

Analyzing glycans cleaved from a biotherapeutic
protein using a combination of ion mobility
spectrometry and cryogenic ion spectroscopy

Présentée le 26 mars 2021

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

Natalia YALOVENKO

Acceptée sur proposition du jury

Prof. H. Girault, président du jury
Prof. T. Rizzo, directeur de thèse
Dr I. Compagnon, rapporteuse
Prof. S. Flitsch, rapporteuse
Dr D. Ortiz, rapporteur

"Sometimes science is more art than science..."

Rick Sanchez

To all curious minds

Abstract

The characterization of biomolecules plays a central role in biomedical research. The development of state-of-the-art technologies is needed to make such analyses faster, more sensitive, and more accurate. While mass spectrometry continues to play a central role in biomolecular analysis, the use of multidimensional approaches that combine mass spectrometry, ion mobility spectrometry, and ion spectroscopy have been rapidly developing. This thesis investigates the application of such an approach for the routine analysis of glycans.

The first part of this thesis describes combining ultrahigh-resolution ion mobility with cryogenic IR spectroscopy and mass spectrometry for analysis of N-linked glycans. This approach is demonstrated by characterizing four N-glycans cleaved from the therapeutic fusion protein Etanercept that differ in abundance from 1% to 22% of the total N-glycan content. Our results demonstrate that the sensitivity, speed, and resolution of vibrational spectra are sufficient for large-scale N-glycan characterization and identification. In addition to the successful demonstration of our approach for identifying known N-glycans, we developed a strategy to synthesize and identify commercially unavailable positional isomers of N-glycans. This work demonstrates the potential of the technique to become a broadly accessible analytical tool for glycan analysis.

In the following part, the ability of our approach to quantify the glycan mixtures is investigated. Spray instability, different ionization efficiency of analytes, and their tagging efficiency complicate the process of quantification. Proposed solutions to these issues are discussed. We also present a study demonstrating that cryogenic infrared spectra of alkali metal complexes of carbohydrates of different complexity, including disaccharides and N-glycans can serve as fingerprints. We showed that the identity of the metal cation has no effect on the IR spectral complexity.

In the last part, cryogenic infrared spectroscopy with isotopic substitution in positive and negative ion modes was used for structural characterization of the reverse peptides GRGDS and SDGRG. The recorded cryogenic spectra represent a useful benchmark for theoretical structural predictions due to their sharp and distinct

features. This work represents the first step towards the structural determination of these peptides, with ongoing efforts focused on quantum chemical calculations.

Keywords: ion mobility spectrometry, high-resolution, cryogenic vibrational spectroscopy, carbohydrates, therapeutic fusion protein, biomolecules, metal complexes, structure, peptides.

Résumé

La caractérisation de biomolécules joue un rôle central dans la recherche biomédicale. Le développement des technologies de l'état de l'art est nécessaire pour rendre ces analyses plus rapides, plus sensibles, et plus précises. Alors que la spectrométrie de masse continue à jouer un rôle central dans l'analyse biomoléculaire, l'utilisation d'approches multidimensionnelles combinant la spectrométrie de masse, la mobilité ionique, et la spectroscopie ionique sont en croissance rapide. Cette thèse examine l'application d'une telle approche pour une analyse routine de glycanes.

La première partie de cette thèse décrit la combinaison de mobilité ionique à ultra-haute résolution avec de la spectroscopie IR cryogénique et la spectrométrie de masse pour l'analyse d'oligosides N-glycosylés. Cette approche est démontrée par la caractérisation de quatre N-glycanes coupés de la protéine thérapeutique de fusion Etanercept qui diffère en abondance de 1% à 22% du contenu total de N-glycanes. Nos résultats démontrent que la sensibilité, la rapidité, et la résolution de spectroscopie vibrationnelle est suffisante pour la caractérisation et l'identification de N-glycanes *connues* à grande échelle. Ce travail démontre le potentiel de cette technique de devenir un outil analytique largement accessible pour l'analyse de glycanes.

Dans la partie suivante, la capacité de notre approche à quantifier les mélanges de glycanes est examinée. L'instabilité de la source, les efficacités d'ionisation de différents analytes, et leurs efficacités de complexation compliquent le processus de quantification. Des solutions à ces problèmes sont proposées. Nous présentons également une étude démontrant que les spectres infrarouges cryogéniques des complexes de métaux alcalins et de glucides de différentes complexités, qui incluent des disaccharides et des N-glycanes, peuvent servir d'empreinte digitale. Nous montrons que l'identité de cation métallique n'a aucune influence sur la complexité spectrale IR.

Dans la troisième partie, la spectroscopie cryogénique infrarouge avec substitution isotopique en mode cationique et anionique était utilisée pour la caractérisation de la

séquence de peptides inversés, GRGDS et SDGRG. Les spectres acquis représentent des points de références utiles pour des prédictions structurales théoriques, due à leurs caractéristiques distincts et biens définis. Ce travail représente le premier pas vers la détermination structurale de ces peptides, avec du travail en cours, axé sur des calculs quantiques.

Mots clés: spectrométrie de mobilité ionique, haute résolution, spectroscopie vibrationnelle cryogénique, hydrates de carbone, protéine de fusion thérapeutique, biomolécules, complexes métalliques, structures, peptides.

List of abbreviations

nESI: nanoelectrospray

DNA: deoxyribonucleic acid

RNA: ribonucleic acid

SNFG: symbol nomenclature for glycans

OST: oligosaccharyltransferase

CDG: congenital disorder of glycosylation

CHO: Chinese hamster ovary

Igs: immunoglobulins

mAb: monoclonal antibody

TNFR: tumor necrosis factor receptor

Fc: fragment crystallizable

CDC: complement-dependent cytotoxicity

PNGaseF: peptide N-glycosidase F

SPE: solid-phase extraction

CE: capillary electrophoresis

LC: liquid chromatography

PGC: porous graphitized carbon

HILIC: hydrophilic interaction liquid chromatography

HPLC: high-performance liquid chromatography

MS: mass spectrometry

IMS: ion mobility spectrometry

m/z: mass-to-charge ratio

DT-IMS: drift tube ion mobility spectrometry

TIMS: trapped ion mobility spectrometry

TWIMS: traveling-wave ion mobility spectrometry

NMR: nuclear magnetic resonance spectroscopy

SLIM: structures for lossless ion manipulations

CCS: collisional cross sections

ATD: arrival time distribution

PCB: printed circuit board

TOF: time-of-flight

MALDI: matrix-assisted laser desorption/ionization

CID: collision-induced dissociation

IRMPD: infrared multiphoton dissociation

DB: database

FUT8: alpha-(1,6)-fucosyltransferase

OPO /A: optical parametric oscillator/ amplifier

LOD: limit of detection

LOQ: limit of quantification

HB: hydrogen bonding

Table of contents

ABSTRACT	II
RÉSUMÉ	IV
LIST OF ABBREVIATIONS.....	VI
TABLE OF CONTENTS	VIII
CHAPTER 1 INTRODUCTION.....	1
1.1 CARBOHYDRATES ARE ONE OF THE MAIN CLASSES OF BIOLOGICAL MACROMOLECULES	1
1.2 INTRINSIC STRUCTURAL COMPLEXITY OF GLYCANS.....	2
1.3 STRUCTURAL NOMENCLATURE OF GLYCANS.....	4
1.4 N-LINKED GLYCOSYLATION AS A POST-TRANSLATIONAL MODIFICATION OF PROTEINS	6
1.5 GLYCOPROTEINS ARE A MAJOR CLASS OF BIOTHERAPEUTIC PROTEINS.....	11
1.5.1 Therapeutic Fc-Fusion Proteins	12
OVERVIEW OF THE THESIS	14
REFERENCES:.....	15
CHAPTER 2 ANALYSIS OF N-GLYCANS	23
2.1 SEPARATION TECHNIQUES	23
2.1.1 Capillary electrophoresis.....	23
2.1.2 Liquid chromatography	24
2.1.3 Ion mobility spectrometry	25
2.2 IDENTIFICATION TECHNIQUES.....	27
2.2.1 Nuclear magnetic resonance spectroscopy.....	27

2.2.2	Mass spectrometry.....	28
2.2.3	Spectroscopy.....	29
2.3	GLYCAN DATABASES.....	30
2.4	OUR NEW APPROACH TO GLYCAN ANALYSIS.....	31
	REFERENCES:.....	33
CHAPTER 3 EXPERIMENTAL APPROACH.....		45
3.1	ION MOBILITY SPECTROMETRY COUPLED WITH CRYOGENIC ION SPECTROSCOPY.....	45
3.1.1	Overview of the instrument.....	45
3.1.2	SLIM-based ion mobility module.....	48
3.1.3	Principles of cryogenic messenger-tagging spectroscopy.....	52
	REFERENCES:.....	55
CHAPTER 4 ANALYZING GLYCANS CLEAVED FROM A BIOTHERAPEUTIC PROTEIN USING ULTRAHIGH-RESOLUTION ION MOBILITY SPECTROMETRY TOGETHER WITH CRYOGENIC ION SPECTROSCOPY *.....		58
4.1	SAMPLE PREPARATION	58
4.1.1	Sample preparation of standards.....	58
4.1.2	N-linked glycans release from Etanercept and clean-up.....	59
4.2	COMPARISON OF THE RELEASED SUGARS WITH THEIR RESPECTIVE STANDARDS.....	61
4.3	COMPARISON OF MATHEMATICAL METHODS TO QUANTIFY THE SIMILARITY BETWEEN IR SPECTRA OF REFERENCE AND CLEAVED GLYCANS	69
4.3.1	Dot product method.....	70
4.3.2	Cross-correlation method.....	72
4.3.3	Correlation coefficient method.....	73
4.4	NEW PERSPECTIVES ON STUDYING RITUXIMAB	75

4.5	CONCLUSIONS AND FUTURE PERSPECTIVES	77
	REFERENCES:.....	79
CHAPTER 5 STRATEGY TO DISTINGUISH POSITIONAL ISOMERS OF N-GLYCANS*		82
5.1	DISTINGUISHING POSITIONAL ISOMERS OF N-GLYCANS IS A CHALLENGE....	82
5.2	CHEMOENZYMATIC SYNTHESIS OF ASYMMETRIC N-GLYCANS	84
5.3	IDENTIFICATION OF POSITIONAL N-GLYCAN ISOMERS.....	88
5.4	CONCLUSIONS AND FUTURE PERSPECTIVES	96
	REFERENCES:.....	98
CONCLUSIONS AND FUTURE PERSPECTIVES		104
	REFERENCES:.....	106
CHAPTER 6 GLYCAN QUANTIFICATION		107
6.1	THE PROBLEM OF GLYCAN QUANTIFICATION.....	107
6.1.1	Parameters for method validation	109
6.2	EXPERIMENTAL METHODS	111
6.2.1	Tandem mass-spectrometer for cryogenic spectroscopy	111
6.2.2	Ion mobility spectrometer coupled to a cryogenic trap and TOF	114
6.3	QUANTIFICATION OF GLYCAN MIXTURES	114
6.3.1	Sample preparation and notation	114
6.3.2	Quantifying vibrational spectra of disaccharide mixtures	115
6.3.3	Quantifying vibrational spectra of trisaccharide mixtures	121
6.3.4	Identifying limiting factors	124
6.3.4.1	<i>Electrospray</i>	124
6.3.4.2	<i>Tagging efficiency</i>	129
6.3.4.3	<i>Linearity of a channeltron</i>	131
6.3.4.4	<i>Other parameters</i>	132
6.3.5	Parameters for method validation	133

6.3.5.1	<i>Precision</i>	133
6.3.5.2	<i>Dynamic range, linearity, LOD and LOQ</i>	136
6.4	CONCLUSIONS AND FUTURE PERSPECTIVES	141
	REFERENCES:.....	143
CHAPTER 7 METAL COMPLEXES OF GLYCANS.....		150
7.1	SAMPLE PREPARATION	150
7.2	METAL COMPLEXES OF DISACCHARIDES	151
7.3	METAL COMPLEXES OF N-GLYCANS.....	156
7.4	CONCLUSIONS AND FUTURE PERSPECTIVES	161
	REFERENCES:.....	163
CHAPTER 8 SPECTROSCOPY OF GRGDS AND SDGRG.		164
8.1	RGD PEPTIDE MOTIF AND ITS BIOLOGICAL RELEVANCE.....	164
8.2	CALIBRATION MIXTURE OF PENTAPEPTIDES GRGDS AND SDGRG FOR ION-MOBILITY SET-UPS	166
8.3	CURRENT METHODS TO CHARACTERIZE THE STRUCTURE OF THE REVERSE PEPTIDES IN THE GAS PHASE.....	167
8.4	EXPERIMENTAL METHODS	167
8.4.1	Tandem mass-spectrometer for cryogenic spectroscopy	167
8.4.2	Sample preparation	168
8.4.3	Isotopic labeling	168
8.5	CRYOGENIC ION SPECTROSCOPY OF GRGDS AND SDGRG.....	169
8.5.1	Vibrational spectra of singly-charged GRGDS and SDGRG.....	169
8.5.2	The assignment of amide NH stretches of singly-charged SDGRG	171
8.5.3	Vibrational spectra of the doubly-charged GRGDS and SDGRG.....	172
8.5.4	The assignment of amide NH stretches of doubly charged SDGRG.....	173
8.5.5	Vibrational spectra of deprotonated GRGDS and SDGRG	175
8.5.6	The assignment of amide NH stretches of deprotonated SDGRG.....	176

8.6 CONCLUSIONS AND FUTURE PERSPECTIVES	178
REFERENCES:.....	179
APPENDIX	184
APPENDIX A	184
APPENDIX B	187
APPENDIX C	196
APPENDIX D	202
ACKNOWLEDGMENTS.....	224
CURRICULUM VITAE.....	228

Chapter 1 Introduction

1.1 Carbohydrates are one of the main classes of biological macromolecules

Carbohydrates, also rereferred to as sugars or glycans, are one of the main classes of biological macromolecules that can be found in all forms of life. For example, ribose and deoxyribose are important building blocks of ribonucleic acid (RNA), deoxyribonucleic acid (DNA), and adenosine triphosphate (ATP). Carbohydrates are also elements in various other molecules, including antibiotics, coenzymes, and toxins. Cellulose is the main constituent of plant cell walls and the most abundant molecule on the Earth.

Glycans play a myriad of roles in biological systems [1–6]. For example, they serve as a source of energy [2,6]. There is also evidence for catalytic properties of certain sugars [7]. Carbohydrates serve as structural and protective elements in the cell walls of bacteria and plants and are found in the connective tissues and cell coats of mammals [1,8]. This layer of glycans (glycocalyx) plays a significant role in cell-to-cell recognition, cell signaling, and immune system responses. Additionally, glycans are involved in some way with nearly every disease [2,9,10]. The severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) virus is no exception – the so-called spike (S) protein responsible for its attachment to cells is heavily decorated with glycans [11,12].

Understanding the structures and functions of glycans is central to understanding life. However, glycomics remains a nascent field of study. This results from two factors: (1) the intrinsic structural complexity of glycans and (2) a lack of tools for deciphering glycosylation patterns. This is unfortunate, since the fields of genomics and proteomics greatly benefit from existing sequencing methods that have led to a greater understanding of many of the basic mechanisms present in biological systems [13–16].

This thesis introduces a promising technique to study sugars released from biological samples which is based upon preliminary work conducted by Masellis *et al* [17,18].

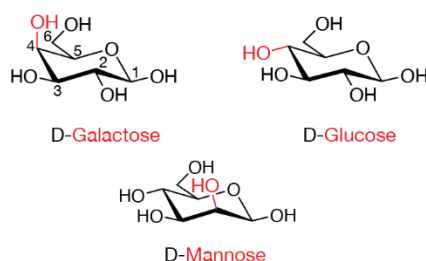
This chapter will first describe fundamental principles of glycans, their synthesis, and their functions. Then, glycosylated therapeutic proteins will be introduced. Finally, an overview of the structure of the thesis will be given.

1.2 Intrinsic structural complexity of glycans

Glycan complexity arises from their intrinsic chemical nature. They are comprised of the simplest sugar building blocks called monosaccharides (mono = one, sacchar = sugar). Monosaccharides can be classified by their number of carbon atoms, for example, trioses, tetroses, pentoses, hexoses, etc.

One level of glycan complexity arises from the isomeric nature of the most common monosaccharides that differ only in the stereochemistry of the chiral centers. Figure 1.1 demonstrates this variety using galactose, glucose, and mannose as an example. Sugars differing in configuration at a single chiral carbon atom are isomers known as epimers. For example, D-galactose and D-glucose are C-4 epimers, whereas D-galactose and D-mannose are C-2 epimers, referred to as diastereoisomers. (Figure 1.1 a). Additionally, carbohydrates occur as α and β anomers depending on the position of OH group at the C1 anomeric carbon (Figure 1.1 b) [1].

a) Epimers



b) Anomers

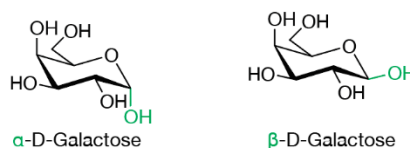


Figure 1.1 Configurations of monosaccharides shown for galactose, glucose and mannose. a) D-galactose and D-glucose are epimers that differ at the C4 carbon, D-galactose and D-mannose are epimers that differ at the C2 carbon; b) α -D-galactose

and β -D-galactose are anomers that differ by the orientation of OH group at the anomeric C1 carbon.

Mammalian carbohydrate structures are primarily comprised of only a few monosaccharide building blocks, such as D-glucose (Glc), D-galactose (Gal), D-mannose (Man), N-acetyl-D-glucosamine (GlcNAc), N-acetyl-D-galactosamine (GalNAc), L-fucose (Fuc) and N-acetylneuraminic acid (NeuAc) (Figure 1.2) [19].

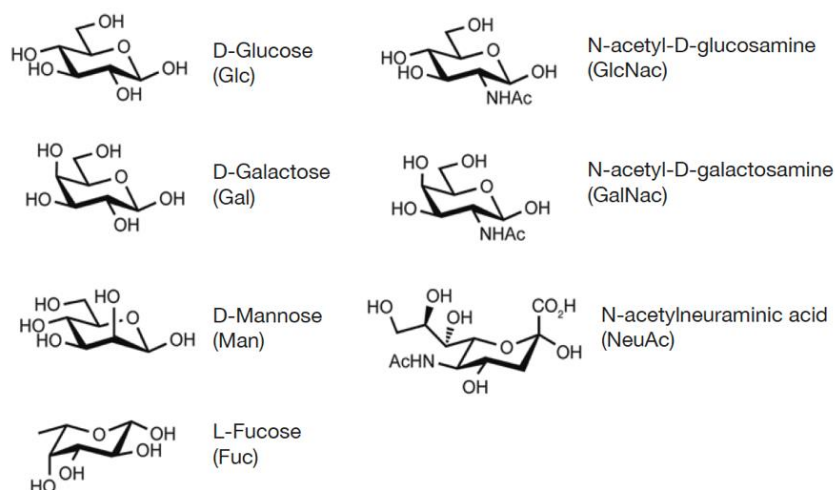


Figure 1.2 Structures of the most abundant monosaccharides discovered in animals.

The diversity and complexity of glycans increase when the monosaccharide units are linked together by characteristic linkages called glycosidic bonds. Whereas in peptides or proteins, amino acids have one possible connection point resulting in a linear arrangement, two monosaccharide units can be connected in numerous ways, giving a rise to an immense variety of glycans. Very large numbers of different glycans can be made by varying number, order and type of monosaccharide units. This arises from the fact that 1) the glycosidic bond can be formed in α - or β -configurations (Figure 1.3 a); 2) each monosaccharide has many OH groups at different positions at which to form a glycosidic bond (Figure 1.3 b); and 3) monosaccharide residues can be involved in two glycosidic bonds, thereby serving as a branching point (Figure 1.3 c). As an example of the isomeric complexity of glycans, an aldohexose ($C_6H_{12}O_6$) and four asymmetric carbon atoms can be described in 16 different isomeric forms.

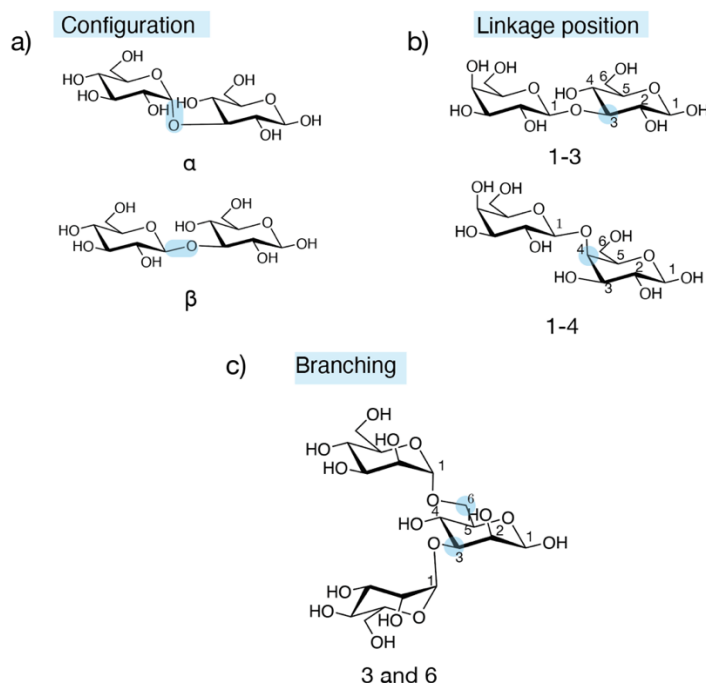


Figure 1.3 Carbohydrate structural diversity. a) two stereoisomers differ in configuration at the anomeric carbon; b) two structural isomers (regioisomers) that differ in the linkage position; c) monosaccharide residue of trisaccharide exhibits two branching points at C3 and C6.

1.3 Structural nomenclature of glycans

Given the intrinsic complexity of carbohydrate structures, it is not always practical to draw each chemical formula, especially if the drawings include multiple structures for the complex glycans. To address this issue, one uses symbols and abbreviations to depict monosaccharides in complex glycans [20]. Historically, there has been no consensus regarding a standard way of sugar representation; however, in 1999, a standardized nomenclature for glycans was first introduced. In 2005, the revised notation received the approval of the Consortium for Functional Glycomics (CFG) and was known as a CFG nomenclature (Figure 1.4 a) [1,21]. In this nomenclature the monosaccharide units are denoted by colored geometric forms. The shape specifies the class of the monosaccharide (hexose, deoxy-hexose) while the color designates the specific monosaccharide, such as mannose, glucosamine, fucose. Anomer

and linkage information is presented as text close to the line of the glycosidic bond. The Oxford Glycobiology Institute introduced an alternative, the UOXF notation [22] that indicates the configuration by the line type (continuous for α -configuration, dashed for β -configuration), while the linkage position is visually represented by the orientation and angle of monosaccharides (Figure 1.4 b).

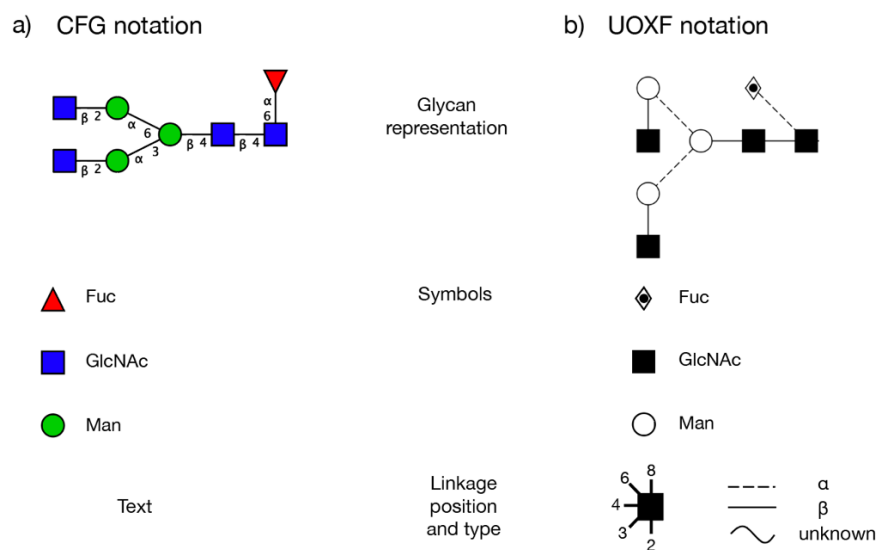


Figure 1.4 The comparison of a) the CFG and b) the UOXF nomenclatures based on the representation of glycans, used symbols, linkage type and position.

Finally, in 2015 the symbol nomenclature for glycans (SNFG) was introduced [23] that combines elements of the CFG nomenclature with those of UOXF. Here, all glycosidic linkages may follow the UOXF notation that inserts both the anomericity and specificity of the monosaccharide linkages. This nomenclature is widely used in the community. Figure 1.5 demonstrates the application of the SNFG to the A2F glycan. To simplify the presentation and discussion of glycan structures, the SNFG is used in this thesis.

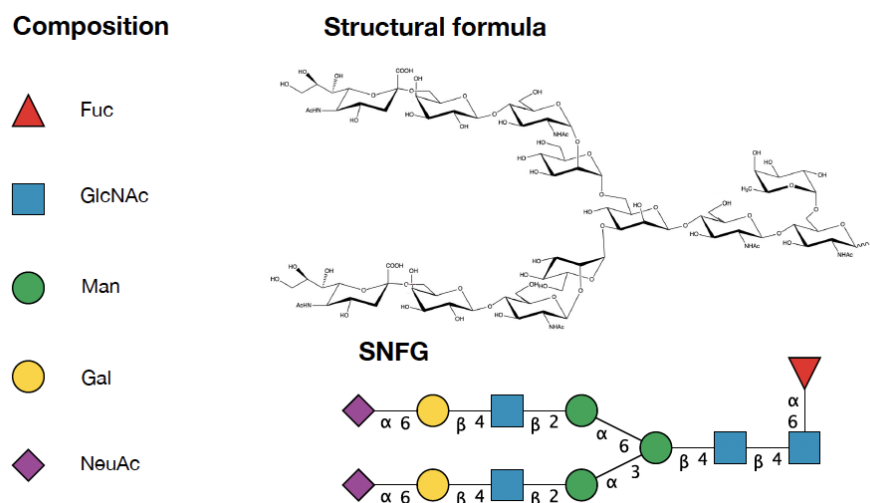


Figure 1.5 The demonstration of the symbol nomenclature for glycans (SNFG). Monosaccharides are depicted using symbols, stereochemistry is omitted for the simplicity. A2F N-glycan is used as an example that drawn as the structural formula and by the SNFG representation.

1.4 N-linked glycosylation as a post-translational modification of proteins

In addition to the intrinsic chemical nature and various structures of glycans, they also can form glycoconjugates. Glycoconjugates are molecules of glycans covalently bonded to other biomolecules, including proteins and lipids, through glycosidic linkages. The main classes of conjugates that cover the outer surface of mammalian cells are shown in Figure 1.6 [1,24,25].

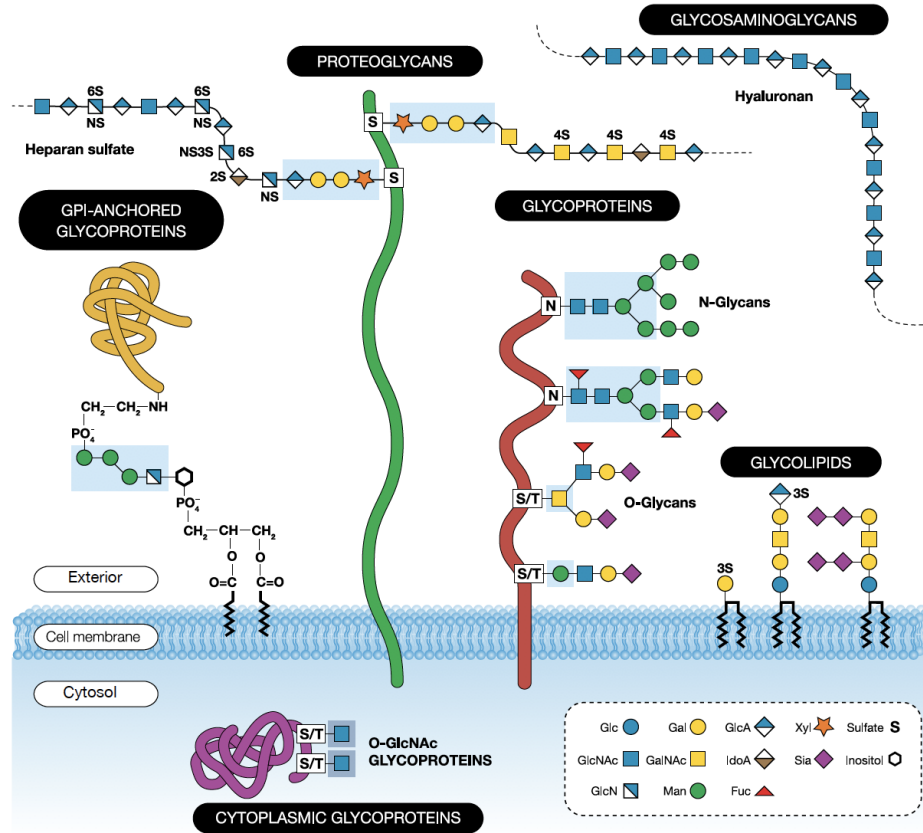


Figure 1.6 A schematic representation of the common glycans and glycoconjugates in mammals. The filled light blue boxes indicate the core structure for each glycan. Glycans are represented according with the SNFG.

The majority of eukaryotic proteins are modified by post-translational attachment of complex oligosaccharides/glycans to generate the most complex modification – protein glycosylation. There are various types and mechanisms of glycosylation, but the most common type is N-glycosylation. In fact, 90% of all glycoproteins in nature are expected to be N-glycosylated, which is not surprising, since N-linked glycosylation is a co-translational event (i.e. a glycan is attached to a protein as the latter is being translated and transported in the endoplasmic reticulum (ER)). N-glycosylation holds the "N" in its name since the glycan moieties contain a GlcNAc bound via a β 1-glycosidic linkage to the nitrogen atom of the amino group of asparagine (Asn, N) at the consensus glycosylation motif Asn-X-Ser/Thr (in which X denotes any amino acid except for Pro) of the nascent recipient protein [1].

Considering that N-glycans cover the cell surface, they are involved in a myriad of cellular and molecular processes, including cell–cell interaction, cell/matrix interaction, cell growth, cell differentiation, cell adhesion, etc [3,4,26]. Bound N-glycans can modulate the functions to proteins and lipids. For example, α 1-6 core fucosylation of N-glycans can alter signaling of transforming growth factors that might lead to abnormal lung development [27]. Additionally, it has been demonstrated that N-glycans play a major role in glycoprotein folding, since the primary steps of protein N-glycosylation, involving the removal of glucose and mannose residues, have a major impact on protein quality control [4]. Moreover, N-glycans are a significant target for glycoprotein-based therapeutics, since these sugars affect structural integrity and glycoprotein function. Glycosylation of therapeutic proteins is considered a critical quality attribute (CQA) that affects their bioactivity, efficacy, solubility, stability, pharmacokinetics, pharmacodynamics, and immunogenicity [28,29]. Monitoring batch-to-batch glycosylation consistency in therapeutic glycoprotein production and comparing biosimilar monoclonal antibodies (mAbs) and Fc-fusion proteins to their originator biologics are required to ensure the product quality [30].

N-glycans can affect immune reactions in humans [31]. The alterations of the glycosylation of tumor cells can lead to immunosuppressive response [32]. Many diseases, such as congenital disorder of glycosylation (CDG), cancer, diseases of blood vessels, and rheumatoid arthritis are associated with N-glycans [10,12]. For example, CDG was initially described as a genetic defect in the N-glycosylation pathway. An overexpression of branched N-glycan is typical for tumor cells. In addition, N-glycans can bind to various bacteria, parasites, and viruses [4,33]. For instance, many viruses comprise the protein hemagglutinin that specifically recognizes the type of host sialic acid, its modifications, and its linkage. Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) recently emerged as a threat to the global society. Consequently, N-linked glycans on the Coronavirus spike glycoprotein that can result in receptor binding, membrane fusion, and vaccine design [12,34] have been extensively studied.

While N-glycans are unquestionably important, they are also exceptionally complex. Two large obstacles in the study of glycans are their aforementioned non-linear complex chemical structure and the absence of a direct genetic template. Unlike

proteins, which are synthesized strictly according to their genetic blueprint, glycan biosynthesis is not template-driven. The biosynthesis of N-glycans, shown schematically in Figure 1.7, can be divided into the following steps [35]: 1) precursor glycan assembly; 2) attachment of the glycan precursor to asparagine residues of the protein; 3) trimming; and 4) maturation. It is important to mention that the steps of the synthetic pathways that take place in the ER are highly conserved within eukaryotes. However, the steps in the Golgi pathway lead to the structural diversity of N-linked glycans, because they require a large number of enzymes and precursors. In fact, the Golgi pathway leads to the principal differences between plants, fungi, and mammals.

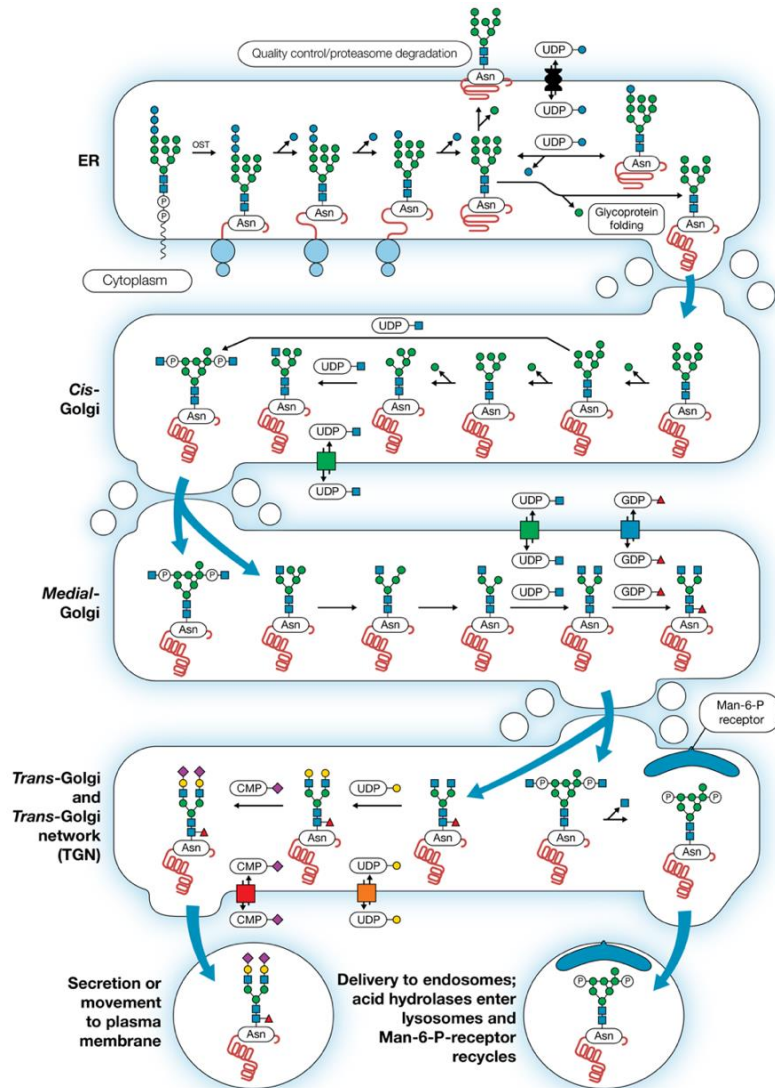


Figure 1.7 A schematic biosynthesis of *N*-glycans.

The resulting glycan structures have a common core $\text{Man}_3\text{GlcNAc}_2$ and can be divided into the stated groups: a) high(oligo)mannose (Figure 1.8 a); b) complex (Figure 1.8 b); hybrid (Figure 1.8 c) [4].

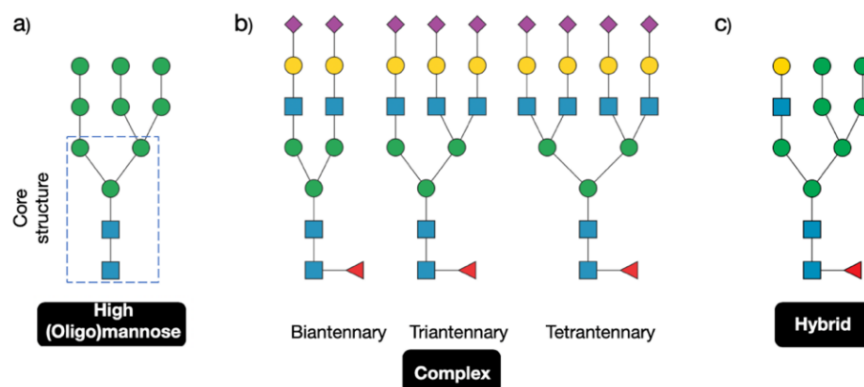


Figure 1.8 Types of *N*-glycans that share a common pentasaccharide core $\text{Man}_3\text{GlcNAc}_2$ (indicated in blue dashed box), a) high(oligo)mannose, b) complex with different antennary forms and c) hybrid. Regio- and stereochemistry is not mentioned to simplify the presentation.

High(Oligo)-mannose glycans contain only mannose residues varying from five to nine in total attached to the $(\text{GlcNAc})_2$. The glycan name abbreviations are derived from the total number of mannose residues in the structure. For example, Man_5 has five mannose residues in total. Complex *N*-glycans contain GlcNAc sequences that are known as “antennae”. Complex glycans exist in bi-, tri- and tetra antennary forms (Figure 1.9 b). Hybrid glycans contain both terminal mannose residues and *N*-acetylglucosamine residues.

The crucial role that *N*-glycans play in numerous life processes together with their complexity were the motivation factors in choosing them as the primary target of study in this thesis.

1.5 Glycoproteins are a major class of biotherapeutic proteins

Recombinant biotherapeutic proteins, or biologics, are produced in living cells. Biologics, including mAbs, fusion proteins, and other recombinant protein products, are accepted worldwide biological drug classes [36]. In 2018, eight of the top ten highest-selling drugs in the world were biologics [37]. This increasing number of biologics highlights their impact in the pharmaceutical industry. Therapeutic recombinant proteins have advantages over small-molecule drugs, such as high specificity, decreased immunogenicity, and fast clinical development and regulatory approval [38,39]. Biologics are commonly used for clinical treatment of a large variety of diseases, such as cancers, chronic inflammatory diseases, metabolic disorders, and other life-threatening conditions [40,41].

The main expression systems used to produce recombinant proteins are mammalian, bacteria, and yeast cells [42]. Due to the differences among them, especially in terms of post-translational modifications, it is hard to guarantee the product quality. Glycosylation is one of the most ubiquitous and by far the most complex post-translational modifications of recombinant proteins [43]. Nearly 70% of all therapeutic recombinant proteins are glycosylated [44]. For example, Erythropoietin, known as Epogen® (Amgen), was the first glycosylated recombinant glycoprotein produced in Chinese hamster ovary (CHO) cells. This drug was approved by the US FDA (Food and Drug Administration) in 1985 [45].

As has been discussed in the previous section, glycosylation is an important quality attribute of a recombinant protein that can alter intracellular trafficking, folding, stability, and solubility of a protein [13,14,34]. Glycosylation heterogeneity is influenced by various factors, such as host cell type and cell culture environment, which includes pH, dissolved oxygen, stirring rate, and temperature [47,48].

The safety, purity, and biological activity of a glycosylated recombinant protein are generally impacted by its formulation during the manufacturing process. Consequently, the primary focus is the control of the manufacturing process to achieve the desired product quality. Additionally, there is a growing market of biosimilars, which are less expensive copies of the licensed biologics. Biosimilars can only be

approved if a manufacturing process demonstrates similarity and batch-to-batch consistency in glycosylation comparing to the originator biologics. As a result, glycosylation is categorized as a critical quality attribute that needs an adequate analytical approach to guarantee product quality. However, the isomeric complexity and branched structure of glycans makes their analysis a significant challenge for the existing techniques. This is why there is a need for a new tool for glycan analysis of biotherapeutics.

1.5.1 *Therapeutic Fc-Fusion Proteins*

The type of biotherapeutics studied in this thesis is therapeutic Fc-fusion proteins. They are also known as Fc chimeric fusion proteins. Therapeutic Fc-fusion proteins are innovative biological entities with potential as valuable therapeutic drugs [49]. They have been broadly used in academic and industrial endeavors. Fc-fusion proteins are used as alternatives to mAbs and other drugs.

Fc-Fusion proteins are made of a specific polypeptide or protein genetically attached to the Fc (fragment crystallizable) part of immunoglobulin G (IgG). The Fc domains have been selected for fusion to polypeptides for therapeutic application since the Fc part contributes to prolonged the half-life of recombinant proteins and peptides [50,51]. A longer half-life of the drug leads to reduced total drug requirements and improved pharmacological exposure.

The most successful Fc-Fusion protein in clinical use is Etanercept, which is a key pharmaceutical component of the drug Enbrel® [52]. This drug is considered as one of the top products in the therapeutic biologics market. It is used to treat autoimmune diseases, including rheumatoid arthritis, juvenile idiopathic arthritis, psoriatic arthritis, and plaque psoriasis [51,52].

Etanercept is a dimeric fusion protein containing the extracellular domain of the human tumor necrosis factor receptor (TNFR) attached to the F portion of human IgG1 by disulfide bonds (and hence its abbreviation TNFR-Fc). A schematic representation is shown in Figure 1.9 [54]. It is produced by recombinant DNA technology in CHO cells. Etanercept functions as a tumor necrosis factor (TNF) inhibitor by binding to TNF, a proinflammatory cytokine, and blocking the inflammatory cellular responses. TNFR-Fc is a glycoprotein with 934 amino acids and

a molecular weight of ~ 128 kDa. The glycans attached to the protein can potentially affect the structure, clearance, signaling, and immunogenicity [54].

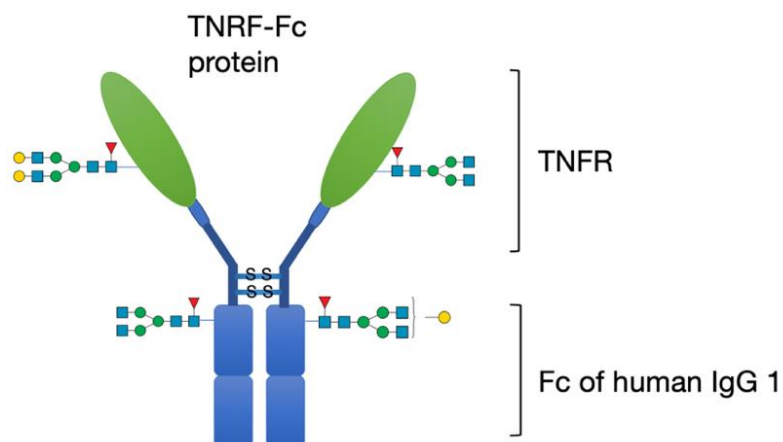


Figure 1.9 A schematic structure of Etanercept. Interchain disulfide bonds within the hinge region are presented.

Overview of the thesis

The main part of this thesis focuses on the analysis of N-glycans cleaved from a therapeutic protein. In Chapter 2, an overview of the existing techniques to analyze released N-glycans from therapeutic proteins is provided. Chapter 3 gives a detailed description of our experimental approach that we used to analyze native glycans cleaved from a biotherapeutic protein in the gas phase. Chapter 4 demonstrates the ability of our technique to identify N-linked glycans cleaved from a therapeutic protein using a database approach, which serves as a proof of principle that the sensitivity, speed, and resolution should be sufficient for large-scale N-glycan characterization and identification. Mathematical methods to quantify the similarity between IR spectra of reference and cleaved glycans are also presented. A promising strategy to synthesize and distinguish positional isomers of N-glycans that are not available on the market is described in Chapter 5. Moreover, we discuss the necessary future experiments which are needed for this strategy to prove its utility. The “Conclusion and future perspectives” section summarizes our key results and gives an overview of future directions.

Another part of this thesis focuses on the quantitation of cryogenic spectra of glycans. In Chapter 6, an overview of the existing methods to quantify glycans is described. Additionally, it provides a description the experimental set-up and the mathematical methods that we used to quantify sugars in the gas phase. The limiting factors are investigated and the future perspectives to improve the current results are given. The effect of metals on the complexity of cryogenic IR spectra of glycans employing our technique is discussed in Chapter 7.

Chapter 8 is not related to glycans but involves a side project applying similar experimental techniques for structural characterization of two reverse sequence peptides GRGDS and SDGRG.

References:

- (1) Varki A, Cummings R, Esko J, Freeze H, Hart G, Marth J. *Essentials of Glycobiology*; Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press, 1999.
- (2) Varki, A.; Lowe, J. B. Biological Roles of Glycans. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., Etzler, M. E., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2009.
- (3) Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27* (1), 3–49. <https://doi.org/10.1093/glycob/cww086>.
- (4) Spearman, M.; Butler, M. Glycosylation in Cell Culture; 2015; Vol. 9, pp 237–258. https://doi.org/10.1007/978-3-319-10320-4_9.
- (5) Dwek, R. A. Glycobiology: Toward Understanding the Function of Sugars. *Chem. Rev.* **1996**, *96* (2), 683–720.
- (6) O'Neill, M. A.; Moon, R. J.; York, W. S.; Darvill, A. G. Glycans in Bioenergy and Materials Science. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2015.
- (7) Del Valle, E. M. M. Cyclodextrins and Their Uses: A Review. *Process Biochem.* **2004**, *39* (9), 1033–1046. [https://doi.org/10.1016/S0032-9592\(03\)00258-9](https://doi.org/10.1016/S0032-9592(03)00258-9).
- (8) Jørgensen, N. O. G. Carbohydrates. In *Encyclopedia of Inland Waters*; Likens, G. E., Ed.; Academic Press: Oxford, 2009; pp 727–742. <https://doi.org/10.1016/B978-012370626-3.00258-1>.
- (9) National Research Council (US) Committee on Assessing the Importance and Impact of Glycomics and Glycosciences. *Transforming Glycoscience: A Roadmap for the Future*; The National Academies Collection: Reports funded by National Institutes of Health; National Academies Press (US): Washington (DC), 2012.

- (10) Kolarich, D.; Lepenies, B.; Seeberger, P. H. Glycomics, Glycoproteomics and the Immune System. *Curr. Opin. Chem. Biol.* **2012**, *16* (1–2), 214–220. <https://doi.org/10.1016/j.cbpa.2011.12.006>.
- (11) Grant, O. C.; Montgomery, D.; Ito, K.; Woods, R. J. Analysis of the SARS-CoV-2 Spike Protein Glycan Shield: Implications for Immune Recognition. *bioRxiv* **2020**. <https://doi.org/10.1101/2020.04.07.030445>.
- (12) Watanabe, Y.; Allen, J. D.; Wrapp, D.; McLellan, J. S.; Crispin, M. Site-Specific Glycan Analysis of the SARS-CoV-2 Spike. *Science* **2020**. <https://doi.org/10.1126/science.abb9983>.
- (13) Lander, E. S.; Linton, L. M.; Birren, B.; Nusbaum, C.; Zody, M. C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; Funke, R.; Gage, D.; Harris, K.; Heaford, A.; Howland, J.; Kann, L.; Lehoczky, J.; LeVine, R.; McEwan, P.; McKernan, K.; Meldrim, J.; Mesirov, J. P.; Miranda, C.; Morris, W.; Naylor, J.; Raymond, C.; Rosetti, M.; Santos, R.; Sheridan, A.; Sougnez, C.; Stange-Thomann, N.; Stojanovic, N.; Subramanian, A.; Wyman, D.; Rogers, J.; Sulston, J.; Ainscough, R.; Beck, S.; Bentley, D.; Burton, J.; Clee, C.; Carter, N.; Coulson, A.; Deadman, R.; Deloukas, P.; Dunham, A.; Dunham, I.; Durbin, R.; French, L.; Grafham, D.; Gregory, S.; Hubbard, T.; Humphray, S.; Hunt, A.; Jones, M.; Lloyd, C.; McMurray, A.; Matthews, L.; Mercer, S.; Milne, S.; Mullikin, J. C.; Mungall, A.; Plumb, R.; Ross, M.; Shownkeen, R.; Sims, S.; Waterston, R. H.; Wilson, R. K.; Hillier, L. W.; McPherson, J. D.; Marra, M. A.; Mardis, E. R.; Fulton, L. A.; Chinwalla, A. T.; Pepin, K. H.; Gish, W. R.; Chissoe, S. L.; Wendl, M. C.; Delehaunty, K. D.; Miner, T. L.; Delehaunty, A.; Kramer, J. B.; Cook, L. L.; Fulton, R. S.; Johnson, D. L.; Minx, P. J.; Clifton, S. W.; Hawkins, T.; Branscomb, E.; Predki, P.; Richardson, P.; Wenning, S.; Slezak, T.; Doggett, N.; Cheng, J.-F.; Olsen, A.; Lucas, S.; Elkin, C.; Uberbacher, E.; Frazier, M.; Gibbs, R. A.; Muzny, D. M.; Scherer, S. E.; Bouck, J. B.; Sodergren, E. J.; Worley, K. C.; Rives, C. M.; Gorrell, J. H.; Metzker, M. L.; Naylor, S. L.; Kucherlapati, R. S.; Nelson, D. L.; Weinstock, G. M.; Sakaki, Y.; Fujiyama, A.; Hattori, M.; Yada, T.; Toyoda, A.; Itoh, T.; Kawagoe, C.; Watanabe, H.; Totoki, Y.; Taylor, T.; Weissenbach, J.; Heilig, R.; Saurin, W.; Artiguenave, F.; Brottier, P.; Bruls, T.; Pelletier, E.; Robert, C.; Wincker, P.; Rosenthal, A.; Platzer, M.; Nyakatura, G.; Taudien, S.; Rump, A.; Smith, D. R.; Doucette-Stamm, L.

Rubinfeld, M.; Weinstock, K.; Lee, H. M.; Dubois, J.; Yang, H.; Yu, J.; Wang, J.; Huang, G.; Gu, J.; Hood, L.; Rowen, L.; Madan, A.; Qin, S.; Davis, R. W.; Federspiel, N. A.; Abola, A. P.; Proctor, M. J.; Roe, B. A.; Chen, F.; Pan, H.; Ramser, J.; Lehrach, H.; Reinhardt, R.; McCombie, W. R.; de la Bastide, M.; Dedhia, N.; Blöcker, H.; Hornischer, K.; Nordsiek, G.; Agarwala, R.; Aravind, L.; Bailey, J. A.; Bateman, A.; Batzoglu, S.; Birney, E.; Bork, P.; Brown, D. G.; Burge, C. B.; Cerutti, L.; Chen, H.-C.; Church, D.; Clamp, M.; Copley, R. R.; Doerks, T.; Eddy, S. R.; Eichler, E. E.; Furey, T. S.; Galagan, J.; Gilbert, J. G. R.; Harmon, C.; Hayashizaki, Y.; Haussler, D.; Hermjakob, H.; Hokamp, K.; Jang, W.; Johnson, L. S.; Jones, T. A.; Kasif, S.; Kasprzyk, A.; Kennedy, S.; Kent, W. J.; Kitts, P.; Koonin, E. V.; Korf, I.; Kulp, D.; Lancet, D.; Lowe, T. M.; McLysaght, A.; Mikkelsen, T.; Moran, J. V.; Mulder, N.; Pollara, V. J.; Ponting, C. P.; Schuler, G.; Schultz, J.; Slater, G.; Smit, A. F. A.; Stupka, E.; Szustakowki, J.; Thierry-Mieg, D.; Thierry-Mieg, J.; Wagner, L.; Wallis, J.; Wheeler, R.; Williams, A.; Wolf, Y. I.; Wolfe, K. H.; Yang, S.-P.; Yeh, R.-F.; Collins, F.; Guyer, M. S.; Peterson, J.; Felsenfeld, A.; Wetterstrand, K. A.; Myers, R. M.; Schmutz, J.; Dickson, M.; Grimwood, J.; Cox, D. R.; Olson, M. V.; Kaul, R.; Raymond, C.; Shimizu, N.; Kawasaki, K.; Minoshima, S.; Evans, G. A.; Athanasiou, M.; Schultz, R.; Patrinos, A.; Morgan, M. J.; International Human Genome Sequencing Consortium; Whitehead Institute for Biomedical Research, C. for G. R.; The Sanger Centre;; Washington University Genome Sequencing Center; US DOE Joint Genome Institute;; Baylor College of Medicine Human Genome Sequencing Center;; RIKEN Genomic Sciences Center;; Genoscope and CNRS UMR-8030;; Department of Genome Analysis, I. of M. B.; GTC Sequencing Center;; Beijing Genomics Institute/Human Genome Center;; Multimegabase Sequencing Center, T. I. for S. B.; Stanford Genome Technology Center;; University of Oklahoma's Advanced Center for Genome Technology;; Max Planck Institute for Molecular Genetics;; Cold Spring Harbor Laboratory, L. A. H. G. C.; GBF—German Research Centre for Biotechnology;; *Genome Analysis Group (listed in alphabetical order, also includes individuals listed under other headings);; Scientific management: National Human Genome Research Institute, U. N. I. of H.; Stanford Human Genome Center;; University of Washington Genome Center;; Department of Molecular Biology, K. U. S. of M.; University of Texas Southwestern Medical Center at Dallas;; Office of Science, U. D. of E.; The Wellcome Trust: Initial Sequencing and Analysis of the

Human Genome. *Nature* **2001**, *409* (6822), 860–921.
<https://doi.org/10.1038/35057062>.

(14) Sheynkman, G. M.; Shortreed, M. R.; Cesnik, A. J.; Smith, L. M. Proteogenomics: Integrating Next-Generation Sequencing and Mass Spectrometry to Characterize Human Proteomic Variation. *Annu. Rev. Anal. Chem. Palo Alto Calif* **2016**, *9* (1), 521–545. <https://doi.org/10.1146/annurev-anchem-071015-041722>.

(15) Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels. *Anal. Chem.* **1996**, *68* (5), 850–858. <https://doi.org/10.1021/ac950914h>.

(16) *Sequencing of Proteins and Peptides*; Elsevier, 2011.

(17) Masellis, C. Cryogenic Ion Spectroscopy of Peptides and Glycans, EPFL, 2017. <https://doi.org/10.5075/epfl-thesis-8086>.

(18) Masellis, C.; Khanal, N.; Kamrath, M. Z.; Clemmer, D. E.; Rizzo, T. R. Cryogenic Vibrational Spectroscopy Provides Unique Fingerprints for Glycan Identification. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (10), 2217–2222. <https://doi.org/10.1007/s13361-017-1728-6>.

(19) Werz, D. B.; Ranzinger, R.; Herget, S.; Adibekian, A.; von der Lieth, C.-W.; Seeberger, P. H. Exploring the Structural Diversity of Mammalian Carbohydrates (“Glycospace”) by Statistical Databank Analysis. *ACS Chem. Biol.* **2007**, *2* (10), 685–691. <https://doi.org/10.1021/cb700178s>.

(20) S. Kornfeld, E. Li, I. Tabas, J. The Synthesis of Complex-Type Oligosaccharides. II. Characterization of the Processing Intermediates in the Synthesis of the Complex Oligosaccharide Units of the Vesicular Stomatitis Virus G Protein. **1978**, *253*, 7771–7778.

(21) Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Marth, J. D.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E. Symbol nomenclature for glycan representation. *PROTEOMICS* **2009**, *9* (24), 5398–5399. <https://doi.org/10.1002/pmic.200900708>.

- (22) Harvey, D. J.; Merry, A. H.; Royle, L.; Campbell, M. P.; Dwek, R. A.; Rudd, P. M. Proposal for a Standard System for Drawing Structural Diagrams of N- and O-Linked Carbohydrates and Related Compounds. *Proteomics* **2009**, *9* (15), 3796–3801. <https://doi.org/10.1002/pmic.200900096>.
- (23) Varki, A.; Cummings, R. D.; Aebi, M.; Packer, N. H.; Seeberger, P. H.; Esko, J. D.; Stanley, P.; Hart, G.; Darvill, A.; Kinoshita, T.; Prestegard, J. J.; Schnaar, R. L.; Freeze, H. H.; Marth, J. D.; Bertozzi, C. R.; Etzler, M. E.; Frank, M.; Vliegthart, J. F.; Lütke, T.; Perez, S.; Bolton, E.; Rudd, P.; Paulson, J.; Kanehisa, M.; Toukach, P.; Aoki-Kinoshita, K. F.; Dell, A.; Narimatsu, H.; York, W.; Taniguchi, N.; Kornfeld, S. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology* **2015**, *25* (12), 1323–1324. <https://doi.org/10.1093/glycob/cwv091>.
- (24) Moremen, K. W.; Tiemeyer, M.; Nairn, A. V. Vertebrate Protein Glycosylation: Diversity, Synthesis and Function. *Nat. Rev. Mol. Cell Biol.* **2012**, *13* (7), 448–462. <https://doi.org/10.1038/nrm3383>.
- (25) Stanley, P. What Have We Learned from Glycosyltransferase Knockouts in Mice? *J. Mol. Biol.* **2016**, *428* (16), 3166–3182. <https://doi.org/10.1016/j.jmb.2016.03.025>.
- (26) Kleene, R.; Schachner, M. Glycans and Neural Cell Interactions. *Nat. Rev. Neurosci.* **2004**, *5* (3), 195–208. <https://doi.org/10.1038/nrn1349>.
- (27) Wang, X.; Inoue, S.; Gu, J.; Miyoshi, E.; Noda, K.; Li, W.; Mizuno-Horikawa, Y.; Nakano, M.; Asahi, M.; Takahashi, M.; Uozumi, N.; Ihara, S.; Lee, S. H.; Ikeda, Y.; Yamaguchi, Y.; Aze, Y.; Tomiyama, Y.; Fujii, J.; Suzuki, K.; Kondo, A.; Shapiro, S. D.; Lopez-Otin, C.; Kuwaki, T.; Okabe, M.; Honke, K.; Taniguchi, N. Dysregulation of TGF- β 1 Receptor Activation Leads to Abnormal Lung Development and Emphysema-like Phenotype in Core Fucose-Deficient Mice. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (44), 15791–15796. <https://doi.org/10.1073/pnas.0507375102>.
- (28) Reusch, D.; Tejada, M. L. Fc Glycans of Therapeutic Antibodies as Critical Quality Attributes. *Glycobiology* **2015**, *25* (12), 1325–1334. <https://doi.org/10.1093/glycob/cwv065>.

- (29) Liu, L. Antibody Glycosylation and Its Impact on the Pharmacokinetics and Pharmacodynamics of Monoclonal Antibodies and Fc-Fusion Proteins. *J. Pharm. Sci.* **2015**, *104* (6), 1866–1884. <https://doi.org/10.1002/jps.24444>.
- (30) Zhang, P.; Woen, S.; Wang, T.; Liao, B.; Zhao, S.; Chen, C.; Yang, Y.; Song, Z.; Wormald, M. R.; Yu, C.; Rudd, P. M. Challenges of Glycosylation Analysis and Control: An Integrated Approach to Producing Optimal and Consistent Therapeutic Drugs. *Drug Discov. Today* **2016**, *21* (5), 740–765. <https://doi.org/10.1016/j.drudis.2016.01.006>.
- (31) Marth, J. D.; Grewal, P. K. Mammalian Glycosylation in Immunity. *Nat. Rev. Immunol.* **2008**, *8* (11), 874–887. <https://doi.org/10.1038/nri2417>.
- (32) Nardy, A. F. F. R.; 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*. <https://doi.org/10.3389/fonc.2016.00054>.
- (33) Balzarini, J. Targeting the Glycans of Glycoproteins: A Novel Paradigm for Antiviral Therapy. *Nat. Rev. Microbiol.* **2007**, *5* (8), 583–597. <https://doi.org/10.1038/nrmicro1707>.
- (34) Walls, A. C.; Park, Y.-J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veasley, D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **2020**. <https://doi.org/10.1016/j.cell.2020.02.058>.
- (35) Bieberich, E. Synthesis, Processing, and Function of N-Glycans in N-Glycoproteins. *Adv. Neurobiol.* **2014**, *9*, 47–70. https://doi.org/10.1007/978-1-4939-1154-7_3.
- (36) Dimitrov, D. S. Therapeutic Proteins. *Methods Mol. Biol. Clifton NJ* **2012**, *899*, 1–26. https://doi.org/10.1007/978-1-61779-921-1_1.
- (37) Walsh, G. Biopharmaceutical Benchmarks 2018. *Nat. Biotechnol.* **2018**, *36* (12), 1136–1145. <https://doi.org/10.1038/nbt.4305>.
- (38) Kesik-Brodacka, M. Progress in Biopharmaceutical Development. *Biotechnol. Appl. Biochem.* **2018**, *65* (3), 306–322. <https://doi.org/10.1002/bab.1617>.

- (39) Carter, P. J. Introduction to Current and Future Protein Therapeutics: A Protein Engineering Perspective. *Exp. Cell Res.* **2011**, *317* (9), 1261–1269. <https://doi.org/10.1016/j.yexcr.2011.02.013>.
- (40) Garrison, L.; McDonnell, N. Etanercept: Therapeutic Use in Patients with Rheumatoid Arthritis. *Ann. Rheum. Dis.* **1999**, *58* (Suppl 1), I65–I69.
- (41) Dumont, J.; Euwart, D.; Mei, B.; Estes, S.; Kshirsagar, R. Human Cell Lines for Biopharmaceutical Manufacturing: History, Status, and Future Perspectives. *Crit. Rev. Biotechnol.* **2016**, *36* (6), 1110–1122. <https://doi.org/10.3109/07388551.2015.1084266>.
- (42) Recombinant Protein and Its Expression Systems - Creative BioMart <https://www.creativebiomart.net/resource/articles-recombinant-protein-and-its-expression-systems-365.htm> (accessed May 4, 2020).
- (43) Berger, M.; Kaup, M.; Blanchard, V. Protein Glycosylation and Its Impact on Biotechnology. In *Genomics and Systems Biology of Mammalian Cell Culture*; Hu, W. S., Zeng, A.-P., Eds.; Advances in Biochemical Engineering Biotechnology; Springer: Berlin, Heidelberg, 2012; pp 165–185. https://doi.org/10.1007/10_2011_101.
- (44) Lee, G. M.; Kildegaard, H. F. *Cell Culture Engineering: Recombinant Protein Production*; John Wiley & Sons, 2019.
- (45) Kalantar-Zadeh, K. History of Erythropoiesis-Stimulating Agents, the Development of Biosimilars, and the Future of Anemia Treatment in Nephrology. *Am. J. Nephrol.* **2017**, *45* (3), 235–247. <https://doi.org/10.1159/000455387>.
- (46) Solá, R. J.; Griebenow, K. Glycosylation of Therapeutic Proteins: An Effective Strategy to Optimize Efficacy. *BioDrugs Clin. Immunother. Biopharm. Gene Ther.* **2010**, *24* (1), 9–21. <https://doi.org/10.2165/11530550-000000000-00000>.
- (47) Dicker, M.; Strasser, R. Using Glyco-Engineering to Produce Therapeutic Proteins. *Expert Opin. Biol. Ther.* **2015**, *15* (10), 1501–1516. <https://doi.org/10.1517/14712598.2015.1069271>.

- (48) Gupta, S. K.; Shukla, P. Sophisticated Cloning, Fermentation, and Purification Technologies for an Enhanced Therapeutic Protein Production: A Review. *Front. Pharmacol.* **2017**, *8*. <https://doi.org/10.3389/fphar.2017.00419>.
- (49) Czajkowsky, D. M.; Hu, J.; Shao, Z.; Pleass, R. J. Fc-Fusion Proteins: New Developments and Future Perspectives. *EMBO Mol. Med.* **2012**, *4* (10), 1015–1028. <https://doi.org/10.1002/emmm.201201379>.
- (50) Meibohm, B.; Zhou, H. Characterizing the Impact of Renal Impairment on the Clinical Pharmacology of Biologics. *J. Clin. Pharmacol.* **2012**, *52* (1 Suppl), 54S–62S. <https://doi.org/10.1177/0091270011413894>.
- (51) Ultee, M. E. Therapeutic Fc-Fusion Proteins. *mAbs* **2014**, *6* (4), 810–811. <https://doi.org/10.4161/mabs.28832>.
- (52) Beck, A.; Reichert, J. M. Therapeutic Fc-Fusion Proteins and Peptides as Successful Alternatives to Antibodies. *mAbs* **2011**, *3* (5), 415–416. <https://doi.org/10.4161/mabs.3.5.17334>.
- (53) Pendley, C.; Schantz, A.; Wagner, C. Immunogenicity of Therapeutic Monoclonal Antibodies. *Curr. Opin. Mol. Ther.* **2003**, *5* (2), 172–179.
- (54) Grossman, J.; Chamow, S. M. Etanercept. In *Therapeutic Fc-Fusion Proteins*; John Wiley & Sons, Ltd, 2013; pp 255–282. <https://doi.org/10.1002/9783527675272.ch10>.

Chapter 2 Analysis of N-glycans

This chapter describes currently used techniques for the analysis of N-glycans released from glycoproteins, highlighting some of their advantages and disadvantages. The approach that we have developed in our laboratory and used in this thesis is then presented, highlighting some of the advantages of this new method.

2.1 Separation techniques

Because of the intrinsic complexity of glycan samples comprising numerous compositional and linkage isomers, glycan analysis requires a separation step in combination with a detection/identification step. There are many different combinations of separation and detection techniques that are commonly used. We first describe here the variety of separation techniques followed by the detection/identification techniques that are typically coupled with them, pointing out some of the advantages and disadvantages of the various approaches.

2.1.1 *Capillary electrophoresis*

Capillary electrophoresis is a powerful tool for glycan separation [1], providing relatively high-speed and high resolution [2,3]. It is based on the electrophoretic mobility of an ion under an applied voltage. The electrophoretic mobility is influenced by the number of ionic charges of the molecule and its size. The rate at which the ions move is directly proportional to the applied electric field. Carbohydrate separation by CE is limited, since many glycan species are uncharged, excluding acidic sugars (e.g., *N*-acetylneuraminic acid, glucuronic acid, etc.). Moreover, most of carbohydrate molecules cannot be detected light absorption in the UV range, since they lack a chromophore. Labeling carbohydrates with a fluorophore that carries a charge helps to overcome these limitations and make CE applicable for glycan separation. The separation relies on electrophoretic transport to emphasize differences in charge-to-size ratio.

Several studies have shown that CE enables high-throughput isomer separation and linkage analysis, and the possibility of glycan quantitation [2-5]. Additionally, the CE method offers a considerably reduced cost per sample. However, limitations include necessity of the glycan label to allow its detection, the lack of a structural database, and difficulties to couple with mass spectrometry (MS) detection technique [4,5], which has led to its limited acceptance by the glycomics community.

2.1.2 *Liquid chromatography*

Liquid chromatography is a broadly applicable separation technique, and in combination with tandem MS it is considered the “gold standard” for glycan analysis. Separation is based on the interactions of the sample in a liquid state with the mobile and stationary phases. The stationary phase is placed within the LC column, and a sample with the analyte is injected onto the column then pumped through with a mobile phase flow. As the sample passes through the column, the mixture of glycans has different affinity with the stationary phase that leads to glycan separation and elution from the column at different times, so-called retention times.

The most common stationary phases for released glycan analysis that can be coupled to MS or used with other detection include high-pH anion-exchange (HPAE) [6,7], hydrophilic interaction liquid chromatography (HILIC) [8,9], porous graphitic carbon (PGC) [10–12], and reversed phase (RP)[13,14].

Non-labeled glycans have been commonly studied by using HPAE with pulsed amperometric detection (PAD), which provides high-sensitivity, high-efficiency separations of linkage and positional isomers based on partial deprotonation of hydroxyl groups at high pH >12 [6,15]. Despite the benefits of HPAE, high-salt eluents that are required for these separations are incompatible with many nebulizer-based detectors, such as MS.

High resolution and selectivity to separate native and derivatized glycan isomers are achieved applying HILIC [6,8,9]. Separation is accomplished due to increasing hydrophilicity of the analyte which increases with N-glycan size. The major drawbacks are that sample preparation can be laborious and mobile phases may lead to ion suppression [16].

Using PGC as a stationary phase is a robust method to separate complex mixtures of isobaric and isomeric glycans [10]. However, the implementation of PGC-based separations in studies of glycans has been limited since system-independent retention values have not been established to normalize technical variation [11]. To address this limitation, Ashwood *et al.* applied an internal standard (hydrolyzed dextran) together with Skyline software to reduce retention time and peak area technical variation in PGC-based glycan analyses [12].

Reversed phase LC separates only derivatized sugars on the basis of their hydrophobicity [13,14]. Moreover, the mobile phases used for RP HPLC are highly MS-compatible.

2.1.3 Ion mobility spectrometry

Recently, ion mobility spectrometry (IMS) has become an appealing and rapid alternative to liquid chromatography for glycan analysis. The main advantages of ion mobility (IM) techniques over chromatography are speed and ease of coupling to other orthogonal techniques, particularly MS. Here, the widely-used IMS techniques coupled to MS for glycan studies are presented.

Ion mobility spectrometry measures the mobility, K , of ions in gas phase under an electric field in the presence of a buffer gas. The mobility depends on the ion charge, shape, and size [17]. Ion mobility spectrometry is similar to capillary electrophoresis in its principle of operation, but because it is carried out in the gas phase, the timescale is significantly shorter (high- μ s to ms). Thus, IMS can be a foundation for high-throughput analysis, which is commonly required for studies of complex samples. The mechanism of ion mobility of glycans in the gas phase varies among instrument classes or ion mobility types. The nature of the electric field to propel the ions through the IM cell differs among these types.

Differential/field asymmetric ion mobility spectrometry (DMS/FAIMS) is a spatially-dispersive IM-based technique. It is based on the difference in mobility between high and low electric fields and can be easily combined with MS. Despite its high resolution, collisional cross sections (CCS) cannot be derived from the mobility data. Nevertheless, FAIMS coupled to tandem MS has been proven to separate glycan isomers differing in linkage positions [18].

Drift tube ion mobility spectrometry (DT-IMS) is the simplest form of IMS. It is the only time-dispersive ion-mobility technique where ion collisional cross-sections can be directly derived from the measured ion drift times [19]. Numerous research groups, including ours, performed studies on complex mixtures of glycans employing DT-IMS-MS [20,21]. As an example, DT-IMS-MS has been applied to characterize serum N-linked glycans from 81 individuals to determine variations among patients with different diseases [22]. One of the drawbacks of using DT-IMS is the limited resolution (50-100). Since the resolution is proportional to the square root of drift length, there is only so much one can extend this length in a practical device.

This problem of limited resolution can be overcome with a different type of IMS, which is known as a traveling wave ion mobility spectrometry (TW-IMS) [23,24]. Another advantage is commercial availability of the TW-IMS that made this IM type widely accepted and attractive within the research community. The first commercial TW-IMS instrument (Synapt HDMS) has been introduced by WatersTM in 2006 [24]. Since then many improved versions have been released. Several studies have demonstrated the importance of using both precursor and fragment ions to determine complex glycan structures by coupling tandem MS with TW-IMS on the Synapt instruments [25–27]. Another study by Pagel and coworkers showed the ability of combination of IM with MS to separate and quantify anomers and regioisomers of tri- and penta-saccharides [28]. Unfortunately, the resolution of these instruments is typically ~ 50 , which is not sufficient for separating most glycan isomers. However, WatersTM has recently introduced the SELECT SERIES Cyclic IMS system with a new cyclic geometry [29] in which ions can be circulated through the same drift path multiple times, increasing the separation path length. The new IMS instrument has a resolution more than 900 [30]. Toraño *et al.* applied cyclic IMS in combination with collision-induced dissociation (CID) to study fucosylated glycans and determine precisely their fucosyl positions [31]. Another study by Ujma and coworkers demonstrated separation of anomers and ring-open forms of a series of three pentasaccharides and their mixtures [32].

Smith and co-workers have pioneered a competing approach called Structures for Lossless Ion Manipulations (SLIM). It creates tracks between electrodes on a “sandwich” of PC boards to create extended paths along which ions are separated by

TW-IMS. By using labyrinthine paths for the ions in which they can turn corners, this approach allows a TW-IMS device to be compact and increase its resolution to ~ 1000 , which can resolve miniscule differences in the arrival time distributions of isomeric ions [33,34]. Additionally, SLIM can accumulate ions on the board with minimal loss, which helps to improve the sensitivity [35]. The SLIM-based IM separations have demonstrated the ability to provide additional information about possible conformers of glycans that was previously not accessible [36]. Moreover, recent studies on isotopologues and isotopomers use SLIM separation paths of more than 1 km [37]. The work performed in this thesis implements SLIM-technology for glycan separation in combination with cryogenic infrared spectroscopy.

Trapped ion mobility spectrometry (TIMS) is a field-dispersive type of device that flows gas through an ion trap to separate ions in space and is able to achieve a resolving power of ~ 300 [38]. Bruker has developed a commercial instrument implementing TIMS technology. Pe *et al.* demonstrated an application of TIMS with electron activated dissociation (ExD) tandem MS to separate isomeric glycans [39].

For TIMS and TWIMS techniques, external calibration must be employed to obtain CCS of the molecules.

2.2 Identification techniques

2.2.1 *Nuclear magnetic resonance spectroscopy*

Nuclear magnetic resonance (NMR) spectroscopy is a non-destructive technique for the characterization of the structure and dynamics of molecules that can be applied to glycans. It provides structural information, including anomeric configuration, stereochemistry, linkage position, and glycan sequence using samples in solution [40]. Moreover, NMR is applicable to study glycan-protein interactions [41]. Its main limitation is that it requires large quantities and high purity of the biological sample, which are often not available. Additionally, it might need some chemical preparation, including tagging with a metal or isotopic labeling. Interpretation of the NMR data requires high-level quantum calculations [42]. Thus, it is difficult to envision using NMR as an online analytical technique for glycan identification.

2.2.2 *Mass spectrometry*

Mass spectrometry, which generates ions and detects them according to their mass-to-charge (m/z) ratios, has been extensively used for glycan analysis alone or in combination with other methods. Glycans are typically produced as gas-phase ions by either Electrospray Ionization (ESI) or Matrix-Assisted Laser Desorption/Ionization (MALDI).

Given that MS method alone cannot distinguish isomeric species, tandem MS techniques (MS^n) are used to distinguish the isomers and obtain detailed information about glycan structure. Tandem MS techniques include low- and high-energy CID [43–48], electron-based dissociation [49–54], infrared multiphoton dissociation (IRMPD) [55–57] or ultraviolet photodissociation (UVPD) methods [58–61].

While CID is the most common dissociation technique used for glycan analysis, N-glycans can undergo structural rearrangements under low energy CID conditions that can complicate structural analysis. In particular, some studies reported unusual rearrangement of fucose and xylose in during the CID experiments of N-glycans [45,46]. In order to prevent fucose migration, glycans are studied as metal adducts [62]. For N-glycans, high-energy CID conditions are applied to produce cross-ring fragments with the aim of obtaining information on the glycosidic linkages [47]. In this case, permethylation of the hydroxyl groups is typically used [48].

Electron-based dissociation techniques are complementary to CID [44–49]. For example, electron capture dissociation (ECD) enhances cross-ring fragmentation over glycosidic bond cleavage, which preserves linkage information [44,45]. While ECD generates fragmentation in the positive ion mode for analysis of neutral glycans, electron detachment dissociation (EDD) is widely used in the negative mode to study acidic glycans [49]. However, these dissociation techniques are implemented on Orbitrap or FT-ICR MS instruments. Currently, the software and databases needed to analyze the extensive data produced by these dissociation techniques does not yet exist.

Photodissociation techniques, such as IRMPD and UVPD are widely employed to obtain branching and linkage information. The group of Lebrilla applied IRMPD with FTICR-MS for the structural characterization of complex sugars [50,51]. Reilly

and coworkers reported extensive fragmentation patterns obtained upon UVPD at 157 nm of permethylated N-glycans [58,59]. They also demonstrated the ability to generate significant cross-ring cleavages that is an advantage of UVPD methods to characterize oligosaccharides.

In summary, tandem MS using a variety of dissociation techniques shows the ability to reveal many structural details of chemically modified glycans. However, the data interpretation is time-consuming and currently lacks extensive databases. Hyphenated techniques that combine any of aforementioned separation methods (see Section 2.1) with MS detection play a crucial role in glycan analysis.

2.2.3 *Spectroscopy*

Gas-phase spectroscopy is a technique that has been gaining popularity in the field of glycomics. Simons and coworkers were one of the first groups to study carbohydrates spectroscopically [63]. They labeled small neutral carbohydrates with an aromatic chromophore and applied UV-IR double resonance spectroscopy in a supersonic molecular beam. The absorption patterns allowed conformational assignments of small sugars.

Gas-phase spectroscopy opens a new door to identification of glycan isomers. Several groups demonstrated this by adding a spectroscopic dimension to existing techniques for isomer separation [64,65]. For example, Hernandez and coworkers coupled differential IMS-MS (DIMS) with IRMPD spectroscopy to study anomeric monosaccharides [64]. Polfer *et al.* combined IMRPD with high-resolution MS to identify disaccharides [66]. Compagnon and coworkers measured IR fingerprints of fragments and compared them to reference standards of GAGs [67]. Later, they discovered that glycans retain memory of the anomericity of the glycosidic bond upon fragmentation [68] to form C_n fragments, which holds great promise for carbohydrate sequencing. In a subsequent publication, the group of Compagnon applied the LC-MS-IRMPD approach to mixtures of monosaccharides and disaccharides [69].

Using IRMPD spectroscopy of room temperature ions to identify relatively large N-glycans runs into the problem of spectral congestion, where peak overlaps and broadening limit the spectral information one can obtain. Moreover, larger glycans

might exist as multiple conformers at room temperature with different absorption patterns.

Cryogenic-IR spectroscopy has been applied to overcome this limitation. Pagel and coworkers have employed a combination of MS and ultra-cold IR spectroscopy in helium nanodroplets to study oligosaccharides [70,71]. Recorded spectra exhibit unambiguous vibrational bands characteristic of each sugar that serve as a unique fingerprint. This study shows a clear advantage of using low temperatures for structural identification. However, it is difficult to imagine commercializing the use of superfluid helium droplets in an analytical technique. As described below, our group has introduced approaches for glycan analysis by combining either MS or IMS-MS with cold messenger tagging spectroscopy [20,72–75].

2.3 Glycan databases

Due to the abundance of glycans in nature and the development of new techniques to study them, considerable amounts of information about glycans is available, and it is important to structure this information to make it accessible and usable. The scientific community would greatly benefit from a universal database (DB) containing all data on glycans and their analysis. Moreover, the database would require a powerful and user-friendly software to access the information. Eventually, this universal database would greatly assist in deciphering the glycome of humans and understanding the relation with other types of life.

Several databases currently exist and can be divided into two groups based on the type of the provided information. The first group includes databases with experimental data for glycan characterization and identification. Among these are GlycoBase 3.2 with LC, MS, and exoglycosidase sequencing data of approximately 650 N- and O- linked glycan structures [76]. Another well-known database is UniCarb-DB, which is hosted by Swiss Institute for Bioinformatics [77] and contains LC-MS/MS data of glycans. The UniCarb KnowledgeBase (UniCarbKB) is a platform that helps researchers access UniCarb-DB, GlycoBase, and GlycosuiteDB, together with the EUROCarbDB software [78,79]. Glycomob is another glycan repository that provides IM-MS data based on precursor glycan mass and CCS [80].

GLYCOSCIENCES.DB contains NMR shifts and 3D structures [81]. Carbohydrate Structure Database (CSDB) provides NMR data [82].

Other databases are devoted to biochemical information, such as glycotransferases, glycan-binding proteins (lectins). The Consortium of Functional Glycomics (CFG) DB contains data from glycan array screening and glycan profiling [83]. Carbohydrate-Active Enzyme (CAZy) DB provides the data on the enzymes for glycan sequencing [84]. The Kyoto Encyclopedia of Genes and Genomes (KEGG) GLYCAN DB gives access to glycan structures together with glyco-enzyme data [85].

Currently existing databases do not include gas-phase IR spectra of glycans. This is because the use of spectroscopic techniques for glycan identification is still relatively new and requires a high degree of specialization to operate the state-of-art instruments. Our ultimate goal is to create a database of mass, mobility, and IR fingerprint spectra for known N-glycans and make our technique widely available.

Over the last decade, the growth of data for glycan analysis has increased in complexity. However, the content of the glycan databases and the representations of glycan sequences in the aforementioned databases are diverse and not standardized. Moreover, the information that one can obtain from the different databases can partially overlap. Additionally, the access to the data is not always user-friendly. Nowadays, data collection about a single glycan structure is a cumbersome procedure. There is a clear need for an easy and standard tool to access and collect all the variety of the information about glycans.

2.4 Our new approach to glycan analysis

All the existing techniques for glycan analysis listed above have their advantages and limitations. In this thesis, we propose a multidimensional approach for monitoring released glycans that combines ultrahigh-resolution IMS and cryogenic vibrational spectroscopy to study N-glycans.

In our approach, ions of glycan adducts are generated by electrospray, their arrival time distribution is measured in a SLIM module, and ions are then injected into a cryogenic ion trap, where we cool them and tag them with N_2 . We then measure

a linear vibrational spectrum by monitoring the loss of the N_2 tag as the molecule absorbs infrared photons. The depletion of the tagged species is detected in a time-of-flight mass spectrometer (TOF-MS) as a function of the infrared frequency.

Comprehensive high-level quantum mechanical approaches to determine glycan structures have been typically limited to disaccharides, mostly because of long computational times and the vast conformational space of such complex molecular ions. Recently, molecular dynamics calculations on large glycans have become tractable yielding insight into their structure and isomeric complexity [70]. Since we focus only on the covalent structure of glycans, including all isomeric forms, rather than comparing our measurements with calculated spectra, we take a totally different approach. Our goal is to identify glycan molecule by recording a mass, arrival time distribution, and vibrational fingerprint spectrum and comparing them later with entries in a database that we will create. This means that our determination of primary structure of large native glycans does not depend on accurate, high-level, quantum-chemical calculations. Thus, this work represents an advance in the analysis of N-linked glycans cleaved from biopharmaceutical proteins that could eventually be used as a tool for monitoring biopharmaceutical glycoforms. The detailed method description, the proposed database approach, and the obtained results will be presented in the next chapters of this thesis.

References:

- (1) Lu, G.; Crieffeld, C. L.; Gattu, S.; Veltri, L. M.; Holland, L. A. Capillary Electrophoresis Separations of Glycans. *Chem. Rev.* **2018**, *118* (17), 7867–7885. <https://doi.org/10.1021/acs.chemrev.7b00669>.
- (2) Oefner, P. J.; Chiesa, C. Capillary Electrophoresis of Carbohydrates. *Glycobiology* **1994**, *4* (4), 397–412. <https://doi.org/10.1093/glycob/4.4.397>.
- (3) Gaunitz, S.; Nagy, G.; Pohl, N. L. B.; Novotny, M. V. Recent Advances in the Analysis of Complex Glycoproteins. *Anal. Chem.* **2017**, *89* (1), 389–413. <https://doi.org/10.1021/acs.analchem.6b04343>.
- (4) Huffman, J. E.; Pučić-Baković, M.; Klarić, L.; Hennig, R.; Selman, M. H. J.; Vučković, F.; Novokmet, M.; Krištić, J.; Borowiak, M.; Muth, T.; Polašek, O.; Razdorov, G.; Gornik, O.; Plomp, R.; Theodoratou, E.; Wright, A. F.; Rudan, I.; Hayward, C.; Campbell, H.; Deelder, A. M.; Reichl, U.; Aulchenko, Y. S.; Rapp, E.; Wuhrer, M.; Lauc, G. Comparative Performance of Four Methods for High-Throughput Glycosylation Analysis of Immunoglobulin G in Genetic and Epidemiological Research. *Mol. Cell. Proteomics MCP* **2014**, *13* (6), 1598–1610. <https://doi.org/10.1074/mcp.M113.037465>.
- (5) Mechref, Y.; Novotny, M. V. Glycomic Analysis by Capillary Electrophoresis-Mass Spectrometry. *Mass Spectrom. Rev.* **2009**, *28* (2), 207–222. <https://doi.org/10.1002/mas.20196>.
- (6) Szabo, Z.; Thayer, J. R.; Agroskin, Y.; Lin, S.; Liu, Y.; Srinivasan, K.; Saba, J.; Viner, R.; Huhmer, A.; Rohrer, J.; Reusch, D.; Harfouche, R.; Khan, S. H.; Pohl, C. In-Depth Analyses of Native N-Linked Glycans Facilitated by High-Performance Anion Exchange Chromatography-Pulsed Amperometric Detection Coupled to Mass Spectrometry. *Anal. Bioanal. Chem.* **2017**, *409* (12), 3089–3101. <https://doi.org/10.1007/s00216-017-0248-3>.
- (7) Grey, C.; Edebrink, P.; Krook, M.; Jacobsson, S. P. Development of a High Performance Anion Exchange Chromatography Analysis for Mapping of

Oligosaccharides. *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* **2009**, *877* (20–21), 1827–1832. <https://doi.org/10.1016/j.jchromb.2009.05.003>.

(8) Alpert, A. J. Hydrophilic-Interaction Chromatography for the Separation of Peptides, Nucleic Acids and Other Polar Compounds. *J. Chromatogr. A* **1990**, *499*, 177–196. [https://doi.org/10.1016/S0021-9673\(00\)96972-3](https://doi.org/10.1016/S0021-9673(00)96972-3).

(9) Wuhrer, M.; de Boer, A. R.; Deelder, A. M. Structural Glycomics Using Hydrophilic Interaction Chromatography (HILIC) with Mass Spectrometry. *Mass Spectrom. Rev.* **2009**, *28* (2), 192–206. <https://doi.org/10.1002/mas.20195>.

(10) Kolarich, D.; Windwarder, M.; Alagesan, K.; Altmann, F. Isomer-Specific Analysis of Released N-Glycans by LC-ESI MS/MS with Porous Graphitized Carbon. *Methods Mol. Biol. Clifton NJ* **2015**, *1321*, 427–435. https://doi.org/10.1007/978-1-4939-2760-9_29.

(11) Abrahams, J. L.; Campbell, M. P.; Packer, N. H. Building a PGC-LC-MS N-Glycan Retention Library and Elution Mapping Resource. *Glycoconj. J.* **2018**, *35* (1), 15–29. <https://doi.org/10.1007/s10719-017-9793-4>.

(12) Ashwood, C.; Pratt, B.; MacLean, B. X.; Gundry, R. L.; Packer, N. H. Standardization of PGC-LC-MS-Based Glycomics for Sample Specific Glycotyping. *Analyst* **2019**, *144* (11), 3601–3612. <https://doi.org/10.1039/C9AN00486F>.

(13) Vreeker, G. C. M.; Wuhrer, M. Reversed-Phase Separation Methods for Glycan Analysis. *Anal. Bioanal. Chem.* **2017**, *409* (2), 359–378. <https://doi.org/10.1007/s00216-016-0073-0>.

(14) Higel, F.; Demelbauer, U.; Seidl, A.; Friess, W.; Sörgel, F. Reversed-Phase Liquid-Chromatographic Mass Spectrometric N-Glycan Analysis of Biopharmaceuticals. *Anal. Bioanal. Chem.* **2013**, *405* (8), 2481–2493. <https://doi.org/10.1007/s00216-012-6690-3>.

(15) Ruhaak, L. R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A. M.; Wuhrer, M. Glycan Labeling Strategies and Their Use in Identification and Quantification. *Anal. Bioanal. Chem.* **2010**, *397* (8), 3457–3481. <https://doi.org/10.1007/s00216-010-3532-z>.

- (16) Royle, L.; Campbell, M. P.; Radcliffe, C. M.; White, D. M.; Harvey, D. J.; Abrahams, J. L.; Kim, Y.-G.; Henry, G. W.; Shadick, N. A.; Weinblatt, M. E.; Lee, D. M.; Rudd, P. M.; Dwek, R. A. HPLC-Based Analysis of Serum N-Glycans on a 96-Well Plate Platform with Dedicated Database Software. *Anal. Biochem.* **2008**, *376* (1), 1–12. <https://doi.org/10.1016/j.ab.2007.12.012>.
- (17) Creaser, C. S.; Griffiths, J. R.; Bramwell, C. J.; Noreen, S.; Hill, C. A.; Thomas, C. L. P. Ion Mobility Spectrometry: A Review. Part 1. Structural Analysis by Mobility Measurement. *Analyst* **2004**, *129* (11), 984–994. <https://doi.org/10.1039/B404531A>.
- (18) Lane, C. S.; McManus, K.; Widdowson, P.; Flowers, S. A.; Powell, G.; Anderson, I.; Campbell, J. L. Separation of Sialylated Glycan Isomers by Differential Mobility Spectrometry. *Anal. Chem.* **2019**, *91* (15), 9916–9924. <https://doi.org/10.1021/acs.analchem.9b01595>.
- (19) May, J. C.; Dodds, J. N.; Kurulugama, R. T.; Stafford, G. C.; Fjeldsted, J. C.; McLean, J. A. Broudscale Resolving Power Performance of a High Precision Uniform Field Ion Mobility-Mass Spectrometer. *Analyst* **2015**, *140* (20), 6824–6833. <https://doi.org/10.1039/C5AN00923E>.
- (20) Masellis C., Khanal N., Kamrath M.Z., Clemmer D.E., C.; Rizzo T.R. Cryogenic Vibrational Spectroscopy Provides Unique Fingerprints for Glycan Identification. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (10), 2217–2222. <https://doi.org/10.1007/s13361-017-1728-6>.
- (21) Zhu, F.; Lee, S.; Valentine, S. J.; Reilly, J. P.; Clemmer, D. E. Mannose7 Glycan Isomer Characterization by IMS-MS/MS Analysis. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (12), 2158–2166. <https://doi.org/10.1007/s13361-012-0491-y>.
- (22) Isailovic, D.; Plasencia, M. D.; Gaye, M. M.; Stokes, S. T.; Kurulugama, Ruwan. T.; Pungpapong, V.; Zhang, M.; Kyselova, Z.; Goldman, R.; Mechref, Y.; Novotny, M. V.; Clemmer, D. E. Delineating Diseases by IMS-MS Profiling of Serum N-Linked Glycans. *J. Proteome Res.* **2012**, *11* (2), 576–585. <https://doi.org/10.1021/pr200777u>.

- (23) Shvartsburg, A. A.; Smith, R. D. Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Anal. Chem.* **2008**, *80* (24), 9689–9699. <https://doi.org/10.1021/ac8016295>.
- (24) Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H. Applications of a Travelling Wave-Based Radio-Frequency-Only Stacked Ring Ion Guide. *Rapid Commun. Mass Spectrom.* **2004**, *18* (20), 2401–2414. <https://doi.org/10.1002/rcm.1641>.
- (25) Both, P.; Green, A. P.; Grey, C. J.; Šardžik, R.; Voglmeir, J.; Fontana, C.; Austeri, M.; Rejzek, M.; Richardson, D.; Field, R. A.; Widmalm, G.; Flitsch, S. L.; Eyers, C. E. Discrimination of Epimeric Glycans and Glycopeptides Using IM-MS and Its Potential for Carbohydrate Sequencing. *Nat. Chem.* **2013**, *6* (1), 65–74. <https://doi.org/10.1038/nchem.1817>.
- (26) Harvey, D. J.; Scarff, C. A.; Edgeworth, M.; Pagel, K.; Thalassinou, K.; Struwe, W. B.; Crispin, M.; Scrivens, J. H. Travelling-Wave Ion Mobility Mass Spectrometry and Negative Ion Fragmentation of Hybrid and Complex N-Glycans. *J. Mass Spectrom.* **2016**, *51* (11), 1064–1079. <https://doi.org/10.1002/jms.3828>.
- (27) Harvey, D. J.; Struwe, W. B. Structural Studies of Fucosylated N-Glycans by Ion Mobility Mass Spectrometry and Collision-Induced Fragmentation of Negative Ions. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (6), 1179–1193. <https://doi.org/10.1007/s13361-018-1950-x>.
- (28) Hofmann, J.; Hahm, H. S.; Seeberger, P. H.; Pagel, K. Identification of Carbohydrate Anomers Using Ion Mobility-Mass Spectrometry. *Nature* **2015**, *526* (7572), 241–244. <https://doi.org/10.1038/nature15388>.
- (29) Giles, K.; Ujma, J.; Wildgoose, J.; Pringle, S.; Richardson, K.; Langridge, D.; Green, M. A Cyclic Ion Mobility-Mass Spectrometry System. *Anal. Chem.* **2019**, *91* (13), 8564–8573. <https://doi.org/10.1021/acs.analchem.9b01838>.
- (30) Ropartz, D.; Fanuel, M.; Ujma, J.; Palmer, M.; Giles, K.; Rogniaux, H. Structure Determination of Large Isomeric Oligosaccharides of Natural Origin through Multipass and Multistage Cyclic Traveling-Wave Ion Mobility Mass

Spectrometry. *Anal. Chem.* **2019**, *91* (18), 12030–12037. <https://doi.org/10.1021/acs.analchem.9b03036>.

(31) Toraño, J. S.; Gagarinov, I. A.; Vos, G. M.; Broszeit, F.; Srivastava, A. D.; Palmer, M.; Langridge, J. I.; Aizpurua-Olaizola, O.; Somovilla, V. J.; Boons, G.-J. Ion-Mobility Spectrometry Can Assign Exact Fucosyl Positions in Glycans and Prevent Misinterpretation of Mass-Spectrometry Data After Gas-Phase Rearrangement. *Angew. Chem. Int. Ed.* **2019**, *58* (49), 17616–17620. <https://doi.org/10.1002/anie.201909623>.

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

(33) Deng, L.; Ibrahim, Y. M.; Hamid, A. M.; Garimella, S. V. B.; Webb, I. K.; Zheng, X.; Prost, S. A.; Sandoval, J. A.; Norheim, R. V.; Anderson, G. A.; Tolmachev, A. V.; Baker, E. S.; Smith, R. D. Ultra-High Resolution Ion Mobility Separations Utilizing Traveling Waves in a 13 m Serpentine Path Length Structures for Lossless Ion Manipulations Module. *Anal. Chem.* **2016**, *88* (18), 8957–8964. <https://doi.org/10.1021/acs.analchem.6b01915>.

(34) Hamid, A. M.; Ibrahim, Y. M.; Garimella, S. V. B.; Webb, I. K.; Deng, L.; Chen, T.-C.; Anderson, G. A.; Prost, S. A.; Norheim, R. V.; Tolmachev, A. V.; Smith, R. D. Characterization of Traveling Wave Ion Mobility Separations in Structures for Lossless Ion Manipulations. *Anal. Chem.* **2015**, *87* (22), 11301–11308. <https://doi.org/10.1021/acs.analchem.5b02481>.

(35) Deng, L.; Garimella, S. V. B.; Hamid, A. M.; Webb, I. K.; Attah, I. K.; Norheim, R. V.; Prost, S. A.; Zheng, X.; Sandoval, J. A.; Baker, E. S.; Ibrahim, Y. M.; Smith, R. D. Compression Ratio Ion Mobility Programming (CRIMP) Accumulation and Compression of Billions of Ions for Ion Mobility-Mass Spectrometry Using Traveling Waves in Structures for Lossless Ion Manipulations (SLIM). *Anal. Chem.* **2017**, *89* (12), 6432–6439. <https://doi.org/10.1021/acs.analchem.7b00189>.

- (36) Nagy, G.; Attah, I. K.; Garimella, S. V. B.; Tang, K.; Ibrahim, Y. M.; Baker, E. S.; Smith, R. D. Unraveling the Isomeric Heterogeneity of Glycans: Ion Mobility Separations in Structures for Lossless Ion Manipulations. *Chem. Commun.* **2018**, 54 (83), 11701–11704. <https://doi.org/10.1039/C8CC06966B>.
- (37) Wojcik, R.; Nagy, G.; Attah, Isaac. K.; Webb, I. K.; Garimella, S. V. B.; Weitz, K. K.; Hollerbach, A.; Monroe, M. E.; Ligare, M. R.; Nielson, F. F.; Norheim, R. V.; Renslow, R. S.; Metz, T. O.; Ibrahim, Y. M.; Smith, R. D. SLIM Ultrahigh Resolution Ion Mobility Spectrometry Separations of Isotopologues and Isotopomers Reveal Mobility Shifts Due to Mass Distribution Changes. *Anal. Chem.* **2019**, 91 (18), 11952–11962. <https://doi.org/10.1021/acs.analchem.9b02808>.
- (38) Michelmann, K.; Silveira, J. A.; Ridgeway, M. E.; Park, M. A. Fundamentals of Trapped Ion Mobility Spectrometry. *J. Am. Soc. Mass Spectrom.* **2015**, 26 (1), 14–24. <https://doi.org/10.1021/jasms.8b04886>.
- (39) Pu, Y.; Ridgeway, M. E.; Glaskin, R. S.; Park, M. A.; Costello, C. E.; Lin, C. Separation and Identification of Isomeric Glycans by Selected Accumulation-Trapped Ion Mobility Spectrometry-Electron Activated Dissociation Tandem Mass Spectrometry. *Anal. Chem.* **2016**, 88 (7), 3440–3443. <https://doi.org/10.1021/acs.analchem.6b00041>.
- (40) Duus, J. Ø.; Gotfredsen, C. H.; Bock, K. Carbohydrate Structural Determination by NMR Spectroscopy: Modern Methods and Limitations. *Chem. Rev.* **2000**, 100 (12), 4589–4614. <https://doi.org/10.1021/cr990302n>.
- (41) Bewley, C. A.; Shahzad-ul-Hussan, S. Characterizing Carbohydrate-Protein Interactions by NMR. *Biopolymers* **2013**, 99 (10). <https://doi.org/10.1002/bip.22329>.
- (42) Fadda, E.; Woods, R. J. Molecular Simulations of Carbohydrates and Protein–Carbohydrate Interactions: Motivation, Issues and Prospects. *Drug Discov. Today* **2010**, 15 (15), 596–609. <https://doi.org/10.1016/j.drudis.2010.06.001>.
- (43) Lemoine, J.; Fournet, B.; Despeyroux, D.; Jennings, K. R.; Rosenberg, R.; de Hoffmann, E. Collision-Induced Dissociation of Alkali Metal Cationized and Permethylated Oligosaccharides: Influence of the Collision Energy and of the Collision

Gas for the Assignment of Linkage Position. *J. Am. Soc. Mass Spectrom.* **1993**, *4* (3), 197–203. [https://doi.org/10.1016/1044-0305\(93\)85081-8](https://doi.org/10.1016/1044-0305(93)85081-8).

(44) Li, B.; An, H. J.; Hedrick, J. L.; Lebrilla, C. B. Collision-Induced Dissociation Tandem Mass Spectrometry for Structural Elucidation of Glycans. In *Glycomics: Methods and Protocols*; Packer, N. H., Karlsson, N. G., Eds.; Methods in Molecular BiologyTM; Humana Press: Totowa, NJ, 2009; pp 133–145. https://doi.org/10.1007/978-1-59745-022-5_10.

(45) Hecht, E. S.; Loziuk, P. L.; Muddiman, D. C. Xylose Migration During Tandem Mass Spectrometry of N-Linked Glycans. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (4), 729–732. <https://doi.org/10.1007/s13361-016-1588-5>.

(46) Brüll, L. P.; Kováčik, V.; Thomas-Oates, J. E.; Heerma, W.; Haverkamp, J. Sodium-Cationized Oligosaccharides Do Not Appear to Undergo “internal Residue Loss” Rearrangement Processes on Tandem Mass Spectrometry. *Rapid Commun. Mass Spectrom. RCM* **1998**, *12* (20), 1520–1532. [https://doi.org/10.1002/\(SICI\)1097-0231\(19981030\)12:20<1520::AID-RCM336>3.0.CO;2-W](https://doi.org/10.1002/(SICI)1097-0231(19981030)12:20<1520::AID-RCM336>3.0.CO;2-W).

(47) Stephens, E.; Maslen, S. L.; Green, L. G.; Williams, D. H. Fragmentation Characteristics of Neutral N-Linked Glycans Using a MALDI-TOF/TOF Tandem Mass Spectrometer. *Anal. Chem.* **2004**, *76* (8), 2343–2354. <https://doi.org/10.1021/ac030333p>.

(48) Morelle, W.; Faïd, V.; Michalski, J.-C. Structural Analysis of Permethylated Oligosaccharides Using Electrospray Ionization Quadrupole Time-of-Flight Tandem Mass Spectrometry and Deutero-Reduction. *Rapid Commun. Mass Spectrom.* **2004**, *18* (20), 2451–2464. <https://doi.org/10.1002/rcm.1640>.

(49) Adamson, J. T.; Håkansson, K. Electron Capture Dissociation of Oligosaccharides Ionized with Alkali, Alkaline Earth, and Transition Metals. *Anal. Chem.* **2007**, *79* (7), 2901–2910. <https://doi.org/10.1021/ac0621423>.

(50) Zhao, C.; Xie, B.; Chan, S.-Y.; Costello, C. E.; O’Connor, P. B. Collisionally Activated Dissociation and Electron Capture Dissociation Provide Complementary Structural Information for Branched Permethylated

Oligosaccharides. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (1), 138–150. <https://doi.org/10.1016/j.jasms.2007.10.022>.

(51) Yu, X.; Jiang, Y.; Chen, Y.; Huang, Y.; Costello, C. E.; Lin, C. Detailed Glycan Structural Characterization by Electronic Excitation Dissociation. *Anal. Chem.* **2013**, *85* (21). <https://doi.org/10.1021/ac402886q>.

(52) Han, L.; Costello, C. E. Electron Transfer Dissociation of Milk Oligosaccharides. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (6), 997–1013. <https://doi.org/10.1021/jasms.8b04041>.

(53) Zhou, W.; Håkansson, K. Electron Capture Dissociation of Divalent Metal-Adducted Sulfated N-Glycans Released from Bovine Thyroid Stimulating Hormone. *J. Am. Soc. Mass Spectrom.* **2013**, *24* (11), 1798–1806. <https://doi.org/10.1021/jasms.8b04414>.

(54) Wolff, J. J.; Amster, I. J.; Chi, L.; Linhardt, R. J. Electron Detachment Dissociation of Glycosaminoglycan Tetrasaccharides. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (2), 234–244. <https://doi.org/10.1021/jasms.8b02854>.

(55) Li, B.; An, H. J.; Hedrick, J. L.; Lebrilla, C. B. Infrared Multiphoton Dissociation Mass Spectrometry for Structural Elucidation of Oligosaccharides. In *Glycomics: Methods and Protocols*; Packer, N. H., Karlsson, N. G., Eds.; Methods in Molecular BiologyTM; Humana Press: Totowa, NJ, 2009; pp 23–35. https://doi.org/10.1007/978-1-59745-022-5_2.

(56) Zhang, J.; Schuboth, K.; Li, B.; Russell, S.; Lebrilla, C. B. Infrared Multiphoton Dissociation of O-Linked Mucin-Type Oligosaccharides. *Anal. Chem.* **2005**, *77* (1), 208–214. <https://doi.org/10.1021/ac0489824>.

(57) Pikulski, M.; Hargrove, A.; Shabbir, S. H.; Anslyn, E. V.; Brodbelt, J. S. Sequencing and Characterization of Oligosaccharides Using Infrared Multiphoton Dissociation and Boronic Acid Derivatization in a Quadrupole Ion Trap. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (12), 2094–2106. <https://doi.org/10.1021/jasms.8b02829>.

(58) Devakumar, A.; Mechref, Y.; Kang, P.; Novotny, M. V.; Reilly, J. P. Laser-Induced Photofragmentation of Neutral and Acidic Glycans inside an Ion-Trap

Mass Spectrometer. *Rapid Commun. Mass Spectrom. RCM* **2007**, *21* (8), 1452–1460. <https://doi.org/10.1002/rcm.2981>.

(59) Devakumar, A.; Mechref, Y.; Kang, P.; Novotny, M. V.; Reilly, J. P. Identification of Isomeric N-Glycan Structures by Mass Spectrometry with 157 Nm Laser-Induced Photofragmentation. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (7), 1027–1040. <https://doi.org/10.1016/j.jasms.2008.03.005>.

(60) Ko, B. J.; Brodbelt, J. S. Comparison of Glycopeptide Fragmentation by Collision Induced Dissociation and Ultraviolet Photodissociation. *Int. J. Mass Spectrom.* **2015**, *377* (1), 385–392. <https://doi.org/10.1016/j.ijms.2014.07.032>.

(61) Morrison, K. A.; Clowers, B. H. Differential Fragmentation of Mobility-Selected Glycans via Ultraviolet Photodissociation and Ion Mobility-Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (6), 1236–1241. <https://doi.org/10.1021/jasms.8b05563>.

(62) Lettow, M.; Mucha, E.; Manz, C.; Thomas, D. A.; Marianski, M.; Meijer, G.; von Helden, G.; Pagel, K. The Role of the Mobile Proton in Fucose Migration. *Anal. Bioanal. Chem.* **2019**, *411* (19), 4637–4645. <https://doi.org/10.1007/s00216-019-01657-w>.

(63) Talbot, F. O.; Simons, J. P. Sugars in the Gas Phase: The Spectroscopy and Structure of Jet-Cooled Phenyl β -D-Glucopyranoside. *Phys. Chem. Chem. Phys.* **2002**, *4* (15), 3562–3565. <https://doi.org/10.1039/B204132D>.

(64) Hernandez, O.; Isenberg, S.; Steinmetz, V.; Glush, G. L.; Maitre, P. Probing Mobility-Selected Saccharide Isomers: Selective Ion–Molecule Reactions and Wavelength-Specific IR Activation. *J. Phys. Chem. A* **2015**, *119* (23), 6057–6064. <https://doi.org/10.1021/jp511975f>.

(65) Mucha, E.; Stuckmann, A.; Marianski, M.; Struwe, W. B.; Meijer, G.; Pagel, K. In-Depth Structural Analysis of Glycans in the Gas Phase. *Chem. Sci.* **2019**, *10* (5), 1272–1284. <https://doi.org/10.1039/C8SC05426F>.

(66) Polfer, N. C.; Valle, J. J.; Moore, D. T.; Oomens, J.; Eyler, J. R.; Bendiak, B. Differentiation of Isomers by Wavelength-Tunable Infrared Multiple-

Photon Dissociation-Mass Spectrometry: Application to Glucose-Containing Disaccharides. *Anal. Chem.* **2006**, *78* (3), 670–679. <https://doi.org/10.1021/ac0519458>.

(67) Schindler, B.; Renois-Predelus, G.; Bagdadi, N.; Melizi, S.; Barnes, L.; Chambert, S.; Allouche, A.-R.; Compagnon, I. MS/IR, a New MS-Based Hyphenated Method for Analysis of Hexuronic Acid Epimers in Glycosaminoglycans. *Glycoconj. J.* **2017**, *34* (3), 421–425. <https://doi.org/10.1007/s10719-016-9741-8>.

(68) Schindler, B.; Barnes, L.; Renois, G.; Grey, C.; Chambert, S.; Fort, S.; Flitsch, S.; Loison, C.; Allouche, A.-R.; Compagnon, I. Anomeric Memory of the Glycosidic Bond upon Fragmentation and Its Consequences for Carbohydrate Sequencing. *Nat. Commun.* **2017**, *8* (1), 973. <https://doi.org/10.1038/s41467-017-01179-y>.

(69) Schindler, B.; Laloy-Borgna, G.; Barnes, L.; Allouche, A.-R.; Bouju, E.; Dugas, V.; Demesmay, C.; Compagnon, I. 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.* **2018**, *90* (20), 11741–11745. <https://doi.org/10.1021/acs.analchem.8b02801>.

(70) Mucha, E.; González Flórez, A. I.; Marianski, M.; Thomas, D. A.; Hoffmann, W.; Struwe, W. B.; Hahm, H. S.; Gewinner, S.; Schöllkopf, W.; Seeberger, P. H.; von Helden, G.; Pagel, K. Glycan Fingerprinting via Cold-Ion Infrared Spectroscopy. *Angew. Chem. Int. Ed.* **2017**, *56* (37), 11248–11251. <https://doi.org/10.1002/anie.201702896>.

(71) Lettow, M.; Grabarics, M.; Mucha, E.; Thomas, D. A.; Polewski, Ł.; Freyse, J.; Rademann, J.; Meijer, G.; von Helden, G.; Pagel, K. IR Action Spectroscopy of Glycosaminoglycan Oligosaccharides. *Anal. Bioanal. Chem.* **2020**, *412* (3), 533–537. <https://doi.org/10.1007/s00216-019-02327-7>.

(72) Khanal, N.; Masellis, C.; Kamrath, M. Z.; Clemmer, D. E.; Rizzo, T. R. Glycosaminoglycan Analysis by Cryogenic Messenger-Tagging IR Spectroscopy

Combined with IMS-MS. *Anal. Chem.* **2017**, *89* (14), 7601–7606. <https://doi.org/10.1021/acs.analchem.7b01467>.

(73) Khanal N., Masellis C., Kamrath M.Z., Rizzo T.R.,; Clemmer D.E. Cryogenic IR Spectroscopy Combined with Ion Mobility Spectrometry for the Analysis of Human Milk Oligosaccharides. *Analyst* **2018**, *143*, 1846–1852. <https://doi.org/10.1039/C8AN00230D>.

(74) Ben Faleh, A.; Warnke, S.; Rizzo, T. R. Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures. *Anal. Chem.* **2019**, *91* (7), 4876–4882. <https://doi.org/10.1021/acs.analchem.9b00659>.

(75) Dyukova, I.; Carrascosa, E.; Pellegrinelli, R. P.; Rizzo, T. R. Combining Cryogenic Infrared Spectroscopy with Selective Enzymatic Cleavage for Determining Glycan Primary Structure. *Anal. Chem.* **2020**, *92* (2), 1658–1662. <https://doi.org/10.1021/acs.analchem.9b04776>.

(76) Campbell, M. P.; Royle, L.; Radcliffe, C. M.; Dwek, R. A.; Rudd, P. M. GlycoBase and AutoGU: Tools for HPLC-Based Glycan Analysis. *Bioinformatics* **2008**, *24* (9), 1214–1216. <https://doi.org/10.1093/bioinformatics/btn090>.

(77) UniCarb-DB <https://unicarb-db.expasy.org/> (accessed Jun 20, 2020).

(78) Cooper, C. A.; Harrison, M. J.; Wilkins, M. R.; Packer, N. H. GlycoSuiteDB: A New Curated Relational Database of Glycoprotein Glycan Structures and Their Biological Sources. *Nucleic Acids Res.* **2001**, *29* (1), 332–335.

(79) Yuriev, E.; Agostino, M. *Structural and Computational Glycobiology: Immunity and Infection*; Frontiers Media SA, 2015.

(80) Struwe, W. B.; Pagel, K.; Benesch, J. L. P.; Harvey, D. J.; Campbell, M. P. GlycoMob: An Ion Mobility-Mass Spectrometry Collision Cross Section Database for Glycomics. *Glycoconj. J.* **2016**, *33* (3), 399–404. <https://doi.org/10.1007/s10719-015-9613-7>.

- (81) Böhm, M.; Bohne-Lang, A.; Frank, M.; Loss, A.; Rojas-Macias, M. A.; Lütteke, T. Glycosciences.DB: An Annotated Data Collection Linking Glycomics and Proteomics Data (2018 Update). *Nucleic Acids Res.* **2019**, *47* (D1), D1195–D1201. <https://doi.org/10.1093/nar/gky994>.
- (82) Russian CSDB <http://csdb.glycoscience.ru/> (accessed Jun 20, 2020).
- (83) Owens, R. J.; Nettleship, J. E. *Functional and Structural Proteomics of Glycoproteins*; Springer Science & Business Media, 2010.
- (84) Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P. M.; Henrissat, B. The Carbohydrate-Active Enzymes Database (CAZy) in 2013. *Nucleic Acids Res.* **2014**, *42* (Database issue), D490–495. <https://doi.org/10.1093/nar/gkt1178>.
- (85) Hashimoto, K.; Goto, S.; Kawano, S.; Aoki-Kinoshita, K. F.; Ueda, N.; Hamajima, M.; Kawasaki, T.; Kanehisa, M. KEGG as a Glycome Informatics Resource. *Glycobiology* **2006**, *16* (5), 63R–70R. <https://doi.org/10.1093/glycob/cwj010>.

Chapter 3 Experimental approach

This chapter describes the general principles of the gas-phase experimental method to analyze native glycans that includes the details of SLIM-based ion mobility module and the principles of cryogenic messenger tagging spectroscopy.

3.1 Ion mobility spectrometry coupled with cryogenic ion spectroscopy

3.1.1 Overview of the instrument

Our experiments have been performed on a home-built instrument (Figure 3.1) that couples a TW SLIM-based IM module with a cryogenic ion trap and a time-of-flight (TOF) mass spectrometer.

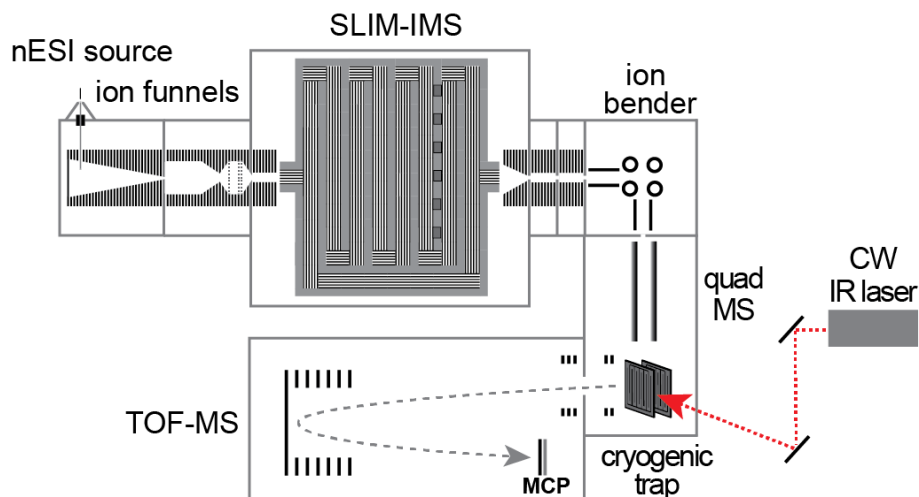


Figure 3.1 Schematic overview of the current experimental set-up [1].

Glycan ions were generated by a nanoelectrospray ionization (nESI) source using commercial borosilicate nano-ESI emitters (ThermoFisher, Germany). The emitters were gold/palladium-coated by the manufacturer to allow electrical contact. The high voltage (~ 1800 V) used to sustain the nESI was applied through the stainless steel

alligator clip by a high voltage power supply (BERTAN Associates Inc., USA). Sample flow rate was facilitated by a backpressure on the liquid in the needle holder. Pressures up to 1 bar above atmospheric pressure were typically applied to achieve optimal spraying conditions. After each needle was loaded and placed in the holder, the needle position, voltage, and pressure were adjusted for maximum ion signal. The generated glycan ions are introduced into the first vacuum stage through a heated flared stainless steel capillary (100°C -150 °C) [2]. After exiting the capillary, the ions enter a dual ion funnel trap (IFT) (Masstech, USA), which was used to focus and accumulate ions in front of a dual grid assembly. Ions were typically accumulated for 100 or 200 ms and then released as short intense packets of ~ 150 μ s duration. The RF frequency and amplitude applied to the IFT was 950 kHz and ~ 200 V_{pp} respectively, and the pressure in the IFT trapping region was typically 4 mbar. The ions exiting the trapping region move through a 2.5 cm long converging region of the IFT before being transferred through ring electrode guide into a TW IMS device based on the SLIM technology originally developed by Smith and coworkers [3–5].

After drifting through the TW SLIM module, the ions are focused and guided using a series of 3D-printed ring electrode ion guides and hexapoles through two differentially pumped regions with pressures 10^{-2} and 10^{-5} mbar [6]. We use an electrostatic steering-lens system to deflect unwanted ions and select the ions of interest with a certain drift time. After moving through the last hexapole guide, the ions are turned 90° by a DC quadrupole bender that facilitates separation from neutral molecules.

The ions are then guided by an octupole to a quadrupole mass filter. For the measurement of IMS arrival time distributions (ATDs), the mass-selected ions passing through the quadrupole mass filter are detected using a channeltron.

To measure vibrational spectra, the channeltron is moved out of the beam path and mobility- and mass-selected ions are injected into a planar, cryogenic trap (Figure 3.2) [7,8]. Trap consists of two sets of 16 gold RF electrodes deposited on a printed circuit board (PCB), each set being connected to opposite phases of the RF. A pair of these PC-board electrodes serve to confine the ions in the z-direction. Four DC plates surround the RF electrodes and confine the ions in the x- and y-directions. The pair of PCBs are mounted onto a copper frame and sandwiched in between two copper

plates, which are attached to the cold head of a helium cryostat (Sumimoto, Germany). The cold trap is electrically isolated from the cold head by sapphire plates.

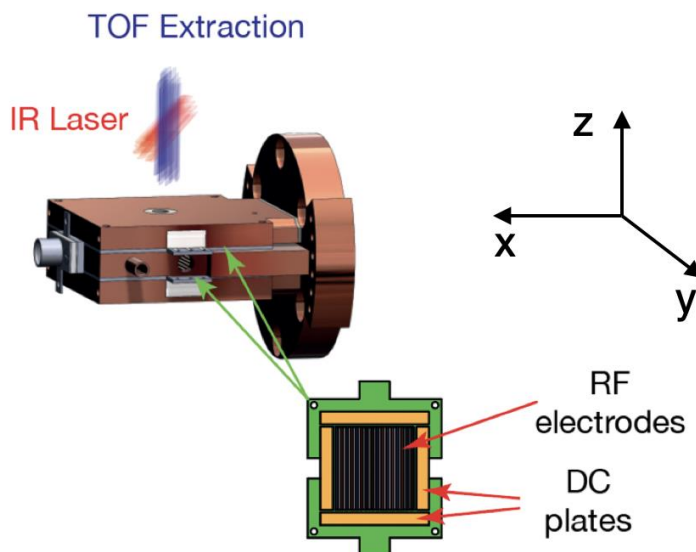


Figure 3.2 Schematic representation of the planar, cryogenic trap and PCB electrodes.

Upon entering the planar trap, the ions are cooled by collisions with a cold buffer gas composed of a He/N₂ (90:10) mixture. The gas is pulsed into the trap, which is maintained at 39 K, using a pulsed valve (Parker Series 9, Germany). At low temperatures, one or more nitrogen molecules attach to the glycan ions, serving as a “messenger tag” for detecting photon absorption. The temperature of the trap is kept constant and monitored using a Lake Shore 350 PID Controller (Lake Shore Cryotronics, Inc; USA) and a cartridge heater.

In this thesis, two laser systems were used to acquire cryogenic infrared (IR) spectra. In early experiments, the tagged ions were irradiated with an IR pulse from a tunable optical parametric oscillator/amplifier system (OPO/OPA) (Laser Vision, USA) pumped by a Nd:YAG laser (Innolas, Germany). Before irradiation with an IR pulse, ions are equilibrated for 50-90 ms with the cold buffer gas. After irradiation, ions are extracted through slots between the electrodes on the upper PCB into the

time-of-flight mass spectrometer (TOF) (Jordan TOF Inc., USA). Extraction is achieved by applying ~ 3.9 kV to the RF electrodes of the bottom PCB board, ~ 2.8 kV to the RF electrodes of the top PCB, and ~ 1.2 kV to the extraction plate on the top (Figure 3.2). Resonant absorption of IR photons by tagged ions followed by intramolecular vibrational energy redistribution leads to the evaporation of the tag(s), which is measured as a depletion of ion signal at the mass of the tagged glycan species. A plot of the number of tagged ions as a function of the laser wavenumber provides a vibrational fingerprint by which we can identify a given glycan. The linewidth of the IR light delivered by the OPO/OPA system is ~ 3 cm^{-1} , and the IR pulse width is 10 ns. In order to increase the IR spectral resolution and acquisition speed, a continuous wave (CW) $\text{Cr}^{2+}/\text{ZnS}/\text{ZnSe}$ mid-IR fiber-pumped laser (IPG Photonics, Germany) was installed, delivering IR linewidth of $\sim 0.3\text{-}2$ cm^{-1} . For spectroscopy with this laser, the ions in the cryogenic trap were subjected to 0.2 W of IR light for 50 ms and subsequently ejected into the TOF mass spectrometer. A chopper system (MC2000B-EC, Thorlabs, Inc., Germany) synchronized with the main experimental trigger is used to produce 50 ms pulses from the CW laser output. The TOF data is acquired using a WaveSurfer MXs-B Oscilloscope (Teledyne LeCroy SA, Switzerland), transferred to a PC, and processed using in-house control software written in Labview.

3.1.2 *SLIM-based ion mobility module*

One of the key components of our experimental approach is to use high-resolution ion mobility separation of glycans using SLIM technology. The SLIM technology is a type of a travelling wave ion mobility that employs a “sandwich” of electrodes patterned on two mirrored printed circuit boards (PCBs). A basic unit of a SLIM module is shown in Figure 3.3.

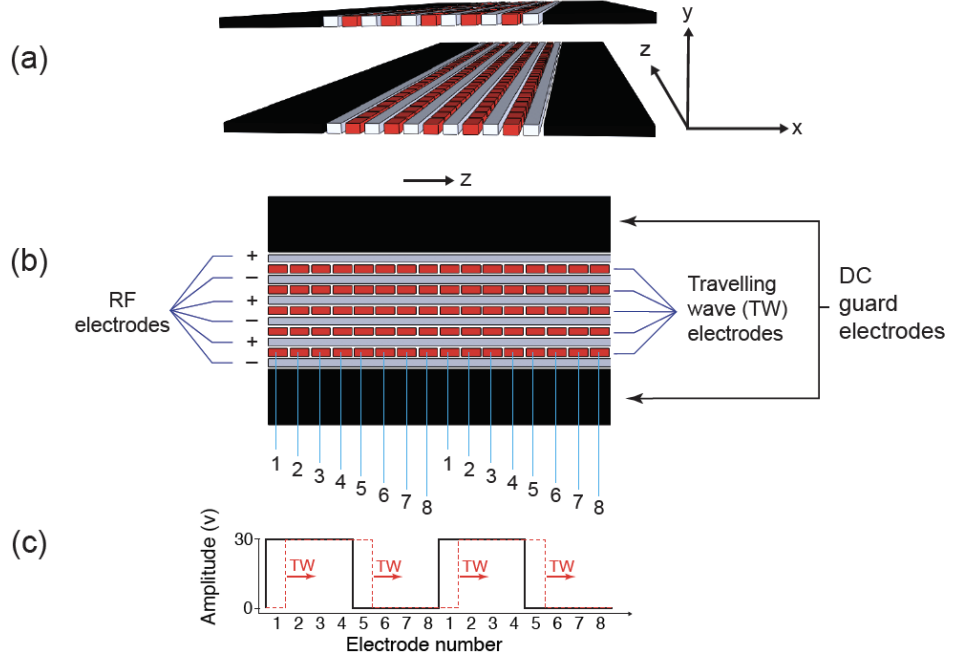


Figure 3.3 Schematic representation of the basic unit of a SLIM module unit (adapted from Ref. 4). a) Two mirrored PCBs that create channels to separate ions; b) SLIM board indicating the application of DC and RF potentials; c) travelling wave potential.

The potentials generated by the two PCBs together make up the SLIM IM channel used to separate ions (Figure 3.3 a). In order to generate the electric fields between the boards, DC and RF potentials are applied to three types of electrodes, the DC guard, RF, and travelling wave (TW) electrodes. The DC voltage applied to the guard electrodes enables the required ion confinement in the x-direction. In addition to these guards, each PCB contains eleven tracks, six of which are RF electrodes (shown in grey in Figure 3.3 b), while five other tracks (shown in red) are TW electrodes. The RF electrodes carry alternate phases of an RF potential and confine the ions in the y-direction. The TW electrodes create a travelling wave that moves ions in the z-direction. The traveling square-well potential is generated by simultaneous application of a DC potential to four sequential electrodes while the next four electrodes are kept at ground potential. The TW then propagates by changing the potential of the first electrode from high to low and that of the fifth electrode from low to high (Figure 3.3 c). This sequence is then shifted one electrode at a time in the

direction of ion motion in a repeating manner across all electrode sets within the TW tracks.

In our initial implementation of the SLIM technique, we used $15\text{ cm} \times 15\text{ cm}$ PCBs to reach a single-pass path length of 1.8 m [9]. The initial design included a serpentine path with 90° turns and a T-shaped switch that enables cycling ions over the same drift path multiple times or ejecting them from the SLIM device for detection [10]. Ion ejection from the board is achieved by applying 40-45 V DC potential to the blocking electrode and the TW potential to the exit electrode (see inset, Figure 3.4). In order to cycle ions, the DC potential on the blocking electrode is switched back to the TW potential and the exit electrode is set to a 45 V DC potential.

This design allows us to achieve high IMS resolution by propelling ions through a 1.8 m serpentine path, and this resolution can be further increased by cycling ions through this path multiple times (Figure 3.4). The resolution scales with \sqrt{n} , where n represents the number of cycles [6, 11].

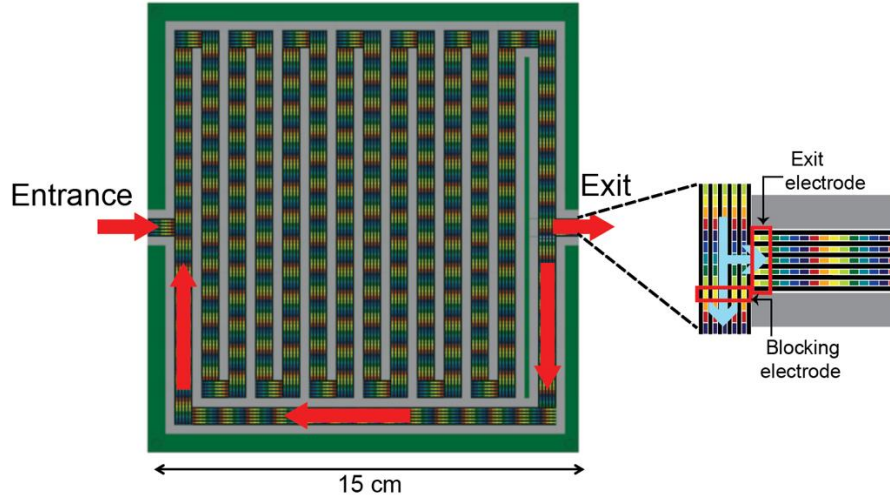


Figure 3.4 Schematic representation of serpentine SLIM board with the option to cycle ions multiple times through the same track before ejecting them for detection or spectroscopy. Adapted from Refs. 1 and 6.

A second SLIM design used in our experiments has two sections that incorporate ion trapping regions (Figure 3.5) onto the board [1]. The separation region is similar

our initial design apart from a decrease in path length from 1.8 m to 1.5 m. The trap sections allow selection and storage of mobility-separated ions. This region is useful for increasing the sensitivity by accumulating several IMS-selected ion packets from the IFT in the on-board trap before detection and spectroscopic analysis. This signal enrichment process for the ions of interest is performed in the following manner. Ion packets (130 μ s wide) are released from the IFT at a repetition rate of 5 Hz to 10 Hz, depending upon the time window that is needed to mobility-separate the ions. After 1.2 m of separation path, the ions of interest are selectively diverted to the on-board trap. This is done by applying a 50 V bias to the blocking electrodes (inset, Figure 3.5) and the TW potential to the trap entrance electrodes, which guide the ions of interest into the trap. Additionally, to anneal conformational distributions and reassure that the ATDs are reproducible, we apply ‘harsh’ and ‘normal’ conditions to operate the trap. Under normal conditions the bias potential of both the trap and normal serpentine track region is maintained at 40 V. Under harsh conditions, 3ms before the stored ions exit the trap, the bias potential of the all electrodes of the trap (incl. DC guard, TW, RF, entrance electrodes) is raised by 65 V, while that of the exit electrode is raised by 100 V. Additionally, a potential difference of 25 V between the trap and normal serpentine track region leads to collisions with the buffer gas, thus activating the ions of interest upon being released from the trap to follow the normal serpentine track. The ions then can either undergo additional cycles of mobility separation or sent to the cryogenic trap for IR spectroscopy. For signal enrichment purposes, ions from the next IFT pulse can be separated on the 1.2 m drift path, selected, and added to the same on-board trap, which typically results in doubling of the number of mobility selected ions of interest. While the on-board trap starts to overfill after 3 IFT packets, an increase in signal is still observed using up to 6 IFT ion packets. The optimal TW amplitude and speed in the SLIM-IMS region were found experimentally for each glycan. Typical IMS parameters used for the experiments were: RF frequency 885 kHz, RF amplitude 150 Vpp, TW amplitude 20 V, TW speed 500 m/s, drift gas pressure 3 mbar (He).

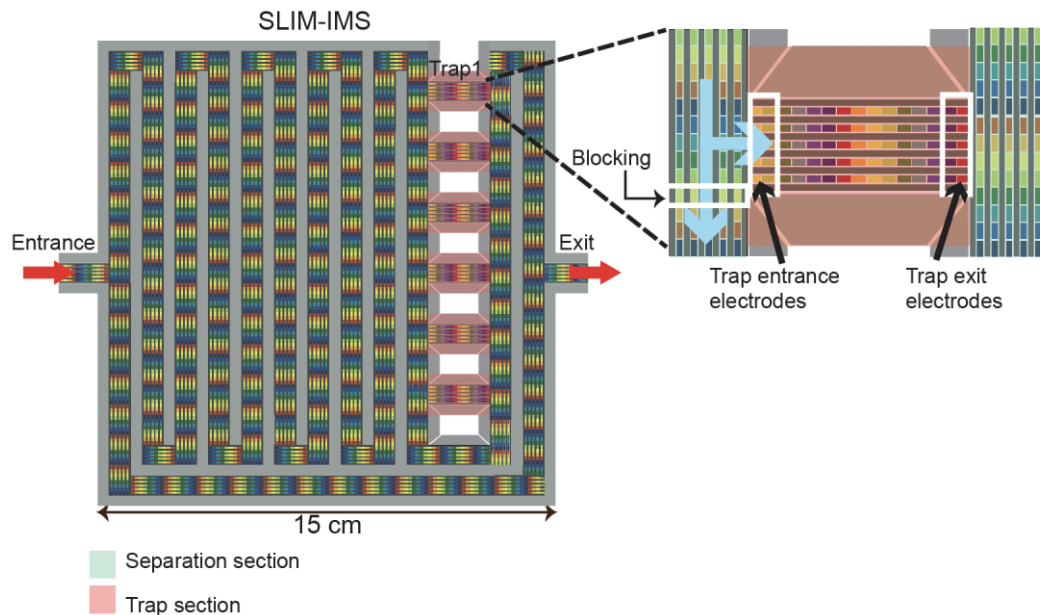


Figure 3.5 Layout of the updated SLIM-IMS device. Left: scheme of the SLIM board. Right: detailed layout of the on-board trap used for signal enrichment. Adapted from Ref. 1.

3.1.3 Principles of cryogenic messenger-tagging spectroscopy

The application of messenger-tagging IR spectroscopy was initially reported by Okumura and Lee in 1985 [12]. Johnson and coworkers at Yale university further developed this technique over many years [13–15].

The fundamental principles of messenger tagging spectroscopy are following: the ions of interest, in our case glycans, form a complex at cryogenic temperatures with one or several chemically inert gas molecules, used as messengers or tags, through weak van der Waals interactions. Upon resonant absorption of a single infrared photon by the analyte ion, the energy is redistributed and the tag is then detached. The ratio of tagged to untagged species is then monitored by a mass spectrometer.

In most applications of messenger-tagging spectroscopy the goal is to obtain vibrational spectra that resemble that of the bare molecule. In this case, one wants to use a light species that will cause minimal perturbation of the IR spectrum of the parent molecule. However, in our case, we simply seek to establish spectroscopic fingerprints by which to identify glycans, and it is not important if the spectrum

resembles that of the untagged molecule. Thus, the main requirement to choose the messenger molecule is to find something that attaches efficiently but not so strongly as to remain after IR excitation. While helium might seem to be an ideal messenger molecule from spectroscopic point of view, it condenses onto molecules at temperatures $\sim 4\text{K}$, which is not practical for an analytical technique. Hydrogen (H_2) and Deuterium (D_2) can form tagged species more easily than He, however the mass shifts by H_2 (2 Da), or D_2 (4 Da), are relatively small. If the N-glycan ions are multiply charged, the mass-shift of the tagged species with respect to the parent is even lower than for singly-charged species in mass spectrum. These smaller, more weakly bound species may lose their tag upon collisional activation during extraction into the time-of-flight mass spectrometer. Heavier tags, such as nitrogen (N_2) (28 Da) and argon (Ar) (40 Da) are thus preferable for our application, given their higher polarizability and their low cost. Moreover, the trap temperature needed to form complexes with N_2 ($\sim 39\text{K}$) is higher, which makes it a good compromise for our experiment. One factor that complicates these experiments is that the tag molecules can freeze out on the walls of the trap. To overcome this, the N_2 tag molecules are usually added as a small percentage of a He gas mix. Helium is used to collisionally cool the ions, while N_2 is easily condensed onto the molecules *via* three-body collisions.

We record a spectrum as a depletion of the tagged ion complexes. Upon absorption of an IR photon, the intensity of untagged ions increases while the tagged ion signal decreases. The signal is normalized following:

$$\textit{spectrum} = \frac{\Sigma \textit{tagged ion complexes}}{\Sigma \textit{tagged ion complexes} + \textit{bare parent ion}} ; (\textit{Equation 3.1})$$

The normalization is used to compensate fluctuations in the ion intensity from the ion source and helps to increase S/N.

A typical recorded mass spectrum using the TOF (Figure 3.6) exhibits clear separation the glycan of interest (blue) from the tagged complexes (green).

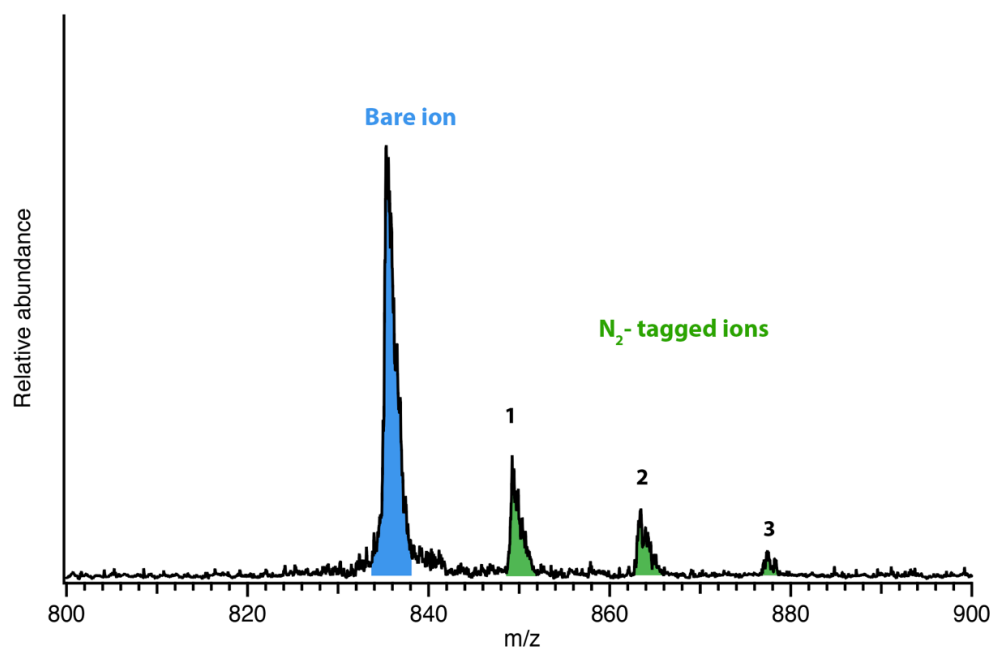


Figure 3.6 Typical mass spectrum acquired with LabView software. The blue-colored peak corresponds to the untagged glycan of interest (bare parent ion). The green-colored peaks correspond to the tagged glycan ion complexes with the molecules.

References:

- (1) Bansal, P.; Yatsyna, V.; AbiKhodr, A. H.; Warnke, S.; Ben Faleh, A.; Yalovenko, N.; Wysocki, V. H.; Rizzo, T. R. Using SLIM-Based IMS-IMS Together with Cryogenic Infrared Spectroscopy for Glycan Analysis. *Anal. Chem.* **2020**. <https://doi.org/10.1021/acs.analchem.0c01265>.
- (2) Pauly, M.; Sroka, M.; Reiss, J.; Rinke, G.; Albarghash, A.; Vogelgesang, R.; Hahne, H.; Kuster, B.; Sesterhenn, J.; Kern, K.; Rauschenbach, S. A Hydrodynamically Optimized Nano-Electrospray Ionization Source and Vacuum Interface. *Analyst* **2014**, *139* (8), 1856–1867. <https://doi.org/10.1039/C3AN01836A>.
- (3) Deng, L.; Ibrahim, Y. M.; Hamid, A. M.; Garimella, S. V. B.; Webb, I. K.; Zheng, X.; Prost, S. A.; Sandoval, J. A.; Norheim, R. V.; Anderson, G. A.; Tolmachev, A. V.; Baker, E. S.; Smith, R. D. Ultra-High Resolution Ion Mobility Separations Utilizing Traveling Waves in a 13 m Serpentine Path Length Structures for Lossless Ion Manipulations Module. *Anal. Chem.* **2016**, *88* (18), 8957–8964. <https://doi.org/10.1021/acs.analchem.6b01915>.
- (4) Hamid, A. M.; Ibrahim, Y. M.; Garimella, S. V. B.; Webb, I. K.; Deng, L.; Chen, T.-C.; Anderson, G. A.; Prost, S. A.; Norheim, R. V.; Tolmachev, A. V.; Smith, R. D. Characterization of Traveling Wave Ion Mobility Separations in Structures for Lossless Ion Manipulations. *Anal. Chem.* **2015**, *87* (22), 11301–11308. <https://doi.org/10.1021/acs.analchem.5b02481>.
- (5) 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.; Smith, R. D. Achieving High Resolution Ion Mobility Separations Using Traveling Waves in Compact Multiturn Structures for Lossless Ion Manipulations. *Anal. Chem.* **2016**, *88* (18), 8949–8956. <https://doi.org/10.1021/acs.analchem.6b01914>.
- (6) Gordon, S. D. S.; Osterwalder, A. 3D-Printed Beam Splitter for Polar Neutral Molecules. *Phys. Rev. Appl.* **2017**, *7* (4), 044022. <https://doi.org/10.1103/PhysRevApplied.7.044022>.

- (7) Masson, A.; Williams, E. R.; Rizzo, T. R. Molecular Hydrogen Messengers Can Lead to Structural Infidelity: A Cautionary Tale of Protonated Glycine. *J. Chem. Phys.* **2015**, *143* (10), 104313. <https://doi.org/10.1063/1.4930196>.
- (8) Masson, A.; Kamrath, M. Z.; Perez, M. A. S.; Glover, M. S.; Rothlisberger, U.; Clemmer, D. E.; Rizzo, T. R. Infrared Spectroscopy of Mobility-Selected H⁺-Gly-Pro-Gly-Gly (GPGG). *J. Am. Soc. Mass Spectrom.* **2015**, *26* (9), 1444–1454. <https://doi.org/10.1007/s13361-015-1172-4>.
- (9) Ben Faleh, A.; Warnke, S.; Rizzo, T. R. Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures. *Anal. Chem.* **2019**, *91* (7), 4876–4882. <https://doi.org/10.1021/acs.analchem.9b00659>.
- (10) Garimella, S. V. B.; Ibrahim, Y. M.; Webb, I. K.; Ipsen, A. B.; Chen, T.-C.; Tolmachev, A. V.; Baker, E. S.; Anderson, G. A.; Smith, R. D. Ion Manipulations in Structures for Lossless Ion Manipulations (SLIM): Computational Evaluation of a 90° Turn and a Switch. *The Analyst* **2015**, *140* (20), 6845–6852. <https://doi.org/10.1039/c5an00844a>.
- (11) Shvartsburg, A. A.; Smith, R. D. Fundamentals of Traveling Wave Ion Mobility Spectrometry. *Anal. Chem.* **2008**, *80* (24), 9689–9699. <https://doi.org/10.1021/ac8016295>.
- (12) Okumura, M.; Yeh, L. I.; Lee, Y. T. The Vibrational Predissociation Spectroscopy of Hydrogen Cluster Ions. *J. Chem. Phys.* **1985**, *83* (7), 3705–3706. <https://doi.org/10.1063/1.449127>.
- (13) Bailey, C. G.; Kim, J.; Dessent, C. E. H.; Johnson, M. A. Vibrational Predissociation Spectra of I[−] · (H₂O): Isotopic Labels and Weakly Bound Complexes with Ar and N₂. *Chem. Phys. Lett.* **1997**, *269* (1), 122–127. [https://doi.org/10.1016/S0009-2614\(97\)00254-6](https://doi.org/10.1016/S0009-2614(97)00254-6).
- (14) Kamrath, M. Z.; Relph, R. A.; Guasco, T. L.; Leavitt, C. M.; Johnson, M. A. Vibrational Predissociation Spectroscopy of the H₂-Tagged Mono- and Dicarboxylate Anions of Dodecanedioic Acid. *Int. J. Mass Spectrom.* **2011**, *300* (2–3), 91–98. <https://doi.org/10.1016/j.ijms.2010.10.021>.

(15) Kamrath, M. Z.; Garand, E.; Jordan, P. A.; Leavitt, C. M.; Wolk, A. B.; Van Stipdonk, M. J.; Miller, S. J.; Johnson, M. A. Vibrational Characterization of Simple Peptides Using Cryogenic Infrared Photodissociation of H₂-Tagged, Mass-Selected Ions. *J. Am. Chem. Soc.* **2011**, *133* (16), 6440–6448. <https://doi.org/10.1021/ja200849g>.

Chapter 4 Analyzing glycans cleaved from a biotherapeutic protein using ultrahigh-resolution ion mobility spectrometry together with cryogenic ion spectroscopy *

This chapter demonstrates that combining ultrahigh-resolution IMS with cryogenic vibrational spectroscopy can provide characteristic fingerprints for N-linked glycans that can be used as a rapid and sensitive means to monitor them. We also discuss how this approach might be used more generally as a tool for glycan analysis of biotherapeutics. Finally, we provide the results of the mathematical methods used to quantify the similarity between IR spectra of standard and released glycans.

4.1 Sample preparation

4.1.1 *Sample preparation of standards*

Synthetically derived N-glycans standards (estimated to be > 85% pure) were purchased from Dextra Laboratories (UK). The standards were then analyzed without any additional purification. HPLC grade solvents were used. Ultrapure water was obtained from a Milli-Q Integral system. All standards were reconstituted in water/acetonitrile (70:30) to a concentration of 80 μ M prior to analysis. The stock solutions were further diluted using water/acetonitrile (70:30) to obtain 5-20 μ M analyte solutions. We added 3-5 μ L of 0.1% formic acid to enhance generation of the protonated N-glycan species.

4.1.2 *N-linked glycans release from Etanercept and clean-up*

Etanercept was kindly provided by Johnathan Paz Montoya from the EPFL Proteomics Facility. It was expressed in CHO cell lines. PNGase F (recombinant) was purchased from Roche (Basel, Switzerland).

The entire sample preparation procedure is shown schematically (Figure 4.1).

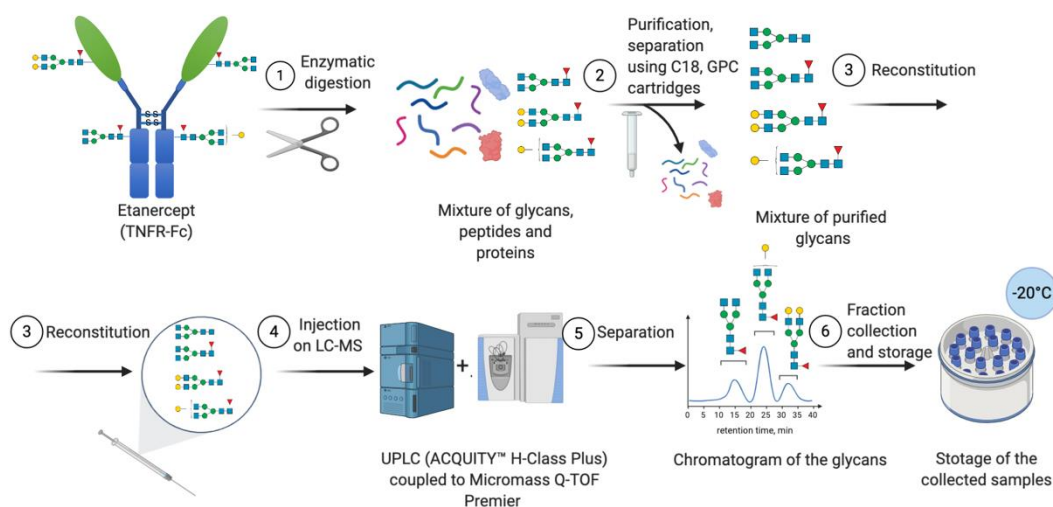


Figure 4.1 *Schematic representation of the sample preparation and fraction collection of the released N-glycans. All the steps are described in detail below.*

In step 1, N-glycans were released from Etanercept using PNGase F. Two 300 µg portions of the protein were reconstituted in 20 mM sodium phosphate buffer (pH 7.2) to a final concentration of 1 µg/µl. The enzymatic digestion was followed by overnight incubation at 37 °C.

In step 2, we first employed C18 cartridges (ThermoFisher, Germany) to separate N-linked glycans from O-glycopeptides, proteins and other contaminants. The procedure was following:

- a) We conditioned a C18 cartridge with 1 ml of methanol (x3) and then with 1 ml of 5% acetic acid (AcOH) (x3).
- b) We loaded the glycan sample onto the C18 column and rinsed the sample container with 100 µL 5% acetic acid 2 times to make sure all the sample was

transferred. This step allows to wash off the non-glycan contaminants.

c) We eluted N-glycans with 1 ml 5% AcOH (x3).

d) We dried the resulting N-glycan sample using speed vacuum (SpeedVac, Eppendorf). This step allows post elution work-up.

In order to remove salts and detergents, we then applied porous graphitic carbon (PGC) cartridges (ThermoFisher, Germany). The steps in the procedure were similar to the one for C18 cartridges, differing in solution composition.

a) Conditioning of the PGC cartridge with 1 ml ACN (x3), followed by 1 ml of 60% ACN (x3).

b) Equilibration of the cartridge with 1 ml of water (x3).

c) Loading of the glycan mixture onto the PGC cartridge, either directly after deglycosylation with PNGase F or after cleanup with C18 cartridges.

d) Washing the glycan mixture twice with 1 ml water, discarding the filtrate.

e) Elution of glycans with 100 μ l of 40% ACN/60% 100 mM ammonium formate pH 4.5.

f) Drying samples using a SpeedVac.

The combined dried residues were then reconstituted in 50 μ l distilled water/acetonitrile (70:30) prior to analysis (step 3). We monitor the efficiency of the enzymatic digestion (step 4) with a UPLC (ACQUITY™ H-Class Plus, Waters, UK) coupled to a Micromass Q-TOF Premier (Waters, UK). In step 5, the digested sample was further separated on an XBridge Glycan BEH Amide Column (3X150 mm, 2.5 μ m) (Waters) at 0.4 mL/min at 60°C. The glycans were eluted by following linear gradients: 22% mobile phase A (ammonium formate buffer, 100mM) and 78% mobile phase B (ACN) for 38.5 min. The mobile phase A was increased from 22% to 44% after 1 min, then from 44% to 100% after an additional 1 min. In the final wash step, mobile phase A was held at 100% for 5 min. Prior to analysis, the column was equilibrated by running 22% mobile phase A and 78% mobile phase B for 2 min. The auto-sampler was kept at 8°C. In the final step (6), each eluted glycan was collected

using a Waters Fraction Collector III, evaporated down, and reconstituted to a total volume of 1 ml. Assuming no loss during the sample preparation, we estimate a final maximum concentration of 0.33 μM for the least abundant glycan (G0) and 7 μM for the most abundant (G0F). Sample solutions were stored at -20°C .

4.2 Comparison of the released sugars with their respective standards

The glycans released from a therapeutic protein Etanercept (TNFR-Fc, see Chapter 1, Section 1.5.1 for details) studied in this work are shown in Figure 4.2. They span the range of relative abundance from 22% for G0F to 1% for G0 [1,2]. The released glycans have been separated and purified prior to analysis as described in Section 4.1. A typical chromatogram for the studied glycans is presented in Appendix (see Appendix A, Figure A.1). The fractions for each glycan were collected for 1 min.

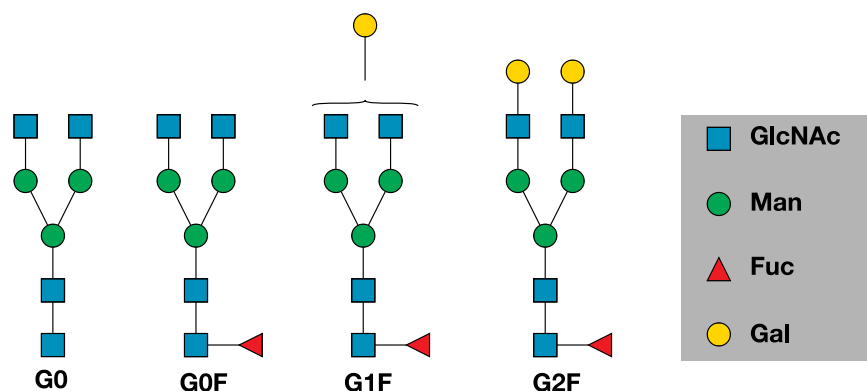


Figure 4.2 Schematic structures of the *N*-glycans studied in this work. Glycans are represented using the *Symbol Nomenclature for Glycans* (SNFG). Note that in the case of *G1F*, we do not distinguish the two positional isomers. However, a strategy to distinguish them is described in Chapter 5.

Here, we compare our different *N*-glycans cleaved from TNFR-Fc, G0, G0F, G1F, G2F, to a targeted database that we created using standards. The database contains mass, ion mobility ATDs, as well as reference IR spectra. Figure 4.3 shows a mass spectrum obtained for the G0F standard compared with that for G0F cleaved from TNFR-Fc (see Section 4.1.2.). The spectrum of the standard shows multiple ion

adduct species, including $[M + H + \text{NH}_4]^{2+}$, $[M + H + \text{Na}]^{2+}$, $[M + H + \text{K}]^{2+}$, $[M + 2\text{Na}]^{2+}$, and $[M + \text{K} + \text{Na}]^{2+}$, whereas the sample cleaved from TNFR-Fc shows primarily the doubly protonated form. The predominance of the latter is related to the clean-up procedure of the cleaved N-glycans performed in acidic (pH 4.5) media. The mass spectra for other glycans are shown in Appendix A (Figures A.2-A.4). For this reason, we compare the cleaved glycans with their corresponding standards in their doubly protonated form. To avoid salt presence, one can consider to treat samples using preparative HPLC prior sample analysis. However, this step can be time-consuming and may cause the sample losses.

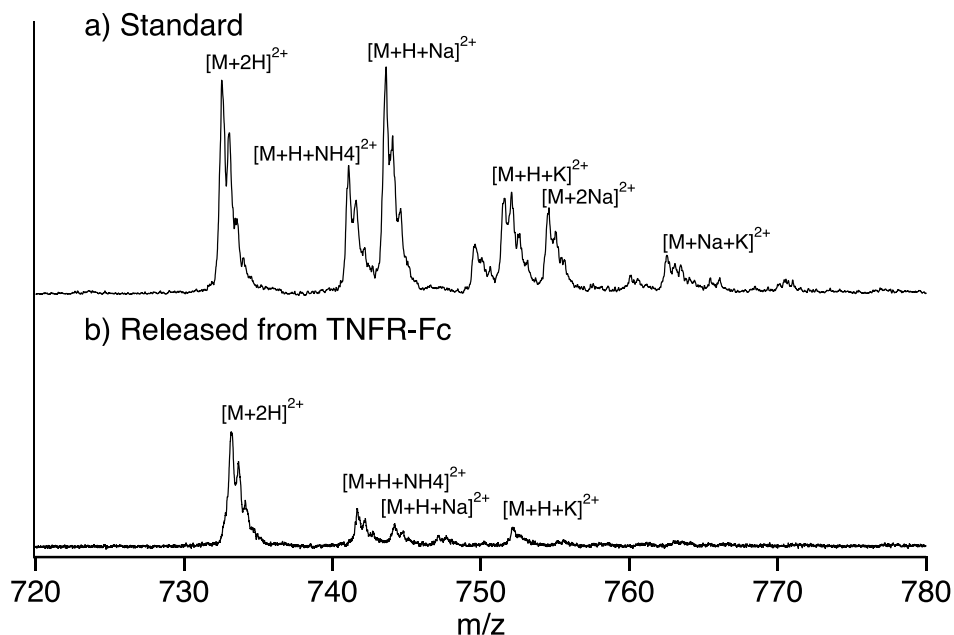


Figure 4.3 Comparison of ion adducts of G0F produced in positive ESI mode. a) standard, b) released of TNFR-Fc.

Figure 4.4 shows ATDs of the G0F standard. The three sharp peaks observed after one cycle on the SLIM board (Figure 4.4 a) separate into multiple peaks after three SLIM cycles (Figure 4.4 b). We clearly observe additional small peaks that correspond to ions having slightly different three-dimensional structure from the major species. Multiple peaks in the ATD could correspond to different isomers (e.g., α and β anomers at the reducing end OH [3–5]) or to multiple conformers of those isomers.

There can also be different protonation sites or different sites to which fucose might migrate [6].

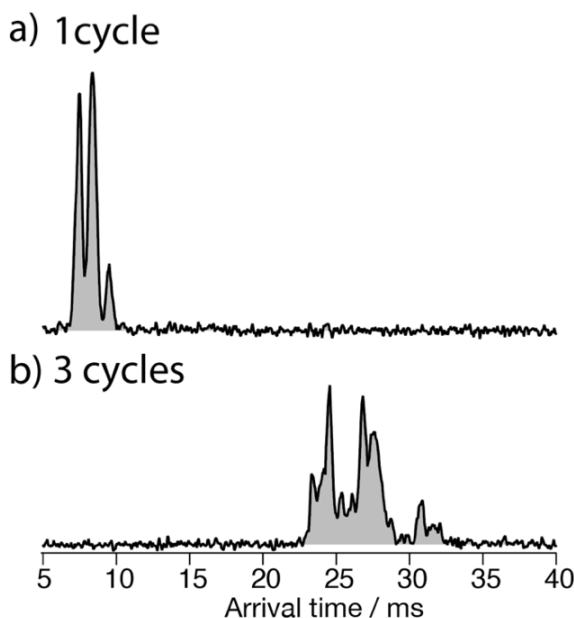


Figure 4.4 Arrival time distribution of doubly protonated G0F after (a) one cycle on the SLIM board (drift length of 1.5 m) and (b) three separation cycles (drift length of 5.27 m).

The ATDs reveal that glycan ions can adopt numerous stable conformations in the gas phase. An important aspect to consider while defining a protocol for the database is to ensure that the observed ATDs are highly reproducible between different experimental runs and do not depend on the nESI conditions. Here, we have used significantly high electric field gradients and RF amplitudes in the first stages of our instrument prior to IMS separation to anneal the conformational distribution. While this does not necessarily imply that we are observing the lowest energy conformers, since some may be kinetically trapped [7–9], it ensures that the observed ATDs are reproducible between different experimental runs and independent of the nESI conditions. In all cases for the measured ATDs, we detected no significant change under acceleration in the trap segment on the SLIM board (see Chapter 3, section 3.1.2 for details). As an example, we show in Figure 4.5 the ATD of doubly protonated G0F, under harsh (Figure 4.5 a) and normal conditions (Figure 4.5 b). As one can

observe, apart from the slight change in some peak intensities that are designated with the dash lines, the ATDs are reproducible.

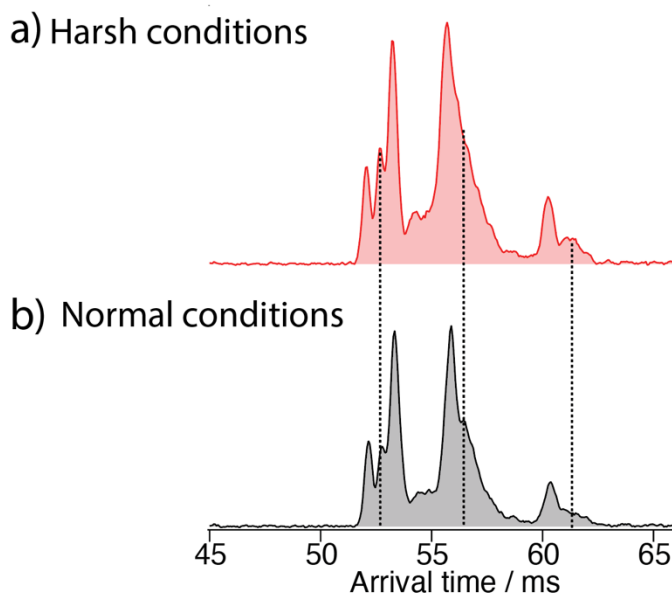


Figure 4.5 Arrival time distribution of the doubly protonated G0F of (m/z 732.7 Th) after three cycles after trapping (5.22 m) on the SLIM board. a) under harsh conditions (see Chapter 3 for details); b) under normal conditions (see Chapter 3 for details).

For our N-glycan database of standards, we recorded cryogenic IR spectra of the most intense peaks in the ATD, shown in Fig. 4.6 for G0F. The infrared spectra exhibit sharp, distinct features in the free OH stretch region ($3580\text{--}3700\text{ cm}^{-1}$) and broad transitions in the weakly hydrogen-bonded OH stretch region ($3450\text{--}3550\text{ cm}^{-1}$), all of which provide a fingerprint that can be used to identify the molecule. While these mobility-selected infrared spectra likely represent subtly different molecular conformations, the reproducibility of the ATD's ensures that they provide a reliable identifying fingerprint.

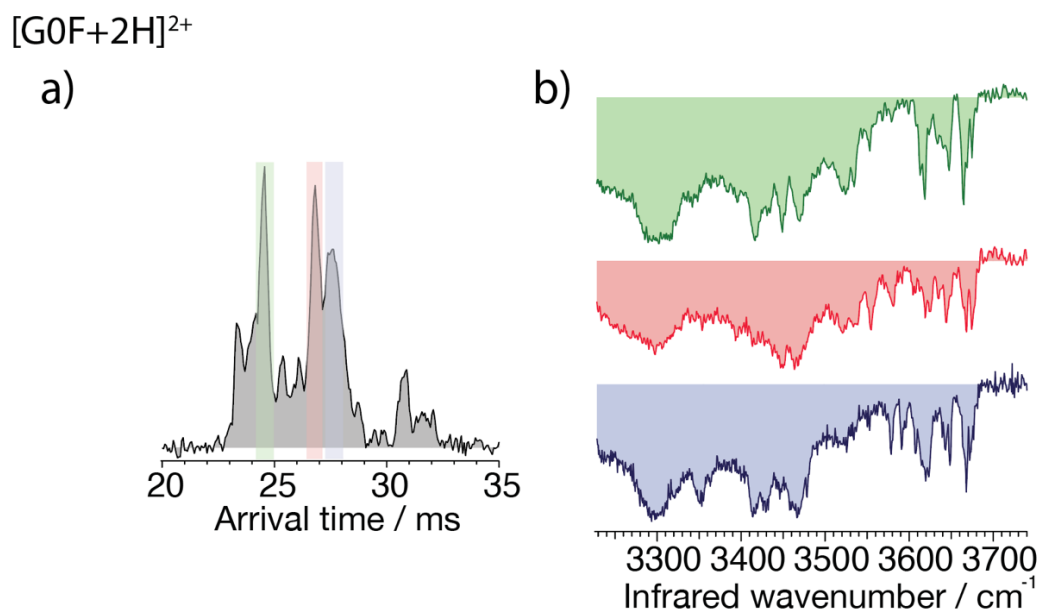


Figure 4.6 Arrival time distribution of the doubly-protonated G0F of m/z 732.7 Th shown on the left. The cryogenic IR spectrum of mobility-separated drift peaks is shown on the right (red, green, and blue trace for first, second and third peak respectively).

The data of Figure 4.7 below highlight the utility of ultrahigh-resolution ion mobility for comparing the released sugars with their respective standards in the doubly protonated form. The ATD for the G0 standard (Figure 4.7 a, top) shows numerous peaks that are not fully resolved after five cycles on the SLM board, while that for the released G0 (Figure 4.7 a, bottom) shows a similar profile, but the intensity distribution differs. For the G0F standard (Figure 4.7 b, top) the ATD has multiple peaks that agree well both in position and intensities with those of released G0F (Figure 4.7 b, bottom). In the case of the G1F standard which is a mixture of two positional isomers (Figure 4.7 c, top), the ATD exhibits various prominent peaks, and while that of the cleaved G1F has the same number of peaks, they differ slightly in their intensities (Figure 4.7 c, bottom). The G2F standard shows two main peaks in the ATD with slight shoulders that are not fully resolved (Figure 4.7 d, top), and the released G2F has a similar shape with slightly different intensities of these two peaks (Figure 4.7 d, bottom). While some minor differences appear in the ATDs of

the released sugars compared to their respective standards, as demonstrated below, their respective cryogenic IR spectra confirm unambiguously our ability to identify them.

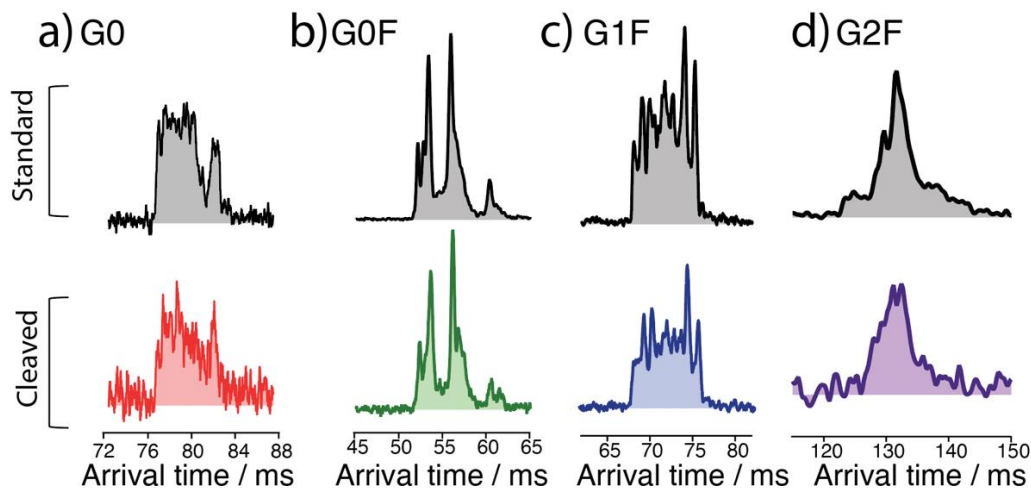


Figure 4.7 Arrival time distributions for the doubly-protonated a) G0 standard (top, grey) and G0 from the Etanercept (red), 5 cycles after enrichment (8.62 m); b) G0F standard (top, grey) and G0F from the Etanercept (green), 3 cycles after enrichment (5.22 m); c) G1F standard (top, grey) and G1F from the Etanercept (blue), 3 cycle after enrichment (5.22 m); d) G2F standard (top, grey) and G2F from the Etanercept (violet), 5 cycles after enrichment (8.62 m). ATDs have been recorded using the different IMS settings for each glycan sample, and consequently cannot be directly compared.

We recorded cryogenic IR spectra of the four cleaved N-glycans and their respective standards without ion mobility selection, shown in Figure 4.8. The spectrum of each N-glycan has unique, well-resolved transitions in the free OH stretch region ($3580\text{--}3700\text{ cm}^{-1}$) that can be used for identification. In addition, G0 and G0F have distinct broad bands in the weakly hydrogen-bonded OH stretch region ($3450\text{--}3550\text{ cm}^{-1}$) that are also distinctive. We chose to use only the $3450\text{--}3750\text{ cm}^{-1}$ region for identification purposes.

The IR spectra of standard and released G0 (Figure 4.8 a) generally match in both the position and intensity of the vibrational bands, with only minor differences

for broader bands (3471 cm^{-1} and 3511 cm^{-1}). Spectra of the reference and cleaved G0F (Figure 4.8 b), G1F (Figure 4.8 c), and G2F (Figure 4.8 d) were nearly identical in both the band positions and intensities. To quantify the similarity between IR spectra of standard and released glycans, we calculated the correlation coefficient between each corresponding pair (see subsection 4.3). The resulting high degree of correlation (95.9 % for G0, 97.7 % for G0F, 98.8 % for G1F, and 98.3 % for G2F) confirms their visual similarity and shows the feasibility of using a spectroscopic database to identify them. The strong correlation is mostly due to the sharpness of the features in the high frequency region of the spectrum, which is the result of the low temperature of the trapped ions, rendering this approach sensitive to the subtlest structural differences. The slight differences between the reference spectra and those of the released glycans likely arises from small differences in the conformer distributions.

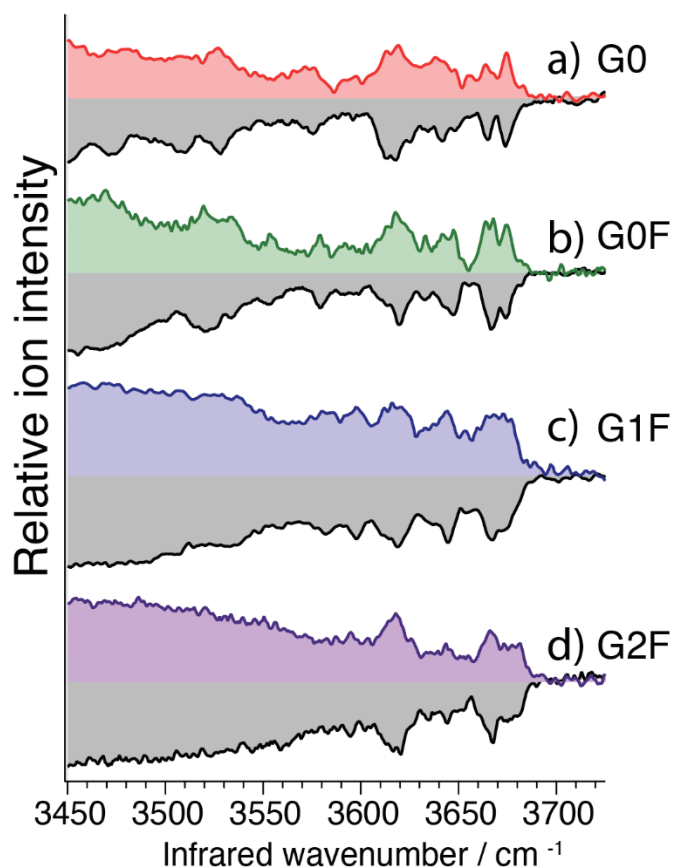


Figure 4.8 *Cryogenic IR-spectra of doubly-protonated glycan standards (grey) and released glycans: (a) G0; (b) G0F; (c) G1F; and (d) G2F. The spectra of G1F represent the mixture of the two positional isomers.*

Here, we did not record IR spectra for individual ion mobility peaks of each released N-glycan, as the total spectra of the sugars were nearly identical to their respective standards and match well both in the position and intensities of the vibrational bands. As demonstrated, mobility-selected IR spectra are slightly sharper and more distinctive and therefore provide a more stringent spectroscopic fingerprint. Nevertheless, this comes at the cost of sensitivity, since we are only sampling a subset of the ions, and hence we do this only when necessary to resolve ambiguous assignments.

To further assess the feasibility of this approach as a general tool for glycan analysis of large-scale production of biotherapeutics, we also need to demonstrate that the amount of sample required is reasonable. We used Etanercept having a mass of ~ 128 kDa, of which, 12-13 kDa can be attributed to N-glycans [10]. Consequently, in the 600 μg sample of Etanercept that we used, ~ 60 μg corresponds to N-glycans. To determine an upper bound for our limit of detection, we assume that we lose none of the sample during preparation and clean-up.

The glycan G0 (~ 1316 Da) has ~ 1 % abundance among N-glycans attached to TNFR-Fc [2]. This represents $\sim 0.73\%$ by mass, which would mean that we would have ~ 0.44 μg of this particular glycan. We typically had ~ 1 ml of the combined fractions together for G0, which gives us ~ 0.33 μM of the cleaved sugar G0. The measurement of one IR spectrum takes on average 3 mins to scan the range from 3400 cm^{-1} to 3750 cm^{-1} , and using an upper limit to the nESI flow rate of 100 nl/min, this means we use 300 nL of solution. From this, we can estimate that a total amount of sample upon which we made our measurement is ~ 100 fmol. If we repeat the same calculation for G0F, which is the most abundant N-glycan from Etanercept (21.8%), this means that our measurement was made on ~ 2 pmol. Any losses in sample during the clean-up procedure, which we neglected, would mean that our measurements were made on smaller amounts than this. While higher sensitivity is always better, our detection limit is in the range that would make this approach feasible for monitoring N-glycosylation in the large-scale production of biotherapeutics [11].

In addition to necessitating adequate sensitivity, a feasible analytical approach must be able to make measurements rapidly. An example of a cryogenic IR spectrum of the G0F standard measured in 55 seconds compared to one measured over a period of 180 s is given in Figure 4.9. Although the signal-to-noise ratio is not as good in the faster scan marked in pink, it still clearly exhibits resolved IR transitions in the OH spectral region that can be used as an identifying fingerprint. The resulting 95.8% degree of correlation between the fast and the slow cryogenic IR spectra confirms their visual similarity. Such a short analysis time scale emphasizes the promise of our approach for rapid N-glycosylation profiling.

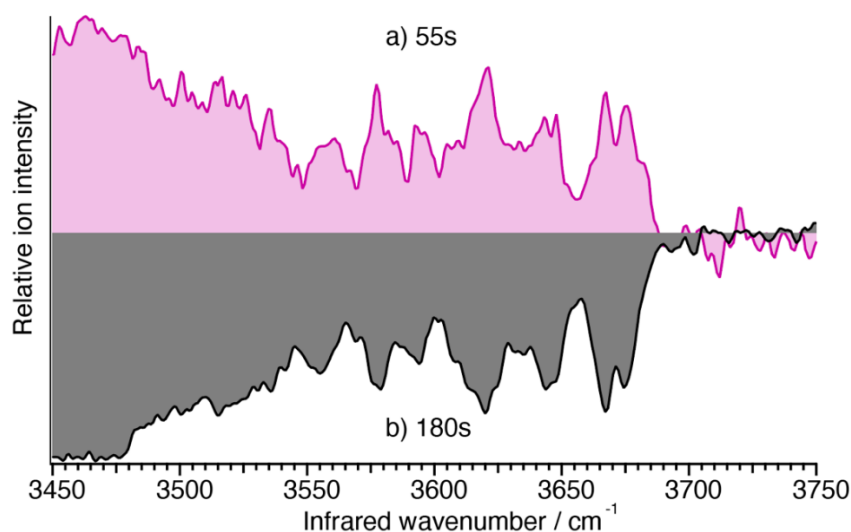


Figure 4.9 Cryogenic IR spectrum of G0F standard recorded in a) 55 s and b) in 180 s.

4.3 Comparison of mathematical methods to quantify the similarity between IR spectra of reference and cleaved glycans

Quantification of spectral similarity is a key aspect of using a spectroscopic database. We need to have a proper method to determine how good the match is between our reference and analyzed compounds. Additionally, the mathematical method should be able to tell apart even minor differences between spectra and should be easily implemented in our future database search engine.

Here, we applied and compared three different mathematical methods to quantify the IR spectral similarity between N-glycans released from TNFR-Fc and their respective standards. All the calculations have been done using Matlab software.

4.3.1 *Dot product method*

The dot product or “the spectral contrast angle” has been routinely applied to quantify the similarity of mass-spectra, but it can also be used for the comparison of IR spectra [12,13].

The dot product treats each spectrum as a vector of ordered peak intensities and measures the cosine of the angle between the vectors. The closer the angle value to zero, the greater the similarity between the vectors. The higher angle value means greater differences between vectors.

The dot product of two vectors \mathbf{x} and \mathbf{y} is defined in the following manner:

$$\cos \theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}; \quad (\text{Equation 4. 1})$$

where \mathbf{x} is the reference spectrum vector and \mathbf{y} is the analyte (cleaved sample) spectrum vector, and θ is the angle between them. As the angle θ approaches 0° , $\cos \theta$ approaches 1.

Note that

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n \mathbf{x}_i \cdot \mathbf{y}_i,$$

where \mathbf{x}_i represents the individual points of the reference spectrum and \mathbf{y}_i represents the individual points for the analyte;

and:

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^n \mathbf{x}_i^2} \text{ and } \|\mathbf{y}\| = \sqrt{\sum_{i=1}^n \mathbf{y}_i^2},$$

where $\|\mathbf{x}\|$ and $\|\mathbf{y}\|$ designate the length of vectors \mathbf{x} and \mathbf{y} . The vectors \mathbf{x} and \mathbf{y} must be normalized to the total number of points.

The dot product has been employed in the “CromAlign” algorithm [14] which is used for LC-MS. Stein *et al.* have made a comparison of the existing algorithms, including the dot product in the NIST/EPA/NIH Mass Spectral Library [15]. Additionally, this method is quick to compute, since the formula is easy to apply.

Table 4.1 presents the similarity coefficients between the spectra of the doubly protonated standard and the cleaved sugars calculated by the dot product method.

Table 4.1 *The results of the dot product method applied to the spectra of each corresponding pair between the reference and the cleaved sugars.*

	G0(cleaved)	G0F(cleaved)	G1F(cleaved)	G2F(cleaved)
G0 (standard)	0.9923	0.9813	0.9784	0.9734
G0F (standard)	0.9697	0.9927	0.9629	0.9691
G1F (standard)	0.9819	0.9811	0.9977	0.9926
G2F (standard)	0.9733	0.9712	0.9933	0.9968

The similarity between the reference and the cleaved sugars, as measured by the dot product method, are the following: 99.2 % for G0, 99.3 % for G0F, 99.8 % for G1F, and 99.7 % for G2F. From Table 4.1 we can also conclude that all our IR spectra are very similar which is true since they belong to one class of molecules and have the same functional groups. Moreover, the data shows that dot product method fails to distinguish clearly between G1F and G2F. To circumvent this, some constraints, such as intensity scaling or mass weighting could be introduced [15]. Moreover, the dot product algorithm might require a peak selection process to reduce the complexity of spectra for comparison.

4.3.2 Cross-correlation method

We also applied the weighted cross-correlation method. It is generally used to compare NMR spectra, where one needs to recognize the similarity of spectra having chemical shifts.

De Gelder and co-authors [16] proposed a general equation (Equation 4. 2) to express spectral similarity, S_{fg} , that is based on a cross-correlation function, c_{fg} , weighted with a weighting function $w(\Delta)$, and normalized with the product of two auto-correlation functions, c_{ff} and c_{gg} :

$$S_{fg} = \frac{\int w(\Delta)c_{fg}(\Delta)d\Delta}{\sqrt{\int w(\Delta)c_{ff}(\Delta)d\Delta \cdot \int w(\Delta)c_{gg}(\Delta)d\Delta}}, \quad (\text{Equation 4. 3})$$

where

$$c_{ff}(\Delta) = \int f(x)f(x + \Delta x) dx;$$

$$c_{gg}(\Delta) = \int g(x)g(x + \Delta x) dx;$$

$$c_{fg}(\Delta) = \int f(x)g(x + \Delta x) dx.$$

When applied to our IR spectral comparison problem, c_{ff} is an auto-correlation function of the reference spectrum, c_{gg} is auto-correlation function of the analyte (cleaved sugars) spectrum, and c_{fg} is the cross-correlation function. The triangular weighting function with a width l was chosen for our analysis.

Table 4.2 presents the cross-correlation method results for the similarity between the spectra of the doubly protonated standard and the cleaved sugars, calculated using a triangular weighting function with a width $l = 2 \text{ cm}^{-1}$. The weighting function width was chosen empirically.

Table 4.2 *The results of the cross-correlation coefficient method applied to the spectra of each corresponding pair between the standard and the cleaved sugars.*

	G0(cleaved)	G0F(cleaved)	G1F(cleaved)	G2F(cleaved)
G0 (standard)	0.9901	0.9598	0.9764	0.9656

G0F (standard)	0.9748	0.9896	0.9791	0.9683
G1F (standard)	0.9738	0.9567	0.9973	0.9915
G2F (standard)	0.9657	0.9613	0.99	0.9966

The similarity between the standard and the cleaved sugars, as measured by the weighted cross correlation method, are the following: 99 % for G0, 98.9 % for G0F, 99.7 % for G1F, and 99.7 % for G2F. Similarly to the dot product, the cross correlation method fails to distinguish between G1F and G2F spectra.

4.3.3 Correlation coefficient method

The correlation coefficient method provides the degree of similarity between two data vectors (Equation 4.3).

The correlation coefficient (r) between two vectors \mathbf{x} and \mathbf{y} is defined as

$$r = \frac{\sqrt{\text{cov}(x,y)}}{\sqrt{\text{var}(x)} \cdot \sqrt{\text{var}(y)}} ; \quad (\text{Equation 4.3})$$

where

$$\text{cov}(x,y) = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y});$$

$$\text{var}(x) = \sum_{i=1}^n (x_i - \bar{x})^2;$$

$$\text{var}(y) = \sum_{i=1}^n (y_i - \bar{y})^2;$$

and:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i ;$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i ;$$

It can be seen that the vectors have to be normalized on the total number of points.

Geometrically, this method can be perceived as the cosine of the angle between the two mean-centered vectors. The correlation method is closely related to the dot product. However, the covariance is the mean corrected dot product, and the squared correlation coefficient (R^2) is the covariance corrected by the standard deviation. This approach is one of the most commonly used methods to determine spectral similarity. Varmuza *et al.* compared different approaches for searches in an IR database, and in his study the correlation coefficient performed the best [17].

Table 4.3 presents the cross-correlation method results that show the similarity between the spectra of the doubly protonated standard and the cleaved sugars.

Table 4.3 *The results of the cross-correlation coefficient method applied to the spectra of each corresponding pair between the standard and the cleaved sugars.*

	G0 (cleaved)	G0F(cleaved)	G1F(cleaved)	G2F(cleaved)
G0 (standard)	0.9599	0.9178	0.8777	0.8627
G0F (standard)	0.8973	0.9772	0.8993	0.9008
G1F (standard)	0.9103	0.9279	0.9884	0.9603
G2F (standard)	0.8649	0.8819	0.9642	0.9831

The similarity between the standard and the cleaved sugars, as measured by the correlation coefficient, are the following: 95.9 % for G0, 97.7 % for G0F, 98.8 % for G1F, and 98.3 % for G2F. It provided the best results for the comparison of spectral similarity among all the sugars in terms of correctly identifying the best match between the spectra of cleaved sugars and references. Nevertheless, the correlation coefficient between the spectra of G1F and G2F is still relatively high (96%) due to

the similarity between their overall spectra that do not contain sufficiently distinctive features in the weakly hydrogen-bonded OH stretch region ($3450\text{--}3550\text{ cm}^{-1}$). However, if we use only the free OH stretch region ($3580\text{--}3700\text{ cm}^{-1}$) for comparison, then the resulting correlation coefficient between G1F and G2F is just 91.2%.

It is important to note that in cases where the spectral correlation between two species is high, the ATD can also be used to distinguish them. The spectra of Figure 4.8 were measured without selecting any part of the ATD which produced significantly complex cryogenic IR spectra that were somewhat difficult to treat with spectra comparison methods. When selecting certain peaks in the ATD, sharper spectra having considerably more structure can be obtained as shown in Figure 4.7. This needs to be done only in cases where non-mobility-resolved spectra are highly correlated

4.4 New perspectives on studying Rituximab

Of ongoing interest is a demonstration that our technique can be applied to glycans from different therapeutic proteins and mAbs. For this reason, we chose to study released glycans from Rituximab, which is a major pharmaceutical component of the drug RituxanTM[18]. Rituximab is a chimeric mAb used to treat autoimmune diseases, including rheumatoid arthritis and types of cancer, such as non-Hodgkin lymphoma, and chronic lymphocytic leukemia [19,20]. Here, we start with a comparison of the G0F glycan. Figure 4.10 shows mass spectra obtained for the G0F standard compared with that for G0F cleaved from TNFR-Fc studied previously and for G0F cleaved from Rituximab. Similarly to the mass-spectrum of the cleaved G0F from TNFR-Fc, the mass spectrum of the cleaved G0F from Rituximab predominantly exhibits the doubly protonated form. However, the mass spectrum itself shows some remaining impurities from peptides.

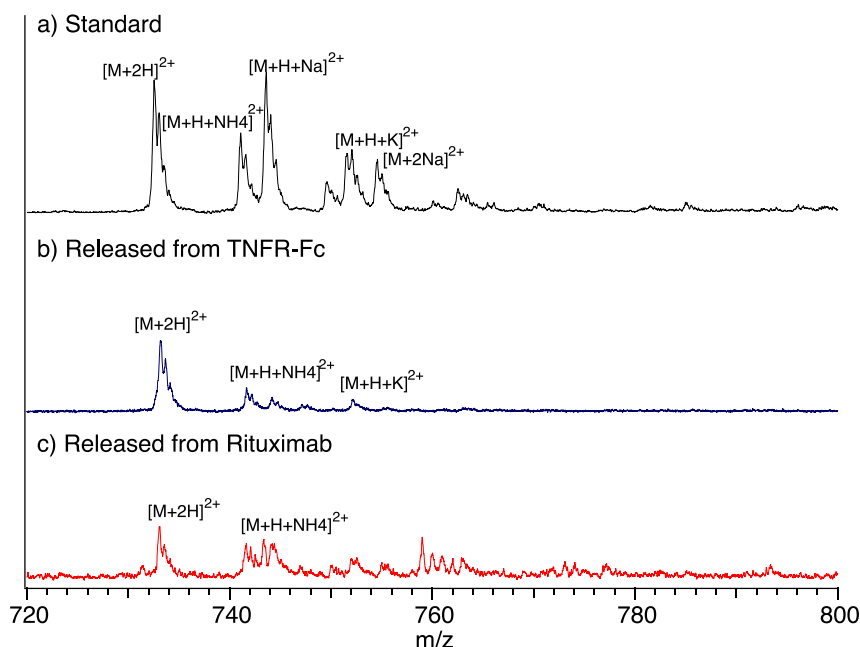


Figure 4.10 Comparison of ion adducts of G0F produced in positive ESI mode. a) standard, b) released of TNFR-Fc, c) released of Rituximab.

We then recorded the cryogenic IR spectrum of the cleaved G0F from Rituximab and compared it to those of the G0F standard and G0F cleaved from TNFR-Fc without ion mobility selection, shown in Figure 4.11. The spectrum of the G0F cleaved from Rituximab has distinctive well-resolved transitions in the free OH stretch region (3580–3700 cm^{-1}) that can be used as a fingerprint for its identification. Spectra of the reference and cleaved G0F have a good agreement in both the band positions and intensities. The correlation coefficient between each corresponding pair (97.7 % for the standard (Figure 4.11 a) and the cleaved sugar from TNFR-Fc (Figure 4.11 b) and 96.6 % for the standard (Figure 4.11 a) and the cleaved sugar from Rituximab (Figure 4.11 c)) confirms the feasibility of using a spectroscopic database to identify them. As we mentioned above, the subtle differences between the reference spectra and those of the released glycans likely arise from slight differences in their conformational distributions.

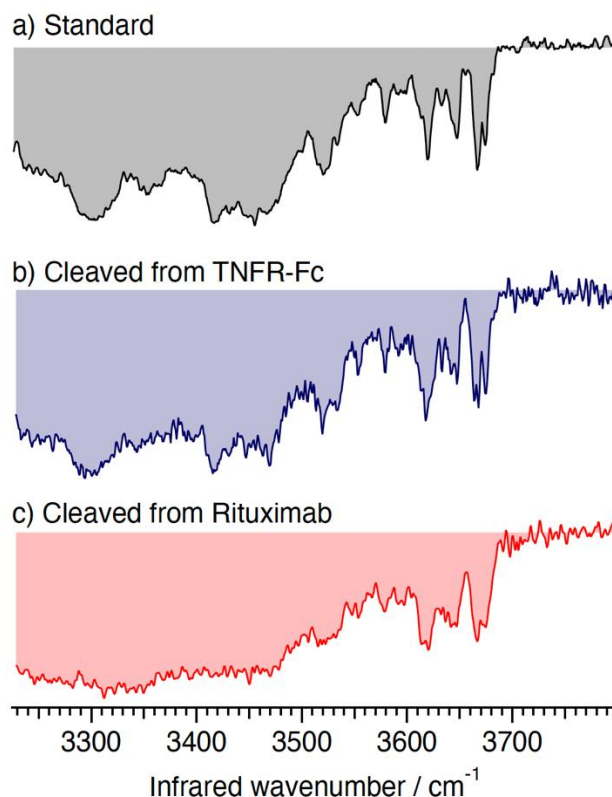


Figure 4.11 Cryogenic IR-spectra of doubly protonated G0F a) standard (grey), b) released G0F from TNFR-Fc (blue), c) from Rituximab (red).

This preliminary data shows a great promise of applicability of our technique to study cleaved sugars from different biologics. We are confident that our research will serve as a base for future studies on the consistency of glycosylation in novel therapeutics.

4.5 Conclusions and future perspectives

In this chapter, we have demonstrated that the combination of ultrahigh-resolution IMS based on SLIM technology together with cryogenic vibrational spectroscopy can unambiguously identify N-linked glycans cleaved from a therapeutic protein using a database approach. This study serves as a proof of principle that the sensitivity, speed, and resolution should be sufficient for large-scale N-glycan characterization and identification. This approach should simplify basic research and process development for novel therapeutics and biosimilar products. Our technique is

very promising and could eventually result in the development of a commercial instrument that is simple to operate.

While our technique is extremely versatile, there are several challenges remaining to be solved. One of the most important tasks is to increase throughput. Multiplexing spectroscopy strategies, such as Hadamard Transform, should make our measurements even more suitable for incorporation into analytical workflows.

References:

- (1) Largy, E.; Cantais, F.; Van Vyncht, G.; Beck, A.; Delobel, A. Orthogonal Liquid Chromatography-Mass Spectrometry Methods for the Comprehensive Characterization of Therapeutic Glycoproteins, from Released Glycans to Intact Protein Level. *J. Chromatogr. A* **2017**, *1498*, 128–146. <https://doi.org/10.1016/j.chroma.2017.02.072>.
- (2) Houel, S.; Hilliard, M.; Yu, Y. Q.; McLoughlin, N.; Martin, S. M.; Rudd, P. M.; Williams, J. P.; Chen, W. N- and O-Glycosylation Analysis of Etanercept Using Liquid Chromatography and Quadrupole Time-of-Flight Mass Spectrometry Equipped with Electron-Transfer Dissociation Functionality. *Anal. Chem.* **2014**, *86* (1), 576–584. <https://doi.org/10.1021/ac402726h>.
- (3) Nagy, G.; Attah, I. K.; Garimella, S. V. B.; Tang, K.; Ibrahim, Y. M.; Baker, E. S.; Smith, R. D. Unraveling the Isomeric Heterogeneity of Glycans: Ion Mobility Separations in Structures for Lossless Ion Manipulations. *Chem. Commun.* **2018**, *54* (83), 11701–11704. <https://doi.org/10.1039/C8CC06966B>.
- (4) Veillon, L.; Huang, Y.; Peng, W.; Dong, X.; Cho, B. G.; Mechref, Y. Characterization of Isomeric Glycan Structures by LC-MS/MS. *Electrophoresis* **2017**, *38* (17), 2100–2114. <https://doi.org/10.1002/elps.201700042>.
- (5) Warnke, S.; Ben Faleh, A.; Scutelnic, V.; Rizzo, T. R. Separation and Identification of Glycan Anomers Using Ultrahigh-Resolution Ion-Mobility Spectrometry and Cryogenic Ion Spectroscopy. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (11), 2204–2211. <https://doi.org/10.1021/jasms.8b06266>.
- (6) Lettow, M.; Mucha, E.; Manz, C.; Thomas, D. A.; Marianski, M.; Meijer, G.; von Helden, G.; Pagel, K. The Role of the Mobile Proton in Fucose Migration. *Anal. Bioanal. Chem.* **2019**, *411* (19), 4637–4645. <https://doi.org/10.1007/s00216-019-01657-w>.
- (7) Stearns, J. A.; Seaiby, C.; Boyarkin, O. V.; Rizzo, T. R. Spectroscopy and Conformational Preferences of Gas-Phase Helices. *Phys Chem Chem Phys* **2009**, *11* (1), 125–132. <https://doi.org/10.1039/B814143F>.

- (8) Rizzo, T. R.; Stearns, J. A.; Boyarkin, O. V. Spectroscopic Studies of Cold, Gas-Phase Biomolecular Ions. *Int. Rev. Phys. Chem.* **2009**, *28* (3), 481–515. <https://doi.org/10.1080/01442350903069931>.
- (9) Masson, A.; Kamrath, M. Z.; Perez, M. A. S.; Glover, M. S.; Rothlisberger, U.; Clemmer, D. E.; Rizzo, T. R. Infrared Spectroscopy of Mobility-Selected H⁺-Gly-Pro-Gly-Gly (GPGG). *J. Am. Soc. Mass Spectrom.* **2015**, *26* (9), 1444–1454. <https://doi.org/10.1007/s13361-015-1172-4>.
- (10) Wohlschlager, T.; Scheffler, K.; Forstenlehner, I. C.; Skala, W.; Senn, S.; Damoc, E.; Holzmann, J.; Huber, C. G. Native Mass Spectrometry Combined with Enzymatic Dissection Unravels Glycoform Heterogeneity of Biopharmaceuticals. *Nat. Commun.* **2018**, *9* (1), 1713. <https://doi.org/10.1038/s41467-018-04061-7>.
- (11) Zhou, S.; Veillon, L.; Dong, X.; Huang, Y.; Mechref, Y. Direct Comparison of Derivatization Strategies for LC-MS/MS Analysis of N-Glycans. *The Analyst* **2017**, *142* (23), 4446–4455. <https://doi.org/10.1039/c7an01262d>.
- (12) Mark, H.; Jr, J. W. *Chemometrics in Spectroscopy*; Elsevier, 2010.
- (13) Wan, K. X.; Vidavsky, I.; Gross, M. L. Comparing Similar Spectra: From Similarity Index to Spectral Contrast Angle. *J. Am. Soc. Mass Spectrom.* **2002**, *13* (1), 85–88. [https://doi.org/10.1016/S1044-0305\(01\)00327-0](https://doi.org/10.1016/S1044-0305(01)00327-0).
- (14) Sadygov, R. G.; Maroto, F. M.; Hühmer, A. F. R. ChromAlign: A Two-Step Algorithmic Procedure for Time Alignment of Three-Dimensional LC-MS Chromatographic Surfaces. *Anal. Chem.* **2006**, *78* (24), 8207–8217. <https://doi.org/10.1021/ac060923y>.
- (15) Stein, S. E.; Scott, D. R. Optimization and Testing of Mass Spectral Library Search Algorithms for Compound Identification. *J. Am. Soc. Mass Spectrom.* **1994**, *5* (9), 859–866. [https://doi.org/10.1016/1044-0305\(94\)87009-8](https://doi.org/10.1016/1044-0305(94)87009-8).
- (16) Gelder, R. de; Wehrens, R.; Hageman, J. A. A Generalized Expression for the Similarity of Spectra: Application to Powder Diffraction Pattern Classification. *J. Comput. Chem.* **2001**, *22* (3), 273–289. [https://doi.org/10.1002/1096-987X\(200102\)22:3<273::AID-JCC1001>3.0.CO;2-0](https://doi.org/10.1002/1096-987X(200102)22:3<273::AID-JCC1001>3.0.CO;2-0).

- (17) Varmuza, K.; Karlovits, M.; Demuth, W. Spectral Similarity versus Structural Similarity: Infrared Spectroscopy. *Anal. Chim. Acta* **2003**, *490* (1), 313–324. [https://doi.org/10.1016/S0003-2670\(03\)00668-8](https://doi.org/10.1016/S0003-2670(03)00668-8).
- (18) Pierpont, T. M.; Limper, C. B.; Richards, K. L. Past, Present, and Future of Rituximab—The World’s First Oncology Monoclonal Antibody Therapy. *Front. Oncol.* **2018**, *8*. <https://doi.org/10.3389/fonc.2018.00163>.
- (19) Randall, K. L. Rituximab in Autoimmune Diseases. *Aust. Prescr.* **2016**, *39* (4), 131–134. <https://doi.org/10.18773/austprescr.2016.053>.
- (20) What is RITUXAN® (rituximab) <https://www.rituxan.com/patient/what-is-rituxan.html> (accessed Sep 7, 2020).

Chapter 5 Strategy to distinguish positional isomers of N-glycans*

This chapter describes preliminary data on a promising strategy to distinguish and identify positional isomers of N-isomeric glycans for which standards are not commercially available. This strategy includes a chemoenzymatic synthesis and a combination of ultra-high resolution ion mobility together with cryogenic ion spectroscopy.

5.1 Distinguishing positional isomers of N-glycans is a challenge

N-glycosylation is one of the most ubiquitous and complex post-translational modifications of eukaryotic cell proteins. N-glycans are involved in numerous pathological and physiological processes, such as cell adhesion, immune response, and tumor invasion [1–4]. Moreover, N-glycans with immense structural variety affect the regulation of the diverse protein functions, including folding, protein-protein interactions protein trafficking, etc [5,6]. The identification of isomeric N-glycans is particularly important in the biopharmaceutical industry since N-glycosylation is a critical quality attribute (CQA) of a drug [7–9].

As discussed in Chapter 1, N-linked glycans are both structurally diverse and complex to analyze. Their structural diversity arises from the intricate glycan biosynthesis that requires a large number of enzymes and precursors [10,11]. Consequently, glycosylation heterogeneity is affected by host cell type and culture environment, including pH, dissolved oxygen, stirring rate, and temperature [12,13]. Glycan complexity comes from diversity in monosaccharide composition, in the linkages between them, and branching.

The Gal monosaccharide unit is particularly important in the pathological response, as it is involved in the activation of inflammatory reactions in rheumatoid

arthritis [14,15]. Moreover, galactosylation is crucial for biopharmaceutical development [16–18]. Recent studies have shown that the terminal Gal residues on G1F(6) and G1F(3) attached to Fc of mAbs are associated with effector functions of mAbs [19]. The terminal Gal residue of the Man α 1-6 arm of Fc-glycans has a significant impact on complement-dependent cytotoxicity (CDC) and binding activities. As a result, the evaluation of the consistency of galactosylation in N-glycan structures is often required to guarantee the quality of mAbs.

The complete discrimination of positional isomers of N-glycans is currently a challenging frontier of glycomics. Several different methods have been implemented to distinguish and identify positional isomers of released N-linked glycans with terminal Gal. They include a combination of selective enzymatic digestion or synthesis together with high-resolution proton NMR spectroscopy and LC-MS/MS [20,21].

Current approaches to distinguish positional isomers of N-glycans include capillary electrophoresis [22,23], and specialized LC-MS techniques [24–27] with a derivatization step in their workflows. Despite the great potential of these hyphenated techniques, they have some drawbacks. Established glycan workflows often involve a derivatization step to label the glycans with a fluorescent tag as they do not contain a natural chromophore or fluorophore. This improves sensitivity and enables quantification, but is expensive and can lead to rearrangement/sialic acid loss. Moreover, without the reference standards, LC and CE methods cannot provide any structural information. As an alternative to chromatographic and electrophoretic technologies, ion mobility spectrometry (IMS) [28,29] is used to study positional isomers of asymmetric N-glycans with different terminal galactose positions. Sugita and coworkers applied IM-MS coupled with hydrophilic interaction liquid chromatography (HILIC) to study positional isomers of N-glycans [28]. Additionally, Pallister *et al.* demonstrated the ability to distinguish isomeric glycans and glycopeptides applying a high-throughput LC-FLR-IMS-MS workflow [30]. However, the relatively low IMS resolution can limit the separation of glycan isomers with subtle structural differences. The identification of unlabeled positional isomers of N-glycans remains a daunting challenge, as the standards are not readily available on the market.

Here, we demonstrate the power of using ultrahigh-resolution IMS and cryogenic, messenger-tagging IR spectroscopy to distinguish positional isomers of N-glycans with

different terminal galactose positions. Since the pure positional isomers are not available on the market, we synthesize G1F(6) from G1 using a human enzyme 1,6-fucosyltransferase (FUT8) [31]. Subsequently, the arrival-time distributions and cryogenic infrared fingerprints of G1F(6) are compared to a commercially available mixture of G1F positional isomers. An analogous set of data was also obtained for the G1 positional isomers.

5.2 Chemoenzymatic synthesis of asymmetric N-glycans

The structures of G1 and G1F glycans are shown in Figure 5.1. For G1 and G1F, the Gal unit can be added either to the α 1,6-arm, forming G1(6), or to the α 1,3-arm, producing G1(3). Different commercial suppliers sell G1 and G1F glycans as a mixture of two positional isomers with a terminal galactose sitting on Man α 1,6- or Man α 1,3-arms. High purity isomers G1F(6) and G1F(3) are generally unavailable for purchase on the market due to the extreme difficulty of their production. The main strategies to produce asymmetrically branched N-linked glycans include releasing them from natural sources and undergoing chemical and/or chemoenzymatic synthesis and purification [32,33]. It is challenging to isolate sufficient amounts of N-glycans from their natural sources due to their low abundance and complexity. Synthetic methods for the production of N-Glycans require advanced synthetic chemistry skills, associated infrastructure and are therefore difficult to scale up to industrial quantities. Positional isomers of G1F and G1 purchased from manufactures who claimed to have high purity samples were found to contain a mixture of multiple isomers, highlighting the challenges faced in procurement of the same.

For this study, G1F and G1 standards were purchased from Dextra Laboratories (UK). Those standards represent a mixture of two positional isomers. All standards were reconstituted in water/acetonitrile (70:30) to a concentration of 80 μ M prior to analysis. The stock solutions were further diluted using water/acetonitrile (70:30) to obtain 5-20 μ M analyte solutions. To promote formation of the sodiated adducts of glycans, we added 3-5 μ L of 46 μ M of sodium acetate.

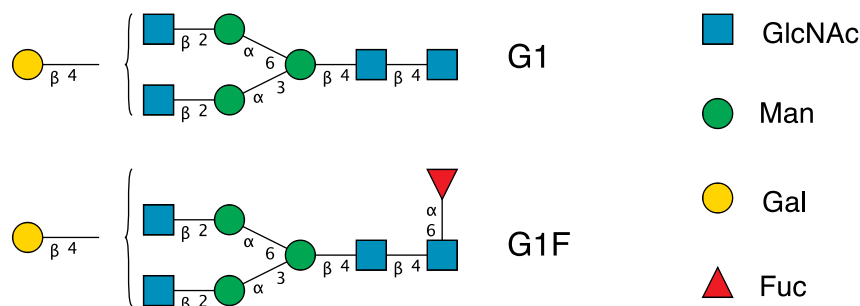


Figure 5.1 Representative structures of *G1* and *G1F* investigated in this thesis.

In this work, we performed the enzymatic synthesis of the positional *G1F*(6) isomer from *G1* using the *FUT8*-catalyzed reaction (Figure 5.2) [31,34]. *FUT8* is a highly specific enzyme that can add α 1,6-linked fucose to N-glycans [35]. The general mechanism to form an α 1,6-linkage is the following: the fucosyl moiety is relocated from the donor sugar nucleotide *GDP*- β -L-Fucose to the 6-hydroxy group of the innermost GlcNAc of the acceptor with inversion of the anomeric configuration. The acceptor, which in our case is *G1*(6), has to contain an unsubstituted GlcNAc residue at the α 1,3-arm to be recognized by the enzyme [31], and thus the reaction will specifically fucosylated the *G1*(6) positional isomer to yield *G1F*(6).

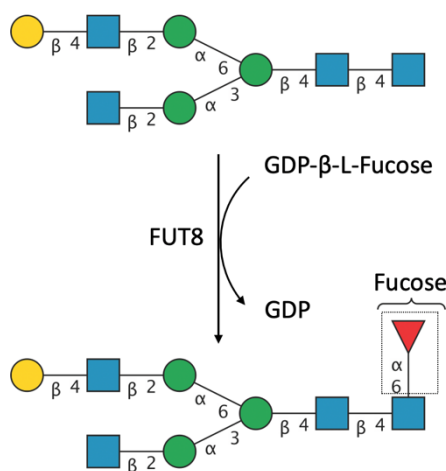


Figure 5.2 Enzymatic reaction catalyzed by *FUT8*.

The chemoenzymatic reaction by *FUT8* was carried out in 50 μ L total volume, containing 25 μ L of MES buffer (100 mM, pH 7.0), *G1* substrate (270.5 μ M), 7.5 μ L

of guanosine 5'-diphospho- β -L-fucose sodium salt (GDP-L-Fucose) (0.54 mM), and 6 μ L of FUT8 (0.12 mg/mL). Reactions were incubated at 37 °C. To determine the optimal incubation time of the reaction, samples were collected at 1, 2, 3, 4, 6, 8, 16 and 48 h. The reactions were then quenched by boiling for 5 min, and subject to LC analysis and purification. For that, we monitored the efficiency of the enzymatic reaction with a UPLC (ACQUITY™ H-Class Plus, Waters, UK) coupled to Micromass Q-TOF Premier (Waters, UK). We used a Waters XBridge BEH amide column (130 Å, 3.5 μ m, 4.6 \times 150 mm) at 0.5 mL/min at 60°. A gradient elution (%A: 36-44% within 20 min) was performed with Solvent A (100 mM ammonium formate, pH 4.5) and Solvent B (Acetonitrile). The conversion efficiency (see Table 5.1) was calculated as $\% = \text{Product peak area} / (\text{Product peak area} + \text{Substrate peak area}) \times 100$.

Table 5.1. *Conversions were monitored by LC.*

	G1F(6)
Time	Conversion (%)
1h	17.1
2h	18.2
3h	19.6
4h	21.5
6h	26.8
8h	30.3
16h	43.7
48h	57.7

Our assumption is that you start with a 50/50 mixture of the positional isomers of G1. Since we are producing only one of the positional isomers, the conversion efficiency should be less than 50%. According to the results in Table 5.1, the best incubation time of the reaction was 16h. It can be seen from Table 5.1 that G1(3) could start reacting to form the G1F(3) impurity, which leads to a larger conversion rate, more than 57%, after 48 h. For this reason, we chose 16h as the optimal incubation time of the reaction, assuming that G1F(6) is exclusively formed.

Figure 5.3 shows typical mass spectra obtained for the G1, G1F standards, and the enzymatic reaction after 16 h using our LC system. These mass spectra demonstrate a predominance of the $[M + H]^+$ and $[M + Na]^+$ forms. The observed predominant forms in LC set-up differ from the main species observed in our set-up due to several factors, including instrumental factors, acid-base chemistry in solution and in the gas phase, molecular conformations and the solvent composition.

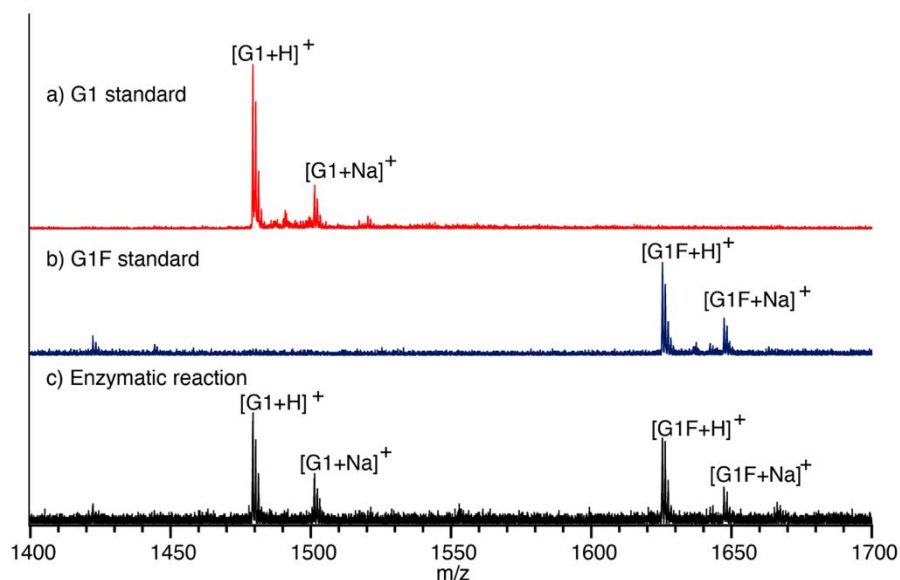


Figure 5.3 Typical mass spectra for a) G1 standard, b) G1F standard, and c) the enzymatic reaction of G1 after 16 h.

A typical chromatogram for the studied glycans is shown in Figure 5.4. As our LC method cannot baseline resolve the eluted peaks, we analyzed them as a mixture of likely remaining donor G1(3) and product G1F(6). The fractions for each glycan were collected for 40 s and were further analyzed with our technique that combines ultrahigh-resolution IMS with cryogenic, messenger-tagging, IR spectroscopy.

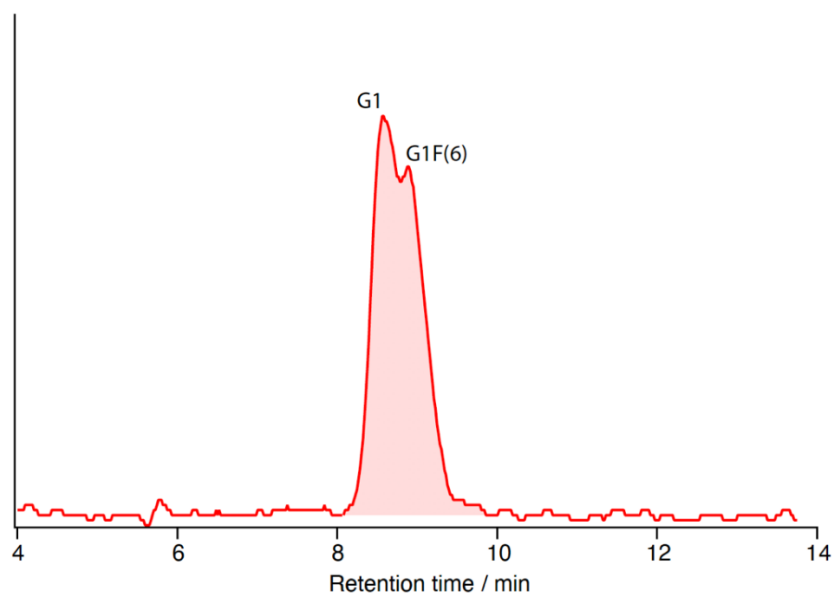


Figure 5.4 Typical chromatogram of the conversion of *G1* to *G1F(6)* after 16h using our LC system. Glycans elute as one peak.

5.3 Identification of positional N-glycan isomers

Here, the power of using ultrahigh-resolution IMS and cryogenic, messenger-tagging, IR spectroscopy is demonstrated by distinguishing isomeric asymmetric N-glycans with different terminal galactose positions that were synthesized. This subchapter contains preliminary results. Further work is being conducted by a subsequent student.

We studied *G1* and *G1F* glycans in their doubly sodiated form, since complexation with sodium tends to simplify the ATD and improves the resolution of the IR spectra (see Appendix D).

Figure 5.5 shows the ATDs and the cryogenic spectra of each ion mobility peak of the *G1F* standard. Five sharp peaks were observed after three SLIM cycles (Figure 5.5 a). These five peaks correspond to isomers with different rotationally averaged cross-sections, which could arise from either of the positional isomers, different conformers of isomers, or *G1F* molecules with different sodium binding sites.

Cryogenic IR spectra of the most intense peaks in the ATD of G1F are shown in Figure 5.5 b. The infrared spectra have sharp, distinct features in the free OH stretch region ($3580\text{--}3700\text{ cm}^{-1}$) and broad transitions in the weakly hydrogen-bonded OH stretch region ($3450\text{--}3550\text{ cm}^{-1}$), except for the first ion mobility peak depicted in red, which has sharp features throughout the spectrum. These spectra constitute a fingerprint by which one can identify G1F, although it is a mixture of positional isomers.

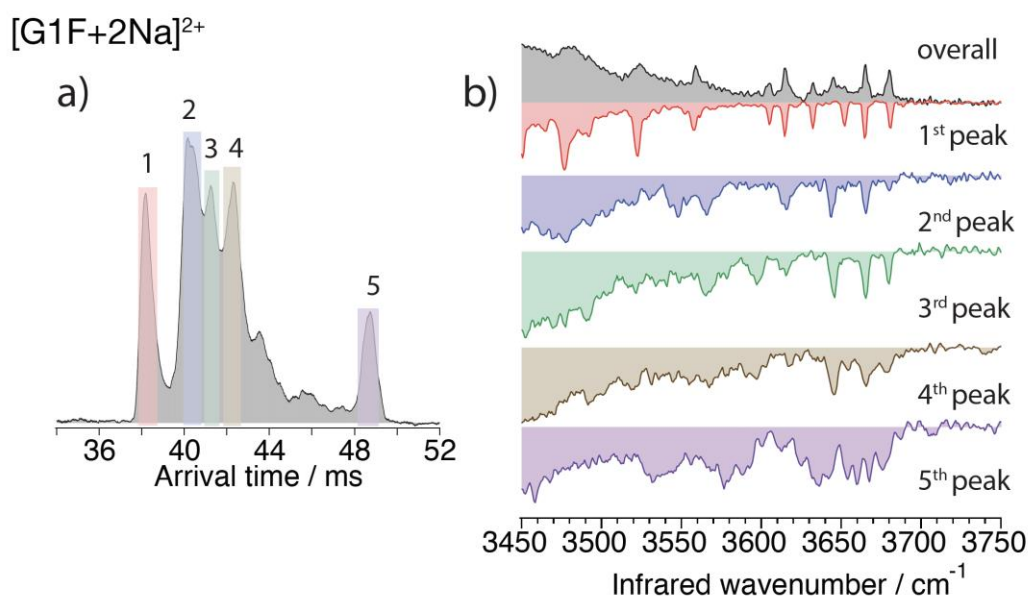


Figure 5.5 Arrival time distribution of doubly sodiated G1F of m/z 835.8 *Th* shown on the left, three cycles (5.27 *m*). The cryogenic IR spectrum of mobility-separated drift peaks is shown on the right (grey, red, blue, green, yellow and violet trace for overall, first, second, third, fourth and fifth peak respectively).

The two positional isomers of G1F could be very different in their mobility. For example, the first peak in the ATD depicted in red differs in drift time from the fifth peak in violet by about 15%. It indicates that the red peak has a more compact shape, while the fifth is more elongated, likely with a different hydrogen-bonding pattern. Studies from Clemmer's group has shown that that folding of α 1-3 arm leads to small ATDs, and the backfolding of the α 1-6 arm to the core chitobiose is commonly observed to lead to elongated structures [36]. This would suggest that the first peak in the ATD might correspond to the G1F(3) isomer and the last to G1F(6). Moreover,

the spectrum of the first ATD peak exhibits very sharp and distinct features comparing to that of the fifth drift peak, which could be rationalized by differences in its binding of the sodium cations and its hydrogen-bonding pattern. For the first peak, the conformations may be locked through interactions with sodium cations, leading to fewer peaks in the free OH region. For the fifth peak, one would expect the reduction of steric hindrance between the monosaccharide subunits and the metal ions to enhance their relative conformational flexibility. This might create a wide range of potentially accessible low energy conformations. In addition to the data from Figure 5.5, the flexibility of the 1–6 linkage has been well documented by other experimental, computational, and structural studies [37–39]. This increased flexibility is due to having three rotatable torsional angles, while other glycosidic linkages contain only two. However, the spectrum of G1F without ion mobility separation strongly resembles that of the most compact structure (red spectrum in Figure 5.5 b). This is due, in part, to the fact that the tagging efficiency of the first peak was greater than for the other species. One might expect that this is linked to the preferential binding of the tag to the sodium sites that are more accessible in the more compact glycan.

The data of Figure 5.6 below compares the ATDs of the G1F standard (grey) with that of what is presumably G1F(6) produced by the selective enzymatic reaction. The ATDs of the G1F standard (Figure 5.6, top) has five peaks, whereas G1F(6) (Figure 5.6, bottom) has three major peaks with some minor peaks. The data suggest that the second, the fourth, and the fifth peaks in the ATD of the mixture correspond to the G1F(6) isomer, while the first and the third peaks feature can be attributed to the G1F(3) isomer. These conclusions are in agreement with previous studies from the Clemmer group on N-glycans [10].

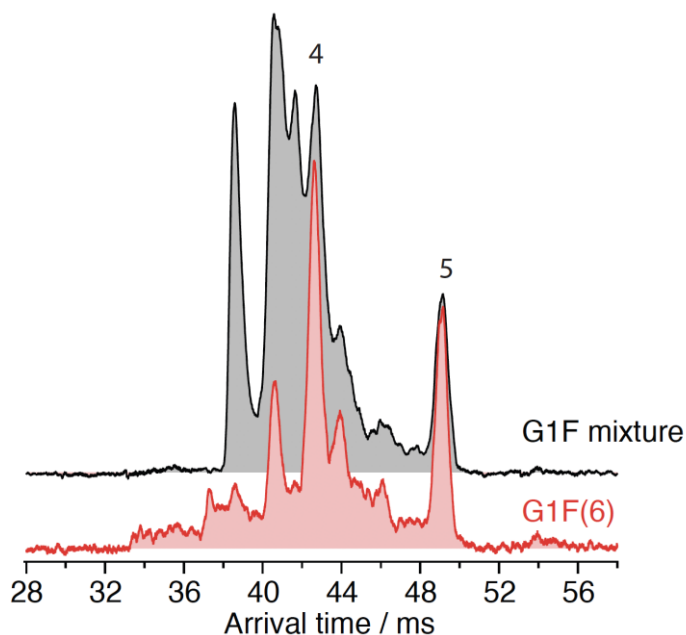


Figure 5.6 Comparison of the ATDs of the G1F standard in grey and the product G1F(6) after the enzymatic reaction in red after three SLIM cycles.

To further assess the identification of the positional isomers, cryogenic spectra without ion mobility selection were recorded for both the G1F standard and G1F(6), shown in Figure 5.7. The spectra are again distinctive in the free OH region. From this data one can conclude that the major contribution to the overall G1F cryogenic spectrum in grey comes from the cryogenic spectrum of the first peak in the ATD (Figure 5.5 b). The cryogenic spectra of the fifth peak in the ATD strongly resembles the overall spectrum of the G1F(6) product.

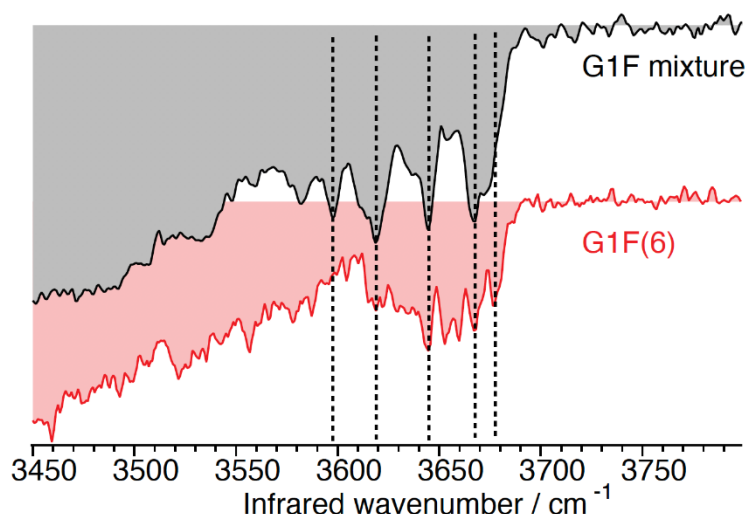


Figure 5.7 Comparison of the cryogenic IR spectra without ion mobility selection of the G1F standard in grey and the formed G1F(6) after the enzymatic reaction in red.

As a next step for the complete assignment of the ATD peaks to the positional isomers of G1F glycan, cryogenic spectra of each peak in the ATD of the synthesized G1F(6) will be recorded and compared with their respective standards. This work is being performed by another student.

While G1F is generally one of the most abundant in mAbs or in therapeutic proteins, G1 is one of the least abundant of these biologics. Our approach can also distinguish G1 positional isomers.

Figure 5.8 demonstrates the ATDs and the cryogenic spectra of each mobility peak of the G1 standard. The five peaks are observed after two SLIM cycles after enrichment (Figure 5.8 a). We detect other minor peaks between the fourth and the fifth peaks related to ions of slightly different three-dimensional structures. Furthermore, the obtained cryogenic IR spectra of the most intense peaks in the ATD for G1 are shown in Figure 5.8 b. The infrared spectra are similar to those for G1F in that they have sharp, distinct features in the free OH stretch region ($3580\text{--}3700\text{ cm}^{-1}$) and broad transitions in the weakly hydrogen-bonded OH stretch region

(3450–3550 cm^{-1}). The more compact form of the glycan corresponding to the green peak in the ATD leads to a sharp vibrational spectrum in both regions.

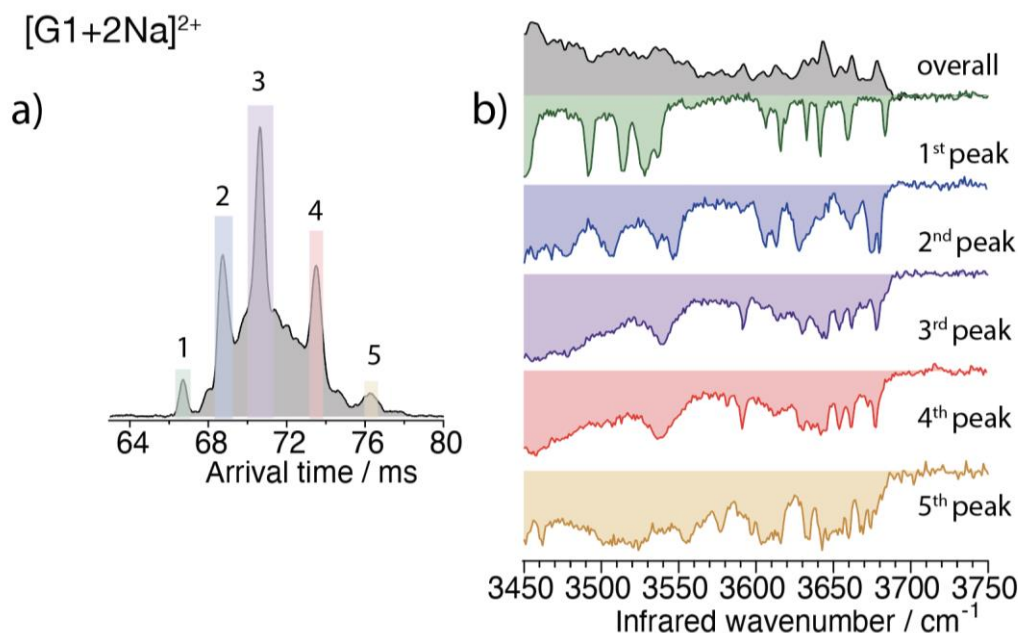


Figure 5.8 Arrival time distribution of doubly sodiated G1 of m/z 762.8 *Th* shown on the left, two cycles after enrichment (2.95 *m*). The cryogenic IR spectrum of mobility-separated drift peaks is shown on the right (grey, green, blue, violet, red and yellow trace for overall, first, second, third, fourth and fifth peak, respectively).

Figure 5.9 demonstrates the advantage of ultrahigh-resolution ion mobility for comparing the G1 standard to the remaining G1 after 16 h of the chemoenzymatic reaction, which could be the G1(3) isomer remnant. The ATD for the G1 standard (Figure 5.9, bottom) exhibits five distinct peaks that are not fully resolved, while that for the remaining G1 (Figure 5.9, top) shows a similar shape, but with only three distinct peaks. This data confirms that the second, the third and the fourth peaks in the ATD might correspond to the G1(3) isomer, and the first peak and the fifth peaks for the G1 standard could correspond to the G1(6) isomer.

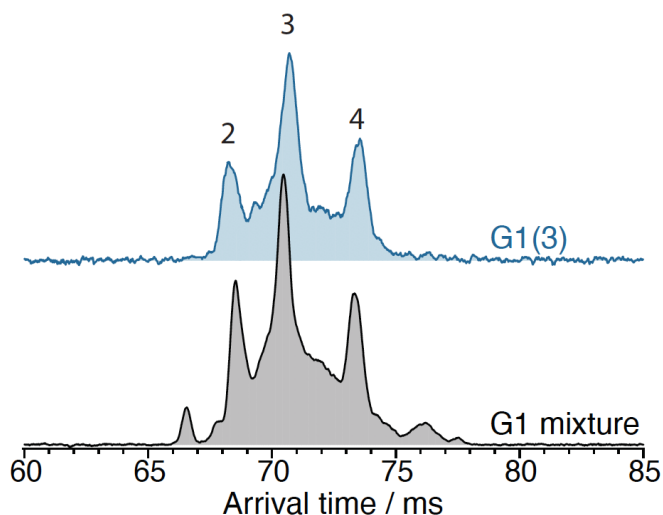


Figure 5.9 Comparison of the ATDs of the G1 standard in grey and the remaining G1(3) after the enzymatic reaction in blue.

To explore this further, we recorded cryogenic IR spectra of the G1 standard and the remaining G1 without ion mobility selection, shown in Figure 5.10. The spectra differ in some vibrational bands, such as 3615 cm^{-1} , 3637 cm^{-1} , 3670 cm^{-1} in the free OH region, and some differences were observed in the weakly hydrogen-bonded region. Our data shows that the spectra are distinctive.

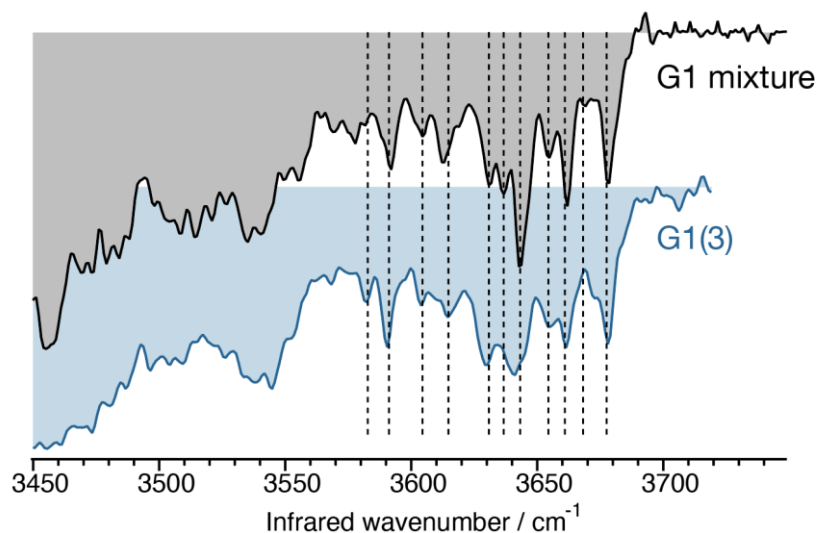


Figure 5.10 *Comparison of the cryogenic spectra of the G1 standard in grey and the remaining G1(3) after the enzymatic reaction in blue.*

We then recorded cryogenic IR spectra of each peak in the ATD of the remaining G1 and compared them to the respective G1 standard, shown in Figure 5.11. The spectrum of each ATD peak has unique well-resolved transitions in the free OH stretch region ($3580\text{--}3700\text{ cm}^{-1}$) that can be used for identification. The IR spectra of the second peak in the ATD of reference and remaining G1 (Figure 5.11 a) generally match in both the position and intensity of the vibrational bands, with only slight differences due to the low signal of the tagged ion intensity for the remaining G1. Spectra of the third peak (Figure 5.11 b) and the fourth peak (Figure 5.11 c) in the ATD of the reference and remaining G1 show a nearly perfect match in both the band positions and intensities. This result showed that the second, the third and the fourth peaks in the ATD correspond to the G1(3) isomer, and the first peak and the fifth peaks for the G1 standard correspond to the G1(6) isomer. However, one question can remain; if these results correspond to the enzymatic reaction after 16h, it might mean that the G1(6) can be present as an impurity in the remaining G1 sample. Future studies, such as performing a similar analysis of the remaining G1 after 48h of chemoenzymatic reaction, are recommended to further validate these findings.

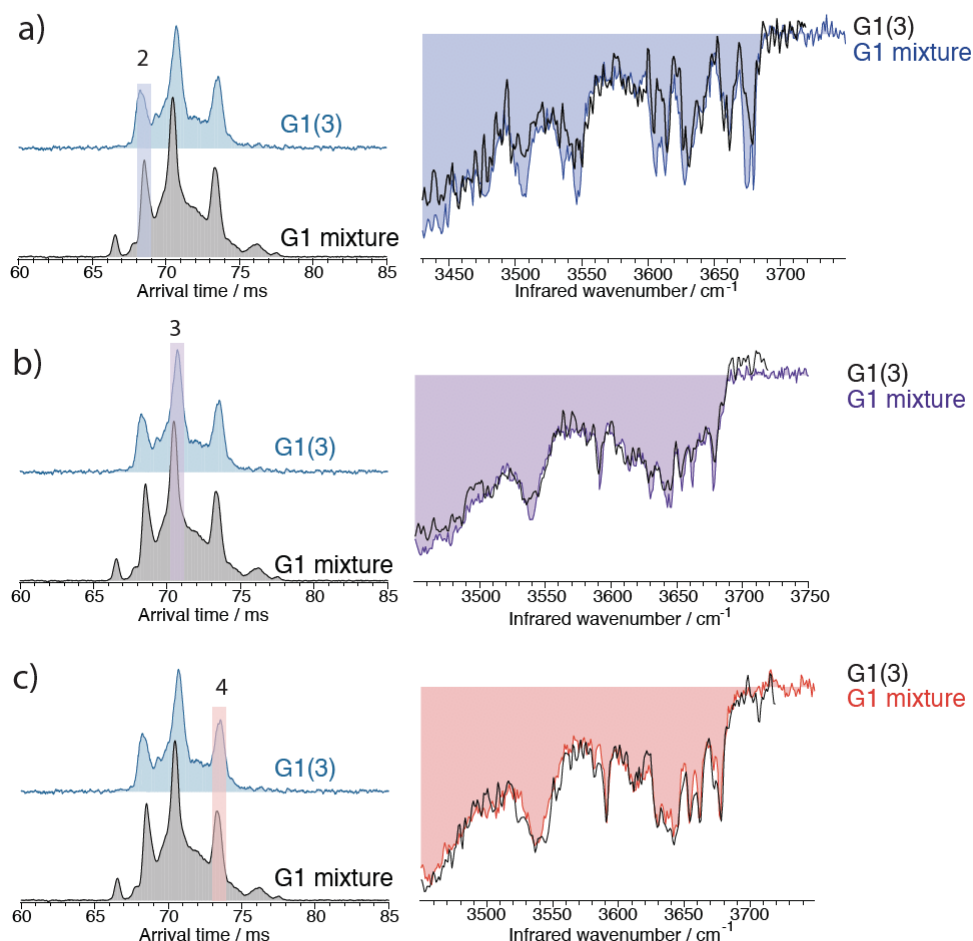


Figure 5.11 Comparison of cryogenic spectra of a) the G1 mixture in blue and the remaining G1(3) in grey after the enzymatic reaction of the first sliced ATD peak, b) the G1 mixture in purple and the remaining G1(3) in grey after the enzymatic reaction of the second sliced ATD peak second, c) the G1 mixture in red and the remaining G1(3) in grey after the enzymatic reaction of the third sliced ATD peak.

5.4 Conclusions and future perspectives

Our results demonstrate that the combination of SLIM-based ion mobility separation with cryogenic vibrational spectroscopy provides diagnostic glycan fingerprints to distinguish positional isomers of glycans. We used a chemoenzymatic approach using the FUT8 enzyme for the selective synthesis of G1F(6), which we employ as a standard for assigning the positional isomers of G1F to peaks. These

results constitute an excellent initial step toward isomeric identification. Work is ongoing to collect the cryogenic spectral signature of each ion mobility peak of the synthesized glycans, which will further strengthen our conclusions. We also plan on applying this powerful technique for determining the difference in the ratio of the individual G1F positional isomers from mAbs that have been produced from two different cell lines: CHO and HEK-293. This work provides an approach to identify positional isomers in standards which are not commercially available, it further allows monitoring of the consistency of galactosylation in N-glycans cleaved of mAbs. These findings pave the way towards a reliable and effective technique for bioanalytical applications.

References:

- (1) Ohtsubo, K.; Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* **2006**, 126 (5), 855–867. <https://doi.org/10.1016/j.cell.2006.08.019>.
- (2) Zhao, Y.-Y.; Takahashi, M.; Gu, J.-G.; Miyoshi, E.; Matsumoto, A.; Kitazume, S.; Taniguchi, N. Functional Roles of N-Glycans in Cell Signaling and Cell Adhesion in Cancer. *Cancer Sci.* **2008**, 99 (7), 1304–1310. <https://doi.org/10.1111/j.1349-7006.2008.00839.x>.
- (3) Corfield, A. P.; Berry, M. Glycan Variation and Evolution in the Eukaryotes. *Trends Biochem. Sci.* **2015**, 40 (7), 351–359. <https://doi.org/10.1016/j.tibs.2015.04.004>.
- (4) Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, 27 (1), 3–49. <https://doi.org/10.1093/glycob/cww086>.
- (5) Helenius, A.; Aebi, M. Intracellular Functions of N-Linked Glycans. *Science* **2001**, 291 (5512), 2364–2369. <https://doi.org/10.1126/science.291.5512.2364>.
- (6) Helenius, A.; Aebi, M. Roles of N-Linked Glycans in the Endoplasmic Reticulum. *Annu. Rev. Biochem.* **2004**, 73 (1), 1019–1049. <https://doi.org/10.1146/annurev.biochem.73.011303.073752>.
- (7) Reusch, D.; Tejada, M. L. Fc Glycans of Therapeutic Antibodies as Critical Quality Attributes. *Glycobiology* **2015**, 25 (12), 1325–1334. <https://doi.org/10.1093/glycob/cwv065>.
- (8) Higel, F.; Seidl, A.; Sörgel, F.; Friess, W. N-Glycosylation Heterogeneity and the Influence on Structure, Function and Pharmacokinetics of Monoclonal Antibodies and Fc Fusion Proteins. *Eur. J. Pharm. Biopharm.* **2016**, 100, 94–100. <https://doi.org/10.1016/j.ejpb.2016.01.005>.
- (9) Zhang, P.; Woen, S.; Wang, T.; Liao, B.; Zhao, S.; Chen, C.; Yang, Y.; Song, Z.; Wormald, M. R.; Yu, C.; Rudd, P. M. Challenges of Glycosylation Analysis and Control: An Integrated Approach to Producing Optimal and Consistent

Therapeutic Drugs. Drug Discov. Today **2016**, 21 (5), 740–765.
<https://doi.org/10.1016/j.drudis.2016.01.006>.

(10) Moremen, K. W.; Tiemeyer, M.; Nairn, A. V. Vertebrate Protein Glycosylation: Diversity, Synthesis and Function. Nat. Rev. Mol. Cell Biol. **2012**, 13 (7), 448–462. <https://doi.org/10.1038/nrm3383>.

(11) Aebi, M. N-Linked Protein Glycosylation in the ER. Biochim. Biophys. Acta BBA - Mol. Cell Res. **2013**, 1833 (11), 2430–2437.
<https://doi.org/10.1016/j.bbamcr.2013.04.001>.

(12) Spearman, M.; Butler, M. Glycosylation in Cell Culture; **2015**; Vol. 9, pp 237–258. https://doi.org/10.1007/978-3-319-10320-4_9.

(13) Li, F.; Vijayasankaran, N.; Shen, A. (Yijuan); Kiss, R.; Amanullah, A. Cell Culture Processes for Monoclonal Antibody Production. mAbs **2010**, 2 (5), 466–477. <https://doi.org/10.4161/mabs.2.5.12720>.

(14) Parekh, R. B.; Dwek, R. A.; Sutton, B. J.; Fernandes, D. L.; Leung, A.; Stanworth, D.; Rademacher, T. W.; Mizuochi, T.; Taniguchi, T.; Matsuta, K.; Takeuchi, F.; Nagano, Y.; Miyamoto, T.; Kobata, A. Association of Rheumatoid Arthritis and Primary Osteoarthritis with Changes in the Glycosylation Pattern of Total Serum IgG. Nature **1985**, 316 (6027), 452–457.
<https://doi.org/10.1038/316452a0>.

(15) Jefferis, R. Rheumatoid Factors, B Cells and Immunoglobulin Genes. Br. Med. Bull. **1995**, 51 (2), 312–331.
<https://doi.org/10.1093/oxfordjournals.bmb.a072963>.

(16) Peschke, B.; Keller, C. W.; Weber, P.; Quast, I.; Lünemann, J. D. Fc-Galactosylation of Human Immunoglobulin Gamma Isotypes Improves C1q Binding and Enhances Complement-Dependent Cytotoxicity. Front. Immunol. **2017**, 8. <https://doi.org/10.3389/fimmu.2017.00646>.

(17) Kiyoshi, M.; Caaveiro, J. M. M.; Tada, M.; Tamura, H.; Tanaka, T.; Terao, Y.; Morante, K.; Harazono, A.; Hashii, N.; Shibata, H.; Kuroda, D.; Nagatoishi, S.; Oe, S.; Ide, T.; Tsumoto, K.; Ishii-Watabe, A. Assessing the Heterogeneity of the

Fc-Glycan of a Therapeutic Antibody Using an Engineered Fc γ Receptor IIIa-Immobilized Column. *Sci. Rep.* **2018**, 8 (1), 1–11. <https://doi.org/10.1038/s41598-018-22199-8>.

(18) Hodoniczky, J.; Zheng, Y. Z.; James, D. C. Control of Recombinant Monoclonal Antibody Effector Functions by Fc N-Glycan Remodeling in Vitro. *Biotechnol. Prog.* **2005**, 21 (6), 1644–1652. <https://doi.org/10.1021/bp050228w>.

(19) Aoyama, M.; Hashii, N.; Tsukimura, W.; Osumi, K.; Harazono, A.; Tada, M.; Kiyoshi, M.; Matsuda, A.; Ishii-Watabe, A. Effects of Terminal Galactose Residues in Mannose A1-6 Arm of Fc-Glycan on the Effector Functions of Therapeutic Monoclonal Antibodies. *mAbs* **2019**, 11 (5), 826–836. <https://doi.org/10.1080/19420862.2019.1608143>.

(20) Takahashi, N.; Ishii, I.; Ishihara, H.; Mori, M.; Tejima, S.; Jefferis, R.; Endo, S.; Arata, Y. Comparative Structural Study of the N-Linked Oligosaccharides of Human Normal and Pathological Immunoglobulin G. *Biochemistry* **1987**, 26 (4), 1137–1144. <https://doi.org/10.1021/bi00378a023>.

(21) Wormald, M. R.; Rudd, P. M.; Harvey, D. J.; Chang, S.-C.; Scragg, I. G.; Dwek, R. A. Variations in Oligosaccharide–Protein Interactions in Immunoglobulin G Determine the Site-Specific Glycosylation Profiles and Modulate the Dynamic Motion of the Fc Oligosaccharides. *Biochemistry* **1997**, 36 (6), 1370–1380. <https://doi.org/10.1021/bi9621472>.

(22) Hamm, M.; Wang, Y.; Rustandi, R. R. Characterization of N-Linked Glycosylation in a Monoclonal Antibody Produced in NS0 Cells Using Capillary Electrophoresis with Laser-Induced Fluorescence Detection. *Pharmaceuticals* **2013**, 6 (3), 393–406. <https://doi.org/10.3390/ph6030393>.

(23) Liu, Y.; Salas-Solano, O.; Gennaro, L. A. Investigation of Sample Preparation Artifacts Formed during the Enzymatic Release of N-Linked Glycans Prior to Analysis by Capillary Electrophoresis. *Anal. Chem.* **2009**, 81 (16), 6823–6829. <https://doi.org/10.1021/ac9010588>.

(24) Melmer, M.; Stangler, T.; Premstaller, A.; Lindner, W. Comparison of Hydrophilic-Interaction, Reversed-Phase and Porous Graphitic Carbon

Chromatography for Glycan Analysis. *J. Chromatogr. A* **2011**, 1218 (1), 118–123. <https://doi.org/10.1016/j.chroma.2010.10.122>.

(25) Higel, F.; Demelbauer, U.; Seidl, A.; Friess, W.; Sörgel, F. Reversed-Phase Liquid-Chromatographic Mass Spectrometric N-Glycan Analysis of Biopharmaceuticals. *Anal. Bioanal. Chem.* **2013**, 405 (8), 2481–2493. <https://doi.org/10.1007/s00216-012-6690-3>.

(26) Hu, Y.; Shihab T.; Shiyue Z.; Kerry W.; Yehia Mechref. LC-MS/MS of Permethylated N-glycans Derived from Model and Human Blood Serum Glycoproteins. *Electrophoresis* **2016**, 37, 1498–1505. <https://doi.org/10.1002/elps.201500560>.

(27) Ashwood, C.; Lin, C.-H.; Thaysen-Andersen, M.; Packer, N. H. Discrimination of Isomers of Released N- and O-Glycans Using Diagnostic Product Ions in Negative Ion PGC-LC-ESI-MS/MS. *J. Am. Soc. Mass Spectrom.* **2018**, 29 (6), 1194–1209. <https://doi.org/10.1021/jasms.8b05838>.

(28) Yamaguchi, Y.; Nishima, W.; Re, S.; Sugita, Y. Confident Identification of Isomeric N-Glycan Structures by Combined Ion Mobility Mass Spectrometry and Hydrophilic Interaction Liquid Chromatography. *Rapid Commun. Mass Spectrom.* **2012**, 26 (24), 2877–2884. <https://doi.org/10.1002/rcm.6412>.

(29) Cosgrave, E. F. J.; Struwe, W. B.; Hayes, J. M.; Harvey, D. J.; Wormald, M. R.; Rudd, P. M. N-Linked Glycan Structures of the Human Fcγ Receptors Produced in NS0 Cells. *J. Proteome Res.* **2013**, 12 (8), 3721–3737. <https://doi.org/10.1021/pr400344h>.

(30) Pallister, E. G.; Choo, M. S. F.; Walsh, I.; Tai, J. N.; Tay, S. J.; Yang, Y. S.; Ng, S. K.; Rudd, P. M.; Flitsch, S. L.; Nguyen-Khuong, T. Utility of Ion-Mobility Spectrometry for Deducing Branching of Multiply Charged Glycans and Glycopeptides in a High-Throughput Positive Ion LC-FLR-IMS-MS Workflow. *Anal. Chem.* **2020**. <https://doi.org/10.1021/acs.analchem.0c01954>.

(31) Calderon, A. D.; Liu, Y.; Li, X.; Wang, X.; Chen, X.; Li, L.; Wang, P. G. Substrate Specificity of FUT8 and Chemoenzymatic Synthesis of Core-Fucosylated

Asymmetric N-Glycans. *Org. Biomol. Chem.* **2016**, 14 (17), 4027–4031. <https://doi.org/10.1039/C6OB00586A>.

(32) Clausen, H.; Wandall, H. H.; Steentoft, C.; Stanley, P.; Schnaar, R. L. Glycosylation Engineering. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2015.

(33) Zhang, Q.; Li, Z.; Song, X. Preparation of Complex Glycans From Natural Sources for Functional Study. *Front. Chem.* **2020**, 8. <https://doi.org/10.3389/fchem.2020.00508>.

(34) Calderon, A. D.; Zhou, J.; Guan, W.; Wu, Z.; Guo, Y.; Bai, J.; Li, Q.; Wang, P. G.; Fang, J.; Li, L. An Enzymatic Strategy to Asymmetrically Branched N-Glycans. *Org. Biomol. Chem.* **2017**, 15 (35), 7258–7262. <https://doi.org/10.1039/C7OB01765K>.

(35) Schneider, M.; Al-Shareffi, E.; Haltiwanger, R. S. Biological Functions of Fucose in Mammals. *Glycobiology* **2017**, 27 (7), 601–618. <https://doi.org/10.1093/glycob/cwx034>.

(36) Plasencia, M. D.; Isailovic, D.; Merenbloom, S. I.; Mechref, Y.; Clemmer, D. E. Resolving and Assigning N-Linked Glycan Structural Isomers from Ovalbumin by IMS-MS. *J. Am. Soc. Mass Spectrom.* **2008**, 19 (11), 1706–1715. <https://doi.org/10.1016/j.jasms.2008.07.020>.

(37) Olson, U.; Säwén, E.; Stenutz, R.; Wildmalm, G. Conformational Flexibility and Dynamics of Two (1→6)-Linked Disaccharides Related to an Oligosaccharide Epitope Expressed on Malignant Tumour Cells. *Chem. Weinh. Bergstr. Ger.* **2009**, 15 (35), 8886–8894. <https://doi.org/10.1002/chem.200900507>.

(38) Wooten, E. W.; Bazzo, R.; Edge, C. J.; Zamze, S.; Dwek, R. A.; Rademacher, T. W. Primary Sequence Dependence of Conformation in Oligomannose Oligosaccharides. *Eur. Biophys. J.* **1990**, 18 (3), 139. <https://doi.org/10.1007/BF02427373>.

- (39) Jo, S.; Qi, Y.; Im, W. Preferred Conformations of N-Glycan Core Pentasaccharide in Solution and in Glycoproteins. *Glycobiology* **2016**, 26 (1), 19–29. <https://doi.org/10.1093/glycob/cwv083>.

Conclusions and future perspectives

This thesis discusses the essential aspects of the database approach that we created using glycan standards. This approach enables glycan identification together with all its possible isomers by the combination of ultrahigh-resolution IMS based on SLIM technology together with cryogenic, messenger-tagging, IR spectroscopy. The results shown in Chapter 4 demonstrate the utility of the approach using numerous examples. Our ultimate goal is to create a user-friendly database that contains mass, ion mobility, and vibrational fingerprint spectra for each existing N-glycan similarly to those measured in Chapter 4. For glycans for which we do not have standards, we can use enzymatic strategies [1] together with collision induced dissociation techniques on the SLIM-IMS platform [2] to reconstruct parent glycan structures and include them in the database.

We showed in a systematic study in Chapter 4 that the combination of ultrahigh-resolution IMS based on SLIM technology together with cryogenic, messenger-tagging, IR spectroscopy can unambiguously identify N-linked glycans released from a therapeutic protein, such as Etanercept, using the targeted database approach. This study serves as a proof of principle that our approach has sufficient sensitivity, speed, and resolution for large-scale N-glycan characterization and identification. This approach should simplify basic research and process development for novel therapeutics and biosimilar products. Also, our technique could eventually lead to the development of a user-friendly, commercial instrument.

While our approach is powerful, several challenges remain that need to be solved. One of the priority tasks is to increase the throughput of our technique, as there are ~82'000 known glycans [3]. Multiplexing spectroscopy strategies, such as the Hadamard Transform, are ways forward and will help make our measurements even more suitable for incorporation into existing analytical workflows. Also, software for autonomous glycan identification using a spectroscopic database will also be necessary to eradicate the need for routine but labor intensive analyses. While Chapter 4

presents a work in progress, it demonstrates that our technique applies to the analysis of therapeutic proteins and mAbs.

In Chapter 5, we addressed the issue of distinguishing positional isomers of N-glycans that are not commercially available. For that, we synthesized one of the N-glycan isomers that is not available on the market using a highly specific enzyme. We developed and adapted an optimized enzymatic strategy for the efficient synthesis of core-fucosylated asymmetric N-glycans. Our results demonstrated that our technique provides diagnostic glycan fingerprints to distinguish the isomeric glycan species. In future experiments, we will record each cryogenic spectral signature of each ion mobility peak of the synthesized glycans to confirm our conclusions about their structures. Moreover, we plan to distinguish the positional isomers of N-glycans that are cleaved from the identical mAb created in CHO and HEK cell lines using our technique.

Even though numerous challenges remain, our technique has great potential to become a broadly accessible technique to facilitate the current glycan analyses.

References:

- (1) Dyukova, I.; Carrascosa, E.; Pellegrinelli, R. P.; Rizzo, T. R. Combining Cryogenic Infrared Spectroscopy with Selective Enzymatic Cleavage for Determining Glycan Primary Structure. *Anal. Chem.* **2020**, *92* (2), 1658–1662. <https://doi.org/10.1021/acs.analchem.9b04776>.
- (2) Bansal, P.; Yatsyna, V.; AbiKhodr, A. H.; Warnke, S.; Ben Faleh, A.; Yalovenko, N.; Wysocki, V. H.; Rizzo, T. R. Using SLIM-Based IMS-IMS Together with Cryogenic Infrared Spectroscopy for Glycan Analysis. *Anal. Chem.* **2020**. <https://doi.org/10.1021/acs.analchem.0c01265>.
- (3) Glycan Repository <https://glytoucan.org/> (accessed Sep 23, 2020)

Chapter 6 Glycan quantification

The development of reliable analytical methods for glycan analysis requires not only identification but also quantification. For example, it is essential for monitoring glycan changes that can be possibly associated with disease progression. However, current techniques can only partially provide the required information.

The section 6.1 summarizes isomeric glycan complexity and provides an overview of current methods to quantify glycans. Moreover, the main validation parameters that need to be considered to become a reliable, quantitative analytical method are described. Section 6.2 describes the experimental set-ups that we have used to attempt the quantification of glycans, including a detailed description of the tandem mass spectrometer for cryogenic ion spectroscopy. The ion mobility spectrometer coupled to a cryogenic trap and TOF is briefly introduced here, having been described in detail in Chapter 3. Section 6.3 outlines the results of the quantitation of glycan mixtures. Here, we describe the attempts to quantify the glycan mixtures and discuss the current limiting factors together with the possible solutions.

6.1 The problem of glycan quantification

As described in Chapter 1, the inherent isomeric complexity of glycans poses a daunting analytical challenge. This challenge arises from the fact that isomeric glycans can strongly vary in monosaccharide composition, connectivity, and branching. It has been estimated that the amount of all possible linear and branched isomers of a hexasaccharide is greater than 1×10^{12} [1].

An ideal analytical technique for glycans should both distinguish the isomeric structure and provide a relative, if not absolute, measure of its concentration in an analyzed sample. We have demonstrated that our technique is sensitive and rapid enough to analyze glycans cleaved of biological samples. However, our long-term goal is to be able to both identify and quantify mixtures of isomeric oligosaccharides.

Current methods to identify glycans have been described in detail in Chapter 2. Here, we cover strategies that are employed for glycan quantification. These strategies normally use separation and mass spectrometric techniques, including capillary electrophoresis, liquid chromatography (different types), MALDI-MS, and ESI-MS[2–4]. Additionally, Hoffman *et al.* demonstrated the possibility of semi-quantitative analysis for relatively small glycans using IMS-MS [5].

Quantification can be divided into two types: absolute and relative/or semi quantitative. Absolute quantification allows one to obtain the precise sample concentrations for individual glycans, whereas relative quantification shows only a percentage/ratio of the normalized glycan content. Absolute quantification is challenging, since samples are usually studied independently by comparison of LC or MS data, where subtle qualitative and/or quantitative differences are frequently neglected. Since the composition of individual glycans has a significant impact on glycan function, absolute glycan quantification is biologically relevant and provides an overview of pathological change in glycosylation pattern.

Glycans generally require derivatization with a UV/ fluorescence label for their detection after separation using chromatographic or electrophoretic separation methods [6,7]. These labeling strategies allow the conversion of the detected signals to the absolute concentration of glycans by using a calibration curve [8]. However, derivatization requires additional experimental steps in the workflow, during which time sample loss and loss of glycan modifications might occur [6-8].

Developments in MS instrumentation reported absolute quantitation [8,9]. While modern MS techniques allow the detection of glycan ions in the low-femtomole range [10], the matrices used in MALDI-MS can strongly affect the ability to quantify glycans [9]. In addition, the formation of multiple metal adducts and the presence of different charge states of the same glycan can make quantitation more difficult.

Relative quantitation using HPLC/CE methods with UV/ fluorescence detection is done by comparison of the relative peak areas, which correspond directly to their relative concentrations [11–13]. However, there are significant drawbacks linked with the derivatization step required for these methods.

Other popular methods of relative glycan quantitation are based on LC/CE-MS. In these methods, the peak areas of extracted ion chromatograms (EIC) of glycan ions can be obtained over the range of identified glycans. Later, their relative abundances can be determined and denoted either as a ratio of the total glycan EIC areas [14] or normalized with respect to the most intense peak [15]. Making the weak assumption that glycans have identical ionization efficiencies over different runs in different samples, and that the relative intensities can be compared. Slight inaccuracies might occur due to discrepancies in the experimental run, for example even minor differences in the glycan elution patterns may skew the relative ionization efficiencies. Stable isotope glycan quantification can be performed to overcome those issues by spiking isotope enriched standards into the sample or by labeling the entire glycan mixture with a heavy-isotope tag. Some well-known labels include incorporating ^{13}C , or ^{15}N hydrazide on the reducing terminus [16–18] or permethylation of the hydroxyl groups (using $^{13}\text{CH}_3\text{I}$ or $^{12}\text{CH}_2\text{DI}$ for example)[19]. However, the resolution of the instrument should be high. The drawback is that the labeling limitation still exists.

6.1.1 *Parameters for method validation*

To become a quantitative analytical method, the method must satisfy certain validation criteria. Although not all of the validation criteria are relevant for all types of experiments, standard validation characteristics to be considered are [20,21]:

- 1) selectivity (specificity),
- 2) linearity,
- 3) dynamic range,
- 4) accuracy,
- 5) precision,
- 6) limit of detection (LOD) and limit of quantitation (LOQ).

Selectivity refers to the ability of the validated method to analyze a certain analyte in any type of mixture without interference with other compounds in the mixture. The selectivity must be validated by showing the absence of interference peaks with respect to impurities and the matrix.

The linearity of an analytical method is the ability to obtain within a certain working range, data that is proportional to the concentration in the sample, either

directly or using a precise mathematical fit. The linearity is typically evaluated by plotting the signal height or peak area with respect to the concentration of standard analyte solutions. The criterion of linearity is normally a high correlation coefficient (r) of 0.99. If r is < 0.99 , an explanation must be provided of how accurate calibration is to be maintained. Additionally, in experiments where a non-linear response is intentionally applied, an explanation must be provided. Linear range can be found from the linearity assessment experiments. The range of an analytical method is the interval from the lower to the upper concentration of analyte with a sufficient level of accuracy, precision, and linearity. The working range is part of the full linear range.

Accuracy is defined as the difference between the expected and the accepted reference value within the confidence intervals, expressed as a percentage. Accuracy and precision together determine the total error of the analysis. For example, accuracy can be estimated by analyzing a sample with a known concentration (control) and comparing this value with the true value of a blank matrix of interest. It should be measured at least nine times over a minimum of three concentrations.

Precision is a measure of statistical variability. It includes any random errors that occur in the experiments [20]. It consists of three parts: repeatability, intermediate precision, and reproducibility. Repeatability is the precision of measurement under the same conditions over a short period of time. Intermediate precision expresses variations within the same laboratory, such as different analysts, different days, etc. Reproducibility (ruggedness) shows the precision between laboratories performing the same experiment.

Precision is normally expressed as the relative standard deviation (RSD) of an analytical results obtained from independently prepared control standards. Precision is a concentration-dependent parameter, it is measured at the lower, middle, and upper parts within the working range.

An important aspect of our experimental approach is to make reproducible measurements under the same conditions with the same distribution of three-dimensional conformers of glycan in the gas phase. Improving these measures will allow us to identify glycans in mixtures by comparing their recorded mass, ATD, and infrared spectrum with our database [22].

Ion mobility spectrometry and cryogenic ion spectroscopy show that ions can have multiple stable gas-phase conformers, some of which can be kinetically trapped [23,24]. To simplify the experiment and exclude the possibility of different conformer distributions in different experiments, all the conformations are annealed in the gas phase to the lowest energy structures so that one is always measuring the ATD and vibrational spectra of the same conformational distribution.

It is important to consider the laser fluence, which is defined as the optical energy delivered per unit area. Because the sensitivity is proportional to the fluence, it should be optimized to provide the highest sensitivity while avoiding saturation of the oscillators with the highest absorption cross-sections. Samples should be irradiated homogeneously to have the same IR fluence at different points of the ion cloud. All this needs to be done in a reproducible manner.

Limit of detection (LOD) is the lowest possible amount of analyte in a sample which can be detected but not quantified. Signal-to-noise (S/N) is 3:1.

Limit of quantification (LOQ) is the lowest amount of analyte in a sample which can be quantified with acceptable accuracy and precision. LOQ is determined by the analysis of samples with known exact concentrations of analyte in each. S/N is 10:1 or better. To identify suitable precision, the calculated relative standard deviation percent of the precision of 6 repetitive injects is plotted with respect to the concentration.

6.2 Experimental methods

This section describes the experimental set-ups that have been used to attempt the quantification of glycans. The chapter provides detailed description of the tandem mass spectrometer for cryogenic ion spectroscopy. The ion mobility spectrometer coupled to a cryogenic trap and TOF is briefly introduced here but was described in detail in Chapter 3.

6.2.1 *Tandem mass-spectrometer for cryogenic spectroscopy*

Experiments have been performed on a home-built tandem mass-spectrometer coupled with a cryogenic octupole ion trap (Figure 6.1) [25]. The main components of

this machine are a nESI source, a cryogenic multipole ion trap, and a mass-selective ion detection system.

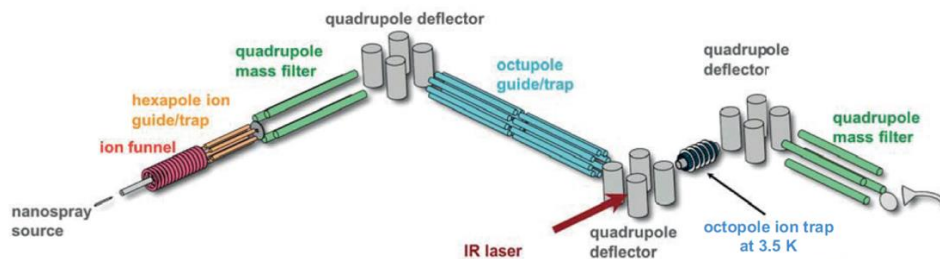


Figure 6.1 Schematic of the tandem mass spectrometer for cryogenic IR spectroscopy.

Singly-charged glycan ions are formed by a nESI source (Proxeon, ES070, Odense, Denmark) by spraying the sample solution in a continuous mode with a spray voltage of 1.2-1.8 kV. The emitters used for the experiments are metalized borosilicate gold/palladium-coated needles (ThermoFisher, Germany). The needle is inserted into a 0.5 mL Eppendorf reservoir filled with ~150 μ L of the sample solution. The details of the sample preparation will be provided in the subsequent sections. The needle is filled with the sample solution by slightly pressurizing it with nitrogen. The electrosprayed ions are then introduced into the first vacuum stage (~2 mbar) through a metal capillary. After exiting the capillary, the ions are focused into an electrodynamic ion funnel, which accumulates them prior to injection into the next pumping stage. The ion funnel is composed of a stack of ring electrodes that gradually decrease in diameter along its length to a minimum of 1.5 mm [3]. Additionally, a small DC electrode, so-called jet disrupter, is installed ~2 cm from the funnel entrance to disrupt the jet of neutral species and large droplets from proceeding directly to the next pumping stage [26]. An RF potential applied to electrodes is superimposed on a DC gradient along the axis of the funnel. This design of the ion funnel makes the signal more stable and allows to focus and accumulate the divergent beam of ions coming from the capillary. However, some collisional activation might occur due to the small aperture at the end of the funnel [27].

After exiting the funnel, ions are accumulated and pre-trapped in a two stage-hexapole ion trap (10^{-2} – 10^{-5} mbar). Trapping in the hexapole allows matching the

continuous nESI with the pulsed trapping in the cryogenic ion trap and the duty cycle of the pulsed laser. To store and release packets of ions, we change the voltage applied to the exit electrode from high to low.

The ions are then ejected from the hexapole and mass-selected by a quadrupole mass-filter (Q1) with an m/z range up to 2000 Th (Extrel, Pittsburgh, PA). After exiting Q1, the ions enter an electrostatic quadrupole deflector (Extrel), which diverts the ions 90° toward an octupole ion guide (Extrel). The quadrupole mass filter and quadrupole deflector are housed in the same differential pumping stage ($\sim 2 \times 10^{-8}$ mbar). The ions are guided through the RF octupole before being deflected 90° using a second quadrupole deflector and moved through a stack of five tube lenses, which are used to focus and decelerate the ions, before they enter a cold octupole ion trap maintained at ~ 4 K by a closed cycle helium cryostat (Sumitomo, SRDK-408D-W71D, Tokyo). The nominal trap temperature is measured by a silicon diode (Lakeshore, DT-670B-CU, Westerville, OH) [28].

Before the ions enter the trap (~ 1 – 2 ms), helium is pulsed into the trap for ~ 200 – 400 μ s. This timing allows for the helium to cool via collisions with the cold trap housing. The incoming ions undergo collisions with the cold He, which decelerates them and cools them to ~ 10 K [29].

In our experiments, we use messenger-tagging spectroscopy previously described in Chapter 3. Here, the He-tagged ions are irradiated every other trapping cycle (5 Hz) by a single IR pulse from a tunable optical parametric oscillator (OPO). When the OPO light is in resonance with a vibrational transition of the analyte, IR radiation is absorbed and then redistributed among its vibrational modes, leading to detachment of the weakly bound tag. The IR spectrum is obtained by monitoring the depletion of the signal at the tagged-ion mass as a function of the OPO wavenumber in a *laser-on/laser-off* experiment, which allows us to normalize the spectra to any long-term fluctuations in the parent ion signal.

Ions ejected from the cryogenic trap are bent 90° once more by a third quadrupole bender to be analyzed by a second quadrupole mass filter (Extrel). Ions are then detected in counting mode by first impinging on a conversion dynode (DeTech, 402-A-H) biased at -5 kV, which generates electrons. These electrons are drawn into a

channel electron multiplier, the entrance of which is biased at -1.9 kV, with the exit is at ground potential. After amplification, the resulting electron pulses are detected, and the data further transferred to our data collection computer.

6.2.2 *Ion mobility spectrometer coupled to a cryogenic trap and TOF*

For our work on quantitation, we have also performed ion mobility experiments. The machine in which these experiments were done has been described in Chapter 3, using the SLIM board shown in Figure 3.5.

Briefly, glycan ions are generated by a nESI source and introduced into the first vacuum stage through a stainless-steel capillary. After exiting the capillary, the ions enter an ion funnel trap (IFT), where they are focused and accumulated into packets. These packets are ejected from the IFT and transported into the SLIM device. Subsequently, m/z selection occurs in a quadrupole mass filter. Arrival time distributions are acquired by detecting the m/z -selected ions on a channeltron detector. The data are acquired using a WaveSurfer MXs-B Oscilloscope (Teledyne LeCroy SA, Switzerland) and processed using in-house control software written in Labview.

6.3 Quantification of glycan mixtures

6.3.1 *Sample preparation and notation*

We performed a series of experiments on disaccharides and trisaccharides to test our ability to do quantitative analysis of glycans. The samples were purchased from Carbosynth Ltd (United Kingdom) and used without further purification. The glycan purity is certified at a minimum of 90%. For each disaccharide and trisaccharide species, we prepared a solution of 50:50 LC grade methanol and nuclease-free water.

The structures of carbohydrates presented in this chapter are shown in Figure 6.2 using the SNFG nomenclature and notation. In panel a) we demonstrate the structures of the isomeric disaccharides that differ in glycosidic linkage type and monosaccharide sequence. Panel b) shows two isomeric trisaccharides that differ by the type glycosidic linkage (α / β) between the first two monosaccharide units. Panel

c) demonstrates the structures of the sugars that have been used to study the ionization efficiency.

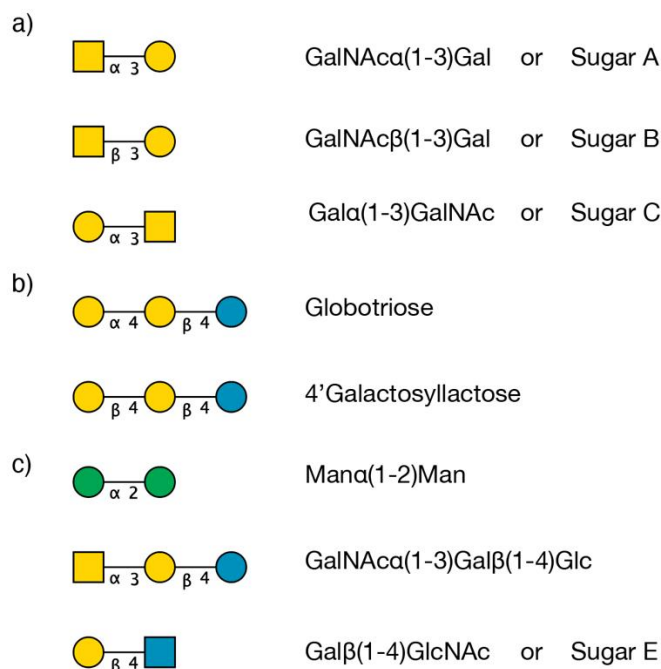


Figure 6.2 Schematic representations of the investigated carbohydrates using the SNFG nomenclature together with the notation. The quantitation experiments on isomeric disaccharides (panel a); The quantitation experiments on isomeric trisaccharides (panel b); The ionization efficiency experiments (panel c).

6.3.2 Quantifying vibrational spectra of disaccharide mixtures

To become a routine analytical technique, our approach should be able to quantify the components in mixtures of various complexity. Initial studies to quantify vibrational spectra were carried out using mixtures of disaccharides (Figure 6.2, panel a). We prepared all combinations of these disaccharides and measured their vibrational spectra, shown in Figures 6.3-6.6. All the measurements have been done under the same conditions on the same day. Moreover, each of the vibrational traces plotted on the figures listed below was an average of three vibrational spectra with identical power of the OPO.

We first measured vibrational spectra of two disaccharides: sugars A and B of Figure 6.3 a and b. Solutions of 75 μM of each sugar with 1 equivalent of NaCl salt were electrosprayed into our setup. We then recorded vibrational spectra of A/B mixtures 1:1 (Figure 6.3 c) and 2:1 (Figure 6.3 g) ratios, respectively. We then compared the experimental spectra to the synthetic ones (Figure 6.3 c and g) constructed from summing the individual spectra weighted by their respective concentrations. We then show the difference spectra (Figure 6.3 d and h).

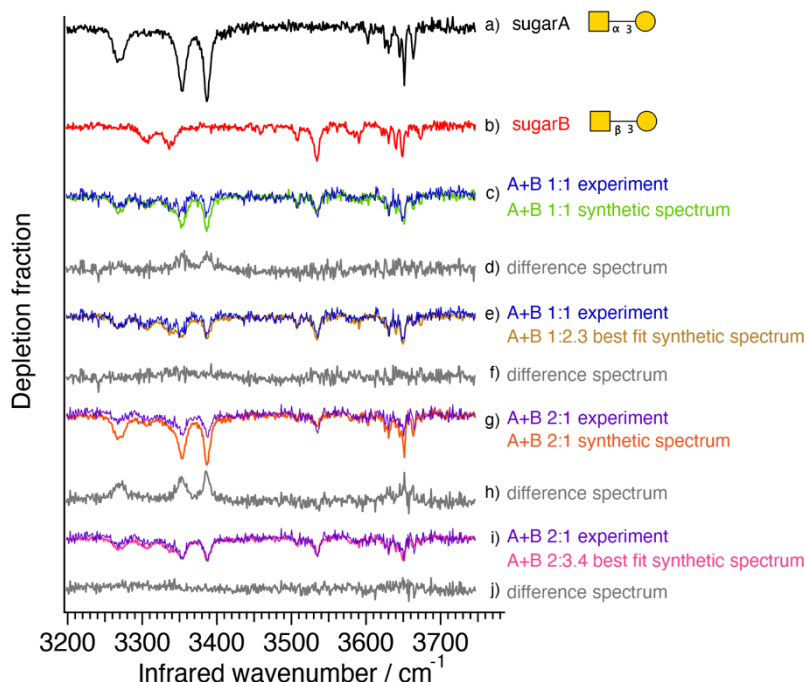


Figure 6.3 Comparison of experimental and synthetic spectra of sugars A and B (see details in the text). The spectra were recorded at 3.5 K using Helium as a tagging gas.

In Figures 6.3 a and b, we show spectroscopic fingerprints of sugars A and B, both at a concentration of 75 μM . Figure 6.3 c shows the experimental spectrum of 1:1 mixture at the same concentration (black trace) along with the 1:1 sum of the individual spectra (green trace). In Figure 6.3 d we display the difference spectrum of a and b. Figure e shows the experimental (blue trace) and best fit synthetic (light-brown trace) spectra of the same 1:1 mixtures, and f displays their difference. Figure 6.3 g displays the experimental spectrum (violet trace) and theoretical one (orange

trace) for a 2:1 mixture (150 μM for sugar A and 75 μM for sugar B), and Figure 6.3 h shows their difference. Figure 6.3 i compares the experimental spectrum (violet trace) and best fit synthetic spectrum (pink trace) for the same 2:1 mixture, and j displays their difference.

These experiments, shown in Figure 6.3 c and g, reveal that the experimental spectra of the mixtures are reproduced by a linear combination of the individual spectra with coefficients not equal to their relative concentration in solution. This is demonstrated by the difference spectra in Figures 6.3 d and h (grey traces), which show distinct features from the sugar A.

To obtain coefficients that do match the experimental spectra of the mixture A and B, we employed an Ordinary Least Squares (OLS) method [33,34]. Briefly, we are expecting to see the changes in the intensities for each band in IR spectra proportionally to the concentration. Our spectrum is a vector (approximately 500 points, 0, 5 step) for each measurement. A general algorithm is following to fit IR spectra:

For example, we have n known substances with $\mathbf{f}_i(\mathbf{v})$ known spectra, i.e. $i=1,2,3,\dots,n$ - number of the spectrum. We make mixtures and obtain \mathbf{g} spectra, which is a linear combination of recorded spectra. We find the fitting coefficients $\mathbf{a}_1, \dots, \mathbf{a}_n$, which can make $\mathbf{g}' = \mathbf{a}_1 \cdot \mathbf{f}_1(\mathbf{v}) + \dots + \mathbf{a}_n \cdot \mathbf{f}_n(\mathbf{v})$; maximally close to \mathbf{g} spectra, expressed as a closeness $\|\mathbf{g}' - \mathbf{g}\|$.

We can solve it by projecting a vector into a subspace defined by a non-orthogonal basis. The following scalar product $\mathbf{a}_{ij} = \int \mathbf{f}_i(\mathbf{v}) \cdot \mathbf{f}_j(\mathbf{v}) d\mathbf{v}$, where \mathbf{a}_{ij} corresponds to all possible known spectra. This scalar product $\mathbf{b}_j = \int \mathbf{g}(\mathbf{v}) \mathbf{f}_j(\mathbf{v}) d\mathbf{v}$, where \mathbf{b}_j corresponds to all our spectra. We then obtain the following equation:

$\mathbf{g} = \sum_i^n \alpha_i \cdot \mathbf{f}_i$, which is equal to

$$\int \mathbf{g}(\mathbf{v}) \cdot \mathbf{f}_i(\mathbf{v}) d\mathbf{v} = \sum_i^n \alpha_i \int \mathbf{f}_i(\mathbf{v}) \mathbf{f}_i(\mathbf{v}) d\mathbf{v} = \sum_i^n \alpha_i \cdot \mathbf{a}_{ij} = \mathbf{b}_j, \text{ where } \sum_i^n \alpha_i \cdot \mathbf{a}_{ij} = \mathbf{b}_j$$

This is a system of our equations what we need to solve to get α_i . Our closeness is expressed by

$$L(\alpha_i^j) = \int (\mathbf{g} - \sum_i^n \alpha_i^j \mathbf{f}_i)^2 d\mathbf{v},$$

where \mathbf{g} -calculated spectrum, α -coefficient to match the synthetic spectrum, \mathbf{f} -the recorded spectrum.

$$\frac{\partial L}{\partial \alpha_i^j} = \left(L(\alpha_1^j, \alpha_2^j, \dots, \alpha_i^j + \delta_\alpha, \dots, \alpha_n^j) - L(\alpha_1^j, \alpha_2^j, \dots, \alpha_i^j, \dots, \alpha_n^j) \right) \frac{1}{\delta_\alpha}, \text{ where}$$

$$\alpha_i^{j+1} = \alpha_i^j + \frac{\partial L}{\partial \alpha_i^j} \cdot \varepsilon.$$

We start from any α_i^0 , calculate for this value $\frac{\partial L}{\partial \alpha_i^0}$. Based on the result we calculate α_i^1 , ... etc. We calculate until $L < \varepsilon$.

The accuracy of the identification between the calculated and the measured IR spectra is evaluated by applying the root-mean-square deviation (RMSD):

$$\text{RMSD} = \sqrt{\frac{1}{N} \cdot \sum_{i=0}^{550} (I_{\text{exp}} - I_{\text{synth}})^2};$$

where N -number of points (550 points), I_{exp} - intensity of the recorded spectrum, I_{synth} - intensity of the calculated spectrum.

The coefficients resulting in the lowest RMSD value were accepted (see Appendix B, Tables B.1-B.2).

The results of the comparison of the experimental and the best fit synthetic spectra are shown in Figure 6.3 e and i. The coefficients for sugar B were significantly higher than that given by its concentration in solution. To illustrate this, Figure 6.3 e demonstrates that the pink trace of the synthetically derived spectrum with an A: B ratio of 1:2.3 instead of 1:1 matches the experimental (violet) trace almost perfectly. The difference spectrum shown in Figure 6.3 f (grey trace) confirms this.

To investigate this further, we repeated the experiment with mixtures of different compositions. Similar to the previous example, we measured individual vibrational spectra of another set of disaccharides: sugar B (Figure 6.4 a) and sugar C (Figure 6.4 b). Solutions of the same 75 μM concentration of each sugar with 1 equivalent of NaCl were used. After recording the individual spectra, we obtained vibrational spectra of mixtures of the sugars B and C with 1:1 (Figure 6.4 c, blue trace) and 2:1 (Figure 6.4 g, red trace) ratios, respectively. We then compared the experimental spectra to the synthetic ones (Figure 6.4 c, black trace; Figure 6.4 g, green trace) and

plotted the difference spectra (Figure 6.4 d and h). Similar to the previous experiment with sugars A and B, the measured spectra for the mixture of sugars B and C do not agree well with the synthetic spectra using coefficients proportional to the concentrations. The coefficient for sugar B in the best fit synthetic spectrum was a factor of three higher than its relative concentration (Figure 6.4 e, orange trace). The difference spectrum shown in Figure 6.4 f (grey trace) highlights this observation. For a 2:1 concentration ratio, the coefficient needed for sugar B to reproduce the spectrum of the mixture was 2.5 times higher than its relative concentration (Figure 6.4 i, light-blue trace) in the best fit synthetic spectrum than in the experimental one (Figure 6.4 i, red trace). Figure 6.4 j demonstrates the difference spectrum between the experiment and the calculated spectrum with the coefficients determined from the least-squares fit, showing a near perfect correspondence.

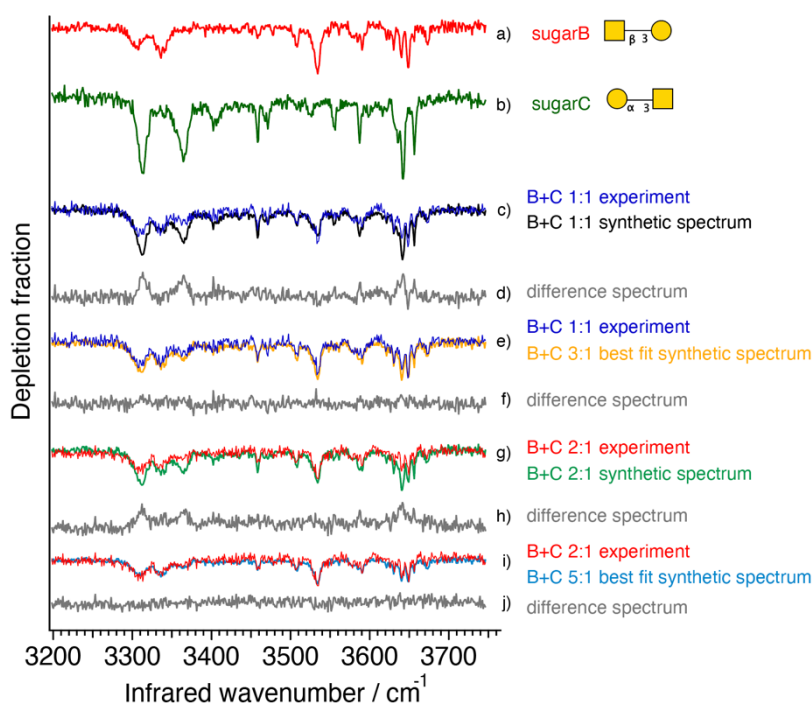


Figure 6.4 Spectroscopic fingerprints of sugars B and C. a) and b) individual spectra of the sugars at a concentration of $75 \mu\text{M}$. c)-f) mixtures at a 1:1 concentration ratio. g)-j) mixtures at a 2:1 concentration ratio, with sugar B at $150 \mu\text{M}$. The spectra were recorded at 3.5 K using Helium as a tagging gas.

The results from figures 6.3 and 6.4 suggest that different sugars have different ionization efficiencies, which need to be explicitly taken into account for quantification purposes. The ionization efficiency together with other possible parameters affecting quantification will be discussed in detail in the following subchapters.

We then studied a mixture of sugars A and C. We first recorded individual vibrational spectra (Figure 6.5 a and 6.5 b) at 75 μM with 1 equivalent of NaCl salt and then spectra of mixtures of each sugar with a 1:1 ratio (Figure 6.5 c, red trace). Figure 6.5 c (light-blue trace) shows the synthetic spectrum using a 1:1 weighting. Figure 6.5 d (grey trace) shows the difference spectrum, which is essentially featureless. This would suggest that the scale factors for sugars A and C are essentially the same. However if this were the case, the relative scale factors B vs. A and B vs. C should be the same, which they are not (see Figures 6.4 and 6.5). Moreover, the difference is greater than the uncertainty in the coefficients generated by the least-squares fit. This indicates that sugars A and C have similar ionization efficiency, while sugar B ionizes differently.

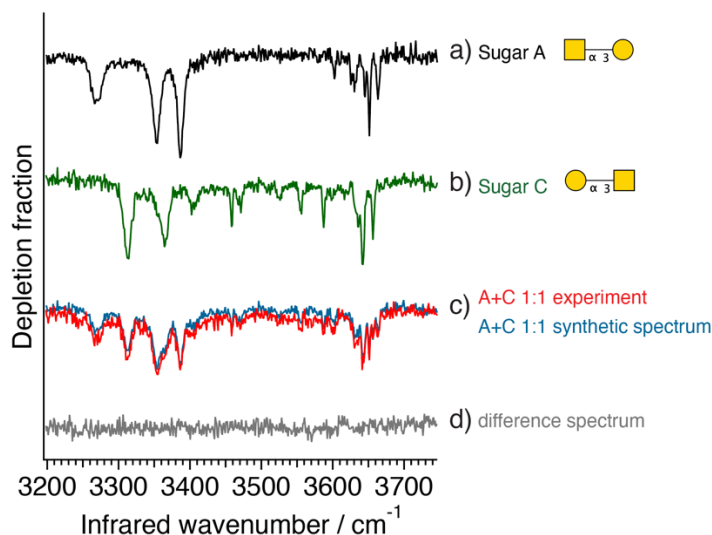


Figure 6.5 Comparison of experimental and synthetic spectra of sugars A and C (see details in the text). The spectra were recorded at 3.5 K using Helium as a tagging gas.

We also investigated a mixture of all three sugars, each at a concentration of 75 μM (Figure 6.6). The measured vibrational spectrum of the 1:1:1 mixture is shown

in Figure 6.6 d. We then compared the experimental spectrum with the synthetic one (Figure 6.6 d, red trace) using a 1:1:1 weighting of the individual spectra. The differences between the two spectra are clearly significant, as shown in Figure 6.6 e. The re-calculated synthetic spectrum (Figure 6.6 f, red trace) using coefficients determined by a least-squares fit of the measured spectrum of the mixture, matches well to the experimental spectrum, as demonstrated by the difference spectrum in Figure 6.6 g. This procedure yields optimum coefficients of 1:5:1 for A: B: C. Analogous to the results from figures 6.3 and 6.4, this experiment highlights the predominance of sugar B over sugars A and C in the mixture. Possible reasons for this observation are discussed below.

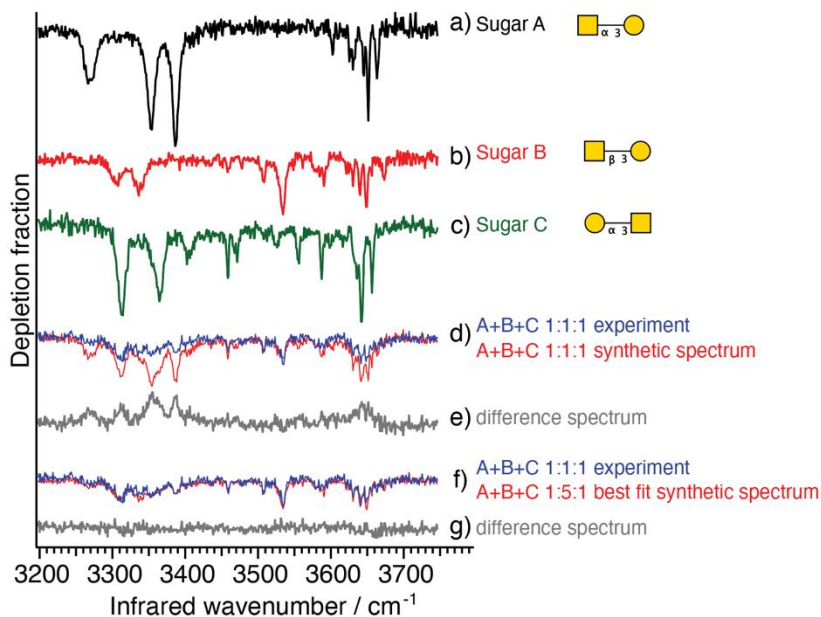


Figure 6.6 Comparison of experimental and synthetic spectra of sugars A, B and C (see details in the text). The spectra were recorded at 3.5 K using Helium as a tagging gas.

6.3.3 Quantifying vibrational spectra of trisaccharide mixtures

In Figure 6.7 a and b, we demonstrate spectroscopic fingerprints of sugars A and B, both at a concentration of 75 μ M. Figure 6.7 c displays the experimental spectrum of 1:1 mixture at the same concentration (blue trace) along with the 1:1 sum of the

individual spectra (green trace). Subsequently, we observed the analogous pattern for trisaccharides that the recorded spectrum of the mixture is a linear sum of their individual spectra with coefficients which differ from the expected experimental values. In fact, some expected bands of globotriose were low in their intensity. In Figure 6.7 d we show the difference spectrum of a and b. Figure 6.7 e shows the experimental (blue trace) and best fit synthetic (light-brown trace) spectra of the same 1:1 mixtures, and f displays their difference. From this observation, we concluded that the contribution of the 4'galactosyllactose, depicted as A, to the mixture was four times more than the contribution of the globotriose. This result might be related to the suppression of globotriose in presence of the 4'galactosyllactose sugar.

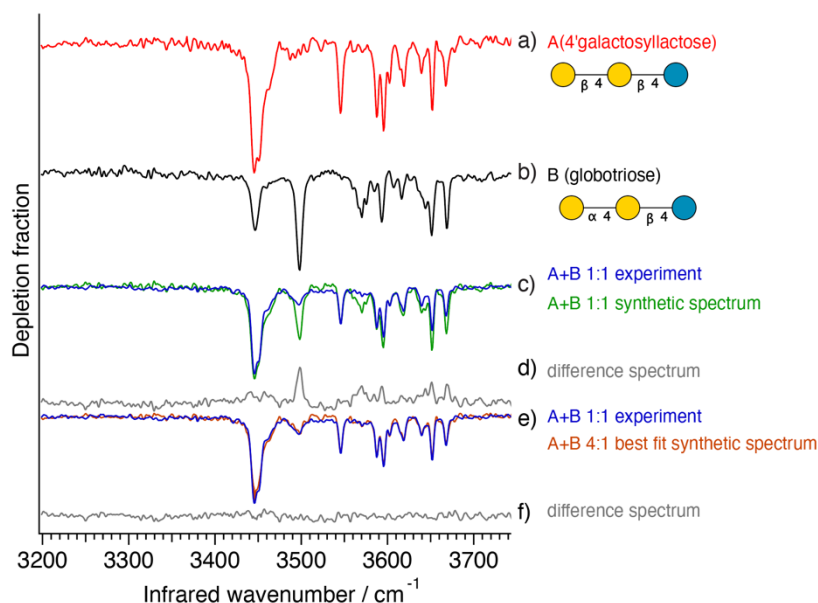


Figure 6.7 Comparison of experimental and synthetic spectra of sugars A and B (see details in the text). The spectra were recorded at 3.5 K using Helium as a tagging gas.

As we observed for mixtures of disaccharides, the similar tendency occurs for trisaccharides. We investigated further a mixture with a ratio 2:1 to understand how strong the suppression of globotriose is.

To investigate this further, we repeated the experiment with same mixtures of different composition. Solutions of the same 75 μ M concentration of each sugar with

1 equivalent of NaCl were used. After recording the individual spectra (Figure 6.8 a and b), we obtained vibrational spectra of mixtures with 2:1 (Figure 6.8 c, blue trace) (Figure 6.8 c, pink trace) ratios, respectively. We then compared the experimental spectrum to the synthetic one (Figure 6.8 c, pink trace) and plotted the difference spectra (Figure 6.8 d, grey trace). Similar to the previous experiment with ratio 1:1 of sugars A and B, the measured spectra for the ratio 2:1 of the mixture A and B do not agree well with the synthetic spectra using coefficients proportional to the concentrations. The coefficient for sugar A in the best fit synthetic spectrum was a factor of four higher than its relative concentration (Figure 6.8 e, light-blue trace). Similar to the aforementioned observations, this experiment underlines the predominance of 4'galactosyllactose over globotriose. Moreover, the same bands have low intensities as in the previously discussed example (Figure 6.8 c, blue trace).

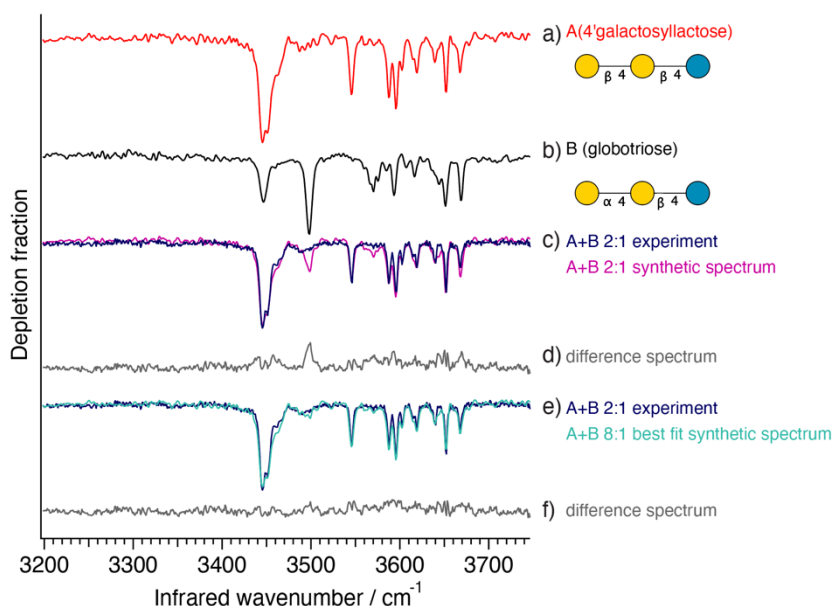


Figure 6.8 Comparison of experimental and synthetic spectra for sugars A and B (see details in the text). The spectra were recorded at 3.5 K using Helium as a tagging gas.

It is worth noting that the experimental spectrum of the mixture with a 10:1 ratio (Appendix B, Figure B.1) was identical to the experimental spectrum of the mixture with a 2:1 ratio. The synthetic spectra with an 8:1 ratio had a good agreement

with both experimental (10:1 and 2:1) spectra. There is a clear indication of the suppression limit of one of the listed sugars since they are both present in the mixture.

Generally speaking, we found that experimental spectra of the mixtures for disaccharides and trisaccharides show a lower than expected contribution of intensities from individual glycans. Sugars that differ in a type of glycosidic linkage (α / β) between the first two monosaccharide units have different ionization efficiency [35–37]. The glycans that have β glycosidic linkage dominate the ionization process over the glycans with α glycosidic linkage (see Figure 6.6-6.8). Also, charge competition between the analytes might take a place [38]. The tagging efficiency might differ for each of the molecules [39,40]. However, the spectrum of the mixture of glycans with different monosaccharide order: A and C is a linear sum with the coefficients proportional to the experimental values. We hypothesize that different monosaccharide order doesn't alter the ionization efficiency, since the primary monosaccharide content is unchanged.

Since our investigations so far have only been applied to a limited number of examples, the findings might not be generalized to all the glycan types. Further experiments on mixtures with different monosaccharide content and linkage should bring more clarity to the overall picture. The next subsections discuss the possible factors which limit the quantitation of the spectra collected on our set-up.

6.3.4 *Identifying limiting factors*

6.3.4.1 Electrospray

a) Ionization efficiency

The formation of the ions via nESI is the first step in our experiments. Figure 6.9 shows mass spectra for disaccharides complexated with 1 eq. of Na^+ : sugar A with three measurements (Figure 6.9, green, blue, and red traces, respectively), sugar B (Figure 6.9, black trace) and their mixtures with ratios of 1:1 (Figure 6.9, pink trace) and 2:1 (Figure 6.9, light-green trace). The intensities of sugar A are considerably lower than for sugar B. As anticipated above, the predominance of the latter might be related to its better ionization efficiency. Moreover, the intensity of the mixture 1:1 is higher than the intensity of the mixture 2:1. This observation can be linked to

the increased amount of the sugar A, which has lower ionization efficiency than the sugar B in the mixture. It is worth highlighting that each MS measurement of the presented three for sugar A varies significantly. The intensity of sugar C (Figure 6.9, brown trace) is lower comparing to sugar A and B. The intensity of the mixture A+C (Figure 6.9, light-blue trace) has approximately same intensities as intensities of the individual A and C sugars. However, the intensities of the B+C (Figure 6.9, violet trace) and A+B+C (Figure 6.9, orange trace) mixtures are higher than for the individual sugars. From these results, it is hard to withdraw the conclusions, given it varies a lot. We could not study the variability of the intensity signal for each sugar individually and in mixture since all species are isomers, and cannot be separated using only MS.

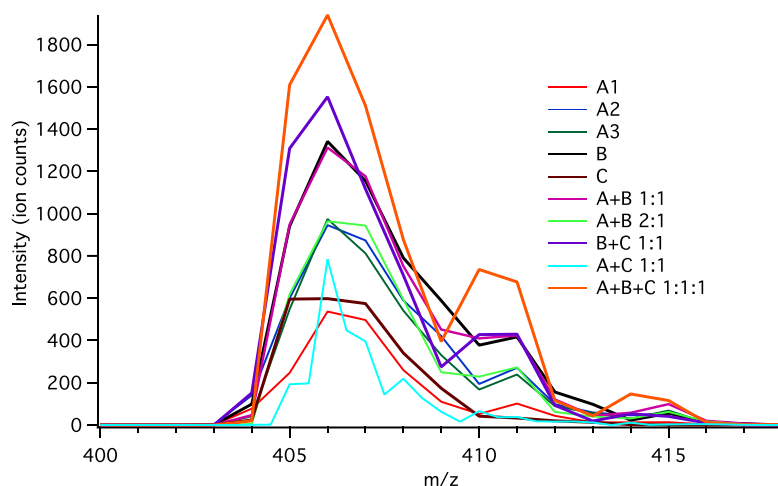
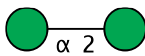

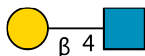


Figure 6.9 Mass spectra of disaccharides: sugar A (three traces in green, blue, and red, respectively); sugar B (black); sugar C (brown); A and B mixtures with 1:1 (pink) and 2:1 (light-green) ratios; B and C mixtures (violet) with 1:1 ratio; A and C mixtures (light blue) with 1:1 ratio; A, B and C mixtures (orange) with 1:1:1 ratio. Mass spectra were acquired with a 350–500 m/z range in positive ESI mode with a 0.5 m/z step size. Each spectrum used was the result of a 10-scan average to reduce the effects of any intensity fluctuations in the ESI.

Table 6.1 summarizes the variability of electrospray ionization for various types of carbohydrates (Figure 6.2, panel c) individually and in the mixtures with the identical concentration (75 μM). This summary shows the clear differences from

expected reduced intensities (ion counts) by 50% for binary mixtures and by 67% for ternary mixtures. Moreover, different ionization efficiencies varying by order of magnitude are observed. For example, B trisaccharide ionizes more efficiently than A and C disaccharides. Additionally, A disaccharide is suppressed while mixed with other carbohydrates. These results can be explained by a chemical structure of A disaccharide which lacks the amide group, leading to lower ionization comparing to other molecules.

Table 6.1 *Intensities of carbohydrates (ion counts) in individual system, binary and ternary mixtures.*

Type of Solution	Carbohydrates (ion counts)		
			
Individual	238	7020	725
In AB mixture	22	3999	
In BC mixture		2083	222
In AC mixture	84		395
In ABC mixture	49	1999	209

We can conclude from the aforementioned observations that a major source of unreliability and ion suppression is nESI ionization, since nESI responses can vary significantly among different sugar analytes with identical solution concentrations. This situation is no different than any mass spectrometry based detection technique.

However, despite numerous efforts to determine the exact mechanism of ion generation in ESI, there is still incomplete knowledge to predict the ionization efficiency of a selected compound in the ESI source. Numerous papers concluded that the complexity of the ESI process is fundamental [19,36,41,42]. The MS signals in general are known to vary from one day to another, due to variations in the ionization source conditions, pressures, cleanliness of the ion optics, etc. And ionization efficiency is affected by numerous factors, including ESI interface design, flow rate, solvent composition, a poorly cut nanospray emitter, and analyte properties (pK_b, logD, etc.),

presence of other compounds that can compete or interfere with the ionization of analyte that results in fluctuations in nESI response [41,42].

It was not possible to investigate the effect of each of the listed factors on our carbohydrate mixtures. However, we created a plan to modify the current set-up that might improve the ionization efficiency and spray stability. We plan on using the SLIM device to separate analytes from other compounds before detection. This implementation will eliminate the presence of undesired compounds before MS detection. Moreover, the most effective nESI operation requires the optimization of the interface for the significantly smaller sample infusion rate, more facile droplet desolvation, and the lower electrospray current [43]. To achieve spray stability and enhance ionization efficiency, we plan to install a commercial ion source together with a heated capillary inlet and the flow and pressure controller.

Several studies suggest using an internal standard (IST) for each of the analytes in the mixtures [44–46]. With an internal standard, one can determine the amount of the analyte by the response factor (intensity ratio of IST and analyte). For example, one of the most popular choices are isotopomers (isotopically labelled with ^{13}C , D, ^{15}N) which have identical ionization efficiency [4,7,47]. These molecules are ideal candidates since we study native carbohydrates.

b) Competition for charge

The charge competition effects on the ESI process happen in the presence of a mixture with high analyte concentrations [48]. To investigate the charge (Na^+) competition, solutions having equimolar concentrations of sugars A and B were prepared with different amounts of the NaCl salt. Figure 6.10 shows mass-spectra and cryogenic spectra of A + B mixture with different amount of Na^+ (1eq, 2q, 3 eq) in the solutions. The most intense peak in mass-scan corresponds to the mixture with 3 eq of Na^+ in the solution. However, the amount of the tag (410 m/z) for this mixture is equal to the mixture with 1 eq of Na^+ . This means that tagging efficiency is the best for the mixture with the least amount of sodium. The least abundant in mass spectrum belongs to the mixture with 2 eq of Na^+ .

From cryogenic spectra, the results demonstrate that the intensities of bands are higher in 3280-3400 cm^{-1} for 2 eq of Na, it might be an indication that the sugar A has

access to the charge. However, the spray is more unstable for 2 eq of Na (Figure 6.10, red trace) than for 1 eq (Figure 6.10, light-green trace) which leads to worse S/N. The mixture with 3 eq of Na^+ (Figure 6.10, violet trace) demonstrates overall lower intensities of the bands comparing to other analytes. This result might indicate that excessive concentrations of salt lead to ion suppression that masks ion signal.

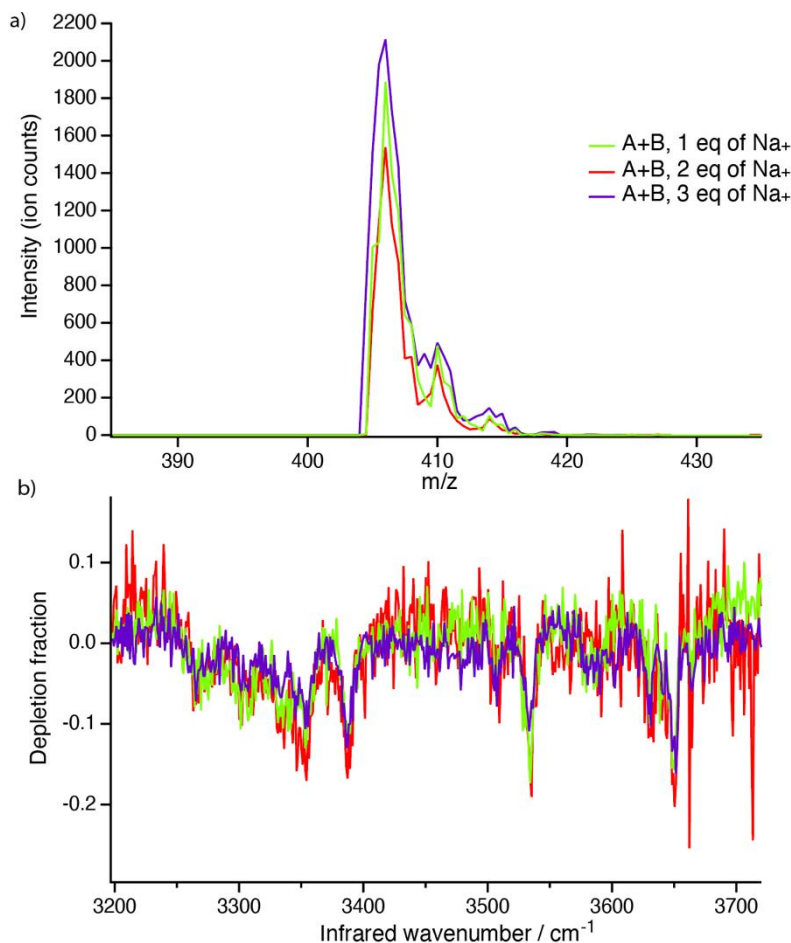


Figure 6.10 a) Mass-spectra and b) spectroscopic fingerprints of mixtures for sugar A and sugar B at a concentration of $75\ \mu\text{M}$ with 1 equivalent (light-green trace), 2 equivalents (red trace), 3 equivalents (violet trace) of NaCl salt. The spectra were recorded at $3.5\ \text{K}$ using Helium as a tagging gas.

We can conclude that better spraying solution, such as the mixture with 3 eq of Na^+ leads to suppression of band intensities. It might be explained by either saturation of the channeltron detector or poor overlap of the ion cloud with the IR beam.

However, it leads to better S/N since a large amount of ion provides a better statistic. Moreover, the low-intensity peaks provide the highest band intensities, but the S/N decreases a lot. Additionally, sodium chloride is a nonvolatile salt. ESI-MS, however, is incompatible with many nonvolatile salts which leads to spray instability. Volatile components such as ammonium acetate/formate and acetic/formic acid can be used instead.

6.3.4.2 *Tagging efficiency*

Tagging efficiency is another aspect to consider while developing a robust technique for bioanalytical applications. The conversion rate of untagged ions into tagged ions is defined mainly by the temperature of the ions and the number density of the tag molecules [39].

Figure 6.11 demonstrates the tagging efficiency of different isomeric glycans. For the sugar A, the tagging efficiency varies from 19 to 28%; for B 31%, and the sugar C 7%. These results suggest that the tagging efficiency depends on an analyte and the ion density in the cold ion trap. The predominance of the efficiency for sugar B can be a reason why we mostly observed sugar B in the IR spectra while mixed with the other two sugars. The observations suggest that the band intensities are inversely proportional to ion count. Thus, the high ion count of the tagged sugar B doesn't overlap with the OPO IR beam efficiently, and the tagged ions from other sugars can be suppressed. Additionally, Scutelnic observed that the He-tagging efficiency monotonously increases as a function of the ion storage time in the octupole trap [32]. Consequently, the density of ions in a trap is limited by the space-charge effect ($\sim 10^7 \text{ cm}^{-3}$ [49]), which does not lead to any measurable depletion of the light flux. Consequently, it is critical not to saturate the trap with ions and record the spectra in the same conditions, including gas and temperature of the trap. Another factor that might be considered is the magnitude of the absorption cross-sections for different types of molecules.

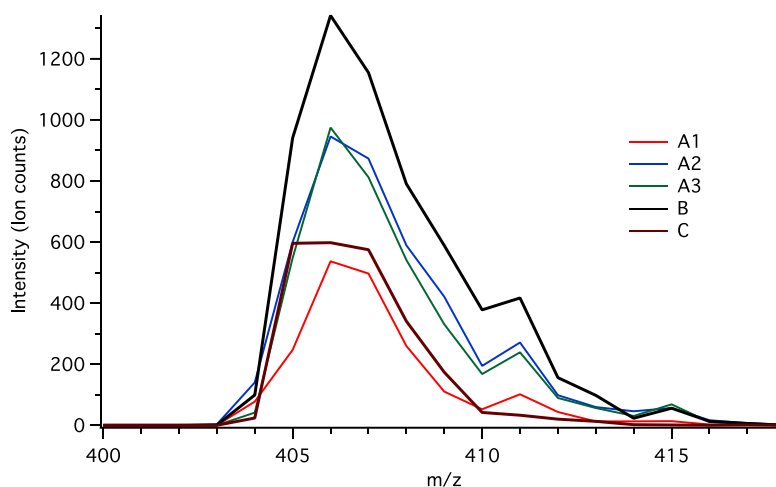


Figure 6.11 Mass spectrum of the disaccharides A (red, green, blue traces), B (black trace), and C (burgundy trace) representing the He tagging efficiency.

Another example of trisaccharides is shown in Figure 6.12. The derived tagging efficiencies are 43.5% for 4'galactosyllactose and 32.3% for globotriose. The higher tagging efficiency for 4'galactosyllactose results in its predominance in the overall spectrum of mixtures containing this sugar (see Figure 6.7-6.8).

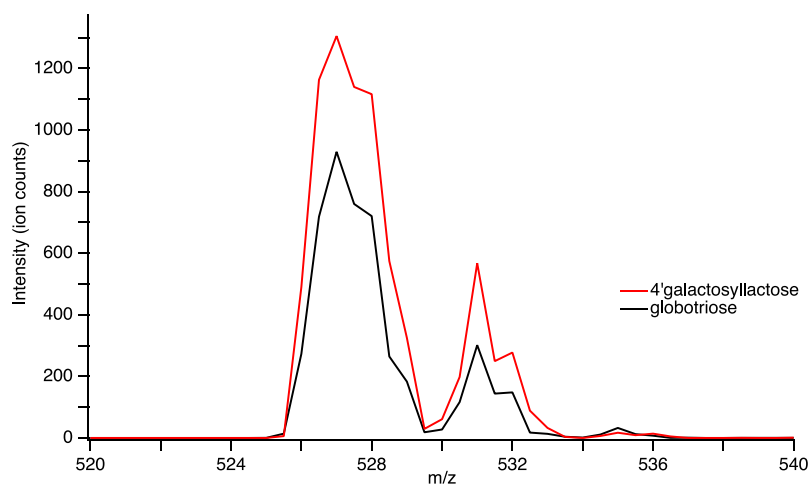


Figure 6.12 Mass spectra of the trisaccharides demonstrating the He tagging efficiency.

Tagging efficiency is one of the factors to consider while creating a glycan database since it can vary among the glycan molecule and can also contribute to a different sugar-dependent scale factor. One should include the tagging efficiency values in the database for quantitation and deconvolution purposes.

For our technique, in case of lacking the separation step priory to the quantification, saturation spectroscopy can be used to get the amounts of different isomers. For that, a unique transition of one isomer can be saturated and depleted fully by laser. From the amount of depletion, the quantity can be obtained. This process can then be repeated for other isomers in the mixture.

6.3.4.3 Linearity of a channeltron

To test the degree of correlation between the observed signal and the linearity of the channeltron response, we measure the ratio of the ^{13}C isotope peak to the monoisotopic peak. The rough estimation of the isotope ratio of the sodiated adduct of the sugar A is provided since the resolution of the quadrupole is relatively low ~ 250 . The ratio of the ^{13}C isotope peak to the monoisotopic peak is consistent with the expected isotope ratio (0.1595) only when the number of detected ions < 100 ions (Figure 6.13). Thus, an ion packet with more than 100 ions leads to greater deviation from linearity and hence a greater deviation from the expected isotopic ratio. Scutelnic reported the same observations for a dipeptide (GlyTrpH^+) in his thesis [32]. This can be a limiting factor for quantitation since we cannot estimate the ion abundances precisely. One needs to correct the measured ion counts while doing quantification. MCP detection coupled with high resolution MS analyzer can improve the experiments abilities with respect to quantification.

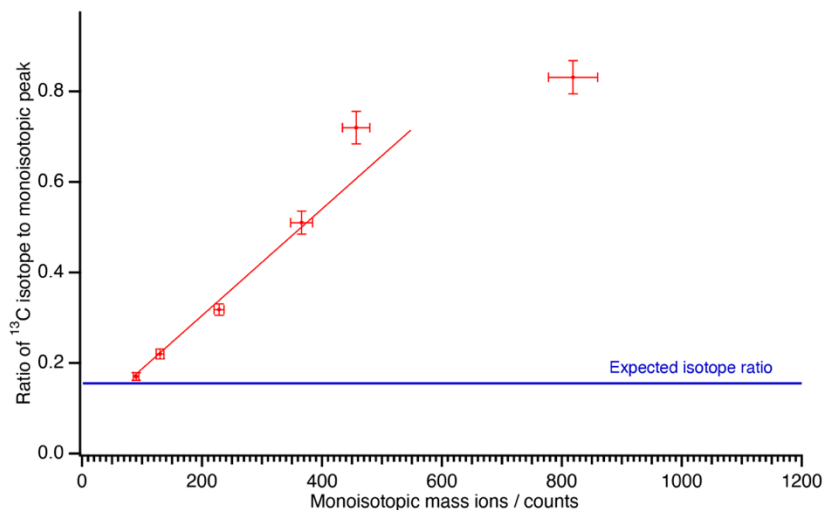


Figure 6.13 Ratio of the measured ^{13}C isotope peak to the monoisotopic peak of $[A + \text{Na}]^+ \text{ sugar}$.

6.3.4.4 Other parameters

All the spectra should be recorded in a regime of linear IR beam power dependence and under the same experimental conditions (power, alignments, etc.). The linearity of the measured IR spectra is maintained due to the low fluence of the IR OPO beam. To exclude the possibility of bands saturation and non-linearity of the IR beam, we measured the IR power dependence of transitions of the sugar A (3386 cm^{-1}) and B (3535 cm^{-1}) in a mixture, shown in Figure 6.14. The linearity is across the entire range of the OPO power. These results highlight the consistence of the conditions for spectroscopy experiments.

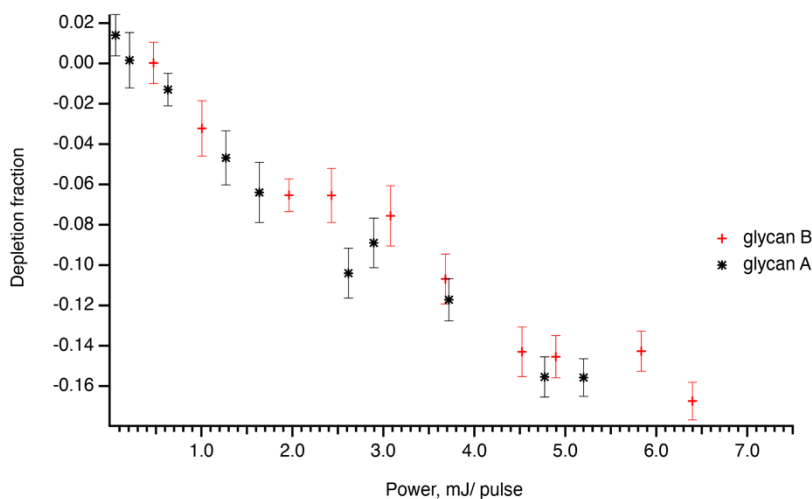


Figure 6.14 The OPO power dependence for transition of sugar A (3386 cm^{-1}) and transition sugar B (3535 cm^{-1}) in a mixture.

Additional parameters, such as discrimination of different masses due to RF frequency/ amplitude do not affect the results since the analytes are isomers.

6.3.5 Parameters for method validation

Knowing the limitations of our approach, we performed several experiments to evaluate few validation characteristics. These results would serve as a reference for future implementations in our multidimensional approach.

6.3.5.1 Precision

Figure 6.15 demonstrates repeatability (intra-day precision) of the experiment on sugar A under the same conditions over a short time. The results demonstrate significant variabilities in mass spectra (Figure 6.15 a) that lead to fluctuations in IR spectra (Figure 6.15 b).

According to ICH guidelines, the accepted RSD for the standards is up to 5%. However, it is strongly dependent on the application, field of studies, the concentration, and the complexity of the analyte.

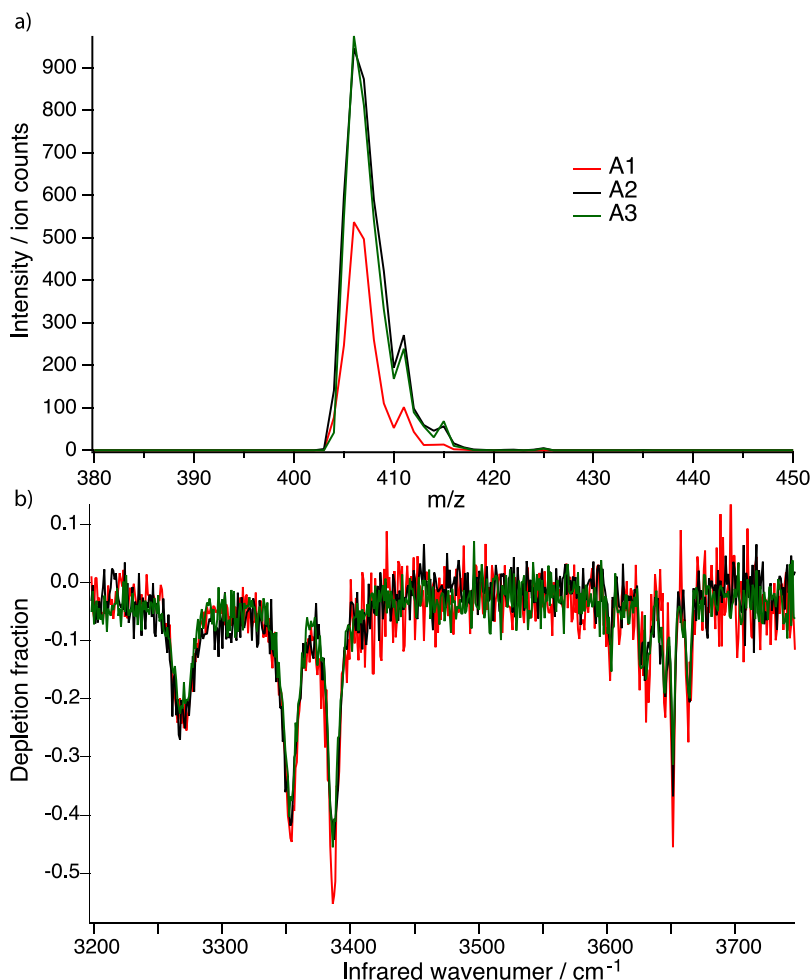


Figure 6.15 The repeatability precision validation for sugar A (75 μM) a) mass spectra b) IR spectra.

Table 6.2 presents the values of RSD of the peak heights in mass spectra that correspond to the untagged ($m/z=406$) and the tagged ($m/z=410$) sugar A molecule; and the values of RSD of the depletion fractions of the band transitions in the IR spectrum. We used three replicate of the experimental run at a test concentration 75 μM . The repeatability varies from 8.5 to 19.8 % in mass spectrum, and from 5 to 18.9 % in the cryogenic spectrum.

Table 6.2 Repeatability precision (%RSD) for the sugar A (75 μM) for IR spectrum and mass spectrum

Infrared wavenumber	Module of the depletion fraction	Average	Std deviation	RSD, %
3267	0.199	0.232	0.036	15.4
	0.225			
	0.270			
3354	0.403	0.423	0.022	5.2
	0.418			
	0.446			
3386	0.553	0.476	0.070	14.6
	0.456			
	0.418			
3651	0.455	0.379	0.072	18.9
	0.368			
	0.313			
m/z	Area	Average	Std deviation	RSD, %
406	0.084	0.106	0.021	19.8
	0.108			
	0.126			
410	0.130	0.142	0.012	8.5
	0.142			
	0.154			

The RSD values of the area of the integrated peak were worse than for the peak heights since S/N and the quadrupole resolution obfuscate the defined peak area (See Appendix B, Table B.3, Table B.4).

Additionally, we evaluated the intermediate precision for the ten experimental measurements of the sugar A (75 μ M), which results in $\sim 65\%$ for MS and 26 % for IR spectroscopy (see Appendix B, Table B.6). According to ICH, RSD $\sim 15\%$ (20% near the limit of quantification) is acceptable. Since the results are unsatisfactory, our laboratory must establish a quality control for all quantitative tests. The quality control can be performed by quantifying a known reagent with a known concentration before an actual experiment. Evaluating each test run in this way allows us to determine if the obtained results are accurate and reliable.

We evaluated the precision for a pentapeptide SDGRG (80 μM) that ionizes more efficiently than the sugar A (see details Appendix, Table B.7). The repeatability (RSD) is 5–9% for mass spectra and 2–5% for IR spectra for the same individual performing the same experiment over a short time. Intermediate precision (RSD) is 55% if different people perform the same experiment on the same day. To improve the RSD of the aforementioned validation parameters, a needle cutting robot coupled to the nESI is required.

6.3.5.2 *Dynamic range, linearity, LOD and LOQ*

a) Tandem mass-spectrometer

To further evaluate our approach, we need to estimate the linearity of calibration curves, including LOD and LOQ. For the estimation, the stock solution (2.6 mM) of the sugar A was prepared in water/methanol (1:1, v:v). Stock solution was diluted to yield the concentrations (conc.) of 5000, 1000, 500, 150, 25, 10, 1, 0.5 μM . To achieve constant experimental conditions, we performed our measurements with constant settings with three repetitions (see Appendix B, Table B. 8). We extracted the 406 m/z ion in the mass spectrum. The peak height is related to the concentration of the sample. Figure 6.16 demonstrates the dynamic range of our approach in a range of concentrations 0.5–1000 μM for sugar A with RSD from 3 to 14%. The results show that saturation occurs after 150 μM . At higher concentrations, the excess charge on the surface of the droplets can become limiting that leads to the loss of linearity. Moreover, the co-eluting compounds can influence the ionization process (the so-called matrix effect) and lead to the decrease or loss of linearity.

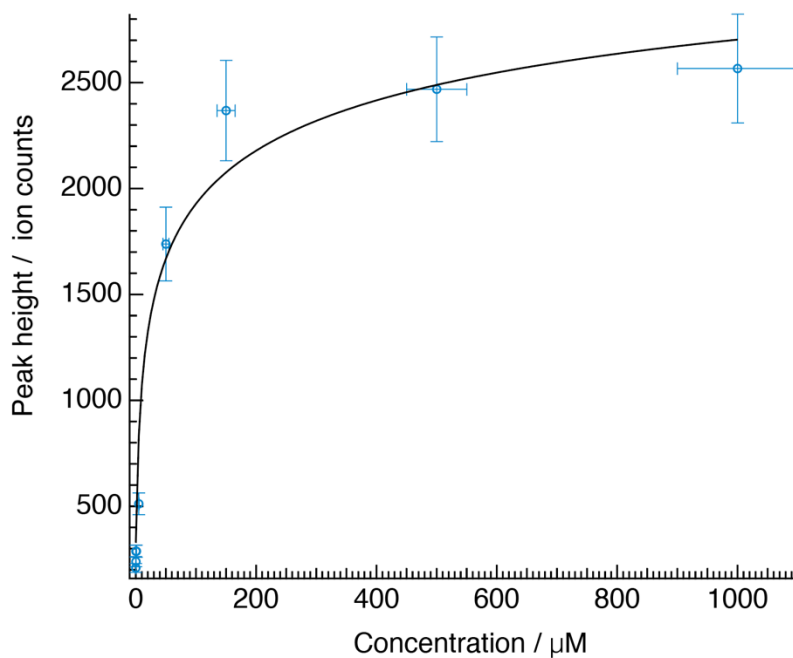


Figure 6.16 *Dynamic range of our approach in a range of 0.5-1000 μM .*

To avoid saturation, Figure 6.17 demonstrates the linear dynamic range from 1 to 150 μM concentration with RSD from 3 to 8%. The linearity of the calibration curve is usually expressed through the coefficient of correlation, R , or coefficient of determination, R^2 . However, the R^2 value for our calibration curve is 0.88979 since our approach provides the signal fluctuations that need to be corrected.

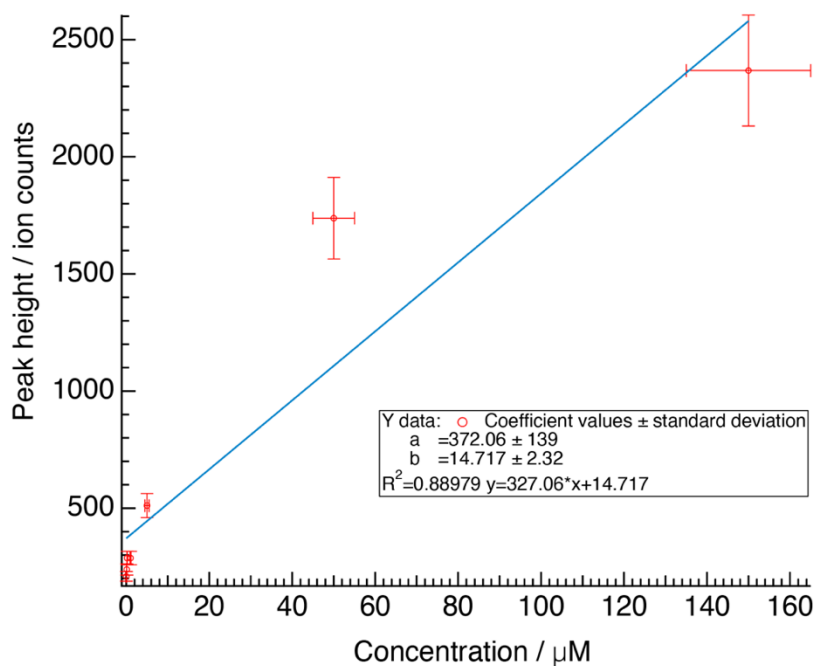


Figure 6.17 Linear dynamic range of our approach in a range of 1-150 μM .

To correct the fluctuations of the signal, we used the proportion of XIC and TIC as a signal response instead of a single XIC in a mass spectrum (see Appendix, Table B.9). Figure 6.18 determines the linear dynamic range from 1 to 150 μM concentration with RSD from 2 to 10%. The R^2 value for the calibration curve is 0.97414 that highlights the improvement in linearity of the curve. These implementations make it broadly applicable for analytical purposes. The obtained values of the proportion of XIC and TIC should be used in the future experiments.

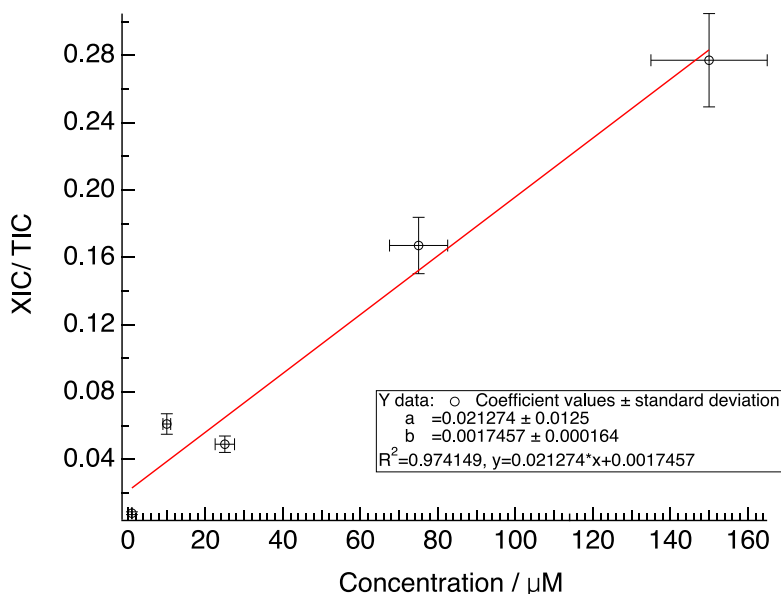


Figure 6.18 Dynamic range of our approach in a range of 1-150 μM using a corrected instrumental response (XIC/TIC).

The LOD value was calculated using a $1/x$ -weighted linear regression of the expected concentration and corresponding peak height counts. Subsequently, the $1/x$ -weighted regression residual standard deviation was multiplied by 3, then subtracted by the y-intercept, and finally divided by the slope. The LOQ was calculated similarly to the LOD, except that the residual standard deviation was multiplied by 10 instead of 3. The estimated LOD was 11 μM , the estimated LOQ is 33 μM for the tandem mass spectrometer [50].

b) The SLIM

We recommended the SLIM technique as a possible solution for the quantitation. Since the previously described results related to the earlier work were done on a tandem mass spectrometer, we performed similar experiments on the linear dynamic range estimation together with LOD and LOQ values. For the estimation, the stock solution (2.6 mM) of the sugar A was prepared in water/methanol (1:1, v:v). Stock solution was diluted to yield concentrations (conc.) of 1000, 500, 150, 25, 10, 1, 0.5 μM (see Appendix B, Figure B.3). To achieve constant experimental conditions, we performed our measurements with constant settings. Extraction of the drift-time

distribution of the 406 m/z ion showed two separate drift times, each of which corresponded to one of the two isomers. The peak height under the drift-time distribution is related to the concentration of the sample. Figure 6.19 demonstrates the dynamic range of the SLIM approach in a range of concentrations 0.5–1000 μM for sugar A with RSD from 2 to 8 %. The results demonstrate the saturation which appears after 150 μM , including the matrix effect.

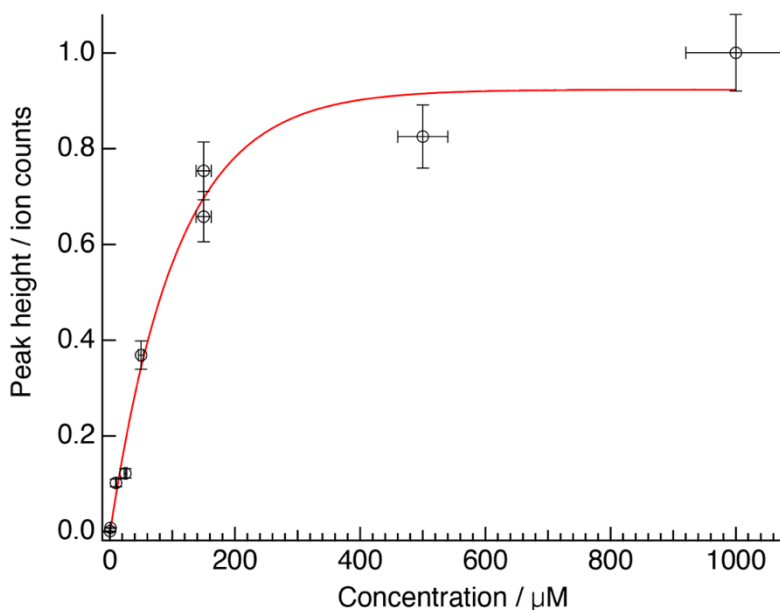


Figure 6.19 *Dynamic range of the SLIM approach in a range of 0.5–1000 μM .*

To estimate the linear dynamic range, Figure 6.20 demonstrates the results. The R^2 value of the calibration curve for the sugar A is 0.968689. These results are comparable to the ones on the tandem mass-spectrometer. The estimated LOD is 2.6 μM , the estimated LOQ is 8.2 μM for the SLIM apparatus. These results show promise that semiquantitative analysis will be possible using an ion mobility spectrometer coupled to a cryogenic trap and TOF.

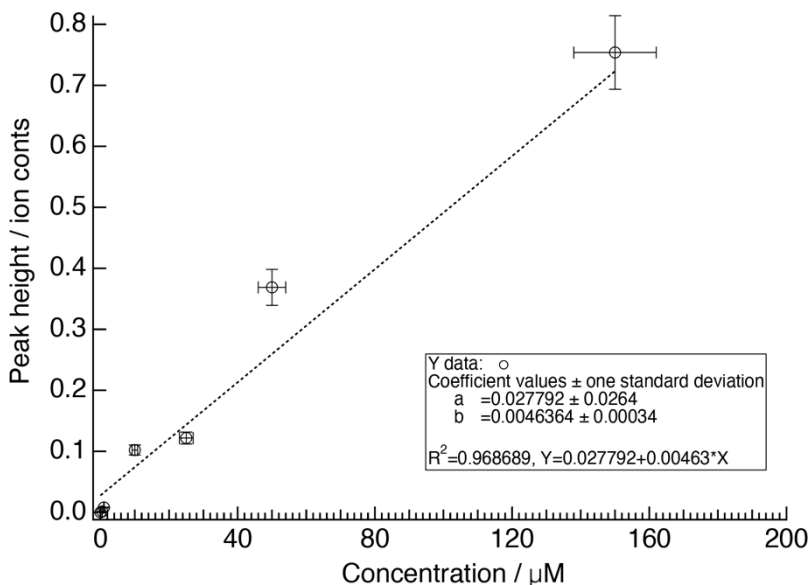


Figure 6.20 Linear dynamic range of our approach in a range of 0.5-150 μM .

6.4 Conclusions and future perspectives

The method demonstrated in this thesis to quantify the mixtures has some limitations. The first is the electrospray, including the spray instability and ionization efficiencies of analytes. The second is tagging efficiency that is revealed to be different for each molecule. Since we are planning on creating a database, we should include the tagging efficiency together with the temperature of the trap and the tagging gas to overcome these limitations. Another limitation is the non-linearity of the channeltron response which requires a new MS detection scheme to address. These limitations are evidence of the difficulty of precise quantitation. Given that our results were only a preliminary attempt to quantify mixtures and precise quantification is difficult, these findings provide insight into how the experimental design might be improved to facilitate it.

To improve the ionization efficiency, the stability of the spray should be provided using a commercial electrospray with an ability to control all the necessary parameters including needle cutting set-up. To minimize the confounding effect of ionization efficiency differences, the quantification of an analyte with MS usually requires comparison to an internal standard that has a similar structure to the analyte (its

stable isotopologue). MS response of the analyte should be corrected to the MS response of the internal standard. The internal standard is added at the earliest step possible during sample preparation and is analyzed simultaneously with the analyte. Additionally, the SLIM module will enable separation of the analyzed mixture from any impurities that might affect analysis. The TOF mass analyzer with a higher mass resolution should help to make fast and more precise measurements of the signal response. Knowing the tagging efficiency of the molecules, these values can be used as coefficients for correcting the resulting IR spectra.

This research will therefore serve as a base for future studies on glycan quantitation once the improvements that were identified in this thesis have been realized. All these findings and the solutions to the issues identified should be taken into account and tried on larger sample sizes.

References:

- (1) Laine, R. A. Invited Commentary: A Calculation of All Possible Oligosaccharide Isomers Both Branched and Linear Yields 1.05×10^{12} Structures for a Reducing Hexasaccharide: The Isomer Barrier to Development of Single-Method Saccharide Sequencing or Synthesis Systems. *Glycobiology* **1994**, 4 (6), 759–767. <https://doi.org/10.1093/glycob/4.6.759>.
- (2) Alvarez-Manilla, G.; Warren, N. L.; Abney, T.; Atwood, J.; Azadi, P.; York, W. S.; Pierce, M.; Orlando, R. Tools for Glycomics: Relative Quantitation of Glycans by Isotopic Permethylation Using $^{13}\text{CH}_3\text{I}$. *Glycobiology* **2007**, 17 (7), 677–687. <https://doi.org/10.1093/glycob/cwm033>.
- (3) Etxebarria, J.; Reichardt, N.-C. Methods for the Absolute Quantification of N-Glycan Biomarkers. *Biochim. Biophys. Acta BBA - Gen. Subj.* **2016**, 1860 (8), 1676–1687. <https://doi.org/10.1016/j.bbagen.2016.03.003>.
- (4) Delafield, D. G.; Li, L. Recent Advances in Analytical Approaches for Glycan and Glycopeptide Quantitation. *Mol. Cell. Proteomics* **2020**. <https://doi.org/10.1074/mcp.R120.002095>.
- (5) Hofmann, J.; Hahm, H. S.; Seeberger, P. H.; Pagel, K. Identification of Carbohydrate Anomers Using Ion Mobility–Mass Spectrometry. *Nature* **2015**, 526 (7572), 241–244. <https://doi.org/10.1038/nature15388>.
- (6) Ruhaak, L. R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A. M.; Wuhrer, M. Glycan Labeling Strategies and Their Use in Identification and Quantification. *Anal. Bioanal. Chem.* **2010**, 397 (8), 3457–3481. <https://doi.org/10.1007/s00216-010-3532-z>.
- (7) Harvey, D. J. Derivatization of Carbohydrates for Analysis by Chromatography; Electrophoresis and Mass Spectrometry. *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* **2011**, 879 (17–18), 1196–1225. <https://doi.org/10.1016/j.jchromb.2010.11.010>.
- (8) Gil, G.-C.; Kim, Y.-G.; Kim, B.-G. A Relative and Absolute Quantification of Neutral N-Linked Oligosaccharides Using Modification with

Carboxymethyl Trimethylammonium Hydrazide and Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Anal. Biochem.* **2008**, 379 (1), 45–59. <https://doi.org/10.1016/j.ab.2008.04.039>.

(9) Harvey, D. J. Analysis of Carbohydrates and Glycoconjugates by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry: An Update for 2009–2010. *Mass Spectrom. Rev.* **2015**, 34 (3), 268–422. <https://doi.org/10.1002/mas.21411>.

(10) Wuhrer, M.; Koeleman, C. A. M.; Hokke, C. H.; Deelder, A. M. Nano-Scale Liquid Chromatography-Mass Spectrometry of 2-Aminobenzamide-Labeled Oligosaccharides at Low Femtomole Sensitivity. *Int. J. Mass Spectrom.* **2004**, 232 (1), 51–57. <https://doi.org/10.1016/j.ijms.2003.11.009>.

(11) Knežević, A.; Bones, J.; Kračun, S. K.; Gornik, O.; Rudd, P. M.; Lauc, G. High Throughput Plasma N-Glycome Profiling Using Multiplexed Labelling and UPLC with Fluorescence Detection. *The Analyst* **2011**, 136 (22), 4670–4673. <https://doi.org/10.1039/c1an15684e>.

(12) Gennaro, L. A.; Salas-Solano, O. On-Line CE–LIF–MS Technology for the Direct Characterization of N-Linked Glycans from Therapeutic Antibodies. *Anal. Chem.* **2008**, 80 (10), 3838–3845. <https://doi.org/10.1021/ac800152h>.

(13) Mariño, K.; Lane, J. A.; Abrahams, J. L.; Struwe, W. B.; Harvey, D. J.; Marotta, M.; Hickey, R. M.; Rudd, P. M. Method for Milk Oligosaccharide Profiling by 2-Aminobenzamide Labeling and Hydrophilic Interaction Chromatography. *Glycobiology* **2011**, 21 (10), 1317–1330. <https://doi.org/10.1093/glycob/cwr067>.

(14) Chik, J. H. L.; Zhou, J.; Moh, E. S. X.; Christopherson, R.; Clarke, S. J.; Molloy, M. P.; Packer, N. H. Comprehensive Glycomics Comparison between Colon Cancer Cell Cultures and Tumours: Implications for Biomarker Studies. *J. Proteomics* **2014**, 108, 146–162. <https://doi.org/10.1016/j.jprot.2014.05.002>.

(15) Pang, P.-C.; Chiu, P. C. N.; Lee, C.-L.; Chang, L.-Y.; Panico, M.; Morris, H. R.; Haslam, S. M.; Khoo, K.-H.; Clark, G. F.; Yeung, W. S. B.; Dell, A. Human Sperm Binding Is Mediated by the Sialyl-Lewis(x) Oligosaccharide on the

Zona Pellucida. *Science* **2011**, 333 (6050), 1761–1764.
<https://doi.org/10.1126/science.1207438>.

(16) Yang, S.; Yuan, W.; Yang, W.; Zhou, J.; Harlan, R.; Edwards, J.; Li, S.; Zhang, H. Glycan Analysis by Isobaric Aldehyde Reactive Tags and Mass Spectrometry. *Anal. Chem.* **2013**, 85 (17), 8188–8195.
<https://doi.org/10.1021/ac401226d>.

(17) Zhang, P.; Zhang, Y.; Xue, X.; Wang, C.; Wang, Z.; Huang, L. Relative Quantitation of Glycans Using Stable Isotopic Labels 1-(D0/D5) Phenyl-3-Methyl-5-Pyrazolone by Mass Spectrometry. *Anal. Biochem.* 2011, 418 (1), 1–9.
<https://doi.org/10.1016/j.ab.2011.07.006>.

(18) Walker, S. H.; Taylor, A. D.; Muddiman, D. C. Individuality Normalization When Labeling with Isotopic Glycan Hydrazide Tags (INLIGHT): A Novel Glycan-Relative Quantification Strategy. *J. Am. Soc. Mass Spectrom.* **2013**, 24 (9), 1376–1384. <https://doi.org/10.1007/s13361-013-0681-2>.

(19) Atwood, J. A.; Cheng, L.; Alvarez-Manilla, G.; Warren, N. L.; York, W. S.; Orlando, R. Quantitation by Isobaric Labeling: Applications to Glycomics. *J. Proteome Res.* **2008**, 7 (1), 367–374. <https://doi.org/10.1021/pr070476i>.

(20) FDA/CDER/"Beers, D. Analytical Procedures and Methods Validation for Drugs and Biologics. **2015**, 18.

(21) Guide to Method Validation for Quantitative Analysis in Chemical Testing Laboratories. **2019**, No. 6, 23.

(22) Masellis C., Khanal N., Kamrath M.Z., Clemmer D.E., C.; Rizzo T.R. Cryogenic Vibrational Spectroscopy Provides Unique Fingerprints for Glycan Identification. *J. Am. Soc. Mass Spectrom.* **2017**, 28 (10), 2217–2222.
<https://doi.org/10.1007/s13361-017-1728-6>.

(23) Rizzo, T. R.; Stearns, J. A.; Boyarkin, O. V. Spectroscopic Studies of Cold, Gas-Phase Biomolecular Ions. *Int. Rev. Phys. Chem.* **2009**, 28 (3), 481–515.
<https://doi.org/10.1080/01442350903069931>.

- (24) Voronina, L.; Masson, A.; Kamrath, M.; Schubert, F.; Clemmer, D.; Baldauf, C.; Rizzo, T. Conformations of Prolyl–Peptide Bonds in the Bradykinin 1–5 Fragment in Solution and in the Gas Phase. *J. Am. Chem. Soc.* **2016**, 138 (29), 9224–9233. <https://doi.org/10.1021/jacs.6b04550>.
- (25) Svendsen, A.; Lorenz, U. J.; Boyarkin, O. V.; Rizzo, T. R. A New Tandem Mass Spectrometer for Photofragment Spectroscopy of Cold, Gas-Phase Molecular Ions. *Rev. Sci. Instrum.* **2010**, 81 (7), 073107. <https://doi.org/10.1063/1.3458014>.
- (26) Kim, T.; Tang, K.; Udseth, H. R.; Smith, R. D. A Multicapillary Inlet Jet Disruption Electrodynamical Ion Funnel Interface for Improved Sensitivity Using Atmospheric Pressure Ion Sources. *Anal. Chem.* **2001**, 73 (17), 4162–4170. <https://doi.org/10.1021/ac010174e>.
- (27) Kelly, R. T.; Tolmachev, A. V.; Page, J. S.; Tang, K.; Smith, R. D. The Ion Funnel: Theory, Implementations, and Applications. *Mass Spectrom. Rev.* **2010**, 29 (2), 294–312. <https://doi.org/10.1002/mas.20232>.
- (28) Boyarkin, O. V.; Kopysov, V. Cryogenically Cooled Octupole Ion Trap for Spectroscopy of Biomolecular Ions. *Rev. Sci. Instrum.* **2014**, 85 (3), 033105. <https://doi.org/10.1063/1.4868178>.
- (29) Aseev, O. Spectroscopic Studies of Peptide Fragments Produced by Collision-Induced Dissociation, EPFL, 2014. <https://doi.org/10.5075/epfl-thesis-6387>.
- (30) Jašík, J.; Gerlich, D.; Roithová, J. Two-Color Infrared Predissociation Spectroscopy of C₆H₆²⁺ Isomers Using Helium Tagging. *J. Phys. Chem. A* **2015**, 119 (11), 2532–2542. <https://doi.org/10.1021/jp5088064>.
- (31) Asvany, O.; Brünken, S.; Kluge, L.; Schlemmer, S. COLTRAP: A 22-Pole Ion Trapping Machine for Spectroscopy at 4 K. *Appl. Phys. B* **2014**, 114 (1), 203–211. <https://doi.org/10.1007/s00340-013-5684-y>.
- (32) Scutelnic, V. Cryogenic Infrared Spectroscopy as a Probe of Structure and Dynamics of Biomolecules, EPFL, 2018. <https://doi.org/10.5075/epfl-thesis-9164>.

- (33) Jørgensen, K.; Segtnan, V.; Thyholt, K.; Næs, T. A Comparison of Methods for Analysing Regression Models with Both Spectral and Designed Variables. *J. Chemom.* **2004**, 18 (10), 451–464. <https://doi.org/10.1002/cem.890>.
- (34) Mohammadi, S.; Zuckerman, N.; Goldsmith, A.; Grama, A. A Critical Survey of Deconvolution Methods for Separating Cell-Types in Complex Tissues. *Proc. IEEE* **2017**, 105 (2), 340–366. <https://doi.org/10.1109/JPROC.2016.2607121>.
- (35) Mandra, V. J.; Kouskoura, M. G.; Markopoulou, C. K. Using the Partial Least Squares Method to Model the Electrospray Ionization Response Produced by Small Pharmaceutical Molecules in Positive Mode. *Rapid Commun. Mass Spectrom.* **2015**, 29 (18), 1661–1675. <https://doi.org/10.1002/rcm.7263>.
- (36) Kiontke, A.; Oliveira-Birkmeier, A.; Opitz, A.; Birkemeyer, C. Electrospray Ionization Efficiency Is Dependent on Different Molecular Descriptors with Respect to Solvent PH and Instrumental Configuration. *PloS One* **2016**, 11 (12), e0167502. <https://doi.org/10.1371/journal.pone.0167502>.
- (37) Hiraoka, K.; Murata, K.; Kudaka, I. Do the Electrospray Mass Spectra Reflect the Ion Concentrations in Sample Solution? *J. Mass Spectrom. Soc. Jpn.* **1995**, 43 (3), 127–138. <https://doi.org/10.5702/massspec.43.127>.
- (38) Li Y, Cole RB. Charge State Distributions in Electrospray and MALDI. In: Cole RB, Editors. *Electrospray and MALDI Mass Spectrometry: Fundamentals, Instrumentation, Practicalities, and Biological Applications*; John Wiley & Sons, Inc.: Hoboken, NJ, USA:, 2010.
- (39) Cismesia, A. P.; Bailey, L. S.; Bell, M. R.; Tesler, L. F.; Polfer, N. C. Making Mass Spectrometry See the Light: The Promises and Challenges of Cryogenic Infrared Ion Spectroscopy as a Bioanalytical Technique. *J. Am. Soc. Mass Spectrom.* **2016**, 27 (5), 757–766. <https://doi.org/10.1007/s13361-016-1366-4>.
- (40) Johnson, C. J.; Wolk, A. B.; Fournier, J. A.; Sullivan, E. N.; Weddle, G. H.; Johnson, M. A. Communication: He-Tagged Vibrational Spectra of the SarGlyH⁺ and H⁺ (H₂O)_{2,3} Ions: Quantifying Tag Effects in Cryogenic Ion Vibrational Predissociation (CIVP) Spectroscopy. *J. Chem. Phys.* **2014**, 140 (22), 221101. <https://doi.org/10.1063/1.4880475>.

- (41) Leito, I.; Herodes, K.; Huopolainen, M.; Virro, K.; Künnapas, A.; Kruve, A.; Tanner, R. Towards the Electrospray Ionization Mass Spectrometry Ionization Efficiency Scale of Organic Compounds. *Rapid Commun. Mass Spectrom.* **2008**, 22 (3), 379–384. <https://doi.org/10.1002/rcm.3371>.
- (42) Tang, Liang.; Kebarle, Paul. Dependence of Ion Intensity in Electrospray Mass Spectrometry on the Concentration of the Analytes in the Electrosprayed Solution. *Anal. Chem.* **1993**, 65 (24), 3654–3668. <https://doi.org/10.1021/ac00072a020>.
- (43) Mora, J. F. D. L.; Loscertales, I. G. The Current Emitted by Highly Conducting Taylor Cones. *J. Fluid Mech.* **1994**, 260, 155–184. <https://doi.org/10.1017/S0022112094003472>.
- (44) Bawazeer, S.; Muhsen Ali, A.; Alhawiti, A.; Khalaf, A.; Gibson, C.; Tusiimire, J.; Watson, D. G. A Method for the Analysis of Sugars in Biological Systems Using Reductive Amination in Combination with Hydrophilic Interaction Chromatography and High Resolution Mass Spectrometry. *Talanta* **2017**, 166, 75–80. <https://doi.org/10.1016/j.talanta.2017.01.038>.
- (45) Alagesan, K.; Silva, D. V.; Seeberger, P. H.; Kolarich, D. A Novel, Ultrasensitive Approach for Quantitative Carbohydrate Composition and Linkage Analysis Using LC-ESI Ion Trap Tandem Mass Spectrometry. *bioRxiv* **2019**, 853036. <https://doi.org/10.1101/853036>.
- (46) Xu, W.; Liang, L.; Zhu, M. Determination of Sugars in Molasses by HPLC Following Solid-Phase Extraction. *Int. J. Food Prop.* **2015**, 18 (3), 547–557. <https://doi.org/10.1080/10942912.2013.837064>.
- (47) Grünwald-Gruber, C.; Thader, A.; Maresch, D.; Dalik, T.; Altmann, F. Determination of True Ratios of Different N-Glycan Structures in Electrospray Ionization Mass Spectrometry. *Anal. Bioanal. Chem.* **2017**, 409 (10), 2519–2530. <https://doi.org/10.1007/s00216-017-0235-8>.
- (48) Tang, K.; Page, J. S.; Smith, R. D. Charge Competition and the Linear Dynamic Range of Detection in Electrospray Ionization Mass Spectrometry. *J. Am.*

Soc. Mass Spectrom. **2004**, 15 (10), 1416–1423.
<https://doi.org/10.1016/j.jasms.2004.04.034>.

(49) Majima, T.; Santambrogio, G.; Bartels, C.; Terasaki, A.; Kondow, T.; Meinen, J.; Leisner, T. Spatial Distribution of Ions in a Linear Octopole Radio-Frequency Ion Trap in the Space-Charge Limit. *Phys. Rev. A* **2012**, 85 (5), 053414.
<https://doi.org/10.1103/PhysRevA.85.053414>.

(50) Vial, J.; Jardy, A. Experimental Comparison of the Different Approaches To Estimate LOD and LOQ of an HPLC Method. *Anal. Chem.* **1999**, 71 (14), 2672–2677. <https://doi.org/10.1021/ac981179n>.

Chapter 7 Metal complexes of glycans

The database approach for glycans has been presented in detail in Chapter 4. The use of cryogenic vibrational spectroscopy for glycan identification requires the development of this spectroscopic database. Our experiments have previously been conducted exclusively on sodiated or protonated carbohydrate species, however, it was unclear if we can use cryogenic infrared spectra of any other metal adducts of glycans for their unambiguous identification. In this chapter, we present a study demonstrating that cryogenic spectra of any alkali metal adducts of oligosaccharides are structured enough to be used as fingerprints for building the database. We present several examples, including disaccharides and N-glycans applying SLIM-based ion mobility and cryogenic ion spectroscopy.

7.1 Sample preparation

Disaccharides were purchased from Carbosynth Ltd (UK) and N-glycans from Dextra Laboratories (UK). The glycan purity is certified at a minimum of 85 %, and all were used without further purification. The metal salts were chlorides of lithium (Li(I)), sodium (Na(I)), potassium (K(I)), rubidium (Rb(I)), cesium (Cs(I)), magnesium (Mg(II)), calcium (Ca(II)), and barium (Ba(II)). All the salts were purchased from Sigma-Aldrich (USA). Aqueous 1 mM stock solutions of disaccharides were diluted with water/methanol (v/v, 50/50) to yield a 150 μ M analyte solution. To ensure metal adduction, 50 μ M of each metal salt was added to each disaccharide sample before analysis. For N-glycans, the solutions were prepared in a 30:70 solution of water/acetonitrile to yield a 20-30 μ M analyte solution. All solvents were HPLC grade.

The SLIM-IMS parameters for the experiments were as follows: RF frequency 880 kHz; RF amplitude 100 V_{pp}; TW amplitude 27 V; TW speed 600 m/s; He buffer gas pressure of 3 mbar.

7.2 Metal complexes of disaccharides

Before discussing the differences between the measured IR spectra, it is worth mentioning how the presence of an alkali metal ion can impact the structure of a disaccharide. The studied disaccharides have multiple OH groups and a single NH group, which can form a hydrogen bonding (HB) network. This network will be altered upon binding of a singly charged alkali metal, and it may change the overall structure (folding and conformation) of the carbohydrate if many hydrogen bonds are affected.

We first describe how metal coordination affects the ion mobility and IR spectra of a disaccharide GalNAc α (1–3)Gal. Figure 7.1 shows the mass spectra obtained for the disaccharide complexed with different metals. The mass spectra have been recorded under the same conditions and on the same day, using the tandem mass-spectrometer described in Chapter 6. A comparison of the relative intensities of mass-spectra for the GalNAc α (1–3)Gal shows the predominance of the singly sodiated species (black trace in Figure 7.1), whereas the singly potassiated species (red trace) are the least abundant. This might be related to the presence of sodium in the sample itself or solvent glass containers. The low abundance of the potassiated species can be linked to the competition between sodium and potassium in the solution (see Appendix C, Figure C.1). We did not observe complexes with the alkaline earth metals except for Ba (burgundy trace in Figure 7.1). Overall, the intensities of the alkali metal complexes were much greater than complexes with doubly charged Ba, likely because of the greater efficiency of desolvation of the singly charged alkali metal complexes with respect to doubly charged alkaline earth complexes in the ESI.

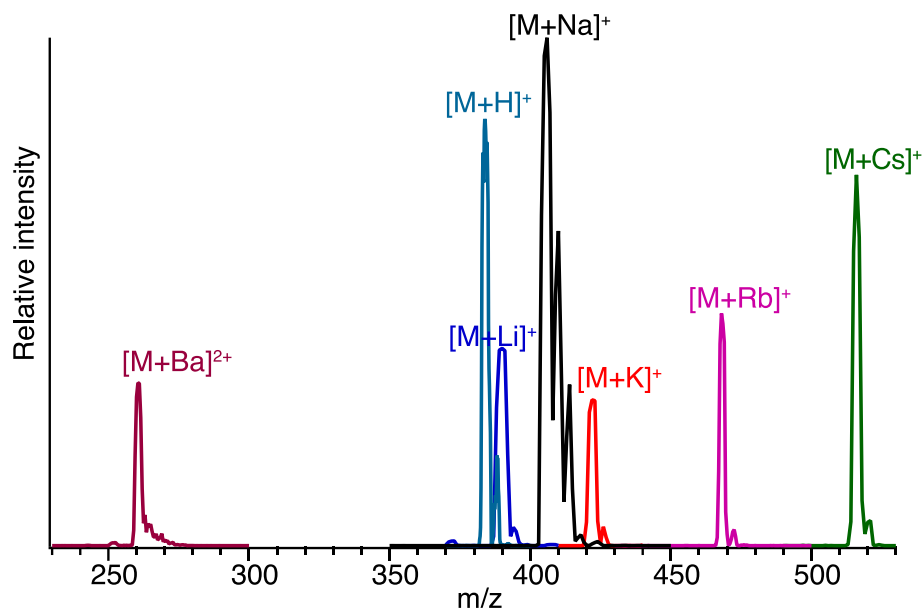


Figure 7.1 Mass spectra of the barium ion adducts (burgundy trace), protonated ion adducts (light blue trace), lithium ion adducts (blue trace), sodium ion adducts (black trace), potassium ion adducts (red trace), rubidium ion adducts (pink trace), and cesium ion adducts (green trace) of GalNAc α (1-3)Gal.

Figure 7.2 shows the ATDs of the alkali adducts of the GalNAc α (1-3)Gal obtained using a single SLIM cycle corresponding to a 1.8 m drift path. All recorded ATDs exhibit one major peak. We can see that the most intense peak corresponds to the sodiated adduct similarly to the mass spectra shown in Figure 7.1. As expected, the size of the metal ion exerts a significant influence on the drift times of the carbohydrate metal ion adducts. As a general trend, the drift times of these ions increased as the size of metal ions increased. Interestingly, the drift time of the lithium adduct of the sugar was shorter than that of singly protonated species. This can be explained by a more compact shape of the carbohydrate with lithium since the sugar can wrap around the metal [1]. Cancilla and coworkers also suggest that the overall size of a metal ions adducts of carbohydrate depends not only on the ionic radius of the metal ion but also on the preferred coordination of the metal combined with the structural flexibility of specific carbohydrate moieties [2].

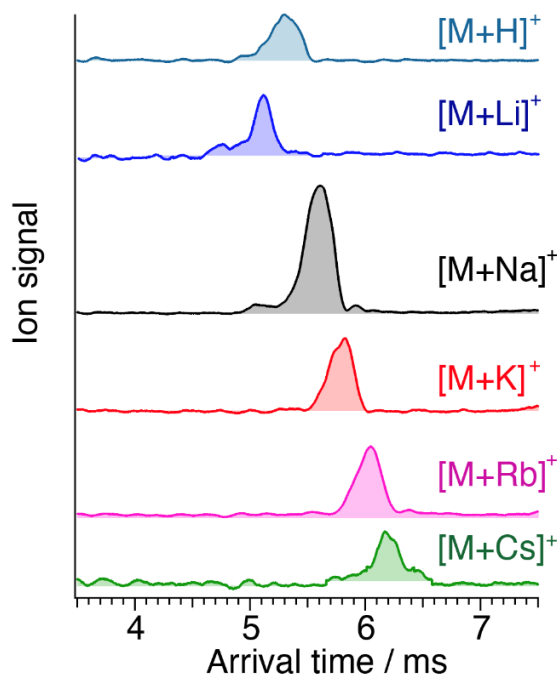


Figure 7.2 Arrival time distributions of the protonated ion adducts (light blue trace), lithium ion adducts (blue trace), sodium ion adducts (black trace), potassium ion adducts (red trace), rubidium ion adducts (pink trace), and cesium ion adducts (green trace) of GalNAc(1-3)Gal, obtained using a single SLIM cycle corresponding to 1.8 m drift path.

Figure 7.3 shows the cryogenic IR spectra for each of the metal ion adducts of GalNAc(1-3)Gal using He as a messenger tag. All infrared spectra are relatively simple and show distinct features in both the hydrogen-bonding and free OH stretch regions. All of these IR spectra can serve as a fingerprint to identify this molecule. Moreover, the IR spectral complexity is not affected by the identity of the metal cation.

The spectral signatures of GalNAc(1-3)Gal coordinated with Li, Na, K, Rb, Cs, and H are similar, suggesting that they have similar structures and most probably the similar metal binding site. Three intense, well-spaced and broadened bands in the hydrogen-bonded region, shown in yellow, are observed for each species together with a cluster of bands in the free OH region, shown in blue. The most notable trend in the alkali metal sequence from Li to Cs is the gradual blue-shifting of the low

wavenumber bands. This gradual shift is most probably due to the effect of different nature (e.g., atomic radius, charge density, polarizability) of each alkali metal ion. Moreover, we hypothesize that compared to the protonated disaccharide (light blue trace in Figure 7.3), the sugar with Li^+ exhibits a strong red-shift of the low wavenumber bands (blue trace in Figure 7.3), most probably related to the weakening of OH covalent bonds due to interaction with alkali metal ion. Preliminary results of the quantum calculations to reveal the structures of the metal adducts of $\text{GalNAc}\alpha(1-3)\text{Gal}$ are presented in Appendix C.

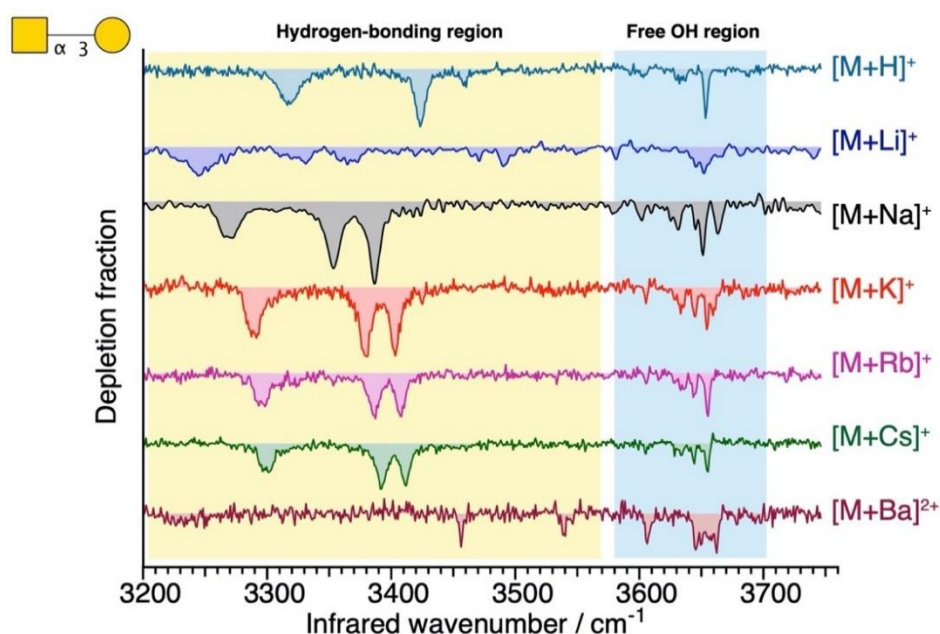


Figure 7.3 Cryogenic infrared spectra of the protonated and metal ion adducts of the sugar $\text{GalNAc}\alpha(1-3)\text{Gal}$.

If we plot the frequencies of the low wavenumber bands versus the ionic radius (Figure 7.4), it becomes clear that the different strength of the interaction between a carbohydrate molecule and each of the five alkali metal ions translates into different frequency shifts for OH and NH stretching modes, with Li and Cs ion adducts introducing the largest and smallest shifts, respectively. This can be rationalized by considering that the higher charge density of Li^+ is more effective in drawing electrons from the carbohydrate molecule that, in turn, weakens the OH and NH bonds and shifts the stretching frequencies more to the red compared to the other alkali ions.

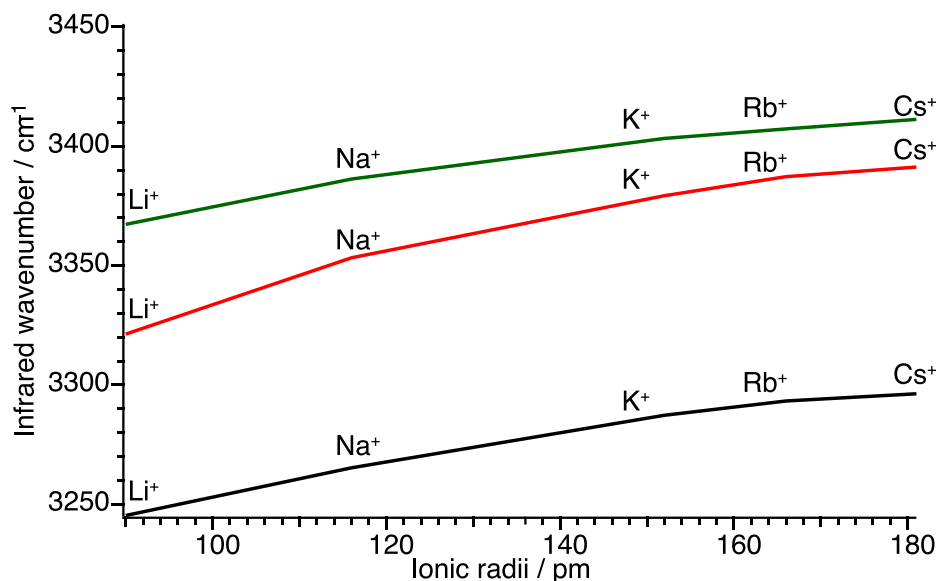


Figure 7.4 Comparative plot of the frequencies of the three low-wavenumber bands versus ionic radii of the alkali metals [3].

The spectrum of GalNAc α (1–3)Gal coordinated with Ba²⁺ (Figure 7.3, burgundy trace) showed two sharp, strongly displaced bands at low wavenumbers together with a cluster of bands lying at higher wavenumbers. We hypothesize that most of the vibrational bands that we observe correspond to free NH (3400 cm^{–1}) and nearly-free OH (3570–3700 cm^{–1}) stretches, while other hydrogen-bonded OH stretches (if any) were strongly red-shifted out of the measured IR frequency region due to strong interaction with the doubly charged metal ion.

Another disaccharide that we studied using cryogenic vibrational spectroscopy is the Gal β (1–4)GlcNAc, which has the same mass as the GalNAc α (1–3)Gal but differs in the sequence, monosaccharide content, and the glycosidic linkage. As expected, all IR signatures for this sugar, shown in Figure 7.5, differ significantly from the isomeric GalNAc α (1–3)Gal. However, similarly to GalNAc α (1–3)Gal, the Li, Na, K, Rb, and Cs ion adducts of the Gal β (1–4)GlcNAc have similar spectral signatures, more specifically, three highly-spaced bands at low wavenumber and multiple free OH stretch bands at higher wavenumber. Our findings appear to have the same trend as with GalNAc α (1–3)Gal: a shift of three bands in the weakly hydrogen-bonding OH

stretch region to the blue going from Li^+ to Cs^+ , suggesting that the alkali metal binding energies are decreasing with an increase of the ionic radius.

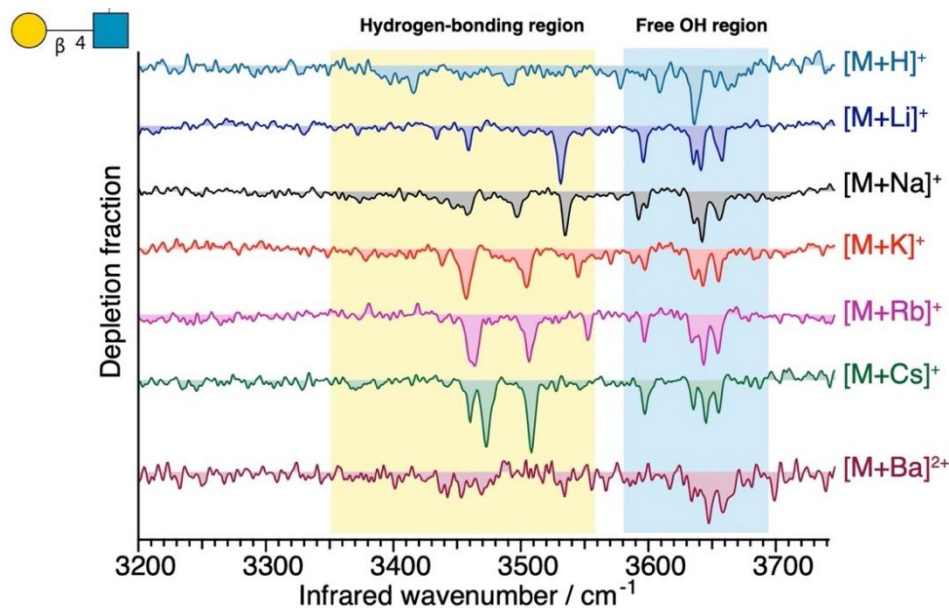


Figure 7.5 Cryogenic infrared spectra of the protonated and metal ion adducts of the $\text{Gal}\beta(1-4)\text{GlcNAc}$.

Our data confirms that all metal ion adducts can be used for carbohydrate unambiguous identification since all the spectra exhibit well-resolved IR signatures. The subsequent section also shows some examples for N-glycan complexed with various metals.

7.3 Metal complexes of N-glycans

As we demonstrated in Chapter 4, multiple ion adducts of N-glycans, such as $[\text{M} + 2\text{H}]^{2+}$, $[\text{M} + \text{H} + \text{Na}]^{2+}$, $[\text{M} + \text{H} + \text{K}]^{2+}$, $[\text{M} + 2\text{Na}]^{2+}$, and $[\text{M} + \text{Na} + \text{K}]^{2+}$ are formed by electrospray without addition of any sodium or potassium. For example, Figure 7.6 shows a typical mass spectrum obtained for the G0 N-glycan standard, for which multiple metal complexes are observed. These observations can be explained by the natural presence of Na and K salts in the sample itself or sample storage in glassware, or using ESI borosilicate emitters for generating ions.

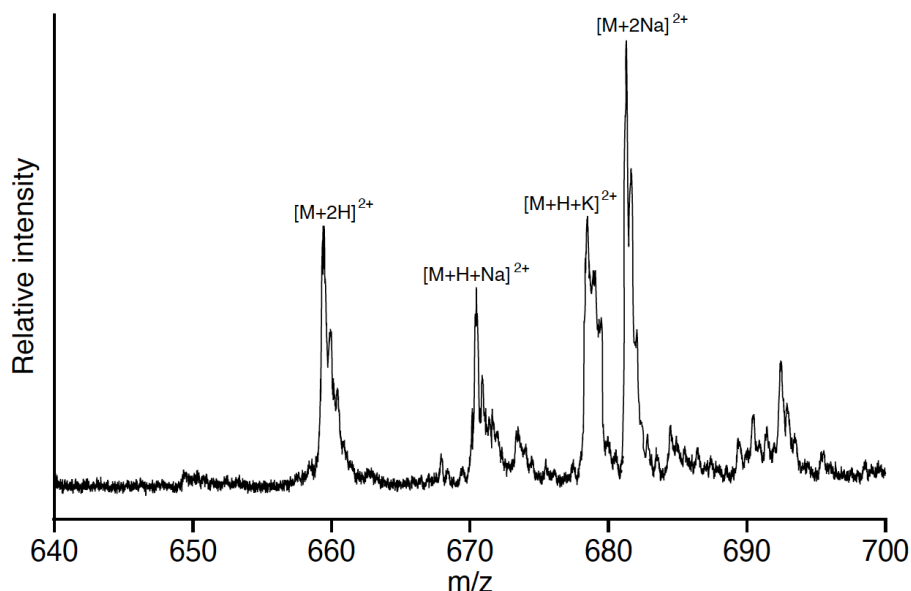


Figure 7.6 Typical mass spectrum of ion adducts of G0 N-glycan standard formed in positive ESI mode.

The cryogenic vibrational spectra of the most intense ion adducts of G0 measured without ion mobility selection are shown in Figure 7.7. Each spectrum in Figure 7.7 shows unique well-resolved transitions in the free OH stretch region ($3580\text{--}3700\text{ cm}^{-1}$) that can be used for identification, regardless of the increased complexity of the spectra compared to disaccharides (Figure 7.3). Moreover, Figure 7.7 demonstrates that the free OH region for these different adducts of the G0 molecule is the most informative. The observed differences between the spectra are most probably due to slightly different geometries induced by protonation and sodium or potassium ion binding.

Mobility-selected IR spectra are slightly sharper and more distinctive and, thus, provide a more stringent spectroscopic fingerprint (see Chapter 4, Figure 4.6). The aim of this chapter, however, was not to provide an exhaustive study of IR spectra of all the possible conformations in the gas phase, and therefore the IMS selection was not applied. Nevertheless, Figure 7.7 shows that even without IMS selection the cryogenic vibrational spectra of various ion adducts of G0 are resolved enough to be used as fingerprints for its unambiguous identification.

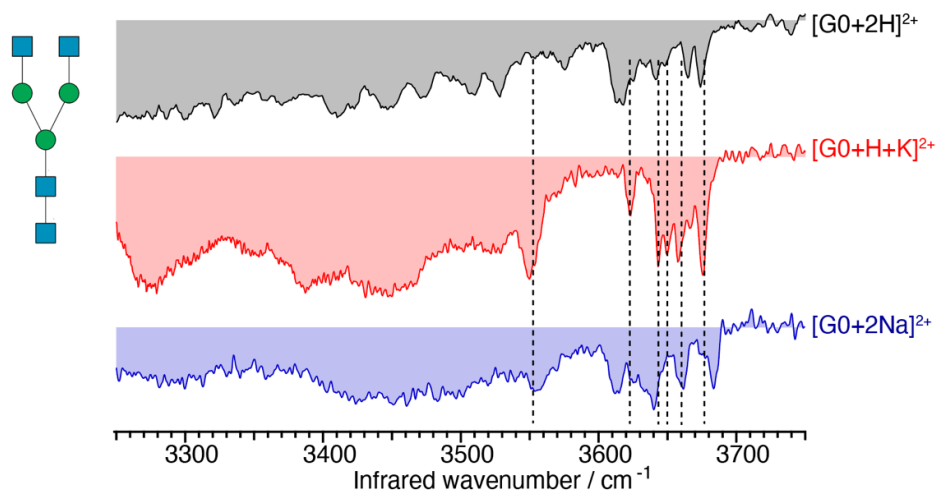


Figure 7.7 Cryogenic IR-spectra of $[G0+2H]^{2+}$ (grey), of $[G0+H+K]^{2+}$ (red), and of $[G0+2Na]^{2+}$ (blue). Dashed lines are used to highlight the differences.

To further assess the applicability of various metal adducts for glycan identification, we also studied Man5 N-glycan. Figure 7.8 shows a mass spectrum of ion adduct species of Man5, including $[M+H+Na]^{2+}$ and $[M+H+K]^{2+}$, produced by nESI. As we can see, Man5 does not form the doubly sodiated species if we do not add extra sodium into the solution. Therefore, recording the IR spectra of other predominant adducts for constructing our database is very useful if we would like to analyze the glycan samples without adding extra sodium.

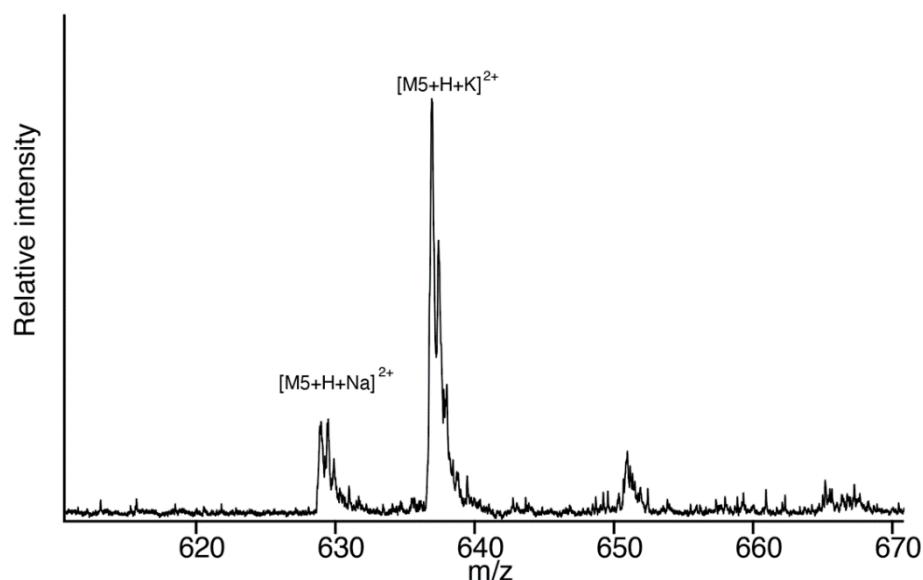


Figure 7.8 Typical mass spectrum of ion adducts of Man5 N-glycan standard formed in positive ESI mode.

The comparison of the cryogenic vibrational spectra between $[H+Na]^{2+}$ and $[H+K]^{2+}$ adducts of Man5 glycan is shown in Figure 7.9. Similarly to G0, the spectra of the $[H+Na]^{2+}$ and $[H+K]^{2+}$ states differ in the intensities and band positions in the free OH stretch region ($3580\text{--}3700\text{ cm}^{-1}$). Moreover, broader bands are observed in the weakly hydrogen-bonding OH stretch region ($3450\text{--}3550\text{ cm}^{-1}$) which are still different between the two species. Each of the IR spectrum can serve as a fingerprint since both cryogenic spectra present a large number of resolved IR transitions.

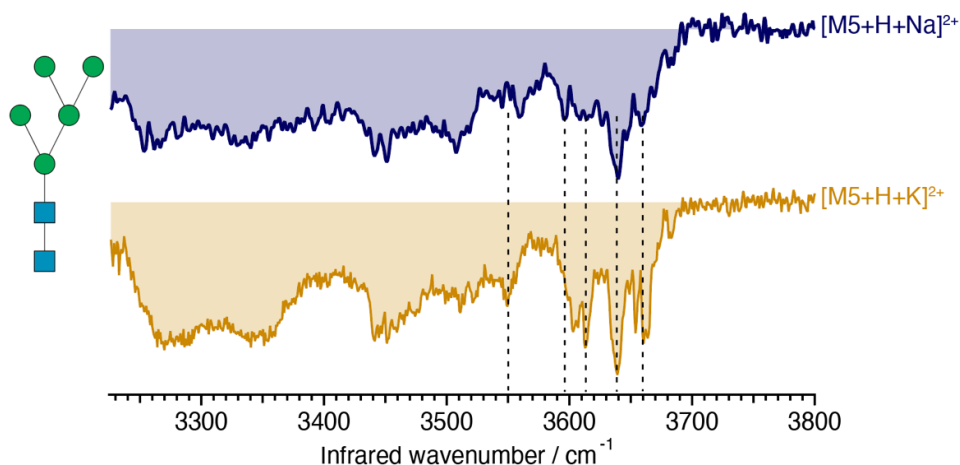


Figure 7.9 Cryogenic IR-spectra of $[M5+H+Na]^{2+}$ (blue) and $[M5+H+K]^{2+}$ (yellow). Dashed lines are used to highlight the differences.

The next example corresponds to Man6 sugar that is larger than Man5 by one mannose residue. Figure 7.10 demonstrates a mass spectrum of the Man6 N-glycan standard. Several metal adducts are formed, including $[M + H + Na]^{2+}$, $[M + H + K]^{2+}$, $[M + 2Na]^{2+}$, and $[M + Na + K]^{2+}$.

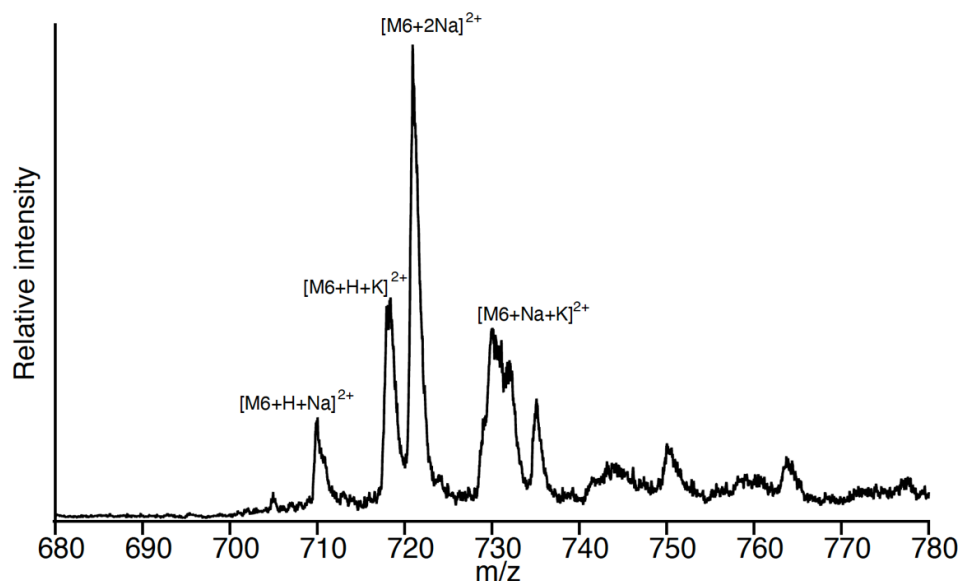


Figure 7.10 Typical mass spectrum of ion adducts of Man6 N-glycan standard formed in positive ESI mode.

We recorded cryogenic IR spectra of the four metal adducts, such as $[M+H+Na]^{2+}$, $[M+H+K]^{2+}$, $[M+2Na]^{2+}$, and $[M+Na+K]^{2+}$ of Man6 without ion mobility selection, and the spectra are shown in Figure 7.11. The spectrum of each metal adduct has unique well-resolved transitions in the free OH stretch region ($3580\text{--}3700\text{ cm}^{-1}$) that can be used for identification of Man6 glycan. The spectrum of $[M+H+K]^{2+}$ has broad bands in the weakly hydrogen-bonded OH stretch region ($3270\text{--}3550\text{ cm}^{-1}$) that are also distinctive. In general, this data clearly demonstrates that any of the cryogenic spectra can be used for unambiguous identification of the sugar regardless of the identity of the cation.

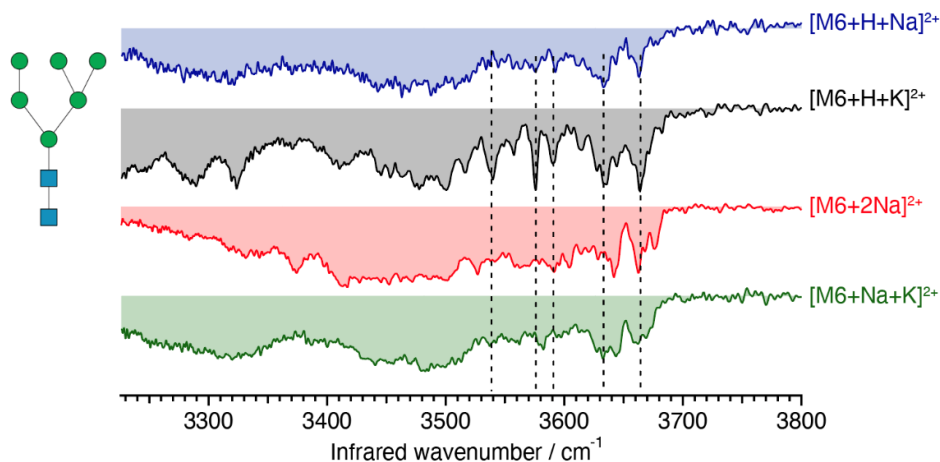


Figure 7.11 Cryogenic IR-spectra of $[M6+H+Na]^{2+}$ (blue), $[M6+H+K]^{2+}$ (grey), $[M6+2Na]^{2+}$ (red) and $[M6+Na+K]^{2+}$ (green). Dashed lines are used to highlight the differences.

7.4 Conclusions and future perspectives

This chapter demonstrated that cryogenic infrared spectra of various alkali metal complexes of carbohydrates of different complexity can serve as fingerprints for their identification. Additionally, we showed that the IR spectral complexity is not affected by the identity of the metal cation.

Cryogenic spectroscopy of disaccharide metal adducts showed that the different strength of the interaction between a carbohydrate molecule and each of the five alkali metal ions translates into different frequency shifts for OH and NH stretching modes. It results in Li and Cs ion adducts having the largest and smallest shifts in vibrational spectra, respectively, when compared to protonated form.

The vibrational spectra of different metal adducts of N-glycans demonstrate the power of cryogenic IR spectroscopy to provide the resolved features even for various metal adducts of the relatively large molecules. The obtained spectra provide sufficient evidence that even with increasing size, we can still obtain a distinct and unique IR band pattern for the most abundant species, and this pattern can be used as a fingerprint for unambiguous identification of the glycan. Moreover, as was shown previously, the IR spectra of large glycans can be significantly simplified using isomer/conformer separation offered by ultrahigh-resolution ion mobility. In this

respect, various metal adducts of glycans can be used for IMS separation in the case when the sodiated species cannot be baseline separated. Thus, the data presented in this chapter provide a good starting point for further investigations on the unambiguous identification of glycans within complex mixtures.

References:

- (1) Fenn, L. S.; McLean, J. A. Structural Resolution of Carbohydrate Positional and Structural Isomers Based on Gas-Phase Ion Mobility-Mass Spectrometry. *Phys. Chem. Chem. Phys.* **2011**, *13* (6), 2196–2205. <https://doi.org/10.1039/C0CP01414A>.
- (2) Cancilla, M. T.; Penn, S. G.; Carroll, J. A.; Lebrilla, C. B. Coordination of Alkali Metals to Oligosaccharides Dictates Fragmentation Behavior in Matrix Assisted Laser Desorption Ionization/Fourier Transform Mass Spectrometry. *J. Am. Chem. Soc.* **1996**, *118* (28), 6736–6745. <https://doi.org/10.1021/ja9603766>.
- (3) Shannon, R. D. Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. *Acta Crystallogr. A* **1976**, *32* (5), 751–767. <https://doi.org/10.1107/S0567739476001551>.

Chapter 8 Spectroscopy of GRGDS and SDGRG

The work presented in this chapter is not in the main theme of this thesis, which is the analysis of glycans. However, it represents work that I performed in the early part of my PhD research.

The study presented in this chapter focuses on the structural characterization of the reverse peptides GRGDS and SDGRG in the gas phase. These peptides have biological relevance in cellular signaling in addition to being used as calibrants in ion mobility. The goal is to determine the actual gas-phase structures of these peptides and understand why they exhibit slightly different drift times and CCS values for the singly and doubly charged species as well as the negatively charged species. The approach, which is based on the combination of cold ion spectroscopy and ^{15}N isotopic labeling to determine the structures of GRGDS and SDGRG, is briefly described. We present here cryogenic ion spectra of these molecules as both positive and negative ions and provide assignments for the NH stretch bands.

To provide a complete structural characterization of these molecules, these highly resolved spectra must be combined with accurate density-functional theory (DFT) calculations of the most stable structures and their corresponding spectra, and this is being done in collaboration with Dr. Vasyl Yatsyna from our group and Prof. Riccardo Spezia from Sorbonne University. The spectra presented here provide stringent selection criteria for the lowest energy structures that will be applied to their calculations once they are completed.

8.1 RGD peptide motif and its biological relevance

Here, two reverse peptides, GRGDS (Gly-Arg-Gly-Asp-Ser) and SDGRG (Ser-Asp-Gly-Arg-Gly), shown in Figure 8.1, are investigated. Both contain the amino acid Arg-Gly-Asp (RGD) sequence. This sequence, known as the RGD-motif, is essential for interactions between cells and the extracellular matrix (ECM) mediated by cell

surface receptors called integrins [1,2]. In this motif, the Arg residue interacts with two Asp residues to form salt-bridge hydrogen bonds in one subunit (α) of the integrins while the Asp residue interacts with the other β subunit. Glycine establishes weak C-H \cdots O=C hydrogen bonds [3]. This specific binding initiates cell-signaling processes.

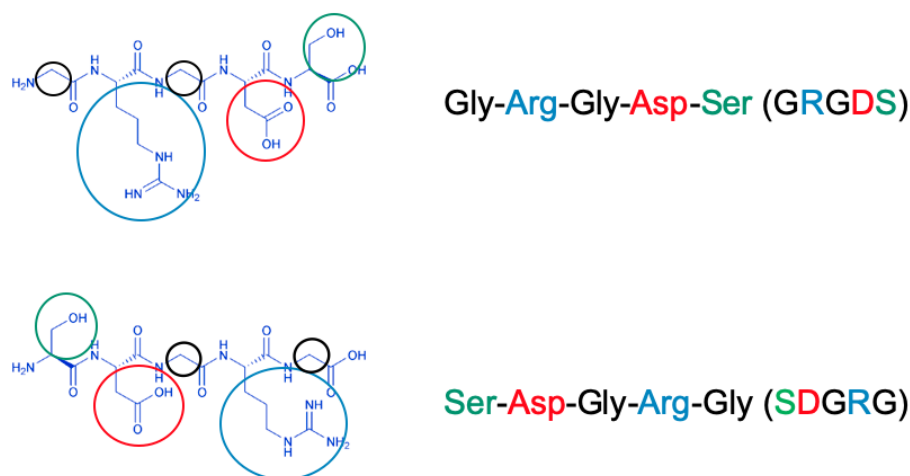


Figure 8.1 Chemical structures of GRGDS and SDGRG. Color circles correspond to the color-coded amino acids residues in sequences.

Originally, Pierschbacher and Ruoslahti discovered the RGD sequence in fibronectin as the minimal integrin-binding motif [4]. Later, this sequence was found in other cell adhesive ECM proteins, such as vitronectin, fibrinogen, and osteopontin [5,6]. The continuing studies lead to the development of RGD-based integrin ligands with different levels of specificity [7]. Moreover, the synthesis of RGD peptides is relatively simple and low cost. It makes peptides containing the RGD sequence a good choice for biomaterials applications. For example, the RGD-based ligands are frequently used in tissue engineering and oncology [8,9]. Consequently, Garanger *et al.* demonstrated that the RGD peptide ligands enable the specific delivery of chemotherapeutics to malignant tissue [8]. Finally, the RGD sequence has a significant role in cell adhesion biology as the prototype adhesion signal and the cell recognition motif.

To assess the impact of the RGD motif in adhesion studies, the GRGDS peptide and the inverse peptide SDGRS are normally evaluated together. The explanation is

that SDGRG does not stimulate integrins and is frequently used as a suitable control peptide [10–12].

8.2 Calibration mixture of pentapeptides GRGDS and SDGRG for ion-mobility set-ups

A mixture of GRGDS and SDGRG is well-known for being used to characterize the resolution and resolving power of ion mobility instruments. Even though the only difference between these two peptides is the reversed amino acid sequence, GRGDS and SDGRG represent an interesting sample set since their collisional cross sections are very similar. The CCS for doubly-charged reverse peptides GRGDS and SDGRG (294.1 m/z) differ in Ω value by 11 Å² (~2.8% difference) [13]. In contrast, the singly charged peptides (491. 2 m/z) differ in Ω value by 3.2 Å² (~1 % difference) [14]. Another advantage of using these pentapeptides as IM standards is their commercial availability.

This reverse peptide sequence provides a convenient benchmark for comparing the IM resolution of different instruments. In 2000, these reverse peptides were first analyzed by the group of Hill on their atmospheric drift tube instrument [15]. They demonstrated fully baseline resolved ATDs for both doubly charged peptides. However, for singly charged species the two peptides exhibited the same collision cross section. In 2009, Kemper and Bowers used these same peptides to showcase their 2 m long high-resolution drift tube instrument [13], in which baseline separation of the doubly charged species was achieved. Then in 2011, Gills and Campuzano demonstrated the use of the reverse peptides GRGDS and SDGRG as a means of measuring and specifying the resolution of the first and second generation travelling wave instruments from Waters [16]. In 2015, Gills and coworkers characterized the cyclic IMS using singly charged reverse peptides [14]. In the same year, Zhang *et al.* used the singly charged species to estimate the resolving power of FAIMS [17]. In 2018, Warnke and coworkers used doubly charged species to characterize the resolution of the SLIM device [18]. In the same year, Chai used isomeric pentapeptides to calibrate the TIMS device [19].

Despite the extensive work on these reverse peptides in positive ion mode, they never been measured in negative ion mode by IM.

8.3 Current methods to characterize the structure of the reverse peptides in the gas phase

Despite their use as an ion mobility calibrant, the structure of GRGDS and SDGRG in the gas phase has not been investigated. The smaller RGD peptide and peptides containing the RGD sequence have been studied theoretically [20,21], and by applying mass spectrometry [22]. Moreover, the structures of protonated RGD were investigated by applying resonant infrared multiphoton dissociation spectroscopy (R-IRMPD) with a free-electron laser [23]. In 2009, Li *et al.* applied a combination of anion photoelectron spectroscopy and quantum chemistry calculations to establish a low-energy stable structure of $[\text{RGD-H}]^-$ [24]. In 2010, Bythell and coworkers investigated the gas-phase dissociation chemistry of the protonated $[\text{RGD+H}]^+$ ion containing Arg at the N-terminus [25]. Additionally, his group studied deprotonated $[\text{RGD-H}]^-$ [26]. They used to apply MS/MS, H/D exchange MS/MS, and IRMPD spectroscopy, and theory to understand the chemistry of this sequence. Here we apply cryogenic infrared spectroscopy as a means to characterize the structure of GRGDS and SDGRG.

8.4 Experimental methods

8.4.1 Tandem mass-spectrometer for cryogenic spectroscopy

Since all the details of the tandem mass spectrometer used in these experiments have been explained in Chapter 6, we briefly summarize them here. Positive and negative ions are produced by a nano-ESI source and injected into an ion funnel. The first quadrupole mass-filter selects the mass of the parent ion. The ions are then guided into a cold (~ 4 K) octupole ion trap, where they are cooled and tagged with helium buffer gas. An infrared spectrum of the mass-selected ions is recorded by monitoring the loss of the tag with a second quadrupole mass-filter as a function of the infrared frequency.

In this work positively and negatively charged ions have been studied. In our tandem mass spectrometer, a channeltron electron multiplier (DeTech, 402-A-H, Extrel) is used as detector of ions. The operating parameters for the positive-ion mode have been described in Chapter 6. To detect the anions, the dynode is kept at +500 V

to guide negatively charged ions to the channeltron entrance maintained at +1.9 kV. The channeltron exit is floated at +3.8 kV with respect to ground.

8.4.2 *Sample preparation*

The pentapeptide GRGDS was purchased from Sigma-Aldrich (USA) with more than 97% purity and its reverse isomer SDGRG was purchased from Bachem (Switzerland) with more than 97.5 % purity. All samples were used without further purification. All solvents were HPLC grade. For the positive-ion mode, peptide solutions were prepared in water: acetonitrile: formic acid (74.9: 25: 0.1) solvent composition at a final analyte concentration of ~ 80 μ M. For the negative-ion mode, peptides were dissolved in in water: 20mM NH₄Ac buffer, pH 8.5 (80: 20) solvent composition to the concentration of peptides of ~ 80 μ M. All prepared solutions were stored in the freezer at -20 °C.

8.4.3 *Isotopic labeling*

To assign the NH vibrational modes to specific amino-acids in the sequence, we performed ¹⁵N isotopic substitutions on SDGRG (Ser-Asp-Gly-Arg-Gly), as shown in Figure 8.2. Three labelled peptides were purchased from Bachem (Switzerland) with >97% isotopic purity:

(1) in the first peptide sequence the amide nitrogen Gly5 was substituted, resulting in a Ser-Asp-Gly-Arg-[¹⁵N]Gly sequence;

(2) in the second one, both amides of Ser and Gly3 was replaced with ¹⁵N, giving a Ser-[¹⁵N]Asp-[¹⁵N]Gly-Arg-Gly;

(3) in the third one, both Asp and Gly3 were replaced with their ¹⁵N counterparts, leading to [¹⁵N]Ser-Asp-[¹⁵N]Gly-Arg-Gly sequence.

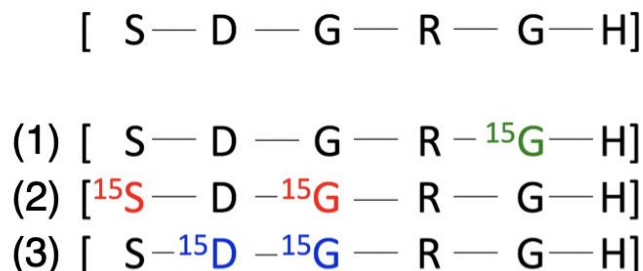


Figure 8.2 ^{15}N isotopic substitutions on SDGRG. (1) the amide nitrogen of Gly5 (green) is substituted; (2) both amides of Ser and Gly3 (red); (3) both of Asp and Gly3 (blue).

The solutions have been prepared in the same manner as non-labeled reagents. The prepared solutions were stored in the freezer at $-20\text{ }^{\circ}\text{C}$.

8.5 Cryogenic ion spectroscopy of GRGDS and SDGRG

8.5.1 Vibrational spectra of singly-charged GRGDS and SDGRG

Figure 8.3 shows vibrational spectra of protonated singly-charged GRGDS and SDGRG, recorded using helium tagging spectroscopy in a cryogenic octupole ion trap. As one can see, the spectra show sharp, distinct transitions, which is crucial for unambiguous structural determination. Even though the collisional cross-sections of singly charged GRGDS and SDGRG are very similar, cryogenic spectroscopy reveals striking differences between these pentapeptides.

The vibrational spectra were recorded in the $3000 - 3700\text{ cm}^{-1}$ region, which includes hydrogen stretching bands (X-H , where X is N, O, or C). These bands are sensitive to hydrogen bonding and provide important structural information. NH stretch vibrations of the NH_3^+ or NH_2^+ groups in protonated peptides typically appear strongly red-shifted into the region of $2900 - 3300\text{ cm}^{-1}$ [27,28], and shift further down upon hydrogen bonding [29,30]. The higher wavenumber region ($3570 - 3700\text{ cm}^{-1}$) covers free (i.e., non-hydrogen bonded) OH stretch vibrations.

The spectrum of SDGRG (Figure 8.3, red) is relatively simple, containing nine sharp bands on top of some broader background absorption. Two of these bands are located at 3573 cm^{-1} and 3580 cm^{-1} , and can be assigned to the free O-H stretches of the carboxylic acid groups (C-terminus and Asp residue) [31]. Additionally, three sharp transitions centered at 3473 cm^{-1} , 3500 cm^{-1} , and 3518 cm^{-1} are observed, which may correspond to NH stretches of the amide or amine group of the Arg (R) residue [32–34]. The remaining four bands are centered at 3334 cm^{-1} , 3341 cm^{-1} , 3355 cm^{-1} , and 3364 cm^{-1} and likely correspond to hydrogen-bonded stretching NH and OH bands.

The spectrum of GRGDS (Figure 8.3, blue) is more complex, having the two bands in the free OH region at 3572 cm^{-1} and 3668 cm^{-1} and a series of overlapping transitions at the hydrogen-bonding region. The former most probably correspond to a free carboxylic acid OH stretch (C-terminus and Asp residue COOH groups), and the latter is almost certainly the serine OH stretch band [33].

The particular spectral pattern for a given peptide molecule depends upon its three-dimensional structure, which can only be determined by comparing the experimental spectrum of the molecule with calculated spectra obtained for various structural candidates determined using quantum chemical calculations. For this purpose, the spectral bands must be sharp and distinctive. Indeed, in each measured spectrum (Figure 8.3), the full-width at half-maximum of the widest bands is $\leq 20\text{ cm}^{-1}$, with some being as narrow as 4 cm^{-1} . These well-resolved peaks provide stringent constraints imposed on the theoretical calculations and help to obtain a reliable structural assignment.

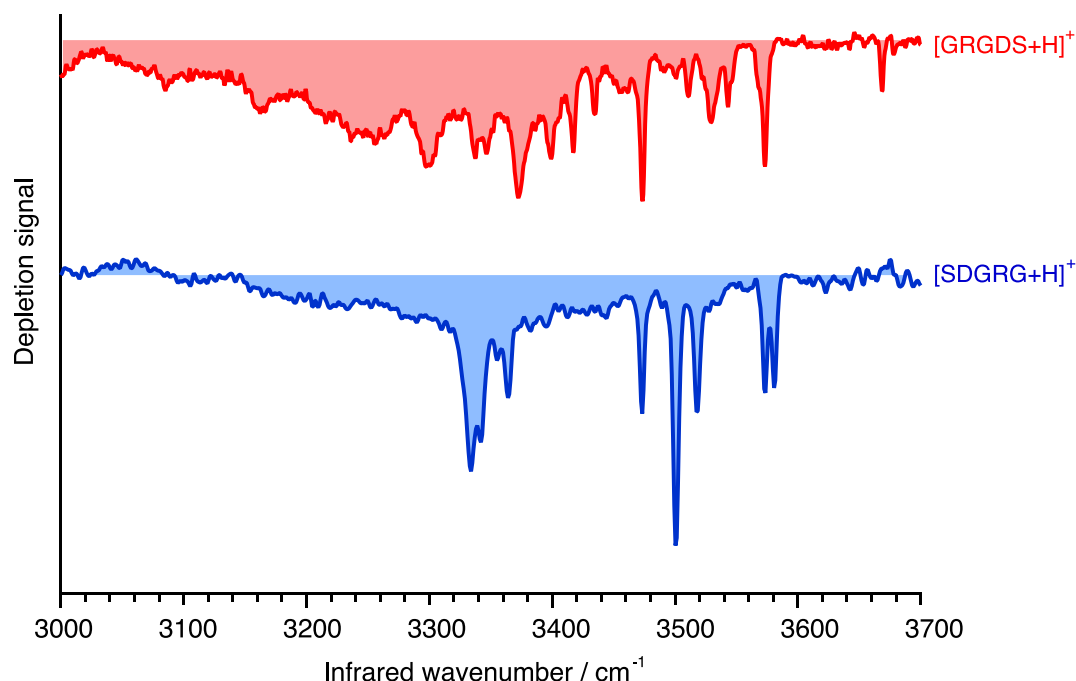


Figure 8.3 Comparison of IR spectra of the singly-charged GRGDS (red trace) and SDGRG (blue trace) at $m/z = 491.2\text{ Th}$.

8.5.2 *The assignment of amide NH stretches of singly-charged SDGRG*

The ultimate goal of this work is to elucidate the structure of these pentapeptides in the gas phase. Since the conformational space of these peptides is large, it is necessary to narrow down the conformational search by obtaining some additional information about possible conformers that can contribute to the experimental spectrum. This can, for example, be achieved by IR band assignment using a systematic isotopic substitution. For instance, the isotopic substitution of the amide nitrogen ($^{14}\text{N} \rightarrow ^{15}\text{N}$) at individual amino acid residues typically leads to a red-shift of $\sim 8\text{ cm}^{-1}$ of the NH stretch bands [35], which can be used for band assignment. The calculated spectra of conformers that agree with the band assignment found by isotopic substitution can be further compared with the overall IR spectrum for precise structural assignment.

In this section, the assignment of the NH-stretch bands in the spectra of the singly-charged SDGRG is discussed. Figure 8.4 compares the cryogenic vibrational spectra of $[\text{SDGRG}+\text{H}]^+$ with natural ^{14}N isotope abundance (solid black curves) to that with ^{15}N -substitution at individual amino acid residues (filled in color). The colored letters on the left side of each spectrum indicate the amino acid residues for which the nitrogen substitution was made.

In a case of the ^{15}N -substitution of the Gly5 (G) residue (Figure 8.4(a)), we did not observe any clear shift of a single band. However, a slight shoulder of the 3334 cm^{-1} band appeared. Moreover, some decrease in the intensity around 3335 cm^{-1} was observed. Both appearance of the shoulder and the dip at 3335 cm^{-1} indicate that there should be a band at around $3334\text{--}3335\text{ cm}^{-1}$ that shifts due to ^{15}N -substitution, but it is not fully resolved in this congested region of the spectrum.

Similarly, ^{15}N -substitution has been done for Gly3 (G) in two substituted peptide samples (Figure 8.4(b), (c)). In both samples the same band at 3364 cm^{-1} was red-shifted by $\sim 8\text{ cm}^{-1}$, confirming that this band corresponds to the NH stretch of the Gly3 residue. By employing the same approach, we can also assign the NH stretches of the D and S residues to the bands at 3331 cm^{-1} and 3440 cm^{-1} , respectively (see Figure 8.4(b) and (c)). However, some NH stretches can be strongly hydrogen bonded

and thus either shifter below 3200 cm^{-1} or broadened, which makes a shift of 8 cm^{-1} difficult to detect.

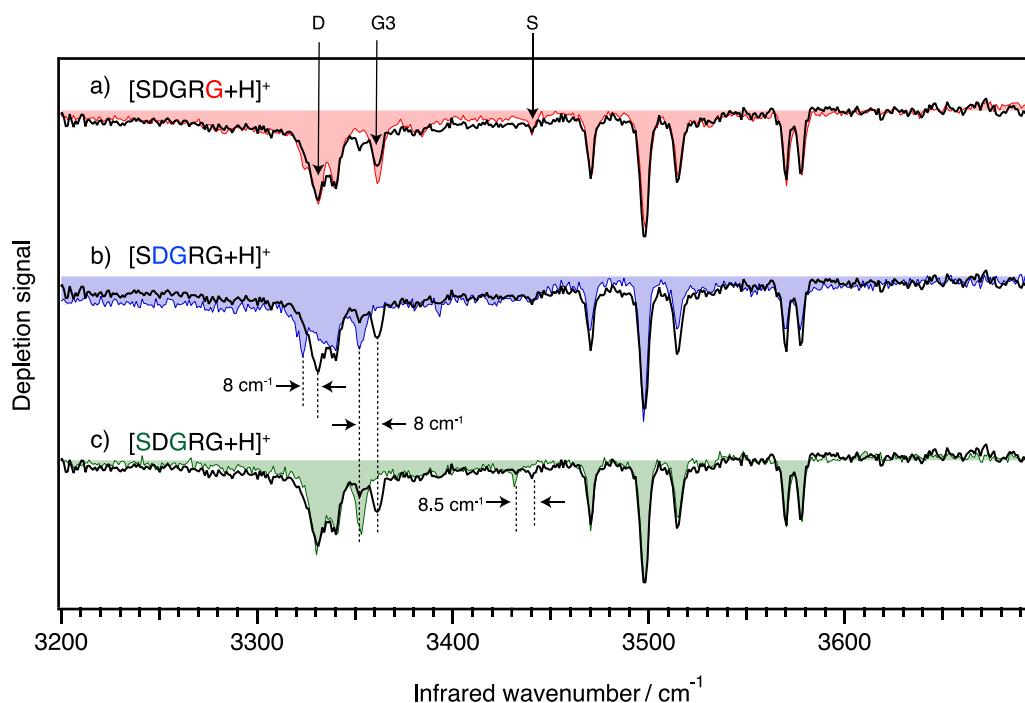


Figure 8.4 Cryogenic spectra of $[SDGRG + H]^+$ with ^{15}N - substitution at individual amino acid residues (filled in color) and with natural isotope ^{14}N abundance (solid black). Colored letters show the residues in which the nitrogen is substituted.

The assignment of the NH-stretch bands presented above provides a valuable benchmark for testing the theoretical structural predictions of the singly-charged SDGRG.

8.5.3 Vibrational spectra of the doubly-charged GRGDS and SDGRG

Figure 8.5 demonstrates the cryogenic vibrational spectroscopic signatures of doubly-charged GRGDS (black) and SDGRG (green). The recorded spectra are also distinct from one another and can be used for identification.

Similar to singly-charged GRGDS, the spectrum of the doubly-charged GRGDS (Figure 8.5, black) has numerous bands in the hydrogen-bonded O-H and N-H

stretching region. The vibrational spectrum of SDGRG (Figure 8.5, green) has bands centered at 3557 cm^{-1} , 3454 cm^{-1} , 3463 cm^{-1} , 3291 cm^{-1} , 3328 cm^{-1} , 3351 cm^{-1} , 3362 cm^{-1} , 3370 cm^{-1} , 3396 cm^{-1} , 3436 cm^{-1} , and 3670 cm^{-1} . These relatively broad bands likely signify the existence of multiple conformers. Both GRGDS and SDGRG spectra have bands at $\sim 3557\text{ cm}^{-1}$ which is characteristic of a free carboxylic acid O-H stretch, which could correspond to the aspartic acid residue or the carboxylic acid O-H stretch of the C-terminus. The band at 3670 is characteristic of a free serine OH stretch. These doubly-charged peptides lack the intense features near $\sim 3000\text{ cm}^{-1}$, suggesting that amine-protonated structures are strongly hydrogen bonded and shifted out of the measured frequency range. We have not yet performed isotopic substitution studies to assign these peaks.

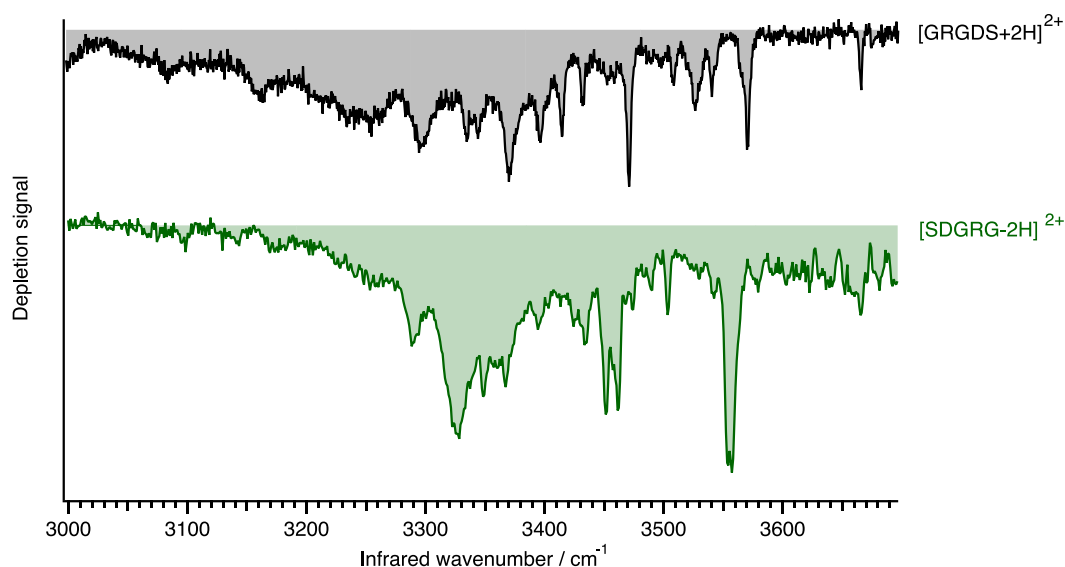


Figure 8.5 Comparison of IR spectra of the doubly-charged GRGDS (black trace) and SDGRG (green trace) at $m/z = 295.1$ Th.

8.5.4 The assignment of amide NH stretches of doubly charged SDGRG

We have used isotopic substitution to assign the NH stretch bands in the spectra of the doubly-charged SDGRG. Figure 8.6 shows the spectra of $[\text{SDGRG}+2\text{H}]^{2+}$, both with natural ^{14}N isotope abundance (solid black curves) and with ^{15}N - substitutions

(filled in color). The colored letters on the left side of each spectrum indicate the labeled amino acid residues in the molecule.

Interestingly, the redshift of the Gly5 (G) residue is not visible in the spectra (see Figure 8.6(a)). This result could possibly be explained by the low intensity of the band and/or high spectral congestion in its vicinity. The NH stretch band associated with the D residue (Figure 8.6(b)) can possibly correspond to a broad band at 3328 cm^{-1} . However, this peak undergoes a shift of $\sim 5\text{ cm}^{-1}$, which is smaller than expected for the NH oscillator.

In a case of the ^{15}N substitution of the Gly3 (G) residue, the redshift of $\sim 8\text{ cm}^{-1}$ was observed for two weak bands (see Figure 8.6 (b) and (c)). These shifts show that 3427 cm^{-1} and 3437 cm^{-1} bands are associated with the NH stretch of Gly3. The fact that two bands shift indicates that there are two conformers contributing to the measured spectra. This is confirmed by the similar trend observed when the S residue was isotopically substituted (see Figure 8.6 (c)), for which red-shifting of both 3351 cm^{-1} and 3370 cm^{-1} bands was detected.

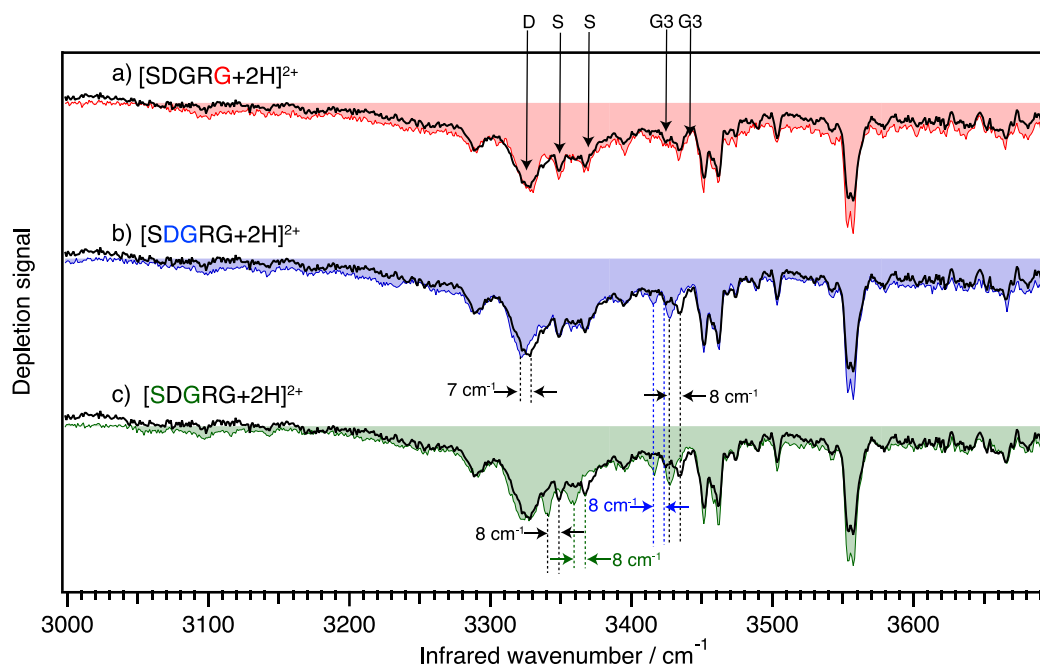


Figure 8.6 *Cryogenic spectra of $[SDGRG + 2H]^{2+}$ with ^{15}N -substitution at individual amino acid residues (filled in color) and with natural isotope ^{14}N abundance (solid black). Colored letters show the residues in which the nitrogen is substituted.*

It was not possible to assign all the NH stretch peaks of the isotopically labeled amino acids due to insufficient band intensity and spectral congestion. However, the assigned bands provide valuable constraints to the quantum chemical calculations which are used to predict the peptide gas-phase structure.

8.5.5 Vibrational spectra of deprotonated GRGDS and SDGRG

The study of the deprotonated peptides offers complementary information to the commonly investigated protonated peptides. Figure 8.7 shows the helium tagged vibrational spectra of the deprotonated GRGDS (purple) and SDGRG (orange). These spectra differ significantly in the intensities and band positions.

The pentapeptides have the guanidinium group on the arginine residue, one carboxylic acid group within the D residue and another carboxylic acid on the C-terminus. As the charge state of the molecules is 1-, we suggest that two possible deprotonation scenarios can exist. In one case, a structure might have only a single site of deprotonation to produce the negative charge. In a second case, a structure can have two sites of deprotonation producing two negative charges combined with a single positively charged guanidinium group at the arginine side chain [26].

The spectrum of GRGDS (Figure 8.7, purple) has broad bands centered at 3060 cm^{-1} , 3154 cm^{-1} and strong bands at 3262 cm^{-1} , 3281 cm^{-1} , 3303 cm^{-1} , 3332 cm^{-1} , 3400 cm^{-1} and 3412 cm^{-1} . All these intense peaks in the $3100 - 3400\text{ cm}^{-1}$ N-H stretch region indicate the presence of hydrogen bonds of various strengths. The symmetric and asymmetric stretches of the NH^{3+} moiety can also contribute to this region. The first two broad peaks (3060 cm^{-1} and 3150 cm^{-1}) might correspond to functional groups that are strongly hydrogen-bonded. No bands are observed in the free OH region. All these observations lead us to conclude that the deprotonation occurs according to the second scenario.

The spectrum of SDGRG (Figure 8.7, orange) contains sharp bands at 3305 cm^{-1} , 3403 cm^{-1} , and numerous broad weaker bands in the region of $3000 - 3270\text{ cm}^{-1}$. The sharp bands could correspond to NH stretches of the amino acid residues. The numerous broad bands likely correspond to hydrogen-bonded OH or NH.

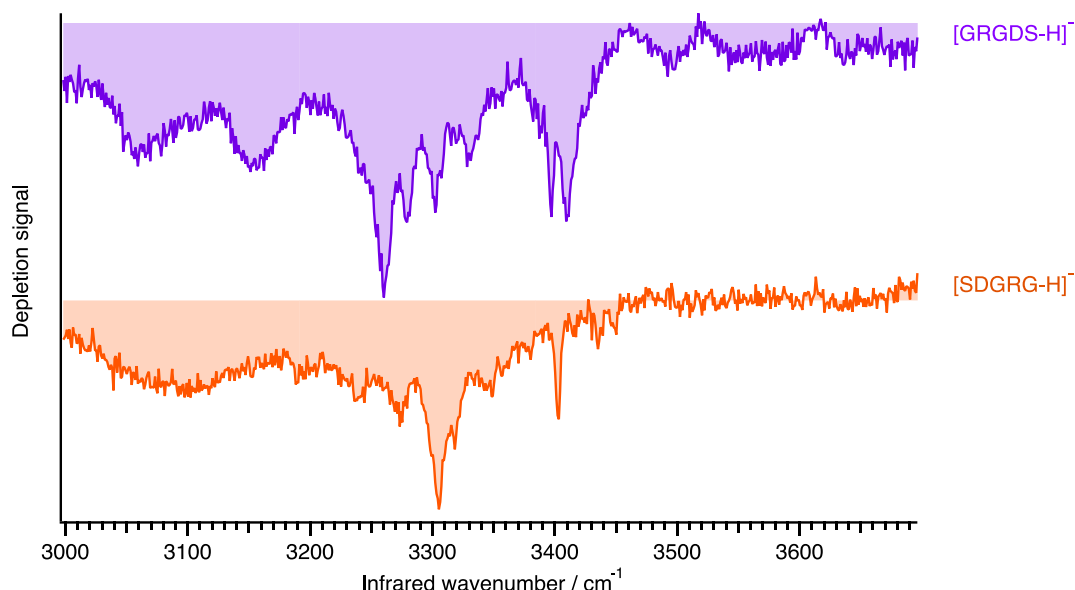


Figure 8.7 Comparison of IR spectra of deprotonated GRGDS (purple trace) and SDGRG (orange trace) at $m/z = 489$ Th.

8.5.6 The assignment of amide NH stretches of deprotonated SDGRG

The assignment of amide NH stretches in the vibrational spectra of deprotonated SDGRG is shown in Figure 8.8. The spectrum for unsubstituted sample (black) is compared to the spectra of samples where NH groups at individual amino acid residues were ^{15}N -substituted (filled in color). The letters in color on the left side of each spectrum designate the labeled amino acid residues in the peptide.

The first example of ^{15}N substitution is made at the C-terminal G residue (Figure 8.8(a)), which introduced three redshifts of the bands at 3403 cm^{-1} , 3435 cm^{-1} and 3451 cm^{-1} by $\sim 8\text{ cm}^{-1}$. In addition to attributing these bands to G5, this observation indicates the presence of three conformers contributing to the measured spectra.

The band at 3275 cm^{-1} can be assigned to the NH stretch of the G3 amino residue (Figure 8.8 (b) and (c)). This assignment is supported by the red-shift of this band by $\sim 8\text{ cm}^{-1}$ in both sequences with the labeled G3.

The next example (Figure 8.8 b) demonstrates a shift of the two bands 3318 cm^{-1} and 3348 cm^{-1} to lower energy by $\sim 7\text{ cm}^{-1}$. These shifts likely correspond to the NH stretches of the D amino acid since they both redshifted only in a single trace. However, the observation of two shifts is not as evident due to high spectral congestion in their vicinity. Even though it is not possible to assign all the NH stretches, one must take into account these assignments when performing quantum chemical calculations to elucidate the structure.

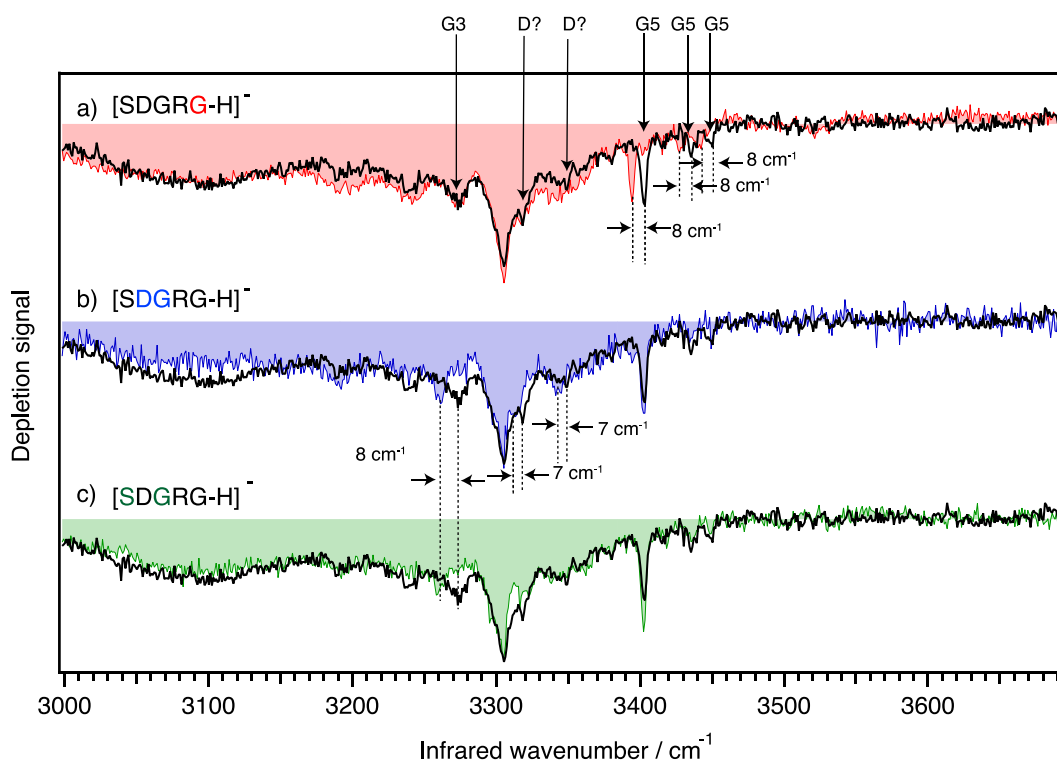


Figure 8.8 Cryogenic spectra of deprotonated SDGRG in the amide NH region with ^{15}N - substitution at individual amino acid residues (filled in color) and with natural isotope ^{14}N abundance (solid black). Colored letters show the residues in which the nitrogen was substituted.

8.6 Conclusions and future perspectives

This chapter describes the structural study of GRGDS and SDGRG using cryogenic messenger tagging spectroscopy. The vibrational spectra were measured for these peptides in the 1+, 1-, and 2+ charge states. Each of the acquired spectra can serve as a distinct fingerprint to identify the molecule.

Vibrational spectra were also recorded for isotopically labeled peptides SDGRG in the 1+, 1-, and 2+ charge states, which enabled us to assign the vibrational bands corresponding to different NH groups within the peptide backbone. The partially-assigned IR spectra with sharp and unique features represent a useful benchmark for theoretical predictions of 3-dimensional structure of these peptides. For this reason, our theoretical collaborators are currently working on conformational search and quantum-chemical calculations, which together with our experimental data can determine the precise structure of SDGRG in the gas phase.

Future experiments might require the assignment of the amide NH stretches of GRGDS by using the ^{15}N isotopic substitution. To reveal the abundance of the individual conformers, experiments using IR-IR double resonance spectroscopy of helium-tagged GRGDS and SDGRG will enable measuring vibrational fingerprints of individual conformers to assist theoretical predictions.

Considering the widespread use and importance of these peptides for ion-mobility calibration purposes, our contribution finally addresses the question of the GRGDS and SDGRG structural properties. This work serves as the first step towards the eventual structural determination of these peptides in the gas phase.

References:

- (1) Gottschalk, K.-E.; Kessler, H. The Structures of Integrins and Integrin–Ligand Complexes: Implications for Drug Design and Signal Transduction. *Angew. Chem. Int. Ed.* **2002**, *41* (20), 3767–3774. [https://doi.org/10.1002/1521-3773\(20021018\)41:20<3767::AID-ANIE3767>3.0.CO;2-T](https://doi.org/10.1002/1521-3773(20021018)41:20<3767::AID-ANIE3767>3.0.CO;2-T).
- (2) Locardi, E.; Mullen, D. G.; Mattern, R.-H.; Goodman, M. Conformations and Pharmacophores of Cyclic RGD Containing Peptides Which Selectively Bind Integrin $\text{Av}\beta 3$. *J. Pept. Sci.* **1999**, *5* (11), 491–506. [https://doi.org/10.1002/\(SICI\)1099-1387\(199911\)5:11<491::AID-PSC218>3.0.CO;2-8](https://doi.org/10.1002/(SICI)1099-1387(199911)5:11<491::AID-PSC218>3.0.CO;2-8).
- (3) Bella, J.; Humphries, M. J. $\text{C}\alpha\text{-H} \cdots \text{O}=\text{C}$ Hydrogen Bonds Contribute to the Specificity of RGD Cell-Adhesion Interactions. *BMC Struct. Biol.* **2005**, *5* (1), 4. <https://doi.org/10.1186/1472-6807-5-4>.
- (4) Pierschbacher, M. D.; Ruoslahti, E. Cell Attachment Activity of Fibronectin Can Be Duplicated by Small Synthetic Fragments of the Molecule. *Nature* **1984**, *309* (5963), 30–33. <https://doi.org/10.1038/309030a0>.
- (5) Felding-Habermann, B.; Cheresh, D. A. Vitronectin and Its Receptors. *Curr. Opin. Cell Biol.* **1993**, *5* (5), 864–868. [https://doi.org/10.1016/0955-0674\(93\)90036-p](https://doi.org/10.1016/0955-0674(93)90036-p).
- (6) Arnaout, M. A.; Mahalingam, B.; Xiong, J.-P. Integrin Structure, Allostery, and Bidirectional Signaling. *Annu. Rev. Cell Dev. Biol.* **2005**, *21*, 381–410. <https://doi.org/10.1146/annurev.cellbio.21.090704.151217>.
- (7) Meyer, A.; Auernheimer, J.; Modlinger, A.; Kessler, H. Targeting RGD Recognizing Integrins: Drug Development, Biomaterial Research, Tumor Imaging and Targeting. *Curr. Pharm. Des.* **2006**, *12* (22), 2723–2747. <https://doi.org/10.2174/138161206777947740>.
- (8) Khan, F.; Tanaka, M.; Ahmad, S. R. Fabrication of Polymeric Biomaterials: A Strategy for Tissue Engineering and Medical Devices. *J. Mater. Chem. B* **2015**, *3* (42), 8224–8249. <https://doi.org/10.1039/C5TB01370D>.

- (9) Garanger, E.; Boturyn, D.; Dumy, P. Tumor Targeting with RGD Peptide Ligands-Design of New Molecular Conjugates for Imaging and Therapy of Cancers. *Anticancer Agents Med. Chem.* **2007**, *7* (5), 552–558. <https://doi.org/10.2174/187152007781668706>.
- (10) Wildering, W. C.; Hermann, P. M.; Bulloch, A. G. M. Neurite Outgrowth, RGD-Dependent, and RGD-Independent Adhesion of Identified Molluscan Motoneurons on Selected Substrates. *J. Neurobiol.* **1998**, *35* (1), 37–52. [https://doi.org/10.1002/\(SICI\)1097-4695\(199804\)35:1<37::AID-NEU4>3.0.CO;2-C](https://doi.org/10.1002/(SICI)1097-4695(199804)35:1<37::AID-NEU4>3.0.CO;2-C).
- (11) Wildering, W. C.; Hermann, P. M.; Bulloch, A. G. M. Rapid Neuromodulatory Actions of Integrin Ligands. *J. Neurosci.* **2002**, *22* (7), 2419–2426. <https://doi.org/10.1523/JNEUROSCI.22-07-02419.2002>.
- (12) Hessel, M. H. M.; Atsma, D. E.; van der Valk, E. J. M.; Bax, W. H.; Schalij, M. J.; van der Laarse, A. Release of Cardiac Troponin I from Viable Cardiomyocytes Is Mediated by Integrin Stimulation. *Pflugers Arch.* **2008**, *455* (6), 979–986. <https://doi.org/10.1007/s00424-007-0354-8>.
- (13) Kemper, P. R.; Dupuis, N. F.; Bowers, M. T. A New, Higher Resolution, Ion Mobility Mass Spectrometer. *Int. J. Mass Spectrom.* **2009**, *287* (1), 46–57. <https://doi.org/10.1016/j.ijms.2009.01.012>.
- (14) Kevin Giles; Jason L Wildgoose; Steven Pringle; David Langridge; Peter Nixon; John; Garside; Peter Carney. Characterising a T-Wave Enabled Multi-Pass Cyclic Ion Mobility Separator; 2015.
- (15) Wu, C.; Siems, W. F.; Klasmeier, J.; Hill, H. H. Separation of Isomeric Peptides Using Electrospray Ionization/High-Resolution Ion Mobility Spectrometry. *Anal. Chem.* **2000**, *72* (2), 391–395. <https://doi.org/10.1021/ac990601c>.
- (16) Giles, K.; Williams, J. P.; Campuzano, I. Enhancements in Travelling Wave Ion Mobility Resolution. *Rapid Commun. Mass Spectrom.* **2011**, *25* (11), 1559–1566. <https://doi.org/10.1002/rcm.5013>.
- (17) Zhang, X.; Ibrahim, Y. M.; Chen, T.-C.; Kyle, J. E.; Norheim, R. V.; Monroe, M. E.; Smith, R. D.; Baker, E. S. Enhancing Biological Analyses with Three

Dimensional Field Asymmetric Ion Mobility, Low Field Drift Tube Ion Mobility and Mass Spectrometry (MFAIMS/IMS-MS) Separations. *The Analyst* **2015**, *140* (20), 6955–6963. <https://doi.org/10.1039/c5an00897b>.

(18) Warnke, S.; Faleh, A. B.; P. Pellegrinelli, R.; Yalovenko, N.; R. Rizzo, T. Combining Ultra-High Resolution Ion Mobility Spectrometry with Cryogenic IR Spectroscopy for the Study of Biomolecular Ions. *Faraday Discuss.* **2019**, *217* (0), 114–125. <https://doi.org/10.1039/C8FD00180D>.

(19) Chai, M.; Young, M. N.; Liu, F. C.; Bleiholder, C. A Transferable, Sample-Independent Calibration Procedure for Trapped Ion Mobility Spectrometry (TIMS). *Anal. Chem.* **2018**, *90* (15), 9040–9047. <https://doi.org/10.1021/acs.analchem.8b01326>.

(20) Krivov, S. V.; Karplus, M. Free Energy Disconnectivity Graphs: Application to Peptide Models. *J. Chem. Phys.* **2002**, *117* (23), 10894–10903. <https://doi.org/10.1063/1.1517606>.

(21) Suvire, F. D.; Rodríguez, A. M.; Mak, M. L.; Papp, J. G.; Enriz, R. D. Binding Mechanism of RGD and Its Mimetics to Receptor GPIIb/IIIa. A Theoretical Study. *J. Mol. Struct. THEOCHEM* **2001**, *540* (1), 257–270. [https://doi.org/10.1016/S0166-1280\(01\)00349-9](https://doi.org/10.1016/S0166-1280(01)00349-9).

(22) Solouki, T.; Fort, R. C.; Alomary, A.; Fattahi, A. Gas Phase Hydrogen Deuterium Exchange Reactions of a Model Peptide: FT-ICR and Computational Analyses of Metal Induced Conformational Mutations. *J. Am. Soc. Mass Spectrom.* **2001**, *12* (12), 1272–1285. <https://doi.org/10.1021/jasms.8b01561>.

(23) Grégoire, G.; Gaigeot, M. P.; Marinica, D. C.; Lemaire, J.; Schermann, J. P.; Desfrancois, C. Resonant Infrared Multiphoton Dissociation Spectroscopy of Gas-Phase Protonated Peptides. Experiments and Car–Parrinello Dynamics at 300 K. *Phys. Chem. Chem. Phys.* **2007**, *9* (24), 3082–3097. <https://doi.org/10.1039/B618094A>.

(24) Li, X.; Wang, H.; Bowen, K. H.; Grégoire, G.; Lecomte, F.; Schermann, J.-P.; Desfrancois, C. The Parent Anion of the RGD Tripeptide: Photoelectron

Spectroscopy and Quantum Chemistry Calculations. *J. Chem. Phys.* **2009**, *130* (21), 214301. <https://doi.org/10.1063/1.3137095>.

(25) Bythell, B. J.; Csonka, I. P.; Suhai, S.; Barofsky, D. F.; Paizs, B. Gas-Phase Structure and Fragmentation Pathways of Singly Protonated Peptides with N-Terminal Arginine. *J. Phys. Chem. B* **2010**, *114* (46), 15092–15105. <https://doi.org/10.1021/jp108452y>.

(26) Guan, S.; Rabus, J. M.; Maître, P.; Bythell, B. J. Gas-Phase Dissociation Chemistry of Deprotonated RGD. *J. Am. Soc. Mass Spectrom.* **2020**. <https://doi.org/10.1021/jasms.0c00074>.

(27) Wassermann, T. N.; Boyarkin, O. V.; Paizs, B.; Rizzo, T. R. Conformation-Specific Spectroscopy of Peptide Fragment Ions in a Low-Temperature Ion Trap. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (6), 1029–1045. <https://doi.org/10.1007/s13361-012-0368-0>.

(28) Kamariotis, A.; Boyarkin, O. V.; Mercier, S. R.; Beck, R. D.; Bush, M. F.; Williams, E. R.; Rizzo, T. R. Infrared Spectroscopy of Hydrated Amino Acids in the Gas Phase: Protonated and Lithiated Valine. *J. Am. Chem. Soc.* **2006**, *128* (3), 905–916. <https://doi.org/10.1021/ja056079v>.

(29) DeBlase, A. F.; Harrilal, C. P.; Lawler, J. T.; Burke, N. L.; McLuckey, S. A.; Zwier, T. S. Conformation-Specific Infrared and Ultraviolet Spectroscopy of Cold [YAPAA+H]⁺ and [YGPA+H]⁺ Ions: A Stereochemical “Twist” on the β -Hairpin Turn. *J. Am. Chem. Soc.* **2017**, *139* (15), 5481–5493. <https://doi.org/10.1021/jacs.7b01315>.

(30) Leavitt, C. M.; DeBlase, A. F.; Johnson, C. J.; van Stipdonk, M.; McCoy, A. B.; Johnson, M. A. Hiding in Plain Sight: Unmasking the Diffuse Spectral Signatures of the Protonated N-Terminus in Isolated Dipeptides Cooled in a Cryogenic Ion Trap. *J. Phys. Chem. Lett.* **2013**, *4* (20), 3450–3457. <https://doi.org/10.1021/jz401681y>.

(31) Yimanouchi, T. *Molecular Vibrational Frequencies*, in: *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, Eds. Linstrom

PJ, Mallard WG,: National Institute of Standards and Technology, Gaithersburg, 2017.

(32) Zabuga, A. V.; Rizzo, T. R. Capping Motif for Peptide Helix Formation. *J. Phys. Chem. Lett.* **2015**, *6* (9), 1504–1508. <https://doi.org/10.1021/acs.jpcllett.5b00407>.

(33) Forbes, M. W.; Bush, M. F.; Polfer, N. C.; Oomens, J.; Dunbar, R. C.; Williams, E. R.; Jockusch, R. A. Infrared Spectroscopy of Arginine Cation Complexes: Direct Observation of Gas-Phase Zwitterions. *J. Phys. Chem. A* **2007**, *111* (46), 11759–11770. <https://doi.org/10.1021/jp074859f>.

(34) Polfer, N. C.; Oomens, J. Vibrational spectroscopy of bare and solvated ionic complexes of biological relevance. *Mass Spectrom. Rev.* **2009**, *28* (3), 468–494. <https://doi.org/10.1002/mas.20215>.

(35) Stearns, J. A.; Seaiby, C.; Boyarkin, O. V.; Rizzo, T. R. Spectroscopy and Conformational Preferences of Gas-Phase Helices. *Phys. Chem. Chem. Phys.* **2008**, *11* (1), 125–132. <https://doi.org/10.1039/B814143F>.

Appendix

Appendix A

Identification of cleaved N-glycans using the targeted database of standards

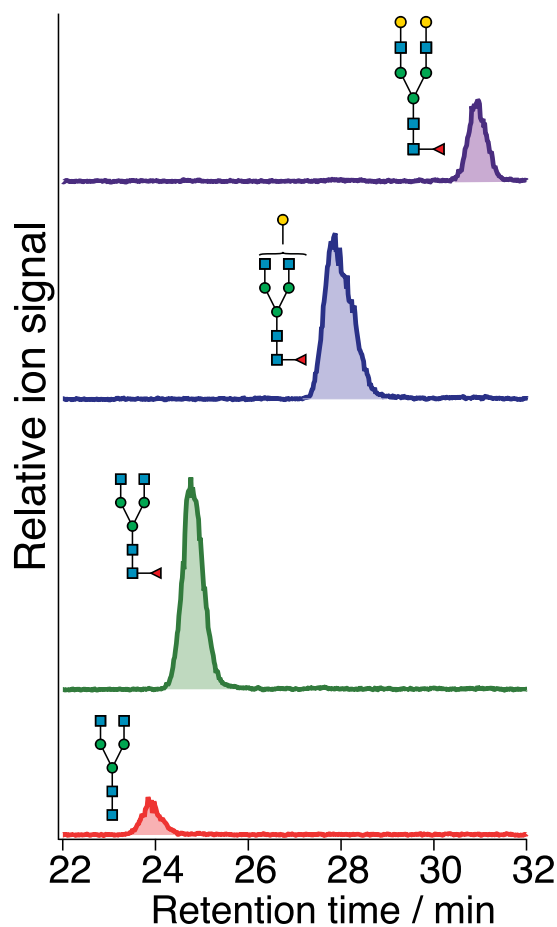


Figure A. 1 Typical chromatogram of fractions of the eluted glycans: *G0*, *G0F*, *G1F* and *G2F*

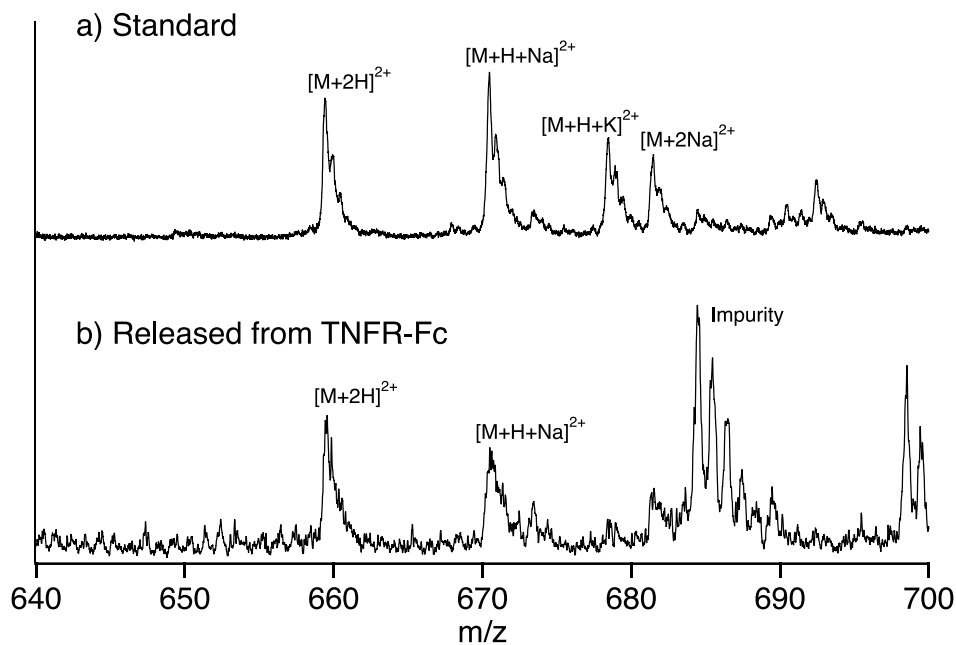


Figure A. 2 Comparison of ion adducts of G0 produced in positive ESI mode. a) standard, b) released of TNFR-Fc.

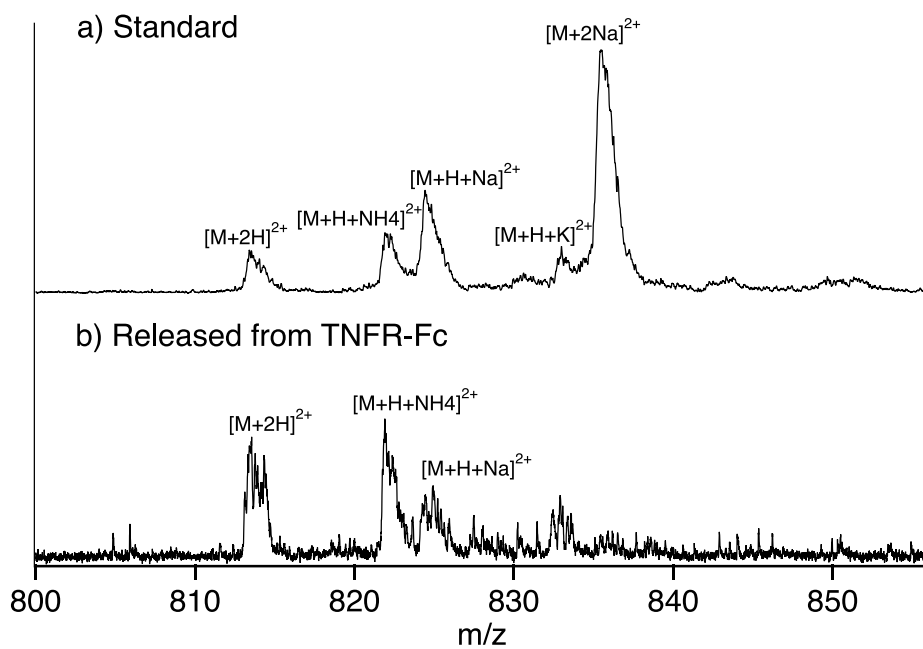


Figure A. 3 Comparison of ion adducts of G1F produced in positive ESI mode. a) standard, b) released of TNFR-Fc.

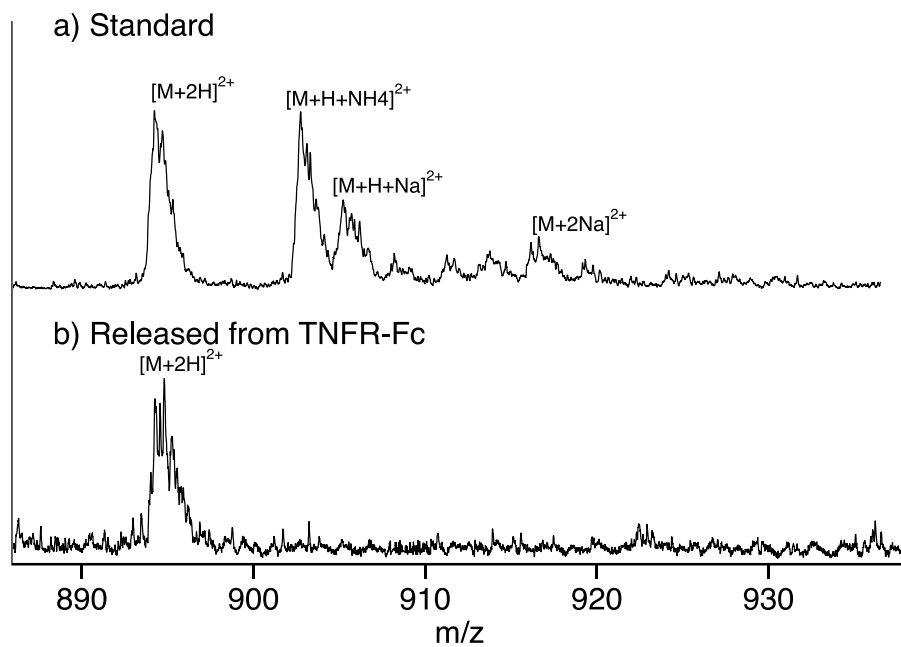


Figure A. 4 Comparison of ion adducts of G2F produced in positive ESI mode. a) standard, b) released of TNFR-Fc.

Appendix B

Deconvolution and quantitation of glycan mixtures

Table B. 1 *RMSD Values in % between the theoretical and experimental traces of the studied mixture. A-sugar A, B-sugar B, C-sugar C, A*-4'galactosyllactose, b-globotriose.*

	A+B 1 : 1	A+B 2 : 1	B+C 1 : 1	B+C 2 : 1	A+C 1 : 1	A+B+C 1 : 1 : 1	A*+B* 1 : 1	A*+B* 2+1
RMSD	29	54	31	48	13	62	40	41

Table B. 2 *RMSD Values in % between the recalculated theoretical and experimental traces of the studied mixture. A-sugar A, B-sugar B, C-sugar C, A*-4'galactosyllactose, b-globotriose.*

	A+B 1 : 1	A+B 2 : 1	B+C 1 : 1	B+C 2 : 1	A+C 1 : 1	A+B+C 1 : 1 : 1	A*+B* 1 : 1	A*+B* 2+1
RMSD	13	9	11	8	13	19	10	17

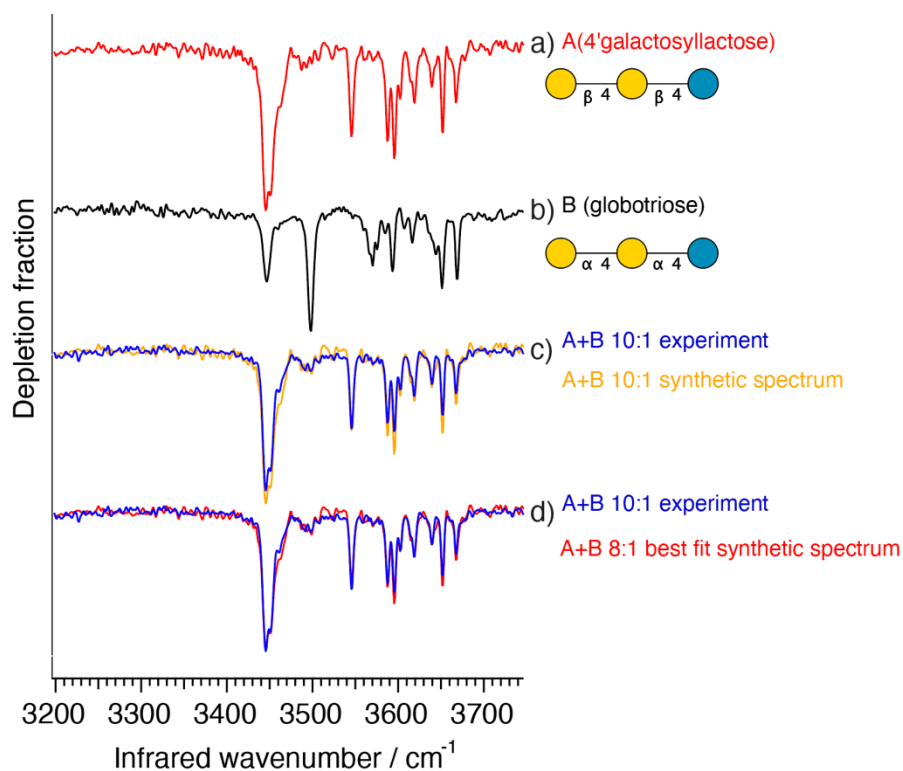


Figure B. 1 Comparison of experimental and synthetic spectra for sugars A and B(10:1), where A 750 μM and B 75 μM .

Table B. 3 Repeatability precision (%RSD, peak height) for the sugar A (75 μM) for IR spectrum normalized on power fluctuations.

Infrared wavenumber	Peak height	Average	STD	RSD, %
3267	-0.0003715	-0.00039951	3.52597E-05	8.82565433
	-0.0004391			
	-0.0003879			
3354	-0.0006812	-0.00073504	5.32785E-05	7.24838339
	-0.0007361			
	-0.0007878			
3386	-0.0006708	-0.00069385	8.35351E-05	12.039418
	-0.0006243			
	-0.0007865			
3651	-0.000841	-0.00098886	0.000159939	16.1741356
	-0.000967			

	-0.0011586			
--	------------	--	--	--

Table B. 4 Repeatability precision (%RSD, area) for the sugar A (75 μ M) for IR spectrum normalized on power fluctuations.

Infrared wavenumber	Area	Average	STD	RSD, %
3267	0.0102836	0.00996811	0.00040153	4.02818295
	0.0101046			
	0.00951612			
3354	0.013057	0.01227557	0.00074443	6.06429848
	0.012195			
	0.0115747			
3386	0.00999052	0.01058539	0.00132899	12.5549725
	0.0121079			
	0.00965774			
3651	0.0076559	0.00645905	0.00146392	22.6646631
	0.00482683			
	0.00689441			

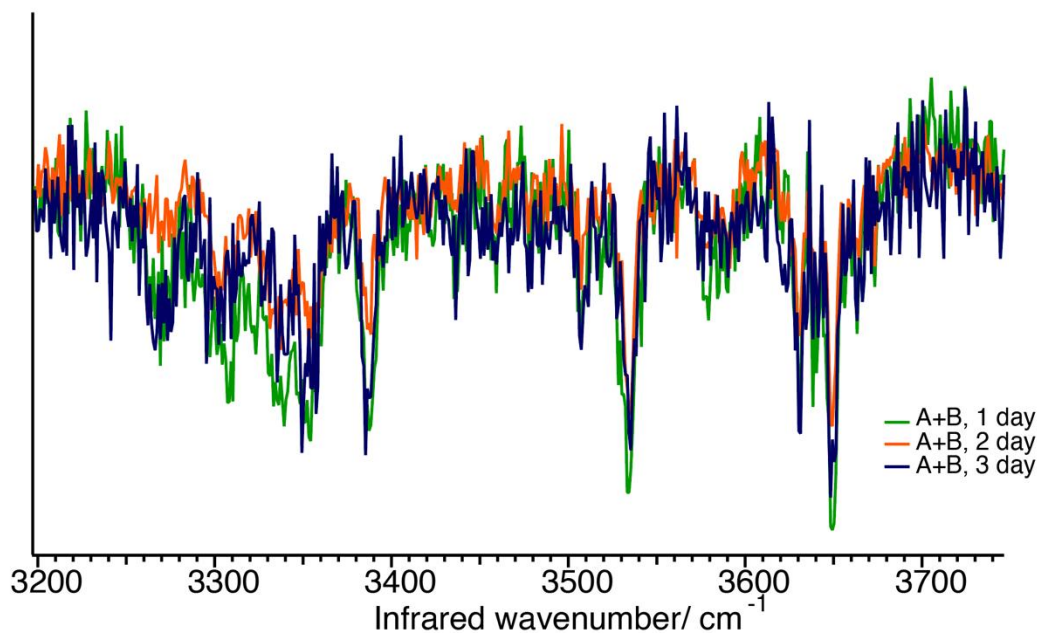


Figure B. 2 The repeatability precision validation for the mixture A and B (1:1)

Table B. 5 Repeatability precision (%RSD) for the mixture A and B (1:1) for IR spectrum.

Infrared wavenumber	Intensity	Average of three	Std deviation	RSD, %
3388	0.084	0.106	0.021	19.8
	0.108			
	0.126			
3535	0.130	0.142	0.012	8.5
	0.142			
	0.154			
3649	0.133	0.158	0.023	14.7
	0.162			
	0.179			

Table B. 6 Intermediate precision for the ten experimental measurements of the sugar A (75 μ M).

	MS, intensity	IR, depletion fraction
	3396	-0.33295
	3300	-0.24532
	3274.5	-0.20162
	3243.5	-0.22913
	3214.5	-0.20343
	3215	-0.20505
	3215	-0.17209
	3234	-0.18419
	3233	-0.18419
	3234	-0.18419
	3230	-0.17608
	2013.5	-0.14789
	2148	-0.17636
	2274.5	-0.1613
	2272.5	-0.1613
	2421.5	-0.14213

	2555.5		-0.33831
	2472.5		-0.34735
	2494.5		-0.31882
	2469		-0.25016
	2472.5		-0.26023
	2272.5		-0.30405
	248		-0.24033
	249		-0.3541
	254		-0.20922
	270		-0.30643
	256		-0.208842
	279		-0.24341
	210		-0.22348
	200		-0.24258
	246		-0.25648
	234		-0.182348
	236		-0.21346
Average	1949.6	Average	-0.2270774
STD	1284.6	STD	0.05949232
RSD,%	65.889	RSD, %	26.199

Table B. 7 Intermediate precision for the ten experimental measurements of a pentapeptide SDGRG (80 μ M).

SDGRG	Person1			Person2	Person3
MS 491 m/z	1 day	2 day	3 day	2 day	3 day
	478	979	1517	784.33	1565.5
	472	796	1961	784.33	1468.5
	452.5	843	1878.7	763	1552
	442.5	960	1654.3	702	1437.5
	400.5	973	1841.3	624.33	1282
	400.5	946.33	1788.3	609	1234.5
	353	934	1817.7	603	1265.55

	387.5	922	1638	586	1207.5
	393.5	850.67	1817.7	562.33	1075.5
	383.5	888.67	1817	573.67	1138.5
	384.5	887.67	1747.3	533.3	1178
	399.5	882.33	1761.3	542	1189
	421.5	870.67	1674	542	1176
	379	870.67	1674	763	1234.5
	377.5	889.67	1676	771	1552
Average	408.366667	899.578667	1750.90667	649.552667	1303.77
STD	36.9466153	52.0110485	112.657822	99.3094848	165.247565
RSD, %	9.04741212	5.78171209	6.4342563	15.2889042	12.6745948
IR 3500 cm ⁻¹	0.64298	0.62232	0.62799	0.65862	0.61348
	0.72088	0.63471	0.64178	0.6385	0.62452
	0.71211	0.72661	0.65049	0.57631	0.64137
	0.69844	0.65968	0.65607	0.6262	0.66392
	0.65433	0.67937	0.65607	0.58391	0.68713
	0.65433	0.69043	0.66923	0.60197	0.65324
	0.66776	0.71581	0.66923	0.61874	0.64377
	0.68322	0.65891	0.66923	0.58291	0.63805
	0.69161	0.66144	0.64178	0.59763	0.64377
	0.69161	0.70581	0.6317	0.60245	0.63805
Average	0.681727	0.675509	0.651357	0.610895	0.64473
STD	0.02609935	0.03433305	0.01534551	0.02957467	0.0203465
RSD, %	3.82841621	5.0825454	2.35592863	4.84120319	3.1558171

Table B. 8 *Dynamic range of our approach in a range of 0.5-1000 μ M for a glycan A.*

Glycan A	Concentration, μ M		Average	St D	RSD
	5000	2358	2263.33333	84.1566001	3.71825921
		2235			
		2197			
	1000	2425.5	2199.5	234.135111	10.6449243
		2215			
		1958			

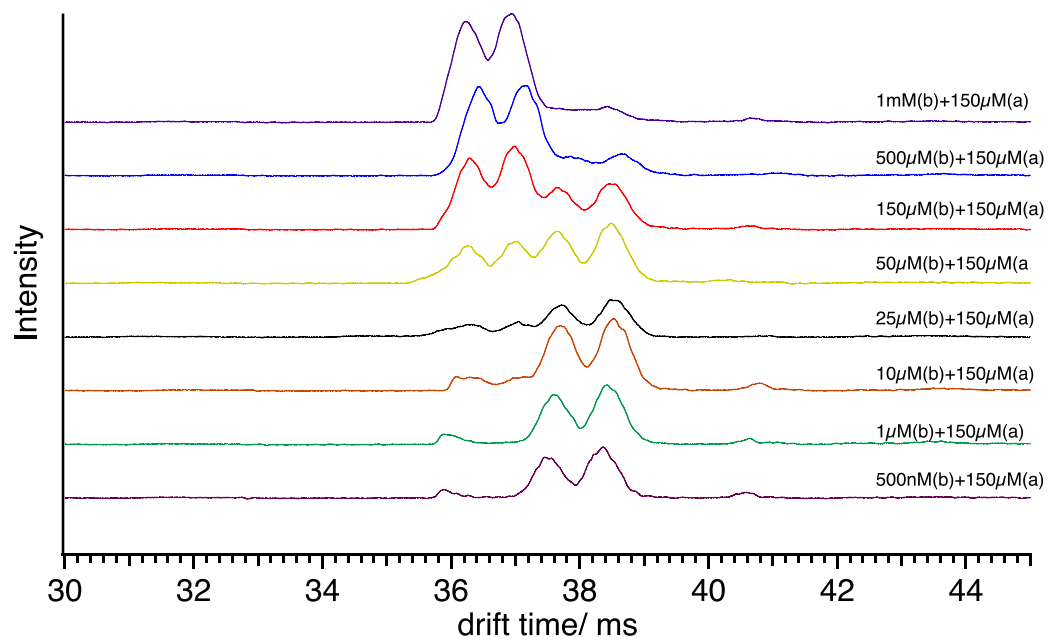
	150	2473.5	2368.33333	95.8127514	4.04557712
		2345.5			
		2286			
	50	1805	1738	58.3866423	3.35941555
		1711			
		1698			
	5	526.5	511.5	19.0984293	3.73380826
		518			
		490			
	1	218	238.666667	34.0783411	14.2786345
		220			
		278			
	0.2	281	287.333333	10.9696551	3.8177454
		281			
		300			
	0.1	310.5	287.666667	24.0849192	8.37250957
		290			
		262.5			
	0.01	188	209.333333	23.1804515	11.0734641
		234			
		206			

Table B. 9 *Dynamic range of our approach in a range of 0.5-1000 μM for a glycan A.*

Concentration, μM	TIC (height)	XIC (height)	TIC/XIC
500	2905	8500	0.34176471
	2753	7500	0.36706667
	1253	5000	0.2506
300	2103	6500	0.32353846
	1387	5000	0.2774
	405	3500	0.11571429
150	1371	3884	0.35298661
	1215.5	3584	0.33914621
	1150	3434	0.33488643
	300.5	2225	0.13505618
	86.5	1284.5	0.06734138

	2755	6068	0.45402109
	1252	4875	0.25682051
	1065	6500	0.16384615
75	1065	6000	0.1775
	643	4000	0.16075
	3	60	0.05
25	3	62	0.0483871
	202	1245	0.1034564
	404	2649	0.15251038
10	13	800	0.01625
	13	790	0.0164557
	11	1654	0.00665054
1	11	1402	0.00784593
	11	1309	0.00840336
	333	1801	0.18489728
0.5	89	922	0.09652928
	10	404	0.02475248
	17	513	0.0331384
BACKGROUND	5	113	0.04424779

Figure B. 3 Series of arrival time distributions after 7 cycles on SLIM board of the mixture A (150 μ M) and B with various concentrations.



Appendix C

Metal complexes of glycans

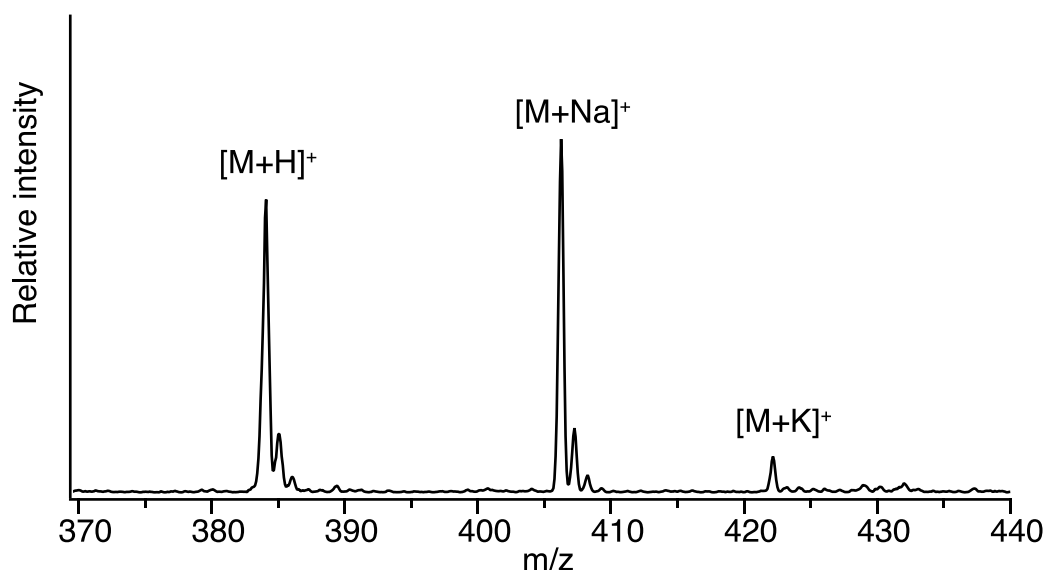


Figure C. 1 Typical mass spectrum of ion adducts of GalNAc α (1–3)Gal standard formed in positive ESI mode.

Theoretical predictions vs. experiment

In order to better understand the trends of the IR spectra observed in Fig. 7.3, Yatsyna performed quantum chemical calculation and obtained tentative structures* of the GalNAc α (1–3)Gal coordinated with lithium, sodium, and potassium (Figure C. 2). As a starting point for the calculations, the structure with sodium tentatively assigned by Scutelnic based on its experimental IR spectrum [1] was used. In this structure, the sodium ion interacts with 4 oxygen atoms of the disaccharide (see Figure C. 2(b)). This structure was submitted to a systematic conformational search using Avogadro and MMFF94 molecular mechanics force field. The selected 22 lowest-energy structures were further optimized with a B97D density functional and def2TZVPP basis set as implemented in Gaussian 16 software. The lowest energy structures were then submitted to a normal mode analysis in Gaussian to obtain harmonic frequencies and intensities.

In order to correct for anharmonicity and density functional errors, the calculated harmonic frequencies were scaled using mode-specific scaling factors [5]. The comparison between the experimental and calculated spectra of the three lowest energy conformers of GalNAc α (1–3)Gal coordinated with sodium is shown in Figure C. 3. The comparison between these spectra as well as the calculated spectra of other conformers shows that the best match is provided by the conformer 3, which corresponds to the structure shown in Figure C. 2(b).

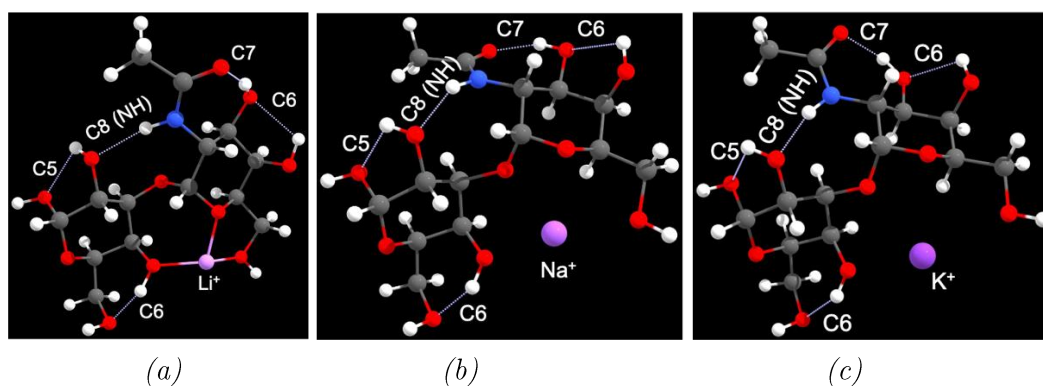


Figure C. 2 Structures of GalNAc α (1–3)Gal (α -anomer at reducing end OH) coordinated with lithium (a), sodium (b), and potassium (c), optimized with B97D density functional and def2TZVPP basis set.

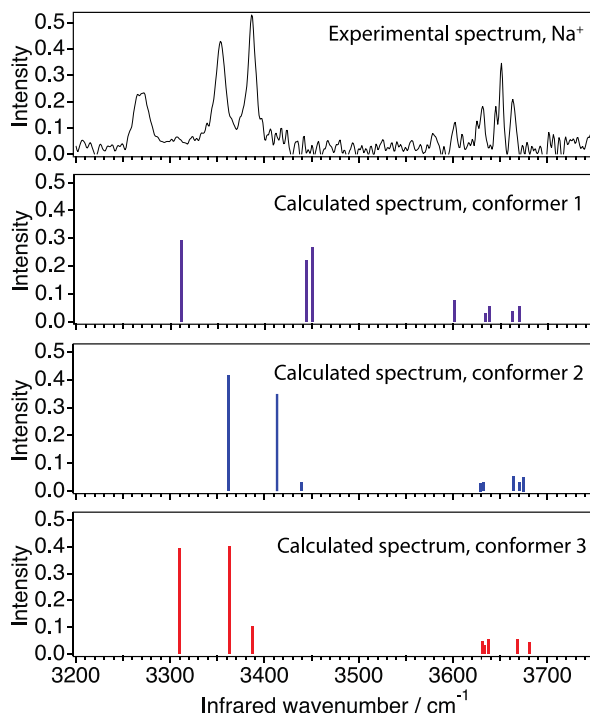


Figure C. 3 Comparison between the experimental spectrum of GalNAc α (1–3)Gal coordinated with sodium, and the spectra of the lowest energy conformers obtained using B97D density functional and def2TZVPP basis set.

Due to the similarity of the IR spectra of GalNAc α (1–3)Gal coordinated with lithium, sodium and potassium (see Figure 7.3), it is safe to assume that they share the same binding site of the alkali metal. Since the conformer 3 structure (see Figure C. 2(b)) matched well with the experimental spectrum of the sodiated disaccharide (Figure C. 3), we assumed that the binding site of lithium and potassium resembles the sodium binding site of the conformer 3. Therefore, the structures of GalNAc α (1–3)Gal coordinated with lithium and potassium were simply obtained by replacing the sodium atom in the conformer 3 structure with lithium and potassium, respectively, and reoptimizing those structures at B97D/Def2TZVPP level of theory. The comparison between the experimental and calculated spectra of the resulting structures is shown in Figure C. 4.

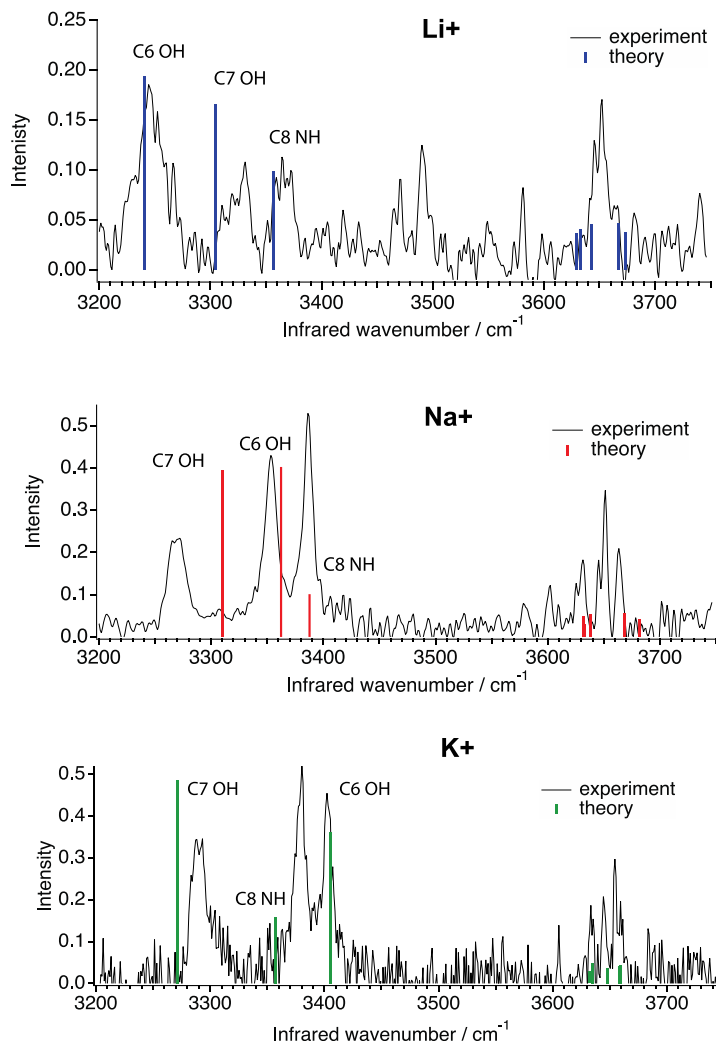


Figure C. 4 Comparison between the experimental and calculated spectra of *GalNAc α (1-3)Gal* coordinated with lithium (Li^+), sodium (Na^+) and potassium (K^+).

At this point it is now worth discussing the hydrogen bonds that are present in the structures and how they are affected by the presence of the alkali atoms. Figure C. 2 shows that all structures have four $\text{OH}\cdots\text{O}$ hydrogen bonds and one $\text{NH}\cdots\text{O}$ hydrogen bond. These bonds were labeled as C5, C6, C7 or C8, depending on the size of the ring that was closed upon the formation of the hydrogen bond. For example, C5 denotes a hydrogen bond closing a 5-membered ring. Among the outlined $\text{OH}\cdots\text{O}$ hydrogen bonds shown in Figure C. 2, C5 are the weakest ones, while C6 and C7 vary in strength among the structures with different metals. This can be seen by examining the hydrogen bond distances $(\text{O})\text{H}\cdots\text{O}$, where stronger bonds typically correspond to shorter distances.

Moreover, the strength of the hydrogen bond is visible in the frequency of the OH stretch vibration associated with the hydrogen bond. For example, the structure with lithium shows lower frequency for the OH stretch associated with C6 bond compared to C7 (see Figure C. 4), which implies that C6 bond is stronger.

Figure C. 4 shows that in all metal adduct structures the same three vibrational bands were affected by the nature of the alkali metal, corresponding to C6 OH, C7 OH and C8 NH stretching transitions. The most pronounced effect is visible for the OH stretch associated with the C6 OH...O hydrogen bond, where the OH group is in the close contact with the metal ion (see Figure C. 2). Indeed, the calculated frequency of this band shifts from 3240 cm⁻¹ in the case of Li⁺ to 3362 cm⁻¹ and 3406 cm⁻¹ in the case of Na⁺ and K⁺, respectively. It is worth noting that the calculated frequency of the C6 OH transition in the structure that does not have any alkali metals is 3565 cm⁻¹, meaning that the neighboring alkali metal introduces a strong blue-shift of the C6 OH stretch band, and the shift decreases with the size of the alkali metal ion. Such behavior correlates with the distance between the alkali ion and the oxygen atom of the OH group associated with C6 OH stretch. This metal-oxygen distance increases from 1.92 Å in Li⁺ structure to 2.34 Å and 2.81 Å in Na⁺ and K⁺ structures, respectively. Moreover, the C6 hydrogen bond distance (O)H...O increases from 1.74 Å in Li⁺ to 1.80 Å and 1.83 Å in Na⁺ and K⁺ structures, respectively, whereas the distance in structure having no metal ions is as large as 2.34 Å. Since the shorter hydrogen bond length is typically associated with the stronger bond, we can conclude that alkali metals drastically increase the strength of the adjacent OH...O hydrogen bonds, and this increase is larger for smaller alkali metals such as Li⁺.

The two other vibrational bands that were affected by the alkali metal nature, C7 OH and C8 NH, are associated with OH and NH groups that are not in direct contact with the alkali metals (see Figure C. 2). Therefore, the frequency variations for these bands are very slight and not so straightforward to explain. The calculations show that they originate from molecular structure variations induced by the alkali metals of varying ionic radii. Indeed, as was already mentioned above, the metal-oxygen distances strongly increase with the increase of the alkali metal ionic radii, and this also leads to the alterations in the hydrogen bond distances and angles even outside of the metal binding site.

To conclude with, our experiments and calculations show that the presence of the alkali metal ions in the disaccharide structure strongly increases the strength of the neighboring hydrogen bonds, and this effect decreases with the size of the alkali metal.

Moreover, slight variations in the strength of the hydrogen bonds that were not in close contact with the alkali metal ion were also observed and predicted, and these originate from slight structure variations due to increase in the metal-oxygen distance in the metal-ion binding site.

References:

- (1) Warnke, S.; Ben Faleh, A.; Scutelnic, S.; and Rizzo T. R. Separation and Identification of Glycan Anomers Using Ultrahigh-Resolution Ion-Mobility Spectrometry and Cryogenic Ion Spectroscopy *J. Am. Soc. Mass Spectrom.* **2019**, 30, 11, 2204–2211. <https://doi.org/10.1021/jasms.8b06266>
- (2) Shannon, R. D. Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. *Acta Crystallogr. A* **1976**, 32 (5), 751–767. <https://doi.org/10.1107/S056773947600155>
- (3) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, 27 (15), 1787–1799. <https://doi.org/10.1002/jcc.20495>.
- (4) Gloaguen, E.; Loquais, Y.; Thomas, J. A.; Pratt, D. W.; Mons, M. Spontaneous Formation of Hydrophobic Domains in Isolated Peptides. *J. Phys. Chem. B* **2013**, 117 (17), 4945–4955. <https://doi.org/10.1021/jp401499x>.

Appendix D

Construction of a targeted database of N-glycans

A database approach that we created using N-glycan standards is described. This database contains mass, ion mobility ATDs, and reference IR spectra. Several examples are listed.

D.1 Targeted database prototype

The previous studies from our group showed that one could use cryogenic spectroscopy to unambiguously identify the isomeric molecules [1–5]. Even for larger glycans, the significant resolving power of vibrational spectroscopy provides a unique spectral fingerprint [6].

Here, we discuss the database approach that provides three types of the information in a one measurement using a single instrument. Our approach could potentially become an ultimate tool for high-throughput glycan analysis.

To perform rapid glycan analysis, we envision a comprehensive database comparable to the existing LC-MS ones [7–9]. Our targeted database contains mass, ion mobility ATDs, and reference IR spectra. The ion mobility ATDs as well as the IR fingerprint of identified sugar are then compared to reference data. First, the IM dimension provides information about the possible amount of all existing isomers and conformers of the molecule. The additional spectroscopic dimension then allows us to collect all characteristic fingerprints for glycan isomers that reflect the subtle structural differences amongst all their isomerism. This method allows one to fully resolve the isomeric glycan complexity. Another advantage of having the database approach is an ability to identify glycan covalent structure instead of 3D structure, for that reason this approach does not require any extensive time-consuming quantum calculations [4].

A simplified prototype of the database is sketched in Figure D. 1 . To identify known glycans present in our targeted database, the workflow of data acquisition is relatively straightforward (Figure D. 1a). The data acquisition part consists of main three steps: 1) m/z measurement of the analyte; 2) measurement of the overall ATD of the selected m/z , 3) collection of its vibrational spectra of the overall ATDs. However, if overall ATD and its spectrum are not enough to identify the glycan specie, slicing of each peak in ATD for this glycan and the following recording of mass- and mobility-selected IR fingerprint will be provided. To simplify Figure D. 1, the described slicing of each peak in ATD has been omitted. Our ultimate goal is to create a database of mass, mobility, and IR fingerprint spectra for known N-glycans in a manner similar to described in Figure D. 1.

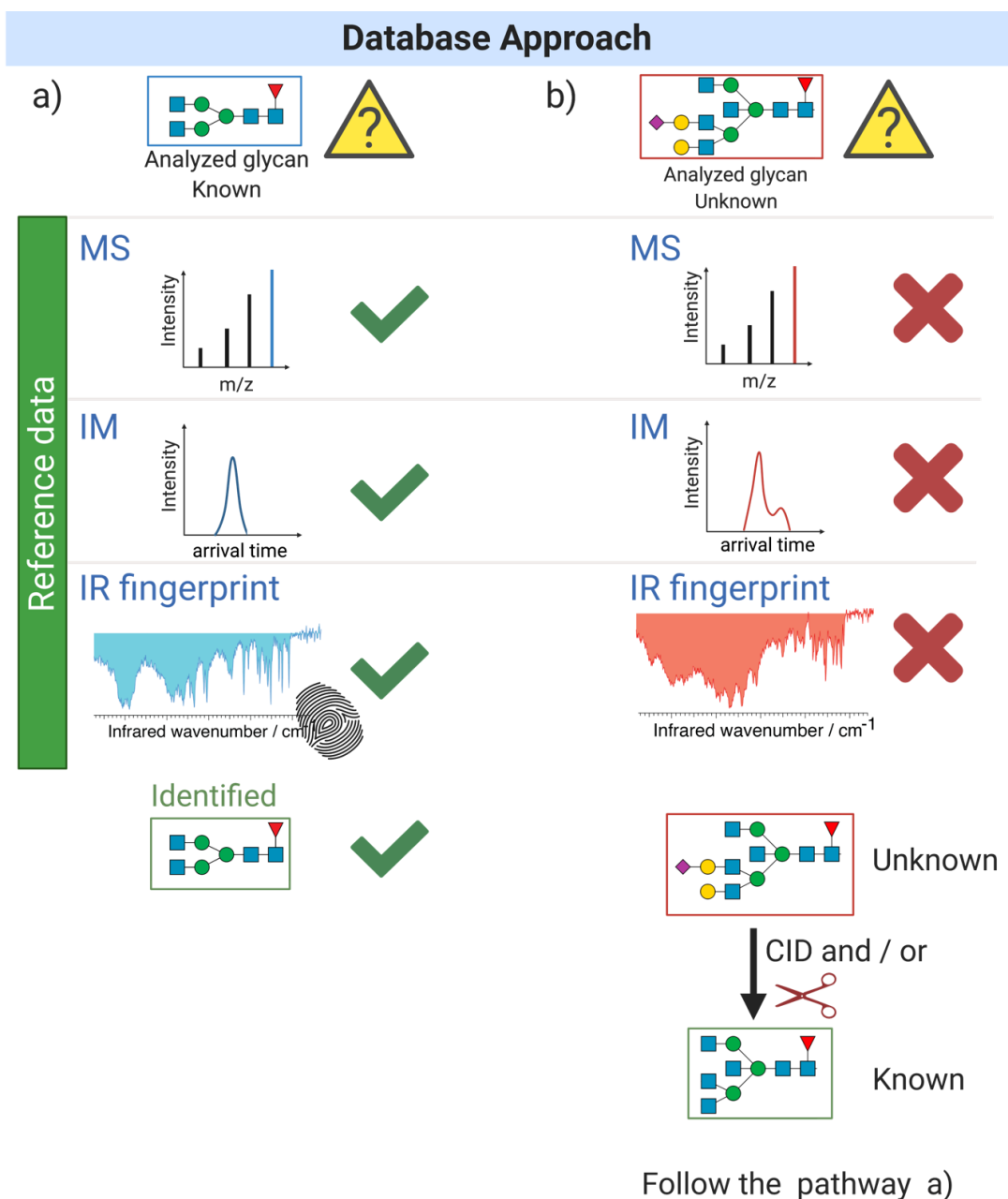


Figure D. 1 A prototype of the database structure for the consistent glycan identification including the m/z values of analytes, their corresponding ATDs and IR fingerprints. a) for known glycans, b) for unknown glycans.

For unknown glycans, we can use collision-induced dissociation techniques on the SLIM-IMS platform [10], as well as enzymatic degradation strategies [6], to determine parent glycan structures and include them in the database (Figure D. 1 b). Eventually, an automated glycan identification analogous to GlycoDeNovo and GIPS for tandem

MS will be developed and applied to make the database searches fast and reliable [11,12].

With a broad scientific audience in mind, we have to guarantee that the provided measurements are highly reproducible and the technique is broadly accessible. In the subsequent sections we provide an experimental proof of principle that our database approach should be sufficient for N-glycan characterization and identification.

D.2 Notation of standards

The majority of samples were purchased from Dextra Laboratories (United Kingdom). Man5-RapiFluor-MS has been purchased from Waters (United Kingdom).

Representative structures and names of the studied N-glycans are depicted using the SNFG notation (Figure D. 1). The heptasaccharide ‘core’ structure is frequently denoted in the literature by the shorthand nomenclature as G0, showing the absence of galactose. G1 and G2 hold the ‘core’ oligosaccharide structure plus one or two galactose residues, respectively. The ‘core’ with fucose is referred to as G0F. G0bF signifies the added fucose and the bisecting N-acetylglucosamine to the ‘core’ structure. Additions of galactose cause G1F, G2F, G1bF, and G2bF forms. Glycans having bisecting GlcNAc in the absence of fucose (G0b, G1b, etc.). As it has been mentioned in Chapter 1, the names of the high mannose type glycans signify the number of mannose residues of the glycan. For example, Man5 has five mannose residues in its formula.

To perform our experiments, we complex our glycans with two sodium cations to form doubly positively charged species. Additionally, we added the formic acid to promote generation of the doubly protonated species. Applying our approach, we don’t know where the exact sodium and proton binding sites are, therefore the glycan structures presented in Figure D. 1 do not carry any charges. Nevertheless, it is important to highlight the fact that we are not interested in the three-dimensional structure, but rather in various isomers, and the locations of the charges are irrelevant as long as the data is reproducible.

SNFG notation	Name
	G0
	G0F
	G1
	G1F
	G2F
	G0b
	Man5
	Man6

Figure D. 2 Representative structures of carbohydrates investigated in this thesis and the corresponding symbol nomenclature for glycans (SNFG) that is used throughout this thesis. The most common names of these sugars are mentioned and used in this thesis.

D.3 Creating the targeted database

D.3.1 *The principle of the database construction*

As we mentioned previously, the targeted database includes mass-to-charge ratio, ion mobility ATDs, and reference vibrational spectra. Here, we demonstrate the principle that can be used to build up the targeted database.

The use of the database requires the development of a robust procedure to generate highly reproducible data independent of the different set-ups and ion sources. One of the crucial points is to ensure that one measures the analytes matching the database entries. A basic question while planning any database related experiment is the choice between label-free versus labeled sugars. Here, we would like to address and discuss the benefits and drawbacks of each of those analyte. As we know that glycans do not contain any chromophore/fluorophore suitable for online detection with common methods of liquid chromatography techniques and mass spectrometry. For this reason, glycans are normally labeled with hydrophobic chromophore prior to their analysis to increase their ionization efficiency. However, the derivatization step is time-consuming and the price of the new labeling agents is high.

After reductive amination with a fluorescent label, the reducing end monosaccharide will show an open ring structure. However, native sugars exhibit predominantly closed-ring structures. To address the impact of labeling on the ATDs and IR spectra of glycans, native species and derivatized with RapiFluor-MS (shown in Figure D.3) Man5 ones were analyzed using our set-up.

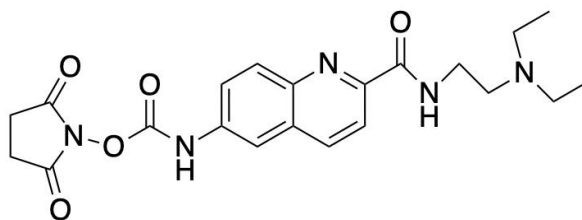


Figure D. 3 *The chemical structure of RapiFluor-MS.*

The ATDs of both molecules are presented in Figure D. 4. The ATD of the labeled Man5 shows one broad peak with a wide shoulder of the small intensity on the left (Figure D. 4 a). The ATD of native Man5 has two main peaks that could

correspond to the major isomers with a slightly different structure (Figure D. 4 b). We observed that ATDs of the native glycans could be more informative since they can adopt multiple isomers, while the relatively big label, such as RapiFluor-MS, can simplify the ATD of Man5 by locking the charge on the tertiary amine in its formula. Additionally, the use of labels can possibly lock one configuration of the molecule.

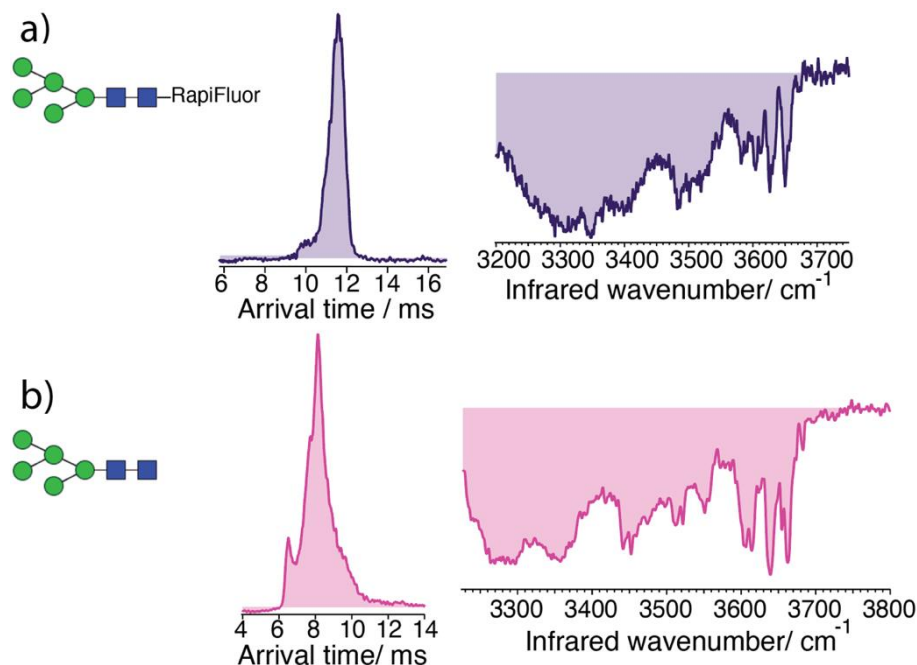


Figure D. 4 a) Arrival time distribution of the doubly protonated Man5-RapiFluor of (m/z 774.3, Th) shown in the center after one cycle (drift length of 1.48 m), the cryogenic spectrum shown on the right; b) Arrival time distribution of the $[Man5+H+K]^{2+}$ of (m/z 637.5, Th) shown in the center after one cycle (drift length of 1.48 m), the cryogenic spectrum shown on the right.

The recorded cryogenic vibrational spectra of Man 5 with no ion mobility selection are shown in Figure D. 5. Each of them was an average of four IR spectra. All the vibrational spectra were well-resolved and exhibit numerous sharp transitions in the free OH stretch region (3580–3700 cm⁻¹) and broad features in the weakly hydrogen-bonded OH stretch region (3450–3550 cm⁻¹). However, the S/N ratio for the native sugar was significantly higher than for the labeled sugar. The additional justification is provided in Figure D. 5. One IR scan that takes ~ 3 mins for both labeled (Figure D. 5 a) and label-free Man 5 (Figure D. 5 b). The tagging efficiency of the label-free Man 5 was higher than of the labeled Man 5. It could be due to the

higher polarizing power of alkali metal that leads to stronger electrostatic interaction with the N2 tag molecule [13]. Figure D. 5 b shows a very resolved spectrum that already can serve as a fingerprint for our technique, whereas the labeled one looks unresolved with a poor S/N and requires more IR scans to obtain a spectrum of descent quality.

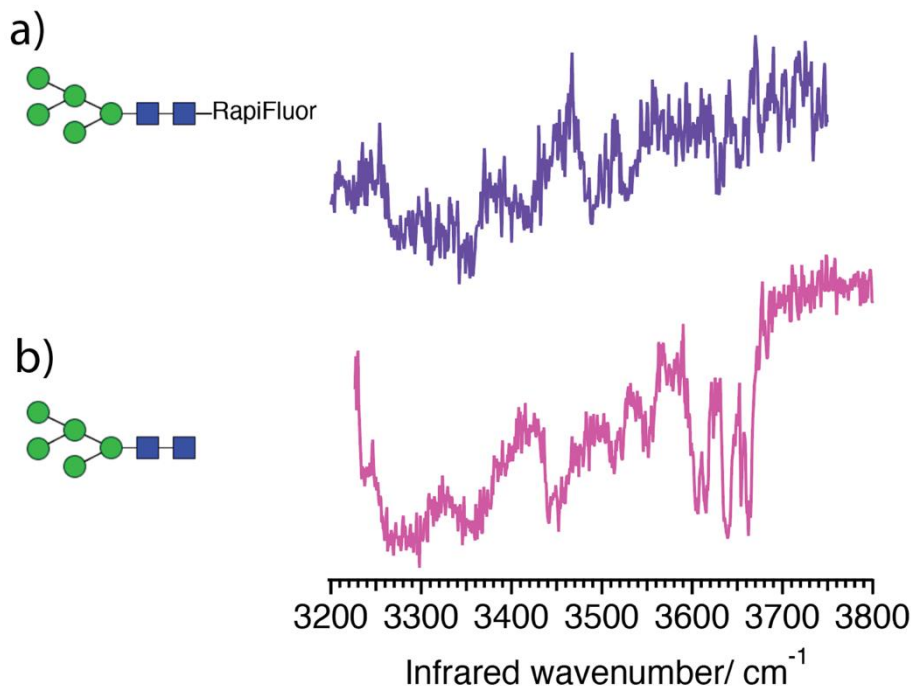


Figure D. 5 One IR scan ~ 3 mins measurement of a) of the doubly protonated Man5-RapiFluor of (m/z 774.3, Th); b) of the $[Man5+H+K]^{2+}$ of (m/z 637.5, Th).

Additionally, we estimated the concentration that we used for our experiments for both samples. The concentration of the labeled Man 5 was 16,6 μM , while the concentration for native Man 5 was 20 μM . The difference in the concentration between samples was minimal, but the additional step in the sample preparation together with new expenses can be a crucial limiting factor to label glycans. For this reason we chose adducts of native glycan for our targeted database, since the benefits of having the analytes labeled were not so critical.

Here, we start from G0F as an example. We studied the doubly protonated and doubly sodiated species of the G0F. ATDs of the doubly protonated G0F standard are presented in Figure D. 6. We observed the three distinct sharp peaks after one cycle on the SLIM board (Figure D.6 a), those three peaks turned into several peaks

after three SLIM cycles (Figure D.6 b). Additionally, we notice extra small peaks that relate to ions of subtly different three-dimensional structures from the major species. We assume that numerous peaks in the obtained ATD could correspond to different isomers, including anomers at the reducing end OH or to several conformers of those isomers. We don't exclude the possibility of having various protonation sites together with different sites to which fucose might migrate [16,17]. For our targeted glycan database of standards, we collected cryogenic vibrational spectra of the most abundant peaks in the ATD of glycan G0F (Figure D.6). The high sensitivity of IR spectroscopy to the smallest structural details is reflected in these spectra. All the vibrational spectra show well-resolved transitions in the free OH stretch region (3580–3700 cm^{-1}) and broad features in the weakly hydrogen-bonded OH stretch region (3450–3550 cm^{-1}), all of which provide a fingerprint that can be used to identify this particular molecule. While these mobility-selected infrared spectra likely represent subtly different molecular conformations, the reproducibility of the ATD's guarantees that they provide a reliable identifying fingerprint.

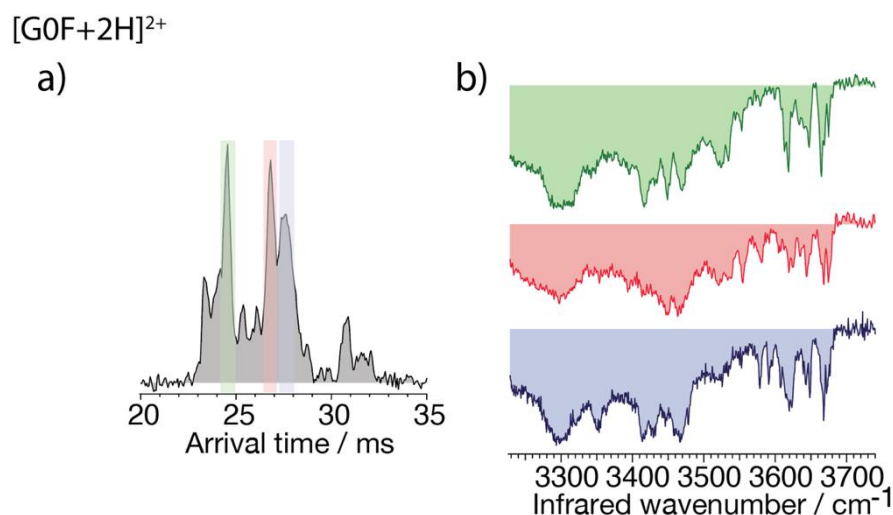


Figure D. 6 a) Arrival time distribution of the doubly protonated G0F of (m/z 732.7, Th) shown on the left after three cycles (drift length of 4.88 m) on the SLIM board; b) The cryogenic IR spectra of each mobility-separated drift peak are shown on the right (green, red, and blue trace for first, second and third peak respectively).

Another example is a doubly sodiated adduct of G0F. ATD of this standard is shown in Figure D. 7. We observed the two distinct sharp peaks after three SLIM cycles (Figure D. 7a). We assume that these two peaks in the obtained ATD could correspond to different isomers. There can be also two sodium binding sites that lock

the conformations together with a possibility of the fucose unit to migrate [17,18]. For the database of standards, we recorded cryogenic vibrational spectra of the two ATD peaks of this glycan (Figure D. 7 b). All the vibrational spectra were distinct and exhibit multiple sharp transitions in the free OH stretch region (3580–3700 cm^{-1}). Additionally, we observed broad bands in the weakly hydrogen-bonded OH stretch region (3450–3550 cm^{-1}). These mobility-selected infrared spectra likely represent different isomers of this molecule.

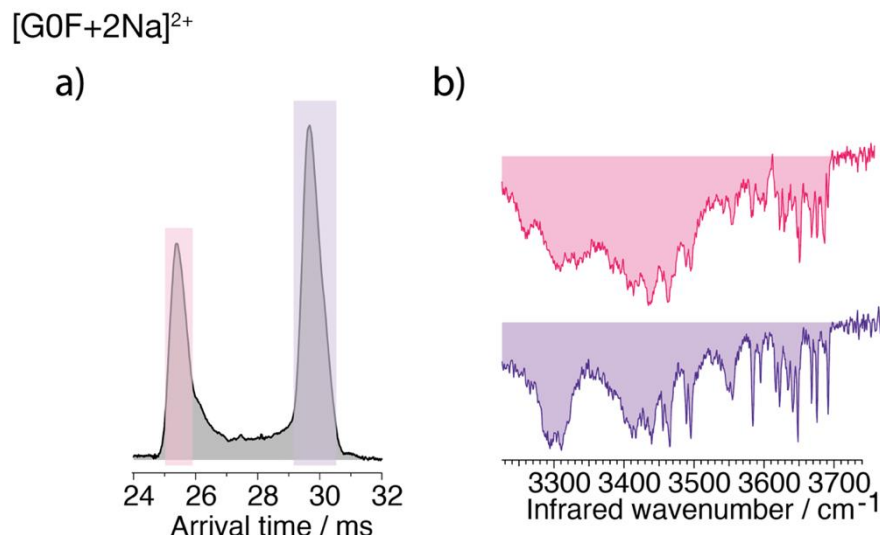


Figure D. 7 a) Arrival time distribution of the doubly protonated G0F of (m/z 754.7, Th) shown on the left after three cycles (drift length of 4.88 m) on the SLIM board; b) The cryogenic IR spectra of each mobility-separated drift peak are shown on the right (pink and violet trace for first and second peak respectively).

Here, we demonstrated the main principle of the targeted database construction. We observed that metal complexation tends to simplify the ATDs of biologically relevant oligosaccharides. However, IR spectra endure the same complexity with sharper transitions for the doubly sodiated species than for the doubly protonated ones. It could be linked to the proton delocalization and the number of OH groups interacting with the metal cation.

D.3.2 Current entries of the targeted N-glycan database

Here, we provided several examples of the data that can be used for the further expansion of the N-glycan targeted database. This data could be a valuable resource for the glycobiology community as it represents a step towards the desired long term

goal of creating a dull database of glycosylation to characterize and categorize glycans for biomedical research.

We recorded the ATD after six cycles on the SLIM together with the IR spectrum with no ion mobility selection for the doubly protonated G0 sugar (Figure D. 8). Since protonated adducts exhibit multiple features in the ATD that could likely correspond to different conformational families of isomers, we omit the taking an IR spectrum of each peak in the ATD for the simplification purpose. However, it might be required if a total spectrum of the studied analyte with no ion mobility selection does not match to a database entry. The ATD of the doubly protonated G0 standard is shown in Figure D. 8 a. We observed the multiple peaks after six cycles on the SLIM board. We assume that numerous peaks in the obtained ATD could correspond to different conformers of two main isomers since the ATD has two main separated peaks. For the database, we recorded an overall spectrum of G0 (Figure D. 8). All the vibrational spectra were distinct and exhibit multiple sharp transitions in the free OH stretch region ($3580\text{--}3700\text{ cm}^{-1}$). Also, we observed broad bands in the weakly hydrogen-bonded OH stretch region ($3450\text{--}3550\text{ cm}^{-1}$). However, this spectrum is very distinct and can be exploited as a fingerprint of this molecule. This spectrum has numerous broad bands that might be due that the protons might form numerous hydrogen bonding with OH and NH groups.

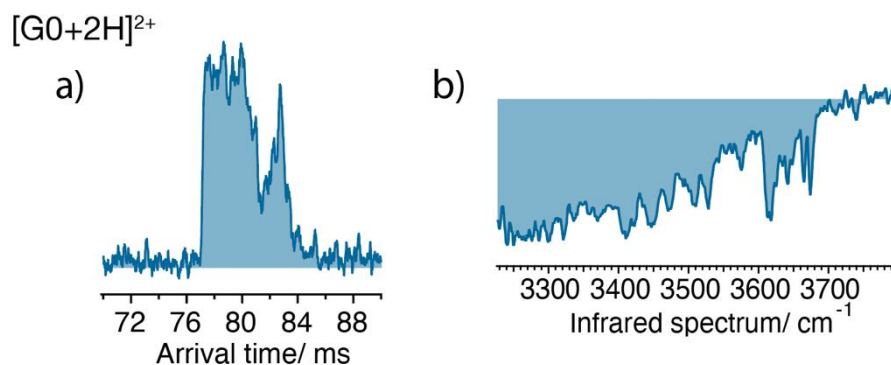


Figure D. 8 a) Arrival time distribution of the doubly protonated G0 of (m/z 659.6, Th) shown on the left, five cycles after enrichment (drift length of 8.62 m) on the SLIM board; b) The cryogenic IR spectrum with no ion mobility selection shown on the right.

Another illustration is the doubly sodiated adduct of the G0. Similarly to the previously discussed G0F glycan, ATD and IR spectra of the G0 differs for various

species since proton and alkali metals bind differently. ATD of the G0 in the doubly sodiated form is presented in Figure D. 9. We see one distinct broad peak after five SLIM cycles (Figure D. 9 a). We assume that this peak could correspond to one main form of this glycan with two binding sodium sites. Moreover, the observation of lower abundance features in ATDs that are estimated to resemble only a single isomer is not uncommon. It could mean that a single isomer has multiple conformations, another isomer might have a different covalent structure of low abundance. We obtained a cryogenic vibrational spectrum with no ion mobility since the ATD has only one major peak (Figure D. 9 b). The vibrational spectrum was clearly distinguishable and showed numerous well-resolved transitions in both the free OH stretch and the weakly hydrogen-bonded OH stretch regions. We think that the sodium cations might bind with some of the glycan hydroxyl groups and shift their vibrational bands outside the frequencies region of the infrared laser, simplifying the IR fingerprint [19]. For that reason, the IR spectra of the doubly sodiated glycans are somewhat sharper than for the doubly protonated forms of the glycans. Another explanation can be related to fewer conformations since the sodium ions lock the conformations making the molecule more rigid.

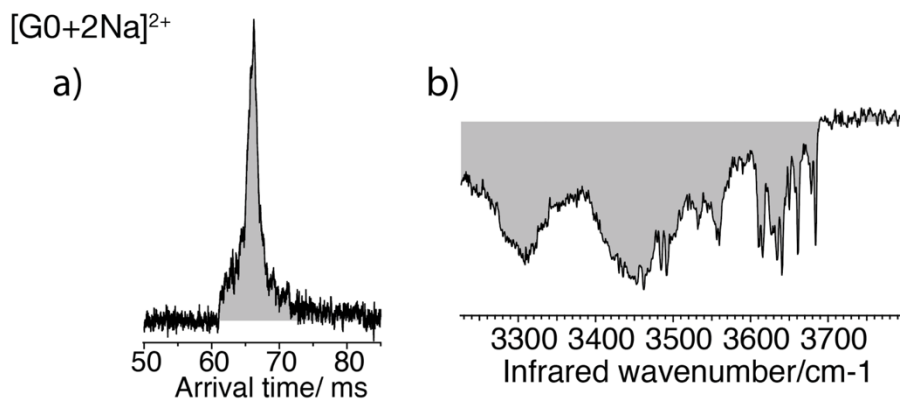


Figure D. 9 a) Arrival time distribution of the doubly sodiated G0 of (m/z 681.6, Th) shown on the left after five cycles (drift length of 7.4 m) on the SLIM board; b) The cryogenic IR spectrum with no ion mobility selection shown on the right.

The example of the G1F listed in our database is particularly interesting example as it exists as a mixture of two positional isomers of the galactose unit at the non-reducing end (Figure D. 10). The detailed studies of this glycan to distinguish these isomers are provided in Chapter 5. Here, we described briefly its characteristic ATDs together with the IR spectra for the doubly protonated and doubly sodiated forms.

The ATD of the doubly protonated form exhibits numerous features after three cycles on the SLIM board that could arise from two positional isomers and their multiple conformers, including fucose migration (Figure D. 10 a). Furthermore, the collected IR spectrum exhibits unique transitions in the free OH stretch region (3580–3700 cm^{-1}) that can be used for identification. However, there are no characteristic features above 3450 cm^{-1} .

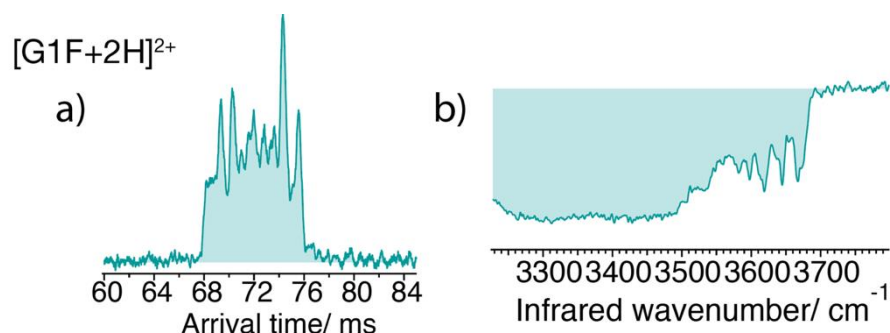


Figure D. 10 a) Arrival time distribution of the doubly protonated G1F of (m/z 813.8, Th) shown on the left, three cycles after enrichment (drift length of 5.27 m) on the SLIM board; b) The cryogenic IR spectrum with no ion mobility selection shown on the right.

The ATD together with the IR fingerprint for the doubly sodiated G1F are presented in Figure D. 11. In contrast to G1F in the doubly protonated form, the ATD of the doubly sodiated G1F has five distinct resolved peaks that differ significantly in their ion mobility (Figure D. 11 a). This observation could possibly suggest globular and elongated forms of the G1F isomers. All the details on the isomers identification are provided in chapter 5. The recorded vibrational spectrum provides highly resolved transitions in the free and in the weakly hydrogen-bonded OH stretch regions (Figure D. 11 b).

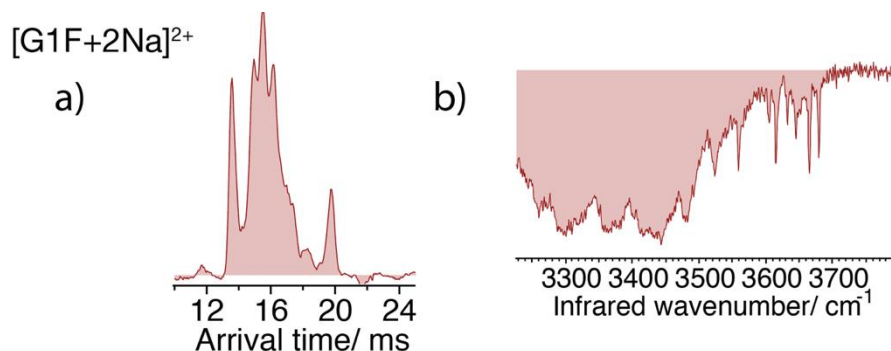


Figure D. 11 a) Arrival time distribution of the doubly sodiated G1F of (m/z 835.8, Th) shown on the left after three cycles (drift length of 4.88 m) on the SLIM board; b) The cryogenic IR spectrum with no ion mobility selection shown on the right.

Another example is G2F in a doubly protonated form, shown in Figure D. 12. ATD demonstrates two main peaks that might correspond to two major conformers (Figure D. 12. a). Some reproducible shoulders that are not fully resolved can correspond to ions with a different structure. Figure D. 12. b shows the spectrum of G2F with unique well-resolved transitions in the free OH stretch region (3580–3700 cm^{-1}). However, it lacks the bands above 3550 cm^{-1} . These observations confirm that the free OH region should be further used as a diagnostic region to expand our database.

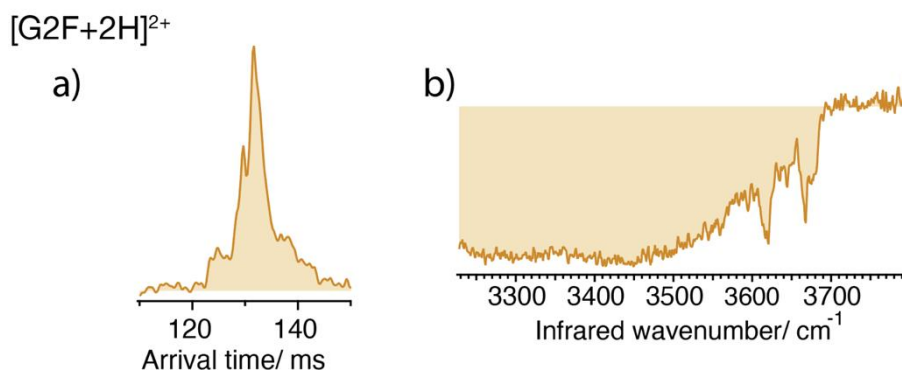


Figure D. 12 a) Arrival time distribution of the doubly protonated G2F of (m/z 894.8, Th) shown on the left, five cycles after enrichment (drift length of 8.62 m) on the SLIM board; b) The cryogenic IR spectrum with no ion mobility selection shown on the right

The ATD and the cryogenic vibrational fingerprint of the doubly sodiated adduct of G2F are presented in Figure D. 13 . The ATD exhibits the main four peaks after three cycles on the SLIM board (Figure D. 13 a). We suppose that these peaks in the ATD could correspond to isomers linked to two sodium sites. Additionally, we include anomers at the reducing end OH or to several conformers of those isomers.

For our targeted glycan database of standards, we collected cryogenic vibrational spectra of the four peaks in the ATD of glycan G2F (Figure D. 13 b), since the m/z of the doubly sodiated molecule is the largest among the glycans. The high sensitivity of IR spectroscopy to the smallest structural details is reflected in these spectra of relatively large molecules. All the vibrational spectra show sharp transitions in the free OH stretch region ($3580\text{--}3700\text{ cm}^{-1}$), all of which provide a fingerprint that can be used to identify this particular molecule. However, broad features in the weakly hydrogen-bonded OH stretch region ($3450\text{--}3550\text{ cm}^{-1}$) did not provide any information. The blue, the green, and the violet traces are quite similar among one another, whereas the red is different from others. This could possibly mean that the red trace corresponds to one of the isomers, while other traces could correspond to different conformers of another isomer of G2F. Even though, the spectral differences among vibrational fingerprints are not significant. Nevertheless, the data of Figure D. 13 b demonstrates the ability of our technique to distinguish between isomeric and/or conformation families, making it a powerful tool for glycan analysis.

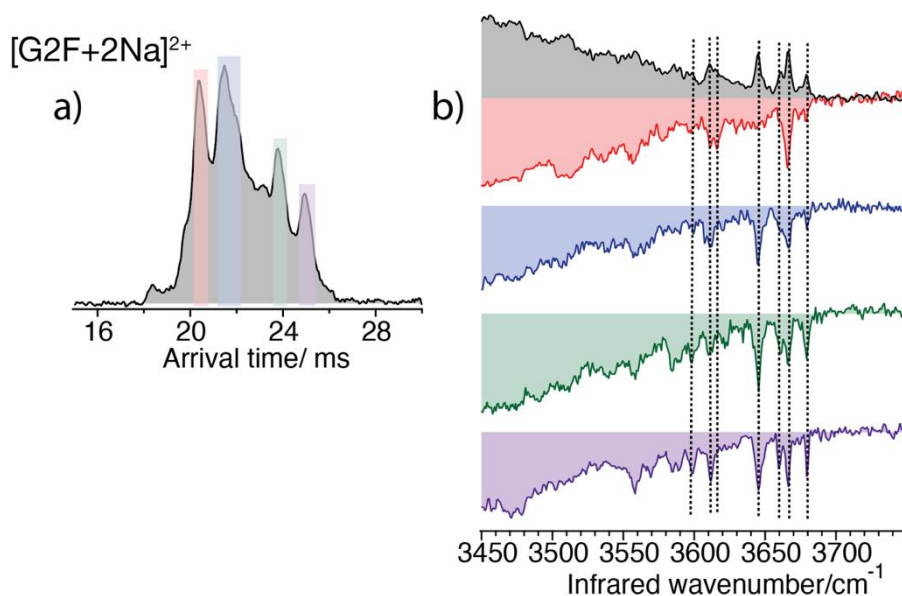


Figure D. 13 a) Arrival time distribution of the doubly sodiated G2F of (m/z 916.8, Th) shown on the left after two cycles (drift length of 3.25 m); b) The cryogenic IR spectra of each mobility-separated drift peak are shown on the right (red, blue, green, and violet trace for first, second, third and fourth peak respectively). The grey traces show the IR spectra of the overall spectra with no ion mobility for comparison. The dash lines represent the differences among the IR spectra.

The following glycans are included in this thesis to demonstrate that ATDs and IR spectra are unique for each sugar. The doubly sodiated adduct of G0b is presented in Figure D. 14. The ATD of this glycan has two distinct peaks after one cycle on the SLIM board (Figure D. 14 a). We assume that these peaks could relate to different isomers of compact globular and rod-like “backfolding” forms. Clemmer group has done studies on doubly sodiated N-glycans and revealed that the backfolding of the α 1-6 arm is usually present in the lowest energy conformers and folding of the α 1-3 arm results in small CCSs [20]. We recorded vibrational fingerprints of each peak in the ATD (Figure D. 14 b). Both spectra exhibit distinct features across the entire scanning region and can be used for identification purposes. The green trace shows less resolved vibrational bands than the yellow one. It could be possibly related to the binding of the OH oscillators with metal ions. We assume that faster ions have a globular shape form due to a more rigid α 1-3 arm. Additionally, this arm has more monosaccharide units that are involved in hydrogen bonding with the sodium ions. Since slower ions have a more elongated shape, it can be linked to the flexible 1-6 linkage with three rotatable torsional angles [21–23].

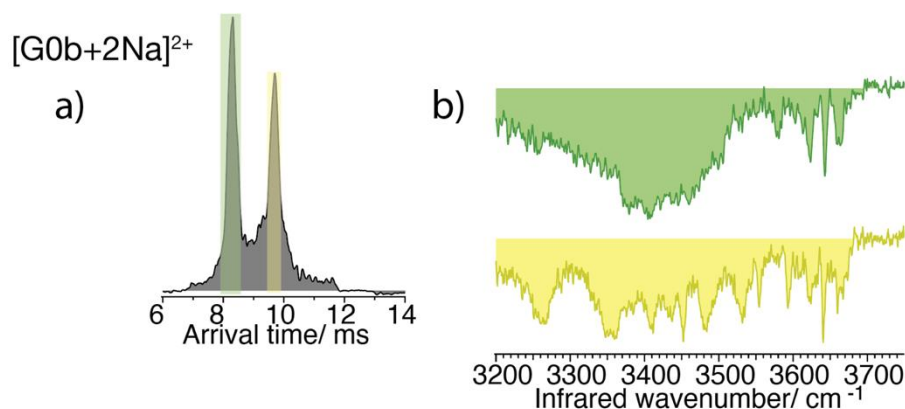


Figure D. 14 a) Arrival time distribution of the doubly sodiated G0b of (m/z 783.2, Th) shown on the left after 1 cycle (drift length of 1.48 m) ; b) The cryogenic

IR spectra of each mobility-separated drift peak are shown on the right (green and yellow trace for first and second peak respectively).

Another example is the doubly sodiated G1 (Figure D.15). Similarly to G1F, it consists of two positional isomers of galactose at non-reducing end. The recorded ATD is presented in Figure D. 16 a. The ATD has five peaks, two of which of the low abundance. All the details will be described in chapter together with G1F. Also, the recorded vibrational spectrum shows highly resolved transitions in the free and in the weakly hydrogen-bonded OH stretch regions (Figure D. 15 b).

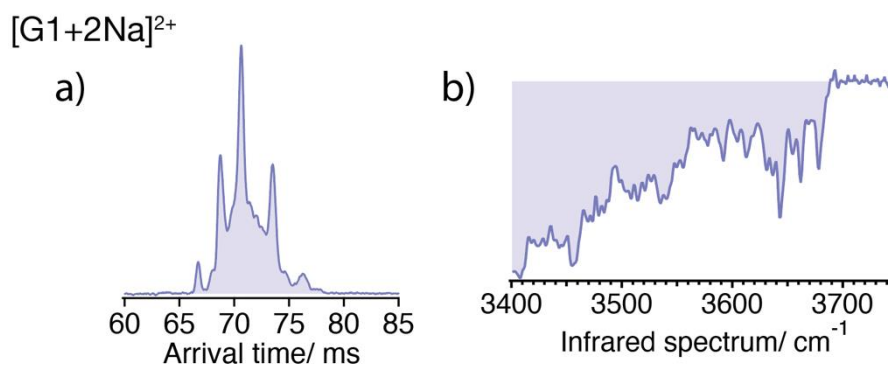


Figure D. 16 a) Arrival time distribution of the doubly sodiated G1 of (m/z 762.7, Th) shown on the left three cycles after enrichment (drift length of 5.27 m) on the SLIM board; b) The cryogenic IR spectrum with no ion mobility selection shown on the right.

And the final example of Man6 in the doubly sodiated form is demonstrated in Figure D. 17. The ATD exhibits one broad peak that could possibly relate to one isomer having multiple conformers (Figure D. 17 a). As it was mentioned previously, the enzymatic release of N-glycans results in the formation of structures with a free reducing end, the presence of anomers of the reducing end N-acetylglucosamine could contribute to that broad peak in the ATD. The cryogenic IR spectrum resembles a unique fingerprint with well-resolved transitions that differ for any spectra (Figure D. 17 b).

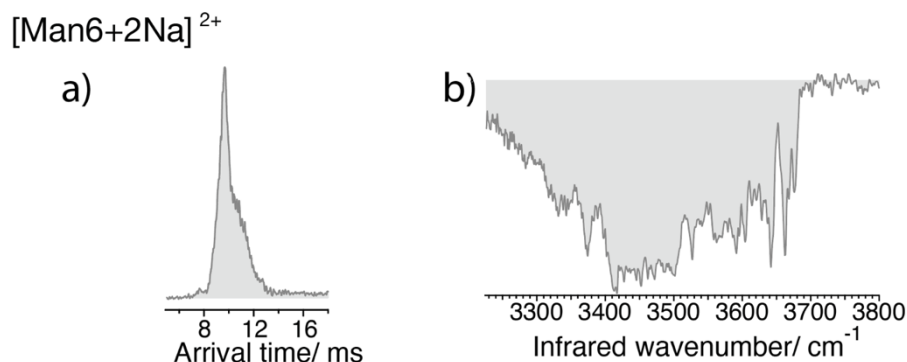


Figure D. 17 a) Arrival time distribution of the doubly sodiated Man6 of (m/z 721.6, Th) shown on the left after 1 cycle (drift length of 1.48 m); b) The cryogenic IR spectrum with no ion mobility selection shown on the right.

D.4 Conclusions and future steps

Our data demonstrated the utility of our database approach of standards. The collected unique cryogenic fingerprints and the ATDs can unambiguously identify the molecule with all its isomers.

The collected ATDs can serve as fingerprints themselves since the ultra-high resolution ion mobility allows them to record very detailed information on the possible structure. However, one dimension is not enough for the unambiguous identification of the molecules. For this reason, we added the dimension of spectroscopy, since even a single flip in the orientation of one stereogenic carbon in the glycan structure may lead to the distortion of the entire network of weakly coupled oscillators. Subsequently, it may result in a completely different spectrum, since IR spectroscopy is exquisitely sensitive to any subtle changes. We demonstrated the power of our technique with various examples.

Our ultimate goal is to create a database of mass, mobility, and vibrational fingerprint spectra for each existing N-glycan in a similar way that we demonstrated in this chapter. For glycans for which we do not have standards, we can use enzymatic strategies [6], as well as CID techniques on the SLIM-IMS platform [10], to determine parent glycan structures and include them in the database.

References:

- (1) Masellis C., Khanal N., Kamrath M.Z., Clemmer D.E., C.; Rizzo T.R. Cryogenic Vibrational Spectroscopy Provides Unique Fingerprints for Glycan Identification. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (10), 2217–2222. <https://doi.org/10.1007/s13361-017-1728-6>.
- (2) Khanal N., Masellis C., Kamrath M.Z., Clemmer D.E., Glycosaminoglycan Analysis by Cryogenic Messenger-Tagging IR Spectroscopy Combined with IMS-MS. *Anal. Chem.* **2017**, *89* (14), 7601–7606. <https://doi.org/10.1021/acs.analchem.7b01467>.
- (3) Khanal N., Masellis C., Kamrath M.Z., Rizzo T.R.,; Clemmer D.E. Cryogenic IR Spectroscopy Combined with Ion Mobility Spectrometry for the Analysis of Human Milk Oligosaccharides. *Analyst* **2018**, *143*, 1846–1852. <https://doi.org/10.1039/C8AN00230D>.
- (4) Ben Faleh, A.; Warnke, S.; Rizzo, T. R. Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures. *Anal. Chem.* **2019**, *91* (7), 4876–4882. <https://doi.org/10.1021/acs.analchem.9b00659>.
- (5) Warnke, S.; Ben Faleh, A.; Scutelnic, V.; Rizzo, T. R. Separation and Identification of Glycan Anomers Using Ultrahigh-Resolution Ion-Mobility Spectrometry and Cryogenic Ion Spectroscopy. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (11), 2204–2211. <https://doi.org/10.1021/jasms.8b06266>.
- (6) Dyukova, I.; Carrascosa, E.; Pellegrinelli, R. P.; Rizzo, T. R. Combining Cryogenic Infrared Spectroscopy with Selective Enzymatic Cleavage for Determining Glycan Primary Structure. *Anal. Chem.* **2020**, *92* (2), 1658–1662. <https://doi.org/10.1021/acs.analchem.9b04776>.
- (7) Campbell, M. P.; Royle, L.; Radcliffe, C. M.; Dwek, R. A.; Rudd, P. M. GlycoBase and AutoGU: Tools for HPLC-Based Glycan Analysis. *Bioinformatics* **2008**, *24* (9), 1214–1216. <https://doi.org/10.1093/bioinformatics/btn090>.

- (8) UniCarb-DB <https://unicarb-db.expasy.org/> (accessed Jun 20, 2020).
- (9) Cooper, C. A.; Harrison, M. J.; Wilkins, M. R.; Packer, N. H. GlycoSuiteDB: A New Curated Relational Database of Glycoprotein Glycan Structures and Their Biological Sources. *Nucleic Acids Res.* **2001**, *29* (1), 332–335.
- (10) Bansal, P.; Yatsyna, V.; AbiKhodr, A. H.; Warnke, S.; Ben Faleh, A.; Yalovenko, N.; Wysocki, V. H.; Rizzo, T. R. Using SLIM-Based IMS-IMS Together with Cryogenic Infrared Spectroscopy for Glycan Analysis. *Anal. Chem.* **2020**. <https://doi.org/10.1021/acs.analchem.0c01265>.
- (11) Hong, P.; Sun, H.; Sha, L.; Pu, Y.; Khatri, K.; Yu, X.; Tang, Y.; Lin, C. GlycoDeNovo – an Efficient Algorithm for Accurate de Novo Glycan Topology Reconstruction from Tandem Mass Spectra. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (11), 2288–2301. <https://doi.org/10.1007/s13361-017-1760-6>.
- (12) Sun, S.; Huang, C.; Wang, Y.; Liu, Y.; Zhang, J.; Zhou, J.; Gao, F.; Yang, F.; Chen, R.; Mulloy, B.; Chai, W.; Li, Y.; Bu, D. Toward Automated Identification of Glycan Branching Patterns Using Multistage Mass Spectrometry with Intelligent Precursor Selection. *Anal. Chem.* **2018**, *90* (24), 14412–14422. <https://doi.org/10.1021/acs.analchem.8b03967>.
- (13) Striegel, A. M.; Piotrowiak, P.; Boué, S. M.; Cole, R. B. Polarizability and Inductive Effect Contributions to Solvent–Cation Binding Observed in Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **1999**, *10* (3), 254–260. [https://doi.org/10.1016/S1044-0305\(98\)00140-8](https://doi.org/10.1016/S1044-0305(98)00140-8).
- (14) Stearns, J. A.; Seaiby, C.; Boyarkin, O. V.; Rizzo, T. R. Spectroscopy and Conformational Preferences of Gas-Phase Helices. *Phys Chem Chem Phys* **2009**, *11* (1), 125–132. <https://doi.org/10.1039/B814143F>.
- (15) Rizzo, T. R.; Stearns, J. A.; Boyarkin, O. V. Spectroscopic Studies of Cold, Gas-Phase Biomolecular Ions. *Int. Rev. Phys. Chem.* **2009**, *28* (3), 481–515. <https://doi.org/10.1080/01442350903069931>.
- (16) Masson, A.; Kamrath, M. Z.; Perez, M. A. S.; Glover, M. S.; Rothlisberger, U.; Clemmer, D. E.; Rizzo, T. R. Infrared Spectroscopy of Mobility-Selected H⁺-Gly-Pro-Gly-Gly (GPGG). *J. Am. Soc. Mass Spectrom.* **2015**, *26* (9), 1444–1454. <https://doi.org/10.1007/s13361-015-1172-4>.

- (17) Mucha, E.; Lettow, M.; Marianski, M.; Thomas, D. A.; Struwe, W. B.; Harvey, D. J.; Meijer, G.; Seeberger, P. H.; von Helden, G.; Pagel, K. Fucose Migration in Intact Protonated Glycan Ions: A Universal Phenomenon in Mass Spectrometry. *Angew. Chem. Int. Ed.* **2018**, *57* (25), 7440–7443. <https://doi.org/10.1002/anie.201801418>.
- (18) Lettow, M.; Mucha, E.; Manz, C.; Thomas, D. A.; Marianski, M.; Meijer, G.; von Helden, G.; Pagel, K. The Role of the Mobile Proton in Fucose Migration. *Anal. Bioanal. Chem.* **2019**, *411* (19), 4637–4645. <https://doi.org/10.1007/s00216-019-01657-w>.
- (19) Masellis, C. Cryogenic Ion Spectroscopy of Peptides and Glycans, EPFL, 2017. <https://doi.org/10.5075/epfl-thesis-8086>.
- (20) Plasencia, M. D.; Isailovic, D.; Merenbloom, S. I.; Mechref, Y.; Clemmer, D. E. Resolving and Assigning N-Linked Glycan Structural Isomers from Ovalbumin by IMS-MS. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (11), 1706–1715. <https://doi.org/10.1016/j.jasms.2008.07.020>.
- (21) U, O.; E, S.; R, S.; G, W. Conformational Flexibility and Dynamics of Two (1→6)-Linked Disaccharides Related to an Oligosaccharide Epitope Expressed on Malignant Tumour Cells. *Chem. Weinh. Bergstr. Ger.* **2009**, *15* (35), 8886–8894. <https://doi.org/10.1002/chem.200900507>.
- (22) Wooten, E. W.; Bazzo, R.; Edge, C. J.; Zamze, S.; Dwek, R. A.; Rademacher, T. W. Primary Sequence Dependence of Conformation in Oligomannose Oligosaccharides. *Eur. Biophys. J.* **1990**, *18* (3), 139. <https://doi.org/10.1007/BF02427373>.
- (23) Jo, S.; Qi, Y.; Im, W. Preferred Conformations of N-Glycan Core Pentasaccharide in Solution and in Glycoproteins. *Glycobiology* **2016**, *26* (1), 19–29. <https://doi.org/10.1093/glycob/cwv083>.

Acknowledgments

The last four years in EPFL have changed my life. During my Ph.D., I learned many things and appreciated all the incredible opportunities that have been given to me. I had a chance to meet many brilliant people from places I had not heard of before, to travel, share the research I have been working on, and more. This thesis would not have been possible without the guidance and the help of people who through the years not only had an influence on my work but also supported me in many ways in this great adventure. It is my pleasure to convey gratitude to all of them in this humble acknowledgment.

First and foremost, I would like to thank my supervisor, Prof. Thomas Rizzo, for providing all the resources and independence, patience through this journey. His insightful feedback pushed me to sharpen my thinking and brought my work to a higher level.

I would like to express my deepest appreciation to my Ph.D. exam committee: Prof. Girault Hubert, Dr. Isabelle Compagnon, Prof. Sabine Flitsch, and Dr. Daniel Ortiz for accepting the invitation and evaluating this thesis.

I am grateful to all the former and current members of the LCPM bio-group. A special thanks to Dr. Valeriu Scutelnic for teaching me how to operate and fix the “second” machine, for always finding time to discuss any results, answer any question, or even to ‘borrow’ his teabags. His incredible enthusiasm for everything he was doing made the days in the lab very special. Thanks as well to Dr. Chiara Masellis for being a “motivator” of the group and bringing some unique atmosphere along with the discussions to the lab. I also thank Dr. Neelam Khanal, the first person whom I met in Switzerland and neighbor by coincidence, for becoming and remaining a friend through time and distance. ‘The family number 9’ experience and the late Greyhound bus will never be forgotten. Particular ‘tanks’ go to Robert Pellegrinelli for the help in the lab, great taste in music, and being the best traveling to the conferences buddy who blindly trusts my navigation skills and walks 3 km to the Chick-fil-A. I also would like to acknowledge Dr. Stephan Warnke and Ahmed ben Faleh for introducing me to the “third” machine and their substantial contributions over the years. I also thank

Irina Dyukova for helping expedite my thesis and for her thoughtful insights. Particular and special thanks go to Dr. Vasyl Yatsyna for being always available by zoom to discuss any topic and for great help with the experiments and giving constructive comments. I would also like to thank Ali K. AbiKhodr “the enricher” for great spirits, Dr. Lei Yue for the fruitful discussions, Dr. Eduardo Carrascosa for assistance, and Pablo for sometimes showing up during the longest weekends and the darkest nights. A big thanks go to Priyanka Bansal whom I am honored to call my friend for all the support and great moments we shared and very delicious Indian food.

I also would like to acknowledge all the members of GGSD, GR-OST, and LND groups for the round tables, group meetings, fun ski weekends, “secret Santa”, Ph.D. defenses, coffee breaks. I want to thank Andreas Osterwalder for hosting great dinners and helping to rescue my friend from the Swiss Alps.

The deepest and warmest appreciation goes to former and present secretaries. Marie-Christine Lehman, I will never forget how you met me with a sign at the airport on Saturday and even bought me some food since nothing works on Sunday in Switzerland. This gesture was one of the kindest. And, of course, Angeles Alarcon for your endless kindness and help with everything from the permits to the bills. And a special thank for your smile that highlights grey days.

I would like to thank Dr. Marcel Drabbles, Dr. Christophe Roussel, Stephane Thonney for having unique experience teaching your courses and nice coffee breaks.

I am particularly grateful for the assistance given by Dr. Jonathan Paz-Montoya. This thesis would not be possible without using their lab facilities in the most crucial times.

I would like to thank the mechanical, the electronical workshops, and the IT team for enormous technical support.

A very special thanks goes to Bruno Credidio, Dr. Tanja Thomsen for many memorable events and constant support. Thanks to Silvia and Oliviero for hosting great Italian dinners with so many yummy courses. A great thanks to Ana, Andrew, Bo-Jung and Harmina for our lunches and coffee breaks.

Thanks to all the people from the societies “ShARE EPFL” and “YoungSCS” I was a part of. Special thanks to Aleksandr Belushkin, Dmitry Vasiliev. It was a great

pleasure to experience many challenges and meet new inspiring people, including the Former Prime ministers of the UK and Ethiopia.

Also I would like to thank all great Russian-speaking people who I met at a bright defense of Dr. Lucy Voronina. It is hard to mention all of them since they are too many. But thanks for all the parties and events, such as Birthdays, Halloweens, BBQs, PhD defenses, Paleo concerts, and many more which I did not attend. A great thanks to Anna Magi for her meme collection for each period in life. In addition, I would like to thank Boris and Vera for our fun outdoor activities.

An enormous thanks go to my good-hearted friends: Ekaterina Belousova, Anna Lukianova, Nadezhda Chupilina, Tatiana Rodina, Eugenia Shabalina, Max Norkin, Eugenii Savushkin (Artur Sobakin) and “the friendship never ends” crew. We are currently quite spread across the globe and do not manage to see each other often, but every time we meet in person, it feels like we were hardly ever separated.

And there is one very special person to me, Dr. Phil. "Cuter" Sean D.S. Gordon. No words can describe how blessed and grateful I am to have you by my side in all the moments of happiness and sadness. Special thanks to him for the morning tea, reading through my thesis, and believing in me. I would also like to extend my deepest gratitude to Gordon's family: Christine and Savvy for their wise counsel and sympathetic ear.

Наконец, я хочу поблагодарить свою семью и родственников. Особенно, свою маму Валентину и бабушку Аллу, кто бескрайне поддерживают меня всю мою жизнь. Моя любовь и уважение к Вам не имеют границ. Спасибо Вам большое за Вас и что я могу всегда на Вас положиться. Я уверена, что мой отец гордился бы мной тоже.

And if you are reading these lines, thanks to you, reader for your genuine interest in this work.

Curriculum vitae

Education

2016 – 2021: PhD student at EPFL, LCPM, PhD advisor: Prof. Thomas Rizzo

2015 – 2016: École Polytechnique, Paris, France, Master's Degree in Molecular Chemistry, Science and Engineering.

2010 – 2015: Lomonosov Moscow State University, Chemistry Specialist Degree with Honours (chemistry department, Laboratory of Analytical Chemistry), GPA: 4.98/5.00

Teaching activities

Spring 2017 – practical course in Physical chemistry for Bachelor students

Autumn 2017 – practical course in Analytical chemistry for Bachelor students

Spring 2018 – practical course in Physical chemistry for Bachelor students

Autumn 2018 – practical course in Analytical chemistry for Bachelor students

Autumn 2019 – substituted some exercise sessions in Mass-spectrometry for Master students

Autumn 2019 – held few practical sessions for the Master's project

Scientific publications

1) Stephan Warnke, Ahmed ben Faleh, Robert P. Pellegrinelli, **Natalia Yalovenko**, and Thomas R. Rizzo, Combining ultra-high resolution ion mobility spectrometry with cryogenic IR spectroscopy for the study of biomolecular ions, Faraday Discuss., 2019,217, 114-125.

2) Priyanka Bansal, Vasyl Yatsyna, Ali H. AbiKhodr, Stephan Warnke, Ahmed Ben Faleh, **Natalia Yalovenko**, Vicki H. Wysocki, and Thomas R. Rizzo, Using

SLIM-Based IMS-IMS Together with Cryogenic Infrared Spectroscopy for Glycan Analysis, *Analytical Chemistry*, 2020, 92 (13), 9079-9085.

3) **Natalia Yalovenko**, Vasyl Yatsyna, Priyanka Bansal, Ali H. AbiKhodr, and Thomas R. Rizzo. Analyzing glycans cleaved from a biotherapeutic protein using ultrahigh-resolution ion mobility spectrometry together with cryogenic ion spectroscopy, *Analyst*, 2020,145, 6493-6499.

4) Irina Dyukova*, **Natalia Yalovenko***, Vasil Yatsyna, Priyanka Bansal and Thomas R. Rizzo. Using high-resolution ion mobility spectrometry with cryogenic infrared spectroscopy for identifying glycan isomers cleaved from monoclonal antibodies, in preparation.