M.; Shashkov, A. S.; Knirel, Y. A. Structure of the O-specific polysaccharide of the lipopolysaccharide of Rahnella aquatilis 95 U003. *Carbohydr. Res.* **2008**, *343*, 2494−2497.
(27) Mu, W.; Li, Q.; Fan, C.; Zhou, C.; Jiang, B. Recent advances on physiological functions and biotechnological production of epilactose. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 1821−1827.
(28) Sizova, O. V.; Ross, A. J.; Ivanova, I. A.; Borodkin, V. S.; Ferguson, M. A. J.; Nikolaev, A. V. Probing elongating and branching b-D-galactosyltransferase activities in Leishmania parasites by making use of synthetic phosphoglycans. *ACS Chem. Biol.* **2011**, *6*, 648−657.
(29) Pol-Fachin, L.; Serrato, R. V.; Verli, H. Solution conformation and dynamics of exopolysaccharides from Burkholderia species. *Carbohydr. Res.* **2010**, *345*, 1922−1931.
(30) Geyer, H.; Wuhrer, M.; Resemann, A.; Geyer, R. Identification and Characterization of Keyhole Limpet Hemocyanin N-Glycans Mediating Cross-reactivity with Schistosoma mansoni. *J. Biol. Chem.* **2005**, *280*, 40731−40748.
(31) Garegg, P. J.; Henrichson, C.; Norberg, T.; Ossowski, P. Syntheses of trisaccharides related to Salmonella serogroup E O-antigenic polysaccharides. *Carbohydr. Res.* **1983**, *119*, 95−100.
(32) Favreau, B. *Chem. Rxiv* **2021**, DOI: 10.26434/chemrxiv.13536767.v1.
(33) Hofmann, J.; Pagel, K. Glycan Analysis by Ion Mobility-Mass Spectrometry. *Angew. Chem., Int. Ed.* **2017**, *56*, 8342−8349.
(34) Warnke, S.; Faleh, A. B.; Pellegrinelli, R. P.; Yalovenko, N.; Rizzo, T. R. Combining Ultra-High Resolution Ion Mobility Spectrometry with Cryogenic IR Spectroscopy for the Study of Biomolecular Ions. *Faraday Discuss.* **2019**, *217*, 114−125.
(35) Gray, C. J.; Compagnon, I.; Flitsch, S. L. Mass Spectrometry Hybridized with Gas-Phase InfraRed Spectroscopy for Glycan Sequencing. *Curr. Opin. Struct. Biol.* **2020**, *62*, 121.
(36) Gabelica, V.; Marklund, E. Fundamentals of Ion Mobility Spectrometry. *Curr. Opin. Chem. Biol.* **2018**, *42*, 51−59.
(37) Hofmann, J.; Hahm, H. S.; Seeberger, P. H.; Pagel, K. Identification of Carbohydrate Anomers Using Ion Mobility-Mass Spectrometry. *Nature* **2015**, *526*, 241−244.
(38) Hernandez, O.; Isenberg, S.; Steinmetz, V.; Glish, G. L.; Maitre, P. Probing Mobility-Selected Saccharide Isomers: Selective Ion−Molecule Reactions and Wavelength-Specific IR Activation. *J. Phys. Chem. A* **2015**, *119*, 6057−6064.
(39) Polfer, N. C.; Oomens, J. Vibrational spectroscopy of bare and solvated ionic complexes of biological relevance. *Mass Spectrom. Rev.* **2009**, *28*, 468−494.
(40) Schindler, B.; Legentil, L.; Allouche, A.-R.; Ferrières, V.; Compagnon, I. Spectroscopic diagnostic for the ring-size of carbohydrates in the gas phase: furanose and pyranose forms of GalNAc. *Phys. Chem. Chem. Phys.* **2019**, *21*, 12460−12467.
(41) Mookherjee, A.; Uppal, S. S.; Murphree, T. A.; Guttman, M. Linkage Memory in Underivatized Protonated Carbohydrates. *J. Am. Soc. Mass Spectrom.* **2021**, *32*, 581−589.
(42) Pellegrinelli, R. P.; Yue, L.; Carrascosa, E.; Warnke, S.; Ben Faleh, A.; Rizzo, T. R. How General Is Anomeric Retention during Collision-Induced Dissociation of Glycans? *J. Am. Chem. Soc.* **2020**, *142*, 5948−5951.
(43) Schindler, B.; Barnes, L.; Renois, G.; Gray, C.; Chambert, S.; Fort, S.; Flitsch, S.; Loison, C.; Allouche, A. R.; Compagnon, I. Anomeric Memory of the Glycosidic Bond upon Fragmentation and Its Consequences for Carbohydrate Sequencing. *Nat. Commun.* **2017**, *8*, 973.
(44) Weber, H. K.; Weber, H.; Kazlauskas, R. J. 'Watching' lipase-catalyzed acylations using1HNMR:competing hydrolysis of vinyl acetate in dry organic solvents. *Tetrahedron: Asymmetry* **1999**, *10*, 2635−2638.
(45) Taylor, M. S. Catalysis Based on Reversible Covalent Interactions of Organoboron Compounds. *Acc. Chem. Res.* **2015**, *48*, 295−305.
(46) Pedersen, C. M.; Olsen, J.; Brka, A. B.; Bols, M. Quantifying the Electronic Effects of Carbohydrate Hydroxy Groups by Using Aminosugar Models. *Chem. - Eur. J.* **2011**, *17*, 7080−7086.
(47) Legentil, L.; Cabezas, Y.; Tasseau, O.; Tellier, C.; Daligault, F.; Ferrières, V. Regioselective galactofuranosylation for the synthesis of disaccharides patterns found in pathogenic microorganisms. *J. Org. Chem.* **2017**, *82*, 7114−7122.
(48) Mandal, P. K.; Misra, A. K. Mild and efficient hydrolysis of thioglycosides to glycosyl hemiacetals using N-iodosaccharin 1. *Synlett* **2007**, *2007*, 1207−1210.
(49) Gorin, P. A. J.; Barreto-Bergter, E. M.; Cruz, F. S. The chemical structure of D-galacto-D-mannan component of Trypanosoma cruzi: 13C-N.M.R. Shift dependence on structure of D-galactose to D-mannose linkage. *Carbohydr. Res.* **1981**, *88*, 177−188.
(50) Marino, C.; Chiocconi, A.; Varela, O.; de Lederkremer, R. M. The glycosyl-aldonolactone approach for the synthesis of -D-Gal*f*-(1,3)-D-Man*p* and 3-deoxy- -D-xylo-hexofuranosyl-(1,3)-D-Man*p*. *Carbohydr. Res.* **1998**, *311*, 183−189.
(51) Krylov, V. B.; Argunov, D. A.; Solovev, A. S.; Petruk, M. I.; Gerbst, A. G.; Dmitrenok, A. S.; Shashkov, A. S.; Latgé, J.-P.; Nifantiev, N. E. Synthesis of oligosaccharides related to galactomannans from Aspergillus fumigatus and their NMR spectral data. *Org. Biomol. Chem.* **2018**, *16*, 1188−1199.
(52) Lankalapalli, R. S.; Baksa, A.; Liliom, K.; Bittman, R. Synthesis and Properties of a Photoactivatable Analogue of Psychosine (*β*-Galactosylsphingosine). *ChemMedChem* **2010**, *5*, 682−686.
(53) Huang, S.; Yu, H.; Chen, X. Disaccharides as Sialic Acid Aldolase Substrates: Synthesis of Disaccharides Containing a Sialic Acid at the Reducing End. *Angew. Chem., Int. Ed.* **2007**, *46*, 2249−2253.
(54) Yu, H.; Chen, X. Aldolase-Catalyzed Synthesis of -D-Galp-(1,9)-D-KDN: A Novel Acceptor for Sialyltransferases. *Org. Lett.* **2006**, *8*, 2393−2396.
(55) Giles, K.; Ujma, J.; Wildgoose, J.; Pringle, S.; Richardson, K.; Langridge, D.; Green, M. A Cyclic Ion Mobility-Mass Spectrometry System. *Anal. Chem.* **2019**, *91*, 8564−8573.
(56) Cho, E.; Riches, E.; Palmer, M.; Giles, K.; Ujma, J.; Kim, S. Isolation of Crude Oil Peaks Differing by m/z U0.1 via Tandem Mass Spectrometry Using a Cyclic Ion Mobility-Mass Spectrometer. *Anal. Chem.* **2019**, *91*, 14268−14274.
(57) Ropartz, D.; Fanuel, M.; Ujma, J.; Palmer, M.; Giles, K.; Rogniaux, H. Structure Determination of Large Isomeric Oligosaccharides of Natural Origin through Multipass and Multistage Cyclic Traveling-Wave Ion Mobility Mass Spectrometry. *Anal. Chem.* **2019**, *91*, 12030−12037.
(58) Ujma, J.; Ropartz, D.; Giles, K.; Richardson, K.; Langridge, D.; Wildgoose, J.; Green, M.; Pringle, S. Cyclic Ion Mobility Mass Spectrometry Distinguishes Anomers and Open-Ring Forms of Pentasaccharides. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 1028−1037.
(59) Richards, M. R.; Bai, Y.; Lowary, T. L. Comparison between DFT- and NMR-based conformational analysis of methyl galactofuranosides. *Carbohydr. Res.* **2013**, *374*, 103−114.
(60) Taha, H. A.; Richards, M. R.; Lowary, T. L. Conformational Analysis of Furanoside-Containing Mono- and Oligosaccharides. *Chem. Rev.* **2013**, *113*, 1851−1876.
(61) Kailemia, M. J.; Ruhaak, L. R.; Lebrilla, C. B.; Amster, I. J. Oligosaccharide Analysis by Mass Spectrometry: A Review of Recent Developments. *Anal. Chem.* **2014**, *86*, 196−212.
(62) Schindler, B.; Barnes, L.; Gray, C. J.; Chambert, S.; Flitsch, S. L.; Oomens, J.; Daniel, R.; Allouche, A. R.; Compagnon, I. IRMPD Spectroscopy Sheds New (Infrared) Light on the Sulfate Pattern of Carbohydrates. *J. Phys. Chem. A* **2017**, *121*, 2114−2120.
(63) Schindler, B.; Joshi, J.; Allouche, A.-R.; Simon, D.; Chambert, S.; Brites, V.; Gaigeot, M.-P.; Compagnon, I. Distinguishing isobaric

phosphated and sulfated carbohydrates by coupling of mass spectrometry with gas phase vibrational spectroscopy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 22131−22138.
(64) Marino, C.; Varela, O.; de Lederkremer, R. M. Reactions of per-O-benzoyl-β-D-Galf isothiocyanate, a chiral resolving agent. *Tetrahedron* **1997**, *53*, 16009−16016.
(65) Danalev, D.; Legentil, L.; Daniellou, R.; Nugier-Chauvin, C.; Ferrières, V. Direct access to new β-D-galactofuranoconjugates: application to the synthesis of galactofuranosyl-L-cysteine and L-serine. *Tetrahedron Lett.* **2011**, *52*, 1121−1123.
(66) Christensen, M.; Meldal, M.; Bock, K. J. Synthesis of mannose 6-phosphate-containing disaccharide threonine building blocks and their use in solid-phase glycopeptide synthesis. *J. Chem. Soc., Perkin Trans. 1* **1993**, 1453−1460.
(67) Gabelica, V.; Shvartsburg, A. A.; Afonso, C.; Barran, P.; Benesch, J. L. P.; Bleiholder, C.; Bowers, M. T.; Bilbao, A.; Bush, M. F.; Campbell, J. L.; Campuzano, I. D. G.; Causon, T.; Clowers, B. H.; Creaser, C. S.; Pauw, E. D.; Far, J.; Fernandez-Lima, F.; Fjeldsted, J. C.; Giles, K.; Groessl, M.; Jr, C. J. H.; Hann, S.; Kim, H. I.; Kurulugama, R. T.; May, J. C.; McLean, J. A.; Pagel, K.; Richardson, K.; Ridgeway, M. E.; Rosu, F.; Sobott, F.; Thalassinos, K.; Valentine, S. J.; Wyttenbach, T. Recommendations for reporting ion mobility Mass Spectrometry measurements. *Mass Spectrom. Rev.* **2019**, *38*, 291−320.
(68) Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V.; Ruotolo, B. T. Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. *Anal. Chem.* **2010**, *82*, 9557−9565.
(69) Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. Ion Mobility-Mass Spectrometry Analysis of Large Protein Complexes. *Nat. Protoc.* **2008**, *3*, 1139−1152.

https://doi.org/10.1021/acs.joc.1c00250
J. Org. Chem. 2021, 86, 6390−6405

Chapter 5 – Investigating the behavior of furanosides upon collision-induced dissociation

A. Context

In the context of the ALGAIMS project, the desired outcome would be to find signatures of the presence of Gal*f* motifs in oligosaccharides from biological sources (such as lichens) using two different analytical strategies (*i.e.*, cIMS and IRMPD) that rely on a library of reference/synthetic disaccharides containing Gal*f*. As a first step of the process, we have established in **Chapter 4** that cIMS allowed to discriminate between furanose/pyranose tautomers and linkage regioisomers in the reference galactomannan disaccharides. In a second step, it was necessary to study synthetic standards of higher DPs and verify that the signature established for disaccharides was preserved upon fragmentation and that it could be used to identify Gal*f* in these larger structures—regardless of the branching type of larger oligosaccharides.

In this Chapter, we thus studied linear and branched trisaccharides of the type $(Gal)_1(Man)_2$. They were synthesized by our partners of the ALGAIMS project and contained, on the one hand, Gal-(β1,4)-Man and Man-(α1,6)-Man linkages and on the other hand, Gal-(β1,6)-Man and Man-(α1,4)-Man linkages (**Figure 48**).The galactofuranosyl residue in these structures raises specific questions. Indeed, Gal*f* are supposedly less stable than their Gal*p* counterparts.[138] This might result in unexpected behavior of galactofuranosides when submitted to gas-phase activation, the knowledge of which would be required to identify the disaccharide signatures in larger oligosaccharides.

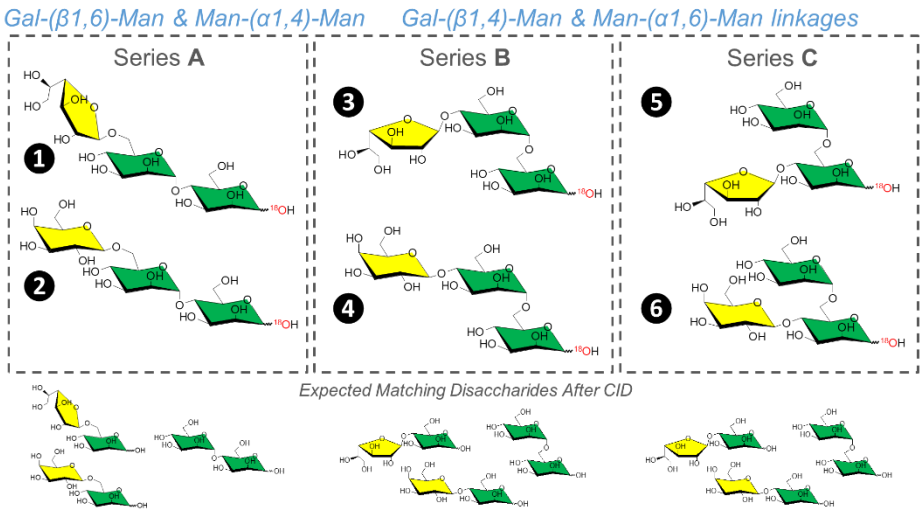


Figure 48. Structures of the studied trisaccharides and reference disaccharides

B. Objectives

In this work, our aim was to explore the fragmentation patterns of Gal*f*-containing trisaccharides, and compare them to their pyranoside counterparts, with two main objectives: (i) check if the DP2-like product ions generated upon CID of these oligosaccharides match the ion mobility of the corresponding disaccharide standards (*i.e.*, if the mobility profile of the fragments matched the distribution observed in **Chapter 4**), and (ii) conduct an in-depth exploration of the arrival time distributions to identify possible gas-phase rearrangements.

C. Summary of the results

To properly distinguish the fragments containing the reducing end of the precursor trisaccharide, the samples were labeled by exchanging the reducing hydroxyl group (^{16}O) with heavy oxygen (^{18}O). The exchange naturally occurs through the mutarotation phenomenon by dissolving the sample in $H_2{}^{18}O$. This induces a +2 Da shift on fragments coming from the reducing end, affording easy discrimination of these fragments with MS.

To start with, we applied the fragmentation strategy to linear trisaccharides. *In most cases, the Gal-Man (and Man-Man) fragments matched well with the corresponding disaccharides.* However, in the case of the Gal*f*-(β1,4)-Man-(α1,6)-Man trisaccharide, we also observed additional peaks in the IM profiles of the fragments. In particular, the *m/z* 351 fragment of this trisaccharide should have presented a single IM peak at a drift time of 22.0 ms, but in fact contained three more peaks (t_d = 20.7, 21.5, and 22.8 ms), as shown in **Figure 49**.

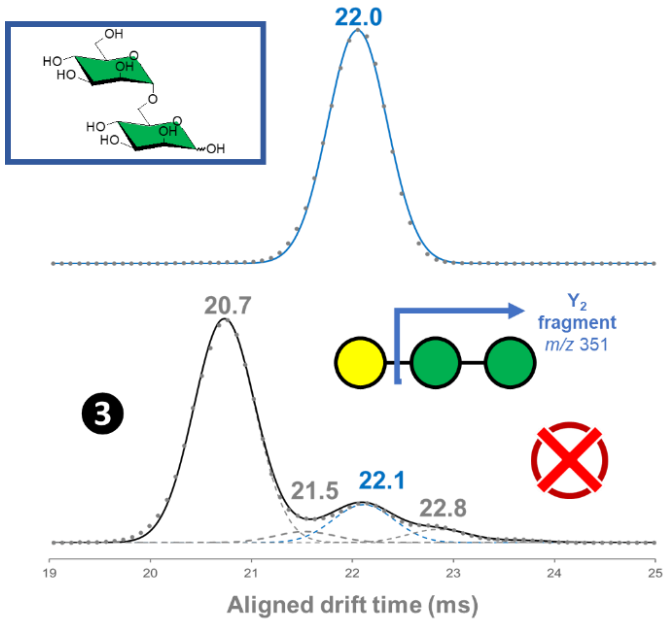


Figure 49. Reference profile of Man-(α1,6)-Man (top) and IM profile of the fragment at *m/z* 351 from Gal*f*-(β1,4)-Man-(α1,6)-Man (bottom)

An in-depth study showed that the additional peaks could be explained through different mechanisms:

- The first two peaks, at 21.5 and 22.8 ms, attested of a structural rearrangement corresponding to a migration of the Gal*f* residue to C-4 of the reducing Man, with an internal residue loss (IRL) of the internal Man subunit. As presented in **Figure 50,** the doublet of peaks is indeed specific to the Gal*f*-(β1,4)-Man*p* standard. The IRL was confidently identified thanks to the ^{18}O-labeling which unequivocally distinguishes the reducing end.

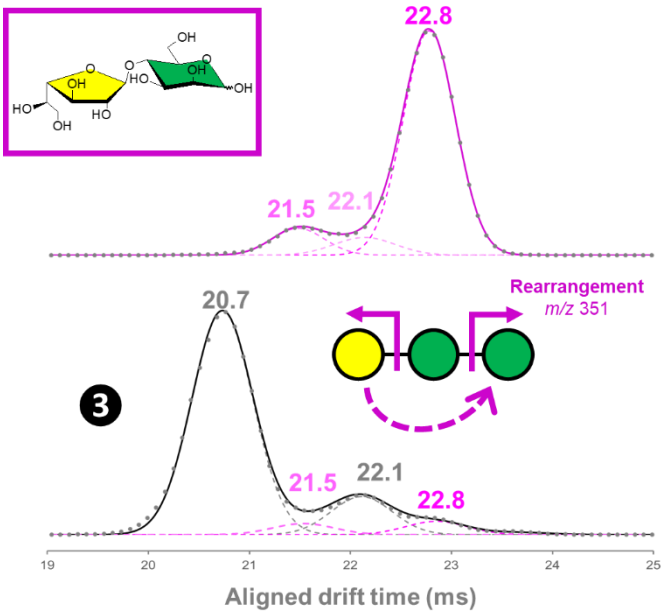


Figure 50. Reference profile of Gal*f*-(β1,4)-Man (top) and IM profile of the fragment at *m/z* 351 from Gal*f*-(β1,4)-Man-(α1,6)-Man (bottom)

- The remaining peak, at 20.7 ms, despite being the main peak of this IM profile, did not match any disaccharide standard in our library. We thus conducted an extensive study to identify this feature. First, we performed a multistage IMS study. We notably used a slicing experiment to select both the expected Y_2 fragment (t_d = 22.1 ms), as well as the unknown peak at t_d = 20.7 ms. We observed that, when activating the ions upon reinjection from the pre-store cell, there was an interconversion between the two populations (induced by collisional activation, see **Figure 51**). This behavior indicates that the population found at 20.7 ms actually consists of one or more conformers of Man-(α1,6)-Man.

Figure 51. Interconversion between the two conformational populations of Man-(α1,6)-Man, as observed using multistage IMS

Finally, we attempted to further validate this hypothesis of conformers, but also attempted to identify the conformers and explain why we observed a peak at 20.7 instead of 22 ms when fragmenting Gal*f*-(β1,4)-Man-(α1,6)-Man. To this end, we performed molecular dynamics (MD) at the PM7 level of theory and identified different conformational families. Finally, we performed theoretical CCS calculations (using the Trajectory Method, Lennard Jones-Induced model) for representative structures from our MD analysis. By comparing these calculated CCSs to the calibrated experimental IM profile (**Figure 52**), we evidenced that the expected conformer for the Y_2 fragment corresponded to the Li^+ counter-ion coordinating with the non-reducing Man subunit, which is occupied by the Gal*f* residue in the studied trisaccharide, explaining why the other conformers are favored.

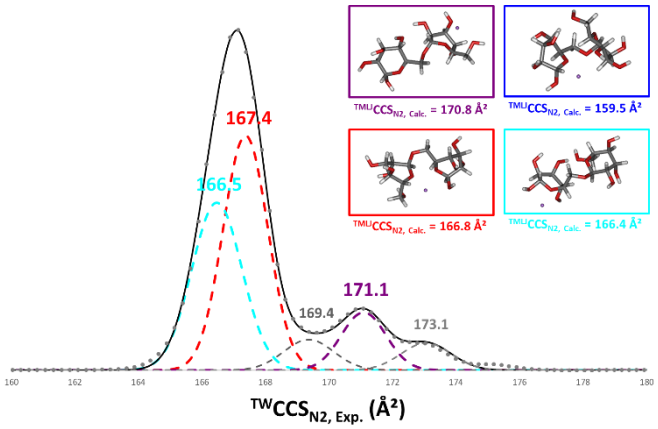


Figure 52. Comparison of the experimental CCS profile of the Y_2 fragment from Gal*f*-(β1,4)-Man-(α1,6)-Man to the conformers calculated by MD analysis of Man-(α1,6)-Man

D. Conclusions

In this Chapter, we have shown that assigning the structure of CID-induced fragments from Gal*f*-containing oligosaccharides using IMS and disaccharide standards is not always straightforward—but still possible.

Notably, our data suggest the occurrence of two phenomena that can happen in the gas phase upon activation and complicate the structural characterization of Gal*f*-containing oligosaccharides. First, we observed a migration of the Gal*f* subunit on linear trisaccharides, with an internal residue loss of a Man subunit. Secondly, we showed that the presence of a Gal*f* at the non-reducing end could also affect the conformation of the mannose backbone to some degree, because of differences in cation coordination. This conformational change is not observed on the Gal*p* counterpart structure in our experimental conditions, indicating that this phenomenon is likely induced by the Gal*f* moiety being more flexible and/or less stable compared to Gal*p*.

E. Scientific Communication

These results will be published in an upcoming article presented as **Publication 2**. The supporting data for this publication are provided as **Appendix 4**. These results have also been presented as a poster in a national conference.

Article in preparation

Ollivier, S.; Legentil, L.; Yeni, O.; David, L.-P.; Ferrières, V.; Compagnon, I.; Rogniaux, H.; Ropartz, D. Gas-phase Behavior Of Galactofuranosides Upon Collisional Fragmentation – A Multistage High-Resolution Ion Mobility Study.

Poster presentation

Ollivier, S.; Legentil, L.; Fanuel, M.; David, L.-P.; Yeni, O.; Compagnon, I.; Ferrières, V.; Ropartz, D.; Rogniaux, H. On the behavior of Gal*f*-containing oligosaccharides upon fragmentation – a multistage high-resolution Ion Mobility study. *Analytics 2022*, Nantes, FR, Sept. 5–8, 2022.

Publication 2 – Gas-Phase Behavior of Galactofuranosides upon Collisional Fragmentation – a Multistage High-Resolution Ion Mobility Study

Simon Ollivier, Laurent Legentil, Oznur Yeni, Louis-Philippe David, Vincent Ferrières, Isabelle Compagnon, Hélène Rogniaux, and David Ropartz*

* Corresponding author

A work in progress

Gas-Phase Behavior of Galactofuranosides upon Collisional Fragmentation – a Multistage High-Resolution Ion Mobility Study

Simon Ollivier,[1,2] Laurent Legentil,[3] Oznur Yeni,[4] Louis-Philippe David,[3] Vincent Ferrières,[3] Isabelle Compagnon,[4] Hélène Rogniaux[1,2] and David Ropartz[1,2*]

[1] INRAE, UR BIA, F-44316 Nantes, France

[2] INRAE, BIBS Facility, F-44316 Nantes, France

[3] Univ Rennes, Ecole Nationale Supérieure de Chimie de Rennes, CNRS, ISCR – UMR 6226, F-35000 Rennes, France

[4] Univ Lyon, Université Claude Bernard Lyon 1, CNRS, Institut Lumière Matière, F-69622 Villeurbanne, France

ABSTRACT: Carbohydrates are ubiquitous in Nature but are among the least conserved biomolecules in Life. These biopolymers pose a particular challenge to analytical chemists because of their high diversity and structural heterogeneity. In addition, they contain many isomerisms that complicate their structural characterization, notably by mass spectrometry. The tautomerism of the constitutive subunits is of particular interest. A given cyclized monosaccharide unit can take two forms: a most common 6-membered ring (pyranose, *p*) and a more flexible 5-membered ring (furanose, *f*). The tautomers impact the biological properties of polysaccharides, resulting in interesting properties of the derived oligosaccharides. From an analytical view, the impact of tautomerism on the gas-phase behavior of ions has scarcely been described in the literature. In this work, we study the behavior of Gal*f*-containing oligosaccharides, ionized as $[M+Li]^+$ species, under collisional dissociation (CID) conditions using high-resolution and multistage ion mobility (IM) on a Cyclic IM platform. In the first part of this work, we studied whether disaccharidic fragments released from Gal*f*-containing $(Gal)_1(Man)_2$ trisaccharides (and their Gal*p* counterpart) would match the corresponding disaccharide standards, and—despite the fragments generally being a good match—we showed the possibility of Gal*f* migrations, and other unidentified alterations in the IM profile. Next, we expended on these unknown features using multistage IMS and molecular dynamics, unveiling the contributions of additional gas-phase conformers in the profile of fragments from a Gal*f*-containing trisaccharide compared to the corresponding disaccharides.

INTRODUCTION

Among other isomerisms of carbohydrates, the furanose/pyranose tautomerism is particularly intriguing. This specific case of tautomerism consists of a given monosaccharide forming either a six-membered ring (pyranose) or a five-membered ring (furanose). In Nature, hexoses and their derivatives most often adopt a pyranose configuration because it is more stable than the furanose form.[1] Yet there are some exceptions, notably with the occurrence of galactofuranose(Gal*f*)-containing polysaccharides. From a biological standpoint, it is interesting to understand why some organisms (including fungi, bacteria and mycobacteria) biosynthesize hexoses in the form of five-membered rings, which are not thermodynamically favored. In addition, there is increasing interest in finding new applications for carbohydrates and Gal*f*-containing carbohydrates are no exception to this general trend. Gal*f* has been reported across the tree of Life in many harmless species/genera such as *Penicillium*, *Agelas* (sponge) or *Rocella* (lichen) but also in highly infectious pathogens like *Mycobacterium tuberculosis*, *Trypanosoma cruzi* or *Shigella dysenteriae*.[2] For such pathogenic organisms, Gal*f*-containing components (including the biosynthetic pathways and involved enzymes) have been targeted to develop new therapeutic strategies.[3] However, a major obstacle to a further exploitation and/or to better understand the biological role of Gal*f*-containing carbohydrates lies in the ability to accurately screen and identify such structures in biological matrices.

Mass spectrometry (MS) offers attractive characteristics to that aim, such as the sensitivity, the selectivity and the quality of the structural information brought by MS and tandem MS experiments. However, the method alone is blind to many kinds of isomerisms, including the furanose/pyranose tautomerism. Gas-phase methods that are sensitive to isomerism—such as Ion Mobility-Mass Spectrometry (IM-MS)[4–17] and InfraRed Multiple Photon Dissociation (IRMPD) spectroscopy[4,18–23]—have been reported recently as promising complements to MS. In the study, both methods achieved a specific signature for Gal*p* or Gal*f* from a library of eight synthetic $(Gal)_1(Man)_1$ disaccharides, composed of all Gal*f*/Gal*p* pairs for the four linkage regioisomerism of the Gal-Man linkage. This proof-of-concept work proposed that both IM-MS and IRMPD could be considered to specifically detect Galf-containing motifs in larger naturally occurring polysaccharides by using a bottom-up strategy involving fragmentation approaches such as CID (collision-induced dissociation).

However, relatively few works have been conducted on the gas-phase behavior of furanosides upon slow thermal activation, such as CID or IRMPD, although this seems an important prerequisite to the validity of the previous strategy. Works by Schindler et al.[24] and Ho et al.[25] have studied the behavior of, respectively, Gal*f*NAc and Gal*f* compared to their pyranose counterparts using IRMPD spectroscopy. Questions remain, however, as to how furanosides, which are less thermodynamically stable than pyranosides, will behave upon slow thermal activation, or whether specific rearrangements in the gas phase can be triggered by activation.

In the present study, we thus aimed at exploring in more depth the gas-phase behavior of Gal*f*-containing trisaccharides by means of ion mobility spectrometry (IMS). Briefly, IMS is a gas-phase method that separates ions according to their interaction with a buffer gas and an electric field; the resulting separation is ultimately related to the tridimensional structure of the ions. As detailed reviews are available elsewhere (see [26–28]), we will not detail all existing IMS methods here. However, one method of particular interest is Traveling Wave IMS (TWIMS) that separates ions using symmetric potential waves in a stationary gas, affording separations over long path length staying under the low-field limit. This is of notable interest because the resolving power (Rp) of TWIMS is a direct function of the square root of the ion path length. Extending the drift path length by cycling the ions for multiple passes around a racetrack has thus recently become a popular strategy to reach ultrahigh-resolution IMS, using Cyclic IMS (Waters Corp., Wilmslow, UK)[29] or Structures for Lossless Ion Manipulation.[30] In particular, the Cyclic IMS (cIMS) technology is equipped with a multifunction array to redirect ions in a given direction affording multi-stage IMS (IMS^n) manipulations including precursor selection, collision-induced dissociation and IMS analysis of the fragments—analogously to MS^n. This platform thus offers many advantages to the study of furanosides in the gas phase and was used in the present work to compare the fragmentation behavior of Gal*f*-trisaccharides to their Gal*p* counterparts.

To this end, we synthesized and analyzed galactomannan trisaccharides from two selected structural families of biological relevance. The first type of biologically relevant Gal-Man linkage consists of a Man-(α1,4)-Man backbone with branching of Gal*f* units in (β1,6). The Gal*f*-(β1,6)-Man*p* motif is notably found in symbiotic organisms (*Roccella decipens*, lichenized Ascomycete), but also in pathogenic organisms such as *Aspergillus* sp. and *Paracoccidioides brasiliensis*.[2] The second type of selected Gal-Man linkage consists of a Man-(α1,6)-Man backbone with branching of Gal*f* units in (β1,4). Interestingly, the Gal*f*-(β1,4)-Man*p* motif is exclusively found in lichens but has been reported in various genera: *Lasallia*, *Gyrophora*, *Usnea*, *Collema*.[31–35]

One additional challenge for such a study is that much of the work found in the literature regarding the fragmentation of oligosaccharides are studies conducted on protonated (or deprotonated) *N*-glycans, but ion mobility separations of neutral oligosaccharides are generally conducted on alkali cation adducts (such as Li^+ or Na^+). This is not only to afford an easier separation of ion populations in IMS,[36] but also to favor their fragmentation[37]—the interest of fragmentation approaches on alkali adducts has long been demonstrated, notably for the characterization of the linkage regioisomerism[38]. On the downside, the documentation regarding the possible rearrangements occurring on Li^+ or Na^+ species is scarce, but such reports are increasing in the recent literature. For instance, Bythell and co-workers recently investigated the fragmentation pathways of alkali-cationized oligosaccharides using molecular dynamics.[39,40] It is also notable that gas-phase rearrangements of oligosaccharides that were previously considered impossible on alkali adducts, such as subunit migration and internal residue loss, were recently described for such ions[41]. This questions on whether other rearrangements described for protonated species can occur. For example, pyranose-to-furanose isomerization was recently evidenced in protonated peptidoglycans.[42]

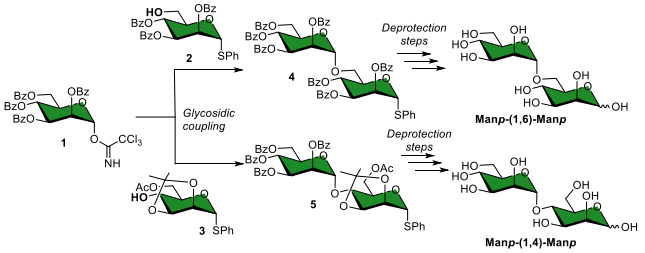
EXPERIMENTAL SECTION

Synthesis of the standards.

Gal-Man disaccharides. The Gal-Man disaccharides used as references in this study were previously used to build a library of IM-MS and IRMPD profiles. Their synthesis and NMR validation are reported in this earlier publication by Favreau *et al.*[4]

Gal-Man-Man trisaccharides. Gal-Man-Man trisaccharides were synthesized for this study using the approach fully described in another article submitted for publication.

Man-Man disaccharides. Man-Man disaccharides were also synthesized for this study to serve as references for the reducing end of the trisaccharides. Shortly, both mannobioses Man*p*-(1,6)-Man*p* and Man*p*-(1,4)-Man*p* were synthesized by coupling donor **1** to acceptor **2** and **3**, respectively, under acidic conditions. The resulting disaccharides were then submitted to standard deprotection steps under appropriated conditions to yield the desired disaccharides (**Scheme 1**). Further details about their synthesis and NMR validation are given in the **Supporting Information**.



Scheme 1. *Principle of the synthetic way used to afford the targeted mannobioses.*

^{18}O-labeling of the reducing end. 50 µL of 1 mg/mL aqueous stock solutions were speed-vacuum-dried and resuspended in the same volume of $H_2{}^{18}O$. Samples were incubated for 10 days at room temperature, away from light, to let the mutarotation exchange the hydroxyl group at the reducing end of the oligosaccharides.

Other chemicals used for IM-MS analysis.

HPLC grade methanol and acetonitrile were purchased from Carlo-Erba (Val de Reuil, France). Ultrapure water

was produced using a Milli-Q system (Millipore, Burlington, MA, USA). Lithium chloride (LiCl) and 97% $H_2{}^{18}O$ were purchased from Sigma-Aldrich (Saint Quentin Fallavier, France).

Cyclic IMS analyses.

Analytical conditions. The analyses were performed on a SELECT SERIES Cyclic IMS (cIMS) Q-IM-TOF spectrometer (Waters, Wilmslow, UK), with the TOF analyzer set to sensitivity mode (V mode). Details about the voltages and pressures applied in the instrument are given in **Supporting Text 1**. Samples were infused at a concentration of 5 µg/mL in 50:50 water/methanol (v/v), at a flow rate of 5 µL/min. Lithium adducts were generated by doping the solvents with 0.1 mM of LiCl. To check the purity of the synthesized trisaccharides, the oligosaccharides were further analyzed using the HPLC gradients described in **Supporting Text 2**—all results presented in the main text were obtained in direct infusion. The trisaccharides were fragmented/activated using collision-induced dissociation (CID) either before IMS analysis (in the "trap" cell) or in-between two rounds of IMS (in so-called IMS/IMS experiments, see next paragraph) using nitrogen as a collision gas. The choice of the fragmentation strategies are detailed throughout the description of the results.

Multistage IMS experiments. IMS experiments were conducted at a pressure of 1.75 mbar in the ion mobility cell, with a traveling wave velocity of 375 m/s and wave height of 22 V. The HR-IMS profiles of the disaccharides and disaccharidic fragments were acquired after a 4-pass separation around the cyclic racetrack (17.5 ms of separation). Two types of multistage IMS experiments were further performed.

The first type of IMS^n experiment that was be performed is the so-called 'head-and-tail' or 'top-and-tail' experiment (**Figure S1**). This experiment mainly serves the purpose of avoiding the wrap-around phenomenon. It is designed to specifically eject undesired ion populations—*i.e.*, ions with higher or lower mobilities than the population of interest—from the ion mobility cell, before they have enough time to catch up to the desired population (for the faster) or before the population of interest catch them up (for the slower). By switching the multifunction array region in the 'eject' direction, the undesired ions are ejected from the mobility cell, thus clearing the IM range. This allows to keep separating the populations of interest without being confronted to superimposition with the undesired population(s), which also makes this mode an interesting option to study the effect of extended exposure to the traveling waves. The main advantage of the 'head-and-tail' experiment is that the ion population of interest stays in the racetrack while ejecting the undesired populations. Therefore, the separation that was already obtained on the species of interest is preserved and can be enhanced by additional passes.

The second type of IMS^n experiment—most useful for structural characterization studies—is the 'slicing' experiment (**Figure S2**). Like the 'head-and-tail' experiment, it empties the mobility cell of the undesired ion populations, but opposite to the 'head-and-tail' experiment, the ions of interest are sent backwards in the pre-array store cell (*i.e.*, they do not remain in the cyclic racetrack). This is achieved by controlling the direction of the traveling wave in the multifunction array region (**Figure S3A**). Then, the ion population of interest is stored in the pre-array store cell for an amount of time defined by the user, after which it can be reinjected in the cyclic IMS racetrack (**Figure S3B**). It is notably possible to activate/fragment the selected population upon reinjection, by increasing the reinjection voltages and making the ions collide with the IMS gas.

IMS data processing. Data were processed using MassLynx 4.2 and Driftscope 2.9 (Waters, Wilmslow, UK). Drift times were recalculated from IMS arrival times by subtracting the times of all events preceding the final separation in the cIMS sequences. To correct for interday variability, the drift times were aligned using the disaccharide standards as a reference (these standards were analyzed in every analytical batch). The final drift time distributions (DTDs) of the disaccharides and disaccharidic fragments were then deconvolved with a gaussian fitting strategy using CIUSuite2,[43] with expected peak widths of 0.6 ± 0.1 ms for Gal*f*-Man and Man-Man, and of 0.58 ± 0.1 for Gal*p*-Man (detection up to > 1 % signal intensity, maximum number of peaks allowed = 6). Collision cross section (CCS) calibration was performed using MajorMix (Waters, Wilmslow, UK) as an external standard, using a logarithmic fit calibration method.[44,45]

Theoretical studies on Man-(α1,6)-Man.

Molecular dynamics. Molecular Dynamics was used to explore the potential energy surface of complexes [Man-(α1,6)-Manα + Li]$^+$ and [Man-(α1,6)-Manβ + Li]$^+$ with the PM7 potential in OpenMopac.[46] For each anomer, a trajectory of 10 ps at 1000 K was ran and yielded 20 000 geometries. After optimization with PM7 and elimination of the identical conformers, around 4100 geometries were obtained for each anomer.

Theoretical CCS calculations. Theoretical CCS values were calculated for representative conformations selected from the MD analysis of [Man-(α1,6)-Man + Li]$^+$ using IMoS Software v1.10. Calculations were performed using the Trajectory Method (4-6-12 Lennard Jones-Induced Dipole model) for N_2 at a pressure of 175 Pa (1.75 mBar) and a temperature of 304 K. The default values for N_2 were applied for the other parameters (M_{gas} 28.00 Da, gas radius 1.5 Å, polarization 1.7 Å^3).

RESULTS AND DISCUSSION

Galactomannan trisaccharides of the type $(Gal)_1(Man)_2$ selected from two structural families of biological relevance were considered in the study. They were chemically synthesized in their Gal*f* and Gal*p* forms, alongside the corresponding Gal-Man and Man-Man disaccharides (**Figure 1**). They are divided in three series: *series A* (trisaccharides ① and ②) represents the Gal-(β1,6)-Man + Man-(α1,4)-Man family, while *series B* (trisaccharides ③ and ④) and *series C* (trisaccharides ⑤ and ⑥) represent the Gal-(β1,4)-Man + Man-(α1,6)-Man family.

Of note, the backbone of galactomannans is made of Man residues. Thus, *series A and B* consist of "linear" trisaccharides whereas in *series C*, the Gal residue is branched laterally on the reducing Man subunit. The odd-numbered

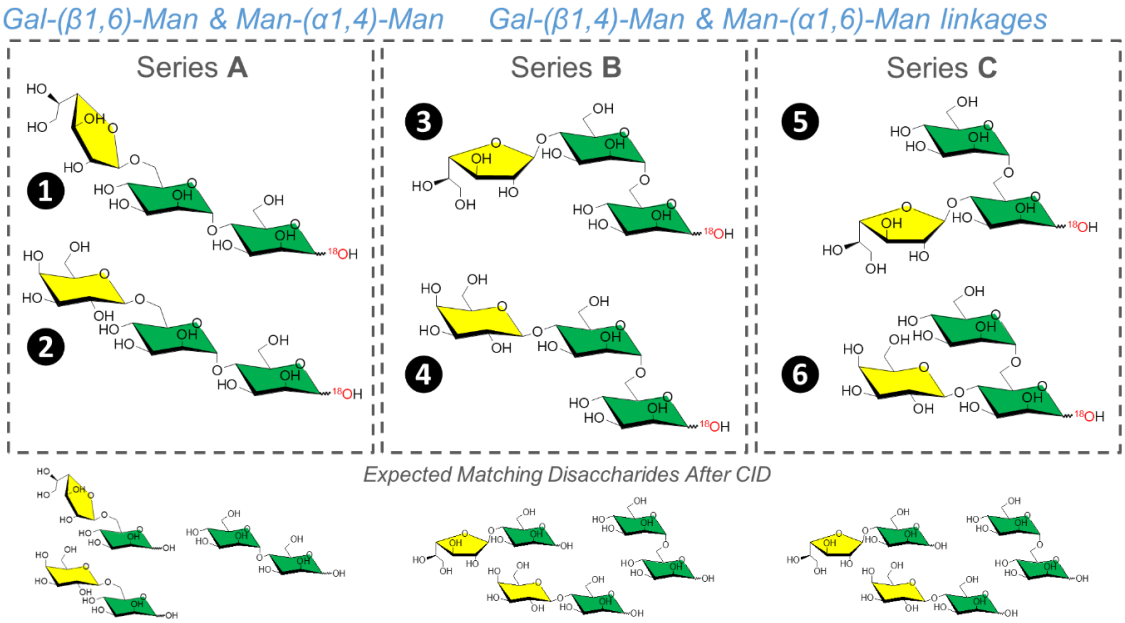


Figure 1. *Galactomannan trisaccharides synthesized for CID-IMS analysis (labeled with ^{18}O at the reducing end), and expected matching disaccharides after CID for each series. Color code for the subunits: yellow = Gal, green = Man.*

species are Gal*f*-containing trisaccharides, and even-numbered species are their Gal*p*-containing counterparts.

Does the IM profile of disaccharidic fragments generated by CID match the corresponding standards?

The initial hypothesis that drove our work was that *the IM profile of disaccharidic fragments should match the corresponding disaccharide references*. We thus aimed at validating this hypothesis using the three series of trisaccharides described above.

Series A.

The first selected series of trisaccharides consists of Gal*f*-(β1,6)-Man-(α1,4)-Man (compound ①) and its Gal*p*-(β1,6)-Man-(α1,4)-Man analogue (compound ②). Of note, LC-IM-MS analysis of trisaccharide ① showed the presence in this sample of a contaminant of unknown nature (**Figure S4**), with a significant contribution to the MS signal of *ca*. 15% (**Figure S5**). Importantly, MS is not quantitative and this does not reflect the actual purity of the sample, which was estimated to be greater than 95% using NMR spectroscopy. Because the contaminant can be partly resolved using cIMS (**Figure S6**), we thus analyzed *series A* using slicing IMS/IMS that allows discarding the contaminant from the IM racetrack. Trisaccharide ① was subjected to an energy of fragmentation of +150 V upon reinjection from the pre-store cell (a remaining contribution from the contaminant is possible but it is reduced compared to trap fragmentation).

The IMS profiles of the disaccharidic fragments from trisaccharides ① and ② are compared to the corresponding reference disaccharides in . The ^{18}O-labeling of the reducing end allows to distinguish the C_2 fragment coming from the non-reducing end (*m/z* 349) and the Y_2 fragment coming from the reducing end (*m/z* 351), as shown in **Figure 2A**. It is noteworthy that the intensity of C_2 ions is always around 100-fold lower compared to the one of Y_2 ions when fragmenting oligosaccharide as $[M+Li]^+$ species using nitrogen as a collision gas. A remaining contribution of ^{16}O-Y_2 in the IM profile of C_2 is due to remaining ^{16}O (around 3%) in the $H_2{}^{18}O$ used for labeling.

Figure 2B presents the profiles of the Gal-(β1,6)-Man reference disaccharides compared to the C_2 fragments of trisaccharides ① and ②. The profile of Gal*f*-(β1,6)-Man exhibited two peaks at drift times of 21.8 and 21.3 ms (pink trace), while Gal*p*-(β1,6)-Man exhibited one major peak at 20.9 ms, and a very minor peak at 21.6 ms (red trace, a y-axis zoom is presented in **Figure S7**). In the IM profile of the C_2 fragment from trisaccharide ①, we observed four peaks: the two major ones at 21.8 and 21.3 ms corresponded, as expected, to the Gal*f*-containing disaccharide. We also observed a small contribution of ^{16}O-Y_2 at 22.3 ms, and the fourth peak with lowest intensity was unidentified. However, this peak was absent when using the approach in LC-IM-MS (**Figure S8**), so we attributed it to a residual trace of the contaminant (see above). In the IM profile of the C_2 fragment from trisaccharide ②, we observed two peaks at 20.9 and 21.6 ms, both matching the Gal*p*-containing disaccharide.

Figure 2C presents the profile of the Man-(α1,4)-Man disaccharide compared to the Y_2 fragments of trisaccharides ① and ②. The profile of the reference mannobiose (blue trace) exhibited two peaks with drift times of 22.3 and 21.7 ms. The Y_2 fragments of both trisaccharides presented, as expected, the same IM profile and these profiles were a good match to the reference disaccharide: two main peaks at 22.3 and 21.6 ms, and a third peak of lower intensity at 21 ms (possibly a minor conformation generated by the activation).

Overall, in the case of *series A*, the IM profiles of the disaccharidic fragments C_2 and Y_2 produced from the Gal*f*- and Gal*p*-containing trisaccharides ① and ② were a good match to the fingerprints of the expected reference disaccharides, suggesting that the Gal*f*/Gal*p* isomerism is maintained upon fragmentation.

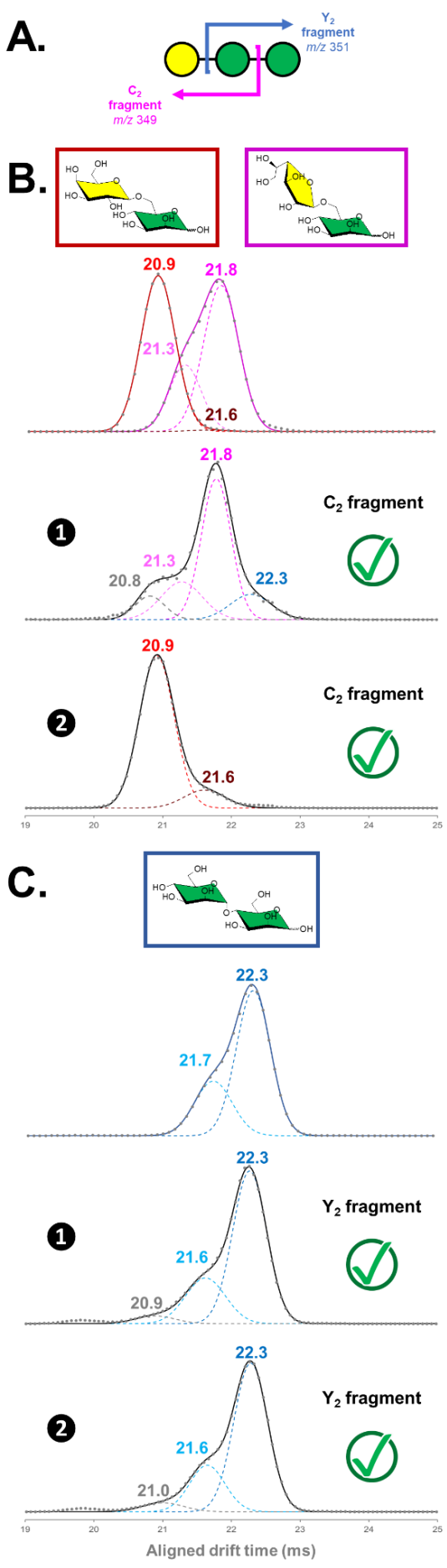


Figure 2. *CID-IMS study of the Gal-(β1,6)-Man-(α1,4)-Man trisaccharides (Series A).* ***A.*** *Types of disaccharidic fragments formed from ^{18}O-labeled linear trisaccharides (nomenclature according to Domon and Costello[47]).* ***B.*** *Comparison of the C_2 fragments (m/z 349) from the non-reducing end of the trisaccharides to the reference unlabeled Gal*f*/p-(β1,6)-Man disaccharides.* ***C.*** *Comparison of the Y_2 fragments (m/z 351) from the reducing end of the trisaccharides to the reference unlabeled Man-(α1,4)-Man disaccharide.*

Series B.

We next selected a second type of trisaccharide with different branching to check whether this "Gal*f*-memory" was always maintained. This series consists of Gal*f*-(β1,4)-Man-(α1,6)-Man (compound ③) and its Gal*p*-(β1,4)-Man-(α1,6)-Man analogue (compound ④). Contrary to *series A*, *series B* did not present any major contribution of contaminants in LC-MS analysis (some contaminants were present in ③, but none over 5% as shown in **Figures S9 and S10**), so we opted to perform CID in the trap cell with a collision energy of 50 eV, for a more efficient fragmentation and enhanced duty cycle.These linear trisaccharides give the same type of C_2 and Y_2 fragments as *Series A* (**Figure 3A**). The IMS profiles of the disaccharidic fragments from trisaccharides ③ and ④ are compared to the corresponding reference disaccharides in **Figure 3**.

Figure 3B presents the profiles of the Gal-(β1,4)-Man disaccharides compared to the C_2 fragments of trisaccharides ③ and ④. The profile of Gal*f*-(β1,4)-Man exhibited a major peak at a drift time of 22.8 ms, and two minor peaks at 21.5 and 22.3 ms (pink trace), while Gal*p*-(β1,4)-Man exhibited a single peak at 20.9 ms (red trace). In the IM profile of the C_2 fragment from trisaccharide ③, only a single peak at 21.5 ms matched that contained in profile of the reference Gal*f*-disaccharide. Otherwise, the profile also contained a feature at 20.7 ms (likely a ^{16}O contribution, see next paragraph), as well as two unidentified peaks at 22.4 (major) and 20.2 ms. This partial match could result from a greater flexibility of the furanose, but is difficult to investigate further because of very low signal intensity. On the other hand, the C_2 fragment of the Gal*p*-containing trisaccharide ④ matched well with the disaccharide standard, presenting a major peak at 20.9 ms, and a minor contribution of ^{16}O-Y_2 at 22.0 ms.

Figure 3C presents the profile of the Man-(α1,6)-Man reference disaccharide compared to the Y_2 fragments of trisaccharides ③ and ④. The profile of the reference mannobiose (blue trace) exhibited a single peak with a drift time of 22.0 ms. The Y_2 fragment of trisaccharide ③ exhibited an unexpected pattern: a main peak was observed at 20.7 ms, while the peak at the expected drift time of Y_2 (22.1 ms) was minor. In addition, we observed a doublet of peaks at 22.8 and 21.5 ms, which are specific of Gal*f*-(β1,4)-Man. During our experiments on Gal*f*-containing oligosaccharides (including our previous building of a disaccharide database[4]), we have not observed this doublet on any other motif (see **Table S1**). This strongly suggests that there is a migration of the Gal*f* residue to C-4 of the reducing Man, with an internal residue loss of a Man subunit. This theory is further supported by the fact that there is no apparent contamination of ③ by a trisaccharide in which

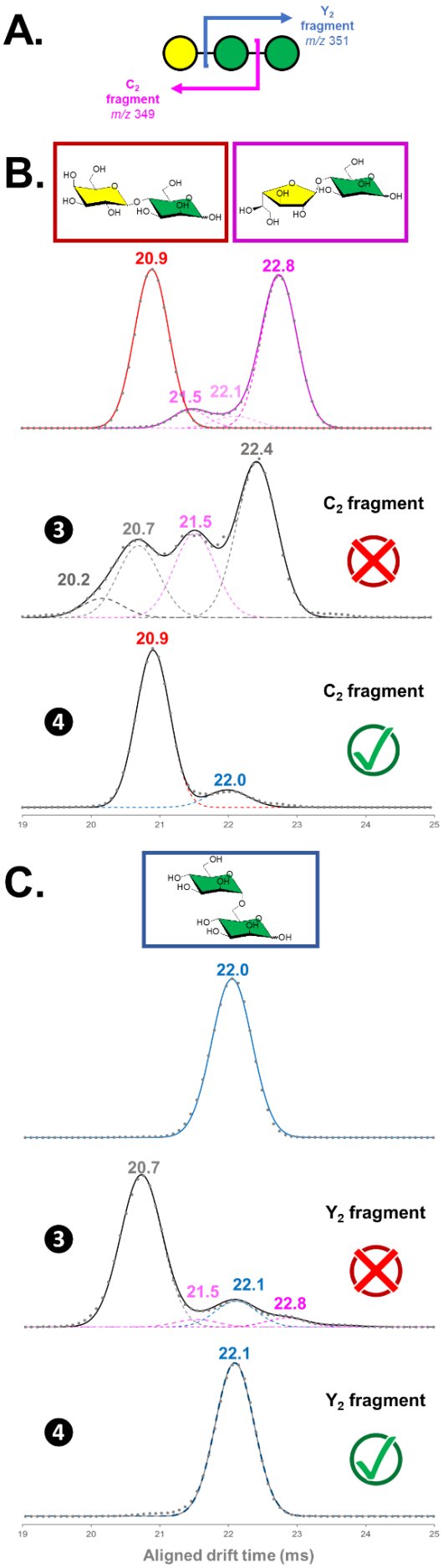


Figure 3. *CID-IMS study of the Gal-(β1,4)-Man-(α1,6)-Man trisaccharides (Series B).* ***A.*** *Types of disaccharidic fragments formed from ^{18}O-labeled linear trisaccharides (nomenclature according to Domon and Costello[47]).* ***B.*** *Comparison of the C_2 fragments (m/z 349) from the non-reducing end of the trisaccharides to the reference unlabeled Gal*f*/p-(β1,4)-Man disaccharides.* ***C.*** *Comparison of the Y_2 fragments (m/z 351) from the reducing end of the trisaccharides to the reference unlabeled Man-(α1,6)-Man disaccharide.*

the Gal*f* residue is positioned on the reducing Man. Such a species was indeed included in the study and analyzed in *series C* (species labeled ⑤) and it was not detected in the LC-MS analyses of ③ (**Figure S11**). We observed that this Gal migration and the unknown feature seemed specific to Gal*f*, because the Y_2 fragment of the Gal*p*-containing analogue (trisaccharide ④) was a good match to the (α1,6)-mannobiose standard, with a single peak at 22.1 ms.

Overall, for *Series B*, we observed that the disaccharidic fragments were a good match to the corresponding reference disaccharides in the case of Gal*p*, but only a partial match for Gal*f*. Although we expected that there could be, to some extent, changes in the Gal*f*-containing C_2 fragments compared to the corresponding disaccharides, the discrepancy in the IM profile of Y_2 ions (notably the main peak at 20.7 ms) was rather unexpected. Our hypothesis is that there is a specific gas-phase behavior for Gal*f*-containing oligosaccharides with a possible migration of the Gal*f* subunit to the reducing Man, as well as another parasitic phenomenon, which remained unidentified at this stage. Since this Y_2 fragment was produced with sufficient intensity, we decided to conduct further investigations.

Series C.

Considering the unexpected behavior of trisaccharide ③, we first wished to see if a similar pattern would be observed for any trisaccharide containing this Gal*f*-(β1,4)-Man motif. We therefore synthesized and analyzed a new series of trisaccharides with the same type of linkages as *Series B*, but in which the Gal residue would be positioned on the reducing Man rather than at the non-reducing end. *Series C* thus consists of Man-(α1,6)-[Gal*f*-(β1,4)-]Man (compound ⑤) and its Man-(α1,6)-[Gal*p*-(β1,4)-]Man analogue (compound ⑥). These trisaccharides were analyzed using the same fragmentation method as *Series B* (CID in the trap cell at 50 eV).

The trisaccharides of *Series C* have a different fragmentation compared to *Series A and B*, because both Gal and Man residues are branched on the reducing-end Man. This results in Gal-Man (Y_1) and Man-Man (Y_α) disaccharidic fragments both containing the ^{18}O and being observed at *m/z* 351 (**Figure 4A**). Of note, we considered the Gal residue to be the lateral one because in galactomannans, the backbone is made of Man while the Gal subunits are in lateral positions.

Figure 4B presents the results of the fragmentation experiments for trisaccharides ⑤ and ⑥. The profiles of the corresponding disaccharides are given for reference on the top trace and are identical to that described in **Figure 3**. In the profile of the disaccharidic fragments generated by CID of compound ⑤, we observe three peaks. The first one had a drift time of 22.1 ms, matching the expected Y_α fragment.

The other two peaks, at 22.8 and 21.5 ms, actually correspond to the expected values for the Gal*f*-containing Y_1 fragment. Regarding compound ⑥, we observed two peaks: one at 22.1 ms that corresponds to Y_α, and one at 21.0 ms that correspond to the Gal*p*-containing Y_1 fragment.

Overall, for *Series C*, both Gal*f*- and Gal*p*-containing trisaccharides exhibited fragmentation profiles that matched very well with the corresponding disaccharides. This suggests that the behavior of compound ③ under CID conditions is not purely related to the presence of Gal*f* but might arise from its positioning at the non-reducing end, and/or the regiochemistry of the mannose backbone in (α1,6).

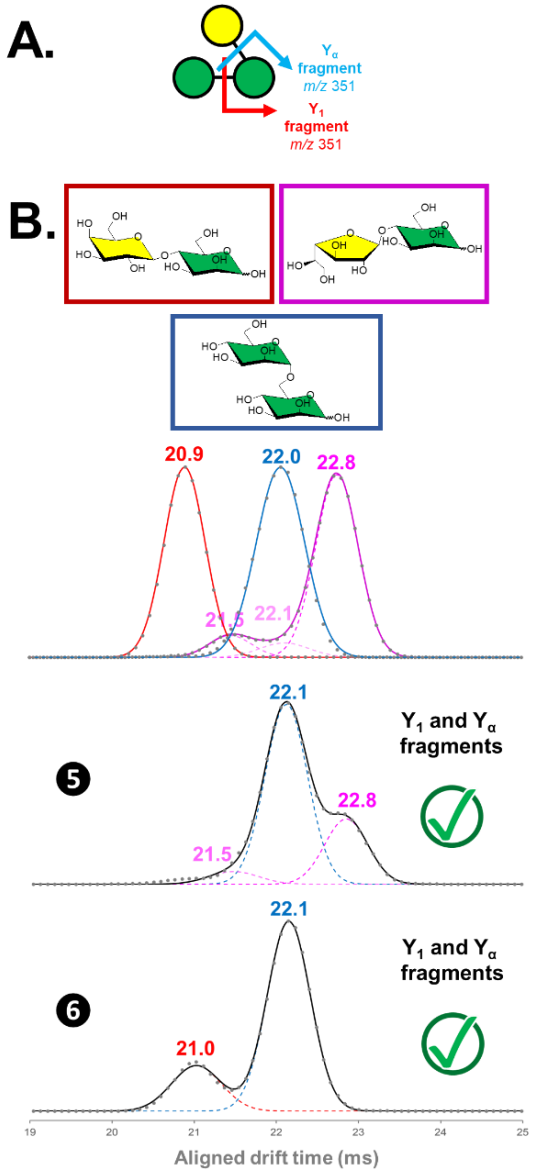


Figure 4. *CID-IMS study of the Man-(α1,6)-[Gal-(β1,4)-]Man trisaccharides (Series C).* ***A.*** *Types of disaccharidic fragments formed from ^{18}O-labeled branched trisaccharides (nomenclature according to Domon and Costello[47]).* ***B.*** *Comparison of the Y_1 and Y_α fragments of the trisaccharides (both containing the reducing end at m/z 351) to the reference unlabeled Gal*f*/p-(β1,4)-Man and Man-(α1,6)-Man disaccharides.*

In-depth investigation of the gas-phase behavior of trisaccharide ③ upon CID.

We focused our interest on the major peak observed at a drift time of 20.7 ms on the IM profiles observed for the Y_2 fragment produced from trisaccharide ③. In fact, when this peak was submitted to a further step of CID after IM separation, it exhibited a pattern of cross-ring fragments typical of a (1,6) linkage (**Figure S12**) but it was absent from our database.

Initially, we suspected that this peak might be generated by inappropriate MS analysis conditions and performed a number of experiments to exclude this hypothesis. First, we performed an energy-resolved IMS (ER-IMS) experiment (**Figure 5**), showing that this peculiar profile was generated independently from the collision energy in the trap cell. Next, we tried to change the counter cation and analyzed trisaccharide ③ as $[M+Na]^+$ (*m/z* 529), but its Y_2 fragment (*m/z* 367) also had higher mobility compared to the (α1,6)-mannobiose standard (**Figure S13**). Finally, since we had observed a migration of the Gal*f* to C-4 of the reducing Man (**Figure 3C**, Y_2 fragment of trisaccharide ③, peaks at 21.5 and 22.8 ms), we considered that the feature at t_d 20.7 ms could be a migration to C-1 ((1,1)-linked disaccharides are absent from our database, and (1,1) linkages can give similar cross-ring fragments compared to (1,6) linkages). We thus analyzed an analogue of trisaccharide ③ protected in C-1 with a thiophenyl group. Yet, the Y_2 fragment generated from this analogue had higher mobility compared to the corresponding protected disaccharide (**Figure S14**). This similar behavior of the thiophenylated analogue compared to trisaccharide ③ indicates that the unknown peak is not a migration to C-1 of the reducing Man.

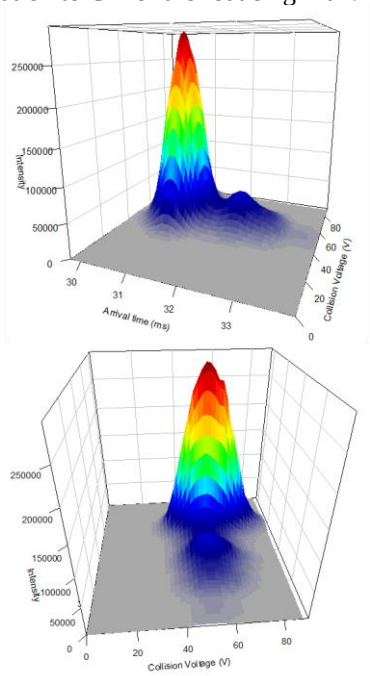


Figure 5. *Energy-resolved IMS (ER-IMS) profile of the Y_2 fragment from trisaccharide ③ (two viewpoints of the same ER-IMS profile are presented for clarity). The oligosaccharide was analyzed in its unlabeled form (^{16}O at the reducing end, Y_2 at m/z 349) to avoid any back exchange with the solvent during the analysis.*

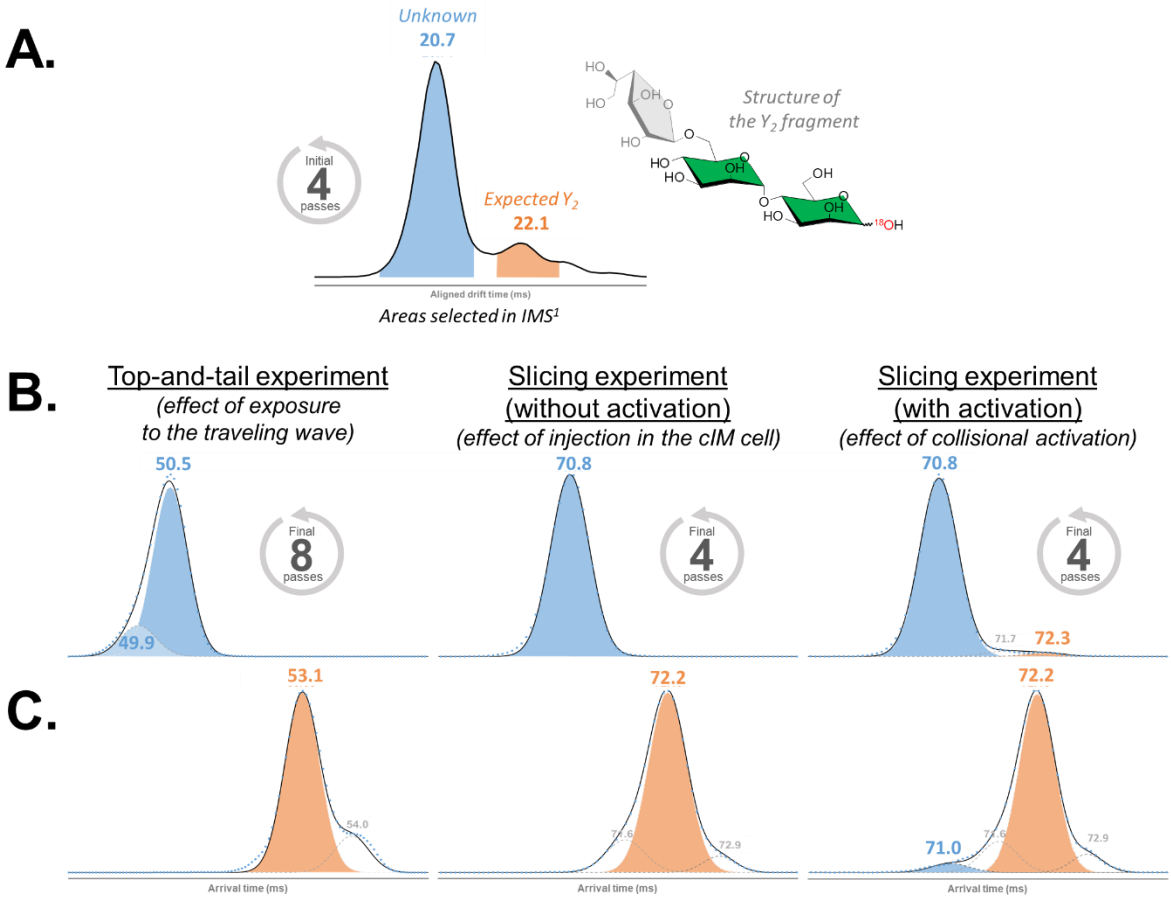


Figure 6. *IMS^n investigation of the interconversion between the expected Y_2 fragment of trisaccharide ③ (t_d = 22.1 ms) and the unknown peak at t_d = 20.7 ms.* ***A.*** *Structure of the Y_2 fragment, and areas selected in IMS^1 for further analysis of the unknown peak (blue, main peak) and of the expected peak (orange).* ***B.*** *Study of the unknown peak under different IMS^2 conditions.* ***C.*** *Study of the expected Y_2 peak under different IMS^2 conditions.*

Gas-phase evidence for a conformational switch upon collisional activation.

The above experiments ultimately led us to think that the peak at 20.7 ms was neither an artifact nor a structural rearrangement, but that the presence of Gal*f* at the non-reducing end had an impact on the conformation of the parent trisaccharide. More specifically, at the level of the Y_2 fragment, the Man-(α1,6)-Man could be subject to a conformational switch because of the high degree of liberty of the (1,6) linkage. A detailed study of the IM profile of (α1,6)-mannobiose (**Figure S15**) indeed revealed a slight shoulder at around 20.7 ms, but that is was of very low intensity (below the detection threshold of CIUSuite for gaussian fitting). However, the intensity of this shoulder could be enhanced with harsh CID conditions (fragmentation with a trap collision energy of 55 eV, *i.e.* 20 eV above our usual fragmentation parameters for disaccharides), which supported the idea that this peak could be a gas-phase conformer of (α1,6)-mannobiose. Similarly, the peak at 20.7 ms in the fragmentation profile of trisaccharide ③ could be a conformer of the expected Y_2 fragment found at 22.1 ms. A series of IMS/IMS experiments were thus conducted on both peaks of the fragmentation profile of trisaccharide ③ to verify if the conformational switch between the two could be induced in the gas phase (**Figure 6**).

First, we performed a top-and-tail IMS/IMS experiment to study the effect of increased exposure to the traveling wave of the IM cell. We doubled the time of exposure (*i.e.*, performed an additional four passes), but we did not observe any interconversion between the two populations. However, this experiment unveiled that the peak observed at a drift time of 20.7 ms after four passes actually contains (at least) two peaks, observed here at arrival times of 49.9 and 50.5 ms.

Second, we moved to slicing IMS/IMS experiments, in which we ejected the populations of interest back to the pre-store cell, then reinjected them and performed a new 4-pass separation. We reinjected the species from the pre-store either without or with energy (+100 V) to activate them. Without energy, we did not observe an interconversion, showing that the transition between the rest of the spectrometer and the higher pressure of the IM cell has no effect. In contrast, under activating conditions, we observed a conformational switch with the appearance of a peak at t_d 72.2 ms when isolating the peak of unknown origin (**Figure 6B**, right), and the appearance of a peak at t_d 70.9 ms when isolating the expected Y_2 peak (**Figure 6C**, right). This shows that the conformational switch can be triggered by collisional activation.

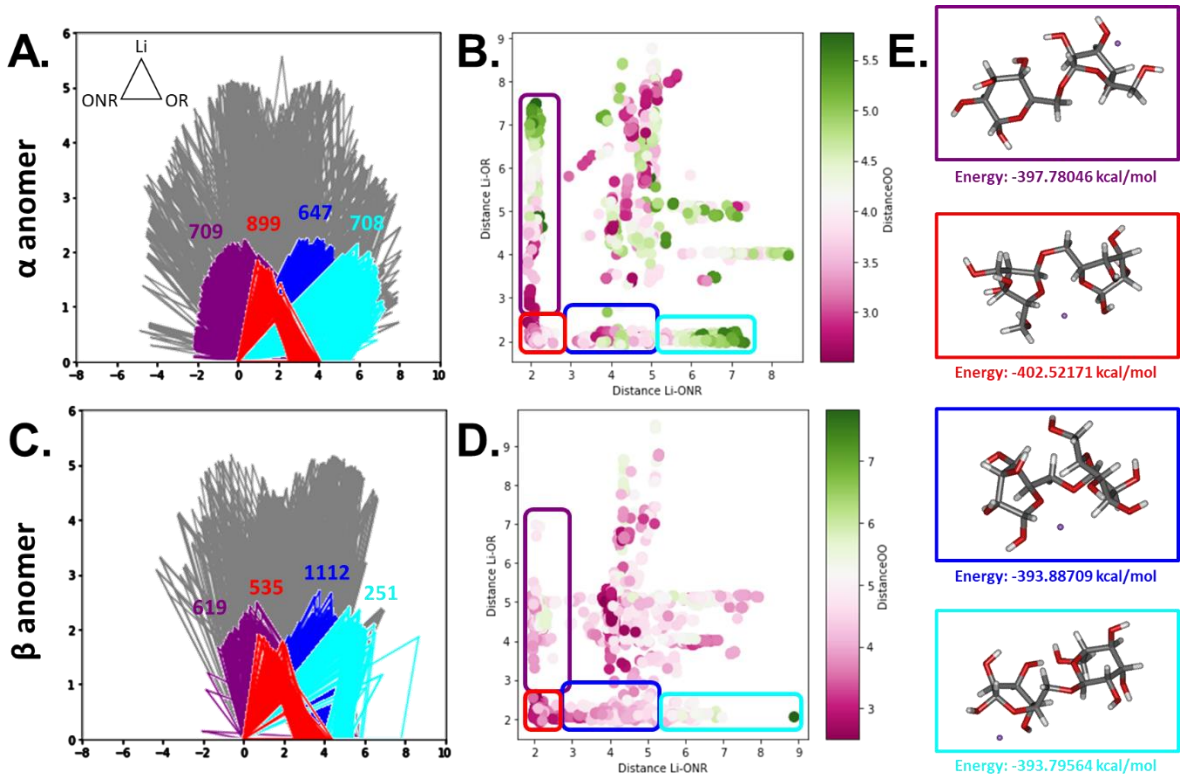


Figure 7. *Geometric analysis of the Man-(α1,6)-Man conformations generated by Molecular Dynamics (MD) for both the α and β reducing-end anomers (**A-B** and **C-D**, respectively). The structures were analyzed by comparing the distances of the Li couter-ion with the intracyclic oxygens from the non-reducing Man (ONR), and from the reducing Man (OR). Two representations are proposed: in the first one (**A** and **C**), each conformation is represented by a triangle representing the distances between Li, ONR, and OR (in Å, anchored at ONR = 0). The second one (**B** and **D**) presents the distance between Li and ONR (x-axis) vs the distance Li-OR (y-axis), with the distance ONR-OR as a color scale. In **E**, representative structures for each identified population (exemplified on the α anomer).*

In summary, we were able to observe an interconversion under CID conditions in the timescale of our IMS/IMS experiment, which validates that the peaks at t_d 20.7 and 22.1 ms are conformers of the same species. However, this interconversion does not occur to an extent that would make it the driving phenomenon for the observation of the major feature at 20.7 ms in the IM profile of the fragment from compound ③. One hypothesis is that, instead of an interconversion, the two conformers would arise from the coordination of the Li^+ cation in the parent trisaccharide.

Theoretical study of the conformational space of the Li-adducted Man-(α1,6)-Man moiety.

To validate this hypothesis, we opted to study the coordination of Li^+ with (α1,6)-mannobiose with quantum mechanical modelling. We thus performed Molecular Dynamics (MD) at the PM7 level of theory on both of the α- and β-reducing end anomers of this disaccharide, the main advantage of this semiempirical exploration being to rapidly generate a high number of different conformations.

To explore the conformations generated by MD, we focused on basic geometric criteria related to our hypothesis that the conformers were related to Li^+ coordination. More specifically, we elected the distances between the Li atom and the intracyclic oxygen atoms of each Man residue (herein coined ONR for the non-reducing end and OR for the reducing end). Indeed, these distances efficiently reflect the distance of Li to the corresponding heterocycles.

For the α-anomer (**Figure 7A-B**), we identified four conformational families by focusing on conformations for which Li was close to either ONR or OR. The four theoretical families represented 2963 out of the 4133 total conformations generated by MD. For two of these families, in red (899 conformations) and blue (647 conformations), Li was located relatively close to both OR and ONR indicating compact structures. Aside from that, we also observed a family for which Li was located at the non-reducing end (violet, 709 conformations), and a family for which Li was located at the reducing end (cyan, 708 conformations). For the β-anomer (**Figure 7C-D**), we observed the same families but the contribution of the red, violet, and cyan families was significantly lower in favor of the blue family, which alone represented 1112 out of 4180 total conformations.

Next, we proceeded to more in-depth analysis of the generated conformations. Representative structures for the four conformational families are presented in **Figure 7E**, exemplified on the α-anomer that is the most abundant for mannose in actual analytical conditions (additional structures are also shown in **Figure S16**). This confirmed that Li was placed at the non-reducing end for the violet population, between the two cycles for the red and blue populations, and at the reducing end for the cyan population.

In an attempt to correlate the conformational families of (α1,6)-mannobiose to the conformers that we observed in the DTD of the Y_2 fragment of trisaccharide ③, we per-

formed (i) a calibration of the DTD to give an experimental $^{TW}CCS_{N2}$ distribution (CCSD) and (ii) theoretical CCS calculations for the four representative structures using the trajectory method (Lennard Jones-induced dipole model). Since we observed that there were actually two peaks under the feature at t_d 20.7 ms (see **Figure 6**), we allowed for more peak overlap when fitting gaussian under the CCSD compared to the earlier part of this paper. These results are presented in **Figure 8**.

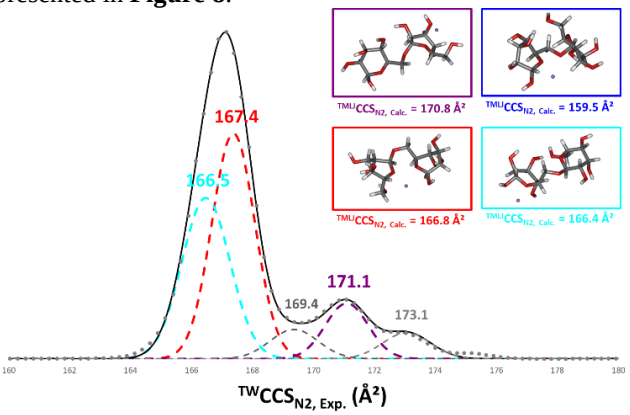


Figure 8. *Comparison between the experimental CCS distribution of the m/z 351 fragment from trisaccharide (3) ($^{TW}CCS_{N2}$), and theoretical CCS for the four model conformations of Man-(α1,6)-Man (see* ***Figure*** *7) calculated with the trajectory method (4-6-12 Lennard Jones-Induced Dipole model, $^{TMLJ}CCS_{N2}$).*

Of note, we obtained really close experimental and theoretical CCS values, leading us to think that the following results are valid despite CCS errors being reported in the range of 1–2% for both CCS calibrations in TWIMS, and theoretical CCS values.[47] However, the lower error when calculating the actual drift time for extended ion path length on the Cyclic, combined to the accuracy of conformation determination by MD methods, should compensate in part for this error.

In the experimental distribution, the expected conformer of the Y_2 fragment (*i.e.*, corresponding to the "normal" conformation of (α1,6)-mannobiose) was measured at ~171 Å², while two peaks were fitted under the major feature at ~166.5 and 167.5 Å². The fourth configuration, theorically calculated at 159.5 Å², was not observed in the experimental spectrum. When compared to our theoretical calculations, this suggests that the expected value corresponds to the Li^+ being located at the non-reducing end (calc. 171 Å²), while the more compact ion populations correspond to Li^+ being located either at the reducing end or between the two cycles.

Considering all of these results, we can explain the unexpected distribution of the Y_2 fragment of trisaccharide ③ as such: the normal cationization site of (α1,6)-mannobiose is located at the non-reducing end, which is occupied by a Gal*f* residue in the case of trisaccharide ③ causing Li^+ to coordinate with other sites instead. This also explains why we did not observe the same phenomenon when studying trisaccharide ⑤, since the non-reducing mannose was free. Interestingly, the phenomenon may be related to the greater flexibility of Gal*f* compared to Gal*p* since this change of conformation was not observed for trisaccharide ④ either.

CONCLUSION

Overall, these results show that (i) disaccharidic fragments matched well with the corresponding disaccharides in most cases, including Gal*f*-containing ones, but that (ii) in some cases the presence of the Gal*f* residue can impact the conformation of the entire oligosaccharide by modifying its coordination with the counter cation. In this study, we worked exclusively with alkali metal adducts that are well known to enhance IMS separations and are also well suited to fragmentation approaches. In the future, it might also be interesting to try different adducts like metals from other groups that might coordinate differently.

Importantly, we also observed a very probable Gal*f* migration to the reducing Man of trisaccharide ③, which had not been previously reported. We note that this rearrangement occurred specifically on the same compound for which we observed a compaction. One hypothesis could be that a compaction around Li^+ would bring the non-reducing Gal*f* close to the reducing Man. The Gal*f*-(β1,4)-Man*p* bond likely being fragile, this would allow the residue to migrate under the effect of collisional activation. Further works will be required to fully validate and understand this rearrangement. On one side, working on compounds where one cycle is entirely labeled with heavy isotopes (e.g., ^{13}C) would certainly help uncover the mechanisms behind the migration. But it would also be interesting to investigate why this phenomenon only occurs in specific cases. For instance, it could be interesting to evaluate the energy of the Gal*f*-Man*p* linkage depending on its regioisomerism.

To conclude, it is apparent that there is still much to understand regarding the gas-phase behavior of furanose, but understanding the behavior of furanosides in MS conditions should help answer another question: Why does Nature sometimes chose to biosynthesize furanosides instead of pyranosides?

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Synthesis and NMR validation of the mannobiose standards; instrumental conditions for cIMS analysis; HPLC analytical conditions; Description of IMS^n strategies; LC-(IM)-MS results for all oligosaccharides; Reference Drift Time values for our entire database of disaccharides; MS^3 spectrum of the unknown peak at *m/z* 351/t_d 20.7 ms; CID-IMS results for [③+Na]$^+$ and [③-SPh+Li]$^+$; IM profile of (α1,6)-mannobiose under different activation conditions. (PDF)

AUTHOR INFORMATION

Corresponding Author

* David Ropartz, david.ropartz@inrae.fr

Author Contributions

S.O., H.R. and D.R. designed the IMS experiments. S.O. performed the IMS experiments. L.L, L.-P.D. and V.F. synthesized the oligosaccharides. O.Y. and I.C. performed the MD calculations. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

Short title: Behavior of galactofuranosides upon CID

ACKNOWLEDGMENT

The authors thank the French National Research Agency for funding the ALGAIMS project (grant ANR-18-CE29-0006, https://algaims-35.webself.net/).

REFERENCES

(1) Taha, H. A.; Richards, M. R.; Lowary, T. L. Conformational Analysis of Furanoside-Containing Mono- and Oligosaccharides. *Chem. Rev.* **2013**, *113* (3), 1851–1876. https://doi.org/10.1021/cr300249c.

(2) Peltier, P.; Euzen, R.; Daniellou, R.; Nugier-Chauvin, C.; Ferrières, V. Recent Knowledge and Innovations Related to Hexofuranosides: Structure, Synthesis and Applications. *Carbohydrate Research* **2008**, *343* (12), 1897–1923. https://doi.org/10.1016/j.carres.2008.02.010.

(3) Ferrières, V.; Nugier-Chauvin, C.; Legentil, L.; Tranchimand, S. Chapter 19: How Recent Knowledge on Furano-Specific Enzymes Has Renewed Interest for the Synthesis of Glycofuranosyl-Containing Conjugates. In *Carbohydrate Chemistry*; 2014; pp 401–417. https://doi.org/10.1039/9781849739986-00401.

(4) Favreau, B.; Yeni, O.; Ollivier, S.; Boustie, J.; Le Dévéhat, F.; Guégan, J.-P.; Fanuel, M.; Rogniaux, H.; Brédy, R.; Compagnon, I.; Ropartz, D.; Legentil, L.; Ferrières, V. Synthesis of an Exhaustive Library of Naturally Occurring Galf-Manp and Galp-Manp Disaccharides. Toward Fingerprinting According to Ring Size by Advanced Mass Spectrometry-Based IM-MS and IRMPD. *J. Org. Chem.* **2021**, *86*, 6390–6405. https://doi.org/10.1021/acs.joc.1c00250.

(5) Manz, C.; Pagel, K. Glycan Analysis by Ion Mobility-Mass Spectrometry and Gas-Phase Spectroscopy. *Current Opinion in Chemical Biology* **2018**, *42*, 16–24. https://doi.org/10.1016/j.cbpa.2017.10.021.

(6) Hofmann, J.; Pagel, K. Glycan Analysis by Ion Mobility-Mass Spectrometry. *Angew. Chem.-Int. Edit.* **2017**, *56* (29), 8342–8349. https://doi.org/10.1002/anie.201701309.

(7) Miller, R. L.; Guimond, S. E.; Schwörer, R.; Zubkova, O. V.; Tyler, P. C.; Xu, Y.; Liu, J.; Chopra, P.; Boons, G.-J.; Grabarics, M.; Manz, C.; Hofmann, J.; Karlsson, N. G.; Turnbull, J. E.; Struwe, W. B.; Pagel, K. Shotgun Ion Mobility Mass Spectrometry Sequencing of Heparan Sulfate Saccharides. *Nat Commun* **2020**, *11* (1), 1–12. https://doi.org/10.1038/s41467-020-15284-y.

(8) Huang, Y.; Dodds, E. D. Ion Mobility Studies of Carbohydrates as Group I Adducts: Isomer Specific Collisional Cross Section Dependence on Metal Ion Radius. *Anal Chem* **2013**, *85* (20), 9728–9735. https://doi.org/10.1021/ac402133f [doi].

(9) Huang, Y. T.; Dodds, E. D. Discrimination of Isomeric Carbohydrates as the Electron Transfer Products of Group II Cation Adducts by Ion Mobility Spectrometry and Tandem Mass Spectrometry. *Analytical Chemistry* **2015**, *87* (11), 5664–5668. https://doi.org/10.1021/acs.analchem.5b00759.

(10) Gray, C. J.; Schindler, B.; Migas, L. G.; Pičmanová, M.; Allouche, A. R.; Green, A. P.; Mandal, S.; Motawia, M. S.; Sánchez-Pérez, R.; Bjarnholt, N.; Møller, B. L.; Rijs, A. M.; Barran, P. E.; Compagnon, I.; Eyers, C. E.; Flitsch, S. L. Bottom-Up Elucidation of Glycosidic Bond Stereochemistry. *Anal. Chem.* **2017**, *89* (8), 4540–4549. https://doi.org/10.1021/acs.analchem.6b04998.

(11) Pagel, K.; Harvey, D. J. Ion Mobility-Mass Spectrometry of Complex Carbohydrates: Collision Cross Sections of Sodiated N-Linked Glycans. *Anal Chem* **2013**, *85* (10), 5138–5145. https://doi.org/10.1021/ac400403d [doi].

(12) Li, H.; Bendiak, B.; Siems, W. F.; Gang, D. R.; Hill, H. H. Carbohydrate Structure Characterization by Tandem Ion Mobility Mass Spectrometry (IMMS)(2). *Analytical Chemistry* **2013**, *85* (5), 2760–2769. https://doi.org/10.1021/ac303273z.

(13) Ujma, J.; Ropartz, D.; Giles, K.; Richardson, K.; Langridge, D.; Wildgoose, J.; Green, M.; Pringle, S. Cyclic Ion Mobility Mass Spectrometry Distinguishes Anomers and Open-Ring Forms of Pentasaccharides. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (6), 1028–1037. https://doi.org/10.1007/s13361-019-02168-9.

(14) Ropartz, D.; Fanuel, M.; Ujma, J.; Palmer, M.; Giles, K.; Rogniaux, H. Structure Determination of Large Isomeric Oligosaccharides of Natural Origin through Multipass and Multistage Cyclic Traveling-Wave Ion Mobility Mass Spectrometry. *Anal. Chem.* **2019**, *91* (18), 12030–12037. https://doi.org/10.1021/acs.analchem.9b03036.

(15) Ollivier, S.; Tarquis, L.; Fanuel, M.; Li, A.; Durand, J.; Laville, E.; Potocki-Veronese, G.; Ropartz, D.; Rogniaux, H. Anomeric Retention of Carbohydrates in Multistage Cyclic Ion Mobility (IMSn): De Novo Structural Elucidation of Enzymatically Produced Mannosides. *Anal. Chem.* **2021**, *93* (15), 6254–6261. https://doi.org/10.1021/acs.analchem.1c00673.

(16) Ollivier, S.; Fanuel, M.; Rogniaux, H.; Ropartz, D. Molecular Networking of High-Resolution Tandem Ion Mobility Spectra: A Structurally Relevant Way of Organizing Data in Glycomics? *Anal. Chem.* **2021**, *93* (31), 10871–10878. https://doi.org/10.1021/acs.analchem.1c01244.

(17) Ropartz, D.; Fanuel, M.; Ollivier, S.; Lissarrague, A.; Benkoulouche, M.; Mulard, L. A.; André, I.; Guieysse, D.; Rogniaux, H. Combination of High-Resolution Multistage Ion Mobility and Tandem MS with High Energy of Activation to Resolve the Structure of Complex Chemoenzymatically Synthesized Glycans. *Anal. Chem.* **2022**, *94* (4), 2279–2287. https://doi.org/10.1021/acs.analchem.1c04982.

(18) Barnes, L.; Schindler, B.; Chambert, S.; Allouche, A. R.; Compagnon, I. Conformational Preferences of Protonated N-Acetylated Hexosamines Probed by InfraRed Multiple Photon Dissociation (IRMPD) Spectroscopy and Ab Initio Calculations. *Int. J. Mass Spectrom.* **2017**, *421*, 116–123. https://doi.org/10.1016/j.ijms.2017.05.005.

(19) Barnes, L.; Allouche, A.-R.; Chambert, S.; Schindler, B.; Compagnon, I. Ion Spectroscopy of Heterogeneous Mixtures: IRMPD and DFT Analysis of Anomers and Conformers of Monosaccharides. *International Journal of Mass Spectrometry* **2020**, *447*, 116235. https://doi.org/10.1016/j.ijms.2019.116235.

(20) Depland, A. D.; Renois-Predelus, G.; Schindler, B.; Compagnon, I. Identification of Sialic Acid Linkage Isomers in Glycans Using Coupled InfraRed Multiple Photon Dissociation (IRMPD) Spectroscopy and Mass Spectrometry. *Int. J. Mass Spectrom.* **2018**, *434*, 65–69. https://doi.org/10.1016/j.ijms.2018.09.005.

(21) Schindler, B.; Barnes, L.; Gray, C. J.; Chambert, S.; Flitsch, S. L.; Oomens, J.; Daniel, R.; Allouche, A. R.; Compagnon, I. IRMPD Spectroscopy Sheds New (Infrared) Light on the Sulfate Pattern of Carbohydrates. *J. Phys. Chem. A* **2017**, *121* (10), 2114–2120. https://doi.org/10.1021/acs.jpca.6b11642.

(22) Yeni, O.; Schindler, B.; Moge, B.; Compagnon, I. Rapid IRMPD (InfraRed Multiple Photon Dissociation) Analysis for Glycomics. *Analyst* **2021**. https://doi.org/10.1039/D1AN01870A.

(23) Yeni, O.; Gharbi, A.; Chambert, S.; Rouillon, J.; Allouche, A.-R.; Schindler, B.; Compagnon, I. O-Acetylated Sugars in the Gas Phase: Stability, Migration, Positional Isomers and Conformation. *Phys. Chem. Chem. Phys.* **2022**, *24* (2), 1016–1022. https://doi.org/10.1039/D1CP04837F.

(24) Schindler, B.; Legentil, L.; Allouche, A. R.; Ferrieres, V.; Compagnon, I. Spectroscopic Diagnostic for the Ring-Size of Carbohydrates in the Gas Phase: Furanose and Pyranose Forms of GalNAc. *Phys. Chem. Chem. Phys.* **2019**, *21* (23), 12460–12467. https://doi.org/10.1039/c8cp04082f.

(25) Ho, J. S.; Gharbi, A.; Schindler, B.; Yeni, O.; Brédy, R.; Legentil, L.; Ferrières, V.; Kiessling, L. L.; Compagnon, I. Dis-

tinguishing Galactoside Isomers with Mass Spectrometry and Gas-Phase Infrared Spectroscopy. *J. Am. Chem. Soc.* **2021**. https://doi.org/10.1021/jacs.0c11919.

(26) Gabelica, V.; Marklund, E. Fundamentals of Ion Mobility Spectrometry. *Curr. Opin. Chem. Biol.* **2018**, *42*, 51–59. https://doi.org/10.1016/j.cbpa.2017.10.022.

(27) Gabelica, V.; Shvartsburg, A. A.; Afonso, C.; Barran, P.; Benesch, J. L. P.; Bleiholder, C.; Bowers, M. T.; Bilbao, A.; Bush, M. F.; Campbell, J. L.; Campuzano, I. D. G.; Causon, T.; Clowers, B. H.; Creaser, C. S.; De Pauw, E.; Far, J.; Fernandez-Lima, F.; Fjeldsted, J. C.; Giles, K.; Groessl, M.; Hogan Jr, C. J.; Hann, S.; Kim, H. I.; Kurulugama, R. T.; May, J. C.; McLean, J. A.; Pagel, K.; Richardson, K.; Ridgeway, M. E.; Rosu, F.; Sobott, F.; Thalassinos, K.; Valentine, S. J.; Wyttenbach, T. Recommendations for Reporting Ion Mobility Mass Spectrometry Measurements. *Mass Spectrometry Reviews* **2019**, *38* (3), 291–320. https://doi.org/10.1002/mas.21585.

(28) Gabelica, V. CHAPTER 1:Ion Mobility–Mass Spectrometry: An Overview. In *Ion Mobility-Mass Spectrometry*; 2021; pp 1–25. https://doi.org/10.1039/9781839162886-00001.

(29) Giles, K.; Ujma, J.; Wildgoose, J.; Pringle, S.; Richardson, K.; Langridge, D.; Green, M. A Cyclic Ion Mobility-Mass Spectrometry System. *Anal. Chem.* **2019**, *91* (13), 8564–8573. https://doi.org/10.1021/acs.analchem.9b01838.

(30) Deng, L. L.; Webb, I. K.; Garimella, S. V. B.; Hamid, A. M.; Zheng, X. Y.; Norheim, R. V.; Prost, S. A.; Anderson, G. A.; Sandoval, J. A.; Baker, E. S.; Ibrahim, Y. M.; Smith, R. D. Serpentine Ultralong Path with Extended Routing (SUPER) High Resolution Traveling Wave Ion Mobility-MS Using Structures for Lossless Ion Manipulations. *Anal. Chem.* **2017**, *89* (8), 4628–4634. https://doi.org/10.1021/acs.analchem.7b00185.

(31) Gorin, P. A. J.; Iacomini, M. Structural Diversity of D-Galacto-d-Mannan Components Isolated from Lichens Having Ascomycetous Mycosymbionts. *Carbohydrate Research* **1985**, *142* (2), 253–267. https://doi.org/10.1016/0008-6215(85)85027-8.

(32) Iacomini, M.; Schneider, C. L.; Gorin, P. A. J. Comparative Studies on the Polysaccharides of Cladonia Alpestris (Reindeer Moss), Cladonia Confusa, and Cladonia Amaurocraea. *Carbohydrate Research* **1985**, *142* (2), 237–251. https://doi.org/10.1016/0008-6215(85)85026-6.

(33) Carbonero, E. R.; Smiderle, F. R.; Gracher, A. H. P.; Mellinger, C. G.; Torri, G.; Ahti, T.; Gorin, P. A. J.; Iacomini, M. Structure of Two Glucans and a Galactofuranomannan from the Lichen Umbilicaria Mammulata. *Carbohydrate Polymers* **2006**, *63* (1), 13–18. https://doi.org/10.1016/j.carbpol.2005.04.010.

(34) Omarsdottir, S.; Petersen, B. O.; Paulsen, B. S.; Togola, A.; Duus, J. Ø.; Olafsdottir, E. S. Structural Characterisation of Novel Lichen Heteroglycans by NMR Spectroscopy and Methylation Analysis. *Carbohydrate Research* **2006**, *341* (14), 2449–2455. https://doi.org/10.1016/j.carres.2006.06.026.

(35) Omarsdottir, S.; Freysdottir, J.; Olafsdottir, E. S. Immunomodulating Polysaccharides from the Lichen Thamnolia Vermicularis Var. Subuliformis. *Phytomedicine* **2007**, *14* (2), 179–184. https://doi.org/10.1016/j.phymed.2006.11.012.

(36) Williamson, D. L.; Bergman, A. E.; Nagy, G. Investigating the Structure of α/β Carbohydrate Linkage Isomers as a Function of Group I Metal Adduction and Degree of Polymerization as Revealed by Cyclic Ion Mobility Separations. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (10), 2573–2582. https://doi.org/10.1021/jasms.1c00207.

(37) Kailemia, M. J.; Ruhaak, L. R.; Lebrilla, C. B.; Amster, I. J. Oligosaccharide Analysis by Mass Spectrometry: A Review of Recent Developments. *Anal. Chem.* **2014**, *86* (1), 196–212. https://doi.org/10.1021/ac403969n.

(38) Hofmeister, G. E.; Zhou, Z.; Leary, J. A. Linkage Position Determination in Lithium-Cationized Disaccharides: Tandem Mass Spectrometry and Semiempirical Calculations. *J. Am. Chem. Soc.* **1991**, *113* (16), 5964–5970. https://doi.org/10.1021/ja00016a007.

(39) Bythell, B. J.; Abutokaikah, M. T.; Wagoner, A. R.; Guan, S.; Rabus, J. M. Cationized Carbohydrate Gas-Phase Fragmentation Chemistry. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (4), 688–703. https://doi.org/10.1007/s13361-016-1530-x.

(40) Abutokaikah, M. T.; Frye, J. W.; Tschampel, J.; Rabus, J. M.; Bythell, B. J. Fragmentation Pathways of Lithiated Hexose Monosaccharides. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (8), 1627–1637. https://doi.org/10.1007/s13361-018-1973-3.

(41) Chizhov, A. O.; Filatov, A. V.; Perepelov, A. V.; Knirel, Y. A. A New Example of Rearrangement Observed in the Tandem Mass Spectra of Oligosaccharides. *J Anal Chem* **2020**, *75* (14), 1842–1845. https://doi.org/10.1134/S1061934820140075.

(42) Guan, S.; Bythell, B. J. Evidence of Gas-Phase Pyranose-to-Furanose Isomerization in Protonated Peptidoglycans. *Phys. Chem. Chem. Phys.* **2021**, *23* (40), 23256–23266. https://doi.org/10.1039/D1CP03842G.

(43) Polasky, D. A.; Dixit, S. M.; Fantin, S. M.; Ruotolo, B. T. CIUSuite 2: Next-Generation Software for the Analysis of Gas-Phase Protein Unfolding Data. *Anal. Chem.* **2019**, *91* (4), 3147–3155. https://doi.org/10.1021/acs.analchem.8b05762.

(44) Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. Ion Mobility–Mass Spectrometry Analysis of Large Protein Complexes. *Nat. Protoc.* **2008**, *3* (7), 1139–1152. https://doi.org/10.1038/nprot.2008.78.

(45) Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V.; Ruotolo, B. T. Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. *Anal. Chem.* **2010**, *82* (22), 9557–9565. https://doi.org/10.1021/ac1022953.

(46) Stewart, J. J. P. MOPAC2012.

(47) Richardson, K.; Langridge, D.; Dixit, S. M.; Ruotolo, B. T. An Improved Calibration Approach for Traveling Wave Ion Mobility Spectrometry: Robust, High-Precision Collision Cross Sections. *Anal. Chem.* **2021**, *93* (7), 3542–3550. https://doi.org/10.1021/acs.analchem.0c04948.

Chapter 6 – Anomeric memory of the glycosidic bond in IMS^n experiments

A. Context

In the previous Chapters, we used cIMS to solve cases of constitutional isomerism. We showed that cIMS can distinguish between linkage regioisomers and furanose/pyranose tautomers. We also showed that cIMS can be used to solve constitutional rearrangements occurring upon CID, that would not have been identified with MS alone.

However, as described in **Chapter 1**, the cases of constitutional isomerism found in carbohydrates are further complicated by stereoisomerism and notably by the α/β anomerism of the glycosidic bond. Previous work showed that it was possible to separate and identify reducing-end anomers (α, β and open chain form) of neutral pentasaccharides using high-resolution cIMS. Ujma *et al.* confirmed that this separation originated from the reducing end by labeling the C-1 carbon with heavy oxygen (^{18}O).[117] Warnke *et al.* were able to assign which IMS peaks corresponded to the α and β reducing-end anomers (on disaccharides) using a unique SLIM-MS instrument coupled with cryogenic gas-phase IR ion spectroscopy, and were also able to observe the mutarotation of glucose.[121]

On another hand, Schindler *et al.* showed by IRMPD spectroscopy the existence of an "anomeric memory" of the glycosidic bond upon fragmentation, meaning that the fragments have retained the anomerism of the precursor ion.[42] Rizzo's group has established the generality of this anomeric retention[122] and extended its applicability to IMS/IMS strategies.[123] However, from those studies, it remained unclear if the "anomeric memory" would be retained after multiple rounds of CID fragmentation. Furthermore, this concept had been applied to standard compounds (*i.e.*, not actual biological samples) and with the support of other spectroscopic methods.

B. Objectives

The present project was a collaboration with the team of Dr Gabrielle Potocki-Veronese at the INRAE-CNRS Toulouse Biotechnology Institute. We characterized *de novo* the reaction products of a novel enzyme—a mannoside synthase named Uhgb_MS classified in the glycoside hydrolase 130 family (GH130). This enzyme has been recently identified from the human gut metagenomic gene catalog[143] and has the potential to synthesize mixed β/α-1,2-linked mannosides—which are biologically relevant structures. Such glycosidic motifs are indeed antigenic epitopes in yeast N-glycans,[144] in particular, in the antigenic factor 6 of the pathogenic yeast *Candida albicans*.[145]

The goal was to establish if Uhgb_MS has the original capability to biosynthesize both α- and β-glycosidic linkages. From an analytical standpoint, the samples were particularly adapted to establish whether (i) we were able to determine the intrachain anomerism of the reaction products of Uhgb_MS using high-resolution IMS-MS (IMS^n analyses) only, with appropriate standards, (ii) the anomerism of the glycosidic bond was preserved upon multiple rounds of CID fragmentation in IMS^n experiments.

C. Summary of the results

We analyzed the medium produced by the reaction catalyzed by Uhgb_MS with α-mannose-1-phosphate as a glycosyl donor and α1,2-mannobiose as an acceptor (**Figure 53**).

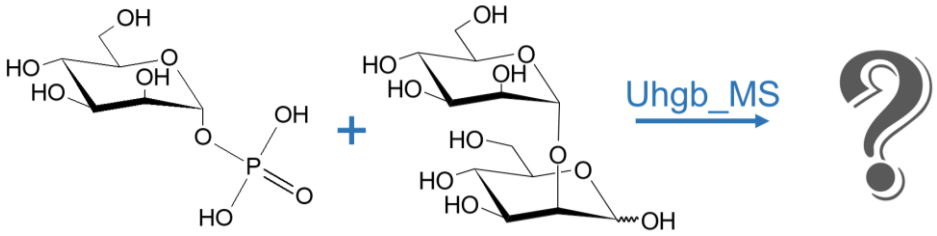


Figure 53. Reaction catalyzed by Uhgb_MS with unknown reaction product(s)

Preliminary HPAEC-PAD and HRMS analyses notably showed the presence of mannotriose and mannotetraose species in the reaction mixture, with a fragmentation pattern indicating pure (1,2) linkages in both cases. For the study, we had at our disposal some pure compounds corresponding to β1,2-mannotriose and α1,2-mannobiose.

We first endeavoured to characterize the mannotriose using IMS^2. First, we acquired reference ATDs for fragments of known (α1,2) and (β1,2) configuration. Then, we submitted the mannotriose to IMS/IMS analysis. We compared the DP2-like fragments coming from both the reducing and non-reducing ends of the mannotriose (Y_2 at *m/z* 351 and B_2 at *m/z* 331, respectively) to the references—they are expected to each contain the anomeric information of a single glycosidic bond (all samples were labeled at the reducing end with ^{18}O). **Figure 54** shows that the fragments of the mannotriose qualitatively match the (α1,2) standard, both at the non-reducing end and at the reducing end. This allowed us to characterize the mannotriose present in the reaction medium as Man-(α1,2)-Man.

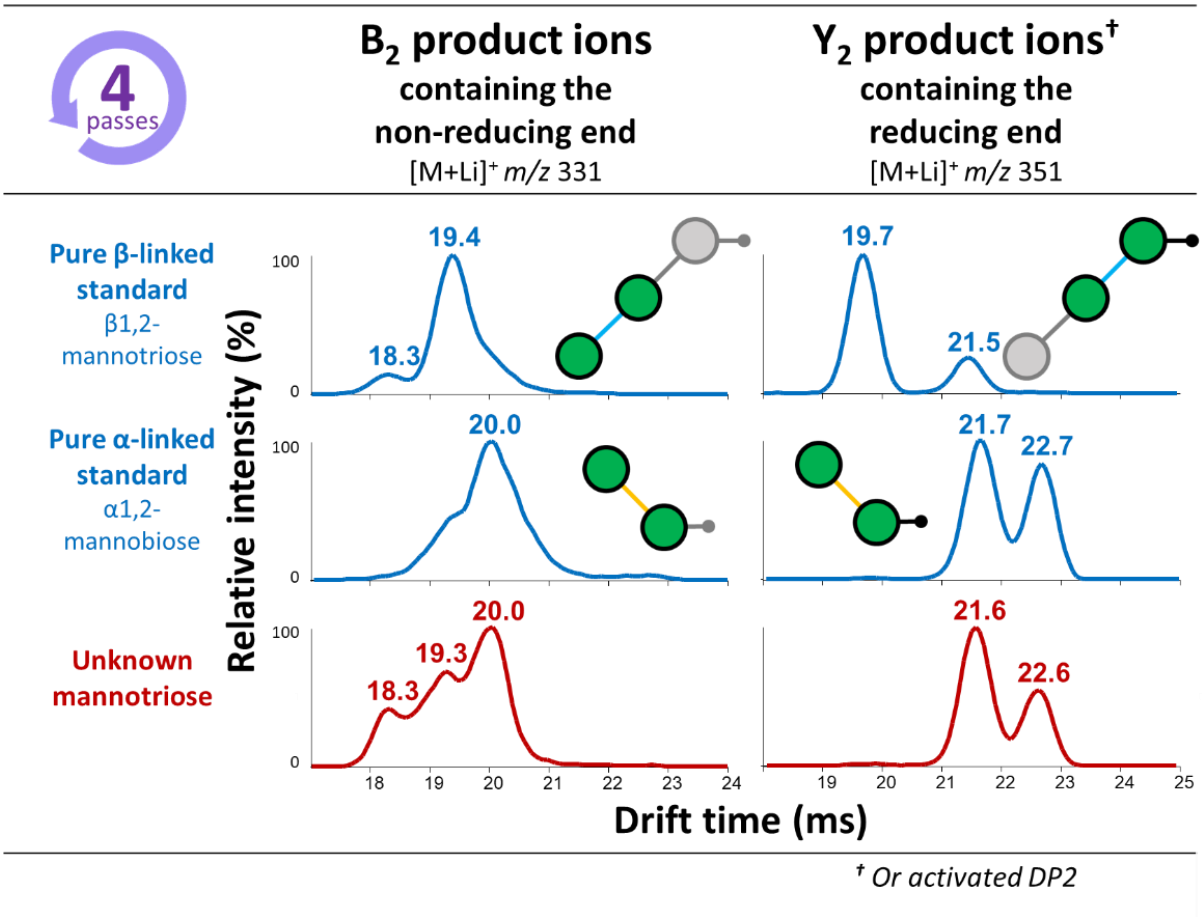


Figure 54. Drift time distributions of the IMS^2 fragments from pure β1,2- and α1,2-linked mannooligosaccharide standards (blue) and of the mannotriose from the reaction medium (red), analyzed as $[M + Li]^+$ adducts. The non-reducing end only yielded B_2 fragments of interest (left column), while the reducing end only yielded Y_2 fragments (right column).

We then moved to the mannotetraose, using a similar strategy. **Figure 55** shows that IMS^2 afforded direct determination of the reducing-end anomerism (as an (α1,2) linkage). The Y_2 fragment, that contains the reducing end, could unambiguously be distinguished in the IMS^2 spectrum thanks to ^{18}O labeling of the C-1 anomeric carbon which shifted the mass of this fragment by 2 Da (*m/z* 351). However, the fragment at *m/z* 331 (corresponding to that expected for B_2) was a mix of (α1,2) and (β1,2) signals, as revealed by the ATD profile in **Figure 55**. This suggests that the ion at *m/z* 331 probably has two origins: the B_2 fragment (containing the non-reducing end linkage) and an internal fragment containing the middle linkage of the mannotetraose. To determine which of the non-reducing end linkage or middle linkage was the one with a (β1,2) configuration, we opted to perform IMS^3 analysis. We selected the Y_3 fragment produced in IMS^2 and refragmented it. Y_3 contains the reducing end linkage and the middle linkage. We were able to confirm that the reducing end linkage was of (α1,2) configuration, and to determine that the middle linkage was also of (α1,2) configuration. This implied that the detected (β1,2) linkage in the ion at *m/z* 331 can only stand at the non-reducing end, and that the mannotetraose is Man-(β1,2)-Man-(α1,2)-Man-(α1,2)-Man.

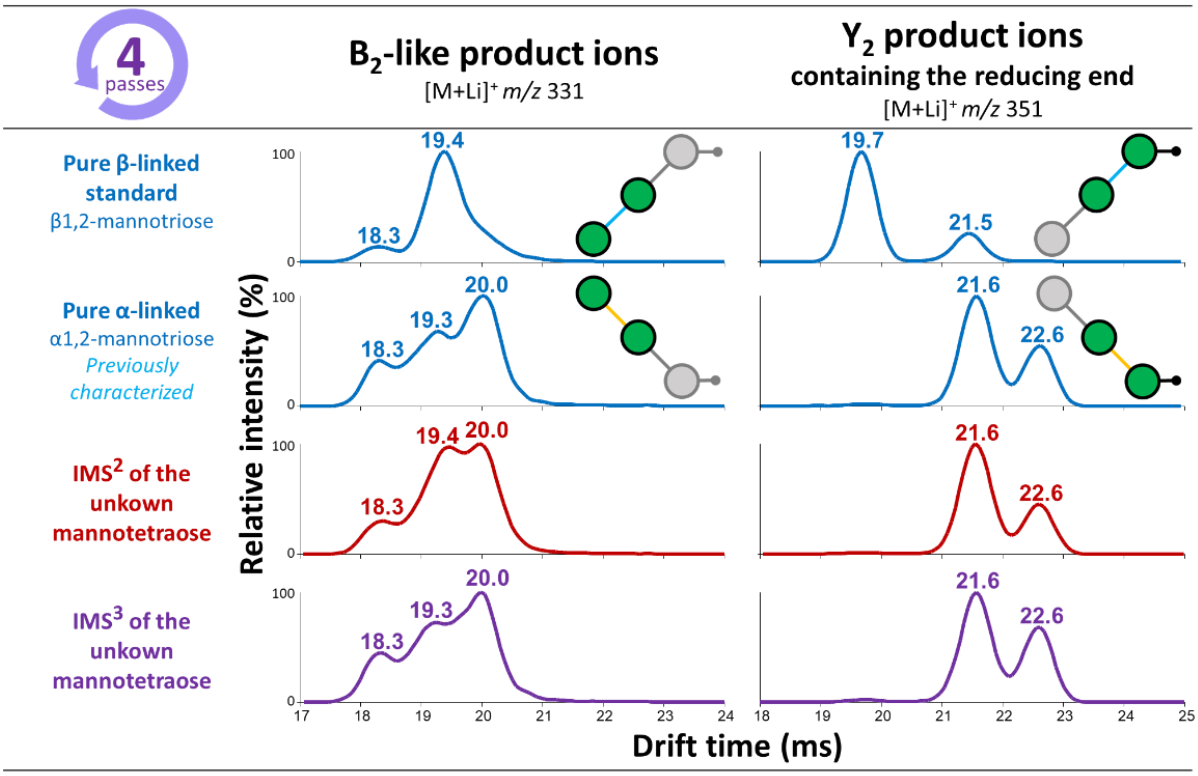


Figure 55. Drift time distributions of the IMS^2 fragments from pure β1,2- and α1,2-linked mannotriose (blue) of the IMS^2 fragments from the mannotetraose produced by Uhgb_MS (red) and of the IMS^3 fragments produced after refragmentation of the Y_3 mannotetraose fragment (violet), analyzed as $[M + Li]^+$ adducts.

D. Conclusions

We have successfully used IMS^n approaches to characterize the intrachain anomerism of two oligosaccharides produced by the Uhgb_MS enzyme. We were notably able to determine the complete intrachain anomerism of a mix-linked β/α-mannotetraose using IMS^3, showing that the ATDs of the fragments produced by CID remain the same after a second round of fragmentation. This allowed us to use IMS^2 fragments of the mannotriose as reference when interpreting IMS^3 results for the mannotetraose. Additionally, we showed in the present work that IMS^n can overcome one of the major drawbacks encountered in the use of CID fragmentation for the structural characterisation of carbohydrates, the occurrence of consecutive fragmentations. This require additional experiments and a supplemental data treatment but this can remove ambiguities or prevent misinterpretations of the data.

E. Scientific Communication

These results have been published in the following article presented as **Publication 3**. The supporting data for this publication are provided as **Appendix 5**. These results have also been presented orally at a national conference.

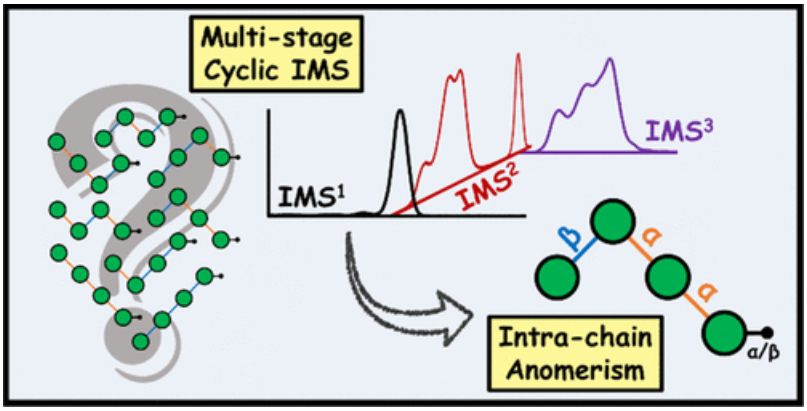
Peer-reviewed article

Ollivier, S.; Tarquis, L.; Fanuel, M.; Li, A.; Durand, J.; Laville, E.; Potocki-Veronese, G.; Ropartz, D.; Rogniaux, H. Anomeric Retention of Carbohydrates in Multistage Cyclic Ion Mobility (IMS^n): De Novo Structural Elucidation of Enzymatically Produced Mannosides. *Anal. Chem.* **2021**, *93* (15), 6254–6261. https://doi.org/10.1021/acs.analchem.1c00673.

Publication 3 – Anomeric Retention of Carbohydrates in Multistage Cyclic Ion Mobility (IMS^n): De Novo Structural Elucidation of Enzymatically Produced Mannosides

Simon Ollivier, Laurence Tarquis, Mathieu Fanuel, Ao Li, Julien Durand, Elisabeth Laville, Gabrielle Potocki-Veronese*, David Ropartz*, and Hélène Rogniaux

* Corresponding authors



Published in: Analytical Chemistry, 2021, 93(15), 6254-6261

analytical chemistry

pubs.acs.org/ac Article

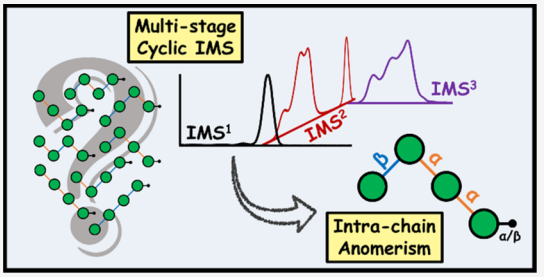
Anomeric Retention of Carbohydrates in Multistage Cyclic Ion Mobility (IMS^n): De Novo Structural Elucidation of Enzymatically Produced Mannosides

Simon Ollivier, Laurence Tarquis, Mathieu Fanuel, Ao Li, Julien Durand, Elisabeth Laville, Gabrielle Potocki-Veronese,* David Ropartz,* and Hélène Rogniaux

Cite This: *Anal. Chem.* 2021, 93, 6254–6261 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Carbohydrates are complex structures that still challenge analysts today because of their different levels of isomerism, notably the anomerism of the glycosidic bond. It has been shown recently that anomerism is preserved upon gas-phase fragmentation and that high-resolution ion mobility (IMS) can distinguish anomers. However, these concepts have yet to be applied to complex biological products. We have used high-resolution IMS on a cyclic device to characterize the reaction products of Uhgb_MS, a novel mannoside synthase of the GH130 family. We designed a so-called IMS^n sequence consisting of (i) separating and isolating specific IMS peaks, (ii) ejecting ions to a pre-array store cell depending on their arrival time, (iii) inducing collisional activation upon reinjection, and (iv) performing multistage IMS analysis of the fragments. First, we applied IMS^2 sequences to purely linked α1,2- and β1,2-mannooligosaccharides, which provided us with reference drift times for fragments of known conformation. Then, we performed IMS^n analyses of enzymatically produced mannosides and, by comparison with the references, we succeeded in determining the intrachain anomerism of a α1,2-mannotriose and a mix-linked β/α1,2-mannotetraose—a first for a crude biological medium. Our results show that the anomerism of glycosides is maintained through multiple stages of collisional fragmentation, and that standalone high-resolution IMS and IMS^n can be used to characterize the intrachain anomerism in tri- and tetrasaccharides in a biological medium. This is also the first evidence that a single carbohydrate-active enzyme can synthesize both α- and β-glycosidic linkages.



Downloaded via INRAE on April 20, 2021 at 07:16:57 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

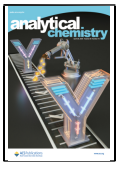
Carbohydrates are composed of individual building blocks—which are most often isomers—and the identification of such monomers is already a challenge in itself. The formation of the glycosidic bond between monomers further complicates the structural elucidation because it is subjected to additional levels of isomerism: (i) the monosaccharides offer a variety of nearly equally reactive positions, resulting in a regioisomerism and (ii) the mutarotation phenomenon and the chirality of the anomeric carbon that ensues result in the existence of α- and β-anomers.[1] The complete elucidation of carbohydrate structures therefore often requires the combination of various analytical methods that have been extensively reviewed by Gray et al.[2]

Of the three aforementioned challenges, the least difficult to address is perhaps the regioisomerism. Regioisomerism can be addressed through tandem mass spectrometry (MS/MS) provided that cross-ring fragments, which are diagnostic of the position at which the osidic units are linked, are produced. A number of activation techniques have been explored recently that considerably improved the diversity of fragments produced from carbohydrates in MS/MS—when compared to the most common collision activation dissociation (CAD, often coined CID for collision-induced dissociation)—including that of cross-ring fragments (for review, see McLuckey and Mentinova[3]). Some of these methods, such as XUVPD[4] and electron-activated dissociation—e.g., EDD,[5] EED,[6] (negative) ETD[7]— or He-CTD,[8,9] are very efficient for the structural characterization of complex carbohydrates and can be coupled with ultra high performance liquid chromatography (UHPLC) for the analysis of complex mixtures.[8,10]

However, despite their effectiveness, MS/MS-based approaches fail to resolve the isomerism of the monosaccharide units and/or the anomerism of the glycosidic bond. In this regard, some approaches such as gas-phase infrared (IR) ion spectroscopy,[2,11] and notably ion-mobility spectrometry (IMS) coupled to MS (IMS-MS), bring promising prospects.[2,12]

IMS-MS is a technique that separates ions depending on their conformation in the gas-phase, the separation originating

Received: February 12, 2021
Accepted: March 30, 2021
Published: April 8, 2021



© 2021 American Chemical Society

https://doi.org/10.1021/acs.analchem.1c00673
Anal. Chem. 2021, 93, 6254–6261

from the balanced interactions of an analyte with an electric field and a buffer gas. There exists three principles of low-field ion mobility available on commercial instruments. The first is drift tube IMS (DTIMS), in which the ions are submitted to a constant drop in potential, resulting in separation. The second is traveling wave IMS (TWIMS), in which the ions travel in a sequence of symmetric potential waves of defined velocity and height. In TWIMS, low-mobility ions surf faster from one wave to the next under the action of the gas, which increases their transit time through the IMS cell. The third is named trapped IMS (TIMS). Ions first get trapped under the action of both an electric field and a high flow of coaxial gas. Then, the electric field is lowered, allowing ions to be released according to their mobilities. IMS can also be performed with alternating low and high fields in a method called field asymmetric IMS (FAIMS) or differential IMS (DIMS). Using an asymmetric periodic field, orthogonal to the ion path, ions are filtered depending on their mobilities in the alternating field.[13,14] IMS-MS has the potential to identify monomers within a carbohydrate sequence, as exemplified by Gray et al. who showed that some fragments of different glucobioses had the same ion mobility as glucose.[15] Furthermore, the fragments of α-linked glucobioses possessed an additional IMS peak compared to β-linked glucobioses. Although the two peaks were poorly resolved in Gray et al.'s study, this finding brings promising perspectives on the capacity of ion mobility to distinguish anomers. Interestingly, it has also been shown by Hofmann et al. that intrachain anomers had different arrival times in TWIMS.[16]

Resolving power in IMS, in the range of $R \approx 50-60$ for most commercial instruments until recently, has long been a limiting factor to encompass the anomeric dimension of carbohydrates. However, lots of instrumental developments have been carried out in recent years and led to high-resolution (HR) IMS systems. Since the resolving power in DTIMS and TWIMS is directly proportional to the square root of the path length of the ions,[14] different groups have endeavored to increase this length. In 2009, the Clemmer group presented an "ion cyclotron mobility" instrument consisting of four serial drift tubes, which was the first instrument able to perform multipass IMS separations.[17] Waters Corp. (Wilmslow, U.K.) improved and commercialized this concept in the form of a Q-cIM-TOF instrument, the SELECT SERIES Cyclic IMS, in which a cyclic TWIMS cell is placed orthogonally to the axis of the mass spectrometer,[18] giving it the unique possibility of performing IMS^n and making it an extremely powerful tool for structural IMS-MS characterization (see the Experimental Section). This design reached a resolving power higher than 900 on carbohydrates.[19] Another high resolution, traveling wave-based IMS design, is the structure for lossless ion manipulation[20] (SLIM), introduced by the group of R. D. Smith. The initial step of their concept was to build a system, allowing minimal loss of ion transmission while increasing the path length. It was followed by a special, serpentine design to optimize the space and path length[21] and, later on, by a multipass SLIM design, making use of ion shutters (SLIM SUPER).[22,23] SLIM devices have been able to reach $R \approx 1800$[22] but have not been launched so far in any commercial instrument.

Having reached such resolving powers, IMS-MS instruments now make it possible to tackle carbohydrate anomerism, as illustrated in several recent studies. Nagy et al. reported two peaks for reducing sugars and only one for nonreducing sugars on a SLIM device and hypothesized the separation of C-1-anomers.[24] Thereafter, Ujma et al. proved that it was possible to separate anomers of neutral carbohydrates using the cyclic TWIMS design, and also confirmed that this separation originated from the reducing end by labeling the C-1 carbon with heavy oxygen (^{18}O).[25] Warnke et al. were able to assign which IMS peaks corresponded to α and β reducing-end anomers using a unique SLIM-MS instrument enabled with cryogenic gas-phase IR ion spectroscopy and were also able to observe the mutarotation of glucose.[26] Interestingly, Schindler et al. by IR ion spectroscopy showed the existence of an "anomeric memory of the glycosidic bond upon fragmentation", i.e. the fragments retained the anomerism of the precursor ion.[27] The Rizzo group has recently established the generality of this anomeric retention[28] and extended its applicability to IMS/IMS strategies.[29]

In this work, our aim was to exploit the unique functionalities of IMS/IMS (IMS^2) and IMS^n on a cyclic TWIMS instrument in an attempt to characterize *de novo* the reaction products of a novel mannoside synthesizing enzyme, named Uhgb_MS (mannoside synthase from an unknown human gut bacterium) and classified in the glycoside hydrolase family 130 (GH130). Its sequence was recently identified from the human gut metagenomic gene catalog[30] and was predicted to encode a mannoside-phosphorylase specific for β-1,2-linkages. Mannoside phosphorylases of the GH130 family breakdown mannosides in the presence of inorganic phosphate to generate α-mannose-1-phosphate (αMan1P) through an inverting phosphorolytic mechanism. They can also act as β-mannoside-synthases from αMan1P and glycosyl acceptors in the so-called reverse-phosphorolysis reaction. In the present work, we tested the potential of Uhgb_MS to synthesize, from αMan1P and α1,2-mannobiose, mixed β/α1,2-linked mannosides. Such glycosidic motifs are indeed antigenic epitopes in yeast *N*-glycans,[31] in particular, in the antigenic factor 6 of the pathogenic yeast*Candida albicans*.[32] We then combined the anomeric resolution of cyclic IMS with the above concept of "anomeric memory" to tackle the challenge of intrachain anomerism determination of the reaction products formed by Uhgb_MS. We succeeded both in resolving their regioisomerism and their intrachain anomerism (in particular, for one mannotriose, Hex3, and one mannotetraose, Hex4). We showed that the enzyme, in fact, produces different mannooligosaccharides, among them a mixed β/α1,2-linked mannotetraose. This is the first evidence that a single carbohydrate-active enzyme can synthesize both β- and α-glycosidic linkages.

EXPERIMENTAL SECTION

Enzyme Production. The Uhgb_MS sequence was identified from the catalog of the human gut microbiome[33] under gene identifier MH0105_GL0062872, locus scaffold 61381_1:8141:9193 (http://meta.genomics.cn/meta/dataTools). The gene encoding Uhgb_MS was synthesized by Biomatik Limited (Cambridge, Ontario, Canada), with optimization of codon usage for expression in *Escherichia coli*. The Uhgb_MS encoding gene was cloned in the pET-23a (+) vector, in fusion with N- and C-terminal (His) 6 tags, and expressed in the *E. coli* BL21-DE3 star strain. The recombinant *E. coli* cells were cultured overnight at 37 °C in a ZYM5052 autoinduction medium, supplemented with 100 μg/mL ampicillin. The recombinant cells were then harvested, resuspended in TALON buffer (20 mM Tris-HCl, pH 7.0,

https://doi.org/10.1021/acs.analchem.1c00673
Anal. Chem. 2021, 93, 6254−6261

300 mM NaCl), and disrupted by sonication. The supernatant was collected by centrifugation and subsequently loaded on a TALON resin loaded with cobalt (GE Healthcare, Chicago, IL) and previously washed with TALON buffer. After washing with TALON buffer + 20 mM imidazole, the purified protein was eluted by TALON buffer and 200 mM imidazole. At last, the purified protein was desalted using a PD-10 column (GE Healthcare) to eliminate imidazole and the buffer was changed to 20 mM Tris-HCl, pH 7.0. The protein concentration was determined by spectrophotometry using the NanoDrop (Thermo Fisher Scientific, Waltham, MA).

Mannooligosaccharide Production. The synthesis of mannooligosaccharides was performed in a reaction volume of 100 μL with 0.1 mg/mL purified enzyme, 10 mM of αM1P and 10 mM of α1,2-mannobiose in 20 mM Tris-HCl buffer, pH 7.0, during 24h at 37 °C.

Chemicals and Oligosaccharide Standards. Man-(α1,2)-Man was purchased from Carbosynth Ltd (Compton, U.K.) and αMan1P from Sigma-Aldrich (Saint Quentin Fallavier, France). Man-(β1,2)-Man-(β1,2)-Man was obtained by purification of the products of the reverse-phosphorolysis reaction performed from 10 mM αM1P and 10 mM Man by the mannoside-phosphorylase Teth514_1788, according to the protocol described in Chiku et al.[34] $H_2{}^{18}O$ was purchased from Sigma-Aldrich. HPLC grade methanol was purchased from Carlo-Erba (Val de Reuil, France). Ultrapure water ($H_2{}^{16}O$) was produced using a Milli-Q system (Millipore, Burlington, MA).

^{18}O Labeling. Fifty microliters of the stock solutions, reaction mixture and standards, were speed-vacuum-dried and resuspended in the same volume of $H_2{}^{18}O$. Samples were incubated for 10 days at room temperature, away from light, to let the mutarotation exchange the hydroxyl group at the reducing end of the oligosaccharides.[35] ^{18}O-labeled samples were diluted in the infusion solvent immediately before analysis.

IMS^n-MS Analyses. The analyses were performed on a SELECT SERIES Cyclic IMS (Waters, Wilmslow, U.K.). The enzymatic reaction medium was diluted 100-fold for infusion. Standards were infused at the appropriate dilution to reach an MS signal-to-noise ratio S/N > 100 on the $[M + Li]^+$ adduct. All samples were infused in a positive ion mode at a flow rate of 5 μL/min on an electrospray ion source. The solvent was composed of $H_2{}^{16}O$/MeOH (50:50) doped at 1 mM with LiCl to promote the detection of $[M + Li]^+$ species. The MS and IMS-MS spectra were recorded using Quartz software (release 6; Waters, Wilmslow, U.K.) on the m/z range of 50−1200, over 1 min (except for the multistage IMS, IMS^n, experiments that were recorded over 2 min). Additional information about the parameters applied on the mass spectrometer are given as the Supporting Information (SI) (Text S1).

First, a single-stage IMS separation was performed on the mannotriose and mannotetraose of interest to separate them from contaminants and minor species, in, respectively, 10 passes and 3 passes around the cyclic IMS cell (1 pass ≈ 1 m). Further details about the configuration of the cyclic IMS cell are given in the SI (Text S2). To characterize the structures using the anomeric memory of fragments, this was followed by an IMS^n experiment (SI, Figures S1 and S2). Briefly, an IMS/IMS (IMS^2) sequence consists of (i) separating peaks in a first IMS step, (ii) isolating specific peaks according to their arrival time (t_A) upon ejection to a prearray store cell, (iii) inducing collisional activation by increasing the kinetic energy of the ions during the reinjection in the T-wave array, and (iv) performing IMS analysis of the fragments. These steps can be repeated on the product ions, leading to so-called IMS^n experiments, analogously to classical MS^n strategies. The complete sequence of events for the anomeric retention experiment on tri- and tetrasaccharides is given in the SI, Table S1.

Data Processing and Analysis. Data were processed using Driftscope v. 2.9 and MassLynx v. 4.2 (both from Waters, Wilmslow, U.K.). The average ion mobility of the first two isotopes was extracted, yielding extracted ion-mobility spectra (EIMs). Arrival time distributions (ATDs) were smoothed in MassLynx 4.2 using the Savitzky−Golay algorithm (2 iterations over 3 bins). To compare the results, the time of all events preceding the final separation step in the cyclic control sequence was subtracted from the gross ion-mobility arrival time to give drift time and drift-time distributions (DTDs) (see Table S1).

Gaussian fitting of the smoothed DTDs was performed using CIUSuite2,[36] with an expected peak width of 0.7 ± 0.3 ms. Fitting of reference experimental distributions of IMS^n fragments was performed by minimizing the difference between the experimental data and a combined profile of reference fragments, as described in the SI, Text S3. ATDs and DTDs are reported according to the recommendations of Gabelica et al.[14]

RESULTS AND DISCUSSION

The products of the reaction catalyzed by Uhgb_MS from αM1P as a glycosyl donor and α1,2-mannobiose as an acceptor were first analyzed by HPAEC chromatography and compared to standards (Figure S3A). This analysis showed that after the reaction, the medium contained αM1P, mannose, α1,2-mannobiose, and β1,2-mannobiose, as well as two unknown oligosaccharides that did not match any of the available standards. The reaction medium was thus analyzed by MS in a discovery strategy (Figure S3B): the mass spectrum confirmed the presence of mannose (Hex1, m/z 187.08, $[M + Li]^+$ ion) and mannobiose (Hex2, m/z 349.13) and showed that the unknown compounds were reducing oligosaccharides at m/z 511.18 and m/z 673.23. These masses match with the masses of mannotriose (Hex3) and mannotetraose (Hex4) and were therefore selected for further structural characterization by IMS^n.

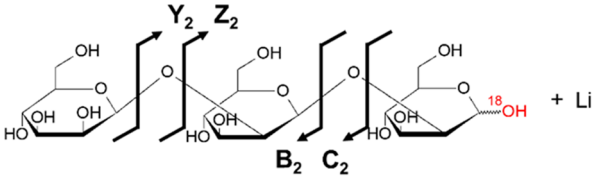
Ion Mobility of the Mannotriose and Mannotetraose. The two mannooligosaccharides resulting from the enzymatic reaction first underwent a single-stage ion mobility analysis to identify isoforms, after mass selection of their lithium adducts in the quadrupole. The mannotriose (Figure S4) showed a major isoform and traces of others: they were completely separated after 10 passes around the cyclic IMS cell, yielding an arrival time of 100 ms for the main mannotriose. The mannotetraose (Figure S5) also showed a major isoform and other traces, separated after three passes; the arrival time of the main mannotetraose peak was 51 ms. No additional IMS peaks were resolved from either mannotriose or mannotetraose when increasing the separation up to 15 passes (data not shown).

Regioisomerism of the Glycosidic Bond of the Produced Mannooligosaccharides. To obtain information about the regioisomerism of the glycosidic bonds of the oligosaccharides produced by Uhgb_MS, we fragmented the mobility separated ions using CID between the IMS cell and the TOF, generating specific fragmentation spectra for the

https://doi.org/10.1021/acs.analchem.1c00673
Anal. Chem. 2021, 93, 6254−6261

major IMS peaks of mannotriose and mannotetraose (Figure S6). To facilitate mass spectral interpretation and validate the approach, samples were previously labeled with heavy oxygen (^{18}O). The labeling leads to a mass shift of 2 Da for all reducing-end-containing species (i.e., precursor and/or product ions) (Scheme 1). This strategy aims at discriminating

Scheme 1. Illustration, for β1,2-Mannotriose, of the Distinction by Heavy Oxygen Labeling of Isobaric Product Ions Hailing from the Non-Reducing and Reducing Ends after Cleavage of the Glycosidic Bonds (Nomenclature According to Domon and Costello[37])



Non-reducing end:	Reducing end:
C_2 fragment *m/z* 349	Y_2 fragment *m/z* 351
B_2 fragment *m/z* 331	Z_2 fragment *m/z* 333

isobaric product ions arising from either end of the structure and was previously shown not to impact the ATDs, neither of the precursor nor of the fragment ions.[25]

For both peaks, a single type of $^{0,2}X$ intracyclic fragment appeared, unequivocally identified thanks to the ^{18}O-labeling. The fragments identified were for mannotriose $^{0,2}X_2$ (loss of 120 Da) at m/z 393.14 and Y_2 + $^{0,2}X_1$ (consecutive fragmentation, labeled by heavy oxygen) at m/z 231.09, and for mannotetraose $^{0,2}X_3$ (loss of 120 Da) at m/z 555.20 and Y_3 + $^{0,2}X_2$ and Y_2 + $^{0,2}X_1$ (consecutive fragmentations, labeled by heavy oxygen) at, respectively, m/z 393.14 and 231.09. These fragments are the specific signature of (1 → 2) linkages,[38] showing that both mannooligosaccharides of interest are linked through (1 → 2) bonds.

Note that, thanks to the ^{18}O-labeling, B_2 and Y_2 product ions could be unequivocally distinguished from Z_2 and C_2 fragments; our experiments indicated that the latter were completely absent from the fragmentation spectra of both oligosaccharides, including when the ions were fragmented in the pre-store cell between two steps of IMS separation (Figure S7).

Dual-stage IMS (IMS^2) Characterization of the Intrachain Anomerism of the Enzymatically Produced Mannotriose. Having characterized the regioisomerism of the enzyme reaction products, we proceeded to characterize the intrachain anomerism of the mannotriose. Only the major isoform separated in IMS (see Figure S4) had a sufficient intensity to be submitted for further IMS^2 analysis of the fragments. The IMS peak was selected according to its arrival time and collisionally fragmented (Figure 1A). The ^{18}O-labeled trisaccharide yielded two fragments at m/z 331 and 351, which uniquely match with B_2 and Y_2 fragments, respectively. Figure 1B compares the drift-time distributions (DTDs) of the B_2 and Y_2 fragments, after four passes around the cIMS cell, to those obtained in a similar experiment performed with two reference (1 → 2)-linked mannooligosaccharides: reducing α1,2-mannotriose was not commercially available so α1,2-mannobiose was used instead; β1,2-mannotriose was produced, as previously described in ref 34.[34]

Regarding the Y_2 fragments, containing the reducing end, the DTDs of the β-linked standard showed a doublet at 19.7 and 21.5 ms, and the α-linked standard showed a different doublet at 21.7 and 22.7 ms. The unknown mannotriose shows a doublet at 21.6 and 22.6 ms, the anomerism of the linkage for the Y_2 fragment is thus obviously of α-type (Man-(α1,2)-Man). Regarding the B_2 fragments, containing the non-reducing end, the distributions were more complex and contained overlapping peaks. The β-linked standard notably showed two clearly defined peaks with drift times of 18.3 and 19.4 ms, while the α-linked standard showed a main peak at 20.0 ms and a shoulder at around 19 ms. At first glance, the B_2 fragment from the unknown mannotriose showed a triplet at 18.3, 19.3, and 20.0 ms, the peak at 20.0 ms being the most intense, which hinted toward an α-type linkage. However, this DTD did not exactly match that of the α-linked standard.

Indeed, because the reducing α1,2-mannotriose was not commercially available (as mentioned before), the α-linked standard that was used is a pseudo-B_2 fragment hailing from the loss of $H_2{}^{18}O$ at the reducing end of α1,2-mannobiose: compared to a real B_2 fragment, intensity variations can be expected amongst the conformations found in the DTD. To verify that the conformations were qualitatively similar, we thus performed a deconvolution of the DTDs by Gaussian fitting (Figure 1C,D). The Gaussian fit distinguished three conformations in the DTD of the pseudo-B_2 fragment from α1,2-mannobiose, at 19.4, 20.1, and 20.6 ms. In the DTD of the unknown mannotriose, four conformations could be distinguished at 18.3, 19.2, 20.1, and 20.7. The last three well matched, both in their width and arrival time, those of the α1,2-mannobiose, albeit with intensity differences. We also noted a shift between the combined Gaussian fit and the experimental distribution for α1,2-mannobiose at around 18 ms, although too low in intensity for the fitting program to suggest one additional conformation; however, a distribution around 18 ms for α1,2-mannobiose is probably present, corresponding to the distribution at 18.3 ms observed for the unknown mannotriose. Thus, because the distributions of the gas-phase conformations between Figure 1C,D are qualitatively highly similar and because the main diagnostic conformation at 20.1 ms excludes a (β1.2) linkage, we characterize the B_2 fragment of the unknown mannotriose as α-linked.

Interestingly, the four Gaussian curves fitting the DTDs in Figure 1D had different widths. Expressed in FWHM (full-width at half maximum), we measured a width of 1.0 ± 0.1 ms for the peaks at 19.2 and 20.7, and 0.6 ± 0.1 ms for the peaks at 18.3 and 20.1 ms. In ion-mobility experiments, gas-phase conformations from the same molecular structure should exhibit peaks of similar width, at least for small molecules with limited flexibility. Our interpretation is that the two conformations at 19.2 and 20.7 ms arise from a first isomer of the B_2 fragment, produced in the gas phase, while the other two—at 18.3 and 20.1 ms—would correspond to a second isomer. The production of isomeric B fragments during CID fragmentation of alkali-cationized carbohydrates has indeed been described: two thermodynamically competing mechanisms yield B fragments of different structures, one with a double bond and the other one with an anhydro bridge (Scheme S1).[39]

In summary, the above structural study of the unknown mannotriose shows that its B_2 and Y_2 fragments are both linked in (α1,2). We thus conclude that Uhgb_MS produced a Man-

https://doi.org/10.1021/acs.analchem.1c00673
Anal. Chem. 2021, 93, 6254−6261

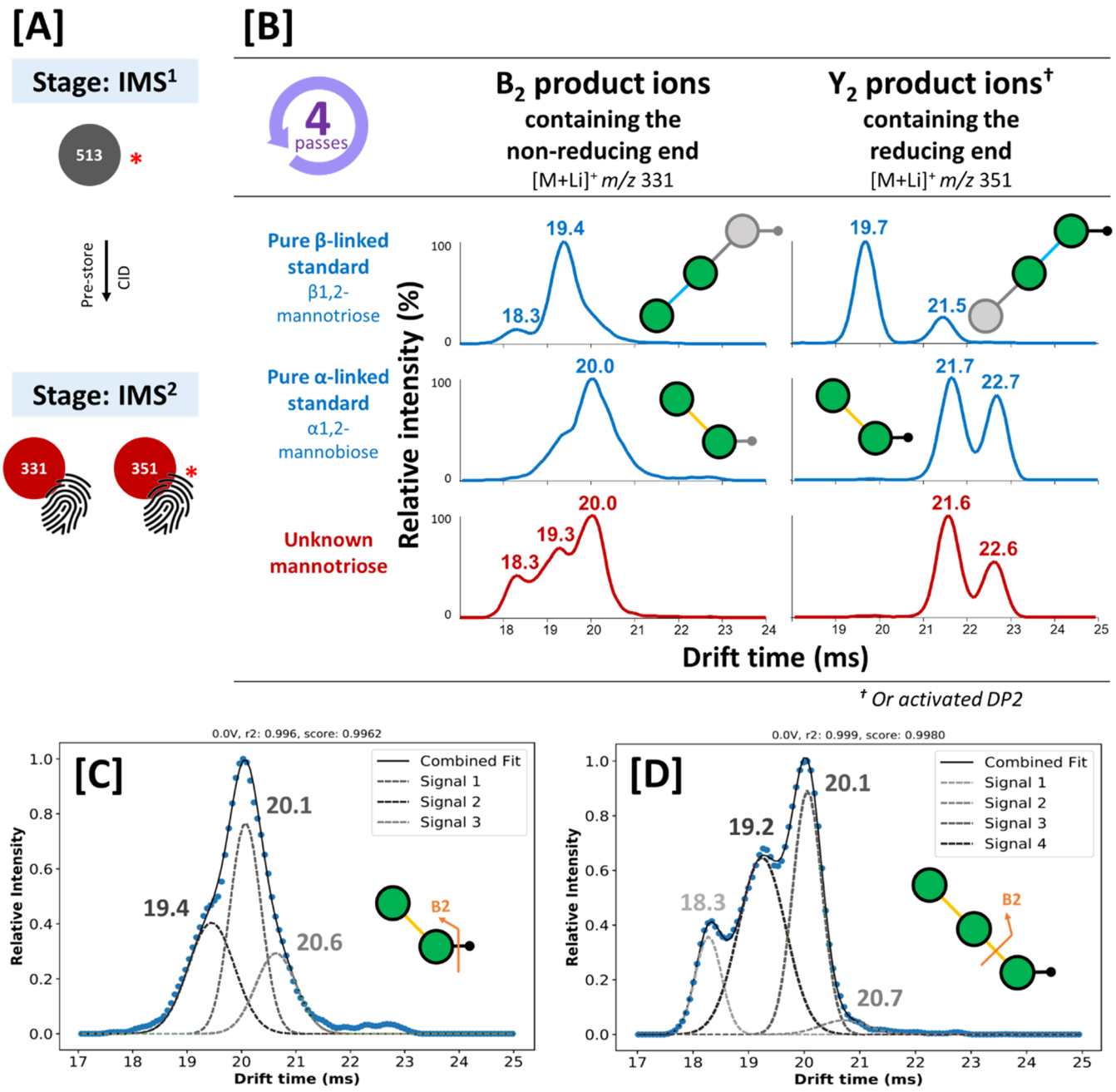


Figure 1. (A) Schematic representation of the IMS^2 selection strategy and of the analyzed mannotriose fragments. Ions bearing the ^{18}O-labeled reducing end are indicated with a red asterisk and fragments used for anomerism determination are indicated with a fingerprint. (B) DTDs of the IMS^2 fragments from pure β1,2- and α1,2-linked mannooligosaccharide standards (blue) and of the mannotriose from the reaction medium (red), analyzed as $[M + Li]^+$ adducts. The non-reducing end only yielded B_2 fragments of interest (left column), while the reducing end only yielded Y_2 fragments (right column). (C) Gaussian deconvolution of the conformations in the DTD of the pseudo-B_2 ions from α1,2-mannobiose. (D) Gaussian deconvolution of the conformations in the DTD of the B_2 ions from the enzymatically produced mannotriose. Insets for (C) and (D) show cleavages yielding m/z 331.

(α1,2)-Man-(α1,2)-Man mannotriose (hereafter called α1,2-mannotriose).

Multistage IMS (IMS^n) Characterization of the Intrachain Anomerism of the Enzymatically Produced Mannotetraose. Because the previous study of mannotriose showed that fragments retained the anomerism of the bond, we performed a similar study for the mannotetraose enzymatic product. Yet, an adaptation of the strategy was required to unravel the anomerism of this Hex4 oligosaccharide, as detailed below. An IMS^n (n = 3) was performed, as schematized in Figure 2A. After selection of the main precursor for the mannotetraose (see Figure S5) and the first step of prestore fragmentation (IMS^2), the fingerprint of the dimeric fragments was recorded a first time.

At this stage, the Y_2 fragment could be identified unequivocally at m/z 351, thanks to the ^{18}O labeling. The DTD of the Y_2 fragment showed a doublet at 21.6 and 22.6 ms (Figure 2B, red trace), which permitted, by comparison to the DTDs measured for the dimeric standards, to identify Y_2 as being α1,2-linked. However, the situation was ambiguous for

https://doi.org/10.1021/acs.analchem.1c00673
Anal. Chem. 2021, 93, 6254−6261

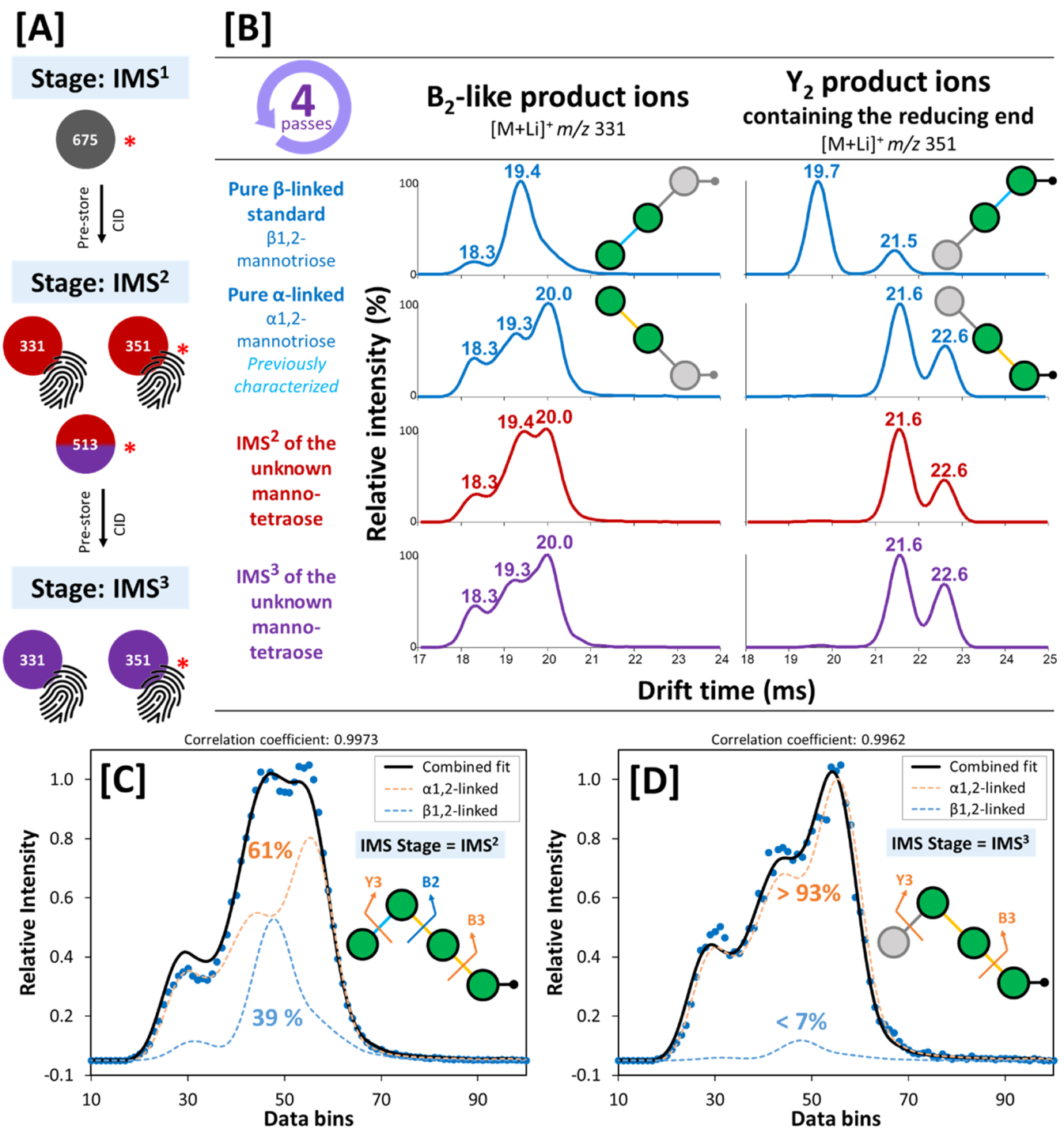


Figure 2. (A) Schematic representation of the IMS^n selection strategy and of the analyzed mannotetraose fragments. Ions bearing the ^{18}O-labeled reducing end are indicated with a red asterisk and fragments used for anomerism determination are indicated with a fingerprint. (B) DTDs of the IMS^2 fragments from pure β1,2- and α1,2-linked mannotriose (blue) of the IMS^2 fragments from the mannotetraose produced by Uhgb_MS (red) and of the IMS^3 fragments produced after refragmentation of the Y_3 mannotetraose fragment (violet), analyzed as $[M + Li]^+$ adducts. The nonreducing end only yielded B_2 fragments of interest (left column), while the reducing end only yielded Y_2 fragments (right column). (C, D) Experimental distribution fits of (C) IMS^2 and (D) IMS^3 fragments of mannotetraose at m/z 331, with the B_2 fragments from α1,2- and β1,2-mannotrioses. The contributions of the α-linked and β-linked fragments are optimized to return the combined fit best matched to the experimental data of mannotetraose. Insets show cleavages yielding m/z 331.

the dimeric fragment at m/z 331, which DTD differed from those of the B_2 fragments of both α1,2- and β1,2-mannotrioses. The DTD of the m/z 331 fragment arising from the mannotetraose presented two major peaks at 19.4 and 20.0 ms and a minor peak at 18.3 ms. Because the peaks at 19.4 and 20.0 ms are the main peaks in the β-linked and α-linked references, respectively, this hinted toward a mixture of α- and β-linked fragments.

To verify this hypothesis, we fitted the reference B_2 fragments from α1,2-mannotriose and β1,2-mannotriose under the distribution of mannotetraose fragments at m/z 331 (Figure 2C). This deconvolution confirmed that there

https://doi.org/10.1021/acs.analchem.1c00673
Anal. Chem. 2021, 93, 6254−6261

were contributions of both types of anomerism in the profile, in a ratio of α/β 60:40. Yet the presence of two Hex4 products in the reaction mixture, bearing different anomerisms at the non-reducing end (i.e., Man-(β1,2)-Man-(α1,2)-Man-(α1,2)-Man and Man-(α1,2)-Man-(β1,2)-Man-(α1,2)-Man), appeared rather unlikely given the sharp IMS peak measured with high resolving power for this compound (see Figure S5). Instead, our hypothesis was that the m/z 331 fragment had two origins: the B_2 fragment of the mannotetraose and a Y_3 + B_3 fragment arising from consecutive fragmentation. Consecutive fragmentations are known to occur during collisional dissociation.

IMS^n allowed us to confirm the above hypothesis: by selecting the Y_3 product ion (m/z 513) of the mannotetraose in the pre-store and re-fragmenting it, we avoided the production of the B_2 fragment and exclusively generated Y_3 + B_3 at m/z 331 (Figure 2B, violet trace), allowing the selective anomeric characterization of the "middle" linkage. The DTD of this Y_3 + B_3 fragment showed three main peaks at 18.3, 19.3, and 20.0 ms, and was quite similar to the B_2 fragment of α1,2-mannotriose. This DTD was also subjected to a distribution fitting for comparison with the reference fragments (Figure 2D). The fit obtained by minimizing the differences between the IMS^3 fragment and the sum of the references showed a correspondence of over 93% with the α1,2-linked B_2; but we also note that when forcing a 100% α-linked composition of the combined fit (Figure S8), the correlation coefficient is identical at the third decimal compared to Figure 2D. The Y_3 + B_3 product ion is thus undoubtedly linked in (α1,2), but with slight differences regarding the conformations' relative intensities between IMS^2 and IMS^3, Figure S8, the intensities of the conformations at 18.3 and 19.3 ms are increased in IMS^3, and this can also be observed on the Y_2 fragment in Figure 2B (the relative intensity of the conformation at 22.6 ms is higher in IMS^3).

We thus established that in mannotetraose, Y_2 is α-linked and Y_3 + B_3 is also α-linked. B_2 is, therefore, the only position left for the β linkage detected in IMS^2. The mannotetraose produced by Uhgb_MS was thus unequivocally characterized as Man-(β1,2)-Man-(α1,2)-Man-(α1,2)-Man.

This product was synthesized by Uhgb_MS by several reactions, involving reverse phosphorolysis and transmannosylation. Enzymes from the GH130 family being described as inverting enzymes, Uhgb_MS, probably used αM1P and the synthesized β1,2-mannobiose as mannosyl donors, respectively, to synthesize β and α linkages. Although the reactional kinetics and mechanisms will have to be further investigated, the present study is the first example of synthesis of mixed-linked oligosaccharides in a one-pot reaction involving one single enzyme.

CONCLUSIONS

The concepts of IMS/IMS of carbohydrates and anomeric retention of glycosidic fragments were previously demonstrated on commercial compounds, typically with the help of gas-phase infrared ion spectroscopy to infer the anomerism.[28,29] We have built our study on those concepts and showed that high-resolution IMS alone can sign, and thus infer, the intrachain anomerism on unknown glycans. The strategy exploits the possibility of performing multiple steps of mobility-based ion isolation followed by collisional fragmentation (IMS^n strategy). We first determined the intrachain anomerism of α1,2-mannotriose using IMS/IMS (IMS^2). Then, we extended the strategy to IMS^3 and discovered a unique mix-linked β1,2-α1,2-α1,2-mannotetraose. Our study demonstrates (i) that glycosidic fragments retain bond anomerism through multiple stages of collisional fragmentation and (ii) the applicability of standalone high-resolution IMS and IMS^n for the *de novo* characterization of tri- and tetrasaccharides produced in a raw medium by a newly discovered enzyme. The characterized glycans show potential interest as antigenic compounds. Furthermore, we found that the enzyme has original catalytic potentialities; in fact, for the first time, a single carbohydrate-active enzyme can synthesize both α- and β-glycosidic linkages. In future, this approach will have to be extended to more complex carbohydrates, with higher degrees of polymerization and/or lateral chains.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.1c00673.

Detailed parameters of the SELECT SERIES cyclic IMS; the experimental distribution fitting method; the flowchart of IMS^n experimental strategies; cIMS event sequences; HPAEC chromatograms; mass spectra (MS and MS/MS); the validation of ^{18}O labeling fragments; IMS^1 separation of oligosaccharides; B fragment structures; and comparison of the IMS^n spectra of fragments (PDF)

AUTHOR INFORMATION

Corresponding Authors

Gabrielle Potocki-Veronese − *TBI, CNRS, INRA, INSAT, Université de Toulouse, F-31077 Toulouse, France*; Email: veronese@insa-toulouse.fr

David Ropartz − *INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France*; orcid.org/0000-0003-4767-6940; Email: david.ropartz@inrae.fr

Authors

Simon Ollivier − *INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France*; orcid.org/0000-0002-7671-1736

Laurence Tarquis − *TBI, CNRS, INRA, INSAT, Université de Toulouse, F-31077 Toulouse, France*

Mathieu Fanuel − *INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France*; orcid.org/0000-0001-8384-8266

Ao Li − *TBI, CNRS, INRA, INSAT, Université de Toulouse, F-31077 Toulouse, France*

Julien Durand − *TBI, CNRS, INRA, INSAT, Université de Toulouse, F-31077 Toulouse, France*

Elisabeth Laville − *TBI, CNRS, INRA, INSAT, Université de Toulouse, F-31077 Toulouse, France*

Hélène Rogniaux − *INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France*; orcid.org/0000-0001-6083-2034

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.1c00673

https://doi.org/10.1021/acs.analchem.1c00673
Anal. Chem. 2021, 93, 6254−6261

Author Contributions

This manuscript was written through the contributions of all authors. All authors have approved the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge financial support from the Agence Nationale de la Recherche (grants ANR-18-CE29-0006 and ANR-16-CE20-0006) and the European Union's Horizon 2020 framework program (LEIT-BIO-2015-685474, Metafluidics). A.L. was supported by INSA Toulouse and the China Scholarship Council under the "CSC—UT/INSA Program".

■ REFERENCES

(1) Laine, R. *Glycobiology* **1994**, *4*, 759−767.
(2) Gray, C. J.; Migas, L. G.; Barran, P. E.; Pagel, K.; Seeberger, P. H.; Eyers, C. E.; Boons, G.-J.; Pohl, N. L. B.; Compagnon, I.; Widmalm, G.; Flitsch, S. L. *J. Am. Chem. Soc.* **2019**, *141*, 14463−14479.
(3) McLuckey, S. A.; Mentinova, M. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 3−12.
(4) Ropartz, D.; Giuliani, A.; Fanuel, M.; Herve, C.; Czjzek, M.; Rogniaux, H. *Anal. Chim. Acta* **2016**, *933*, 1−9.
(5) Leach, F. E.; Wolff, J. J.; Laremore, T. N.; Linhardt, R. J.; Amster, I. J. *Int. J. Mass Spectrom.* **2008**, *276*, 110−115.
(6) Yu, X.; Huang, Y.; Lin, C.; Costello, C. E. *Anal. Chem.* **2012**, *84*, 7487−7494.
(7) Wolff, J. J.; Leach, F. E.; Laremore, T. N.; Kaplan, D. A.; Easterling, M. L.; Linhardt, R. J.; Amster, I. J. *Anal. Chem.* **2010**, *82*, 3460−3466.
(8) Ropartz, D.; Li, P.; Jackson, G. P.; Rogniaux, H. *Anal. Chem.* **2017**, *89*, 3824−3828.
(9) Buck-Wiese, H.; Fanuel, M.; Liebeke, M.; Le Mai Hoang, K.; Pardo-Vargas, A.; Seeberger, P. H.; Hehemann, J.-H.; Rogniaux, H.; Jackson, G. P.; Ropartz, D. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 1249−1259.
(10) Wei, J.; Tang, Y.; Bai, Y.; Zaia, J.; Costello, C. E.; Hong, P.; Lin, C. *Anal. Chem.* **2019**, *92*, 782−791.
(11) Barnes, L.; Allouche, A.-R.; Chambert, S.; Schindler, B.; Compagnon, I. *Int. J. Mass Spectrom.* **2020**, *447*, No. 116235.
(12) Hofmann, J.; Pagel, K. *Angew. Chem., Int. Ed.* **2017**, *56*, 8342−8349.
(13) Gabelica, V.; Marklund, E. *Curr. Opin. Chem. Biol.* **2018**, *42*, 51−59.
(14) Gabelica, V.; Shvartsburg, A. A.; Afonso, C.; Barran, P.; Benesch, J. L. P.; Bleiholder, C.; Bowers, M. T.; Bilbao, A.; Bush, M. F.; Campbell, J. L.; Campuzano, I. D. G.; Causon, T.; Clowers, B. H.; Creaser, C. S.; De Pauw, E.; Far, J.; Fernandez-Lima, F.; Fjeldsted, J. C.; Giles, K.; Groessl, M.; Hogan Jr, C. J.; Hann, S.; Kim, H. I.; Kurulugama, R. T.; May, J. C.; McLean, J. A.; Pagel, K.; Richardson, K.; Ridgeway, M. E.; Rosu, F.; Sobott, F.; Thalassinos, K.; Valentine, S. J.; Wyttenbach, T. *Mass Spectrom. Rev.* **2019**, *38*, 291−320.
(15) Gray, C. J.; Schindler, B.; Migas, L. G.; Pičmanová, M.; Allouche, A. R.; Green, A. P.; Mandal, S.; Motawia, M. S.; Sánchez-Pérez, R.; Bjarnholt, N.; Møller, B. L.; Rijs, A. M.; Barran, P. E.; Compagnon, I.; Eyers, C. E.; Flitsch, S. L. *Anal. Chem.* **2017**, *89*, 4540−4549.
(16) Hofmann, J.; Hahm, H. S.; Seeberger, P. H.; Pagel, K. *Nature* **2015**, *526*, 241−244.
(17) Merenbloom, S. I.; Glaskin, R. S.; Henson, Z. B.; Clemmer, D. E. *Anal. Chem.* **2009**, *81*, 1482−1487.
(18) Giles, K.; Ujma, J.; Wildgoose, J.; Pringle, S.; Richardson, K.; Langridge, D.; Green, M. *Anal. Chem.* **2019**, *91*, 8564−8573.
(19) Ropartz, D.; Fanuel, M.; Ujma, J.; Palmer, M.; Giles, K.; Rogniaux, H. *Anal. Chem.* **2019**, *91*, 12030−12037.
(20) Tolmachev, A. V.; Webb, I. K.; Ibrahim, Y. M.; Garimella, S. V. B.; Zhang, X.; Anderson, G. A.; Smith, R. D. *Anal. Chem.* **2014**, *86*, 9162−9168.
(21) Deng, L.; Ibrahim, Y. M.; Baker, E. S.; Aly, N. A.; Hamid, A. M.; Zhang, X.; Zheng, X.; Garimella, S. V. B.; Webb, I. K.; Prost, S. A.; Sandoval, J. A.; Norheim, R. V.; Anderson, G. A.; Tolmachev, A. V.; Smith, R. D. *ChemistrySelect* **2016**, *1*, 2396−2399.
(22) Deng, L. L.; Webb, I. K.; Garimella, S. V. B.; Hamid, A. M.; Zheng, X. Y.; Norheim, R. V.; Prost, S. A.; Anderson, G. A.; Sandoval, J. A.; Baker, E. S.; Ibrahim, Y. M.; Smith, R. D. *Anal. Chem.* **2017**, *89*, 4628−4634.
(23) Deng, L.; Garimella, S. V. B.; Hamid, A. M.; Webb, I. K.; Attah, I. K.; Norheim, R. V.; Prost, S. A.; Zheng, X.; Sandoval, J. A.; Baker, E. S.; Ibrahim, Y. M.; Smith, R. D. *Anal. Chem.* **2017**, *89*, 6432−6439.
(24) Nagy, G.; Attah, I. K.; Garimella, S. V. B.; Tang, K.; Ibrahim, Y. M.; Baker, E. S.; Smith, R. D. *Chem. Commun.* **2018**, *54*, 11701−11704.
(25) Ujma, J.; Ropartz, D.; Giles, K.; Richardson, K.; Langridge, D.; Wildgoose, J.; Green, M.; Pringle, S. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 1028−1037.
(26) Warnke, S.; Faleh, A. B.; Scutelnic, V.; Rizzo, T. R. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 2204−2211.
(27) Schindler, B.; Barnes, L.; Renois, G.; Gray, C.; Chambert, S.; Fort, S.; Flitsch, S.; Loison, C.; Allouche, A. R.; Compagnon, I. *Nat. Commun.* **2017**, *8*, No. 973.
(28) Pellegrinelli, R. P.; Yue, L.; Carrascosa, E.; Warnke, S.; Ben Faleh, A.; Rizzo, T. R. *J. Am. Chem. Soc.* **2020**, *142*, 5948−5951.
(29) Bansal, P.; Yatsyna, V.; AbiKhodr, A. H.; Warnke, S.; Ben Faleh, A.; Yalovenko, N.; Wysocki, V. H.; Rizzo, T. R. *Anal. Chem.* **2020**, *92*, 9079−9085.
(30) Li, A.; Laville, E.; Tarquis, L.; Lombard, V.; Ropartz, D.; Terrapon, N.; Henrissat, B.; Guieysse, D.; Esque, J.; Durand, J.; Morgavi, D. P.; Potocki-Veronese, G. *Microb. Genomics* **2020**, *6*, No. mgen000404.
(31) Shibata, N.; Kobayashi, H.; Okawa, Y.; Suzuki, S. *Eur. J. Biochem.* **2003**, *270*, 2565−2575.
(32) Kobayashi, H.; Shibata, N.; Suzuki, S. *Infect. Immun.* **1992**, *60*, 2106−2109.
(33) Li, J.; Jia, H.; Cai, X.; Zhong, H.; Feng, Q.; Sunagawa, S.; Arumugam, M.; Kultima, J. R.; Prifti, E.; Nielsen, T.; Juncker, A. S.; Manichanh, C.; Chen, B.; Zhang, W.; Levenez, F.; Wang, J.; Xu, X.; Xiao, L.; Liang, S.; Zhang, D.; Zhang, Z.; Chen, W.; Zhao, H.; Al-Aama, J. Y.; Edris, S.; Yang, H.; Wang, J.; Hansen, T.; Nielsen, H. B.; Brunak, S.; Kristiansen, K.; Guarner, F.; Pedersen, O.; Doré, J.; Ehrlich, S. D.; Bork, P.; Wang, J. *Nat. Biotechnol.* **2014**, *32*, 834−841.
(34) Chiku, K.; Nihira, T.; Suzuki, E.; Nishimoto, M.; Kitaoka, M.; Ohtsubo, K.; Nakai, H. *PLoS One* **2014**, *9*, No. e114882.
(35) Ropartz, D.; Lemoine, J.; Giuliani, A.; Bittebiere, Y.; Enjalbert, Q.; Antoine, R.; Dugourd, P.; Ralet, M.-C.; Rogniaux, H. *Anal. Chim. Acta* **2014**, *807*, 84−95.
(36) Polasky, D. A.; Dixit, S. M.; Fantin, S. M.; Ruotolo, B. T. *Anal. Chem.* **2019**, *91*, 3147−3155.
(37) Domon, B.; Costello, C. E. *Glycoconjugate J.* **1988**, *5*, 397−409.
(38) Zhang, H.; Brokman, S. M.; Fang, N.; Pohl, N. L.; Yeung, E. S. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 1579−1586.
(39) Bythell, B. J.; Abutokaikah, M. T.; Wagoner, A. R.; Guan, S.; Rabus, J. M. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 688−703.

https://doi.org/10.1021/acs.analchem.1c00673
Anal. Chem. 2021, 93, 6254−6261

Chapter 7 – Separation of underivatized monosaccharide stereoisomers

A. Context

In **Chapter 6**, we used cIMS to characterize *de novo* a case of stereoisomerism within oligosaccharides (α/β anomerism of the linkage) by exploiting a comparison to disaccharide standards. However, this strategy appears difficult to generalize because of the very high number of possible disaccharides, because of the difficulty—if not the impossibility—to have all relevant standards available as pure compounds, and because of the amount of work that would be required to assemble an *ad hoc* database.

On the other hand, there are much fewer monosaccharides and building a monomer database seems attainable. It also seems like an ideal way to meet our intentions of developing a cIMS-based sequencing method. As described in **Chapter 1**, the monosaccharide composition of poly/oligosaccharides is often determined using GC approaches, but this requires hydrolyzing the saccharide and the connectivity between the building blocks is lost. It would thus be interesting to develop a sequencing approach in which the monosaccharides would be sequentially cleaved from the oligosaccharide directly in the gas phase, thereby keeping the information of their connectivity.

However, a prerequisite to such cIM-based sequencing approach is *to be able to establish specific signatures for each monosaccharide*. Furthermore, the conditions in which these monosaccharides' signatures are recorded must be compatible with fragmentation conditions.

Previous attempts have been made at generating fingerprints of monosaccharides using cIMS. McKenna *et al.*[146] proposed a double-derivatization approach with 3-carboxy-5-nitrophenyl boronic acid[q] (3C5NBA), that afforded fingerprints of D-hexoses isomers after three IM passes. More recently, Williamson and Nagy[128] worked on hexoses adducted to α-cyclodextrin[r] (αCD) as $[M + 2\alpha CD - 2H]^{2-}$ and achieved near baseline separation of the D/L enantiomers after 25 IM passes. The use of chemical groups such as 3C5NBA or αCD increases the mass of monosaccharides up to several times, facilitating their analysis under IMS conditions (heavier species are more stable compared to lighter species), and helps separate the species. However, this strategy is not ideal for a *de novo* sequencing approach in which the use of CID-IMS appears to be essential, because chemical groups could be detached under CID

[q] 3-carboxy-5-nitrophenyl boronic acid: molecular formula = $C_7H_6BNO_6$, monoisotopic mass = 211.0288
[r] αCD is a cyclic hexasaccharide: molecular formula = $C_{36}H_{60}O_{30}$, monoisotopic mass = 972.3169

conditions. The use of adduction with “big” metal ions (K, Ca, Mg, etc.) is likewise limited in that context because the metal adduct must remain stable under CID conditions.

B. Objectives

With the intention to develop a cIMS-based *de novo* sequencing method as part of my PhD, the present work aimed at providing IM separation conditions for some of the most common monosaccharides from plant or bacterial poly- and oligosaccharides.

The study was performed on underivatized monosaccharides, which are subject to mutarotation and thus potentially contain all α/β or open chain forms, as well as pyranose/furanose forms. Therefore, we did not try to achieve baseline separation of the isomers but rather *to obtain distinctive fingerprints for each monosaccharide*.

It is noteworthy that the context of the study implies some restrictions on the TW parameters. A compromise must be reached between optimal separation of the monosaccharides and retaining the ability to separate larger oligosaccharides for use of IMS^n sequences in the further sequencing strategy.

C. Methods

1. Chemicals

The study was performed using analytical grade monosaccharide standards, all above 98 % purity. D-galactose (Gal), D-glucose (Glc), D-mannose (Man), L-arabinose (Ara), L-ribose (Rib), D-xylose (Xyl), L-fucose (Fuc), D-galacturonic acid (GalA), D-glucuronic acid (GlcA), N-acetyl-D-galactosamine (GalNAc), N-acetyl-D-glucosamine (GlcNAc), and lithium chloride (LiCl) were purchased from Sigma-Aldrich (Saint Quentin Fallavier, France). L-rhamnose (Rha) was purchased from Alfa Aesar (Kandel, Germany). HPLC grade methanol was purchased from Carlo-Erba (Val de Reuil, France). Ultrapure water was produced using a Milli-Q system (Millipore, Burlington, MA, USA).

2. IMS-MS analyses

Pure samples were infused at a concentration of 1µg/mL in 50:50 water/methanol, using a syringe pump set to a flow rate of 5 µL/min. Lithium adducts were generated by doping the solvents with 0.1 mM of LiCl, except in the case of hexoses for which 0.01 mM LiCl was used. Ionization in the Z-spray ESI source was performed with the following parameters: capillary voltage 3.0 kV, cone voltage 60 V, source temperature 100 °C, desolvation gas temperature 300 °C, nebulizer pressure 6 Bars. Details about the voltages and pressures applied in the other parts of the mass spectrometer are available in **Appendix 6, Supplementary text 1**. The optimized traveling wave parameters were: (i) for classical separations, a TW wave height

(WH) of 16 V at a wave velocity (WV) of 350 m/s, (ii) for TW ramps an initial WH of 10 V with an increament of 0.2 V/ms up to a maximum of 22 V, still at a WV of 350 m/s.

MS and IMS-MS spectra were recorded on the cIMS instrument using the Quartz software (Waters Embedded Analyzer release 6, Waters, Wilmslow, UK). Data were processed using MassLynx 4.2 and Driftscope 2.9 (both from Waters, Wilmslow, UK). ATDs were smoothed using the Savitzky-Golay algorithm (window size (scan) ± 1, number of smooths: 2).

D. Results and discussion

1. Design of IMS strategies for separating monosaccharides

The first step in designing sequences to separate monosaccharides was to optimize the traveling wave parameters. Key TW parameters notably include the wave velocity (WV) and wave height (WH). We chose to keep the WV at 350 m/s, which is efficient for small and medium-sized ions (*i.e.*, mono- and oligosaccharides). The main parameter we played with to optimize the separation was thus the WH.

The cIMS setup offers two different functionalities to set up the WH in the mobility cell:

- The first option, the static WH, is the most used by the TWIMS community. As shown in **Figure 56**, a single optimized WH voltage is chosen and the WH is kept constant during the entire duration of the IMS separation. This static mode offers many advantages, especially in cIMS instruments. Because the separation parameters remain the same throughout the time the ions are submitted to the TW, their velocity is directly related to the ion path length. In cIMS, this translates to a constant single-pass drift time, *i.e.*, the time separating the 1-pass arrival time from the 2-pass arrival time is the same as that separating the 2-pass and 3-pass arrival times. This allows predicting the separation time required to reach a given number of passes, and to precisely recalculate the single-pass drift time.

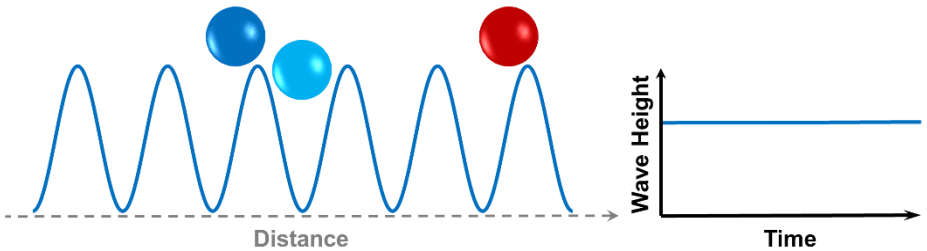


Figure 56. Static traveling wave separation. The WH remains constant over time.

- The second option is to ramp the traveling wave height. The initial WH voltage is kept relatively low, and the WH will be increased every millisecond using a user-defined

slope (**Figure 57**) until a maximum WH value is reached. This option is especially useful for separating species that are not separated using a static TW, either because the WH is too low (the different populations diffuse in the TWIMS without being trapped in waves, so they are not separated) or because the WH is too high (the different populations are trapped in the same wave). By ramping the WH voltages, it is almost certain that an efficient WH value will be reached for the different species at some point. This mode, however, has some disadvantages, clearly visible in a cIMS setup. Because the TW parameters keep changing, the velocity at which the ions undergo roll-over varies over time, resulting in every pass around the cyclic racetrack taking a different time. The change in TW parameters over time can also result in slightly asymmetrical peaks. This makes it impossible to fit the ATD data with gaussians curves as we did in the previous Chapters.

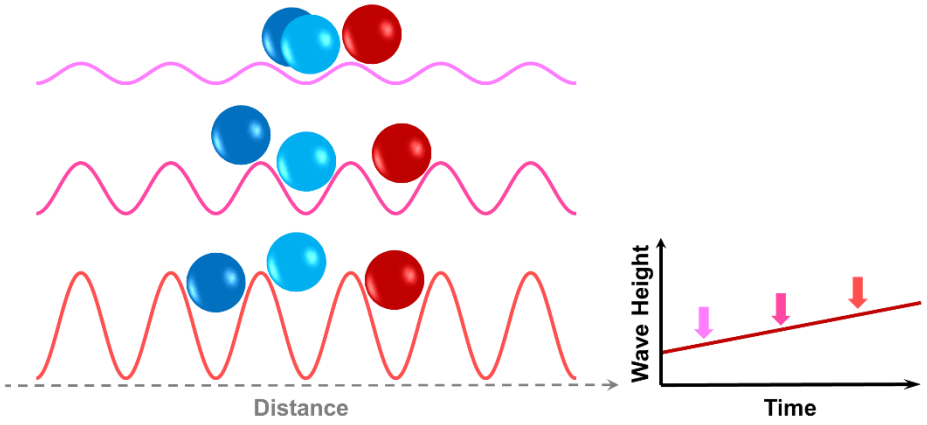


Figure 57. Separation with a ramping of the traveling wave height. Different WH are applied over time.

In a preliminary test conducted on hexoses, we evaluated the above two strategies using single-pass and multi-pass approaches. To evaluate the separation, we chose as the criterion that peak-picking algorithms must be able to distinguish at least one specific peak for each isomer. As mentioned above, gaussian fitting is not suited to separation with ramping WH, so we used the peak detection feature of Driftscope, which accounts for peak tailing. In the case of hexoses, the results obtained using either static or ramping traveling wave were not satisfactory (data not shown). For instance in static mode, some features of the ATDs were well resolved but others were not separated at all—likely because the oligosaccharide-compatible WH value was too high and the ions were trapped in the same wave. In the ramping mode, most of the features were separated but not enough to be distinguishable in a mixture. We thus endeavored to develop a method that could give better-resolved profiles for hexoses.

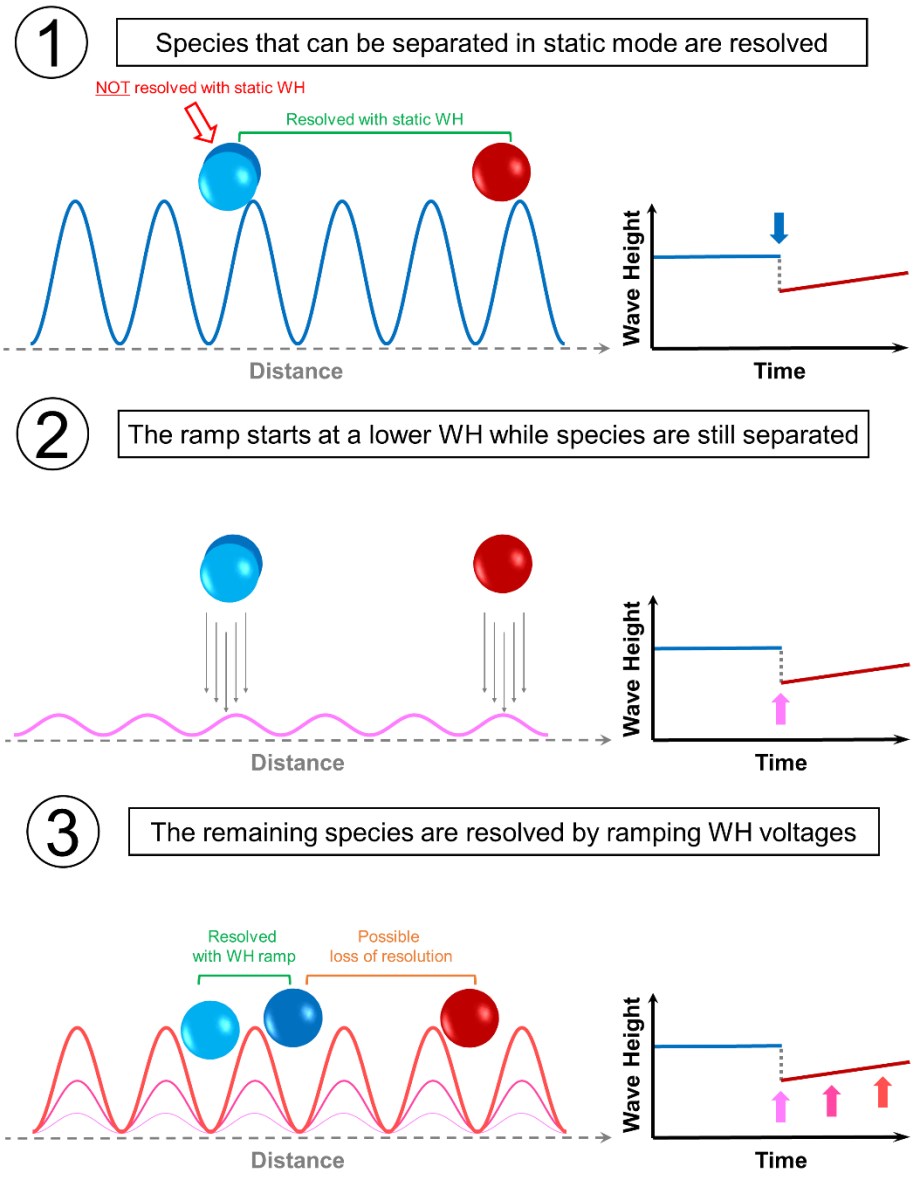


Figure 58. Schematic of the strategy combining static and ramping wave height separations to resolve additional populations

This novel method, described in **Figure 58**, is a combination of the static and ramping WH sequences. In a first step of separation, a static WH is applied until the best separation is obtained on the peaks that can be separated (red vs blue). This is a crucial step because the following step with ramping WH is likely to lower that initial separation. However, it is important not to use an excessive separation here so that ions do not wrap-around during the ramping passes. After the first static WH separation is performed, the WH is lowered to a level at which the unseparated ions (light vs dark blue) are not trapped in the same wave anymore. A WH ramp is then used to gradually get to WH levels were the blue ions are efficiently separated. At the end, this gives an ATD in which all species are resolved, with (i) the blue populations

being separated thanks to the WH ramp and (ii) the blue and red populations being separated with a slightly worse—but still satisfactory—resolution compared to static WH strategies.

2. Separation of lithiated monosaccharide stereoisomers

In the next step, we acquired IM fingerprints for a series of 13 monosaccharides representative of plant or bacterial polysaccharides, which will be presented by isomeric family (**Figure 59**). Different types of ions were investigated. Indeed, it is well documented that using different metal counter-ions yields distinctive ATDs, related to the size of the counter ion.[113,114,147] However, considering our aim of using these fingerprints for sequencing purposes, we determined which of Li^+ and Na^+ counter ions yielded the highest proportion of detached monosaccharide upon fragmentation. CID experiments on a glucose homo-oligomer (cellopentaose, **Appendix 6, Figure S1**) showed that Li^+ adducts yielded more monosaccharidic fragment compared to Na^+ adducts (which released practically no monosaccharide). We thus opted to work on lithium adducts in the following experiments.

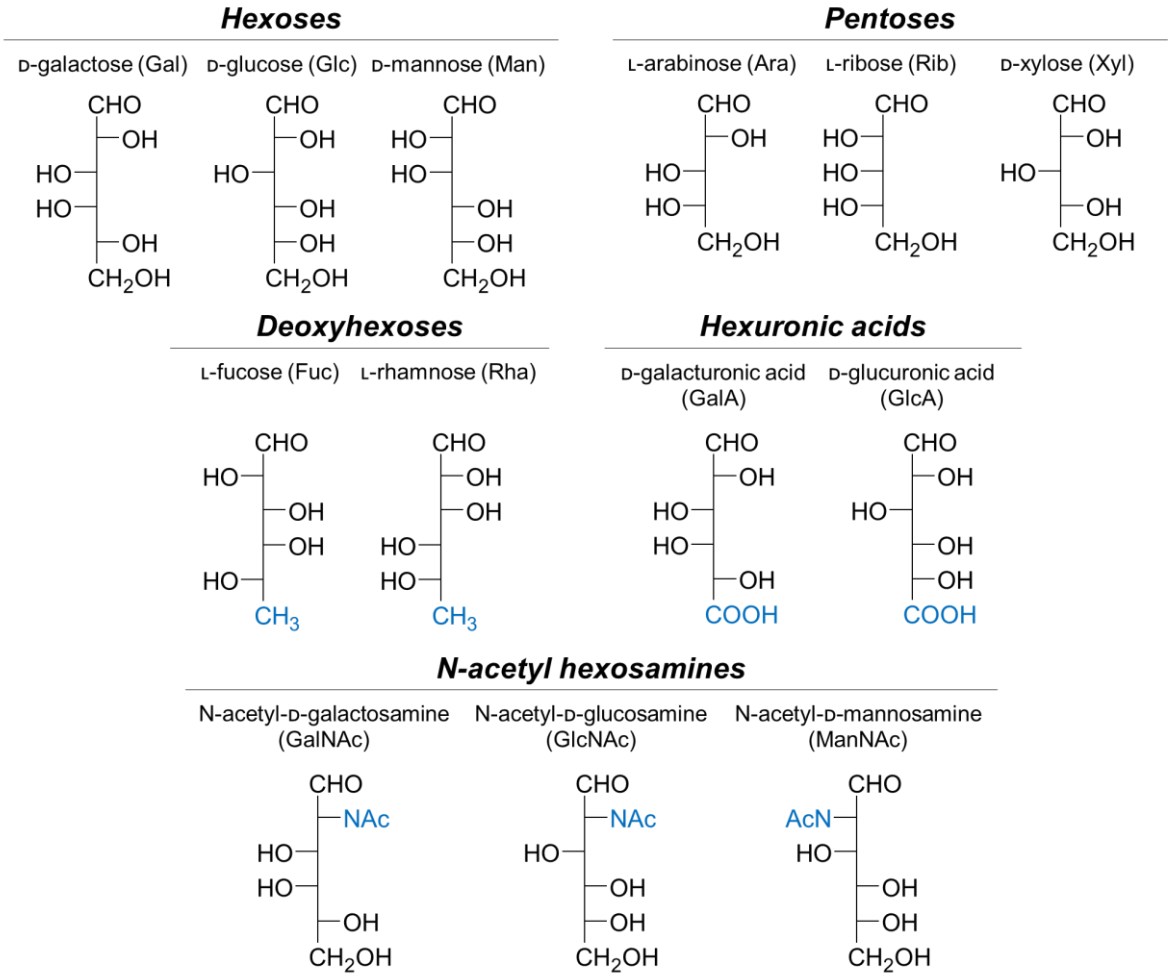


Figure 59. Fischer representation of the 13 studied monosaccharides, sorted by isomeric family

a) Hexoses[s]

We studied three hexoses monosaccharides: glucose (Glc), mannose (Man) and galactose (Gal). It is noteworthy that Man and Gal are both epimers of glucose, in respectively C-2 and C-4. We established specific profiles for these species using a 4-pass static WH separation, followed by a further 2-pass separation with a ramping WH.

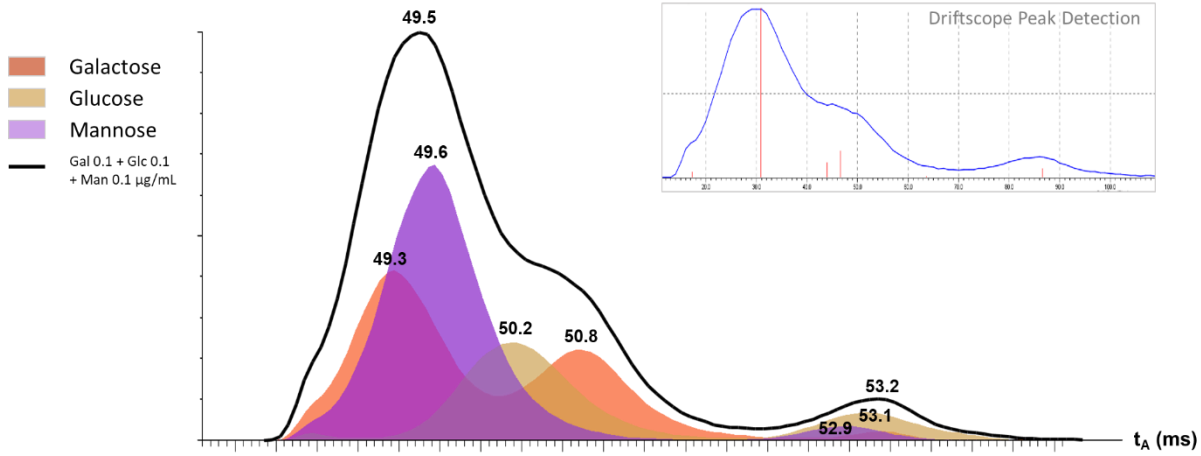


Figure 60. Arrival time distributions for the Glc, Gal and Man monosaccharides (colored), and for an equimolar mixture of the three hexoses at 0.1 µg/mL (solid black line). Inset: result of the Drifscope peak detection for the mixture.

The ATDs of the hexoses are presented in **Figure 60**. All three species present specific bimodal distributions. Glc exhibits two peaks at 50.2 and 53.1 ms, Man exhibits two peaks at 49.6 and 52.9 ms, and Gal exhibits two peaks at 49.3 and 50.8 ms.

The equimolar mixture of these monosaccharides exhibited a multimodal distribution (black trace) in which the Driftscope peak-picking algorithm was able to distinguish the specific features at 49.6, 50.2 and 50.8 ms (inset at the top right). The algorithm also picked a peak at 53 ms but this is not a distinctive feature as this peak is shared between Glc and Man.

In summary, the sequence that we developed by combining steps of static WH and WH ramping yielded at least one distinctive peak for each hexose monosaccharide.

[s] For a hexose: molecular formula $C_6H_{12}O_6$, monoisotopic mass for $[M + Li]^+$ 187.0788

b) Pentoses[t]

In the case of pentoses, we analyzed arabinose (Ara), xylose (Xyl) and ribose (Rib). Ara and Rib are epimers in C-2. Fingerprinting of the pentose species was done using a 4-pass ramping separation.

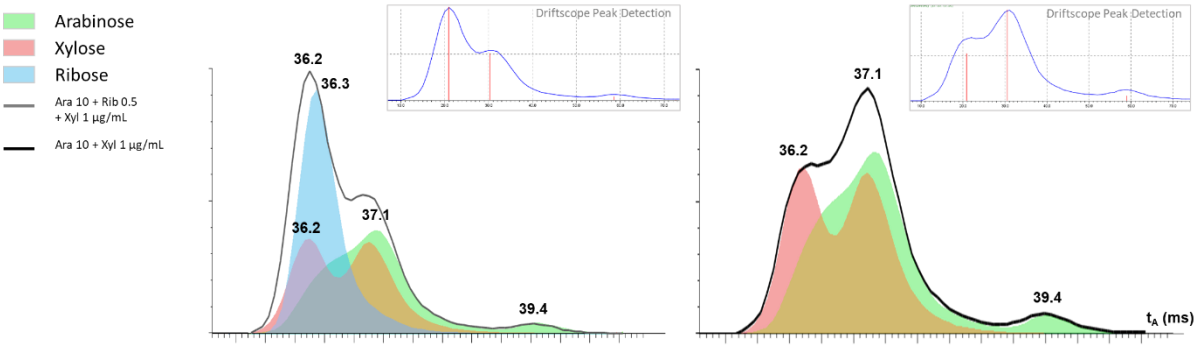


Figure 61. Arrival time distributions for the Ara, Xyl and Rib monosaccharides (colored), and for mixtures of Ara 10 µg/mL, Xyl 1 µg/mL and Rib 0.5 µg/mL (left panel) or only Ara 10 µg/mL and Xyl 1 µg/mL (right panel) (solid black lines). Inset: results of the Drifscope peak detections for the mixtures.

The ATDs are presented in **Figure 61**. Contrary to hexoses, we did not observe a systematic bimodal distribution. Rib exhibited a single peak at 36.3 ms, Xyl exhibited a bimodal distribution at 36.2 and 37.1 ms, and Ara exhibited a multimodal pattern with peaks at 37.1 and 39.4 ms and a shoulder around 36.5 ms.

It is also noteworthy that the three pentoses were likely subject to ionization competition, forbidding us from analyzing an equimolar mixture of the three. Instead, we analyzed non-equimolar mixtures with a notably higher concentration of Ara compared to Rib and Xyl. In the mixture containing all three monosaccharides, the Driftscope peak-picking algorithm detected features at 36.2, 37.1 and 39.4 ms. Of these, only the feature at 39.4 is distinctive of Ara, the feature at 36.3 being shared between Rib and Xyl and that at 37.1 between Ara and Xyl.

However, Rib is rarely found in polysaccharides in conjunction with Ara and Xyl and plant polysaccharides almost exclusively contain Ara and Xyl. When considering a mixture containing only Ara and Xyl, the peak detection algorithm points one peak at 36.2 ms, which is specific of Xyl, one peak at 37.1 ms, which is nonspecific, and one peak at 39.4 ms, which is specific of Ara. We thus considered these conditions were satisfactory for our purposes. In fact, they allow to clearly distinguish Ara and Xyl, which can be found together within plant polysaccharides.

[t] For a pentose: molecular formula $C_5H_{10}O_5$, monoisotopic mass for $[M + Li]^+$ 157.0683

c) Deoxyhexoses (a.k.a. methylpentoses)[u]

In the case of deoxyhexoses, we studied fucose (Fuc) and rhamnose (Rha). Like for the pentoses, fingerprinting was performed using a 4-pass ramping separation.

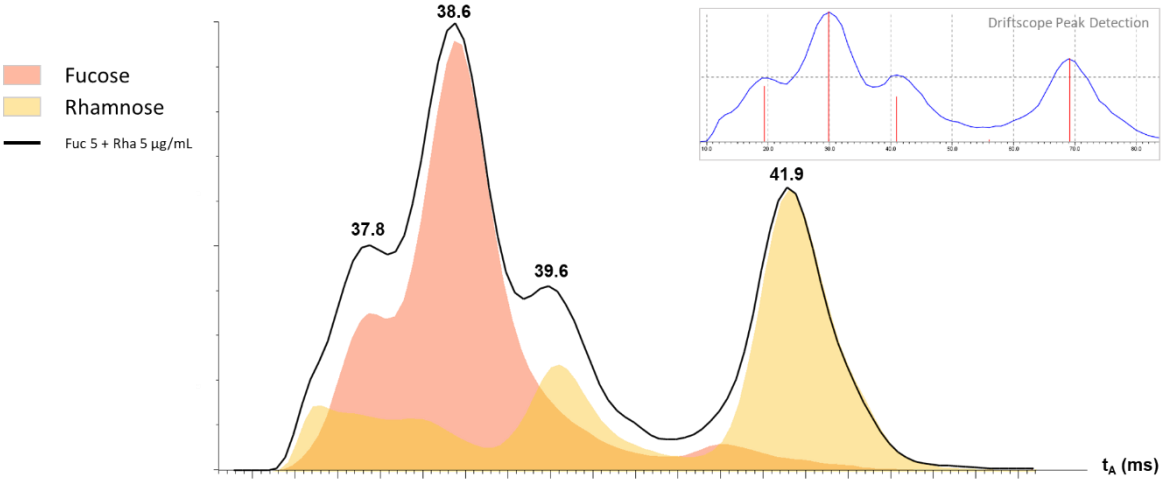


Figure 62. Arrival time distributions for the Fuc and Rha monosaccharides (colored), and for an equimolar mixture of the two deoxyhexoses at 5 µg/mL (solid black line). Inset: result of the Drifscope peak detection for the mixture.

The ATDs of the deoxyhexoses are presented in **Figure 62**. Both Fuc and Rha exhibited multimodal ATDs, with two major peaks each. For Fuc, the major peaks had arrival times of 37.8 and 38.6 ms. For Rha, they had arrival times of 39.6 and 41.9 ms.

The equimolar mixture of Fuc and Rha exhibited a multimodal distribution, but the different components can be clearly distinguished and the Driftscope peak picking had no trouble pointing out the four major features (*i.e.*, detected features at 37.8, 38.6, 39.6, and 41.9 ms) corresponding to the two monosaccharides.

d) Hexuronic acids (HexA)[v]

In the case of hexuronic acids, we studied glucuronic acid (GlcA) and galacturonic acid (GalA). Hexuronic acid are derivatives of hexoses in which the hydroxymethyl group in C-5 is oxidized, giving a carboxyl group. Like Glc and Gal, GlcA and GalA are epimers in C-4. Fingerprinting of the HexA species was done using a 4-pass ramping separation.

[u] For a deoxyhexose: molecular formula $C_6H_{12}O_5$, monoisotopic mass for $[M + Li]^+$ 171.0839
[v] For a hexuronic acid: molecular formula $C_6H_{10}O_7$, monoisotopic mass for $[M + Li]^+$ 201.0581

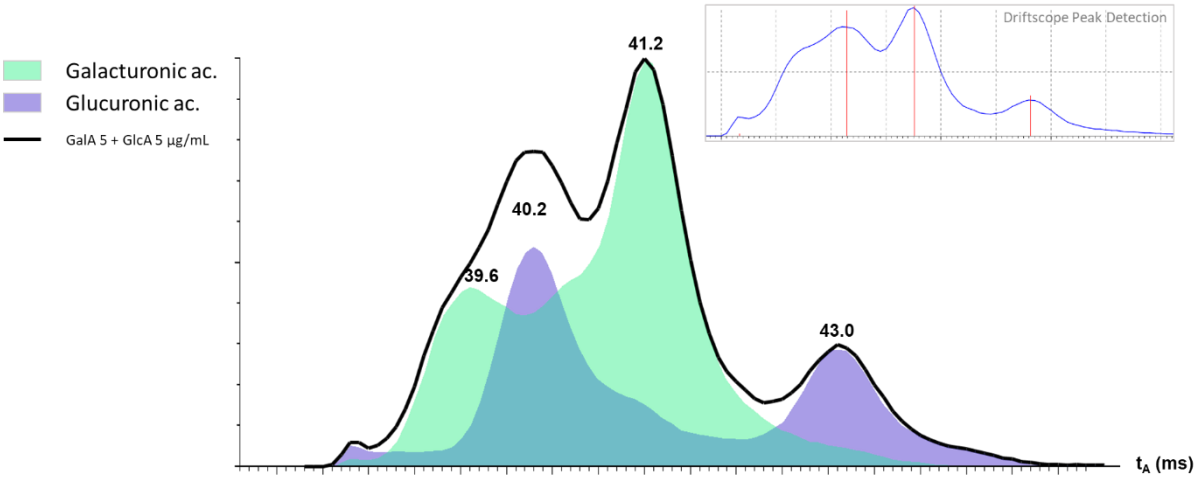


Figure 63. Arrival time distributions for the GlcA and GalA monosaccharides (colored), and for an equimolar mixture of the two HexA at 5 µg/mL (solid black line). Inset: result of the Drifscope peak detection for the mixture.

The ATDs of the HexA species are presented in **Figure 63**. GalA exhibited a multimodal distribution with two major features at 39.6 and 41.2 ms. GlcA also exhibited a multimodal distribution with two major features at 40.2 and 43.0 ms.

In the equimolar mixture of GalA and GlcA, the Driftscope peak picking was able to detect the feature at 41.2 ms for GalA, and both features at 40.2 and 43.0 ms for GlcA.

e) *N*-Acetyl hexosamines (HexNAc)[w]

N-Acetyl hexosamines are hexoses derivatives for which the hydroxyl group in C-2 is replaced by an *N*-acetylation, they are common in bacterial polysaccharides. *N*-acetyl glucosamine (GlcNAc), *N*-acetyl mannosamine (ManNAc) and *N*-acetyl galactosamine (GalNAc) were included in the study. Both ManNAc and GalNAc are epimers of GlcNAc in C-2 and C-4, respectively. HexNAc were the only family of monosaccharides for which we managed to separate the species with a static WH separation (5 passes).

[w] Molecular formula of an N-acetyl hexosamine: $C_8H_{15}NO_6$. Monoisotopic mass: $[M + Li]^+$ 228.1054

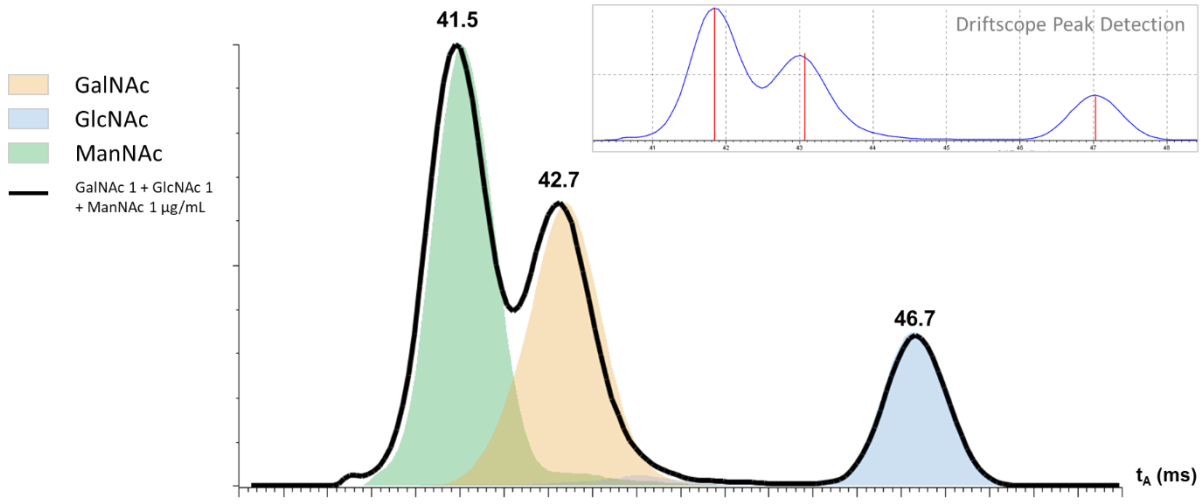


Figure 64. Arrival time distributions for the GlcNAc, ManNAc and GalNAc monosaccharides (colored), and for an equimolar mixture of the three HexNAc at 1 µg/mL (solid black line). Inset: result of the Drifscope peak detection for the mixture.

The ATDs of the HexNAc species are presented in **Figure 64**. Interestingly, despite being reducing saccharides, the HexNAc monosaccharides only exhibited one major peak each. This observation is unlike what we observed on all the previous types of reducing monosaccharides. One hypothesis could be the existence of an ultrapreferential coordination of the Li^+ cation involving the *N*-acetyl group, resulting in the observation of a single conformer. These major peaks were located at 46.7 ms for GlcNAc, 42.7 ms for ManNAc, and 41.5 ms for ManNAc.

All three species were relatively well resolved, and GlcNAc was notably baseline-resolved from GalNAc and ManNAc. They were thus readily distinguished in an equimolar mixture.

E. Conclusions

Table 1. Summary of the cIMS sequences used to separate the monosaccharides

Monosaccharide family	Separation type		Total separation time (ms)
	Static WH	Ramping WH	
Hexoses	4 passes (18.5 ms)	2 passes (19.1 ms)	37.6 ms
Pentoses	Ø	4 passes (24.5 ms)	24.5 ms
Deoxyhexoses	Ø	4 passes (26.1 ms)	26.1 ms
Hexuronic acids	Ø	4 passes (27.5 ms)	27.5 ms
N-acetyl hexosamines	5 passes (29.5 ms)	Ø	29.5 ms

In this Chapter, using different types of TWIMS configuration, including innovative sequences that combined static and ramping traveling waves, we were able to establish specific signatures for five isomeric families of monosaccharides, among the most common in plant and bacterial polysaccharides. Importantly, these signatures were obtained for underivatized monosaccharides. An overview of the sequences used to separate the different monosaccharide families (as $[M + Li]^+$ adducts) is given in **Table 1**.

Of note, most of the IM signature profiles were multimodal, but while multimodal distribution of alkali-cationized oligosaccharides have been previously linked to the α/β anomerism,[147] it seems unlikely in the case of our studied monosaccharides. Indeed, all of them were reducing saccharides (*i.e.*, subject to the mutarotation phenomenon thus to the α/β and pyranose/furanose equilibriums) but some exhibited a single IM peak, or more generally profiles that did not match their reported mutarotation equilibrium. For instance, ribose only presented one major peak whereas the reported abundances of its different forms are: α-ribopyranose 21,5%, β-ribopyranose 58.5%, α-ribofuranose 6,5%, and β-ribofuranose 13.5%.[148] This hints towards the observed profiles being more likely related to conformers rather than anomers—although a link between conformation and anomerism cannot be fully excluded.

It is noteworthy that, when purely considering the IM profiling of monosaccharides, the strategies that we developed in this Chapter can be extended to other types of adducts than $[M + Li]^+$ that we selected in this work because of its compatibility with sequencing approaches. For instance, IM profiles obtained on $[M + Na]^+$ adducts of hexoses and hexuronic acids are shown in **Appendix 6, Figures S2-3**.

We further showed that the signature peaks identified in the ATDs could be identified by processing software in mixtures of stereoisomeric monosaccharides, which might prove useful to deconvolve data generated in sequencing approaches (likely to contain the signatures of multiple monosaccharides). Fitting of experimental arrival time distributions similar to what we used in **Chapter 6** could also be envisioned (*i.e.*, non-gaussian fitting). After establishing the IM signatures for all of these monosaccharides, the next step will be to verify if monosaccharidic fragments generated from larger oligosaccharides match the drift time of these references.

F. Scientific Communication

These results will be described in an upcoming article, alongside our ongoing work on *de novo* sequencing of carbohydrates (see **Chapter 8**).

Part II – Conclusions

Cyclic IMS and the associated experiments (notably IMS^n) are still relatively recent approaches in analytical chemistry. Therefore, before we could address the question of the structural characterization of complex oligosaccharides, we ensured in **Part II** that cIMS could resolve both constitutional isomers and stereoisomers. We have demonstrated this capability through several studies that, together, have covered all types of isomerisms (**Figure 65**).

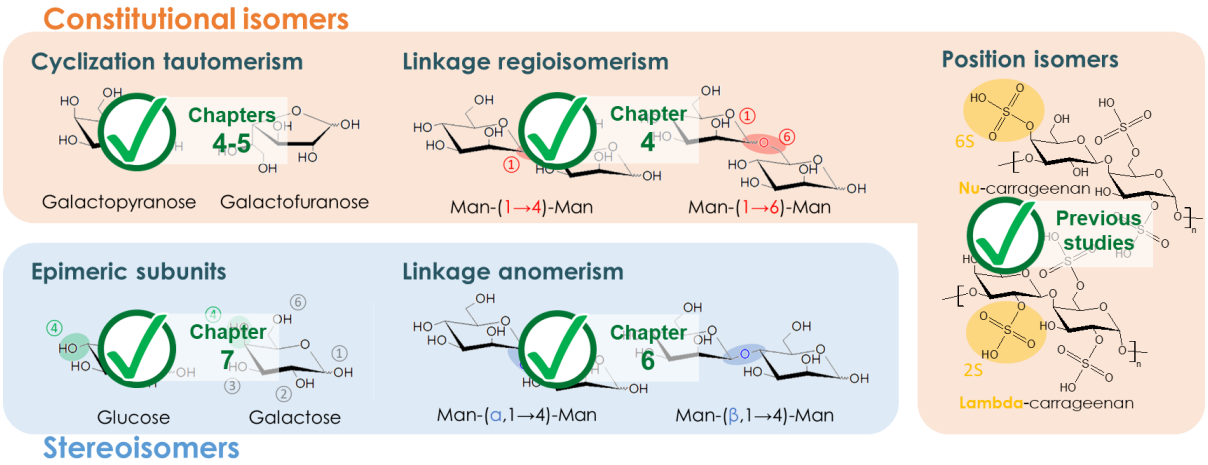


Figure 65. Cyclic IMS can retain and resolve all types of isomerism in oligosaccharides

First, I addressed constitutional isomerism through the study of a comprehensive library of Gal-Man disaccharides, notably containing Gal*f* units. We showed that cIMS could differentiate between all cases of linkage regioisomerism and furanose/pyranose tautomerism. However, an extended study of the less stable Gal*f*-containing oligosaccharides evidenced some structural rearrangements, which may complicate the structural study of Gal*f*-containing oligosaccharides. Importantly, the isomeric fragments hailing from the rearrangements could only be solved thanks to the high resolving power of cIMS.

Second, I addressed the case of stereoisomerism. I showed that the anomerism of the glycosidic bond was preserved upon multiple rounds of CID in IMS^n experiments and exploited this “anomeric memory” to perform the complete anomeric characterization of the products of a mannoside synthase, which was a first from a biological reaction medium. Then, I designed IMS sequences to separate underivatized, alkali-cationized monosaccharides, including monosaccharide epimers (e.g., Glc/Gal or Glc/Man). Sequences were designed for the main families of subunits found in plant or bacterial carbohydrates (hexoses, pentoses, deoxyhexoses, hexuronic acids, *N*-acetyl hexosamines). This work laid the foundation for the carbohydrate *de novo* sequencing strategy that we intended to develop in this PhD, which will be described in the next Part.

Part III – Can we achieve a complete structural determination of complex oligosaccharides using cIMS?

In **Part II** of the manuscript, we were able to assert that *cIMS—and the different experiments that can be associated with it, such as CID—are capable of retaining and resolving all types of isomerism that can be found in oligosaccharides.* This gave us the incentive to move on to another central question of this work: ***Can we achieve a complete structural determination of complex oligosaccharides using cIMS?***

The following Part of my PhD thesis therefore focuses on determining whether it is possible to characterize the full structure of oligosaccharides using cIMS (up to the stereochemistry of the subunits and of the glycosidic bond). This notably consists in exploring the properties of oligosaccharides upon fragmentation to assess if the fragments generated in IMS/IMS or IMS^n contain all of the required structural information.

Part III is divided into two Chapters that present the different strategies that I explored to try to resolve the complete structures of oligosaccharides. These Chapters each introduce a different strategy we attempted to use: (i) a targeted sequencing approach to resolve the structure of a specific oligosaccharide, and (ii) an untargeted networking approach to organize glycomics datasets. A final Chapter provides a synthesis and conclusion of these studies.

Chapter 8. **First insights towards complete oligosaccharide sequencing with cIMS**

Chapter 9. **Molecular Networking of tandem IMS data to organize glycomics datasets**

Synthesis and Conclusions of Part III

Chapter 8 – First insights towards complete oligosaccharide sequencing with cIMS

A. Context

By "complete sequencing" of an oligosaccharide, we mean not only the determination of the monosaccharide composition but also the manner (linkage regioisomerism and anomerism) and order in which these monosaccharides are connected in the molecule. Today, there is no satisfactory method to achieve this goal and a combination of several methods (including (NMR, GC-MS, etc.) is often necessary as described in **Chapter 1**. It should be noted that tandem MS, although widely used, has limitations in this question related to the possible isomerism of the fragments, which prevents in many cases the complete resolution of the sequence.

In **Chapter 7**, however, we established that IMS was able to produce discriminant signatures for several biological monosaccharides, which is a promising prerequisite for its integration into a sequencing strategy. We would also like to highlight the shotgun IMS-MS sequencing ($SIMMS^2$) method introduced by the group of Kevin Pagel[130] (that we already mentioned in **Chapter 2**), because it is highly relevant to our own work. In this work, Miller *et al*. used DTIMS profiles of heparan sulfate oligosaccharides (and their fragments) comparing them to fragments from larger oligosaccharides (notably the Y fragments containing the reducing end and the B fragments containing the non-reducing end). Starting with a comparison of small fragments with the corresponding oligosaccharides, and sequentially increasing the size of the studied references/fragments, they were able to attain good sequence coverage. In some cases, they were even able to use B_1 fragments (*i.e.*, dehydrated monosaccharides) in their assignments. The novel structure would then be added to the reference database.

Nevertheless, heparans are a very specific type of polysaccharide that contains a limited number of subunits (despite the various sulfation patterns). In the context of plant carbohydrates, given the large number of possible subunits, a sequencing method starting directly from the monosaccharides would be desirable. Indeed, building a database of even disaccharide signatures would represent a tremendous effort.

B. Objectives

In this Chapter, our aim was to develop a method of oligosaccharides sequencing following a principle similar to that described by Pagel and coll.[130] but relying directly on the database of

HR-IMS profiles of monosaccharides constructed in **Chapter 7**. The workflow we have envisioned is schematized in **Figure 66**.

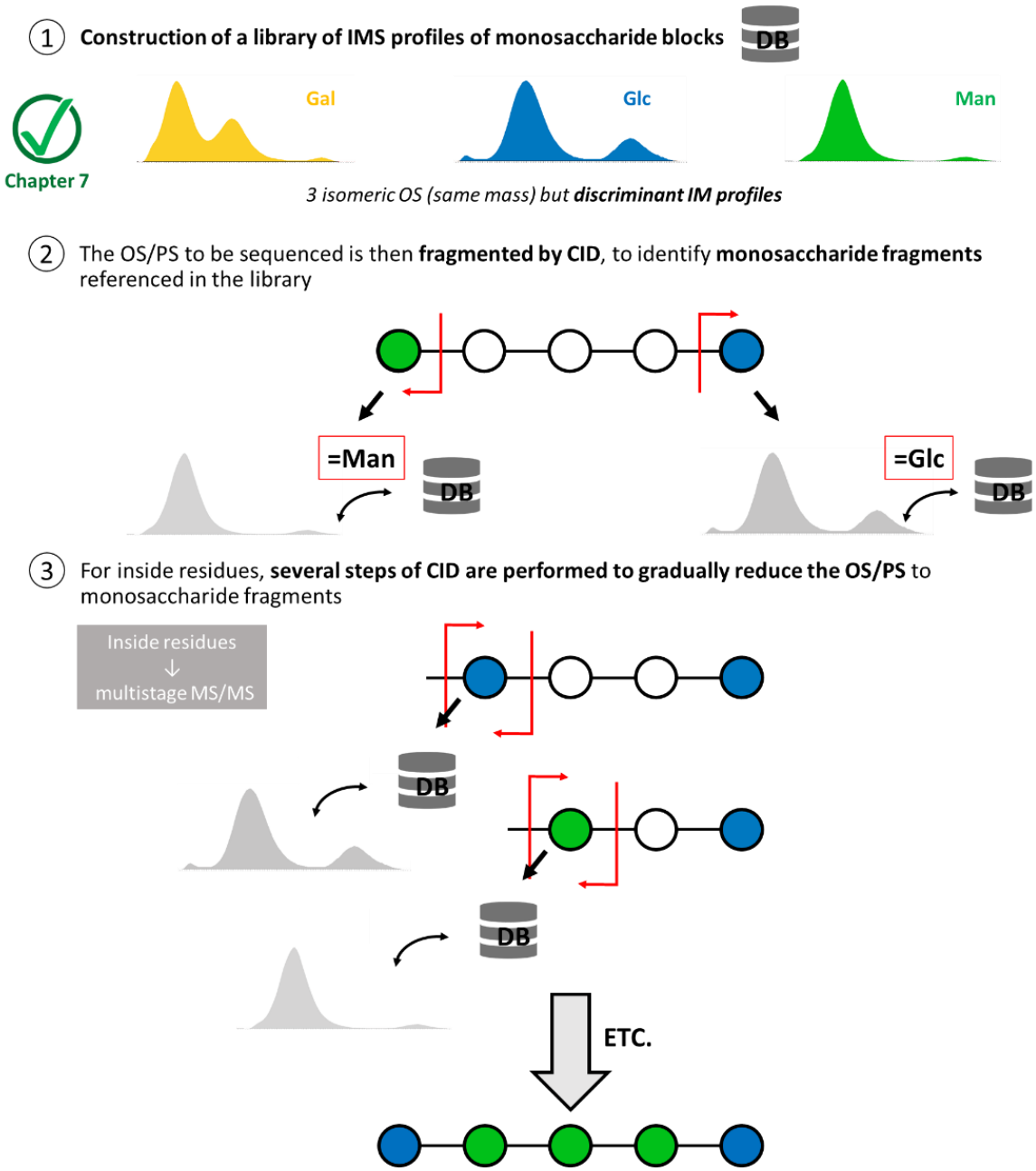


Figure 66. Conceptual scheme of the sequencing of oligosaccharides by IMS^n

The first step is the establishment of a database of HR-IMS profiles for the relevant monosaccharides, which we did already in **Chapter 7** for the most common monosaccharides found in plant or bacterial carbohydrates. The second step would be to identify the residues at both the reducing and non-reducing ends of the saccharide, which can be facilitated by ^{18}O-labeling of the reducing end, as described in the previous Chapters. Then, through multiple steps of IMS^n, the residues would be detached sequentially. In other terms, a fragment missing

one of the extremities would be re-selected using IMS and re-fragmented to identify the following residue, and so on.

The most obvious option would be to build the sequence starting from the non-reducing end—the reducing end fragments containing ^{18}O are easy to identify as targets for re-selection—but ideally oligosaccharides should be sequenced both ways for cross-validation.

Among the encouraging results we have already established is the finding that, with few exceptions, HR-IMS profiles of disaccharides generally matched the corresponding fragments as shown in **Chapters 5 and 6**. However, the whole approach may be compromised in case of significant rearrangements or conformational changes upon CID, such as those observed for Gal*f*-containing oligosaccharides in **Chapter 5**. Finally, it is important to check if CID approaches allow identifying the position at which the subunits get detached from the oligosaccharides. Indeed many other cases of rearrangement occurring during CID of glycans have been reported in the literature, mostly on $[M + H]^+$ and $[M + NH_4]^+$ adducts[149,150] but also on metal adducts.[151]

In the following studies, in order to assess the feasibility of the strategy, several critical steps were examined including: (i) verifying that the appropriate fragments are generated with sufficient intensity, (ii) verifying that the monosaccharidic fragments released by an oligosaccharide upon CID match the database of monosaccharides' IMS profiles, in other terms, that the IM signatures are maintained upon fragmentation, (iii) checking the applicability in IMS^n experiments, and (iv) attempting to actually sequence an oligosaccharide.

C. Methods

1. Chemicals

The study was performed using pure monosaccharide and oligosaccharide standards, either from commercial sources or provided by our partners. *Monosaccharides*: D-galactose (Gal), D-glucose (Glc), D-mannose (Man), and D-xylose (Xyl) were purchased from Sigma-Aldrich (Saint Quentin Fallavier, France). *Oligosaccharides*: Maltopentaose (α1,4-glucan), cellopentaose (β1,4-glucan), laminaripentaose (β1,3-glucan), linear mannopentaose (β1,4-mannan), xylohexaose (β1,4-xylan), and XXXG xyloglucan were purchased from Megazyme (Wicklow, Ireland). Dextran 1000 (α1,6-glucan in mixture) was purchased from Sigma-Aldrich (Saint Quentin Fallavier, France). Branched mannopentaose (α1,6-mannan branched in α1,3) was purchased from Dextra Laboratories (Reading, UK). Galactopentaose (β1,4-galactan) was kindly provided by Dr Estelle Bonnin (INRAE, UR BIA). HPLC grade methanol was purchased from Carlo-Erba (Val de Reuil, France). Ultrapure water was produced using a Milli-Q system

(Millipore, Burlington, MA, USA). Lithium chloride (LiCl) was purchased from Sigma-Aldrich (Saint Quentin Fallavier, France).

2. IMS-MS analyses

Pure samples were infused at a concentration of 1µg/mL in 50:50 water/methanol, using a syringe pump set to a flow rate of 5 µL/min. Lithium adducts were generated by doping the solvents with 0.1 mM of LiCl, except in the case of gluco-oligosaccharides for which 0.01 mM LiCl was used. The MS and IMS parameters were kept identical compared to that described in **Chapter 7**. IMS^n experiments were performed according to the methodologies described in **Chapter 3**, the initial steps of separation on precursor oligosaccharides and "large" fragments being done with a WH of 16 V and WV of 350 m/s. The separation sequences designed in **Chapter 7** were then applied at the end of the IMS^n sequence.

MS and IMS-MS spectra were recorded on the cIMS instrument using the Quartz software (Waters Embedded Analyzer release 6, Waters, Wilmslow, UK). Data were processed using MassLynx 4.2 (Waters, Wilmslow, UK). ATDs were smoothed using the Savitzky-Golay algorithm (window size (scan) ± 1, number of smooths: 2). For cases in which a different smoothing was applied, it is specified in the Figure with the notation *SGxYY* (*YY* representing the number of smooths).

D. Results and Discussion

1. Does CID of oligosaccharides produce monosaccharidic fragments with sufficient yield?

Figure 67 presents the CID fragmentation spectra of four gluco-oligosaccharides (Glc-OS) of DP5, with different linkage anomerism and regioisomerism ((α1,6), (α1,4), (β1,4) and (β1,3) linkages). The monosaccharidic fragments (C_1 and/or Y_1) produced from the $(Glc)_5$ precursor, isolated at *m/z* 835.3 as $[M + Li]^+$ species, are expected at *m/z* 187 ($[M + Li]^+$ species of $(Glc)_1$). These examples show that the intensity of the monosaccharidic fragments (*m/z* 187.1, highlighted in purple) can vary between the four Glc-OSs. For the dextran (α1,6-Glc-OS) and laminarin (β1,3-Glc-OS), the *m/z* 187.1 is of satisfactory intensity. However, for the cellopentaose (β1,4-Glc-OS) and maltopentaose (α1,4-Glc-OS), it is of low intensity. This could be a major problem, because the intensity will further decrease after each step of IMS^n.

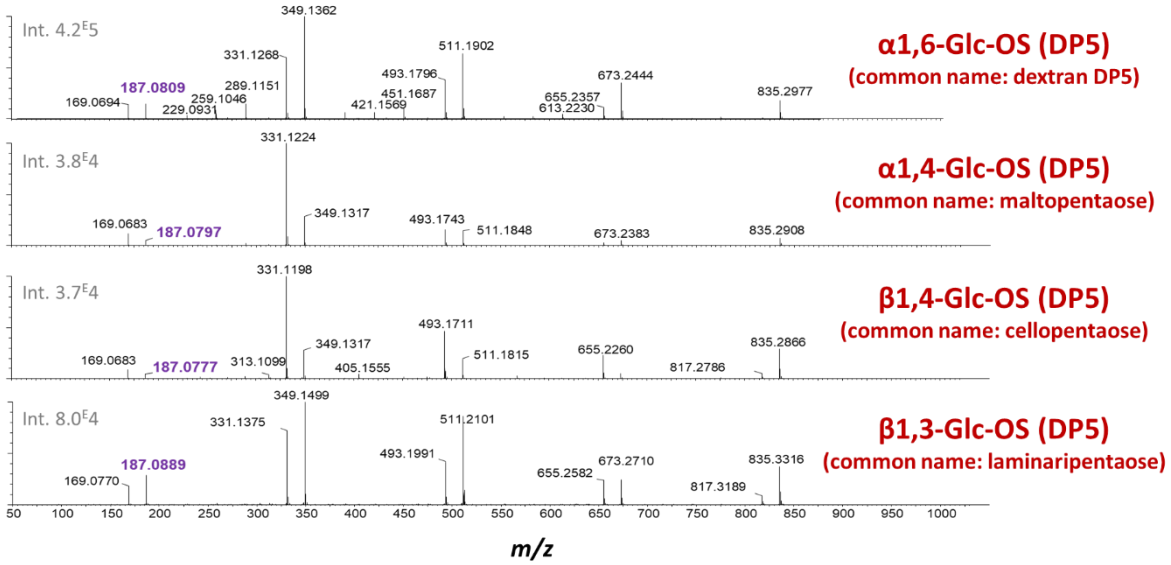


Figure 67. Examples of CID fragmentation spectra produced by activation upon reinjection from the pre-store (+190 V) for four different Glc-OS DP5, displaying different linkage anomerism and regioisomerism. Parent ion ($(Glc)_5$) was isolated at *m/z* 835.3 as $[M + Li]^+$. The monosaccharidic fragments of interest (C_1 or Y_1, corresponding to $(Glc)_1$) are highlighted in purple at *m/z* 187.1.

2. Monosaccharidic fragments from homo-oligosaccharides match the profile of monosaccharides

Next, we verified that the monosaccharidic fragments released upon CID of an oligosaccharide have the same IM profile compared to the corresponding monosaccharide. We conducted this step of validation by fragmenting homo-oligomers of hexoses and pentoses, which are the easiest to procure.

a) Glucosides

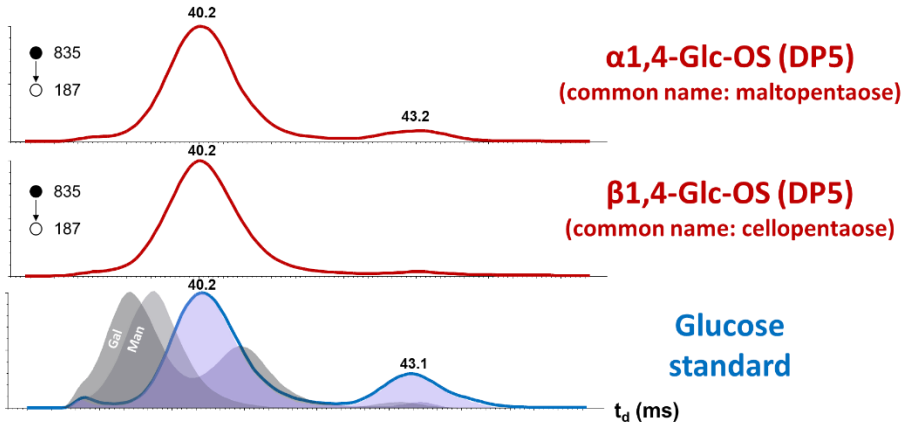


Figure 68. Monosaccharidic fragments released by fragmenting the $[M + Li]^+$ species isolated at *m/z* 835.3 of maltopentaose (α1,4-glucan) and cellopentaose (β1,4-glucan) in red, are compared to the reference profile of Glc in blue.

First, in the case of glucosides presented in **Figure 68**, we observed that the monosaccharidic fragments (fragment at $[M + Li]^+$ *m/z* 187.1) produced for both α1,4- and β1,4-Glc-OSs of DP5 were a good match to the IMS profile of pure glucose. Notably, the major Glc peak at a drift time t_d = 40.2 ms remained the main feature of the IMS profile of the fragments. On the other hand, the feature at 43.1 ms was observed with less intensity on the fragments compared to pure Glc. Interestingly, maltopentaose (α1,4-glucan) exhibited a more intense peak at 43.1 ms compared to cellopentaose (β1,4-glucan). This is a first sign that the anomeric configuration of the glycosidic bond may have some impact on the IM profile of fragments. While we previously exploited anomeric memory in **Chapter 6**, what we observe here is different. Previously, we compared the profile of disaccharidic fragments and verified that it matched the corresponding standards (*i.e.*, the anomerism of the bond is retained within the fragment). Here, what we observe would rather be an *imprint of the anomerism left on the IM profile after cleavage of the glycosidic bond*. This is furthermore in agreement with a previous report by Gray *et al.*,[110] who studied the (low-resolution) IM profile of monosaccharidic fragments generated from disaccharides using CID, and also observed an additional peak for α-Glc-OSs compared to β-Glc-OSs. This will be studied in more detail in the later parts of this Chapter.

b) Mannosides

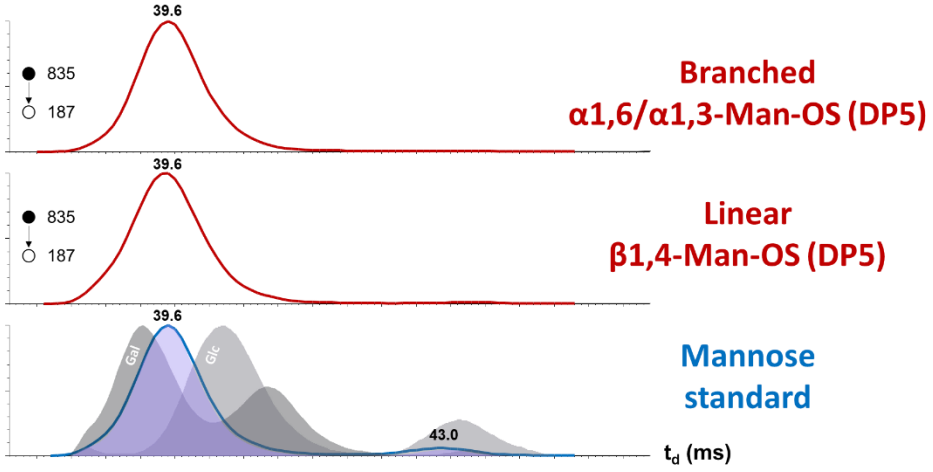


Figure 69. Monosaccharidic fragments released by fragmenting the $[M + Li]^+$ species isolated at *m/z* 835.3 of branched and linear mannopentaose in red, are compared to the reference profile of Man in blue.

Second, in the case of mannosides presented in **Figure 69**, we likewise observed that the monosaccharidic fragments released at *m/z* 187 ($(Man)_1$ as $[M + Li]^+$ species) from two different types of manno-oligosaccharides of DP5 (Man-OS DP5) were a good match to the profile of pure Man. The IM profiles unambiguously exhibited a major feature at 39.6 ms, which was the signature peak of Man. Of note, we did not observe any significant contribution of the feature

at 43.0 ms, visible on the IM profile of the pure Man, either for the branched α-mannopentaose or the linear β-mannopentaose.

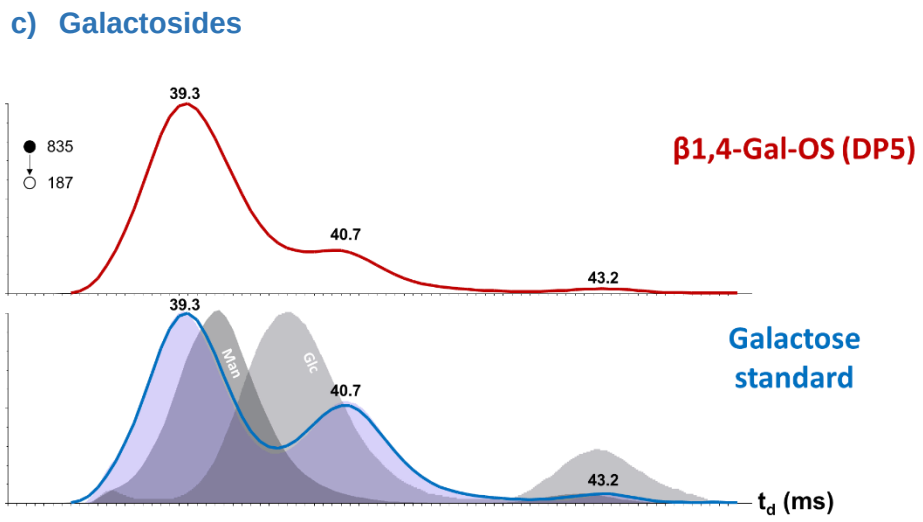


Figure 70. The monosaccharidic fragment released by fragmenting the $[M + Li]^+$ species isolated at *m/z* 835.3 of galactopentaose (β1,4-galactan) in red, is compared to the reference profile of Gal in blue.

The final case of hexosides that we studied was a β1,4 galacto-OS of DP5 (Gal-OS DP5), presented in **Figure 70**. The profile of the monosaccharidic fragment ($(Gal)_1$, detected at *m/z* 187.1 as $[M + Li]^+$ ion) released from this Gal-OS DP5 was a near-perfect match to that of pure Gal. Indeed, we detected all three peaks of Gal in the IM distribution at t_d = 39.3, 40.7, and 43.2 ms, respectively (albeit a slightly lower intensity for the 40.7 ms feature in the fragment compared to the reference).

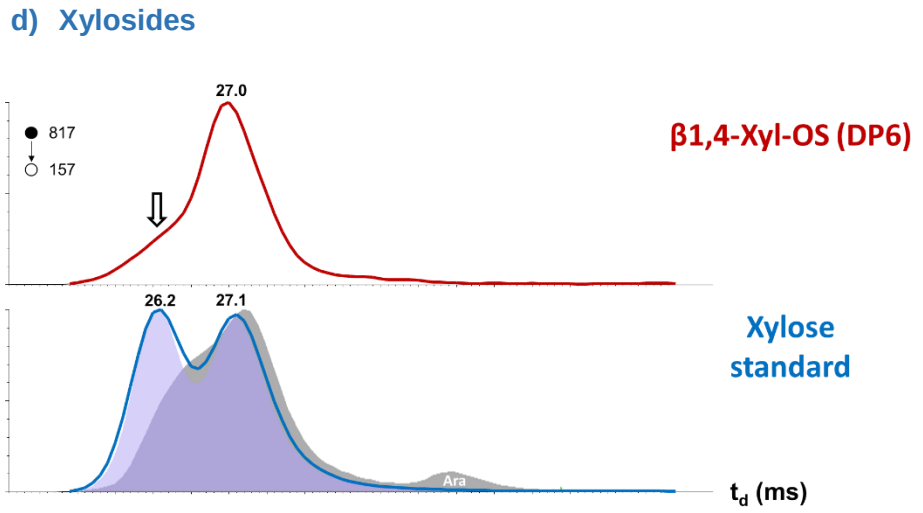


Figure 71. The monosaccharidic fragment released by fragmenting the $[M + Li]^+$ species isolated at *m/z* 817.3 of xylohexaose (β1,4-xylan) in red, is compared to the reference profile of Xyl in blue.

Finally, we looked at some pentosides (xylo-oligosaccharides, Xyl-OS) for further validation that the IM profiles of fragments are comparable to the reference monosaccharides. In the case presented in **Figure 71**, the monosaccharidic fragment ($(Xyl)_1$, detected as $[M + Li]^+$ ion at *m/z* 157.1) released from the Xyl-OS DP6 matched the profile of the reference Xyl. We note that the main peak in the fragment's IM profile is t_d = 27.0 ms, which is also close to a peak found in the profile of Ara. Nevertheless, the feature at 26.2 ms is also observed in the form of a shoulder of significant intensity, and the start/end times of the drift time distribution are identical to the standard.

Overall, we have shown that in all the cases we tested, the IM profile of the monosaccharidic fragments released from larger oligosaccharides were a relatively good match to the IM profile of the corresponding monosaccharide standards. In some cases, differences in the intensity of certain features were observed. In addition, a potential effect of the anomericity of the glycosidic bond on the IM profile, notably in the case of glucosides, was observed.

3. Does the linkage anomerism in the parent oligosaccharide change the IM profile of the monosaccharidic fragments?

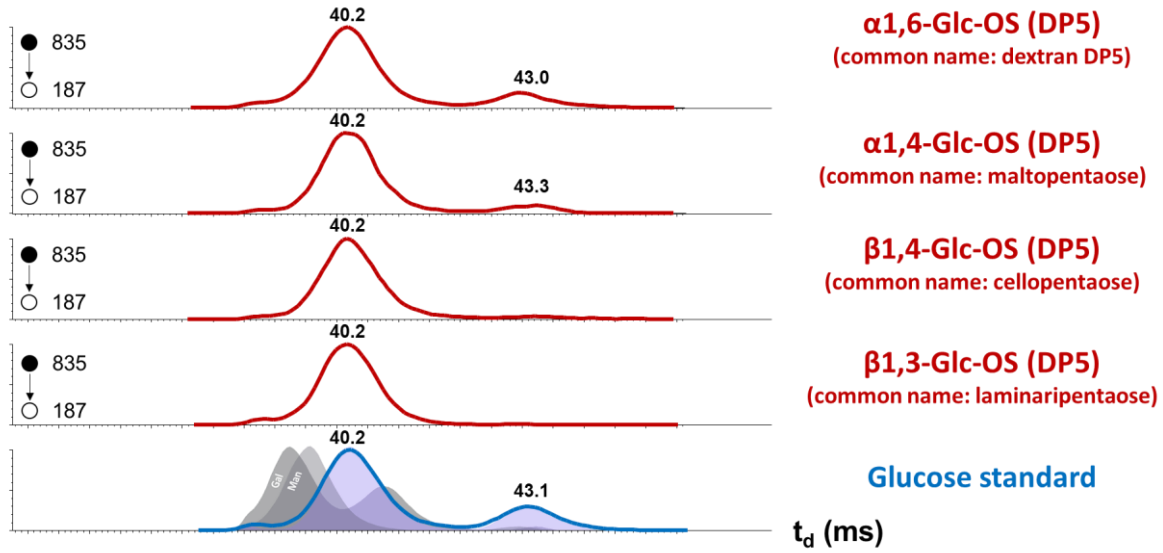


Figure 72. Comparison of the monosaccharidic fragments released from different gluco-oligosaccharides isolated at *m/z* 835.3 as $[M + Li]^+$ with α (DP5 dextran in (α1,6) and maltopentaose in (α1,4)) or β linkages (cellopentaose in (β1,4) and laminaripentaose in (β1,3)). In blue, the IM profile of pure glucose is given for reference.

To check whether there was an impact of the anomeric configuration of the glycosidic bond (or if there was also an impact of the regiochemistry), we opted to study various types of gluco-oligosaccharides (Glc-OS). Indeed, since glucosides are abundant and ubiquitous in Nature, they offer a wide variety of linkage isomers, most of them being available as commercial standards of good purity. In **Figure 72**, we present the IM profiles for four Glc-OS of DP5: two

α-linked Glc-OS (dextran and maltopentaose, linked in (α1,6) and (α1,4) respectively), and two β-linked Glc-OS (cellopentaose and laminarin, linked in (β1,4) and (β1,3) respectively). In all cases, the main peak in the IM profile of the monosaccharidic fragment at *m/z* 187.1 was that at 40.2 ms. However, we indeed observed that for the two α-linked glucans, the feature at 43 ms was more intense compared to the two β-linked glucans.

Overall, it appears that some features of the ATD are likely influenced by the anomeric configuration of the linkage. However, the two features at 40.2 and 43.1 ms do not seem to directly correspond to the α and β anomers either, as we observe the 40.2 ms feature in both α and β-linked Glc-OS, and we observe a slight trace of the 43 ms feature in the profile of cellopentaose (and not laminarin). Still, this is an interesting property, as it could help assign the intrachain anomerism during the sequencing of oligosaccharides.

4. Can we work on B-type fragment ions?

a) Impact of the glycosidic bond (regioisomerism and anomerism) for a given monosaccharide

When studying the CID spectra of glucooligosaccharides (**Figure 67**), we noted that the CID of these oligosaccharides yielded major amounts of B-type ions.[x] These notably included B_1 ions, corresponding to dehydrated monosaccharides and observed at *m/z* 169. We thus aimed at determining if it was possible to study B_1 ions to complement the complete monosaccharidic fragments.

To see if it was possible to work on dehydrated monosaccharidic fragments instead of complete monosaccharidic fragments, we first evaluated if the IM profiles of such fragments would efficiently reflect the composing building blocks (*i.e.*, afford identification of the subunit).

Figure 73 presents the HR-IMS profiles of the B_1 fragments (*m/z* 169.1) of the four Glc-OSs of DP5 after three IM passes. Despite a few intensity differences for some features, we observed very similar profiles between the two α-glucans on one hand, and the two β-glucans on the other hand. However, very different profiles were observed between α and β glucans. Indeed, both α-linked species exhibit a multiplexed profile containing features at 68.5, 69.4, 71.1, and 72.8 ms, while both β-linked species exhibit a simpler profile with a major peak at 69.3 ms and minor features at 68.5 and 71.2 ms.

This finding is interesting because it suggests that IM profiles of the dehydrated monosaccharides (B_1 fragments) sign the anomerism of the bond.

[x] Without the ^{18}O labeling, these could also be Z ions, but during the different studies conducted in **Part II** we observed a predominance of B ions over Z ions. Thus, we will herein consider these to be B ions.

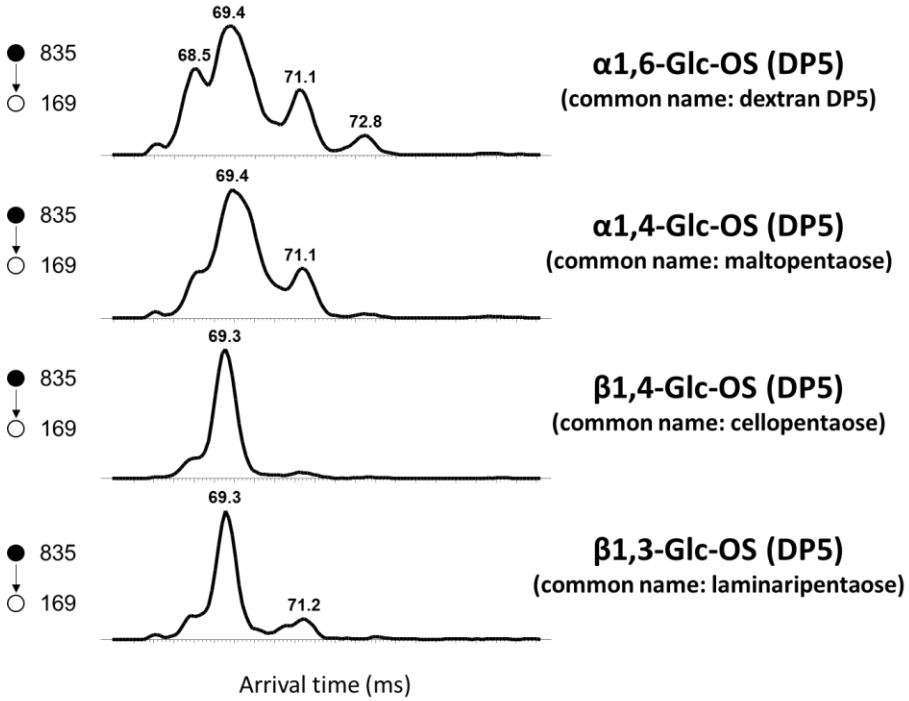


Figure 73. Comparison of the dehydrated monosaccharidic fragments (*m/z* 169.1) released from different Glc-OS of DP5 with α or β linkages. Parent ion was isolated at *m/z* 835.3 as $[M + Li]^+$. IM profiles of the dehydrated monosaccharidic fragment are presented (B_1 fragment, detected at *m/z* 169.1 as $[M + Li]^+$ ion). The presented ATDs were acquired after three passes around the IMS cell (static WH conditions).

b) Impact of the epimerism of the subunit on the IM profile of B_1 fragments

However, it was important to assess whether the IM signatures of B_1 fragments were unique to the subunit (in the previous case, Glc) or would be shared between different types of subunits. Indeed, since B_1 fragments are dehydrated monosaccharides, there is a possibility that the dehydratation of different stereoisomers (such as Glc and Gal, or Glc and Man) could give the exact same fragment structure.

We thus analyzed three hexosides of DP5 composed either of Glc, Man or Gal subunits, all linked with (β1,4) bonds (**Figure 74)**. The dehydrated monosaccharidic fragment produced by the Gal-OS DP5 exhibited a very distinctive IM profile composed of many features (at 68.0, 69.6, 70.7 and 71.5 ms). However, the IM profile of the *m/z* 169 fragment produced from the Man-OS DP5 was strikingly similar to that produced from the Glc-OS DP5 (described in the previous paragraph).

To our understanding, this is actually the result of Glc and Man being C-2-epimers. Upon dissociation of the glycosidic bond and dehydration, a double bond can form between C-2 and O-2 (see fragmentation scheme in **Figure 75**), resulting in the B_1 fragment losing the stereospecificity that would allow to distinguish Glc and Man.

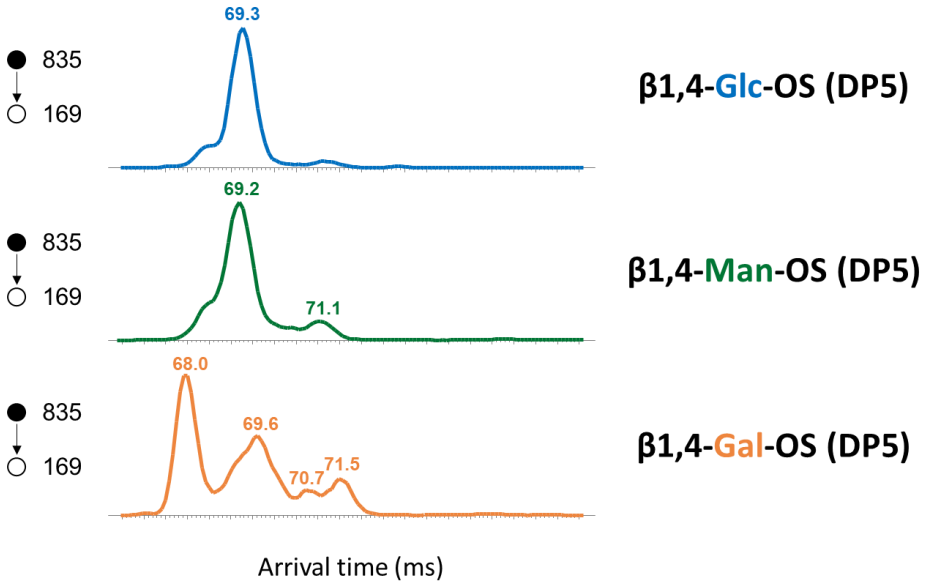


Figure 74. Comparison of the dehydrated monosaccharidic fragments (B_1, *m/z* 169.1 as $[M + Li]^+$ species) released from Glc-OS DP5, Man-OS DP5 and Gal-OS DP5. Parent ion was isolated at *m/z* 835.3 ($(Hex)_5$ as $[M + Li]^+$). The presented ATDs were acquired after three passes around the IMS cell (static WH conditions).

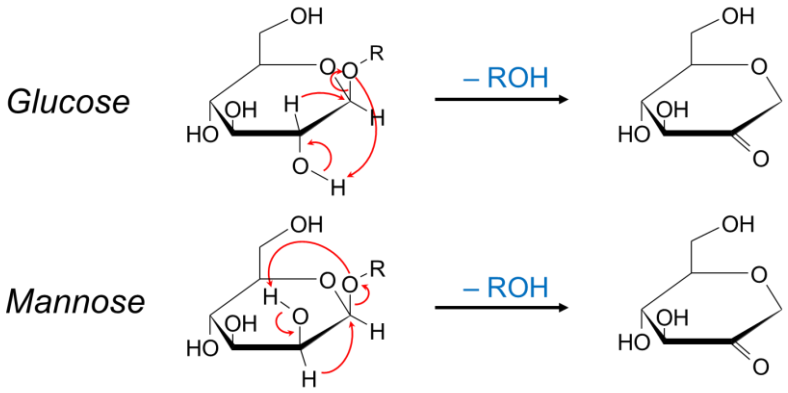


Figure 75. Fragmenting a Glc-(β1,4)-Glc or a Man-(β1,4)-Man bond gives the same B_1 fragment. The involved dehydration mechanism is presented for both Glc and Man subunits (R = oligosaccharidic chain).

In summary, despite the potential interest of dehydrated monosaccharidic fragments to sign the intrachain anomerism of the oligosaccharide, using them alone is not an option since they do not allow identifying the subunits with 100% certainty. Nevertheless, they could still be an interesting complement to the other monosaccharidic fragments (*i.e.*, C_1/Y_1) for cross validation of the sequence.

Overall, in the above paragraphs, we established that the HR-IMS profiles of the monosaccharidic fragments reflect the composition of the precursor oligosaccharides, and even possibly the stereochemistry of the glycosidic bond.

5. Applicability of the approach in IMS^n experiments

The next step towards achieving sequencing of oligosaccharides was to verify if the strategy could be applied in IMS^n experiments. To this end, we performed an IMS^3 study of α1,4-Glc-OS DP5 (maltopentaose). In fact, this oligosaccharide was among those that yielded a low intensity of monosaccharidic fragment at *m/z* 187.1 (**Figure 67**). We estimated that if the IMS^n strategy was successful on this compound, it should be successful on most other oligosaccharides.

The results are presented in **Figure 76**. The first stage of fragmentation (IMS^2) corresponds to the profiles that we already presented, and in which the IM profile of the monosaccharidic fragment produced at *m/z* 187.1 matches the pure Glc standard, which reveals that a Glc residue is standing on at least one end of the oligosaccharide. To perform IMS^3 and reach more internal residues of the sequence, we focused on the Y_4 fragment at *m/z* 673.2[y] and selected it with a slicing experiment. After 6 passes (110 ms of separation) the ATD of Y_4 exhibited two peaks (① and ②). Peak ② did not overlap with any other fragment and could be directly selected for a second round of fragmentation (IMS^3) to give the monosaccharide profile and the main feature in this profile (t_d 40 ms) matched with the profile of Glc, indicating a Glc standing as end-residue of Y_4. On the other hand, even after 6 passes in IMS^2, peak ① overlapped with that of the fragment at *m/z* 655.2. While the IMS^3 analysis of peak ② could have been sufficient for α1,4-Glc-OS DP5, for other oligosaccharides there might be cases in which all Y ions overlap with B ions. So, we also evaluated the applicability of the approach on peak ①. It took another 35 passes (594 ms of separation) to resolve peak ① and the peak of *m/z* 655 enough to slice half of peak ① and give the IMS^3 monosaccharide profile. After these additional passes, the profile of the monosaccharide fragment arising from peak ① (two features at 40 and 43 ms) also fitted with one Glc residue standing at the extremity of Y_4.

In the end, we were able to obtain the IMS^3 profiles of the monosaccharides release after CID of both Y_4 peaks. Furthermore, both signatures matched the expected reference monosaccharide (*i.e.*, Glc) with a peak at 40 ms, but only peak ① presented the feature at 43 ms. Our hypothesis is that this feature may be linked to the occurrence of gas-phase conformers.

[y] Without the ^{18}O labeling, this *m/z* could also be a C_4 ion, but during the different studies conducted in **Part II** we observed a predominance of Y over C ions. Thus, we will herein consider this ion to be Y_4.

Overall, this first attempt to sequence an oligosaccharide in IMS^n was a success but also evidenced some major setbacks. The selection in IMS^n is only based on the arrival time, not the *m/z*, so it can require extensive efforts to isolate a pure fragment for multistage analysis. The resulting IMS^n sequences can extend to hundreds of milliseconds (probably even seconds when going further than IMS^3), which causes poor duty cycle and requires acquisition over several minutes.

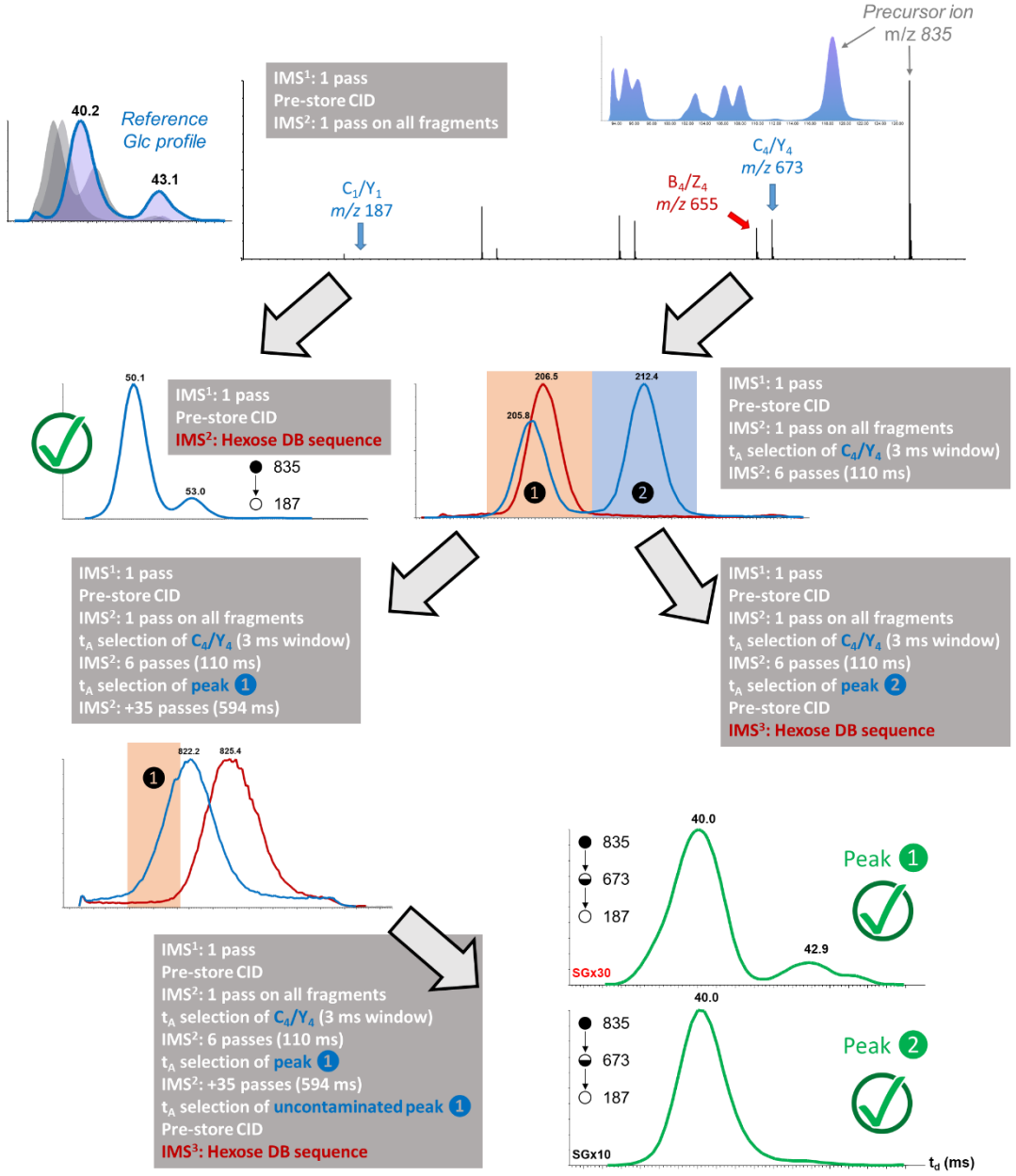


Figure 76. IMS^n study of maltopentaose.[z] The selection steps and separation times are given for each stage of the sequences.

[z] In this Figure, the *n* in IMS^n is only incremented after a round of CID, disregarding simple arrival time selections.

6. One does not simply walk through the sequence

Up to this point, despite issues in terms of duty cycle, the strategy seems applicable. The key remaining point that has to be addressed is: *can we accurately identify the position from which a given monosaccharide was detached?*

At this stage, we were faced with one of the major challenges of using CID for glycans: the occurrence of consecutive fragmentations. Indeed, in CID, it is possible for a fragment (e.g., with one monosaccharide removed) to be re-fragmented by further collision with the gas particles. This fragment would thus lose one (or more) other residues, resulting in the glycan breaking down in an order that cannot be anticipated from its sequence. Knowing that these consecutive fragmentations happen in CID, and combined with the possible symmetry of the molecule and isomerism of the subunits, it is difficult—if not impossible in certain situations—to ascertain the origin of many CID fragments observed on the MS/MS spectra. This poses a major difficulty in the sequencing strategy. Indeed, in order to attain the internal residues of the sequence, it is necessary, in the IMS^3 step and the following ones, to re-select the parts of the structure that undoubtedly carry the residue to be released next.

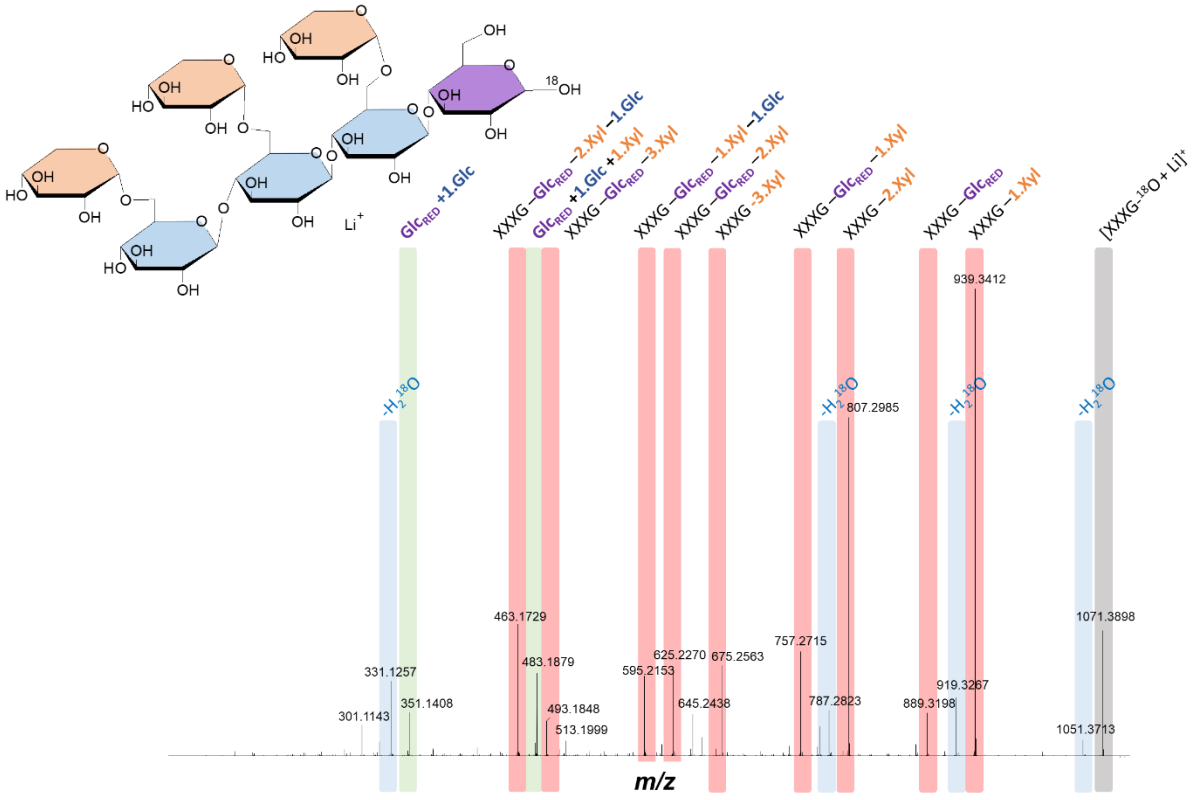


Figure 77. CID fragmentation spectrum of ^{18}O-labeled XXXG xyloglucan (+210 V upon reinjection in the mobility cell), for the $[M + Li]^+$ species at *m/z* 1071.4. Color code for the annotation of the fragments: peak highlighted in green = unambiguous fragment, in red = ambiguous fragment, in blue = loss of $H_2{}^{18}O$ from another fragment.

This phenomenon is exemplified on a xyloglucan (DP7) in **Figure 77**. This structure consists of a backbone made of four β1,4-linked Glc units, each Glc being branched with a Xyl in position 6 except the reducing one. In this spectrum, two of the major peaks (*m/z* 939.3 and *m/z* 807.3) correspond to losses of 1.Xyl and 2.Xyl, respectively. These losses can occur on any of the three Xyl residues and thus leave many possibilities for the remaining structures (3 possible structures for the loss of 1.Xyl and 3 structures for the loss of 2.Xyl). If those structures, which are undifferentiated by their *m/z*, are selected for further steps of IMS^n, it will become complicated to deduce which end releases which residue.

The ^{18}O labeling of the reducing end helps in assigning ambiguous fragments of an MS/MS spectrum, the Y (and Z) fragments supposedly being shifted by 2 Da. Yet, for cost reasons, this labeling cannot be envisaged on all samples. Furthermore, it has also limitations in the context of consecutive fragmentations as a non-labeled peak may arise from a Y (or Z) fragment having undergone a subsequent B/C fragmentation. Another efficient way of overcoming the ambiguity of assignment in MS/MS spectra of carbohydrates is to use high-energy activation methods for fragmentation, as was documented by our group with XUV-DPI or He-CTD.[40,41] These methods were shown to induce limited consecutive fragmentations and in addition, they produce a significant proportion of cross-ring fragments that are less ambiguous than glycosidic bond cleavages. The implementation of the ExD cell on the cIMS platform is promising in that direction.

Finally, another direction would be to proceed to bottom-up sequencing in a shotgun approach like Miller *et al*., but to start directly from the monosaccharides. For instance, in a mixture of oligosaccharides obtained from the degradation of a polysaccharide, the disaccharides would first be characterized using monosaccharide fingerprints, then the trisaccharides would be characterized using the mono- and disaccharides, and so on.

E. Conclusions and Perspectives

In this Chapter presenting our vision of carbohydrate sequencing by IMS^n, we were able to validate several critical steps. First, we confirmed that the monosaccharidic fragments released upon CID of oligosaccharides match the IM profiles of monosaccharides. Second, we validated that these signatures were preserved even after multiple steps of CID in IMS^n experiments. We further showed that B_1 ions could possibly complement the information obtained from monosaccharidic fragments, but with the restriction that some epimers could give identical B_1 fragments. Finally, our results suggest that the intrachain anomerism could possibly be inferred from the IMS profiles of the fragments.

On the other hand, we also identified some limiting factors such as the poor duty cycle in IMS^n experiments and the consecutive fragmentations that occur upon CID of glycans, which makes

the attribution of some fragments ambiguous. In this liminal study on our path towards carbohydrate sequencing, we worked on highly symmetrical oligosaccharides that amplify the latter difficulty. We have recently opted for less symmetrical structures, containing more types of subunits. This should help us rationalize the behavior of oligosaccharides in IMS^n, and consolidate our protocol. To this end, we have initiated a collaboration with the team of Dr Agata Zykwinska[aa] about the bottom-up characterization of a bacterial exopolysaccharide (coined AT1219) that contains both hexoses, hexuronic acids, and *N*-acetyl hexosamines.

F. Scientific Communication

These results will be described in an upcoming article, alongside our work on IMS profiling of monosaccharides (see **Chapter 7**).

[aa] Ifremer, Laboratoire Ecosystèmes Microbiens et Molécules Marines pour les Biotechnologies, F-44311 Nantes, France

Chapter 9 – Molecular Networking of tandem IMS data to organize glycomics datasets

A. Context

In **Chapter 8**, we highlighted that the fragments generated by collisional dissociation of oligosaccharides have potential for sequencing approaches based on IMS^n. Despite our efforts, the sequencing strategy suffers from notable setbacks that make it complicated to select the proper fragment ion for further stages of fragmentation (consecutive fragmentations, generation of isomeric fragments, etc.) and is still under development in the laboratory. At the current stage, the approach still requires extensive manipulation and interpretation time by a specialist, limiting the application to targeted/already known oligosaccharides. However, when looking at complex biological samples, an untargeted strategy able to rapidly assess the chemical diversity of oligosaccharidic structures would be beneficial.

Amidst our sequencing attempts, we noticed that the IMS/IMS[bb] fragmentation spectra of oligosaccharides shared common features between closely related structures, but that even small modifications (like the intrachain anomerism) induced important differences. As an illustration, **Figure 78** presents the IMS/IMS spectra of cello-oligosaccharides of DPs 4 and 5 (β1,4-glucans) *versus* malto-oligosaccharides of the same DPs (α1,4-glucans). When comparing the two spectra of oligosaccharides with the same anomerism (e.g. the two spectra of the β1,4-glucans), the fragments of the DP4 are also found in the spectra of the DP5. However, when comparing the α1,4- *vs* β1,4-glucans there are notable difference between the IM profiles of the fragments (notably in the 15–25 ms drift time region, corresponding to DP3-like of DP4-like fragments). We thus hypothesized that these IMS/IMS spectra could provide interesting fingerprints of oligosaccharide structures and that these could potentially be used in untargeted approaches to structure elucidation (in the same way as MS/MS spectra in metabolomics or proteomics).

One notable untargeted method used in metabolomics to organize MS/MS spectral data is molecular networking. Molecular networks (MN) have been used for data organization since the 2010s[152] and, in 2016, the Global Natural Products Social MN platform (GNPS)[cc] was formally introduced to the community.[153] To build an MN, pairwise comparisons of MS/MS

[bb] Here, we use the term IMS/IMS spectra to designate the IMS profile obtained after (i) IMS slicing of the precursor ion, (ii) collisional fragmentation upon reinjection in the IMS cell, and (iii) single-pass IMS analysis of all of the fragments.
[cc] The GNPS platform is accessible at https://gnps.ucsd.edu

spectra are done through a cosine similarity score: the higher the cosine score, the more similar the fragmentation spectra. If this cosine score for a pair of spectra is above a defined threshold, the spectra are connected by an edge in the network. In this type of MN, only the edges are significant and the distances between clusters hold no particular meaning. The molecular networking strategy is generally accompanied by a dereplication of the network's spectra against databases,[154] which helps identify known ions and propagate the structural information to related compounds.

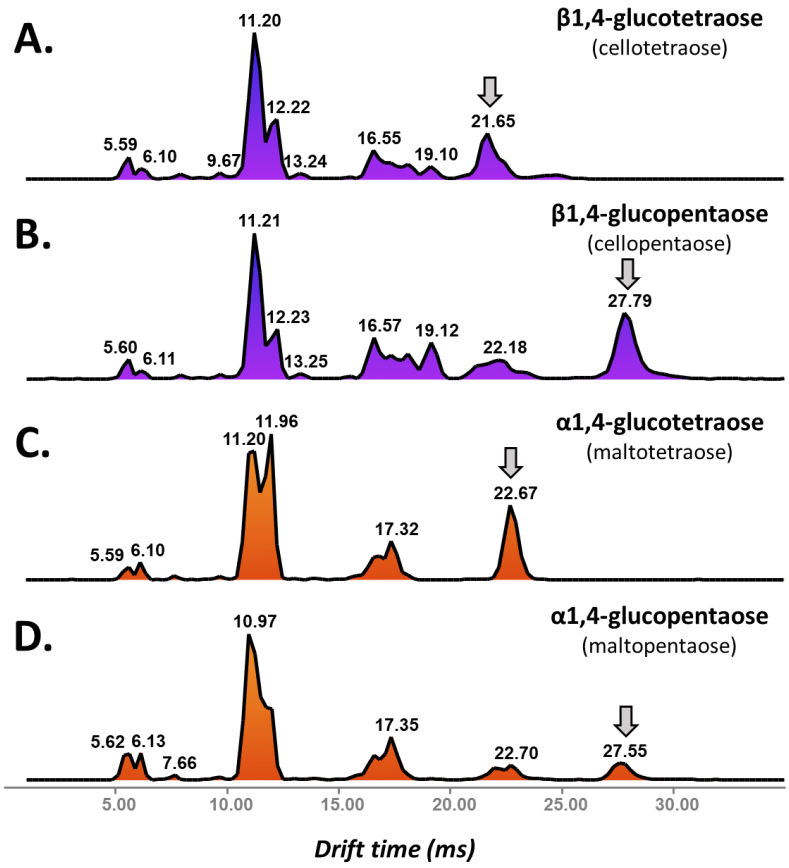


Figure 78. IMS/IMS fragmentation spectra of β1,4-glucans (**A**: cellotetraose, and **B**: cellopentaose) *vs*. α1,4-glucans (**C**: maltotetraose, and **D**: maltopentaose). The precursor ion is identified by an arrow in each spectrum.

B. Objectives

In this study, we aimed to evaluate the interest of molecular networking for the study of oligosaccharides. We evaluated the strategy using first the "traditional" method with MS/MS spectra. Secondly, we conceived and evaluated an original method—that we called ion mobility-molecular networking (IM-MN)—which uses IMS/IMS spectra instead of MS/MS spectra. To this end, we selected *ca.* 30 oligosaccharide standards representative of plant carbohydrates. Our set of compounds notably included linkage regioisomers, linkage anomers, and various types of osidic subunits including epimers (**Figure 79**). After the initial assessment

of the interest of IM-MN, we also aimed to understand the factors that drive the clustering in IM-MN.

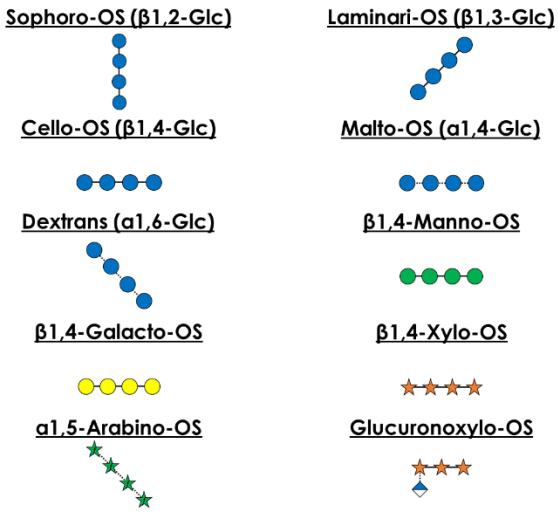


Figure 79. Families of oligosaccharides included in the set of compounds used to evaluate the interest of molecular networking (DP4 are shown as example)

C. Summary of the results

We first tried the "traditional" molecular networking approach with MS/MS spectra recorded for the 30 oligosaccharide standards. As shown in **Figure 80** this is of little interest, because the only clusters formed are directly related to the hexose/pentose nature of the subunits. This was expected with regard to the various levels of isomerism involved, which results in many fragments in these structures sharing the same mass.

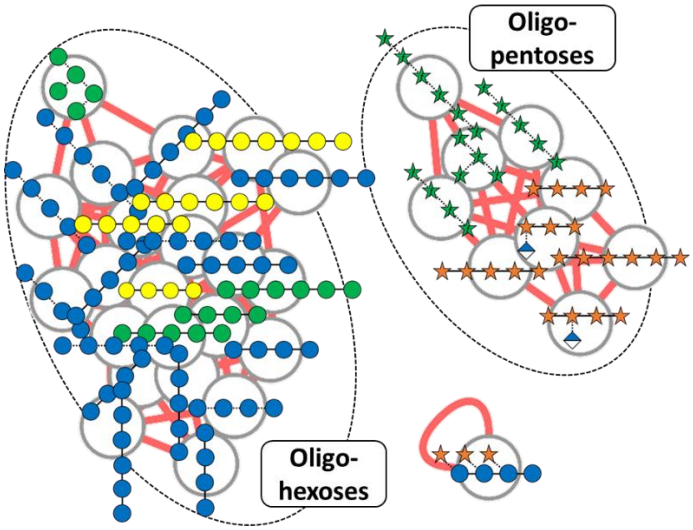


Figure 80. MS-based molecular network for the set of standard oligosaccharides

Next, we investigated the IM-MN network, where the mass information in the fragmentation spectra is replaced by the IM information. This new network, presented in **Figure 81**, contains

various clusters that appear much more structurally relevant than the MS-MN clusters. A closer look into the structures behind the clusters suggests that the clustering might be driven by the structural characteristics of the backbone of the oligosaccharides. For instance, the α1,4-glucans and the β1,4-glucans are split between **Clusters I** and **II** respectively, the different regioisomers of β-glucans are split between **Clusters II–IV**, xylans and arabinans are separated in **Clusters V** and **VI**.

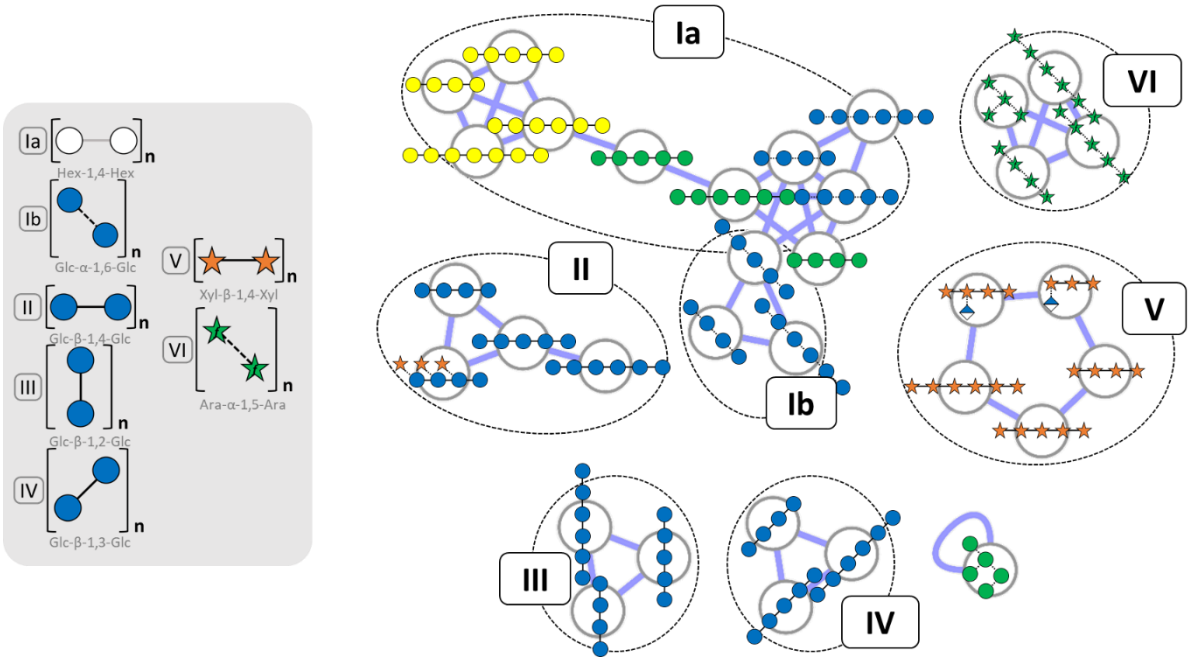


Figure 81. IM-based molecular network for the set of standard oligosaccharides

To validate the hypothesis that the backbone of an oligosaccharide is the main driver of cluster formation, we challenged the IM-MN with more complex oligosaccharides consisting of (i) mixed-linked glucans that contain two linkage regioisomerisms, and (ii) arabinoxylans, containing both Ara and Xyl subunits (**Figure 82**).

We can see that the first mixed-linked glucan, mainly linked in (α1,4) with one (α1,6) bond, clusters with both α1,4- and α1,6-glucans—but these were already linked (**Clusters Ia** and **Ib**). On the other hand, the second mixed-linked glucan, mainly linked in (β1,4) (two bonds) with one (β1,3) bond, clusters with the β1,4-glucans but not the β1,3 (**Cluster II**). Likewise, arabinoxylans containing Ara subunits in lateral positions and xylose in the backbone cluster with xylans (**Cluster V**) but not with arabinans (**Cluster VI**). This supports the hypothesis that cluster formation is driven by the backbone of oligosaccharides.

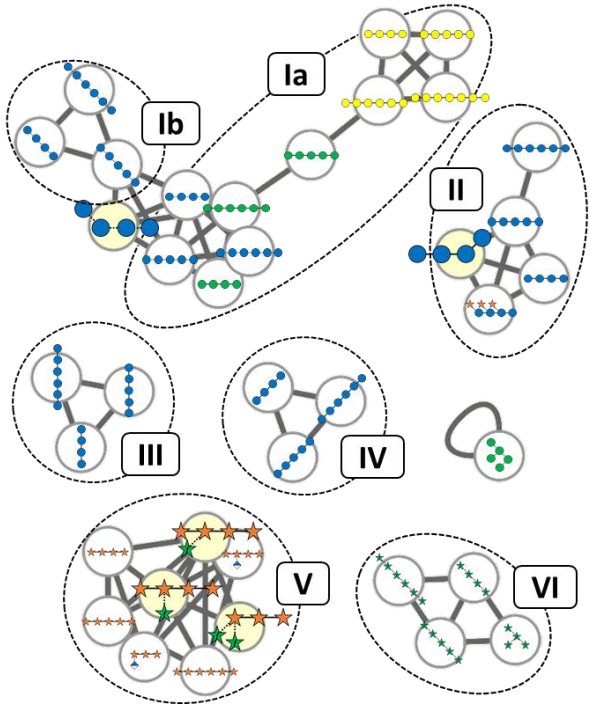


Figure 82. Extended IM-MN in which the initial IM-MN was challenged with structures containing moieties (subunits or linkages) from multiple standards in the initial network. These structures are highlighted in yellow.

D. Conclusions and Perspectives

In this work, we have proposed a novel strategy for structural glycomics: molecular networking of high-resolution IMS/IMS spectra. This strategy combines two forefront approaches in structural analysis for data acquisition and data processing (IMS/IMS and molecular networking, respectively). We have demonstrated that—when it comes to oligosaccharides and their numerous levels of isomerisms—mass-deconvolved IMS/IMS spectra are, without comparison, more powerful than MS/MS spectra to sort and organize molecules with a high degree of structural relevance.

Currently, we are working on coupling the strategy with upfront chromatographic separation. A setback is that the current software implemented on the SELECT Series Cyclic IMS does not allow automation of the IMS/IMS experiments. However, the IMS/IMS spectra can to some extend be replaced by CID-IMS data (*i.e.*, with fragmentation in the trap collision cell of the instrument instead of upon reinjection from the array to the IMS cell). We have started applying this approach to the biological samples produced in the ALGAIMS project (see **Chapter 4**). The samples consist of oligosaccharides produced by acidic hydrolysis of polysaccharides extracted from the lichen *Lasallia pustulata*). As shown in **Figure 83**, the IM-MN strategy uncovered three clusters, suggesting the presence in the sample of several polysaccharides with different structural characteristics. One cluster notably included pentosylated

oligohexoses that have not been described so far as part of these lichen samples, but these species have yet to undergo complete structural characterization.

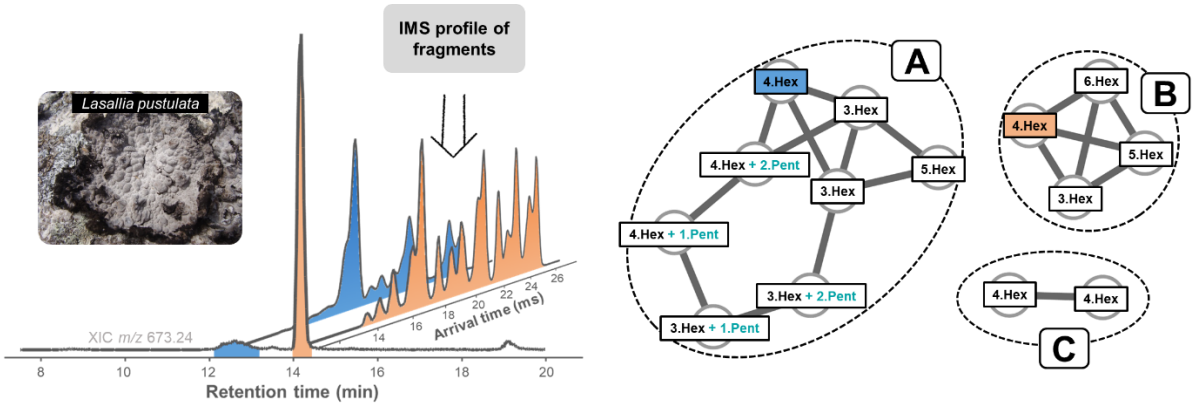


Figure 83. Illustration of the LC-CID-IMS acquisitions, and IM-MN for the *Lasallia pustulata* digest sample.

One current limit of the approach is that the GNPS platform was designed to handle only a single spectral dimension (*i.e.*, the *m/z*). Thus, IMS/IMS data had to be "degraded" before computing the cosine score to keep only the CCS values, *i.e.*, the *m/z* dimension of the IMS/IMS spectra was lost in the process. In the future, it will be interesting to develop new scoring systems that retain the multidimensionality of IMS/IMS data. Molecular networks can then be expected to achieve greater accuracy and relevance when accounting for both m/z and CCS, as this information is not structurally redundant. Therefore, *m/z* and CCS should together provide an enriched and unequivocal structural signature of each fragment ion leading to a unique carbohydrate fingerprint.

E. Scientific Communication

The proof of concept on standard oligosaccharides has been published in the following article presented as **Publication 4**. The supporting data for this publication are provided as **Appendix 7**. Additionally, we also published a protocol explaining how to generate the IMS/IMS spectra for IM-MN analysis, which is presented as **Publication 5** and also includes a video tutorial. The supporting data for this protocol are provided as **Appendix 8**. These results have finally been presented as a poster in an international conference.

Peer-reviewed articles

Ollivier, S.; Fanuel, M.; Rogniaux, H.; Ropartz, D. Molecular Networking of High-Resolution Tandem Ion Mobility Spectra: A Structurally Relevant Way of Organizing Data in Glycomics? *Anal. Chem.* **2021**, *93* (31), 10871–10878. https://doi.org/10.1021/acs.analchem.1c01244.

Ollivier, S.; Fanuel, M.; Rogniaux, H.; Ropartz, D. Using a Cyclic Ion Mobility Spectrometer for Tandem Ion Mobility Experiments. *J. Vis. Exp.* **2022**, *179*, e63451. https://doi.org/10.3791/63451.

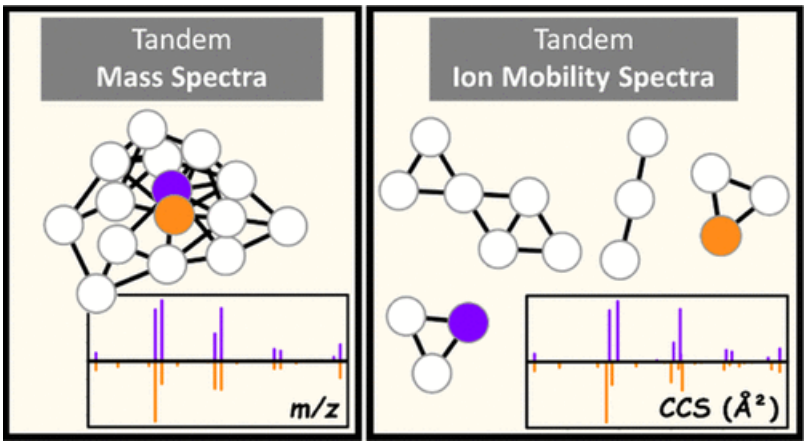
Poster presentation

Ollivier, S.; Fanuel, M.; Guitteny, C.; Le Dévéhat, F.; Rogniaux, H.; Ropartz, D. From tandem ion mobility on a Cyclic IMS platform to ion mobility-molecular networking of glycans. *70th ASMS Conference on Mass Spectrometry and Allied topics*, Minneapolis, MN, USA, June 5–9, 2022.

Publication 4 – Molecular Networking of High-Resolution Tandem Ion Mobility Spectra: A Structurally Relevant Way of Organizing Data in Glycomics?

Simon Ollivier, Mathieu Fanuel, Hélène Rogniaux*, and David Ropartz

* Corresponding author



Published in: Analytical Chemistry, 2021, 93(31), 10871-10878

analytical chemistry

pubs.acs.org/ac Article

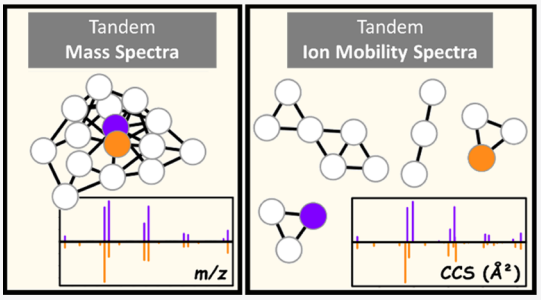
Molecular Networking of High-Resolution Tandem Ion Mobility Spectra: A Structurally Relevant Way of Organizing Data in Glycomics?

Simon Ollivier, Mathieu Fanuel, Hélène Rogniaux,* and David Ropartz

Cite This: *Anal. Chem.* 2021, 93, 10871−10878 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Data organization through molecular networks has been used in metabolomics over the past years as a way to efficiently mine the massive amount of structural information produced by tandem mass spectrometry (MS). However, glycomics lags a step behind: carbohydrate structures involve numerous levels of isomerism, making MS and tandem MS blind to many key structural features of glycans. This roadblock can in part be alleviated with gas-phase ion mobility spectrometry (IMS), a method highly sensitive to isomerism. In this work, we propose a novel strategy for structural glycomics: molecular networking of high-resolution IMS/IMS spectra. We combine the cutting-edge strategies of tandem IMS and molecular networking of spectral data. We demonstrate that—when it comes to oligosaccharides and their numerous levels of isomerisms—molecular networks based on IMS/IMS spectra are widely superior to MS/MS-based networks to sort and organize molecules with a high degree of structural relevance.

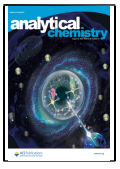


In recent years, "omics" sciences—among which proteomics and metabolomics—have become increasingly popular in life sciences. Mass spectrometry (MS) has established itself at the center of the analytical strategy in both fields, thanks to its ability to provide insight into chemical structures: ions can be mass-selected and collisionally fragmented to obtain signature product ions (MS/MS). These fragments can then notably be used to reconstruct the parent structures,[1,2] or match experimental and *in silico* spectral databases.[3−5] However, because MS-based "omics" produce huge amounts of highly multiplexed data, manual curation of the results is hardly an option. Methods have thus been developed to organize and mine the results—especially in metabolomics where an important diversity of chemical structures coexist. Molecular networks (MN) have been used for data organization since the 2010s[6] and, in 2016, the Global Natural Products Social MN platform (GNPS; https://gnps.ucsd.edu) was formally introduced to the community.[7] To build an MN, pairwise comparisons of MS/MS spectra are done through a cosine similarity score: the higher the cosine score, the more similar the fragmentation spectra. If this cosine score for a pair of spectra is above a defined threshold, the spectra are linked by an edge in the network—in this type of MN, only the edges are significant and the distances between clusters hold no particular meaning.

In glycomics, MS analysis faces many pitfalls due to various levels of isomerism that the technique—which relies on mass differences—fails to resolve. Indeed, many monosaccharide units are (stereo)isomers and are also subject to a cyclization tautomerism (forming furanosides or pyranosides). The glycosidic bond adds a level of complexity by allowing several forms of anomers and regioisomers; these levels of isomerism are reflected on ions produced in MS/MS. Added to the lack of structurally informative fragments (i.e., cross-ring fragments) and the occurrence of consecutive fragmentations during the standard MS/MS process by collision-induced dissociation (CID), these difficulties make MS blind to many structural variations in carbohydrates (for a review, see Gray et al.).[8] The situation can be partially solved using other means of activation with higher energies (e.g., XUVPD,[9] ETD,[10] He-CTD[11,12]). Perderivatization (or other labeling methods) is another alternative. Perderivatization promotes ionization and fragmentation of glycans and disrupts the equivalence between hydroxyl groups in the structure; it thus helps identify product ions and infer the regiochemistry of the glycosidic bond.[13] However, when it comes to natural samples containing widely unknown and structurally diverse compounds with—sometimes—a minute amount of material, such reaction is rather tricky or not efficient (e.g., for acidic polysaccharides) in addition to using harmful reagents.

While MS alone presents some limitations for the fine structural characterization of oligosaccharides, the technique

Received: March 22, 2021
Accepted: July 16, 2021
Published: July 29, 2021



Downloaded via INRAE on August 10, 2021 at 08:54:03 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ACS Publications © 2021 American Chemical Society

https://doi.org/10.1021/acs.analchem.1c01244
Anal. Chem. 2021, 93, 10871−10878

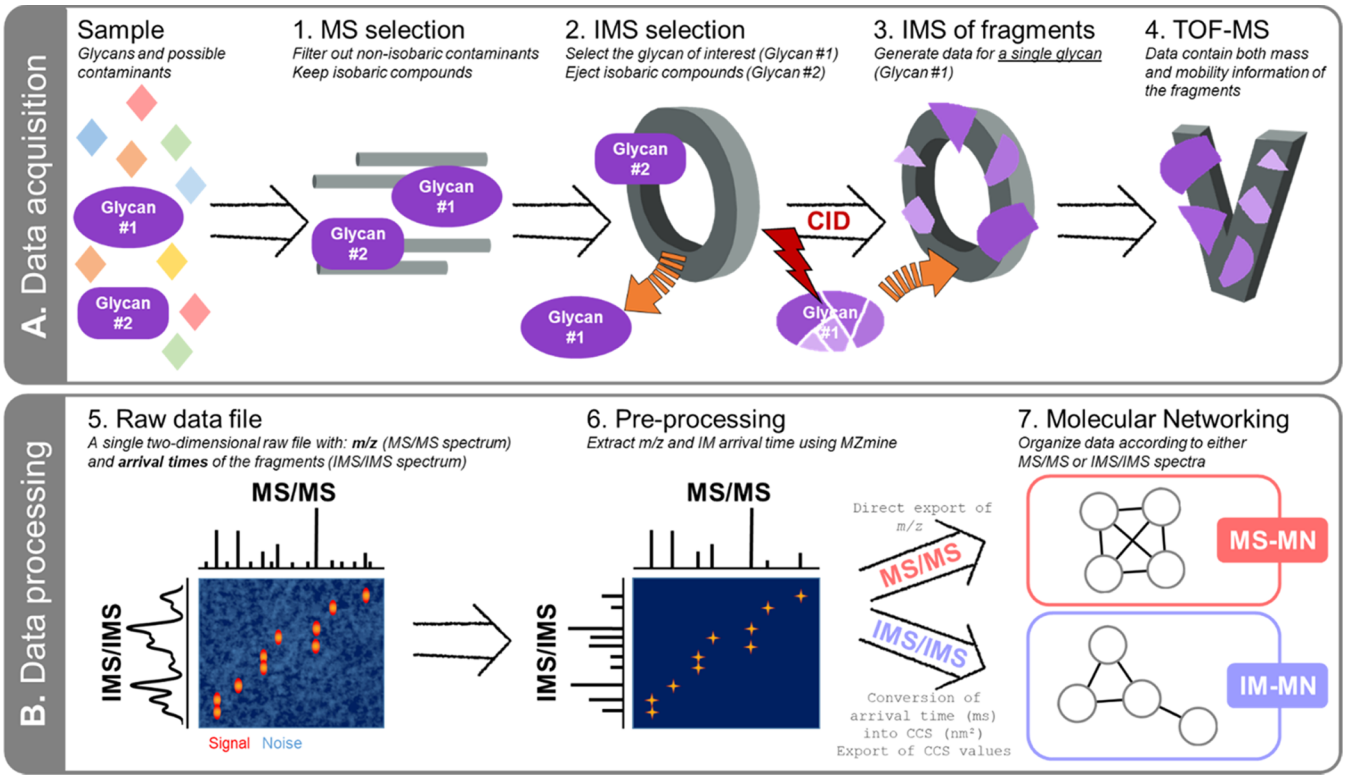


Figure 1. General data acquisition and processing strategy used to build the two networks. (A) Acquisition steps, resulting in a single raw file with two dimensions (m/z of fragments and fragments arrival time). The selection of the precursor before fragmentation is carried out from both the m/z and arrival time of the ion of interest. (B) Raw data file processing steps, leading to input data for MS-MN and IM-MN networks. Additional details about data acquisition and data processing workflow are available in Supplementary Figures 1 and 2, respectively.

can readily and efficiently be combined with upfront separation methods to help deconvolve isomerism; among these, liquid- or gas-phase chromatography or, more recently, gas-phase ion analysis methods such as ion mobility spectrometry (IMS). IMS separates ions according to their gas-phase conformation, depending on their interactions with a buffer gas and an electric field. Commercial instruments typically operate at alternating high and low fields in field asymmetric IMS (FAIMS) or in the low-field limit using drift tube IMS (DTIMS, operating with a constantly decreasing potential), traveling wave IMS (TWIMS, operating using potential waves of defined velocity and height) or the more resolutive trapped IMS (TIMS, operating by trapping the ions in a flowing buffer gas using electric fields); these last three techniques allow access to a collision cross section (CCS, in $Å^2$ or nm^2), either through direct calculation (DTIMS) or through calibration (TWIMS and TIMS).[14] CCS is a repeatable charge-independent value that can be reproduced accurately between different laboratories.[15]

IMS-MS raised interest from the glycomics community as it was notably able to discriminate intrachain anomers,[16] or match monomeric fragments with the corresponding monosaccharide.[17] IMS has recently gained more momentum in carbohydrate analysis for two main reasons: (i) the demonstration that major structural features of carbohydrates are retained during MS/MS, such as the linkage[18] or the anomericity of the glycosidic bond,[19−21] and (ii) the development of ultrahigh-resolution IMS technologies, among which the TWIMS-based Cyclic IMS[22] and the multipass Structures for Lossless Ion Manipulation (SLIM SUPER).[23] Both possess the ability to increase the IMS resolving power almost infinitely by allowing multiple passes in the ion mobility cell (therefore increasing the separation length), as well as to perform tandem IMS/IMS experiments, pending some modifications for the SLIM.[24] Analogously to MS/MS, an ion is selected based on its mobility, then subjected to CID fragmentation, and finally, the IMS spectrum of the fragments is recorded (for a review on the development of tandem IMS, see Eldrid and Thalassinos[25]). As of today and to our best knowledge, the Cyclic IMS is the only commercially available instrument equipped with the multistage IMS technology (SELECT SERIES Cyclic IMS, Waters, Wilmslow, UK); the length of the IMS cell is about 1 m, yielding a single-pass CCS/δCCS peak resolving power of about 100 for carbohydrates.[26]

Because it is now possible to perform IMS/IMS analyses, and because the structural properties of carbohydrates are retained by their fragments (as mentioned above), the hypothesis that guided our work is that *IMS/IMS analysis of carbohydrates should bring reproducible information about the parent structures, and structural relationships could be evidenced by molecular networking of IMS/IMS spectra*. We thus generated, for the first time, molecular networks based on IMS/IMS (which we called IM-MN) from 33 oligosaccharides of known structures, and compared them to MN based on MS/MS (herein coined MS-MN). We show that IM-MN offers much better insight into carbohydrate structures compared to MS-MN—as all edges in the network can be rationalized at the isomeric level. We then challenged our network by analyzing five additional oligosaccharides, having structural characteristics shared with some compounds of the initial network. We confirmed that they clustered with oligosaccharides of similar backbone while being efficiently separated from their isomers.

https://doi.org/10.1021/acs.analchem.1c01244
Anal. Chem. 2021, 93, 10871−10878

EXPERIMENTAL SECTION

Chemicals. Hexo-, pento-oligosaccharides and derivatives were used for the analyses; their provenances and references are given in Supplementary Tables 1 and 2.

IMS/IMS Experiments. Acquisitions were performed on a SELECT SERIES Cyclic IMS (Waters, Wilmslow, UK) in the m/z 50−2000 range using MassLynx 4.2 (Waters, Wilmslow, UK). Samples were infused at 10 μg/mL in methanol/water 50:50 doped with 0.5 mM LiCl at a flow rate of 10 μL/min. Details about the voltages applied in the mass spectrometer and the configuration of the cIMS cell are given in the SI (Supplementary Texts 1 and 2). $^{TW}CCS_{N2}$ calibration was performed post-acquisition using MajorMix (Waters, Wilmslow, UK) as calibrant.

First, the $[M + Li]^+$ adducts (or, if not available, the $[M + 2.Li]^{++}$ adducts) of the species of interest were selected using the quadrupole. Then, IMS/IMS analyses were done with a traveling wave height of 16 V and velocity of 350 m/s. The following sequence was used: (i) ions of interest were separated from eventual contaminant ions using IMS, (ii) ions of interest were ejected backward to a pre-array store cell, (iii) they were reinjected in the IMS cell at increased kinetic energies to induce CID, and (iv) the ion mobility (and high-resolution m/z) of all product ions was acquired (see Figure 1A and Supplementary Figure 1). Although the instrument normally allows us to perform multiple passes around the cyclic IMS cell (to increase IMS resolving power), a so-called ″wrap around″ phenomenon can occur when doing so (i.e., the faster fragments catch up to the slower ones).[22] Thus, in our experiments, IM separation was done with a single pass around the cyclic IMS cell. The detailed cIMS sequences and CID parameters are given in Supplementary Table 3.

Data Preprocessing. Similarly to LC-MS or GC-MS data, IMS-MS and IMS/IMS-MS data are bidimensional, i.e., contain time and m/z. Because of this similarity, some tools meant for chromatography, such as MZmine 2,[27] can easily be diverted for processing IMS-MS and IMS/IMS-MS data. We thus designed a workflow largely based on preexisting software and procedures (Figure 1B, Supplementary Figure 2).

First, the arrival time (AT) vs mass-to-charge ratio (m/z) dimension was extracted using MassLynx 4.2, interfaced with Driftscope 2.9 (both from Waters, Wilmslow, UK). The resulting RAW data files were directly imported in MZmine (v. 2.51)[27] for preprocessing.

In MZmine, mass detection was performed using the "wavelet transform" algorithm with a scale level of 1 and wavelet window size of 100%. Then, extracted ion mobility spectra (EIMs) were built using "ADAP chromatogram builder"[28] with a minimum group size of 5 bins, and an m/z tolerance of 0.2 Da. Isotopes were grouped using "Isotopic peak grouper" with a tolerance of 0.015 m/z and 0.5 ms. EIM deconvolution was then performed using the "Wavelets (ADAP)" algorithm in the "chromatogram deconvolution module" (S/N threshold = 1, coefficient/area threshold = 0.5, RT wavelet range: 0−0.5). These steps yielded centroid EIM spectra that were grouped in a single table using the "join aligner" module and exported as CSV including: m/z, arrival time (AT), charge (z), and fragment peak intensity by EIM.

Further processing was done before molecular networking using an R script,[29] to prepare noise-free MS/MS and IMS/IMS spectra (fragments above 0.2% relative intensity were kept):

- Background ions were removed from MS/MS spectra using the fragments' EIMs from MZmine. The EIMs were grouped based on the MS dimension. The MS intensities of peaks that differed by less than 0.1 Da were summed, giving MS/MS spectra containing *m/z−intensity pairs*.
- IMS/IMS spectra were prepared by (i) correcting the AT into drift time (DT; this consists in subtracting the time of all cIMS events preceding the final separation stage from the AT) and (ii) applying a log CCS calibration[30,31] using the Waters MajorMix as calibrant ions, yielding $^{TW}CCS_{N2}$ values in nm^2. CCS values in nm^2 are more adapted to the binning parameters of the GNPS platform. In fact, the maximum "fragment ion tolerance" on the GNPS is 0.99, which—if the CCS is expressed in Å^2—limits the tolerance to <1% (CCS in Å^2 are of the order of 100−300). A similar value will yield a larger tolerance if the CCS is expressed in nm^2. This process gave IMS/IMS spectra containing *CCS−intensity pairs*. The mean CCS deviation from reference values was 1.2%, corresponding to a CCS error of about 1.5−2 Å^2 (Supplementary Figure 3)—which was thus considered as the uncertainty on IMS/IMS fragments for IM-MN.
- Finally, all MS/MS and IMS/IMS spectra were exported as MGF files for molecular networking on GNPS.

IMS/IMS Molecular Networking (IM-MN). An MN was created using the online workflow (https://ccms-ucsd.github.io/GNPSDocumentation/) on the GNPS website. The precursor ion mass tolerance was set to 0.1 Da, and the IMS/IMS fragment ion $^{TW}CCS_{N2}$ tolerance to 0.015 nm^2 (1.5 Å^2). The tolerance was set with regard to both the uncertainty of the calibration and the binning parameters used during the acquisition of IMS data. A network was then created where edges were filtered to have a cosine score above 0.75 and more than 8 matched peaks. The output was visualized using Cytoscape 3.7.[32] Other values for the cosine score and the number of shared peaks were explored, and their impact was assessed on the networking patterns (results available in the Supplementary Information).

MS/MS Molecular Networking (MS-MN). A matching MS-MN was created with the same workflow as for IM-MN. The precursor ion mass tolerance was set to 0.1 Da, and the MS/MS fragment ion mass tolerance was set to 0.02 Da. A network was then created where edges were filtered to have a cosine score above 0.75 and more than 8 matched peaks.

RESULTS

Set of Oligosaccharides Used to Build the Initial Network. Thirty-three commercially available oligosaccharides of known structures were selected to build the initial networks. The structures are detailed in Supplementary Table 1 and are represented according to the symbol nomenclature for glycans (SNFG, Figure 2).[33,34] We selected both linear and branched structures, but with no overlapping structural moieties (e.g., mixed linkages or lateral branching by subunits that are in the backbone of other structures). We considered structures made of either hexosides or pentosides or both, among which several situations of regioisomerism, anomeric configuration of the glycosidic bond, and isomers arising from the epimerism of their subunits (e.g., Gal vs Glc or Glc vs Man). The 33 structures, with degrees of polymerization (DP) ranging from 4 to 7, are fairly typical of the oligosaccharides that can be produced during the

https://doi.org/10.1021/acs.analchem.1c01244
Anal. Chem. 2021, 93, 10871−10878

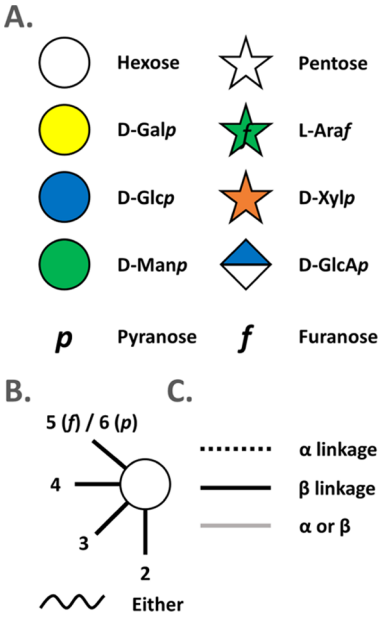


Figure 2. Symbol nomenclature for glycans with the representations of (A) the different osidic units found in the oligosaccharides used to build the network, (B) the regioisomerism of the glycosidic bond (*p*: pyranose; *f*: furanose), and (C) the anomericity of the glycosidic bond.

depolymerization of certain natural plant polysaccharides, such as cellulose, xylans, arabinoxylans, arabinans, galactans, etc.

General Data Acquisition and Processing Strategy Used to Build the MS-MN and IM-MN Networks. The different steps followed for the acquisition and processing of the data are shown schematically in Figure 1 (further details are available in Supplementary Figures 1 and 2, as well as in the Experimental Section). The acquisition comprises several steps: the first consists of a mass selection of the precursor of interest (i.e., one of the glycans). This m/z selection can, however, leave isobaric compounds, which would be co-ionized with the ion of interest. All of the isobaric compounds are thus rejected by a first stage of IMS. At the end of this step, a single isomer is selected according to its ion mobility and moved backward in the so-called "pre-array store" cell of the mass spectrometer. This procedure limits the occurrence of mixture of glycans for the following steps. The next step is the collisional fragmentation of the selected isomer in the pre-array store. The fragments produced then undergo a single-pass IMS separation in the cyclic mobility cell and are ejected to the TOF analyzer for mass measurement.

The result is a single multidimensional raw file, from which one can extract separately: the m/z profile of fragments and their mobility profile. Note that a preprocessing of mobility profile is done to obtain drift times (DT) and to convert DT into CCS (see the Experimental Section). The workflow results in two separate spectra, which we will call MS/MS spectrum and IMS/IMS spectrum throughout the manuscript. They are the input data of, respectively, the MS-MN and IM-MN networks. It is important to note that these steps imply that: (i) the input data for the IM-MN network do not contain any m/z information of the fragments; (ii) the input data for both networks relate to the fragments; and (iii) the data, for both networks, are related to a single isomer of glycan.

Comparison of IM-MN to Classical MS-MN. To assess the usefulness of MS/MS and IMS/IMS spectra for data organization of carbohydrates, we first created MS-MN and IM-MN of the 33 oligosaccharides, on the basis of their MS/MS or IMS/IMS signature, respectively. Figure 3 presents these molecular networks.

Regarding MS-MN built with MS/MS spectra, the network contains two main clusters and one isolated structure (Figure 3A). The two clusters consist of hexosides (**cluster 1**) and pentosides (**cluster 2**). Most hexosides are interconnected, and the associated cluster does not give further insight into the actual structures: there is no distinction between mannosides, galactosides, or the important variety of analyzed glucosides. For pentosides, arabinofuranosides are not distinguished from xylopyranosides either. Nevertheless, MS-MN accurately linked xylo-oligosaccharides branched with a GlcA with pure xylo-oligosaccharides sharing the same backbone in **cluster 2**. However, the analyzed xyloglucan (β-1,4-glucotetraose branched with 3 Xyl) is not linked to the other hexosides.

The IM-MN built with IMS/IMS spectra contains much more clusters: five small clusters and a larger one—which can be subdivided into two subclusters—and one isolated structure (Figure 3B). These clusters are more structurally informative compared to MS-MN—the oligosaccharides are efficiently grouped according to the backbone subunits and glycosidic bond isomerism. Indeed, the hexosides are separated in four clusters (**clusters I–IV**). **Cluster I** is a large cluster composed of two subclusters that contain (i) **subcluster Ia**: (1,4)-linked hexosides (α-1,4-glucans, β-1,4-mannans, and β-1,4-galactans) and (ii) **subcluster Ib**: α-1,6-glucans a.k.a. dextrans. **Subcluster Ib** is connected to **subcluster Ia** through the α-1,4-glucans—they are both α-linked gluco-oligosaccharides. **Cluster II** contains structures with β-1,4-glucan backbones. Notably, with the IM-MN approach, the only heteroglycan, consisting of a backbone composed of β-1,4-linked glucose branched with xylose, groups with pure β-1,4-glucans a.k.a. cello-oligosaccharides. β-1,2-glucans and β-1,3-glucans each form a different cluster: **clusters III and IV**, respectively. The branched α-linked mannopentaose, which appears disconnected, is an expected outlier since it shares no common substructure with any other oligosaccharide in the network. A similar substructure-related separation can also be observed for pentosides, which seem to be sorted according to the monosaccharide subunits composing the backbone. Indeed, xylo-oligosaccharides—linear or substituted by GlcA—cluster separately from arabino-oligosaccharides, respectively, in **clusters V and VI**.

Behavior of Complex Oligosaccharides in IM-MN. We challenged the IM-MN strategy with five additional oligosaccharides, highlighted as yellow-filled nodes in the extended network presented in Figure 4. The detailed structures of the five oligosaccharides are given in Supplementary Table 2. Briefly, they consist of two mixed-linked gluco-oligosaccharides and three arabinoxylan oligosaccharides (AXOS). All five share structural motifs with other oligosaccharides of the network.

The two mix-linked gluco-oligosaccharides cluster differently. The α-1,6/α-1,4-glucotetraose (one bond in α-1,6 and two bonds in α-1,4) is connected to both malto-oligosaccharides (α-1,4) and dextran (α-1,6) in the network (i.e., to both **cluster Ia** and **cluster Ib**); however, it is hard to conclude whether its IMS/IMS signature contains distinctive features of α-1,4 and α-1,6 linkages, since malto-oligosaccharides and dextrans were already connected in the initial IM-MN (Figure 3B) (IMS/IMS spectra for two representative structures of **clusters Ia** and **Ib** are available in Supplementary Figure 4). On the other hand, the mixed β-1,4/β-1,3-glucotetraose (two bonds in β-1,4 and one in β-1,3) is directly linked to **cluster II** (cello-oligosaccharides, β-

https://doi.org/10.1021/acs.analchem.1c01244
Anal. Chem. 2021, 93, 10871−10878

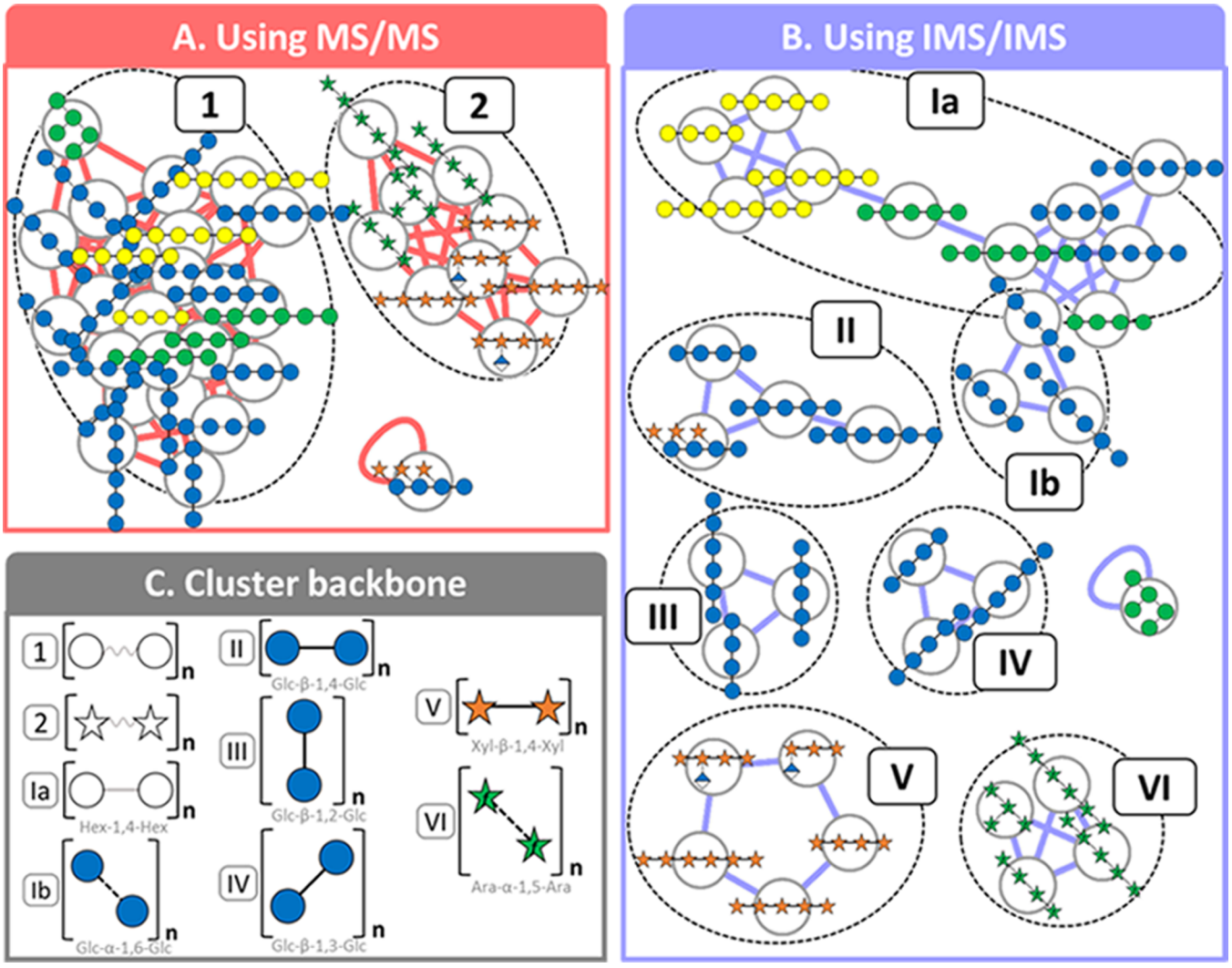


Figure 3. Molecular networks obtained for the 33 oligosaccharides of known structures. (A) MS-MN (top left) and (B) IM-MN (right). Both molecular networks show the clusters formed with a cosine threshold of 0.75 and a minimum of 8 shared pics required to pair 2 fragmentation spectra. (C) Backbone of the oligosaccharides found in each cluster.

1,4), but not to the β-1,3-gluco-oligosaccharides, which cluster apart (**cluster IV**). This observation suggests that the parameters optimized for the construction of the IM-MN favor the main backbone linkage in the clusters; for the latter example, the Glc-(β-1,4)-Glc linkage, which is prominent in the structure of the mix-linked glucotetraose, was presumably the driving feature for the clustering.

Finally, the three AXOS all cluster with xylo-oligosaccharides and glucuronoxylo-oligosaccharides in **cluster V** at a cosine threshold of 0.75, but not with arabino-oligosaccharides (**cluster VI**)—even for the one containing two arabinosides as substituents. This corroborates the hypothesis of IM-MN being able to accurately organize IMS/IMS spectral data in priority according to the structural properties of the backbone of the carbohydrates, at a defined cosine score.

DISCUSSION

Molecular networks were constructed for 33 oligosaccharides of known structures, either based on their MS/MS spectra (MS-MN)—according to a classic approach used in metabolomics, or based on the IMS/IMS spectra (IM-MN)—following a new approach, made possible by the instrumental setup of the recently released SELECT SERIES mass spectrometer. Our results show that MS-MN (built solely from the m/z dimension of the fragmentation data) bring minimal insight into carbohydrate structures: some oligosaccharides can be clustered according to the backbone subunits, but this ability seems limited as not all oligosaccharides can be connected through their backbone (i.e., xyloglucan is not linked to the other hexosides). Overall, MS-MN does not see beyond the mass of the hexose or pentose subunits; this was yet expected with regard to the various levels of isomerism involved.

In contrast, IM-MN appears widely superior for the organization of oligosaccharide structures, with the ability to classify oligosaccharides according to the various levels of isomerism involved. Figure 5 shows the MS/MS spectra and IMS/IMS spectra of two structures (cellopentaose and maltopentaose), which are connected in **cluster 1** in the MS-MN network but separated in **clusters II** and **Ia** in IM-MN. It is evident that the MS/MS spectra of the two pentasaccharides share a high level of similarity (Figure 5A). The similarity arises from the isobaric fragments between cellopentaose and maltopentaose, which differ only in their intrachain anomerism, and because MS is inherently incapable of resolving isobaric

https://doi.org/10.1021/acs.analchem.1c01244
Anal. Chem. 2021, 93, 10871−10878

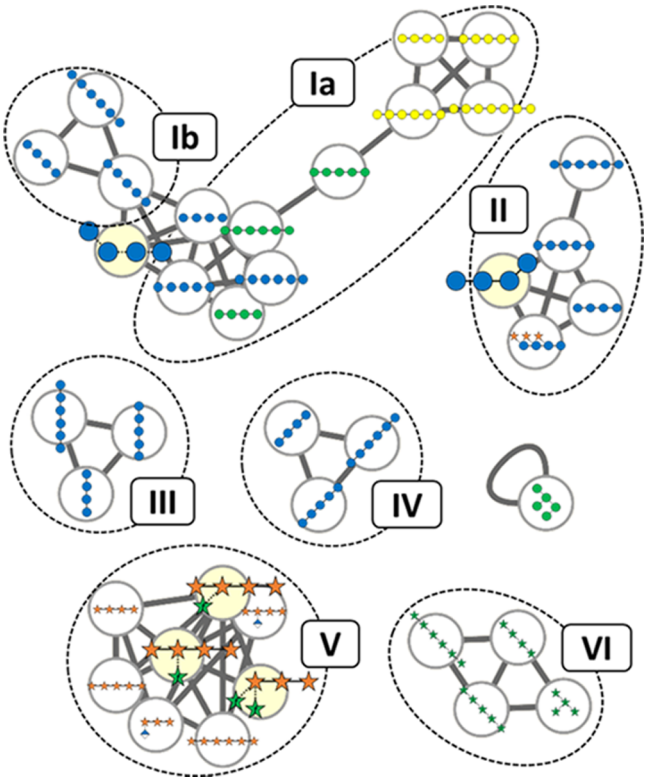


Figure 4. Extended IM-MN in which the initial IM-MN (Figure 3B) was challenged with structures containing moieties (subunits or linkages) from multiple standards in the initial network. These structures—arabinoxylan oligosaccharides and mix-linked gluco-oligosaccharides—are highlighted in yellow.

fragments. On the other hand, the IMS/IMS spectra reveal clear differences between the two glycans (Figure 5B), thanks to the ability of ion mobility to distinguish between isomeric fragments, including fragments differing by their intrachain anomerism[21] (*NB.* the ATD of the pentasaccharides before any processing are available in Supplementary Figure 5). The different nature and specific characteristics of the input data of MS-MN and IM-MN thus drive the observed differences between the two networks in terms of data organization.

While most groupings and separations are easy to interpret, some require more in-depth analysis: as illustrated above for two features of these clusters, it is noteworthy that **subcluster Ia** is separated from **cluster II** at our defined cosine score threshold, even though they both contain 1,4-linked (gluco)-oligosaccharides. Considering that malto-oligosaccharides, β-1,4-mannans, and β-1,4-galactans (grouped in **subcluster Ia**) are all epimers of cello-oligosaccharides (**cluster II**) in, respectively, C-1, C-2, and C-4, this could indicate that the gas-phase conformations, thus the IM profiles, are influenced by the presence of a hydroxyl group in axial position (i.e., cello-oligosaccharides cluster apart from other 1,4-linked oligosaccharides because all of their hydroxyl groups are in equatorial position). All clusters and subclusters of the IM-MN can therefore be rationalized according to the regioisomerism, epimerism, and even anomerism of the oligosaccharides—which are all key elements in the structural description of carbohydrates. Our next questions were: (i) will this organization retain significance if more complex structures are included? And (ii) how will shared structural patterns impact the existing clusters?

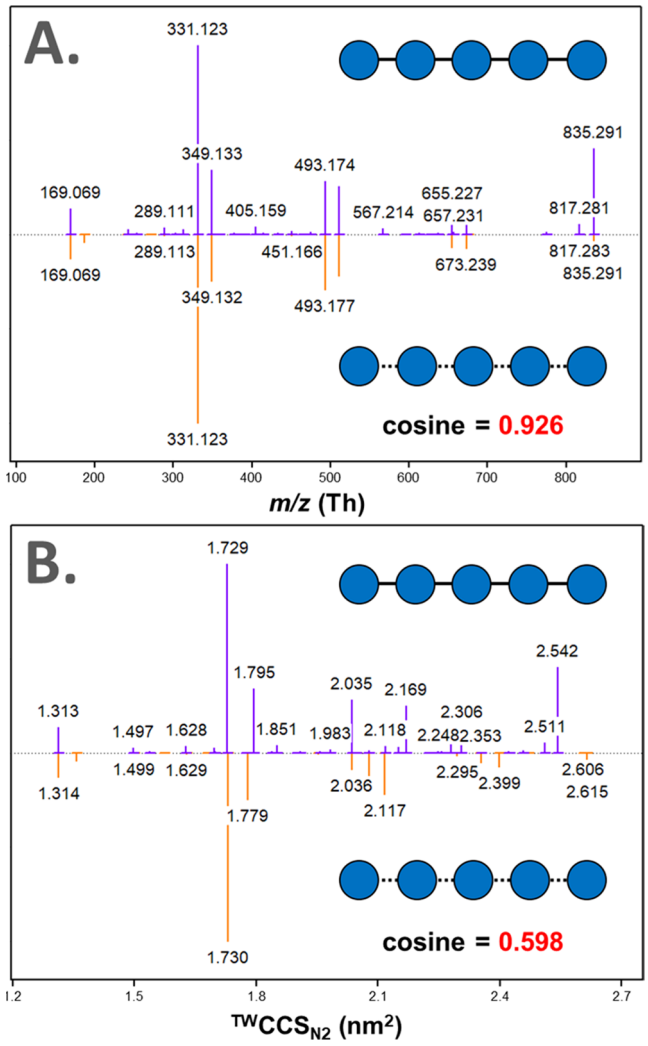


Figure 5. Mirror view comparing the centroided fragmentation spectra of cellopentaose (violet) and maltopentaose (orange): (A) MS/MS spectra and (B) IMS/IMS spectra. The arrival time distribution of the fragments corresponding to the IMS/IMS spectra in (B) is shown in Supplementary Figure 5. These pentasaccharides differ only by their intrachain anomerism, which is a key structural property.

To answer the above questions, the IM-MN strategy was challenged with oligosaccharides of higher complexity sharing structural motifs with different structures from the initial IM-MN: mix-linked β-glucans and AXOS. The results for both types of oligosaccharides suggest that, in IM-MN, the clustering is mainly driven by the backbone of the oligosaccharides—including the anomeric configuration of the bond—but not by secondary motifs. Interestingly, the addition of five novel structures did not significantly change the topology of the network, and the extended network kept a high structural significance as described above.

In this proof of concept, we have not addressed the question of the impact of cosine score threshold in clusterization, as it would require much larger datasets to establish general rules—varying the cosine score could however be a way to get additional information. For instance, when changing the cosine threshold of our extended IM-MN to lower cosine thresholds, the isomeric specificity of the clusters would gradually be lost, whereas at higher cosine thresholds, structurally related compounds would ultimately not be grouped in the same

https://doi.org/10.1021/acs.analchem.1c01244
Anal. Chem. 2021, 93, 10871−10878

cluster (Supplementary Figure 6). In contrast, the number of matched peaks has little impact on the topology of the network, except if this parameter is set too high compared to the degree of polymerization of the oligosaccharides (Supplementary Figure 7).

CONCLUSIONS

In this work, we propose a novel strategy for structural glycomics: molecular networking of high-resolution IMS/IMS spectra. This strategy combines two forefront approaches in structural analysis for data acquisition and data processing. We demonstrate that—when it comes to oligosaccharides and their numerous levels of isomerisms—mass-deconvolved IMS/IMS spectra are more powerful than MS/MS spectra to sort and organize molecules with a high degree of structural relevance.

One current limit of the approach is that the GNPS platform was designed to handle only a single spectral dimension (i.e., the m/z). Thus, IMS/IMS data had to be "degraded" before computing the cosine score to keep only the CCS values, i.e., the m/z dimension of the IMS/IMS spectra was lost in the process. In the future, it will be interesting to develop new scoring systems that retain the multidimensionality of IMS/IMS data. Molecular networks can then be expected to achieve greater accuracy and relevance when accounting for both m/z and CCS—as this information is not structurally redundant. Therefore, they should together provide an enriched and unequivocal structural signature of each fragment ion leading to a unique carbohydrate fingerprint. In addition, it will be interesting to include upstream chromatographic separation before IMS/IMS analysis, for example, to explore complex biological systems with a great diversity of glycans, and to compare the current approach to workflows that make the most of LC-MS analyses such as feature-based molecular networking.[35]

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.1c01244.

General mass spectrometry and ion mobility parameters; acquisition and processing of IMS/IMS-MS data; CCS calibration plot; additional IMS/IMS spectra; evolution of the IM-MN depending on the thresholds for the cosine score and the number of shared peaks (PDF)

Detailed description of the oligosaccharide standards; and individual MS/MS and IMS/IMS parameters for each oligosaccharide standard (XLSX)

AUTHOR INFORMATION

Corresponding Author

Hélène Rogniaux − *INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France;* orcid.org/0000-0001-6083-2034; Email: helene.rogniaux@inrae.fr

Authors

Simon Ollivier − *INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France;* orcid.org/0000-0002-7671-1736

Mathieu Fanuel − *INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France;* orcid.org/0000-0001-8384-8266

David Ropartz − *INRAE, UR BIA, F-44316 Nantes, France; INRAE, BIBS Facility, F-44316 Nantes, France;* orcid.org/0000-0003-4767-6940

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.1c01244

Author Contributions

S.O., M.F., H.R., and D.R. conceived the study. S.O. performed the experiments. S.O., M.F., H.R., and D.R. interpreted the results. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

The R code used for preprocessing is available at: https://github.com/siollivier/IM-MN.

The MZmine batch parameters (XML format), the IMS/IMS and MS/MS spectra (MGF format), the metadata, example raw files, and example input files for the R notebook are available at https://osf.io/ne7wd/. The raw data for the full dataset are available from the authors upon reasonable request. Molecular networks can be downloaded at: MS-MN: https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=feded1b950bc49cdba71b12d6552105a IM-MN: https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=6d4b6c001e6e456492ecdb0260870db6 Extended IM-MN: https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=ab9c5a446d42471d895f38fc5d1d23a1

ACKNOWLEDGMENTS

The authors acknowledge financial support from the French Agence Nationale de la Recherche (grant ANR-18-CE29-0006) for funding the salary of S.O. They are grateful to Estelle Bonnin (INRAE, UR BIA, F-44316 Nantes, FR) for providing the β-1,4-galactooligosaccharides.

REFERENCES

(1) David, M.; Fertin, G.; Rogniaux, H.; Tessier, D. *J. Proteome Res.* **2017**, *16*, 3030−3038.

(2) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. *Nat. Methods* **2019**, *16*, 299.

(3) Ramos, A. E. F.; Le Pogam, P.; Alcover, C. F.; N'Nang, E. O.; Cauchie, G.; Hazni, H.; Awang, K.; Bréard, D.; Echavarren, A. M.; Frédérich, M.; Gaslonde, T.; Girardot, M.; Grougnet, R.; Kirillova, M. S.; Kritsanida, M.; Lémus, C.; Ray, A.-M. L.; Lewin, G.; Litaudon, M.; Mambu, L.; Michel, S.; Miloserdov, F. M.; Muratore, M. E.; Richomme-Peniguel, P.; Roussi, F.; Evanno, L.; Poupon, E.; Champy, P.; Beniddir, M. A. *Sci. Data* **2019**, *6*, 1−6.

(4) Olivier-Jimenez, D.; Chollet-Krugler, M.; Rondeau, D.; Beniddir, M. A.; Ferron, S.; Delhaye, T.; Allard, P.-M.; Wolfender, J.-L.; Sipman, H. J. M.; Lücking, R.; Boustie, J.; Le Pogam, P. *Sci. Data* **2019**, *6*, No. 294.

(5) Allard, P.-M.; Péresse, T.; Bisson, J.; Gindro, K.; Marcourt, L.; Pham, V. C.; Roussi, F.; Litaudon, M.; Wolfender, J.-L. *Anal. Chem.* **2016**, *88*, 3317−3323.

(6) Watrous, J.; Roach, P.; Alexandrov, T.; Heath, B. S.; Yang, J. Y.; Kersten, R. D.; Voort, M.; van der Pogliano, K.; Gross, H.; Raaijmakers, J. M.; Moore, B. S.; Laskin, J.; Bandeira, N.; Dorrestein, P. C. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E1743−E1752.

(7) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W.-T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C.-C.;

Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C.-C.; Yang, Y.-L.; Humpf, H.-U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; P, C. A. B.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P.-M.; Phapale, P.; Nothias, L.-F.; Alexandrov, T.; Litaudon, M.; Wolfender, J.-L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D.-T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. O.; Pogliano, K.; Linington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. *Nat. Biotechnol.* **2016**, *34*, 828−837.

(8) Gray, C. J.; Migas, L. G.; Barran, P. E.; Pagel, K.; Seeberger, P. H.; Eyers, C. E.; Boons, G.-J.; Pohl, N. L. B.; Compagnon, I.; Widmalm, G.; Flitsch, S. L. *J. Am. Chem. Soc.* **2019**, *141*, 14463−14479.

(9) Ropartz, D.; Giuliani, A.; Fanuel, M.; Hervé, C.; Czjzek, M.; Rogniaux, H. *Anal. Chim. Acta* **2016**, *933*, 1−9.

(10) Wolff, J. J.; Leach, F. E.; Laremore, T. N.; Kaplan, D. A.; Easterling, M. L.; Linhardt, R. J.; Amster, I. J. *Anal. Chem.* **2010**, *82*, 3460−3466.

(11) Ropartz, D.; Li, P.; Fanuel, M.; Giuliani, A.; Rogniaux, H.; Jackson, G. P. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 1614−1619.

(12) Buck-Wiese, H.; Fanuel, M.; Liebeke, M.; Le Mai Hoang, K.; Pardo-Vargas, A.; Seeberger, P. H.; Hehemann, J.-H.; Rogniaux, H.; Jackson, G. P.; Ropartz, D. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 1249−1259.

(13) Morelle, W.; Slomianny, M.-C.; Diemer, H.; Schaeffer, C.; Dorsselaer, A.; van Michalski, J.-C. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 2637−2649.

(14) Gabelica, V.; Marklund, E. *Curr. Opin. Chem. Biol.* **2018**, *42*, 51−59.

(15) Hernández-Mesa, M.; D'Atri, V.; Barknowitz, G.; Fanuel, M.; Pezzatti, J.; Dreolin, N.; Ropartz, D.; Monteau, F.; Vigneau, E.; Rudaz, S.; Stead, S.; Rogniaux, H.; Guillarme, D.; Dervilly, G.; Le Bizec, B. *Anal. Chem.* **2020**, *92*, 5013−5022.

(16) Hofmann, J.; Hahm, H. S.; Seeberger, P. H.; Pagel, K. *Nature* **2015**, *526*, 241−244.

(17) Gray, C. J.; Schindler, B.; Migas, L. G.; Pičmanová, M.; Allouche, A. R.; Green, A. P.; Mandal, S.; Motawia, M. S.; Sánchez-Pérez, R.; Bjarnholt, N.; Møller, B. L.; Rijs, A. M.; Barran, P. E.; Compagnon, I.; Eyers, C. E.; Flitsch, S. L. *Anal. Chem.* **2017**, *89*, 4540−4549.

(18) Mookherjee, A.; Uppal, S. S.; Murphree, T. A.; Guttman, M. *J. Am. Soc. Mass Spectrom.* **2020**, *32*, 581−589.

(19) Schindler, B.; Barnes, L.; Renois, G.; Gray, C.; Chambert, S.; Fort, S.; Flitsch, S.; Loison, C.; Allouche, A. R.; Compagnon, I. *Nat. Commun.* **2017**, *8*, No. 973.

(20) Pellegrinelli, R. P.; Yue, L.; Carrascosa, E.; Warnke, S.; Ben Faleh, A.; Rizzo, T. R. *J. Am. Chem. Soc.* **2020**, *142*, 5948−5951.

(21) Ollivier, S.; Tarquis, L.; Fanuel, M.; Li, A.; Durand, J.; Laville, E.; Potocki-Veronese, G.; Ropartz, D.; Rogniaux, H. *Anal. Chem.* **2021**, *93*, 6254−6261.

(22) Giles, K.; Ujma, J.; Wildgoose, J.; Pringle, S.; Richardson, K.; Langridge, D.; Green, M. *Anal. Chem.* **2019**, *91*, 8564−8573.

(23) Deng, L. L.; Webb, I. K.; Garimella, S. V. B.; Hamid, A. M.; Zheng, X. Y.; Norheim, R. V.; Prost, S. A.; Anderson, G. A.; Sandoval, J. A.; Baker, E. S.; Ibrahim, Y. M.; Smith, R. D. *Anal. Chem.* **2017**, *89*, 4628−4634.

(24) Bansal, P.; Yatsyna, V.; AbiKhodr, A. H.; Warnke, S.; Ben Faleh, A.; Yalovenko, N.; Wysocki, V. H.; Rizzo, T. R. *Anal. Chem.* **2020**, *92*, 9079−9085.

(25) Eldrid, C.; Thalassinos, K. *Biochem. Soc. Trans.* **2020**, *48*, 2457−2466.

(26) Ropartz, D.; Fanuel, M.; Ujma, J.; Palmer, M.; Giles, K.; Rogniaux, H. *Anal. Chem.* **2019**, *91*, 12030−12037.

(27) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. *BMC Bioinf.* **2010**, *11*, No. 395.

(28) Smirnov, A.; Jia, W.; Walker, D. I.; Jones, D. P.; Du, X. *J. Proteome Res.* **2018**, *17*, 470−478.

(29) R Core Team. *R: A Language and Environment for Statistical Computing.*; R Foundation for Statistical Computing: Vienna, Austria, 2018.

(30) Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V.; Ruotolo, B. T. *Anal. Chem.* **2010**, *82*, 9557−9565.

(31) Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. *Nat. Protoc.* **2008**, *3*, 1139−1152.

(32) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. *Genome Res.* **2003**, *13*, 2498−2504.

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

(34) Neelamegham, S.; Aoki-Kinoshita, K.; Bolton, E.; Frank, M.; Lisacek, F.; Lütteke, T.; O'Boyle, N.; Packer, N. H.; Stanley, P.; Toukach, P.; Varki, A. *Glycobiology* **2019**, *29*, 620−624.

(35) Nothias, L.-F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H.; Fleischauer, M.; Aicheler, F.; Aksenov, A. A.; Alka, O.; Allard, P.-M.; Barsch, A.; Cachet, X.; Caraballo-Rodriguez, A. M.; Da Silva, R. R.; Dang, T.; Garg, N.; Gauglitz, J. M.; Gurevich, A.; Isaac, G.; Jarmusch, A. K.; Kameník, Z.; Kang, K. B.; Kessler, N.; Koester, I.; Korf, A.; Le Gouellec, A.; Ludwig, M.; H, C. M.; McCall, L.-I.; McSayles, J.; Meyer, S. W.; Mohimani, H.; Morsy, M.; Moyne, O.; Neumann, S.; Neuweger, H.; Nguyen, N. H.; Nothias-Esposito, M.; Paolini, J.; Phelan, V. V.; Pluskal, T.; Quinn, R. A.; Rogers, S.; Shrestha, B.; Tripathi, A.; van der Hooft, J. J. J.; Vargas, F.; Weldon, K. C.; Witting, M.; Yang, H.; Zhang, Z.; Zubeil, F.; Kohlbacher, O.; Böcker, S.; Alexandrov, T.; Bandeira, N.; Wang, M.; Dorrestein, P. C. *Nat Methods* **2020**, *17*, 905−908.

https://doi.org/10.1021/acs.analchem.1c01244
Anal. Chem. 2021, 93, 10871−10878

Publication 5 – Using a Cyclic Ion Mobility Spectrometer for Tandem Ion Mobility Experiments

Simon Ollivier, Mathieu Fanuel, Hélène Rogniaux*, and David Ropartz

* Corresponding author

Published in: Journal of Visualized Experiments (JoVE), 2022, N° e63451

jove

Using a Cyclic Ion Mobility Spectrometer for Tandem Ion Mobility Experiments

Simon Ollivier[1,2], **Mathieu Fanuel**[1,2], **Hélène Rogniaux**[1,2], **David Ropartz**[1,2]

[1] INRAE, UR BIA, F-44316 Nantes, France [2] INRAE, BIBS Facility, F-44316 Nantes, France

Corresponding Author

Hélène Rogniaux

helene.rogniaux@inrae.fr

Citation

Ollivier, S., Fanuel, M., Rogniaux, H., Ropartz, D. Using a Cyclic Ion Mobility Spectrometer for Tandem Ion Mobility Experiments. *J. Vis. Exp.* (179), e63451, doi:10.3791/63451 (2022).

Date Published

January 20, 2022

DOI

10.3791/63451

URL

jove.com/video/63451

Abstract

Accurate characterization of chemical structures is important to understand their underlying biological mechanisms and functional properties. Mass spectrometry (MS) is a popular tool but is not always sufficient to completely unveil all structural features. For example, although carbohydrates are biologically relevant, their characterization is complicated by numerous levels of isomerism. Ion mobility spectrometry (IMS) is an interesting complement because it is sensitive to ion conformations and, thus, to isomerism.

Furthermore, recent advances have significantly improved the technique: the last generation of Cyclic IMS instruments offers additional capabilities compared to linear IMS instruments, such as an increased resolving power or the possibility to perform tandem ion mobility (IMS/IMS) experiments. During IMS/IMS, an ion is selected based on its ion mobility, fragmented, and reanalyzed to obtain ion mobility information about its fragments. Recent work showed that the mobility profiles of the fragments contained in such IMS/IMS data can act as a fingerprint of a particular glycan and can be used in a molecular networking strategy to organize glycomics datasets in a structurally relevant way.

The goal of this protocol is thus to describe how to generate IMS/IMS data, from sample preparation to the final Collision Cross Section (CCS) calibration of the ion mobility dimension that yields reproducible spectra. Taking the example of one representative glycan, this protocol will show how to build an IMS/IMS control sequence on a Cyclic IMS instrument, how to account for this control sequence to translate IMS arrival time into drift time (i.e., the effective separation time applied to the ions), and how to extract the relevant mobility information from the raw data. This protocol is designed to clearly explain the critical points of an IMS/IMS experiment and thus help new Cyclic IMS users perform straightforward and reproducible acquisitions.

Copyright © 2022 JoVE Journal of Visualized Experiments

Introduction

The complete chemical characterization of biomolecules is key to understanding their underlying biological and functional properties. To this end, "omics" sciences have developed in recent years, aiming for the large-scale characterization of chemical structures at biological concentrations. In proteomics and metabolomics, MS has become a core tool to unravel the structural heterogeneity found in biological media-notably thanks to its sensitivity and ability to provide structural information through tandem MS (MS/MS). In MS/MS strategies, an ion is selected according to its mass, then fragmented, and finally, the masses of its fragments are acquired to establish a fingerprint of the molecule. MS/MS spectra can, in particular, be used to match spectral databases[1,2], or tentatively reconstruct the parent structures[3,4]. Under the assumption that similar spectra belong to similar compounds, MS/MS data can also be used to build molecular networks (MNs) connecting related species through a similarity score[5,6].

However, because of the inherent property of MS to detect the mass-to-charge ratio (*m/z*) of ions, the technique is blind to a number of structural features that fall within the range of (stereo)isomerism. For example, carbohydrates are made of several monosaccharide subunits, many of which are stereoisomers or even epimers (e.g., Glc vs. Gal or Glc vs. Man). These subunits are linked by glycosidic bonds, which can differ by the position of the linkage (regioisomerism) and the steric configuration of the anomeric carbon (anomerism). These characteristics make it difficult for standalone MS to distinguish between carbohydrate isomers[7], and only regioisomerism can be addressed using high-energy activation methods[8,9,10]. Although derivatization is an option to disrupt the equivalence of stereoisomeric groups[11], it requires extensive sample preparation. Another, more straightforward option is to couple MS with an analytical dimension sensitive to isomerism, such as IMS.

Because this protocol is designed for users who are already familiar with the basic concepts of IMS, and because detailed reviews are available elsewhere[12,13], only a brief overview of the principles of IMS is given here. IMS is a gas-phase separation method that relies on the interaction of ions with a buffer gas and an electric field, ultimately separating ions according to their gas-phase conformations. Different principles of IMS coupled to MS can be found on commercial instruments: some operate at alternating high and low electric fields (field asymmetric IMS, FAIMS), while most operate within the low field limit—notably drift tube IMS (DTIMS, linearly decreasing electric field), traveling wave IMS (TWIMS, symmetric potential waves), and trapped IMS (TIMS, high flow of buffer gas trapping ions against electric fields)[13]. The low-field methods allow access to a so-called CCS, a property of the ion-gas pair that represents the surface (in $Å^2$ or nm^2) of the ion that interacts with the buffer gas during the separation. CCS is theoretically instrument-independent and is thus useful to generate data that can be reproduced between different laboratories[14]. Ion mobility separations can be impacted by various parameters and, notably, by fluctuations of the gas pressure and gas temperature in the mobility cell. The CCS calibration is a way to remedy this, as both the calibrant and the species of interest will be similarly affected[13]. However, it is mandatory to install the instrument in a temperature-controlled room and to have a reliable gas pressure control system.

Copyright © 2022 JoVE Journal of Visualized Experiments

An interesting evolution of IMS is IMS/IMS, which was first introduced in 2006 by Clemmer's group as an analog of MS/MS[15,16]. In IMS/IMS, an ion of interest is selectively isolated based on its ion mobility; it is then activated (until possible fragmentation), and a new IMS analysis of the activated ion or fragments is performed. In the first instrumental design, two IMS cells were put in series, separated by an ion funnel where the activation stood. Since then, although a number of IMS/IMS setups were proposed (for a review, see Eldrid and Thalassinos[17]), the first commercial mass spectrometer with IMS/IMS capability only became available in 2019[18]. This instrument substantially improved the initial concept by combining it with another technological breakthrough: a cyclic design of the IMS cell.

The cyclic IMS cell theoretically allows increasing near-infinitely the drift path length and, thus, the resolving power of the instrument[19]. This was achieved by means of a particular instrument geometry, where the cyclic TWIMS cell is placed orthogonally to the main ion optical axis. A multifunction array region at the entrance of the IMS cell allows controlling the direction of the ion path: (i) sending ions sideways for IMS separation, (ii) forward for MS detection, or (iii) backward from the IMS cell to be stored in a prearray cell. From this prearray store cell, the ions can be activated and the fragments reinjected in the IMS cell for ion mobility measurement, an approach that has been successfully used to characterize stereoisomers[20]. Ultimately, the collected data contain ion mobility and *m/z* information for the precursor and its fragments.

In a recent publication that used this cyclic design for glycan analyses (Ollivier et al.[21]), we showed that the mobility profile of the fragments contained in such IMS/IMS data acts as a fingerprint of a biomolecule that can be used in a molecular networking strategy. The resulting network, called IM-MN, led to the organization of glycomics datasets in a structurally relevant way, whereas the network built solely from MS/MS data (MS-MN) revealed little information. To complement this publication and help Cyclic IMS users implement this workflow, this protocol provides a complete description of the protocol used to collect the data. This protocol focuses only on the generation of the IMS/IMS data that users can then use to build IM-MN networks (see[21])—or for any other application of their choice. Building of IM-MN will not be considered herein, as protocols for molecular networking are already available[22]. The crucial points that must be followed to generate valuable and reproducible IMS/IMS acquisitions are highlighted. Taking the example of one of the oligosaccharides studied by Ollivier *et al.*[21], the following steps are detailed: (i) sample preparation, (ii) tuning of the Cyclic IMS instrument, (iii) automated peak-picking of the data, and (iv) CCS calibration.

Protocol

NOTE: An overview of the protocol is provided in **Figure 1**. The parameters used for the experiments described in the present protocol can be found in **Supplemental Table S1** and **Supplemental Table S2**.

1. Preparation of the sample solution

NOTE: The protocol is described using an arabinoxylan pentasaccharide (2^3-α-L-arabinofuranosyl-xylotetraose or XA^2XX; see the **Table of Materials**) as an example.

1. Preparation of the solvent: 500 µM LiCl in 50:50 H_2O:MeOH (vol./vol.).
    1. Prepare a 100 mM stock solution of lithium chloride (LiCl) in H_2O by weighing 212 mg of LiCl and add 50

mL of high-purity deionized water (H_2O) in a 50 mL polypropylene conical tube. Shake until completely dissolved.

NOTE: The solvent is doped with a lithium salt to promote the formation of $[M+Li]^+$ adducts in the ion source of the spectrometer, as it usually yields better-quality fragmentation spectra compared to other alkali adducts. The use of LiCl is recommended because organic acids (and thus their salts) have previously been found to impact IMS profiles[23].

2. In a glass bottle, dilute the LiCl stock solution 200x: to 250 µL of the stock solution, add 24.75 mL of H_2O. Add 25 mL of methanol (MeOH) to reach a final concentration of LiCl of 500 µM in 50:50 H_2O:MeOH (v/v). Sonicate for 2 min to degas the solvent.

NOTE: MeOH presents a health hazard (H225, H301, H311, H331, H370); manipulate under an extractor hood wearing a lab coat, gloves, and eye protection. A proportion of 50:50 MeOH/H_2O (v/v) appears to be the best solvent for the ionization of oligosaccharides; however, MeOH can be substituted by acetonitrile (ACN) if needed.

2. In a 1.5 mL polypropylene tube, weigh 1 mg of the carbohydrate. Dissolve with an appropriate volume of 500 µM LiCl to reach a concentration of 1 mg/mL. Dilute to a final concentration of 10 µg/mL in 50:50 MeOH/H_2O + 500 µM LiCl. Store at 4 °C.

NOTE: The concentration of 10 µg/mL was chosen to optimize the signal over all fragment ions during IMS/IMS-MS (this is for a pure compound; increase the concentration when working on mixtures). For the acquisition of reference IMS/IMS spectra, do not dilute the sample further: saturation of the MS detector prior to fragmentation is expected, although the instrument offers options to correct it (see step 3.2.).

2. Tuning of the Cyclic IMS mass spectrometer

NOTE: Software-related instructions (windows, menus, and commands) are highlighted in **bold**.

1. Open the instrument console from the instrument control software (**MS tune** page, see the software details in the **Table of Materials**), and put the instrument in **Operate** mode. Wait for at least 3 h for the high voltages to stabilize in the IMS cell.

NOTE: For the best reproducibility, the voltages in the IMS cell need to be completely stabilized. Turn on the high voltages and let the instrument stabilize overnight before any cyclic IMS analysis. Furthermore, the pressure and temperature in the ion mobility cell must be kept as constant as possible. Although a readback for the pressure is available in the **Vacuum** tab, no readback is available for the temperature. Keep the instrument in a thermostated laboratory. The instrument used in this work operates at 1.75 mbar in a laboratory thermostated at 20 °C.

2. Cyclic IMS instrument setup

NOTE: Standard solutions must be infused using the built-in fluidics system for the instrument setup.

1. Place the fluidics containers filled with the appropriate manufacturer-provided standards on the fluidics system: Reservoir B ('Lockmass'): 10 pg/µL leucine enkephaline (LEU ENK) in 50:50 ACN/H_2O + 0.1% formic acid; Reservoir C ('Calibrant'): MajorMix.

Copyright © 2022 JoVE Journal of Visualized Experiments

NOTE: In this protocol, the MajorMix calibration solution will be used to calibrate both the *m/z* and CCS dimensions. For practical reasons, an external CCS calibration will be performed (see step 5 of the protocol); hence, it is also possible to use an in-house calibrant mixture for the CCS and another calibrant for the *m/z* (e.g., sodium formate or sodium iodide).

2. On the **Tune** page of the Quartz console, go to the **Fluidics** tab. Set the sample fluidics to reservoir C and the reference fluidics to reservoir B. Infuse both solutions consecutively in the ion source to check the MS signal.

3. Perform the ADC setup, detector setup (using LEU ENK), and mass calibration (see the **Table of Materials** for the calibration solution) from the **Instrument Setup** page according to the manufacturer's instructions.

3. Record an IMS acquisition of the calibration solution with a single pass separation (use this for external IMS calibration).

NOTE: The ion source and the traveling wave (TW) parameters (static wave height and wave velocity) must be kept constant during all the acquisitions (calibration and acquisitions). If the user does not have prior knowledge of the optimal parameters for his sample, this step can be performed after step 3 of the protocol (for $[M+Li]^+$ adducts of neutral oligosaccharides, the representative results use a TW height of 16 V and TW velocity of 350 m/s, which give the best results).

1. From the **Fluidics** tab, select baffle position **Sample** and infuse the calibrant (see the **Table of Materials**) in the ion source (using the built-in fluidics system) through the 'Sample' probe at a flow rate of 10 µL/min.

2. Set up a single-pass IMS sequence. From the **Tune** page, put the instrument in **Mobility** mode, and open the **Cyclic Sequence Control** window. Select **Advanced** mode. From the **Cyclic Functions** tab of this new window, select **Add Bundle**, then **Single/Multipass**. Wait for a sequence of mobility events to appear in the **Sequence** tab of the same window.

NOTE: To activate the real-time display, the user must apply the instrument parameters: click on **Tune** in **TOF** mode or **Run** in **Mobility** mode. Before switching the instrument between **TOF** and **Mobility** modes, it is necessary to **Abort** any running acquisition (including the **Tune** page display). The relative abundance of ions may vary between **TOF** mode and **Mobility** mode because of changes in the ion transmission parameters.

3. Adapt the sequence so that all the calibrant ions make a single pass around the cyclic IMS racetrack. Do not change the **Inject** time or the **Eject and acquire** time; however, lower the **Separate** time to 1 ms (in the **Sequence** tab). If some ions of the calibration mixture do not fit in the displayed arrival time window, change the synchronization of the IMS with the pusher of the orthogonal acceleration TOF analyzer by increasing the number of **Pushes Per Bin** in the **ADC Settings** tab.

NOTE: The times in the control sequence only control the multifunction array for ion gating. As long as the ions are engaged in their first (or n^{th}) pass around the racetrack, they will finish said pass even if the direction of the TW has changed in the array in the meantime. Lowering the separation time to

1 ms means that the array will switch to ejection mode after 1 ms. This ensures that the faster ions will not have enough time to pass through the array and engage in a second pass before the slower ions finish their first pass. Therefore, all ions will be subjected to the same number of passes (i.e., one pass), which is necessary to perform IMS calibration.

4. Record a 2-min acquisition. In the **Cyclic Sequence Control** window, click on **Acquire** to open the **Acquisition Settings** popup window. Input the **Filename**, **Description**, and **Length of Acquisition (mins)** and click **Save**.

4. Record another 2 min acquisition of the calibration solution under the same conditions as step 2.3 (use this to check the quality of the CCS calibration). In the **Cyclic Sequence Control** window, click on **Acquire** to open the **Acquisition Settings** popup window. Input the **Filename**, **Description**, and **Length of Acquisition (mins)** and click **Save**.

5. Thoroughly wash the fluidics system with 50:50 H_2O/ACN to avoid crystallization of the calibrant in the peek tubing.

3. IMS/IMS-MS acquisition

1. Using a syringe pump, infuse the (lithium-doped) sample at 10 µg/mL through the sample probe at a flow rate of 10 µL/min.

2. Switch the instrument to **TOF** mode (from the **MS tune** page) to check the stability of the signal. Record a full MS acquisition (1 min) of the sample, which will be useful to check the isotopic pattern and the presence of potential contaminants.

NOTE: Because the sample concentration is chosen to obtain a good ion signal for the fragments, a TOF saturation may be observed at this step. TOF saturation can be identified using the following artifacts: (i) an artificially increased MS resolution, (ii) a change in isotopic ratios, and (iii) a multitude of low-abundance peaks in-between isotopes. Use the **DRE Lens** (Dynamic Range enhancement, **Quad/MS Profile/DRE** tab of the main **Tune** page) to attenuate the transmission of ions and discard the saturation in **TOF** mode (**Figure 2A,B**).

3. Put the instrument in **MSMS** mode (**Quad/MS Profile** tab of the main **Tune** page) and select the mass of the targeted ion in the **MSMS Mass** field for isolation in the quadrupole (in the example: *m/z* of 685.2, corresponding to the $[M+Li]^+$ ionic species of the arabinoxylan pentasaccharide). Record a 1 min acquisition to check the precursor isolation when processing the data.

NOTE: Lithium adducts have an isotope at -1 Da of the monoisotopic peak, which needs to be removed from the MS/MS selection window so that it will not interfere with the processing steps. It can be removed by narrowing the selection range using the **LM Resolution** and **HM Resolution** parameters in the **Quad/MS Profile** tab (**Figure 2C**).

4. Set up a "slicing" IMS sequence to perform a mobility-based selection of the isomer of interest.

1. Switch the instrument to **Mobility** mode (see step 2.3.2). In the **Cyclic Sequence Control** window, from the **Cyclic Functions** Tab, select **Add bundle** and then **Slicing**. Wait for a complex sequence of mobility events to appear in the **Sequence** tab (**Figure 3**).

NOTE: It is possible to visualize each step of the IMS/IMS process: click on the **Eject and Acquire**

event in the **Sequence** tab. Once highlighted in red, move it to the appropriate position within the sequence using the **Up** and **Down** buttons.

2. Position the **Eject and Acquire** event right after the first **Separate** event (i.e., move it at row 3 instead of row 8 in the sequence as displayed in **Figure 3**) and then click **Run**. Look for the results of the initial separation to be displayed in real time. Increase the duration of the first **Separate** event for a multipass separation by changing the time value for this event in the sequence until the resolution of the IMS peaks is satisfactory. Record a 1 min acquisition for reference.

   NOTE: Take note of the **ADC Start Delay** value in the **ADC Setup** tab: it will be useful to check the quality of the isolation.

3. Click **Pause**. Note that the results of the initial separation are displayed, although modifications in the control sequence will not be applied until the user clicks **Run** again. Position the **Eject and Acquire** event below the **Eject**, **Eject to Pre-Store**, and **Hold and Eject** events. Adjust the duration of the events so that the targeted peak is in the **Eject to Pre-Store** region, and any other ion is either in the **Eject** or **Hold and Eject** region.

   NOTE: The duration of these three events compared to the arrival time distributions (ATDs) can be visualized using the color-coded bar below the mobility spectrum in the **Mobilogram** tab (**Figure 3**).

4. Position the **Eject and Acquire** event at the end of the sequence, below the **Reinject from Pre-Store** and the second **Separate** events. Click **Run** to display the selected population.

   NOTE: Because the selected population has left the IMS cell, all previous separation has been lost, and it is back to a single-pass separation (which is desired).

5. Check the quality of the isolation. To verify that only the peak of interest has been selected, perform the same separation after reinjection as before reinjection (i.e., same **Separate** time) as shown in **Figure 4**. Record a 1 min acquisition for reference.

   NOTE: The users are encouraged to check the ejected population; the **Eject to Pre-Store** time window should be baseline level (**Figure 4B**). To check this, put the **ADC Start Delay** in **Manual** mode in the **ADC Settings** tab, and enter the delay time noted in step 3.4.2. Record a 1 min acquisition for reference.

6. In the **Sequence** tab, in the column next to the user-defined event times (the **Time Abs** column, highlighted in red), look for the summed times of all events. Take note of the **Time Abs** found on the line of the **Reinject from Pre-Store** event for performing the CCS calibration.

5. Fragment the targeted peak between the two rounds of IMS. Change the voltages of the reinjection step to increase the kinetic energy of the ions, and fragment them upon collision with the ion mobility gas.

   1. Set the duration of the **Separate** event directly preceding the **Eject and Acquire** to 1 ms (see explanation in step 2.3.3).

   2. On the **Reinject from Pre-Store** line, tick the **Enable Activation** box, and optimize the fragmentation with the built-in control. If the

spectrum is satisfactory (e.g., the base peak is a fragment), proceed directly to step 3.5.4.

NOTE: When enabling activation, three voltages on the line will turn grey: these are the voltages that the user needs to change if manual optimization of the voltages is required (see the next step). These three voltages (**Pre-Array Gradient, Pre-Array Bias,** and **Array Offset**) form the gradient used to activate the ions. The kinetic energy of the ions will increase with the slope between the **Pre-Array Bias** and **Array Offset** (see **Figure 5**). The default values of the Gradient → Bias → Offset values are: without activation 85 → 70 → 45 V; maximum activation of the built-in function 185 → 170 → -5 V (+150 V). After fragmentation, do not forget to readjust the ion transmission using the **DRE lens** (decrease the attenuation of the signal) (see step 3.2.).

3. If the fragmentation is not satisfactory with the built-in control, uncheck the **Enable Activation** box and proceed to manually optimize the reinjection voltages. Increase the **Pre-Array Gradient** voltage (the **Pre-Array Bias** voltage must always be kept 15 V below the **Pre-Array Gradient**), and lower the **Array Offset** voltage (which can be set as negative) until the results are satisfactory.

   NOTE: When manually tuning the voltages of the multifunction array, the user can switch from the 'Mobilogram' view to interactive schematics of the voltages applied in the multifunction array (**PE diagram**) to better visualize the voltage settings (**Figure 5A**).

4. Record a 2 min acquisition. In the acquisition pop-up window, tick the **Retain Drift Time** option to generate a file containing only the arrival times vs *m/z* (the acquisition time used for chromatographic analyses-the retention time-is removed from the file). Note that this file is labeled ***_dt.RAW**.

   NOTE: If the user forgets to check the **Retain Drift Time** option, it is still possible to extract the IMS dimension using the Driftscope 2.9 software (**File | Export to MassLynx | Retain Drift Time**).

6. Turn the instrument back to **TOF** mode in the main **Tune** page, and thoroughly rinse the system with 50:50 MeOH/H_2O before proceeding with the next sample.

4. IMS/IMS-MS processing with MZmine 2[24]

NOTE: MZmine 2 is available from the URL given in the **Table of Materials**. The use of MZmine 2.51 is recommended. At the time of preparation of this manuscript, the later versions cannot open RAW files from Cyclic IMS instruments because of a change in the import function.

1. Import the raw file(s) containing only the IMS and *m/z* dimensions (***_dt.RAW**) using **Raw data methods | Raw data import**.

   NOTE: Raw files will appear on the left side of the main MZmine window. Do not import the original ***.RAW** files that still contain the retention time dimension. MZmine does not distinguish retention time from IMS arrival time, and the data points of both dimensions would overlap.

2. Optimize the workflow parameters on a representative file by selecting it in the **Raw data files** list.

   1. Evaluate the noise level in the data. Right-click on the file in the **Raw data files** list, select **Show TIC** and display the base peak "chromatogram" (BPC). Double-click on the smallest peak observable by eye to display its mass spectrum. Consider the noise level in the data to be around that of the second

isotope of the base peak in this spectrum, and use this same value for all the intensity thresholds in the following processing steps.

NOTE: The data were acquired using quadrupole isolation and are thus considered by MZmine as MS/MS. Throughout the entire MZmine processing, be sure to work at an **MS level = 2**.

2. Perform the mass detection using **Raw data methods | Feature detection | Mass detection**. For data acquired in profile mode, use the **Wavelet transform** algorithm. To set up the parameters of the algorithms in MZmine, click on the **[...]** button next to the algorithm and use the **Show preview** option to visualize the data while optimizing the parameters.

   NOTE: At this stage, the peaks selected by the algorithm will appear in red in the preview window. When using the **wavelet transform** algorithm on proprietary RAW files, MZmine will sometimes mistake the profile data points for centroided peaks. The software will display a message stating that the user is running a profile algorithm on centroided spectra: ignore this message and click **OK**.

3. Reconstruct the extracted ion mobility spectra (EIM) for each fragment mass using **Raw data methods | Feature detection | ADAP Chromatogram builder** on the 'masses' mass list generated by the previous step. As the **m/z tolerance** input at this stage is a scan-to-scan tolerance, be sure to leave it at least 3-4 times higher than the overall expected accuracy.

4. As the previous step does not have a preview option, check the quality of the peak picking directly using the **Feature list** that appeared on the right panel of the main MZmine window. Open the **Feature list**, select all rows, right-click, and select **Show/XIC (dialog)**. Click **All** to display all ions on the mobility spectrum. Inspect the picked peaks that appear in color to ensure that there are no obvious missed peaks.

5. Deconvolve the EIMs to split the *m/z* that contain different peaks in multiple features. Use **Feature list methods | Feature detection | Chromatogram deconvolution**, and choose the **Wavelets (ADAP)** algorithm. Optimize the algorithm for the data using the **Show preview** option and the following key parameters: **S/N threshold, coefficient/area threshold**, and **RT wavelet range**.

   NOTE: Checking the aspect of the deconvolved spectrum is recommended. Use the chromatogram visualization tool, as described in step 4.2.4. The deconvolved peaks will appear in color, and peaks of the same mass should be split, as presented in **Figure 6A**.

6. Deisotope the deconvolved EIMs using **Feature list methods | Isotopes | Isotopic peaks grouper**. Use the expected accuracy of the instrument for the *m/z* tolerance value, and set the arrival time tolerance to 0.1 ms (displayed in MZmine as **Retention time tolerance 0.1 min**), as isotopes are not resolved during the IMS separation. Check the feature list: if any isotopes remain, increase the tolerance values.

   NOTE: Although the deisotoping can theoretically be performed at any moment of the feature list processing, it is important to do it last so that the charge values can be exported (the algorithms used for the other steps will sometimes remove the charge state information).

3. If processing multiple IMS/IMS-MS spectra, repeat the processing with these optimized parameters. Keep the same parameters for all spectra.

4. In the case of multiple spectra, group them in a single table to export them; if not, skip directly to step 4.5. To group the spectra, use **Feature list methods | Alignment | Join aligner**. Because the objective is not to actually align the peaks, use restrictive tolerance values for both *m/z* and arrival time. Give the same weight to both dimensions.

5. Export the final feature list to a ***.csv** file. Use **Feature list methods | Export/Import | Export to CSV file** and export the following values: **Export row m/z, Export row retention time** (the actual IMS arrival time), **Peak m/z**, and **Peak height**. Use a comma as a field separator.

5. $^{TW}CCS_{N2}$ of the centroided IMS/IMS spectra

NOTE: In this protocol, a logarithmic fit calibration[25,26] will be used, which tends to give better results than linear calibration and is easy to implement in a spreadsheet or an in-house processing script. An in-house script (written in R) is available at the URL given in the **Table of Materials**.

1. Pick the reference arrival time values from the calibrant acquisition (see step 2.3). Do this manually using the constructor software (see the **Table of Materials**) to check the aspect of all IMS calibrant peaks.
    1. In the **Chromatogram** window, open the ***_dt.RAW** file corresponding to the calibrant.
    2. For each calibration point, generate the EIM using the **Display | Mass** option.
    3. Check the profile of the EIMs. If some are poorly defined, smooth them using the **Process | Smooth** option (as the best results are typically obtained with the Savitzky-Golay algorithm, smooth 2 times over 3 bins). Report the apex values in a spreadsheet.

    NOTE: Because the reference points are generally acquired using low-resolution DTIMS devices, some multimodal distributions may appear in Cyclic IMS depending on the calibrants. Remove any peak presenting such a distribution from the calibration list.

2. Calculate the logarithmic fit parameters from the calibrants.
    1. For all calibration points, calculate the following.
        1. Calculate the drift time using Eq (**1**):

        $$t_d = t_A - t_{inj} \quad (\mathbf{1})$$

        with t_d the drift time, t_A the measured arrival time, and t_{inj} the time of injection in the IMS cell (all in ms).

        NOTE: For small molecules, such as oligosaccharide fragments, the dead time (flight time between the exit from the IMS cell and the detector) variation between different masses is within the error range of the CCS calibration and can be ignored.

        2. Calculate the neutral mass of the ions using Eq (**2**):

        $$m_{neutral} = (m/z) \times |z| - m_{ion} \quad (\mathbf{2})$$

        with z the charge state of the ion, and m_{ion} the mass of the counter-ion (in Da). Use exact masses to avoid introducing uncertainty. If there is an atom loss instead of a counter-ion, use negative m_{ion} values (e.g., for $[M-H]^-$ $m_{neutral}$

= *(m/z)* * |z| - *(- 1.007276)* = *(m/z)* * |z| + *1.007276*).

3. Calculate the CCS' parameter using Eq (**3**):

$$CCS' = \frac{CCS}{|z| \times \left(\frac{1}{m_{neutral}} + \frac{1}{m_{gas}}\right)^{0.5}} \quad (3)$$

with *CCS* the reference drift tube $^{DT}CCS_{N2}$ value (in nm^2), and m_{gas} the mass of the drift gas (in Da; ex. for nitrogen: m_{gas} = 28.01 Da).

4. Calculate the t_d' parameter using Eq (**4**):

$$t'_d = t_d - 0.001 \times d \times (m/z)^{0.5} \quad (4)$$

with *d* the detector start delay used experimentally to correct for dead time (typically ~1.5 ms).

5. Calculate logarithm of the above parameters: ln (*CCS'*) and ln (*t'd*)

2. Perform a linear regression to determine the R^2 coefficient and the *x* and *y* parameters of the logarithmic fit (with *x* the slope and *ln(y)* the intercept) using Eq (**5**):

$$\ln(CCS') = x \times \ln(t'_d) + \ln(y) \quad (5)$$

NOTE: The user can plot the ln(CCS') vs ln(t_d') values to visually check the results of the calibration, although this is optional.

3. Apply the calibration to the experimental data to calibrate the peaks picked by MZmine for every IMS/IMS spectrum exported to the *.csv file. For each point, calculate the following.

1. Calculate the drift time using Eq (**6**):

$$t_d = t_A - t_{seq} \quad (6)$$

with t_{seq} the time preceding the final IMS separation (the 'Time Abs' value noted in step 3.4.6).

NOTE: If calibrating multiple IMS/IMS spectra acquired with different sequences, carefully check the t_{seq} values.

2. Calculate the neutral mass of the ions using Eq (**7**):

$$m_{neutral} = (m/z) \times |z| - m_{ion} \quad (7)$$

3. Calculate the t_d' and t_d'' parameters using Eq (**8**) and Eq (**9**):

$$t'_d = t_d - 0.001 \times d \times (m/z)^{0.5} \quad (8)$$

$$t''_d = (t'_d)^x \times |z| \times \left(\frac{1}{m_{neutral}} + \frac{1}{m_{gas}}\right)^{0.5} \quad (9)$$

4. Calculate the final calibrated CCS values ($^{TW}CCS_{N2}$ in nm^2) using Eq (**10**):

$$^{TW}CCS_{N2} = t''_d \times y \quad (10)$$

NOTE: Although step 5.2.2. gives *ln(y)* as the intercept, *y* must be used to obtain the final CCS value. Do not forget to apply an exponential function.

4. Check the accuracy of the calibration by applying the calibration to the second acquisition of the calibration solution acquired in step 2.4.

NOTE: The calibration should yield results with an error of ~1-2%.

Representative Results

An arabinoxylan pentasaccharide, XA^2XX, was chosen as an example to illustrate this protocol. This compound is commercially available, but only as a mixture with another arabinoxylan pentasaccharide, XA^3XX (pure XA^3XX is also commercially available). The structures of XA^2XX and XA^3XX are given in **Supplemental Figure S1**. As the ratio of XA^2XX and XA^3XX in the commercial mixture is ~50:50, a solution at 20 µg/mL of the mixture was prepared to reach

an XA^2XX concentration of ~10 µg/mL in 50:50 MeOH/H_2O + 500 µM LiCl.

First, an MS analysis of the XA^2XX + XA^3XX mixture was performed using high-resolution MS. As the two compounds are isomers, a single peak was observed at $[M+Li]^+$ *m/z* 685.24. This MS peak was selected with the quadrupole and the selection window adjusted to remove the -1 Da lithium isotope, which could be mistaken as the monoisotopic peak by processing algorithms (**Figure 2**).

The $[M+Li]^+$ adducts of the pentasaccharides were then submitted to the first stage of IMS separation: after 3 passes around the cyclic IMS cell, 3 peaks were separated with arrival times of 83, 90, and 94 ms. This profile was compared to that of pure XA^3XX (infused at 10 µg/mL), showing that the peaks at 83 and 90 ms corresponded to XA^3XX, while the peak at 94 ms corresponded to XA^2XX (**Figure 4A**). The peak at 94 ms was selected for IMS/IMS analysis: the ions belonging to XA^3XX were ejected (**Figure 4B**), and the peak of interest was sent to the prearray store cell. A 3-pass separation was performed after reinjecting the ion without activation to ensure that only the XA^2XX peak remained after the selection (arriving at 199 ms in **Figure 4C**).

Then, the ion was fragmented upon reinjection from the prestore area, and a single-pass IMS separation was performed on all the fragments. Two different activations were tried: the maximum setting of the built-in prestore activation function was first tried (**Figure 5B,C**); however, the precursor remained the base peak of the spectrum. This is not desired because, for reference spectra, fragments below a certain intensity threshold would typically be removed. Thus, a manually defined **prearray gradient** → **pre-array bias** → **array offset** voltage gradient was chosen (**Figure 5D,E**).

The generated IMS/IMS-MS data were deconvolved with MZmine 2.51, using the arrival time and *m/z* dimensions (**Figure 6A**), to give IMS/IMS spectra containing only the mobility information of the fragments. The peaks above 0.2% relative intensity were exported for CCS calibration (the detailed MZmine parameters are given in **Supplemental Table S1**). The CCS calibration was performed using the calibration solution (R^2 = 0.995, mean absolute deviation of control = 1.63%, see **Supplemental Table S3**). This processing finally afforded a centroided, CCS-calibrated, IMS/IMS spectrum (**Figure 6B**).

Copyright © 2022 JoVE Journal of Visualized Experiments

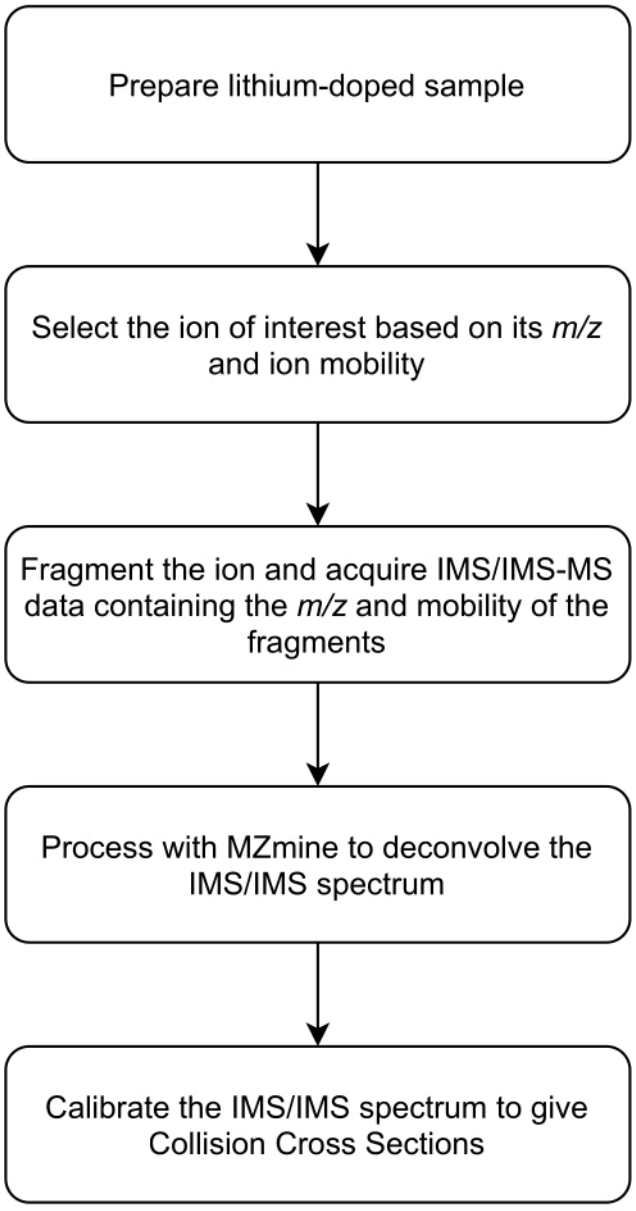


Figure 1: Overview of the IMS/IMS data generation process. Abbreviations: IMS = ion mobility spectrometry; IMS/IMS = tandem IMS. Please click here to view a larger version of this figure.

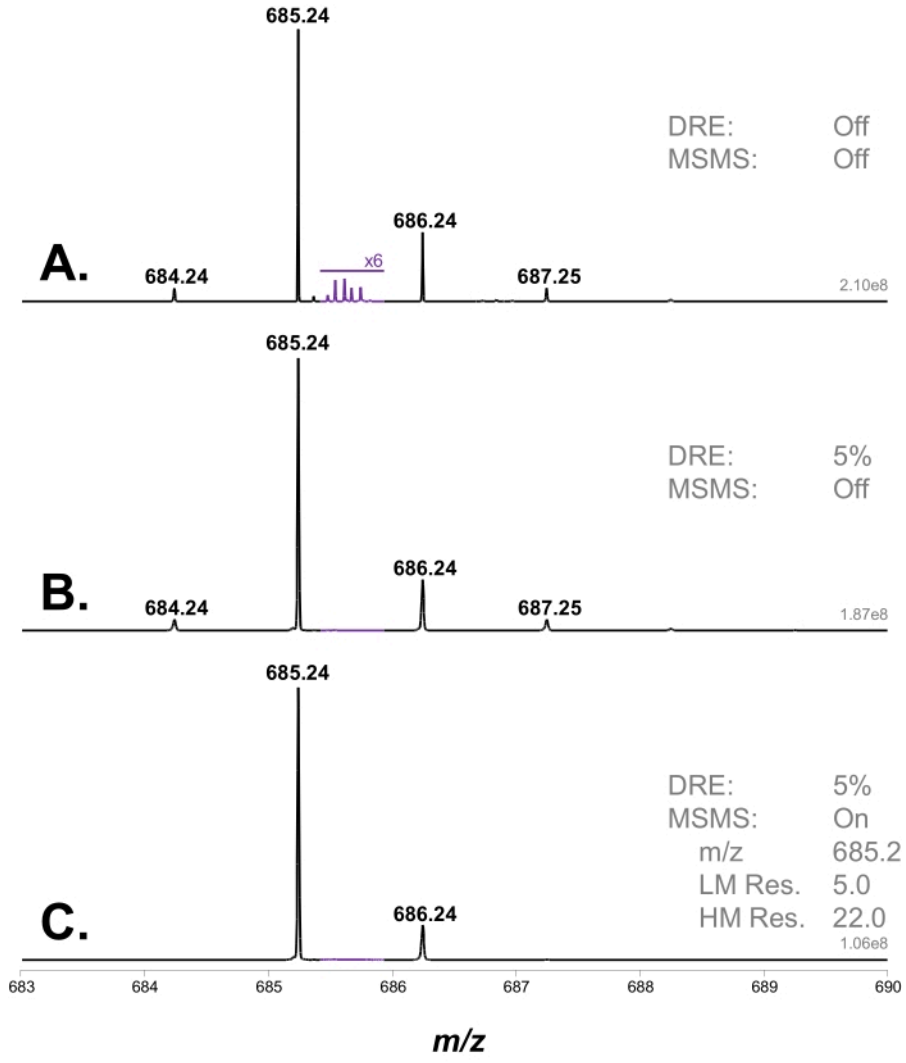


Figure 2: Isotopic pattern of an XA^{3}XX + XA^{2}XX arabinoxylan pentasaccharide mixture. (**A**) Saturated signal without DRE; (**B**) signal corrected using DRE with a 5% ion transmission (i.e., 95% attenuation); and (**C**) profile after quadrupole selection to remove the -1 Da peak corresponding to a lithium isotope. In purple: the region where artifact peaks can appear due to saturation is magnified 6 times. Abbreviations: DRE = dynamic range enhancement; MS = mass spectrometry; MSMS = tandem MS; LM Res = low-mass resolution; HM Res = high-mass resolution. Please click here to view a larger version of this figure.

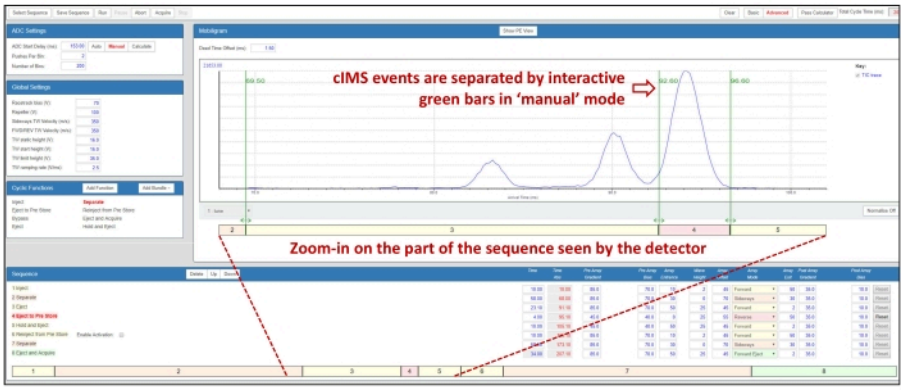


Figure 3: Overview of the Cyclic IMS control window, in which the user defines the IMS/IMS sequence. The sequence displayed shows how to check the quality of the isolation in IMS/IMS, with a selection of XA^2XX after 3 passes (the spectrum displayed corresponds to the setting of the selection events after the first-stage separation). The sequence consists in running a first 3-pass IMS separation over 58 ms, then ejecting the two faster isoforms from the IMS cell (segment 3), ejecting the slower isoform (ATD between 92 and 96 ms) in the prestore (segment 4), reinjecting it in the IMS cell without activation (segment 6), allowing the ions to undergo a further 3-pass (58-ms) separation (segment 7), then ejecting ions from the IMS cell, and acquiring data (segment 8). Abbreviations: IMS = ion mobility spectrometry; cIMS = cyclic IMS; IMS/IMS = tandem IMS; ADC = analog-to-digital converter; TW = traveling wave; PE = potential energy; ATD = arrival time distribution.

Please click here to view a larger version of this figure.

Copyright © 2022 JoVE Journal of Visualized Experiments

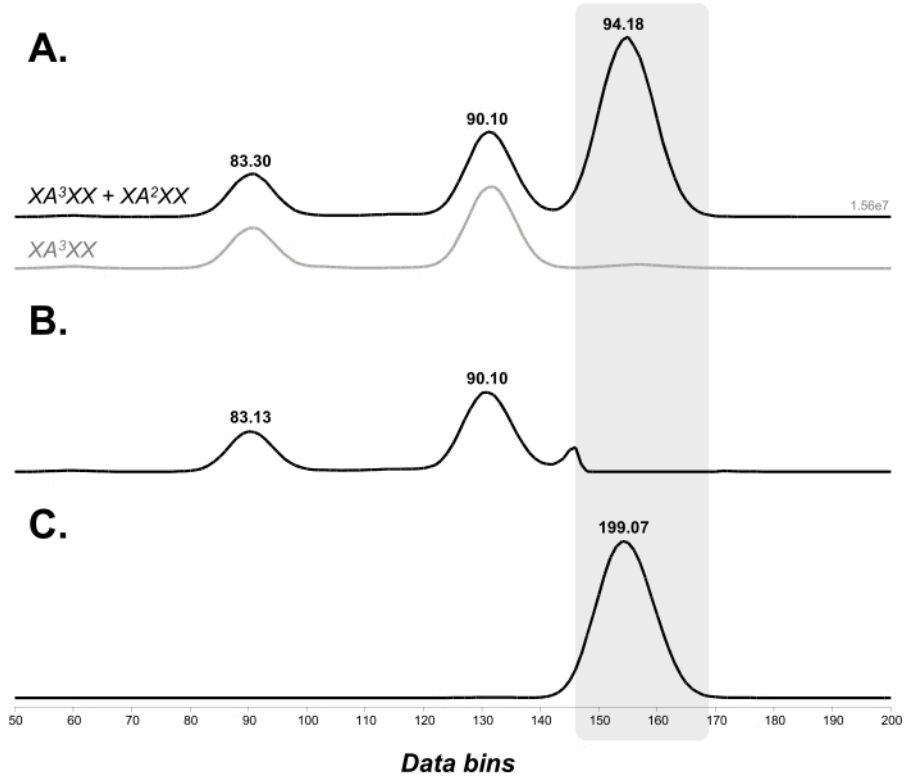


Figure 4: Selection of XA^2XX from the mixture of XA^2XX and XA^3XX. (**A**) Separation of the arabinoxylan pentasaccharides, XA^3XX and XA^2XX, after 3 passes (corresponding to a separation time set at 58 ms) around the Cyclic IMS cell. (**B**) Fraction ejected directly after the first stage of IMS separation. (**C**) Fraction selected for IMS/IMS on which another 3-pass separation was performed after reinjection. The XA^2XX peak of interest is highlighted in gray. The ion mobility spectra are shown in data bins and annotated with their arrival time (ms). Abbreviations: IMS = ion mobility spectrometry; IMS/IMS = tandem IMS. Please click here to view a larger version of this figure.

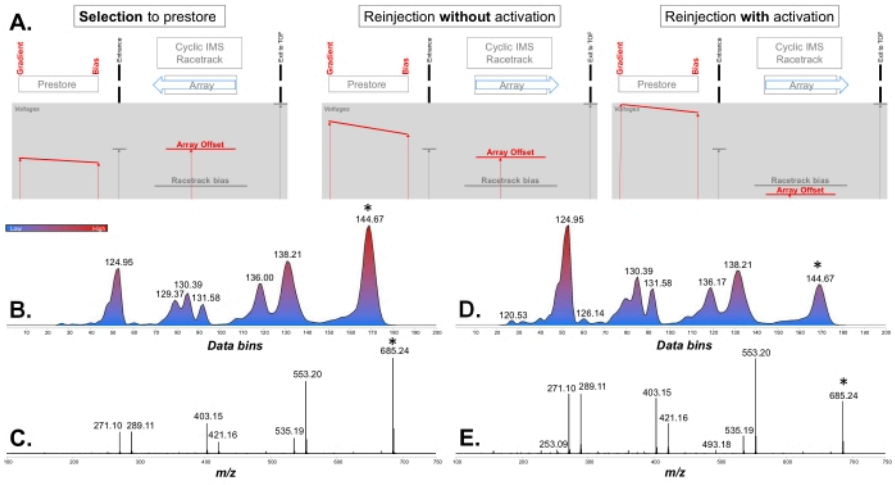


Figure 5: Principles of collision-induced dissociation using the prearray store area. (**A**) Schematics of the multifunction array region detailing key voltages (in red) used for the selection, the reinjection, and the activation during IMS/IMS experiments. The blue arrows show the direction of the traveling wave in the multifunction array. (**B, C**) IMS/IMS and MS/MS spectra obtained for XA^2XX using the built-in prestore activation function (+150 V). The color bar represents the ion intensity scale (blue = low; red = high). (**D, E**) IMS/IMS and MS/MS spectra obtained for XA^2XX with manual optimization of the voltages (prearray gradient 195 V, prearray bias 180 V, array offset -10 V). Precursor ions are indicated by asterisks on the spectra. The ion mobility spectra are shown in data bins and annotated with their arrival time (ms). Abbreviations: IMS = ion mobility spectrometry; IMS/IMS = tandem IMS; TOF = time-of-flight. Please click here to view a larger version of this figure.

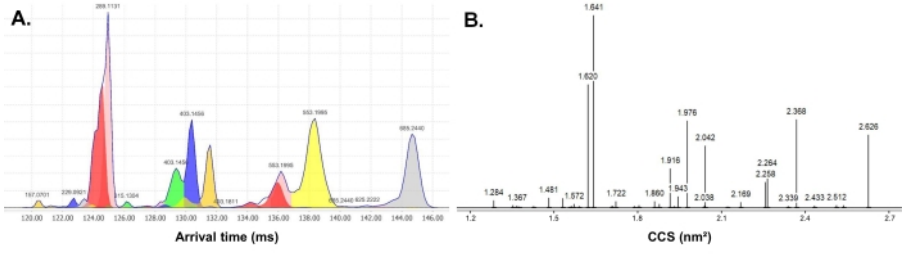


Figure 6: Illustration of the processing steps. Results of (**A**) the MZmine peak picking and (**B**) the CCS calibration of arabinoxylan pentasaccharide XA^2XX. **A** shows the mass deconvolution of the IMS/IMS spectrum through a color code. **B** shows the final IMS/IMS spectrum after centroiding and CCS calibration. Abbreviations: IMS = ion mobility spectrometry; IMS/IMS = tandem IMS; CCS = collision cross section. Please click here to view a larger version of this figure.

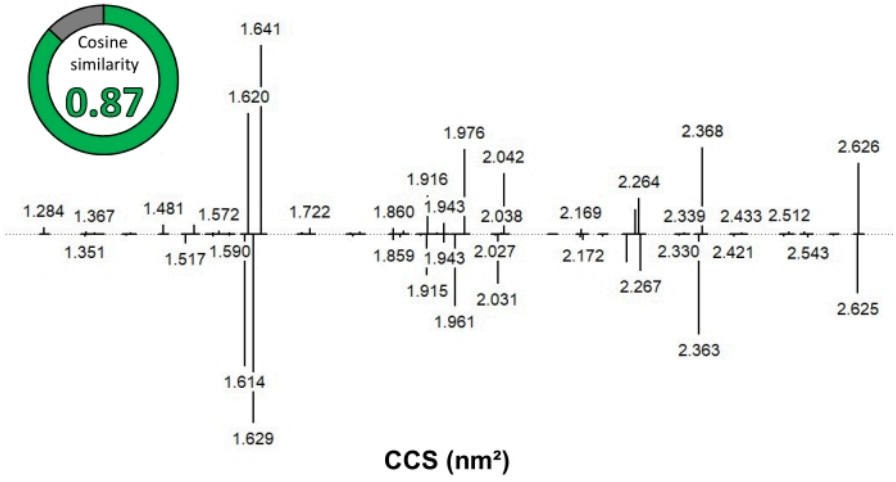


Figure 7: A comparison of two IMS/IMS spectra of XA^{2}XX illustrates the reproducibility of the method. The final calibrated spectrum from this paper (top) is compared to the spectrum from the work by Ollivier et al.[21] (bottom, flipped). Abbreviations: IMS = ion mobility spectrometry; IMS/IMS = tandem IMS; CCS = collision cross section. Please click here to view a larger version of this figure.

Supplemental Figure S1: Structures of the XA^{2}XX and XA^{3}XX arabinoxylan pentasaccharides. Please click here to download this File.

Supplemental Figure S2: Evaluation of the interday repeatability using XA^{2}XX. IMS/IMS acquisitions were repeated at Day 1 (top) and Day 95 (bottom). Abbreviations: IMS = ion mobility spectrometry; IMS/IMS = tandem IMS. Please click here to download this File.

Supplemental Table S1: Detailed MZmine parameters. Please click here to download this Table.

Supplemental Table S2: Instrument parameters changed to evaluate the reproducibility. Abbreviation: ESI = electrospray ionization. Please click here to download this Table.

Supplemental Table S3: Control of the CCS calibration using a second acquisition of calibration solution. Please click here to download this Table.

Discussion

The SELECT SERIES Cyclic IMS is a powerful tool that allows selecting a defined ion population—of a given *m/z* and ion mobility—without the need for upstream chromatographic separation. The instrument affords the possibility of generating a bidimensional fragmentation map of this ion population, from which both MS/MS and IMS/IMS spectra can be extracted. However, the user must note several critical points that require attention during the experimental process.

First, the user should carefully check the MS isolation window for the presence of possible isobaric contaminants. Indeed, the isolation window of a quadrupole is relatively wide, and

knowing ions of a slightly different *m/z* that may be coselected in the quadrupole will help the user properly assign the peak of interest in ion mobility.

Second, when performing the initial separation, the user must ensure that all ions undergo the same number of passes around the Cyclic ion mobility cell. This is an important and tricky aspect of ion mobility separation in a cyclic device. An erroneous evaluation of the number of passes for a given ion may lead to an improper identification and interpretation of the peaks. Controlling the number of passes of different ion populations can be tricky owing to the relatively short single-pass length (~1 m), and species with very different mobilities can quickly overlap.

In particular, a peak can split between two different passes if the array switches direction when this ion population passes through (this is relatively easy to identify: the split peak will appear sharper with a population right at the beginning of the **Eject and Acquire** event). To properly set the number of passes, the user should start with a short separation time (1-5 ms) that will give the 1-pass profile. Then, the user should gradually increase the separation time until the entire population has moved to higher arrival times, which will give the 2-pass profile. The 2-pass profile should look similar to the 1-pass profile but with better-resolved peaks. The time a given ion population takes to make one pass around the cyclic cell is a constant that the user can use to calculate the number of passes as a function of the separation time. For instance, if there is a 10 ms difference between the first and second pass, there will also be a 10 ms difference between the second and the third.

Third, during the IMS selection stage, the user should carefully check the quality of the isolation, as demonstrated in **Figure 4**. It is especially important to check the reinjected profile because if the TW height and velocity settings are too low, the ejection of the other populations might not be complete. Advanced users can correct this by adjusting the **Driftcell RF** radiofrequency voltage in the **RF** tab of the **Tune** page.

Fourth, the user should be careful in generating the fragmentation spectrum and, notably, in selecting the appropriate collision energy, especially if the voltages are tuned manually. Indeed, excessively lowering the **Array Offset** voltage can negatively impact the overall ion intensity by hindering the reinjection. In addition, the precursor and fragments might span over a wide range of mobilities. Thus, they will rapidly undergo a different number of passes if the final separation time is high, so it is important to keep a 1-ms **Separate** event as explained in protocol step 2.3.3. This is a major limitation since the single-pass length is relatively short, limiting the single-pass resolving power to ~100 for oligosaccharides[27]. In this respect, an increased path length in a single pass would be beneficial (i.e., the TWIMS-based Structures for Lossless Ion Manipulation or SLIM, with a path length of 13 m[28]). The SLIM setup was launched commercially very recently[29].

Finally, the user should be careful in defining the final acquisition mobility range using the **Pushes per bin** command, particularly if working on multiply charged ions. Ion mobility is indeed a function of the charge[12], and, for example, singly-charged fragments generated from a doubly-charged precursor are likely to be slower than the precursor (although they are smaller compounds).

A major limitation of using only MS and IMS separations to select the precursor (and not, for instance, an upstream step of chromatography) is that a given *m/z* can yield multiple peaks in IMS and that multiple peaks can come from the

same compound. This is illustrated by the distributions of *m/z* 685.2, for both the XA^2XX+XA^3XX mixture and pure XA^3XX, in **Figure 4A**. Multimodal IMS distributions of a single *m/z* result from different gas-phase conformations. For species analyzed as cation adducts (in positive mode), the different conformations possibly arise from differences in coordination with the counter-ion[30,31,32].

For oligosaccharides, they can also arise from the separation of reducing-end anomers, although separating reducing-end anomers typically requires higher IMS resolving power than what is used here[33,34,35]. In the present case, the multimodal IMS distribution in **Figure 4A** results partly from the individual contributions of XA^2XX and XA^3XX. It is, however, notable that XA^3XX yields two peaks, which are likely cation-coordination conformers. It was easy to identify which peak corresponded to XA^2XX (i.e., the species of interest) because pure XA^3XX is available commercially, and its mobility profile could be recorded separately. To work on complex mixtures such as biological media, it may be important to consider adding a chromatographic separation stage.

Two points must be noted regarding the processing workflow used to obtain mass-deconvolved IMS/IMS spectra. First, in this protocol, it is proposed to use MZmine 2[24] to deconvolve the IMS/IMS spectrum using the MS dimension and, notably, use the ADAP algorithm[36] to split the EIMs into different peaks. Although it gives fairly good results, as illustrated in **Figure 6**, the ADAP algorithm was designed for chromatographic analyses and thus accounts for asymmetry factors inherent to liquid phase chromatography, such as peak tailing. Therefore, the ADAP algorithm might result in not identifying some of the features when applied to IMS peaks (e.g., shoulders). In essence, IMS data are simpler than chromatographic data: because there is no chemical interaction of the compounds with a column, IMS data acquired under appropriate conditions (i.e., without saturating the IMS cell) are expected to follow gaussian distributions[37,38]. Ideally, the ADAP deconvolution step would be best replaced by a gaussian fitting function, such as that used by software destined for IMS like CUISuite 2[39]. However, as it stands, gaussian deconvolution was not directly adapted to the complete chain of treatment described in this protocol. Therefore, using the free, open-source software MZmine appeared to be a good compromise for end-users.

The second part of the processing that warrants discussion is the CCS calibration. This protocol proposes using a logarithmic fit calibration[25,26] and a commercial calibrant mixture from the same provider as that of the spectrometer (see the **Table of Materials**). This procedure is the most straightforward to implement in the lab. Regarding the choice of the calibrant mixture, the user should consider that, as mentioned in several studies, the accuracy of the CCS calibration is improved when using calibrants of the same molecular class and charge state as the analyte[26,40,41]. The error introduced when calibrating with relatively similar types of compounds (e.g., carbohydrates vs. peptides) is moderate. However, it is recommended not to use very different ions such as, for example, using salt clusters as calibrants when measuring carbohydrates[41]. Regarding the choice of the calibration method, Richardson et al.[42] recently reported a new calibration method that takes into account the physics of TWIMS to improve the accuracy of the calibration (with a provided software). However, the approach requires the evaluation of highly specific parameters through the analysis of various types of compounds—ranging from metabolites to native proteins. Because no mixture of such a variety of

Copyright © 2022 JoVE Journal of Visualized Experiments

compounds can be found commercially, this method was not implemented in the present protocol.

Finally, to evaluate the reproducibility of the method, we evaluated the interday reproducibility by repeating the IMS/IMS-MS acquisition at day 1 and day 95 (**Supplemental Figure S2**). The experiment showed that IMS/IMS-MS data are highly reproducible, with no IMS peak shifting by more than 0.2 ms over this extended period. The IMS/IMS spectrum generated in this work was further compared to another spectrum of XA^2XX acquired under different conditions for previous work on ion mobility-molecular networking[21]. Some instrumental parameters that can impact the ion structure and ion mobility profile[13] were deliberately changed—the source parameters and the activation voltage gradient (a comparison of the varying instrumental conditions is given in **Supplemental Table S2**). Then, the two spectra were compared using the cosine similarity score—which is popular for the comparison of MS/MS spectra in metabolomics—on the GNPS platform[5] (CCS tolerance for matching fragments = 0.015 nm^2).

The comparison showed a cosine similarity score of 0.87 (**Figure 7**), which can be considered high with regard to the important instrumental variations applied. This leads to the idea that IMS/IMS spectral libraries could be used to dereplicate glycans in complex mixtures with a high level of confidence, which would not be the case with MS/MS spectra. Note that although the current approach only uses the CCS dimension of the fragmentation spectrum, the IMS/IMS-MS data also contains MS information, which is not redundant with the CCS. To optimize the dereplicative power of IMS/IMS, a bidimensional scoring system must be developed.

Disclosures

The authors have no conflict of interest to disclose.

Acknowledgments

S.O. is thankful to the French National Research Agency for funding his Ph.D. (grant ANR-18-CE29-0006).

References

1. Allard, P. -M. et al. Integration of molecular networking and in-silico MS/MS fragmentation for natural products dereplication. *Analytical Chemistry.* **88** (6), 3317-3323 (2016).
2. Wang, M. et al. Mass spectrometry searches using MASST. *Nature Biotechnology.* **38** (1), 23-26 (2020).
3. David, M., Fertin, G., Rogniaux, H., Tessier, D. SpecOMS: a full open modification search method performing all-to-all spectra comparisons within minutes. *Journal of Proteome Research.* **16** (8), 3030-3038 (2017).
4. Dührkop, K. et al. SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nature Methods.* **16** (4), 299-302 (2019).
5. Wang, M. et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology.* **34** (8), 828-837 (2016).
6. Nothias, L.-F. et al. Feature-based molecular networking in the GNPS analysis environment. *Nature Methods.* **17** (9), 905-908 (2020).
7. Gray, C. J. et al. Advancing solutions to the Carbohydrate Sequencing Challenge. *Journal of the*

American Chemical Society. **141** (37), 14463-14479 (2019).

8. Ropartz, D. et al. Online coupling of high-resolution chromatography with extreme UV photon activation tandem mass spectrometry: Application to the structural investigation of complex glycans by dissociative photoionization. *Analytica Chimica Acta.* **933**, 1-9 (2016).
9. Wolff, J. J. et al. Negative electron transfer dissociation of glycosaminoglycans. *Analytical Chemistry.* **82** (9), 3460-3466 (2010).
10. Ropartz, D. et al. Charge transfer dissociation of complex oligosaccharides: comparison with collision-induced dissociation and extreme ultraviolet dissociative photoionization. *Journal of the American Society for Mass Spectrometry.* **27** (10), 1614-1619 (2016).
11. Morelle, W. et al. Fragmentation characteristics of permethylated oligosaccharides using a matrix-assisted laser desorption/ionization two-stage time-of-flight (TOF/TOF) tandem mass spectrometer. *Rapid Communications in Mass Spectrometry.* **18** (22), 2637-2649 (2004).
12. Gabelica, V., Marklund, E. Fundamentals of ion mobility spectrometry. *Current Opinion in Chemical Biology.* **42**, 51-59 (2018).
13. Gabelica, V. et al. Recommendations for reporting ion mobility mass spectrometry measurements. *Mass Spectrometry Reviews.* **38** (3), 291-320 (2019).
14. Hernandez-Mesa, M. et al. Interlaboratory and interplatform study of steroids collision cross section by traveling wave ion mobility spectrometry. *Analytical Chemistry.* **92** (7), 5013-5022 (2020).
15. Koeniger, S. L. et al. An IMS−IMS analogue of MS−MS. *Analytical Chemistry.* **78** (12), 4161-4174 (2006).
16. Merenbloom, S. I., Koeniger, S. L., Valentine, S. J., Plasencia, M. D., Clemmer, D. E. IMS−IMS and IMS−IMS −IMS/MS for separating peptide and protein fragment ions. *Analytical Chemistry.* **78** (8), 2802-2809 (2006).
17. Eldrid, C., Thalassinos, K. Developments in tandem ion mobility mass spectrometry. *Biochemical Society Transactions.* **48** (6), 2457-2466 (2020).
18. Giles, K. et al. A cyclic ion mobility-mass spectrometry system. *Analytical Chemistry.* **91** (13), 8564-8573 (2019).
19. Merenbloom, S. I., Glaskin, R. S., Henson, Z. B., Clemmer, D. E. High-resolution ion cyclotron mobility spectrometry. *Analytical Chemistry.* **81** (4), 1482-1487 (2009).
20. Ollivier, S. et al. Anomeric retention of carbohydrates in multistage cyclic ion mobility (IMSn): de novo structural elucidation of enzymatically produced mannosides. *Analytical Chemistry.* **93** (15), 6254-6261 (2021).
21. Ollivier, S., Fanuel, M., Rogniaux, H., Ropartz, D. Molecular networking of high-resolution tandem ion mobility spectra: a structurally relevant way of organizing data in glycomics? *Analytical Chemistry.* **93** (31), 10871-10878 (2021).
22. Aron, A. T. et al. Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nature Protocols.* **15** (6), 1954-1991 (2020).
23. McKenna, K. R., Li, L., Krishnamurthy, R., Liotta, C. L., Fernández, F. M. Organic acid shift reagents for the discrimination of carbohydrate isobars by ion mobility-mass spectrometry. *The Analyst.* **145** (24), 8008-8015 (2021).

Copyright © 2022 JoVE Journal of Visualized Experiments

24. Pluskal, T., Castillo, S., Villar-Briones, A., Orešič, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics.* **11**, 395 (2010).

25. Ruotolo, B. T., Benesch, J. L. P., Sandercock, A. M., Hyung, S. -J., Robinson, C. V. Ion mobility-mass spectrometry analysis of large protein complexes. *Nature Protocols.* **3** (7), 1139-1152 (2008).

26. Bush, M. F., Hall, Z., Giles, K., Hoyes, J., Robinson, C. V., Ruotolo, B. T. Collision cross sections of proteins and their complexes: a calibration framework and database for gas-phase structural biology. *Analytical Chemistry.* **82** (22), 9557-9565 (2010).

27. Ropartz, D. et al. Structure determination of large isomeric oligosaccharides of natural origin through multipass and multistage cyclic traveling-wave ion mobility mass spectrometry. *Analytical Chemistry.* **91** (18), 12030-12037 (2019).

28. Tolmachev, A. V. et al. Characterization of ion dynamics in structures for lossless ion manipulations. *Analytical Chemistry.* **86** (18), 9162-9168 (2014).

29. Arndt, J. R. et al. High-resolution ion-mobility-enabled peptide mapping for high-throughput critical quality attribute monitoring. *Journal of the American Society for Mass Spectrometry.* **32** (8), 2019-2032 (2021).

30. Le Fèvre, A., Dugourd, P., Chirot, F. Exploring conformational landscapes using trap and release tandem ion mobility spectrometry. *Analytical Chemistry.* **93** (9), 4183-4190 (2021).

31. Ohshimo, K., He, X., Ito, R., Misaizu, F. Conformer separation of dibenzo-crown-ether complexes with Na^+ and K^+ ions studied by cryogenic ion mobility-mass spectrometry. *The Journal of Physical Chemistry A.* **125** (17), 3718-3725 (2021).

32. Purves, R. W., Barnett, D. A., Ells, B., Guevremont, R. Gas-phase conformers of the [M + 2H]2+ ion of bradykinin investigated by combining high-field asymmetric waveform ion mobility spectrometry, hydrogen/deuterium exchange, and energy-loss measurements. *Rapid Communications in Mass Spectrometry.* **15** (16), 1453-1456 (2001).

33. Ujma, J. et al. Cyclic ion mobility mass spectrometry distinguishes anomers and open-ring forms of pentasaccharides. *Journal of the American Society for Mass Spectrometry.* **30** (6), 1028-1037 (2019).

34. Warnke, S., Faleh, A. B., Scutelnic, V., Rizzo, T. R. Separation and identification of glycan anomers using ultrahigh-resolution ion-mobility spectrometry and cryogenic ion spectroscopy. *Journal of The American Society for Mass Spectrometry.* **30** (11), 2204-2211 (2019).

35. Williamson, D. L., Bergman, A. E., Nagy, G. Investigating the structure of α/β carbohydrate linkage isomers as a function of group I metal adduction and degree of polymerization as revealed by cyclic ion mobility separations. *Journal of the American Society for Mass Spectrometry.* **32** (10), 2573-2582 (2021).

36. Myers, O. D., Sumner, S. J., Li, S., Barnes, S., Du, X. One step forward for reducing false positive and false negative compound identifications from mass spectrometry metabolomics data: new algorithms for constructing extracted ion chromatograms and detecting chromatographic peaks. *Analytical Chemistry.* **89** (17), 8696-8703 (2017).

Copyright © 2022 JoVE Journal of Visualized Experiments

37. Marchand, A., Livet, S., Rosu, F., Gabelica, V. Drift tube ion mobility: how to reconstruct collision cross section distributions from arrival time distributions? *Analytical Chemistry.* **89** (23), 12674-12681 (2017).

38. Davis, D. M. et al. Analysis of ion mobility spectra for mixed vapors using Gaussian deconvolution. *Analytica Chimica Acta.* **289** (3), 263-272 (1994).

39. Polasky, D. A., Dixit, S. M., Fantin, S. M., Ruotolo, B. T. CIUSuite 2: next-generation software for the analysis of gas-phase protein unfolding data. *Analytical Chemistry.* **91** (4), 3147-3155 (2019).

40. Salbo, R. et al. Traveling-wave ion mobility mass spectrometry of protein complexes: accurate calibrated collision cross-sections of human insulin oligomers. *Rapid Communications in Mass Spectrometry.* **26** (10), 1181-1193 (2012).

41. Gelb, A. S., Jarratt, R. E., Huang, Y., Dodds, E. D. A study of calibrant selection in measurement of carbohydrate and peptide ion-neutral collision cross sections by traveling wave ion mobility spectrometry. *Analytical Chemistry.* **86** (22), 11396-11402 (2014).

42. Richardson, K., Langridge, D., Dixit, S. M., Ruotolo, B. T. An improved calibration approach for traveling wave ion mobility spectrometry: robust, high-precision collision cross sections. *Analytical Chemistry.* **93** (7), 3542-3550 (2021).

Copyright © 2022 JoVE Journal of Visualized Experiments

jove

Materials List for

Using a Cyclic Ion Mobility Spectrometer for Tandem Ion Mobility Experiments

Simon Ollivier[1,2], **Mathieu Fanuel**[1,2], **Hélène Rogniaux**[1,2], **David Ropartz**[1,2]

[1]INRAE, UR BIA, F-44316 Nantes, France [2]INRAE, BIBS Facility, F-44316 Nantes, France

Corresponding Author

Hélène Rogniaux

helene.rogniaux@inrae.fr

Citation

Ollivier, S., Fanuel, M., Rogniaux, H., Ropartz, D. Using a Cyclic Ion Mobility Spectrometer for Tandem Ion Mobility Experiments. *J. Vis. Exp.* (), e63451, doi:10.3791/63451 (2022).

Date Published

January 20, 2022

DOI

10.3791/63451

URL

jove.com/video/63451

Materials

Name	Company	Catalog Number	Comments
33-α-L- plus 23-α-L-Arabinofuranosyl-xylotetraose (XA3XX/XA2XX) mixture	Megazyme Ltd., Wicklow, Ireland	O-XAXXMIX	XA^2XX + XA^3XX mixture
33-α-L-Arabinofuranosyl-xylotetraose (XA3XX)	Megazyme Ltd., Wicklow, Ireland	O-XA3XX	Pure XA^3XX standard
Eppendorf Safe-Lock Tubes, 1.5 mL, Eppendorf Quality, colorless, 1,000 tubes	Eppendorf, Hamburg, Germany	0030120086	Used to prepare the carbohydrate stock solution and dilution
FALCON 50 mL Polypropylene Conical Tube 30 x 115 mm	Corning Science México S.A. de C.V., Reynosa, Tamaulipas, Mexico	352070	Used to prepare the aqueous stock solution of 100 mM LiCl
Lithium Chloride (ACS reagent, ≥99 %)	Sigma-Aldrich Inc., Saint Quentin Fallavier, France	310468	Used to dope the sample with lithium
Major Mix IMS/Tof Calibration Kit	Waters Corp., Wilmslow, UK	186008113	Calibration solution for MS and IMS
MassLynx 4.2 SCN1016 Release 6 (Waters Embedded Analyser Platform for Cyclic IMS 2.9.1 Release 9)	Waters Corp., Wilmslow, UK	721022377	Cyclic IMS vendor software for instrument control and data processing
Methanol for HPLC PLUS Gradient grade	Carlo-Erba Reagents, Val de Reuil, France	412383	High-purity solvent
MS Leucine Enkephaline Kit	Waters Corp., Wilmslow, UK	700002456	Reference compound used for tuning of the mass spectrometer
SCHOTT DURAN 100 mL borosilicate glass bottle	VWR INTERNATIONAL, Radnor, Pennsylvania, US	218012458	Used to prepare the solution of 500 µM LiCl in 50:50 MeOH/Water
SELECT SERIES Cyclic IMS	Waters Corp., Wilmslow, UK	186009432	Ion mobility-mass spectrometer equipped with a cylic IMS cell
Website: http://mzmine.github.io/	MZmine Development Team	-	Link to download the MZmine software
Website: https://github.com/siollivier/IM-MN	INRAE, UR BIA, BIBS Facility, Nantes, France	-	Link to an in-house R script containing a CCS calibration function

Copyright © 2022 JoVE Journal of Visualized Experiments

Part III – Conclusions and Perspectives

In this third part, I investigated the ability of cIMS to characterize the complete structure of oligosaccharides.

First, I attempted to perform total sequencing of oligosaccharides using IMS^n. The first results were promising, showing that the monosaccharidic fragments released from oligosaccharides upon CID matched the corresponding monosaccharides (using the separation strategy designed in **Chapter 7**). When working with homo-oligomers, this monosaccharide fingerprint was even preserved under IMS^n conditions. There were, however, some setbacks. For instance, when fragmenting $[M + Li]^+$ species (that we found to give the best compromise between fragmentation efficiency and IMS separation), the non-reducing end mainly gives B-type ions.[dd] This introduces a double-bond in the structure that can potentially negate the difference between epimeric subunits (*i.e.*, Mannose and Glucose, epimers at C-2). The occurrence of consecutive fragmentations in CID is also a problem, complicating the identification of which monosaccharide was detached from the parent structure. Our next step will be to gain more insight into the behavior of large oligosaccharides in cIMS and IMS^n conditions by switching to a model containing more types of non-isomeric subunits, which should be a first step towards the sequencing of even more complex oligosaccharides.

Because of the difficulties met when trying to develop direct sequencing approaches, I second aimed at developing a more general strategy to organize the glycomics datasets, which would be helpful to gain a first level of structural information. I proposed ion mobility-molecular networking (IM-MN), a strategy inspired from the molecular networks used in metabolomics but that uses IM information of the fragments instead of MS. It was evidenced that this strategy accurately grouped oligosaccharides according to their backbone, and preliminary results indicate that IM-MN coupled to LC can give informative results on real biological samples. While some bioinformatics development are still required to make the most of IMS/IMS data, IM-MN will undoubtedly be useful to organize glycomics datasets and dereplicate oligosaccharides using IMS/IMS spectral databases.

Overall, the results I presented in **Part III** suggest that cIMS will be useful to achieve a complete characterization of oligosaccharides, but some methodological development is still required before such sequencing or IM-MN approaches can be applied routinely.

[dd] Reminder: B ions correspond to a dehydrated mono/oligosaccharide

Part IV – Can cIMS be used to address the polysaccharide challenge?

In **Part III** of the manuscript, I proposed two strategies based on cIMS that have the potential to characterize the complete structure of oligosaccharides. Deciphering the structure at the oligosaccharide level is highly relevant for structural glycosciences, which notably aim at elucidating the structure-function relationship of polysaccharides. In fact, polysaccharides—which can weigh up to megadaltons and adopt tertiary and quaternary structures—are most often studied in a "bottom-up" manner meaning that they are depolymerized into smaller oligosaccharides for characterization purposes. The polysaccharide motif is then reconstructed from the oligosaccharidic blocks.

Yet, the reconstruction step can be a piece of work and depolymerization is not necessarily straightforward either. The preferred method is to use enzymes that specifically cleave some glycosidic bonds (depending notably on the subunits and linkage regioisomerism, but also on lateral branching and on chemical modifications). In that situation, knowledge of the enzyme specificity will be of great help when reconstructing the structure of the polysaccharide. However, this strategy is limited to the availability of relevant enzymes and requires prior knowledge of the backbone of the polysaccharide. In other situations, it may be necessary to use nonspecific methods like chemical hydrolysis.

As an illustration of the above difficulties, in the context of the ALGAIMS project, our partners attempted to use an exo-mannosidase and an endo-mannanase to depolymerize a lichen polysaccharide that was described in the literature as a galactomannan with an α1,6-mannan backbone.[155] Either due to inappropriate specificity or conditions of use, these enzymes did not afford the production of oligosaccharides, which led us to turn to chemical methods of depolymerization, notably acid hydrolysis and the redox-based FITDOG method.[156] Using MALDI-TOF-MS and other methods, we showed that controlling these chemical methods was not straightforward and that they could generate undesired byproducts. This study, although not presented in the present manuscript, will be published and a manuscript is currently being drafted.

Knowing from our own experience how challenging it could be to hydrolyze a polysaccharide into oligosaccharides, we estimated that it could be beneficial to consider developing methods for direct IM-MS analysis of polysaccharides. The following Part of my PhD work therefore focuses on assessing the ability of cIMS and associated methods to provide structural information on polysaccharides. We defined two preliminary questions to evaluate the potential of cIMS for the analysis of intact polysaccharides: (i) ***Up to what size can we visualize polysaccharides?*** and (ii) ***Can we obtain structural information on a polysaccharide over 10 kDa?***

Part IV consists of a single Chapter that presents our efforts in developing IM-MS-based strategies for the direct analysis of polysaccharides, for which we largely took inspiration in the methods used in native MS of proteins. Especially, we explore the potential of coupling cIMS with upstream size exclusion chromatography. A final Chapter provides a synthesis and conclusion of these studies.

Chapter 10. Coupling size exclusion chromatography with cIMS for the analysis of intact polysaccharides

Synthesis and Conclusions of Part IV

Chapter 10 – Coupling size exclusion chromatography with cIMS for the analysis of intact polysaccharides

A. Context

The use of ESI for large (bio-)molecules was introduced by Fenn *et al*. in 1989,[25] and has since been applied to increasingly large molecules and macromolecular complexes in a field called native MS (nMS). The term "native" refers to the analysis of non-denatured proteins and/or of protein complexes, starting with protein/ligand complexes[157,158] and followed by protein/protein interactions[159] and more complex systems. Current approaches in nMS afford the analysis of supramolecular complexes up to tens of MDa such as viruses.[160,161]

With the aim of transposing techniques used for large protein complexes to polysaccharides, we developed a collaboration with the team of Dr Sarah Cianférani in Strasbourg,[ee] which specializes in proteomics and notably on nMS and its native IM-MS (nIM-MS) variant. As part of the PhD work of Dr Evolène Deslignère, we evaluated the interest of cIMS for protein complexes in comparison to low-resolution linear TWIMS instruments.

In a first study, we showed that cIMS analysis of monoclonal antibodies (mAbs) afforded profiles that reflected the conformational space of these immunoglobulins G (IgG). We were able to distinguish three IgG subclasses (IgG1, IgG2 and IgG4) based on their nIM-MS profiles and with the support of collisional activation-based approaches. This publication pending revision is presented in **Appendix 9, Publication 6**. In a second study, we compared different IM-based approaches to study a trispecific antibody, including cIMS and a SEC-IM-MS approach that was developed in the LSMBO team.[24,162,163] The existence of two conformers of the trispecific antibody had been evidenced using SEC but could not be resolved with low-resolution IM. We demonstrated the ability of cIMS to separate these two conformers and evidenced an interconversion between the conformers using IM^n experiments. This study was the subject of a publication presented as **Appendix 10, Publication 7**.

Overall, we demonstrated that cIMS was superior to low-resolution IMS for the classification and structural study of monoclonal antibodies. These encouraging results gave us the incentive to proceed with polysaccharides, the main biopolymers of interest for my PhD work. Indeed, these studies on proteins—which are already quite well controlled—have made it possible to grasp the settings and behavior of high molecular weight assemblies in cIMS,

[ee] Laboratoire de Spectrométrie de Masse BioOrganique (LSMBO), Université de Strasbourg, IPHC, CNRS UMR 7178, 67087 Strasbourg, France

whereas it would have been difficult to optimize the IM parameters directly on polysaccharides because the samples are too heterogeneous.

B. Objectives

Since polysaccharides are subject to high levels of polydispersity and molecular heterogeneity, we have opted in this Chapter for a SEC-cIMS coupling, similar to that used by our partners in Strasbourg on monoclonal antibodies, in order to decomplexify the samples and the data.

To our best knowledge, the literature does not reference any study of the sort and most of the SEC-MALS studies on polysaccharides use solvent systems that are not compatible with mass spectrometry (e.g., $NaNO_3$ as buffer), as well as columns operating at much lower pressures compared to our own system (~ 1–5 bars *vs* ~ 50–70 bars). The optimization of the SEC separation was thus a major step, which was done during the bachelor's internship of Alice Pétrau.[ff] Our objective was to obtain conditions of SEC separation that are compatible with IM-MS and able to separate polysaccharides in a straightforward manner.

First, we evaluated the occurrence of secondary interactions of polysaccharides with the SEC column and optimized the solvent system to ensure that the chromatographic separation resulted purely from a size exclusion mechanism. This optimization was performed with a set of pullulan standards of defined molecular weight (MW). We further assessed whether MS, in these conditions, was able to afford MW and structural information on the polysaccharides. Then, we moved to biological fractions with more complex structures (wheat arabinoxylans), for which we attempted to deconvolve the highly multiplexed signal using IM-MS.

C. Methods

1. Chemicals

Commercial wheat arabinoxylan of low viscosity (P-WAXYL) was used in preliminary tests and was purchased from Megazyme (Bray, Ireland). Commercial protein standards were used to assess the performance of the SEC columns: β-lactoglobuline from bovine milk (β-LG, CAS 9045-23-2), alcohol dehydrogenase from *Saccharomyces cerevisiae* (ADHy, CAS 9031-72-5), and pyruvate kinase from rabbit muscle (PK, CAS 9001-59-6) were purchased from Sigma-Aldrich (Saint Quentin Fallavier, France). The pullulan standards series (P-82 pullulan kit) used to optimize the SEC-MS workflow for polysaccharides were purchased from Shodex (Munich, Germany). Deacetylated wheat arabinoxylan fractions were kindly provided by Dr Luc Saulnier (INRAE, UR BIA). These fractions and their complete characterization have been previously described by Dervilly-Pinel *et al*. in reference [164]. Ammonium acetate ($AcONH_4$) was purchased

[ff] L3 "Sciences, Technologies, Santé mention Sciences pour la santé", La Rochelle Université

from Sigma-Aldrich. HPLC grade isopropanol was purchased from Carlo-Erba (Val de Reuil, France). Ultrapure water was produced using a Milli-Q system (Millipore, Burlington, MA, USA).

2. Buffer exchange

For preliminary experiments in which the samples were infused in the ESI source, an upstream step of buffer exchange was performed to remove the non-volatile salts from the samples (non-volatile salts are known to adversely affect the signal in MS). Samples were desalted against 100 mM $AcONH_4$ with eight cycles of centrifugal microconcentrator (Amicon Ultra, 30 kDa cutoff, Merck Millipore, Tullagreen, Ireland). Manual buffer exchange was not performed ahead of SEC analysis, as SEC has been described as an efficient mean to perform online desalting.[163]

3. SEC parameters

Two SEC columns of different porosity were evaluated: an ACQUITY UPLC Protein BEH SEC 1.7 µm, 4.6 x 150 mm column of 200 Å porosity and an ACQUITY UPLC BEH450 SEC 2.5 µm, 4.6 x 150 mm column of 450 Å porosity. Both columns were equipped with a matching 10 mm guard column (all from Waters, Wilmslow, UK). Analysis were performed using an Acquity UPLC H-class+ system (with biological grade platinum inerting) controlled using MassLynx 4.2 (Waters, Wilmslow, UK). Data were recorded using MassLynx 4.2 in the *m/z* range 50–8000 with a 1 s scan time.

The samples were injected at a concentration of 1.5 mg/mL for polysaccharides (or 5 µM for β-LG and ADHy, 8 µM for PK), with an injection volume of 5µL. The separation was performed at a flow rate of 100 µL/min in isocratic mode, with different solvent systems as described throughout the results. The final selected system was 100 mM aqueous $AcONH_4$ (pH 6.9). The operating pressures were *ca.* 90 bars on the 200 Å column and *ca.* 50 bars on the 450 Å column.

4. IM-MS parameters

Ionization in the Z-spray ESI source was performed with the following parameters: capillary voltage 2.0 kV, cone voltage 80 V, source temperature 100 °C, desolvation gas temperature 300 °C, nebulizer pressure 4 Bars. Details about the voltages and pressures applied in the other parts of the mass spectrometer are available in **Appendix 11, Supplementary text 1**. Notable changes in the parameters compared to oligosaccharides include: ion guide TW increased to 5V at 280 m/s; ion guide gas turned on at a flow rate of 35 mL/min; collision gas flow rate increased to 7 mL/min in the trap and transfer cells; Helium cell entrance, bias, and exit set to 10 V, 30 V, and 20 V respectively; TW height set to 5 V and velocity to 600 m/s in the array; ion guide RF and transfer RF voltages increased to 700 Vpp and 800 Vpp,

respectively. Finally, signal attenuation for *m/z* < 2,000 was performed using the quadrupole to remedy the low-mass bias in ESI.

IMS analysis was done using a TW height of 45 V and TW velocity of 900 m/s in the IM racetrack. Injection in and ejection from the racetrack were done at a TW velocity of 600 m/s. Separation of the ions was done for 10 ms before ejection to the TOF analyzer. In order to work on polysaccharides, the voltages in the multifunction array region, which allows performing fine manipulation of the ion populations during IM experiments, were further modified as described in **Table 2**.

Table 2. Voltages used to set the direction of the ions in the multifunction array region of the Cyclic IMS spectrometer for each type of cIM event, with the default settings (top) and the settings adapted for polysaccharides (bottom). The voltages changed for analysis of high molecular weight biomolecules are highlighted in red.

cIM mobility event	Voltages in the multifunction array region (V)				
	Pre-Array Gradient	Pre-Array Bias	Array Entrance	Array Wave Height	Array Offset
			Default		
Inject / Reinject from Pre-Store	85	70	10	2	45
Separate	85	70	30	0	70
Eject and Acquire	85	70	50	25	45
			High molecular weight biomolecules		
Inject / Reinject from Pre-Store	**95**	**85**	10	**8**	**70**
Separate	**95**	**80**	30	0	**75**
Eject and Acquire	85	70	50	25	**60**

5. Data processing

Data were processed using MassLynx 4.2 and Driftscope 2.9 (both from Waters, Wilmslow, UK). When required, the SEC chromatograms were smoothed using the Savitzky-Golay algorithm (window size (scan): 30, number of smooths: 3). When required, the mass spectra were smoothed using the Savitzky-Golay algorithm (smooth window (channels): 30, number of smooths: 10).

Zero charge mass spectra for the pullulan standards were obtained using bayesian deconvolution in the UniChrom2 module of UniDec 4.4.1.[165] The SEC peaks were split in 10 s slices. The slices were then deconvolved with the following parameters (given *pro forma*): expected mass range: 1 to 500 kDa; sample mass every: 10 Da; Smooth Charge State Distributions: Checked; Use Automatic *m/z* Peak Width, Smooth Nearby Points: Some, Suppress Artifacts: None; Peak Shape Function: Split Gaussian/Lorentzian; *m/z* to Mass Transformation: Smart, Maximum number of Iterations: 50. Peak picking was then performed for each slice with the following parameters: Picking Range: 100 Da; Picking Threshold: 0.1; Peak Normalization: None.

D. Results and Discussion

1. Why SEC-IM-MS?

Following our measurements on monoclonal antibodies, in the case of which the high-resolving power of cIMS afforded MS analysis in direct infusion, we first evaluated what happens when infusing a polysaccharide in ESI. It was apparent that the polydispersity of polysaccharides was a major setback, inducing a vastly different ESI behavior of polysaccharides compared to the proteins that we had analyzed previously.

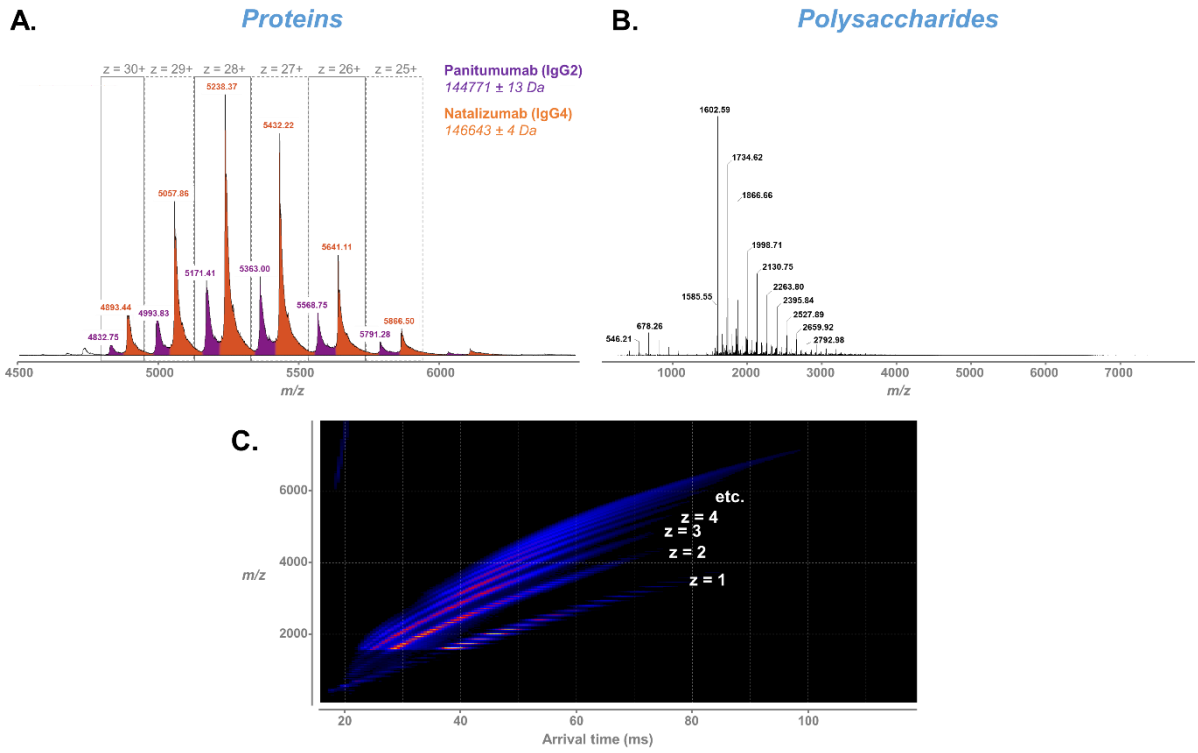


Figure 84. Examples of mass spectra when infusing **A.** proteins and **B.** polysaccharides in electrospray. **C.** Separation of charge states in spectrum B using IMS.

Upon ESI ionization, proteins give an envelope of well-resolved charge states, as illustrated in **Figure 84A** with some of the deglycosylated mAbs that were studied in **Publication 6**. In this example, the relative simplicity of the signal allows to distinguish two mAbs using only the MS dimension, and to accurately measure their masses around 150 kDa.

In contrast, polysaccharides give a more complex MS spectrum with a few peaks of high intensity but also numerous minor peaks, as illustrated in **Figure 84B** for a commercial low molecular weight (LMW) polysaccharide (wheat arabinoxylan low viscosity, P-WAXYL). This polysaccharide has an expected MW of 57 kDa, but at first glance, the MS signal seems to only contain singly- or doubly-charged oligosaccharides of m/z < 3,000. This is in fact not the case, as clarified by the IMS dimension (**Figure 84C**). The MS spectrum actually contains

polysaccharide signals at different charge states and up to *m/z* > 6,000. However, due to ionization competition, the smaller oligosaccharides overwhelm the larger species in terms of signal intensity.

Since most polysaccharides—even those that weigh up to MDa—will contain proportions of smaller species, ionization competition and low-mass bias are major concerns for the study of polysaccharides by IM-MS. One way to address ionization competition and to avoid that smaller oligosaccharides recover all available charges, is to bring higher DPs species to the ion source separately from lower DPs. We therefore turned to SEC-IM-MS coupling.

As briefly described in **Chapter 1**, SEC separates molecules according to their hydrodynamic volume, which is somewhat related to their molecular weight. Indeed, for compounds *of the same molecular family* (e.g., proteins or polysaccharides) the species of high molecular weight will typically have a more important hydrodynamic volume—unless their three-dimensional structure adopts a specific conformation.

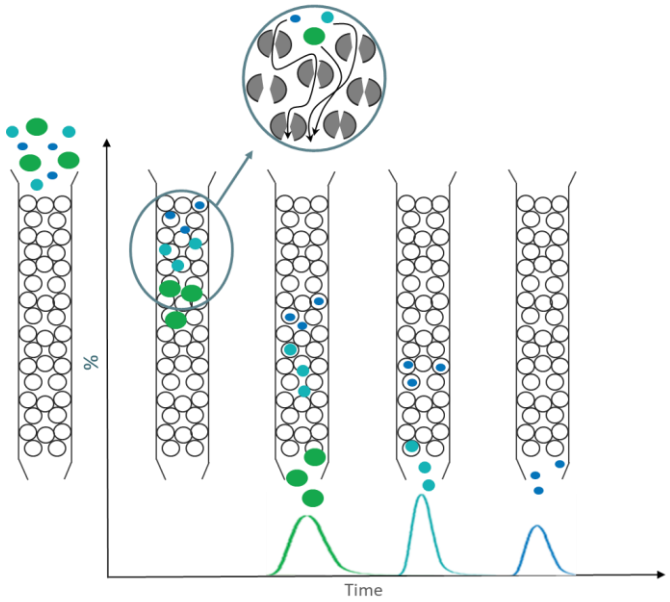


Figure 13. Principle of SEC separation (Reprinted from **Chapter 1**)

The stationary phase in SEC comprises pores centered on a defined pore width (porosity in Å), but some pores will be slightly wider or smaller. SEC separations theoretically purely rely on the path of molecules through the porous stationary phase, as illustrated in **Figure 13**. The species with the most important hydrodynamic volume, in green, will only go through the wider pores and elute first, while the species with smaller hydrodynamic volumes will go through more pores, increasing their path length through the column and they will thus elute later. The porosity of the stationary phase is thus a key parameter to reach the desired separation.

Because the phenomenon occurring in SEC is purely physical (rather than based on chemical interactions), in SEC, the addition of an organic co-solvent should not affect the elution of the analyte. As such, SEC separations are performed in isocratic mode. In practice, there can be secondary interactions with the material of the stationary phase, mostly hydrophobic interactions and/or electrostatic interactions, and the composition of the solvent can be adjusted to reduce these undesirable interactions. Likewise, the temperature does not affect the separation, except in the case where heat would denature the conformation of the polymer and impact its 3D structure.[166]

2. Implementation of SEC-IM-MS for polysaccharides

a) Selection of the column porosity

The first step towards building a SEC-IM-MS workflow was to select a column of appropriate porosity, compatible with UPLC-MS systems. We opted for two columns of the Acquity BEH SEC series from Waters Corp., with porosities of 200 Å and 450 Å. These columns are designed to work on proteins, respectively in ranges of 10–450 kDa and 100–1,500 kDa. However, these MW ranges cannot be transposed directly to polysaccharides, which are more extended polymers compared to proteins (*i.e.*, their hydrodynamic volume for a similar mass is considerably more important).

Because it would have been challenging to evaluate the separation abilities of the two columns on polysaccharides without any *a priori* about the behavior of polysaccharides in SEC operating at over 50 bars, we opted to do this first step of evaluation on native proteins. We selected proteins that cover a relatively large range of MW: β-LG monomer at *ca.* 20 kDa[gg], ADHy tetramer at *ca.* 150 kDa[hh], and PK tetramer at *ca*. 240 kDa[ii].

Figure 85 presents the separation of the three proteins on both columns. On the 200 Å column, the three proteins are relatively well resolved, including ADHy and PK that have the closest MW. However, we observed that the three proteins were already distributed over most of the effective retention time range (after removing the excluded volume from the column). On the other hand, with the 450 Å column, the peak separation was not as good (ADHy and PK almost overlapping) but it left a wider range of available retention time for more extended species. Considering the larger hydrodynamic volume of polysaccharides, we thus selected the 450 Å column.

[gg] Monomer mass = 19,883 Da (for UniProt entry P02754)
[hh] Monomer mass = 36,849 Da (for UniProt entry P00330)
[ii] Monomer mass = 58,048 Da (for UniProt entry P11974)

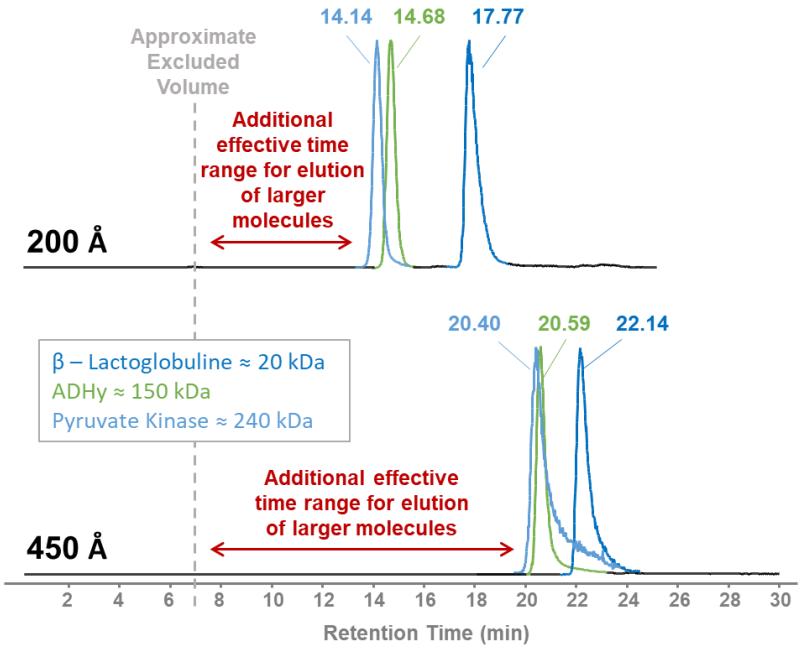


Figure 85. SEC separation of three native protein standards (β-lactoglobuline, yeast alcohol dehydrogenase and pyruvate kinase) on two columns of 200 Å (top) and 450 Å porosity (bottom).

b) Optimization of the solvent system and secondary interactions

We next optimized the solvent system. In our case, the optimum solvent must be compatible with MS, *i.e.*, be volatile and afford good ionization of the polysaccharide (sufficient signal on the polysaccharide without generating too much noise due to salts, etc.). Ideally, the solvent system should also reduce potential secondary interactions with the stationary phase.

We opted to use $AcONH_4$ as a buffer because it is well described in the context of nMS and SEC-MS analysis of proteins.[24] Our starting condition for the optimization was 50 mM aqueous $AcONH_4$. The results of our optimization experiments will herein be exemplified using the P-100 pullulan standard (MW ≈ 100 kDa).

The first secondary interaction that we evaluated was the hydrophobic interactions of the analyte with hydrophobic sites on the packing material. Such interactions increase the retention time of the analyte. They can be disrupted by addition of an organic modifier. In our case, we opted to use isopropanol (IPA) which is compatible with MS (**Figure 86A**). We increased the proportion of IPA in the mobile phase up to 5% (v/v), but not further because polysaccharides tend to precipitate in the presence of alcohols.

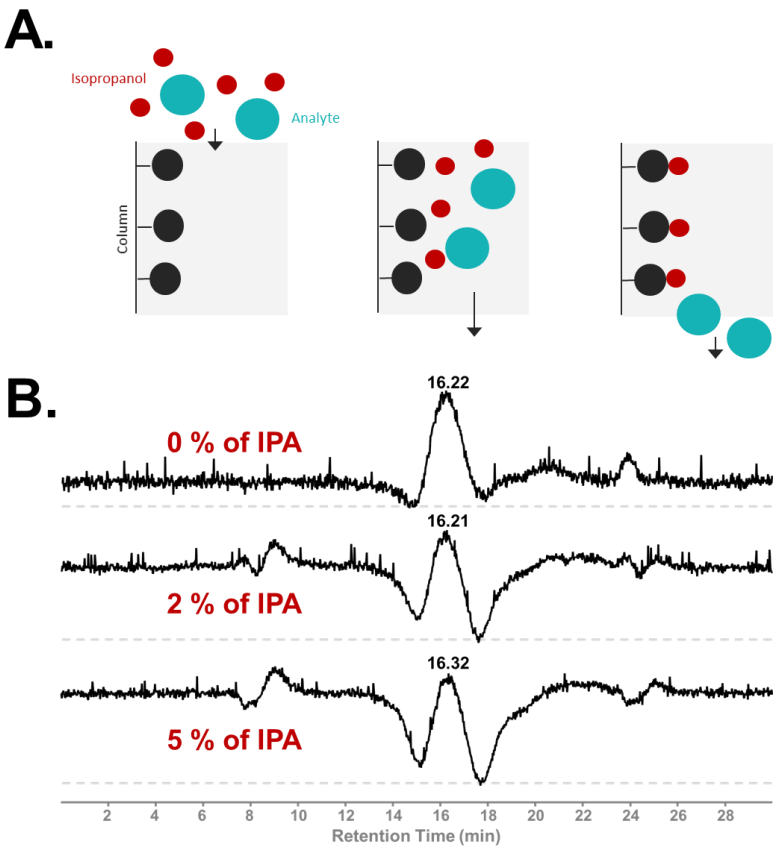


Figure 86. A. Isopropanol disrupts hydrophobic interactions between the analyte and the SEC material; **B.** Total ion chromatogram for the P-100 pullulan standard (*ca.* 100 kDa) at varying concentrations of isopropanol.

The total ion chromatograms (TIC) for the varying concentrations of IPA are presented in **Figure 86B**. We observed no time shift in the retention of the P-100 pullulan standard, indicating that there is no significant hydrophobic interaction of polysaccharides with the SEC support. This was rather expected considering the highly hydrophilic nature of carbohydrates. Of note, the P-100 standard is eluted around 16.2 min, nearly 4.5 min earlier than the PK protein despite being around half its MW. This illustrates what we mentioned above concerning the compactness of polysaccharides, *i.e.*, the expected higher hydrodynamic volume of these macromolecules compared to proteins of similar MW.

On the other hand, when adding IPA, we observed an increase in the baseline level in the parts of the chromatogram where no polysaccharide was eluted. The “noise” is produced by the ionization of the solvent, mainly composed of $AcONH_4$. The increase in noise level is possibly related to a change in spray properties and desolvation of $AcONH_4$ with an higher proportion of IPA. Interestingly, there was a part of the chromatogram around the P-100 peak where the “noise” dropped down, resulting in a “well” in the TIC profile. Hypothetically, this could correspond to the NH_4^+ charges being captured in source by the polysaccharide. There would be a delay between the moment when the polysaccharides start entering the ion source

and capturing charges and the moment when polysaccharide ions are generated in sufficient amounts to be detected. This “ionization well” phenomenon should be attenuated with higher $AcONH_4$ concentrations, since more charges would be available to the polysaccharide.

In a second step, we assessed the occurrence of electrostatic interactions between the analyte and the stationary phase. If the charges of the analyte and the SEC support are the same, the analyte will be excluded from the SEC pores by repulsion effects and thus be eluted faster. On the contrary, if their charges are opposite, the analyte will be adsorbed on the stationary phase. So, increasing the ionic strength of the mobile phase will help attenuate the electrostatic effect by displacing the analyte from the support (**Figure 87A**).

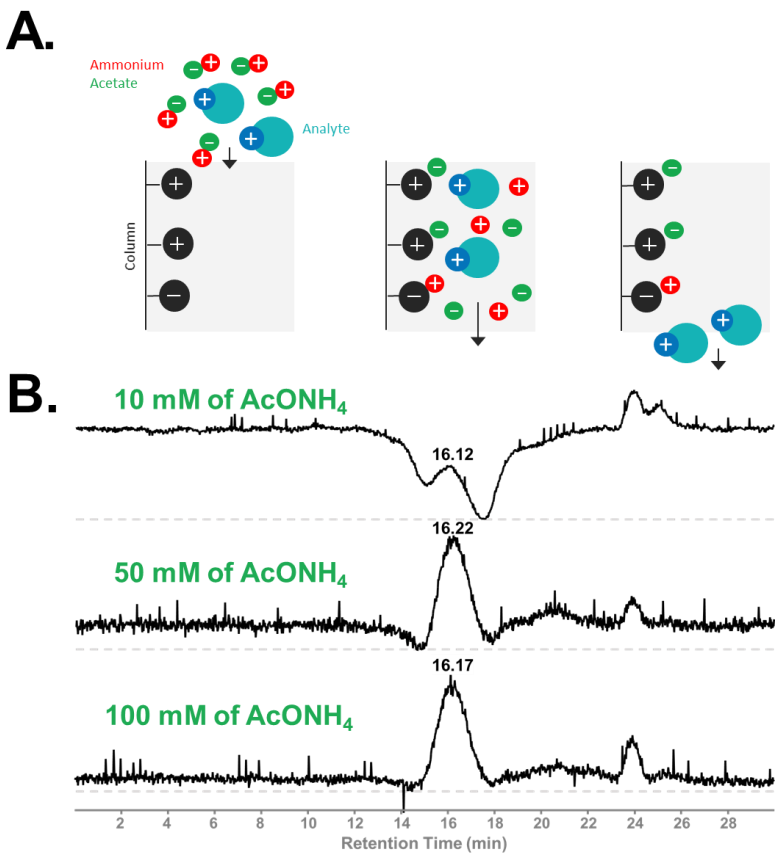


Figure 87. A. Salt disrupts electrostatic interactions between the analyte and the SEC material; **B.** Total ion chromatogram for the P-100 pullulan standard (*ca.* 100 kDa) at varying concentrations of $AcONH_4$.

The TICs for various concentrations of $AcONH_4$ are presented in **Figure 87B**. Like with IPA, we did not observe any shift in the retention time of the P-100 standard with the variation of the $AcONH_4$ concentration. This indicates that polysaccharides do not appear to be subject to electrostatic interactions with the SEC column. The “well” in the TIC profile that we observed with IPA also occurred at lower $AcONH_4$ concentrations, and disappeared at 100 mM of $AcONH_4$. This supports our hypothesis that increased concentrations of NH_4^+ cations help

ionize and detect a greater fraction of the polysaccharide. However, some electric spikes were observed in the signal at 100 mM $AcNH_4$ so we did not raise the concentration of $AcONH_4$ further.

Overall, our results demonstrate that polysaccharides do not develop significant hydrophobic or electrostatic interactions with the SEC support. We can therefore assume that the separation in the following parts of this Chapter will result purely from the size exclusion phenomenon, and reflect the actual hydrodynamic volumes of the polysaccharides. Given the relative signal level of the polysaccharides and the background, we chose 100 mM aqueous $AcONH_4$ as the mobile phase for the following experiments.

c) Separation of pullulan standards

We applied these separation conditions (100 mM $AcONH_4$ at a flow rate of 100 µL/min with the 450 Å SEC column) to multiple pullulan standards: P-200 (MW ≈ 186 kDa), P-100, P-50 (MW ≈ 48 kDa), and P-20 (MW ≈ 23 kDa).

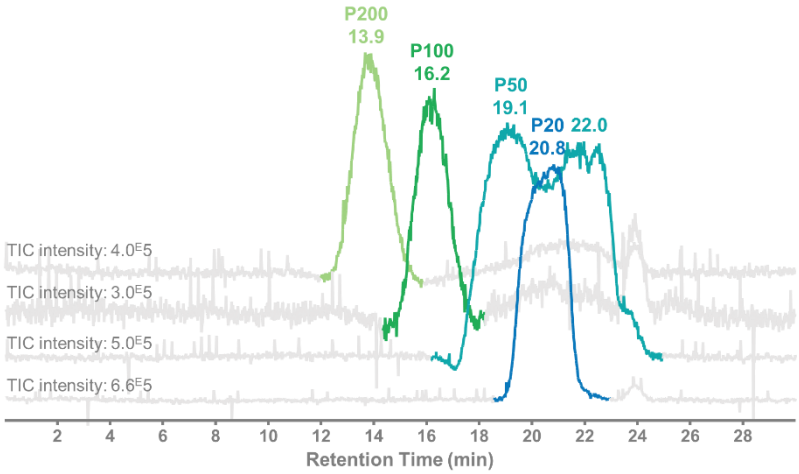


Figure 88. Total ion chromatograms for the pullulan standards P-200, P-100, P-50 and P-20.

Figure 88 presents the MS chromatograms for the four pullulans. The method afforded the SEC-MS detection of these polysaccharides, even at 200 kDa. Furthermore, the different species were relatively well resolved.

Of note, the HMW standards (P-200 and P-100) gave a single gaussian peak, but the profiles of the LMW standards were more complex. For P-50, we noted a main peak at 19.1 min, followed by a valley in the signal and then a second wide peak around 22–23 min. For P-20, we noted a single asymmetrical peak, corresponding exactly to the valley in P-50. The manufacturer does not provide any details regarding the method used to produce these standards, but these profiles may possibly result from a single polysaccharide being

fractionated, which would explain the presence in the P-50 fraction of species with a lower hydrodynamic volume compared to P-20.

3. Does the SEC profile relate to the structure?

Thanks to the MS dimension, we were able to dive deeper into the structures behind the two peaks observed in P-50. We integrated the mass spectra for the peak corresponding to the "expected" retention time (by comparison to the other standards) (blue, RT 19 min), and for the peak with a lower hydrodynamic volume than expected (pink, RT 22 min) as shown in **Figure 89A**.

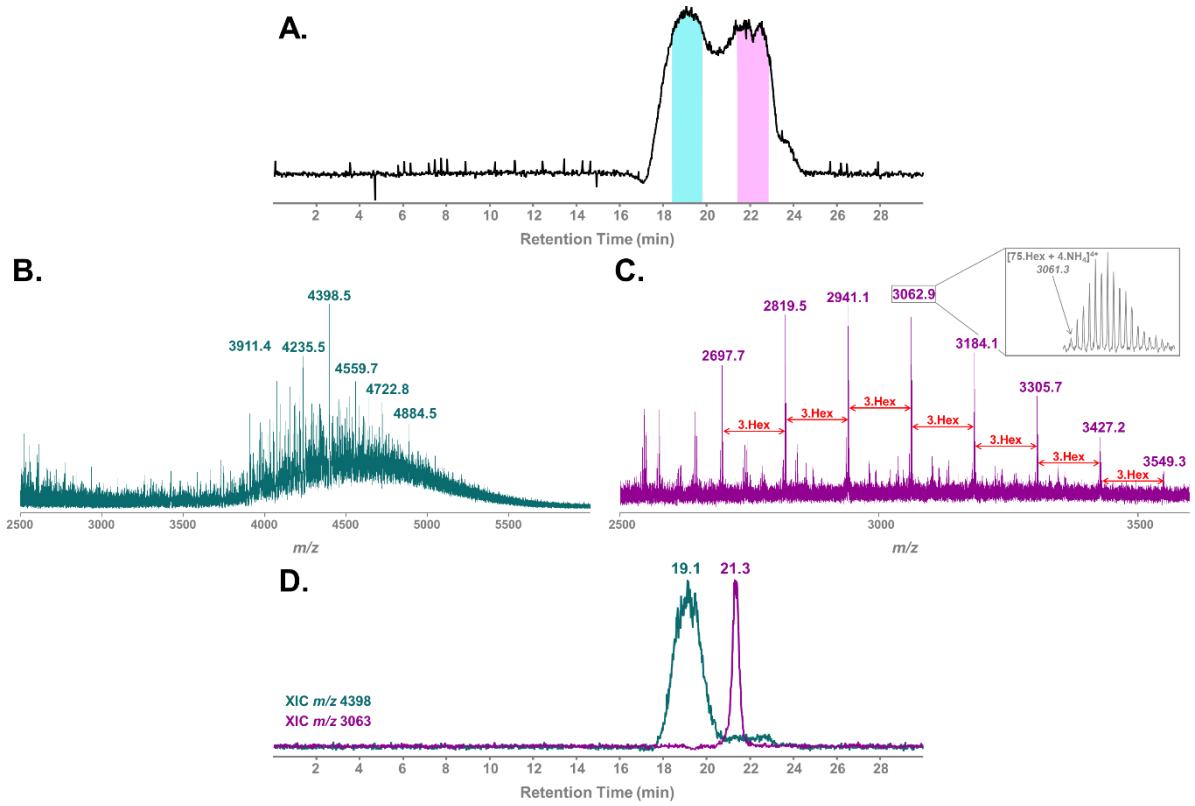


Figure 89. A. Regions integrated in the TIC of the P-50 fraction; **B.** Mass spectrum for the peak at 19 min (blue); **C.** Mass spectrum for the peak at 22 min (pink); **D.** Extracted ion chromatograms for the base peaks of spectra B and C (respectively, *m/z* 4398 and 3063).

The mass spectrum of the peak at 19 min (**Figure 89B**) exhibits a wide MS signal between *m/z* 4000 and 5500, with a few peaks presenting an *m/z* difference of 162 (corresponding to the mass of an hexose, e.g. *m/z* 4398 and 4560), the base peak being *m/z* 4398. The complexity of the signal could be attributed to the MS profile of medium to high MW polysaccharides. The Δ*m/z* 162 pattern likely constitutes an interesting trend, but at this stage, the signal was too convoluted to rationalize the wide MS signal of polysaccharides (this will be studied later in this Chapter on a simpler example).

On the other hand, the mass spectrum of the peak eluted around 22 min (**Figure 89C**) exhibited an envelope of isotopically resolved peaks, the base peak at *m/z* 3062.9 corresponding to the isotopic profile of a $[75.Hex + 4.NH_4]^{4+}$ species (*i.e.*, a small polysaccharide). The peaks in this envelope were separated by a mass difference corresponding to three hexoses. This is in agreement and fairly typical of the structure of pullulans, which are made of maltotriose units (*i.e.*, trisaccharides) linked together by (α1,6) bonds (**Figure 90**). The enzyme that assembles these trisaccharide building blocks can therefore be expected to produce growing polymers separated by the mass of three hexoses.

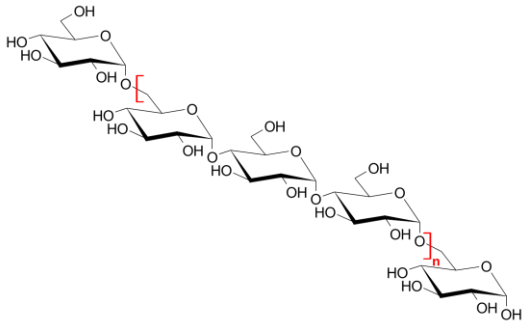


Figure 90. Backbone pattern of pullulans.

To verify that the wide polysaccharide signal and the isotopically resolved species were actually representative of their related SEC peaks, we generated the extracted ion chromatograms (XICs) for *m/z* 4398 and 3063 (**Figure 89D**). We confirmed that the peak from the wide signal was only produced under the “expected” polysaccharide peak (RT = 19.1 min), while the small species of DP 75 was only found in the later fraction (RT = 21.3 min). Interestingly, integrating the isolated DP 75 gave a very sharp peak, but the XIC for *m/z* 4398 remained wide. This is a first sign that there may still be multiples species behind this *m/z* value.

Overall, we show a clear added value of coupling SEC with MS instead of direct infusion ESI. In addition, the use of MS instead of light scattering-based detectors like MALS allowed to obtain precious structural information by identifying the specific trisaccharidic pattern of the pullulan envelope. The smaller polymers that we observe with significant intensity are probably overrepresented in abundance because of the ionization bias towards low-masses as discussed above.

4. Attempts to deconvolve the zero charge mass spectrum

Considering that the data obtained for the pullulan standards were of good quality and contained relevant structural information, we wanted to see if it was possible to use the SEC-MS profiles to trace back to the actual mass of the polysaccharide.

We thus generated zero charge mass spectra using UniDec.[165] UniDec is an academic software that has grown popular in recent years in nMS and nIM-MS, because it allows to deconvolute the mass of protein complexes through a bayesian identification of their charge states. Unfortunately, UniDec is not presently able to deconvolve data using both LC and IM information at the same time. We opted to perform the deconvolution by dividing the chromatogram into thin slices, like conventionally done with MALS (**Figure 91A, D, G, J**).

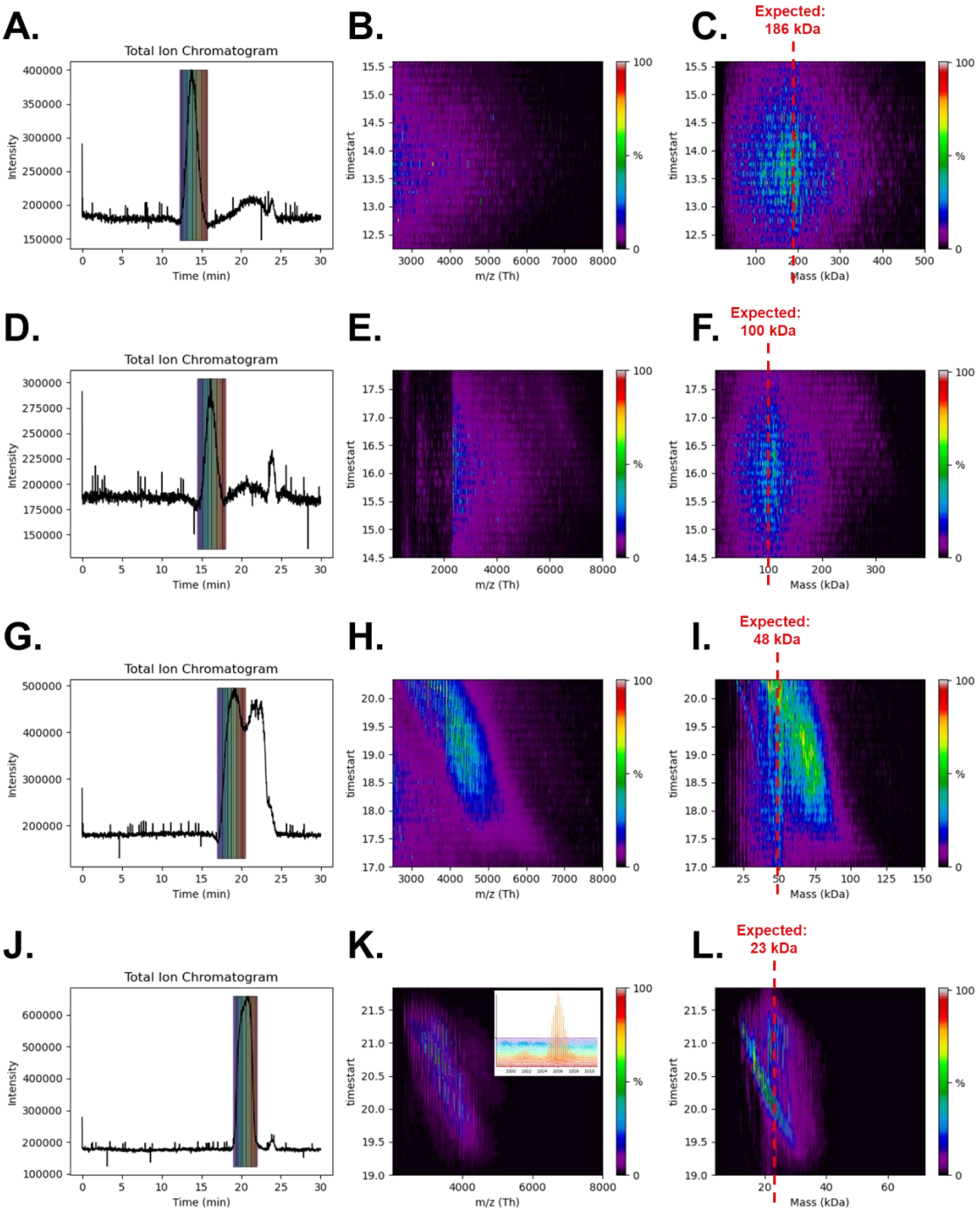


Figure 91. A–C. P-200; **D–F.** P-100; **G–I.** P-50; **J–L.** P-20 standards. **A,D,G,J.** Slices done in the TIC to perform deconvolution; **B,E, H, K.** *m/z* vs retention time plots; **C, F, I, L.** Deconvolved Zero Charge Mass Spectra vs retention time plots.

Of note, a few key parameters had to be changed for each sample, as described in **Table 3**.

Table 3. Specific deconvolution parameters changed in UniDec to deconvolve the Zero Charge Mass Spectrum of each standard.

Sample	Selected time range (min)	Low-*m/z* cutoff	Maximum allowed charge
P200	12.25 – 15.75	2500	*80*
P100	14.50 – 18.00	Ø	*50*
P50	17.00 – 18.75	2500	*20*
P20	19.00 – 22.00	2000	*10*

The Zero Charge Mass Spectra presented in **Figure 91** show that with accurate deconvolution parameters, it is possible to determine the mass of the polysaccharide from the SEC-MS data. Interestingly, P-200 and P-100, which presented gaussian peaks in SEC, showed only one major distribution in their deconvolved mass profile, relatively well centered on the expected masses of 186 and 100 kDa, respectively. However, P-50 and P-20, which presented asymmetrical distributions in SEC presented multiple populations in their mass spectra. Still, these populations remained globally centered on the expected masses of 48 and 23 kDa.

Importantly, the main parameter that had to be adapted for the deconvolved mass to match the expected value was the maximum number of allowed charges. Indeed, considering the level of convolution of the data, it can be conceived that the algorithms will not succeed in accurately estimating the charge state of all species.

To some extent, IMS can help determine the charge state. However, as the charge state increases, the IMS separation will gradually degrade (see **Figure 84C**). It would therefore be worth considering other state-of-the-art MS-based approaches such as charge detection mass spectrometry (CDMS). CDMS now allows to determine the charge state of large molecules with near-perfect accuracy, with no upper mass limit.[167]

5. Analysis of biological fractions of arabinoxylans

The results on the pullulan standards prompted us to proceed with more complex samples such as biological extracts, as well as to evaluate the specific contribution of the IM dimension to the SEC-MS analysis of polysaccharides.

We selected deacetylated arabinoxylan (AX) fractions, kindly provided by Dr Luc Saulnier from our research unit. The chosen fractions have the notable advantage that they underwent complete characterization by the conventional means, and are described in detail in Dervilly-Pinel *et al*. (reference [164]).

In their original publication, Dervilly-Pinel *et al*.[164] evidenced a semi-flexible conformation for these AX, and stated that *"Conformational characteristics […] were similar among all subfractions and fitted with a semi-flexible conformation, whatever their structural characteristics. Substitution degree of the xylan backbone by arabinose residues has no influence on the conformation of arabinoxylans."*

This original publication dating back to 20 years ago, we wished to see if our modern instrumentation could evidence conformational differences based on the structural properties of the arabinoxylans. We therefore selected four fractions of similar MW to see if the SEC-IM profiles would relate to the structural characteristics of the polysaccharides, among which an important characteristics is the proportion of Arabinose to Xylose residues. In fact, AX are made of a Xylose backbone that can be laterally substituted by Arabinose. For a given number of pentose residues (*i.e.*, a given DP and a given MW), polysaccharides with higher Ara/Xyl ratio will be more substituted, with a maximum of two Ara substitution for a single Xyl subunit.

These characteristics are described in **Table 4**.

Table 4. Structural characteristics of the selected arabinoxylan fractions.

Fraction	Molecular Weight (kDa)	Ara/Xyl ratio	Unsubstituted Xyl (%)	Mono-substituted Xyl (%)	Doubly substituted Xyl (%)
F30.8	125	0.38	72	17.9	10.1
F40.6	125	0.56	65.1	13.7	21.2
F50.7	130	0.60	65.5	9.0	25.5
F60.8	125	0.63	63.8	9.4	26.8

Of note, these fractions are labeled according to the fractionation method that was used by Dervilly-Pinel *et al*. First, the AX were fractionated by consecutive precipitations with different concentrations of ethanol (polysaccharide species with a lower A/X ratio precipitate with lower concentrations of ethanol). Second, the fractions produced in the first step were re-fractionated using preparative (low pressure) SEC. The fractions are labeled F*XX*.*Y* where *XX* is the percentage of ethanol for precipitation, and *Y* is the number of the SEC fraction.

The TIC for these four fractions are presented in **Figure 92**, alongside 2D heatmaps of IM arrival time vs SEC retention time. If looking only at the TIC, all fractions seem to present a single peak of extreme width (*ca.* 8 min) with no resolved populations. However, the 2D heatmap evidences the presence of different populations for each fraction, emphasizing the major interest of combining the SEC and IMS dimensions.

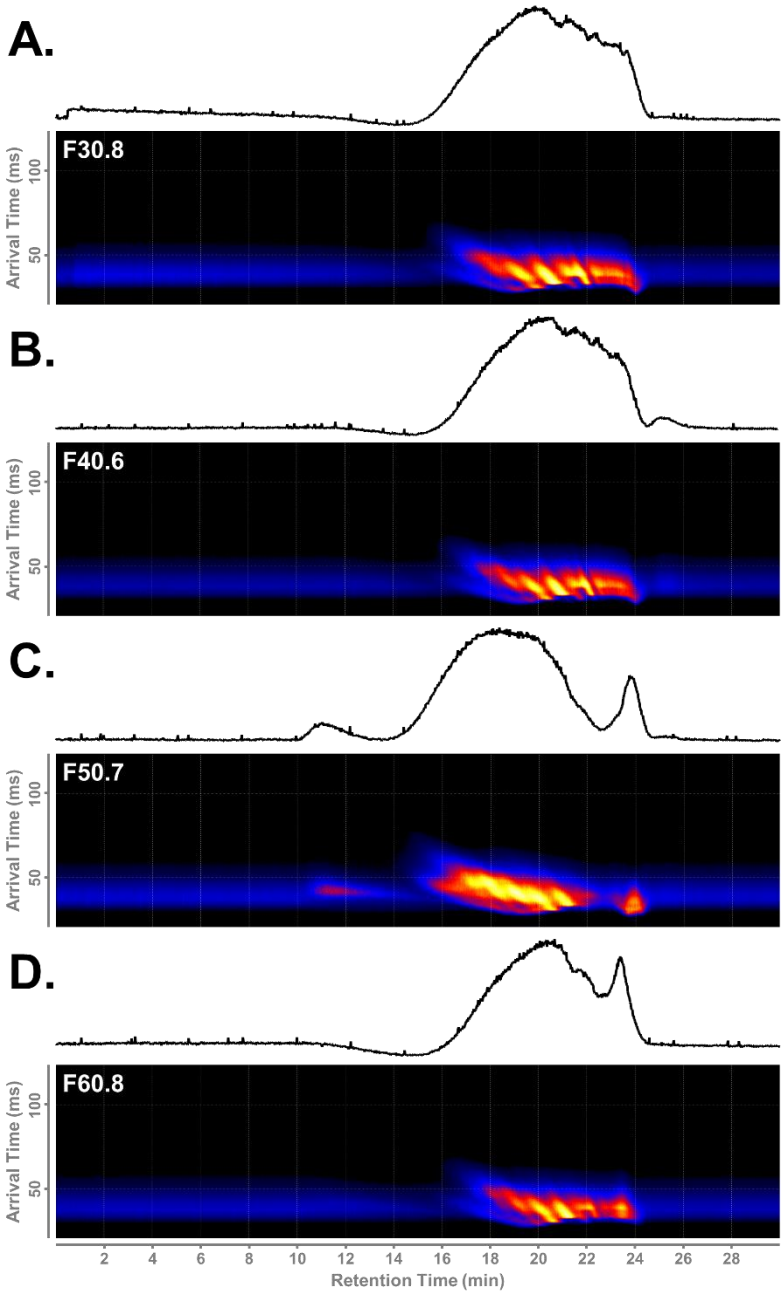


Figure 92. Total ion chromatograms and 2D retention time *vs* arrival time for fractions **A.** F30.8, **B.** F40.6, **C.** F50.7, and **D.** F60.8.

On the other hand, the separated populations do not seem to correlate with the properties of the fractions described in **Table 4**, and are similar for the four fractions. Fraction F50.7 seems to have an overall higher hydrodynamic volume, in line with its slightly higher MW (130 kDa *vs* 125 kDa for the others).

Because of this apparent lack of correlation between the properties of the fractions and the SEC-IM profiles, we further explored the mass spectra of the different SEC-IM populations found in the AX fractions. This is exemplified in **Figure 93** for the five main populations identified in fraction F50.7.

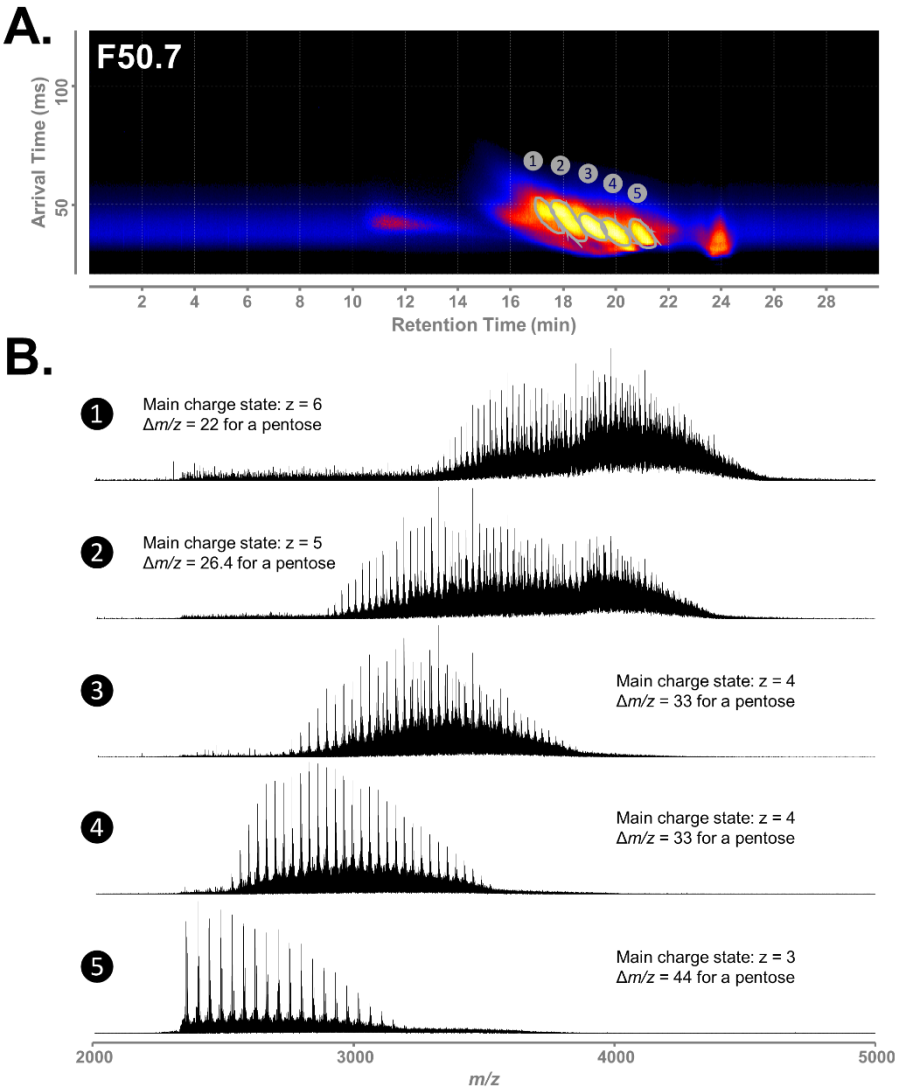


Figure 93. A. Selected SEC-IM regions for the F50.7 AX fraction, **B.** Mass spectra for the selected regions.

First, we observed that the populations that were eluted first from the SEC (higher hydrodynamic volume) had overall higher *m/z* ratios. This confirms that the SEC separation of the AX reflects their MW. Second, we observed one main charge state for each population (z = 6 for population ①, z = 5 for ②, z = 4 for ③ and ④, and z = 3 for ⑤)—with a significant signal above *m/z* 4000 for ① and ②, which suggests underlying higher charge states. Considering that all the populations have different mass spectra, but sometimes similar charges (*i.e.*, populations ③ and ④), this hints towards the SEC-IM separation reflecting an actual heterogeneity within the fraction, possibly conformational.

Overall, we conclude that our own data did not exhibit any significant difference between the SEC-IM populations observed in the different AX fractions. These results are thus in line with the previous observations of Dervilly-Pinel *et al.*, and suggest that our SEC-IM separation may

actually reflect the reported semi-flexible conformation of AX (that we also observed to be independent from the substitution of the xylan backbone). This would be of major interest in the context of polysaccharide characterization: the 3D conformation of polysaccharides being related to their structural family, we envision that the SEC-IM profile could act as a fingerprint of the polysaccharide family. In future experiments, it would be interesting to compare the SEC-IM profiles of arabinoxylans to other types of flexible polysaccharides to further explore this theory.

6. The "Emerald City" pattern in SEC-IM-MS of polysaccharides

The above results demonstrate that SEC-cIMS can be a powerful platform for the analysis of polysaccharides, which should be able to bring complementary information compared to other existing approaches. However, despite the good quality of the data that are generated, they are extremely multiplexed and in-depth interpretation is complicated. *Among the hundreds—if not thousands—of individual features (retention times, arrival times and m/z), what should we look at?* Indeed, a polysaccharide contains hundreds of degrees of polymerization, each likely to contain multiple structures, each giving multiple charge states upon ESI ionization.

In this paragraph, we would like to introduce a phenomenon that occurs in MS of (homo)polymers[jj] and that the group of Scott McLuckey recently discussed.[168] Basically, when analysing polymers, different degrees of polymerization will have close *m/z* values depending on their charge state. For example, a DP 100 charged 10 times will have an *m/z* close to a DP 200 charged 20 times, or a DP 300 charged 30 times. As an illustration, for an ammonium-cationized polypentose, the *most intense isotope* of these DPs would be detected in MS at $[100.Pent + 10.NH_4]^{10+}$ *m/z* 1340.9, $[200.Pent + 20.NH_4]^{20+}$ *m/z* 1340.1, and $[300.Pent + 30.NH_4]^{30+}$ *m/z* 1339.8. In all of these cases, the peaks will have a DP-to-charge ratio (DP/z) of 10. This results in sharp peaks within the convolved mass spectra, occurring where these charge states align, that Chao *et al.* refer to as an "***Emerald City***". The other charge states of the same degrees of polymerization will not be aligned, but rather scattered around the "Emerald City" in what Chao *et al*. refer to as "***wings***". Chao *et al*. included some low MW polysaccharides in their study describing the Emerald City phenomenon (specifically, dextrans centered on 5 and 12 kDa), analyzed using nano-ESI. Our polysaccharides of interest have much higher MW and are more heterogeneous—and thus require adding the SEC dimension—but we still observed this specific pattern.

[jj] The term "homopolymer" refers to a polymer made of only identical subunits. In the case of polysaccharides made of isomeric subunits like AX, we can expect that they will behave similarly to homopolymers.

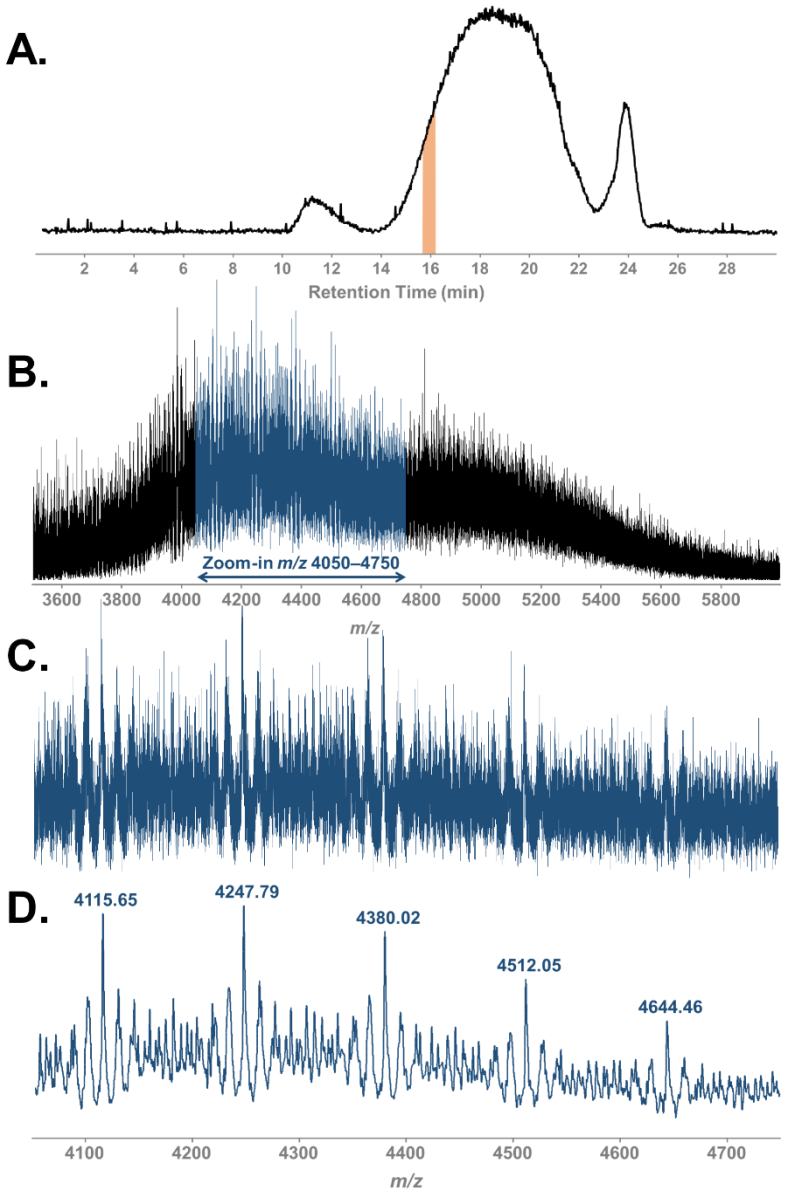


Figure 94. A. TIC of the F50.7 AX fraction, with the selected region highlighted in orange; **B.** Raw mass spectrum for the orange region (in the *m/z* 3500–6000 range); **C.** Zoom-in on the *m/z* 4050–4750 range of the raw spectrum; **D.** Smoothed mass spectrum for the *m/z* 4050–4750 range.

We came across this phenomenon while exploring the mass spectra in the F50.7 AX fraction, more specifically the higher MW region around 16 min (**Figure 94A**). In this region, we do not observe any isotopically resolved peaks, but rather a large MS signal that seems to present a periodic pattern (**Figure 94B**). When zooming-in on this periodic pattern and performing a smooth on this mass spectrum (**Figure 94C** and **D**), we identify a series of sharp peaks separated by Δ*m/z* 132 (corresponding to the mass of a pentose subunit). These peaks, annotated in **Figure 94D**, are “Emerald Cities” and the multitude of undefined peaks that surround them are the “wings”. Looking back, the Δ*m/z* 162 pattern that we observed for

pullulan in **Figure 89B**, are also very likely Emerald Cities, but not as well defined as for the AX which made the phenomenon more difficult to identify on this previous sample.

This Emerald City pattern seems like an interesting option to study polysaccharides, because it helps focus on a single (or at most a few) defined charge state(s) for each DP, and also because Emerald Cities are relatively easy to identify within the extremely multiplexed data generated during MS analysis of polysaccharides.

To evaluate the potential of the Emerald City pattern in decomplexifying the SEC-MS and SEC-IM-MS data, we focused on the smoothed mass spectrum for our AX F50.7 fraction (**Figure 95A**).

In this spectrum, we identified Emerald Cities for DP/z 28 – 38. The XICs of SEC-MS (without using the IM dimension) for all of these DP/z values are displayed in **Figure 95B**. Firstly, when generating XICs for the Emerald Cities, we obtained relatively well defined peaks with a FWHM similar to the pullulan standards. The general trend was that the higher DP/z were eluted faster, indicating a correlation of DP/z with the hydrodynamic volume and thus likely with the MW of the polysaccharide. Secondly, integrating the DP/z values allowed us to evidence at least two populations in the AX fraction, the main population ① eluting earlier than the minor population ②. These two SEC populations could possibly be related to different structures, or more likely to the conformational flexibility described by Dervilly-Pinel *et al*.

Finally, by adding the IM dimension of the data we were able to further deconvolve each DP/z peak (**Figure 95C**). For species of lower DP/z, we were able to split the DP/z SEC profile into charge-defined DPs using the information obtained from HR-IMS. For instance, in the case of DP/z 28 we identified charges from z = 4 to z = 9 (*i.e.*, DP 112 – 252, or MW 15 – 33 kDa). The charge increased with DP/z, becoming increasingly difficult to separate in IMS, until the point where we were unable to accurately determine any charge state for DP/z 38. Nevertheless, for the lower DP/z values, the IM dimension helped us obtain a lot more information compared to simple SEC-MS. As an illustration, the mass spectrum for DP/z 28 of population ① isolated with the SEC dimension is presented in **Figure 95D**. This mass spectrum is not informative, as it is impossible to distinguish the contribution of each DP—we only observe peaks separated by 1 mass unit, that we interprete as an isotopic version of the Emerald City (every 1 m.u. some ^{13}C isotopes of the various DPs in presence would overlap). On the other hand, adding an ion mobility filter to separate the charge states allowed us to isolate a specific DP and to obtain isotopically-resolved MS spectra, as illustrated for DP 112 and DP 140 (DP/z = 28 with z = 4 and z = 5, respectively) in **Figure 95E**.

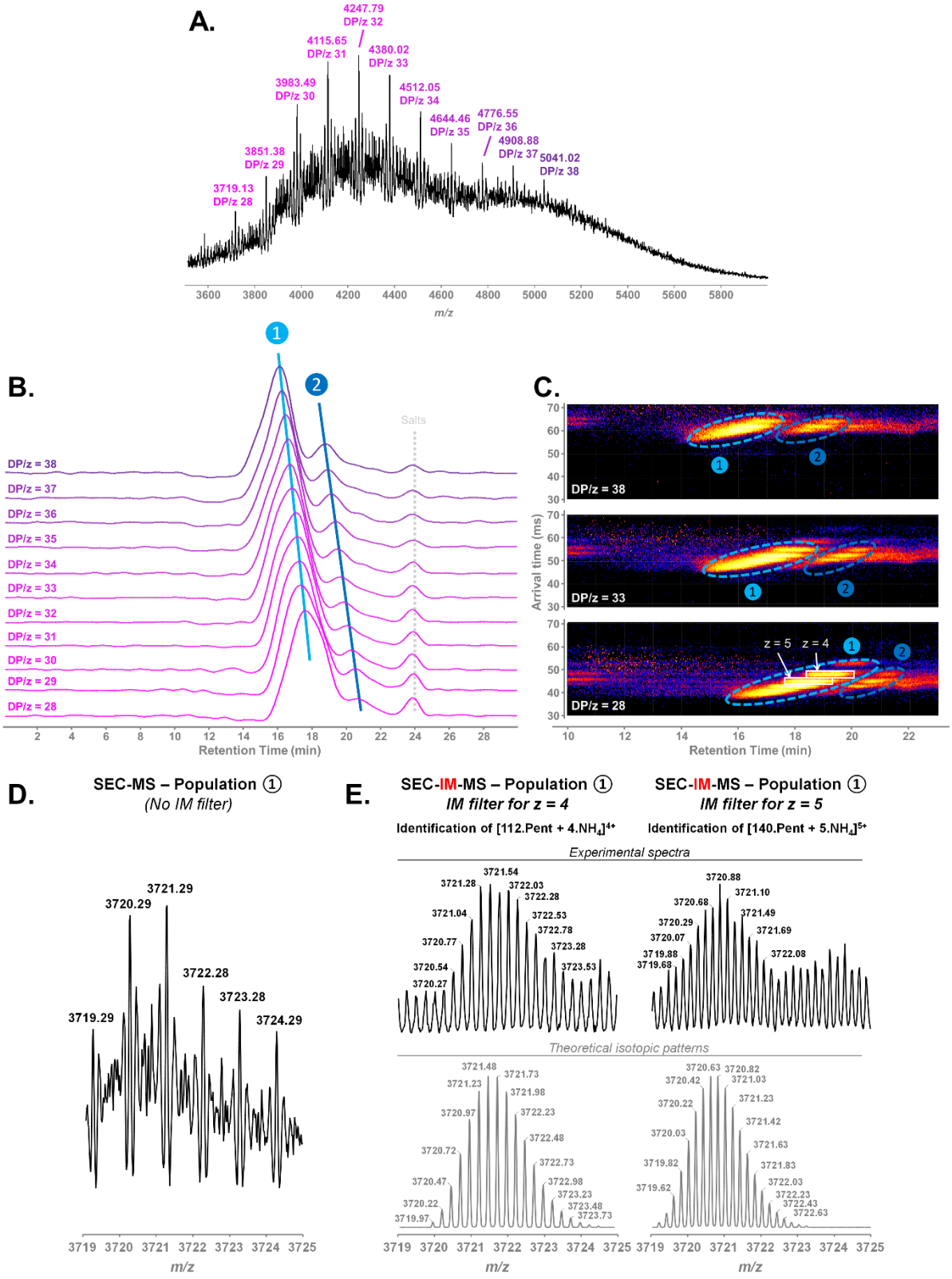


Figure 95. A. Smoothed mass spectrum for the F50.7 AX fraction (see **Figure 94B**); **B.** Extracted ion chromatograms for DP/z 28–38 (smoothed); **C.** Examples of 2D retention time vs arrival time for DP/z 28, 33, and 38. **D.** Mass spectrum of DP/z 28 in population ① using only the SEC dimension as a filter. **E.** Example mass spectra for DP112 and DP140, isolated at DP/z 28 in population ① using SEC-IM filtering (the theoretical isotopic patterns for these species are given in gray for reference).

Overall, this Emerald City pattern seems like a promising way to tackle the challenge of interpreting the highly multiplexed data generated by SEC-IM-MS analysis of HMW polysaccharides, as it allows to identify hotspots in the MS spectra than can be targeted to interprete the results. Furthermore, adding the IM dimension to the SEC-MS workflow is key, because it affords to isolate individual DPs in the very complex polysaccharide matrices.

E. Conclusions and Perspectives

In this Chapter, we have developed and optimized a SEC-cIM-MS setup for the direct analysis of intact polysaccharides. Our setup demonstrated the ability to separate and detect polysaccharides ranging from 20 to 200 kDa. We were also able to obtain a first level of structural information on the analyzed polysaccharides, like by identifying the trisaccharidic motif of pullulan or obtaining the mass of the building blocks through the Emerald City pattern. Our setup finally proved to be able to work directly on highly fractionated biological products, highlighting the good quality of the generated data.

The main drawbacks that we could identify in this approach mostly concern the data interpretation and processing. With regards to the high number of DPs, structures, and charge states, the high multiplexing of the data was to be expected. The Emerald City pattern appeared to be an interesting way to address this complexity by focusing the interpretation on certain values of DP/z. A major limitation remains the difficulty to accurately identify the charge state, even with the help of IMS although the charge state is a key information to deconvolve the *m/z* dimension into MW. Of note, it was not possible to move to multipass cIMS strategies to increase the resolution as too many species would wrap around for an unknown number of passes. As discussed above, CDMS might be an alternative to consider.

F. Scientific Communication

These results will be published in an upcoming article.

Part IV – Conclusions and Perspectives

In this fourth Part, I set up and optimized the coupling of size exclusion chromatography to directly analyze intact polysaccharides on the cIMS platform.

The method was successful in separating and measuring pullulans up to 200 kDa, as well as some complex arabinoxylans of around 125 kDa. Of note, the IM resolving power of the cIMS afforded the distinction of polysaccharide populations in otherwise unresolved SEC profiles. This illustrates the complementarity and orthogonality of SEC and IM, SEC providing information related to the liquid-phase hydrodynamic volume of the polysaccharides and IM to the gas-phase cross section. The SEC-IM-MS coupling provided a first level of structural information on the polysaccharides (mass of the constituting monosaccharides, specific trisaccharidic motif), which is a substantial advantage compared to other existing SEC-based methods like SEC-MALS.

Overall, these results open the way to the characterization of intact polysaccharides by MS, which has scarcely been reported in the literature. Earlier attempts to use SEC-MS for polysaccharides only allowed to detect species a little in excess of 9 kDa.[169] However, we note a recent method that could potentially complement our own approach, in which polysaccharides up to over 3 MDa were separated using SEC and then fragmented with in-source CID before structural analysis—this is equivalent to an in-line middle-down approach.[170]

However, computing the MW of the polysaccharide was challenging, notably because of difficulties in accurately identifying the charge states. Even if the Emerald City phenomenon showed promise for identifying hotspots in SEC-IM-MS data and assigning lower charge states, charge state determination will be an important point in future developments of the SEC-IM-MS workflow for polysaccharides. In this respect, CDMS will likely be an interesting option to consider. Furthermore, the lack of fully adapted processing software is a major limitation: to our best knowledge, there are no existing software adapted to so highly multiplexed data containing thousands of features. In the future, it will likely be important to improve processing software to be able to readily interprete SEC-IM-MS data of polysaccharides.

General Conclusions and Perspectives

My PhD work revolved around the potential offered by new high-resolution IM technologies, specifically in my case the cIMS platform, for the structural characterization of oligo- and polysaccharides. In this context, I conducted several studies, which aimed to gradually address **three fundamental issues** connected to the structural characterization of oligo and polysaccharides:

- **Resolving the different types of isomerism found in oligosaccharides:** In the first stages of my PhD, I endeavored to show that the different structural features (*i.e.*, isomerisms) found in carbohydrates could be resolved using cIMS. Previous studies from our laboratory had already showed that cIMS could resolve the reducing-end anomerism[117] and the position isomerism of functional groups with post-IMS fragmentation.[103] I have continued and extended these initial studies by demonstrating that cIMS could resolve all other possible cases of isomerisms. Precisely, I demonstrated that cIMS could determine other situations of constitutional isomerism such as cyclization tautomerism and linkage regioisomerism. I established that the IMS signature profiles for the different isomers were—in most cases—retained upon collisional fragmentation. I then moved to stereoisomers and demonstrated that cIMS could determine the intrachain anomerism of the glycosidic bond, even in mix-linked species. Finally, I designed cIMS sequences that were able to generate specific profiles for isomeric monosaccharides of different families, notably including cases of epimeric monosaccharides (e.g., Glc vs Man that are C-2-epimers, or Glc vs Gal that are C-4-epimers).
- **Achieving a complete structural determination of oligosaccharides:** The demonstration that cIMS could provide “fingerprints” that were both discriminating of all types of isomerism in oligosaccharides and retained upon fragmentation gave us the incentive to tackle one main challenge, the complete structural determination of (complex) oligosaccharides. First, I attempted to create a method to sequence oligosaccharides right from the IM fingerprint of the constituting monosaccharides. We validated key points on our path towards sequencing: (i) the mobilities of monosaccharidic fragments match with the corresponding monosaccharides, (ii) these signature profiles are preserved during multistage IMS^n experiments. However, we also

faced notable limitations, such as the poor duty cycle of IMS^n experiments and more importantly, the consecutive fragmentations (and lack of structurally informative cross-ring fragments) upon CID. While this sequencing project is still ongoing, these difficulties have led me to propose, as an alternative, an untargeted approach that allows to organize glycomics datasets and obtain a first level of structural insight. To this end, we drew inspiration from untargeted metabolomics and created an ion mobility-molecular networking (IM-MN) strategy that classifies oligosaccharides using the CCS of their fragments, with a high level of structural relevance.

- **Analyzing intact polysaccharides in IM-MS:** The field of structural glycosciences is still much hampered by the lack of analytical approaches capable of dealing with the intact polysaccharide, coupled with the difficulty in successfully depolymerizing a polysaccharide of interest due to the lack of enzymes or sufficient information on the overall structure of the polymer. This led us to envisage a direct analysis of polysaccharides by cIMS. We thus built a SEC-cIMS setup (which, to the best of our knowledge, is a first), and were successful in analyzing polysaccharides up to 200 kDa. While the SEC-MS dimension proved useful to gain some level of structural insight (mass of the constituting monosaccharides, trisaccharidic motif of pullulans), the IM dimension proved complementary to SEC and MS and was useful to resolve populations that were not separated using SEC alone in the case of arabinoxylans. Altogether, the SEC-cIMS coupling has the potential to unveil information that was previously unattainable using a single platform.

The various studies that I conducted during my PhD thesis undoubtedly support the potential of cIMS (and more generally, of high-resolution IM) for the structural characterization of carbohydrates—from monosaccharides weighting as low as 150 Da to polysaccharides weighting hundreds of kDa. However, we also identified several challenges that will require further studies and/or developments over the next few years:

- **Gas-phase rearrangements of oligosaccharides:** We encountered our first major setback during our work on carbohydrate isomerism, in the place of rearrangements in galactomannan oligosaccharide structures. Our HR-IMS results suggest that there can be *both conformational and structural rearrangements* in the gas phase, although this was only noted on Gal*f*-containing compounds, which means in a highly specific case. Nonetheless, these results suggest a probable migration of the Gal*f* residue on $[M + Li]^+$ species (accompanied by internal residue loss of a Man), which will lead us to remain

cautious in the interpretation of the observed species especially in experiments involving an activation of ions.

Although these migrations phenomena have been known since the 90's,[171–173] they were long believed to occur only on $[M + H]^+$ or $[M + NH_4]^+$ species,[149] to the point where some authors proposed to use metal adducts to avoid this migration phenomenon.[174] Migrations on alkali metal adducts of oligosaccharides were only recently described on $[M + Na]^+$ species.[151] Furthermore, the migration of Gal had not been described previously (to our best knowledge), the literature mainly containing description of fucose, xylose or rhamnose migration.[171,173,175] Hexose rearrangements have been previously reported,[176] but such descriptions are rare. We believe this may be because such rearrangements are difficult to identify with MS alone—whereas the mass of a Fuc residue, for instance, is easier to detect in Hex (or HexNAc, etc) backbones. In the future, it seems important to us to conduct further studies to extend our knowledge on glycan rearrangements. The contribution of gas-phase methods such as IMS (but also IRMPD[177]) may help uncover new types of rearrangements or new mechanisms.

- **Consecutive fragmentations and lack of intracyclic fragments in CID:** Another major point that will require future studies and development is the limited information provided by CID fragments. Indeed, CID mostly gives glycosidic bond cleavages and few cross-ring fragments. While glycosidic bond cleavages are interesting because entire subunits are released from the precursor oligosaccharide, that information is rendered ambiguous because of consecutive fragmentations. This also limits the assignment of linkage/functionalization regioisomers. A known way to address this issue is to use alternative fragmentation methods to generate the more informative cross-ring fragments. The literature contains many examples in which high-energy fragmentation methods were used to characterize carbohydrates. These methods notably include electron-based methods, such as ECD[178] or EID.[179] However, until very recently, these particular methods were restricted to FTICR instruments.

  High-energy activation methods have been one of the focuses of our laboratory over the past few years. After promising results obtained using photon-based activation (XUVPD using synchrotron radiation),[39] our lab collaborated with the group of Prof. Glen Jackson at West Virginia University to implement helium-charge transfer dissociation (He-CTD), which can be mounted on a benchtop instrument. He-CTD is an ion-ion interaction-based fragmentation method yielding radical-induced fragments. Our group showed that He-CTD gives similar fragments compared to XUVPD[40] and has exploited these performances on several occasions to characterize complex oligosaccharides.[22,41,180–182] Notably, in a recent study, we combined the high peak

capacity of LC-cIMS (performed on the cIMS instrument) with LC-CTD (performed on another MS platform) to separate and decipher the structure of eight glucosylation isomers in a raw enzymatic reaction medium. This study is presented in **Appendix 12, Publication 8**. This study clearly highlights the interest of a single platform coupling cIMS with high-energy fragmentation. We thus recently invested in an ExD cell[kk] to *perform high-energy fragmentation (ECD and EID) directly on our cIMS platform*. In the near future, it will be interesting to test the capabilities of the ExD cell to generate cross-ring fragments in both pre-IMS and post-IMS strategies.

- **Challenges related to the application of IM-MN to biological samples:** The IM-MN strategy that we developed proved very promising when applied to a large set of standard oligosaccharides representative of plant carbohydrates. Yet there are still many points of optimization that need to be addressed before a potential adoption of this tool by the glycomics community. First and foremost, MN strategies in metabolomics are designed to be used on *LC*-MS data to increase specificity. We have conducted some validation steps regarding the compatibility of IM-MN with LC, but we were restricted to the use of CID-IMS approaches instead of IMS/IMS, which is a significant degradation of the approach. This limitation is directly related to the control software for the cIMS instrument, which at the time of writing this thesis does not offer the possibility to perform data dependent acquisitions (DDA) for IMS/IMS. The implementation of DDA for the cIMS, which is expected to be launched soon, will undoubtedly be a major step forward.

  A second critical point that will require further development is the scoring system used to build the IM-MN networks. In our initial approach, we built the network using only the CCS dimension of the fragmentation spectra, sacrificing the *m/z* dimension in the process. It would be highly beneficial to develop a bidimensional scoring system that accounts for both the CCS and the *m/z*, because the information contained in both dimensions is non-redundant. However, even when working on a single dimension like we did, there is room for improvement. Indeed, we used the most classical cosine similarity score but a number of other scores have been proposed to improve the performances of MS-based molecular networks. At least a few of the best known scoring systems—like the modified cosine[183] or Spec2Vec[184], for example—deserve to be tested for our IM-MN approach.

  Once the IM-MN method is optimized to a satisfactory point, it will also be key to build IMS/IMS spectral databases that will be usable to dereplicate glycans in complex

[kk] Reminder: the ExD cell is described in **Chapter 3**.

mixtures. Ideally, such databases should be built upon global community efforts and freely accessible, like the GNPS initiative for metabolomics.[153]

- **Analytical challenges related to the analysis of HMW polysaccharides:** During my PhD work, we obtained preliminary yet promising results using SEC-cIMS for intact polysaccharides. Nonetheless, interpreting the highly multiplexed data generated by these analyses was a major challenge. This is in part due to the lack of adapted processing software, as intact polysaccharides are made of up to thousands of DPs, each giving multiple charge states. We anticipate that further software development will help process such data more easily and more accurately. Indeed, increasingly powerful options have recently emerged to process MS-based data of biopolymers. For instance, the UniDec software, originally designed to only deconvolve zero charge mass spectra from MS or IM-MS data of intact proteins,[165] has since then been improved to support LC-MS data (UniChrom), and to perform batch processing (MetaUniDec).[185]

  Another critical point that we identified is the necessity for an accurate evaluation of the charge state of polysaccharides. This is a key point to deduce the actual mass of polysaccharides and it has proved difficult to precisely estimate *z* using the MS and IM dimensions—even with the increased resolving power of the cIMS instrument. A first option to consider would be to turn to even higher IM resolving power configurations to better resolve the charge states. This is difficult with cIMS because of the wrap-around phenomenon, but SLIM have previously been used to analyze native proteins.[186] However, this would likely imply the removal of the SEC dimension because SLIM as they stand are difficult to couple with LC because of extended separation times (to our knowledge, the LC-SLIM coupling has only been reported on one occasion with nano-RPLC[187]). Another—and possibly more attractive—option is CDMS that can directly detect both *z* and *m/z* (*z* is determined for individual ions as they pass through a conductive cylinder, because *z* impacts the amplitude of the induced current).[188] CDMS thus affords the analysis of complexes up to several dozen MDa,[189] and also notably of highly heterogeneous samples such as heavily glycosylated SARS-CoV-2 spike protein isoforms (which is challenging because of glycan heterogeneity).[190] Of note, an adaptation of the UniDec software for CDMS data (UniDecCD) was implemented recently,[191] which makes this option even more appealing. However, it must also be noted that since CDMS is a single ion method, the acquisitions are also lengthy and difficult to combine with chromatographic approaches—but developments are currently performed to monitor multiple ions at once and thus reduce the acquisition time.[192]

To conclude, following the example of what has been developed in proteomics, we foresee two potential directions that structural glycosciences could take with the possibilities opened up by the developments that I conducted during my thesis. The first would be the *de novo* determination of the fine structure of polysaccharides using ***bottom-up sequencing***, where the polysaccharide is hydrolysed (chemically or enzymatically) before determining the structure of the generated oligosaccharides and then reassembling the motif of the polysaccharide. The main advantage of this method would obviously be that it relies on the analytical strategies that I developped during my PhD and which have become very efficient for the determination of oligosaccharides structures—as we have shown throughout this manuscript. However, independently of the analytical tools, the hydrolysis of the polysaccharide leads, on the one hand, to the loss of information regarding the organization of the oligosaccharides blocks within the polymer chain and on the other hand, makes any structural heterogeneity (size, presence of different populations, etc.) undetectable. The second option would be the development of methods that directly address intact polysaccharides and head towards ***top-down structural analysis*** with MS-based techniques. The main advantages of such approaches would be the preservation of polysaccharide populations, as well as the ability to come back to their intact mass, while also uncovering their structural features in a single analysis. On the other side, despite the contribution demonstrated by the use of the SEC-IMS-MS coupling, it appears quite difficult to obtain very detailed structural information from intact polysaccharides at the present stage of performance of analytical methods. ***Middle-down strategies*** could offer a good compromise of the two approaches. Indeed, larger parts of a polysaccharide (e.g., 10–20 kDa) would be more representative of the complexity and topology of the polymer, compared to the oligosaccharides we usually work with. A first line of work would be to apply a strategy similar to that developed by Jiang *et al.*,[170] *i.e.*, to couple SEC with a first step of in-source decay to reduce the size of the polysaccharide and to increment the workflow with our own cIMS methodologies for structural characterization. We envision that using the ExD cell to fragment the polysaccharides with ECD rather than CID could also help with the structural characterization step.

Finally, although much work is still needed to fully develop the bottom-up and top-down strategies—and perhaps explore the middle-down approach as well—this PhD work proves that they are now within our reach thanks to cIMS.

References

(1) Laine, R. A Calculation of All Possible Oligosaccharide Isomers Both Branched and Linear Yields 1.05x10(12) Structures for a Reducing Hexasaccharide - the Isomer-Barrier to Development of Single-Method Saccharide Sequencing or Synthesis Systems. *Glycobiology* **1994**, *4* (6), 759–767. https://doi.org/10.1093/glycob/4.6.759.

(2) de Souza, M. A.; Vilas-Boas, I. T.; Leite-da-Silva, J. M.; Abrahão, P. do N.; Teixeira-Costa, B. E.; Veiga-Junior, V. F. Polysaccharides in Agro-Industrial Biomass Residues. *Polysaccharides* **2022**, *3* (1), 95–120. https://doi.org/10.3390/polysaccharides3010005.

(3) Xu, S.-Y.; Huang, X.; Cheong, K.-L. Recent Advances in Marine Algae Polysaccharides: Isolation, Structure, and Activities. *Mar. Drugs* **2017**, *15* (12), 388. https://doi.org/10.3390/md15120388.

(4) Khan, M. S.; Zhang, X.; You, L.; Fu, X.; Abbasi, A. M. Structure and Bioactivities of Fungal Polysaccharides. In *Polysaccharides*; Ramawat, K. G., Mérillon, J.-M., Eds.; Springer International Publishing: Cham, 2015; pp 1851–1866. https://doi.org/10.1007/978-3-319-16298-0_28.

(5) Naik, S. N.; Goud, V. V.; Rout, P. K.; Dalai, A. K. Production of First and Second Generation Biofuels: A Comprehensive Review. *Renew. Sustain. Energy Rev.* **2010**, *14* (2), 578–597. https://doi.org/10.1016/j.rser.2009.10.003.

(6) Ferdouse, F.; Holdt, S. L.; Smith, R.; Murúa, P.; Yang, Z. *The Global Status of Seaweed Production, Trade and Utilization*; Globefish Research Programme; Food and Agriculture Organization of the United Nations, 2018.

(7) Khalid, A.; Asim-ur-Rehman; Ahmed, N.; Chaudhery, I.; Al-Jafary, M. A.; Al-Suhaimi, E. A.; Tarhini, M.; Lebaz, N.; Elaissari, A. Polysaccharide Chemistry in Drug Delivery, Endocrinology, and Vaccines. *Chem. – Eur. J.* **2021**, *27* (33), 8437–8451. https://doi.org/10.1002/chem.202100204.

(8) Tanaka, S.; Kojić, D.; Tsenkova, R.; Yasui, M. Quantification of Anomeric Structural Changes of Glucose Solutions Using Near-Infrared Spectra. *Carbohydr. Res.* **2018**, *463*, 40–46. https://doi.org/10.1016/j.carres.2018.04.012.

(9) Varki, A.; Cummings, R. D.; Aebi, M.; Packer, N. H.; Seeberger, P. H.; Esko, J. D.; Stanley, P.; Hart, G.; Darvill, A.; Kinoshita, T.; Prestegard, J. J.; Schnaar, R. L.; Freeze, H. H.; Marth, J. D.; Bertozzi, C. R.; Etzler, M. E.; Frank, M.; Vliegenthart, J. F.; Lütteke, T.; Perez, S.; Bolton, E.; Rudd, P.; Paulson, J.; Kanehisa, M.; Toukach, P.; Aoki-Kinoshita, K. F.; Dell, A.; Narimatsu, H.; York, W.; Taniguchi, N.; Kornfeld, S. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **2015**, *25* (12), 1323–1324. https://doi.org/10.1093/glycob/cwv091.

(10) Neelamegham, S.; Aoki-Kinoshita, K.; Bolton, E.; Frank, M.; Lisacek, F.; Lütteke, T.; O'Boyle, N.; Packer, N. H.; Stanley, P.; Toukach, P.; Varki, A.; Woods, R. J.; The SNFG Discussion Group. Updates to the Symbol Nomenclature for Glycans Guidelines. *Glycobiology* **2019**, *29* (9), 620–624. https://doi.org/10.1093/glycob/cwz045.

(11) Fry, S. C.; York, W. S.; Albersheim, P.; Darvill, A.; Hayashi, T.; Joseleau, J.-P.; Kato, Y.; Lorences, E. P.; Maclachlan, G. A.; McNeil, M.; Mort, A. J.; Grant Reid, J. S.; Seitz, H. U.; Selvendran, R. R.; Voragen, A. G. J.; White, A. R. An Unambiguous Nomenclature for Xyloglucan-Derived Oligosaccharides. *Physiol. Plant.* **1993**, *89* (1), 1–3. https://doi.org/10.1111/j.1399-3054.1993.tb01778.x.

(12) Willats, W. G. T.; McCartney, L.; Mackie, W.; Knox, J. P. Pectin: Cell Biology and Prospects for Functional Analysis. *Plant Mol. Biol.* **2001**, *47* (1), 9–27. https://doi.org/10.1023/A:1010662911148.

(13) Buffetto, F.; Ropartz, D.; Zhang, X. J.; Gilbert, H. J.; Guillon, F.; Ralet, M.-C. Recovery and Fine Structure Variability of RGII Sub-Domains in Wine (Vitis Vinifera Merlot). *Ann. Bot.* **2014**, *114* (6), 1327–1337. https://doi.org/10.1093/aob/mcu097.

(14) Busse-Wicher, M.; Li, A.; Silveira, R. L.; Pereira, C. S.; Tryfona, T.; Gomes, T. C. F.; Skaf, M. S.; Dupree, P. Evolution of Xylan Substitution Patterns in Gymnosperms and Angiosperms: Implications for Xylan Interaction with Cellulose. *Plant Physiol.* **2016**, *171* (4), 2418–2431. https://doi.org/10.1104/pp.16.00539.
(15) Caffall, K. H.; Mohnen, D. The Structure, Function, and Biosynthesis of Plant Cell Wall Pectic Polysaccharides. *Carbohydr. Res.* **2009**, *344* (14), 1879–1900. https://doi.org/10.1016/j.carres.2009.05.021.
(16) Ferreira, S. S.; Passos, C. P.; Madureira, P.; Vilanova, M.; Coimbra, M. A. Structure–Function Relationships of Immunostimulatory Polysaccharides: A Review. *Carbohydr. Polym.* **2015**, *132*, 378–396. https://doi.org/10.1016/j.carbpol.2015.05.079.
(17) Gray, C. J.; Migas, L. G.; Barran, P. E.; Pagel, K.; Seeberger, P. H.; Eyers, C. E.; Boons, G.-J.; Pohl, N. L. B.; Compagnon, I.; Widmalm, G.; Flitsch, S. L. Advancing Solutions to the Carbohydrate Sequencing Challenge. *J. Am. Chem. Soc.* **2019**, *141* (37), 14463–14479. https://doi.org/10.1021/jacs.9b06406.
(18) Duus, J. Ø.; Gotfredsen, C. H.; Bock, K. Carbohydrate Structural Determination by NMR Spectroscopy: Modern Methods and Limitations. *Chem. Rev.* **2000**, *100* (12), 4589–4614. https://doi.org/10.1021/cr990302n.
(19) Yao, H.-Y.-Y.; Wang, J.-Q.; Yin, J.-Y.; Nie, S.-P.; Xie, M.-Y. A Review of NMR Analysis in Polysaccharide Structure and Conformation: Progress, Challenge and Perspective. *Food Res. Int.* **2021**, *143*, 110290. https://doi.org/10.1016/j.foodres.2021.110290.
(20) Singh, K.; Jacquemmoz, C.; Giraudeau, P.; Frydman, L.; Dumez, J.-N. Ultrafast 2D 1 H–1 H NMR Spectroscopy of DNP-Hyperpolarised Substrates for the Analysis of Mixtures. *Chem. Commun.* **2021**, *57* (65), 8035–8038. https://doi.org/10.1039/D1CC03079E.
(21) Kuberan, B.; Lech, M.; Zhang, L.; Wu, Z. L.; Beeler, D. L.; Rosenberg, R. D. Analysis of Heparan Sulfate Oligosaccharides with Ion Pair-Reverse Phase Capillary High Performance Liquid Chromatography-Microelectrospray Ionization Time-of-Flight Mass Spectrometry. *J. Am. Chem. Soc.* **2002**, *124* (29), 8707–8718. https://doi.org/10.1021/ja0178867.
(22) Mendis, P. M.; Sasiene, Z. J.; Ropartz, D.; Rogniaux, H.; Jackson, G. P. Ultra-High-Performance Liquid Chromatography Charge Transfer Dissociation Mass Spectrometry (UHPLC-CTD-MS) as a Tool for Analyzing the Structural Heterogeneity in Carrageenan Oligosaccharides. *Anal. Bioanal. Chem.* **2022**, *414* (1), 303–318. https://doi.org/10.1007/s00216-021-03396-3.
(23) West, C.; Elfakir, C.; Lafosse, M. Porous Graphitic Carbon: A Versatile Stationary Phase for Liquid Chromatography. *J. Chromatogr. A* **2010**, *1217* (19), 3201–3216. https://doi.org/10.1016/j.chroma.2009.09.052.
(24) Deslignière, E.; Ley, M.; Bourguet, M.; Ehkirch, A.; Botzanowski, T.; Erb, S.; Hernandez-Alba, O.; Cianférani, S. Pushing the Limits of Native MS: Online SEC-Native MS for Structural Biology Applications. *Int. J. Mass Spectrom.* **2021**, *461*, 116502. https://doi.org/10.1016/j.ijms.2020.116502.
(25) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* **1989**, *246* (4926), 64–71. https://doi.org/10.1126/science.2675315.
(26) Karas, Michael.; Hillenkamp, Franz. Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10,000 Daltons. *Anal. Chem.* **1988**, *60* (20), 2299–2301. https://doi.org/10.1021/ac00171a028.
(27) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. Protein and Polymer Analyses up to m/z 100 000 by Laser Ionization Time-of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1988**, *2* (8), 151–153. https://doi.org/10.1002/rcm.1290020802.
(28) Schwamborn, K.; Caprioli, R. M. Molecular Imaging by Mass Spectrometry — Looking beyond Classical Histology. *Nat. Rev. Cancer* **2010**, *10* (9), 639–646. https://doi.org/10.1038/nrc2917.
(29) Arnaud, B.; Durand, S.; Fanuel, M.; Guillon, F.; Méchin, V.; Rogniaux, H. Imaging Study by Mass Spectrometry of the Spatial Variation of Cellulose and Hemicellulose Structures

in Corn Stalks. *J. Agric. Food Chem.* **2020**, *68* (13), 4042–4050. https://doi.org/10.1021/acs.jafc.9b07579.
(30) Fanuel, M.; Ropartz, D.; Guillon, F.; Saulnier, L.; Rogniaux, H. Distribution of Cell Wall Hemicelluloses in the Wheat Grain Endosperm: A 3D Perspective. *Planta* **2018**, *248* (6), 1505–1513. https://doi.org/10.1007/s00425-018-2980-0.
(31) Domon, B.; Costello, C. E. A Systematic Nomenclature for Carbohydrate Fragmentations in FAB-MS MS Spectra of Glycoconjugates. *Glycoconj. J.* **1988**, *5* (4), 397–409. https://doi.org/10.1007/bf01049915.
(32) Bayat, P.; Lesage, D.; Cole, R. B. Tutorial: Ion Activation in Tandem Mass Spectrometry Using Ultra-High Resolution Instrumentation. *Mass Spectrom. Rev.* **2020**, *39* (5–6), 680–702. https://doi.org/10.1002/mas.21623.
(33) Leach, F. E.; Wolff, J. J.; Laremore, T. N.; Linhardt, R. J.; Amster, I. J. Evaluation of the Experimental Parameters Which Control Electron Detachment Dissociation, and Their Effect on the Fragmentation Efficiency of Glycosaminoglycan Carbohydrates. *Int. J. Mass Spectrom.* **2008**, *276* (2), 110–115. https://doi.org/10.1016/j.ijms.2008.05.017.
(34) Yu, X.; Huang, Y.; Lin, C.; Costello, C. E. Energy-Dependent Electron Activated Dissociation of Metal-Adducted Permethylated Oligosaccharides. *Anal. Chem.* **2012**, *84* (17), 7487–7494. https://doi.org/10.1021/ac301589z.
(35) Wolff, J. J.; Leach, F. E.; Laremore, T. N.; Kaplan, D. A.; Easterling, M. L.; Linhardt, R. J.; Amster, I. J. Negative Electron Transfer Dissociation of Glycosaminoglycans. *Anal. Chem.* **2010**, *82* (9), 3460–3466. https://doi.org/10.1021/ac100554a.
(36) Lee, S.; Valentine, S. J.; Reilly, J. P.; Clemmer, D. E. Analyzing a Mixture of Disaccharides by IMS-VUVPD-MS. *Int. J. Mass Spectrom.* **2012**, *309*, 161–167. https://doi.org/10.1016/j.ijms.2011.09.013.
(37) Brodbelt, J. S.; Morrison, L. J.; Santos, I. Ultraviolet Photodissociation Mass Spectrometry for Analysis of Biological Molecules. *Chem. Rev.* **2020**, *120* (7), 3328–3380. https://doi.org/10.1021/acs.chemrev.9b00440.
(38) Ropartz, D.; Lemoine, J.; Giuliani, A.; Bittebière, Y.; Enjalbert, Q.; Antoine, R.; Dugourd, P.; Ralet, M.-C.; Rogniaux, H. Deciphering the Structure of Isomeric Oligosaccharides in a Complex Mixture by Tandem Mass Spectrometry: Photon Activation with Vacuum Ultra-Violet Brings Unique Information and Enables Definitive Structure Assignment. *Anal. Chim. Acta* **2014**, *807*, 84–95. https://doi.org/10.1016/j.aca.2013.11.018.
(39) Ropartz, D.; Giuliani, A.; Fanuel, M.; Hervé, C.; Czjzek, M.; Rogniaux, H. Online Coupling of High-Resolution Chromatography with Extreme UV Photon Activation Tandem Mass Spectrometry: Application to the Structural Investigation of Complex Glycans by Dissociative Photoionization. *Anal. Chim. Acta* **2016**, *933*, 1–9. https://doi.org/10.1016/j.aca.2016.05.036.
(40) Ropartz, D.; Li, P.; Fanuel, M.; Giuliani, A.; Rogniaux, H.; Jackson, G. P. Charge Transfer Dissociation of Complex Oligosaccharides: Comparison with Collision-Induced Dissociation and Extreme Ultraviolet Dissociative Photoionization. *J. Am. Soc. Mass Spectrom.* **2016**, *27* (10), 1614–1619. https://doi.org/10.1021/jasms.8b05150.
(41) Buck-Wiese, H.; Fanuel, M.; Liebeke, M.; Hoang, K. L. M.; Pardo-Vargas, A.; Seeberger, P. H.; Hehemann, J.-H.; Rogniaux, H.; Jackson, G. P.; Ropartz, D. Discrimination of Beta-1,4-and Beta-1,3-Linkages in Native Oligosaccharides via Charge Transfer Dissociation Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (6), 1249–1259. https://doi.org/10.1021/jasms.0c00087.
(42) Schindler, B.; Barnes, L.; Renois, G.; Gray, C.; Chambert, S.; Fort, S.; Flitsch, S.; Loison, C.; Allouche, A.-R.; Compagnon, I. Anomeric Memory of the Glycosidic Bond upon Fragmentation and Its Consequences for Carbohydrate Sequencing. *Nat. Commun.* **2017**, *8* (1), 973. https://doi.org/10.1038/s41467-017-01179-y.
(43) Schindler, B.; Legentil, L.; Allouche, A. R.; Ferrieres, V.; Compagnon, I. Spectroscopic Diagnostic for the Ring-Size of Carbohydrates in the Gas Phase: Furanose and Pyranose Forms of GalNAc. *Phys. Chem. Chem. Phys.* **2019**, *21* (23), 12460–12467. https://doi.org/10.1039/c8cp04082f.

(44) Rutherford, E. LIV. *The Velocity and Rate of Recombination of the Ions of Gases Exposed to Röntgen Radiation*. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1897**, *44* (270), 422–440. https://doi.org/10.1080/14786449708621085.
(45) Thomson, J. J. XIX. *Further Experiments on Positive Rays*. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1912**, *24* (140), 209–253. https://doi.org/10.1080/14786440808637325.
(46) Langevin, P. L'ionization Des Gaz. *Ann Chim Phys* **1903**.
(47) Langevin, P. Recombinaison et Mobilités Des Ions Dans Les Gaz. *Ann Chim Phys* **1903**.
(48) Mason, E. A.; Schamp, H. W. Mobility of Gaseous Lons in Weak Electric Fields. *Ann. Phys.* **1958**, *4* (3), 233–270. https://doi.org/10.1016/0003-4916(58)90049-6.
(49) McDaniel, E. W. *Collision Phenomena in Ionized Gases*; 1964.
(50) McDaniel, E. W.; Martin, D. W.; Barnes, W. S. Drift Tube-Mass Spectrometer for Studies of Low-Energy Ion-Molecule Reactions. *Rev. Sci. Instrum.* **1962**, *33* (1), 2–7. https://doi.org/10.1063/1.1717656.
(51) Barnes, W. S.; Martin, D. W.; McDaniel, E. W. Mass Spectrographic Identification of the Ion Observed in Hydrogen Mobility Experiments. *Phys. Rev. Lett.* **1961**, *6* (3), 110–111. https://doi.org/10.1103/PhysRevLett.6.110.
(52) McAfee, K. B.; Edelson, D. Identification and Mobility of Ions in a Townsend Discharge by Time-Resolved Mass Spectrometry. *Proc. Phys. Soc.* **1963**, *81* (2), 382–384. https://doi.org/10.1088/0370-1328/81/2/125.
(53) Cohen, M. J.; Karasek, F. W. Plasma Chromatography—A New Dimension for Gas Chromatography and Mass Spectrometry. *J. Chromatogr. Sci.* **1970**, *8* (6), 330–337. https://doi.org/10.1093/chromsci/8.6.330.
(54) Shaffer, S. A.; Prior, D. C.; Anderson, G. A.; Udseth, H. R.; Smith, R. D. An Ion Funnel Interface for Improved Ion Focusing and Sensitivity Using Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **1998**, *70* (19), 4111–4119. https://doi.org/10.1021/ac9802170.
(55) Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. Naked Protein Conformations: Cytochrome c in the Gas Phase. *J. Am. Chem. Soc.* **1995**, *117* (40), 10141–10142. https://doi.org/10.1021/ja00145a037.
(56) Dugourd, Ph.; Hudgins, R. R.; Clemmer, D. E.; Jarrold, M. F. High-Resolution Ion Mobility Measurements. *Rev. Sci. Instrum.* **1997**, *68* (2), 1122–1129. https://doi.org/10.1063/1.1147873.
(57) Kemper, P. R.; Bowers, M. T. A Hybrid Double-Focusing Mass Spectrometer—High-Pressure Drift Reaction Cell to Study Thermal Energy Reactions of Mass-Selected Ions. *J. Am. Soc. Mass Spectrom.* **1990**, *1* (3), 197–207. https://doi.org/10.1016/1044-0305(90)85036-L.
(58) von Helden, G.; Wyttenbach, T.; Bowers, M. T. Conformation of Macromolecules in the Gas Phase: Use of Matrix-Assisted Laser Desorption Methods in Ion Chromatography. *Science* **1995**, *267* (5203), 1483–1485. https://doi.org/10.1126/science.267.5203.1483.
(59) Wu, C.; Siems, W. F.; Asbury, G. R.; Hill, H. H. Electrospray Ionization High-Resolution Ion Mobility Spectrometry−Mass Spectrometry. *Anal. Chem.* **1998**, *70* (23), 4929–4938. https://doi.org/10.1021/ac980414z.
(60) Wittmer, Doug.; Chen, Y. Hong.; Luckenbill, B. K.; Hill, H. H. Electrospray Ionization Ion Mobility Spectrometry. *Anal. Chem.* **1994**, *66* (14), 2348–2355. https://doi.org/10.1021/ac00086a021.
(61) Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H. Applications of a Travelling Wave-Based Radio-Frequency-Only Stacked Ring Ion Guide. *Rapid Commun. Mass Spectrom.* **2004**, *18* (20), 2401–2414. https://doi.org/10.1002/rcm.1641.
(62) Pringle, S. D.; Giles, K.; Wildgoose, J. L.; Williams, J. P.; Slade, S. E.; Thalassinos, K.; Bateman, R. H.; Bowers, M. T.; Scrivens, J. H. 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.* **2007**, *261* (1), 1–12. https://doi.org/10.1016/j.ijms.2006.07.021.

(63) *Web of Science Core Collection*. https://www.webofscience.com/wos/woscc/analyze-results/3e1305fb-e212-4a82-aee8-8c0ac92279ec-46c9f811 (accessed 2022-08-02).
(64) Gabelica, V.; Marklund, E. Fundamentals of Ion Mobility Spectrometry. *Curr. Opin. Chem. Biol.* **2018**, *42*, 51–59. https://doi.org/10.1016/j.cbpa.2017.10.022.
(65) Gabelica, V.; Shvartsburg, A. A.; Afonso, C.; Barran, P.; Benesch, J. L. P.; Bleiholder, C.; Bowers, M. T.; Bilbao, A.; Bush, M. F.; Campbell, J. L.; Campuzano, I. D. G.; Causon, T.; Clowers, B. H.; Creaser, C. S.; De Pauw, E.; Far, J.; Fernandez-Lima, F.; Fjeldsted, J. C.; Giles, K.; Groessl, M.; Hogan, C. J.; Hann, S.; Kim, H. I.; Kurulugama, R. T.; May, J. C.; McLean, J. A.; Pagel, K.; Richardson, K.; Ridgeway, M. E.; Rosu, F.; Sobott, F.; Thalassinos, K.; Valentine, S. J.; Wyttenbach, T. Recommendations for Reporting Ion Mobility Mass Spectrometry Measurements. *Mass Spectrom. Rev.* **2019**, *38* (3), 291–320. https://doi.org/10.1002/mas.21585.
(66) May, J. C.; Goodwin, C. R.; Lareau, N. M.; Leaptrot, K. L.; Morris, C. B.; Kurulugama, R. T.; Mordehai, A.; Klein, C.; Barry, W.; Darland, E.; Overney, G.; Imatani, K.; Stafford, G. C.; Fjeldsted, J. C.; McLean, J. A. Conformational Ordering of Biomolecules in the Gas Phase: Nitrogen Collision Cross Sections Measured on a Prototype High Resolution Drift Tube Ion Mobility-Mass Spectrometer. *Anal. Chem.* **2014**, *86* (4), 2107–2116. https://doi.org/10.1021/ac4038448.
(67) Eiceman, G. A.; Karpas, Z.; Jr, H. H. H. *Ion Mobility Spectrometry, Third Edition*; CRC Press, 2013.
(68) Revercomb, H. E.; Mason, E. A. Theory of Plasma Chromatography/Gaseous Electrophoresis. Review. *Anal. Chem.* **1975**, *47* (7), 970–983. https://doi.org/10.1021/ac60357a043.
(69) BIPM. *1st Meeting. Com. Cons. Quantité de Matière*; 1995; p Q5.
(70) May, J. C.; McLean, J. A. The Influence of Drift Gas Composition on the Separation Mechanism in Traveling Wave Ion Mobility Spectrometry: Insight from Electrodynamic Simulations. *Int. J. Ion Mobil. Spectrom.* **2013**, *16* (2), 85–94. https://doi.org/10.1007/s12127-013-0123-7.
(71) Richardson, K.; Langridge, D.; Dixit, S. M.; Ruotolo, B. T. An Improved Calibration Approach for Traveling Wave Ion Mobility Spectrometry: Robust, High-Precision Collision Cross Sections. *Anal. Chem.* **2021**, *93* (7), 3542–3550. https://doi.org/10.1021/acs.analchem.0c04948.
(72) Fernandez-Lima, F.; Kaplan, D. A.; Suetering, J.; Park, M. A. Gas-Phase Separation Using a Trapped Ion Mobility Spectrometer. *Int. J. Ion Mobil. Spectrom.* **2011**, *14* (2), 93–98. https://doi.org/10.1007/s12127-011-0067-8.
(73) Fernandez-Lima, F. A.; Kaplan, D. A.; Park, M. A. Note: Integration of Trapped Ion Mobility Spectrometry with Mass Spectrometry. *Rev. Sci. Instrum.* **2011**, *82* (12), 126106. https://doi.org/10.1063/1.3665933.
(74) Michelmann, K.; Silveira, J. A.; Ridgeway, M. E.; Park, M. A. Fundamentals of Trapped Ion Mobility Spectrometry. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (1), 14–24. https://doi.org/10.1007/s13361-014-0999-4.
(75) Hernandez, D. R.; DeBord, J. D.; Ridgeway, M. E.; Kaplan, D. A.; Park, M. A.; Fernandez-Lima, F. Ion Dynamics in a Trapped Ion Mobility Spectrometer. *Analyst* **2014**, *139* (8), 1913–1921. https://doi.org/10.1039/C3AN02174B.
(76) Meier, F.; Brunner, A.-D.; Koch, S.; Koch, H.; Lubeck, M.; Krause, M.; Goedecke, N.; Decker, J.; Kosinski, T.; Park, M. A.; Bache, N.; Hoerning, O.; Cox, J.; Räther, O.; Mann, M. Online Parallel Accumulation–Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer *. *Mol. Cell. Proteomics* **2018**, *17* (12), 2534–2545. https://doi.org/10.1074/mcp.TIR118.000900.
(77) Rosell-Llompart, J.; Loscertales, I. G.; Bingham, D.; Fernández de la Mora, J. Sizing Nanoparticles and Ions with a Short Differential Mobility Analyzer. *J. Aerosol Sci.* **1996**, *27* (5), 695–719. https://doi.org/10.1016/0021-8502(96)00016-X.
(78) Flagan, R. C. On Differential Mobility Analyzer Resolution. *Aerosol Sci. Technol.* **1999**, *30* (6), 556–570. https://doi.org/10.1080/027868299304417.

(79) Santos, J. P.; Hontañón, E.; Ramiro, E.; Alonso, M. Performance Evaluation of a High-Resolution Parallel-Plate Differential Mobility Analyzer. *Atmospheric Chem. Phys.* **2009**, *9* (7), 2419–2429. https://doi.org/10.5194/acp-9-2419-2009.
(80) Guha, S.; Li, M.; Tarlov, M. J.; Zachariah, M. R. Electrospray–Differential Mobility Analysis of Bionanoparticles. *Trends Biotechnol.* **2012**, *30* (5), 291–300. https://doi.org/10.1016/j.tibtech.2012.02.003.
(81) Pease, L. F. Physical Analysis of Virus Particles Using Electrospray Differential Mobility Analysis. *Trends Biotechnol.* **2012**, *30* (4), 216–224. https://doi.org/10.1016/j.tibtech.2011.11.004.
(82) Guevremont, R. High-Field Asymmetric Waveform Ion Mobility Spectrometry: A New Tool for Mass Spectrometry. *J. Chromatogr. A* **2004**, *1058* (1), 3–19. https://doi.org/10.1016/j.chroma.2004.08.119.
(83) Guevremont, R.; Thekkadath, G.; Hilton, C. K. Compensation Voltage (CV) Peak Shapes Using a Domed FAIMS with the Inner Electrode Translated to Various Longitudinal Positions. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (6), 948–956. https://doi.org/10.1016/j.jasms.2005.03.001.
(84) Mabrouki, R.; Kelly, R. T.; Prior, D. C.; Shvartsburg, A. A.; Tang, K.; Smith, R. D. Improving FAIMS Sensitivity Using a Planar Geometry with Slit Interfaces. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (9), 1768–1774. https://doi.org/10.1016/j.jasms.2009.05.019.
(85) Delvaux, A.; Rathahao-Paris, E.; Alves, S. Different Ion Mobility-Mass Spectrometry Coupling Techniques to Promote Metabolomics. *Mass Spectrom. Rev.* **2021**, *n/a* (n/a). https://doi.org/10.1002/mas.21685.
(86) Li, H.; Smith, B. K.; Márk, L.; Nemes, P.; Nazarian, J.; Vertes, A. Ambient Molecular Imaging by Laser Ablation Electrospray Ionization Mass Spectrometry with Ion Mobility Separation. *Int. J. Mass Spectrom.* **2015**, *377*, 681–689. https://doi.org/10.1016/j.ijms.2014.06.025.
(87) Sukumar, H.; Stone, J. A.; Nishiyama, T.; Yuan, C.; Eiceman, G. A. Paper Spray Ionization with Ion Mobility Spectrometry at Ambient Pressure. *Int. J. Ion Mobil. Spectrom.* **2011**, *14* (2), 51. https://doi.org/10.1007/s12127-011-0069-6.
(88) Izquierdo-Sandoval, D.; Fabregat-Safont, D.; Lacalle-Bergeron, L.; Sancho, J. V.; Hernández, F.; Portoles, T. Benefits of Ion Mobility Separation in GC-APCI-HRMS Screening: From the Construction of a CCS Library to the Application to Real-World Samples. *Anal. Chem.* **2022**, *94* (25), 9040–9047. https://doi.org/10.1021/acs.analchem.2c01118.
(89) Sabo, M.; Matejčík, Š. Corona Discharge Ion Mobility Spectrometry with Orthogonal Acceleration Time of Flight Mass Spectrometry for Monitoring of Volatile Organic Compounds. *Anal. Chem.* **2012**, *84* (12), 5327–5334. https://doi.org/10.1021/ac300722s.
(90) Yamaguchi, S.; Asada, R.; Kishi, S.; Sekioka, R.; Kitagawa, N.; Tokita, K.; Yamamoto, S.; Seto, Y. Detection Performance of a Portable Ion Mobility Spectrometer with 63Ni Radioactive Ionization for Chemical Warfare Agents. *Forensic Toxicol.* **2010**, *28* (2), 84–95. https://doi.org/10.1007/s11419-010-0092-z.
(91) Saba, J.; Bonneil, E.; Pomiès, C.; Eng, K.; Thibault, P. Enhanced Sensitivity in Proteomics Experiments Using FAIMS Coupled with a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer. *J. Proteome Res.* **2009**, *8* (7), 3355–3366. https://doi.org/10.1021/pr801106a.
(92) Robinson, E. W.; Garcia, D. E.; Leib, R. D.; Williams, E. R. Enhanced Mixture Analysis of Poly(Ethylene Glycol) Using High-Field Asymmetric Waveform Ion Mobility Spectrometry Combined with Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2006**, *78* (7), 2190–2198. https://doi.org/10.1021/ac051709x.
(93) Ibrahim, Y. M.; Garimella, S. V. B.; Prost, S. A.; Wojcik, R.; Norheim, R. V.; Baker, E. S.; Rusyn, I.; Smith, R. D. Development of an Ion Mobility Spectrometry-Orbitrap Mass Spectrometer Platform. *Anal. Chem.* **2016**, *88* (24), 12152–12160. https://doi.org/10.1021/acs.analchem.6b03027.
(94) Ridgeway, M. E.; Wolff, J. J.; Silveira, J. A.; Lin, C.; Costello, C. E.; Park, M. A. Gated Trapped Ion Mobility Spectrometry Coupled to Fourier Transform Ion Cyclotron

Resonance Mass Spectrometry. *Int. J. Ion Mobil. Spectrom.* **2016**, *19* (2), 77–85. https://doi.org/10.1007/s12127-016-0197-0.

(95) Benigni, P.; Fernandez-Lima, F. Oversampling Selective Accumulation Trapped Ion Mobility Spectrometry Coupled to FT-ICR MS: Fundamentals and Applications. *Anal. Chem.* **2016**, *88* (14), 7404–7412. https://doi.org/10.1021/acs.analchem.6b01946.

(96) Benigni, P.; Porter, J.; Ridgeway, Mark. E.; Park, Melvin. A.; Fernandez-Lima, F. Increasing Analytical Separation and Duty Cycle with Nonlinear Analytical Mobility Scan Functions in TIMS-FT-ICR MS. *Anal. Chem.* **2018**, *90* (4), 2446–2450. https://doi.org/10.1021/acs.analchem.7b04053.

(97) Poltash, M. L.; McCabe, J. W.; Shirzadeh, M.; Laganowsky, A.; Clowers, B. H.; Russell, D. H. Fourier Transform-Ion Mobility-Orbitrap Mass Spectrometer: A Next-Generation Instrument for Native Mass Spectrometry. *Anal. Chem.* **2018**, *90* (17), 10472–10478. https://doi.org/10.1021/acs.analchem.8b02463.

(98) Dodds, J. N.; May, J. C.; McLean, J. A. Investigation of the Complete Suite of the Leucine and Isoleucine Isomers: Toward Prediction of Ion Mobility Separation Capabilities. *Anal. Chem.* **2017**, *89* (1), 952–959. https://doi.org/10.1021/acs.analchem.6b04171.

(99) Adams, K. J.; Montero, D.; Aga, D.; Fernandez-Lima, F. Isomer Separation of Polybrominated Diphenyl Ether Metabolites Using NanoESI-TIMS-MS. *Int. J. Ion Mobil. Spectrom.* **2016**, *19* (2), 69–76. https://doi.org/10.1007/s12127-016-0198-z.

(100) Shvartsburg, A. A.; Smith, R. D. Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Anal. Chem.* **2008**, *80* (24), 9689–9699. https://doi.org/10.1021/ac8016295.

(101) Merenbloom, S. I.; Glaskin, R. S.; Henson, Z. B.; Clemmer, D. E. High-Resolution Ion Cyclotron Mobility Spectrometry. *Anal. Chem.* **2009**, *81* (4), 1482–1487. https://doi.org/10.1021/ac801880a.

(102) Giles, K.; Ujma, J.; Wildgoose, J.; Pringle, S.; Richardson, K.; Langridge, D.; Green, M. A Cyclic Ion Mobility-Mass Spectrometry System. *Anal. Chem.* **2019**, *91* (13), 8564–8573. https://doi.org/10.1021/acs.analchem.9b01838.

(103) Ropartz, D.; Fanuel, M.; Ujma, J.; Palmer, M.; Giles, K.; Rogniaux, H. Structure Determination of Large Isomeric Oligosaccharides of Natural Origin through Multipass and Multistage Cyclic Traveling-Wave Ion Mobility Mass Spectrometry. *Anal. Chem.* **2019**, *91* (18), 12030–12037. https://doi.org/10.1021/acs.analchem.9b03036.

(104) Tolmachev, A. V.; Webb, I. K.; Ibrahim, Y. M.; Garimella, S. V. B.; Zhang, X.; Anderson, G. A.; Smith, R. D. Characterization of Ion Dynamics in Structures for Lossless Ion Manipulations. *Anal. Chem.* **2014**, *86* (18), 9162–9168. https://doi.org/10.1021/ac502054p.

(105) Deng, L.; Ibrahim, Y. M.; Hamid, A. M.; Garimella, S. V. B.; Webb, I. K.; Zheng, X.; Prost, S. A.; Sandoval, J. A.; Norheim, R. V.; Anderson, G. A.; Tolmachev, A. V.; Baker, E. S.; Smith, R. D. Ultra-High Resolution Ion Mobility Separations Utilizing Traveling Waves in a 13 m Serpentine Path Length Structures for Lossless Ion Manipulations Module. *Anal. Chem.* **2016**, *88* (18), 8957–8964. https://doi.org/10.1021/acs.analchem.6b01915.

(106) Deng, L.; Webb, I. K.; Garimella, S. V. B.; Hamid, A. M.; Zheng, X.; Norheim, R. V.; Prost, S. A.; Anderson, G. A.; Sandoval, J. A.; Baker, E. S.; Ibrahim, Y. M.; Smith, R. D. Serpentine Ultralong Path with Extended Routing (SUPER) High Resolution Traveling Wave Ion Mobility-MS Using Structures for Lossless Ion Manipulations. *Anal. Chem.* **2017**, *89* (8), 4628–4634. https://doi.org/10.1021/acs.analchem.7b00185.

(107) Clowers, B. H.; Cabrera, E.; Anderson, G.; Deng, L.; Moser, K.; Van Aken, G.; DeBord, J. D. Masked Multiplexed Separations to Enhance Duty Cycle for Structures for Lossless Ion Manipulations. *Anal. Chem.* **2021**, *93* (14), 5727–5734. https://doi.org/10.1021/acs.analchem.0c04799.

(108) Ibrahim, Y. M.; Hamid, A. M.; Cox, J. T.; Garimella, S. V. B.; Smith, R. D. Ion Elevators and Escalators in Multilevel Structures for Lossless Ion Manipulations. *Anal. Chem.* **2017**, *89* (3), 1972–1977. https://doi.org/10.1021/acs.analchem.6b04500.

(109) Kinlein, Z. R.; Anderson, G. A.; Clowers, B. H. Accelerating Prototyping Experiments for Traveling Wave Structures for Lossless Ion Manipulations. *Talanta* **2022**, *244*, 123446. https://doi.org/10.1016/j.talanta.2022.123446.
(110) Gray, C. J.; Schindler, B.; Migas, L. G.; Pičmanová, M.; Allouche, A. R.; Green, A. P.; Mandal, S.; Motawia, M. S.; Sánchez-Pérez, R.; Bjarnholt, N.; Møller, B. L.; Rijs, A. M.; Barran, P. E.; Compagnon, I.; Eyers, C. E.; Flitsch, S. L. Bottom-Up Elucidation of Glycosidic Bond Stereochemistry. *Anal. Chem.* **2017**, *89* (8), 4540–4549. https://doi.org/10.1021/acs.analchem.6b04998.
(111) Hofmann, J.; Hahm, H. S.; Seeberger, P. H.; Pagel, K. Identification of Carbohydrate Anomers Using Ion Mobility–Mass Spectrometry. *Nature* **2015**, *526* (7572), 241–244. https://doi.org/10.1038/nature15388.
(112) Struwe, W. B.; Pagel, K.; Benesch, J. L. P.; Harvey, D. J.; Campbell, M. P. GlycoMob: An Ion Mobility-Mass Spectrometry Collision Cross Section Database for Glycomics. *Glycoconj. J.* **2016**, *33* (3), 399–404. https://doi.org/10.1007/s10719-015-9613-7.
(113) Huang, Y.; Dodds, E. D. Ion Mobility Studies of Carbohydrates as Group I Adducts: Isomer Specific Collisional Cross Section Dependence on Metal Ion Radius. *Anal. Chem.* **2013**, *85* (20), 9728–9735. https://doi.org/10.1021/ac402133f.
(114) Huang, Y.; Dodds, E. D. Discrimination of Isomeric Carbohydrates as the Electron Transfer Products of Group II Cation Adducts by Ion Mobility Spectrometry and Tandem Mass Spectrometry. *Anal. Chem.* **2015**, *87* (11), 5664–5668. https://doi.org/10.1021/acs.analchem.5b00759.
(115) Deng, L. L.; Ibrahim, Y. M.; Baker, E. S.; Aly, N. A.; Hamid, A. M.; Zhang, X.; Zheng, X. Y.; Garimella, S. V. B.; Webb, I. K.; Prost, S. A.; Sandoval, J. A.; Norheim, R. V.; Anderson, G. A.; Tolmachev, A. V.; Smith, R. D. Ion Mobility Separations of Isomers Based upon Long Path Length Structures for Lossless Ion Manipulations Combined with Mass Spectrometry. *Chemistryselect* **2016**, *1* (10), 2396–2399. https://doi.org/10.1002/slct.201600460.
(116) Nagy, G.; Attah, I. K.; Garimella, S. V. B.; Tang, K.; Ibrahim, Y. M.; Baker, E. S.; Smith, R. D. Unraveling the Isomeric Heterogeneity of Glycans: Ion Mobility Separations in Structures for Lossless Ion Manipulations. *Chem. Commun.* **2018**, *54* (83), 11701–11704. https://doi.org/10.1039/C8CC06966B.
(117) Ujma, J.; Ropartz, D.; Giles, K.; Richardson, K.; Langridge, D.; Wildgoose, J.; Green, M.; Pringle, S. Cyclic Ion Mobility Mass Spectrometry Distinguishes Anomers and Open-Ring Forms of Pentasaccharides. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (6), 1028–1037. https://doi.org/10.1007/s13361-019-02168-9.
(118) Ben Faleh, A.; Warnke, S.; Rizzo, T. R. Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures. *Anal. Chem.* **2019**, *91* (7), 4876–4882. https://doi.org/10.1021/acs.analchem.9b00659.
(119) Dyukova, I.; Faleh, A. B.; Warnke, S.; Yalovenko, N.; Yatsyna, V.; Bansal, P.; Rizzo, T. R. A New Approach for Identifying Positional Isomers of Glycans Cleaved from Monoclonal Antibodies. *Analyst* **2021**, *146* (15), 4789–4795. https://doi.org/10.1039/D1AN00780G.
(120) Abikhodr, A. H.; Yatsyna, V.; Ben Faleh, A.; Warnke, S.; Rizzo, T. R. Identifying Mixtures of Isomeric Human Milk Oligosaccharides by the Decomposition of IR Spectral Fingerprints. *Anal. Chem.* **2021**, *93* (44), 14730–14736. https://doi.org/10.1021/acs.analchem.1c03190.
(121) Warnke, S.; Faleh, A. B.; Scutelnic, V.; Rizzo, T. R. Separation and Identification of Glycan Anomers Using Ultrahigh-Resolution Ion-Mobility Spectrometry and Cryogenic Ion Spectroscopy. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (11), 2204–2211. https://doi.org/10.1007/s13361-019-02333-0.
(122) Pellegrinelli, R. P.; Yue, L.; Carrascosa, E.; Warnke, S.; Ben Faleh, A.; Rizzo, T. R. How General Is Anomeric Retention during Collision-Induced Dissociation of Glycans? *J. Am. Chem. Soc.* **2020**, *142* (13), 5948–5951. https://doi.org/10.1021/jacs.0c00264.
(123) Bansal, P.; Yatsyna, V.; AbiKhodr, A. H.; Warnke, S.; Ben Faleh, A.; Yalovenko, N.; Wysocki, V. H.; Rizzo, T. R. Using SLIM-Based IMS-IMS Together with Cryogenic

Infrared Spectroscopy for Glycan Analysis. *Anal. Chem.* **2020**, *92* (13), 9079–9085. https://doi.org/10.1021/acs.analchem.0c01265.

(124) Bansal, P.; Faleh, A. B.; Warnke, S.; Rizzo, T. R. Identification of N-Glycan Positional Isomers by Combining IMS and Vibrational Fingerprinting of Structurally Determinant CID Fragments. *Analyst* **2022**, *147* (4), 704–711. https://doi.org/10.1039/D1AN01861B.

(125) Warnke, S.; Ben Faleh, A.; Rizzo, T. R. Toward High-Throughput Cryogenic IR Fingerprinting of Mobility-Separated Glycan Isomers. *ACS Meas. Sci. Au* **2021**, *1* (3), 157–164. https://doi.org/10.1021/acsmeasuresciau.1c00018.

(126) Yatsyna, V.; Abikhodr, A. H.; Ben Faleh, A.; Warnke, S.; Rizzo, T. R. High-Throughput Multiplexed Infrared Spectroscopy of Ion Mobility-Separated Species Using Hadamard Transform. *Anal. Chem.* **2022**, *94* (6), 2912–2917. https://doi.org/10.1021/acs.analchem.1c04843.

(127) Williamson, D. L.; Bergman, A. E.; Nagy, G. Investigating the Structure of α/β Carbohydrate Linkage Isomers as a Function of Group I Metal Adduction and Degree of Polymerization as Revealed by Cyclic Ion Mobility Separations. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (10), 2573–2582. https://doi.org/10.1021/jasms.1c00207.

(128) Williamson, D. L.; Nagy, G. Evaluating the Utility of Temporal Compression in High-Resolution Traveling Wave-Based Cyclic Ion Mobility Separations. *ACS Meas. Sci. Au* **2022**, *2* (4), 361–369. https://doi.org/10.1021/acsmeasuresciau.2c00016.

(129) Cavallero, G. J.; Zaia, J. Resolving Heparan Sulfate Oligosaccharide Positional Isomers Using Hydrophilic Interaction Liquid Chromatography-Cyclic Ion Mobility Mass Spectrometry. *Anal. Chem.* **2022**, *94* (5), 2366–2374. https://doi.org/10.1021/acs.analchem.1c03543.

(130) Miller, R. L.; Guimond, S. E.; Schwörer, R.; Zubkova, O. V.; Tyler, P. C.; Xu, Y.; Liu, J.; Chopra, P.; Boons, G.-J.; Grabarics, M.; Manz, C.; Hofmann, J.; Karlsson, N. G.; Turnbull, J. E.; Struwe, W. B.; Pagel, K. Shotgun Ion Mobility Mass Spectrometry Sequencing of Heparan Sulfate Saccharides. *Nat. Commun.* **2020**, *11* (1), 1481. https://doi.org/10.1038/s41467-020-15284-y.

(131) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray Ionization-Principles and Practice. *Mass Spectrom. Rev.* **1990**, *9* (1), 37–70. https://doi.org/10.1002/mas.1280090103.

(132) Fenn, J. B. Ion Formation from Charged Droplets: Roles of Geometry, Energy, and Time. *J. Am. Soc. Mass Spectrom.* **1993**, *4* (7), 524–535. https://doi.org/10.1016/1044-0305(93)85014-O.

(133) Hogan, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L. A Combined Charged Residue-Field Emission Model of Macromolecular Electrospray Ionization. *Anal. Chem.* **2009**, *81* (1), 369–377. https://doi.org/10.1021/ac8016532.

(134) Douglas, D. j. Linear Quadrupoles in Mass Spectrometry. *Mass Spectrom. Rev.* **2009**, *28* (6), 937–960. https://doi.org/10.1002/mas.20249.

(135) McLuckey, S. A.; Mentinova, M. Ion/Neutral, Ion/Electron, Ion/Photon, and Ion/Ion Interactions in Tandem Mass Spectrometry: Do We Need Them All? Are They Enough? *J Am Soc Mass Spectrom* **2011**, *22* (1), 3–12. https://doi.org/10.1007/s13361-010-0004-9 [doi].

(136) Boesl, U. Time-of-Flight Mass Spectrometry: Introduction to the Basics. *Mass Spectrom. Rev.* **2017**, *36* (1), 86–109. https://doi.org/10.1002/mas.21520.

(137) Ayet San Andrés, S.; Hornung, C.; Ebert, J.; Plaß, W. R.; Dickel, T.; Geissel, H.; Scheidenberger, C.; Bergmann, J.; Greiner, F.; Haettner, E.; Jesch, C.; Lippert, W.; Mardor, I.; Miskun, I.; Patyk, Z.; Pietri, S.; Pihktelev, A.; Purushothaman, S.; Reiter, M. P.; Rink, A.-K.; Weick, H.; Yavor, M. I.; Bagchi, S.; Charviakova, V.; Constantin, P.; Diwisch, M.; Finlay, A.; Kaur, S.; Knöbel, R.; Lang, J.; Mei, B.; Moore, I. D.; Otto, J.-H.; Pohjalainen, I.; Prochazka, A.; Rappold, C.; Takechi, M.; Tanaka, Y. K.; Winfield, J. S.; Xu, X. High-Resolution, Accurate Multiple-Reflection Time-of-Flight Mass Spectrometry for Short-Lived, Exotic Nuclei of a Few Events in Their Ground and Low-Lying Isomeric States. *Phys. Rev. C* **2019**, *99* (6), 064313. https://doi.org/10.1103/PhysRevC.99.064313.

(138) Taha, H. A.; Richards, M. R.; Lowary, T. L. Conformational Analysis of Furanoside-Containing Mono- and Oligosaccharides. *Chem. Rev.* **2013**, *113* (3), 1851–1876. https://doi.org/10.1021/cr300249c.
(139) Peltier, P.; Euzen, R.; Daniellou, R.; Nugier-Chauvin, C.; Ferrières, V. Recent Knowledge and Innovations Related to Hexofuranosides: Structure, Synthesis and Applications. *Carbohydr. Res.* **2008**, *343* (12), 1897–1923. https://doi.org/10.1016/j.carres.2008.02.010.
(140) Ferrières, V.; Nugier-Chauvin, C.; Legentil, L.; Tranchimand, S. Chapter 19: How Recent Knowledge on Furano-Specific Enzymes Has Renewed Interest for the Synthesis of Glycofuranosyl-Containing Conjugates. In *Carbohydrate Chemistry*; 2014; pp 401–417. https://doi.org/10.1039/9781849739986-00401.
(141) Cho, E.; Riches, E.; Palmer, M.; Giles, K.; Ujma, J.; Kim, S. Isolation of Crude Oil Peaks Differing by m/z ~0.1 via Tandem Mass Spectrometry Using a Cyclic Ion Mobility-Mass Spectrometer. *Anal. Chem.* **2019**, *91* (22), 14268–14274. https://doi.org/10.1021/acs.analchem.9b02255.
(142) Richards, M. R.; Bai, Y.; Lowary, T. L. Comparison between DFT- and NMR-Based Conformational Analysis of Methyl Galactofuranosides. *Carbohydr. Res.* **2013**, *374*, 103–114. https://doi.org/10.1016/j.carres.2013.03.030.
(143) Li, A.; Laville, E.; Tarquis, L.; Lombard, V.; Ropartz, D.; Terrapon, N.; Henrissat, B.; Guieysse, D.; Esque, J.; Durand, J.; Morgavi, D. P.; Potocki-Veronese, G. 2020. Analysis of the Diversity of the Glycoside Hydrolase Family 130 in Mammal Gut Microbiomes Reveals a Novel Mannoside-Phosphorylase Function. *Microb. Genomics* **2020**, *6* (10), e000404. https://doi.org/10.1099/mgen.0.000404.
(144) Shibata, N.; Kobayashi, H.; Okawa, Y.; Suzuki, S. Existence of Novel β-1,2 Linkage-Containing Side Chain in the Mannan of Candida Lusitaniae, Antigenically Related to Candida Albicans Serotype A. *Eur. J. Biochem.* **2003**, *270* (12), 2565–2575. https://doi.org/10.1046/j.1432-1033.2003.03622.x.
(145) Kobayashi, H.; Shibata, N.; Suzuki, S. Evidence for Oligomannosyl Residues Containing Both Beta-1,2 and Alpha-1,2 Linkages as a Serotype A-Specific Epitope(s) in Mannans of Candida Albicans. *Infect. Immun.* **1992**, *60* (5), 2106–2109.
(146) McKenna, K. R.; Li, L.; Baker, A. G.; Ujma, J.; Krishnamurthy, R.; Liotta, C. L.; Fernández, F. M. Carbohydrate Isomer Resolution via Multi-Site Derivatization Cyclic Ion Mobility-Mass Spectrometry. *Analyst* **2019**, *144* (24), 7220–7226. https://doi.org/10.1039/C9AN01584A.
(147) Williamson, D. L.; Bergman, A. E.; Nagy, G. Investigating the Structure of α/β Carbohydrate Linkage Isomers as a Function of Group I Metal Adduction and Degree of Polymerization as Revealed by Cyclic Ion Mobility Separations. *J. Am. Soc. Mass Spectrom.* **2021**, *32* (10), 2573–2582. https://doi.org/10.1021/jasms.1c00207.
(148) Collins, P. M. *Dictionary of Carbohydrates*; CRC Press, 2005.
(149) Grabarics, M.; Lettow, M.; Kirschbaum, C.; Greis, K.; Manz, C.; Pagel, K. Mass Spectrometry-Based Techniques to Elucidate the Sugar Code. *Chem. Rev.* **2022**, *122* (8), 7840–7908. https://doi.org/10.1021/acs.chemrev.1c00380.
(150) Moge, B.; Schindler, B.; Yeni, O.; Compagnon, I. Fucose Migration: Where To ? **2022**. https://doi.org/10.26434/chemrxiv-2022-m22ws.
(151) Chizhov, A. O.; Filatov, A. V.; Perepelov, A. V.; Knirel, Y. A. A New Example of Rearrangement Observed in the Tandem Mass Spectra of Oligosaccharides. *J. Anal. Chem.* **2020**, *75* (14), 1842–1845. https://doi.org/10.1134/S1061934820140075.
(152) Watrous, J.; Roach, P.; Alexandrov, T.; Heath, B. S.; Yang, J. Y.; Kersten, R. D.; Voort, M. van der; Pogliano, K.; Gross, H.; Raaijmakers, J. M.; Moore, B. S.; Laskin, J.; Bandeira, N.; Dorrestein, P. C. Mass Spectral Molecular Networking of Living Microbial Colonies. *Proc. Natl. Acad. Sci.* **2012**, *109* (26), E1743–E1752. https://doi.org/10.1073/pnas.1203689109.
(153) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W.-T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-

Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C.-C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C.-C.; Yang, Y.-L.; Humpf, H.-U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; P, C. A. B.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P.-M.; Phapale, P.; Nothias, L.-F.; Alexandrov, T.; Litaudon, M.; Wolfender, J.-L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D.-T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. O.; Pogliano, K.; Linington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34*, 828–837. https://doi.org/10.1038/nbt.3597.

(154) Wang, M.; Jarmusch, A. K.; Vargas, F.; Aksenov, A. A.; Gauglitz, J. M.; Weldon, K.; Petras, D.; da Silva, R.; Quinn, R.; Melnik, A. V.; van der Hooft, J. J. J.; Caraballo-Rodríguez, A. M.; Nothias, L. F.; Aceves, C. M.; Panitchpakdi, M.; Brown, E.; Di Ottavio, F.; Sikora, N.; Elijah, E. O.; Labarta-Bajo, L.; Gentry, E. C.; Shalapour, S.; Kyle, K. E.; Puckett, S. P.; Watrous, J. D.; Carpenter, C. S.; Bouslimani, A.; Ernst, M.; Swafford, A. D.; Zúñiga, E. I.; Balunas, M. J.; Klassen, J. L.; Loomba, R.; Knight, R.; Bandeira, N.; Dorrestein, P. C. Mass Spectrometry Searches Using MASST. *Nat. Biotechnol.* **2020**, *38* (1), 23–26. https://doi.org/10.1038/s41587-019-0375-9.

(155) Carbonero, E. R.; Smiderle, F. R.; Gracher, A. H. P.; Mellinger, C. G.; Torri, G.; Ahti, T.; Gorin, P. A. J.; Iacomini, M. Structure of Two Glucans and a Galactofuranomannan from the Lichen Umbilicaria Mammulata. *Carbohydr. Polym.* **2006**, *63* (1), 13–18. https://doi.org/10.1016/j.carbpol.2005.04.010.

(156) Amicucci, M. J.; Nandita, E.; Galermo, A. G.; Castillo, J. J.; Chen, S.; Park, D.; Smilowitz, J. T.; German, J. B.; Mills, D. A.; Lebrilla, C. B. A Nonenzymatic Method for Cleaving Polysaccharides to Yield Oligosaccharides for Structural Analysis. *Nat. Commun.* **2020**, *11* (1), 3963. https://doi.org/10.1038/s41467-020-17778-1.

(157) Ganem, B.; Li, Y. T.; Henion, J. D. Detection of Noncovalent Receptor-Ligand Complexes by Mass Spectrometry. *J. Am. Chem. Soc.* **1991**, *113* (16), 6294–6296. https://doi.org/10.1021/ja00016a069.

(158) Katta, V.; Chait, B. T. Observation of the Heme-Globin Complex in Native Myoglobin by Electrospray-Ionization Mass Spectrometry. *J. Am. Chem. Soc.* **1991**, *113* (22), 8534–8535. https://doi.org/10.1021/ja00022a058.

(159) Smith, R. D.; Light-Wahl, K. J.; Winger, B. E.; Loo, J. A. Preservation of Non-Covalent Associations in Electrospray Ionization Mass Spectrometry: Multiply Charged Polypeptide and Protein Dimers. *Org. Mass Spectrom.* **1992**, *27* (7), 811–821. https://doi.org/10.1002/oms.1210270709.

(160) Snijder, J.; Rose, R. J.; Veesler, D.; Johnson, J. E.; Heck, A. J. R. Studying 18 MDa Virus Assemblies with Native Mass Spectrometry. *Angew. Chem. Int. Ed.* **2013**, *52* (14), 4020–4023. https://doi.org/10.1002/anie.201210197.

(161) Shoemaker, G. K.; Duijn, E. van; Crawford, S. E.; Uetrecht, C.; Baclayon, M.; Roos, W. H.; Wuite, G. J. L.; Estes, M. K.; Prasad, B. V. V.; Heck, A. J. R. Norwalk Virus Assembly and Stability Monitored by Mass Spectrometry *. *Mol. Cell. Proteomics* **2010**, *9* (8), 1742–1751. https://doi.org/10.1074/mcp.M900620-MCP200.

(162) Ehkirch, A.; Hernandez-Alba, O.; Colas, O.; Beck, A.; Guillarme, D.; Cianférani, S. Hyphenation of Size Exclusion Chromatography to Native Ion Mobility Mass

Spectrometry for the Analytical Characterization of Therapeutic Antibodies and Related Products. *J. Chromatogr. B* **2018**, *1086*, 176–183. https://doi.org/10.1016/j.jchromb.2018.04.010.

(163) Deslignière, E.; Ehkirch, A.; Botzanowski, T.; Beck, A.; Hernandez-Alba, O.; Cianférani, S. Toward Automation of Collision-Induced Unfolding Experiments through Online Size Exclusion Chromatography Coupled to Native Mass Spectrometry. *Anal. Chem.* **2020**, *92* (19), 12900–12908. https://doi.org/10.1021/acs.analchem.0c01426.

(164) Dervilly-Pinel, G.; Thibault, J.-F.; Saulnier, L. Experimental Evidence for a Semi-FLexible Conformation for Arabinoxylans. *Carbohydr. Res.* **2001**, *330*, 365–372.

(165) Marty, M. T.; Baldwin, A. J.; Marklund, E. G.; Hochberg, G. K. A.; Benesch, J. L. P.; Robinson, C. V. Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Anal. Chem.* **2015**, *87* (8), 4370–4376. https://doi.org/10.1021/acs.analchem.5b00140.

(166) Hong, P.; Koza, S.; Bouvier, E. S. P. A Review Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and Their Aggregates. *J. Liq. Chromatogr. Relat. Technol.* **2012**, *35* (20), 2923–2950. https://doi.org/10.1080/10826076.2012.743724.

(167) Keifer, D. Z.; Shinholt, D. L.; Jarrold, M. F. Charge Detection Mass Spectrometry with Almost Perfect Charge Accuracy. *Anal. Chem.* **2015**, *87* (20), 10330–10337. https://doi.org/10.1021/acs.analchem.5b02324.

(168) Chao, H.-C.; Lee, K. W.; Shih, M.; McLuckey, S. A. Characterization of Homopolymer Distributions via Direct Infusion ESI-MS/MS Using Wide Mass-to-Charge Windows and Gas-Phase Ion/Ion Reactions. *J. Am. Soc. Mass Spectrom.* **2022**, *33* (4), 704–713. https://doi.org/10.1021/jasms.2c00001.

(169) Deery, M. J.; Stimson, E.; Chappell, C. G. Size Exclusion Chromatography/Mass Spectrometry Applied to the Analysis of Polysaccharides. *Rapid Commun. Mass Spectrom.* **2001**, *15* (23), 2273–2283. https://doi.org/10.1002/rcm.458.

(170) Jiang, Q.; Wang, Y.; Li, H.; Chen, D. D. Y. Combining Online Size Exclusion Chromatography and Electrospray Ionization Mass Spectrometry to Characterize Plant Polysaccharides. *Carbohydr. Polym.* **2020**, *246*, 116591. https://doi.org/10.1016/j.carbpol.2020.116591.

(171) Kováčik, V.; Hirsch, J.; Kováč, P.; Heerma, W.; Thomas-Oates, J.; Haverkamp, J. Oligosaccharide Characterization Using Collision-Induced Dissociation Fast Atom Bombardment Mass Spectrometry: Evidence for Internal Monosaccharide Residue Loss. *J. Mass Spectrom.* **1995**, *30* (7), 949–958. https://doi.org/10.1002/jms.1190300704.

(172) Brüll, L. P.; Heerma, W.; Thomas-Oates, J.; Haverkamp, J.; Kovácik, V.; Kovác, P. Loss of Internal 1 → 6 Substituted Monosaccharide Residues from Underivatized and Per-O-Methylated Trisaccharides. *J. Am. Soc. Mass Spectrom.* **1997**, *8* (1), 43–49. https://doi.org/10.1016/S1044-0305(96)00134-1.

(173) Ernst, B.; Müller, D. R.; Richter, W. J. False Sugar Sequence Ions in Electrospray Tandem Mass Spectrometry of Underivatized Sialyl-Lewis-Type Oligosaccharides. *Int. J. Mass Spectrom. Ion Process.* **1997**, *160* (1), 283–290. https://doi.org/10.1016/S0168-1176(96)04487-4.

(174) Brüll, L. P.; Kovácik, V.; Thomas-Oates, J. E.; Heerma, W.; Haverkamp, J. Sodium-Cationized Oligosaccharides Do Not Appear to Undergo 'Internal Residue Loss' Rearrangement Processes on Tandem Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1998**, *12* (20), 1520–1532. https://doi.org/10.1002/(SICI)1097-0231(19981030)12:20<1520::AID-RCM336>3.0.CO;2-W.

(175) Hecht, E. S.; Loziuk, P. L.; Muddiman, D. C. Xylose Migration During Tandem Mass Spectrometry of N-Linked Glycans. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (4), 729–732. https://doi.org/10.1007/s13361-016-1588-5.

(176) Wuhrer, M.; Koeleman, C. A. M.; Deelder, A. M. Hexose Rearrangements upon Fragmentation of N-Glycopeptides and Reductively Aminated N-Glycans. *Anal. Chem.* **2009**, *81* (11), 4422–4432. https://doi.org/10.1021/ac900278q.

(177) Moge, B.; Schindler, B.; Yeni, O.; Compagnon, I. Fucose Migration: Where To ? ChemRxiv September 9, 2022. https://doi.org/10.26434/chemrxiv-2022-m22ws.

(178) Adamson, J. T.; Hakansson, K. Electron Capture Dissociation of Oligosaccharides Ionized with Alkali, Alkaline Earth, and Transition Metals. *Anal Chem* **2007**, *79* (7), 2901–2910. https://doi.org/10.1021/ac0621423 [doi].
(179) Wolff, J. J.; Laremore, T. N.; Aslam, H.; Linhardt, R. J.; Amster, I. J. Electron-Induced Dissociation of Glycosaminoglycan Tetrasaccharides. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (10), 1449–1458. https://doi.org/10.1016/j.jasms.2008.06.024.
(180) Sasiene, Z. J.; Mendis, P. M.; Ropartz, D.; Rogniaux, H.; Jackson, G. P. The Influence of Na/H Exchange on the Charge Transfer Dissociation (CTD) Spectra of Mannuronic Acid Oligomers. *Int. J. Mass Spectrom.* **2021**, *468*, 116634. https://doi.org/10.1016/j.ijms.2021.116634.
(181) Sasiene, Z. J.; Ropartz, D.; Rogniaux, H.; Jackson, G. P. Charge Transfer Dissociation of a Branched Glycan with Alkali and Alkaline Earth Metal Adducts. *J. Mass Spectrom.* **2021**, *56* (7), e4774. https://doi.org/10.1002/jms.4774.
(182) Ropartz, D.; Li, P.; Jackson, G. P.; Rogniaux, H. Negative Polarity Helium Charge Transfer Dissociation Tandem Mass Spectrometry: Radical-Initiated Fragmentation of Complex Polysulfated Anions. *Anal. Chem.* **2017**, *89* (7), 3824–3828. https://doi.org/10.1021/acs.analchem.7b00473.
(183) Bittremieux, W.; Schmid, R.; Huber, F.; van der Hooft, J. J. J.; Wang, M.; Dorrestein, P. C. Comparison of Cosine, Modified Cosine, and Neutral Loss Based Spectrum Alignment For Discovery of Structurally Related Molecules. *J. Am. Soc. Mass Spectrom.* **2022**, *33* (9), 1733–1744. https://doi.org/10.1021/jasms.2c00153.
(184) Huber, F.; Ridder, L.; Verhoeven, S.; Spaaks, J. H.; Diblen, F.; Rogers, S.; Hooft, J. J. J. van der. Spec2Vec: Improved Mass Spectral Similarity Scoring through Learning of Structural Relationships. *PLOS Comput. Biol.* **2021**, *17* (2), e1008724. https://doi.org/10.1371/journal.pcbi.1008724.
(185) Reid, D. J.; Diesing, J. M.; Miller, M. A.; Perry, S. M.; Wales, J. A.; Montfort, W. R.; Marty, M. T. MetaUniDec: High-Throughput Deconvolution of Native Mass Spectra. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (1), 118–127. https://doi.org/10.1007/s13361-018-1951-9.
(186) Allen, S. J.; Eaton, R. M.; Bush, M. F. Analysis of Native-Like Ions Using Structures for Lossless Ion Manipulations. *Anal. Chem.* **2016**, *88* (18), 9118–9126. https://doi.org/10.1021/acs.analchem.6b02089.
(187) Chouinard, C. D.; Nagy, G.; Webb, I. K.; Shi, T.; Baker, E. S.; Prost, S. A.; Liu, T.; Ibrahim, Y. M.; Smith, R. D. Improved Sensitivity and Separations for Phosphopeptides Using Online Liquid Chromotography Coupled with Structures for Lossless Ion Manipulations Ion Mobility–Mass Spectrometry. *Anal. Chem.* **2018**, *90* (18), 10889–10896. https://doi.org/10.1021/acs.analchem.8b02397.
(188) Keifer, D. Z.; Pierson, E. E.; Jarrold, M. F. Charge Detection Mass Spectrometry: Weighing Heavier Things. *Analyst* **2017**, *142* (10), 1654–1671. https://doi.org/10.1039/C7AN00277G.
(189) Miller, L. M.; Bond, K. M.; Draper, B. E.; Jarrold, M. F. Characterization of Classical Vaccines by Charge Detection Mass Spectrometry. *Anal. Chem.* **2021**, *93* (35), 11965–11972. https://doi.org/10.1021/acs.analchem.1c01893.
(190) Miller, L. M.; Barnes, L. F.; Raab, S. A.; Draper, B. E.; El-Baba, T. J.; Lutomski, C. A.; Robinson, C. V.; Clemmer, D. E.; Jarrold, M. F. Heterogeneity of Glycan Processing on Trimeric SARS-CoV-2 Spike Protein Revealed by Charge Detection Mass Spectrometry. *J. Am. Chem. Soc.* **2021**, *143* (10), 3959–3966. https://doi.org/10.1021/jacs.1c00353.
(191) Kostelic, M. M.; Zak, C. K.; Liu, Y.; Chen, V. S.; Wu, Z.; Sivinski, J.; Chapman, E.; Marty, M. T. UniDecCD: Deconvolution of Charge Detection-Mass Spectrometry Data. *Anal. Chem.* **2021**, *93* (44), 14722–14729. https://doi.org/10.1021/acs.analchem.1c03181.
(192) Harper, C. C.; Elliott, A. G.; Oltrogge, L. M.; Savage, D. F.; Williams, E. R. Multiplexed Charge Detection Mass Spectrometry for High-Throughput Single Ion Analysis of Large Molecules. *Anal. Chem.* **2019**, *91* (11), 7458–7465. https://doi.org/10.1021/acs.analchem.9b01669.

Résumé en Français

Les carbohydrates sont les biopolymères les plus abondants sur Terre : ils sont omniprésents dans les organismes vivants et se retrouvent aussi bien chez les animaux, les plantes, les algues, les champignons que les micro-organismes. D'un point de vue phylogénétique, ce sont les biomolécules les moins conservées du Vivant. En effet, selon l'organisme producteur et même selon le stade de vie, l'organe ou le tissu, leurs structures et compositions varient fortement. Leurs structures peuvent même être influencées par des conditions environnementales ou abiotiques, par exemple les marées ont un impact sur le schéma de sulfatation des polysaccharides d'algues. Les carbohydrates remplissent diverses fonctions biologiques allant de la signalisation cellulaire à des rôles structurels clés (par exemple, dans la paroi cellulaire). Ils trouvent également de nombreuses applications *ex vivo*, notamment dans les industries alimentaire, cosmétique et pharmaceutique et en bioraffinerie. Cependant, les propriétés biologiques des carbohydrates – ainsi que les applications industrielles qui en découlent – dépendent en grande partie de leur structure. De nombreuses caractéristiques subtiles de la structure des carbohydrates sont essentielles à leurs propriétés mais restent compliquées à déchiffrer par les méthodes analytiques actuelles. En conséquence, de nombreux carbohydrates restent structurellement inconnus, ce qui limite l'étude de leur relation structure-fonction et fait des carbohydrates une ressource probablement sous-exploitée.

Pour d'autres types de biopolymères tels que les protéines, la spectrométrie de masse (MS) s'est imposée comme un outil essentiel permettant de caractériser jusqu'à des milliers de protéines par heure, y compris de déterminer avec précision leurs modifications potentielles. Cependant, dans le domaine des glycosciences structurales, la MS est nettement en retard. La prépondérance des formes isomériques dans ces composés constitue un obstacle majeur à l'utilisation de la MS pour la caractérisation des oligo- et polysaccharides. Tout d'abord, de nombreux types de sous-unités isomériques existent parmi les principales familles de carbohydrates (c'est-à-dire parmi les hexoses, les pentoses, les acides uroniques, etc.). Chacune de ces familles comprend en fait différents stéréoisomères, qui peuvent même parfois correspondre à l'orientation d'une seule fonction hydroxyle (épimères). Une complexité supplémentaire provient de la coexistence de tautomères de cyclisation, un même type de monosaccharide pouvant former des cycles à 6 chaînons (forme pyranose) et/ou à 5 chaînons (forme furanose). La formation de la liaison glycosidique entre les sous-unités augmente la complexité en ajoutant deux types d'isomérie supplémentaires : (i) la régioisomérie de la liaison, qui peut se produire sur n'importe quelle fonction hydroxyle, et (ii) l'anomérie α/β de la liaison, résultant des anomères α et β existant pour chaque monosaccharide branché.

Tous les facteurs ci-dessus induisent une extrême diversité structurale des carbohydrates. On a calculé par exemple qu'il existe plus de 12 millions de structures possibles pour un oligosaccharide construit à partir de six sous-unités d'hexose. Enfin, diverses modifications

chimiques (acétylation, *N*-acétylation, sulfatation, ponts anhydro, méthyl-étherification, etc.) peuvent se produire à différentes positions disponibles des sous-unités, amplifiant encore les possibilités d'isomérie.

Face à ces cas d'isomérie, la spectrométrie de mobilité ionique (IMS) est un complément intéressant à la MS. L'IMS est une technique d'analyse en phase gazeuse, qui sépare les populations d'ions par le biais d'une interaction avec un gaz tampon et un champ électrique. La séparation qui en résulte dépend notamment de la structure tridimensionnelle des ions, ce qui la rend sensible à toutes sortes d'isomères et de conformères en phase gazeuse (à condition que l'appareil ait une résolution suffisante en IMS). L'IMS peut être facilement intégrée dans les spectromètres de masse et est maintenant disponible dans le commerce avec des résolutions élevées. Une configuration notable est la mobilité ionique cyclique de Waters Corp. (Wilmslow, Royaume-Uni), technologie qui a été commercialisée en 2019 sous le nom de SELECT SERIES Cyclic IMS (cIMS). Cet instrument offre une résolution IMS ultrahaute mais aussi une grande flexibilité concernant la manipulation des ions, permettant des expériences de dissociation induite par collision (CID)-IMS, IMS-CID ainsi que des expériences IMS multiétapes (IMS^n). **L'objectif de ma thèse était d'explorer les potentialités de l'instrument cIMS pour la caractérisation structurale des carbohydrates.**

A. La mobilité ionique cyclique peut-elle préserver et résoudre tous les types d'isomérie dans les oligosaccharides ?

Au cours de mon travail de doctorat, la première question principale à laquelle nous avons tenté de répondre était la suivante : ***Pouvons-nous procéder à une détermination structurale complète d'oligosaccharides complexes en utilisant la cIMS ?*** (avec l'hypothèse implicite que l'IMS est combinée avec la MS).

Pour répondre à cette question, nous avons défini une question intermédiaire qui est un préalable à la première : ***La cIMS – et les différentes expériences qui peuvent lui être associées, comme la MS/MS – est-elle capable de retenir et de résoudre tous les types d'isomérie que l'on peut trouver dans les oligosaccharides ?*** Les différents types d'isomérie comprennent l'isomérie de constitution (tautomérie de cyclisation, régioisomérie de liaison, isomérie de position des groupes fonctionnels) et la stéréoisomérie (stéréoisomérie des sous-unités, y compris épimérie, et anomérie de la liaison glycosidique).

Ce travail vise donc d'abord à établir que la cIMS est capable de discriminer sans équivoque les différents types d'isoméries. Une étude précédente de notre groupe a déjà démontré que l'isomérie de position des groupes fonctionnels pouvait être résolue par des approches de fragmentation post-IMS. Mon travail a consisté à aller plus en profondeur en explorant la

résolution des isoméries constitutionnelles, la régioisomérie de liaison et la tautomérie furanose/pyranose, ainsi que l'anomérie de liaison et la distinction de sous-unités épimériques.

Tout d'abord, j'ai abordé l'isomérie constitutionnelle par l'étude d'une bibliothèque complète de disaccharides Gal-Man, contenant notamment des unités Gal*f*. Nous avons montré que la cIMS pouvait différencier tous les cas de régioisomérie de liaison et de tautomérie furanose/pyranose. Cependant, une étude approfondie des oligosaccharides contenant des unités Gal*f*, moins stables, a mis en évidence certains réarrangements structurels, ce qui peut compliquer l'étude structurale des oligosaccharides contenant des unités Gal*f*. Il est important de noter que les fragments isomères issus de ces réarrangements n'ont pu être résolus que grâce au grand pouvoir résolutif de la cIMS.

Deuxièmement, j'ai abordé le cas de la stéréoisomérie. J'ai montré que l'anomérie de la liaison glycosidique était préservée après de multiples cycles de CID dans les expériences IMS^n et j'ai exploité cette rétention, ou "mémoire anomérique", pour effectuer la caractérisation anomérique complète de produits de réaction enzymatique, ce qui était une première dans un milieu biologique. Ensuite, j'ai conçu des séquences IMS pour séparer des monosaccharides non dérivatisés et cationisés par des métaux alcalins, y compris des monosaccharides épimères (par exemple, Glc/Gal ou Glc/Man). Des séquences ont été conçues pour les principales familles isomériques de sous-unités présentes dans les carbohydrates végétaux ou bactériens (hexoses, pentoses, désoxyhexoses, acides hexuroniques, *N*-acétyl hexosamines). Ce travail a posé les bases de la stratégie de séquençage *de novo* des carbohydrates que nous avions l'intention de développer dans cette thèse, et que nous avons abordée dans un second temps.

B. Peut-on réaliser une détermination structurale complète d'oligosaccharides complexes en utilisant la cIMS ?

Dans la partie précédente de ce travail, nous avons pu affirmer que la cIMS – et les différentes expériences qui peuvent lui être associées, comme la CID – sont capables de retenir et de résoudre tous les types d'isomérie que l'on peut trouver dans les oligosaccharides. Cela nous a incités à passer à une autre question centrale de ce travail : ***Pouvons-nous réaliser une détermination structurelle complète d'oligosaccharides complexes en utilisant la cIMS ?***

La suite de ma thèse vise donc à déterminer s'il est possible de caractériser la structure complète des oligosaccharides à l'aide de la cIMS (jusqu'à la stéréochimie des sous-unités et de la liaison glycosidique). Il s'agit notamment d'explorer les propriétés des oligosaccharides

lors de leur fragmentation pour évaluer si les fragments générés en IMS/IMS ou IMS^n contiennent toutes les informations structurales requises.

La présente section présente les différentes stratégies que j'ai explorées pour tenter de résoudre les structures complètes des oligosaccharides : (i) une approche de séquençage ciblé pour résoudre la structure d'un oligosaccharide spécifique, et (ii) une approche non ciblée de mise en réseau pour organiser les jeux de données glycomiques.

Tout d'abord, j'ai tenté d'effectuer un séquençage total des oligosaccharides en utilisant l'IMS^n. Les premiers résultats étaient prometteurs, montrant que les fragments monosaccharidiques libérés par les oligosaccharides lors de la CID correspondaient aux monosaccharides appropriés (en utilisant la stratégie de séparation conçue précédemment). Lorsque l'on travaille avec des homo-oligomères, cette empreinte monosaccharidique est même préservée dans des conditions IMS^n. Il y a cependant eu certains revers. Par exemple, lors de la fragmentation des espèces $[M + Li]^+$ (que nous avons jugé être le meilleur compromis entre l'efficacité de la fragmentation et la séparation IMS), l'extrémité non réductrice donne principalement des ions de type B. Cela introduit une double liaison dans la structure de la molécule. Ce phénomène peut potentiellement annuler la différence entre les sous-unités épimères (notamment le Mannose et le Glucose, épimères en C-2). L'apparition de fragmentations consécutives en CID est également un problème, compliquant l'identification de quel monosaccharide a été détaché à partir de la structure initiale. Une prochaine étape consistera à mieux comprendre le comportement des grands oligosaccharides dans les conditions de la cIMS et de l'IMS^n en choisissant un modèle contenant plus de types de sous-unités non isomériques, ce qui devrait être un premier pas vers le séquençage d'oligosaccharides encore plus complexes.

En raison des difficultés rencontrées lors du développement d'approches de séquençage direct, j'ai cherché à développer une stratégie plus générale pour organiser les jeux de données glycomiques, dans l’objectif d’obtenir un premier niveau d'information structurale. J'ai ainsi proposé la stratégie IM-MN (ion mobility-molecular networking), inspirée des réseaux moléculaires utilisés en métabolomique, mais qui utilise les informations de mobilité ionique des fragments au lieu de celle de leur masse. Nous avons constaté que cette stratégie regroupait avec précision les oligosaccharides en fonction de leur structure, et des résultats préliminaires indiquent que la stratégie IM-MN couplée à la chromatographie liquide peut donner des résultats informatifs sur des échantillons biologiques réels. Certains développements bioinformatiques sont encore nécessaires pour tirer le meilleur parti des données IMS/IMS, mais la méthode IM-MN sera sans aucun doute utile pour organiser les

jeux de données glycomiques et dérépliquer les oligosaccharides en utilisant des bases de données spectrales de type IMS/IMS.

Dans l'ensemble, les résultats que j'ai présentés dans cette partie suggèrent que la cIMS sera utile pour parvenir à une caractérisation complète des oligosaccharides, mais certains développements méthodologiques sont encore nécessaires avant que de telles approches de séquençage ou d'IM-MN puissent être appliquées en routine.

C. La cIMS peut-elle être utilisée pour faire face au défi de la caractérisation structurale à l'échelle des polysaccharides ?

Déchiffrer la structure à l'échelle de l'oligosaccharide est d'ores et déjà pertinent pour les glycosciences structurales, qui visent notamment à élucider la relation structure-fonction des polysaccharides. En effet, les polysaccharides - qui peuvent peser jusqu'à plusieurs mégadaltons et adopter des structures tertiaires et quaternaires - sont le plus souvent étudiés de manière " bottom-up ", c'est-à-dire qu'ils sont dépolymérisés en oligosaccharides plus petits à des fins de caractérisation. Le motif polysaccharidique est ensuite reconstruit à partir des blocs oligosaccharidiques.

Or, l'étape de reconstruction peut être un travail de longue haleine et la dépolymérisation n'est pas nécessairement aisée non plus. La méthode privilégiée est d'utiliser des enzymes qui clivent spécifiquement certaines liaisons glycosidiques (en fonction notamment des sous-unités et de la régioisomérie des liaisons, mais aussi des ramifications latérales et des modifications chimiques). Dans cette situation, la connaissance de la spécificité de l'enzyme est d'une grande aide lors de la reconstruction de la structure du polysaccharide. Cependant, cette stratégie est limitée par la disponibilité d'enzymes appropriées et nécessite une connaissance préalable du squelette du polysaccharide. Dans d'autres situations, il peut se révéler nécessaire d'utiliser des méthodes non spécifiques comme l'hydrolyse chimique.

Pour illustrer ces difficultés, dans le cadre de mon travail de thèse, nos partenaires ont tenté d'utiliser une exo-mannosidase et une endo-mannanase pour dépolymériser un polysaccharide de lichen décrit dans la littérature comme un galactomannane avec un squelette de type α1,6-mannane. Soit en raison d'une spécificité ou de conditions d'utilisation inappropriées, ces enzymes n'ont pas permis la production d'oligosaccharides, ce qui nous a conduit à nous tourner vers des méthodes chimiques de dépolymérisation, notamment l'hydrolyse acide et la méthode FITDOG basée sur l'oxydoréduction. En utilisant le MALDI-TOF-MS et d'autres méthodes, nous avons montré que le contrôle de ces méthodes chimiques n'était pas simple et qu'elles pouvaient générer des coproduits indésirables.

Connaissant par expérience le défi que représente l'hydrolyse d'un polysaccharide en oligosaccharides, nous avons estimé qu'il pourrait être bénéfique d'envisager le ***développement de méthodes d'analyse directe des polysaccharides par IM-MS***. La partie suivante de mon travail de thèse porte donc sur l'évaluation de la capacité de la cIMS et des méthodes associées à fournir des informations structurales sur les polysaccharides. Nous avons défini deux questions préliminaires pour évaluer le potentiel de la cIMS pour l'analyse des polysaccharides intacts : ***(i) Jusqu'à quelle taille peut-on visualiser les polysaccharides ? et (ii) Peut-on obtenir des informations structurelles sur un polysaccharide de plus de 10 kDa ?***

J'ai mis en place et optimisé le couplage de la chromatographie d'exclusion stérique (SEC) pour l'analyse directe de polysaccharides intacts sur la plateforme cIMS.

La méthode a permis de séparer et de mesurer des pullulanes jusqu'à 200 kDa, ainsi que certains arabinoxylanes complexes d'environ 125 kDa. Il convient de noter que le pouvoir de résolution IM du cIMS a permis de distinguer des populations de polysaccharides dans des profils SEC autrement non résolus. Ceci illustre la complémentarité et l'orthogonalité de la SEC et de l'IM, la SEC fournissant des informations liées au volume hydrodynamique en phase liquide des polysaccharides et l'IM à la section efficace en phase gazeuse. Le couplage SEC-IM-MS a fourni un premier niveau d'information structurale sur les polysaccharides (masse des monosaccharides constitutifs, motif trisaccharidique spécifique), ce qui constitue un avantage substantiel par rapport aux autres méthodes existantes basées sur la SEC, comme la SEC-MALS.

Dans l'ensemble, ces résultats ouvrent la voie à la caractérisation des polysaccharides intacts par MS, ce qui a rarement été rapporté dans la littérature. De précédentes tentatives d'utilisation de la SEC-MS pour les polysaccharides n'ont permis de détecter que des espèces dépassant légèrement 9 kDa. Cependant, nous avons noté une méthode récente qui pourrait potentiellement compléter notre propre approche, dans laquelle des polysaccharides allant jusqu'à plus de 3 MDa ont été séparés en utilisant la SEC, puis fragmentés par CID dans la source du spectromètre avant de procéder à une analyse structurale – ce qui équivaut à une approche « middle-down » en ligne.

Cependant, le calcul du poids moléculaire du polysaccharide s'est avéré difficile, notamment en raison des difficultés à identifier avec précision les états de charge. La visualisation d'*Emerald Cities* s'est avérée prometteuse pour identifier les massifs pertinents dans les données SEC-IM-MS et attribuer les états de charge les plus faibles, mais la détermination exacte de l'état de charge sera un point important dans les développements futurs du workflow SEC-IM-MS pour les polysaccharides. À cet égard, la CDMS (*charge detection mass*

spectrometry) sera probablement une option intéressante à considérer. Par ailleurs, il n'existe pas à notre connaissance de logiciel adapté à des données aussi fortement multiplexées contenant des milliers d'éléments. À l'avenir, il sera probablement important d'améliorer les logiciels pour pouvoir interpréter facilement les données SEC-IM-MS obtenues sur les polysaccharides.

D. Perspectives

Les différentes études que j'ai menées au cours de ma thèse soutiennent sans aucun doute le potentiel de la cIMS (et plus généralement, de l'IM à haute résolution) pour la caractérisation structurale des carbohydrates - depuis les monosaccharides pesant aussi peu que 150 Da jusqu'aux polysaccharides pesant des centaines de kDa. Cependant, nous avons également identifié plusieurs défis qui nécessiteront des études et/ou des développements supplémentaires au cours des prochaines années :

- **Réarrangements en phase gazeuse des oligosaccharides :** Nous avons rencontré une première difficulté majeure au cours de nos travaux sur l'isomérie des carbohydrates, à savoir les réarrangements structurels des oligosaccharides de type galactomannane. Nos résultats HR-IMS suggèrent qu'il peut y avoir à la fois des réarrangements conformationnels et structurels en phase gazeuse, bien que cela n'ait été noté que sur les composés contenant du Gal*f*, donc dans un cas très spécifique. Néanmoins, ces résultats indiquent une migration probable du résidu Galf sur les espèces $[M + Li]^+$ (accompagnée d'une perte de résidu interne d'un Man), ce qui amènera à rester prudent dans l'interprétation des espèces observées notamment dans les expériences impliquant une activation des ions.

  Bien que ces phénomènes de migration soient connus depuis les années 90, on a longtemps pensé qu'ils ne se produisaient que sur des espèces $[M + H]^+$ ou $[M + NH_4]^+$, au point que certains auteurs ont proposé d'utiliser des adduits métalliques pour éviter ce phénomène de migration. Les migrations sur des adduits de métaux alcalins à des oligosaccharides n'ont été décrites que récemment sur des espèces $[M + Na]^+$. En outre, la migration de Gal n'avait pas été décrite auparavant (à notre connaissance), la littérature contenant principalement des descriptions de la migration du fucose, du xylose ou du rhamnose. Des réarrangements d'hexose ont été signalés précédemment, mais ces descriptions sont rares. Nous pensons que cela peut s'expliquer par le fait que ces réarrangements sont difficiles à identifier avec la MS seule, alors que la masse d'un résidu Fuc, par exemple, est plus facile à détecter dans les squelettes Hex (ou HexNAc, etc.). A l'avenir, il nous semble important de mener d'autres études pour étendre nos connaissances sur les réarrangements de glycanes. La contribution des méthodes en

phase gazeuse telles que l'IMS (mais aussi l'IRMPD) pourrait aider à découvrir de nouveaux types de réarrangements ou de nouveaux mécanismes.

- **Fragmentations consécutives et manque de fragments intracycliques en CID :** Un autre point majeur qui nécessitera des études et des développements futurs est l'information limitée fournie par les fragments en CID. En effet, la CID donne principalement des clivages de liaisons glycosidiques et peu de fragments intracycliques. Bien que les clivages de liaisons glycosidiques soient intéressants car des sous-unités entières sont libérées de l'oligosaccharide précurseur, cette information est rendue ambiguë à cause des fragmentations consécutives. Cela limite également l'attribution des régioisomères de liaison/fonctionnalisation. Une façon connue de résoudre ce problème est d'utiliser des méthodes de fragmentation alternatives pour générer des fragments intracycliques plus informatifs. La littérature contient de nombreux exemples dans lesquels des méthodes de fragmentation à haute énergie ont été utilisées pour caractériser les glucides. Ces méthodes comprennent notamment les méthodes basées sur les électrons, comme l'ECD ou l'EID. Cependant, jusqu'à très récemment, ces méthodes particulières étaient limitées aux instruments FTICR.
  Les méthodes d'activation à haute énergie ont été un des centres d'intérêt de notre laboratoire au cours des dernières années. Après des résultats prometteurs obtenus avec l'activation basée sur des photons (XUVPD utilisant un rayonnement synchrotron), notre laboratoire a collaboré il y a quelques années avec le groupe du professeur Glen Jackson de l'Université de Virginie Occidentale pour mettre en place la dissociation par transfert de charge de l'hélium (He-CTD), qui peut être montée sur un instrument de paillasse. La He-CTD est une méthode de fragmentation basée sur des interactions ion-ion qui produit des fragments radicalaires. Notre groupe a montré que la He-CTD donne des fragments similaires à ceux de la XUVPD et a exploité ces performances à plusieurs reprises pour caractériser des oligosaccharides complexes. Notamment, dans une étude récente, nous avons combiné la capacité de pic élevée de la LC-cIMS (réalisée sur l'instrument cIMS) avec la LC-CTD (réalisée sur un autre instrument) pour séparer et déchiffrer la structure de huit isomères de glucosylation dans un milieu de réaction enzymatique brut. Cette étude met clairement en évidence l'intérêt d'une plateforme unique couplant la cIMS et la fragmentation à haute énergie. Nous avons donc récemment investi dans une cellule ExD pour réaliser la fragmentation à haute énergie (ECD et EID) directement sur notre plateforme cIMS. Dans un futur proche, il sera intéressant de tester les capacités de la cellule ExD à générer des fragments intracycliques dans les stratégies pré-IMS et post-IMS.

- **Défis liés à l'application de l'IM-MN aux échantillons biologiques :** La stratégie IM-MN que nous avons développée s'est avérée très prometteuse lorsqu'elle a été appliquée à un large ensemble d'oligosaccharides modèles, représentatifs des carbohydrates végétaux. Cependant, de nombreux points d'optimisation doivent encore être abordés avant une adoption potentielle de cet outil par la communauté glycomique. Tout d'abord, les stratégies MN en métabolomique sont conçues pour être utilisées sur des données LC-MS afin d'augmenter la spécificité. Nous avons mené quelques étapes de validation concernant la compatibilité de la stratégie IM-MN avec la LC, mais nous avons été limités à l'utilisation de méthodes CID-IMS au lieu d'IMS/IMS, ce qui constitue une dégradation importante de l'approche. Cette limitation est directement liée au logiciel de contrôle de l'instrument cIMS, qui, au moment de la rédaction de cette thèse, n'offre pas la possibilité d'effectuer des acquisitions dépendantes des données (DDA) pour l'IMS/IMS. L'implémentation de la DDA pour le cIMS, qui devrait être disponible prochainement, constituera sans aucun doute une avancée majeure.

  Le système de score utilisé pour construire les réseaux IM-MN est un deuxième point critique qui devra être amélioré. Dans notre approche initiale, nous avons construit le réseau en utilisant uniquement la dimension CCS (section efficace de collision) des spectres de fragmentation, en sacrifiant la dimension *m/z* au cours du processus. Il serait très utile de développer un système de notation bidimensionnel qui tienne compte à la fois de la CCS et du *m/z*, car les informations contenues dans ces deux dimensions ne sont pas redondantes. Cependant, même en travaillant sur une seule dimension comme nous l'avons fait, des améliorations sont possibles. En effet, nous avons utilisé le score de similarité Cosine le plus classique mais un certain nombre d'autres scores ont été proposés pour améliorer les performances des réseaux moléculaires basés sur la MS. Au moins quelques-uns des systèmes de notation les plus connus – comme le Cosine modifié ou Spec2Vec, par exemple – méritent d'être testés pour notre approche IM-MN. Une fois la méthode IM-MN optimisée à un point satisfaisant, il sera également essentiel de construire des bases de données spectrales IMS/IMS qui seront utilisables pour dérépliquer les glycanes dans des mélanges complexes. Idéalement, ces bases de données devraient être construites à partir des efforts de la communauté mondiale et être librement accessibles, comme l'initiative GNPS pour la métabolomique.
- **Défis analytiques liés à l'analyse des polysaccharides de haut poids moléculaire :** Au cours de mon travail de doctorat, nous avons obtenu des résultats préliminaires mais prometteurs en utilisant la SEC-cIMS pour les polysaccharides intacts. Néanmoins, l'interprétation des données hautement multiplexées générées était un défi majeur. Cela est dû en partie au manque de logiciel de traitement adapté, car les polysaccharides intacts sont composés de milliers de DP, chacun donnant de multiples états de charge.

Nous pensons que le développement de nouveaux logiciels permettra de traiter ces données plus facilement et avec plus de précision. En effet, des options de plus en plus puissantes sont apparues récemment pour traiter les données MS des biopolymères. Par exemple, le logiciel UniDec, conçu à l'origine pour déconvoluer uniquement les spectres de masse à charge nulle (“zero charge mass spectra”) à partir de données MS ou IM-MS de protéines intactes, a depuis été amélioré pour prendre en charge les données LC-MS (UniChrom) et pour effectuer un traitement par batch (MetaUniDec).

Un autre point critique que nous avons identifié est la nécessité d'une évaluation précise de l'état de charge des polysaccharides. C'est un point clé pour déduire la masse réelle des polysaccharides et il s'est avéré difficile d'estimer précisément *z* en utilisant les dimensions MS et IM – même avec le pouvoir de résolution accru de l'instrument cIMS. Une première option à envisager serait de se tourner vers des configurations à résolution IM encore plus élevée pour mieux résoudre les états de charge. Cela est difficile avec la cIMS en raison du phénomène de « wrap-around », mais les SLIM ont déjà été utilisées pour analyser des protéines natives. Cependant, cela impliquerait probablement la suppression de la dimension SEC car les SLIM, telles qu'elles sont conçues, sont difficiles à coupler avec la LC en raison des temps de séparation prolongés (à notre connaissance, le couplage LC-SLIM n'a été signalé qu'à une seule occasion avec une nano-RPLC). Une autre option, peut-être plus attrayante, est la CDMS qui peut détecter directement à la fois *z* et *m/z* (*z* est déterminé pour chaque ion lorsqu'il traverse un cylindre conducteur, car *z* a un impact sur l'amplitude du courant induit). La CDMS permet donc l'analyse de complexes allant jusqu'à plusieurs dizaines de MDa, et aussi notamment d'échantillons très hétérogènes tels que les isoformes de la protéine Spike du SARS-CoV-2 fortement glycosylées (ce qui est un défi en raison de l'hétérogénéité des glycanes). Il est à noter qu'une adaptation du logiciel UniDec pour les données CDMS (UniDecCD) a été récemment implémentée, ce qui rend cette option encore plus attrayante. Cependant, il faut également noter que, la CDMS étant une méthode analysant des ions individualisés, les acquisitions sont également longues et difficiles à combiner avec les approches chromatographiques – mais des développements sont actuellement réalisés pour suivre plusieurs ions à la fois et ainsi réduire le temps d'acquisition.

Pour conclure, à l'instar de ce qui a été développé en protéomique, nous envisageons deux directions potentielles que pourraient prendre les glycosciences structurales compte tenu des possibilités ouvertes par les développements que j'ai menés au cours de ma thèse. La première serait la détermination *de novo* de la structure fine des polysaccharides par ***séquençage bottom-up***, où le polysaccharide est hydrolysé (chimiquement ou

enzymatiquement) avant de déterminer la structure des oligosaccharides générés puis de réassembler le motif du polysaccharide. Le principal avantage de cette méthode est évidemment qu'elle s'appuie sur les stratégies analytiques que j'ai développées au cours de ma thèse et qui sont devenues très efficaces pour la détermination de la structure des oligosaccharides, comme nous l'avons montré au fil de ce manuscrit. Cependant, indépendamment des outils analytiques, l'hydrolyse du polysaccharide conduit, d'une part, à la perte d'informations concernant l'organisation des blocs d'oligosaccharides au sein de la chaîne polymère et d'autre part, rend indétectable toute hétérogénéité structurale à l'échelle du polymère ou de l'échantillon (taille, présence de différentes populations, présence de régions structurales particulières, etc.). La deuxième option serait le développement de méthodes qui s'adressent directement aux polysaccharides intacts et se dirigent vers une ***analyse structurale top-down*** avec des techniques basées sur la MS. Les principaux avantages de ces approches seraient la préservation des populations de polysaccharides, ainsi que la possibilité de revenir à leur masse intacte, tout en découvrant leurs caractéristiques structurales en une seule analyse. D'un autre côté, malgré la contribution démontrée du couplage SEC-IMS-MS, il semble assez difficile d'obtenir des informations structurales très détaillées à partir de polysaccharides intacts au stade actuel de performance des méthodes analytiques. Les ***stratégies middle-down*** pourraient offrir un bon compromis entre les deux approches. En effet, des parties plus grandes d'un polysaccharide (par exemple, 10-20 kDa) seraient plus représentatives de la complexité et de la topologie du polymère, par rapport aux oligosaccharides avec lesquels nous travaillons habituellement. Une première ligne de travail serait d'appliquer une stratégie similaire à celle développée par Jiang *et al.*, c'est-à-dire de coupler la SEC avec une première étape de fragmentation en source pour réduire la taille du polysaccharide et de compléter le workflow avec nos propres méthodologies cIMS pour la caractérisation structurale. Nous envisageons que l'utilisation de la cellule ExD pour fragmenter les polysaccharides avec l'ECD plutôt qu'avec la CID pourrait également contribuer à l'étape de caractérisation structurale.

Enfin, bien que beaucoup de travail soit encore nécessaire pour développer pleinement les stratégies bottom-up et top-down – et peut-être explorer également l'approche middle-down – ce travail de thèse prouve qu'elles sont désormais à notre portée grâce au système cIMS.

DOCTORAT ECOLOGIE BRETAGNE GEOSCIENCES LOIRE AGRONOMIE ALIMENTATION

Nantes Université

Titre : Exploration de la mobilité ionique haute résolution couplée à la spectrométrie de masse pour les glycosciences structurales

Mots clés : Spectrométrie de Masse, Spectrométrie de Mobilité Ionique, Cyclic IMS, Glycosciences Structurales, Carbohydrates, Isomérie

Résumé : La complexité inhérente aux carbohydrates reste un défi pour les sciences analytiques. Aucune méthode n'est actuellement en mesure de résoudre complètement les diverses isoméries présentes dans les carbohydrates, bien qu'il s'agisse de caractéristiques structurales déterminantes. La spectrométrie de masse (MS) seule est aveugle à de nombreux cas d'isomérie, et donne donc des informations incomplètes. Au cours des dix dernières années, le couplage de la spectrométrie de mobilité ionique (IMS) avec la MS n'a cessé de prendre de l'ampleur. En effet, l'IMS est sensible aux conformations des molécules en phase gazeuse et donc aux (stéréo)isoméries. Dans ce contexte, ce travail de thèse explore le potentiel de l'IMS haute résolution pour les glycosciences structurales.
La nouvelle géométrie d'IMS cyclique (cIMS), qui permet des manipulations fines des ions en phase gazeuse, est étudiée.
Premièrement, nous démontrons par une série d`études que la cIMS peut résoudre les différentes isoméries rencontrées dans les carbohydrates. Sur la base de ces résultats, nous développons ensuite des moyens pour caractériser la structure fine des oligosaccharides en utilisant la cIMS multi-étapes. Ces développements comprennent notamment une stratégie ciblée de séquençage et une approche non ciblée de réseaux moléculaires informés par IMS. Enfin, nous couplons la cIMS avec la chromatographie d'exclusion stérique pour travailler directement sur des polysaccharides de l'ordre de la centaine de kilodalton.
Dans l'ensemble, nos résultats soulignent le potentiel de l'IM-MS haute résolution à devenir une méthode pivot dans les glycosciences structurales, et ouvrent la voie à l'étude de structures polysaccharidiques plus grandes, voire intactes.

Title: Exploration of high-resolution ion mobility-mass spectrometry for structural glycosciences

Keywords: Mass Spectrometry, Ion Mobility Spectrometry, Cyclic IMS, Structural Glycosciences, Carbohydrates, Isomerism

Abstract: The inherent complexity of carbohydrates remains a challenge for analytical sciences. No single method is currently able to fully resolve the various isomerisms present within the subunits and in glycosidic bonds, although these are determining structural features of carbohydrates. Mass spectrometry (MS) alone is blind to many cases of isomerism, and thus gives incomplete information for carbohydrates. Over the last ten years, the coupling of ion mobility spectrometry (IMS) with MS has kept gaining momentum. In fact, IMS is sensitive to the gas-phase conformations of molecules and thus, to (stereo)isomerisms. In that context, this PhD work explores the potential of high-resolution IMS approaches for structural glycosciences.
The state-of-the-art Cyclic IMS (cIMS) geometry, which affords fine gas-phase manipulations of ions, is studied.
In a first part of our work, we demonstrate through a series of case studies that cIMS can resolve the various cases of isomerism found in carbohydrates. Based on these findings, we then endeavor to develop means to characterize the fine structure of oligosaccharides using multistage IMS. These developments notably include a targeted sequencing strategy and an untargeted IMS-based approach relying on molecular networks. In the last part, we couple cIMS with size exclusion chromatography with the aim to work directly on polysaccharides in the hundred-kilodalton range.
Altogether, our results emphasize the potential of high-resolution IM-MS to settle as a pivotal method in structural glycosciences, and open the way to the study of larger if not intact polysaccharide structures.