EPFL

Thèse n° 10 329

Advancing Glycan Analysis: Cryogenic Ion Spectroscopy for Database-Driven Identification

Présentée le 8 décembre 2023

Faculté des sciences de base
Laboratoire de chimie physique moléculaire
Programme doctoral en chimie et génie chimique

pour l'obtention du grade de Docteur ès Sciences

par

Ali H ABIKHODR

Acceptée sur proposition du jury

Prof. R. Beck, président du jury
Prof. T. Rizzo, directeur de thèse
Prof. K. Pagel, rapporteur
Prof. J. Oomens, rapporteur
Prof. S. Gerber, rapporteuse

École polytechnique fédérale de Lausanne

2023

Abstract

Glycans play a pivotal role in both physiological processes; however, the field of glycobiology remains relatively understudied within biochemistry. The intricate complexity of glycans, coupled with their abundance of isomeric forms, presents formidable challenges in their analysis, thereby impeding a comprehensive analysis of their functional significance. The first step towards understanding the precise function of each glycan is the identification of its primary structure. To tackle this, substantial research efforts have been directed towards the development of methods for glycan structural elucidation.

Gas phase techniques offer a significant advantage for the detailed characterization of glycans, primarily due to their ability to isolate distinct isomeric forms. However, traditional analytical approaches still fall short as they do not provide complete structural information. This necessitates the employment of new tools to unravel the intricacies of glycans.

This thesis demonstrates the use of ion mobility, mass spectrometry, and cryogenic IR spectroscopy for the comprehensive analysis and unambiguous identification of glycans. The initial focus of this research is the enhancement of the sensitivity and speed of these techniques. Subsequently, we introduce a spectroscopic database approach that enables the discrimination of isomeric glycans that have been separated through ion mobility. Additionally, we present the application of spectral decomposition to resolve glycans with overlapping ion mobility profiles.

Next, we illustrate the utilization of Hadamard transform to advance the throughput of our technique. This approach allows for simultaneous acquisition of the IR spectra of diverse glycans separated by ion mobility within a single laser scan.

After establishing a robust identification procedure for glycans against our database, we showcase the combination of ion mobility spectrometry, collision-induced dissociation, and cryogenic IR spectroscopy for identifying isomeric human milk oligosaccharides not represented in our database. This method involves fragmenting the glycans, identifying the resulting fragments, and subsequently reconstructing the structure of the precursor molecule. This innovative approach expands our glycan database without the necessity for additional standards.

The final part of this research focuses on the online integration of liquid chromatography with cryogenic IR spectroscopy. We demonstrate that the advances in acquisition speed facilitate real-time identification of peaks as they elute during liquid chromatographic analysis.

Résumé

Les glycanes jouent un rôle central dans deux processus physiologiques ; cependant, le domaine de la glycobiologie reste relativement peu étudié en biochimie. La complexité intrinsèque des glycanes, associée à leur abondance de formes isomères, présente d'énormes défis dans leur analyse, ce qui entrave une compréhension complète de leur signification fonctionnelle. La première étape vers la compréhension de la fonction précise de chaque glycan consiste en l'identification de sa structure primaire. Pour résoudre ce problème, d'importants efforts de recherche ont été dirigés vers le développement de méthodes d'élucidation structurale des glycanes.

Les techniques en phase gazeuse offrent un avantage significatif pour la caractérisation détaillée des glycanes, principalement en raison de leur capacité à isoler des formes isomères distinctes. Cependant, les approches analytiques traditionnelles demeurent insuffisantes car elles ne fournissent pas des informations structurelles complètes. Cela nécessite l'adoption de nouveaux outils pour dévoiler les subtilités des glycanes.

Cette thèse démontre l'utilisation de la mobilité ionique, de la spectrométrie de masse et de la spectroscopie IR cryogénique pour une analyse complète et une identification sans ambiguïté des glycanes. L'objectif initial de cette recherche est d'améliorer la sensibilité et la rapidité de ces techniques. Par la suite, nous introduisons une approche de base de données spectroscopique qui permet de distinguer les glycanes isomères séparés par mobilité ionique. De plus, nous présentons l'application de la décomposition spectrale pour résoudre les glycanes dont les profils de mobilité ionique se chevauchent.

Nous illustrons ensuite l'utilisation de la transformée de Hadamard pour augmenter le débit de notre technique. Cette approche permet l'acquisition simultanée des spectres IR de divers glycanes séparés par mobilité ionique au sein d'un seul balayage laser.

Après avoir établi une procédure d'identification robuste des glycanes par rapport à notre base de données, nous présentons la combinaison de la spectrométrie de mobilité ionique, de la dissociation induite par collision et de la spectroscopie IR cryogénique pour identifier les oligosaccharides isomères du lait maternel qui ne sont pas représentés dans notre base de données. Cette méthode consiste à fragmenter les glycanes, à identifier les fragments résultants, puis à reconstruire la structure de la molécule précurseur. Cette approche innovante élargit notre base de données sur les glycanes sans nécessiter de normes supplémentaires.

La dernière partie de cette recherche se concentre sur l'intégration en ligne de la chromatographie liquide avec la spectroscopie IR cryogénique. Nous démontrons que les progrès en matière de vitesse d'acquisition facilitent l'identification en temps réel des pics lors de leur élution lors de l'analyse chromatographique liquide.

Table of Contents

Chapter 1 Introduction

1.1 The importance of glycans

Glycans, also known as carbohydrates, oligosaccharides, or sugars, are a diverse class of biopolymers comprised of covalently linked monosaccharide units. As either free molecules or attachments to proteins, lipids, and nucleic acids,[1, 2] these complex sugars fulfill crucial functions in various biological processes, including cell recognition, cell signaling, and immune response. They provide a mechanism for cells to identify and interact with one another, [3, 4] enabling them to differentiate between self and non-self.[5, 6] This is particularly important in the development and function of the immune system, where the glycans on the surface of white blood cells act as receptors, allowing them to recognize and respond to pathogens.[6-8] Glycans also serve as receptors for signaling molecules,[9-11] enabling cells to communicate with one another, which is critical for multiple processes such as cell growth, differentiation, and the immunosuppressive respone.[12, 13] Glycans further influence cell signaling pathways, by modulating the activity of enzymes and other proteins.[14-17]

Glycans also play a significant role in various physiological processes.[18] For example, they are essential for developing and maintaining the extracellular matrix,[19, 20] which provides structural support for cells and tissues. They influence the regulation of blood coagulation and the formation of the blood-brain barrier.[21-24] Moreover, alterations in glycosylation patterns have been linked to various types of disease. For example, cancer cells often display abnormal glycans on their surface, which can be used as markers for early detection and diagnosis.[25-27] Pathogenic microorganisms, such as viruses and bacteria, have evolved to exploit host glycans as receptors for entry and infection.[28-30] Figure 1.1 shows how pathogens interact with different glycans for recognition and adhesion.

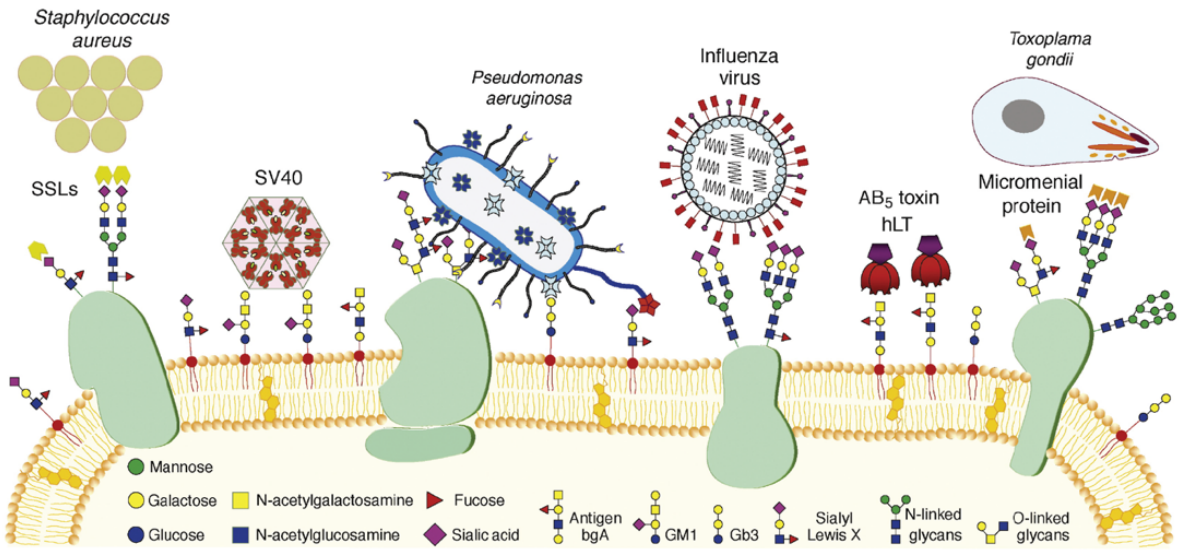


Figure 1.1. Various strategies used by pathogens for glycan recognition and adhesion. Adapted from Imberty et. al.[31]

More recently, glycans have played an integral role in the development of therapeutic strategies.[32, 33] For instance, monoclonal antibodies that target glycans have been developed as therapeutics for cancer, autoimmune diseases, and other conditions.[34, 35] Glycans have also been used as scaffolds for developing new drugs and vaccines.[36, 37] Their role in disease pathology and their potential for therapeutic strategies emphasizes the need for analytical tools that can elucidate their structure. .

1.2 Glycan isomeric complexity

Elucidating the primary structure of glycans is not trivial, due to their isomeric complexity.[2] Several factors contribute to this complexity. First, many monosaccharide building blocks are isomers of one another, differing by the stereochemistry of the asymmetric carbon atoms to which hydroxy groups are attached. For example, Figure 1.2 shows how changing the stereochemistry of C4 of glucose from an (S) to an (R) configuration transforms it into galactose, whereas switching C2 from an (R) to an (S) configuration would give mannose.

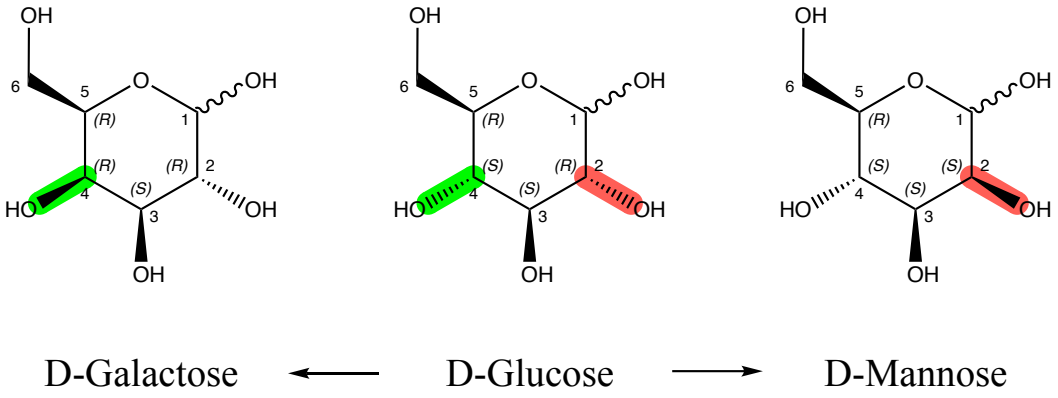
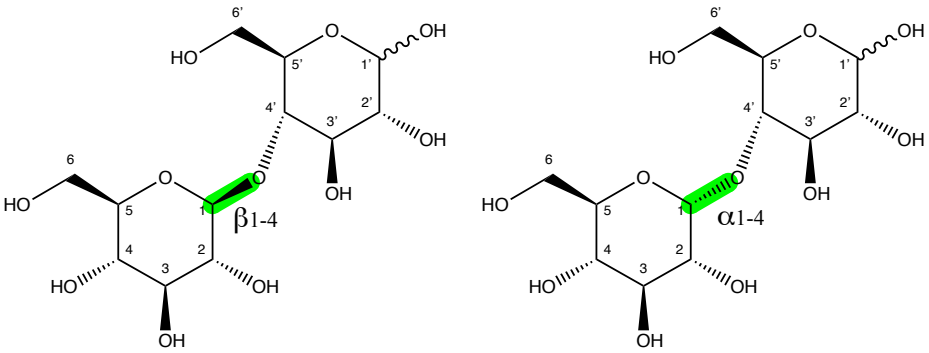


Figure 1.2. Stereochemistry differences between three monosaccharide isomers

The isomeric complexity of glycans is thus further compounded by the fact that each monosaccharide can exist in one of two anomeric forms. One form is referred to as the β anomer in which the hydroxyl groups on C1 and C5 of the monosaccharide are in the cis configuration, while the other form is referred to as the α anomer, where these two groups are in the trans configuration. When a glycosidic bond is formed between the anomeric carbon (C1) of a monosaccharide and a hydroxy group of another, it can thus give rise to two isomeric disaccharides. Figure 1.3 illustrates how the only structural difference between the isomers cellobiose and maltose is the configuration of the hydroxy group on the C1 carbon involved in the glycosidic linkage.



Cellobiose (Glu β1-4’ Glu) Maltose (Glu α1-4’ Glu)

Figure 1.3. The effect the anomericity of C1 on glycan structure

Reducing sugars have a free hydroxy group at the C1 carbon of a monosaccharide unit, and this defines the *reducing end* of the oligosaccharide. The reducing end monosaccharide can undergo a process called mutarotation, in which the OH group can switch between the α and β configurations by the ring opening and reclosing, which occurs only in solution. Non-reducing sugars do not have a free C1 hydroxy group and cannot undergo this mutarotation process.

Common examples of reducing sugars include maltose, and lactose, while sucrose and trehalose serve as examples of non-reducing sugars. (Figure 1.4)

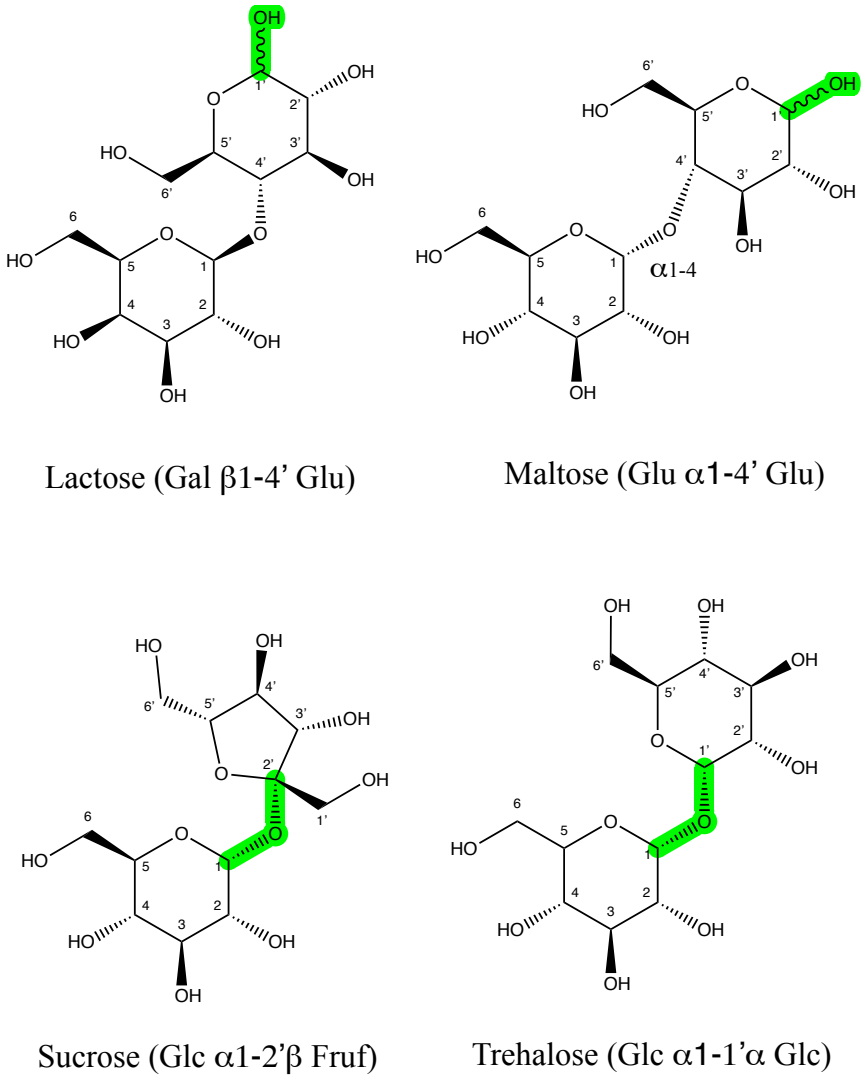
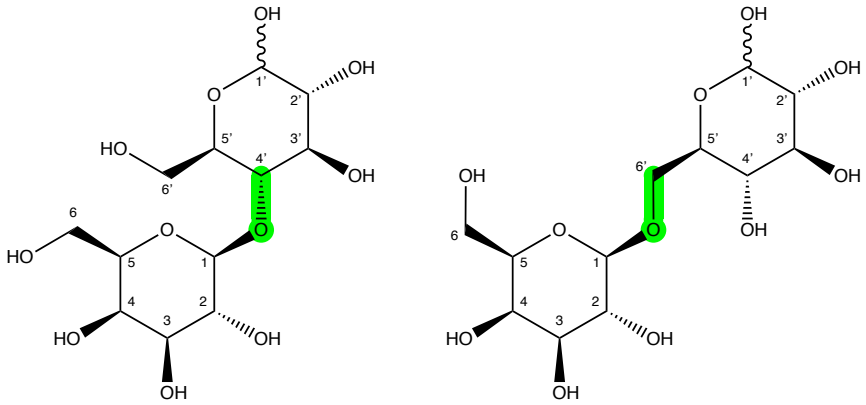


Figure 1.4. Reducing sugars Lactose and Maltose with the reducing end OH highlighted in green, and non-reducing sugars Sucrose and Trehalose highlighting the reducing ends in green.

The abundance of hydroxy groups on each monosaccharide unit also gives rise to many possible linkage isomers, also called regio-isomers. Figure 1.5 shows an example of two isomeric disaccharides (lactose and allolactose) that differ only in the linkage position between the respective monosaccharides.



Lactose (Gal β1-4 Glu) Allolactose (Gal β1-6 Glu)

Figure 1.5. Glycan regio-isomers

In addition, the multiple linkage sites available on each monosaccharide building block can lead to branched structures, which further increases their isomeric complexity. Furthermore, substitutions at the hydroxy groups, such as phosphates, sulfates, and acetates, result in a complex array of possible isomeric structures that grows exponentially with glycan size. All these types of isomerization make the separation and identification of glycans extremely challenging.

1.3 Glycan nomenclature and structure

As described above, glycans are composed of monosaccharide units linked together in various ways. Throughout this thesis, the Symbolic Nomenclature for Glycans (SNFG) will be used to represent their primary structure.[38] SNFG is a systematic way of naming glycans that is widely used in the glycomics community, as it allows easy comparison and communication of glycan structures. The notation uses a combination of shapes, letters and numbers to describe the type and position of each monosaccharide and the glycosidic linkages between them. Each monosaccharide unit is represented by a shape for a particular type of monosaccharide, with different colors representing different isomers. For example, blue and green circles represent the glucose and mannose, respectively, whereas a red triangle represents fucose. The linkage between the sugars is represented as α or β followed by a number, indicating the anomericity and the position of the glycosidic bond. Table 1 shows the SNFG notation of various monosaccharides.

SHAPE	White	Blue	Green	Yellow	Orange	Pink	Purple	Light Blue	Brown	Red
Filled Circle	Hexose	Glc	Man	Gal	Gul	Alt	All	Tal	Ido	
Filled Square	HexNAc	GlcNAc	ManNAc	GalNAc	GulNAc	AltNAc	AllNAc	TalNAc	IdoNAc	
Crossed Square	Hexosamine	GlcN	ManN	GalN	GulN	AltN	AllN	TalN	IdoN	
Divided Diamond	Hexuronate	GlcA	ManA	GalA	GulA	AltA	AllA	TalA	IdoA	
Filled Triangle	Deoxyhexose	Qui	Rha		6dGul	6dAlt		6dTal		Fuc
Divided Triangle	DeoxyhexNAc	QuiNAc	RhaNAc			6dAltNAc		6dTalNAc		FucNAc
Flat Rectangle	Di-deoxyhexose	Oli	Tyv		Abe	Par	Dig	Col		
Filled Star	Pentose		Ara	Lyx	Xyl	Rib				
Filled Diamond	Deoxynonulosonate		Kdn				Neu5Ac	Neu5Gc	Neu	Sia
Flat Diamond	Di-deoxynonulosonate		Pse	Leg		Aci		4eLeg		
Flat Hexagon	Unknown	Bac	LDManHep	Kdo	Dha	DDManHep	MurNAc	MurNGc	Mur	
Pentagon	Assigned	Api	Fruc	Tag	Sor	Psi				

Table 1: Symbols Nomenclature for Glycans representation

There are several different types of glycans, each with their unique structures. Some common types include:

- **Human Milk Oligosaccharides (HMOs)**: Human milk oligosaccharides (HMOs) are complex carbohydrates present in human breast milk. They are composed of short chains of simple monosaccharides, such as glucose, galactose, and N-acetyl glucosamine, and are one of the most abundant components in human milk.[39] These chains can be decorated by different numbers of fucose and sialic acid residues. Figure 1.5 shows the generic structure of an HMO molecule. HMOs are not digested or absorbed by infants, but instead pass through the gastrointestinal tract and act as prebiotics, promoting beneficial gut bacteria growth, such as bifidobacteria and lactobacilli.[40-42] HMOs also act as a barrier against pathogenic bacteria and viruses that can harm infants.[43-45] By binding to the receptors of pathogens, they can prevent them from adhering to the gut wall, thus reducing the risk of infection. HMOs are currently produced in large quantities and incorporated into infant formula to make it emulate the properties human milk.

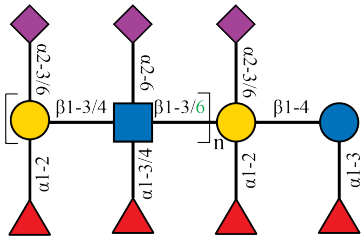


Figure 1.5. Generic structure of HMOs indicating possible bond types and fucose/sialic acid positions. The GlcNac with a β1-6 linkage induces the possibility of branching. [46]

- **N-linked glycans**: These are attached to the nitrogen atom of an asparagine residue in proteins. N-linked glycans share a common core structure of five monosaccharide units as shown in Figure 1.6. This structure can be extended by additional mannose, GlcNac, and galactose residues. Like HMOs, they are often decorated by fucose and sialic acid. N-glycans are found on almost all secreted and membrane-bound proteins and are essential for their proper functioning. They play a role in protein folding and stability by helping shield the protein from the environment and providing recognition sites for other molecules.[47-49]

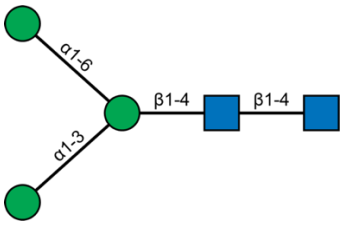


Figure 1.6. Core structure of N-linked glycans.

- **O-linked glycans**: O-linked glycans are also found on proteins. Unlike N-glycans, which are attached to the amino acid asparagine, O-glycans are connected to the hydroxy group of serine or threonine. They play a role in many biological processes, such as cell adhesion, migration, and signaling.[50-52] O-glycans are also implicated in the development of certain diseases, including cancer, and they can be used as tumor markers and therapeutic targets.[53-56] Figure 1.7 shows the eight main core structures of mammalian O-glycans, with cores 1 and 2 being the most abundant.

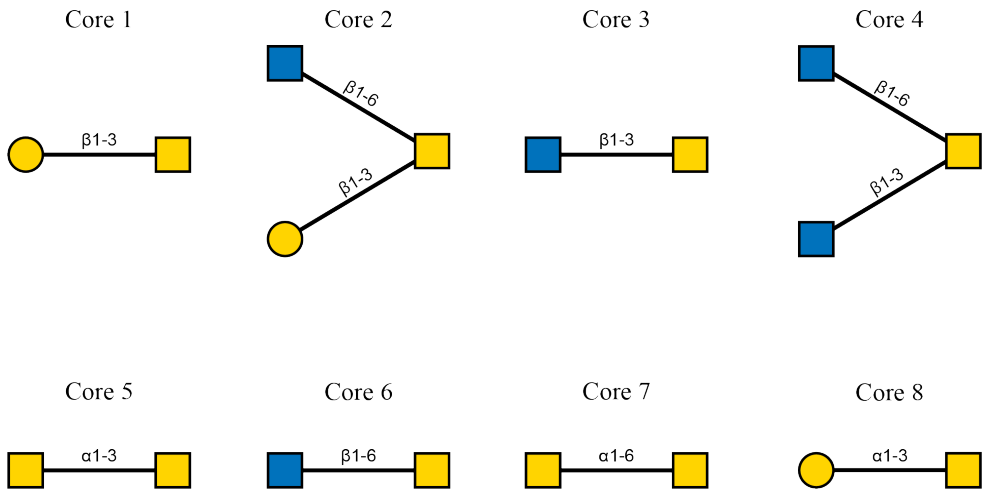


Figure 1.7. Core structures of O-linked glycans

- **Glycosaminoglycans (GAGs)**: Glycosaminoglycans (GAGs) are long chains of repeating disaccharide units found in various tissues throughout the body, including the skin, cartilage, and connective tissue. GAGs are highly hydrophilic, which gives them their characteristic viscosity and gel-like properties.[57] They are often found in the extracellular matrix, where they act as lubricants and shock absorbers, helping maintain the tissue's structural integrity.[58-60]

Some other less common glycans include S-linked and C-linked glycans. S-linked refers to oligosaccharide attachment to the sulfur atom of cysteine, whereas C-linked glycans are those where the mannose is covalently attached to a tryptophan residue. It is important to note that while these types of glycans are less common, they are still essential to consider when studying the functions and potential applications of glycans. Figure 1.8 shows how these different classes of glycans are represented using the SNFG notation.

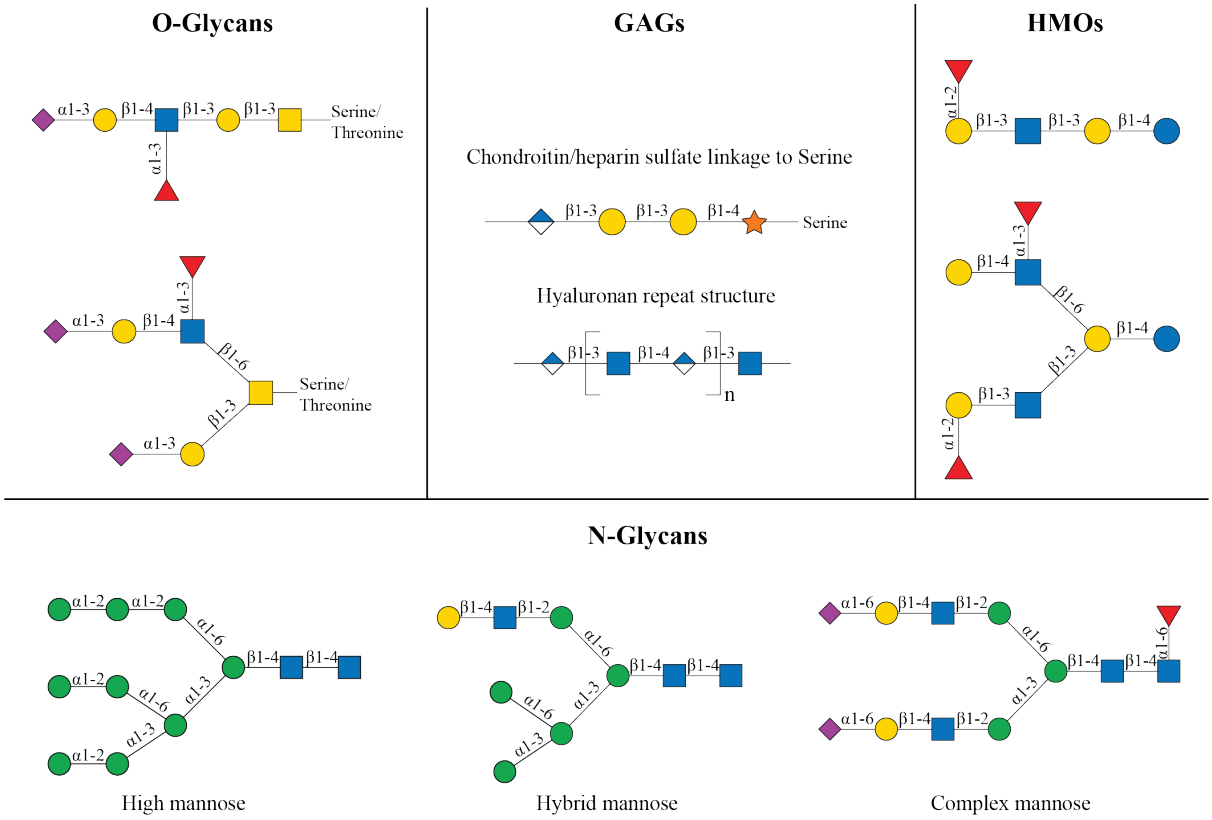


Figure 1.8. Different glycan types represented using SNFG notation

1.4 Thesis overview

The complexity of glycans makes it challenging for a single analytical technique to fully characterize their structure. This thesis work is focused on developing and demonstrating a multidimensional approach, combining orthogonal techniques for glycan analysis. We use ion mobility spectrometry (IMS) or liquid chromatography (LC) for isomeric separation followed by cryogenic infrared (IR) spectroscopy together with mass spectrometry (MS) for identification.[61-63] Cryogenic IR spectroscopy probes structure-sensitive intramolecular interactions and provides a molecular fingerprint that is highly unique and specific to the molecule.

Chapter 2 of this thesis surveys the techniques that are currently used for glycan analysis, some of which are incorporated into our multidimensional approach.

Chapter 3 provides a comprehensive overview of the technology employed in this work. We also highlight significant advancements in sensitivity and acquisition speed achieved in the latest generation instrument utilized in our experiments.

Chapter 4 discusses the challenge of identifying glycans using IMS due to their isomeric heterogeneity. We demonstrate the ability of infrared (IR) spectroscopy to distinguish isomers from a glycan mixture. To do this, we built a small database of IR spectra and show that even when IMS cannot separate isomers, a spectrum can still be used for identification.

Chapter 5 introduces a novel multiplexing method based on Hadamard transform, which enhances the signal-to-noise ratio compared to traditional signal averaging approaches. We demonstrate the method's utility for cryogenic ion spectroscopy of peptides and glycan mixtures.

Chapter 6 describes our approach using collision-induced dissociation (CID) to generate ion fragments that are separated by IMS and identified using the vibrational fingerprints of a few select standards. By identifying these fragments, we can determine the precursor structure and incorporate its vibrational fingerprint into our database. We demonstrate how this strategy can be used to determine the structure of mobility-separated isomers found in pooled human milk.

Chapter 7 demonstrates the direct linking of liquid chromatography with cryogenic infrared spectroscopy. We show how an IR spectrum can be obtained within the time that an LC peak elutes. This shows how IR spectroscopy can be eventually streamlined to becoming a routine analysis technique together with liquid chromatography.

References:

1. Schnaar, R. L., Glycobiology simplified: diverse roles of glycan recognition in inflammation. *J Leukoc Biol* **2016,** *99* (6), 825-38.

2. Laine, R. A., A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05 x 10(12) structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* **1994,** *4* (6), 759-67.

3. Kleene, R.; Schachner, M., Glycans and neural cell interactions. *Nat Rev Neurosci* **2004,** *5* (3), 195-208.

4. Parker, R. B.; Kohler, J. J., Regulation of intracellular signaling by extracellular glycan remodeling. *ACS Chem Biol* **2010,** *5* (1), 35-46.

5. Pereira, M. S.; Alves, I.; Vicente, M.; Campar, A.; Silva, M. C.; Padrao, N. A.; Pinto, V.; Fernandes, A.; Dias, A. M.; Pinho, S. S., Glycans as Key Checkpoints of T Cell Activity and Function. *Front Immunol* **2018,** *9*, 2754.

6. Dennis, J. W.; Lau, K. S.; Demetriou, M.; Nabi, I. R., Adaptive regulation at the cell surface by N-glycosylation. *Traffic* **2009,** *10* (11), 1569-78.

7. Unanue, E. R.; Turk, V.; Neefjes, J., Variations in MHC Class II Antigen Processing and Presentation in Health and Disease. *Annu Rev Immunol* **2016,** *34* (1), 265-97.

8. Cabral, J.; Hanley, S. A.; Gerlach, J. Q.; O'Leary, N.; Cunningham, S.; Ritter, T.; Ceredig, R.; Joshi, L.; Griffin, M. D., Distinctive Surface Glycosylation Patterns Associated With Mouse and Human CD4(+) Regulatory T Cells and Their Suppressive Function. *Front Immunol* **2017,** *8*, 987.

9. Demetriou, M.; Granovsky, M.; Quaggin, S.; Dennis, J. W., Negative regulation of T-cell activation and autoimmunity by Mgat5 N-glycosylation. *Nature* **2001,** *409* (6821), 733-9.

10. Dias, A. M.; Pereira, M. S.; Padrao, N. A.; Alves, I.; Marcos-Pinto, R.; Lago, P.; Pinho, S. S., Glycans as critical regulators of gut immunity in homeostasis and disease. *Cell Immunol* **2018,** *333*, 9-18.

11. Fujii, H.; Shinzaki, S.; Iijima, H.; Wakamatsu, K.; Iwamoto, C.; Sobajima, T.; Kuwahara, R.; Hiyama, S.; Hayashi, Y.; Takamatsu, S.; Uozumi, N.; Kamada, Y.; Tsujii, M.; Taniguchi, N.; Takehara, T.; Miyoshi, E., Core Fucosylation on T Cells, Required for Activation of T-Cell Receptor Signaling and Induction of Colitis in Mice, Is Increased in Patients With Inflammatory Bowel Disease. *Gastroenterology* **2016,** *150* (7), 1620-1632.

12. Daniels, M. A.; Hogquist, K. A.; Jameson, S. C., Sweet 'n' sour: the impact of differential glycosylation on T cell responses. *Nat Immunol* **2002,** *3* (10), 903-10.

13. Garcia, G. G.; Berger, S. B.; Sadighi Akha, A. A.; Miller, R. A., Age-associated changes in glycosylation of CD43 and CD45 on mouse CD4 T cells. *Eur J Immunol* **2005,** *35* (2), 622-31.

14. Rosenau, J.; Grothaus, I. L.; Yang, Y.; Kumar, N. D.; Ciacchi, L. C.; Kelm, S.; Waespy, M., N-glycosylation modulates enzymatic activity of Trypanosoma congolense trans-sialidase. *J Biol Chem* **2022,** *298* (10), 102403.

15. Skropeta, D., The effect of individual N-glycans on enzyme activity. *Bioorg Med Chem* **2009,** *17* (7), 2645-53.

16. Chavaroche, A.; Cudic, M.; Giulianotti, M.; Houghten, R. A.; Fields, G. B.; Minond, D., Glycosylation of a disintegrin and metalloprotease 17 affects its activity and inhibition. *Anal Biochem* **2014,** *449*, 68-75.

17. Howard, S. C.; Wittwer, A. J.; Welply, J. K., Oligosaccharides at each glycosylation site make structure-dependent contributions to biological properties of human tissue plasminogen activator. *Glycobiology* **1991,** *1* (4), 411-8.

18. Sackstein, R.; Stowell, S. R.; Hoffmeister, K. M.; Freeze, H. H.; Varki, A., *Glycans in Systemic Physiology*. 4th ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY): 2022.

19. Barallobre-Barreiro, J.; Baig, F.; Fava, M.; Yin, X.; Mayr, M., Glycoproteomics of the Extracellular Matrix: A Method for Intact Glycopeptide Analysis Using Mass Spectrometry. *J Vis Exp* **2017,** (122).

20. Rebelo, A. L.; Chevalier, M. T.; Russo, L.; Pandit, A., Sweet tailoring of glyco-modulatory extracellular matrix-inspired biomaterials to target neuroinflammation. *Cell Reports Physical Science* **2021,** *2* (2), 100321.

21. Bourin, M. C.; Lindahl, U., Glycosaminoglycans and the regulation of blood coagulation. *Biochem J* **1993,** *289 (Pt 2)* (Pt 2), 313-30.

22. Ma, C.; Liu, D.; Li, D.; Zhang, J.; Xu, X. Q.; Zhu, H.; Wan, X. F.; Miao, C. H.; Konkle, B. A.; Onigman, P.; Xiao, W.; Li, L., Comprehensive N- and O-glycosylation mapping of human coagulation factor V. *J Thromb Haemost* **2020,** *18* (8), 1884-1892.

23. Finke, J. M.; Ayres, K. R.; Brisbin, R. P.; Hill, H. A.; Wing, E. E.; Banks, W. A., Antibody blood-brain barrier efflux is modulated by glycan modification. *Biochim Biophys Acta Gen Subj* **2017,** *1861* (9), 2228-2239.

24. Chiricozzi, E.; Di Biase, E.; Lunghi, G.; Fazzari, M.; Loberto, N.; Aureli, M.; Mauri, L.; Sonnino, S., Turning the spotlight on the oligosaccharide chain of GM1 ganglioside. *Glycoconj J* **2021,** *38* (1), 101-117.

25. Nardy, A. F.; Freire-de-Lima, L.; Freire-de-Lima, C. G.; Morrot, A., The Sweet Side of Immune Evasion: Role of Glycans in the Mechanisms of Cancer Progression. *Front Oncol* **2016,** *6*, 54.

26. Taniguchi, N.; Kizuka, Y., Glycans and cancer: role of N-glycans in cancer biomarker, progression and metastasis, and therapeutics. *Adv Cancer Res* **2015,** *126*, 11-51.

27. Dube, D. H.; Bertozzi, C. R., Glycans in cancer and inflammation--potential for therapeutics and diagnostics. *Nat Rev Drug Discov* **2005,** *4* (6), 477-88.

28. Li, Y.; Liu, D.; Wang, Y.; Su, W.; Liu, G.; Dong, W., The Importance of Glycans of Viral and Host Proteins in Enveloped Virus Infection. *Front Immunol* **2021,** *12*, 638573.

29. Miller, N. L.; Clark, T.; Raman, R.; Sasisekharan, R., Glycans in Virus-Host Interactions: A Structural Perspective. *Front Mol Biosci* **2021,** *8*, 666756.

30. Dugan, A. E.; Peiffer, A. L.; Kiessling, L. L., Advances in glycoscience to understand viral infection and colonization. *Nat Methods* **2022,** *19* (4), 384-387.

31. Imberty, A.; Varrot, A., Microbial recognition of human cell surface glycoconjugates. *Curr Opin Struct Biol* **2008,** *18* (5), 567-76.

32. Dammen-Brower, K.; Epler, P.; Zhu, S.; Bernstein, Z. J.; Stabach, P. R.; Braddock, D. T.; Spangler, J. B.; Yarema, K. J., Strategies for Glycoengineering Therapeutic Proteins. *Front Chem* **2022,** *10*, 863118.

33. Edwards, E.; Livanos, M.; Krueger, A.; Dell, A.; Haslam, S. M.; Mark Smales, C.; Bracewell, D. G., Strategies to control therapeutic antibody glycosylation during bioprocessing: Synthesis and separation. *Biotechnol Bioeng* **2022,** *119* (6), 1343-1358.

34. Costa, A. F.; Campos, D.; Reis, C. A.; Gomes, C., Targeting Glycosylation: A New Road for Cancer Drug Discovery. *Trends Cancer* **2020,** *6* (9), 757-766.

35. Lu, R. M.; Hwang, Y. C.; Liu, I. J.; Lee, C. C.; Tsai, H. Z.; Li, H. J.; Wu, H. C., Development of therapeutic antibodies for the treatment of diseases. *J Biomed Sci* **2020,** *27* (1), 1.

36. Vong, K.; Yamamoto, T.; Tanaka, K., Artificial Glycoproteins as a Scaffold for Targeted Drug Therapy. *Small* **2020,** *16* (27), e1906890.

37. Valverde, P.; Arda, A.; Reichardt, N. C.; Jimenez-Barbero, J.; Gimeno, A., Glycans in drug discovery. *Medchemcomm* **2019,** *10* (10), 1678-1691.

38. Neelamegham, S.; Aoki-Kinoshita, K.; Bolton, E.; Frank, M.; Lisacek, F.; Lutteke, T.; O'Boyle, N.; Packer, N. H.; Stanley, P.; Toukach, P.; Varki, A.; Woods, R. J.; Group, S. D., Updates to the Symbol Nomenclature for Glycans guidelines. *Glycobiology* **2019,** *29* (9), 620-624.

39. Hegar, B.; Wibowo, Y.; Basrowi, R. W.; Ranuh, R. G.; Sudarmo, S. M.; Munasir, Z.; Atthiyah, A. F.; Widodo, A. D.; Supriatmo; Kadim, M.; Suryawan, A.; Diana, N. R.; Manoppo, C.; Vandenplas, Y., The Role of Two Human Milk Oligosaccharides, 2'-Fucosyllactose and Lacto-N-Neotetraose, in Infant Nutrition. *Pediatr Gastroenterol Hepatol Nutr* **2019,** *22* (4), 330-340.

40. Bode, L., Human milk oligosaccharides: every baby needs a sugar mama. *Glycobiology* **2012,** *22* (9), 1147-62.

41. Duijts, L.; Jaddoe, V. W.; Hofman, A.; Moll, H. A., Prolonged and exclusive breastfeeding reduces the risk of infectious diseases in infancy. *Pediatrics* **2010,** *126* (1), e18-25.

42. Walsh, C.; Lane, J. A.; van Sinderen, D.; Hickey, R. M., Human milk oligosaccharides: Shaping the infant gut microbiota and supporting health. *J Funct Foods* **2020,** *72*, 104074.

43. Moore, R. E.; Xu, L. L.; Townsend, S. D., Prospecting Human Milk Oligosaccharides as a Defense Against Viral Infections. *ACS Infect Dis* **2021,** *7* (2), 254-263.

44. Morrow, A. L.; Ruiz-Palacios, G. M.; Jiang, X.; Newburg, D. S., Human-milk glycans that inhibit pathogen binding protect breast-feeding infants against infectious diarrhea. *J Nutr* **2005,** *135* (5), 1304-7.

45. Morozov, V.; Hansman, G.; Hanisch, F. G.; Schroten, H.; Kunz, C., Human Milk Oligosaccharides as Promising Antivirals. *Mol Nutr Food Res* **2018,** *62* (6), e1700679.

46. Ayechu-Muruzabal, V.; van Stigt, A. H.; Mank, M.; Willemsen, L. E. M.; Stahl, B.; Garssen, J.; Van't Land, B., Diversity of Human Milk Oligosaccharides and Effects on Early Life Immune Development. *Front Pediatr* **2018,** *6*, 239.

47. Shental-Bechor, D.; Levy, Y., Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc Natl Acad Sci U S A* **2008,** *105* (24), 8256-61.

48. Molinari, M., N-glycan structure dictates extension of protein folding or onset of disposal. *Nat Chem Biol* **2007,** *3* (6), 313-20.

49. Roth, J.; Zuber, C.; Park, S.; Jang, I.; Lee, Y.; Kysela, K. G.; Le Fourn, V.; Santimaria, R.; Guhl, B.; Cho, J. W., Protein N-glycosylation, protein folding, and protein quality control. *Mol Cells* **2010,** *30* (6), 497-506.

50. Fukuda, M., Roles of mucin-type O-glycans in cell adhesion. *Biochim Biophys Acta* **2002,** *1573* (3), 394-405.

51. Chia, J.; Tham, K. M.; Gill, D. J.; Bard-Chapeau, E. A.; Bard, F. A., ERK8 is a negative regulator of O-GalNAc glycosylation and cell migration. *Elife* **2014,** *3*, e01828.

52. Freitas, D.; Campos, D.; Gomes, J.; Pinto, F.; Macedo, J. A.; Matos, R.; Mereiter, S.; Pinto, M. T.; Polonia, A.; Gartner, F.; Magalhaes, A.; Reis, C. A., O-glycans truncation modulates gastric cancer cell signaling and transcription leading to a more aggressive phenotype. *EBioMedicine* **2019,** *40*, 349-362.

53. Takakura, D.; Ohashi, S.; Kobayashi, N.; Tokuhisa, M.; Ichikawa, Y.; Kawasaki, N., Targeted O-glycoproteomics for the development of diagnostic markers for advanced colorectal cancer. *Front Oncol* **2023,** *13*, 1104936.

54. Chandler, K.; Goldman, R., Glycoprotein disease markers and single protein-omics. *Mol Cell Proteomics* **2013,** *12* (4), 836-45.

55. Radhakrishnan, P.; Dabelsteen, S.; Madsen, F. B.; Francavilla, C.; Kopp, K. L.; Steentoft, C.; Vakhrushev, S. Y.; Olsen, J. V.; Hansen, L.; Bennett, E. P.; Woetmann, A.; Yin, G.; Chen, L.; Song, H.; Bak, M.; Hlady, R. A.; Peters, S. L.; Opavsky, R.; Thode, C.; Qvortrup, K.; Schjoldager, K. T.; Clausen, H.; Hollingsworth, M. A.; Wandall, H. H., Immature truncated O-glycophenotype of cancer directly induces oncogenic features. *Proc Natl Acad Sci U S A* **2014,** *111* (39), E4066-75.

56. Thompson, N.; Wakarchuk, W., O-glycosylation and its role in therapeutic proteins. *Biosci Rep* **2022,** *42* (10).

57. Lujan, T. J.; Underwood, C. J.; Jacobs, N. T.; Weiss, J. A., Contribution of glycosaminoglycans to viscoelastic tensile behavior of human ligament. *J Appl Physiol (1985)* **2009,** *106* (2), 423-31.

58. Lovekamp, J. J.; Simionescu, D. T.; Mercuri, J. J.; Zubiate, B.; Sacks, M. S.; Vyavahare, N. R., Stability and function of glycosaminoglycans in porcine bioprosthetic heart valves. *Biomaterials* **2006,** *27* (8), 1507-18.

59. Santarella, F.; Simpson, C. R.; Lemoine, M.; McGrath, S.; Cavanagh, B.; Smith, A.; Murphy, C. M.; Garlick, J. A.; O'Brien, F. J.; Kearney, C. J., The lubricating effect of iPS-reprogrammed fibroblasts on collagen-GAG scaffolds for cartilage repair applications. *J Mech Behav Biomed Mater* **2021,** *114*, 104174.

60. Hayes, A. J.; Melrose, J., Glycosaminoglycan and Proteoglycan Biotherapeutics in Articular Cartilage Protection and Repair Strategies: Novel Approaches to Visco-supplementation in Orthobiologics. *Advanced Therapeutics* **2019,** *2* (8), 1900034.

61. Warnke, S.; Ben Faleh, A.; 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* (0), 114-125.

62. 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.

63. Ben Faleh, A.; Warnke, S.; Van Wieringen, T.; Abikhodr, A. H.; Rizzo, T. R., New Approach for the Identification of Isobaric and Isomeric Metabolites. *Anal Chem* **2023**.

Chapter 2 Current approaches for glycan analysis

The first step to understanding the function of glycans is to identify their primary structure. Various analytical approaches have been used to deal with their inherent isomeric complexity. The main focus of this thesis is on techniques used for determining the primary structure of released glycans – that is, those cleaved from proteins and lipids (glycoconjugates) or free glycans such as found in human milk. In this chapter, we first explain methods for releasing glycans from proteins before surveying the experimental approaches currently used for glycan analysis.

2.1 Glycan release from proteins

There are two main strategies for releasing bound glycans: enzymatic and chemical. Enzymatic release of glycans is a critical step in glycomic analysis, as it provides specific and complete sugar removal under mild conditions. The most common method for N-glycan release is peptide-N-glycosidase F (PNGase F), which results in the full release of N-glycans from glycoproteins or glycopeptides in complex mixtures. Other enzymes, such as PNGase A and PNGase H+, are available with varying levels of specificity and efficiency. Endoglycosidases, enzymes that cleave between the two GlcNac units, are also available with specificity for different types of glycans, including high-mannose, complex, and hybrid N-glycans (Fig. 1.2).[1, 2] Enzymatic release of O-glycans is more challenging, due to the diversity of core structures and the lack of enzymes with broad substrate specificity. Thus, for these glycans, chemical release is more suitable. This involves using various methods, such as reductive β-elimination, nonreductive β-elimination, and hydrazinolysis.[3-5] These methods can result in the release of N- and O-glycans for subsequent labeling or analysis using MS-based methods. However, chemical release methods often cause peeling of the released glycans, resulting in altered compositions and poor reproducibility. Chemical release can also introduce degradation of the polypeptide chain, making it impossible to analyze the deglycosylated proteins for glycoproteomic study. A recent approach called oxidative release of natural glycans (ORNG) has been reported to liberate all types of glycans from glycoproteins and glycosphingolipids.[6]

Released and free glycan samples from mammalian systems are often in a matrix consisting mainly of proteins, peptides, and lipids. The most widely used glycan isolation and purification

method is solid phase exchange (SPE), specifically C18-SPE and porous graphitic carbon (PGC)-SPE cartridges. The C18 cartridge is mainly used because of its ability to retain non-polar molecules such as proteins and peptides, allowing their separation from glycans. After removing the proteins, a glycan sample can be cleaned using a PGC cartridge, which removes all the salts and small molecules that may have been introduced from the modifications or cleavage chemistry applied to the sample. [7]

2.2 Glycan separation methods

The similarity and complexity of isomeric glycan structures often impede one's ability to discriminate and identify them. This complexity can be reduced by using a method that can separate isomers before identifying them. The most widely used separation techniques are liquid chromatography (LC), capillary electrophoresis (CE), and ion mobility spectrometry (IMS).

2.2.1 Liquid chromatography

Liquid chromatography (LC) is a powerful analytical technique that is used to separate, identify, and quantify chemical compounds in complex mixtures. It is based on separating analytes carried in a mobile liquid phase through their difference in interaction with a stationary phase, based on their chemical and physical properties such as size, shape, polarity, or charge.[8, 9] The stationary phase is typically a solid material, such as silica or alumina, packed into a column. The mobile phase is a single solvent or a mixture of solvents. As the sample passes through the column, the components of the sample are separated and identified by the time it takes for them to elute, allowing for the selective isolation and purification of individual components of a complex mixture.[10]

The three main chromatographic methods used for glycan analysis are Hydrophilic Interaction Liquid Interaction Chromatography (HILIC), Reversed Phase Liquid Chromatography (RPLC), and Porous Graphitized Carbon Chromatography (PGC).[11]

- Hydrophilic Interaction Liquid Interaction Chromatography [12]

Hydrophilic Interaction Liquid Chromatography (HILIC) excels in analyzing complex glycan samples due to its ability to separate hydrophilic compounds. Recent advancements in HILIC materials, such as amide-based stationary phases, have increased selectivity for sialylated glycopeptides and retention of neutral ones, boosting its overall efficiency. Derivatization of glycans with 2-aminobenzoic acid (2-AA)/2-aminobenzamide (2-AB) or 8-aminonaphthalene-

1,3,6-trisulfonic acid (ANTS) is typically employed to enhance their detectability and improve their retention behavior in the HILIC column.

- Reversed Phase Liquid Chromatography (RPLC)[13]

Reversed-Phase Liquid Chromatography (RPLC) has limited selectivity for glycans due to interactions with the aromatic label, requiring shallow gradients and resulting in low peak capacity. However, its grouping ability and compatibility with UHPLC materials make it suitable for quality control, particularly for determining galactosylation, sialylation, fucosylation, and mannosylation in mAb glycan samples. RPLC is robust, reproducible, and versatile, with no restrictions on aqueous sample injection volume. Compatibility with mass spectrometry is improved by using a weaker organic buffer, while derivatization is often needed for glycan analysis, similar to HILIC.

- Porous Graphitized Carbon Chromatography (PGC)[11]

Porous Graphitized Carbon (PGC) chromatography offers higher resolution for glycan isomers compared to HILIC, but its separation mechanism remains unclear, hindering the prediction of glycan retention times using in silico models. PGC's mixed retention properties involve hydrophobic, polar, and ionic interactions, providing increased selectivity for sialylated glycans and superior peak capacity with steeper gradients. It's valuable for high-resolution separation of glycan structures and discovering novel ones. However, limited implementation is due to the lack of established retention values for normalizing technical variation. PGC is a versatile tool for various applications, including glycan analysis, metabolomics, and proteomics.

2.2.2 Capillary electrophoresis

Capillary electrophoresis (CE) is an analytical separation technique often paired with mass spectrometry (MS) for glycan analysis. *[14]* CE separates glycans based on charge and size within a capillary tube and has three main types: capillary zone electrophoresis (CZE), capillary gel electrophoresis (CGE), and capillary isoelectric focusing (CIEF).*[15-17]* The main separation mechanisms in CE are electroosmotic flow (EOF) and electrophoretic mobility, where electrophoretic mobility depends on the volume-to-charge ratio of the analytes, and EOF is the flow of the background electrolyte from the anode to the cathode. The separations are achieved through the superposition of these two phenomena. Carbohydrate separations can be challenging, due to their generally uncharged nature and lack of a chromophore, but labeling

with a fluorophore overcomes these limitations. When combined with MS, CE offers valuable insights into glycan structure.

2.2.3 Ion mobility spectrometry (IMS)

Ion mobility spectrometry (IMS) is an analytical technique used for the separation, identification, and quantification of ions in the gas phase. It is a widely used method in fields such as environmental monitoring, forensic analysis, and detection of hazardous substances. The principle of ion mobility spectrometry is based on the differential movement of ions in a gas under the influence of an electric field.[18]

One way to classify mobility measurement approaches is based on how ions are physically separated - either in time or in space. Time-dispersive techniques separate ions based on the time they spend in the mobility analyzer. Two examples of this are drift tube IMS (DTIMS) and traveling wave IMS (TWIMS). On the other hand, spatially dispersive approaches give ions different trajectories based on their mobility and are often used as filters. An example of this is field-asymmetric waveform IMS.[19] Trapped ion mobility spectrometry separates ions spatially in a trapping region, then releases them based on their mobility.[20]

The collisional cross section (CCS) is an important parameter in IMS, as it is often used as a means of identification. The CCS of an ion is a measure of its size and shape and represents the effective area that the ion presents to a gas molecule during a collision, expressed in units of Å². Calibration of different IMS instruments is performed by measuring the drift times or arrival times of reference standards of known size and shape. The accurate determination of CCS values is critical for the interpretation and comparison of IMS data and its use as an analytical technique. The measurement and calculation of CCS values is an active area of research, and thus various experimental and theoretical methods have been developed for this purpose.[21, 22]

The technique that is best suited to determining collision cross sections is DTIMS. In DTIMS, the mobility (K) of an ion is defined as the ratio of the ion's drift velocity (v_d) to the electric field strength (E) applied to the traveling ions.

$$K = \frac{v_d}{E} \qquad \text{Equation 2.1}$$

The mobility is usually expressed in units of $cm^2\ V^{-1}\ s^{-1}$. To more uniformly represent the mobility across different DTIMS instruments under different conditions, the reduced mobility

(K_0) is typically used. It is defined as the mobility of an ion under standard conditions of temperature (T) and pressure (P), usually at 1 atm and 25°C.

$$K_0 = \frac{P}{P_0}\frac{T_0}{T}K \qquad \text{Equation 2.2}$$

The measured mobility is then converted into a CCS using the Mason-Schamp equation:

$$\Omega = \frac{3}{16N}\left(\frac{2\pi}{\mu k_b T}\right)^{\frac{1}{2}}\frac{ze}{K_0} \qquad \text{Equation 2.3}$$

where N is the buffer gas density, ze is the charge of the ion, μ is the reduced mass of the ion and buffer gas collision, and k_b is the Boltzmann constant. Collisional cross sections are tabulated in databases and used to identify compounds.

In the following I introduce the general workings of some of the most widely used IMS techniques:

1. Drift tube ion mobility spectrometry (DTIMS) [23-25]

   In DTIMS, a packet of ions are introduced into a drift tube filled with a stationary buffer gas and subjected to a weak uniform electric field. The electric field causes ions to experience a net drift velocity that is proportional to their mobility, allowing ions with different mobilities to be separated. The drift velocity (v_d) required for the mobility calculation, and subsequently CCS calculation, is defined as the quotient of the length of the drift tube (l) and the time it takes the ions to traverse this length (t_d, drift time). This measurement of detected ions versus time gives an arrival time distribution (ATD). (Figure 2.1)

   The primary benefit of utilizing DTIMS lies in its intrinsic operational properties-specifically, the implementation of a uniform electric field and a relatively uncomplicated calibration protocol. These attributes enhance its ability to generate precise CCS values. Nevertheless, since the resolving power is proportional to the square root of the drift length, achieving higher resolution using DTIMS presents certain challenges, necessitating the exploration of alternative strategies for resolution enhancement.

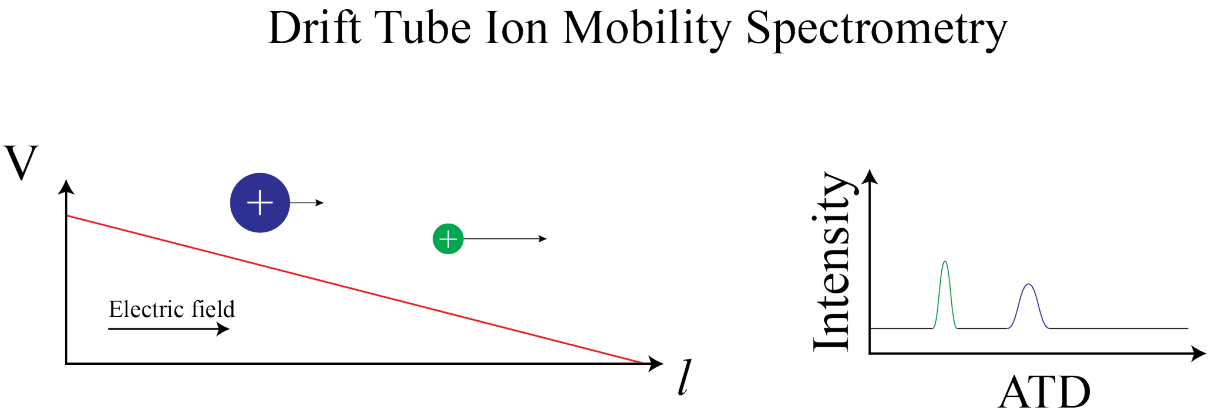


Figure 2.1. Schematic representation of the DTIMS platform.

2. Traveling wave ion mobility spectrometry (TWIMS)

In TWIMS, ions are introduced into a drift region filled with a buffer gas and subjected to a series of high-frequency, low-amplitude traveling waves that traverse the length of the drift region.[26, 27] These waves are created using a set of DC and RF voltages that confine and push the ions along a set path. The mode of separation here is also based on collisions with the buffer gas, where ions with smaller CCS encounter fewer collisions and can thus ride the wave, whereas ions with larger CCS experience more collisions, thus impeding their travel, causing them to 'roll over' the waves, and consequently, extending their drift time to the detector. (Figure 2.2)

Obtaining CCS values using TWIMS experiments is not as straightforward as in DTIMS. This is primarily because the trajectories ions take while riding the travelling wave are difficult to predict. Consequently, CCS determination necessitates the calibration of the instrument utilizing molecules of known mobility, preferably belonging to a class of molecules analogous to those being analyzed.

The ability to induce directional changes of ions through the traveling wave allows for ions to be manipulated along drift lengths that can go from centimeters to hundreds of meters with minimal ion loss, greatly increasing the resolving power. There are two main technologies that utilize this approach to develop ultrahigh resolution IMS devices.

Waters Corporation has developed a commercial cyclic IMS device (cIMS) that is based on TWIMS and can achieve long path lengths by having ions go around a circular electrode setup.[28] Smith and coworkers at Pacific Northwest National Laboratory (PNNL) developed

a device called structures for lossless ion manipulation (SLIM), also based on TWIMS, which make ions follow a serpentine path. [29, 30] Since the work undertaken in this thesis uses a home-built SLIM instrument for glycan separation, a more detailed description of SLIM devices is provided in the following chapter.

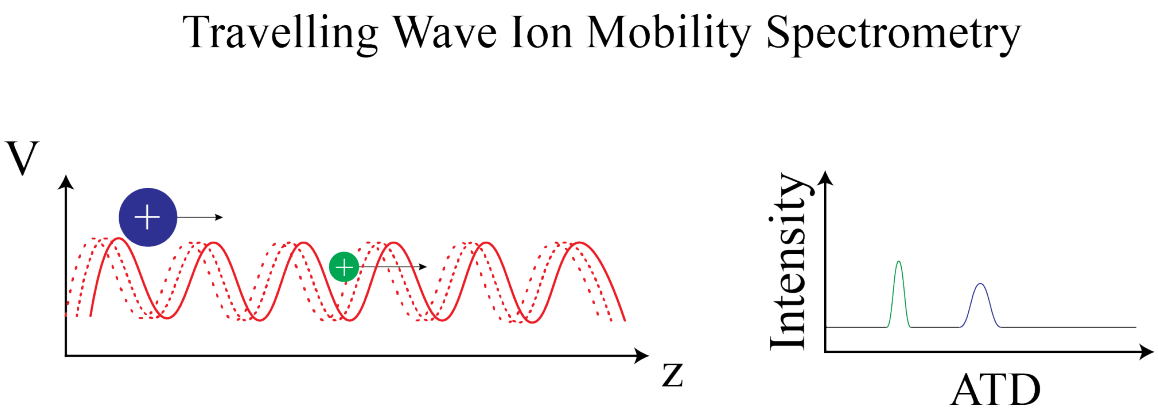


Figure 2.2. Schematic representation of the TWIMS platform.

3. Trapped ion mobility spectrometry (TIMS)

In TIMS, ions are introduced into a device with a series of electrodes that trap the ions in a confined space through a combination of RF potentials in the radial direction and a DC voltage ramp in the axial direction along which separation occurs. As opposed to DTIMS and TWIMS where the buffer gas is stationary, in TIMS, a gas flow propels the ions in the opposite direction to the DC electric field ramp, separating them in space along the axial direction. The ions are then released by lowering the height of the ramp, as shown schematically in in Figure 2.3.[20, 31] Similar to TWIMS, the extraction of CCS values from TIMS requires the use of a calibrant.

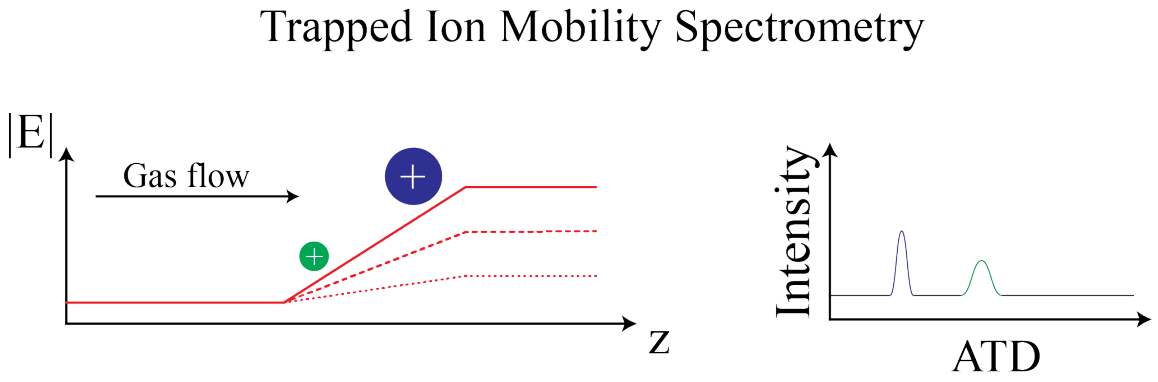


Figure 2.3. Schematic representation of the TIMS platform.

4. High-field asymmetric waveform ion mobility spectrometry (FAIMS)

In FAIMS, illustrated in Figure 2.4, ions are introduced into a gas-filled chamber kept close to atmospheric pressure with a non-uniform electric field that oscillates between opposite polarities of different amplitude. The electric field causes ions to undergo a transverse velocity that is proportional to their difference in mobility between high and low fields, allowing them to be separated based on their different trajectories through the chamber, as shown in Figure 2.4(a).[32, 33] A compensation voltage (CV) is then overlaid onto the oscillating asymmetric waveform. Depending on the CV, the trajectory of only certain ions will permit their release while eliminating others.(Figure 2.4 (b)). Since FAIMS operates at high electric fields compared to other mobility methods, a CCS cannot be obtained. Thus, this technique is used primarily as ion filter.

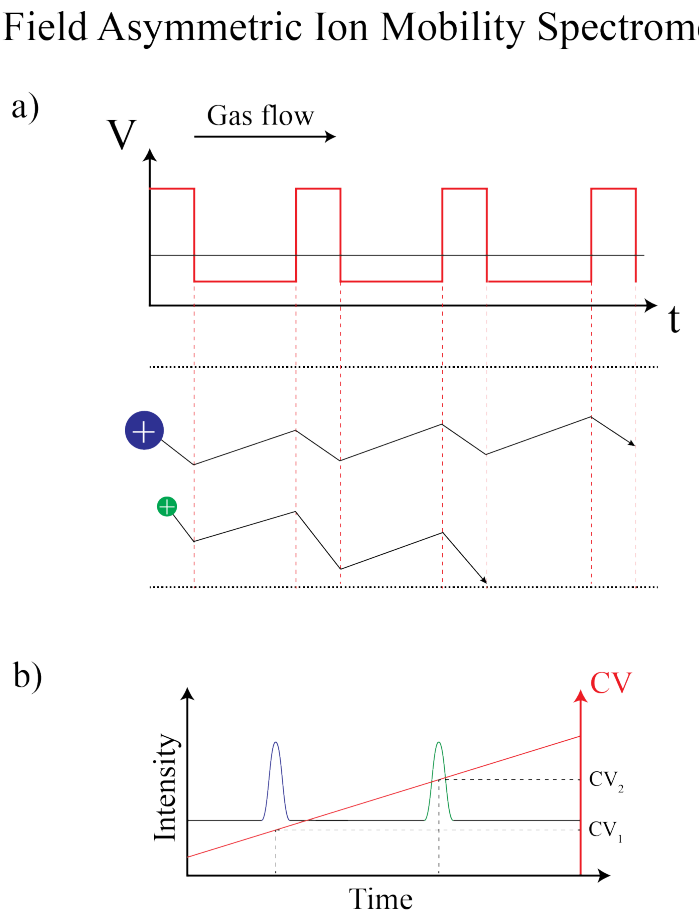


Figure 2.4. (a) Schematic representation of the FAIMS platform. (b) graph representing how molecules are released as a function of compensation voltage (CV) applied.

The potential of IMS in biomolecular analysis has been accelerated by the development of ultrahigh-resolution devices that allow the separation of isomeric species. However, the CCS is not an intrinsic property of an ion but is dependent on the drift gas used as well as the conditions of the experiment. While the increase in resolution allows the separation of isomers with only a slight difference in mobility, identification of complex molecules based upon a single number (i.e., the CCS) can be problematic. Moreover, even when high resolution IMS techniques can be calibrated with molecules of known CCS, the ability to calculate the CCS to the same resolution is beyond current computational capabilities. For this reason, in our experiments we use high-resolution IMS only as a separation tool and not an identification method.

2.3 Glycan identification

While the separation techniques described above are also used in many cases to identify glycans , combining these separation methods with spectrometric or spectroscopic methods can provide a much more powerful means of glycan identification. We briefly describe here several such methods that can be used for glycan identification either alone or in combination with separation techniques.

2.3.1 Nuclear magnetic resonance (NMR)

Nuclear magnetic resonance (NMR) is a powerful technique for glycan structural characterization.[34] The assignment of ^{1}H- and ^{13}C-NMR spectra is necessary for complete structural elucidation of a glycan. This is achieved by combining two-dimensional NMR methods that allow for the assignment of individual ^{1}H signals of monosaccharide residues, such as correlation spectroscopy (COSY) and total correlation spectroscopy (TOCSY).[35, 36] Afterward, the assignment of the ^{13}C spectrum can be extended using heteronuclear single-quantum coherence (HSQC) techniques.[37] Two-dimensional heteronuclear multiple-bond correlation (HMBC) experiments, which find connections between the proton on the anomeric carbon and the carbon atom on the other side of the glycosidic linkage, is crucial for sequencing.[38] NMR experiments are performed on samples in the liquid phase and is non-destructive. However, large quantities of pure glycan samples are often required for NMR identification. Since these quantities are often unavailable, more sensitive (gas-phase) methods are usually preferred.

2.3.2 Mass spectrometry

The most sensitive and widely used approach for glycan analysis is mass spectrometry (MS). The first step in glycan analysis by MS is to ionize the molecules. Electrospray ionization (ESI) and matrix assisted laser desorption/ionization (MALDI) are the most frequently used ionization techniques for oligosaccharides.[39, 40] Compared to earlier techniques for ionizing carbohydrates, such as electron impact ionization or fast atom bombardment, they apply little energy to the sample, leading to less fragmentation during the ionization process. Depending on the sample, ions can be produced in either a positive or negative ion mode. Negative ion mode facilitates the analysis of oligosaccharides with acidic groups (sulfate, carboxylate, or phosphate). For native oligosaccharides, both ionization modes are typically utilized.

To produce ions by MALDI, the sample is dried on a metal target while surrounded by a chromophoric matrix, causing matrix crystals to form that contain trapped sample molecules. The dry mixture patch is then exposed to an ultraviolet or infrared laser pulse, and the matrix absorbs part of the energy and transmits it to the analyte, which ionizes.[41, 42] With ESI, analyte ions are produced at atmospheric pressure by passing a solution through a needle directed towards a nearby capillary, at a voltage of 1-4 kV. The potential difference between the capillary and the tip produces finely charged droplets.[43] When the ions are sucked into the mass spectrometer, heated drying gas evaporates the solvent from the ions. Using this technique, and depending on the solution conditions, both positive and negative ions can be produced. ESI works well with online liquid flow separation techniques like liquid chromatography (LC) or capillary electrophoresis (CE).[44, 45] With nanoelectrospray (nESI), the smaller droplets produced increases sensitivity and demonstrates increased tolerance to salts and other contaminants.[46, 47]

After ionization, single stage mass spectrometry (MS) can identify the number of the various types of monosaccharides through an accurate mass measurement. However, it remains blind to structural details such as the specific monosaccharide composition, sequence, branching, and stereochemistry. It is possible to gain more information through multistage MS (i.e., MS^n). Here, ions are mass selected in a first stage, after which they undergo activation and dissociation, followed by mass analysis of fragments. These fragments can act as a fingerprint, often containing information about the monosaccharide linkage, sequence, and other possible modifications. Many methods have been employed for fragmenting glycans, the most widely used of which is collisional induced dissociation (CID).[48]

In CID, the molecule is activated through the collisions with an inert buffer gas, usually nitrogen or argon. When an ion is collisionally activated in a CID process, the energy transferred is distributed among the various vibrational degrees of freedom of the molecule, leading to fragmentation. (Figure 2.5)

The two main types of CID can be distinguished.

1. **High-energy CID**. In this case, ions are accelerated to high translational energy and activated by a single collision. The substantial energy transferred in high-energy CID can allow the molecule to overcome multiple energy barriers on the PES, leading to broad range of fragment ions.

2. **Low-energy CID**: In this approach, molecules are activated by multiple lower energy collisions. In this case, the energy tends to be statistically distributed among the internal vibrational modes of the ions, leading to dissociation via the lowest energy pathway.

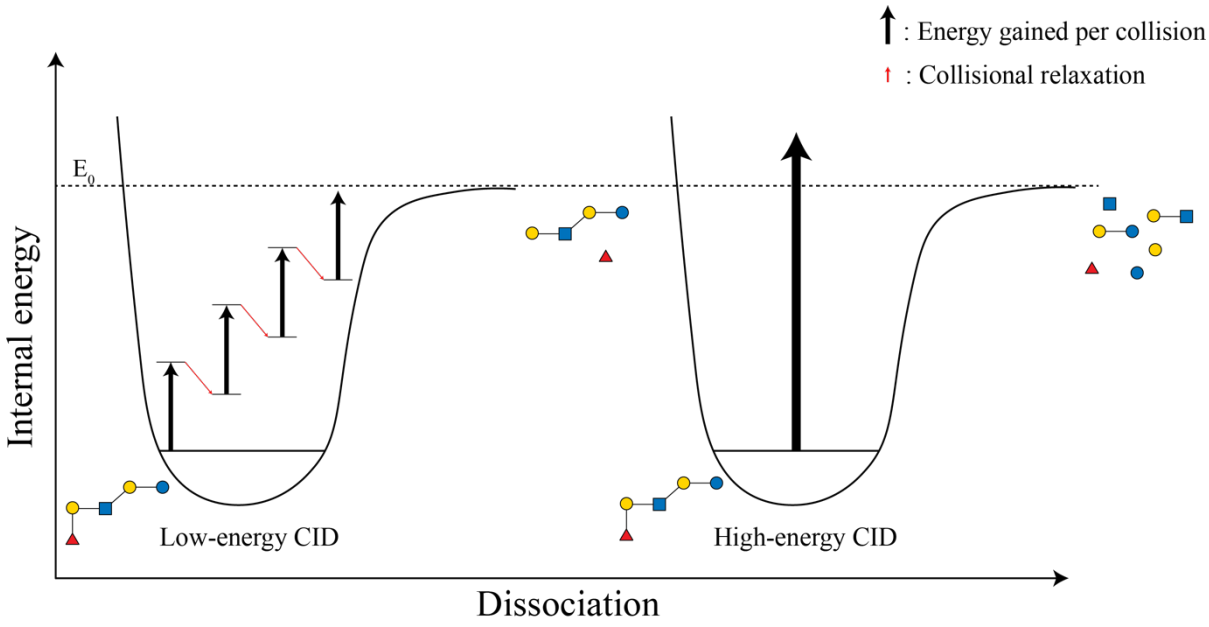


Figure 2.5: Schematic representing the mechanism of the collisional activation of a molecule by low-energy and high-energy CID.

The specific fragmentation pathway during a collision-induced dissociation (CID) process depends upon the molecule's potential energy surface (PES), the excitation energy, and the redistribution of internal energy among its various degrees of freedom.

In a mass spectrometer, the collision energy is predominantly determined by the precursor ion's translational energy prior to the collision, which depends upon the ion acceleration voltage and

the distance it can be accelerated before having a collision. The ion's mean free path (λ), or the average distance travelled between collisions with a gas molecule, thus plays a significant role in the fragmentation process.

The mean free path (λ) of an ion is given by:

$$\lambda = \frac{k_B T}{\sqrt{2}\pi d^2 N_A P} \qquad \text{Equation 2.4}$$

where k_B is the Boltzmann constant, T is the temperature of the gas, d is the diameter of the molecule, N_A is Avogadro's number, and P is the pressure of the buffer gas in the collision cell.

When an ion of charge q is accelerated through a potential difference, the kinetic energy gained by the ion is given by the equation:

$$E = |qV| \qquad \text{Equation 2.5}$$

The energy E in electron volts (eV) is equal to the potential difference V (in volts) if the charge q is a single charge (negative or positive).

In a disaccharide molecule, the activation energy requisite for the cleavage of the glycosidic bond is around 59.7 kcal/mol,[49] or equivalently 2.6 electron volts (eV). Consequently, for the disaccharide to undergo fragmentation, the ion must be supplied with a minimum of 2.6 eV energy to overcome this activation barrier. This energy is delivered by modulating the parameters of the collision cell, specifically pressure and voltage acceleration.

As will be detailed in Chapter 3, the region of our instrument employed for CID maintains an average pressure of 3 mbar and applies an accelerating voltage of approximately 250 V across a grid separation of 0.8 mm. This results in an accelerating electric field of ~ 310 V/mm, which corresponds to a collisional energy of 0.3 eV.

Considering a disaccharide ion with a diameter (d) of 2 nm, its mean free path, as calculated from equation 2.4, is approximately 1 μm. Therefore, to accumulate the necessary 2.6 eV to cleave the glycosidic bond, multiple collisions are required, which puts it in the category of low-energy CID. High-energy CID conditions can be achieved by lowering the pressure to lengthen the mean free path.

CID of glycans typically cleaves the weakest bonds, which are the glycosidic bonds that link monosaccharides together. As a result, the fragmentation primarily yields smaller saccharide subunits. In this identification procedure, the mass-to-charge ratio of both the precursor molecule and the derived fragments aids in elucidating the structure and sequence of the glycan under study.

Depending on the molecule's charge state, CID can sometimes lead to rearrangement of the monosaccharide sequence. For example, in the protonated charge state, fragments formed by CID can undergo fucose migration. leading to a wrong assignment of the precursor structure.[50] However, this rearrangement is avoided if the glycan ions are formed with metal adducts. Collision-induced dissociation of positive and negative ions can lead to different fragmentation pathways that can complement and enhance structural assignment.

To obtain more detailed information about structural branching and residue connectivity, more energetic fragmentation methods that cause cross-ring fragmentation are typically be used. This can be achieved using electron based methods, such as electron capture dissociation (ECD),[51] or through photodissociation methods, such as ultraviolet photodissociation (UVPD).[52] ECD is performed by directing multiply charged ions into a trap where they are subjected to low energy electrons. Capturing an electron promotes them to a highly excited electronic state, leading to rapid dissociation, often including cross-ring fragments that can hold information about the molecule's connectivity and stereochemistry. Similarly, cross ring fragmentation can be achieved using UVPD, since the higher energy imparted to the precursor molecule can lead to the breaking of multiple chemical bonds.

Since many monosaccharide units just differ by the orientation of the hydroxy groups (Fig. 1.3), fragmentation approaches, even (MS^n), have difficulty distinguishing subtly different isomers.

Liquid chromatography (LC) combined with tandem mass spectrometry (LC-MS^n) is the most widely used method for glycan analysis. The LC dimension allows the separation of isomeric species and provides a retention time metric that can aid in identification before analysis is performed in by MS. However, most LC-based separation methods require that the glycan be derivatized before analysis.[53]

2.3.3 Vibrational spectroscopy

The basis of vibrational spectroscopy lies in the quantum mechanical nature of molecules. The fact that molecular vibrations occur at specific resonant frequencies reflects the quantization of the vibrational energy, and these frequencies are determined by the shape of the molecule's potential energy surface. The collection of resonant frequencies comprises a vibrational spectrum, which can serve as a unique molecular fingerprint that reflects functional groups present in a molecule, their interactions, and their arrangements in space.

The nature of the spectrum can be strongly influenced by the environment in which it is measured. In solution phase spectroscopy, the molecules of interest are dissolved in a suitable solvent, but interactions with the solvent often leads to line shifts and broadening. This reduces the resolution achievable, which can wash out the uniqueness of a spectrum and its ability to serve as a molecular fingerprint. On the other hand, vibrational spectra of isolated molecules in the gaseous state are free from these types of broadening mechanisms and thus can serve as a robust molecular identifier. As discussed below, the resolution is even further increased if the analyte molecules are cooled to low temperature, as this removes thermal broadening of vibrational transitions.

When a molecule absorbs a photon at its resonant frequency, it is excited from a lower vibrational energy level to a higher one, and there are different ways to detect this absorption event and generate a spectrum. One approach is to measure the amount of light directly absorbed by a sample at different frequencies. This method often requires a significant concentration of the species of interest. For this reason, it is not possible to perform direct absorption spectroscopy for gas-phase ions, for example, where the density is extremely low. In this case, one uses indirect methods to detect the absorption in which the effect of light absorption on some physical or chemical property (i.e., the *action*) rather than the absorption itself. The *action* can be various things, such as fluorescence or dissociation. Because one only detects signals from the absorbing molecules, it has much higher sensitivity than direction absorption methods.

Infrared Multiphoton Dissociation (IRMPD) is a dissociation action spectroscopy technique that can be used in mass spectrometry for the fragmentation and structural characterization of glycans.[54-58] In IRMPD, when an IR photon has a frequency resonant with a molecular vibrational transition, it is absorbed. This absorption is thought to be an incoherent process,[59] where the energy gained is quickly dissipated into the vibrational degrees of freedom of the

molecule *via* intramolecular vibrational energy redistribution (IVR).[60] Figure 2.6 illustrates this process where each sequential absorption of a photon increases the internal energy of the molecule, until reaching a certain energy threshold where fragmentation occurs. It should be noted that the energy is stored in the bath of vibrational degrees of freedom of the molecule, which means that the photons absorbed effectively lead to a heating of the ions.[61]

Instead of following a sequential, "ladder-climbing" process (i.e., $v_0 \rightarrow v_1 \rightarrow v_2 \rightarrow \ldots$), the problem of anharmonicity bottleneck is circumvented due to IVR. However, at increased internal energies, the normal modes lose their complete orthogonality, introducing cross-anharmonicities. Although the bands undergo a less severe red-shift, they still move out of resonance with the laser frequency. This shift, however, is mitigated by a considerable increase in the density of states related to the internal energy, resulting in the broadening of the absorption lines. As the internal energy increases further, a quasi-continuum state is reached, where the absorption cross-section becomes significant across all wavelengths.

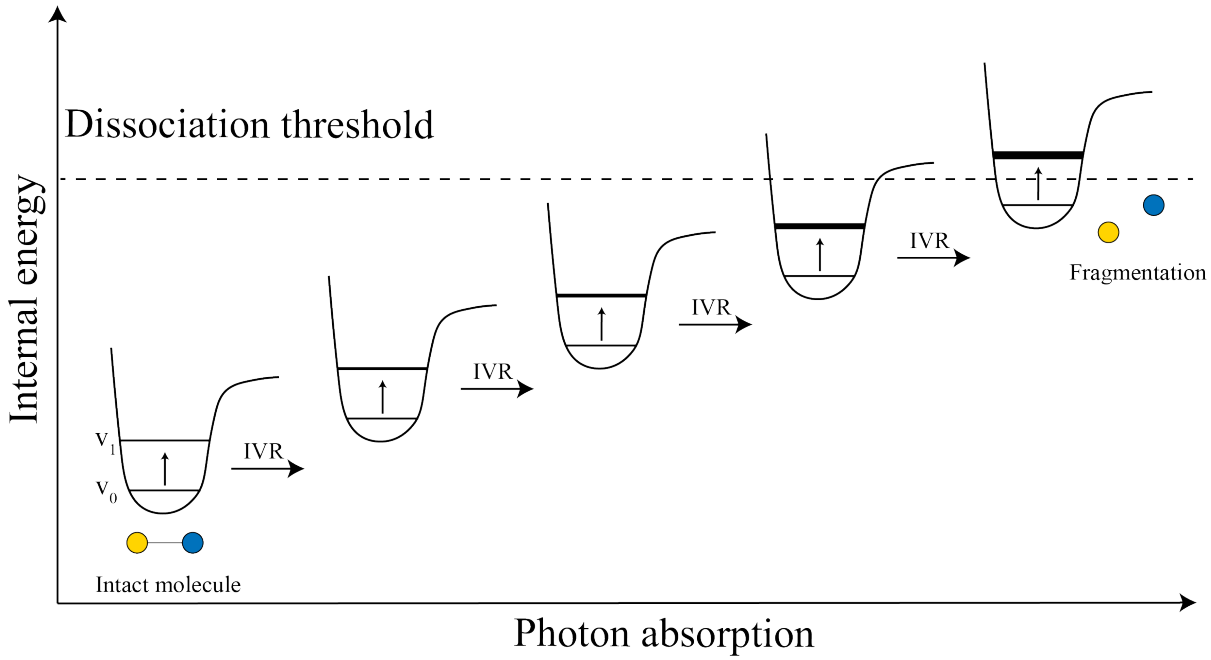


Figure 2.6. Infrared multiple-photon dissociation (IRMPD) mechanism. The internal energy of the molecule increases with the sequential absorption of multiple photons through the IVR process. Once the dissociation threshold is exceeded, the ion undergoes dissociation. The broadening of the v_1 linewidth shows the increase in the density of states.

Although the fragmentation itself can provide information on glycan structures, a vibrational spectrum of an ion can also be obtained by monitoring the formation of fragments as a function of IR laser wavenumber. This spectrum is highly unique and characteristic of the precursor

molecule. Using this technique, Compagnon et al., showed that monosaccharides produced from fragmenting disaccharides retain the stereochemistry of glycosidic bond.[62]

With increasing oligosaccharide size, IRMPD spectra become too congested due to the inhomogeneous broadening of the spectral lines at room temperature. The increasing molecule size often also leads to the existence of multiple conformations, which adds to the complexity of the spectrum obtained.

Cryogenic spectroscopy techniques can eliminate many of these issues, as at lower temperatures, inhomogeneous broadening is mostly eliminated, since the population of higher vibrational states decreases. As a result, the absorption bands become sharper and more highly resolved, which makes them a better molecular identifier.

One method to perform cryogenic spectroscopy for glycan analysis is to embed them in superfluid liquid helium droplets, which can encapsulate a wide range of molecules and ions.[63] Helium droplets are generated by allowing high-pressure helium to expand into a vacuum via a tiny, cooled nozzle. Once in vacuum, the helium droplets reach a temperature of 0.38 K achieved through evaporative cooling. These droplets can then traverse an ion trap containing biomolecules. As they do so, they capture these ions and bring them to a temperature of 0.38 K, utilizing the same principle of evaporative cooling. The encapsulated ions can then be probed using IR radiation. Absorption of an IR photon leads to ion ejection from the droplet.[64] The IR spectrum is recorded as the number of ions ejected from the helium droplet as a function of IR wavenumber. Pagel and coworkers have combined MS and ultra-cold IR spectroscopy in helium droplets to obtain well-resolved spectra for identifying oligosaccharides. [65, 66]

Another method for action spectroscopy of ions is to use messenger tagging. This is done by first cooling ions of interest until a neutral gas molecule weakly binds to them. This weakly bound messenger molecule then detaches when an IR photon is absorbed, and all this can be detected in the mass spectrometer. This is the approach used in this thesis, and thus a more detailed explanation will be given in the following chapter.

2.3.4 Hybrid approaches

Mass spectrometry (MS) is the primary technique used in glycan research; however, it falls short of complete structural elucidation. This gap is filled by a range of hybrid MS techniques, which offer a comprehensive characterization of carbohydrates. By integrating MS with

separation procedures, ionization efficiency is significantly improved, and mass spectra are simplified – an imperative for examining complex mixtures.

2.3.4.1 LC-MS

Liquid chromatography (LC) combined with tandem mass spectrometry (LC-MS^n) is the most widely used method for glycan analysis. The LC dimension allows the separation of isomeric species and provides a retention time metric that can aid in identification before analysis is performed in the MS. However, most LC-based separation methods require that the glycan be derivatized before analysis.[53]

2.3.4.2 IMS-MS

Multidimensional platforms that combine IMS with MS further enrich glycan analysis.[67-69] In recent years, ion mobility spectrometry (IMS) has rapidly been acknowledged as a robust analytical tool capable of distinguishing numerous glycan isomers within milliseconds. IMS-MS enables the identification of isomeric structures by utilizing both the ion mobility-derived CCS and tandem-MS measurements of mass-to-charge ratios. Nevertheless, this technique still faces challenges when it comes to analyzing isomers that exhibit only slight differences in their CCS values.

2.3.4.3 Combining LC-MS and gas phase IR spectroscopy

The combination of LC-MS with gas-phase infrared (IR) spectroscopy is a powerful analytical tool that allows for simultaneous identification and quantification of complex mixtures. However, direct coupling of these two techniques can be challenging due to the vast difference in the time required for each technique. While a typical LC peak has a width of seconds as it elutes, acquiring high-quality IR spectra has traditionally required tens of minutes. Multiple methods have been employed to address this issue.

Compagnon and coworkers have utilized a stop-flow technique to perform spectroscopy of analytes as they elute from an LC column.[70] This approach involves temporarily halting the flow of the mobile phase during the LC separation process when a peak of interest is detected. By stopping the flow, the target analyte is held in the column, allowing sufficient time for gas-phase IR spectroscopy analysis. The analyte from the liquid phase is transferred to the gas

phase using electrospray ionization. These ions are introduced into the gas phase IR spectrometer for the analysis of their molecular structure and composition.

Oomens and coworkers devised an approach that incorporates flow injection with gas-phase IR spectroscopy and mass spectrometry for analyzing compounds eluting from a liquid chromatograph.[71] This method involves the automated collection of fractions of eluted analytes at specific time intervals or peak regions. The collected LC fractions are introduced into the flow injection system consisting of a syringe pump, a sample loop, and a connecting tube that transports the sample to the ionization source. The generated ions are then subjected to gas-phase infrared spectroscopy, allowing for the acquisition of vibrational spectra providing detailed structural information about the analyte, such as functional groups and molecular conformation.

The fusion of LC and gas-phase IR spectroscopy offers numerous advantages, including enhanced sensitivity, selectivity, and structural characterization. The technique allows for the simultaneous identification and quantification of complex mixtures, enabling a deeper understanding of molecular composition and effective differentiation between isomers and structurally similar compounds. The successful online coupling of these techniques can have a significant analytical impact as LC is traditionally the most widely used analytical technique. However, the approach also presents challenges, such as developing a suitable interface for efficient analyte transfer, and the ability to process complex mixtures with diverse polarities and volatilities.

2.4 The approach of this thesis

The complexity of glycans makes it challenging for a single analytical technique to fully characterize their structure. To overcome this limitation, we employ a multidimensional approach, combining high-resolution separation techniques with spectroscopic identification for glycan analysis. We use IMS or LC to separate isomers and cryogenic infrared (IR) spectroscopy together with mass spectrometry for identification.[72-74] IR spectroscopy probes structure-sensitive intramolecular interactions and provides a molecular fingerprint that is highly unique and specific to the molecule.

In the following chapter, we provide a detailed description of the techniques that we employ in our multidimensional approach.

References:

1. Grunow, D., and Blanchard, V. (2019) Enzymatic Release of Glycoprotein N-Glycans and Fluorescent Labeling, *Methods Mol Biol 1934*, 43-49.
2. Weng, Y., Sui, Z., Jiang, H., Shan, Y., Chen, L., Zhang, S., Zhang, L., and Zhang, Y. (2015) Releasing N-glycan from peptide N-terminus by N-terminal succinylation assisted enzymatic deglycosylation, *Sci Rep 5*, 9770.
3. Goso, Y., Tsubokawa, D., and Ishihara, K. (2009) Evaluation of conditions for release of mucin-type oligosaccharides from glycoproteins by hydrazine gas treatment, *J Biochem 145*, 739-749.
4. Iyer, R. N., and Carlson, D. M. (1971) Alkaline borohydride degradation of blood group H substance, *Arch Biochem Biophys 142*, 101-105.
5. Halfinger, B., Sarg, B., and Lindner, H. H. (2011) Evaluation of non-reductive beta-elimination/Michael addition for glycosylation site determination in mucin-like O-glycopeptides, *Electrophoresis 32*, 3546-3553.
6. Song, X., Ju, H., Lasanajak, Y., Kudelka, M. R., Smith, D. F., and Cummings, R. D. (2016) Oxidative release of natural glycans for functional glycomics, *Nat Methods 13*, 528-534.
7. Zhang, Q., Li, H., Feng, X., Liu, B. F., and Liu, X. (2014) Purification of derivatized oligosaccharides by solid phase extraction for glycomic analysis, *PLoS One 9*, e94232.
8. Karger, B. L. (1997) HPLC: Early and recent perspectives, *Journal of Chemical Education 74*, 45-48.
9. Pitt, J. J. (2009) Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry, *Clin Biochem Rev 30*, 19-34.
10. Dorsey, J. G., Cooper, W. T., Siles, B. A., Foley, J. P., and Barth, H. G. (1996) Liquid chromatography: Theory and methodology, *Analytical Chemistry 68*, R515-R568.
11. Ashwood, C., Pratt, B., MacLean, B. X., Gundry, R. L., and Packer, N. H. (2019) Standardization of PGC-LC-MS-based glycomics for sample specific glycotyping, *Analyst 144*, 3601-3612.
12. Buszewski, B., and Noga, S. (2012) Hydrophilic interaction liquid chromatography (HILIC)--a powerful separation technique, *Anal Bioanal Chem 402*, 231-247.

13. Vreeker, G. C., and Wuhrer, M. (2017) Reversed-phase separation methods for glycan analysis, *Anal Bioanal Chem 409*, 359-378.
14. Zaia, J. (2013) Capillary electrophoresis-mass spectrometry of carbohydrates, *Methods Mol Biol 984*, 13-25.
15. Rustandi, R. R., Anderson, C., and Hamm, M. (2013) Application of capillary electrophoresis in glycoprotein analysis, *Methods Mol Biol 988*, 181-197.
16. Schwarzer, J., Rapp, E., and Reichl, U. (2008) N-glycan analysis by CGE-LIF: profiling influenza A virus hemagglutinin N-glycosylation during vaccine production, *Electrophoresis 29*, 4203-4214.
17. Makrydaki, E., Kotidis, P., Polizzi, K. M., and Kontoravdi, C. (2021) Hitting the sweet spot with capillary electrophoresis: advances in N-glycomics and glycoproteomics, *Curr Opin Biotechnol 71*, 182-190.
18. Dodds, J. N., and Baker, E. S. (2019) Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead, *J Am Soc Mass Spectrom 30*, 2185-2195.
19. Kailemia, M. J., Park, M., Kaplan, D. A., Venot, A., Boons, G. J., Li, L., Linhardt, R. J., and Amster, I. J. (2014) High-field asymmetric-waveform ion mobility spectrometry and electron detachment dissociation of isobaric mixtures of glycosaminoglycans, *J Am Soc Mass Spectrom 25*, 258-268.
20. Michelmann, K., Silveira, J. A., Ridgeway, M. E., and Park, M. A. (2015) Fundamentals of trapped ion mobility spectrometry, *J Am Soc Mass Spectrom 26*, 14-24.
21. Plante, P. L., Francovic-Fontaine, E., May, J. C., McLean, J. A., Baker, E. S., Laviolette, F., Marchand, M., and Corbeil, J. (2019) Predicting Ion Mobility Collision Cross-Sections Using a Deep Neural Network: DeepCCS, *Anal Chem 91*, 5191-5199.
22. Glaskin, R. S., Khatri, K., Wang, Q., Zaia, J., and Costello, C. E. (2017) Construction of a Database of Collision Cross Section Values for Glycopeptides, Glycans, and Peptides Determined by IM-MS, *Anal Chem 89*, 4452-4460.
23. Zhu, F., Lee, S., Valentine, S. J., Reilly, J. P., and Clemmer, D. E. (2012) Mannose7 glycan isomer characterization by IMS-MS/MS analysis, *J Am Soc Mass Spectrom 23*, 2158-2166.

24. Masellis, C., Khanal, N., Kamrath, M. Z., Clemmer, D. E., and Rizzo, T. R. (2017) Cryogenic Vibrational Spectroscopy Provides Unique Fingerprints for Glycan Identification, *J Am Soc Mass Spectrom 28*, 2217-2222.

25. 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., and McLean, J. A. (2014) 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 86*, 2107-2116.

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

27. Shvartsburg, A. A., and Smith, R. D. (2008) Fundamentals of traveling wave ion mobility spectrometry, *Anal Chem 80*, 9689-9699.

28. Giles, K., Ujma, J., Wildgoose, J., Pringle, S., Richardson, K., Langridge, D., and Green, M. (2019) A Cyclic Ion Mobility-Mass Spectrometry System, *Anal Chem 91*, 8564-8573.

29. Deng, L., Ibrahim, Y. M., Hamid, A. M., Garimella, S. V., Webb, I. K., Zheng, X., Prost, S. A., Sandoval, J. A., Norheim, R. V., Anderson, G. A., Tolmachev, A. V., Baker, E. S., and Smith, R. D. (2016) Ultra-High Resolution Ion Mobility Separations Utilizing Traveling Waves in a 13 m Serpentine Path Length Structures for Lossless Ion Manipulations Module, *Anal Chem 88*, 8957-8964.

30. Nagy, G., Attah, I. K., Garimella, S. V. B., Tang, K., Ibrahim, Y. M., Baker, E. S., and Smith, R. D. (2018) Unraveling the isomeric heterogeneity of glycans: ion mobility separations in structures for lossless ion manipulations, *Chem Commun (Camb) 54*, 11701-11704.

31. Wei, J., Tang, Y., Ridgeway, M. E., Park, M. A., Costello, C. E., and Lin, C. (2020) Accurate Identification of Isomeric Glycans by Trapped Ion Mobility Spectrometry-Electronic Excitation Dissociation Tandem Mass Spectrometry, *Anal Chem 92*, 13211-13220.

32. Shvartsburg, A. A., Li, F., Tang, K., and Smith, R. D. (2006) High-resolution field asymmetric waveform ion mobility spectrometry using new planar geometry analyzers, *Anal Chem 78*, 3706-3714.

33. Fang, P., Ji, Y., Silbern, I., Viner, R., Oellerich, T., Pan, K. T., and Urlaub, H. (2021) Evaluation and Optimization of High-Field Asymmetric Waveform Ion-Mobility Spectrometry for Multiplexed Quantitative Site-Specific N-Glycoproteomics, *Anal Chem 93*, 8846-8855.

34. Fontana, C., and Widmalm, G. (2023) Primary Structure of Glycans by NMR Spectroscopy, *Chem Rev 123*, 1040-1102.

35. H, V. (2012) Unravelling Glycobiology by NMR Spectroscopy, In *Glycosylation* (Stefana, P., Ed.), p Ch. 4, IntechOpen, Rijeka.

36. Speciale, I., Laugieri, M. E., Noel, E., Lin, S., Lowary, T. L., Molinaro, A., Duncan, G. A., Agarkova, I. V., Garozzo, D., Tonetti, M. G., Van Etten, J. L., and De Castro, C. (2020) Chlorovirus PBCV-1 protein A064R has three of the transferase activities necessary to synthesize its capsid protein N-linked glycans, *Proc Natl Acad Sci U S A 117*, 28735-28742.

37. Sassaki, G. L., Guerrini, M., Serrato, R. V., Santana Filho, A. P., Carlotto, J., Simas-Tosin, F., Cipriani, T. R., Iacomini, M., Torri, G., and Gorin, P. A. (2014) Monosaccharide composition of glycans based on Q-HSQC NMR, *Carbohydr Polym 104*, 34-41.

38. Mocsai, R., Kaehlig, H., Blaukopf, M., Stadlmann, J., Kosma, P., and Altmann, F. (2021) The Structural Difference of Isobaric N-Glycans of Two Microalgae Samples Reveals Taxonomic Distance, *Front Plant Sci 12*, 643249.

39. Han, L., and Costello, C. E. (2013) Mass spectrometry of glycans, *Biochemistry (Mosc) 78*, 710-720.

40. Ruhaak, L. R., Xu, G., Li, Q., Goonatilleke, E., and Lebrilla, C. B. (2018) Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses, *Chem Rev 118*, 7886-7930.

41. Karas, M., Bachmann, D., and Hillenkamp, F. (1985) Influence of the Wavelength in High-Irradiance Ultraviolet-Laser Desorption Mass-Spectrometry of Organic-Molecules, *Analytical Chemistry 57*, 2935-2939.

42. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., Yoshida, T., and Matsuo, T. (1988) Protein and polymer analyses up tom/z 100 000 by laser ionization time-of-flight mass spectrometry, *Rapid Communications in Mass Spectrometry 2*, 151-153.

43. Ho, C. S., Lam, C. W., Chan, M. H., Cheung, R. C., Law, L. K., Lit, L. C., Ng, K. F., Suen, M. W., and Tai, H. L. (2003) Electrospray ionisation mass spectrometry: principles and clinical applications, *Clin Biochem Rev 24*, 3-12.

44. Bandu, R., Ahn, H. S., Lee, J. W., Kim, Y. W., Choi, S. H., Kim, H. J., and Kim, K. P. (2015) Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometric (LC/ESI-MS/MS) Study for the Identification and Characterization of In Vivo Metabolites of

Cisplatin in Rat Kidney Cancer Tissues: Online Hydrogen/Deuterium (H/D) Exchange Study, *PLoS One 10*, e0134027.

45. Britz-McKibbin, P. (2011) Capillary electrophoresis-electrospray ionization-mass spectrometry (CE-ESI-MS)-based metabolomics, *Methods Mol Biol 708*, 229-246.

46. Kourtchev, I., Szeto, P., O'Connor, I., Popoola, O. A. M., Maenhaut, W., Wenger, J., and Kalberer, M. (2020) Comparison of Heated Electrospray Ionization and Nanoelectrospray Ionization Sources Coupled to Ultra-High-Resolution Mass Spectrometry for Analysis of Highly Complex Atmospheric Aerosol Samples, *Anal Chem 92*, 8396-8403.

47. Wilm, M., and Mann, M. (1996) Analytical properties of the nanoelectrospray ion source, *Anal Chem 68*, 1-8.

48. An, H. J., and Lebrilla, C. B. (2011) Structure elucidation of native N- and O-linked glycans by tandem mass spectrometry (tutorial), *Mass Spectrom Rev 30*, 560-578.

49. Bythell, B. J., Abutokaikah, M. T., Wagoner, A. R., Guan, S., and Rabus, J. M. (2017) Cationized Carbohydrate Gas-Phase Fragmentation Chemistry, *J Am Soc Mass Spectrom 28*, 688-703.

50. Mucha, E., Lettow, M., Marianski, M., Thomas, D. A., Struwe, W. B., Harvey, D. J., Meijer, G., Seeberger, P. H., von Helden, G., and Pagel, K. (2018) Fucose Migration in Intact Protonated Glycan Ions: A Universal Phenomenon in Mass Spectrometry, *Angew Chem Int Ed Engl 57*, 7440-7443.

51. Zhao, C., Xie, B., Chan, S. Y., Costello, C. E., and O'Connor, P. B. (2008) Collisionally activated dissociation and electron capture dissociation provide complementary structural information for branched permethylated oligosaccharides, *J Am Soc Mass Spectrom 19*, 138-150.

52. Devakumar, A., Mechref, Y., Kang, P., Novotny, M. V., and Reilly, J. P. (2008) Identification of isomeric N-glycan structures by mass spectrometry with 157 nm laser-induced photofragmentation, *J Am Soc Mass Spectrom 19*, 1027-1040.

53. Zhou, S., Veillon, L., Dong, X., Huang, Y., and Mechref, Y. (2017) Direct comparison of derivatization strategies for LC-MS/MS analysis of N-glycans, *Analyst 142*, 4446-4455.

54. Maitre, P., Scuderi, D., Corinti, D., Chiavarino, B., Crestoni, M. E., and Fornarini, S. (2020) Applications of Infrared Multiple Photon Dissociation (IRMPD) to the Detection of Posttranslational Modifications, *Chem Rev 120*, 3261-3295.

55. Li, B., Russell, S. C., Zhang, J., Hedrick, J. L., and Lebrilla, C. B. (2011) Structure determination by MALDI-IRMPD mass spectrometry and exoglycosidase digestions of O-linked oligosaccharides from Xenopus borealis egg jelly, *Glycobiology 21*, 877-894.
56. Pikulski, M., Hargrove, A., Shabbir, S. H., Anslyn, E. V., and Brodbelt, J. S. (2007) Sequencing and characterization of oligosaccharides using infrared multiphoton dissociation and boronic acid derivatization in a quadrupole ion trap, *J Am Soc Mass Spectrom 18*, 2094-2106.
57. Schindler, B., Barnes, L., Gray, C. J., Chambert, S., Flitsch, S. L., Oomens, J., Daniel, R., Allouche, A. R., and Compagnon, I. (2017) IRMPD Spectroscopy Sheds New (Infrared) Light on the Sulfate Pattern of Carbohydrates, *J Phys Chem A 121*, 2114-2120.
58. Renois-Predelus, G., Schindler, B., and Compagnon, I. (2018) Analysis of Sulfate Patterns in Glycosaminoglycan Oligosaccharides by MS(n) Coupled to Infrared Ion Spectroscopy: the Case of GalNAc4S and GalNAc6S, *J Am Soc Mass Spectrom 29*, 1242-1249.
59. Lupo, D. W., and Quack, M. (1987) Ir-Laser Photochemistry, *Chemical Reviews 87*, 181-216.
60. Grant, E. R., Schulz, P. A., Sudbo, A. S., Shen, Y. R., and Lee, Y. T. (1978) Is Multiphoton Dissociation of Molecules a Statistical Thermal Process?, *Physical Review Letters 40*, 115-118.
61. Polfer, N. C. (2011) Infrared multiple photon dissociation spectroscopy of trapped ions, *Chem Soc Rev 40*, 2211-2221.
62. Schindler, B., Barnes, L., Renois, G., Gray, C., Chambert, S., Fort, S., Flitsch, S., Loison, C., Allouche, A. R., and Compagnon, I. (2017) Anomeric memory of the glycosidic bond upon fragmentation and its consequences for carbohydrate sequencing, *Nat Commun 8*, 973.
63. Filsinger, F., Ahn, D. S., Meijer, G., and von Helden, G. (2012) Photoexcitation of mass/charge selected hemin+, caught in helium nanodroplets, *Phys Chem Chem Phys 14*, 13370-13377.
64. Smolarek, S., Brauer, N. B., Buma, W. J., and Drabbels, M. (2010) IR spectroscopy of molecular ions by nonthermal ion ejection from helium nanodroplets, *J Am Chem Soc 132*, 14086-14091.

65. Mucha, E., Gonzalez Florez, A. I., Marianski, M., Thomas, D. A., Hoffmann, W., Struwe, W. B., Hahm, H. S., Gewinner, S., Schollkopf, W., Seeberger, P. H., von Helden, G., and Pagel, K. (2017) Glycan Fingerprinting via Cold-Ion Infrared Spectroscopy, *Angew Chem Int Ed Engl 56*, 11248-11251.

66. Lettow, M., Grabarics, M., Greis, K., Mucha, E., Thomas, D. A., Chopra, P., Boons, G. J., Karlsson, R., Turnbull, J. E., Meijer, G., Miller, R. L., von Helden, G., and Pagel, K. (2020) Cryogenic Infrared Spectroscopy Reveals Structural Modularity in the Vibrational Fingerprints of Heparan Sulfate Diastereomers, *Anal Chem 92*, 10228-10232.

67. Fenn, L. S., and McLean, J. A. (2013) Structural separations by ion mobility-MS for glycomics and glycoproteomics, *Methods Mol Biol 951*, 171-194.

68. Hofmann, J., and Pagel, K. (2017) Glycan Analysis by Ion Mobility-Mass Spectrometry, *Angew Chem Int Ed Engl 56*, 8342-8349.

69. Chen, Z., Glover, M. S., and Li, L. (2018) Recent advances in ion mobility-mass spectrometry for improved structural characterization of glycans and glycoconjugates, *Curr Opin Chem Biol 42*, 1-8.

70. Schindler, B., Laloy-Borgna, G., Barnes, L., Allouche, A. R., Bouju, E., Dugas, V., Demesmay, C., and Compagnon, I. (2018) Online Separation and Identification of Isomers Using Infrared Multiple Photon Dissociation Ion Spectroscopy Coupled to Liquid Chromatography: Application to the Analysis of Disaccharides Regio-Isomers and Monosaccharide Anomers, *Anal Chem 90*, 11741-11745.

71. Martens, J., Koppen, V., Berden, G., Cuyckens, F., and Oomens, J. (2017) Combined Liquid Chromatography-Infrared Ion Spectroscopy for Identification of Regioisomeric Drug Metabolites, *Anal Chem 89*, 4359-4362.

72. Warnke, S., Ben Faleh, A., Pellegrinelli, R. P., Yalovenko, N., and Rizzo, T. R. (2019) Combining ultra-high resolution ion mobility spectrometry with cryogenic IR spectroscopy for the study of biomolecular ions, *Faraday Discuss 217*, 114-125.

73. Ben Faleh, A., Warnke, S., and Rizzo, T. R. (2019) Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures, *Anal Chem 91*, 4876-4882.

74. Ben Faleh, A., Warnke, S., Van Wieringen, T., Abikhodr, A. H., and Rizzo, T. R. (2023) New Approach for the Identification of Isobaric and Isomeric Metabolites, *Anal Chem*.

Chapter 3 Experimental approach

In this chapter, we present an overview of the different experimental techniques used in this work. We first focus on some the major technologies that we employ: liquid chromatography, high-resolution ion mobility, and cryogenic IR spectroscopy. We then describe the instruments into which we incorporate these techniques and use for the experiments described in this thesis.

3.1 Liquid chromatography

All our LC work is performed on an AQUITY UPLC H-Class Plus system (Waters).

Figure 3.1. AQUITY UPLC H-Class Plus System

We employ porous graphitized carbon (PGC) as the stationary phase, as it is known for its high efficiency in separating polar compounds, especially oligosaccharides.[1] PGC is a conductive crystalline material made of intertwined graphite ribbons. The graphitized sheets are made of sp^2 hybridized carbon atoms arranged in a hexagonal pattern, bonded together by covalent bonds. These sheets are held together by van der Waals forces, providing PGC with mechanical stability.[2] Figure 3.2(a,b) shows high-resolution electron micrographs of the amorphous carbon formed when carbon is heated at 1000°C and the graphitization that occurs after heating at 2340°C to obtain PGC. Unlike three-dimensional graphite, the layers in PGC are not aligned in a regular manner, making it a "two-dimensional" form of graphite as shown in Figure 3.2 (c).[3]

PGC is a highly effective stationary phase for analyzing and separating complex mixtures due to its unique retention and selectivity properties. The retention mechanisms in PGC chromatography are multifaceted and involve a combination of different interactions which are dispersive, electronic, and steric.[4] Dispersive interactions are similar to those observed in reversed phase chromatography and are heavily influenced by the mobile phase strength and

the properties of the analyte. This interaction mode includes solvent-solute interactions that can be dipole-dipole, hydrogen bonding, hydrophobic-hydrophilic repulsion interactions, and solvent and solute dispersive forces with graphite.[4-6]

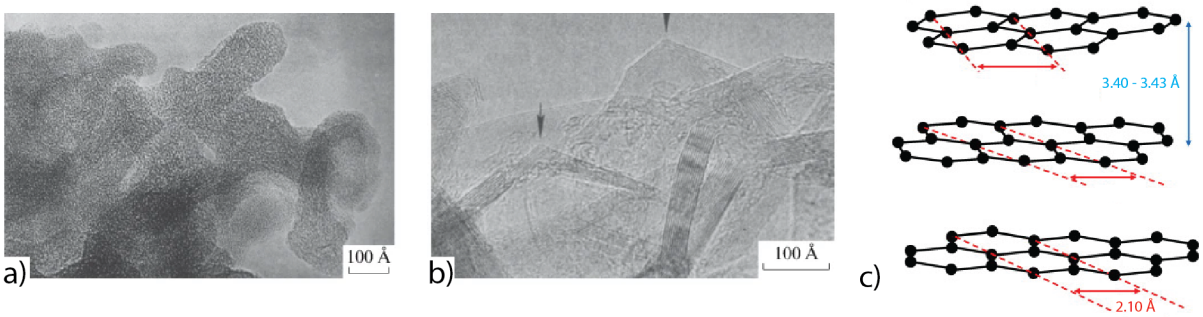


Figure 3.2. (a) Amorphous carbon obtained at 1000°C, (b) PGC obtained after graphitization at 2340°C (c) Warren structure of two-dimensional turbostratic graphite with no layer registration (PGC).[3]

Electronic interactions between polarizable groups of the analyte and the surface of graphite, known as the polar retention effect on graphite (PREG), are not as well understood but are thought to stem from some form of charge induction or electron lone pair donating/accepting interactions with the π-cloud of graphite.[7-10] This effect can be significant if there is π-π overlap due to the unsaturation or aromaticity of the analyte. The stereochemistry of the analyte and the location of the polar functional groups with respect to the graphitic plane also play an important role in electronic interactions.

Finally, steric interactions are also essential and involve the size and shape of the analyte.[11] Thus, increased surface area contact of the analyte with the surface of graphite results in longer retention times. This trait becomes especially noticeable with aromatic compounds, where π-π overlap is significant. By considering these different modes of PGC interactions, it is possible to optimize the separation and analysis of complex mixtures.

For this work, the advantage of using PGC is its high efficiency at separating isomeric glycans, which can be challenging to achieve with other stationary phases. Additionally, PGC can be used with various mobile phase systems, including polar organic solvents and ammonium acetate buffers, making it a versatile stationary phase that is compatible with electrospray sources on mass spectrometers.

However, the use of PGC for glycan analysis also has some limitations. One is that PGC is prone to oxidation, which can lead to increased baseline noise and reduced column lifetime. Another is that optimizing the mobile phase for specific glycan structures can be difficult, as

different glycans can have other retention mechanisms on the PGC stationary phase. Furthermore, the porous structure of PGC can lead to sample adsorption, resulting in lower recovery of glycans.

In Chapter 7 we describe in more detail the coupling of our UPLC system to our home built instrument.

3.2 Ion mobility using structures for lossless ion manipulations (SLIM)

Among the various types of ion mobility spectrometry mentioned in the previous chapter, the one that has demonstrated the highest resolving power is called structures for lossless ion manipulations (SLIM) developed by Smith and coworkers.[12, 13] SLIM is a traveling-wave ion mobility technique that uses two mirrored printed circuit boards (PCBs) to create a path for ion separation. As shown in Figure 3.3(a), a SLIM module has three types of electrodes: DC guard electrodes, radio frequency (RF) electrodes, and traveling-wave (TW) electrodes. Both PCBs of the mirrored pair consists of 6 RF and 5 TW tracks, all surrounded by 2 DC guard electrodes (Figure 3.3(a)). Guard electrodes confine ions laterally (x-axis), while RF electrodes confine them between the planar surfaces (y-axis). Traveling-wave electrodes apply DC potentials in a sequence to produce a square wave, driving ion motion and enabling mobility separation of molecules along the z-axis. The travelling-wave is applied on a series of 8 repeating electrodes across the entire SLIM length, 4 of which are set to high potential while the other four are set to low potential, forming a repeating square wave. For the wave to propagate, the first high potential electrode is set to low potential, the first low potential electrode is set to high, as shown in Figure 3.3(b). This happens consecutively, producing a traveling square wave potential that pushes the ions through the separation path.

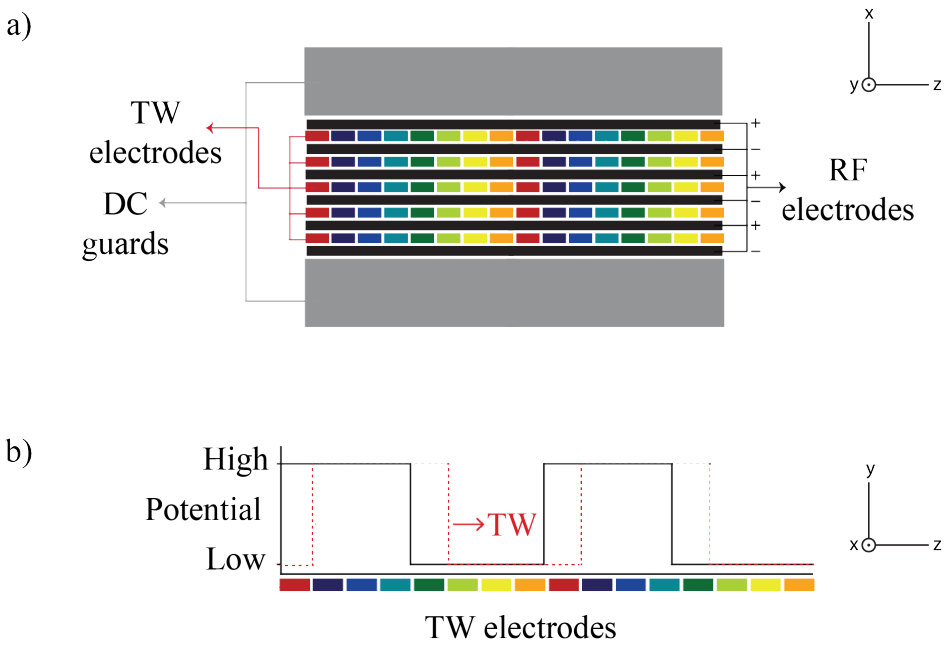


Figure 3.3. (a) Schematic of the basic unit of SLIM; (b) how the traveling wave is made to propagate.

Smith and coworkers demonstrated the potential of SLIM tracks to maneuver ions around 90-degree turns,[12-14] as shown in Figure 3.4(a), enabling the creation of serpentine paths that extend the total separation length within a moderately sized device. For instance, a 13-meter separation path can be achieved on a 45.9 cm x 32.5 cm PCB. This results in a significant increase in IMS resolving power (R), which scales with $L^{1/2}$ in TWIMS, where L is the total path length. Additionally, they introduced T-shaped switches (Figure 3.4(b)) capable of efficiently rerouting ions from a linear course to a perpendicular one. By switching a set of electrodes at the exit of the SLIM separation region, ions can either be directed 90 degrees towards the next instrument section or pushed towards a return path, allowing ions to traverse multiple cycles, further enhancing the path length and hence the resolving power. A resolving power exceeding 1800 can be achieved with path lengths surpassing a kilometer through this versatile design.[15] The TW-SLIM module's notable features include ion trapping, storage, and ejection capabilities, allowing ion packets to be strategically positioned at designated locations along the separation path and discharged as needed.

a) SLIM 90° turn b) SLIM T-switch

Figure 3.4. (a) A SLIM 90-degree turn employs a particular arrangement of DC pads to redirect ions perpendicular to their original motion direction. (b) A SLIM switch is achieved by manipulating specific rows of DC pads (depicted in pink) to guide the ions' trajectory.

We have utilized this characteristic to design a novel SLIM device equipped with collision-induced dissociation (CID) capabilities. Each SLIM board we have developed contains trapping regions where ions can be captured and subsequently fragmented as required. This is accomplished by integrating the electrode configuration depicted in Figure 3.5 between the serpentine paths on the SLIM device.

Trapping is achieved by simultaneously increasing the bias on the serpentine path blocking electrodes and decreasing it on the trap entrance blocking electrodes, while applying the board's traveling wave to the latter electrodes. This directs the ions into the trap where they are maintained by the traveling wave and the elevated bias on the trap exit blocking electrodes. To subsequently release these ions, the bias on the trap exit blocking electrodes is reduced while simultaneously subjecting them to the board's traveling wave.

On our prototype instrument, to execute CID on the trapped ions, the bias of all the electrodes in the trap assembly (DC guards, TW, and RF) is increased so that the ions are propelled towards the serpentine path. This is followed by a reduction of the trap exit blocking electrodes to release the ions.

As a result, the ions undergo sufficient collisions with the buffer gas inducing their fragmentation as they exit the SLIM trap. These fragments can then be cycled again on the SLIM board for further mobility separation of potential isomeric fragments, which can then be trapped again for multiple fragmentation rounds (IMS^n).

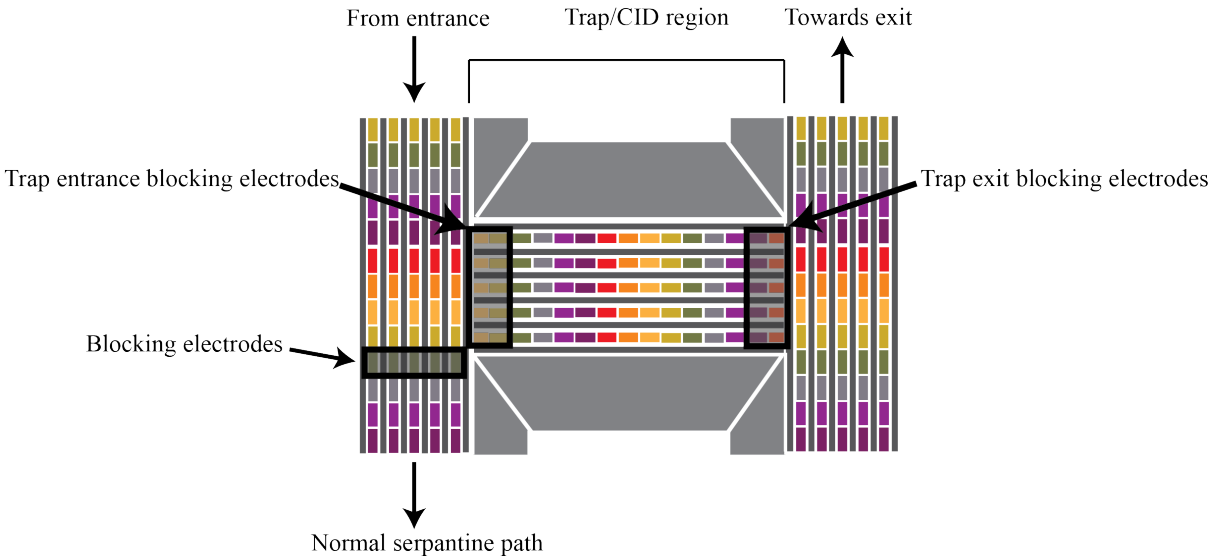


Figure 3.5. Ion trap assembly on the SLIM device.

In later iterations of our SLIM devices, two metallic grids – separated by 0.8mm – were installed at the entrance of the trapping region. Here a uniform electric field can be applied to the trapped ions causing them to fragment inside the trap.[16]

3.3 Cryogenic messenger-tagging IR spectroscopy

Due to the low density of ions in the gas phase, measuring the amount of light they absorb becomes impractical. For this reason, action spectroscopy is usually performed when working with gas-phase ions. Instead of directly measuring the light attenuation, some change in the molecule induced by photon absorption is monitored.

In this work, we employ cryogenic infrared (IR) messenger-tagging spectroscopy for the analysis of gas-phase ions. The first step in this method is trapping the ions within a thermally conductive container cooled to the desired temperature. Then a buffer gas, which in our case is composed of 90% helium and 10% nitrogen, is pulsed into the trap, where it can readily achieve thermal equilibrium with the surface of the cold enclosing box. We cool the trap to around 45 Kelvin using a commercial closed-cycle, dual-stage helium refrigerator.[17] Upon collision with the cold gas, the ions are cooled and form weakly bound clusters with N_2.

While different molecules or atoms can be used as a tag, our choice of N_2 is based on its favorable mass shift (+28 Da) and the fact that its clusters with ions are sufficiently strongly bound that cooling is only necessary to about 40 K – tagging by either helium or hydrogen requires much lower temperatures. Generally, when choosing a tag, it should be IR inactive so that it does not interfere with the IR absorption of the molecule itself, and this is achievable

with both atomic tags, such as helium and argon, as well as homonuclear diatomic molecules, since they have no allowed infrared absorptions.

The cold clusters are then irradiated by an IR laser, and when the laser frequency is in resonance with the vibration of a functional group in the molecule, a photon is absorbed. The absorption of IR radiation is followed by a redistribution of vibrational energy, ultimately leading to the detachment of N_2 (see Figure 3.6).[18-21]

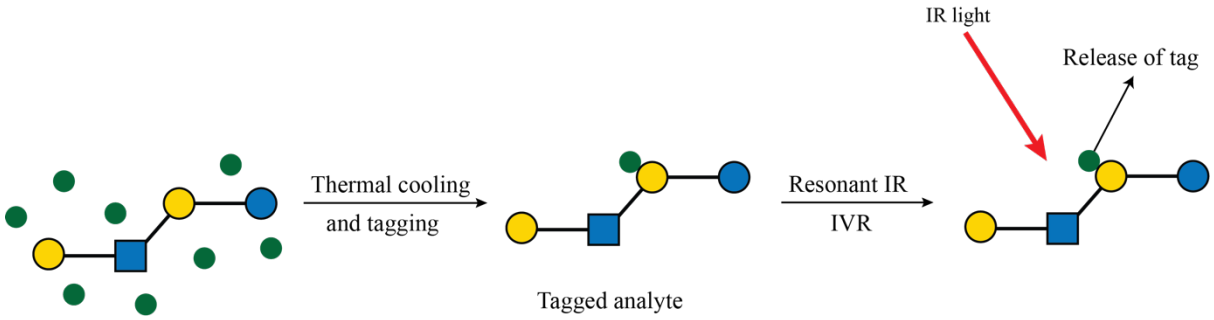


Figure 3.6. Schematic overview of messenger-tagging spectroscopy.

To then obtain the IR spectrum from the recorded mass spectra (Figure 3.7), the fraction of tagged ions is plotted as a function of the laser wavenumber:

$$IR\ spectrum\ (cm^{-1}) = \frac{\sum tagged\ clusters}{\sum Precursor+tagged\ clusters} \qquad \text{equation 3.1}$$

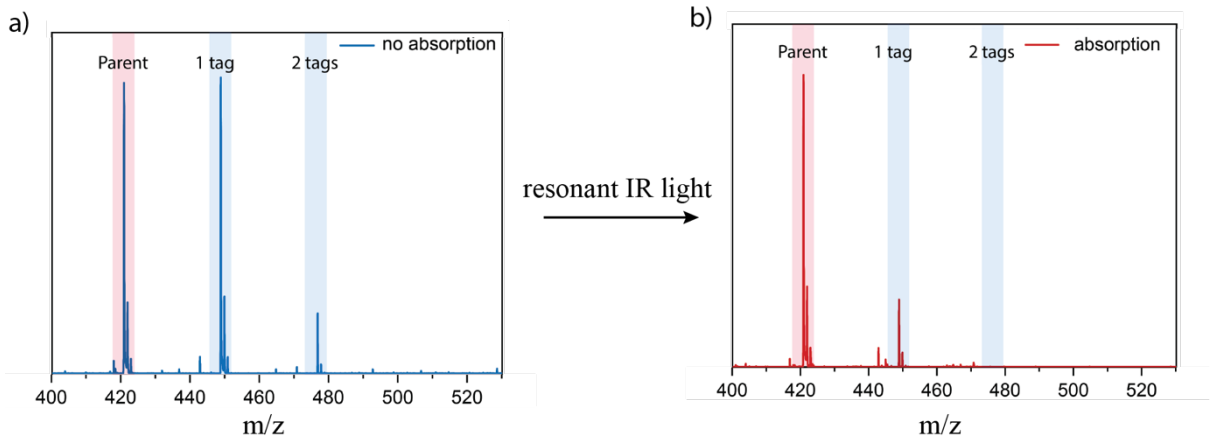


Figure 3.7. (a) a mass spectrum of a tagged molecule, (b) the same spectrum after absorption and depletion occurs.

3.4 Experimental setup: Enhancing sensitivity and speed

In various scientific domains, such as medical diagnostics and pharmaceutical quality control, a substantial volume of samples is usually analyzed. Consequently, there is a demand for

analytical tools that can provide both rapid and accurate results. Therefore, we explore increasing our instrument sensitivity so that lower concentration species can be analyzed in addition to decreasing our IR spectral acquisition time so that it is more compatible with the timescale of other analytical techniques such as liquid chromatography.

3.4.1 CW laser installation and calibration using a photoacoustic cell

In the context of this research, we have employed a continuous-wave (CW) infrared (IR) laser system from IPG Photonics as an alternative to the pulsed optical parametric oscillators (OPOs) typically used in this wavenumber region. This particular CW laser system uses a chromium-doped ZnS/Se crystal pumped by a thulium fiber laser, spanning the spectral range from 3230 cm^{-1} to 3980 cm^{-1}. One advantage of this CW laser system is its continuous operation, which allows us to run the experiment at a repetition rate determined by the time it takes for mobility separation of the ions. Furthermore, this laser system can emit consistent power output across all wavelengths by adjusting the power of the fiber pump laser during a scan, enabling the generation of highly reproducible IR depletion spectra.

Accurate laser wavelength calibration is essential for reproducibly obtaining spectra and building a reliable database. One method to do this uses the photoacoustic effect to measure the accurately know transitions of water vapor.

The photoacoustic effect is a phenomenon in which the absorption of laser light by a sample, water vapor in this case, generates acoustic waves subsequent to collisional relaxation of the vibrationally excited molecules.[22] By using a sensitive hearing-aid microphone to measure this acoustic signal as a function of laser stepper motor position, and comparing the transitions to the accurately known absorptions of water, one can determine the laser wavelength with a high degree of precision.[23]

The photoacoustic cell, shown schematically in Figure 3.8, is designed as a closed container with a transparent window through which the laser beam enters. A small drop of water is placed inside the cell, which is otherwise at atmospheric pressure. When the laser beam interacts with the water vapor, it generates an acoustic wave, which is detected by a sensitive microphone positioned in the center of the cell. To enhance the sensitivity of the measurement, an optical chopper is placed between the laser and the cell, modulating the laser at the resonant frequency of the cylindrical cell, which in our case was determined experimentally to be around 330 Hz. This modulation results in a significant increase in signal-to-noise ratio and sensitivity.

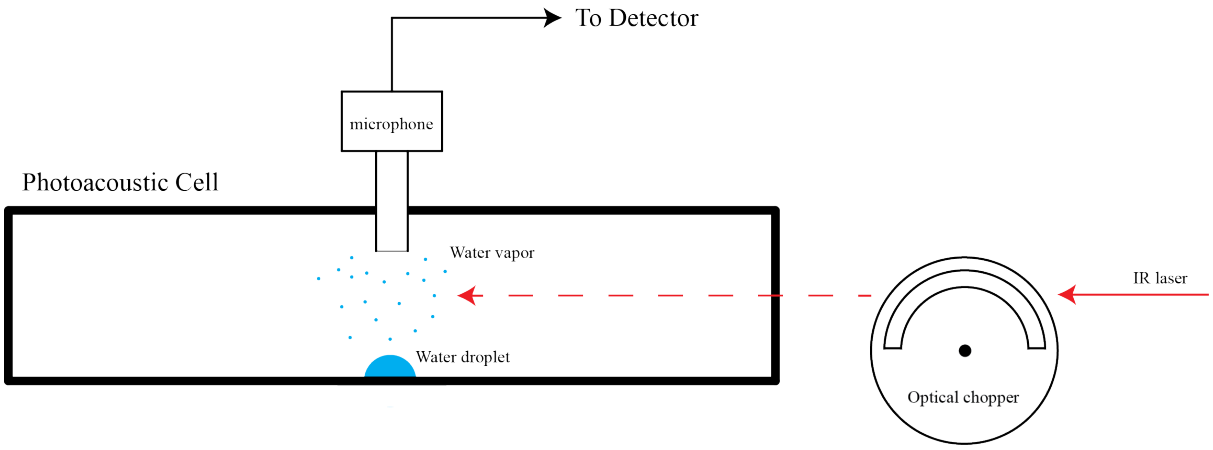
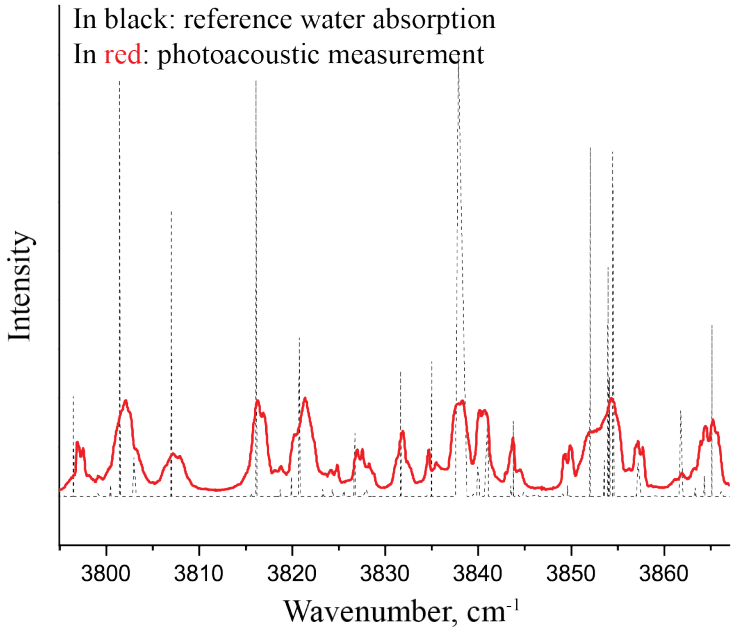


Figure 3.8. Schematic of the photoacoustic cell used in this work.

To calibrate the laser wavelength, we record the acoustic signal as a function of the position of the laser stepper motor , and the acquired spectrum is compared and calibrated by reference to the water absorption spectrum from the Hitran database,[24] as shown in Figure 3.9 (a). Multiple sample points are selected across the entire scan region and are then fitted using a modified calibration equation provided by the laser manufacturer.

This then correlates the stepper motor position to the wavelength being emitted by the laser as shown in Figure 3.9 (b).

a)



b)

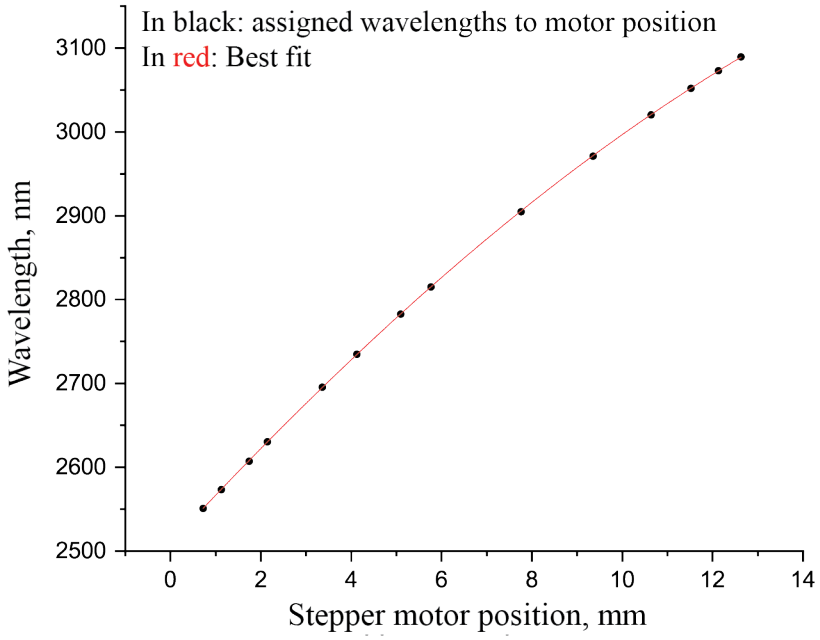


Figure 3.9. (a) Figure showing overlap of measured photoacoustic frequency of water vapor fitted to a reference spectrum, (b) Curve fit of 14 datapoints sampled from 3.9 (a).

3.4.2 SLIM-IR Prototype

The prototype instrument, shown schematically in Figure 3.8, features a SLIM-based traveling-wave ion mobility device, a quadrupole, a planar trap, and mass analysis by time-of-flight (TOF) mass analyzer.[25] Ions generated by nanoelectrospray ionization (nESI) are introduced into the instrument through a flared stainless-steel capillary, which is maintained at a temperature of 150°C. Upon exiting the capillary, the ions enter a dual ion funnel trap (MassTech, USA), where they are temporarily stored for tens of milliseconds. These stored ions are subsequently released in pulses of 100-150 μs duration thorough a ring-electrode guide into the SLIM-IMS region. Another ring electrode guide is positioned at the SLIM exit to ensure efficient unloading of the mobility-separated species. After this, the ions are guided through a series of ring-electrode and hexapole guides into two differentially-pumped stages (I and II), which are maintained at pressures of 10^{-2} and 10^{-5} mbar, respectively. A lens assembly, composed of four steering plates, installed after the hexapole guide, deflects unwanted ions and allows only the ions of interest within a chosen drift-time range to advance towards the quadrupole bender. Here, the ions are turned 90° by a DC quadrupole bender towards an octupole guide and a quadrupole mass filter. The ion arrival time distributions (ATDs) can be measured on a channeltron detector, with or without m/z mass selection in the quadrupole mass filter.

After the quadrupole mass filter, the ions enter a planar cryogenic trap composed of two PCBs featuring four DC electrodes that surround a set of radio-frequency (RF) electrodes.[26] This trap assembly is affixed to a copper housing, which is attached to a helium cold head maintained at 40 K. Prior to the ions' entrance into the trap, a cold buffer gas composed of a 90:10 He/N2 mixture is introduced using a pulsed valve (Parker, Series 9). The ions are confined and cooled as they collide with the cold buffer gas, and become tagged with N_2 during the cooling process, which act as messenger tags for IR absorption detection. To obtain an IR spectrum, the N_2-tagged ions are irradiated for 50 ms with a continuous-wave (CW) mid-IR laser (IPG Photonics) operating at 0.2 W output power. The ions are subsequently extracted through the slots in the top PCB of the cryogenic trap into the time-of-flight (ToF) mass spectrometer.[27] All experiments described in chapter 4 and some of those in chapter 5 were performed using this instrument.

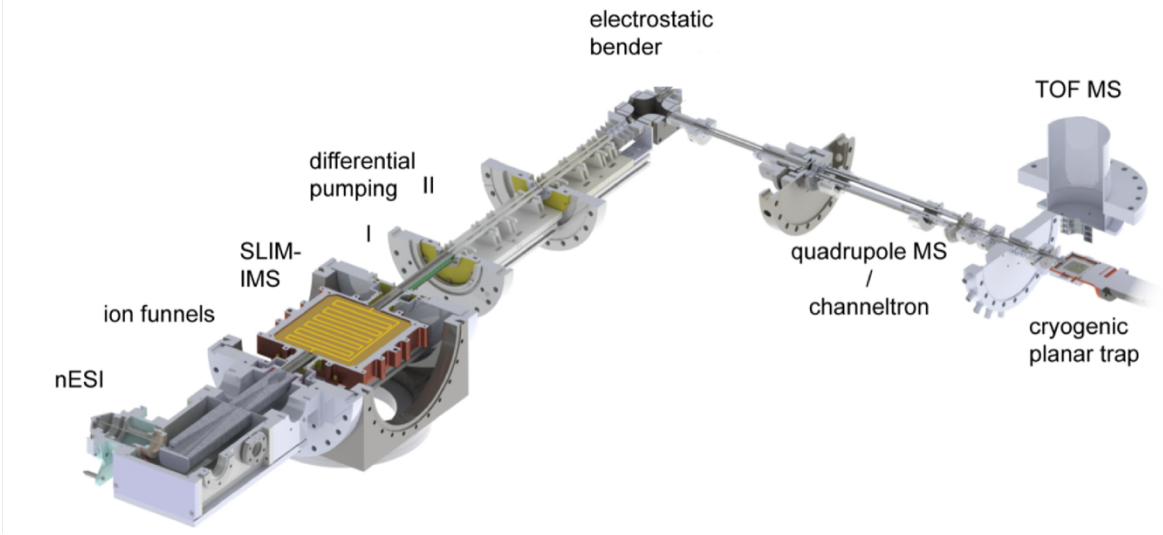


Figure 3.8. A schematic overview of our prototype instrument combining SLIM-IMS with messenger-tagging spectroscopy[25]

3.4.3 SLIM-IMS

Figure 3.9 shows details of the SLIM device used in this prototype.[28] The serpentine path on the SLIM board in this instrument is approximately 1.5 m in length. After reaching the exit of the SLIM, ions can be either sent to the cryogenic trap for detection or cycled back around the serpentine path for additional separation. In addition, this board has small ion trap segments (17 mm in length) used for confining mobility separated ions, which can be stored for a prolonged amount of time with minimal loss in signal intensity and used for CID as described in Section 3.2. The resolving power of this board is approximately 60 after one pass and reaches approximately 300 after 8 cycles (~13 m path length).

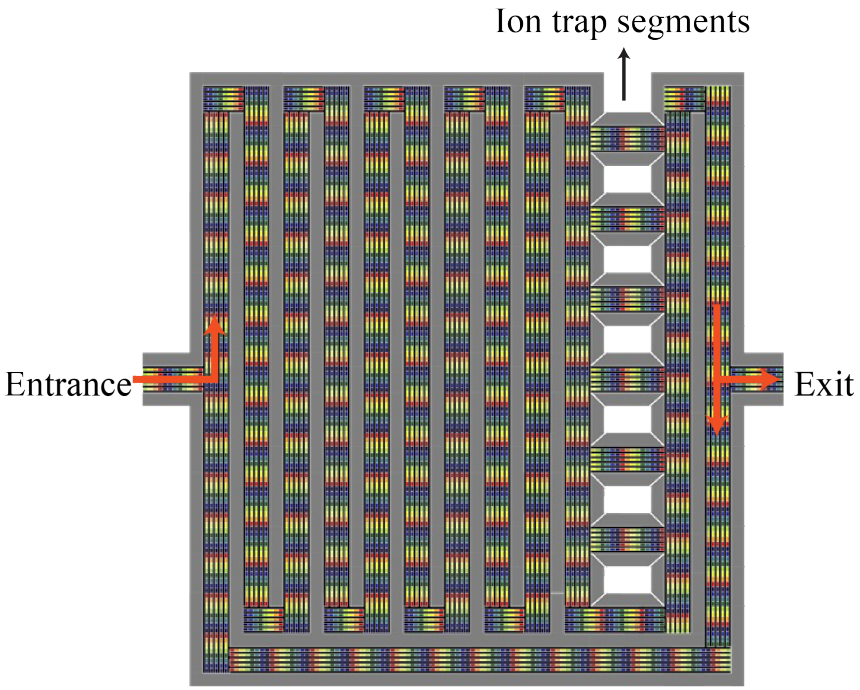


Figure 3.9. Schematic of the SLIM-PCB used in the prototype.

3.4.4 Improving sensitivity

In normal operation, an ion funnel trap (IFT) is used to introduce pulses of ions into the instrument. However, when the sprayed solution contains multiple ionic species, the IFT can quickly overfill, resulting in discrimination against higher mass ions and a decreased signal for potential ions of interest, negatively impacting sensitivity. By using the on-board traps mentioned in section 3.2, ion packets from multiple machine cycles can be accumulated in onboard traps, enhancing ion signal intensity. Additionally, ions can be mobility separated before trapping, ensuring only the signal of the target ions is enriched.

Enhancing the ion signal is performed by varying the repetition rates of the IFT and SLIM trap exit. For instance, by setting the IFT at a 20 Hz repetition rate, ions are released every 50 ms, and the ions of interest are trapped and stored. These ions are retained until the next ion packet arrives from the IFT, which is then added to the trap. We then release the ions from the SLIM trap at a 10 Hz repetition rate. The SLIM traps can be run at even lower repetition rates for further enrichment while maintaining the IFT at 20 Hz. Figure 3.10 demonstrates the signal increase for a 100 µM disaccharide solution using the SLIM ion traps. The increase in ion signal is not linear, which is likely due to overfilling the trap and subsequent ion loss. While in principle, these traps have the same problem as the IFT itself (overfilling due to space charge effects), in this case, any overfilling is due to mobility selected ions of interest and no other ions in the electrospray plume. The capacity of these on-board traps can in principle be increased by making them longer.

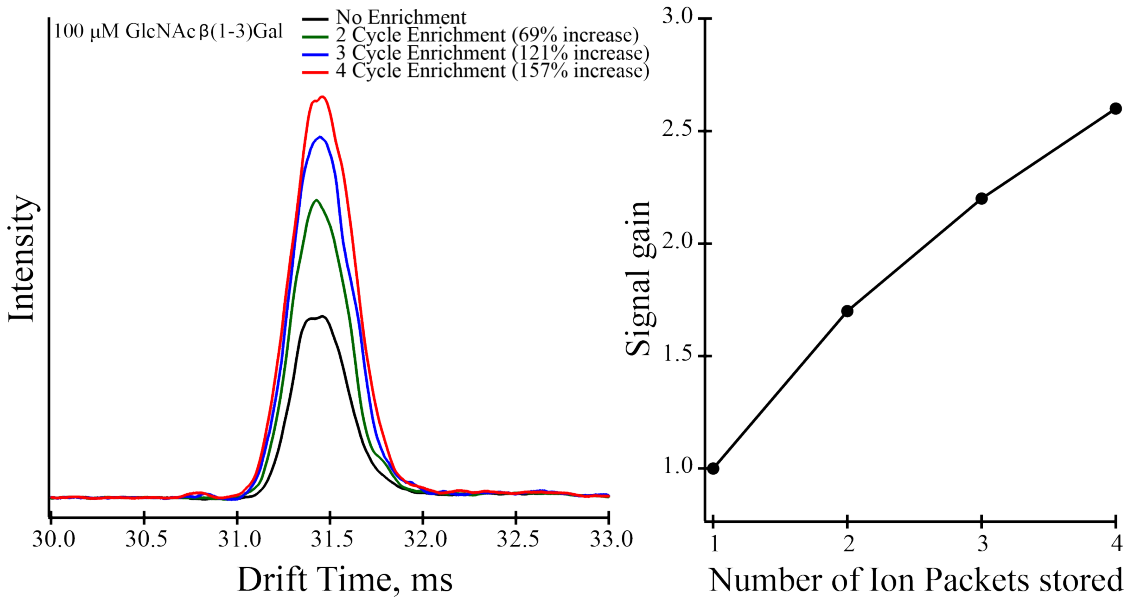


Figure 3.10. showing the increase of signal as a function of the number of enrichment cycles applied

To showcase the effectiveness of on-board enrichment of mobility selected ions, a 1 µM solution of the disaccharide GlcNAc β (1-3) Gal was used to measure its IR spectrum as a function of the number of enrichment cycles. As seen in Figure 3.11, without enrichment, the IR spectrum (depicted in black) is noisy with peaks that are difficult to discern. However, applying the enrichment technique considerably improves the signal-to-noise ratio and, consequently, the sensitivity.

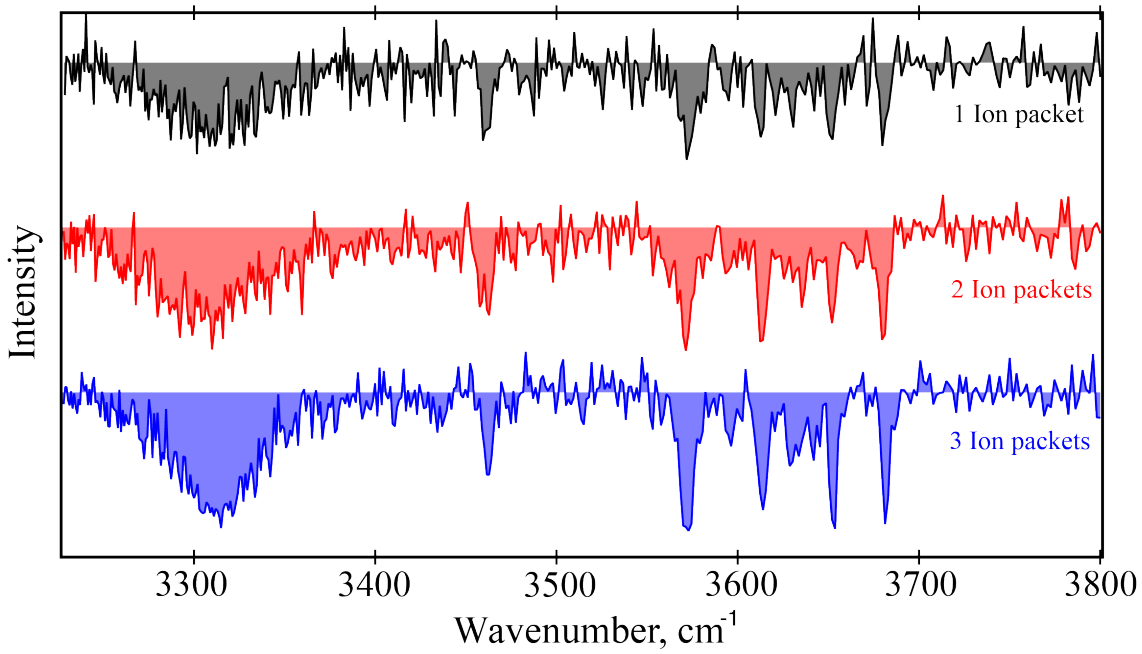


Figure 3.11. showing the IR spectrum of a 1 µM solution of GlcNAc β (1-3) Gal obtained before and after applying the ion packet enrichment procedure.

3.4.5 Improving acquisition speed

With the OPO system, acquiring a single spectrum required 10-15 minutes, and the shot-to-shot pulse energy fluctuations required multiple scans to achieve a decent signal-to-noise ratio. In contrast, the CW laser is not limited by a repetition rate, and we can repeat ion injects as fast as possible given the drift time during mobility separation. Unless stated otherwise stated, all spectra have been acquired at 0.1 W of power and an irradiation time of 50 ms. This laser also has the advantage that it can maintain a constant power output across all wavelengths by adjusting the power of the fiber pump laser during a scan to compensate for the laser gain curve. This is done by a one-time calibration of the pump power required to produce a 0.1 W output as a function of laser wavenumber.

Figure 3.12 shows the ability of this CW laser to reduce the spectral acquisition time. The 86-second scan taken shows a good signal-to-noise ratio (SNR) and was performed using a running average (3 averages per point). The 57-second and 43-second scans appear to have similar resolution, suggesting that the data may have been oversampled in the former. The resolution starts to decrease in the 34-second scan. Because the laser is scanned a fixed speed without stopping, scanning it too quickly while implementing a running average leads to a loss of resolution due to averaging spectra recorded at slightly different wavenumbers.

To address this issue, we programmed the laser to stop at constant wavenumber increments instead of scanning at a constant speed. This approach allows data collection at each wavenumber while the laser remains stationary, which optimally uses the inherent resolution of the laser, which is approximately 1 cm^{-1}. Furthermore, stepping in constant wavenumber increments ensures consistent point sampling across the entire spectral region.

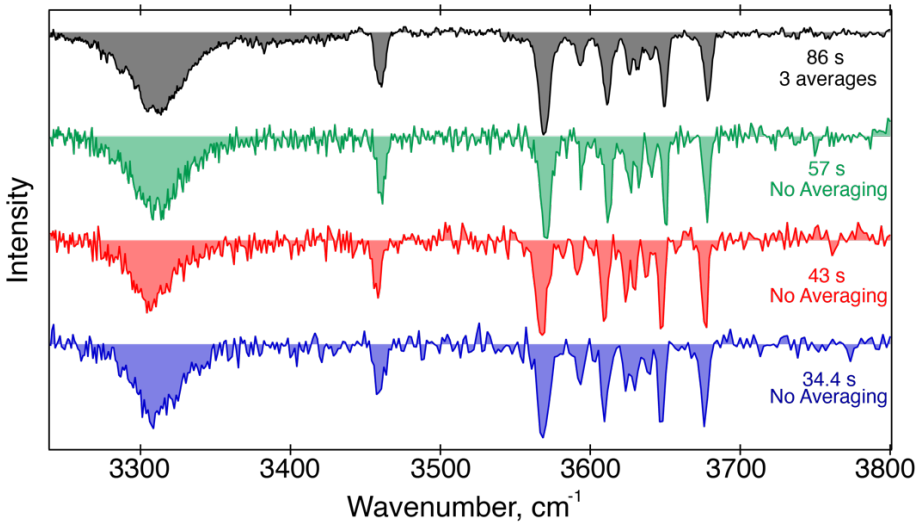


Figure 3.12. showing the Spectrum of a 100-µM solution of GlcNAcβ(1-3)Gal taken at different laser scanning speeds.

3.5 *Overview of the second-generation instrument*

The next-generation instrument, depicted in Figure 3.13, was designed to enhance the IMS resolving power by incorporating a longer SLIM-IMS region,[16] which enables a 10 m separation path length compared to the 1.5 m path length of the prototype machine. In addition, this instrument features a cryogenic multi-trap, facilitating the simultaneous acquisition of infrared spectra from multiple mobility selected species. Chapters 6 and 7 of this thesis were conducted using this instrument.

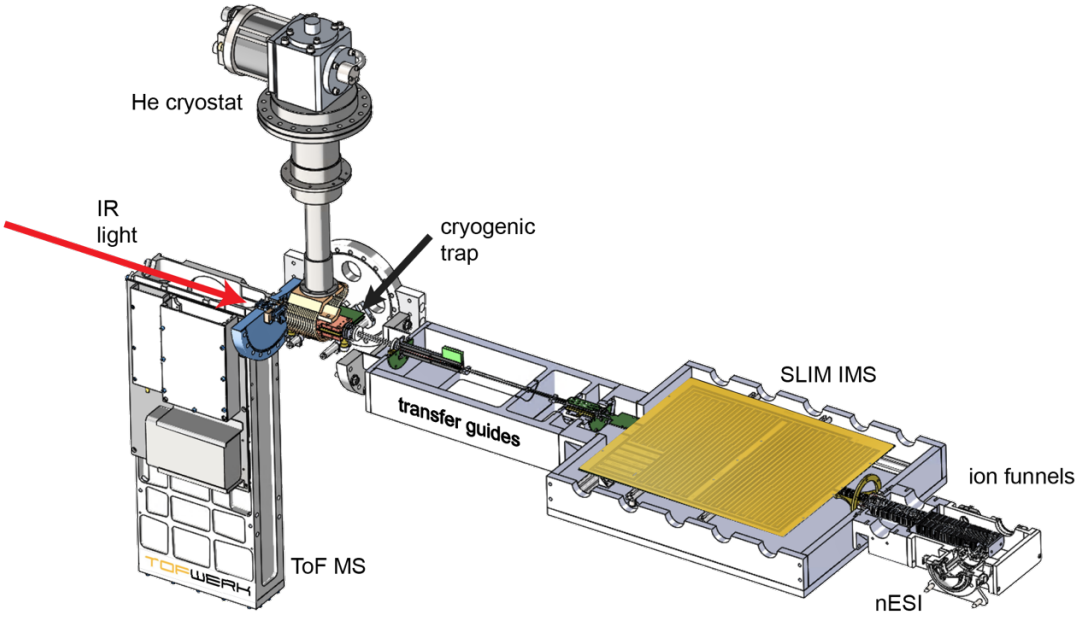


Figure 3.13. showing a schematic overview of our prototype instrument combining SLIM-IMS with messenger-tagging spectroscopy

Similar to the prototype, ions produced by nESI are sent through a dual-stage ion funnel source assembly (MassTech, USA) and a ring electrode guide, leading to the SLIM module, which is maintained at a pressure of 2.2 mbar. Upon entry onto the SLIM module, ions are accumulated in a 2 m storage section and subsequently released into the separation section as pulses ranging from microseconds to a few milliseconds in duration. A second ring-electrode guide, positioned at the exit of the SLIM device, transfers ions through a conductance limit and a series of multipole guides, transporting them through differential pumping stages (10^{-2}, 10^{-4}, and 10^{-6} mbar) into a high vacuum stage. The first pumping stage (10^{-2} mbar) incorporates another ring-electrode guide, followed by quadrupole and octupole guides in the second stage (10^{-4} mbar). The incoming ions are then guided by two multipole guides into the third stage, which is maintained at 10^{-6} mbar. Subsequently, an octupole guide and an einzel lens assembly steer ions into a planar cryogenic multi-trap. The planar trap is affixed to two copper plates to ensure homogeneous cooling. These plates are connected to a helium (He) cryostat (RDK 408

Sumitomo, Japan), and maintained at 45 K. Prior to the arrival of incoming ions, an 80:20 He:N_2 mixture is pulsed into the trap to aid in their confinement and cooling. During this phase, the ions form weakly-bound clusters with one or two N_2 molecules, which are used to perform messenger-tagging infrared (IR) spectroscopy. The tagged ions are irradiated for 50 ms with a continuous-wave (CW), mid-infrared (IR), fiber-pumped laser (IPG Photonics), operating at 0.5 - 1 W output power. Finally, the ions are detected by a commercial time-of-flight (ToF) mass spectrometer (Tofwerk).

3.5.1 SLIM-IMS

Figure 3.14 shows the SLIM device used in the second-generation instrument. The serpentine path on the SLIM board deployed here is approximately 10 m in length, with a single-pass resolving power of approximately 200. This device has an onboard collection region from which the ions are pulsed. After reaching the exit of the SLIM, the ions can be either sent for detection or rerouted through the serpentine path for additional separation. In addition, this board has ion trap segments used for confining mobility separated ions, which can be trapped for prolonged periods of time with minimal loss and fragmented by CID. In this case, CID is performed by inserting two metallic grids separated by 0.8 mm across which we apply a potential of a few hundred volts.[16] Ions are accelerated as they pass through these grids and fragmented by collisions with the nitrogen carrier gas.

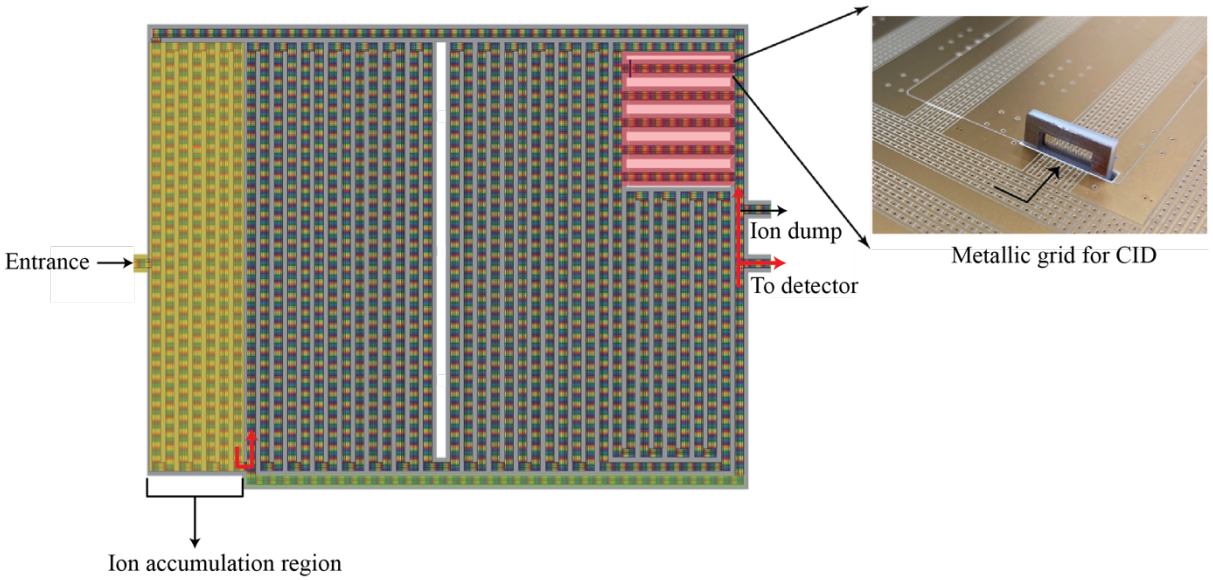


Figure 3.14. Schematic of the SLIM-PCB used in the second-generation instrument. It features an on-board accumulation region highlighted in yellow, a bypass section to avoid separation, highlighted in green, and ion/CID traps highlighted in red. A picture of the metallic grids is shown in the inset.

References:

1. Melmer, M., Stangler, T., Premstaller, A., and Lindner, W. (2011) Comparison of hydrophilic-interaction, reversed-phase and porous graphitic carbon chromatography for glycan analysis, *J Chromatogr A 1218*, 118-123.
2. West, C., Elfakir, C., and Lafosse, M. (2010) Porous graphitic carbon: a versatile stationary phase for liquid chromatography, *J Chromatogr A 1217*, 3201-3216.
3. Knox, J. H., Kaur, B., and Millward, G. R. (1986) Structure and Performance of Porous Graphitic Carbon in Liquid-Chromatography, *Journal of Chromatography 352*, 3-25.
4. Pereira, L. (2008) Porous graphitic carbon as a stationary phase in HPLC: Theory and applications, *Journal of Liquid Chromatography & Related Technologies 31*, 1687-1731.
5. Goncharova, E. N., Statkus, M. A., Tsizin, G. I., and Zolotov, Y. A. (2020) Porous Graphitized Carbon for the Separation and Preconcentration of Hydrophilic Substances, *Journal of Analytical Chemistry 75*, 423-442.
6. Kříž, J., Adamcová, E., Knox, J. H., and Hora, J. (1994) Characterization of adsorbents by high-performance liquid chromatography using aromatic hydrocarbons Porous graphite and its comparison with silica gel, alumina, octadecylsilica and phenylsilica, *Journal of Chromatography A 663*, 151-161.
7. Knox, J. H., and Ross, P. (1997) Carbon-based packing materials for liquid chromatography - Structure, performance, and retention mechanisms, *Advances in Chromatography, Vol 37 37*, 73-119.
8. Möckel, H. J., Braedikow, A., Melzer, H., and Aced, G. (2006) A Comparison of the Retention of Homologous Series and Other Test Solutes on an Ods Column and a Hypercarb Carbon Column, *Journal of Liquid Chromatography 14*, 2477-2498.
9. Bassler, B. J., Kaliszan, R., and Hartwick, R. A. (1989) Retention Mechanisms on Metallic Stationary Phases, *Journal of Chromatography 461*, 139-147.
10. West, C., and Lesellier, E. (2005) Effects of modifiers in subcritical fluid chromatography on retention with porous graphitic carbon, *J Chromatogr A 1087*, 64-76.
11. Gaudin, K., Chaminade, P., and Baillet, A. (2002) Eluotropic strength in non-aqueous liquid chromatography with porous graphitic carbon, *J Chromatogr A 973*, 61-68.

12. Deng, L., Ibrahim, Y. M., Hamid, A. M., Garimella, S. V., Webb, I. K., Zheng, X., Prost, S. A., Sandoval, J. A., Norheim, R. V., Anderson, G. A., Tolmachev, A. V., Baker, E. S., and Smith, R. D. (2016) Ultra-High Resolution Ion Mobility Separations Utilizing Traveling Waves in a 13 m Serpentine Path Length Structures for Lossless Ion Manipulations Module, *Anal Chem 88*, 8957-8964.
13. Hamid, A. M., Garimella, S. V. B., Ibrahim, Y. M., Deng, L., Zheng, X., Webb, I. K., Anderson, G. A., Prost, S. A., Norheim, R. V., Tolmachev, A. V., Baker, E. S., and Smith, R. D. (2016) Achieving High Resolution Ion Mobility Separations Using Traveling Waves in Compact Multiturn Structures for Lossless Ion Manipulations, *Anal Chem 88*, 8949-8956.
14. 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., and Smith, R. D. (2017) Serpentine Ultralong Path with Extended Routing (SUPER) High Resolution Traveling Wave Ion Mobility-MS using Structures for Lossless Ion Manipulations, *Anal Chem 89*, 4628-4634.
15. Hollerbach, A. L., Conant, C. R., Nagy, G., and Ibrahim, Y. M. (2022) Implementation of Ion Mobility Spectrometry-Based Separations in Structures for Lossless Ion Manipulations (SLIM), *Methods Mol Biol 2394*, 453-469.
16. Warnke, S., Ben Faleh, A., and Rizzo, T. R. (2021) Toward High-Throughput Cryogenic IR Fingerprinting of Mobility-Separated Glycan Isomers, *ACS Meas Sci Au 1*, 157-164.
17. Rizzo, T. R., and Boyarkin, O. V. (2015) Cryogenic Methods for the Spectroscopy of Large, Biomolecular Ions, In *Gas-Phase IR Spectroscopy and Structure of Biological Molecules* (Rijs, A. M., and Oomens, J., Eds.), pp 43-97, Springer International Publishing, Cham.
18. Okumura, M., Yeh, L. I.-C., Myers, J. D., and Lee, Y. P. (1986) Infrared spectra of the cluster ions H7O+3·H2 and H9O+4·H2, *Journal of Chemical Physics 85*, 2328-2329.
19. Kamrath, M. Z., Garand, E., Jordan, P. A., Leavitt, C. M., Wolk, A. B., Van Stipdonk, M. J., Miller, S. J., and Johnson, M. A. (2011) Vibrational characterization of simple peptides using cryogenic infrared photodissociation of H2-tagged, mass-selected ions, *J Am Chem Soc 133*, 6440-6448.
20. Cismesia, A. P., Bailey, L. S., Bell, M. R., Tesler, L. F., and Polfer, N. C. (2016) Making Mass Spectrometry See the Light: The Promises and Challenges of Cryogenic Infrared Ion Spectroscopy as a Bioanalytical Technique, *J Am Soc Mass Spectrom 27*, 757-766.

21. Khanal, N., Masellis, C., Kamrath, M. Z., Clemmer, D. E., and Rizzo, T. R. (2017) Glycosaminoglycan Analysis by Cryogenic Messenger-Tagging IR Spectroscopy Combined with IMS-MS, *Anal Chem 89*, 7601-7606.
22. Palzer, S. (2020) Photoacoustic-Based Gas Sensing: A Review, *Sensors (Basel) 20*.
23. Bozóki, Z., Sneider, J., Gingl, Z., Mohácsi, Á., Szakáll, M., Bor, Z., and Szabó, G. (1999) A high-sensitivity, near-infrared tunable-diode-laser-based photoacoustic water-vapour-detection system for automated operation, *Measurement Science and Technology 10*, 999-1003.
24. Gordon, I. E., Rothman, L. S., Hargreaves, R. J., Hashemi, R., Karlovets, E. V., Skinner, F. M., Conway, E. K., Hill, C., Kochanov, R. V., Tan, Y., Wcisło, P., Finenko, A. A., Nelson, K., Bernath, P. F., Birk, M., Boudon, V., Campargue, A., Chance, K. V., Coustenis, A., Drouin, B. J., Flaud, J. M., Gamache, R. R., Hodges, J. T., Jacquemart, D., Mlawer, E. J., Nikitin, A. V., Perevalov, V. I., Rotger, M., Tennyson, J., Toon, G. C., Tran, H., Tyuterev, V. G., Adkins, E. M., Baker, A., Barbe, A., Canè, E., Császár, A. G., Dudaryonok, A., Egorov, O., Fleisher, A. J., Fleurbaey, H., Foltynowicz, A., Furtenbacher, T., Harrison, J. J., Hartmann, J. M., Horneman, V. M., Huang, X., Karman, T., Karns, J., Kassi, S., Kleiner, I., Kofman, V., Kwabia–Tchana, F., Lavrentieva, N. N., Lee, T. J., Long, D. A., Lukashevskaya, A. A., Lyulin, O. M., Makhnev, V. Y., Matt, W., Massie, S. T., Melosso, M., Mikhailenko, S. N., Mondelain, D., Müller, H. S. P., Naumenko, O. V., Perrin, A., Polyansky, O. L., Raddaoui, E., Raston, P. L., Reed, Z. D., Rey, M., Richard, C., Tóbiás, R., Sadiek, I., Schwenke, D. W., Starikova, E., Sung, K., Tamassia, F., Tashkun, S. A., Vander Auwera, J., Vasilenko, I. A., Vigasin, A. A., Villanueva, G. L., Vispoel, B., Wagner, G., Yachmenev, A., and Yurchenko, S. N. (2022) The HITRAN2020 molecular spectroscopic database, *Journal of Quantitative Spectroscopy and Radiative Transfer 277*, 107949.
25. Ben Faleh, A., Warnke, S., and Rizzo, T. R. (2019) Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures, *Anal Chem 91*, 4876-4882.
26. Masson, A., Kamrath, M. Z., Perez, M. A., Glover, M. S., Rothlisberger, U., Clemmer, D. E., and Rizzo, T. R. (2015) Infrared Spectroscopy of Mobility-Selected H+-Gly-Pro-Gly-Gly (GPGG), *J Am Soc Mass Spectrom 26*, 1444-1454.

27. Lorenz, U. J., and Rizzo, T. R. (2011) Planar multipole ion trap/time-of-flight mass spectrometer, *Anal Chem 83*, 7895-7901.
28. Bansal, P., Yatsyna, V., AbiKhodr, A. H., Warnke, S., Ben Faleh, A., Yalovenko, N., Wysocki, V. H., and Rizzo, T. R. (2020) Using SLIM-Based IMS-IMS Together with Cryogenic Infrared Spectroscopy for Glycan Analysis, *Anal Chem 92*, 9079-9085.

Chapter 4 Identifying mixtures of isomeric human milk oligosaccharides by the decomposition of IR spectral fingerprints[1]

The analysis of glycans presents a significant challenge that arises from their isomeric heterogeneity. While high-resolution ion mobility spectrometry (IMS) has shown the ability to resolve subtly different glycan isomers, their unambiguous assignment remains difficult. Here, we demonstrate an infrared (IR) spectroscopic approach for identifying isomers in a glycan mixture. To display the feasibility of this approach, we have constructed a small database of cryogenic spectra of five lacto-N-fucopentaose (LNFP) and six disaccharide isomers and demonstrated that in the cases where they cannot be separated by IMS, we can use a cryogenic IR spectrum to identify the isomeric components.

4.1 Introduction

This study aims to demonstrate that cryogenic IR spectra can be used in a database approach for the unambiguous identification of glycan isomers in a mixture. Database approaches are used for identifying glycans on the basis of tandem MS[1-5] as well as collision cross sections determined by ion mobility.[6] In our approach, we create a spectroscopic database using high-resolution ion mobility to separate isomers and then measure an IR spectrum of the isomerically pure species. For the initial creation of the IR reference database, high ion mobility resolving power is necessary, but once these spectra are measured, any separation technique (LC, IMS, etc.) can be used in conjunction with IR spectroscopy for identification and quantification of the isomeric composition. Even components that are not fully separated by ion mobility or liquid chromatography can be identified by decomposition of the cryogenic IR spectra of the

[1] This chapter is based on the publication "Abikhodr, A. H., Yatsyna, V., Ben Faleh, A., Warnke, S., and Rizzo, T. R., Identifying Mixtures of Isomeric Human Milk Oligosaccharides by the Decomposition of IR Spectral Fingerprints." *Anal Chem* **2021,** *93* (44), 14730-14736. The content and figures were adapted with permission from the American Chemical Society.

overlapping isomers, as long as the corresponding isomerically pure reference spectra have been added to the database.

To illustrate the feasibility of this approach, we use a set of five isomeric human milk oligosaccharides (HMOs), lacto-N-fucopentaoses (LNFPs), from which we construct a small database of cryogenic IR spectra. We then use this database to determine the presence of these glycans in mixtures.

4.2 Experimental Methods

4.2.1 Sample Preparation

Human milk oligosaccharides (HMOs) LNFP I, II, III, and V were acquired from Carbosynth Ltd, and LNFP VI was purchased from Dextra Laboratories Ltd. An HMO mixture that includes 60% of penta- and hexasaccharides was purchased from Carbosynth Ltd and made into a 0.1 mg/ml solution. All samples were prepared in a 50/50 mixture of water/methanol and used without further purification. Molecules were analyzed in the sodiated charge state and the protonated-potassiated charge state.

4.2.2 Instrumentation

These experiments were performed on our prototype instrument described in Chapter 3. Briefly, oligosaccharide solutions with concentrations between 10 and 25 μM are introduced into the instrument *via* a heated stainless-steel capillary (170° C) using a nano-electrospray (nESI) source. Initially, ions pass through a series of ion funnels and accumulate in an ion funnel trap (IFT). Next, short, intense ion packets (approximately 130 μs wide) are released from the IFT and directed to the SLIM-IMS region filled with helium at 3 mbar, where the molecules are separated based on their drift time.

Our previously described SLIM-IMS device was designed following the work of Smith and colleagues.[7-9] It consists of a pair of mirrored 15 x 15 cm printed circuit boards, featuring a 1.5 m single-cycle serpentine path for ion mobility separation and six individually accessible traps for storing ions with selected mobility. Long path lengths are achieved by allowing ions to cycle through the serpentine path multiple times.

Infrared spectral decomposition was performed using MATLAB. We used the fminsearch function to iteratively minimize the root-mean-square deviation (RMSD) between the spectrum of a mixture and a synthetic spectrum constructed from all normalized reference spectra. This

spectral decomposition process yields an isomer ratio that corresponds to the best fit of the measured spectrum. The resulting isomer ratios are then linearly corrected by the relative tagging efficiencies at 40 K of all isomers to give a more accurate representation of the mixture composition.

4.3 Results and discussion

4.3.1 Sodiated LNFP isomers

Our initial investigations concentrated on singly sodiated isomers of LNFP I, II, III, V, and VI (m/z = 876.3 Th). We analyzed the pure isomeric form of each species to obtain a collection of IR reference spectra (Figure 4.1), which function as identifying fingerprints. In the 3250-3700 cm^{-1} spectral region, which encompasses free and hydrogen-bonded OH and NH stretch vibrations, each isomer exhibits a distinct spectrum that is easily distinguishable from the others.

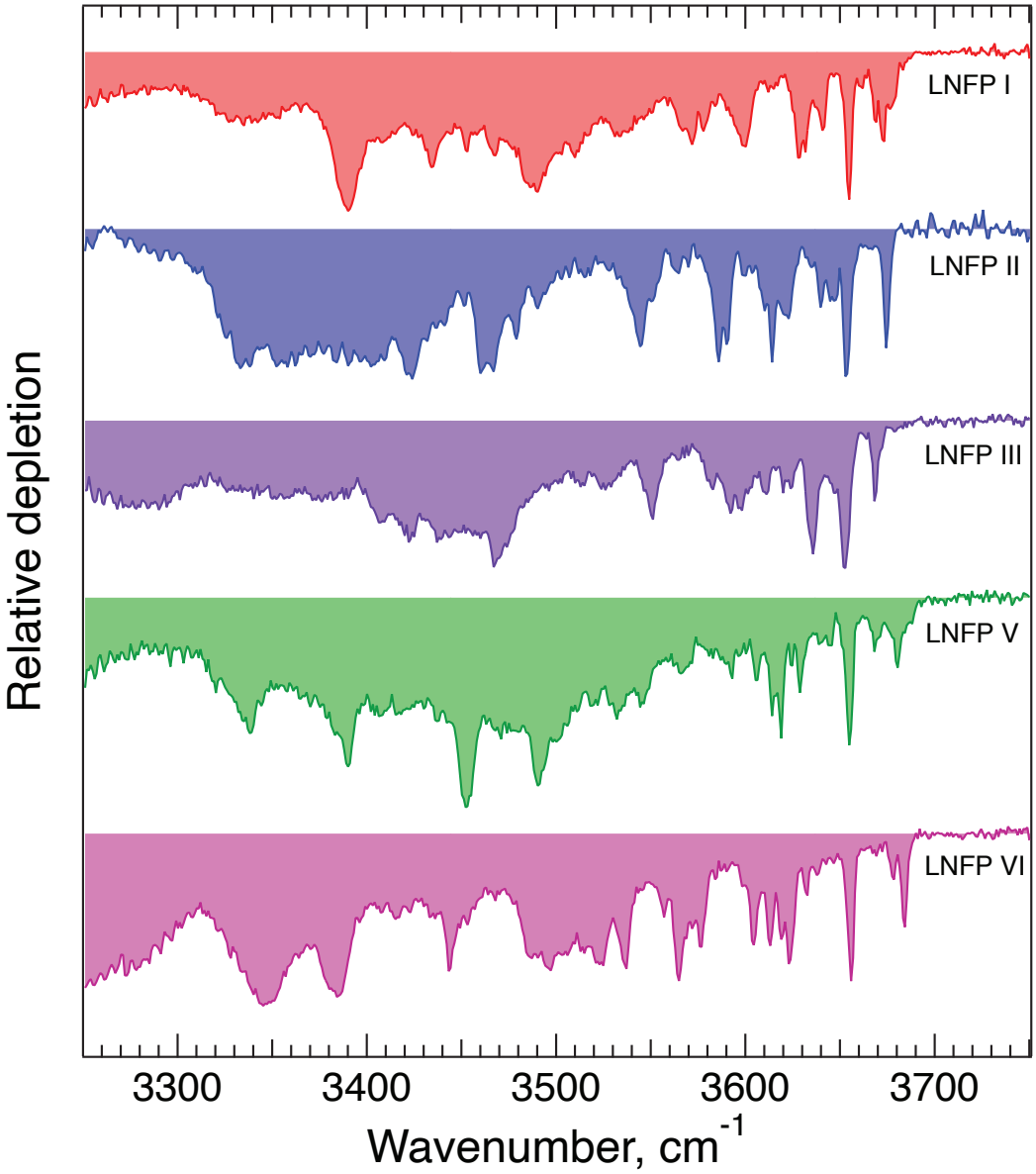


Figure 4.1. Messenger tagging IR depletion spectra of LNFP isomers in $[M+Na]^+$ state

The arrival-time distribution of each isomer sample displayed a single primary peak, although the line shapes suggest additional unresolved features in most instances (Figure 4.2). This is likely due to the overlap of reducing-end anomers, which are in equilibrium in solution before electrospray and do not interconvert in the gas phase.[10] In many cases, we can resolve these anomers and assign them, for example, by comparing with spectra of O-methylated species that are locked in a specific configuration.[11] Since we do not resolve individual reducing-end anomers in this case, all reference spectra acquired for the $[M+Na]^+$ charge state correspond to a mixture of these species.

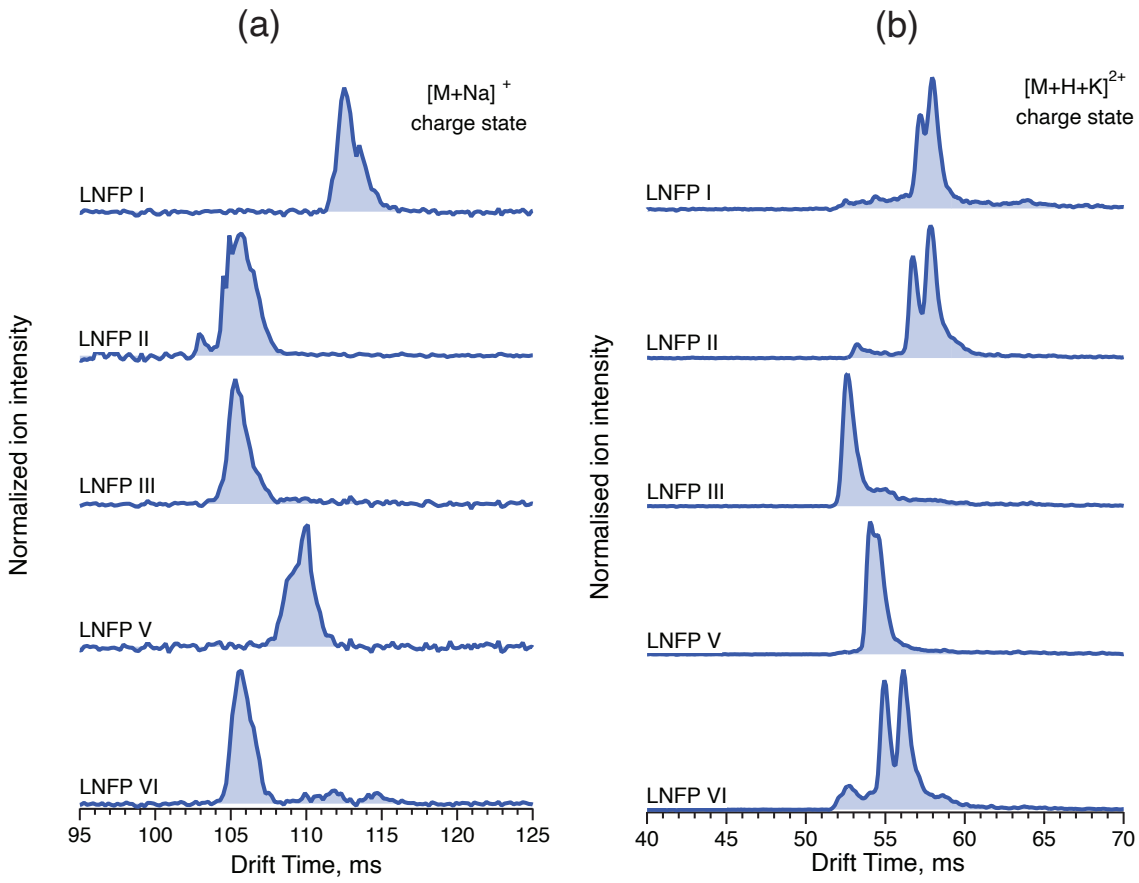


Figure 4.2. (a) ATD of LNFP I, II, III, V and VI isomers ($[M+Na]^+$ charge state) analyzed individually using similar IMS separation conditions obtained after 7 SLIM cycles, ≈ 10.5 m drift length. (b) ATD of LNFP I, II, III, V and VI isomers ($[M+H+K]^+$ charge state) analyzed individually using similar IMS separation conditions obtained after 8 SLIM cycles, ≈ 12 m drift length.

Having measured the reference spectra, we then prepared an equimolar mixture containing all five isomers, introduced them into our instrument, and used SLIM-IMS to separate them. Figure 4.3 (a) shows the ATD of the mixture after 7 cycles on the SLIM separation track (~10.5 m drift length), which exhibits three major peaks. Figure 4.3 (b,c) displays the IR spectra in colors corresponding to each of these arrival time peaks. Comparison of these spectra to our database allowed us to attribute the second and third peaks in the ATD of Figure 4.3 (a) to LNFP V and I (Figure 4.3 (b)), respectively, and the first peak to an overlap of LNFP II, III, and VI (Figure 4.3 (c,d)). However, the percentages we obtain (26%, 23.3%, and 48.7%, for LNFP II, III, and VI, respectively) do not reflect the expected isomer ratios of the equimolar solution. The reasons for this are discussed below.

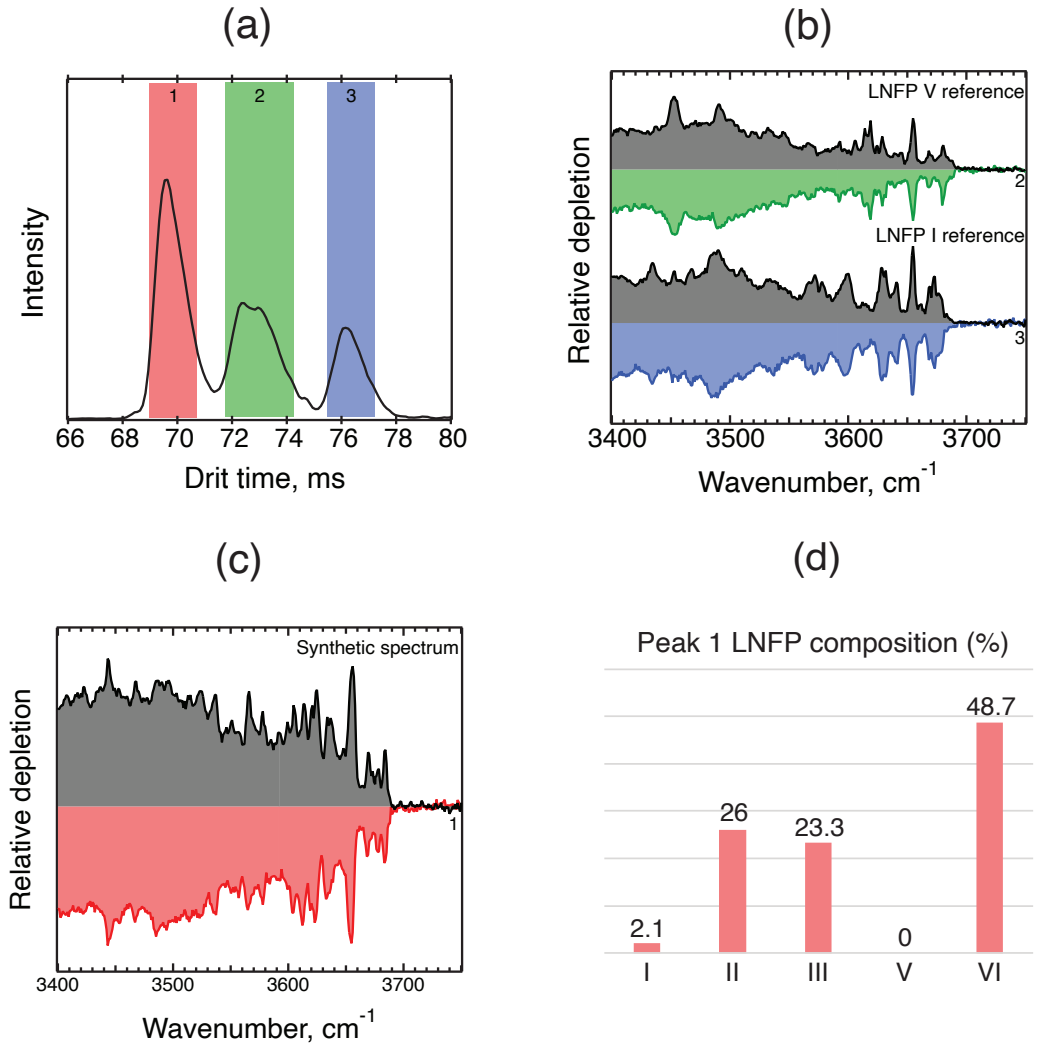


Figure 4.3. (a) Arrival time distribution of the mixture including LNFP I, II, III, V and VI ($[M+Na]^+$ charge state) obtained after 7 SLIM cycles, ≈ 10.5 m drift length; (b) messenger tagging IR depletion spectra corresponding to individual peaks 2 (green) and 3 (blue) shown in panel (a) compared with the reference spectra identifying each (gray); (c) messenger tagging IR depletion spectra of the species found in peak 1 (red) shown in panel (a) compared with the best fit synthetic spectrum identifying the isomers present (gray); (d) Bar plot showing percentages of LNFP isomers in the best fit synthetic spectrum of panel (c).

4.3.2 Protonated-potassiated LNFP isomers

To address the issue of overlapping isomers in the ATD of the sodiated species, we also investigated the protonated−potassiated doubly charged species of the same LNFP isomers (m/z = 446.6 Th). As before, we acquired a reference IR spectrum for each pure isomer (Figure 4.4), which serves as an identifying fingerprint.

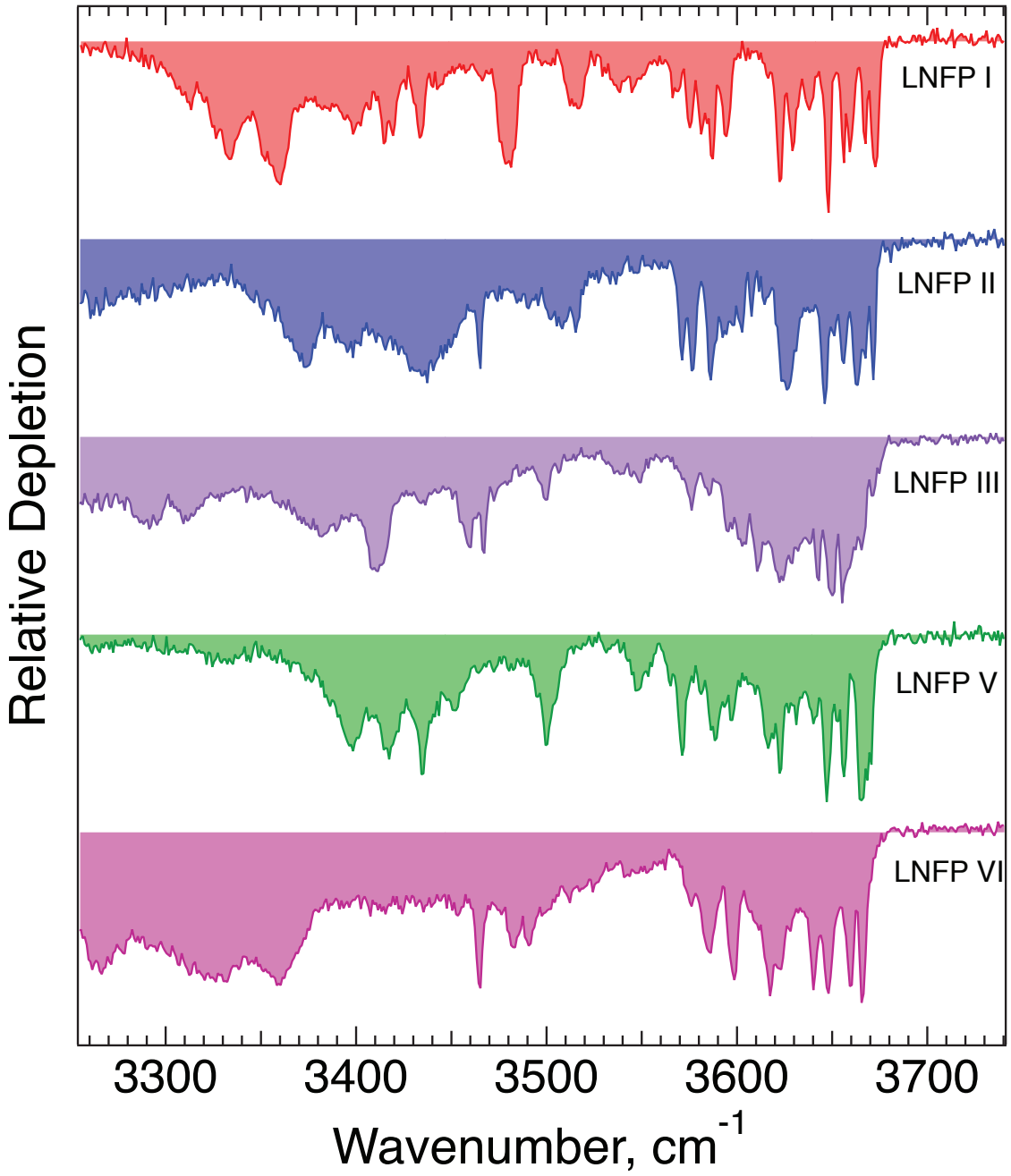


Figure 4.4. Messenger tagging IR depletion spectra of LNFP isomers in $[M+H+K]^{2+}$ state

Unlike the sodiated species, some of the LNFP isomers in the $[M + H + K]^{2+}$ charge state clearly display two major peaks in their ATD (Figure 4.2 (b)), most probably originating from the two reducing-end anomers,[11-13] which further complicates complete IMS separation of all isomers in the mixture. Examples of anomer-specific spectra are shown in Figure 4.5; these were also added to our small database.

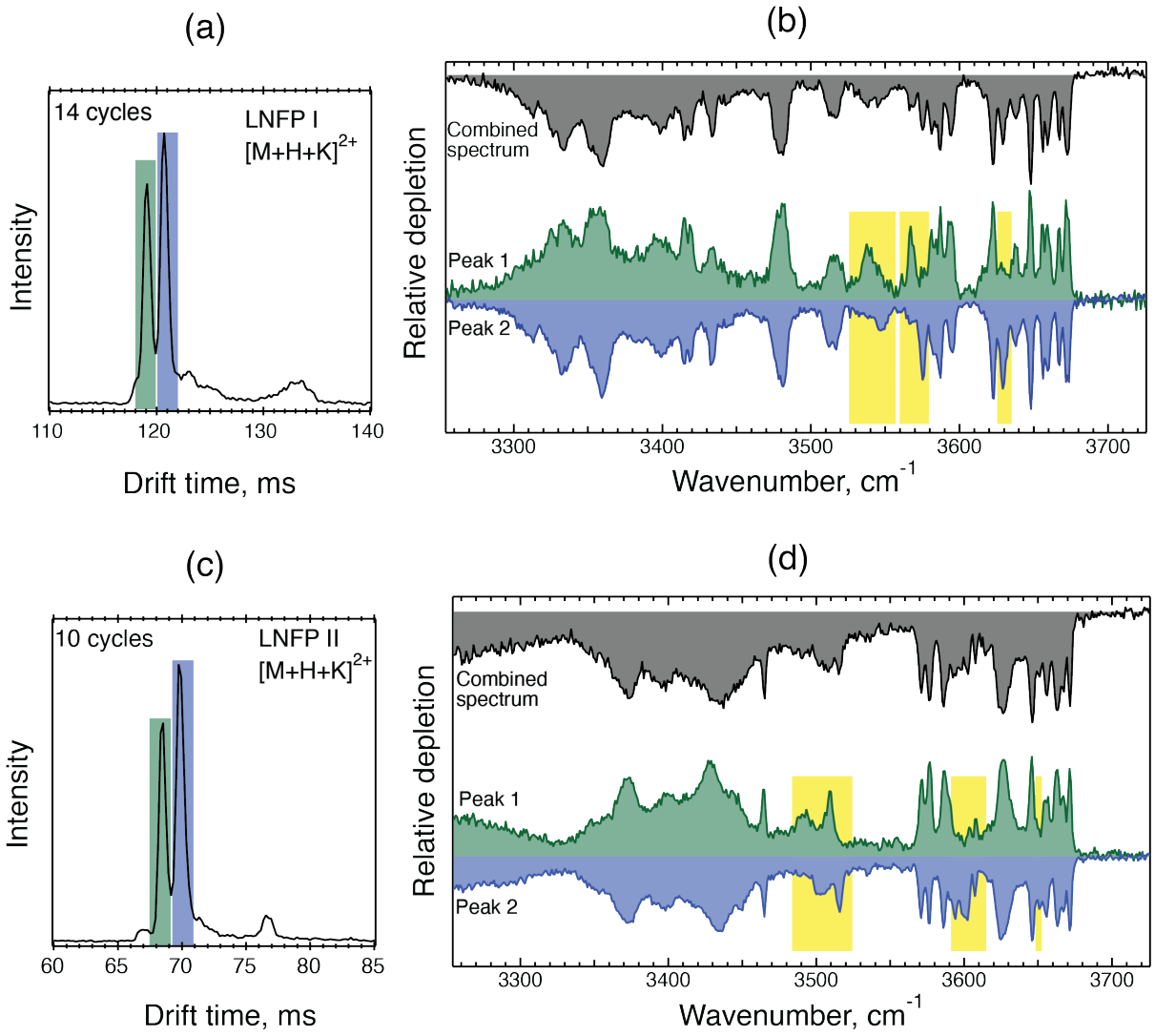


Figure 4.5. (a) Arrival time distribution of LNFP I ($[H+K]^{+2}$ charge state) obtained after 14 SLIM cycles, ≈ 21 m drift length; (b) messenger tagging IR depletion spectrum of LNFP I (in black) and depletion spectra of each highlighted region in panel (a) with yellow highlights on the spectra to show the major differences between them. (c) Arrival time distribution of LNFP II ($[H+K]^{+2}$ charge state) obtained after 10 SLIM cycles, ≈ 15 m drift length; (d) messenger tagging IR depletion spectrum of LNFP I (in black) and depletion spectra of each highlighted region in panel (c) with yellow highlights on the spectra to show the major differences between them.

Figure 4.6 shows the arrival time distribution of the mixture of all five LNFP isomers in the $[M + H + K]^{2+}$ charge state after eight separation cycles (12 m), where the identity of individual drift peaks is assigned based on their drift times (Figure 4.2 (b)). Even though LNFP II, III, and VI show partial separation, the overlap between other isomers is observed due to multiple drift peaks appearing for the individual compounds in the $[M + H + K]^{2+}$ charge state. Thus, we

cannot completely separate and identify all isomers of LNFPs present in this mixture using IMS alone, even when using both the singly sodiated and potassiated–protonated charge states. If we were to cycle the ions further to increase the resolution and separate the individual peaks, the ions would begin to overlap in the 1.5 m mobility separation region of our SLIM device.

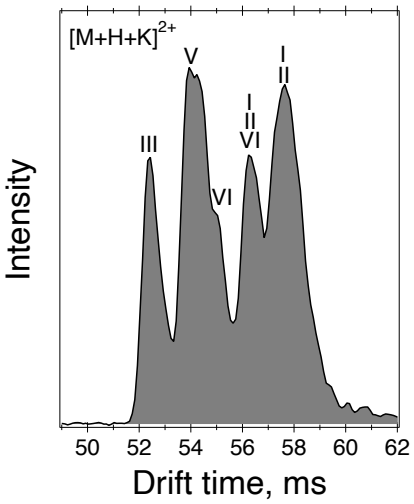


Figure 4.6. Arrival time distribution of LNFP I, II, III, V and VI mixture ($[M+H+K]^{2+}$ charge state) obtained after eight SLIM cycles (12 m drift length). Each peak is denoted with its isomer content.

4.3.3 Isomer identification using spectral decomposition

Figure 4.7 shows the optimal synthetic spectrum (red) for the equimolar mixture of LNFP I and II in the $[M + H + K]^{2+}$ charge state obtained from the MATLAB decomposition algorithm, compared to the measured spectrum (gray). Even though these species almost completely overlap in their ATD after a 12 m separation drift length, the algorithm correctly identifies them as major constituents of the mixture (43.9% LNFP I and 55.5% LNFP II), demonstrating the sensitivity of the cryogenic IR spectra to the slightest structural differences.

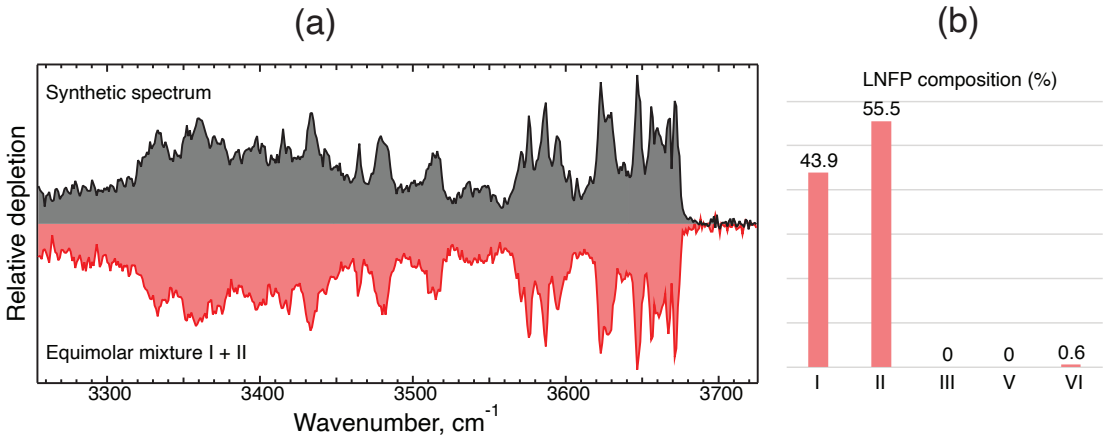


Figure 4.7. (a) Messenger tagging IR depletion spectra of an equimolar mixture LNFP I and II in the $[M+H+K]^{2+}$ charge state (red) compared with the best fit synthetic spectrum (gray); (b) a bar plot showing percentages corresponding to the best fit synthetic spectrum.

To further assess the performance of spectral decomposition, the reference cryogenic IR spectra of a series of six isomeric disaccharides were measured in the $[M + Na]^+$ charge state (Figure 4.8).

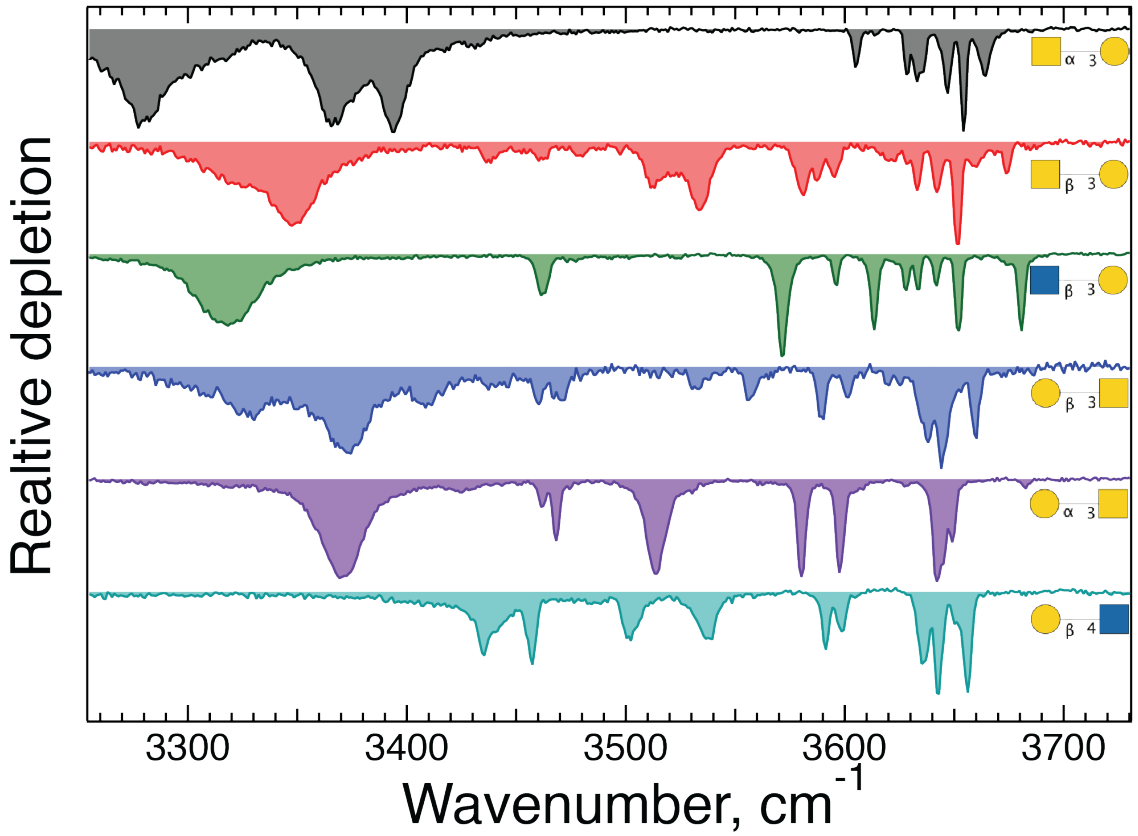


Figure 4.8. Messenger tagging IR depletion spectra of disaccharide isomers ($[M+Na]^+$ state)

Various mixtures of these disaccharides were then prepared, and the IR spectrum of each mixture without IMS separation was measured and analyzed. The decomposition algorithm successfully identified the constituent isomer makeup of each spectrum (Figures 4.9–4.15), with a detection limit of 2.9% (2σ) in the mixtures we examined. This value was determined by considering all of the percentages of falsely identified isomers. This method is thus well suited for solving the problem of overlapping ion mobility peaks and can give reliable identification of the constituent isomer makeup in complex mixtures, although as discussed more fully below, our ability to perform a quantitative analysis is still limited.

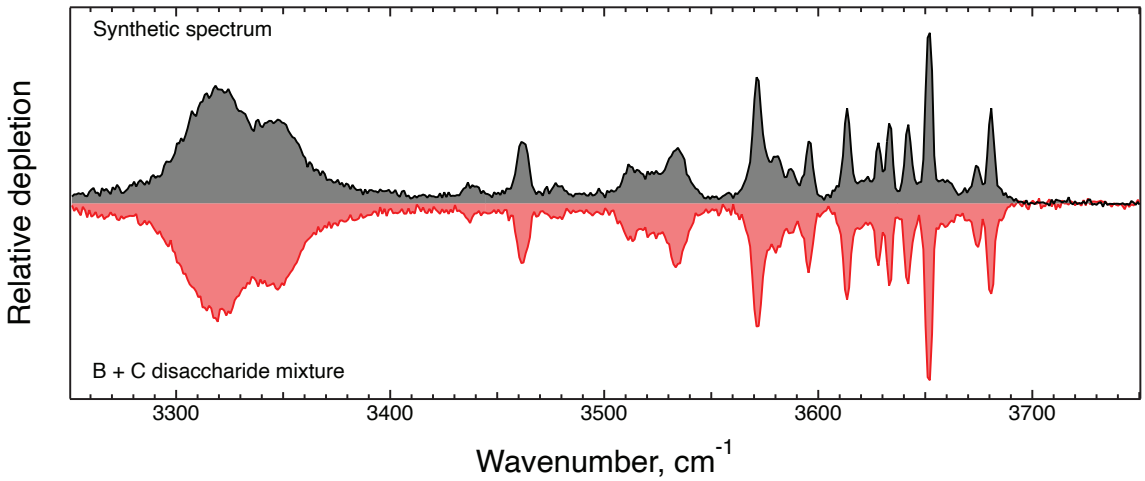


Figure 4.9. Messenger tagging IR depletion spectra of equimolar disaccharide isomers B and C in $[M+Na]^+$ state (red) compared with the optimized synthetic spectrum (gray). Decomposition: **B (49.9%), C (49.3%)** and F (0.8%).

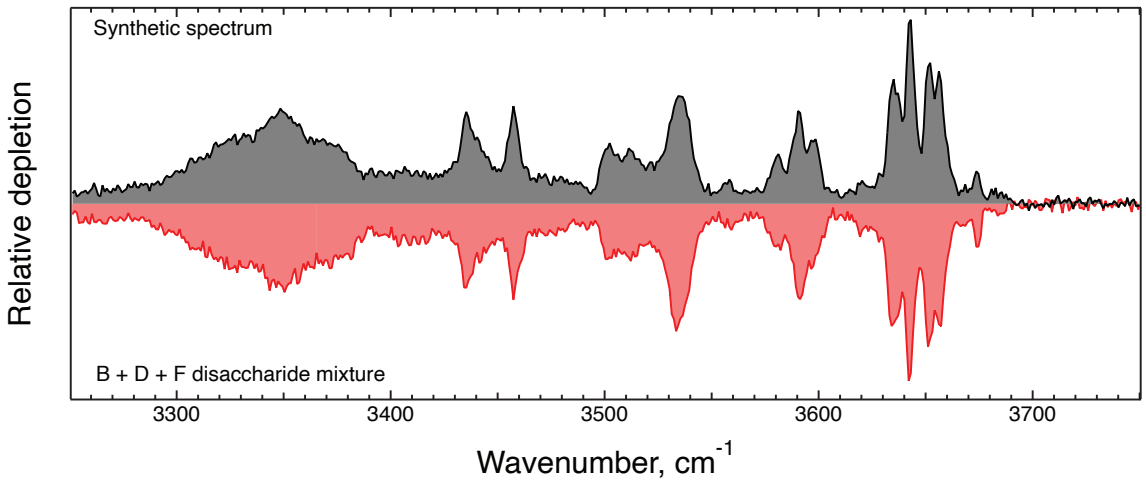


Figure 4.10. Messenger tagging IR depletion spectra of disaccharide isomers B, D and F (15, 7, and 15 μM respectively) in $[M+Na]^+$ state (red) compared with the optimized synthetic spectrum (gray). Decomposition: A (0.6%), **B (34.1%)**, C (2.1%), **D (26.5%)**, E (0%), and **F (36.7%)**.

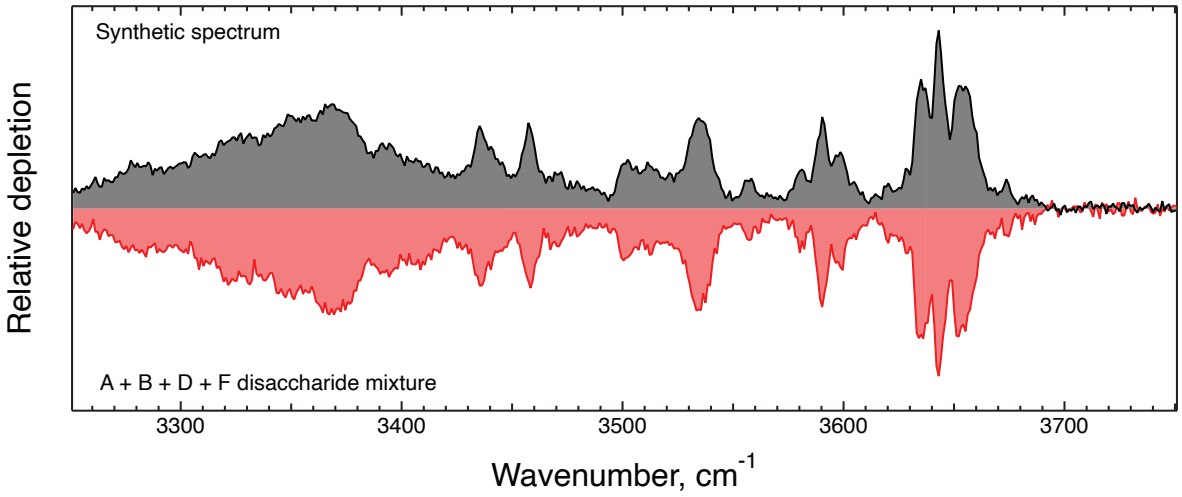


Figure 4.11. Messenger tagging IR depletion spectra of equimolar disaccharide isomers A,B,D and F in $[M+Na]^+$ state (red) compared with the optimized synthetic spectrum (gray). Decomposition: **A (7.1%)**, **B (23.4%),** C (1.2%), **D (41.9%)**, E (0%), and **F (26.4%)**.

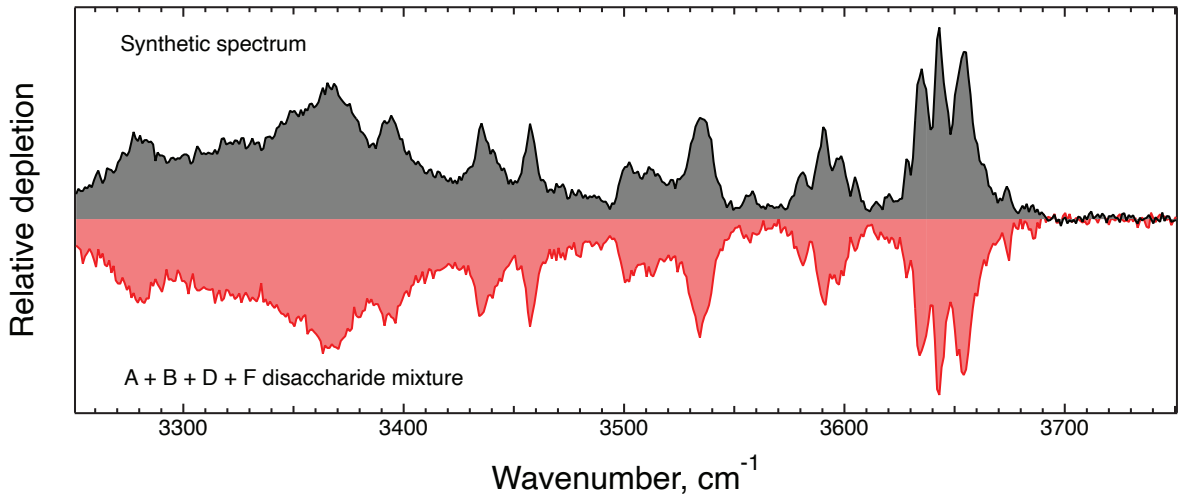


Figure 4.12. Messenger tagging IR depletion spectra of disaccharide isomers A, B, D and F (30, 15, 7, 15 μM respectively) in $[M+Na]^+$ state (red) compared with the optimized synthetic spectrum (gray). Decomposition: **A (15.2%)**, **B (26.2%)**, C (1.4%), **D (28.7%)**, E (0.2%), and **F (28.3%)**.

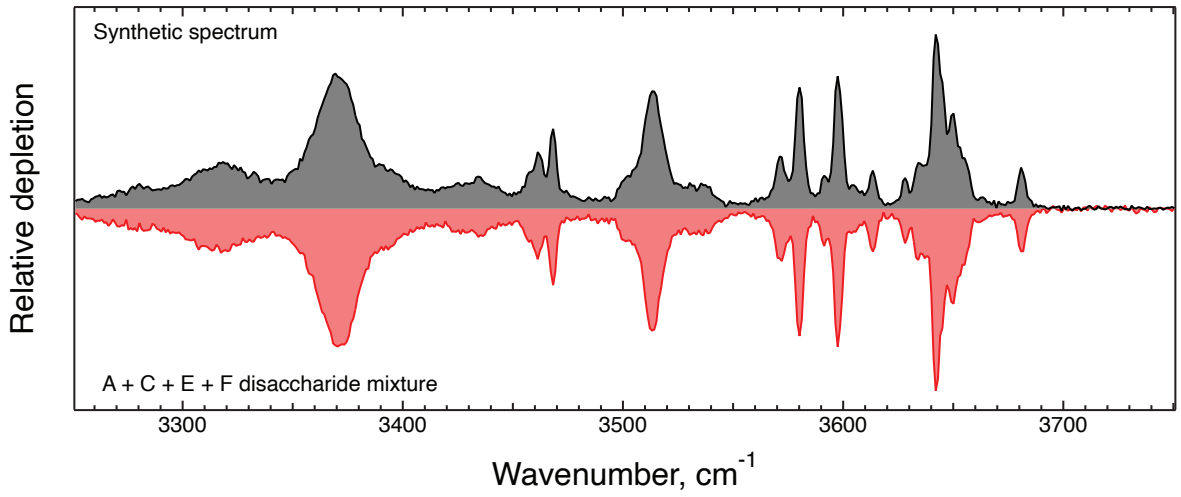


Figure 4.13. Messenger tagging IR depletion spectra of equimolar disaccharide isomers A, C, E, and F in $[M+Na]^+$ state (red) compared with the optimized synthetic spectrum (gray). Decomposition: **A (6%)**, B (0%), **C (22.4%)**, D (0%), **E (48%)**, and **F (23.6%)**.

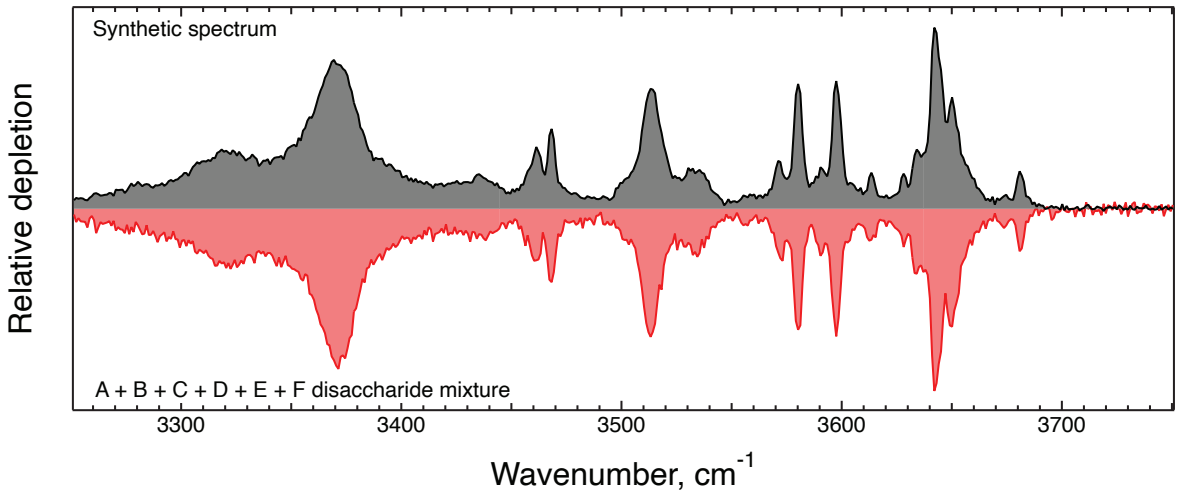


Figure 4.14. Messenger tagging IR depletion spectra of all six disaccharide isomers (equimolar) in $[M+Na]^+$ state (red) compared with the optimized synthetic spectrum (gray). Decomposition: **A (4.2%), B (13.3%), C (14.1%), D (22.2%), E (32.3%), and F (13.9%).**

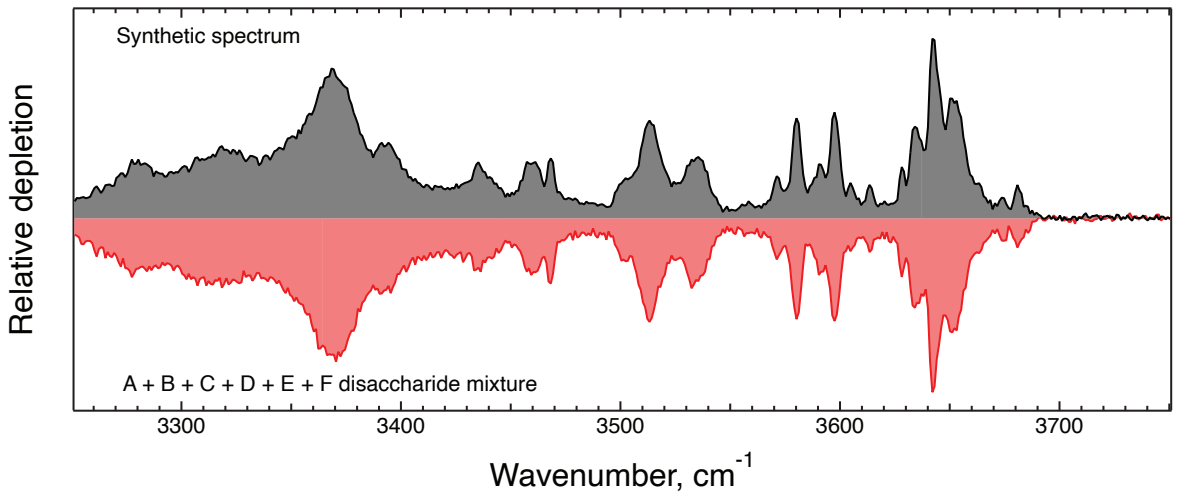


Figure 4.15. Messenger tagging IR depletion spectra of all six disaccharide isomers (30, 15, 7, 7, 7, 15 μM respectively, starting with sugar A) in $[M+Na]^+$ state (red) compared with the optimized synthetic spectrum (gray). Decomposition: **A (13.5%), B (21.6%), C (9.8%), D (16.7%), E (17.5%), and F (20.9%).**

4.3.4 Application to an unknown mixture

Having applied our database approach to mixtures of known concentration, we proceeded to analyze a sample obtained from Carbosynth Ltd., which contains an unknown mixture of HMOs including lacto-*N*-fucopentaoses. According to the supplier, the raw material for this sample was obtained from a panel of subjects and purified using liquid chromatography. Figure 5a shows the ATD of sodiated LNFP found in this sample after 7 cycles on our SLIM board. Visual comparison of this ATD to that in Figure 4.3 (a) suggests the presence of at least one of LNFP II, III, and VI in the first peak, the possibility of LNFP V in the middle region, and LNFP I in the second, well-resolved peak. It is important to note that the ATDs shown in Figures 2 and 5 correspond to the same drift length, however, the drift times differ due to different separation conditions used in each experiment (different traveling wave pulse voltage and speed). A small peak appears in the ATD at about 114 ms, which is not present in Figure 4.3 (a). We believe this is an artifact, perhaps from a solvent cluster, since it disappears when the drift gas is changed from He to N_2 (Figure 4.16).

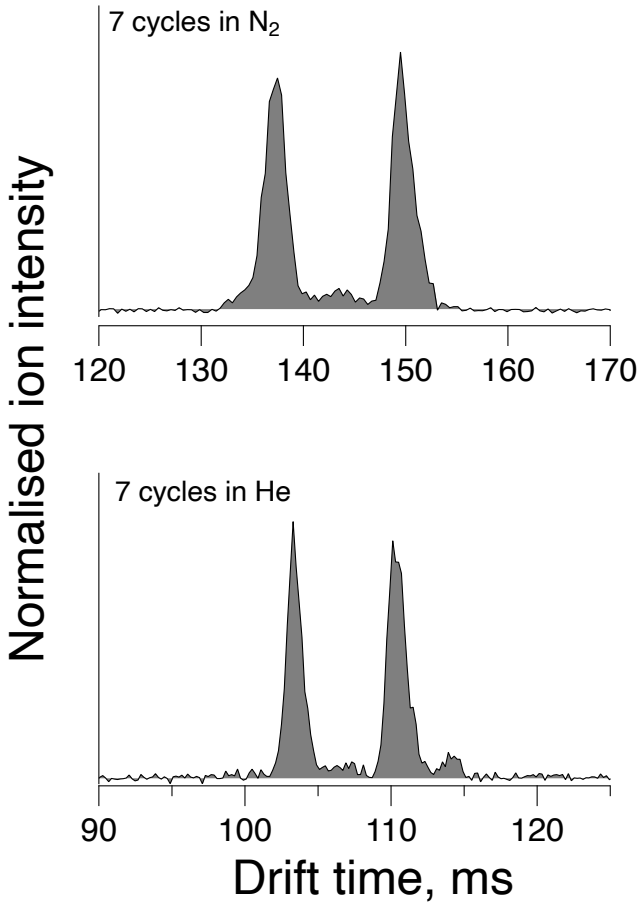


Figure 4.16. Arrival time distribution of LNFP mixture found in the Carbosynth HMO sample ($[M+Na]^+$ state) obtained after 7 SLIM cycles (~10.5 m drift length) using nitrogen (top) and helium (bottom) in the SLIM-IMS region.

The cryogenic IR spectrum of the first ATD region was measured and analyzed using the spectral decomposition method as described above. Indeed, Figure 4.17 (b) shows that the synthetic spectrum using the percentages returned by the decomposition algorithm (gray trace) is virtually identical to the measured spectrum (red trace), with its major components being LNFP II (61.8%) and III (37.5%). The algorithm also shows a minor contribution of LNFP VI (0.7%), which is below the limit of detection using this approach. Thus, we conclude that the first peak in the ATD of Figure 4.17 (a) contains primarily LNFP II and III. The IR spectrum of region 2, shown in Figure 4.17 (c), indicates an exact match with LNFP I, thus confirming its presence.

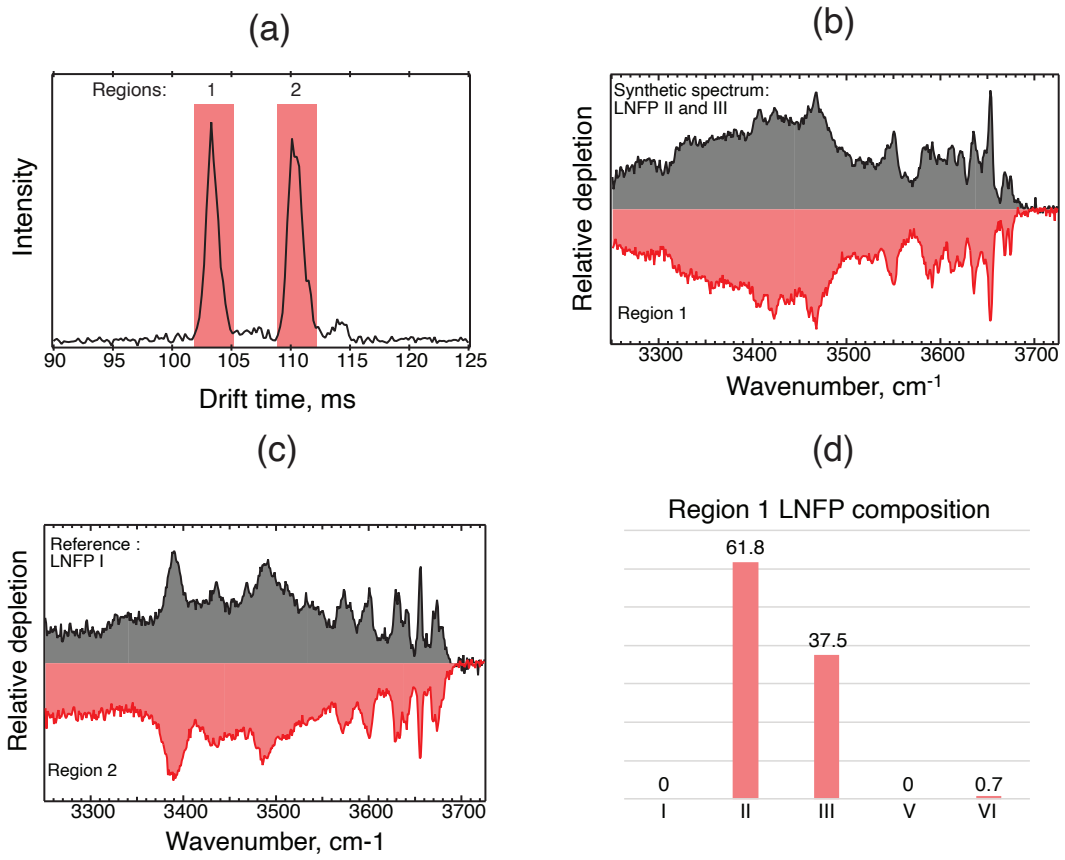


Figure 4.17. (a) Arrival time distribution of LNFP found in the Carbosynth HMO sample ($[M+Na]^+$ charge state) obtained after 7 SLIM cycles (~10.5 m drift length); (b) Messenger tagging IR spectra of species found in the ATD region 1 (red) , plotted against the best fit synthetic spectra identifying the isomers present (gray); (c) region 2 spectrum identifying LNFP I; (d) Bar plot showing percentages of LNFP isomers in the best fit synthetic spectrum of panel (b).

We also found that the spectrum of the ions arriving between the two main peaks in the ATD corresponds to that of LNFP V in our database (Figure 4.18).

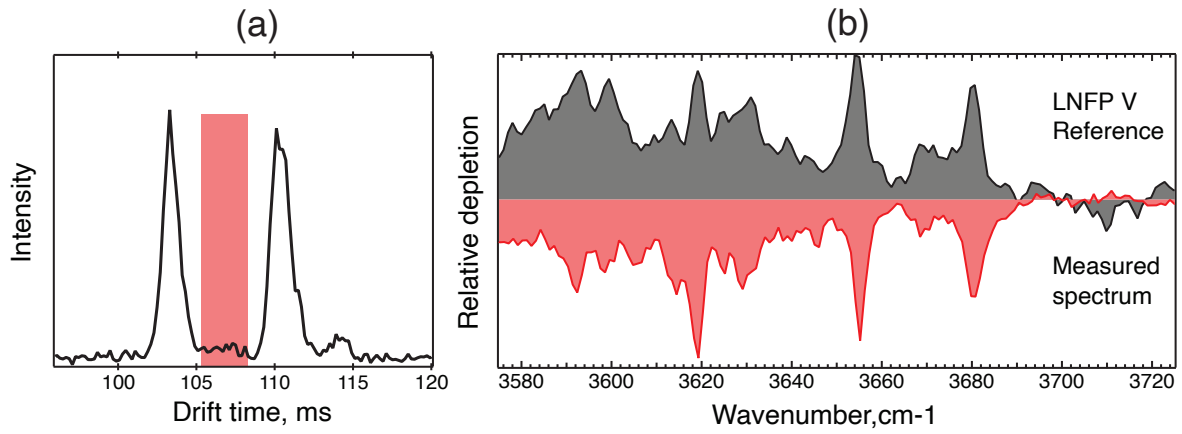


Figure 4.18. (a) Arrival time distribution of LNFP mixture found in the Carbosynth HMO sample ($[M+Na]^+$ state) obtained after 7 SLIM cycles(~10.5 m drift length). (b) Messenger tagging IR depletion spectra of highlighted region in (a) (red) compared with the reference spectrum identifying it as LNFP V (gray).

4.3.5 Issues related to quantitation

While we have demonstrated that spectral decomposition is well suited for the identification of isomers present in the mixtures, the precise quantitation of the isomer content in the mixture solution is complicated by several issues that are intrinsic to ESI-MS analysis.[14, 15] These issues include possible differences in the ionization efficiencies and charge-state distributions of different isomers, as well as charge competition upon ESI.[16] For example, the spectral decomposition of an equimolar mixture of LNFP I and II, as well as some of the binary mixtures of disaccharides that we studied, showed nearly equal ratios between each isomer, whereas in equimolar solutions with more than two isomers, at least one isomer was underestimated (Figures 4.10–4.15). It is important to emphasize that this issue is not related to the spectral decomposition itself since this method accurately determines the gas-phase isomer ratios even though they might differ from the solution ratios due to the aforementioned effects.

4.4 Separating overlapping drift peaks through borohydride reduction of glycans

The reduction of glycans using sodium borohydride is a well-established reaction that plays an important role in glycan analysis. It is primarily carried out to convert carbonyl groups (aldehydes and ketones) present in the glycans into alcohol groups (primary alcohols). This

then freezes the glycan in the open-ring form, significantly changing the collisional cross section.

Figure 4.19 shows the reduction mechanism glycans undergo in the presence of sodium borohydride ($NaBH_4$). The activated sodium borohydride, typically as BH_4^-, attacks the carbonyl group of the glycan (open-ring), resulting in the reduction of the carbonyl carbon to a hydroxyl group. The reaction produces an alkoxide ion intermediate. The alkoxide ion is then protonated by water or other acidic species present in the reaction mixture. The result is the conversion of the carbonyl group into a hydroxyl group, leading to the reduced glycan. The reduced glycan will differ by a m/z ratio of +2, because two hydrogen atoms have been added.

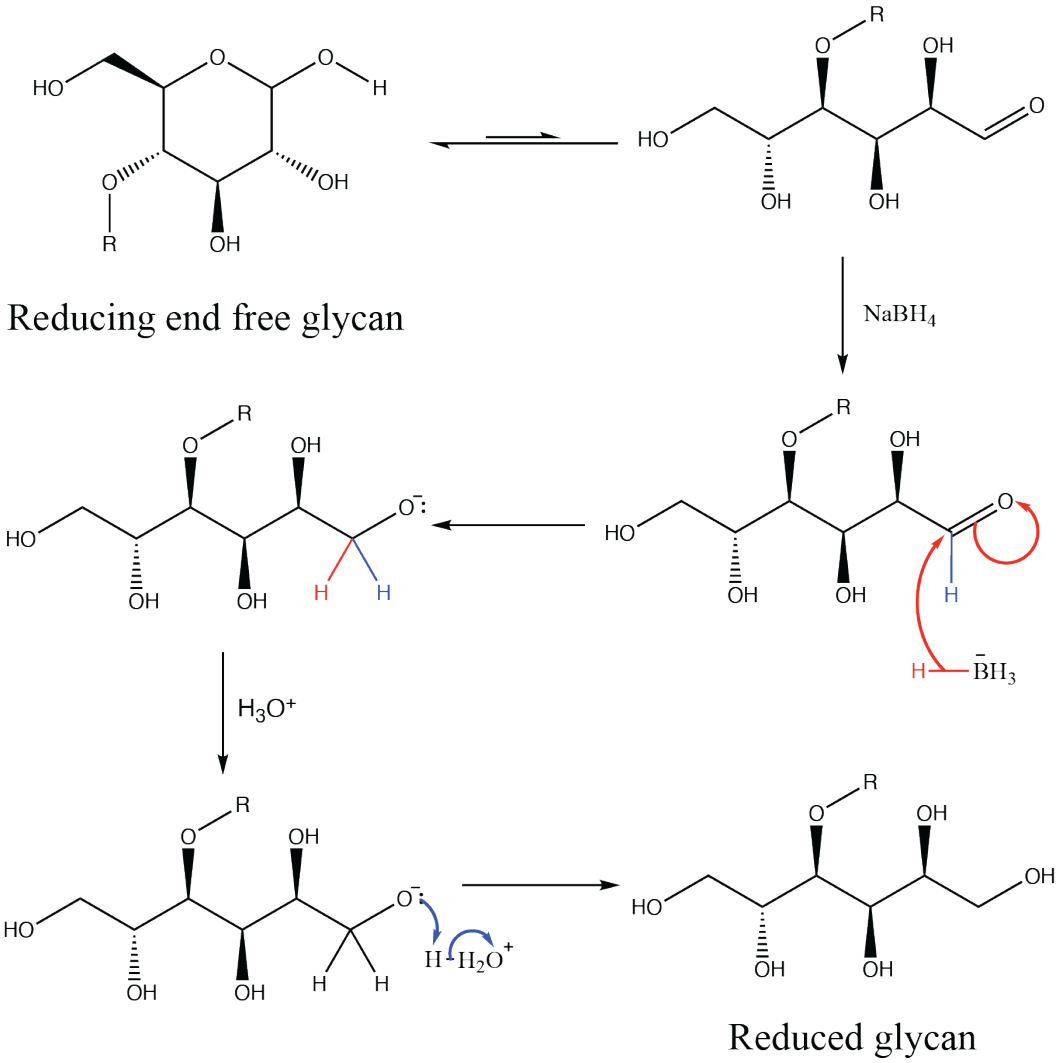


Figure 4.19. Schematic showing the mechanism of reducing oligosaccharides in the presence of $NaBH_4$

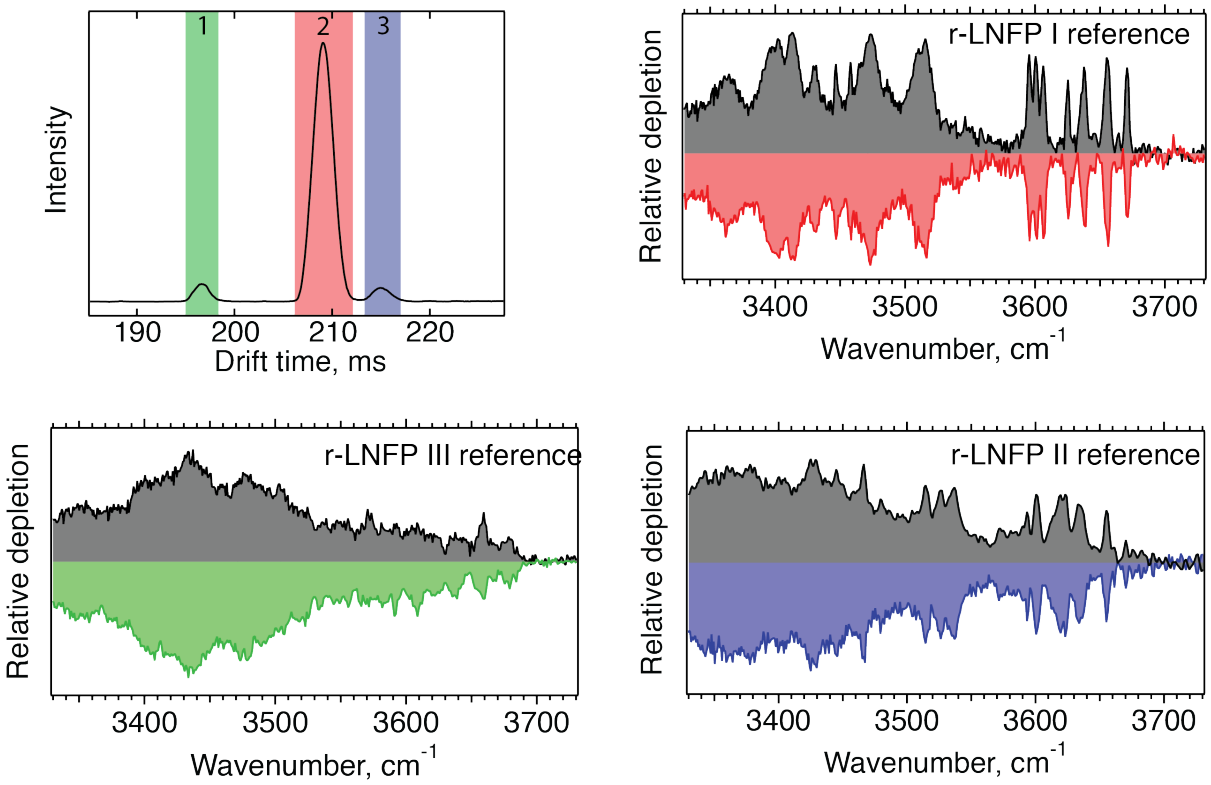
The reaction conditions applied in this work were as follows:

1. Oligosaccharide sample: 100 µL of a 50 µM sample.
2. Sodium borohydride sample: 100 µL of 2 M $NaBH_4$.
3. Mix the two samples together, place in a heated bath at 60 °C for 15 minutes.
4. Quench the reaction with formic acid until bubbling stops. (Caused by release of H_2)

After this, the resulting reduced glycans were extracted and cleaned from the excess salts using a PGC solid phase exchange cartridge.

By reducing the carbonyl groups to hydroxyl groups, the overall polarity and structure of the glycans change. This alteration in molecular structure can lead to differences in CCS, affecting the mobility of the glycan ions in gas-phase ion mobility spectrometry (IMS) or other gas-phase separation techniques. Additionally, the reduction of specific glycan isomers may result in distinct CCS values, enabling the separation and identification of isomeric glycan species, which is often challenging by other means.

We demonstrate this by applying it to the HMO sample discussed in section 4.3.4 Figure 4.17, where we observed that LNFP II and LNFP III overlap in mobility. One can see in Figure 4.20 that upon reduction we can separate three major peaks to baseline and identify them to be LNFP III, LNFP I, and LNFP II respectively.



(a) Arrival time distribution of reduced LNFP (r-LNFP) found in the Carbosynth HMO sample ($[M+Na]^+$ charge state) obtained; (b) Messenger tagging IR spectra of species found in the ATD region 1 (green) identifying LNFP III , (c) region 2 (red) spectrum identifying LNFP I; (d) region 3 (blue) spectrum identifying LNFP II.

4.5 Conclusion

In this work, we have demonstrated the principle that a library of cryogenic IR spectra can be used for identification of isomeric glycans in mixtures. We emphasize that high-resolution ion mobility is only necessary for obtaining isomerically pure IR fingerprints to build the database. Once the database is available, any moderate resolution IMS device could be used for partial separation and reduction of the number of overlapping species for reliable IR spectra decomposition and isomer identification.

It is worth noting that under moderate separation conditions, it would be almost impossible to perform isomer identification using ion mobility alone or even in combination with fragmentation techniques, since this would require unique fragment peaks for each isomer. In contrast, the linear nature of IR spectra and their high sensitivity to even minor structural variations between isomers make identification by spectral decomposition feasible. The identification process could be accelerated by shortening the IR scanning range, thereby

reducing the IR spectral acquisition time. We also demonstrate how preparative chemistry through reducing the glycans can help in resolving overlapping mobility peaks. However, this requires database spectra of the reduced species to also be measured.

In addition, with an extensive database, machine learning methods could be introduced to identify unique regions of the isomer spectra that would further reduce the scan range and acquisition time. The spectral measurement process can also be multiplexed, which will allow the acquisition of spectra of many species (including isomers) within a single laser scan. This will be demonstrated in the following chapter.

References:

1. Wei, J., Tang, Y., Ridgeway, M. E., Park, M. A., Costello, C. E., and Lin, C. (2020) Accurate Identification of Isomeric Glycans by Trapped Ion Mobility Spectrometry-Electronic Excitation Dissociation Tandem Mass Spectrometry, *Anal Chem 92*, 13211-13220.

2. Morrison, K. A., and Clowers, B. H. (2017) Differential Fragmentation of Mobility-Selected Glycans via Ultraviolet Photodissociation and Ion Mobility-Mass Spectrometry, *J Am Soc Mass Spectrom 28*, 1236-1241.

3. Mank, M., Welsch, P., Heck, A. J. R., and Stahl, B. (2019) Label-free targeted LC-ESI-MS(2) analysis of human milk oligosaccharides (HMOS) and related human milk groups with enhanced structural selectivity, *Anal Bioanal Chem 411*, 231-250.

4. Bansal, P., Yatsyna, V., AbiKhodr, A. H., Warnke, S., Ben Faleh, A., Yalovenko, N., Wysocki, V. H., and Rizzo, T. R. (2020) Using SLIM-Based IMS-IMS Together with Cryogenic Infrared Spectroscopy for Glycan Analysis, *Anal Chem 92*, 9079-9085.

5. Remoroza, C. A., Liang, Y., Mak, T. D., Mirokhin, Y., Sheetlin, S. L., Yang, X., San Andres, J. V., Power, M. L., and Stein, S. E. (2020) Increasing the Coverage of a Mass Spectral Library of Milk Oligosaccharides Using a Hybrid-Search-Based Bootstrapping Method and Milks from a Wide Variety of Mammals, *Anal Chem 92*, 10316-10326.

6. Struwe, W. B., Pagel, K., Benesch, J. L., Harvey, D. J., and Campbell, M. P. (2016) GlycoMob: an ion mobility-mass spectrometry collision cross section database for glycomics, *Glycoconj J 33*, 399-404.

7. Hamid, A. M., Garimella, S. V. B., Ibrahim, Y. M., Deng, L., Zheng, X., Webb, I. K., Anderson, G. A., Prost, S. A., Norheim, R. V., Tolmachev, A. V., Baker, E. S., and Smith, R. D. (2016) Achieving High Resolution Ion Mobility Separations Using Traveling Waves in Compact Multiturn Structures for Lossless Ion Manipulations, *Anal Chem 88*, 8949-8956.

8. Deng, L., Ibrahim, Y. M., Hamid, A. M., Garimella, S. V., Webb, I. K., Zheng, X., Prost, S. A., Sandoval, J. A., Norheim, R. V., Anderson, G. A., Tolmachev, A. V., Baker, E. S., and Smith, R. D. (2016) Ultra-High Resolution Ion Mobility Separations Utilizing

Traveling Waves in a 13 m Serpentine Path Length Structures for Lossless Ion Manipulations Module, *Anal Chem 88*, 8957-8964.

9. 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., and Smith, R. D. (2017) Serpentine Ultralong Path with Extended Routing (SUPER) High Resolution Traveling Wave Ion Mobility-MS using Structures for Lossless Ion Manipulations, *Anal Chem 89*, 4628-4634.

10. Fraschetti, C., Guarcini, L., Zazza, C., Mannina, L., Circi, S., Piccirillo, S., Chiavarino, B., and Filippi, A. (2018) Real time evolution of unprotected protonated galactosamine probed by IRMPD spectroscopy, *Phys Chem Chem Phys 20*, 8737-8743.

11. Warnke, S., Ben Faleh, A., Scutelnic, V., and Rizzo, T. R. (2019) Separation and Identification of Glycan Anomers Using Ultrahigh-Resolution Ion-Mobility Spectrometry and Cryogenic Ion Spectroscopy, *J Am Soc Mass Spectrom 30*, 2204-2211.

12. Ujma, J., Ropartz, D., Giles, K., Richardson, K., Langridge, D., Wildgoose, J., Green, M., and Pringle, S. (2019) Cyclic Ion Mobility Mass Spectrometry Distinguishes Anomers and Open-Ring Forms of Pentasaccharides, *J Am Soc Mass Spectrom 30*, 1028-1037.

13. Nagy, G., Attah, I. K., Garimella, S. V. B., Tang, K., Ibrahim, Y. M., Baker, E. S., and Smith, R. D. (2018) Unraveling the isomeric heterogeneity of glycans: ion mobility separations in structures for lossless ion manipulations, *Chem Commun (Camb) 54*, 11701-11704.

14. Grunwald-Gruber, C., Thader, A., Maresch, D., Dalik, T., and Altmann, F. (2017) Determination of true ratios of different N-glycan structures in electrospray ionization mass spectrometry, *Anal Bioanal Chem 409*, 2519-2530.

15. Kirschbaum, C., Greis, K., Mucha, E., Kain, L., Deng, S., Zappe, A., Gewinner, S., Schollkopf, W., von Helden, G., Meijer, G., Savage, P. B., Marianski, M., Teyton, L., and Pagel, K. (2021) Unravelling the structural complexity of glycolipids with cryogenic infrared spectroscopy, *Nat Commun 12*, 1201.

16. Tang, K., Page, J. S., and Smith, R. D. (2004) Charge competition and the linear dynamic range of detection in electrospray ionization mass spectrometry, *J Am Soc Mass Spectrom 15*, 1416-1423.

Chapter 5 High-throughput multiplexed infrared spectroscopy of ion mobility separated species using Hadamard transform[2]

Coupling vibrational ion spectroscopy with high-resolution ion mobility separation offers a promising approach for detailed analysis of biomolecules in the gas phase. Improvements in the ion mobility technology have made it possible to separate isomers with minor structural differences, and their interrogation with a tunable infrared laser provides vibrational fingerprints for unambiguous database-enabled identification. Nevertheless, wide analytical application of this technique requires high-throughput approaches for acquisition of vibrational spectra of all species present in complex mixtures. In this chapter, we present a novel multiplexed approach and demonstrate its utility for cryogenic ion spectroscopy of peptides and glycans in mixtures. Since the method is based on Hadamard transform multiplexing, it yields infrared spectra with an increased signal-to-noise ratio compared to a conventional signal averaging approach.

5.1 Introduction

Separation and identification of isomers are long-standing problems in analytical chemistry, particularly crucial in fields like glycomics,[1-3] metabolomics,[4, 5] and lipidomics,[6, 7] where high isomeric complexity poses significant analytical challenges. Mass spectrometry (MS) offers rapid analysis with high sensitivity[8] but struggles to distinguish between isomers, so it is often combined with orthogonal separation techniques like liquid chromatography (LC),[2, 4, 9] capillary electrophoresis,[10] or gas chromatography[11]. However, adding an orthogonal separation method often extends the analysis time. Ion mobility spectrometry (IMS)[12-14] can

[2] This chapter is based on the publication "Yatsyna, V., Abikhodr, A. H., Ben Faleh, A., Warnke, S., and Rizzo, T. R., High-Throughput Multiplexed Infrared Spectroscopy of Ion Mobility-Separated Species Using Hadamard Transform." *Anal Chem* **2022,** *94* (6), 2912-2917. The content and figures were adapted with permission from the American Chemical Society.

separate isomers on a millisecond timescale and is easily coupled to MS, showing promise for high-throughput analysis of complex isomeric mixtures.

We recently demonstrated that a combination of cryogenic IR spectroscopy with ultrahigh-resolution IMS offers unambiguous identification of isomers with subtle structural differences.[15-21] While adding a spectroscopic dimension to IMS–MS increases the sample analysis time, we have developed a scheme in which IR fingerprints of multiple, mobility-separated species can be obtained quickly.[22] However, fast spectroscopic analysis of all species in a complex mixture still presents a challenge.

In this chapter, we present a novel multiplexed spectroscopic technique based on Hadamard transform, which can simultaneously acquire the IR spectra of all species present in a mixture in a single laser scan. This approach can be easily implemented on any IMS–MS instrument equipped with ion trapping and laser irradiation and is well-suited for spectroscopic studies of complex mixtures with a broad range of ion mobilities. We demonstrate the method by recording the cryogenic IR spectra of all species separated by SLIM–IMS for several peptide and glycan mixtures, showing that the spectral signal-to-noise ratio (SNR) increases in agreement with the theoretical multiplex advantage.

5.2 Experimental Methods

5.2.1 Sample Preparation

Disaccharide GalNAc-α(1–3)-Gal was purchased from Dextra (UK), whereas Gal-β(1–3)-GalNAc and human milk oligosaccharide samples were purchased from Carbosynth (UK). Bovine serum albumin (BSA) protein digest was purchased from Thermo Fisher Scientific. All samples were used without further purification. For electrospray, the disaccharide samples were dissolved in water/methanol (v/v 50/50) to a final concentration of 5–50 μM, and the BSA protein digest was dissolved in water to yield a 0.5 μM concentration.

5.2.2 Instrumentation

These experiments were performed on our prototype instrument described in chapter 3. Briefly, ions generated by nanoelectrospray are introduced into the instrument through a heated stainless-steel capillary (170 °C) and accumulated in a dual-stage ion funnel assembly (MassTech) used as a trap. Short, intense ion packets (150 μs duration) are released into a compact SLIM ion mobility device (15 × 15 cm) filled with helium at 3 mbar, where molecules are separated based on their shape (i.e., their rotationally averaged CCS).

The cold tagged ions are then irradiated with a tunable continuous-wave mid-IR laser (IPG Photonics) and mass-analyzed using a reflectron-type TOF mass spectrometer. When the laser frequency is resonant with a vibrational transition of the tagged molecule, it absorbs an IR photon, leading to intramolecular vibrational energy redistribution and dissociation of the nitrogen tag(s). Plotting the ratio between the tagged and total ion signals of the same species as a function of the laser wavenumber produces highly resolved IR spectra that can serve as fingerprints for molecular identification.

Arrival time distributions (ATDs) of ions following SLIM-IMS separation are obtained using a channeltron detector after m/z selection by the quadrupole mass filter. Alternatively, two-dimensional IMS-MS profiles covering all m/z values of complex samples can be obtained by scanning a narrow transmission window (e.g., 300 μs) in the arrival time dimension and acquiring mass spectra corresponding to this window using TOF-MS.

5.2.3 Multiplexed spectroscopy approach

Multiplexed spectroscopy based on the Hadamard transform measures the IR response of multiple known combinations of IMS peaks in a single laser scan, allowing for the spectrum of individual IMS peaks to be obtained with increased SNR, known as the multiplex or Fellgett advantage.[23-26] The basics of the multiplexing approach applicable to cryogenic ion spectroscopy combined with IMS is described below.

First, the IMS-MS profile of the sample is measured, and the window covering all IMS peaks of interest is selected and divided into N bins. Multiplexing is performed using a Simplex matrix S(N × N), a cyclic binary matrix derived from Hadamard matrices. The first row of S corresponds to a binary pseudorandom sequence of ones and zeros,[26] while the next rows correspond to a cyclic shift of the previous row one place to the left (Figure 5.1 (a)). Typically, N equals 15, 19, 23, or 31 bins of width chosen so that two IMS peaks with the same m/z do not overlap in one bin.

During multiplexed spectroscopy scans, ions in the IMS bins corresponding to zeros in each row of the S matrix are deflected, while ions corresponding to ones are sent to the cryogenic trap, where they are cooled, tagged, irradiated with the IR laser, and mass-analyzed using TOF-MS. This multiplexed IMS-MS data is then decoded to obtain the mass spectra of individual IMS peaks at each laser wavelength λ, which can be used to obtain their messenger-tagging IR spectra as described above.

Ion deflection corresponding to zeros in the S matrix is performed by increasing the voltage on one of the elements of the ion steering lens and is controlled by the digital output module of a PCIe-6351 card (National Instruments) using a LabView program. Ion deflection in this setup allows for rectangular-shaped pulses with rise and fall times of around 3 μs. If needed, the minimum bin length for multiplexing can be as short as approximately 20 microseconds.

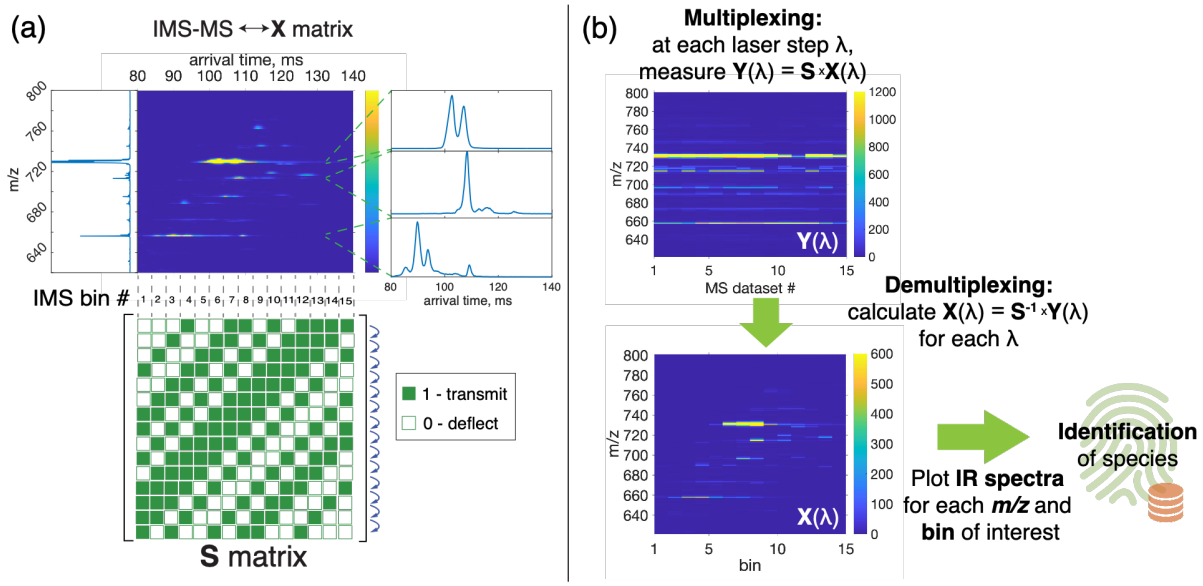


Figure 5.1. (a) An example of a relatively complex IMS-MS profile with multiple species separated using SLIM-IMS (human milk oligosaccharide mixture, total drift length 10 m). The IMS-MS profile can be split into *N* bins and represents the quantity of interest **X** upon multiplexing with the Simplex matrix **S** (*N*×*N*). (b) Multiplexing and demultiplexing procedure, showing the example of multiplexed data **Y** = **S** × **X,** where **S** is a (15×15) Simplex matrix and **X** is the IMS-MS profile from panel (a).

In mathematical terms, multiplexing means that at each laser wavelength λ, we measure the encoded matrix Y(λ) = S × X(λ), where S is an (N × N) Simplex matrix, X(λ) is an (N × k) matrix containing the individual mass spectra of N bins in the chosen ATD window (Figure 5.1 (a)), and k represents the length of the MS data vector. The measured Y(λ) matrices consist of various known combinations of mass spectra from different IMS bins (Figure 5.1 (b)). During data analysis, we multiply the Y(λ) matrices with the inverse of the S matrix to obtain the individual MS traces in each IMS bin [i.e., we compute X(λ) = S–1 × Y(λ) at each λ]. This demultiplexing step is fast and can be performed while scanning the laser wavelength.

In this approach, the theoretical Fellgett advantage (i.e., the increase in the SNR using multiplexed spectroscopy compared to normal spectroscopy of single species) can be estimated as:[25]

$$(G_{SNR})_i = \frac{SNR_{multiplexed}}{SNR_{normal}} = \left(\frac{x_i}{2\bar{x}}\right)^{1/2} \quad (1)$$

where, x_i represents the signal intensity of the m/z channel of interest in the i^{th} IMS bin, and $\bar{x}$ is defined as the average signal across all the IMS bins in the m/z channel of interest. Equation 1 applies to shot-noise limited conditions, which are inherent in TOF–MS measurements. It suggests that an increased SNR will be observed for all IMS bins with an intensity twice higher than the average intensity across all bins for the m/z value of interest. This typically occurs for sparse ATDs with a few IMS peaks in each m/z channel, often obtained using ultrahigh-resolution SLIM–IMS separations of complex mixtures with a broad range of mobilities. Equation 1 also indicates that the SNR of weak IMS peaks will decrease upon multiplexing when strong peaks with the same m/z value are present. Such weak IMS peaks can be identified during the initial ATD inspection and can be analyzed individually if necessary.

It's worth mentioning that Hadamard transform multiplexing has been previously applied in ion mobility instruments to increase their duty cycle when coupled with continuous ion sources such as ESI.[27-31] In this case, multiplexing was achieved by sending multiple identical ion packets for separation by the IMS device in a single experimental cycle, leading to an increased SNR and resolution in the ion mobility dimension. However, this study employs a significantly different approach, as the goal is to obtain IR spectra of ion packets with different ion mobility.

5.3 Results and discussion

We first demonstrate our multiplexed spectroscopy approach by analyzing a mixture of two isomeric disaccharides, GalNAc-α(1–3)-Gal and Gal-β(1–3)-GalNAc, which we separate after twelve SLIM–IMS roundtrips (total drift length of 18 m). The ATD corresponding to $[M + Na]^+$ ions (*m/z* 406) is shown in black in Figure 5.2 (a), where four major peaks can be observed. More specifically, each disaccharide exhibits two ion mobility peaks, which likely originate from reducing-end anomers, as was shown in previous work from our laboratory.[15-17] For multiplexing, the ATD was split into 31 bins of 1.4 ms each, as shown in red in Figure 5.2 (a), and the IMS–MS data was encoded with an $\mathbf{S}(31 \times 31)$ matrix at each laser wavelength step. The peaks of interest that represent the two disaccharides can be found within bins 3, 4, 7, and 8 (Figure 5.2 (a)). The spectra corresponding to these bins are shown in Figures 5.2 (b–e) and were obtained by demultiplexing the encoded IMS–MS data for each laser wavelength

and plotting the ratio between the messenger-tagged and total ion signal of the disaccharide species.

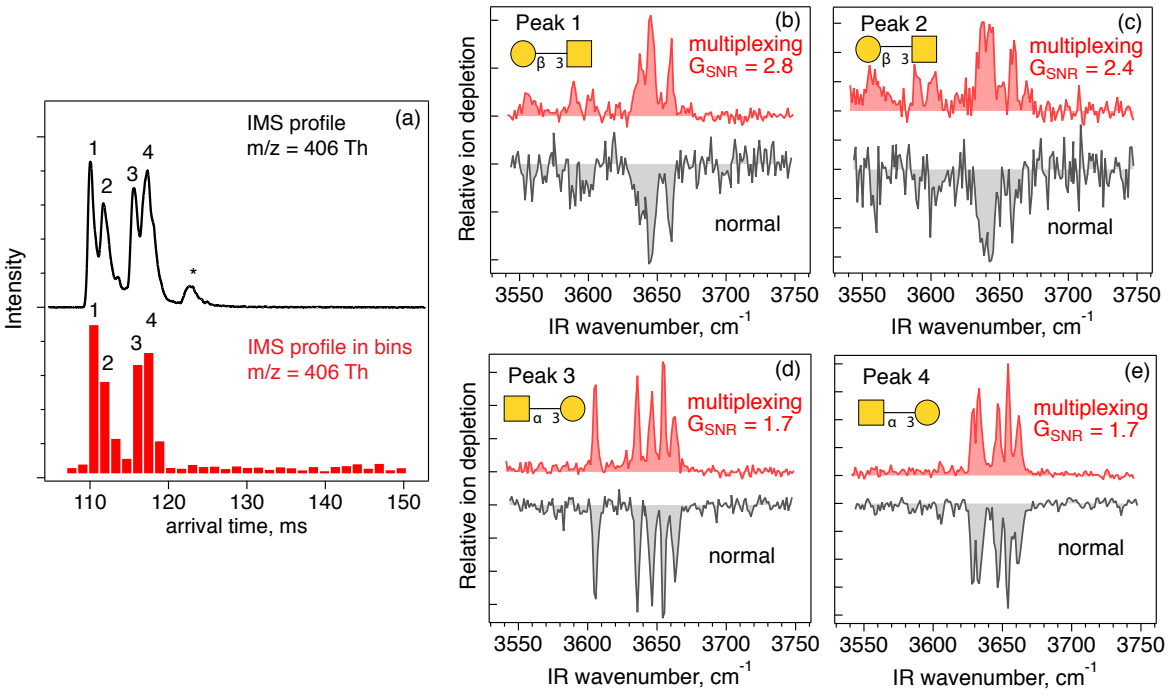


Figure 5.2. (a) The arrival time distribution of a mixture of two disaccharides, GalNAc-α(1-3)-Gal and Gal-β(1-3)-GalNAc, in the [M+Na]+ state, acquired after twelve SLIM-IMS separation cycles (18 m drift length) using the quadrupole mass filter and the channeltron detector, compared to the IMS profile recorded in 31 bins, 1.4 ms/bin using TOF-MS. The peak marked with the asterisk is due to an impurity that was not filtered by the quadrupole. (b-e) Cryogenic IR spectra of individual peaks separated by SLIM-IMS obtained using our multiplexed approach (red) compared with the spectra obtained under normal non-multiplexed measurements (grey). Multiplexed IR spectroscopy was performed using -S(31×31) in the arrival time window of 107-150 ms, which was split into 31 bins of 1.4 ms duration. Non-multiplexed data was acquired directly after the multiplexed laser scan under the same experimental conditions and using the same width and the number of bins.

Infrared spectra in the free OH stretch region obtained in this way were compared to the reference spectra acquired using isomerically pure samples, allowing us to assign peaks 1–2 in the ATD to Gal-β(1–3)-GalNAc and peaks 3–4 to GalNAc-α(1–3)-Gal.

For the purpose of comparing the SNR of multiplexed and normal scans, the grey traces in Figure 5.2 (b–e) show the corresponding laser scans acquired with the same settings but without multiplexing (i.e., **S** equals the identity matrix). One can see that the SNR improves upon multiplexing, allowing us to obtain high quality spectra in a shorter period of time. In order to estimate the SNR gain quantitively, we have analyzed the noise levels of multiplexed

and non-multiplexed measurements by determining the standard deviation of data within the baseline of each spectrum where no depletion peaks occur (3680–3750 cm^{-1}). We find that the observed gain in SNR, G_{SNR}, reaches values between 1.7 and 2.8, depending on the IMS peak of interest (see Figure 5.2 (b–e)), and agrees well with the theoretical estimation using eq 1. On average, an SNR increase of 2.14 is observed, and this corresponds to a reduction in analysis time by a factor of 4.6.

Next, we demonstrate the high-throughput nature of this multiplexed spectroscopy approach by analyzing a BSA protein digest mixture. Figure 5.3 (a) shows the IMS–MS profile after a single-cycle separation using our compact SLIM–IMS device (1.5 m drift length). Even though in this case we do not separate peptide isomers or conformers, rapid separation by SLIM–IMS reduces the sample complexity for spectroscopic analysis. Indeed, under cryogenic conditions (40 K), every peak visible in Figure 5.3 (a) splits into several *m/z* channels due to the nitrogen tagging process, and SLIM–IMS separation helps to eliminate the possible crosstalk between the peptide IR spectra due to the overlapping *m/z* channels under the limited mass resolution of our TOF analyzer. Our multiplexed spectroscopy analysis covered the arrival-time window of 5–36 ms (see Figure 5.3 (a)) that was split into 23 bins of 1.35 ms each for multiplexing using $\mathbf{S}(23 \times 23)$. As a result, we have obtained high-quality cryogenic IR spectra of 37 peptide species in a single 24 min-long laser scan that covered the diagnostic frequency range of 3300–3700 cm^{-1}, probing vibrations corresponding to free and hydrogen-bonded OH/NH stretching. Figure 5.3 (b–e) shows several examples of spectra obtained using multiplexing (red) compared to the non-multiplexed data (grey) obtained under identical experimental conditions (e.g., same arrival time window, bin width, and bin number). The comparison clearly indicates a significant improvement in the SNR thanks to the multiplex advantage. More specifically, we have obtained an average SNR gain across all species, $\langle G_{SNR} \rangle$, equal to 1.8 ± 0.2. This value is slightly lower than the estimated theoretical value of 2.1 ± 0.3 obtained using eq 1, possibly due to slightly decreased messenger-tagging efficiency upon multiplexing due to the presence of much larger ion populations in the cryogenic ion trap compared to non-multiplexed analysis. Nevertheless, the obtained SNR gain of 1.8 implies that without multiplexing one has to average data ∼3 times longer in order to reach the same SNR. The comparison presented in Figure 5.3 (e) also demonstrates that improved SNR can in some cases lead to a better resolved peak pattern, which can provide a more reliable identification of species based on their database spectra. These results clearly demonstrate the power of high-

throughput multiplexed spectroscopy when spectra of all species in a complex mixture need to be acquired with high accuracy in a relatively short time.

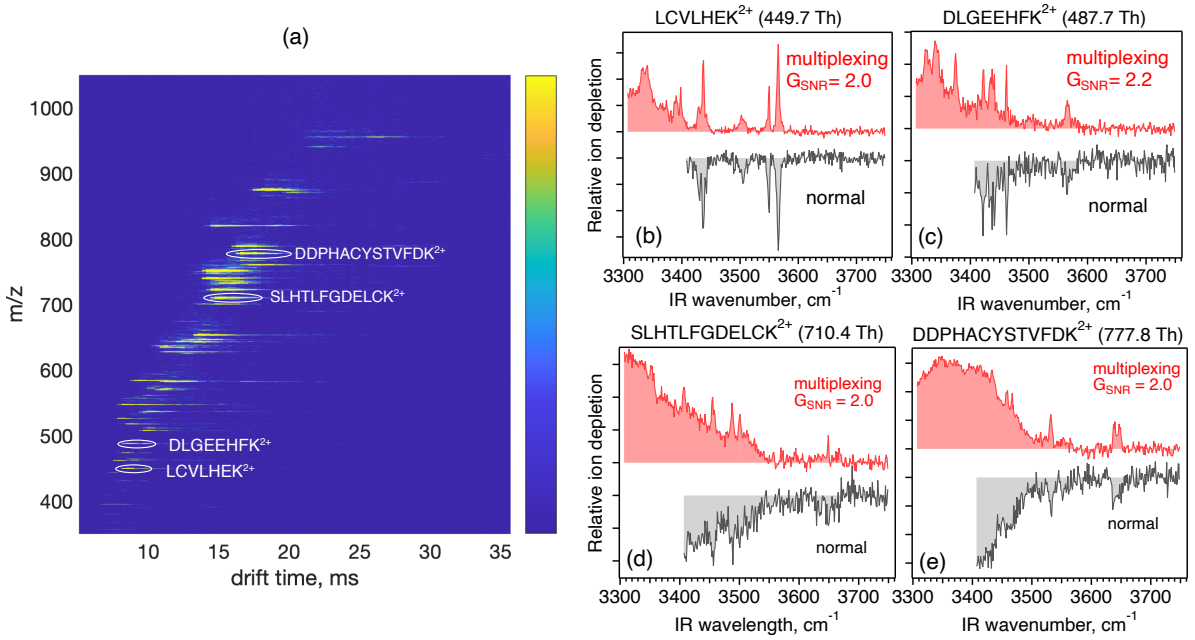


Figure 5.3. (a) IMS–MS profile acquired for the BSA protein digest sample after a single-cycle SLIM–IMS separation (drift length of 1.5 m). (b–e) Examples of peptide cryogenic IR spectra obtained using our multiplexed approach (red) compared with the spectra obtained using a normal, non-multiplexed approach (gray) under the same experimental conditions. Multiplexed IR spectroscopy was performed using **S**(23 × 23) in the arrival time window of 5–36 ms, which was split into 23 bins of 1.35 ms each. Normal (non-multiplexed) data were acquired directly after the multiplexed laser scan using the same arrival time window and number of bins.

5.4 Conclusion

In this work, we have demonstrated a Hadamard transform multiplexing approach that allows measuring the IR spectra of all species present in a complex mixture in a single laser scan. This method is particularly promising for the analysis of complex mixtures with high isomeric complexity and can easily be implemented in various IMS–MS-spectroscopy setups. For example, in addition to its use in cryogenic IR spectroscopy employed in this work, the method can be used with IRMPD and UV spectroscopy. Moreover, one can use this method for multiplexed fragmentation of mobility selected species as well as for IMS^n analysis, which can provide alternative identification means for isomers.[18, 32, 33]

By merging Hadamard transform with our cryogenic multi-trap approach, we substantially increase the throughput of spectroscopic analysis for unambiguous identification of molecular isomers in highly complex mixtures. This was shown in our latest work.

References:

1. Han, L., and Costello, C. E. (2013) Mass spectrometry of glycans, *Biochemistry (Mosc) 78*, 710-720.

2. Gutierrez Reyes, C. D., Jiang, P., Donohoo, K., Atashi, M., and Mechref, Y. S. (2021) Glycomics and glycoproteomics: Approaches to address isomeric separation of glycans and glycopeptides, *J Sep Sci 44*, 403-425.

3. Peng, W., Gutierrez Reyes, C. D., Gautam, S., Yu, A., Cho, B. G., Goli, M., Donohoo, K., Mondello, S., Kobeissy, F., and Mechref, Y. (2023) MS-based glycomics and glycoproteomics methods enabling isomeric characterization, *Mass Spectrom Rev 42*, 577-616.

4. Lu, W., Su, X., Klein, M. S., Lewis, I. A., Fiehn, O., and Rabinowitz, J. D. (2017) Metabolite Measurement: Pitfalls to Avoid and Practices to Follow, *Annu Rev Biochem 86*, 277-304.

5. Opialla, T., Kempa, S., and Pietzke, M. (2020) Towards a More Reliable Identification of Isomeric Metabolites Using Pattern Guided Retention Validation, *Metabolites 10*, 457.

6. Harris, R. A., Leaptrot, K. L., May, J. C., and McLean, J. A. (2019) New Frontiers in Lipidomics Analyses using Structurally Selective Ion Mobility-Mass Spectrometry, *Trends Analyt Chem 116*, 316-323.

7. Porta Siegel, T., Ekroos, K., and Ellis, S. R. (2019) Reshaping Lipid Biochemistry by Pushing Barriers in Structural Lipidomics, *Angew Chem Int Ed Engl 58*, 6492-6501.

8. Smith, R. D., Shen, Y., and Tang, K. (2004) Ultrasensitive and quantitative analyses from combined separations-mass spectrometry for the characterization of proteomes, *Acc Chem Res 37*, 269-278.

9. Nagy, G., Peng, T., and Pohl, N. L. B. (2017) Recent Liquid Chromatographic Approaches and Developments for the Separation and Purification of Carbohydrates, *Anal Methods 9*, 3579-3593.

10. Lu, G., Crihfield, C. L., Gattu, S., Veltri, L. M., and Holland, L. A. (2018) Capillary Electrophoresis Separations of Glycans, *Chem Rev 118*, 7867-7885.

11. Halket, J. M., Waterman, D., Przyborowska, A. M., Patel, R. K., Fraser, P. D., and Bramley, P. M. (2005) Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS, *J Exp Bot 56*, 219-243.

12. Delvaux, A., Rathahao-Paris, E., and Alves, S. (2022) Different ion mobility-mass spectrometry coupling techniques to promote metabolomics, *Mass Spectrom Rev 41*, 695-721.

13. Garcia, X., Sabate, M. D., Aubets, J., Jansat, J. M., and Sentellas, S. (2021) Ion Mobility-Mass Spectrometry for Bioanalysis, *Separations 8*, 33.

14. Chen, Z., Glover, M. S., and Li, L. (2018) Recent advances in ion mobility-mass spectrometry for improved structural characterization of glycans and glycoconjugates, *Curr Opin Chem Biol 42*, 1-8.

15. Warnke, S., Ben Faleh, A., Scutelnic, V., and Rizzo, T. R. (2019) Separation and Identification of Glycan Anomers Using Ultrahigh-Resolution Ion-Mobility Spectrometry and Cryogenic Ion Spectroscopy, *J Am Soc Mass Spectrom 30*, 2204-2211.

16. Ben Faleh, A., Warnke, S., and Rizzo, T. R. (2019) Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures, *Anal Chem 91*, 4876-4882.

17. Warnke, S., Ben Faleh, A., Pellegrinelli, R. P., Yalovenko, N., and Rizzo, T. R. (2019) Combining ultra-high resolution ion mobility spectrometry with cryogenic IR spectroscopy for the study of biomolecular ions, *Faraday Discuss 217*, 114-125.

18. Pellegrinelli, R. P., Yue, L., Carrascosa, E., Warnke, S., Ben Faleh, A., and Rizzo, T. R. (2020) How General Is Anomeric Retention during Collision-Induced Dissociation of Glycans?, *J Am Chem Soc 142*, 5948-5951.

19. Dyukova, I., Ben Faleh, A., Warnke, S., Yalovenko, N., Yatsyna, V., Bansal, P., and Rizzo, T. R. (2021) A new approach for identifying positional isomers of glycans cleaved from monoclonal antibodies, *Analyst 146*, 4789-4795.

20. Yalovenko, N., Yatsyna, V., Bansal, P., AbiKhodr, A. H., and Rizzo, T. R. (2020) Analyzing glycans cleaved from a biotherapeutic protein using ultrahigh-resolution ion mobility spectrometry together with cryogenic ion spectroscopy, *Analyst 145*, 6493-6499.

21. Khanal, N., Masellis, C., Kamrath, M. Z., Clemmer, D. E., and Rizzo, T. R. (2017) Glycosaminoglycan Analysis by Cryogenic Messenger-Tagging IR Spectroscopy Combined with IMS-MS, *Anal Chem 89*, 7601-7606.

22. Warnke, S., Ben Faleh, A., and Rizzo, T. R. (2021) Toward High-Throughput Cryogenic IR Fingerprinting of Mobility-Separated Glycan Isomers, *ACS Meas Sci Au 1*, 157-164.

23. Zare, R. N., Fernandez, F. M., and Kimmel, J. R. (2003) Hadamard transform time-of-flight mass spectrometry: more signal, more of the time, *Angew Chem Int Ed Engl 42*, 30-35.

24. Brock, A., Rodriguez, N., and Zare, R. N. (1998) Hadamard Transform Time-of-Flight Mass Spectrometry, *Analytical Chemistry 70*, 3735-3741.

25. Larson, N. M., Crosmun, R., and Talmi, Y. (1974) Theoretical comparison of singly multiplexed hadamard transform spectrometers and scanning spectrometers, *Appl Opt 13*, 2662-2668.

26. Harwit, M., and Sloane, N. J. A. (1979) An Introduction to Optical Multiplexing Techniques, In *Hadamard Transform Optics* (Harwit, M., and Sloane, N. J. A., Eds.), pp 1-19, Academic Press.

27. Clowers, B. H., Siems, W. F., Hill, H. H., and Massick, S. M. (2006) Hadamard transform ion mobility spectrometry, *Anal Chem 78*, 44-51.

28. Belov, M. E., Buschbach, M. A., Prior, D. C., Tang, K., and Smith, R. D. (2007) Multiplexed ion mobility spectrometry-orthogonal time-of-flight mass spectrometry, *Anal Chem 79*, 2451-2462.

29. Ibrahim, Y. M., Baker, E. S., Danielson, W. F., 3rd, Norheim, R. V., Prior, D. C., Anderson, G. A., Belov, M. E., and Smith, R. D. (2015) Development of a New Ion Mobility (Quadrupole) Time-of-Flight Mass Spectrometer, *Int J Mass Spectrom 377*, 655-662.

30. Clowers, B. H., Cabrera, E., Anderson, G., Deng, L., Moser, K., Van Aken, G., and DeBord, J. D. (2021) Masked Multiplexed Separations to Enhance Duty Cycle for Structures for Lossless Ion Manipulations, *Anal Chem 93*, 5727-5734.

31. Reinecke, T., Naylor, C. N., and Clowers, B. H. (2019) Ion multiplexing: Maximizing throughput and signal to noise ratio for ion mobility spectrometry, *Trac-Trends in Analytical Chemistry 116*, 340-345.

32. Bansal, P., Yatsyna, V., AbiKhodr, A. H., Warnke, S., Ben Faleh, A., Yalovenko, N., Wysocki, V. H., and Rizzo, T. R. (2020) Using SLIM-Based IMS-IMS Together with Cryogenic Infrared Spectroscopy for Glycan Analysis, *Anal Chem 92*, 9079-9085.

33. Peterson, T. L., and Nagy, G. (2021) Toward Sequencing the Human Milk Glycome: High-Resolution Cyclic Ion Mobility Separations of Core Human Milk Oligosaccharide Building Blocks, *Anal Chem 93*, 9397-9407.

Chapter 6 Identification of human milk oligosaccharide positional isomers by combining IMS-CID-IMS and cryogenic IR spectroscopy[3]

High-resolution ion mobility spectrometry (IMS) coupled with cryogenic infrared spectroscopy has proven to be a powerful technique for the identification of oligosaccharides. However, the need for an extensive database, combined with the scarcity of pure standards, remains a significant barrier to the broad application of this approach. To solve this issue, we demonstrate a method in which ion fragments produced by collision-induced dissociation (CID) are separated using IMS and identified using the vibrational fingerprints of only a few standards. Identification of the fragments allows us to determine the structure of the precursor molecule, the vibrational fingerprint of which is then added to our database. We then show how we can use this approach to identify the structure of mobility separated isomers found in pooled human milk.

6.1 Introduction

In our laboratory, we couple high-resolution IMS with cryogenic infrared (IR) spectroscopy to measure IR fingerprints of separated isomers and use these spectra as identifiers *via* a database approach.[1-3] The advantages of using an IR fingerprint for identification are twofold: (1) it is extremely sensitive to the slightest difference in structure and hence unique to a given isomer; and (2) it is an intrinsic property of the molecule, which makes it a robust metric.[4] However, this approach relies on an extensive IR database of analytical standards, which can be scarce and challenging to synthesize.

---

[3] This chapter is based on the publication "Abikhodr, A. H., Ben Faleh, A., Warnke, S., Yatsyna, V., and Rizzo, T. R. (2023) Identification of human milk oligosaccharide positional isomers by combining IMS-CID-IMS and cryogenic IR spectroscopy, *Analyst*." The content and figures were adapted with permission from the Royal Society of Chemistry.

To tackle this issue, we have developed a procedure based on tandem IMS in combination with cryogenic IR fingerprinting[5] to identify glycans for which we do not have standards.[6] In brief, isomeric fragments generated by CID are separated by IMS and identified by their IR fingerprints, using a limited number of standards. Knowledge of the structure of these fragments is then used to reconstruct the primary structure of the precursor molecule. Once the precise isomeric form of the precursor is identified, its IR spectrum can be used to expand the IR database. This approach is especially well suited for human milk oligosaccharides due to the nature of their biosynthesis in which larger species are built off a lactose core by enzymes with high regio- and stereoselectivity. [7-9]

To illustrate the feasibility of this approach, we first apply it to a fucosylated human milk oligosaccharide, where we identify the position and linkage of the fucose residues. While many of these oligosaccharides can be identified by LC-MS/MS, which combines information from the retention time and MS/MS fragmentation data,[10, 11] their identification is column- and technique-specific, which means that the identification process must be repeated if the LC/fragmentation method is changed. In contrast, because a vibrational spectrum is an intrinsic property of the analyte molecule, the identification procedure described above has only to be performed once to acquire the parent IR fingerprint, after which IR spectroscopy alone is sufficient for identification. We then use this method to identify oligosaccharides extracted from a pooled human milk sample for which we had no entries in our infrared spectral database. We demonstrate how we can assign peaks in a complex arrival-time distribution (ATD) to specific isomers.

6.2 Experimental Methods

6.2.1 Sample Preparation

The human milk oligosaccharide standards employed in this work (LNFP I, LNFP II, LNH, LNDFH I, MFLNH I, and DFLNH b) were acquired from Carbosynth Ltd and Dextra Laboratories Ltd. All standard solutions were prepared in a 50/50 mixture of water/methanol and used without further purification at a concentration of 10–25 μM. All samples were analyzed in the sodiated charge state to avoid fucose migration, which occurs in the protonated state.[12]

A pooled human milk sample (0.5 mL lyophilized), purchased from Chemie Brunschwig AG, was reconstituted in 0.5 mL of water. We then centrifuged 50 μL of this sample at 13 000 rpm

for 30 min at 4 °C to de-fat the solution. The resulting aqueous fraction was mixed with ethanol (1 : 1) and held at −80 °C for 1 h, followed by 30 minutes of centrifugation at 13 000 rpm and 4 °C to precipitate the proteins. The supernatant containing HMOs was then diluted 60 times before analysis on our instrument.

6.2.2 Instrumentation

All experiments in this study were conducted on the Next-Gen instrument described in chapter 3. It includes high-resolution ion mobility separation using structures for lossless ion manipulations (SLIM) and a cryogenic trap coupled to a time-of-flight (TOF) mass spectrometer (TOFWERK) for IR spectroscopy. Briefly, ions are produced by nano-electrospray (nESI) and introduced into the instrument through a heated stainless-steel capillary (130° C). Initially, the ions are guided through a set of ion funnels and accumulated in an 2m long accumulation segment on the SLIM device operating with traveling-wave frequency of 9900 Hz with an amplitude of 14 V_{p-p}. Ion packets of 1 ms duration are then released and guided through the SLIM-IMS separation region, which is held at a nitrogen pressure of 2.2 mbar. The SLIM device we designed has a single-pass path length of 10 m and can be operated in cyclic mode to increase the resolution. It also includes a trapping section in which we perform CID on mobility selected ions. While the conditions for optimal separation depend on the *m*/*z* we are analyzing, the traveling-wave frequency is on the order of 19 800 Hz with and amplitude of 42 V_{p-p}.

Following separation, mobility-selected ions are guided through multiple differential pumping stages to the cryogenic trap, which is maintained at 45 K in order to perform messenger-tagging IR spectroscopy. Trapped ions are cooled *via* collisions with a He/N_2 (80 : 20) gas mixture and form weakly bound clusters with N_2 molecules. The N_2-tagged ions are irradiated with a continuous-wave mid-IR laser (IPG Photonics) for 50 ms and analyzed using the TOF mass spectrometer. Upon absorption of an IR photon, energy is redistributed among the vibrational modes, leading to dissociation of the nitrogen tag(s). We obtain an IR fingerprint spectrum of the tagged species by measuring the ratio between the number of tagged and untagged ions at each laser wavenumber step. While we can measure a spectrum in as little as 7 seconds, in this work we typically average two individual 7-minute scans to improve the S/N.

6.2.3 HMO IR fingerprint database

Infrared spectra of LNFP I and II have been previously measured and added to our database (Chapter 4). For the work reported here, we measured IR spectra of three additional standards, LNH, DFLNH b (and its fragments), and MFLNH I. This database allows us to directly detect the presence of these species in complex mixtures as intact molecules or as fragments of larger unknown molecules.

6.2.4 Spectral comparison

In this study, we use the Pearson correlation coefficient (PCC) to compare measured infrared (IR) spectra with those in our database.[13-15] The PCC method measures the linear relationship between two vector variables, and then provides a value ranging from −1 to 1. A coefficient close to 1 indicates a strong positive correlation, while a value close to −1 signifies a strong negative correlation, and a value around 0 implies no correlation. By calculating the PCC for our measured IR spectrum with each of our reference spectra, we can determine the degree of similarity between them and hence the best match. In this work, all identified spectra had a PCC ranging between 0.92 and 0.97, which indicates very good agreement.

6.3 Results and discussion

6.3.1 IR Identification of CID Fragments

As a test to demonstrate the principle of our approach for identifying human milk oligosaccharides, we first investigate lacto-*N*-difucohexaose I (LNDFH I, *m*/*z* 1021). Figure 6.1(a) shows the CID fragment mass spectrum of an LNDFH I standard. The major fragment (*m*/*z* 876) corresponds to the loss of a single fucose, leaving three hexoses, one GlcNAc, and one fucose behind. The ATD of this fragment shows two major peaks (Figure 6.1(b)), each of which we direct into our cryogenic ion trap and measure their respective IR spectrum, shown in Figure 6.1(c and d). The good agreement between the spectrum of each mobility peak with those found in our database allows us to identify the two isomeric fragments as LNFP II and LNFP I.

The presence of the LNFP II fragment reveals the position of a fucose on the *N*-acetyl glucosamine (GlcNAc) *via* an (α1-4) bond, whereas the LNFP I fragment identifies a fucose bound to the terminal galactose (Gal) *via* an (α1-2) bond. Moreover, both fragments share the same backbone. Combining this information allows us to correctly deduce the structure of the precursor molecule, LNDFH 1.

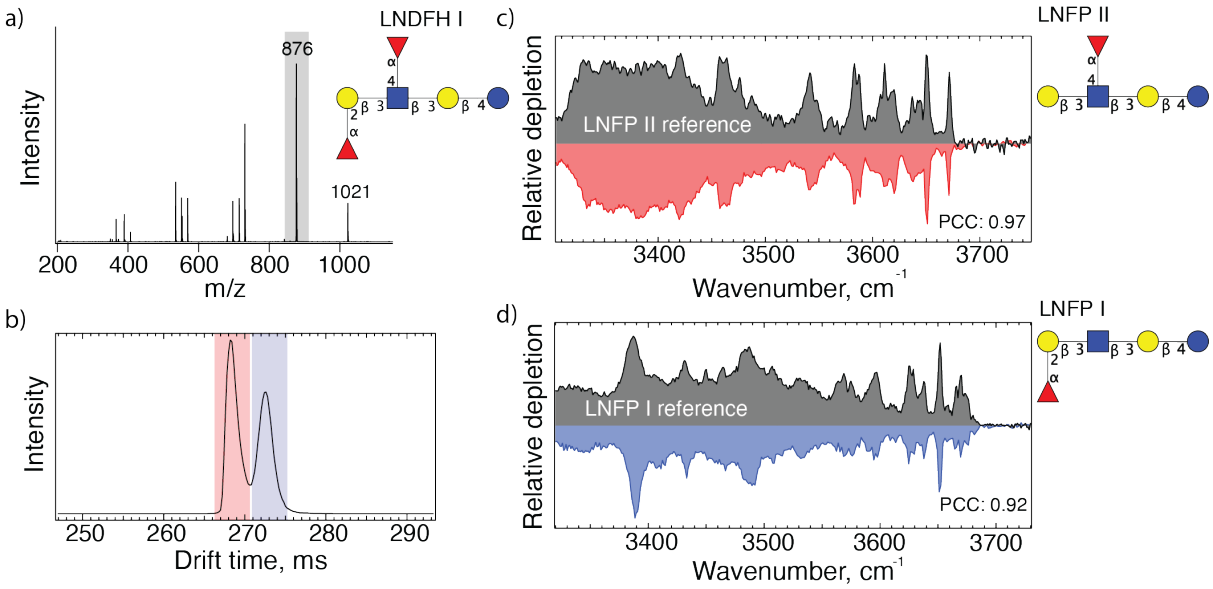


Figure 6.1. (a) MS/MS spectrum of sodiated Lacto-N-difucohexaose I (LNDFH I structure shown); (b) Arrival time distribution of 876 *m/z* fragments after 10 m SLIM-IMS separation; (c,d) messenger- tagging IR spectra corresponding to individual arrival time peaks shaded red and blue in panel (b) compared with database spectra.

6.3.2 Characterization of a selection of oligosaccharides extracted from human milk

Having demonstrated the principle of our fragment-based identification approach on an LNDFH I standard where we already knew the structure, we now show its applicability to a complex pooled human milk sample from which we extracted the glycan content as described above. The arrival time distribution of glycans of *m/z* 1241 after three separation cycles (30 m drift length), shown in Figure 6.2(a), consists of at least six different features, emphasizing the isomeric complexity of HMOs. As shown in Figure 6.2(b), we can directly identify the isomer responsible for the peak arriving at 514 ms *via* spectral comparison with our database to be MFLNH 1, the structure of which is shown as an inset in Figure 6.2(a).

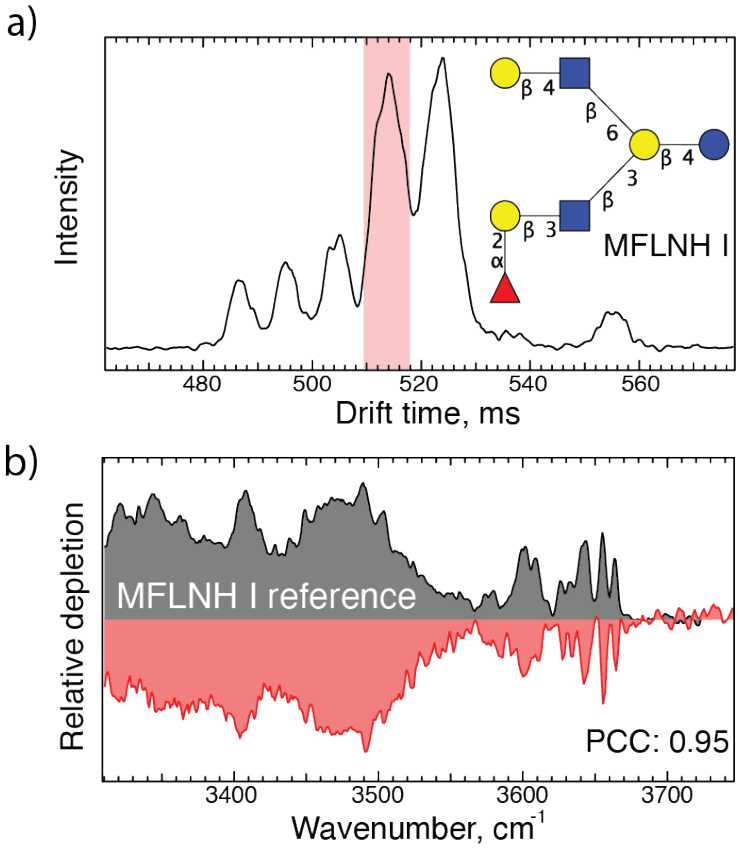


Figure 6.2. (a) Arrival-time distribution of an HMO with *m/z* 1241 from human milk after 30 m SLIM-IMS separation; (b) IR fingerprint of the highlighted peak in panel (a) shown in red, and the corresponding database spectrum of MFLNH I shown in grey.

While this identification was straightforward, the IR fingerprint of the isomer arriving at 525 ms (shown in Figure 6.3(b)) did not match any entry in our database. We therefore selectively performed CID to generate diagnostic fragments, the mass spectrum of which is displayed in Figure 6.3(c). Using our IR database, we can assign the *m/z* 1095 fragment, which corresponds to the loss of one fucose, to the oligosaccharide LNH, shown in Figure 6.3(d). Furthermore, upon mobility separation of the *m/z* 876 fragment, which is composed of three hexoses, one GlcNAc and one fucose, we observe two isomeric structures, as shown in Figure 6.3(e). Based on matches of their IR spectra with entries in our database, shown in Figure 6.3(f), we identify these isomers as the two reducing-end anomers of the glycan Gal(β1-4)[(Fuc(α1-3)]GlcNAc(β1-6)Gal(β1-4)Glc. Putting all this information together, we can assign the structure of the precursor (*m/z* 1241) as having an LNH backbone with a fucose α1-3 linked to the top branch GlcNac, which is MFLNH III. We thus add its IR fingerprint (shown in Figure 6.3(b)), which we assigned without the need for a pure analytical standard for this isomer, to our database.

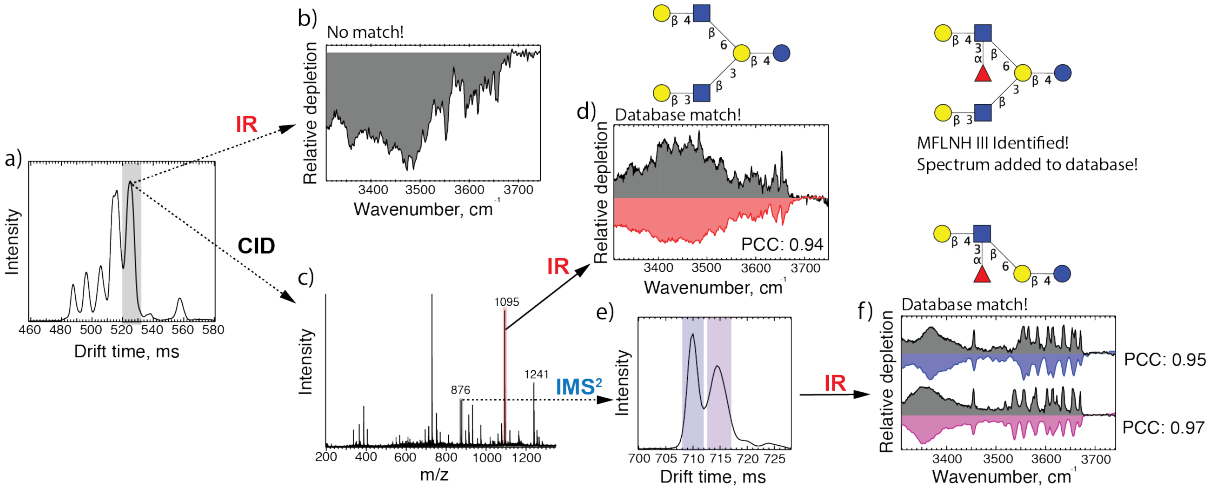


Figure 6.3. (a)ATD of *m/z* 1241 from human milk; (b) IR fingerprint of the highlighted peak in panel (a); (c) fragment mass spectrum of the highlighted peak in panel (a); (d) IR fingerprint of *m/z* 1095; (e) ATD of *m/z* 876 fragment after 10 m drift path; (f) IR fingerprint of both peaks in panel (e) (in color) compared with spectra from our IR database (in grey).

Figure 6.4(a) displays the ATD of another HMO (*m/z* 1387) from the pooled milk sample, which after two cycles (20 m) of separation appears to consist of three major isomers. The IR spectrum of the intense peak arriving at 403 ms (Figure 6.4(b)) did not match those already in our database, and thus required application of our fragmentation scheme. Figure 6.4(c) shows the fragment mass spectrum of the mobility selected peak, and Figure 6.4(d) displays the ATD (one separation cycle) of the *m/z* 876 fragment, where we observe three features. Using cryogenic IR spectroscopy, we can assign the first two peaks in the ATD to the reducing-end anomers of the oligosaccharide Gal(β1-4)[(Fuc(α1-3)]GlcNAc(β1-6)Gal(β1-4)Glc and the third peak as LNFP I (*i.e.*, Fuc(α1-2)Gal(β1-3)GlcNAc(β1-3)Gal(β1-4)Glc), the structures of which are shown in Figure 6.4(e and f), respectively. Both of these oligosaccharides feature a common lactose core (Gal(β1-4)Glc). The one shown in Figure 6.4(f) has Fuc(α1-2)Gal(β1-3)GlcNAc connected to the non-reducing end of lactose *via* a β1-3 bond. This intact oligosaccharide is not found in human milk, but we had generated it as a fragment from a larger HMO and measured its IR spectrum. The structure in Figure 6.4(e) indicates a Gal(β1-4)[(Fuc(α1-3)]GlcNAc branched from lactose *via* a β1-6 bond. Putting these together, we identify the precursor structure to be (Fuc(α1-2)Gal(β1-3)GlcNAc(β1-3)[Gal(β1-4)[(Fuc(α1-3)]GlcNAc(β1-6)]Gal(β1-4)Glc), which is called DFLNH a and shown in the upper right hand corner of Figure 6.4. We thus add its IR fingerprint (Figure 6.4(b)) to our database without having had the need of a pure analytical standard for this species.

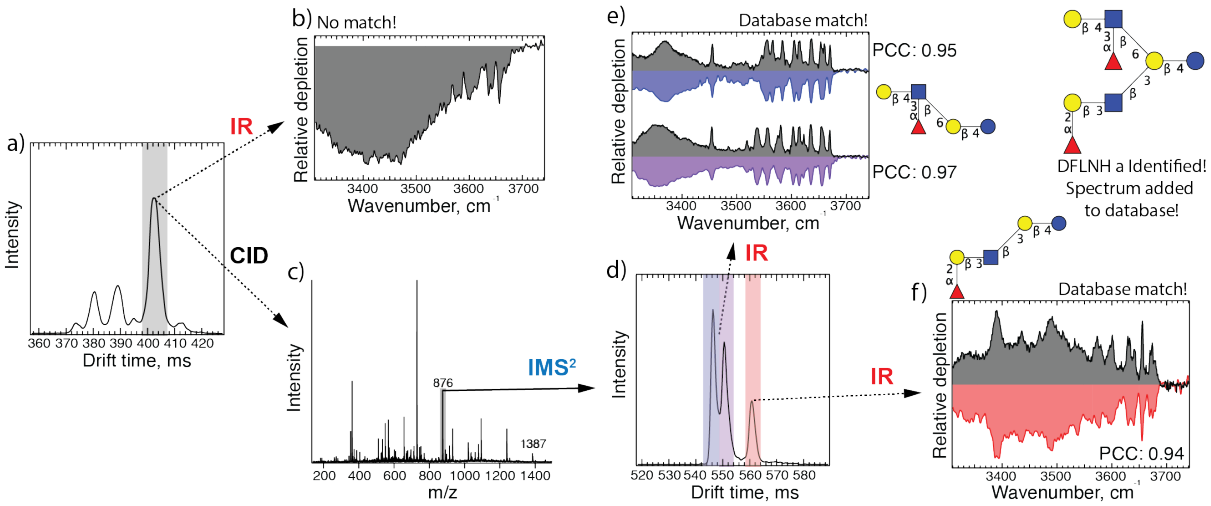


Figure 6.4. (a) Arrival time distribution of the HMO with *m/z* 1387 extracted from human milk after 20 m SLIM-IMS separation; (b) IR fingerprint of the highlighted peak in panel (a); (c) fragment mass spectrum of the highlighted peak in panel (a); (d) ATD of *m/z* 876 fragment after 10 m IMS separation; (e) IR fingerprint of the first two peaks in panel (d) (in color) and their comparison with database spectra (in grey); (f) IR fingerprint of the third peak in panel (e) (red) and its comparison with database spectrum (grey).

6.4 Conclusion

We have combined IMS-CID-IMS with cryogenic IR spectroscopy to identify peaks in the arrival-time distributions of HMO isomers without having had analytical standards for the precursor molecules. We first applied our approach to LNDFH I to verify that we can determine the correct location of fucose residues. We then assigned features in the ATDs of HMOs from a pooled human milk sample with *m/z* 1241 and 1387 as corresponding to the isomers MFLNH III and DFLNH a, respectively, and expanded our IR database with their spectra. Future encounter of these molecules can then be identified by IR spectroscopy alone. The database is also expanded by fragmenting these oligosaccharide structures and acquiring the IR fingerprints of the fragments.

References:

1. Warnke, S., Ben Faleh, A., Pellegrinelli, R. P., Yalovenko, N., and Rizzo, T. R. (2019) Combining ultra-high resolution ion mobility spectrometry with cryogenic IR spectroscopy for the study of biomolecular ions, *Faraday Discuss 217*, 114-125.

2. Ben Faleh, A., Warnke, S., and Rizzo, T. R. (2019) Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures, *Anal Chem 91*, 4876-4882.

3. Warnke, S., Ben Faleh, A., Scutelnic, V., and Rizzo, T. R. (2019) Separation and Identification of Glycan Anomers Using Ultrahigh-Resolution Ion-Mobility Spectrometry and Cryogenic Ion Spectroscopy, *J Am Soc Mass Spectrom 30*, 2204-2211.

4. Pellegrinelli, R. P., Yue, L., Carrascosa, E., Warnke, S., Ben Faleh, A., and Rizzo, T. R. (2020) How General Is Anomeric Retention during Collision-Induced Dissociation of Glycans?, *J Am Chem Soc 142*, 5948-5951.

5. Bansal, P., Yatsyna, V., AbiKhodr, A. H., Warnke, S., Ben Faleh, A., Yalovenko, N., Wysocki, V. H., and Rizzo, T. R. (2020) Using SLIM-Based IMS-IMS Together with Cryogenic Infrared Spectroscopy for Glycan Analysis, *Anal Chem 92*, 9079-9085.

6. Bansal, P., Ben Faleh, A., Warnke, S., and Rizzo, T. R. (2022) Identification of N-glycan positional isomers by combining IMS and vibrational fingerprinting of structurally determinant CID fragments, *Analyst 147*, 704-711.

7. Wopereis, S., Lefeber, D. J., Morava, E., and Wevers, R. A. (2006) Mechanisms in protein O-glycan biosynthesis and clinical and molecular aspects of protein O-glycan biosynthesis defects: a review, *Clin Chem 52*, 574-600.

8. Kellman, B. P., Richelle, A., Yang, J. Y., Chapla, D., Chiang, A. W. T., Najera, J. A., Liang, C., Furst, A., Bao, B., Koga, N., Mohammad, M. A., Bruntse, A. B., Haymond, M. W., Moremen, K. W., Bode, L., and Lewis, N. E. (2022) Elucidating Human Milk Oligosaccharide biosynthetic genes through network-based multi-omics integration, *Nat Commun 13*, 2455.

9. Zeuner, B., Teze, D., Muschiol, J., and Meyer, A. S. (2019) Synthesis of Human Milk Oligosaccharides: Protein Engineering Strategies for Improved Enzymatic Transglycosylation, *Molecules 24*.

10. Remoroza, C. A., Mak, T. D., De Leoz, M. L. A., Mirokhin, Y. A., and Stein, S. E. (2018) Creating a Mass Spectral Reference Library for Oligosaccharides in Human Milk, *Anal Chem 90*, 8977-8988.

11. Ninonuevo, M. R., Park, Y., Yin, H., Zhang, J., Ward, R. E., Clowers, B. H., German, J. B., Freeman, S. L., Killeen, K., Grimm, R., and Lebrilla, C. B. (2006) A strategy for annotating the human milk glycome, *J Agric Food Chem 54*, 7471-7480.

12. Mucha, E., Lettow, M., Marianski, M., Thomas, D. A., Struwe, W. B., Harvey, D. J., Meijer, G., Seeberger, P. H., von Helden, G., and Pagel, K. (2018) Fucose Migration in Intact Protonated Glycan Ions: A Universal Phenomenon in Mass Spectrometry, *Angew Chem Int Ed Engl 57*, 7440-7443.

13. Zapata, F., and Garcia-Ruiz, C. (2018) The discrimination of 72 nitrate, chlorate and perchlorate salts using IR and Raman spectroscopy, *Spectrochim Acta A Mol Biomol Spectrosc 189*, 535-542.

14. Vrancic, C., and Petrich, W. (2011) Effective fragment potential study of the influence of hydration on the vibrational spectrum of glucose, *J Phys Chem A 115*, 12373-12379.

15. Tan, X. F., Chen, X. L., and Song, S. Z. (2017) A computational study of spectral matching algorithms for identifying Raman spectra of polycyclic aromatic hydrocarbons, *J Raman Spectrosc 48*, 113-118.

Chapter 7 Combining liquid chromatography and cryogenic IR spectroscopy in real time for the analysis of oligosaccharides[4]

While the combination of liquid chromatography (LC) and mass spectrometry (MS) serves as a robust approach for oligosaccharide analysis, it has difficulty distinguishing the smallest differences between isomers. The integration of infrared (IR) spectroscopy as an additional analytical dimension can effectively address this limitation by providing a molecular fingerprint that is unique to each isomer. However, the direct interfacing of LC-MS with IR spectroscopy presents a technical challenge arising from the mismatch in the operational timescale of each method. In previous studies, this temporal incompatibility was mitigated by employing strategies designed to slow down or broaden the LC elution peaks of interest, but this workaround is only applicable for a few species at a time, necessitating multiple LC runs for comprehensive analysis. In the current work, we directly couple LC with IR spectroscopy by acquiring a spectrum in as little as 10 seconds. We have successfully demonstrated this approach on a commercially available human milk oligosaccharide product, acquiring spectral information of the eluting peaks in real time, thereby identifying both the specified constituents and non-specified product impurities.

7.1 Introduction

Traditional mass spectrometry-based approaches typically rely on a prior isomer separation step using liquid chromatography (LC),[1, 2] ion mobility spectrometry (IMS),[3, 4] or capillary electrophoresis (CE).[5, 6] While the coupling of these techniques allows for the separation of the

[4] This chapter is based on the recently submitted publication "Abikhodr, A. H., Warnke, S., Ben Faleh, A., and Rizzo, T. R., Combining liquid chromatography and cryogenic IR spectroscopy in real time for the analysis of oligosaccharides." *Anal Chem*. The content and figures were adapted with permission from the American Chemical Society.

components of complex mixtures, the challenge is in assigning their precise isomeric forms. The currently preferred strategy for analysis and identification often integrates these separation techniques with tandem-MS, which relies on the observation of isomer-specific fragments.[7-11] However, the capacity to discriminate between subtly different isomers remains a challenge, because their fragmentation patterns often exhibit only small changes in relative intensities. Such differences are often not universal and can depend on instrument conditions and platforms.

To address this issue, a number of groups[12-16] have explored the addition of a new analytical dimension – that of an infrared (IR) spectrum, which provides a unique molecular fingerprint for each molecule that can distinguish isomers. Because of the richness of information contained in an IR spectrum, particularly when performed at cryogenic temperatures, an analyte can be confidently identified when matched with a database spectrum without the need of database entries for all plausible isomeric forms. We have previously demonstrated the combination of cryo-IR spectroscopy with IMS-MS.[17, 18] Because of the ubiquity of LC workflows and their importance for minimizing ion suppression affects in complex mixtures, the combination of LC with IR spectroscopy would be an important addition to the oligosaccharide analysis toolbox. However, the integration of IR spectroscopy with LC-MS has been fraught with technical challenges, predominantly due to the disparity in operational timescales between the methods. In previous implementations, recording an IR spectrum required tens of minutes, which is much greater than the widths of peaks eluting from a liquid chromatograph. To deal with this temporal incompatibility, past methods have either extended the elution time of LC peaks of interest or used online fractionation and re-injection at lower flow rates.[19, 20] While the latter approach is promising, it is limited in its applicability, as it necessitates multiple LC runs for a comprehensive analysis, which can be impractical for large sample sets.

Here we propose an approach that directly couples LC-MS with IR spectroscopy by reducing IR spectral data acquisition time to as little as 10 seconds, which allows obtaining the valuable molecular information while maintaining the high throughput of LC-MS. We demonstrate this technique by applying it to a commercially available nutritional supplement for children that is specified to contain five human milk oligosaccharides. Within a single LC run, we have used IR spectra to identify the five expected commercial glycans in addition to impurities not specified by the supplier.

7.2 Experimental Methods

7.2.1 Sample preparation

A commercial nutritional supplement containing five human milk oligosaccharides (2'-FL, LNT, LNnT, 3SL, and 6SL) for children age 1+ was dissolved in water as instructed on the packaging. All samples were analyzed in the sodiated charge state to avoid fucose migration, which occurs in protonated species.[21]

Prior to injection into the LC system, we diluted the sample 60-fold. Subsequently, the oligosaccharides were extracted using a two-step process involving C18 and porous graphitic carbon (PGC) solid phase exchange. For the C18 step, the cartridge was activated and equilibrated sequentially with 1 ml of methanol and water. Following activation, 1 ml of the sample was loaded onto the cartridge and the flow through was collected. Subsequently, the cartridge was subjected to three washes using 0.5 ml of water, with each wash fraction combined with the previously collected eluent. Moving on to the PGC solid phase exchange, the PGC cartridge was activated and equilibrated sequentially with 1 ml of 60% acetonitrile/water and water. The sample was then loaded onto the cartridge and washed three times with 0.5 ml of water. Following the washing steps, the glycans were released from the cartridge using 0.3 ml of a solution consisting of 50% acetonitrile, and this process was repeated three times. Finally, 50 µl of the resulting sample was directly injected into the LC system for analysis.

7.2.2 Instrumentation

All experiments were performed using an Acquity Premier UPLC system (Waters Corp.) coupled to our homebuilt instrument, shown in Figure 7.1.

The PGC UPLC separation was conducted using a Hypercarb™ column (150 × 0.5 mm, 5 µm, Thermo Scientific) at a flow rate of 600 µl/min. The sample was eluted from the column using a binary solvent system composed of Solvent A (0.1% formic acid in water) and Solvent B (0.1% formic acid in acetonitrile) at a temperature of 70°C. For separation of the glycan mixtures, an optimized gradient elution method was employed. The elution gradient proceeded as follows: a gradient of 100% to 40% of Solvent A was used for the first 60 minutes. Subsequently, the composition was gradually changed to 100% Solvent A over the next 10 minutes.

Our homebuilt instrument includes a cryogenic ion trap for performing infrared (IR) messenger-tagging spectroscopy which is connected to a time-of-flight (TOF) mass spectrometer (TOFWERK). The process begins with the molecules eluting from the PGC column, after which a separate flow of sodium acetate (10 μM) is added before the generation of ions through electrospray ionization (ESI) to facilitate the formation of sodiated species. They are then introduced into the instrument *via* a heated stainless-steel capillary maintained at 150°C. Initially, the ions are directed through a series of ion funnels and accumulated. Afterward, packets of ions with a duration of 1 ms are released and guided through multiple stages of differential pumping until they reach the cryogenic trap, which is maintained at a temperature of 45 K. The trapped ions are cooled by undergoing collisions with a mixture of helium and nitrogen in an 80:20 ratio, forming weakly bound clusters with N_2. A continuous-wave mid-IR laser (IPG Photonics) is then employed to irradiate the N_2-tagged ions for a duration of 50 ms, at which point they are released and analyzed using the TOF mass spectrometer. Upon absorption of an IR photon, the energy is redistributed among the vibrational modes of the ions, leading to the dissociation of the nitrogen tag(s). By measuring the ratio between tagged and untagged ions at each laser wavenumber step, an IR fingerprint spectrum of the tagged species is obtained. A spectrum over the range of 3500-3700 cm^{-1} can be recorded in approximately 10 seconds.

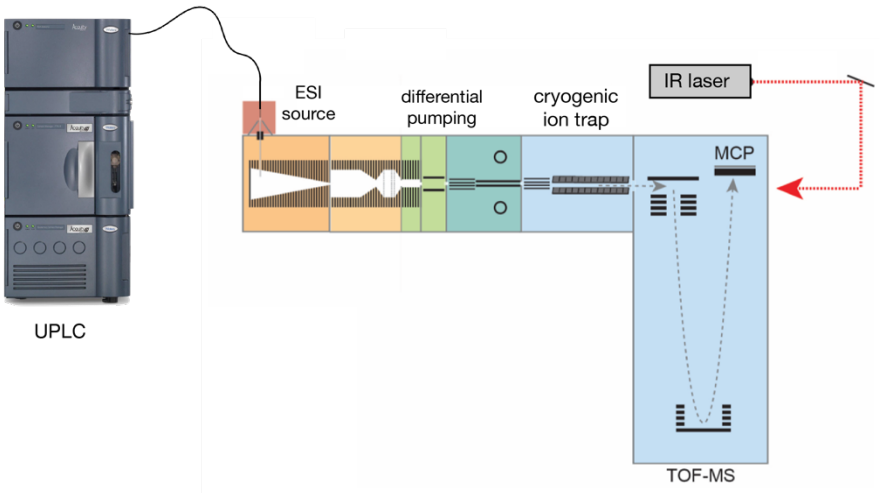


Figure 7.1. Schematic of the experimental setup.

7.2.3 Spectral comparison with an HMO IR fingerprint database

We have previously measured infrared spectra of over 35 human milk oligosaccharides (HMOs), forming a database that allows us to directly detect the presence of these species in complex mixtures as intact molecules or as fragments of larger unknown molecules.

We employ the Pearson correlation coefficient (PCC) to assess the similarity between the measured infrared (IR) spectra and those contained in our database. The PCC is a statistical method utilized to evaluate the linear relationship between two vector variables, and it yields a value within the range of -1 to 1. A coefficient approaching 1 indicates a robust positive correlation, while a value close to -1 suggests a strong negative correlation. A coefficient approximately equal to 0 suggests no substantial correlation between the variables. To determine the extent of similarity between our measured IR spectrum and each of the reference spectra within our database, we compute the PCC for each pair, which enables us to identify the most suitable match.

7.3 Results and Discussion

Identifying the main components

A chromatogram of the major elution peaks of the commercial HMO sample is shown in Figure 7.2. From the mass information one can determine the monosaccharide content of the major elution peaks (see figure caption for notation), and based on the specified product constituents, one can assume that elution peak 2 is 2'-FL, peaks 4 and 5 are some combination of LNT/LNnT, and peaks 6 and 7 are that of 3'-SL and 6'-SL. Other oligosaccharides which are not mentioned on the label are also present.

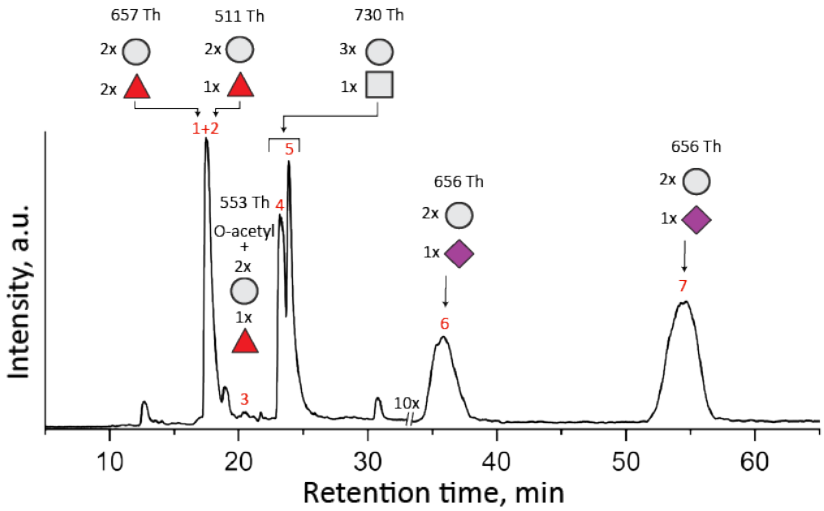


Figure 7.2. Total ion count (TIC) chromatogram of glycans eluting after PGC separation of the commercial HMO mixture. Above each peak of interest is the mass to charge ratio and the general monosaccharide makeup, with hexose units denoted by circles, deoxyhexoses shown as triangles, HexNAcs depicted as squares, and sialic acids are denoted as diamonds.

To assign the retention peaks unambiguously, we acquired an IR spectrum of each of the major peaks. From comparison with the database spectra shown in gray, the elution peaks 2, 5, 6, and 7 can be identified as 2'-FL, LNT, 3'-SL, and 6'-SL respectively, with their measured spectra shown in red. Here four of the five specified oligosaccharides were identified.

While visual inspection would be sufficient to assign these spectra correctly, determining the Pearson correlation coefficient provides an objective quantitative measure of the assignment. If, for example, we were to take the spectrum of 3'-SL and compare it to the database spectrum for 6'-SL, it would give a PCC of 0.62, considerably lower than 0.8 for the assignment of Figure 7.3. One expects a certain degree of correlation between spectra of isomeric HMOs because they possess the same functional groups, in this case differing only by the orientation of the terminal sialic acid. However, if one compares the spectrum of 2'-FL with our database spectrum of 3'-FL, which differ by the attachment site of the fucose residue, it gives a PCC of 0.22. The calculated PCC thus seems to be a robust measure of the spectral assignment.

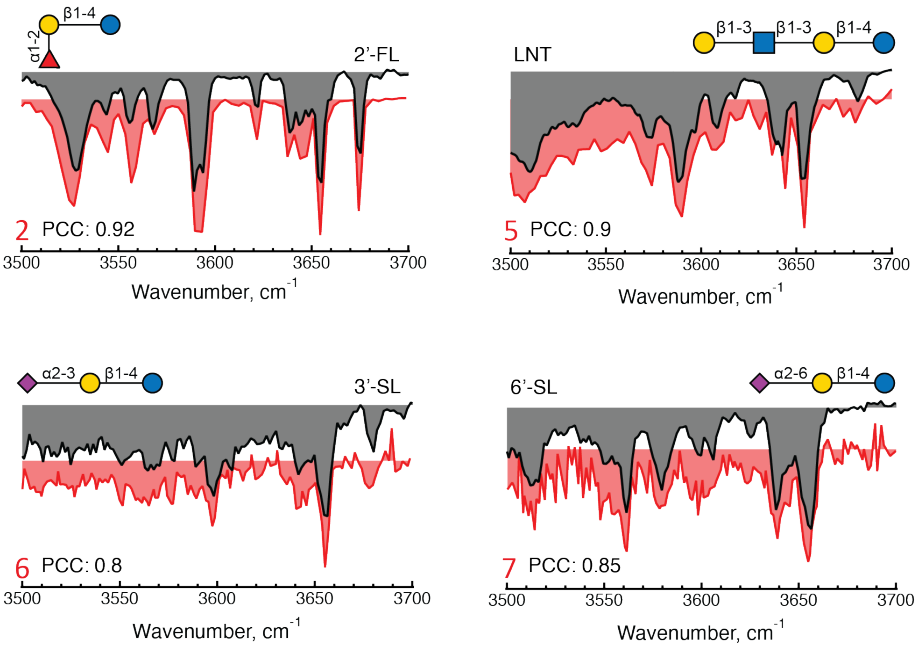


Figure 7.3. Messenger tagging IR depletion spectra of elution peaks 2, 5, 6, and 7 (in red) compared with the best identified fit.

As for the final specified component, LNnT, peak 4 exhibits the correct mass, but as shown in Figure 7.4(a) it is broader than peak 5, which was identified as LNT. We thus measured two

IR spectra for this elution peak, one at the very beginning, highlighted in red, and one towards the end, highlighted in green. The two measured spectra shown in Figure 7.4(b) correspond to the α and β reducing-end anomers of the molecule LNnT, as suggested in previous work.[12, 22] Note that comparing the spectrum of anomer 1 with the database spectrum of anomer 2 would yield a PCC of 0.37 rather than 0.75 for the assigned spectrum. This demonstrates the extreme sensitivity of the infrared spectrum to primary structure.

The spectroscopic match of these LNnT isomers further demonstrates that: 1) the timescale for the mutarotation reaction between the reducing end anomers is greater than the elution time; and 2) a PGC column has sufficient resolution to separate these isomers, which differ only by the orientation of the reducing-end hydroxy group. While this could be anticipated from our previous anomer-resolved infrared spectroscopy measurements,[23] the ability to separate such subtly distinct isomers by liquid chromatograph is noteworthy.

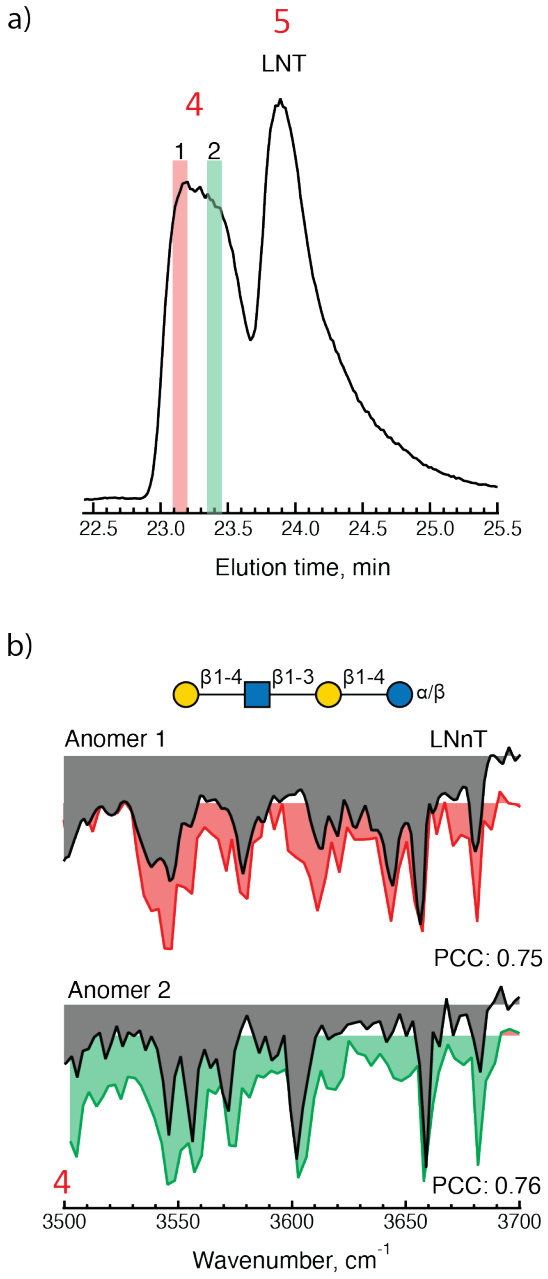


Figure 7.4. (a) Elution peak 4 on an expanded scale, highlighting the regions where a messenger tagging IR spectra were taken (red and green). (b) Messenger tagging IR depletion spectra of highlighted regions compared with the best identified fit using PCC identifying the constituent of each peak.

In addition to the HMOs that were specified components of the commercial product, two additional molecules were identified. As shown in Figure 7.5, we can identify peaks 1 and 3 of the chromatogram of Figure 7.2 as difucosyllactose (DFL) and O-acetylated 2'-fucosyllactose, with the acetylation being on the sixth position of the glucose ring. While DFL is a human milk oligosaccharide, 6-O-acetylated 2'-FL is not found in human milk.

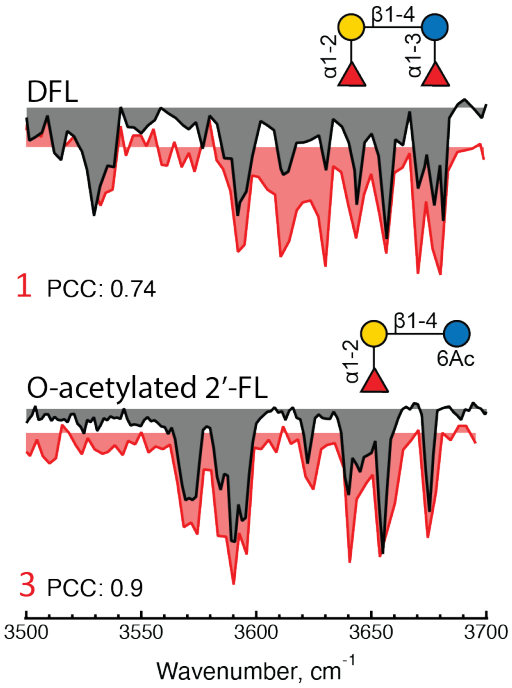


Figure 7.5. Messenger tagging IR depletion spectra of elution peaks 1, and 3 (in red) compared with the best identified fit using PCC identifying the constituent of each peak.

Regarding the remaining small features observed in the chromatogram of Figure 7.2, no corresponding matches were found within our database. These peaks could, in principle, be identified using a CID-IR method as previously demonstrated as shown in the previous chapter.[18, 24, 25]

7.4 Conclusion

We have combined liquid chromatography with cryogenic infrared spectroscopy to facilitate the analysis of biomolecular structures. In this study, we demonstrate the utility of this approach through the analysis of a commercially available human milk oligosaccharide product. The identification of its constituent species was accomplished within a single liquid chromatography run, using real-time cryogenic IR spectroscopy together our expanding infrared spectroscopic database.

The efficiency of this analytical procedure was significantly enhanced by eliminating the need for calibration of elution times *via* external standards. Moreover, we leveraged an automatic method that assigns the experimentally determined spectra by comparing them with all recorded spectra within our database to determine the most suitable match.

References:

1. Bones, J., Mittermayr, S., O'Donoghue, N., Guttman, A., and Rudd, P. M. (2010) Ultra performance liquid chromatographic profiling of serum N-glycans for fast and efficient identification of cancer associated alterations in glycosylation, *Anal Chem 82*, 10208-10215.
2. Kalay, H., Ambrosini, M., van Berkel, P. H., Parren, P. W., van Kooyk, Y., and Garcia Vallejo, J. J. (2012) Online nanoliquid chromatography-mass spectrometry and nanofluorescence detection for high-resolution quantitative N-glycan analysis, *Anal Biochem 423*, 153-162.
3. Fenn, L. S., and McLean, J. A. (2011) Structural resolution of carbohydrate positional and structural isomers based on gas-phase ion mobility-mass spectrometry, *Phys Chem Chem Phys 13*, 2196-2205.
4. Fenn, L. S., and McLean, J. A. (2013) Structural separations by ion mobility-MS for glycomics and glycoproteomics, *Methods Mol Biol 951*, 171-194.
5. Sarkozy, D., and Guttman, A. (2022) CEandCE–MSApproaches for Glycan Analysis, In *Capillary Electrophoresis-Mass Spectrometry for Proteomics and Metabolomics*, pp 313-334.
6. Lageveen-Kammeijer, G. S. M., de Haan, N., Mohaupt, P., Wagt, S., Filius, M., Nouta, J., Falck, D., and Wuhrer, M. (2019) Highly sensitive CE-ESI-MS analysis of N-glycans from complex biological samples, *Nat Commun 10*, 2137.
7. Ashline, D., Singh, S., Hanneman, A., and Reinhold, V. (2005) Congruent strategies for carbohydrate sequencing. 1. Mining structural details by MSn, *Anal Chem 77*, 6250-6262.
8. Veillon, L., Huang, Y., Peng, W., Dong, X., Cho, B. G., and Mechref, Y. (2017) Characterization of isomeric glycan structures by LC-MS/MS, *Electrophoresis 38*, 2100-2114.
9. Zhou, S., Veillon, L., Dong, X., Huang, Y., and Mechref, Y. (2017) Direct comparison of derivatization strategies for LC-MS/MS analysis of N-glycans, *Analyst 142*, 4446-4455.
10. Leymarie, N., and Zaia, J. (2012) Effective use of mass spectrometry for glycan and glycopeptide structural analysis, *Anal Chem 84*, 3040-3048.
11. Xu, G., Goonatilleke, E., Wongkham, S., and Lebrilla, C. B. (2020) Deep Structural Analysis and Quantitation of O-Linked Glycans on Cell Membrane Reveal High

Abundances and Distinct Glycomic Profiles Associated with Cell Type and Stages of Differentiation, *Anal Chem 92*, 3758-3768.

12. Bansal, P., Ben Faleh, A., Warnke, S., and Rizzo, T. R. (2023) Multistage Ion Mobility Spectrometry Combined with Infrared Spectroscopy for Glycan Analysis, *J Am Soc Mass Spectrom 34*, 695-700.
13. Greis, K., Kirschbaum, C., von Helden, G., and Pagel, K. (2022) Gas-phase infrared spectroscopy of glycans and glycoconjugates, *Curr Opin Struct Biol 72*, 194-202.
14. Manz, C., and Pagel, K. (2018) Glycan analysis by ion mobility-mass spectrometry and gas-phase spectroscopy, *Curr Opin Chem Biol 42*, 16-24.
15. Moge, B., Yeni, O., Infantino, A., and Compagnon, I. (2023) CO2 laser enhanced rapid IRMPD spectroscopy for glycan analysis, *International Journal of Mass Spectrometry 490*, 117071.
16. Ho, J. S., Gharbi, A., Schindler, B., Yeni, O., Bredy, R., Legentil, L., Ferrieres, V., Kiessling, L. L., and Compagnon, I. (2021) Distinguishing Galactoside Isomers with Mass Spectrometry and Gas-Phase Infrared Spectroscopy, *J Am Chem Soc 143*, 10509-10513.
17. Ben Faleh, A., Warnke, S., and Rizzo, T. R. (2019) Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures, *Anal Chem 91*, 4876-4882.
18. Abikhodr, A. H., Ben Faleh, A., Warnke, S., Yatsyna, V., and Rizzo, T. R. (2023) Identification of human milk oligosaccharide positional isomers by combining IMS-CID-IMS and cryogenic IR spectroscopy, *Analyst*.
19. Schindler, B., Laloy-Borgna, G., Barnes, L., Allouche, A. R., Bouju, E., Dugas, V., Demesmay, C., and Compagnon, I. (2018) Online Separation and Identification of Isomers Using Infrared Multiple Photon Dissociation Ion Spectroscopy Coupled to Liquid Chromatography: Application to the Analysis of Disaccharides Regio-Isomers and Monosaccharide Anomers, *Anal Chem 90*, 11741-11745.
20. Martens, J., Koppen, V., Berden, G., Cuyckens, F., and Oomens, J. (2017) Combined Liquid Chromatography-Infrared Ion Spectroscopy for Identification of Regioisomeric Drug Metabolites, *Anal Chem 89*, 4359-4362.
21. Mucha, E., Lettow, M., Marianski, M., Thomas, D. A., Struwe, W. B., Harvey, D. J., Meijer, G., Seeberger, P. H., von Helden, G., and Pagel, K. (2018) Fucose Migration in Intact Protonated Glycan Ions: A Universal Phenomenon in Mass Spectrometry, *Angew Chem Int Ed Engl 57*, 7440-7443.

22. Warnke, S., Ben Faleh, A., and Rizzo, T. R. (2021) Toward High-Throughput Cryogenic IR Fingerprinting of Mobility-Separated Glycan Isomers, *ACS Meas Sci Au 1*, 157-164.
23. Pellegrinelli, R. P., Yue, L., Carrascosa, E., Ben Faleh, A., Warnke, S., Bansal, P., and Rizzo, T. R. (2022) A New Strategy Coupling Ion-Mobility-Selective CID and Cryogenic IR Spectroscopy to Identify Glycan Anomers, *J Am Soc Mass Spectrom 33*, 859-864.
24. Bansal, P., Ben Faleh, A., Warnke, S., and Rizzo, T. R. (2022) Identification of N-glycan positional isomers by combining IMS and vibrational fingerprinting of structurally determinant CID fragments, *Analyst 147*, 704-711.
25. Ben Faleh, A., Warnke, S., Bansal, P., Pellegrinelli, R. P., Dyukova, I., and Rizzo, T. R. (2022) Identification of Mobility-Resolved N-Glycan Isomers, *Anal Chem 94*, 10101-10108.

Chapter 8 Summary and future perspective

The combination of ion mobility spectrometry with mass spectrometry has proven to be a powerful tool in glycan analysis. Adding a spectroscopic dimension to the analysis allows to generate isomer-specific IR fingerprints that can be used to build a database to unambiguously identify glycan isomers present in given mixture. In this work we aimed at increasing speed and throughput of this technique, in addition to increasing compatibility of gas-phase IR spectroscopy with liquid chromatography.

This thesis first explores the effective identification of isomeric glycans within complex mixtures using a library of cryogenic IR spectra. High-resolution ion mobility is essential for creating a pure IR fingerprint database, but once established, moderate resolution IMS devices can aid in partial separation, reducing overlapping species for reliable IR spectral decomposition and isomer identification. The linear nature of IR spectra obtained by the messenger tagging technique and their high sensitivity to structural variations enable feasible identification. Shortening the IR scanning range can accelerate the process, and machine learning techniques could automate it.

Next, the Hadamard-transform multiplexing approach presented here is particularly promising for analyzing complex mixtures with high isomeric complexity and can be implemented in various IMS–MS-spectroscopy setups, including IRMPD and UV spectroscopy. It enhances throughput for identifying molecular isomers in complex mixtures. Combining IMS-CID-IMS with cryogenic IR spectroscopy allows the identification of isomers even without analytical standards for precursor molecules. The database is expanded by fragmenting these structures and acquiring IR fingerprints of the fragments.

Finally, the combination of liquid chromatography with cryogenic infrared spectroscopy streamlines biomolecular structure analysis, exemplified by studying a human milk oligosaccharide product. With our ability to measure a spectrum in approximately 10 seconds, identification was accomplished within a single liquid chromatography run *via* an automatic matching method based on our expanding IR spectroscopic database.

Future directions of this research could encompass multiple aspects. Initially, expanding the infrared spectral database will be necessary to make this a generally useful analytical tool. Although our fragmentation identification approach, described in Chapter 6, is robust and

shows potential, this approach assumes the availability of the building block molecules within our database, and the gas-phase fragmentation techniques' ability to yield those building blocks found in our database.

In this work we have used two methods for the automated identification of molecules *via* their IR spectra—Root Mean Square Deviation (RMSD) and the Pearson Correlation Coefficient. While these methods have been applied "manually" by a user, the integration of machine learning algorithms and user-friendly interfaces could completely automate this process. This will be not too dissimilar from comparing retention times and mass-to-charge ratios to an established database as commercial LC-MS software currently does, however as an intrinsic fingerprint of the molecule, the IR spectrum provides a more robust identifier, which is insensitive to experimental conditions.

Similarly, our fully online LC-MS-IR analysis currently requires manual intervention, as user input is required to start the spectrum acquiring once a peak of interest arrives. This could be incorporated into existing LC-MS software that can monitor signal intensity, and once a certain threshold is reached, a command for scanning the laser is issued. This will allow for the integration of this technique into routine LC-MS work flows.

Lastly, the spectroscopic range used within this thesis primarily spans the mid-IR domain, extending from 3200 cm^{-1} to 3800 cm^{-1}. This constraint leaves us incapable of identifying molecules that have no vibrational bands in this region. Extending our wavelength region to cover 5-10 μm (1000-2000 cm^{-1}) would enable the application of our technique to a wide array of metabolite molecule classes.

This thesis underscores the feasibility of utilizing cryogenic IR spectra for isomer identification, with potential applications in various analytical techniques, and highlights the successful development of methodologies for accurate and efficient identification of complex mixtures.

Acknowledgements

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

It feels like just yesterday when I arrived in Lausanne to embark on my PhD journey in May 2019. What once felt like an insurmountable challenge is now drawing to a close. The hardships I encountered along the way have transformed into cherished memories and invaluable friendships that I hold dear.

Leaving behind everything and everyone I had known and grown up with to pursue my PhD at EPFL was no easy decision. However, fate smiled upon me as I met my now wife, and found an incredible research group to collaborate with. These circumstances allowed me to build a family while pursuing my dream of attaining a PhD.

As I stand at the end of this academic road, I must acknowledge that none of this would have been possible without the unwavering support of my parents, *Hassan* and *Sama*, my younger sister, *Esraa*, my brother-in-law, *Oussama*, and my baby sister, *Amar*, though she is no longer a baby. You have all been my pillars of strength, consistently encouraging me to chase my dreams and surmount every obstacle while remaining grounded in faith. Your unconditional love and unwavering support have been the bedrock of my journey, and for this, I am eternally grateful. I owe you my past, my present, and my future.

I want to extend my deepest gratitude to my second family, Fawaz and Suzanne, as well as their children, Ali, Zeina, and Lea. You opened your hearts to me and offered comfort, particularly during those times when I felt homesick. Living with you all during the challenges of the COVID-19 pandemic was a source of great support and stability, especially those nights we would stay up playing games. You were my guiding lights here in Switzerland, and together we explored all that this beautiful country has to offer.

To my uncle *Ramez* and his wife *Afkar*, with whom I spent most of my undergraduate years, you have been my second family. I cherish the memories we created especially with my cousins, *Zahraa*, *Amir*, *Mohammad*, and *Mahdi*. Your unwavering support has molded me into the person I am today. Special gratitude goes to my college roommate, *Amir*, whose belief in my potential has been a constant source of motivation.

To my uncle *Ali* and your son, my dear big brother, *Mohammad Ali*, I hold close all that I've learned from you. You have been the steadfast backbone behind my achievements thus far. I carry a deep sense of indebtedness for your boundless generosity and will forever hold gratitude in my heart for everything you've done for me. *Hamoudy*, you are the true driving force behind my accomplishments. No words can adequately express the depth of my gratitude and sincerity towards you.

To my uncle *Mohammad*, I want to express my heartfelt gratitude for all the gifts you've been surprising me with since the day I was born. I'll never forget the joyous days we spent together at grandma's house with the whole family, and it's all thanks to your thoughtful gestures and warm-heartedness.

To my beloved grandma, *Zeinab*, I wish you were still with us to witness this accomplishment. You reside forever in my heart and prayers, until the day we reunite. May you be blessed by God.

My gratitude extends to my extended paternal family—uncles, aunts, and cousins—for the days of joy, laughter, and the invaluable role you played in my upbringing. Friday nights at grandma's were blessings we sometimes took for granted. I love you all dearly.

To my extended maternal family in Lebanon and my grandparents, *Hassan* and *Nova*, thank you for the wonderful summers and memorable adventures. You taught me the importance of being true to myself and voicing my thoughts. Auntie *Amal*, you are a second mother to me, and your son *Ali*, will forever be my brother.

Lastly, to you, my cherished wife, *Alina*, my partner in life.. My life transformed the day I met you, and you became my inner voice and my sanctuary. Your presence was a significant factor in my decision to embark on this academic journey in Switzerland, to be closer to you. The twinkle in your eyes and the warmth of your smile when we are together are the highlights of my day. My love for you knows no bounds, especially now that we have *Ivy,* our beautiful baby girl. You are my anchor during turbulent times and the driving force behind my best efforts. This thesis journey is as much yours as it is mine, as you have been by my side every step of the way. Your love is my constant.

I am deeply grateful to my undergraduate professor, Mary T. Rodgers, who introduced me to the fascinating world of analytical/physical chemistry and mass spectrometry. Her exceptional

personality and scientific guidance played an instrumental role in bolstering my confidence and inspiring me to pursue a PhD.

I owe a debt of gratitude to my esteemed professor, *Thomas Rizzo*, whose mentorship has been a wellspring of inspiration. Your unwavering dedication to pushing the boundaries of knowledge and your unwavering support in the lab have been transformative. I remember moments when experiments didn't go as planned, and your guidance and words of encouragement rekindled our determination to persevere. I am incredibly fortunate to have had the privilege of learning from you. Your kindness, and support have made our time together unforgettable. Thank you for your generosity in welcoming us to your home with your wife, *Karen*, for delightful BBQ gatherings.

On my first day of this PhD journey, I joined forces with *Vasyl Yatsyna*. He is perhaps the most kind-hearted person one could ever meet, and our connection evolved from strangers to colleagues to friends. He was instrumental in my success during this thesis, imparting invaluable knowledge on problem-solving, experimental design, and patience during experiments. His willingness to address my numerous questions with meticulous attention to detail has my deepest appreciation. I extend my thanks and wish him a bright future in his new endeavors.

Next, I must acknowledge *Stephan Warnke* and *Ahmed Ben Faleh*, the dynamic lab duo who were my constant companions throughout most of my thesis. Our professional relationship quickly grew into a personal friendship, and I learned a great deal from both of you. Your brilliance made my lab experience enjoyable and exhilarating. Thank you for your guidance and assistance throughout my thesis, and I eagerly anticipate the next chapter in our shared journey.

I am profoundly grateful to the remaining members of my lab at LCPM: *Eduardo, Irina, Robert*, *Priyanka*, *Natalia*, and *Lei*. You all played pivotal roles in shaping my research experience, and your collective wisdom, collaborative spirit, and unwavering support were essential to the successful completion of this thesis.

To *Angeles Alarcon*, who not only handled administrative matters but also knew how to bring a smile to my face during stressful times, I extend my gratitude. Your presence was a source of comfort.

My thanks also goes to the esteemed members of my thesis jury: Prof. *Rainer Beck* and Prof. *Sandrine Gerber* from EPFL, Prof. *Kevin Pagel* from Freie Universität Berlin, and Prof. *Jos Oomens* from Radboud University Nijmegen. Your insightful feedback and constructive criticism have been invaluable in refining my work and enhancing its academic rigor.

This thesis represents the culmination of the collective efforts, support, and encouragement of all those mentioned above, and I am deeply appreciative of each and every one of you.