

# Combining ion mobility spectrometry and cryogenic ion spectroscopy with enzymatic digestion/synthesis for the study of N-linked glycans

Présentée le 27 janvier 2023

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

pour l'obtention du grade de Docteur ès Sciences

par

**Irina MÄRKI**

Acceptée sur proposition du jury

Prof. J.-E. Moser, président du jury  
Prof. T. Rizzo, directeur de thèse  
Prof. A. Rijs, rapporteuse  
Prof. B. Clowers, rapporteur  
Prof. S. Gerber, rapporteuse

# Abstract

Glycans play an essential role in numerous physiological and pathological processes of living organisms. Despite their significant biological relevance, glycobiology remains one of the least explored fields of biochemistry. The intrinsic isomeric complexity of glycans poses a great analytical challenge that limits the study of their functions. For this reason, increasing efforts have been directed toward developing new methods that provide a fast, sensitive and accurate analysis of glycans. In this work, we demonstrate a new multidimensional approach for glycan identification based on their unique infrared fingerprints. Our experiments rely on state-of-the-art technology that combines cryogenic messenger-tagging infrared (IR) spectroscopy and ultra-high resolution ion mobility spectrometry (IMS). Adding a spectroscopic dimension to the current databases allows differentiating even subtle structural details between glycan isomers.

The first part of this thesis describes a spectroscopic database approach, in particular, the mechanism for assigning IR spectra to structures that are not contained in our database. The combined implementation of selective enzymatic digestion and collision-induced dissociation (CID) with cryogenic IR spectroscopy provides an unambiguous identification of the primary structure of parent glycan molecules. Each time we add a new species to the database, we can identify it from a mixture based on its spectral fingerprint and obtain a new core structure that can serve to identify larger glycans.

In the next part, we demonstrate the combination of ultra-high resolution ion mobility with IR spectroscopy for the identification of isomeric glycans. We used a selective chemoenzymatic approach to synthesize a pure glycan standard that could not be obtained commercially and then used its IR fingerprint spectra to assign mobility-separated positional isomers. Using these results, we then investigated the impact of the host cell line on the glycan profile of a monoclonal antibody (mAb) at the isomer level. We demonstrated that our technique can monitor glycosylation patterns in mAbs and can be used to complement, or even replace, existing methods for establishing the similarity of glycan profiles between biological drugs and their biosimilars.

In the last part, we apply our IMS-IR approach to study sialylated N-linked glycans. One part of this work is focused on searching for new glycan biomarkers in blood serum samples obtained from female donors with different stages of breast cancer. Exploiting the differences in glycosylation between healthy and sick individuals, we identified a potential glycan biomarker candidate. Moreover, we showed that adding a CID dimension to IMS-IR provides a way for isomer identification without the need for glycan standards. In the second part of this work, we described the relative quantification of sialylated glycans released from follicle-stimulating hormone (FSH) using ion mobility spectrometry. This method allows a fast and reliable analysis of the content of glycans in their native form directly from a complex mixture.

**Keywords:** cryogenic ion spectroscopy, ion mobility spectrometry, N-glycosylation, enzymatic cleavage/synthesis, glycan isomers, IR database of glycans, collision-induced dissociation, glycan biomarkers, liquid chromatography.

# Résumé

Les glycanes jouent un rôle essentiel dans de nombreux processus physiologiques et pathologiques des organismes vivants. Malgré leur importance biologique significative, la glycobiologie reste l'un des domaines les moins explorés de la biochimie. La complexité isomérique intrinsèque des glycanes représente un grand défi analytique qui limite l'étude de leurs fonctions. Pour cette raison, des efforts croissants ont été dirigés vers le développement de nouvelles méthodes qui fournissent une analyse des glycanes rapide, sensible et précise. Dans ce travail, nous présentons une nouvelle approche multidimensionnelle pour l'identification des glycanes basée sur leurs empreintes infrarouges uniques. Nos expériences s'appuient sur une technologie de pointe qui combine la spectroscopie infrarouge (IR) cryogénique et la spectrométrie de mobilité ionique (IMS) à ultra-haute résolution. L'ajout d'une dimension spectroscopique aux bases de données actuelles permet de distinguer même des isomères de glycanes qui diffèrent seulement par des détails structurels subtils.

La première partie de cette thèse décrit une approche de base de données spectroscopique, en particulier, le mécanisme d'attribution des spectres IR aux structures qui ne sont pas contenues dans notre base de données. La combinaison de la digestion enzymatique sélective et de la dissociation induite par collision (CID) avec la spectroscopie infrarouge cryogénique fournit une identification sans ambiguïté de la structure primaire des molécules du glycanes parent. Chaque fois que nous ajoutons une nouvelle espèce à la base de données, nous pouvons l'identifier à partir d'un mélange basé sur son empreinte spectrale et obtenir une nouvelle structure de base qui peut servir à identifier des glycanes plus grands.

Dans la partie suivante, nous présentons la combinaison de la mobilité ionique à ultra-haute résolution avec la spectroscopie IR pour l'identification de glycanes isomères. Nous avons utilisé une approche chimio-enzymatique sélective pour synthétiser un échantillon standard de glycanes pur qui ne pouvait pas être obtenu dans le commerce, puis nous avons utilisé ses spectres d'empreintes infrarouges pour identifier des isomères positionnels séparés par leur mobilité. En utilisant ces résultats, nous avons ensuite étudié l'impact de la lignée cellulaire hôte sur le profil glycanes d'un anticorps monoclonal (mAb) au niveau des isomères. Nous avons démontré que notre technique peut surveiller les modes de glycosylation dans les mAbs et peut



être utilisée pour compléter, ou même remplacer, les méthodes existantes pour établir la similitude des profils de glycanes entre les médicaments biologiques et leurs biosimilaires.

Dans la dernière partie, nous appliquons notre approche IMS-IR pour étudier les glycanes sialylés. Une partie de ce travail est axée sur la recherche de nouveaux biomarqueurs glycanes dans des échantillons de sérum sanguin provenant de donneuses présentant différents stades de cancer du sein. En exploitant les différences de glycosylation entre les individus sains et malades, nous avons identifié un candidat potentiel de biomarqueur glycane. De plus, nous avons montré que l'ajout d'une dimension CID à l'IMS-IR fournit un moyen d'identification des isomères sans avoir besoin d'un glycane étalon. Dans la deuxième partie de ce travail, nous avons décrit la quantification relative des glycanes sialylés libérés par l'hormone folliculo-stimulante (FSH) à l'aide de la spectrométrie de mobilité ionique. Cette méthode permet une analyse rapide et fiable de la teneur en glycanes sous leur forme native directement à partir d'un mélange complexe.

**Mots clés :** spectroscopie ionique cryogénique, spectrométrie de mobilité ionique, N-glycosylation, clivage/synthèse enzymatique, isomères de glycanes, base de données IR de glycanes, dissociation induite par collision, biomarqueurs glycanes, chromatographie liquide.

# Table of contents

Abstract .....	i
Résumé .....	iii
Table of contents .....	v
List of abbreviations .....	ix
<b>Chapter 1. Introduction</b> .....	11
1.1. The importance of glycan analysis .....	11
1.2. Challenges in glycan analysis .....	13
1.3. Glycan nomenclature .....	16
1.4. Synthesis of N-linked glycans in eukaryotes .....	17
1.5. Traditional methods for glycan analysis .....	19
1.5.1. Mass spectrometry .....	19
1.5.2. Liquid chromatography.....	20
1.5.3. Capillary electrophoresis .....	21
1.5.4. Nuclear magnetic resonance spectroscopy .....	22
1.6. Ion mobility separation .....	23
1.6.1. Drift tube ion mobility spectrometry (DTIMS) .....	24
1.6.2. Traveling wave ion mobility spectrometry (TWIMS).....	25
1.6.3. Trapped ion mobility spectrometry (TIMS) .....	26
1.6.4. Field asymmetric waveform ion mobility spectrometry (FAIMS).....	26
1.7. Gas-phase infrared spectroscopy of ions .....	27
1.7.1. IR multiple photon dissociation (IRMPD).....	28
1.7.2. Cold ion spectroscopy .....	29
1.8. Multidimensional techniques .....	31
1.9. Our approach for glycan analysis .....	32
1.10. Outline of the thesis .....	33
References.....	35
<b>Chapter 2. Experimental approach</b> .....	45
2.1. Experimental setup.....	46
2.1.1. Generation I: the cold ion spectroscopy (CIS) instrument.....	46
2.1.2. Generation II: ion mobility spectrometry coupled to cold ion spectroscopy (IMS-CIS I instrument).....	47

2.1.3. Generation III: ion mobility spectrometry coupled to cold ion spectroscopy version II (IMS-CIS II instrument) .....	48
2.2. SLIM-based ion mobility .....	49
2.2.1. Basic principles of the SLIM platform .....	49
2.2.2. The ion mobility module of the IMS-CIS I instrument. ....	51
2.2.3. The ion mobility module of the IMS-CIS II instrument.....	52
2.3. Cryogenic messenger-tagging spectroscopy.....	54
2.4. Ultra-performance liquid chromatography (UPLC) coupled to a Q-TOF mass spectrometer.....	57
References.....	58

### **Chapter 3. Combining enzymatic cleavage and collision-induced dissociation with cryogenic IR spectroscopy for determining the primary structure of glycans.....**

3.1. Database approaches for the identification of glycans .....	61
3.2. Approaches for the identification of unknown glycans .....	64
3.2.1. Selective enzymatic digestion.....	64
3.2.2. Collision-induced dissociation.....	66
3.3. Combining cryogenic IR with selective enzymatic cleavage for determining the primary structure of glycans .....	68
3.3.1. Exoglycosidase digestion of N-linked glycans.....	68
3.3.2. Infrared spectroscopy of glycan standards and smaller glycans after selective enzymatic cleavage .....	70
3.4. Combination of CID and enzymatic cleavage for the identification of isomeric glycan fragments.....	75
Conclusions.....	79
References.....	80

### **Chapter 4. Identification of positional isomers of glycans cleaved from monoclonal antibodies.....**

4.1. Introduction.....	83
4.2. Ion-mobility-selective IR spectroscopy of glycan G1F isomers.....	86
4.3. Selective enzymatic synthesis of the glycan G1( $\alpha$ 1,6)F.....	88
4.4. Identification of the positional isomers of the G1F based on the ion-mobility-selective IR fingerprints.....	92
4.5. Impact of the host cell line on the glycan profile.....	94

4.6. IR spectral comparison using PCA analysis .....	97
Conclusions.....	99
References.....	100
<b>Chapter 5. Searching for new N-glycan biomarkers .....</b>	<b>107</b>
5.1. Motivation and objectives.....	107
5.2. Sialylated N-glycans from human blood serum.....	109
5.3. Identification of glycans released from the serum samples using the spectroscopic database.....	110
5.4. Comparison of arrival time distributions of glycans collected from healthy and sick donors.....	113
5.5. Scheme for isomer identification of sialylated N-glycans.....	116
Conclusions.....	121
References.....	122
<b>Chapter 6. Relative quantification of sialylated N-glycans.....</b>	<b>125</b>
6.1. Introduction.....	125
6.2. Releasing sialylated N-glycans from r-hFSH .....	126
6.3. Ion mobility for relative quantification of glycans .....	127
6.3.1. Relative quantification with nESI needles.....	128
6.3.2. Relative quantification with ESI source .....	129
Conclusions.....	132
References.....	133
<b>Chapter 7. Summary and future perspectives.....</b>	<b>135</b>
Appendix I .....	139
Acknowledgements.....	141
Curriculum Vitae .....	143



# List of abbreviations

PMT: post-translational modification  
SNFG: symbol nomenclature for glycans  
UOXF: Oxford nomenclature for glycan notation  
MS: mass spectrometry  
MS<sup>n</sup>: tandem mass spectrometry  
ESI: electrospray ionization  
MALDI: matrix-assisted laser desorption/ionization  
CID: collision-induced dissociation  
IRMPD: infrared multiphoton dissociation  
ETD: electron transfer dissociation  
ECD: electron capture dissociation  
UVPD: ultraviolet photodissociation  
UPLC: ultra-performance liquid chromatography  
RP-LC: reversed-phase liquid chromatography  
NP-LC: normal-phase liquid chromatography  
HPAEC: high-performance anion-exchange chromatography  
PGC: porous graphitized carbon  
2-AB: 2-aminobenzamide  
HILIC: hydrophilic interaction liquid chromatography  
CE: capillary electrophoresis  
NMR: nuclear magnetic resonance  
IMS: ion mobility separation  
ATD: arrival time distribution  
CCS: collision cross-section  
DTIMS: drift tube ion mobility spectrometry  
TWIMS: traveling wave ion mobility spectrometry  
SLIM: structures for lossless ion manipulations  
TIMS: trapped ion mobility spectrometry  
FAIMS: field asymmetric waveform ion mobility spectrometry  
IR: infrared spectroscopy  
IVR: intramolecular vibrational redistribution  
Nd:YAG: neodymium-doped yttrium aluminum garnet

OPO/OPA: optical parametric oscillator/optical parametric amplifier

CIS: cold ion spectroscopy instrument

RF: radio frequency

DC: direct current

Q-TOF: tandem quadrupole time-of-flight mass spectrometer

nESI: nano-electrospray ionization

IFT: ion funnel trap

PCB: printed circuit board

DB: database

GT: glycosyltransferase

mAb: monoclonal antibody

IgG: immunoglobulin G

ADCC: antibody-dependent cell-mediated cytotoxicity

CDC: complement-dependent cytotoxicity

CHO: Chinese hamster ovary

HEK-293: human embryonic kidney 293

FUT8:  $\alpha$ 1,6-fucosyltransferase

PCA: principal component analysis

DTT: dithiothreitol

SA: sialic acid

BMI: body mass index

FSH: follicle-stimulating hormone

# Chapter 1. Introduction

## 1.1. The importance of glycan analysis

Glycans (also known as saccharides, carbohydrates or simply sugars) are one of the four essential components of all living cells (along with oligonucleotides, proteins and lipids) and represent the most abundant and diverse class of biopolymers.<sup>1</sup> In contrast to oligonucleotides and proteins that have only one basic type of linkage between either nucleic or amino acids, respectively, glycan monomers can be attached to one another by many different linkage positions, leading to an almost unlimited number of structures.<sup>2</sup> Glycobiology is a relatively young field of natural sciences that was first mentioned in the late 1980s.<sup>1</sup> It focuses on the structure, chemistry, biosynthesis, and biological function of glycans and their derivatives. The majority of carbohydrates present in cells are attached to proteins or lipids and form two big groups of glycoconjugates: glycoproteins and glycolipids.

Glycosylation is one of the most common and diverse post-translational modifications (PTMs) that introduces significant complexity to a protein structure.<sup>3</sup> It is estimated that about 70% of human proteins are glycosylated.<sup>4</sup> There are three different types of glycans that can be attached to proteins: N-linked glycans, O-linked glycans and glycosaminoglycans (GAGs). N-linked glycans are attached to the nitrogen atom of the amino group of the asparagine (Asn) residue at the glycosylation motif Asn-X-Ser/Thr of a protein (X is any amino acid except Pro). O-linked glycans are attached to the oxygen atom of the hydroxyl groups of serine (Ser) or threonine (Thr) residues. GAGs are attached to the hydroxyl group of the Ser in the Ser-Gly-X-Gly amino acid motifs of a protein (Figure 1.1).<sup>5</sup>

Glycoproteins are known to be involved in many important biological processes, including cell growth, immune response, inflammation, fertilization, viral replication, parasitic infection and blood hemostasis.<sup>1, 6-9</sup> The glycosylation pattern of a protein has a crucial impact on its intracellular and extracellular functions. Oligosaccharides ensure protein stability, solubility, correct folding, and protease resistance. Their major role is modulating the interactions of glycoconjugates with other molecules, allowing cell-to-cell recognition and signal transduction (Figure 1.2).



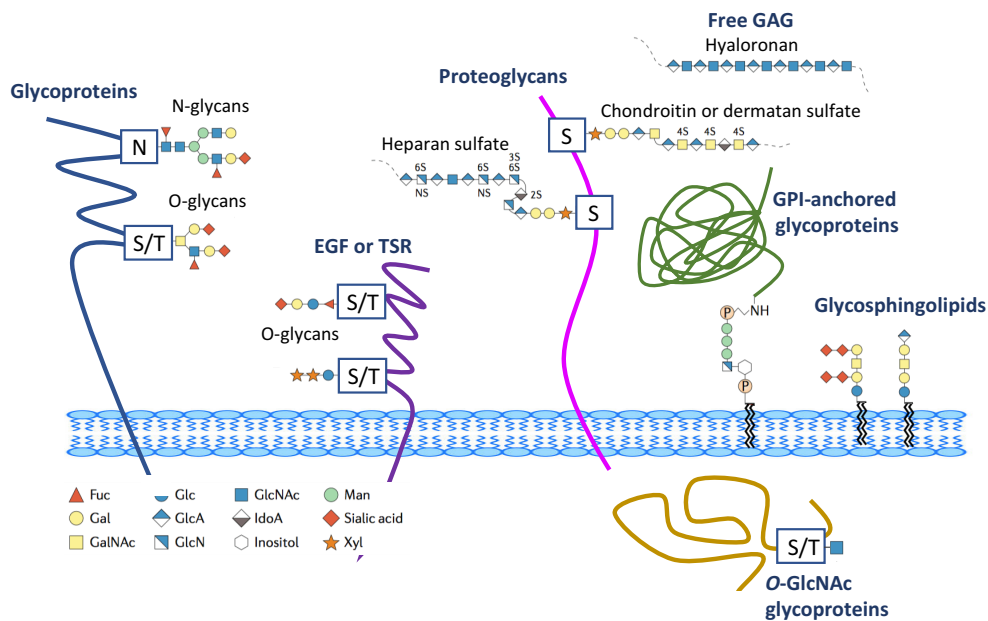


Figure 1.1. Major types of glycosylation in humans (Redrawn from Fig. 1 of Ref. 5).

Specific changes in glycan composition are directly linked to human diseases, many of which are life-threatening. These include diabetes mellitus,<sup>10</sup> Parkinson's,<sup>11</sup> autoimmunity<sup>12</sup> and chronic inflammation,<sup>13</sup> cardiovascular<sup>14</sup> and Alzheimer's<sup>15</sup> diseases. Changes in glycosylation were observed during the progression of liver<sup>16</sup> and kidney diseases.<sup>17, 18</sup> Moreover, glycosylation patterns were one of the first biomarkers of various types of cancer (lungs, breast, prostate).<sup>19-21</sup> The glycan composition changes together with the evolution of cancer cells at different stages. Glycans represent a promising class of non-invasive biomarkers that can reveal diseases with high specificity and sensitivity.<sup>22</sup>

A better understanding of glycan functions and their involvement in disease processes would allow to identify therapeutic targets. The glycosylation pattern of biotherapeutic drugs directly impacts their efficacy and safety.<sup>23,24</sup> Recently, progress in HIV-1 vaccine was enabled by a better understanding of the HIV-1 envelope glycoprotein and the effects of its glycan composition on immune response and evasion.<sup>25</sup> Furthermore, the study of glycans is extremely important for organ xenotransplantation, where the nature of carbohydrate antigens in donors needs to be understood to prevent immune rejection.<sup>26,27</sup>

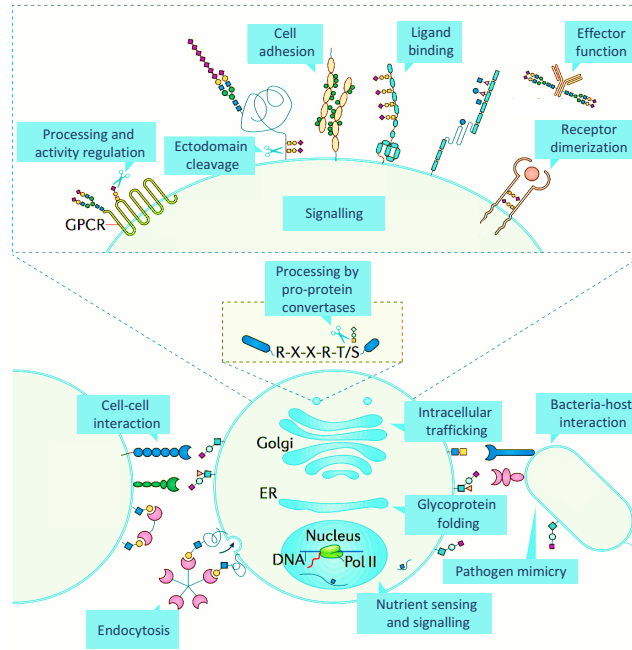


Figure 1.2. Different roles of glycans in regulating protein functions (Adapted from Fig. 4a of Ref. 8).

## 1.2. Challenges in glycan analysis

Oligosaccharides, or glycans, consist of monosaccharide building blocks that have an empirical formula of  $(C_x(H_2O)_n)$   $n=3-9$ . Monosaccharides exist in solution as an equilibrium mixture of acyclic and cyclic forms. The ratio depends on the sugar structure; the acyclic form usually represents less than 0.01% in an equilibrium mixture. The cyclic structure is formed by the reaction of one of the hydroxy groups with the C-1 aldehyde or ketone. The C1 carbon becomes an asymmetric center called the *anomeric carbon*.<sup>2</sup> The hydroxy group linked to this C1 can have two possible orientations axial ( $\alpha$  anomer) or equatorial ( $\beta$  anomer). Depending on the solution conditions, these two stereoisomers can interconvert via the process of *mutarotation* (Figure 1.3).

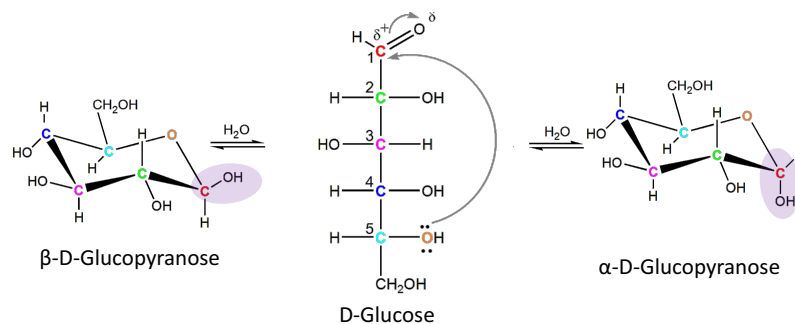


Figure 1.3. Conversion from  $\beta$ -D-glucopyranose to  $\alpha$ -D-glucopyranose through ring opening (Fisher projection) and closing.

The polymerization of monosaccharides occurs by the formation of a glycosidic bond between the anomeric carbon of one monosaccharide and an OH-group of another. The monosaccharide that keeps the anomeric OH-group is the reducing end, while the other represents the non-reducing end. Figure 1.4 illustrates seven of the most abundant monosaccharides found in mammals. It includes glucose (Glu) 2.5%\*, galactose (Gal) 24.8%, mannose (Man) 18.9%, fucose (Fuc) 7.2%, N-acetyl-glucosamine (GlcNAc) 31.8%, N-acetyl-galactosamine (GalNAc) 4.8% and N-acetyl-neuraminic acids (Neu5Ac) 8.3% (\*relative abundances).<sup>24</sup>

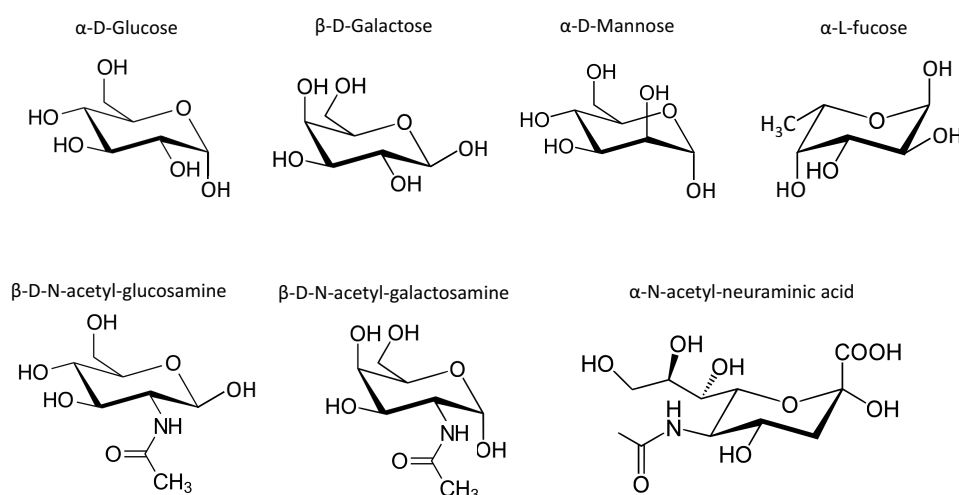


Figure 1.4. Structures of the seven most abundant monosaccharides found in mammalian cells.<sup>24</sup>

The monosaccharide building blocks are often isomeric, displaying subtle structural differences. For example, the stereochemistry of a single asymmetric carbon atom is the only difference between hexose saccharides (Figure 1.5a). Moreover, the diversity of ways in which monosaccharides can be linked to each other dramatically increases the complexity of higher-order structures. The linkage orientation can be  $\alpha$  or  $\beta$  (Figure 1.5b), and two monosaccharides can have different attachment positions, leading to different regioisomers (Figure 1.5c). In addition, monosaccharide units can form linear chains as well as branched structures with the same monosaccharide content (Figure 1.5d). These factors come from the intrinsic chemical nature of glycans and allow a vast number of possible combinations of monosaccharides. Taking into account all types of isomerism present in glycans, it was calculated that the number of possible structures for reducing hexasaccharides comprised of D hexoses alone could be more than  $10^{12}$ .<sup>28</sup>

From a biological standpoint, though the known biosynthetic pathways limit the variety of structures of physiologically relevant carbohydrates, it is not a template-driven process, unlike nucleic acids and proteins. Over 250 enzymes participate in glycan synthesis, and the resulting

glycan structures are determined by their concentration, level of activity, expression and localization, as well as the substrate/donor specificity and availability.<sup>29</sup> All these factors contribute to functionally significant glycan heterogeneity, which cannot be predicted *a priori*.

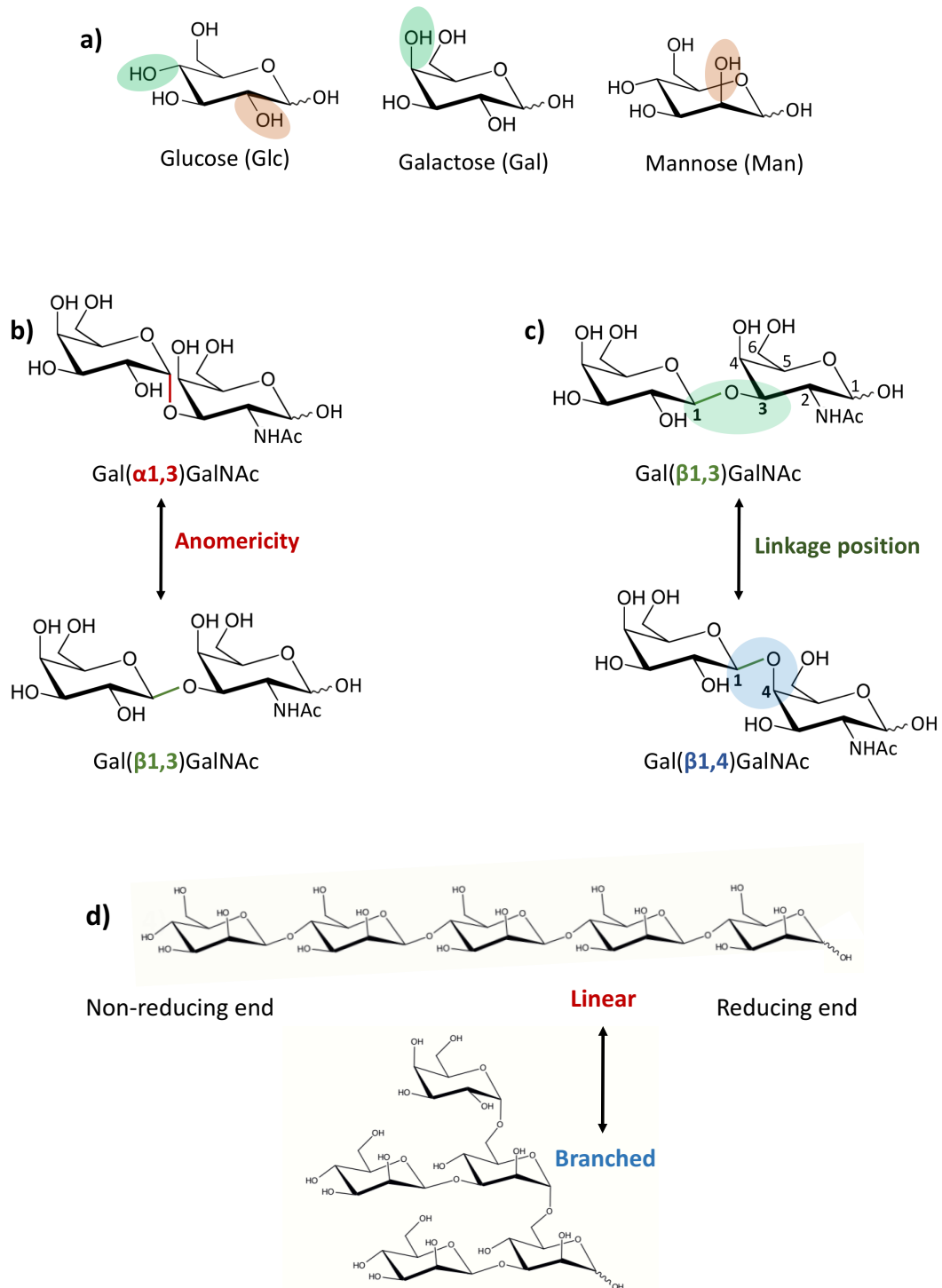
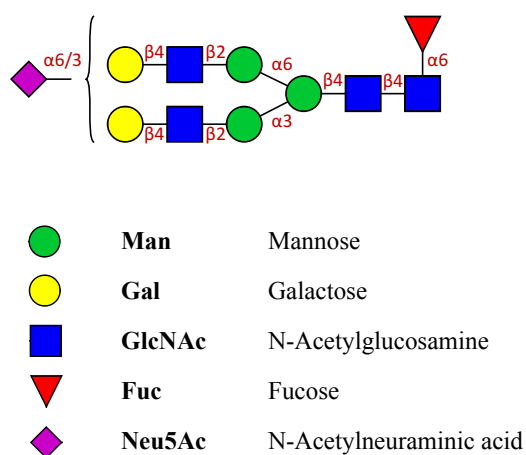


Figure 1.5. Types of glycan isomerism: a) isomeric monosaccharide building blocks; b) anomericity of the glycosidic linkage; c) different linkage positions; d) different connectivities (linear or branched structures).

### 1.3. Glycan nomenclature

Until the 1970s, the scientific community used either a drawing method or the standard IUPAC abbreviations for glycan representation. However, the subsequent development of the glycomics field and the discovery of more complex glycan structures made these two methods impractical. In 1978, Kornfeld and colleagues introduced a new simple and elegant way to depict glycan structures using a set of geometric shapes and colors.<sup>30</sup> This system became popular, though it had some limitations and inconsistencies. Over the following four decades, the system has undergone many transformations and standardizations based on the inputs and recommendations from scientists involved in this field. The most advanced and complete version was published in *Essentials of Glycobiology* in 2015.<sup>31</sup> Editors advised the representation to be called Symbol Nomenclature For Glycans (SNFG). Currently, the SNFG is hosted by the National Center for Biotechnology Information (NCBI) in the US. Figure 1.6 (left) shows SNFG rules for depicting glycans on the example of the N-linked glycan A1F. On the other hand, the Oxford nomenclature for glycan notation (UOXF) was independently developed by the researchers from the Oxford Glycobiology Institute in 2009.<sup>32</sup> The UOXF uses black and white for monosaccharide symbols (Figure 1.6, right).

#### Symbol Nomenclature for Glycans (SNFG)



#### Oxford Nomenclature (UOXF)

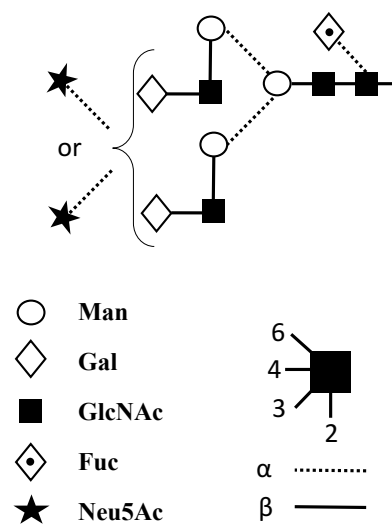


Figure 1.6. Comparison of the symbol nomenclature for glycans (SNGF) and the Oxford nomenclature (UOXF) for glycan notation on the example of the N-linked glycan A1F.

## 1.4. Synthesis of N-linked glycans in eukaryotes

The biosynthesis of N-linked glycans occurs via three major steps: (I) synthesis of a dolichol-linked precursor oligosaccharide, (II) transfer of the precursor oligosaccharide to a protein and (III) processing of the oligosaccharide.<sup>33-35</sup> Steps I and II occur in the endoplasmic reticulum (ER) and have highly conserved pathways. Thus, the spectrum of glycoforms remains rather uniform until step III, which is carried out in the Golgi apparatus, where structural diversification is introduced through a series of nonuniform species-, cell type-, protein-, and even site-specific modifications.

The dolichol-linked precursor is represented by the  $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$  oligosaccharide attached to the dolichol (a lipid molecule composed of repeating isoprene units) through a phosphate linkage (Figure 1.7). The dolichol is attached to the ER and originally carries one phosphate group. The oligosaccharide chain is then extended through the addition of the 2 GlcNAc, 9 Man and 3 Glc residues. Once the precursor oligosaccharide is assembled, it is transferred by oligosaccharyltransferase (OST) to the Asn in a receptive Asn-X-Ser/Thr sequon in a protein in the lumen of the ER membrane.

The first processing steps still take place in the ER and correspond to the trimming of the three Glc residues and the terminal  $\alpha$ 1-2Man from the central arm from the precursor. The formed structure is a key intermediate to high-mannose type N-glycans. In the *cis*-Golgi, the trimming continues to produce  $\text{Man}_5\text{GlcNAc}_2$ , a key intermediate of hybrid N-glycans. This is followed by addition of a GlcNAc residue to the core structure. In the *medial*-Golgi, the terminal  $\alpha$ 1-3Man and  $\alpha$ 1-6Man residues are trimmed and a second GlcNAc is added to the core to form the precursor for all biantennary, complex N-glycans. A further enzymatic modification of N-glycan branches takes place in the *trans*-Golgi. It includes the extension with different building blocks such as sialic acids, Fuc, Gal, GlcNAc, and sulfate, all of which lead to a vast variety of complex N-glycans that differ in branch number, composition, length and isomeric content.

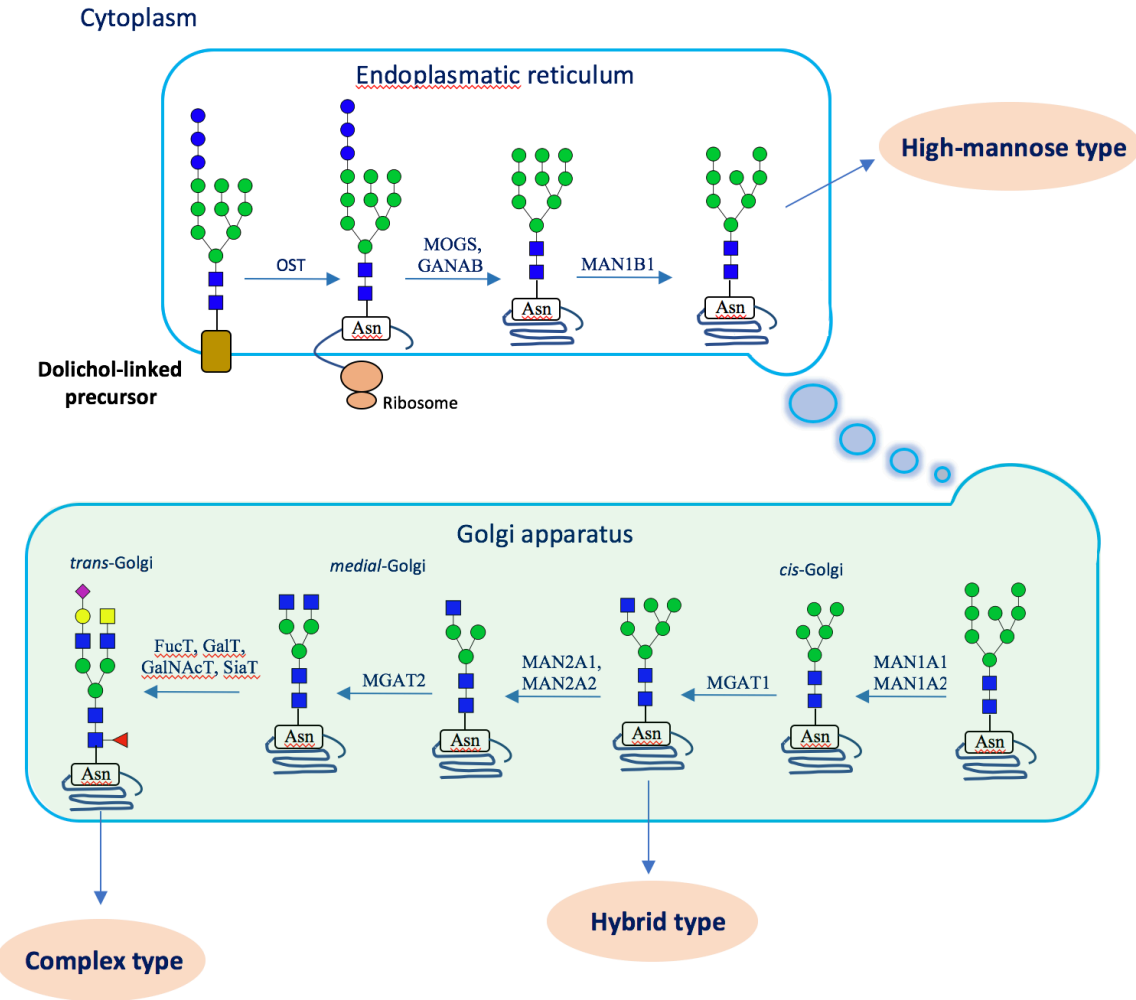


Figure 1.7. Scheme of the biosynthesis of the N-linked core oligosaccharide in eukaryotes.

All eukaryotic N-glycans share a common Man-3 core structure,  $\text{Man}\alpha 1\text{-3}(\text{Man}\alpha 1\text{-6})\text{Man}\beta 1\text{-4GlcNAc}\beta 1\text{-4GlcNAc}\beta 1$ , and are divided into three types depending on the core extension: (1) *high-mannose* (2) *hybrid* and (3) *complex* (Figure 1.8).<sup>35</sup>

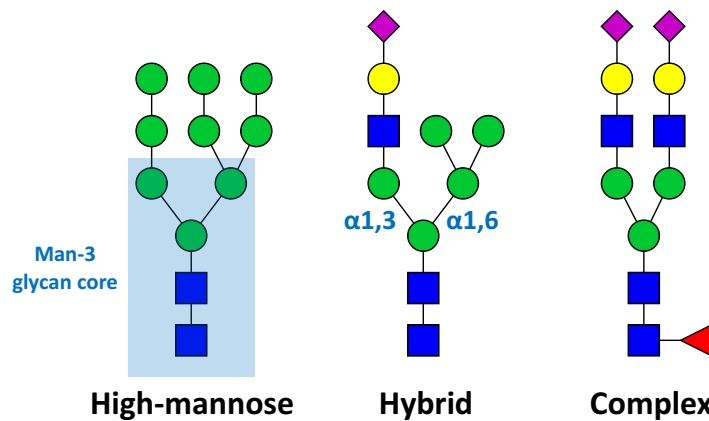


Figure 1.8. Three major types of N-linked glycans.

## 1.5. Traditional methods for glycan analysis

### 1.5.1. Mass spectrometry (MS)

Mass spectrometry is based on the production and detection of ions according to their mass to charge ( $m/z$ ) ratio. This is the only analytical method that can provide information on the primary structure of glycans at femtomolar and even attomolar levels with a wide dynamic range. Glycans contain labile residues such as fucose and sialic acid that might easily be fragmented during ionization. For this reason, MS analysis of glycans became possible only after the development of soft ionization techniques, which impart little excess energy and thus generate intact biomolecular ions. To obtain an overall glycan profile, the two most widely used techniques include Electrospray Ionization (ESI) and Matrix Assisted Laser Desorption/Ionization (MALDI).<sup>36, 37</sup>

While MS is a key technique for glycan analysis, when used alone, it cannot provide structural information, including glycan isomerism. However, the ability to employ controlled dissociation steps and tandem mass spectrometry allows MS methods to overcome this limitation. Low/high energy dissociation and enzymatic cleavage methods are employed for this purpose. The most common low energy techniques are collision-induced dissociation (CID)<sup>38</sup> and infrared multiphoton dissociation (IRMPD).<sup>39</sup> These techniques result in primarily glycosidic bond cleavages that provide information on glycan sequence and connectivity. The high energy CID methods include high-energy CID, electron transfer dissociation (ETD),<sup>40, 41</sup> electron capture dissociation (ECD)<sup>42</sup> and ultraviolet photodissociation (UVPD).<sup>43</sup> They lead to more abundant cross-ring cleavages, which provide more detailed information on linkage positions and bond anomericity, complementing low energy methods. However, they struggle with issues related to interpreting a huge amount of generated MS data and the lack of databases.

Enzymatic digestion employs a set of exoglycosidases, which selectively cleave monosaccharides from the non-reducing end. Knowing the specificity of the applied enzymes, one can infer the primary structure of the unknown parent glycan.<sup>44</sup>



## 1.5.2. Liquid chromatography (LC)

High-performance anion-exchange chromatography (HPAEC) in combination with pulsed amperometric detection (PAD) is a widely used method for glycan separation and quantification.<sup>45-47</sup> It is based on the separation of negative analyte ions according to their affinity to the cationic stationary phase. The main advantage of HPAEC-PAD is that it does not require a derivatization step. Characterization of sialic acid content in monoclonal antibodies is an important application of HPAEC. Even though HPAEC demonstrates a great promise in glycan analysis, this method suffers from challenges associated with coupling it to ESI-MS due to high salt gradients, high pH and the corrosive nature of NaOH in the mobile phase. In addition, HPAEC exhibits issues with the noise level and the baseline.

Porous graphitized carbon (PGC) liquid chromatography demonstrates high selectivity towards linkage and positional glycan isomers.<sup>48, 49</sup> Its separation mechanism is based on the hydrophobic interactions of the stationary phase with the analytes. The packing material exhibits several advantages: excellent mechanical, thermal and chemical stability, applicable in the pH range of 0–14; well-defined pores; no swelling nor shrinking.<sup>50, 51</sup> Carbohydrates are typically derivatized to make them more hydrophobic and to increase affinity to the stationary phase. The most commonly used labels for N-glycans include a 2-aminobenzamide (2-AB) and a 2-aminobenzoic acid (2-AA) with ultraviolet (UV)- and fluorescence (FL)-based detection. Derivatization with a hydrophobic label also increases glycan ionization efficiency. In addition, the mobile phase (an acetonitrile-water solution) used in PGC chromatography allows it to be compatible with ESI-MS.

Among all LC-based approaches, hydrophilic interaction liquid chromatography (HILIC) combined with FL, UV or evaporative light scattering (ELSD) detection is considered to be the method of choice for analysis of released N-linked glycans, as it is especially efficient at separating highly polar compounds such as carbohydrates. HILIC employs polar stationary phases like normal-phase chromatography (NP-LC) and the mobile phase is similar to the reversed-phase (RP) LC type. Typical HILIC stationary phases consist of bare silica or silica gels modified with polar functional groups (amide, diol, amino, cyano, pentalol).<sup>52, 53</sup>

In HILIC, the components of the analyte mixture are separated by differential distribution between two liquid layers: an organic mobile phase and a water layer adsorbed onto the hydrophilic stationary phase (Figure 1.9). More hydrophilic molecules elute later from a column as their partitioning equilibrium is more shifted towards the adsorbed water layer on

the stationary phase. The separation mechanism takes into account the polarities of the separated compounds and their degree of solvation, as well as hydrogen bonding, dipole-dipole and electrostatic interactions. Each of these parameters can dominate depending on the type of a stationary phase and mobile phase, including the ratio between water and organic solvents, salt concentration and pH.

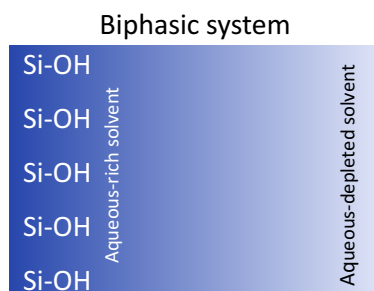


Figure 1.9. Scheme of the HILIC conditions.<sup>52</sup>

Though it is capable of separating both native and modified glycans, a typical workflow includes glycan labeling with fluorophores.<sup>54, 55</sup> A typical mobile phase includes a high content of a water-miscible polar organic solvent such as methanol or acetonitrile with a volatile buffer solution, which allows coupling with MS.<sup>56</sup> Elution is achieved by increasing the aqueous buffer concentrations.

Since LC systems often suffer from run-to-run variations, the retention time is usually measured relative to glucose units (dextran ladder), enabling oligosaccharide identification.

### 1.5.3. Capillary electrophoresis (CE)

In capillary electrophoresis, the separation mechanism is based on the charge-to-size ( $c/s$ ) ratio of analytes in an electric field. It is carried out in a silica capillary filled with a background electrolyte. Electroosmotic flow (EOF) and electrophoretic mobility (EM) are two factors that affect the transport of analytes in CE. CE with laser-induced fluorescent (LIF) detection is a powerful analytical method that has been widely used for glycan analysis in free solutions (zone electrophoresis) or gel-filled columns (gel electrophoresis). It has demonstrated its ability to distinguish between certain linkage and positional glycan isomers at relatively high speed and separation efficiency.<sup>57-60</sup> In addition, it requires a low amount of sample (much  $< 1 \mu\text{L}$ ) for protein analysis.

Despite its advantages, oligosaccharide analysis in CE is limited by several factors. The separation is complicated by the lack of charge on glycan molecules, except for glycans with

acidic sugars. In addition, carbohydrate molecules do not have a chromophore and cannot be detected by light adsorption in the UV region. To facilitate the glycan analysis in CE, labeling the carbohydrates with a charged fluorophore 8-aminopyrene-1,3,6-trisulfonate (APTS) is used. Coupling CE to MS is challenging due to the presence of buffers in CE that are not compatible with MS and due to the low flow rates of CE. Only recent studies have demonstrated successful hyphenation of CE with ESI-MS for rapid and sensitive glycan analysis. In addition, similarly to LC, CE migration times should be normalized by internal standards, and the corresponding glucose-unit (GU) are used for identification.

#### 1.5.4. Nuclear magnetic resonance (NMR) spectroscopy

NMR has proven to be a suitable tool for de novo glycan structural characterization.<sup>61, 62</sup> Due to its non-destructive nature, NMR leaves the sample intact for further analysis (by e.g., LC, MS). The main approach for determining a glycan sequence is 2-D heteronuclear multiple-bond correlation (HMBC), which detects the coupling between the anomeric H and the C atom on the opposite side of the glycosidic linkage.

Though NMR provides a way to unravel structures of completely novel sugars, this method shows low sensitivity compared to other spectroscopic and spectrometric techniques. To perform an NMR analysis, a large amount (mg-g range) of pure samples with predefined isotope labeling is required. The total amount of released glycoprotein glycans is not sufficient, as it is usually available at nM– $\mu$ M levels. Moreover, the limitations of NMR include the cost of spectrometers and the high level of expertise required for the interpretation of the NMR spectra.<sup>63, 64</sup>

## 1.6. Ion mobility separation (IMS)

Ion mobility spectrometry is based on the separation of ions in an inert gas (buffer gas) under the influence of an electric field.<sup>65-68</sup> This idea originates from the works of Thomson and Rutherford in 1896, before it was introduced as an analytical tool in the 1960s and applied for the analysis of peptides and proteins in the 1990s. The mobility of an ion in a gas ( $K$ ) is calculated as the quotient of the drift velocity ( $v_d$ ) and the applied uniform electric field ( $E$ ) (Equation 1.1).

$$K = \frac{v_d}{E} \quad (1.1)$$

Ions are separated either in a time-dispersive or space-dispersive manner (the ions with different mobilities are separated in time or in space, respectively) based on the particular IMS method. The experimentally derived  $K$  is normalized to standard conditions in order to obtain the reduced mobility ( $K_0$ ) that is used under different experimental conditions depending on the temperature ( $T$ ) and pressure ( $P$ ) of the buffer gas (Equation 1.2).

$$K_0 = K \frac{P}{760} \frac{273.2}{T} \quad (1.2)$$

The measured mobility is usually converted into the calculated ion-neutral collision cross-section value (CCS or  $\Omega$ ) using the Mason-Schamp equation (Equation 1.3).<sup>69</sup> The CCS represents the effective area (typically denoted in units of  $\text{\AA}^2$ ) of the ion that interacts with the drift gas under precise experimental conditions and gives information about the conformation of the ion.

$$\Omega = \frac{3}{16N} \left( \frac{2\pi}{\mu k_b T} \right)^{\frac{1}{2}} \frac{ze}{K_0} \quad (1.3)$$

This equation includes the following parameters:  $N$  - buffer gas density;  $ze$  - charge state of an ion ( $e = 1.6 \times 10^{-19}$  C);  $\mu$  - reduced mass of an ion ( $m$ ) and buffer gas ( $M$ ): [ $\mu = mM/(m + M)$ ];  $k_b$  - Boltzmann's constant ( $1.38 \times 10^{-23}$  J/K). CCS is used for the construction of databases, calibration of methods and the prediction of molecular structures. Helium and nitrogen are two of the most widely used inert drift gases. The resolving power ( $R_p = \Omega/\Delta\Omega$ ) of typical IMS instruments is in the order of 50–400. An IMS analysis can be conducted on

different available platforms and the choice depends on the experimental objectives, taking into account their individual advantages and disadvantages.

### 1.6.1. Drift tube ion mobility spectrometry (DTIMS)

Drift tube ion mobility spectrometry is the simplest form of IMS that is widely used for glycan analysis.<sup>70-72</sup> A DTIMS setup includes a tube filled with a buffer gas. The tube is composed of a series of electrodes generating a weak uniform electric field ( $\sim 5\text{--}15\text{ V/cm}$ ) that propels ions through the drift region (Figure 1.10). Ions move against neutral gas and are separated in a time-dispersive manner. The drift velocity ( $v_d$ ) in DTIMS is defined by  $L/t_d$  ( $L$  - length of the drift tube,  $t_d$  - drift time). The mobility of ions depends on the number of collisions with the buffer gas. For a given charge, compact ions drift faster than elongated ones because they undergo fewer collisions with the buffer gas. The electric field must be maintained low to obtain a correlation between mobility and shape/size. At the exit of the drift tube, ions are detected as a function of time, leading to an arrival time distribution (ATD).

The most significant advantage of DTIMS is its ability to directly obtain highly accurate CCS values without calibration. Other IMS methods require a calibrant with a CCS value previously obtained from a calibration curve based on DTIMS. Moreover, DTIMS instruments do not require RF for ion confinement which minimizes ion heating effects. On the other hand, DTIMS analyses have a very low duty cycle. In many experiments, less than 10 % of ions in each pulse are used. Increasing the resolving power of DTIMS devices represents an additional challenge.

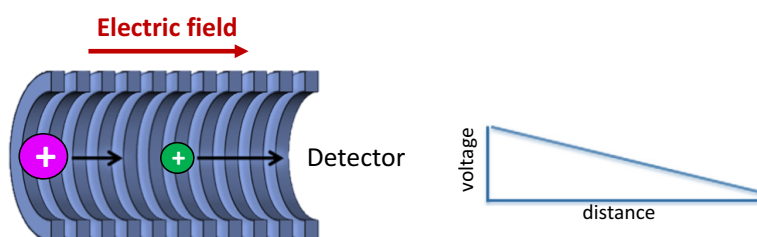


Figure 1.10. Schematic representation of the DTIMS platform (Adapted from Fig. 1 of Ref. 67).

## 1.6.2. Traveling wave ion mobility spectrometry (TWIMS)

Traveling wave ion mobility spectrometry was first described in 2004 and commercialized in 2006 with the Synapt HDMS systems from Waters. TWIMS is also a time-dispersive method and is very similar to DTIMS. It consists of stacked ring electrodes and uses pulsed ion packets (Figure 1.11). However, while DTIMS applies a uniform electric field to move ions through the full length of the mobility region, TWIMS uses a combination of RF and DC voltages to create a wave that induces the migration of ions. Ions with a smaller CCS experience fewer collisions with the drift gas and surf on the wave, while ions with a larger CCS undergo more collisions and slip more often behind the waves, which makes their migration slower. RF potential provides radial ion confinement and prevents them from diffusion, allowing for a higher ion signal. The mobility separation in TWIMS is defined by the traveling wave (TW) amplitude and its speed, as well as the pressure of the drift gas.

Due to the incomplete understanding of ion trajectories in TWIMS, the CCS cannot be determined directly from the drift time but only after a calibration using molecules with known CCS values, ideally of a similar class to the studied analytes. The main advantage of the TWIMS technology is that it allows to manipulate ions along extremely long paths to enhance the mobility separation (tens to hundreds of meters in length) without significant ion loss. This has been implemented in the circular IMS device from Waters and in the Structures for Lossless Ion Manipulations (SLIM) technology developed in the Smith group (PNNL, Richland, WA).<sup>73, 74</sup> Moreover, TWIMS-based devices are more compact and more flexible due to the possibility of coupling a fragmentation step before and after the mobility region and the much greater duty cycle in comparison with DTIMS platforms.<sup>75-77</sup>

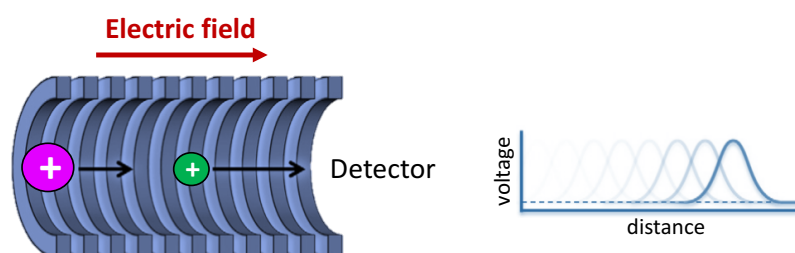


Figure 1.11. Schematic representation of the TWIMS platform (Adapted from Fig. 1 of Ref. 67).

### 1.6.3. Trapped ion mobility spectrometry (TIMS)

In contrast to DTIMS or TWIMS, where the drift gas has no directional flow, TIMS uses a high flow of nitrogen buffer gas that propagates ions towards the detector. An electric field is applied in the opposite direction to the gas flow and traps ions in different regions of the TIMS mobility region depending on the  $\Omega/z$  (Figure 1.12).<sup>78, 79</sup> Then the electric field is gradually decreased, ejecting ions from high to low  $\Omega/z$ . In DTIMS and TWIMS, all ions are studied under the same experimental conditions, while TIMS requires changes to the experimental parameters to be able to see all ions.

Similarly to DTIMS and TWIMS, TIMS uses pulsed ion packets and has some losses in its duty cycle. In addition, TIMS also operates at a low electric field that enables the extraction of the CCS values via calibration.<sup>80</sup> Despite the fact that TIMS appeared very recently, this method has already demonstrated several advantages. These include high resolving powers (200-400 K/ $\Delta$ K), compactness (5–10 cm) of the TIMS devices and the flexibility to incorporate several TIMS separations with fragmentation steps.

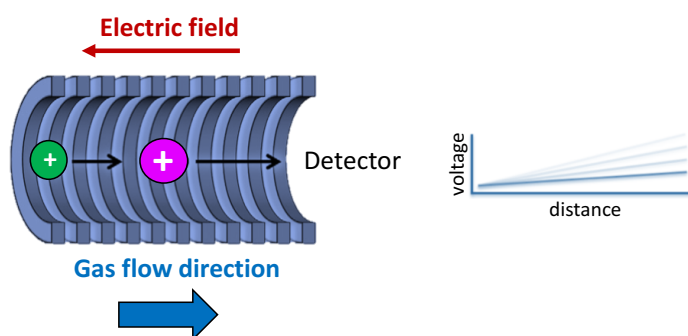


Figure 1.12. Schematic representation of the TIMS platform (Adapted from Fig. 1 of Ref. 67).

### 1.6.4. Field asymmetric waveform ion mobility spectrometry (FAIMS)

Field asymmetric waveform ion mobility spectrometry is a space-dispersive IMS method that is also known as differential mobility spectrometry (DMS) and differential ion mobility spectrometry (DIMS). In FAIMS, ions are carried by a drift gas at reduced or atmospheric pressure between two electrodes applying a high asymmetric waveform, up to 10 000 V/cm (Figure 1.13).<sup>81</sup> A compensation voltage (CV) is then superimposed on the waveform. By varying the CV, the trajectory of the ions of interest can be changed to release them while all other ions are eliminated.

Unlike all previously described methods, FAIMS operates at high electric fields and the CCS values cannot be calculated from the mobility data. For this reason, FAIMS is used as a

filtering technique rather than a method for structural characterization. Several groups have demonstrated the application of FAIMS for the study of glycans.<sup>82, 83</sup>

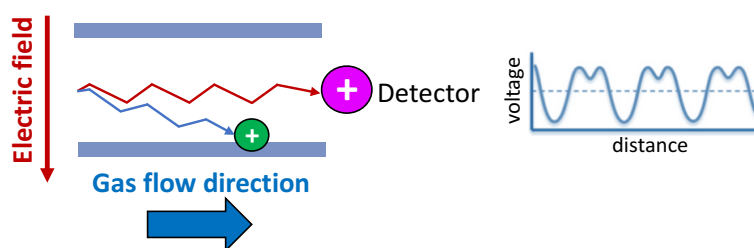


Figure 1.13. Schematic representation of the FAIMS platform.

## 1.7. Gas-phase infrared (IR) spectroscopy of ions

Gas-phase IR spectroscopy is a relatively new tool for glycan analysis; however, this area of research is rapidly growing. The interest is driven by two factors: the increasing accessibility of tunable IR laser sources and the additional information provided by vibrational spectra.

In conventional vibrational spectroscopy, where chemical substances are studied in liquid or solid forms, the sample directly adsorbs IR radiation and the relation between the incident ( $I$ ) and transmitted light beams ( $I_0$ ) measured as a function of the frequency of the IR radiation ( $\nu$ ) is described by the Lambert–Beer law:

$$I(\nu) = I_0 e^{-\sigma(\nu)Ln}, \quad (1.4)$$

$\sigma(\nu)$  - absorption cross section,  $L$  - path length,  $n$  - number density of the sample. However, in the gas phase, the number density is so small ( $< 10^8 \text{ cm}^{-3}$ ) that the difference between  $I(\nu)$  and  $I_0$  is not detectable. In contrast, *action spectroscopy* measures the absorption-induced change in the sample. The fraction of molecules affected can be defined as:

$$N(\nu) = N_0 e^{-\sigma(\nu)\Phi(\nu)}, \quad (1.5)$$

$\Phi(\nu)$  - the number of photons per second per unit area experienced by the sample. The argument of the exponent has values of the order of 1 for typical laser systems. Measuring the IR spectra eliminates matrix interferences revealed in liquid and solid phase vibrational spectroscopy. The main types of action spectroscopy implemented for carbohydrate analysis include IR multiple photon dissociation and cold ion spectroscopy techniques.



### 1.7.1. IR multiple photon dissociation (IRMPD)

The concept of IRMPD is based on the irradiation of a mass-selected ion of interest with a tunable IR laser at a particular frequency. If the laser frequency is in resonance with a vibrational transition of the molecule, the ion adsorbs one or several IR photons leading to its dissociation. IRMPD represents a non-coherent process where the absorbed energy is quickly dissipated between different vibrational degrees of freedom of the molecule via a process called *intramolecular vibrational redistribution* (IVR) (Figure 1.14). The sequential absorption of multiple photons increases the internal energy of the ion, until it becomes high enough for the dissociation to occur. This process causes a change in the mass spectrum where the precursor ion intensity is decreased while the ion signal of the dissociation product appears. IR spectra are constructed by plotting the fragmentation yield as a function of the laser frequency. Multiple photon processes require powerful IR light sources such as free electron lasers (FELs) or optical parametric oscillators (OPOs).

Using IRMPD spectroscopy, Compagnon and coworkers recently demonstrated that the anomericity of glycosidic linkages can be retained in the dissociation process, which is important structural information.<sup>84</sup> In addition, this spectroscopic approach does not require chemical tags. However, common issues in IRMPD spectroscopy are spectral congestion and low resolution, since spectra are recorded at room temperature.<sup>85, 86</sup> The spectral congestion arises from the coexistence of several conformers with different adsorptions patterns at room temperature and from the molecular heating of the ions during the absorption of multiple photons. Nevertheless, IRMPD remains a powerful method for the analysis of small systems such as mono- and disaccharides.

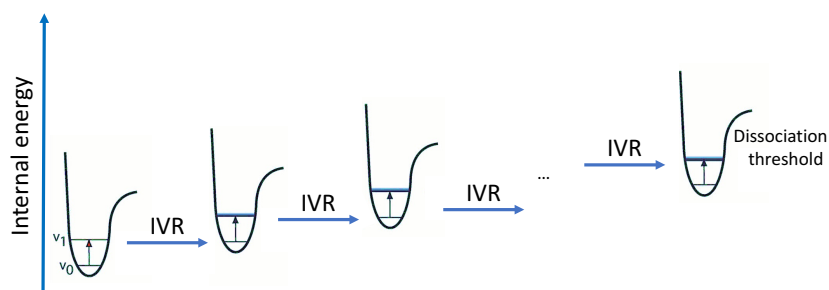


Figure 1.14. Scheme of the IRMPD mechanism. The sequential absorption of multiple photons increases the internal energy of the ion through the IVR process. When the dissociation threshold is exceeded, the ion undergoes fragmentation.

## 1.7.2. Cold ion spectroscopy

Thermal congestion in the vibrational spectra can be fully eliminated by confining all the ions in a single rovibrational state. In order to suppress the broadening, the method of cryogenic cooling of rovibrational degrees of freedom is used and widely represented in different experimental realizations. Figure 1.15 demonstrates a significant difference between the UV spectra of gas-phase protonated tyrosine recorded at room temperature (red) and at 12 K (blue). The room temperature spectrum is very broad and does not provide any information on the molecule, while the cryogenic spectrum exhibits highly-resolved transitions and can be used both for calculations and as a unique identifier.

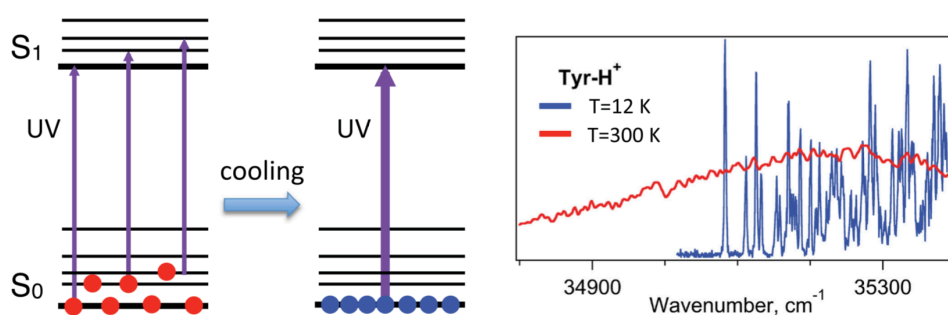


Figure 1.15. The left part illustrates the suppression of thermal congestion by cryogenic cooling; the right part compares the UV spectra of gas-phase protonated tyrosine measured at room temperature (300 K) and at 12 K.<sup>87</sup>

## HeDrop spectroscopy

Mucha et al. used superfluid helium nanodroplets to achieve ion cooling to  $\sim 0.37$  K.<sup>88</sup> Ions produced by nESI and mass-selected were encapsulated in He droplets inside a linear ion trap. The doped droplets were then irradiated with IR photons from a tunable FEL, leading to the resonant absorption of photons followed by the ion ejection from the droplets. The absorption does not lead to ion heating in He droplets due to evaporative cooling. Scanning the IR wavelength enables the plotting of an IR spectrum. The spectra of six extremely similar isomeric trisaccharides displayed unprecedented resolving power. Each spectrum demonstrated highly resolved bands, leading to a unique vibrational fingerprint that allows unambiguous distinction between all possible types of glycan isomerism.

Despite the remarkable spectral resolution, only very few groups use this method for glycan analysis. An extremely complex experimental setup and limited accessibility to a free-electron laser are two major drawbacks that leave HeDrop spectroscopy far from the broad application.

### **Messenger-tagging spectroscopy**

Tagging spectroscopy is a more widely used technique to perform cold ion spectroscopy. Its concept is based on forming weakly bound complexes between a precursor analyte and inert ‘tag’ atoms or molecules.<sup>89</sup> The adsorption of a single photon induces a tag detachment, therefore this method requires less powerful light sources (typically OPO laser systems). Although tagging spectroscopy cannot reach temperatures below 1 K, the accessible range is sufficiently low (from ~ 10 K) to obtain highly-resolved IR fingerprints of glycans.<sup>90</sup> Moreover, the experimental setups are more user-friendly than in HeDrop IR and will possibly be commercially available in the near future. This type of spectroscopy will be discussed in more detail in Chapter 2.

### **IR-IR double-resonance spectroscopy**

In order to obtain spectra of individual glycan conformers, cryogenic IR-IR double-resonance spectroscopy can be implemented. In an IR-IR scheme, one IR laser is fixed at a specific resonance frequency associated with a particular conformer and depletes its population while the second IR laser spectroscopically interrogates the other conformer, yielding its IR spectrum.<sup>91, 92</sup> The main limitation of this approach is the necessity to identify conformer-specific transitions that can be time-consuming and simply impossible in case of large glycan structures where the adsorptions of different conformers are overlapping.

## 1.8. Multidimensional techniques

While mass spectrometry is the main technique for glycan analysis, when used alone, it cannot provide full structural and quantitative information. A variety of MS-based hyphenated techniques have been developed for a more complete characterization of carbohydrates. Coupling MS with separation methods dramatically improves the ionization efficiency and leads to simplified mass spectra, which is crucial for the analysis of complex mixtures. Moreover, MS cannot distinguish all different types of glycan isomerism and needs orthogonal techniques for their elucidation. The most widely used example that already became routine is the combination of MS with liquid chromatography (LC-MS).<sup>93</sup> It has proven to be a powerful method for glycan separation and characterization. The development of sheath-liquid interfaces made it also possible to combine capillary electrophoresis with mass spectrometry (CE-MS).<sup>94</sup>

95

Recently, ion mobility spectrometry was introduced as a promising orthogonal technique for glycan analysis. In a relatively short time, it has been established as a robust analytical method that can resolve many of the glycan isomers on the timescale of milliseconds. Nearly all major MS companies have already implemented different types of IMS into their instruments, including DTIMS (Agilent), TWIMS (Waters), FAIMS (Thermo) and TIMS (Bruker), making IMS-MS more accessible for wide applications.

Other examples include three-dimensional platforms where IM-MS is coupled to LC/CE in order to enhance glycan analysis. One other possibility is adding a spectroscopic dimension to IMS-MS for further isomer discrimination. Several groups demonstrated promising results on combining IMS-MS with IRMPD spectroscopy.<sup>96</sup> However, IRMPD spectra of glycans are recorded at room temperature and suffer from spectral congestion and low resolution that limit IRMPD spectroscopy to mono- and disaccharides. In order to overcome this limitation, cryogenic-ion IR spectroscopy was recently combined with IM-MS for glycan analysis as it provides highly resolved IR fingerprints that are unique even for complex molecules.<sup>90</sup>

## 1.9. Our approach for glycan analysis

Despite the huge progress in glycan analysis associated with the hyphenation of different analytical techniques, even the most powerful of them (represented by coupling IMS or LC to MS) are still blind to many of the structural details in isomeric glycans. Moreover, the TW-based IMS devices that now provide the highest resolving power for glycan separation cannot be used alone for identification and database construction. Arrival time distributions determined with this technique depend sensitively on the traveling wave speed and amplitude as well as the temperature and pressure of the drift gas and hence need to be calibrated to obtain CCSs. Given a typical deviation of reported CCSs of approximately 2%, it would be tenuous to identify isomers with a difference in CCSs less than this value by IMS alone, even if done carefully. In order to overcome these limitations, an orthogonal technique needs to be implemented.

In our laboratory, we have recently demonstrated that cryogenic vibrational spectroscopy provides unique fingerprints of isolated glycans that are extremely sensitive to the slightest structural differences.<sup>72, 97-99</sup> The reason of this sensitivity stands behind glycans' nature itself. Glycans contain many highly polar OH groups (each monosaccharide block contains 4–5 OH groups) that are strongly coupled to one another. A change in orientation, even in one of them, affects the entire network of interactions which is directly reflected in the vibrational spectrum. In addition, performing an experiment at cryogenic temperatures allows us to obtain highly resolved spectra, which would not be possible with room temperature spectroscopic techniques.

One very important advantage of IR spectroscopy is that the vibrational spectrum is an intrinsic property of the molecule that is insensitive to external parameters. The IR laser needs to be calibrated with a standard wavemeter, and sufficiently low temperatures should be reached to perform tagging spectroscopy. Adding cryogenic IR to IMS-MS provides an additional dimension for glycan isomer discrimination.

We have developed a novel method for analyzing the primary structure of glycans, which combines state-of-the-art ultra-high resolution ion mobility spectrometry, cryogenic vibrational spectroscopy and time-of-flight mass spectrometry into a single instrument. The measured data includes the mass and a cryogenic vibrational spectrum of each species, while the IMS dimension is used as a separation tool. All obtained data can be tabulated in a database and used for glycan identification.

It is also important to note that a wide analytical application of vibrational spectroscopy became possible due to recent progress in solid-state IR laser sources. Powerful fiber-pumped lasers allow to reduce the scanning time from 20 to 2 min in comparison with measurements done on the OPO systems and make IR spectroscopy competitive with other analytical techniques.

## 1.10. Outline of the thesis

*Chapter 2* introduces experimental techniques that we employed in this work to study N-linked glycans. The first part provides an overview of the three generations of home-build instruments for the analysis of glycans in the gas phase using cold ion spectroscopy. We also present a detailed description of the SLIM-based ion mobility modules incorporated into the instruments and the basic principles of cold messenger-tagging spectroscopy. The second part describes ultra-performance liquid chromatography (UPLC) coupled to a Q-TOF mass spectrometer in a HILIC mode.

*Chapter 3* demonstrates our results on developing a new approach for determining the primary structure of glycans based on the combination of cryogenic IR with two fragmentation techniques. By combining the protocols for selective exoglycosidase digestion and collision-induced dissociation with our home-built setup, we demonstrate them as powerful methods for the identification of unknown glycan structures and for expanding our spectroscopic database. We also show the complementarity of these two methods by enzymatically generating standards for some isomeric CID glycan fragments.

*Chapter 4* focuses on the separation and identification of positional isomers of N-glycans using ion-mobility-selective IR fingerprints and their comparison with a machine-learning approach. For this purpose, we implemented a chemoenzymatic synthesis of an asymmetric N-glycan standard that is difficult to obtain commercially. In addition, we studied positional isomers of glycans cleaved from monoclonal antibodies and demonstrated the impact of the host cell line on their glycan profile. We showed that our IMS-MS-IR approach can complement existing methods for establishing the similarity of glycan profiles between biological drugs and their biosimilars.

*Chapter 5* describes the search of new glycan biomarkers in human blood serum using our approach. We studied sialylated glycans released from samples of healthy female donors and patients diagnosed with different stages of breast cancer. Exploiting the differences in

glycosylation between diseased and healthy individuals, we have defined a preliminary potential glycan candidate that will be used for more detailed screening. In addition, we demonstrate a scheme for isomer identification of sialylated glycans combining the IMS-MS-IR approach with CID.

In *Chapter 6*, we describe experiments in which we performed a relative quantification for mono, doubly, triply and quadruply sialylated glycans released from follicle stimulating hormone (FSH) using ion mobility spectrometry, since their content is critical to guarantee drug quality. We demonstrate the comparison of the results obtained using two types of electrospray sources and the reproducibility between measurements.

*Chapter 7* summarizes the results and proposes future work.

# References

1. Ohtsubo, K.; Marth, J. D., Glycosylation in cellular mechanisms of health and disease. *Cell* **2006**, *126* (5), 855-67.
2. Seeberger, P. H., Monosaccharide Diversity. In *Essentials of Glycobiology*, rd; 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 (NY), 2015; pp 19-30.
3. Nilsson, J.; Halim, A.; Grahn, A.; Larson, G., Targeting the glycoproteome. *Glycoconj J* **2013**, *30* (2), 119-36.
4. Apweiler, R.; Hermjakob, H.; Sharon, N., On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Bba-Gen Subjects* **1999**, *1473* (1), 4-8.
5. Reily, C.; Stewart, T. J.; Renfrow, M. B.; Novak, J., Glycosylation in health and disease. *Nat Rev Nephrol* **2019**, *15* (6), 346-366.
6. Marth, J. D.; Grewal, P. K., Mammalian glycosylation in immunity. *Nat Rev Immunol* **2008**, *8* (11), 874-87.
7. Dwek, R. A., Glycobiology: Toward Understanding the Function of Sugars. *Chem Rev* **1996**, *96* (2), 683-720.
8. Schjoldager, K. T.; Narimatsu, Y.; Joshi, H. J.; Clausen, H., Global view of human protein glycosylation pathways and functions. *Nat Rev Mol Cell Biol* **2020**, *21* (12), 729-749.
9. Varki, A.; Gagneux, P., Biological Functions of Glycans. In *Essentials of Glycobiology*, rd; 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 (NY), 2015; pp 77-88.
10. Heinemann, L., Biosimilar Insulin and Costs: What Can We Expect? *J Diabetes Sci Technol* **2015**, *10* (2), 457-62.
11. Russell, A. C.; Simurina, M.; Garcia, M. T.; Novokmet, M.; Wang, Y.; Rudan, I.; Campbell, H.; Lauc, G.; Thomas, M. G.; Wang, W., The N-glycosylation of immunoglobulin G as a novel biomarker of Parkinson's disease. *Glycobiology* **2017**, *27* (5), 501-510.
12. Zhou, J. Y.; Cobb, B. A., Glycans in Immunologic Health and Disease. *Annu Rev Immunol* **2021**, *39*, 511-536.
13. Dube, D. H.; Bertozzi, C. R., Glycans in cancer and inflammation--potential for therapeutics and diagnostics. *Nat Rev Drug Discov* **2005**, *4* (6), 477-88.



14. Menni, C.; Gudelj, I.; Macdonald-Dunlop, E.; Mangino, M.; Zierer, J.; Besic, E.; Joshi, P. K.; Trbojevic-Akmacic, I.; Chowienczyk, P. J.; Spector, T. D.; Wilson, J. F.; Lauc, G.; Valdes, A. M., Glycosylation Profile of Immunoglobulin G Is Cross-Sectionally Associated With Cardiovascular Disease Risk Score and Subclinical Atherosclerosis in Two Independent Cohorts. *Circ Res* **2018**, *122* (11), 1555-1564.
15. Kizuka, Y.; Kitazume, S.; Taniguchi, N., N-glycan and Alzheimer's disease. *Biochim Biophys Acta Gen Subj* **2017**, *1861* (10), 2447-2454.
16. Kuno, A.; Ikehara, Y.; Tanaka, Y.; Ito, K.; Matsuda, A.; Sekiya, S.; Hige, S.; Sakamoto, M.; Kage, M.; Mizokami, M.; Narimatsu, H., A serum "sweet-doughnut" protein facilitates fibrosis evaluation and therapy assessment in patients with viral hepatitis. *Sci Rep* **2013**, *3*, 1065.
17. van den Born, J.; van den Heuvel, L. P.; Bakker, M. A.; Veerkamp, J. H.; Assmann, K. J.; Weening, J. J.; Berden, J. H., Distribution of GBM heparan sulfate proteoglycan core protein and side chains in human glomerular diseases. *Kidney Int* **1993**, *43* (2), 454-63.
18. Kanwar, Y. S.; Linker, A.; Farquhar, M. G., Increased permeability of the glomerular basement membrane to ferritin after removal of glycosaminoglycans (heparan sulfate) by enzyme digestion. *J Cell Biol* **1980**, *86* (2), 688-93.
19. Arnold, J. N.; Saldova, R.; Galligan, M. C.; Murphy, T. B.; Mimura-Kimura, Y.; Telford, J. E.; Godwin, A. K.; Rudd, P. M., Novel glycan biomarkers for the detection of lung cancer. *J Proteome Res* **2011**, *10* (4), 1755-64.
20. Ruhaak, L. R.; Miyamoto, S.; Lebrilla, C. B., Developments in the identification of glycan biomarkers for the detection of cancer. *Mol Cell Proteomics* **2013**, *12* (4), 846-55.
21. Scott, E.; Munkley, J., Glycans as Biomarkers in Prostate Cancer. *Int J Mol Sci* **2019**, *20* (6).
22. Hu, M.; Lan, Y.; Lu, A.; Ma, X.; Zhang, L., Glycan-based biomarkers for diagnosis of cancers and other diseases: Past, present, and future. *Prog Mol Biol Transl Sci* **2019**, *162*, 1-24.
23. Mimura, Y.; Katoh, T.; Saldova, R.; O'Flaherty, R.; Izumi, T.; Mimura-Kimura, Y.; Utsunomiya, T.; Mizukami, Y.; Yamamoto, K.; Matsumoto, T.; Rudd, P. M., Glycosylation engineering of therapeutic IgG antibodies: challenges for the safety, functionality and efficacy. *Protein Cell* **2018**, *9* (1), 47-62.
24. Zhang, P.; Woen, S.; Wang, T.; Liau, 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), 740-65.

25. Kwong, P. D.; Mascola, J. R., HIV-1 Vaccines Based on Antibody Identification, B Cell Ontogeny, and Epitope Structure. *Immunity* **2018**, *48* (5), 855-871.

26. Byrne, G. W.; Stalboerger, P. G.; Du, Z.; Davis, T. R.; McGregor, C. G., Identification of new carbohydrate and membrane protein antigens in cardiac xenotransplantation. *Transplantation* **2011**, *91* (3), 287-92.

27. Lutz, A. J.; Li, P.; Estrada, J. L.; Sidner, R. A.; Chihara, R. K.; Downey, S. M.; Burlak, C.; Wang, Z. Y.; Reyes, L. M.; Ivary, B.; Yin, F.; Blankenship, R. L.; Paris, L. L.; Tector, A. J., Double knockout pigs deficient in N-glycolylneuraminic acid and galactose alpha-1,3-galactose reduce the humoral barrier to xenotransplantation. *Xenotransplantation* **2013**, *20* (1), 27-35.

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

29. Varki, A.; Kornfeld, S., Historical Background and Overview. In *Essentials of Glycobiology*, 3rd ed.; 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 (NY), 2015; pp 1-18.

30. Kornfeld, S.; Li, E.; Tabas, I., 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. *J Biol Chem* **1978**, *253* (21), 7771-8.

31. 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.; Lutteke, 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), 1323-4.

32. 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), 3796-801.

33. Helenius, A.; Aebi, M., Intracellular functions of N-linked glycans. *Science* **2001**, *291* (5512), 2364-9.
34. Tivey, T. R.; Parkinson, J. E.; Mandelare, P. E.; Adpressa, D. A.; Peng, W.; Dong, X.; Mechref, Y.; Weis, V. M.; Loesgen, S., N-Linked Surface Glycan Biosynthesis, Composition, Inhibition, and Function in Cnidarian-Dinoflagellate Symbiosis. *Microb Ecol* **2020**, *80* (1), 223-236.
35. Stanley, P.; Taniguchi, N.; Aebi, M., N-Glycans. In *Essentials of Glycobiology*, rd; 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 (NY), 2015; pp 99-111.
36. Ruhaak, L. R.; Xu, G. G.; Li, Q. Y.; Goonatilleke, E.; Lebrilla, C. B., Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses. *Chemical Reviews* **2018**, *118* (17), 7886-7930.
37. Han, L.; Costello, C. E., Mass spectrometry of glycans. *Biochemistry-Moscow+* **2013**, *78* (7), 710-720.
38. An, H. J.; Lebrilla, C. B., Structure elucidation of native N- and O-linked glycans by tandem mass spectrometry (tutorial). *Mass Spectrom Rev* **2011**, *30* (4), 560-78.
39. Polfer, N. C., Infrared multiple photon dissociation spectroscopy of trapped ions. *Chem Soc Rev* **2011**, *40* (5), 2211-21.
40. Ford, K. L.; Zeng, W.; Heazlewood, J. L.; Bacic, A., Characterization of protein N-glycosylation by tandem mass spectrometry using complementary fragmentation techniques. *Front Plant Sci* **2015**, *6*, 674.
41. Han, L.; Costello, C. E., Electron transfer dissociation of milk oligosaccharides. *J Am Soc Mass Spectrom* **2011**, *22* (6), 997-1013.
42. 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), 138-50.
43. 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), 1027-40.
44. Marino, K.; Bones, J.; Kattla, J. J.; Rudd, P. M., A systematic approach to protein glycosylation analysis: a path through the maze. *Nat Chem Biol* **2010**, *6* (10), 713-723.

45. Rohrer, J. S.; Basumallick, L.; Hurum, D., High-performance anion-exchange chromatography with pulsed amperometric detection for carbohydrate analysis of glycoproteins. *Biochemistry (Mosc)* **2013**, *78* (7), 697-709.
46. Hajba, L.; Csanky, E.; Guttman, A., Liquid phase separation methods for N-glycosylation analysis of glycoproteins of biomedical and biopharmaceutical interest. A critical review. *Anal Chim Acta* **2016**, *943*, 8-16.
47. Kaur, H., Characterization of glycosylation in monoclonal antibodies and its importance in therapeutic antibody development. *Crit Rev Biotechnol* **2021**, *41* (2), 300-315.
48. 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* **2015**, *1321*, 427-35.
49. Hua, S.; An, H. J.; Ozcan, S.; Ro, G. S.; Soares, S.; DeVere-White, R.; Lebrilla, C. B., Comprehensive native glycan profiling with isomer separation and quantitation for the discovery of cancer biomarkers. *Analyst* **2011**, *136* (18), 3663-71.
50. West, C.; Elfakir, C.; Lafosse, M., Porous graphitic carbon: a versatile stationary phase for liquid chromatography. *J Chromatogr A* **2010**, *1217* (19), 3201-16.
51. Pereira, L., Porous graphitic carbon as a stationary phase in HPLC: Theory and applications. *J Liq Chromatogr R T* **2008**, *31* (11-12), 1687-1731.
52. Buszewski, B.; Noga, S., Hydrophilic interaction liquid chromatography (HILIC)--a powerful separation technique. *Anal Bioanal Chem* **2012**, *402* (1), 231-47.
53. Molnarova, K.; Cokrtova, K.; Tomnikova, A.; Krizek, T.; Kozlik, P., Liquid chromatography and capillary electrophoresis in glycomic and glycoproteomic analysis. *Monatsh Chem* **2022**.
54. Melmer, M.; Stangler, T.; Schiefermeier, M.; Brunner, W.; Toll, H.; Rupprechter, A.; Lindner, W.; Premstaller, A., HILIC analysis of fluorescence-labeled N-glycans from recombinant biopharmaceuticals. *Analytical and Bioanalytical Chemistry* **2010**, *398* (2), 905-914.
55. Ahn, J.; Bones, J.; Yu, Y. Q.; Rudd, P. M.; Gilar, M., Separation of 2-aminobenzamide labeled glycans using hydrophilic interaction chromatography columns packed with 1.7  $\mu\text{m}$  sorbent. *J Chromatogr B* **2010**, *878* (3-4), 403-408.
56. Wuhler, 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.

57. Lu, G.; Carihfield, C. L.; Gattu, S.; Veltri, L. M.; Holland, L. A., Capillary Electrophoresis Separations of Glycans. *Chemical Reviews* **2018**, *118* (17), 7867-7885.
58. Gattu, S.; Carihfield, C. L.; Lu, G.; Bwanali, L.; Veltri, L. M.; Holland, L. A., Advances in enzyme substrate analysis with capillary electrophoresis. *Methods* **2018**, *146*, 93-106.
59. Guttman, A., Capillary electrophoresis in the N-glycosylation analysis of biopharmaceuticals. *Trac-Trend Anal Chem* **2013**, *48*, 132-143.
60. Camperi, J.; Pichon, V.; Delaunay, N., Separation methods hyphenated to mass spectrometry for the characterization of the protein glycosylation at the intact level. *J Pharmaceut Biomed* **2020**, *178*.
61. Unione, L.; Arda, A.; Jimenez-Barbero, J.; Millet, O., NMR of glycoproteins: profiling, structure, conformation and interactions. *Curr Opin Struc Biol* **2021**, *68*, 9-17.
62. Pauli, G. F.; Chen, S. N.; Lankin, D. C.; Bisson, J.; Case, R. J.; Chadwick, L. R.; Godecke, T.; Inui, T.; Kronic, A.; Jaki, B. U.; McAlpine, J. B.; Mo, S. Y.; Napolitano, J. G.; Orjala, J.; Lehtivarjo, J.; Korhonen, S. P.; Niemitz, M., Essential Parameters for Structural Analysis and Dereplication by H-1 NMR Spectroscopy. *J Nat Prod* **2014**, *77* (6), 1473-1487.
63. Fellenberg, M.; Behnken, H. N.; Nagel, T.; Wiegandt, A.; Baerenfaenger, M.; Meyer, B., Glycan analysis: scope and limitations of different techniques-a case for integrated use of LC-MS(/MS) and NMR techniques. *Analytical and Bioanalytical Chemistry* **2013**, *405* (23), 7291-7305.
64. Mulloy, B.; Dell, A.; Stanley, P.; J, H. P., Structural Analysis of Glycans. In *Essentials of Glycobiology*, rd; 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 (NY), 2015; pp 639-652.
65. D'Atri, V.; Causon, T.; Hernandez-Alba, O.; Mutabazi, A.; Veuthey, J. L.; Cianferani, S.; Guillarme, D., Adding a new separation dimension to MS and LC-MS: What is the utility of ion mobility spectrometry? *J Sep Sci* **2018**, *41* (1), 20-67.
66. Mu, Y.; Schulz, B. L.; Ferro, V., Applications of Ion Mobility-Mass Spectrometry in Carbohydrate Chemistry and Glycobiology. *Molecules* **2018**, *23* (10).
67. Dodds, J. N.; Baker, E. S., Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead. *J Am Soc Mass Spectrom* **2019**, *30* (11), 2185-2195.

68. Gray, C. J.; Thomas, B.; Upton, R.; Migas, L. G.; Eyers, C. E.; Barran, P. E.; Flitsch, S. L., Applications of ion mobility mass spectrometry for high throughput, high resolution glycan analysis. *Biochim Biophys Acta* **2016**, *1860* (8), 1688-709.
69. Mason, E. A.; Schamp, H. W., Mobility of Gaseous Ions in Weak Electric Fields. *Ann Phys-New York* **1958**, *4* (3), 233-270.
70. Plasencia, M. D.; Isailovic, D.; Merenbloom, S. I.; Mechref, Y.; Novotny, M. V.; Clemmer, D. E., Resolving and assigning N-linked glycan structural isomers from ovalbumin by IMS-MS. *J Am Soc Mass Spectrom* **2008**, *19* (11), 1706-15.
71. 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), 2158-66.
72. 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 Spectr* **2017**, *28* (10), 2217-2222.
73. 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.
74. Deng, L.; Ibrahim, Y. M.; Hamid, A. M.; Garimella, S. V.; Webb, I. K.; Zheng, X.; Prost, S. A.; Sandoval, J. A.; Norheim, R. V.; Anderson, G. A.; Tolmachev, A. V.; Baker, E. S.; 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), 8957-64.
75. 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), 2196-205.
76. Hofmann, J.; Struwe, W. B.; Scarff, C. A.; Scrivens, J. H.; Harvey, D. J.; Pagel, K., Estimating collision cross sections of negatively charged N-glycans using traveling wave ion mobility-mass spectrometry. *Anal Chem* **2014**, *86* (21), 10789-95.
77. Harvey, D. J.; Scarff, C. A.; Edgeworth, M.; Crispin, M.; Scanlan, C. N.; Sobott, F.; Allman, S.; Baruah, K.; Pritchard, L.; Scrivens, J. H., Travelling wave ion mobility and negative ion fragmentation for the structural determination of N-linked glycans. *Electrophoresis* **2013**, *34* (16), 2368-78.

78. Fernandez-Lima, F.; Kaplan, D. A.; Suetering, J.; Park, M. A., Gas-phase separation using a trapped ion mobility spectrometer. *Int J Ion Mobil Spectrom* **2011**, *14* (2-3).
79. 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.
80. 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), 3440-3.
81. Kolakowski, B. M.; Mester, Z., Review of applications of high-field asymmetric waveform ion mobility spectrometry (FAIMS) and differential mobility spectrometry (DMS). *Analyst* **2007**, *132* (9), 842-64.
82. Gabryelski, W.; Froese, K. L., Rapid and sensitive differentiation of anomers, linkage, and position isomers of disaccharides using High-Field Asymmetric Waveform Ion Mobility Spectrometry (FAIMS). *J Am Soc Mass Spectr* **2003**, *14* (3), 265-277.
83. Pathak, P.; Baird, M. A.; Shvartsburg, A. A., High-Resolution Ion Mobility Separations of Isomeric Glycoforms with Variations on the Peptide and Glycan Levels. *J Am Soc Mass Spectr* **2020**, *31* (7), 1603-1609.
84. Schindler, B.; Barnes, L.; Renois, G.; Gray, C.; Chambert, S.; Fort, S.; Flitsch, S.; Loison, C.; Allouche, A. R.; Compagnon, I., Anomeric memory of the glycosidic bond upon fragmentation and its consequences for carbohydrate sequencing. *Nat Commun* **2017**, *8* (1), 973.
85. Schindler, B.; Barnes, L.; Gray, C. J.; Chambert, S.; Flitsch, S. L.; Oomens, J.; Daniel, R.; Allouche, A. R.; Compagnon, I., IRMPD Spectroscopy Sheds New (Infrared) Light on the Sulfate Pattern of Carbohydrates. *J Phys Chem A* **2017**, *121* (10), 2114-2120.
86. Tan, Y.; Zhao, N.; Liu, J.; Li, P.; Stedwell, C. N.; Yu, L.; Polfer, N. C., Vibrational Signatures of Isomeric Lithiated N-acetyl-D-hexosamines by Gas-Phase Infrared Multiple-Photon Dissociation (IRMPD) Spectroscopy. *J Am Soc Mass Spectrom* **2017**, *28* (3), 539-550.
87. Boyarkin, O. V., Cold ion spectroscopy for structural identifications of biomolecules. *Int Rev Phys Chem* **2018**, *37* (3-4), 559-606.
88. Mucha, E.; Gonzalez Florez, A. I.; Marianski, M.; Thomas, D. A.; Hoffmann, W.; Struwe, W. B.; Hahm, H. S.; Gewinner, S.; Schollkopf, W.; Seeberger, P. H.; von Helden, G.; Pagel, K., Glycan Fingerprinting via Cold-Ion Infrared Spectroscopy. *Angew Chem Int Ed Engl* **2017**, *56* (37), 11248-11251.

89. 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 H<sub>2</sub>-tagged, mass-selected ions. *J Am Chem Soc* **2011**, *133* (16), 6440-8.
90. 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.
91. Voss, J. M.; Kregel, S. J.; Fischer, K. C.; Garand, E., IR-IR Conformation Specific Spectroscopy of Na<sup>(+)</sup>(Glucose) Adducts. *J Am Soc Mass Spectrom* **2018**, *29* (1), 42-50.
92. Scutelnic, V.; Rizzo, T. R., Cryogenic Ion Spectroscopy for Identification of Monosaccharide Anomers. *J Phys Chem A* **2019**, *123* (13), 2815-2819.
93. Peng, W.; Gutierrez Reyes, C. D.; Gautam, S.; Yu, A.; Cho, B. G.; Goli, M.; Donohoo, K.; Mondello, S.; Kobeissy, F.; Mechref, Y., MS-based glycomics and glycoproteomics methods enabling isomeric characterization. *Mass Spectrom Rev* **2021**.
94. Haselberg, R.; de Jong, G. J.; Somsen, G. W., CE-MS for the analysis of intact proteins 2010-2012. *Electrophoresis* **2013**, *34* (1), 99-112.
95. Stolz, A.; Jooss, K.; Hocker, O.; Romer, J.; Schlecht, J.; Neuss, C., Recent advances in capillary electrophoresis-mass spectrometry: Instrumentation, methodology and applications. *Electrophoresis* **2019**, *40* (1), 79-112.
96. Hernandez, O.; Isenberg, S.; Steinmetz, V.; Glish, G. L.; Maitre, P., Probing Mobility-Selected Saccharide Isomers: Selective Ion-Molecule Reactions and Wavelength-Specific IR Activation. *J Phys Chem A* **2015**, *119* (23), 6057-64.
97. 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.
98. Dyukova, I.; Ben Faleh, A.; Warnke, S.; Yalovenko, N.; Yatsyna, V.; Bansal, P.; Rizzo, T. R., A new approach for identifying positional isomers of glycans cleaved from monoclonal antibodies. *Analyst* **2021**, *146* (15), 4789-4795.
99. Ben Faleh, A.; Warnke, S.; Bansal, P.; Pellegrinelli, R. P.; Dyukova, I.; Rizzo, T. R., Identification of Mobility-Resolved N-Glycan Isomers. *Anal Chem* **2022**, *94* (28), 10101-10108.





## Chapter 2. Experimental approach

The first part of this chapter introduces three generations of home-built instruments on which we performed all our experiments. It begins with an overview of a first-generation tandem mass spectrometer for cryogenic ion spectroscopy (CIS). We then discuss a second-generation instrument that incorporates a SLIM-based ion mobility module (referred to the IMS-CIS I) and a third-generation instrument that is an upgraded version of IMS-CIS I to overcome some of its limitations (referred to the IMS-CIS II) (Figure 2.1).

The second part of this chapter describes in detail and compares the high resolution SLIM devices of the IMS-CIS I and II instruments. It also explains the principles of cryogenic messenger-tagging spectroscopy that was implemented to obtain vibrational fingerprints of glycans. The last section of this chapter presents glycan analysis with ultra-performance liquid chromatography (UPLC) with a HILIC column coupled to a Q-TOF mass spectrometer that we used in some experiments.

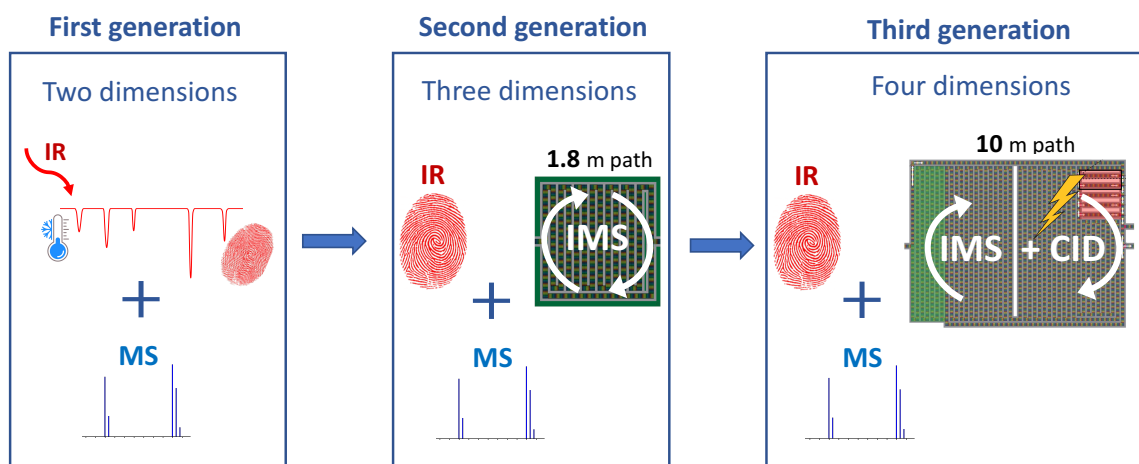


Figure 2.1. Overview of the three generations of instruments used in this work.

## 2.1. Experimental setup

### 2.1.1. Generation I: the cold ion spectroscopy (CIS) instrument

A home-built tandem mass spectrometer coupled to a cryogenic ion trap was used to perform the experiments discussed in the first part of Chapter 3.

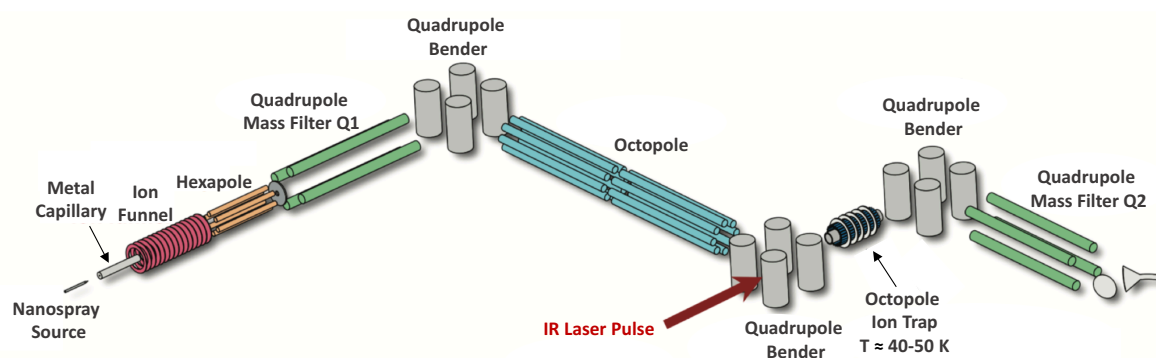


Figure 2.2. Scheme of the first-generation cryogenic home-built tandem mass spectrometer.

Figure 2.2 shows a schematic overview of the instrument.<sup>1</sup> We produce cationized glycan ions *via* a nanoelectrospray ionization (nESI) source from a 50:50 acetonitrile-water solution. The ions are transferred into the instrument through a metal capillary in a continuous fashion and then focused in an ion funnel. The ion funnel is equipped with a jet disruptor that prevents the propagation of large droplets from the ESI into the next stages of the setup.<sup>2, 3</sup> The ions are then pre-trapped in a hexapole ion guide and released as short pulses matching the 10 Hz duty cycle of the laser. Accumulation in the hexapole region reduces the fluctuations of the ion signal due to the instability of the electrospray. The resulting ion packets then pass through a quadrupole mass filter (Q1), which selects those with a particular  $m/z$ . The transmitted ions are then turned by  $90^\circ$  using a static quadrupole deflector, guided through an octopole, and deflected an additional  $90^\circ$  before passing through a set of decelerating lenses. After the set of lenses, the ions enter a cold octopole ion trap that is maintained at 40–50 K by a closed cycle helium cryostat.<sup>4</sup> Before the ion packet arrives in the trap, we inject a pulse of helium and nitrogen gases (90:10 mixture), which cools in collisions with the trap wall. Once in the trap, the ions collide with this cold gas mixture, cooling them and forming weakly-bound complexes with nitrogen, which we use to perform messenger-tagging spectroscopy.

The nitrogen-tagged ions are then irradiated with infrared pulses generated from a Nd:YAG pumped, tunable optical parametric oscillator (LaserVison). When the OPO wavenumber is tuned in resonance with a vibrational transition, IR radiation is absorbed and then internally

redistributed, causing the weakly bound  $N_2$  molecule to evaporate. Monitoring the depletion of the tagged ions as a function of the IR wavenumber allows us to obtain an infrared absorption spectrum. All the ions are then ejected from the trap, bent  $90^\circ$  one more time by the third quadrupole bender to be analyzed by a second quadrupole mass filter (Q2) and detected by a channeltron detector.

### 2.1.2. Generation II: ion mobility spectrometry coupled to cold ion spectroscopy (IMS-CIS I instrument)

A home-built instrument, which we refer to as IMS-CIS I, couples ion mobility and cold ion spectroscopy. This instrument was used to perform the experiments discussed in the first part of Chapter 4.

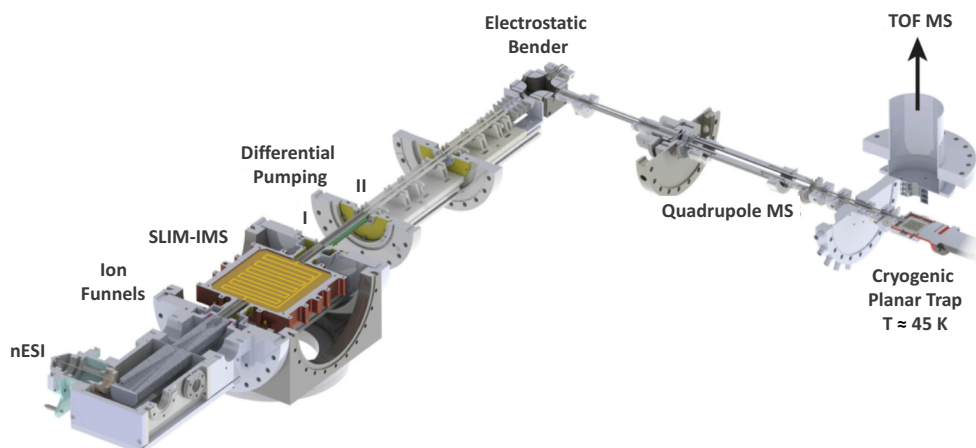


Figure 2.3. Schematic overview of the second-generation instrument combining ultra-high resolution SLIM-based IMS with cryogenic IR spectroscopy.

Figure 2.3 shows a schematic overview of the IMS-CIS I instrument.<sup>5</sup> The major differences between this machine and CIS instrument include the addition of a traveling-wave ion mobility device based on structures for lossless ion manipulations (SLIM) that was appended to an existing drift tube instrument.<sup>6</sup> Moreover, the octopole ion trap is replaced by a planar trap, mass analysis is done by TOF rather than a quadrupole and the OPO system was replaced with a tunable continuous-wave mid-IR laser.

Doubly sodiated glycan ions are produced *via* nESI and transferred into the instrument through a stainless-steel capillary heated to  $160\text{--}170^\circ\text{C}$ . A dual-stage ion funnel trap (IFT) assembly (MassTech, United States) focuses and stores ions before releasing them in short packets ( $\sim 150\ \mu\text{s}$ ) into the SLIM ion-mobility region (described in detail below). In the SLIM-

IMS device (1.8 m single-cycle path length), which is filled with helium drift gas at 3 mbar, ions are separated based on their respective collision cross section. They are then directed through differential pumping stages using ring-electrode ion guides and hexapoles. A set of steering electrodes mounted prior to the electrostatic ion bender gives the possibility to selectively send ions of a particular drift time for further investigation and electrostatically deflect unwanted ions. After the  $m/z$  selection in a quadrupole mass filter, the ions enter the planar cryogenic ion trap (45 K), where they are stored and cooled through collisions with a He : N<sub>2</sub> mixture (90:10).<sup>7</sup> The nitrogen-tagged ions are then spectroscopically interrogated for 50 ms by a continuous-wave, mid-IR laser (IPG Photonics) operated at 0.2 W output power. An IR spectrum is generated by monitoring the wavelength-dependent ion signal of the N<sub>2</sub>-tagged species in the TOF mass spectrum.

### 2.1.3. Generation III: ion mobility spectrometry coupled to cold ion spectroscopy version II (IMS-CIS II instrument)

The IMS-CIS II instrument was designed as an upgrade to the IMS-CIS I instrument to increase the IMS resolving power and introduce the possibility of multiple-stage IMS. A schematic overview of the instrument is shown in Figure 2.4. It was used to conduct the experiments described in the second parts of Chapters 3 and 4, as well as all the experiments in Chapters 5 and 6.<sup>8</sup>

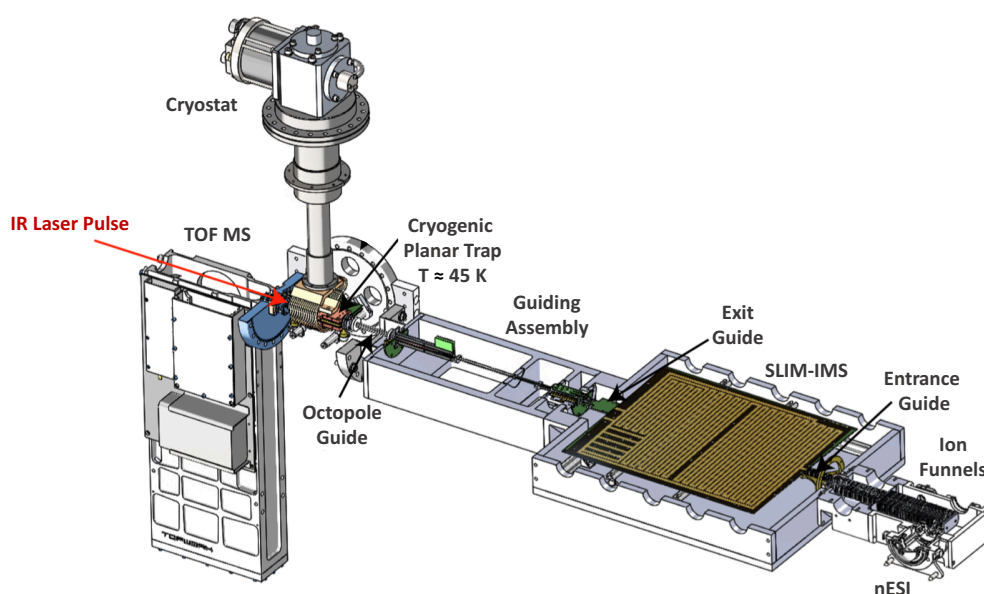


Figure 2.4. Schematic overview of the third-generation instrument combining ultra-high resolution SLIM IMS with cryogenic IR spectroscopy.

Ions are generated in a nESI source, transferred through a stainless-steel capillary into the instrument and guided *via* a dual ion funnel assembly (MassTech, USA) toward the SLIM-IMS module (10 m single-cycle path length) filled with nitrogen drift gas with a pressure of 2.2 mbar. We then accumulate ions in a 2 m storage section of the SLIM device and release them in pulses from hundreds of  $\mu\text{s}$  to 10 ms in duration into the separation region where ions follow a serpentine path. After mobility separation, ions are sent through differential pumping stages into a cryogenic ion trap maintained at 45 K to perform messenger-tagging spectroscopy.

## 2.2. SLIM-based ion mobility

### 2.2.1. Basic principles of the SLIM platform

Recently R. Smith and co-workers introduced the Structures for Lossless Ion Manipulations (SLIM) technology to perform high resolution IMS.<sup>9-11</sup> SLIM is a type of traveling-wave ion mobility that utilizes two mirrored printed circuit boards (PCBs) to create a path in which the ions move (see Figure 2.5). The SLIM module employs three different types of electrodes: guard electrodes, radio frequency (RF) electrodes and traveling-wave (TW) electrodes. Each board consists of six RF and five TW tracks (Figure 2.5b). The DC guard electrodes confine the ions laterally (i.e. the x-direction in Figure 2.5a). The RF electrodes are located between the guard electrodes and serve to confine ions between the two planar surfaces (i.e., the y-direction in Figure 2.5a). The traveling wave electrodes have DC potentials applied by them in such a way as to produce a square wave – in this case by a repeating sequence of eight DC pads over the length of the module, on both sides. The TW sequence can be defined as 11110000, meaning that at each period of time a high DC potential is applied to four individual pads while the other four stay at a low potential. The square wave “travels” by propagating the set of four “high” electrodes one pad at a time (i.e., 11110000, 01111000, 00111100, 00011110, etc.). The TW drives the ion motion and enables mobility separation of the molecules.

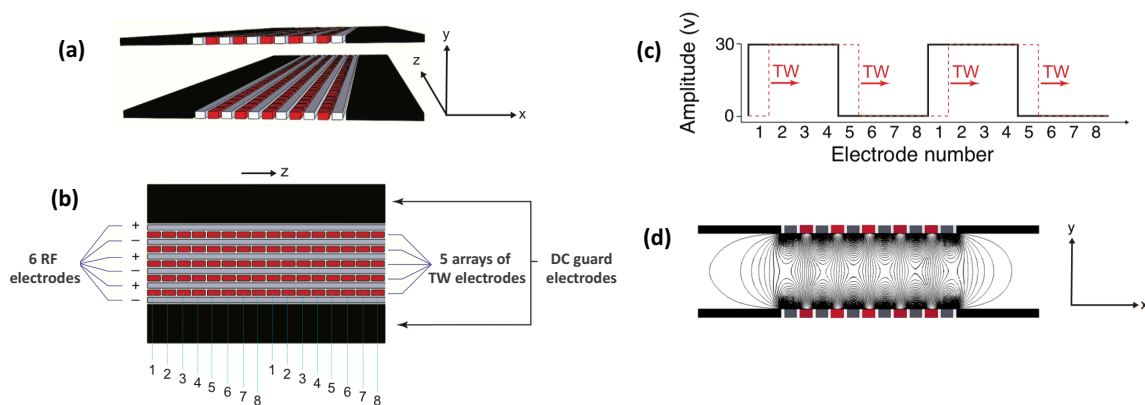


Figure 2.5. Scheme of the main parts of a SLIM module. (a) Two PC boards that form the separation path; (b) the RF, traveling wave and guard electrodes; (c) the traveling wave potential (d) the field created by the RF electrodes in the x-y plane. (Adapted from Fig. 1 of Ref. 12)

The arrangement of the electrodes allows two key features of the SLIM-board (illustrated in Figure 2.6b, c). The possibility of 90° turns in the ion path without ion loss over a wide  $m/z$  range and degradation in the resolution permits extremely long path lengths in a small footprint. In addition, T-shaped switches provide flexibility in ion manipulation, allowing one to divert them from their separation path. These can be used, for example, to eject ions from the SLIM module or to divert them into on-board traps. One can also use them to direct ions along the serpentine path multiple times to achieve higher resolution. A great advantage of these PCB-based SLIM platforms is that new designs can be rapidly and easily fabricated at a low cost.

The initial design of the SLIM module created in the Smith's group had a size of 12" x 18" and consisted of linear paths connected via 44 turns to create a longer path of ~13 m. The performance of the short linear path (44 cm) was comparable to the separation performed on a 1 m drift tube for two reverse peptides (GAGAS and SAGAG). However, a 13 m long TW SLIM device was able to achieve baseline separation. The difference between DT and TW SLIM devices was more significant in the separation of various pentasaccharide isomers.<sup>13</sup> While a 1 m drift tube was able to distinguish only three conformers, the 13 m long path length identified at least five conformers. To evaluate the resolving power in the SLIM SUPER IM mode separations, the Agilent tuning mix was subjected to 40 passes (~540 m) and a resolution of ~400 was achieved. The corresponding resolving power was ~1870.<sup>9</sup>

## 2.2.2. The ion mobility module of the IMS-CIS I instrument

To perform ultra-high resolution ion mobility separation, we implemented SLIM-based traveling-wave ion mobility spectrometry into our second-generation instrument IMS-CIS I. We used two printed circuit boards of size  $15 \times 15$  cm placed 2.75 mm apart to obtain a single-pass serpentine ion path length of 1.8 m (Figure 2.6a). An on-board T-switch is used either to send ions toward the cryogenic ion trap or to cycle them multiple times over the same serpentine path to increase the resolving power (Figure 2.6c).

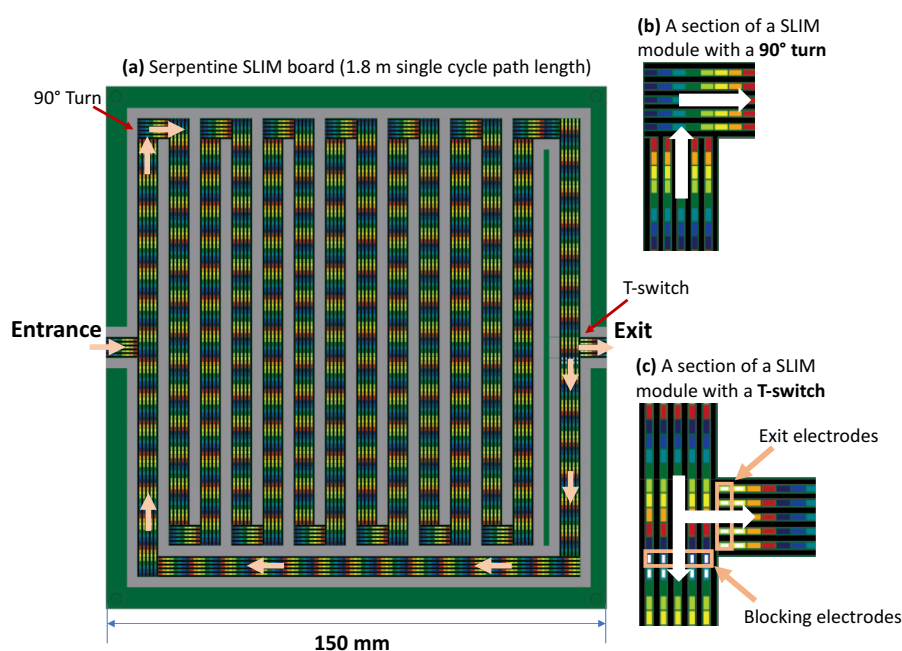


Figure 2.6. (a) Schematic view of the SLIM board with a 1.8 m serpentine path incorporated into the IMS-CIS I instrument; (b) detailed layout of an on-board section with a  $90^\circ$  turn that enables a serpentine path; (c) detailed layout of an on-board T-switch that enables cycling IMS. (Adapted from Fig. 3.9 of Ref. 14)

We measured the resolving power of this device on the  $[M+2H]^{2+}$  ions of the ion mobility standard reverse-sequence pair of peptides GRGDS/SDGRG.<sup>5</sup> Figure 2.7 shows the resulting resolving power  $R$  as a function of the number of cycles  $n$ . It reaches approximately 500 after 28 cycles and can be increased by further cycling. However, one must take into account the fact that the separated packets of ions with high mobility will start overlapping with less mobile ions at the highest values of  $n$ . The instrument demonstrated a loss of ions of up to 3% per cycle.



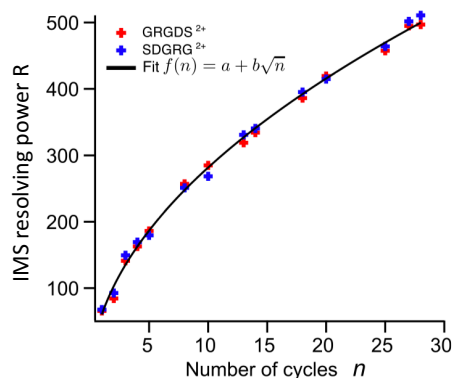


Figure 2.7. Resolving power of the SLIM board with a 1,8 m (IMS-CIS II) single cycle path length.

### 2.2.3. The ion mobility module of the IMS-CIS II instrument

While the electrode design is the same as in the SLIM module of the IMS-CIT I instrument, the ion mobility device in the third-generation apparatus includes several new features (Figure 2.8). The larger size of the board (50 cm x 40 cm) enables a single-path of 10 m, which allows to perform an increased number of separation cycles and offers a much higher peak capacity, preventing ions from overlapping in the separation path.

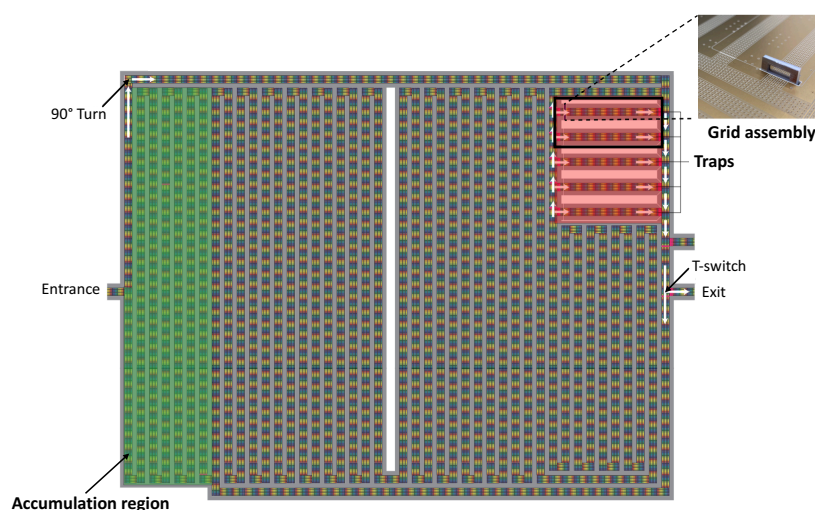


Figure 2.8. Schematic view of the SLIM board with a 10 m serpentine path incorporated in the IMS-CIS II instrument. The board offers a 2 m accumulation region (green section) and 5 traps (red section). One of the traps also includes a CID grid assembly. (Adapted from Fig. 7.5 of Ref. 14)

Figure 2.9 shows the resolving power of this board as a function of the drift length using a pair of reverse sequence peptides GRGDS/SDGRG in their singly sodiated form.<sup>8</sup> A single separation cycle reaches a resolving power of approximately 200, and almost 1000 after 200 m or 20 separation cycles.

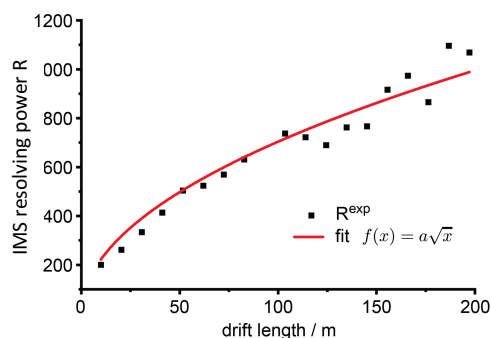


Figure 2.9. Resolving power of the SLIM board with a 10 m (IMS-CIS II) single cycle path length.

In addition, the device contains a 2 m accumulation region that is constantly filled from the ion source. The high-volume of the accumulation region allows for an increased ion utilization efficiency compared to previous implementations that were using ion funnel traps for storage prior to the mobility separation. Five on-board trapping regions based on the design of the SLIM T-switches can be used for intermediate ion storage and enrichment of low abundance analyte ions. Two of the five SLIM traps (Figure 2.8) offer the possibility to perform fragmentation via collision-induced dissociation. One of these CID traps also includes a grid assembly that consists of two metallic grids separated by a Teflon spacer of 0.8 mm. After being separated by mobility along the serpentine path, parent ions are extracted into the trapping region, which is held at a lower bias voltage than in the rest of the IMS device (a typical voltage step is  $\sim 100$  V). Passing through the grid system at the trap entrance, ions experience a homogenous electric field of up to 3300 V/cm, which induces them to collide with  $N_2$  molecules, leading to their dissociation. After the bias voltage of the trapping region is raised to the level of the separation region, the resulting fragments can exit the trap and follow the serpentine path to either exit the SLIM device or undergo additional mobility separation before being directed to the cryogenic ion trap, hence allowing for the acquisition of IR spectra of mobility separated fragments.

## 2.3. Cryogenic messenger-tagging spectroscopy

In all three generations of instruments, we use the cryogenic messenger-tagging spectroscopy technique to acquire the IR spectra of molecules. This method was originally proposed by Y.T. Lee's group in 1980's,<sup>15,16</sup> then it was further developed and employed by M. A. Johnson and coworkers.<sup>17,18</sup> Messenger-tagging spectroscopy is based on the following principle: the ion of interest forms a weakly bound cluster with one or more neutral molecules (messengers or tags) through van der Waals interactions at sufficiently low temperatures. When the laser wavenumber is tuned in resonance with a vibrational transition, a resonant photon is absorbed and then its energy redistributed *via* IVR, causing tag molecules to evaporate. Monitoring the depletion of tagged ions as a function of the IR wavenumber allows one to obtain an infrared absorption spectrum (Figure 2.10).

In the gas phase, two molecules can stick upon colliding only if their binding energy finds a way to be dissipated. Usually, clusters are formed as a result of a three-body collision.<sup>19</sup> An excited two-body cluster of the ion of interest with a tag molecule collides with a second tag molecule to dissipate the excess energy and form a stable complex.

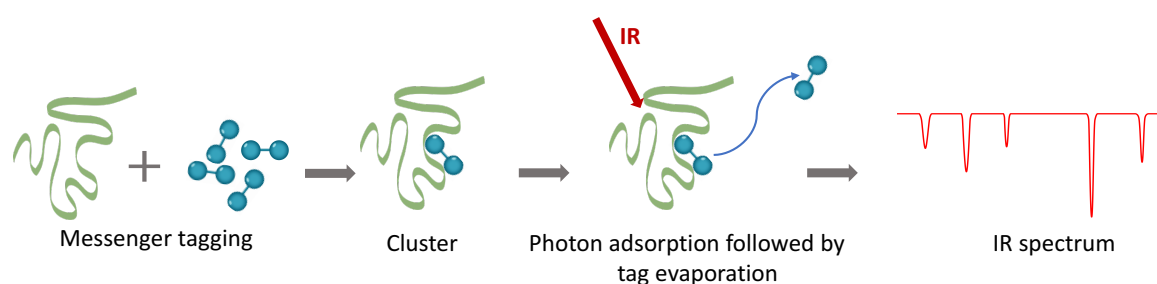


Figure 2.10. Schematic overview of messenger-tagging spectroscopy.

To be applicable to this type of experiment, a tag molecule must fulfill the following requirements:

- 1) The tag molecule should not have any absorption in the IR region. Monoatomic and homonuclear diatomic molecules fulfil this requirement, and thus H<sub>2</sub>, D<sub>2</sub>, He, N<sub>2</sub>, Ne and Ar have all been used in the past as tag molecules.<sup>20-25</sup>
- 2) The sticking efficiency of the tag molecule should be high enough to obtain a detectable number of complexes with a parent molecule. Heavier tags such as Ar, Ne and N<sub>2</sub> condense more easily and at higher temperatures than the light ones due to their higher polarizability. On the other hand, if one aims to obtain information about the 3D

structure of a molecule, it is preferable to work with H<sub>2</sub>, D<sub>2</sub>, He as it was demonstrated in many studies that Ar, Ne and N<sub>2</sub> can significantly perturb the vibrational spectrum of a bare molecule.

- 3) The mass shift between complexes and a bare ion should be easily detectable. In this regard, H<sub>2</sub> and D<sub>2</sub> reveal a disadvantage, as even for singly charged ions their mass shifts (H<sub>2</sub> (2 Da), D<sub>2</sub> (4 Da)) are relatively small and they become even smaller for multiply charged ions. While modern mass spectrometers have no problem resolving the peaks of the tagged species, for large molecules peaks of the tagged species will overlap with the isotope peaks of the parent species.

In our approach, we use the IR spectra as unique glycan identifiers. We do not aim to unravel their 3D structures based on the analysis of the IR spectrum but to identify primary structures using the comparison of IR fingerprints with a database. For this reason, requirement 2 is not an issue using larger tag molecules, as long as the database is constructed with the same tag.

Regarding requirement 3, N-linked glycans have high masses and are present in the gas phase mostly in multiply charged states. We observed this from the N-linked glycan core structure, Man-3 (910 Da), which was studied in a doubly charged state, up to the glycan A4F (3681 Da), the biggest glycan in this work, which was studied in a quadruply charged state. For these reasons, N<sub>2</sub> was chosen as the optimal compromise for a tag molecule in our experiments. In addition, N<sub>2</sub> tags condense into an ion of interest at ~ 45 K, which makes them more practical as an analytical tool compared to He, which requires cooling to 4 K to produce a sufficient number of complexes.

We obtain the IR spectrum by plotting the number of clusters divided by the total number of clusters and bare ions as follows (Figure 2.11):

$$IR\ spectrum = \frac{\sum clusters}{\sum (clusters + bare\ ions)}$$

After the absorption of a photon, the number of clusters decreases while the number of bare ions increases by the same amount. The cluster depletion is normalized to take into account a possible ion signal fluctuation during the recording of the spectrum, mostly coming from electrospray instability.

The infrared radiation for our measurements was generated by two laser systems. All the experiments with the first-generation CIS instrument were performed using an OPO/OPA (Laser Vision) system pumped by a Nd:YAG (Innolas) laser system, while the second- and

third-generation instruments (IMS-CIS I / II) use a continuous-wave, mid-IR laser (IPG Photonics).

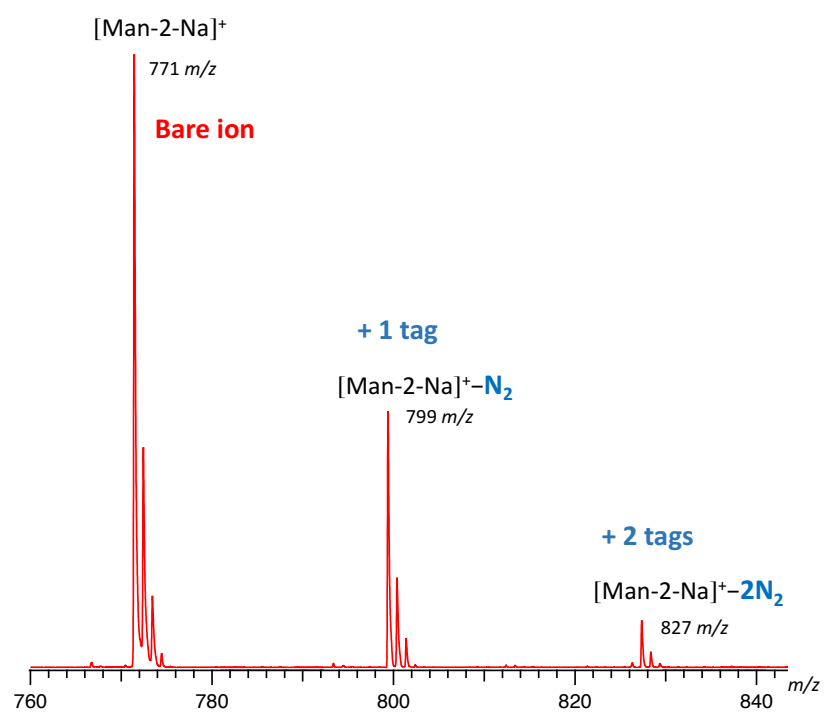


Figure 2.11. Typical recorded mass spectrum of parent and tagged species using the example of glycan Man-2.

## 2.4. Ultra-performance liquid chromatography (UPLC) coupled to a Q-TOF mass spectrometer

In this work, we developed a new approach for glycan analysis that combines enzymatic digestion and synthesis with our IMS-MS-IR technique. Before being introduced into the instruments, glycan samples were analyzed and purified using a commercial LC-MS setup. This sample preparation step was performed on the ACQUITY UPLC H-Class Plus System (Waters) coupled to a Q-TOF mass spectrometer (Waters Premier) (Figure 2.12). We used an HILIC XBridge Glycan BEH Amide Column (130 Å, 3.5 μm, 4.6 × 150 mm) from Waters (kept at 60 °C) with a bisolvent system as a mobile phase – acetonitrile and ammonium formate buffer solution (100 mM, pH = 4–5; flow rate - 0.4 ml/min). For some applications, glycan fractions were collected using a Waters Fraction Collector III coupled to the UPLC-MS system and the solvents were evaporated to concentrate the samples up to 15–20 μM final concentration, calculated assuming a 100% elution of the samples from the column. To be able to perform the detection and collection at the same time, we installed a T-shape splitter right after the column to divide the eluate into two parts: 20 % of the flow is directed to the Q-TOF mass detector and 80% to the fraction collector.

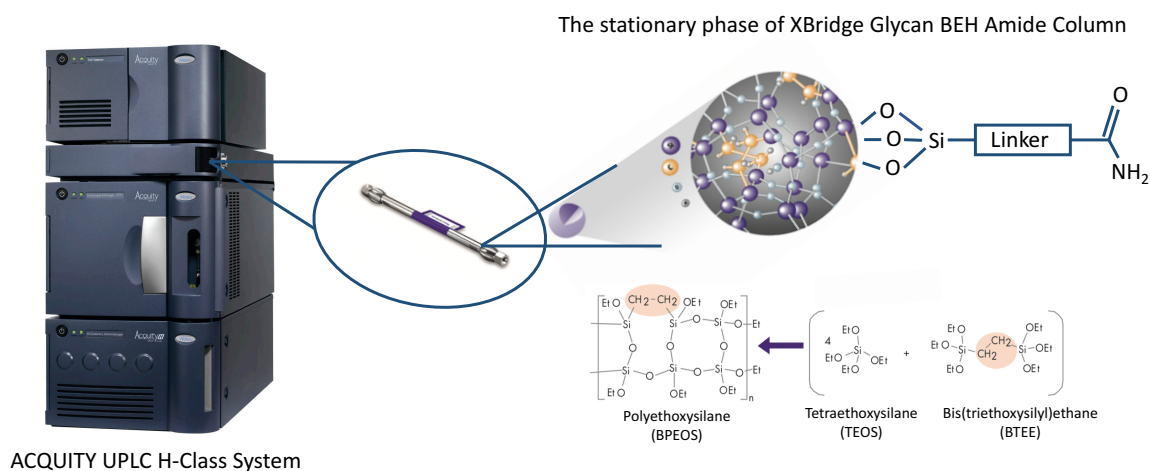


Figure 2.12. ACQUITY UPLC H-Class Plus System and an XBridge Glycan BEH Amide Column with detailed representation of the mobile phase. (<https://www.waters.com/nextgen/us/en.html>).

## References

1. 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.
2. Kim, T.; Tolmachev, A. V.; Harkewicz, R.; Prior, D. C.; Anderson, G.; Udseth, H. R.; Smith, R. D., Design and implementation of a new electrodynamic ion funnel. *Anal Chem* **2000**, *72* (10), 2247-55.
3. Kim, T.; Tang, K. Q.; Udseth, H. R.; Smith, R. D., A multicapillary inlet jet disruption electrodynamic ion funnel interface for improved sensitivity using atmospheric pressure ion sources. *Analytical Chemistry* **2001**, *73* (17), 4162-4170.
4. Boyarkin, O. V.; Kopysov, V., Cryogenically cooled octupole ion trap for spectroscopy of biomolecular ions. *Review of Scientific Instruments* **2014**, *85* (3).
5. Ben Faleh, A.; Warnke, S.; Rizzo, T. R., Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures. *Analytical Chemistry* **2019**, *91* (7), 4876-4882.
6. 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<sup>+</sup>-Gly-Pro-Gly-Gly (GPGG). *J Am Soc Mass Spectr* **2015**, *26* (9), 1444-1454.
7. Lorenz, U. J.; Rizzo, T. R., Planar Multipole Ion Trap/Time-of-Flight Mass Spectrometer. *Analytical Chemistry* **2011**, *83* (20), 7895-7901.
8. Warnke, S.; Ben Faleh, A.; Rizzo, T. R., Toward High-Throughput Cryogenic IR Fingerprinting of Mobility-Separated Glycan Isomers. *ACS Meas Sci Au* **2021**, *1* (3), 157-164.
9. Deng, L. L.; Webb, I. K.; Garimella, S. V. B.; Hamid, A. M.; Zheng, X. Y.; Norheim, R. V.; Prost, S. A.; Anderson, G. A.; Sandoval, J. A.; Baker, E. S.; Ibrahim, Y. M.; Smith, R. D., Serpentine Ultralong Path with Extended Routing (SUPER) High Resolution Traveling Wave Ion Mobility-MS using Structures for Lossless Ion Manipulations. *Analytical Chemistry* **2017**, *89* (8), 4628-4634.
10. Hamid, A. M.; Garimella, S. V. B.; Ibrahim, Y. M.; Deng, L. L.; Zheng, X. Y.; 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. *Analytical Chemistry* **2016**, *88* (18), 8949-8956.

11. Deng, L. L.; Ibrahim, Y. M.; Hamid, A. M.; Garimella, S. V. B.; Webb, I. K.; Zheng, X. Y.; 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. *Analytical Chemistry* **2016**, *88* (18), 8957-8964.
12. Hamid, A. M.; Ibrahim, Y. M.; Garimella, S. V.; 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), 11301-8.
13. Deng, L. L.; Ibrahim, Y. M.; Baker, E. S.; Aly, N. A.; Hamid, A. M.; Zhang, X.; Zheng, X. Y.; Garimella, S. V. B.; Webb, I. K.; Prost, S. A.; Sandoval, J. A.; Norheim, R. V.; Anderson, G. A.; Tolmachev, A. V.; Smith, R. D., Ion Mobility Separations of Isomers based upon Long Path Length Structures for Lossless Ion Manipulations Combined with Mass Spectrometry. *Chemistryselect* **2016**, *1* (10), 2396-2399.
14. Ben Faleh, A. A gas phase approach for glycan analysis: Combining ultrahigh-resolution ion mobility spectrometry with cryogenic vibrational spectroscopy. EPFL, 2021.
15. Okumura, M.; Yeh, L. I.; Myers, J. D.; Lee, Y. T., Infrared-Spectra of the Cluster Ions  $H_7O_3^+ \cdot H_2$  and  $H_9O_4^+ \cdot H_2$ . *J Chem Phys* **1986**, *85* (4), 2328-2329.
16. Lisy, J. M., Infrared studies of ionic clusters: The influence of Yuan T. Lee. *J Chem Phys* **2006**, *125* (13).
17. 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 H<sub>2</sub>-tagged, mass-selected ions. *J Am Chem Soc* **2011**, *133* (16), 6440-8.
18. Leavitt, C. M.; Wolk, A. B.; Kamrath, M. Z.; Garand, E.; Van Stipdonk, M. J.; Johnson, M. A., Characterizing the intramolecular H-bond and secondary structure in methylated GlyGlyH<sup>+</sup> with H<sub>2</sub> predissociation spectroscopy. *J Am Soc Mass Spectrom* **2011**, *22* (11), 1941-52.
19. Croft, J. F. E.; Bohn, J. L., Non-sticking of helium buffer gas to hydrocarbons. *Phys Rev A* **2015**, *91* (3).
20. DePalma, J. W.; Kelleher, P. J.; Tavares, L. C.; Johnson, M. A., Coordination-Dependent Spectroscopic Signatures of Divalent Metal Ion Binding to Carboxylate Head Groups: H<sub>2</sub>- and He-Tagged Vibrational Spectra of M<sub>2</sub><sup>+</sup>center dot RCO<sub>2</sub><sup>-</sup> (M = Mg and Ca, R = -CD<sub>3</sub>, -CD<sub>2</sub>CD<sub>3</sub>) Complexes. *J Phys Chem Lett* **2017**, *8* (2), 484-488.



21. Leavitt, C. M.; Wolk, A. B.; Fournier, J. A.; Kamrath, M. Z.; Garand, E.; Van Stipdonk, M. J.; Johnson, M. A., Isomer-Specific IR-IR Double. Resonance Spectroscopy of D-2-Tagged Protonated Dipeptides Prepared in a Cryogenic Ion Trap. *J Phys Chem Lett* **2012**, 3 (9), 1099-1105.
22. 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. *Analytical Chemistry* **2017**, 89 (14), 7601-7606.
23. 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 Spectr* **2017**, 28 (10), 2217-2222.
24. Brunken, S.; Lipparini, F.; Stoffels, A.; Jusko, P.; Redlich, B.; Gauss, J.; Schlemmer, S., Gas-Phase Vibrational Spectroscopy of the Hydrocarbon Cations l-C<sub>3</sub>H<sup>+</sup>, HC<sub>3</sub>H<sup>+</sup>, and c-C<sub>3</sub>H<sub>2</sub><sup>+</sup>: Structures, Isomers, and the Influence of Ne-Tagging. *J Phys Chem A* **2019**, 123 (37), 8053-8062.
25. Li, J. W.; Morita, M.; Takahashi, K.; Kuo, J. L., Features in Vibrational Spectra Induced by Ar-Tagging for H<sub>3</sub>O(+)Arm, m = 0-3. *J Phys Chem A* **2015**, 119 (44), 10887-92.

# **Chapter 3. Combining enzymatic cleavage and collision-induced dissociation with cryogenic IR spectroscopy for determining the primary structure of glycans**

## **3.1. Database approaches for the identification of glycans**

Given the biological relevance and intrinsic structural complexity of glycans, many different approaches have been employed for their analysis. In the past decades, with progress in separation and mass spectrometry-based techniques, the amount of glycan data has increased significantly. As a result, the development of databases and bioinformatics tools was greatly needed to provide all the available information on glycans in a user-friendly way. In this chapter, we discuss the main advantages and drawbacks of current databases and demonstrate the necessity of adding a spectroscopic dimension to classic techniques. Moreover, we show a new approach for determining the primary structure of unknown glycans based on the combination of cryogenic IR with two fragmentation techniques.

The first glycan database with public access was the Complex Carbohydrate Structural Database (CCSD), which is also known as the CarbBank.<sup>1</sup> This DB was constructed by the Complex Carbohydrate Research Center of the University of Georgia (USA) in the 1980s. It collected all published carbohydrate structures larger than disaccharides that were known at the time, together with citations and supplementary information, providing more than 40 000 entries. However, the funding needed for further maintenance of the CarbBank was discontinued in 1997. Since then, and inspired by this work, several different databases have been developed that incorporate the data from CCSD. Current major glycan databases are listed in Table 3.1. They provide data for glycan identification and characterization as well as biochemical information obtained using different approaches and techniques.

Table 3.1. Glycan databases

Abbreviation	Database name	Type of data
CFG	Consortium for Functional Glycomics	Data on glycan structures, glycan-binding proteins and glycosyltransferases. <sup>2</sup> <a href="http://www.functionalglycomics.org">http://www.functionalglycomics.org</a>
JCGGDB	Japan Consortium for Glycobiology and Glycotechnology	MS data, including tandem MS <sup>n</sup> spectra, glycoprotein and lectin array data. <sup>3</sup> <a href="http://jcgddb.jp/database_en.html">http://jcgddb.jp/database_en.html</a>
–	UniCarbKB (composed of GlycoSuiteDB, GlycoBase, and EUROCarbDB)	Glycosylation site information as well as MS, HPLC and NMR glycan data; incorporates bioinformatics tools. <sup>4</sup> <a href="http://www.unicarbkb.org">http://www.unicarbkb.org</a>
KEGG	Kyoto Encyclopedia of Genes and Genomes	Glycan biosynthesis and metabolism pathways. <sup>5</sup> <a href="https://www.genome.jp/kegg/">https://www.genome.jp/kegg/</a>
–	GLYCOSCIENCES.de	Glycan 3D structures; links glycomics and proteomics data. <sup>6</sup> <a href="http://www.glycosciences.de">http://www.glycosciences.de</a>
–	UniCarb-DB	Collection of LC–MS/MS glycan fragments released from glycoproteins. <sup>7</sup> <a href="https://unicarb-db.expasy.org">https://unicarb-db.expasy.org</a>
–	GlyTouCan	Provides a globally unique accession number to each glycan structure. <sup>8</sup> <a href="https://glytoucan.org">https://glytoucan.org</a>
–	GlycoStore	U/HPLC/PGC chromatography and capillary electrophoresis data. <sup>9</sup> <a href="https://www.glycostore.org">https://www.glycostore.org</a>
CSDB	Carbohydrate Structure DataBase	NMR data of carbohydrates in bacteria, archaea, fungi and plants. <sup>10</sup> <a href="http://csdb.glycoscience.ru/database/">http://csdb.glycoscience.ru/database/</a>
–	GlycoMob	Ion mobility-mass spectrometry data; glycan identification based on the mass and CCS value. <sup>11</sup> <a href="http://www.glycomob.org">http://www.glycomob.org</a>

Though current glycan DBs represent powerful tools for fundamental research in chemistry and molecular biology as well as for the pharmaceutical industry, they still have several drawbacks as bioinformatics services. The content of different DBs is partially overlapping and not standardized. Moreover, the DBs are not interconnected and have incompatible formats, which make it impossible to obtain an overview of all saved structures. Most importantly, there is still a big gap in the field of isomeric glycan structures. Even the

combination of such powerful techniques as IMS or LC with MS does not allow one to determine many of the subtle structural details that distinguish glycan isomers.

In the last few years, vibrational spectroscopy has been proven to provide unique structural fingerprints of isolated glycans. We aim to complement currently existing databases by adding a vibrational fingerprint for each glycan species. Though our method is based on spectroscopic measurements, the philosophy of our experiment is different from standard spectroscopic approaches, which determine ion structures based on the comparison of measured spectra with those of structures derived from high-level quantum chemical calculations.<sup>12-14</sup> Our approach is to separate isomeric glycans using ultra-high resolution ion mobility and record IR spectra of each mobility-selected peak. The database will include glycan mass and isomer-specific IR fingerprints (Figure 3.1). We will then identify an unknown glycan structure by measuring these parameters and comparing them with the database.

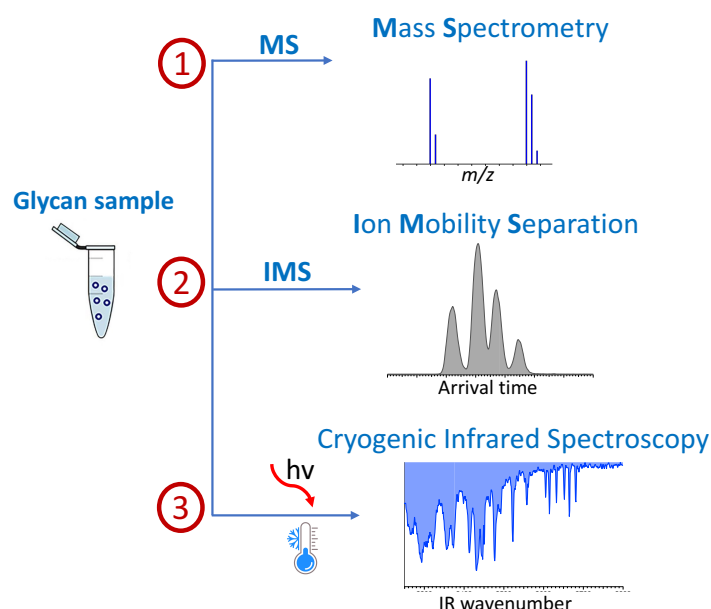


Figure 3.1. Scheme of the three-dimensional approach for glycan analysis developed in our laboratory.

It is worth noting that the quantum chemical calculations for determining glycan structure are extremely complicated for those consisting of more than a few monosaccharide units. This is mainly due to the challenges associated with determining precise band positions for large flexible molecules with many interacting chromophores, such as N-linked glycans.<sup>15</sup> On the other hand, performing spectroscopic measurements at cryogenic temperatures allows us to obtain highly structured glycan ion spectra that serve as unique identifiers even for large glycans. Even in cases where IMS is not able to separate glycan isomers fully and the recorded

spectrum is a superposition, we can still derive individual IR fingerprints using the spectral deconvolution method.<sup>16</sup>

One big advantage of using an IR spectrum as an identifying fingerprint is that it is an intrinsic property of the molecule, determined only by quantum mechanics. As such, it is a robust identifier that is insensitive to the experimental conditions, depending only relatively weakly on the temperature, making it easy to reproduce from one laboratory to the next. This removes the need for continual calibration, as is required when using LC or IMS for molecular identification. In this thesis, the comparison between measured spectra and those in the database are made visually, although eventually, quantitative metrics should be applied to estimate the level of certainty of the assignments.

While this spectroscopic database will be constructed initially using standard/known glycans, one needs a mechanism to associate a vibrational fingerprint with glycan structures that are presently unknown or those for which standards are not easily obtained. For the identification of unknown glycans, we propose a bottom-up approach using cryogenic spectroscopy to identify specifically cleaved glycan fragments. Glycan structures can be obtained from their fragments using two approaches: *enzymatic digestion* and *collision-induced dissociation*. The aim of this work is to evaluate the feasibility of combining exoglycosidase digestion and CID techniques with cryogenic IR-spectroscopy as a new tool for glycan identification.

## 3.2. Approaches for the identification of unknown glycans

### 3.2.1. Selective enzymatic digestion

The enzymatic digestion approach allows the identification of complex glycans through a series of bond-specific cleavages. Unknown glycans can be sequenced by a set of exoglycosidases that selectively cleave monosaccharides from the non-reducing glycan end.<sup>17</sup> This class of enzymes is highly specific to the monosaccharide type, the anomeric configuration ( $\alpha/\beta$  anomer) and the linkage position. Some of them are also specific to the saccharide branching. The mechanism of the cleavage is based on the hydrolysis of the glycosidic bond. In acidic media, protonation of the glycosidic O atom leads to the cleavage of the glycosidic bond between two monosaccharides, forming an anomeric carbocation, which

then reacts with water to yield a mixture of  $\alpha,\beta$ -anomers of the removed monosaccharide from the non-reducing end and the smaller intact glycan (Figure 3.2).<sup>18</sup>

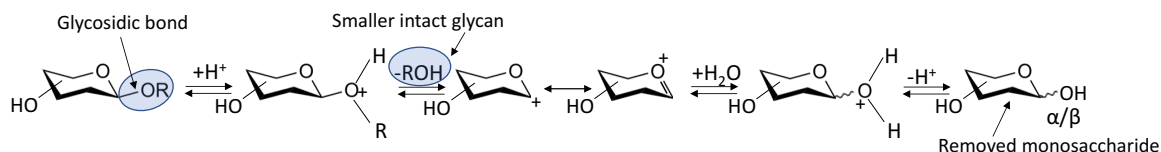


Figure 3.2. Simplified reaction mechanism of the glycosidic bond hydrolysis.

Figure 3.3 shows a scheme of the exoglycosidase sequencing of a complex glycan where arrows of different colors correspond to specific enzymes that can be applied in a single or an array mode.<sup>19</sup> The first method involves the removal of a single monosaccharide from the non-reducing end of its parent carbohydrate by one specific enzyme at each stage. This procedure is repeated until one obtains a known glycan. In the second technique, the oligosaccharide sample is divided into aliquots that are incubated with a precisely defined mixture of exoglycosidases. Recombining the products of each incubation and performing a single analysis on the product pool allow to obtain the glycan structure. Both of these methods are frequently applied for the identification of the primary structure of glycans. The first method is less ambiguous but takes much time in the case of relatively large glycans, as one has to repeat the isolation and determination of the product prior to each incubation. The second approach is much faster; however, it can be challenging to interpret the results unambiguously. In a standard procedure, exoglycosidase digestion is normally used in combination with LC-MS, where the enzymatic treatment is followed by a chromatographic analysis and a detection of the shifts in the retention time. Moreover, there are several bioinformatics tools that help to design enzymatic mixtures and simulate their action on oligosaccharides.<sup>20, 21</sup>

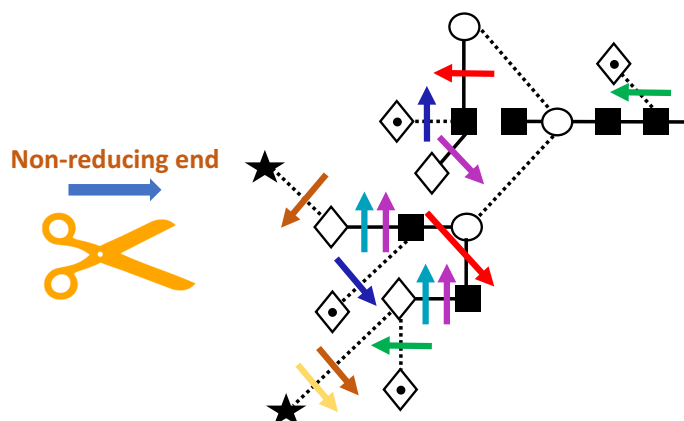


Figure 3.3. Scheme of exoglycosidase sequencing of a complex glycan from the non-reducing end. Arrows of different colors correspond to specific enzymes. (Redrawn from Fig. 4 of Ref. <sup>19</sup>)

While we will initially add as many known glycan structures as possible into our database, we will certainly encounter species that do not have references. In part 3.3.1. of this chapter, we show the results of combining the exoglycosidase digestion with IR spectroscopy as proof of a new principle for the identification of unknown glycan structures. Figure 3.4 illustrates our approach on the example of G2, which was used as a model for an unknown glycan. We degraded a parent glycan G2 through unknown glycan G0 until glycan Man-3 whose spectrum is already present in our database. Knowing the isomeric specificity of the applied enzymes, we can reconstruct the primary structure of both unknown parent glycan G2 and intermediate glycan G0 and then add their vibrational spectra to the database. This way, we identify unknown glycans and expand our spectroscopic database.

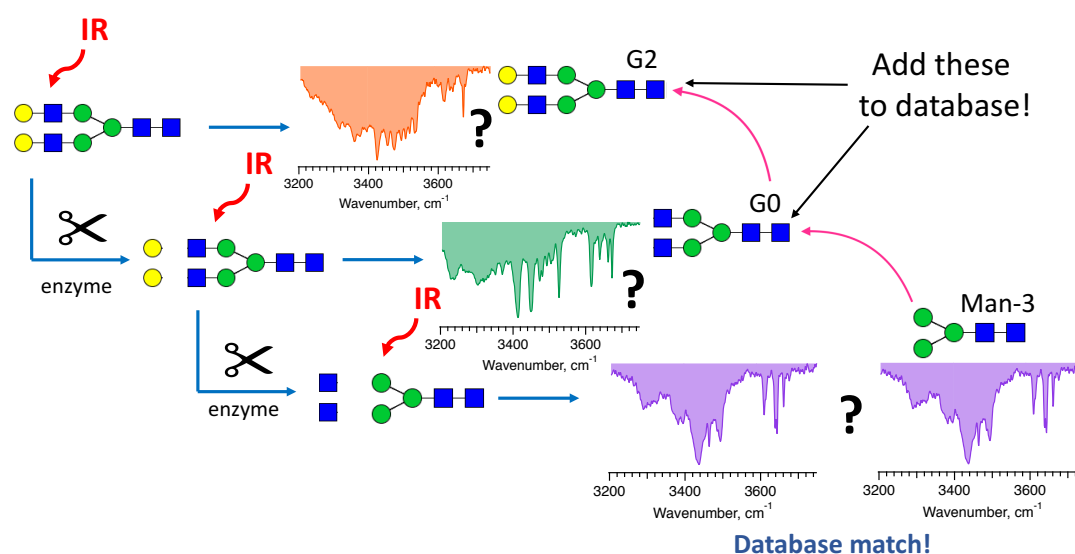


Figure 3.4. Scheme of database usage and construction based on the combination of IR spectroscopy and enzymatic degradation techniques.

### 3.2.2. Collision-induced dissociation

Tandem mass spectrometry ( $MS^n$ ), including low- and high-energy collision-induced dissociation, is a powerful method for glycan structure identification. After being isolated by a mass filter, the ion of interest undergoes one or more collisions with an inert buffer gas. The kinetic energy from this process is converted into an internal energy, which causes bonds in the molecule to break.<sup>22-25</sup> The resulting fragments are then analyzed by MS/MS to reconstruct the molecular structure of the precursor molecule. Low-energy CID leads to the cleavage of glycosidic bonds and occurs between monosaccharide units ( $Y_m/B_n$  and  $Z_m/C_n$  fragments in Figure 3.5). It reveals details on the glycan sequence and branching. High-energy CID induces cross-ring cleavages that involve the breaking of multiple bonds within the sugar rings,

resulting in  $A_n$  and  $X_m$  fragments. These fragments provide valuable information regarding linkages and branching. According to the nomenclature, fragments  $A_n$ ,  $B_n$  and  $C_n$  contain the non-reducing end, while  $X_m$ ,  $Y_m$  and  $Z_m$  keep the reducing end. The subscripts correspond to the position in the glycan chain relative to the reducing end and the superscripts denote the bonds that were cleaved during the cross-ring fragmentation.

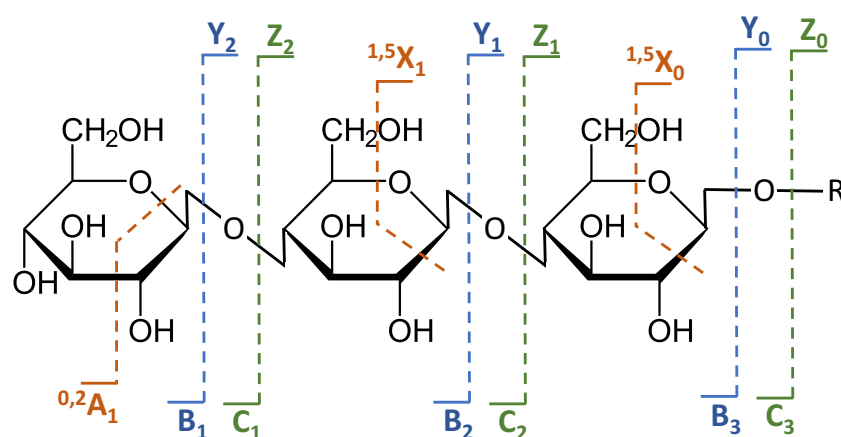


Figure 3.5. Nomenclature for oligosaccharide fragments introduced by B. Domon and C. Costello in 1988.<sup>26</sup>

In part 3.3.2. of the present chapter, we show our results on combining IR-IMS-MS with on-board CID as a four-dimensional approach for glycan structure identification. Moreover, we demonstrate the complementarity of exoglycosidase digestion and CID as both techniques have drawbacks and can only provide sufficient information for an unambiguous interpretation of the parent structure when implemented together. In order to identify unknown structures, we use C- and Y-fragments that represent intact glycans.



### 3.3. Combining cryogenic IR with selective enzymatic cleavage for determining the primary structure of glycans

The content of this section is adapted with the permission of the authors from the article: I. Dyukova, E. Carrascosa, R. P. Pellegrinelli, and T. R. Rizzo. “Combining cryogenic infrared spectroscopy with selective enzymatic cleavage for determining glycan primary structure”. *Anal. Chem.* **2020**, 92, 1658–1662.<sup>27</sup>

#### 3.3.1. Exoglycosidase digestion of N-linked glycans

In this work, we monitor the enzymatic degradation of N-linked glycans using UPLC-MS. For our experiment, the following set of glycan standards was used: Man-1, Man-3, G0, and G2 (Figure 3.6).

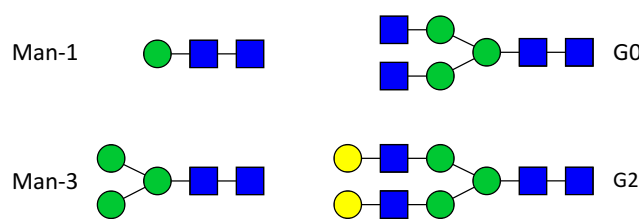


Figure 3.6. Schematic structures of the N-linked glycans used in this study.

All glycan reference compounds were purchased from Dextra (UK) and exoglycosidases ( $\beta(1-3,4)$ -galactosidase (Bovine testis),  $\alpha(1-2,3,6)$ -mannosidase (Jack bean),  $\beta$ -N-acetylhexosaminidase (Jack bean)) were purchased from Prozyme (Denmark). The glycan purity is certified to be of minimum 85 %.

To demonstrate our approach, we used an N-linked glycan, G2, as a model for an unknown glycan. To generate smaller glycans, we performed single enzymatic digestions for Man-3, G0 and G2, and an array of multiple digestion for G2. All reactions were carried out in a total glycan concentration of 0.4 mg/mL. A solution of oligosaccharide and exoglycosidase in a 50  $\mu$ L of buffer solution was incubated at 37 °C for various periods of time. Man-3 was incubated with 1  $\mu$ L of  $\alpha(1-2,3,6)$ -mannosidase (10 U) in a sodium acetate buffer solution (pH = 4-5, 100 mM) for 1 hour; G0 was incubated with 1  $\mu$ L of  $\beta$ -N-acetylhexosaminidase (5 U) and a sodium citrate/phosphate buffer solution (pH = 4-5, 100 mM) for 18 hours; the single cleavage of G2 was performed with 1  $\mu$ L of  $\beta(1-3,4)$ -galactosidase (0.5 U) in a sodium citrate/phosphate buffer

solution (pH = 5-6, 100 mM) for 18 hours; the multiple cleavage of G2 was carried out by the enzymatic mixture consisting of 1  $\mu$ l of  $\beta$ (1-3,4)-galactosidase (0.5 U) and 1  $\mu$ l of  $\beta$ -N-acetylhexosaminidase (5 U) in an ammonium acetate buffer solution (pH = 5, 100 mM) for 18 hours. In order to prevent column contamination, the enzymes were always removed using Sep-Pak C18 (1 cc, 50 mg) cartridges prior to the UPLC-MS analysis of the products of each cleavage reaction.

To control the products of each reaction, we first obtained the retention times of all four standards, including Man-1, Man-3, G0, and G2, by doing an LC run of their mixture (Figure 3.7A). During the incubation period of Man-3 with  $\alpha$ (1-2,3,6)-mannosidase, two mannose residues were removed from the non-reducing end of Man-3, leaving the glycan Man-1. The chromatogram (Figure 3.7B) of the mixture after cleavage confirms the full conversion of Man-3 to Man-1, showing a single peak for Man-1 with a retention time of 6.07 min. During the incubation period of G0 with  $\beta$ -N-acetylhexosaminidase, two N-acetylglucosamine monosaccharides were removed from the non-reducing end of G0 and the resulting smaller glycan was Man-3. The chromatogram (Figure 3.7C) of the mixture after cleavage confirms the full conversion of G0 to Man-3, showing a single peak with a retention time of 10.82 min. The single cleavage of G2 by  $\beta$ (1-3,4)-galactosidase led to the removal of two galactose units and the chromatogram (Figure 3.7D) of the mixture after cleavage confirms the full conversion of G2 to G0, showing a single peak for G0 with a retention time of 13.22 min. After the multiple exoenzyme cleavage of G2 using a combination of  $\beta$ (1-3,4)-galactosidase and  $\beta$ -N-acetylhexosaminidase, the same chromatogram (Figure 3.7C) was obtained as after a single exoenzyme cleavage of G0. The observation of a single peak with a retention time of 10.82 min confirms the full conversion of G2 to Man-3 in which two galactose units and two N-acetylglucosamine monosaccharides were removed from the non-reducing end.

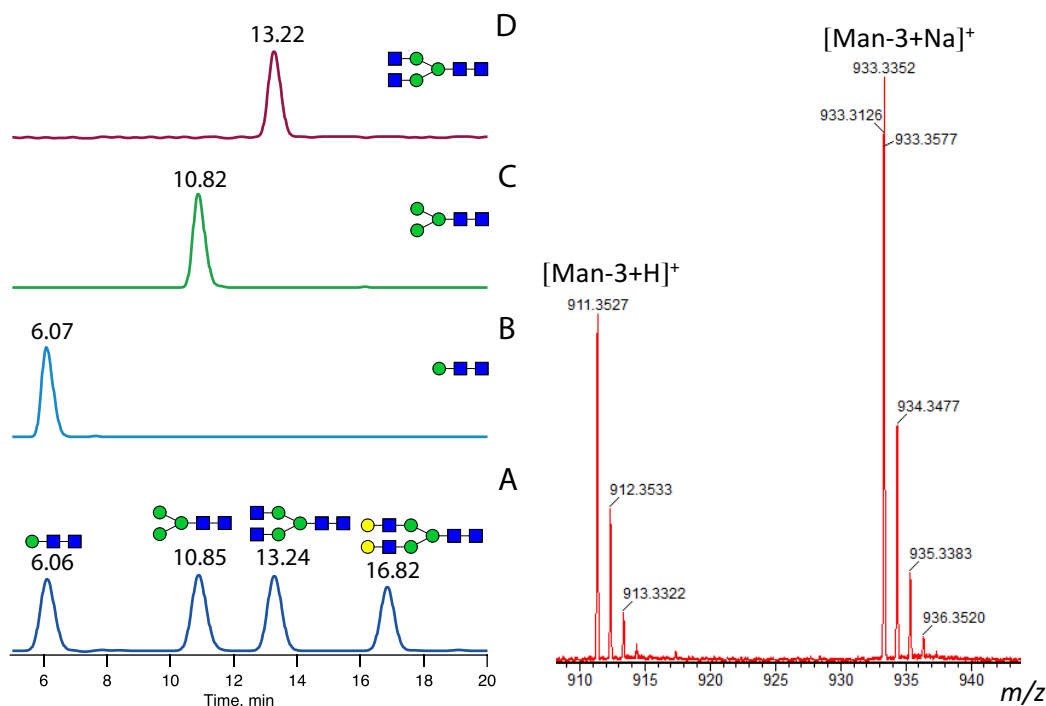


Figure 3.7. Left panel – chromatograms of a mixture of glycan standards (A) and smaller glycans after enzymatic cleavage (B–D). B – conversion of Man-3 to Man-1; C – conversion of G0 and G2 to Man-3; D – conversion of G2 to G0. Right panel – a typical mass spectrum recorded by the Q-TOF instrument using the Man-3 glycan standard.

After digestion, smaller glycans were purified by the LC-MS and collected using a fraction collector. It is important to note that here we do not follow the standard procedure for glycan analysis that includes labeling followed by UV or fluorescence detection. Using a Q-TOF mass spectrometer as a detector allows us to simplify and shorten the sample preparation procedure as we work with free glycans.

### 3.3.2. Infrared spectroscopy of glycan standards and smaller glycans after enzymatic cleavage

Before introducing the collected samples into the CIS instrument, purified glycans were concentrated and reconstituted in a 50:50 water/acetonitrile solution up to a final concentration of 10–15  $\mu\text{M}$ . The flow rate for nESI was approximately 1–100  $\mu\text{l}/\text{min}$  (depending on the needle and pressure), leading to 0.3–41 pmoles of the total sample amount for this experiment. The acquisition time using the Nd:YAG laser pumped OPO with 15 averages for each IR-spectra was  $\sim 30$  min for the range 3200–3750  $\text{cm}^{-1}$ . In the present work, we did not aim to show the relative sensitivity of our machine in comparison with another method for glycan

characterization. We were primarily focusing on the combination of two techniques and proving their compatibility as a new mechanism for the identification of unknown glycan species.

Due to the presence of at least two N-acetylglucosamine monosaccharides in the core structure, all N-linked glycans can be easily protonated in acidic medium. Since we used a buffer solution with low pH (4–5) during the clean-up procedure after enzymatic cleavage, the corresponding protonated glycans were the predominant species observed in all mass spectra of our collected samples. On the other hand, the glycan standards were mainly present in a sodiated form. In order to obtain a number of protonated species sufficient for measurements, acetic acid was added to the glycan standards before spraying. Based on the procedure described in Chapter 2, the IR spectra of protonated glycan standards and glycans after enzymatic cleavage were measured using the messenger-tagging technique.

### **Man-1 IR-spectra**

Figure 3.8A shows the measured IR-spectra of protonated Man-1 obtained by enzymatic cleavage of Man-3 along with that from the Man-1 reference compound. Both consist of a number of overlapping transitions in the range 3450–3550  $\text{cm}^{-1}$  and four distinct lines in the “free” OH-stretch region (3550–3700  $\text{cm}^{-1}$ ). The good agreement in both the position and intensities of the vibrational bands of the reference and cleaved Man-1 confirms the identity of the enzymatic degradation product. Figure 3.8B and 3.8C show the corresponding spectra of the sodiated and ammoniated Man-1. To obtain a number of sodiated and ammoniated adducts sufficient for the measurements, a sodium acetate was added to all collected glycans, and an ammonium acetate was added to glycan standards before spraying. As in the case of the protonated species, the spectra of the enzymatic digestion product matches that of the reference compound unambiguously. One expects the spectra of the protonated, sodiated and ammoniated species to differ, since the proton, sodium and ammonium cations will bind differently, leading to a different geometry and hence different vibrational frequencies. The spectral matches of the three species with their respective standards provide independent information with which to identify the enzymatic digestion product.

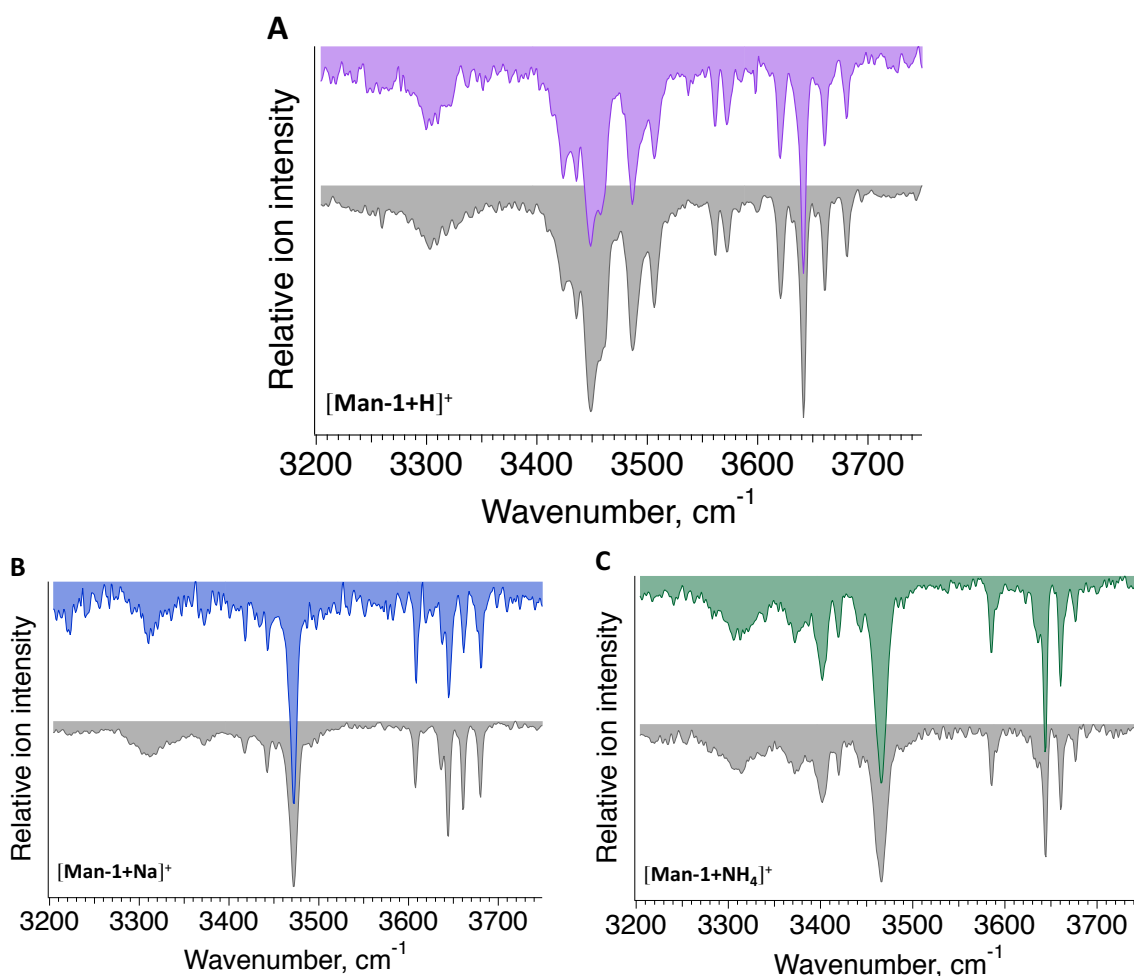


Figure 3.8. A, B, and C correspond to IR-spectra of the protonated, sodiated and ammoniated Man-1, respectively. Purple, blue and green spectra correspond to the reference Man-1; grey spectra correspond to Man-1 after enzymatic cleavage of Man-3.

All these spectra can further serve as glycan identifiers and be included in a spectroscopic database. Since the choice of solvent and concentration, as well as the specific enzymatic digestion procedure, will affect the relative intensities of these adducts, it is useful to include all these spectra in a database. This will enhance the probability of glycan identification irrespective of the experimental method.

### Man-3 IR-spectra

Man-3 was obtained by both the single digestion of G0 with  $\beta$ -N-acetylhexosaminidase and multiple digestion of G2 with a mixture of  $\beta$ (1-3,4)-galactosidase and  $\beta$ -N-acetylhexosaminidase. We measured the IR-spectra of these cleaved Man-3 oligosaccharides and compared them with the reference Man-3 spectra for protonated, sodiated and ammoniated adducts (Figure 3.9). The spectra of Man-3 after the single enzymatic digestion of G0 and

multiple digestion of G2 are in excellent agreement with the reference spectra of Man-3. These results demonstrate that the spectra of smaller glycans do not depend on the means by which they are produced and reinforce the practicality of using a spectroscopic database to identify them.

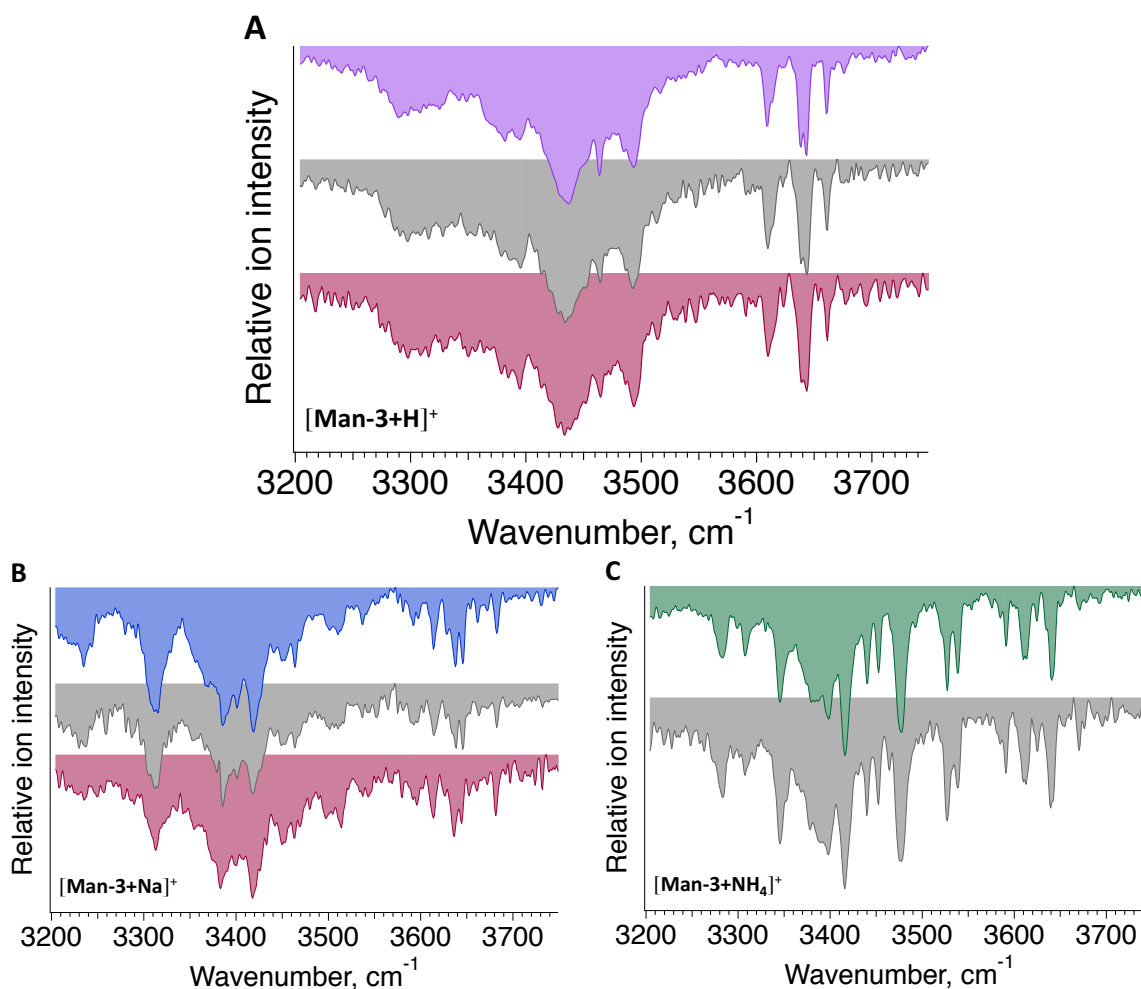


Figure 3.9. A, B, and C correspond to IR-spectra of the protonated, sodiated and ammoniated Man-3, respectively. Purple, blue and green spectra correspond to the reference Man-3; grey spectra correspond to Man-3 after enzymatic cleavage of G0; red spectra correspond to Man-3 after multiple enzymatic cleavage of G2.

## G0 IR-spectra

For G2, we performed the single digestion with  $\beta(1-3,4)$ -galactosidase and obtained the smaller glycan G0. Figure 3.10 shows the IR-spectra of the cleaved species together with that of the standard for the protonated and sodiated glycan G0.

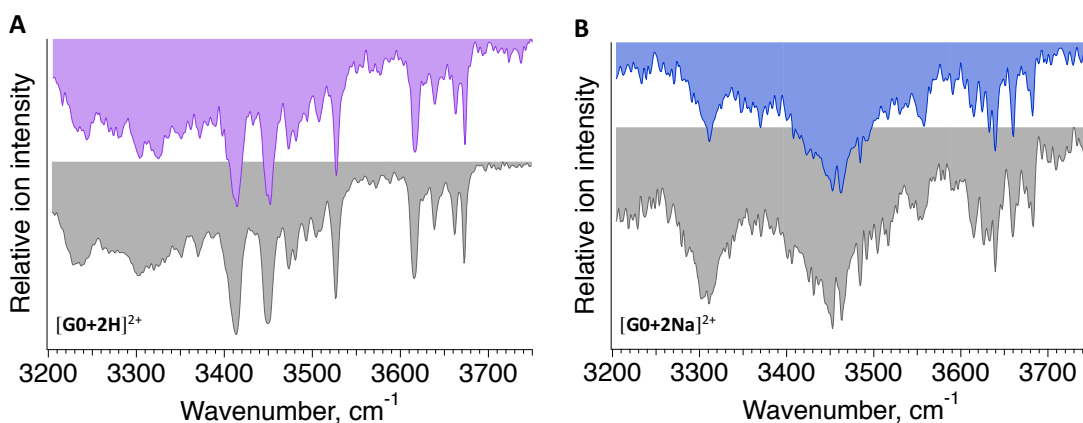


Figure 3.10. A and B correspond to IR-spectra of the protonated and sodiated G0, respectively. Purple and blue spectra correspond to the reference G0; grey spectra correspond to G0 after enzymatic cleavage of G2.

## G2 IR-spectra

The goal of our approach is to identify either unknown glycans or known glycans for which we cannot obtain standards. In the present case, we have used G2 as a model for an unknown glycan to demonstrate our procedure. Once we have identified the target parent glycan, we can then add its cryogenic IR spectrum to our database and identify it subsequently by its spectroscopic fingerprint. The spectrum of the G2 parent, shown in Figure 3.11, is sufficiently structured that it would serve well as such a fingerprint. Moreover, it provides the basis for identifying larger glycans that we can enzymatically degrade to G2.

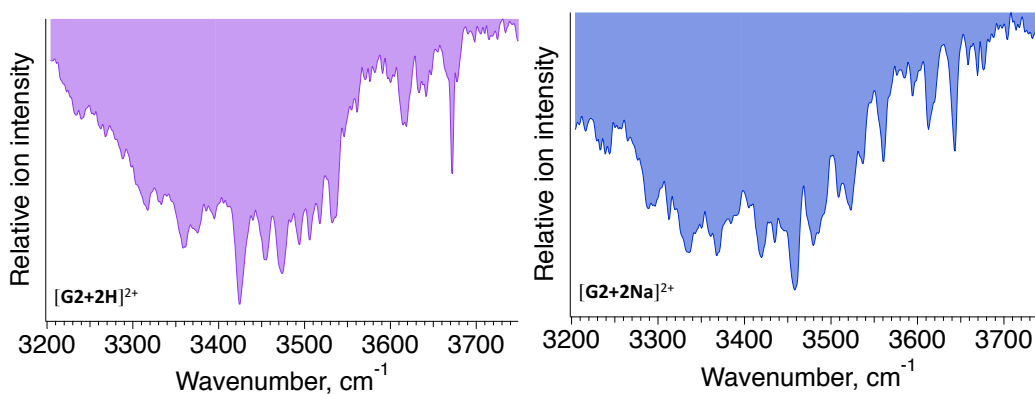


Figure 3.11. IR-spectrum of the protonated (purple) and the sodiated (blue) G2 reference.

### 3.4. Combination of CID and enzymatic cleavage for the identification of isomeric glycan fragments

As collision-induced dissociation is an intrinsically non-specific technique, the multitude of generated isomeric fragments can lead to an ambiguous interpretation of the parent glycan structure. In contrast, the enzymatic digestion approach allows a series of bond-specific cleavages that generate strictly defined structures; however, the variety of possible fragments is limited by the specificity of applied exoglycosidases. This part of the project aimed to show the complementarity of exoglycosidase digestion and CID for the identification of isomeric CID fragments that would further serve as the basis for the unravelling of unknown glycan structures.

We demonstrate our strategy using the example of two groups of isomeric fragments that were generated by an on-board CID of a triantennary N-linked glycan core (NGA3F)  $[\text{NGA3F}+2\text{Na}]^{2+}$  with  $m/z = 856$  on the third-generation IMS-CIS II instrument. The first group of fragments corresponds to the mass  $m/z = 754$  and to the loss of one N-acetylglucosamine monosaccharide, namely fragment  $[\text{G0F}+2\text{Na}]^{2+}$ . The second group corresponds to the mass  $m/z = 771$ , which involves the loss of three N-acetylglucosamine, one mannose monosaccharide and one fucose residues from NGA3F, leaving  $[\text{Man-2}+\text{Na}]^+$ . Figure 3.12 shows possible positional isomers for fragments with these masses.

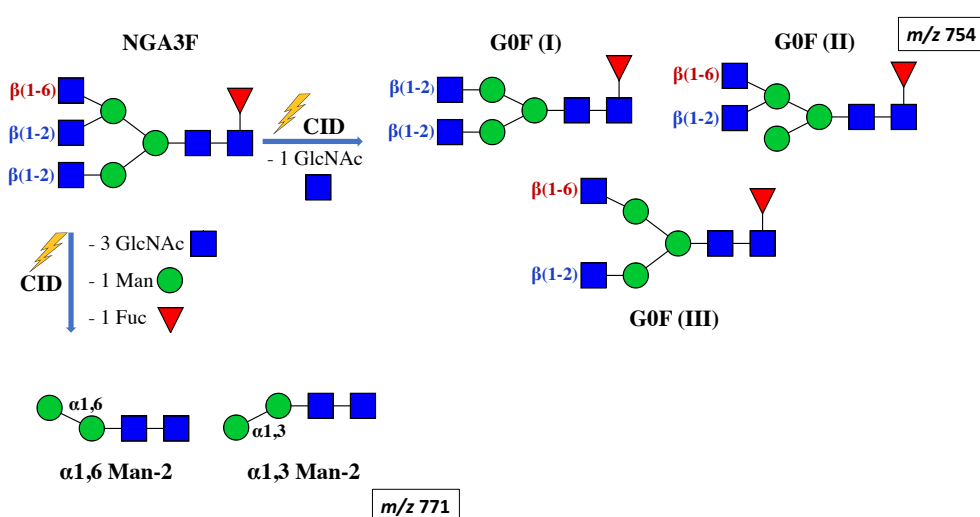


Figure 3.12. Possible “G0F” and “Man-2” isomeric structures after fragmentation of NGA3F.

After 3 cycles of separation (30 m path length) of the CID fragments  $[\text{G0F}+2\text{Na}]^{2+}$  on the SLIM board, we observed two distinct peaks (Figure 3.13A). Figure 3.13A\* depicts their IR



fingerprints. A symmetrical G0F (I) glycan standard with two  $\beta$ 1,2 bonds between the non-reducing end N-acetylglucosamine and a mannose was purchased from Dextra (UK). Pure standards of the other two structures were not found on the market due to the complexity of their synthesis and separation. To be able to assign those peaks, we implemented the enzymatic digestion approach to generate one of the standards. We applied  $\beta$ -N-acetylhexosaminidase (Agilent) to NGA3F, which specifically removes an N-acetylglucosamine monosaccharide from the  $\alpha$ 1,3-arm, leading to the G0F (II) isomer shown in Figure 3.12. The substrate (10  $\mu$ g) was incubated in a 50  $\mu$ l Tris-HCl buffer solution (20 mM, pH = 7.5) with 1,5  $\mu$ l of the enzyme (1.6 U) at 37  $^{\circ}$ C for 1 hour. After 3 cycles of separation, the purchased isomer G0F (I) displayed two peaks in the ATD that most likely correspond to two reducing-end anomers (Figure 3.13B). Under the same separation conditions, the enzymatically generated isomer G0F (II) showed only a single peak that we presume contains both reducing-end anomers (Figure 3.13C). The second IMS peak of the CID fragments overlaps with the peak 1\* of the G0F (I) and the peak of the G0F (II).

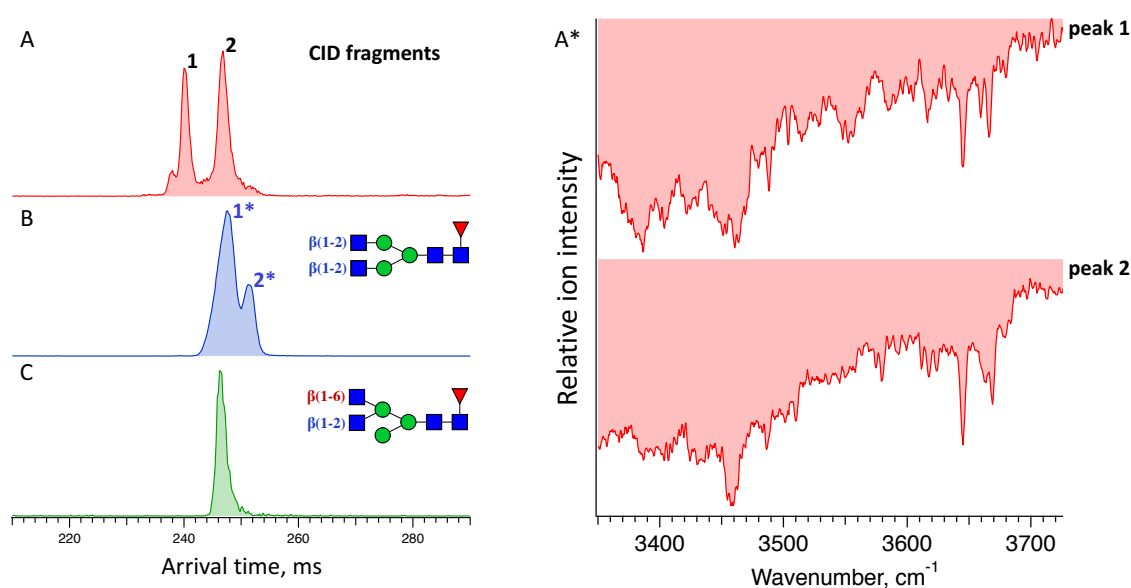


Figure 3.13. A, B, C – ATDs after 3 cycles separation of the  $[G0F+2Na]^{2+}$  isomeric fragments after an on-board fragmentation, the G0F(I) and G0F(II) isomers, respectively. A\* – IR fingerprints of the CID generated fragments.

We then used spectroscopy for further isomer discrimination. Based on the comparison of the IR fingerprints (Figure 3.14), we found that the spectrum of peak 2 contains features from both G0F (II) and peak 1\* of G0F (I). To determine the ratio between the two isomers, we performed a decomposition of the spectrum of peak 1 in MATLAB. The `fminsearch` function was used to minimize the root-mean-square deviation (RMSD) between the spectrum of the

mixture and a synthetic spectrum constructed from spectra of peak 1\* of G0F (I) and the lone peak of G0F (II) within the range of 3380–3686  $\text{cm}^{-1}$ . The best fit for the measured spectrum corresponds to 68.8 % of G0F (II) isomer and 31.2 % of G0F (I) isomer with RMSD = 4.7 %. While this approach provides qualitative information, ion suppression and charge competition limit its ability to be truly quantitative. The ratio between G0F (II) and G0F (I) isomers calculated above represents an approximate composition.

Peak 1 of the CID fragments could be matched neither to the arrival time nor to the IR fingerprint of any reference compounds. By exclusion, this peak must correspond to the third possible fragment (G0 (III)) for which we did not have a reference. This is an example where we can determine the primary structure of glycans and expand our database by adding vibrational fingerprints of those that cannot be purchased or generated enzymatically.

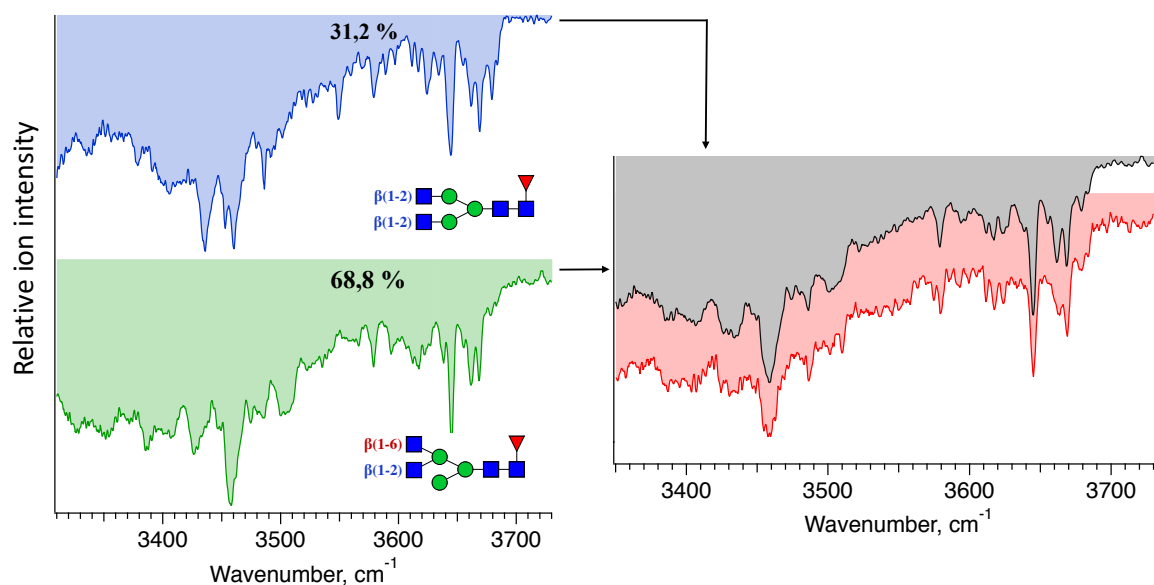


Figure 3.14. Isomer selective IR spectra of peak 1\* of G0F (I) (blue) and the lone peak of G0F (II) (green); synthetic spectrum (grey) constructed from spectra of peak 1\* of G0F (I) (31.2 %) and G0F (II) (68.8 %), peak II after CID (red).

For the identification of the second group of CID fragments  $[\text{Man-2}+\text{Na}]^+$ , we enzymatically synthesized one of the standards by applying  $\alpha(1-2,3,6)$ -mannosidase to the glycan Man-3. The  $\alpha(1-2,3,6)$ -mannosidase is known to completely cleave the Man-Man bond on the  $\alpha 1,3$ -arm first and then on the  $\alpha 1,6$ -arm. Man-3 was incubated with 1  $\mu\text{l}$  of  $\alpha(1-2,3,6)$ -mannosidase (10 U) in a sodium acetate buffer solution (pH = 4-5, 100 mM) for different periods of time within 1 hour (10, 30, 60 min) (Figure 3.15). Monitoring the kinetics of the enzymatic reaction, a 10

min incubation period was defined as optimal to produce  $\alpha$ 1,6 Man-2 isomer before the full conversion of Man-3 to Man-1 happens.

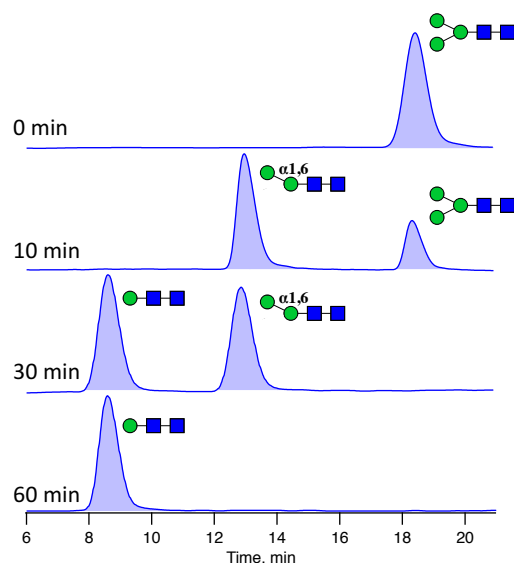


Figure 3.15. Kinetics of the conversion of a Man-3 substrate to Man-1 through the intermediate product Man-2 with  $\alpha$ (1-2,3,6)-mannosidase.

After 3 cycles separation (30 m path length) of the CID fragments  $[\text{Man-2}+\text{Na}]^+$  on the SLIM board, we observed two peaks (Figure 3.16). Based on the comparison between the IR fingerprints of these two peaks and the only peak of the  $\alpha$ 1,6 Man-2 isomer, we found that peak 2 of the CID fragments corresponds to the  $\alpha$ 1,6 Man-2 isomer and peak 1 must correspond to  $\alpha$ 1,3 Man-2. We can now add the spectra of these two structures ( $\alpha$ 1,6 Man-2 and  $\alpha$ 1,3 Man-2) to our database, even though we only had one standard.

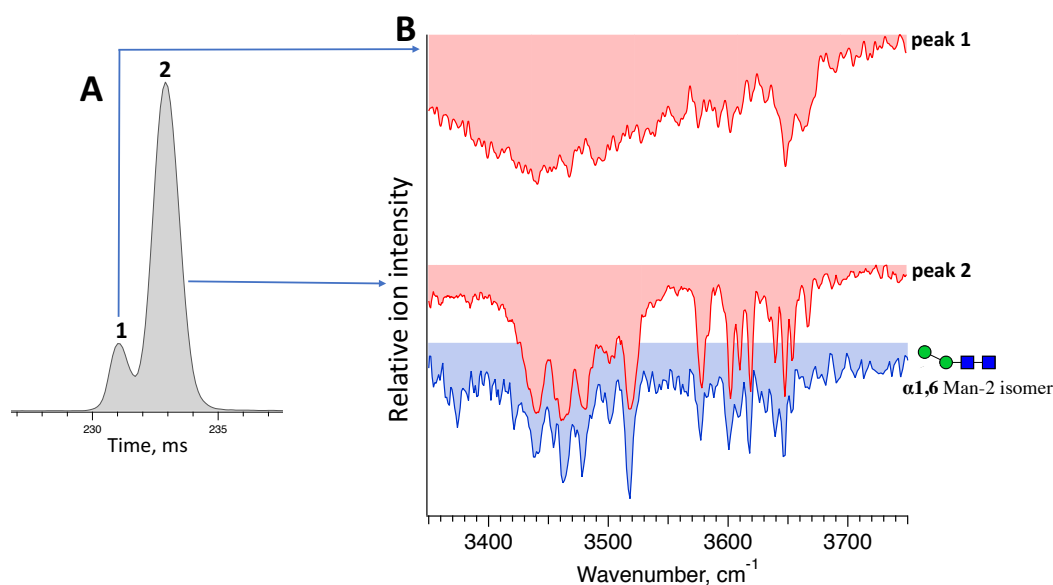


Figure 3.16. A – ATD of  $[\text{Man-2}+\text{Na}]^+$  isomeric fragments after an on-board CID and 3 cycles separation. B – isomer selective IR spectra of peaks I and II after CID (red);  $\alpha$ 1,6 Man-2 isomer (blue).

# Conclusions

The present work demonstrates the feasibility of combining enzymatic digestion and collision-induced dissociation with cryogenic IR-spectroscopy for the identification of unknown glycan using our spectroscopic database. For the development of this database, we will initially measure the spectra of as many known glycan structures as possible. Yet, one major goal is to analyze glycans from biological samples, which will certainly include species that are not (at least initially) contained in our database. By combining the existing protocols for exoglycosidase cleavage and CID with cryogenic infrared spectroscopy, we have developed a mechanism by which to determine unknown glycan primary structures and add them to the database. The IR-spectra of enzymatically cleaved and CID-generated glycans are in very good agreement with those of reference standards. The fact that G2 is the biggest glycan demonstrated in this work does not mean that our method is limited to this size/mass. G2 consists of nine monosaccharide units and we believe that, while obviously showing a broader and more complex vibrational spectrum than the spectra of trisaccharide Man-1, it possesses well-structured IR fingerprints in the region of OH stretching vibrations. To our knowledge, the spectra presented here show the best resolved infrared spectra of such a complex glycan entity to date. We strongly believe that even increasing the structural complexity of the molecules, we will be able to identify them at least in the “free” OH-stretch region (3550-3700  $\text{cm}^{-1}$ ).

These results serve as a starting point for constructing a glycan database and can be useful to benchmark future calculations of the structure of increasingly complex glycans. It is important to note that the methodology described in this work represents a procedure for constructing rather than using the database. We only have to go through this time-consuming process once for each unknown species. Each time we add a new species to the database, we gain the ability to identify it from a mixture based on its spectral fingerprint and obtain a new core structure that can be used in identifying still larger glycans that are built upon it. By initially focusing on N-linked glycans, we will construct a functional database that will progressively expand. While G2 will serve as a minimal motif for larger unknown glycans, these new species will serve as new minimal motifs in the stepwise construction of the future database. This process will continue until all possible glycan structures are added to the database.

# References

1. Doubet, S.; Bock, K.; Smith, D.; Darvill, A.; Albersheim, P., The Complex Carbohydrate Structure Database. *Trends Biochem Sci* **1989**, *14* (12), 475-7.
2. Raman, R.; Venkataraman, M.; Ramakrishnan, S.; Lang, W.; Raguram, S.; Sasisekharan, R., Advancing glycomics: Implementation strategies at the consortium for functional glycomics. *Glycobiology* **2006**, *16* (5), 82r-90r.
3. Maeda, M.; Fujita, N.; Suzuki, Y.; Sawaki, H.; Shikanai, T.; Narimatsu, H., JCGGDB: Japan Consortium for Glycobiology and Glycotechnology Database. *Methods Mol Biol* **2015**, *1273*, 161-79.
4. Campbell, M. P.; Peterson, R.; Mariethoz, J.; Gasteiger, E.; Akune, Y.; Aoki-Kinoshita, K. F.; Lisacek, F.; Packer, N. H., UniCarbKB: building a knowledge platform for glycoproteomics. *Nucleic Acids Res* **2014**, *42* (Database issue), D215-21.
5. Kanehisa, M.; Goto, S., KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **2000**, *28* (1), 27-30.
6. Lutteke, T.; Bohne-Lang, A.; Loss, A.; Goetz, T.; Frank, M.; von der Lieth, C. W., GLYCOSCIENCES.de: an Internet portal to support glycomics and glycobiology research. *Glycobiology* **2006**, *16* (5), 71R-81R.
7. Hayes, C. A.; Karlsson, N. G.; Struwe, W. B.; Lisacek, F.; Rudd, P. M.; Packer, N. H.; Campbell, M. P., UniCarb-DB: a database resource for glycomic discovery. *Bioinformatics* **2011**, *27* (9), 1343-4.
8. Tiemeyer, M.; Aoki, K.; Paulson, J.; Cummings, R. D.; York, W. S.; Karlsson, N. G.; Lisacek, F.; Packer, N. H.; Campbell, M. P.; Aoki, N. P.; Fujita, A.; Matsubara, M.; Shinmachi, D.; Tsuchiya, S.; Yamada, I.; Pierce, M.; Ranzinger, R.; Narimatsu, H.; Aoki-Kinoshita, K. F., GlyTouCan: an accessible glycan structure repository. *Glycobiology* **2017**, *27* (10), 915-919.
9. Zhao, S.; Walsh, I.; Abrahams, J. L.; Royle, L.; Nguyen-Khuong, T.; Spencer, D.; Fernandes, D. L.; Packer, N. H.; Rudd, P. M.; Campbell, M. P., GlycoStore: a database of retention properties for glycan analysis. *Bioinformatics* **2018**, *34* (18), 3231-3232.
10. Toukach, P. V.; Egorova, K. S., Carbohydrate structure database merged from bacterial, archaeal, plant and fungal parts. *Nucleic Acids Res* **2016**, *44* (D1), D1229-36.
11. Struwe, W. B.; Pagel, K.; Benesch, J. L.; 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.

12. 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<sup>+</sup>-Gly-Pro-Gly-Gly (GPPG). *J Am Soc Mass Spectr* **2015**, *26* (9), 1444-1454.
13. Aseev, O.; Perez, M. A.; Rothlisberger, U.; Rizzo, T. R., Cryogenic Spectroscopy and Quantum Molecular Dynamics Determine the Structure of Cyclic Intermediates Involved in Peptide Sequence Scrambling. *J Phys Chem Lett* **2015**, *6* (13), 2524-9.
14. Nagornova, N. S.; Guglielmi, M.; Doemer, M.; Tavernelli, I.; Rothlisberger, U.; Rizzo, T. R.; Boyarkin, O. V., Cold-ion spectroscopy reveals the intrinsic structure of a decapeptide. *Angew Chem Int Ed Engl* **2011**, *50* (23), 5383-6.
15. 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.
16. Abikhodr, A. H.; Yatsyna, V.; Ben Faleh, A.; Warnke, S.; Rizzo, T. R., Identifying Mixtures of Isomeric Human Milk Oligosaccharides by the Decomposition of IR Spectral Fingerprints. *Anal Chem* **2021**, *93* (44), 14730-14736.
17. Royle, L.; Radcliffe, C. M.; Dwek, R. A.; Rudd, P. M., Detailed structural analysis of N-glycans released from glycoproteins in SDS-PAGE gel bands using HPLC combined with exoglycosidase array digestions. *Methods Mol Biol* **2006**, *347*, 125-43.
18. Kamerling, J. P.; Gerwig, G. J., Strategies for the Structural Analysis of Carbohydrates. *Comprehensive Glycoscience* **2007**, *2*, 1-68.
19. Marino, K.; Bones, J.; Kattla, J. J.; Rudd, P. M., A systematic approach to protein glycosylation analysis: a path through the maze. *Nat Chem Biol* **2010**, *6* (10), 713-723.
20. Walsh, I.; Nguyen-Khuong, T.; Wongtrakul-Kish, K.; Tay, S. J.; Chew, D.; Josee, T.; Taron, C. H.; Rudd, P. M., GlycanAnalyzer: software for automated interpretation of N-glycan profiles after exoglycosidase digestions. *Bioinformatics* **2019**, *35* (4), 688-690.
21. Gotz, L.; Abrahams, J. L.; Mariethoz, J.; Rudd, P. M.; Karlsson, N. G.; Packer, N. H.; Campbell, M. P.; Lisacek, F., GlycoDigest: a tool for the targeted use of exoglycosidase digestions in glycan structure determination. *Bioinformatics* **2014**, *30* (21), 3131-3133.
22. Reinhold, V. N.; Reinhold, B. B.; Costello, C. E., Carbohydrate molecular weight profiling, sequence, linkage, and branching data: ES-MS and CID. *Anal Chem* **1995**, *67* (11), 1772-84.
23. Tsai, S. T.; Liew, C. Y.; Hsu, C.; Huang, S. P.; Weng, W. C.; Kuo, Y. H.; Ni, C. K., Automatic Full Glycan Structural Determination through Logically Derived Sequence Tandem Mass Spectrometry. *Chembiochem* **2019**, *20* (18), 2351-2359.

24. Gray, C. J.; Schindler, B.; Migas, L. G.; Picmanova, M.; Allouche, A. R.; Green, A. P.; Mandal, S.; Motawia, M. S.; Sanchez-Perez, R.; Bjarnholt, N.; Moller, B. L.; Rijs, A. M.; Barran, P. E.; Compagnon, I.; Evers, C. E.; Flitsch, S. L., Bottom-Up Elucidation of Glycosidic Bond Stereochemistry. *Analytical Chemistry* **2017**, *89* (8), 4540-4549.
25. Harvey, D. J., Collision-induced fragmentation of underivatized N-linked carbohydrates ionized by electrospray. *J Mass Spectrom* **2000**, *35* (10), 1178-1190.
26. Domon, B.; Costello, C. E., A Systematic Nomenclature for Carbohydrate Fragmentations in Fab-MS MS Spectra of Glycoconjugates. *Glycoconjugate J* **1988**, *5* (4), 397-409.
27. 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.

# Chapter 4. Identification of positional isomers of glycans cleaved from monoclonal antibodies

## 4.1. Introduction

The content of this chapter is adapted with the permission of the authors from the article: I. Dyukova, A. Ben Faleh, S. Warnke, N. Yalovenko, V. Yatsyna, P. Bansal, and T. R. Rizzo. “A new approach for identifying positional isomers of glycans cleaved from monoclonal antibodies”. *Analyst* **2021**, 146, 4789–4795.<sup>1</sup>

Biologics or biological drugs are products made from living systems. Biosimilars represent copies of licensed biologics made by different manufacturers.<sup>2</sup> Biologics and biosimilars are used to treat the deadliest diseases such as various types of cancer,<sup>3</sup> immune-mediated disorders<sup>4</sup> and diabetes.<sup>5</sup> According to the guidelines of the European Medicines Agency's (<https://www.ema.europa.eu/en>), “biosimilars can only be approved as long as there are no clinically meaningful differences in their safety and effectiveness in comparison with the original drug”.

Due to the high cost of production and characterization of originator biologics, there is a huge interest in developing biosimilars, which are much cheaper since they do not need to undergo clinical trials.<sup>6</sup> The typical decrease in price is 15–45 % and in some cases can even reach 80 % (e.g. biosimilars of Humira from AbbVie Inc.).<sup>7, 8</sup> Cost savings allow to increase patient access to the therapy and help to implement the treatment already at an early stage of the disease. Moreover, the competition between biologics and biosimilars stimulates the development of next-generation products. In Europe, by 2018, 34 biologics became patent-free and the patents of 15 more will expire in the next 5 years.<sup>9, 10</sup>

Monoclonal antibodies (mAbs) represent one of the biggest groups of all biologics. Most of the current therapeutic mAbs are humanized or human immunoglobulins G (IgGs), produced as recombinant glycoproteins in eukaryotic cells.<sup>11</sup> IgGs are about 150 kDa in size and composed of two identical heavy chains of ~50 kDa and two identical light chains of ~25 kDa (Figure 4.1).<sup>12</sup>



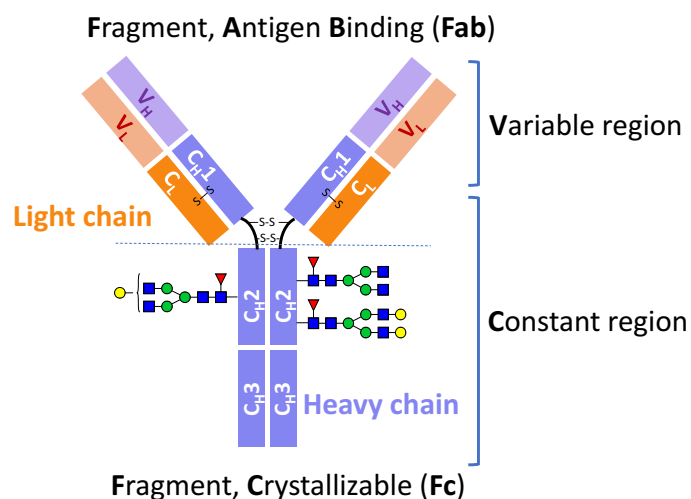


Figure 4.1. Schematic structure of Immunoglobulin G antibody.

The N-terminal domain of an IgG consists of a variable region with the complementarity-determining region that binds to a specific epitope of an antigen. The other domains within the IgG make up the constant regions. The structure of an IgG is divided into the Fab (2/3) and the Fc regions (1/3). Each of the two identical Fabs consists of light and heavy chains linked by an interchain disulfide (S–S) bond. The hinge region connects the Fabs to the Fc via a double disulfide linkage between the heavy chains.<sup>13</sup>

In contrast to small molecules (150–600 Da), it is challenging to reproduce mAbs precisely, as they are very large and complex systems exhibiting micro-heterogeneities. To produce the same biosimilar product, suppliers use different cell lines and manufacturing processes that may affect the molecular structure and, as a result, the quality and safety of the final medicine.<sup>2</sup> Moreover, biologics might also have a micro-heterogeneity problem that comes from the batch-to-batch variability. Therefore, all parties (innovators and biosimilar manufacturers) are interested in comprehensively analyzing their products.

Immunoglobulin G molecules are glycosylated in the C<sub>H</sub>2 domains of the Fc region (Figure 4.1), with glycans being covalently attached to the Asn297 residue. The N-glycans of the Fc region contribute approximately 2–3% to the total mass of the IgG protein.<sup>11, 12</sup> Despite this low percentage, the N-glycan moieties can have a significant impact on the effector functions of antibodies, such as the antibody-dependent cell-mediated cytotoxicity (ADCC) and the complement-dependent cytotoxicity (CDC).<sup>14, 15</sup> For example, it has been established that the absence of core fucose (Fuc) residues in the N-glycans of the Fc region substantially increases the ADCC activity. Moreover, a high sialic acid content reduces ADCC activity but at the same

time plays an important role in anti-inflammatory responses.<sup>16, 17</sup> Terminal galactose is well known to enhance CDC activity and its impact on ADCC activity has also been reported.<sup>18-23</sup>

It has only recently been demonstrated that the terminal galactose position (i.e., on either the core mannose (Man)  $\alpha$ 1,6 or  $\alpha$ 1,3 branch) has a significant effect on the effector functions of mAbs. Aoyama et al. have shown that the G1( $\alpha$ 1,6)F mAb has higher complement component 1q (C1q)- and Fc gamma receptor (Fc $\gamma$ R)-binding activities and CDC activity than the G1( $\alpha$ 1,3)F mAb because of the greater involvement of the galactose on the  $\alpha$ 1,6 branch in the structural stability of the C<sub>H</sub>2 domain.<sup>24</sup> It is important to note that mAbs exhibit micro-heterogeneities that can lead to the presence/absence or different ratios between the N-glycans in the Fc region with terminal Gal on the Man  $\alpha$ 1,6 and  $\alpha$ 1,3 arms. Effective tools are thus needed to analyze protein glycoforms, even at the isomer level, for both biological mAbs and biosimilars.<sup>25-27</sup>

Several methods have been implemented to distinguish and identify positional isomers of released N-linked glycans with terminal Gal ( $\alpha$ 1,6/ $\alpha$ 1,3). These include tandem mass spectrometry,<sup>28</sup> ion mobility spectrometry,<sup>29</sup> and various combinations of selective enzymatic digestion or synthesis with NMR or liquid chromatographic analysis.<sup>30-33</sup> The most commonly used method currently combines HILIC and mass spectrometry, where the chromatographic peak assignment is based on the previously published work indicating that the glycan with a terminal galactose on the upper Man ( $\alpha$ 1,6) arm elutes prior to that with galactose on the lower Man ( $\alpha$ 1,3) arm.<sup>34</sup> Despite the potential of this hybrid technique, glycan LC workflows typically involve a derivatization step to label the glycans with a fluorescent tag. While this improves sensitivity and facilitates quantification, it complicates the workflow, and the labels can be expensive.

In the present work, we use a combination of ultra-high resolution IMS with cryogenic infrared spectroscopy as a rapid and reliable technique for glycan isomer identification. Our approach allows one to obtain highly resolved, isomer-specific vibrational spectra, even of larger, more complex glycan ions. We have implemented a chemoenzymatic approach to synthesize selectively the glycan isomer G1( $\alpha$ 1,6)F and characterized it by IMS and vibrational spectroscopy.<sup>35, 36</sup> We then demonstrated the impact of the host cell line (CHO and HEK-293) on the ratio of G1F isomers within the glycan profile of IgG.

## 4.2. Ion-mobility-selective IR spectroscopy of glycan G1F isomers

For the first time, we demonstrate a comprehensive analysis of N-linked glycan G1F positional isomers using a combination of IMS and cryogenic IR spectroscopy. Commercially available G1F (Dextra Laboratories) was provided as a mixture of two positional isomers, with the terminal galactose sitting on either the Man  $\alpha$ 1,6 or Man  $\alpha$ 1,3 arms. The solutions were prepared by dissolving the respective powder in a 50/50 H<sub>2</sub>O/acetonitrile mixture for a concentration of 15–20  $\mu$ M. We performed our experiments for the doubly sodiated G1F glycan standard ( $[G1F+2Na]^{2+}$ ,  $m/z = 835$ ) on the second-generation IMS-CIS I instruments. Four peaks were observed after three cycles on the SLIM board (4.87 m drift path) (Figure 4.2). We compared our results with a separation by LC. Two positional isomers elute from the HILIC column as one single peak regardless of the separation conditions.

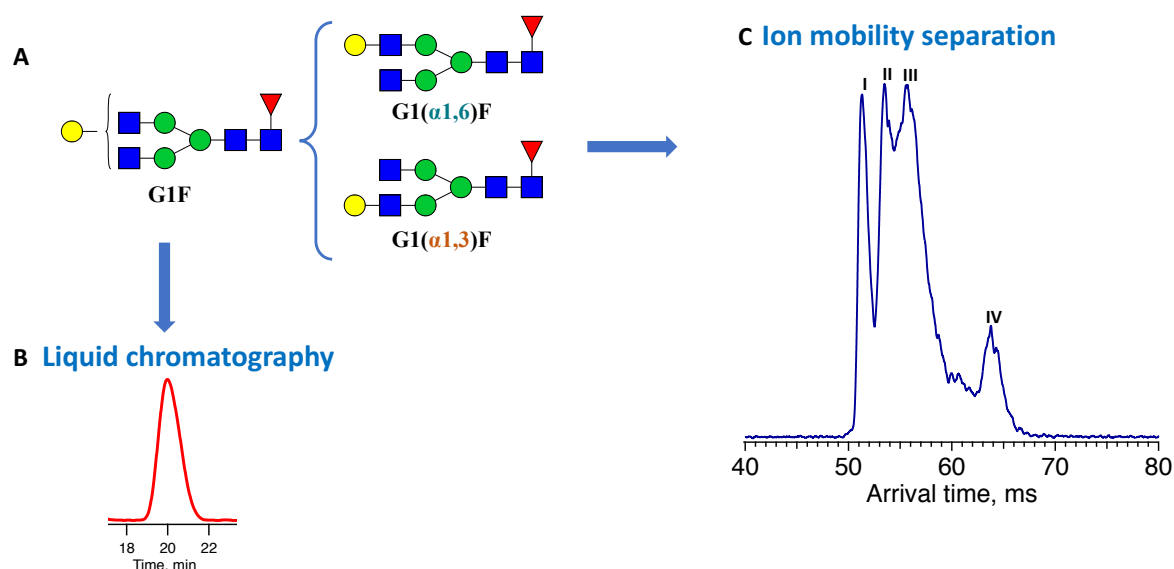


Figure 4.2. A – Positional isomers of glycan G1F: G1( $\alpha$ 1,6)F and G1( $\alpha$ 1,3)F; B – one peak separation of the G1F mixture with HILIC after 22 min; C – arrival time distribution of the G1F mixture after 3 cycles on the SLIM board of IMS-CIS I.

In order to achieve a better separation of the peaks II and III (Figure 4.2C), we would need to send ions for more cycles along the traveling path. However, after 4 cycles, the ATD starts overlapping with a previous cycle on the SLIM board of the IMS-CIS I instrument. In addition, we were not able to obtain a pure ATD for doubly sodiated species as it has a very close arrival time with  $[G1F+H+K]^{2+}$  ( $m/z = 832$ ) species. The quadrupole mass filter could not select such a narrow window that transmits  $m/z = 835$  without  $m/z = 832$ . Therefore, to increase mobility

and mass resolving power, all the experiments for this project were transferred to the third-generation IMS-CIS II instrument.

Figure 4.3 shows four distinct peaks of the G1F mixture after six cycles on the SLIM board (60 m drift path) of the IMS-CIS II instrument. Since two neighboring peaks (II and III) are just resolved, we can estimate the difference in their CCS values to be 0.2–0.3%. For the two positional isomers, we observed four distinct peaks on the ATD. These peaks could arise from the two reducing-end anomers, which we have shown can be separated by ultra-high resolution IMS,<sup>37</sup> or from different conformers for each isomer.

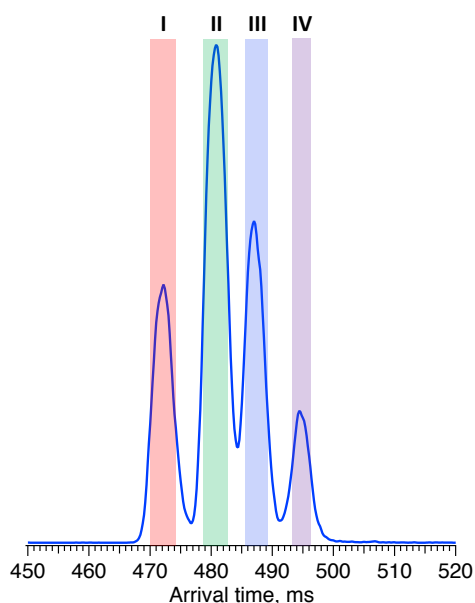


Figure 4.3. Arrival time distribution of G1F after a 60 m IMS separation (i.e., six separation cycles) on the SLIM board of IMS-CIS II, representing a mixture of the G1( $\alpha$ 1,6)F and G1( $\alpha$ 1,3)F positional isomers.

Figure 4.4 displays the cryogenic IR spectra of each peak in the ATD of G1F. Even though G1F consists of nine monosaccharide units, each spectrum exhibits a highly distinct spectral fingerprint that can serve as a unique and robust identifier. We spectroscopically confirmed that regardless of the experimental conditions affecting the arrival time distribution, the same species corresponds to each peak in the ATD.

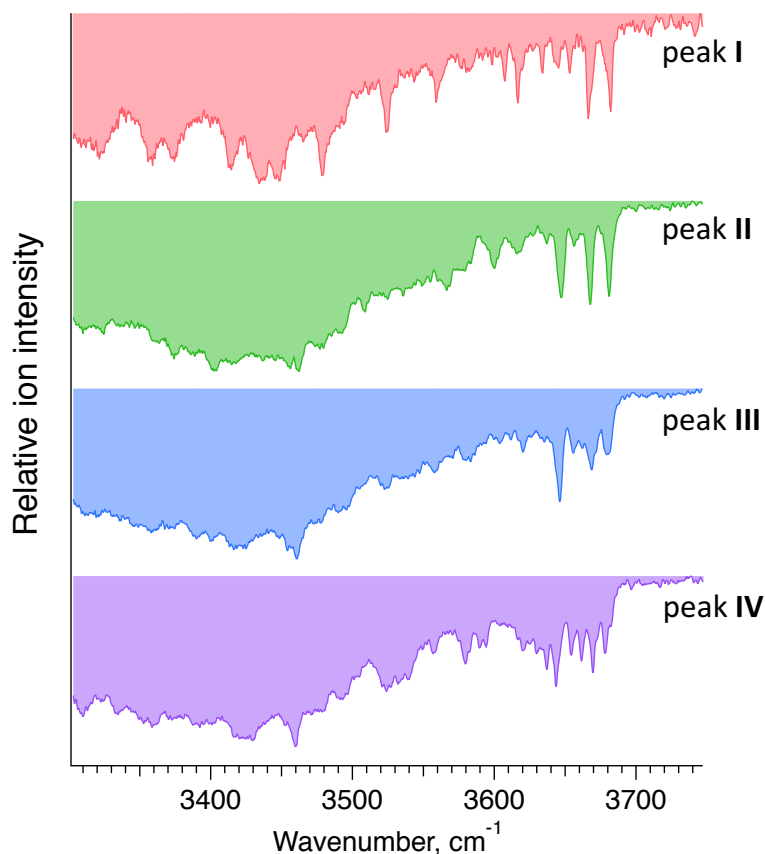


Figure 4.4. Cryogenic IR spectra of peaks I–IV observed in the ATD of the G1F glycan shown in Fig.4.3.

Each IR spectrum was acquired within ~10 min to allow for averaging of three scans. However, many fewer spectral data points are needed to uniquely identify each species. Moreover, one can see that the 3600–3700  $\text{cm}^{-1}$  range exhibits sufficiently unique absorption bands in every one of the four spectra to facilitate compound identification. Taken together, this would allow us to identify each of these glycans in less than one minute.

### 4.3. Selective enzymatic synthesis of the glycan G1( $\alpha$ 1,6)F

In order to assign each peak in the ATD of G1F to one of two positional isomers, we needed pure standards. However, pure isomers G1( $\alpha$ 1,6)F and G1( $\alpha$ 1,3)F are difficult to obtain commercially. The samples we purchased from the few manufacturers that offer them were of unsatisfactory isomeric purity, exhibiting the same ATD profile as for the G1F mixture.

There are three main strategies to produce the standards of N-linked glycans.

## 1. Releasing from natural sources

Naturally occurring N-linked glycans are extremely micro-heterogeneous due to the processes by which they are assembled in biosystems, and they are often found in low concentrations. Moreover, isolated glycans represent a mixture of compounds with close structures that are difficult to separate by current analytical techniques. All these factors make it challenging to obtain adequate amounts of N-glycans with specific structures from natural sources.<sup>38, 39</sup>

## 2. Chemical synthesis

Chemical synthesis has significantly contributed to the construction of libraries for oligosaccharide standards. Regioselective protection/deprotection methods for multiple hydroxy groups of sugars and efficient glycosylation reactions have provided tools for the synthesis of a large variety of N-glycans, even with complex structures such as tri- and tetra-antennary oligosaccharides. However, this type of synthesis is a difficult and time-consuming task that usually requires many steps, including purification of stereoselective glycosidic linkages. Even biantennary complex-type oligosaccharides are synthesized in at least 20 steps.<sup>40-43</sup> Moreover, the chemical synthesis of glycans represents non-trivial work that relies on the solid synthetic experience of the researchers.

## 3. Chemoenzymatic synthesis

The surge in the study of glycans over the last two decades has stimulated the development of the chemoenzymatic method, which has been proven to be a fast and reliable technology to produce glycan standards. The main principle of chemoenzymatic synthesis is based on the modification of synthetic glycan precursors by one or a set of glycosyltransferases (GTs) to yield derivatives of higher complexity. GTs are a class of enzymes that catalyze the formation of glycosidic bonds by transferring a donor sugar block to an acceptor substrate.<sup>44</sup> Glycosyltransferases demonstrate stereo- and regiospecificity, and they transfer donors with either retention or inversion of the configuration at their anomeric carbon (Figure 4.5). The name of each enzyme includes the name of the donor used in the reaction, the position of the bond formation on a substrate and the stereochemistry of the transfer. Initially, such an approach was exclusively focused on symmetrically branched oligosaccharides,<sup>45, 46</sup> and only

recently, a few groups demonstrated a strategy for chemoenzymatic synthesis of asymmetrical glycans.<sup>36, 47-49</sup>

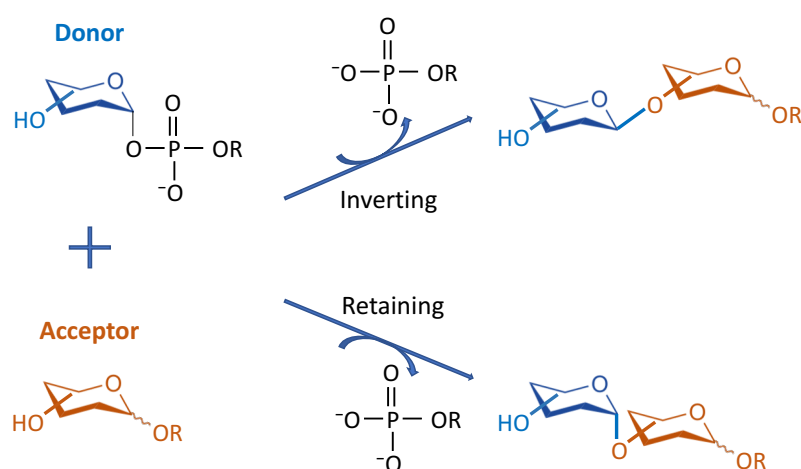


Figure 4.5. Reaction mechanism of glycosyltransferases upon inversion or retention of the anomeric glycosidic bond.

FUT8 ( $\alpha$ 1,6-fucosyltransferase) is the only enzyme that catalyzes core-fucosylation in mammals. Previous studies have demonstrated that galactosylation of the GlcNAc residue at the  $\alpha$ 1,3-mannose branch and the bisecting of N-glycans prevent FUT8 activity. For this reason, it has been widely used in the synthesis of asymmetrical N-glycans. We have thus used a chemoenzymatic approach to synthesize the G1( $\alpha$ 1,6)F isomer selectively. The enzyme catalyzes the transfer of L-fucose to form an  $\alpha$ 1,6-linkage to the reducing-end GlcNAc of G1( $\alpha$ 1,6), leading to the selective synthesis of the G1( $\alpha$ 1,6)F isomer (Figure 4.6).<sup>35, 36, 49</sup>

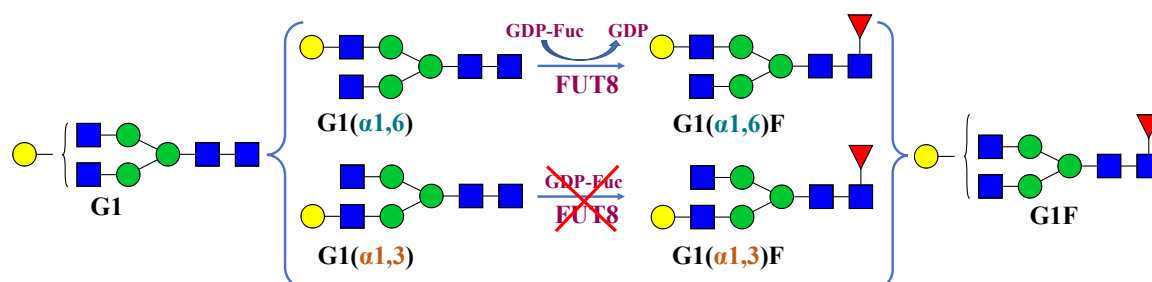


Figure 4.6. Positional isomers of G0 and G1F glycans with a scheme of chemoenzymatic synthesis of the G1( $\alpha$ 1,6)F isomer using human  $\alpha$ 1,6-fucosyltransferase FUT8.

The synthesis of G1( $\alpha$ 1,6)F was performed in a total volume of 50  $\mu$ l containing 0.27 mM of an acceptor G1 N-linked glycan (Dextra Laboratories, UK), 0.54 mM of GDP-Fuc (guanosine 5'-diphospho- $\beta$ -L-fucose sodium salt) (Sigma-Aldrich), a MES (2-(N-

morpholino)ethanesulfonic acid) buffer solution (100 mM, pH = 7.0) and 0.12 mg/mL of FUT8 (Creative BioMart, USA). The reaction was incubated overnight at 37 °C. In order to determine the optimal reaction time, samples were incubated for 1, 2, 3, 4, 6, 8, 16 and 48 h. The reactions were then quenched by adding 30  $\mu$ l of ice-cold acetonitrile to the mixture and analyzed by LC-MS. The conversion efficiency was calculated as  $\% = \text{Product peak area} / (\text{Product peak area} + \text{Substrate peak area}) \times 100$  (Table 4.1).

Table 4.1. Conversion efficiency for different incubation times of the reaction

	G1( $\alpha$ 1,6)F
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

Figure 4.7 depicts the liquid chromatogram of the reaction mixture after 8h of incubation along with a typical MS spectrum. The final product, the G1( $\alpha$ 1,6)F isomer, was collected after 16 h, with a yield of ~50%, using a fraction collector, and the solvents were evaporated to concentrate the sample up to 15–20  $\mu$ M (50/50 H<sub>2</sub>O/acetonitrile mixture). Our assumption was that we start with a 50/50 mixture of the positional isomers of G<sub>0</sub>. Since we were producing only one of the positional isomers, the conversion efficiency should be less than 50%. According to the results, the best incubation time of the reaction was 16h. It can be seen from Table 4.1 that a conversion rate of more than 50% was observed after 48 h (i.e., 57%), possibly due to the reaction of G1( $\alpha$ 1,3) and the formation of the G1( $\alpha$ 1,3)F isomer. For this reason, we chose 16h as the optimal incubation time of the reaction, assuming that G1( $\alpha$ 1,6)F is exclusively formed.



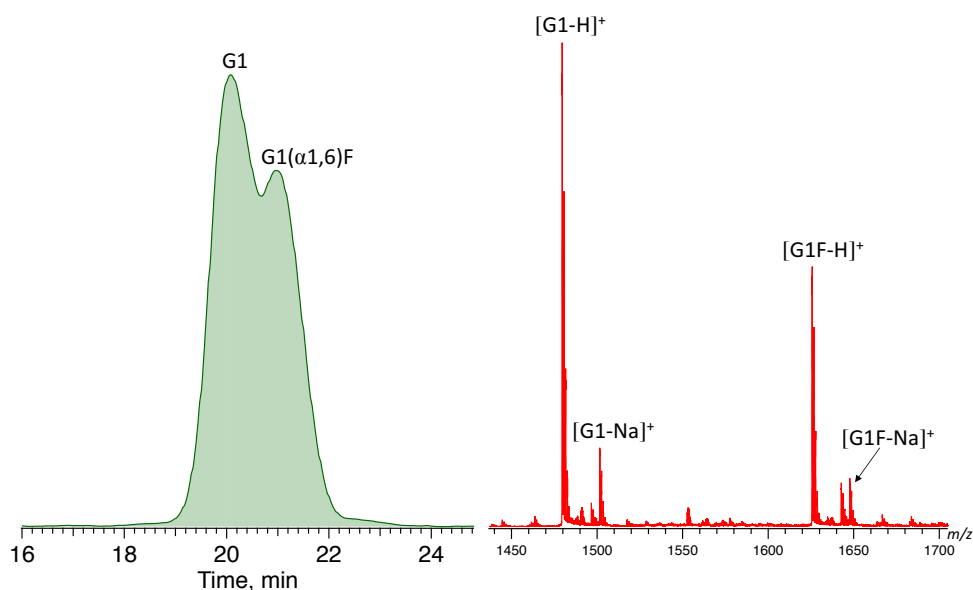


Figure 4.7. Liquid chromatogram and MS spectrum of the reaction mixture after 8h of incubation.

#### 4.4. Identification of the positional isomers of the G1F based on the ion-mobility-selective IR fingerprints

Figure 4.8 displays the ATD obtained for the doubly sodiated G1F glycan mixture (purple) and the synthesized G1(α1,6)F isomer (green) performed under the same experimental conditions after 6 cycles (60 m) of separation by SLIM-IMS. One can see that the ATD of the G1(α1,6)F isomer is simpler, exhibiting only two distinct peaks (**III\*** and **IV\***) that match the drift times of peaks **III** and **IV** of the G1F mixture, respectively.

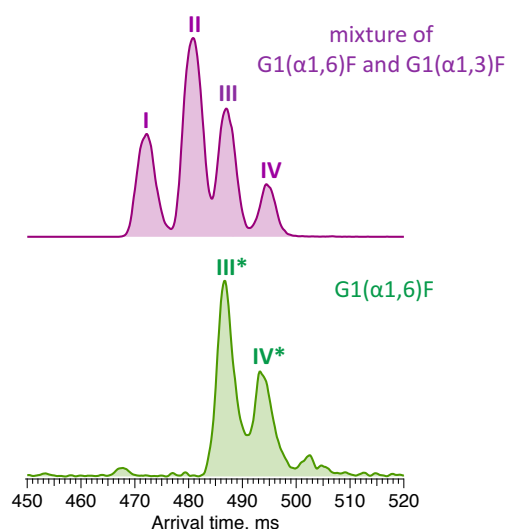


Figure 4.8. ATDs of the G1F glycan representing a mixture of G1(α1,6)F and G1(α1,3)F positional isomers (purple) and the synthesized G1(α1,6)F isomer (green) after a 60 m SLIM-IMS separation.

To confirm that peaks **III\*** and **IV\*** in the ATD of G1( $\alpha$ 1,6)F correspond to peaks **III** and **IV** in the ATD of the G1F mixture, we measured and compared their respective IR spectra. The good agreement between the IR fingerprints of **III\*/III** and **IV\*/IV** shown in Figure 4.9 indeed confirms our assignment.

Based on the selectivity of the chemoenzymatic synthesis of the G1( $\alpha$ 1,6)F isomer, these results demonstrate that peaks **III** and **IV** of the ATD of the G1F mixture correspond to G1( $\alpha$ 1-6)F, and thus peaks **I** and **II** must correspond to G1( $\alpha$ 1,3)F. It is worth noting that in the ATD measured here by ion mobility, G1( $\alpha$ 1,3)F arrives prior to G1( $\alpha$ 1,6)F, while using HILIC, the opposite is observed.<sup>34</sup> This difference results from the different mechanisms of separation. In HILIC, G1( $\alpha$ 1,3)F has a higher affinity to the stationary phase and elutes after G1( $\alpha$ 1,6)F. In IMS, G1( $\alpha$ 1,3)F has a more compact structure and therefore undergoes fewer collisions with the buffer gas, which makes it arrive sooner than G1( $\alpha$ 1,6)F.<sup>50</sup>

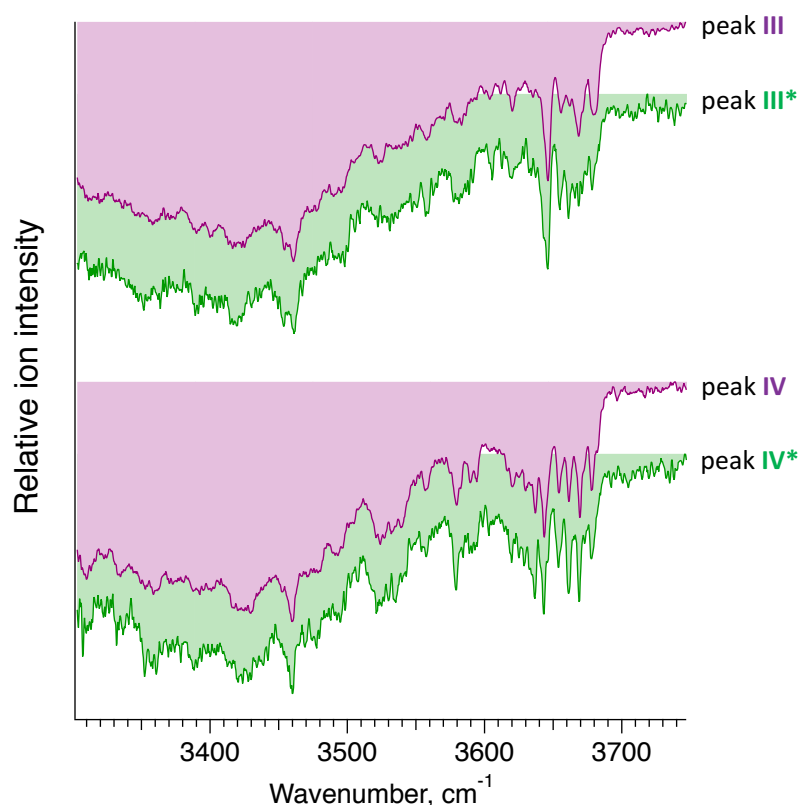


Figure 4.9. Cryogenic IR spectra of peaks **III** vs **III\*** and **IV** vs **IV\*** observed in the ATDs of G1F, representing a mixture of the G1( $\alpha$ 1,6)F and G1( $\alpha$ 1,3)F positional isomers (purple), and the synthesized G1( $\alpha$ 1,6)F isomer (green) (Figure 4.8).

## 4.5. Impact of the host cell line on the glycan profile

Different biosimilar suppliers use independently derived cell lines and different manufacturing processes that affect the glycosylation profile and thus potentially the quality, safety, and efficacy of the final protein product. In the case of G1F, we demonstrate the impact of the host cell line (CHO and HEK-293) on its ATD profile, which reflects the ratio of positional isomers.

Recombinant IgG glycoproteins from CHO and HEK-293 cell lines were produced and purified at the EPFL protein production and structure core facility. One hundred  $\mu\text{g}$  of lyophilized mAb was diluted in 16  $\mu\text{l}$  of  $\text{H}_2\text{O}$  and incubated with 4  $\mu\text{l}$  of Rapid PNGase F buffer and 1  $\mu\text{l}$  of PNGaseF enzyme (BioConcept, Allschwil, Switzerland) at 50 °C for 10 minutes to detach the N-linked glycans.

To remove peptides, protein and other contaminants present in the glycan samples after digestion, we used Sep-Pak C18 cartridges. The cartridges were first conditioned with 1 mL of MeOH, followed by 1 mL of 5% acetic acid three times before the glycan samples were applied directly after the deglycosylation step. Then the cartridges were washed three times with 1 mL of 5% acetic acid. The collected fractions with released N-linked glycans were lyophilized and diluted in 80  $\mu\text{l}$  of a 30/70  $\text{H}_2\text{O}$ /acetonitrile mixture. The mixtures of released N-linked glycans from IgG produced in CHO and HEK-293 cell lines were analyzed by UPLC-MS (Figure 4.10). The obtained glycoprofiles for both mAbs are consistent with previously published data. They include three major glycans: G0F, G1F and G2F.<sup>51, 52</sup> The glycan G1F appears as a single peak consisting of two positional isomers: G1( $\alpha$ 1,6)F and G1( $\alpha$ 1,3)F. Using the fraction collection procedure explained above, the peak corresponding to G1F cleaved from four portions of 100  $\mu\text{g}$  of each mAb was collected and concentrated. The sample solutions were stored at -20 °C. We added sodium acetate to all collected fractions to observe the doubly sodiated adduct as the predominant species for ion mobility and spectroscopy measurements.

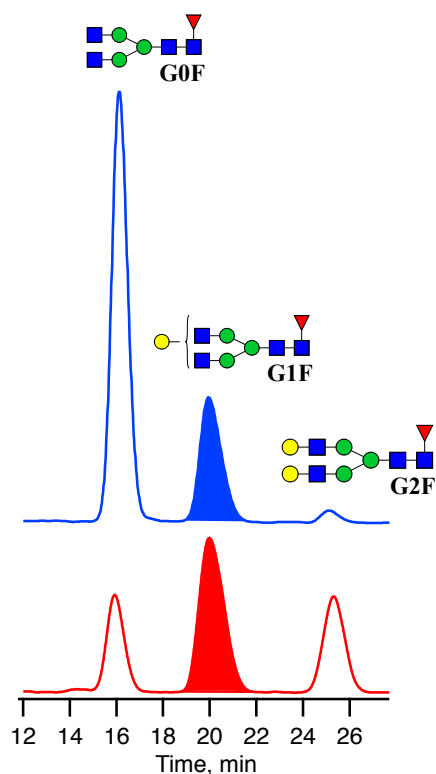


Figure 4.10. Chromatograms of three major N-linked glycans released from IgGs produced in CHO (blue) and HEK-293 (red) cell lines.

Figure 4.11 shows the ATD profile of the G1F standard together with those of G1F released from IgG antibodies produced in either CHO or HEK-293 cell lines. While the ATDs look quite similar in the number of features and their respective drift times, demonstrating the presence of both positional isomers in all three samples, their ratio changes. Based upon the integrated area of each pair of peaks corresponding to the respective positional isomers, the ratio of G1( $\alpha$ 1,3)F to G1( $\alpha$ 1,6)F increases in the following order: standard (65/35) < HEK-293 (70/30) < CHO (75/25) (Figure 4.12, Table 4.2). In addition to the different ratios of positional isomers, we observe a difference in the relative intensities of the two peaks in the ATD assigned to each isomer. As discussed above, these peaks could arise from the two reducing-end anomers or from different conformations for each isomer.

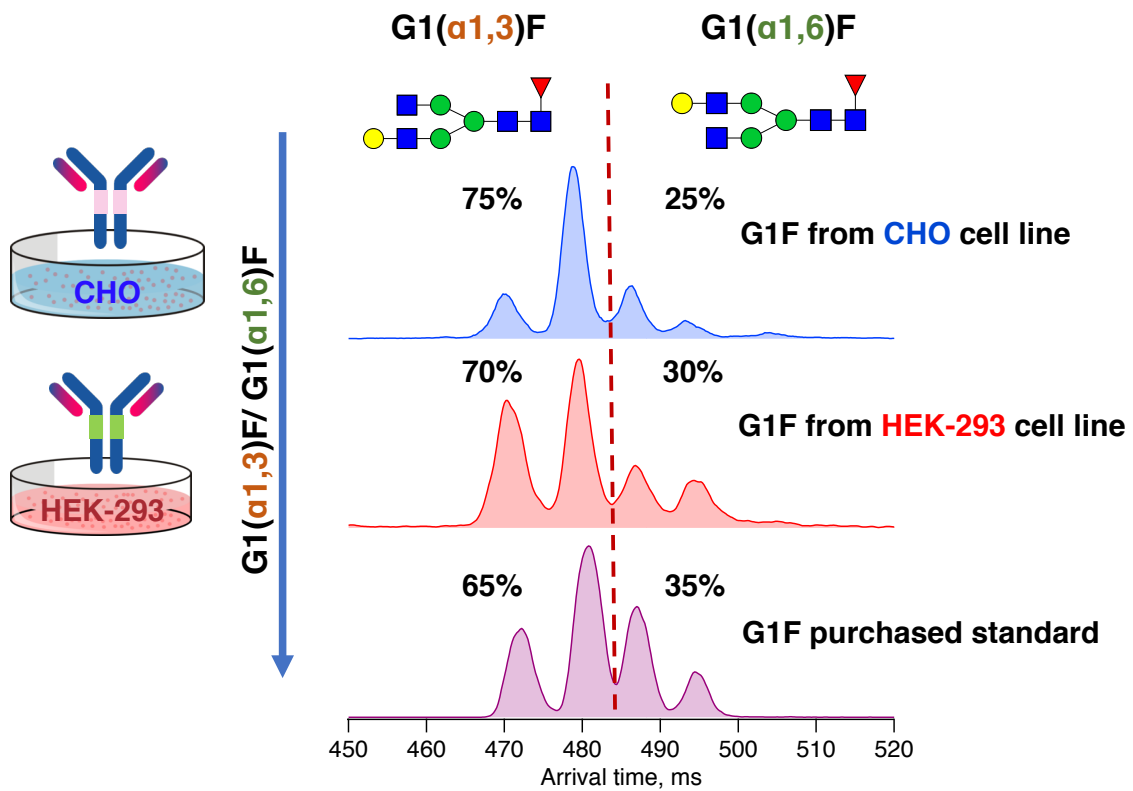


Figure 4.11. ATDs of the G1F standard (purple) and G1F released from a recombinant IgG antibody produced in the HEK-293 (red) cell lines and CHO (blue) after a 60 m SLIM-IMS separation.

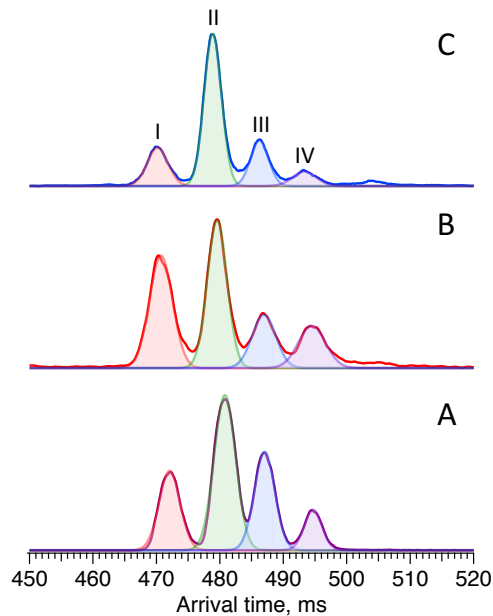


Figure 4.12. Integrated peak areas after gaussian peak fitting of each signal in the ATDs of the G1F standard (A), G1F from the HEK-293 cell line (B) and G1F from the CHO cell line (C).

Table 4.2. Integrated and normalized peak areas of ATD signals of the G1F standard, G1F from the HEK-293 cell line and G1F from the CHO cell line

G1F	Integrated peak areas, %				G1( $\alpha$ 1,3)F/G1( $\alpha$ 1,6)F
	G1( $\alpha$ 1,3)F		G1( $\alpha$ 1,6)F		
	Peak I	Peak II	Peak III	Peak IV	
Standard	22	43	26	9	65/35
HEK-293	33	37	16	14	70/30
CHO	17	58	18	7	75/25

## 4.6. IR spectral comparison using PCA analysis

The IR spectra were compared using an algorithm that combines principal component analysis (PCA) and automatic cluster detection to identify and assign the positional isomers of G1F. The algorithm makes use of the ‘scikit-learn’ python software library using both the ‘PCA’ fit function from the decomposition library and the ‘kmeans’ fit and predict functions from the cluster library.<sup>53</sup> As a first step, PCA is used to reduce the dimensionality of the considered IR spectra to a few principal components that sufficiently describe the original data set (in this case 3 principal components). Following this, the different reference spectra are automatically identified and classified into so-called ‘clusters’ in the principle-component space. Finally, the predict function automatically assigns the analyte spectra to the corresponding references. This approach allows for rapid and accurate identification and assignment of the IR spectra to the different isomers of the considered molecule. Moreover, no user interpretation is required, making the analysis fully automated and thus putting machine learning at the service of glycan analysis.

The PCA and clustering algorithm correctly assigns the different IR spectra of the references (3 spectra per isomer) to the same cluster, depicted in the same color in Figure 4.13. The spectra in the principle-component space are represented by colored circles where peaks 1 and 2 of the ATD of G1( $\alpha$ 1,3)F isomer are depicted in different shades of red and peaks 1 and 2 of the G1( $\alpha$ 1,6)F are depicted in different shades of blue. The explained variance ratios of the principle components describe the percentage of variance that is attributed to each of the components and are shown in Table 4.3.

Table 4.3. The variance ratios of the principle components

Principle components	First	Second	Third
Explained variance ratio	0.55	0.28	0.08

It is important to note that the different colors corresponding to different isomers in Figure 4.13 were automatically assigned by the algorithm through the ‘kmeans’ *fit* function, which classifies the data set into 4 separate clusters.

Finally, the analytes, depicted by a black star for peak 1 of the ATD of the synthesized molecule and by a black triangle for peak 2 of the same ATD are assigned to the corresponding references using the ‘kmeans’ *predict* function. The resulting assignment confirms that the observed ATD peaks correspond indeed to the G1( $\alpha$ 1,6)F isomers. The probability that “Analyte peak ATD peak 1” belongs to the cluster of “Isomer G1( $\alpha$ 1,6)F peak 1” is 92% and “Analyte peak ATD peak 2” belongs to the cluster of “Isomer G1( $\alpha$ 1,6)F peak 2” is 89%. This will depend slightly on the number of spectra that are used for each database entry.

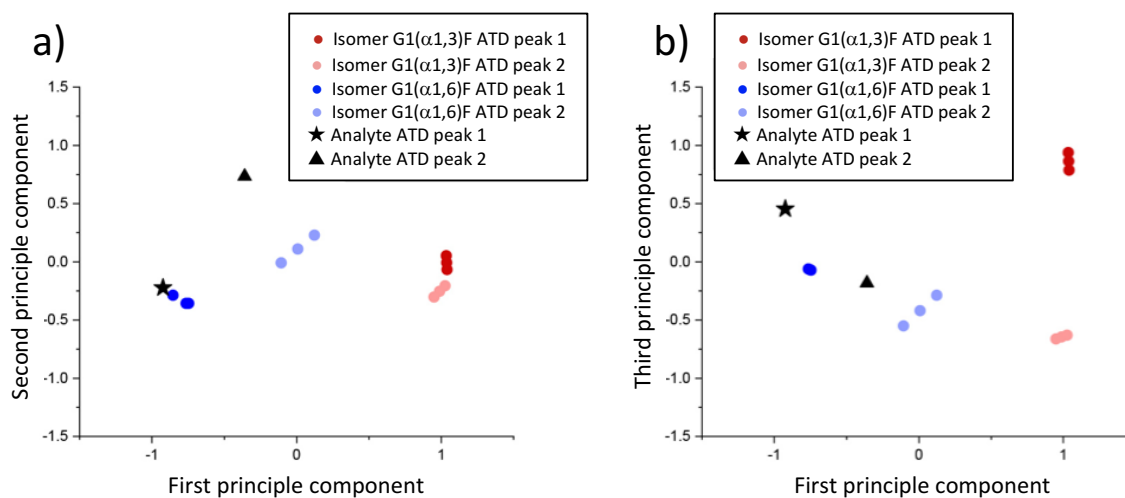


Figure 4.13. Scatter plot of the PCA and clustering results for the different isomers of G1F. (a) First and second principle components, (b) first and third principle components.

## Conclusions

Our results demonstrate that the combination of ultra-high resolution ion mobility separation with cryogenic vibrational spectroscopy represents a rapid and reliable analytical method to distinguish positional isomers of glycans. We used a chemoenzymatic approach with the FUT8 enzyme for the selective synthesis of G1( $\alpha$ 1,6)F which we employed as a standard for assigning the mobility-separated positional isomers of G1F based on their unique IR fingerprint spectra. With this assignment, we can now identify G1F positional isomers in subsequent experiments based on their vibrational fingerprints alone. One should note that the observed isomers differ by as little as 0.2% in their collision cross section, emphasizing the difficulty to distinguish them by IMS-MS methods alone. We then applied this technique for determining the difference in the ratio of the individual G1F positional isomers from mAbs that have been produced in two different cell lines: CHO and HEK-293. The higher content of the G1( $\alpha$ 1,6)F isomer in the IgG from the HEK-293 cell line may lead to its higher complement-dependent cytotoxicity compared to the IgG from the CHO cell line. This work demonstrates the power of our approach for monitoring subtle differences in galactosylation of N-glycans cleaved from mAbs.



# References

1. Dyukova, I.; Ben Faleh, A.; Warnke, S.; Yalovenko, N.; Yatsyna, V.; Bansal, P.; Rizzo, T. R., A new approach for identifying positional isomers of glycans cleaved from monoclonal antibodies. *Analyst* **2021**, *146* (15), 4789-4795.
2. Patel, P. K.; King, C. R.; Feldman, S. R., Biologics and biosimilars. *J Dermatol Treat* **2015**, *26* (4), 299-302.
3. Henry, D.; Taylor, C., Pharmacoeconomics of cancer therapies: considerations with the introduction of biosimilars. *Semin Oncol* **2014**, *41 Suppl 3*, S13-20.
4. Smolen, J. S.; Landewe, R.; Bijlsma, J.; Burmester, G.; Chatzidionysiou, K.; Dougados, M.; Nam, J.; Ramiro, S.; Voshaar, M.; van Vollenhoven, R.; Aletaha, D.; Aringer, M.; Boers, M.; Buckley, C. D.; Buttgerit, F.; Bykerk, V.; Cardiel, M.; Combe, B.; Cutolo, M.; van Eijk-Hustings, Y.; Emery, P.; Finckh, A.; Gabay, C.; Gomez-Reino, J.; Gossec, L.; Gottenberg, J. E.; Hazes, J. M. W.; Huizinga, T.; Jani, M.; Karateev, D.; Kouloumas, M.; Kvien, T.; Li, Z.; Mariette, X.; McInnes, I.; Mysler, E.; Nash, P.; Pavelka, K.; Poor, G.; Richez, C.; van Riel, P.; Rubbert-Roth, A.; Saag, K.; da Silva, J.; Stamm, T.; Takeuchi, T.; Westhovens, R.; de Wit, M.; van der Heijde, D., EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* **2017**, *76* (6), 960-977.
5. Heinemann, L., Biosimilar Insulin and Costs: What Can We Expect? *J Diabetes Sci Technol* **2015**, *10* (2), 457-62.
6. Simoens, S.; Jacobs, I.; Popovian, R.; Isakov, L.; Shane, L. G., Assessing the Value of Biosimilars: A Review of the Role of Budget Impact Analysis. *Pharmacoeconomics* **2017**, *35* (10), 1047-1062.
7. Heredia, E.; Ribeiro, A., Discounts Offered by First and Subsequent Biosimilars in the Us, Eu and Latam: Impact Trends of Originator Starting Price, Market Dynamics and Regulations. *Value Health* **2018**, *21*, S103-S104.
8. Moorkens, E.; Godman, B.; Huys, I.; Hoxha, I.; Malaj, A.; Keuerleber, S.; Stockinger, S.; Mortenhuber, S.; Dimitrova, M.; Tachkov, K.; Voncina, L.; Palcevski, V. V.; Achiotou, G.; Slaby, J.; Popelkova, L.; Kohoutova, K.; Bartels, D.; Laius, O.; Martikainen, J. E.; Selke, G. W.; Kourafalos, V.; Magnusson, E.; Einarsdottir, R.; Adams, R.; Joppi, R.; Allocati, E.; Jakupi, A.; Viksna, A.; Greiciute-Kuprijanov, I.; Vella Bonanno, P.; Suttorp, V.; Melien, O.; Plisko, R.; Mardare, I.; Meshkov, D.; Novakovic, T.; Furst, J.;

- Zara, C.; Markovic-Pekovic, V.; Grubisa, N.; Befrits, G.; Puckett, R.; Vulto, A. G., The Expiry of Humira((R)) Market Exclusivity and the Entry of Adalimumab Biosimilars in Europe: An Overview of Pricing and National Policy Measures. *Front Pharmacol* **2020**, *11*, 591134.
9. Derbyshire, M., Patent expiry dates for biologicals: 2017 update. *Gabi J* **2018**, *7* (1), 29-34.
  10. Derbyshire, M.; Shina, S., Patent expiry dates for biologicals: 2018 update. *Gabi J* **2019**, *8* (1), 24-31.
  11. Beck, A.; Wagner-Rousset, E.; Bussat, M. C.; Lokteff, M.; Klinguer-Hamour, C.; Haeuw, J. F.; Goetsch, L.; Wurch, T.; Van Dorsselaer, A.; Corvaia, N., Trends in Glycosylation, Glycoanalysis and Glycoengineering of Therapeutic Antibodies and Fc-Fusion Proteins. *Curr Pharm Biotechno* **2008**, *9* (6), 482-501.
  12. Wright, A.; Morrison, S. L., Effect of glycosylation on antibody function: implications for genetic engineering. *Trends Biotechnol* **1997**, *15* (1), 26-32.
  13. Awwad, S.; Angkawinitwong, U., Overview of Antibody Drug Delivery. *Pharmaceutics* **2018**, *10* (3).
  14. Raju, T. S., Terminal sugars of Fc glycans influence antibody effector functions of IgGs. *Curr Opin Immunol* **2008**, *20* (4), 471-8.
  15. Dashivets, T.; Thomann, M.; Rueger, P.; Knaupp, A.; Buchner, J.; Schlothauer, T., Multi-Angle Effector Function Analysis of Human Monoclonal IgG Glycovariants. *PLoS One* **2015**, *10* (12), e0143520.
  16. Scallon, B. J.; Tam, S. H.; McCarthy, S. G.; Cai, A. N.; Raju, T. S., Higher levels of sialylated Fc glycans in immunoglobulin G molecules can adversely impact functionality. *Mol Immunol* **2007**, *44* (7), 1524-34.
  17. Kaneko, Y.; Nimmerjahn, F.; Ravetch, J. V., Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. *Science* **2006**, *313* (5787), 670-3.
  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), 1644-52.
  19. 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.

20. Dekkers, G.; Plomp, R.; Koeleman, C. A.; Visser, R.; von Horsten, H. H.; Sandig, V.; Rispens, T.; Wuhrer, M.; Vidarsson, G., Multi-level glyco-engineering techniques to generate IgG with defined Fc-glycans. *Sci Rep* **2016**, *6*, 36964.
21. 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 Fcγ3R-Immobilized Column. *Sci Rep* **2018**, *8* (1), 3955.
22. Houde, D.; Peng, Y.; Berkowitz, S. A.; Engen, J. R., Post-translational modifications differentially affect IgG1 conformation and receptor binding. *Mol Cell Proteomics* **2010**, *9* (8), 1716-28.
23. Thomann, M.; Reckermann, K.; Reusch, D.; Prasser, J.; Tejada, M. L., Fc-galactosylation modulates antibody-dependent cellular cytotoxicity of therapeutic antibodies. *Mol Immunol* **2016**, *73*, 69-75.
24. 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 alpha1-6 arm of Fc-glycan on the effector functions of therapeutic monoclonal antibodies. *MAbs* **2019**, *11* (5), 826-836.
25. Marino, K.; Bones, J.; Kattla, J. J.; Rudd, P. M., A systematic approach to protein glycosylation analysis: a path through the maze. *Nat Chem Biol* **2010**, *6* (10), 713-23.
26. Beck, A.; Sanglier-Cianferani, S.; Van Dorsselaer, A., Biosimilar, biobetter, and next generation antibody characterization by mass spectrometry. *Anal Chem* **2012**, *84* (11), 4637-46.
27. Geist, B. J.; Davis, D.; McIntosh, T.; Yang, T. Y.; Goldberg, K.; Han, C.; Pendley, C.; Davis, H. M., A novel approach for the simultaneous quantification of a therapeutic monoclonal antibody in serum produced from two distinct host cell lines. *MAbs* **2013**, *5* (1), 150-61.
28. 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), 1370-80.
29. 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-FRLR-IMS-MS Workflow. *Anal Chem* **2020**, *92* (23), 15323-15335.

30. Barb, A. W.; Prestegard, J. H., NMR analysis demonstrates immunoglobulin G N-glycans are accessible and dynamic. *Nat Chem Biol* **2011**, *7* (3), 147-53.
31. Mizuochi, T.; Hamako, J.; Titani, K., Structures of the sugar chains of mouse immunoglobulin G. *Arch Biochem Biophys* **1987**, *257* (2), 387-94.
32. 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), 1137-44.
33. Yamaguchi, Y.; Kato, K.; Shindo, M.; Aoki, S.; Furusho, K.; Koga, K.; Takahashi, N.; Arata, Y.; Shimada, I., Dynamics of the carbohydrate chains attached to the Fc portion of immunoglobulin G as studied by NMR spectroscopy assisted by selective <sup>13</sup>C labeling of the glycans. *J Biomol NMR* **1998**, *12* (3), 385-94.
34. Reusch, D.; Habegger, M.; Maier, B.; Maier, M.; Kloseck, R.; Zimmermann, B.; Hook, M.; Szabo, Z.; Tep, S.; Wegstein, J.; Alt, N.; Bulau, P.; Wuhler, M., Comparison of methods for the analysis of therapeutic immunoglobulin G Fc-glycosylation profiles--part 1: separation-based methods. *MAbs* **2015**, *7* (1), 167-79.
35. 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), 4027-31.
36. 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.
37. Bansal, P.; Ben Faleh, A.; Warnke, S.; Rizzo, T. R., Identification of N-glycan positional isomers by combining IMS and vibrational fingerprinting of structurally determinant CID fragments. *Analyst* **2022**, *147* (4), 704-711.
38. Rudd, P. M.; Dwek, R. A., Glycosylation: heterogeneity and the 3D structure of proteins. *Crit Rev Biochem Mol Biol* **1997**, *32* (1), 1-100.
39. Stanley, P.; Taniguchi, N.; Aebi, M., N-Glycans. In *Essentials of Glycobiology*, rd; 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 (NY), 2015; pp 99-111.
40. Boltje, T. J.; Buskas, T.; Boons, G. J., Opportunities and challenges in synthetic oligosaccharide and glycoconjugate research. *Nat Chem* **2009**, *1* (8), 611-22.

41. Walczak, M. A.; Hayashida, J.; Danishefsky, S. J., Building biologics by chemical synthesis: practical preparation of di- and triantennary N-linked glycoconjugates. *J Am Chem Soc* **2013**, *135* (12), 4700-3.
42. Unverzagt, C.; Gundel, G.; Eller, S.; Schuberth, R.; Seifert, J.; Weiss, H.; Niemietz, M.; Pischl, M.; Raps, C., Synthesis of multiantennary complex type N-glycans by use of modular building blocks. *Chemistry* **2009**, *15* (45), 12292-302.
43. Wu, B.; Hua, Z. H.; Warren, J. D.; Ranganathan, K.; Wan, Q.; Chen, G.; Tan, Z. P.; Chen, J. H.; Endo, A.; Danishefsky, S. J., Synthesis of the fucosylated biantennary N-glycan of erythropoietin. *Tetrahedron Lett* **2006**, *47* (31), 5577-5579.
44. Palcic, M. M., Glycosyltransferases as biocatalysts. *Curr Opin Chem Biol* **2011**, *15* (2), 226-233.
45. Schmaltz, R. M.; Hanson, S. R.; Wong, C. H., Enzymes in the synthesis of glycoconjugates. *Chem Rev* **2011**, *111* (7), 4259-307.
46. Unverzagt, C.; Kajihara, Y., Chemical assembly of N-glycoproteins: a refined toolbox to address a ubiquitous posttranslational modification. *Chem Soc Rev* **2013**, *42* (10), 4408-20.
47. Wang, Z.; Chinoy, Z. S.; Ambre, S. G.; Peng, W.; McBride, R.; de Vries, R. P.; Glushka, J.; Paulson, J. C.; Boons, G. J., A general strategy for the chemoenzymatic synthesis of asymmetrically branched N-glycans. *Science* **2013**, *341* (6144), 379-83.
48. Gagarinov, I. A.; Li, T. H.; Torano, J. S.; Caval, T.; Srivastava, A. D.; Kruijtzter, J. A. W.; Heck, A. J. R.; Boons, G. J., Chemoenzymatic Approach for the Preparation of Asymmetric Bi-, Tri-, and Tetra-Antennary N-Glycans from a Common Precursor. *Journal of the American Chemical Society* **2017**, *139* (2), 1011-1018.
49. Li, L.; Liu, Y. P.; Ma, C.; Qu, J. Y.; Calderon, A. D.; Wu, B. L.; Wei, N.; Wang, X.; Guo, Y. X.; Xiao, Z. Y.; Song, J.; Sugiarto, G.; Li, Y. H.; Yu, H.; Chen, X.; Wang, P. G., Efficient chemoenzymatic synthesis of an N-glycan isomer library. *Chem Sci* **2015**, *6* (10), 5652-5661.
50. Re, S.; Watabe, S.; Nishima, W.; Muneyuki, E.; Yamaguchi, Y.; MacKerell, A. D., Jr.; Sugita, Y., Characterization of Conformational Ensembles of Protonated N-glycans in the Gas-Phase. *Sci Rep* **2018**, *8* (1), 1644.
51. Nallet, S.; Fornelli, L.; Schmitt, S.; Parra, J.; Baldi, L.; Tsybin, Y. O.; Wurm, F. M., Glycan variability on a recombinant IgG antibody transiently produced in HEK-293E cells. *N Biotechnol* **2012**, *29* (4), 471-6.
52. Monteil, D. T.; Juvet, V.; Paz, J.; Moniatte, M.; Baldi, L.; Hacker, D. L.; Wurm, F. M., A comparison of orbitally-shaken and stirred-tank bioreactors: pH modulation and

bioreactor type affect CHO cell growth and protein glycosylation. *Biotechnol Prog* **2016**, *32* (5), 1174-1180.

53. Stow, S. M.; Causon, T. J.; Zheng, X.; Kurulugama, R. T.; Mairinger, T.; May, J. C.; Rennie, E. E.; Baker, E. S.; Smith, R. D.; McLean, J. A.; Hann, S.; Fjeldsted, J. C., An Interlaboratory Evaluation of Drift Tube Ion Mobility-Mass Spectrometry Collision Cross Section Measurements. *Anal Chem* **2017**, *89* (17), 9048-9055.



# Chapter 5. Searching for new N-glycan biomarkers

## 5.1. Motivation and objectives

Cancer is a major cause of death worldwide. The statistics reveal that the highest cancer mortality levels are found among patients with breast, prostate, lung, colon, stomach or liver cancers.<sup>1</sup> Breast cancer is the leading cause of cancer death (16% of all cancer cases) among women in Europe ([www.europadonna.org](http://www.europadonna.org)). The discovery of reliable biomarkers for various cancers will allow early-stage detection, monitoring of disease progression and efficient therapy. All these factors, taken together, will enhance the survival rate.

Despite the importance of early-stage detection, screening rates are still low. Current tests based on traditional tissue biopsies suffer from their invasive and uncomfortable nature, their high cost, and their lack of specificity or sensitivity. Body fluids (urine, sputum, blood serum and plasma) offer a minimally invasive source of new molecular biomarkers that promise to overcome the main issues of the traditional screening methods.

Changes in glycosylation have been proven to be one of the hallmarks of cancer. Glycans are involved in numerous pathological processes, including cancer progression and metastasis.<sup>2-6</sup> The N-linked serum glycome has been shown to be different between normal and cancer cells due to dysregulated glycosyltransferases and glycosidases. These differences are reflected in the degree of glycan branching, truncation, mannose content, core fucosylation and sialylation levels. N-glycan structures from serum glycoproteins of breast,<sup>7-9</sup> prostate,<sup>10</sup> ovarian,<sup>11</sup> colorectal,<sup>12</sup> liver,<sup>13</sup> and lung<sup>14</sup> cancer patients have already been characterized. N-glycan markers from serum represent a promising, non-invasive group of novel cancer biomarkers.

Sialic acids (SAs) are an important family of monosaccharides when it comes to cancer-associated glycan changes. They include more than 50 forms and the hydroxyl groups at the positions 4, 7, 8, and 9 can be methylated or esterified, leading to an even larger variety of forms. Sialic acid is also the name used for the most abundant member of this group, Neu5Ac (Figure 5.1). These negatively charged and hydrophilic units are usually attached to galactose residues at the end of the glycan through an  $\alpha$ -2,3 or  $\alpha$ -2,6 glycosidic linkage in humans. Due to their special location and wide distribution, sialylated N-glycans are involved in a variety of



physiological and pathological processes.<sup>15</sup> SAs interact with carbohydrate-binding proteins with high specificity and selectivity. Changes in the expression of the  $\alpha$ -2,3- and  $\alpha$ -2,6-linkage isomers are recognized to be signs of different stages of cancer.<sup>16-19</sup> Therefore, the analysis of sialyl-linkage isomers can be used to monitor the progression of the disease.

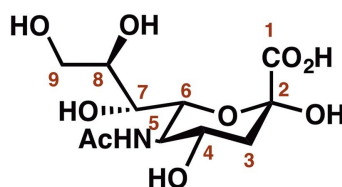


Figure 5.1. Structure of the most common sialic acid monosaccharide – N-acetylneuraminic acid (Neu5Ac).


Classical methods for analyzing sialylated glycans include exoglycosidase digestion and lectin binding. While these methods have proven to be of value, both also have several drawbacks. Enzymatic cleavage is a time-consuming method that requires a large amount of purified sample. In turn, lectins often suffer from low binding affinity and limited specificity. In recent years, the combination of mass spectrometry with HILIC or CE has been implemented for the analysis of sialyl-linkage glycan isomers. Glycans are analyzed either derivatized or in their native form. Both methods are widely used to study the sialylated glycan content, and the choice depends on the experiential setup of the laboratory. Permethylated is used with MALDI ionization to prevent the loss of labile sialic acid groups and to improve the sensitivity of MS. However, the process of permethylation can lead to glycan degradation, and partially methylated products can produce a chemical noise.

Thus, the identification and quantification of sialylated glycans from serum at the isomer level remain challenging due to their structural complexity, low abundance and high heterogeneity. Therefore, the development of a new robust method is greatly needed to determine sialyl-linkage and positional isomers. We demonstrate a method for the analysis of native sialylated N-glycans and for finding new glycan biomarkers at the isomer level based on a four-dimensional approach, which combines MS, IMS, IR and on-board CID techniques.

## 5.2. Sialylated N-glycans from human blood serum

In this study, we analyzed the N-glycans in serum samples from healthy female donors (control samples) and from patients diagnosed with different stages of breast cancer, specifically invasive ductal carcinoma. The samples were purchased from the Central BioHub GmbH (Germany). It is well known that glycan structures are affected by both pathological and physiological conditions in the human body. In order to minimize the difference in glycosylation profiles associated with the physiological characteristics of each individual, we selected healthy and sick donors with the same principal parameters, including race, age and body mass index (BMI) (Table 5.1). The glycan standards were purchased from Dextra (UK), TheraProteins (Portugal) and Agilent (Switzerland). All samples were stored at  $-20\text{ }^{\circ}\text{C}$ .

Table 5.1. Serum samples



Serum donor	Stage	Treatment	Gender	Diagnosis	Race	Age (y.o.)	BMI (kg/m <sup>2</sup> )
<b>1</b>	<b>Control</b>	Healthy	Female	Invasive ductal carcinoma	Caucasian	54	30
<b>2</b>							
<b>I</b>	<b>0</b>	Treated					
<b>II</b>	<b>IA</b>	Not treated					
<b>III</b>	<b>IIA</b>						
<b>IV</b>	<b>IIB</b>						

Ten  $\mu\text{l}$  of the serum sample from each donor were denatured by incubation with ammonium carbonate ( $55\text{ }\mu\text{l}$ ,  $100\text{ mM}$ ) and dithiothreitol ( $2\text{ }\mu\text{l}$ ,  $1\text{ M}$ ) for 1h at  $60\text{ }^{\circ}\text{C}$ . The sample was then cooled down to room temperature and incubated with iodoacetamide ( $3\text{ }\mu\text{l}$ ,  $1\text{ M}$ ) in a dark place for 1h. Finally, the reaction mixture was digested with  $1\text{ }\mu\text{l}$  of PNGaseF enzyme (BioConcept, Allschwil, Switzerland) at  $50\text{ }^{\circ}\text{C}$  for 2 h and left overnight at  $37\text{ }^{\circ}\text{C}$  to allow for complete detachment of the N-linked glycans. After the enzymatic reaction, the mixture was centrifuged, and the precipitated part with proteins and other contaminants contained in the serum was discarded. In turn, the solution with the released glycan pool was subjected to a cleaning procedure with Amicon centrifuge filters ( $10\text{ kDa}$  MW cut-off) followed by Sep-Pak C18 cartridges. The mixture of the released N-linked glycans was then analyzed by UPLC-MS.

N-glycans containing sialic acid were found to be the most abundant oligosaccharides in the serum samples. Using a short LC method and the fraction collection procedure explained above, glycans with 1–3 terminal sialic acid monosaccharides (Table 5.2) were collected together and concentrated for further measurements on the IMS-CIS II instrument. The serum

N-glycan profiles are dominated by glycans extracted from the most abundant glycoproteins contained in the blood serum. These include IgG, transferrin, haptoglobin, IgA, alpha-1-antitrypsin and alpha-2-macroglobulin.

Table 5.2. Structures of the most abundant sialylated N-linked glycans released from the serum samples

	Glycan structure	Charge state	m/z
A1		[A1-H] <sup>-</sup>	1930
A1F		[A1F-H] <sup>-</sup>	2078
A2		[A2-2H] <sup>2-</sup>	1111
A2F		[A2F-2H] <sup>2-</sup>	1184
A3(2)		[A2-2H] <sup>2-</sup>	1294
A3(2)F		[A2F-2H] <sup>2-</sup>	1367
A3(3)		[A3-2H] <sup>2-</sup>	1439
A3(3)F		[A3F-2H] <sup>2-</sup>	1512

### 5.3. Identification of glycans released from the serum samples using the spectroscopic database

Sialylated glycans (standards and released from the serum) were reconstituted in a 50:50 water/methanol solution up to a final concentration of 5  $\mu$ M. Since sialic acid can easily deprotonate, measurements were performed in the negative ion mode. This helps to improve sensitivity and to simplify the mass spectrum. Each glycan appears in a single negatively charged state (depending on the number of sialic acid units in the molecule and the pH of the medium) and is not distributed over different charge states as we typically observe in positive ion mode. A crucial factor in observing such large glycan structures (Table 5.2) was a long time (5 s) in the accumulation region of the SLIM device.

The glycan A2 represents the most abundant sialylated glycan released from the serum samples. We thus use it to demonstrate the workflow for identification based on our

spectroscopic database. Glycan A2 has four linkage isomers: two symmetrical structures with sialic acid linked to galactose through an  $\alpha 2,3$  (Figure 5.2A) or  $\alpha 2,6$  bond (Figure 5.2B) on both the  $\alpha 1,6$  and  $\alpha 1,3$  branches of the glycan core; two asymmetrical structures with sialic acid linked through an  $\alpha 2,3$  bond to galactose on the  $\alpha 1,6$  branch and through an  $\alpha 2,6$  bond on the  $\alpha 1,3$  branch (Figure 5.2C), and the mirrored opposite (Figure 5.2D). Our database contains two symmetrical A2 isomers (A and B). After separation by 2 cycles on the SLIM board, each isomer exhibits two peaks, the IR fingerprints of which are illustrated in Figure 5.3. Each spectrum displays a series of unique absorptions across the entire region that are sufficient for identification. The two peaks in the ATD of each linkage isomer could arise from the two reducing-end anomers, which we have shown can be separated by ultra-high resolution IMS,<sup>20</sup> or from two different conformers.

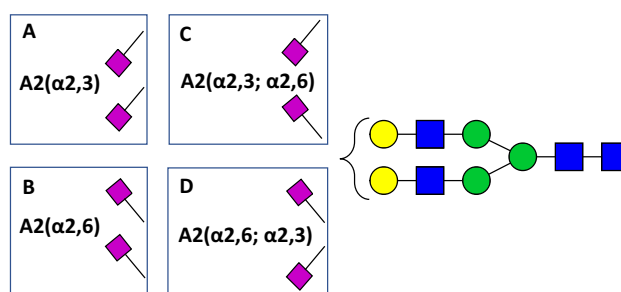


Figure 5.2. Linkage isomers of the glycan A2.

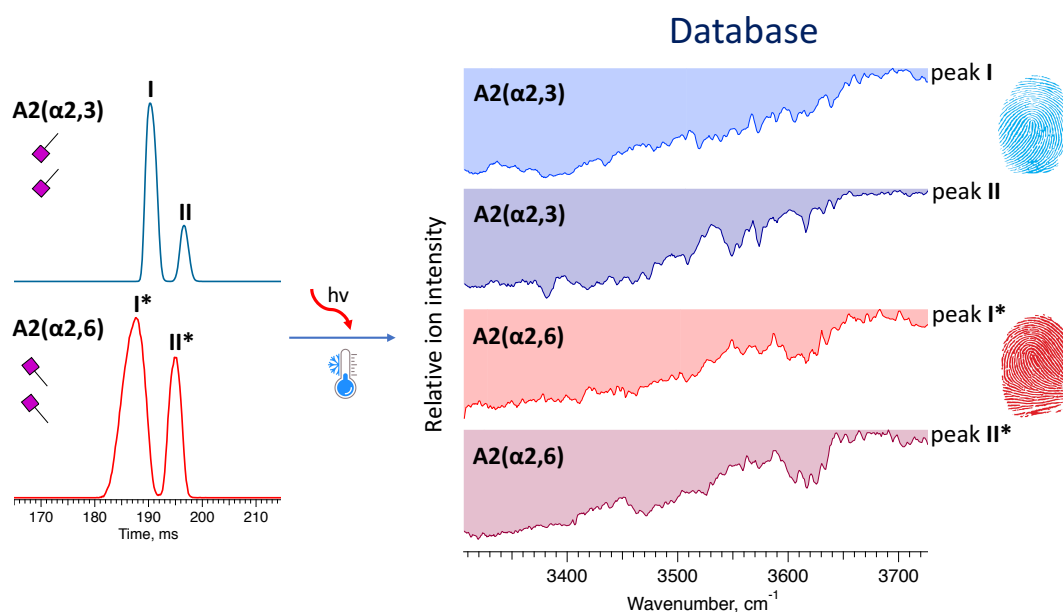


Figure 5.3. Left part – 2-cycle separation of the A2( $\alpha 2,3$ ) and A2( $\alpha 2,6$ ) standards on the SLIM board of IMS-CIS II. Right part – database IR fingerprints of peaks I, II and I\*, II\* in the ATD of the A2( $\alpha 2,3$ ) and A2( $\alpha 2,6$ ) standards, respectively.

Glycans extracted from real biological samples can contain one isomer or a mixture of isomers. Before we applied our procedure to the glycan A2 released from a serum sample, we determined the optimal separation conditions using a synthetic mixture of the A2( $\alpha$ 2,3) and A2( $\alpha$ 2,6) standards. Four peaks of the mixture (two peaks from each isomer) are well resolved after 4 cycles separation on the SLIM board. Based on the comparison of the IR fingerprints of each peak in the mixture with the database spectra of the individual A2( $\alpha$ 2,3) and A2( $\alpha$ 2,6) isomers, we identified peaks I, III to be the A2( $\alpha$ 2,6) isomer and peaks II, IV to correspond to the A2( $\alpha$ 2,3) isomer (Figure 5.4). Even though the spectra look relatively congested and similar (specifically in the case of peaks I and III of A2( $\alpha$ 2,6), the high reproducibility of the measurements allows for precise identification.

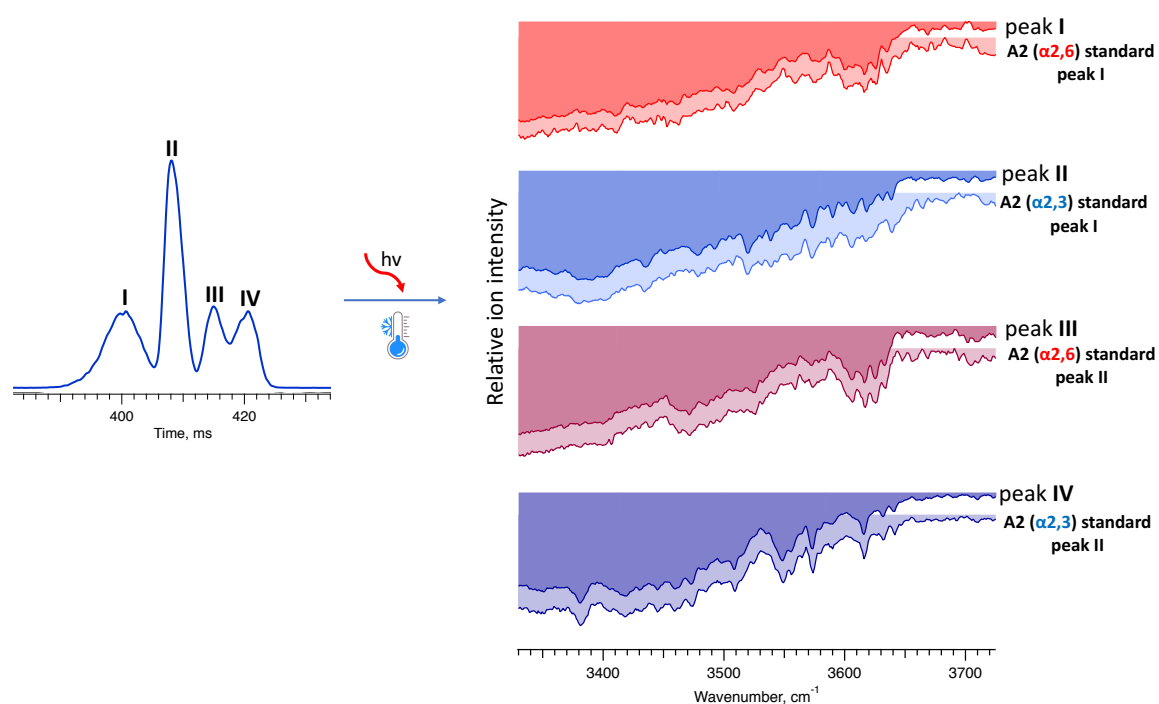


Figure 5.4. Left part – 4 cycles separation of the synthetic mixture of the A2( $\alpha$ 2,3) and A2( $\alpha$ 2,6) standards on the SLIM board of IMS-CIS II. Right part – comparison of the IR fingerprints of each peak in the synthetic mixture with the database spectra of the individual A2( $\alpha$ 2,3) and A2( $\alpha$ 2,6) isomers.

Under the same separation conditions, the glycan A2 released from the Control 1 sample exhibited two peaks in the ATD, and the IR spectra of peaks I\* and II\* correspond to peaks I and II of the A2( $\alpha$ 2,6) isomer, respectively (Figure 5.5). We identified A2 from Control 1 and found that it is only present in the form of one A2( $\alpha$ 2,6) isomer. It has previously been reported that the pathway of A2 biosynthesis in humans is restricted to symmetrical structures and that the  $\alpha$ -2,6-sialylated N-glycans are the predominant species.<sup>21, 22</sup> However, a complete dysregulation of enzymes and a deviation from the usual pathway could lead to isomeric

structures of A2 that are not contained in our database. In this case, we would implement our method for the identification of isomeric structures without the need for standards, which will be explained later in this chapter. Additionally, it is important to note that the same identification procedure can be applied to any glycan (known or unknown) released from different biological samples, including serum, urine, saliva etc.

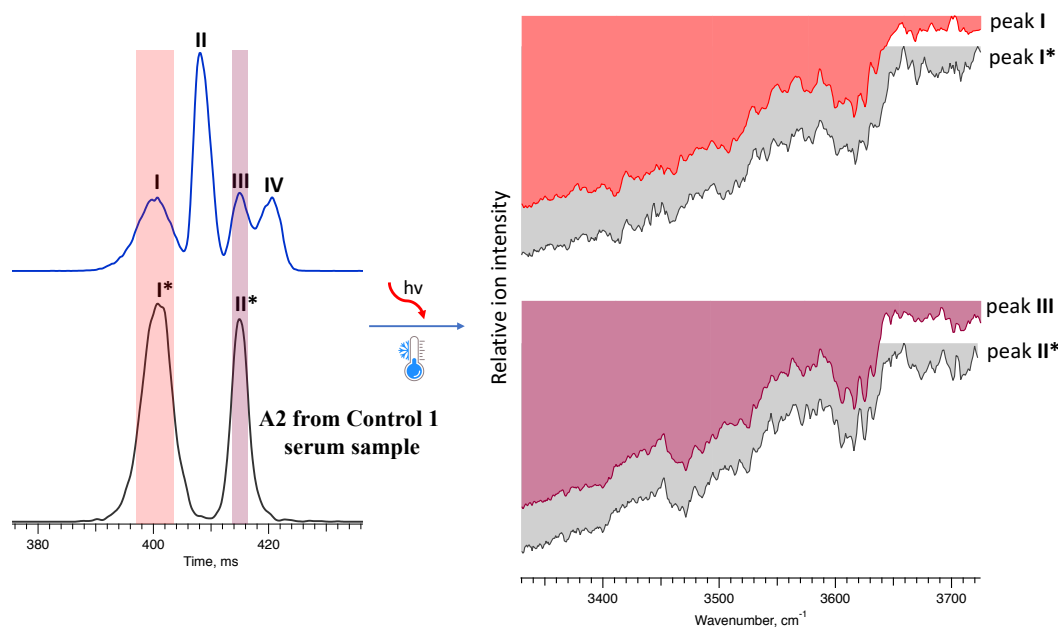


Figure 5.5. Left part – 4 cycles separation of the synthetic mixture of the A2( $\alpha$ 2,3) and A2( $\alpha$ 2,6) standards (blue) and the A2 released from the Control 1 serum sample (grey) on the SLIM board of IMS-CIS II. Right part – the comparison of the IR fingerprints of peaks I and II in the synthetic mixture with the IR spectra of peaks I\* and II\* in the ATD of A2 released from the Control 1 serum sample, respectively.

## 5.4. Comparison of arrival time distributions of glycans collected from healthy and sick donors

We used our ultra-high resolution IMS to separate isomers of each observed glycan and monitor differences in glycosylation profiles between healthy and sick donors at the isomer level. The ATD profiles of A1, A1F and A2 (Table 5.2) showed no difference between all studied samples; these glycans were not subjected to further investigations. Figure 5.6 compares the ATDs of glycans A3(2) and A3(3) extracted from Control 1 and from sick individuals I, II and III. The separation on the SLIM board only revealed the slight differences between samples, and these can come from signal fluctuations. On the other hand, glycan A2F from donor II and glycan A3(2)F from donor III (Figure 5.7) showed significant differences in comparison with the control experiment (highlighted with check marks). Although the number and the location of the features on the ATD did not change in both cases, their relative

intensities are notably different. Nevertheless, the differences were observed only in one donor for each glycan (A2F from Donor II and A3(2)F from Donor III), and there was no consistency between all samples.

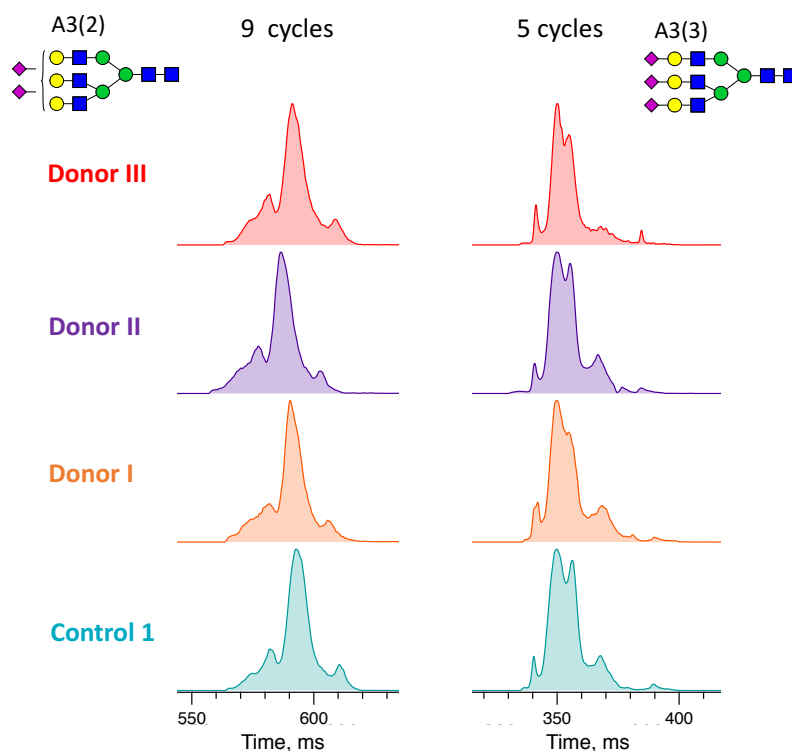


Figure 5.6. Comparison (vertically) of ATDs of glycans A3(2) (left) and A3(3) (right) extracted from the Control 1 sample and the sick individuals I–III.

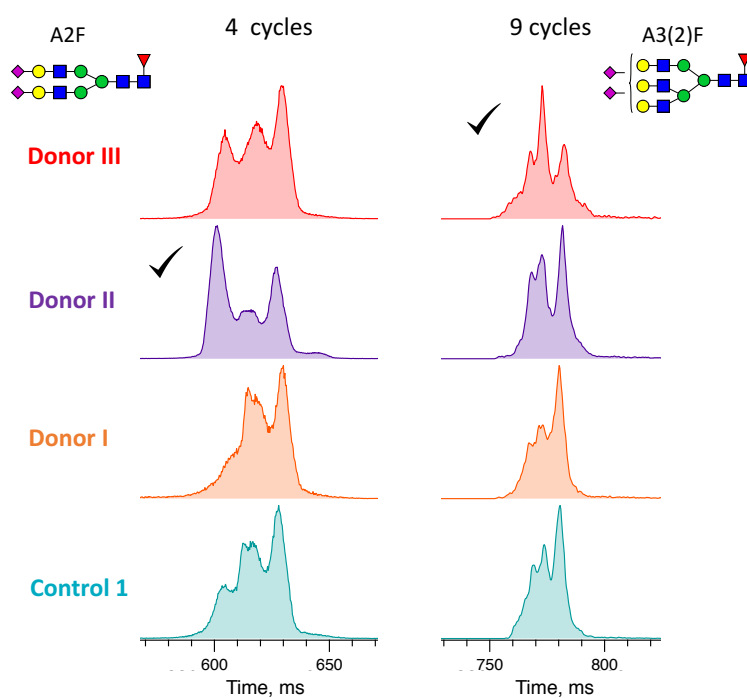


Figure 5.7. Comparison (vertically) of ATDs of glycans A2F (left) and A3(2)F (right) extracted from the Control 1 sample and the sick individuals I–III.

Glycan A3(3)F has up to 16 possibilities for linkage and positional isomers. ATD profiles of A3(3)F released from sick individuals with stage IA (donor II) and stage IIA (donor III) cancer exhibit a much higher content in the first part of the ATD (from the left; specifically, the sharp peak\* at the very beginning) than the A3(3)F from the healthy donor (Control 1) (Figure 5.8). In order to confirm the consistency of our preliminary results, we performed the experiment using two more serum samples: Control 2 and donor IV with stage IIB cancer. We obtained two groups of ATD profiles: the first group includes Control 1, Control 2 and donor I (stage 0), while the second group includes donors II–IV with stages IA, IIA and IIB, respectively. The results demonstrate that the isomeric content of A3(3)F starts to undergo a modification at stage IA. The glycan A3(3)F is found to be a promising candidate for a biomarker of ductal carcinoma that will be used for more detailed screening on a large group of patients with this diagnosis at different stages.

Upcoming studies will allow for a complete conclusion and outline prospects regarding this glycan. Thus, we have proven that ultra-high resolution IMS can be used as an effective tool for screening glycosylation profiles at the isomer level. This technique allows to select potential biomarker candidates by exploiting the difference between the ATDs of glycans extracted from healthy and sick donors.

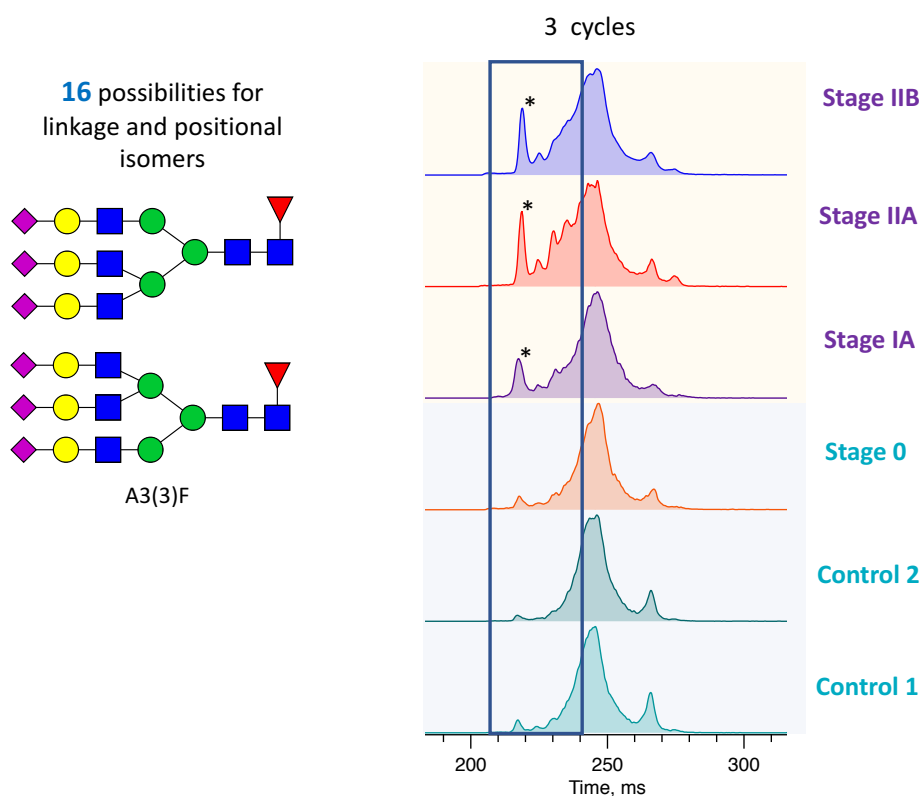


Figure 5.8. Comparison of ATDs of the glycan A3(3)F released from the Control 1, 2 samples and the samples of the sick individuals I–IV.



## 5.5. Scheme for isomer identification of sialylated N-glycans

Another essential part in the search for glycan biomarkers is the identification of the isomeric structures standing behind the differences between glycans extracted from healthy and sick donors. In the case of A3(3)F, it is impossible to purchase standards of all 16 isomers due to the complexity of their synthesis and the difficulty of their isolation; A3(3)F( $\alpha$ 2,3;  $\alpha$ 1,6) (Figure 5.9) is the only available isomer on the market. For this reason, we propose an alternative way to identify isomers using the combination of CID, IMS and IR. The scheme is based on the fragmentation of the parent molecule in order to generate diagnostic C fragments that keep sialic acid. The comparison of the isomer-selective IR fingerprints of the fragments with their standards contained in our database will allow for the identification of the parent molecule.

To be able to identify isomers of glycan A3(3)F we needed to create a small database of tri- and tetrasaccharide diagnostic fragments (Figure 5.9) that will serve to elucidate sialic acid linkage and branching, respectively. For this, we purchased and used the following pure glycan isomers: A3(3)F( $\alpha$ 2,3;  $\alpha$ 1,6), A3(3)( $\alpha$ 2,6;  $\alpha$ 1,3), A2( $\alpha$ 2,3) and A2( $\alpha$ 2,6). In addition, a CID experiment of tri-antennary core standards (A3(3)F( $\alpha$ 2,3;  $\alpha$ 1,6) and A3(3)( $\alpha$ 2,6;  $\alpha$ 1,3)) will provide an overall fragmentation pattern of the A3(3)F isomers extracted from the serum samples, assuming it will be similar between the isomers.

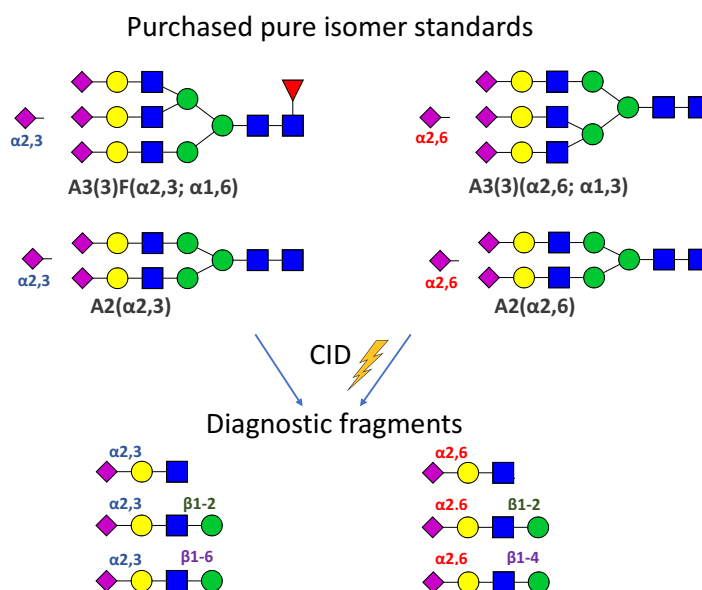


Figure 5.9. Scheme for the generation of diagnostic CID fragments for the identification of sialylated N-glycan isomers.

## Trisaccharide fragments

We performed on-board CID of the glycan A3(3)F( $\alpha$ 2,3;  $\alpha$ 1,6) followed by ion mobility separation of the C<sub>3</sub>( $\alpha$ 2,3) fragment with  $m/z = 673$ . Breaking any of the three GlcNAc–Man bonds in the molecule leads to the same C<sub>3</sub>( $\alpha$ 2,3) fragment. This fragment displayed one single peak in the ATD even after multiple cycling. We generated the same C<sub>3</sub>( $\alpha$ 2,3) fragment from the glycan A2( $\alpha$ 2,3), which also exhibited one peak on the ATD. The IR spectra of the C<sub>3</sub>( $\alpha$ 2,3) fragments from A3(3)F( $\alpha$ 2,3;  $\alpha$ 1,6) and from A2( $\alpha$ 2,3) matched well (Figure 5.10). These results prove that the IR fingerprints of fragments do not depend on the structure of the parent molecule they were generated from, demonstrating the flexibility of our approach.

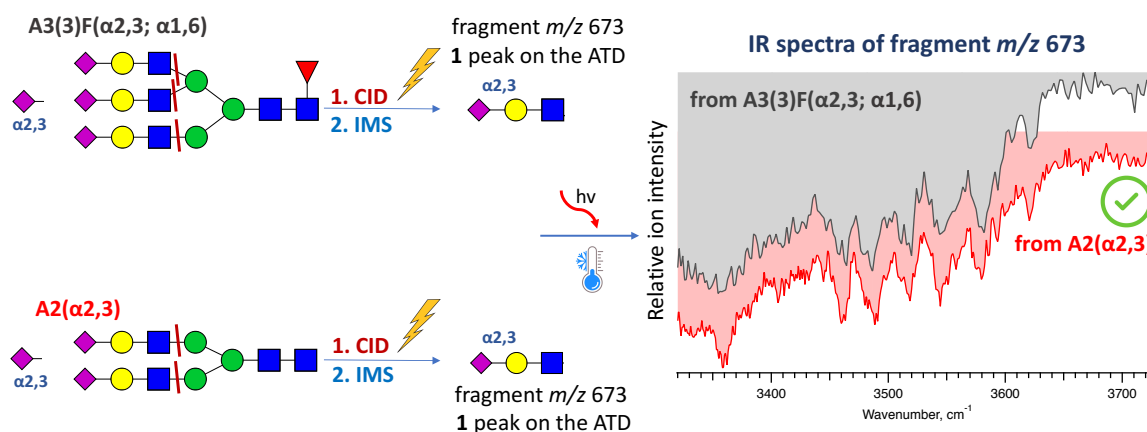


Figure 5.10. IR spectra of the C<sub>3</sub>( $\alpha$ 2,3) fragments obtained by CID of A3(3)F( $\alpha$ 2,3;  $\alpha$ 1,6) (grey) and of A2( $\alpha$ 2,3) (red).

To obtain the spectrum of the C<sub>3</sub>( $\alpha$ 2,6) fragment, we performed CID of A3(3)( $\alpha$ 2,6;  $\alpha$ 1,3). After 2 cycles of separation, the ATD of the C<sub>3</sub>( $\alpha$ 2,6) fragment unexpectedly displayed two peaks. However, under the same separation conditions, the C<sub>3</sub>( $\alpha$ 2,6) fragment from A2( $\alpha$ 2,6) showed only one peak. The IR spectra of peak I of the C<sub>3</sub>( $\alpha$ 2,6) fragment from A3(3)( $\alpha$ 2,6;  $\alpha$ 1,3) and the single peak of C<sub>3</sub>( $\alpha$ 2,6) from A2( $\alpha$ 2,6) displayed a good match (Figure 5.11). The origin of peak II was initially unclear.

The comparison between the IR fingerprints of peak II of the C<sub>3</sub>( $\alpha$ 2,6) fragment generated from A3(3)( $\alpha$ 2,6;  $\alpha$ 1,3), from A3(3)F( $\alpha$ 2,3;  $\alpha$ 1,6) and from A2( $\alpha$ 2,3) displayed a good match (Figure 5.12). These results revealed that the purchased A3(3)( $\alpha$ 2,6;  $\alpha$ 1,3) was not a pure  $\alpha$ 2,6 standard but contained both  $\alpha$ 2,6 and  $\alpha$ 2,3 linkages, which was spectroscopically confirmed. The ratio between peaks approximately I and II was 3:2.

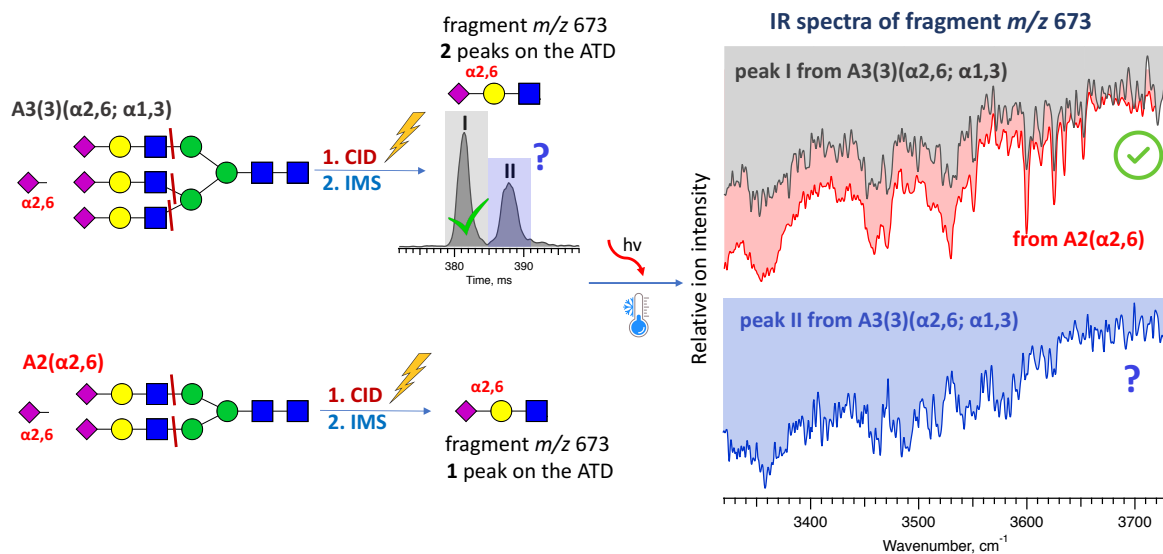


Figure 5.11. IR spectra of the  $C_3(\alpha2,6)$  fragments obtained by CID of  $A3(3)(\alpha2,6; \alpha1,3)$  (peak I – grey; peak II – blue) and  $A2(\alpha2,6)$  (red).

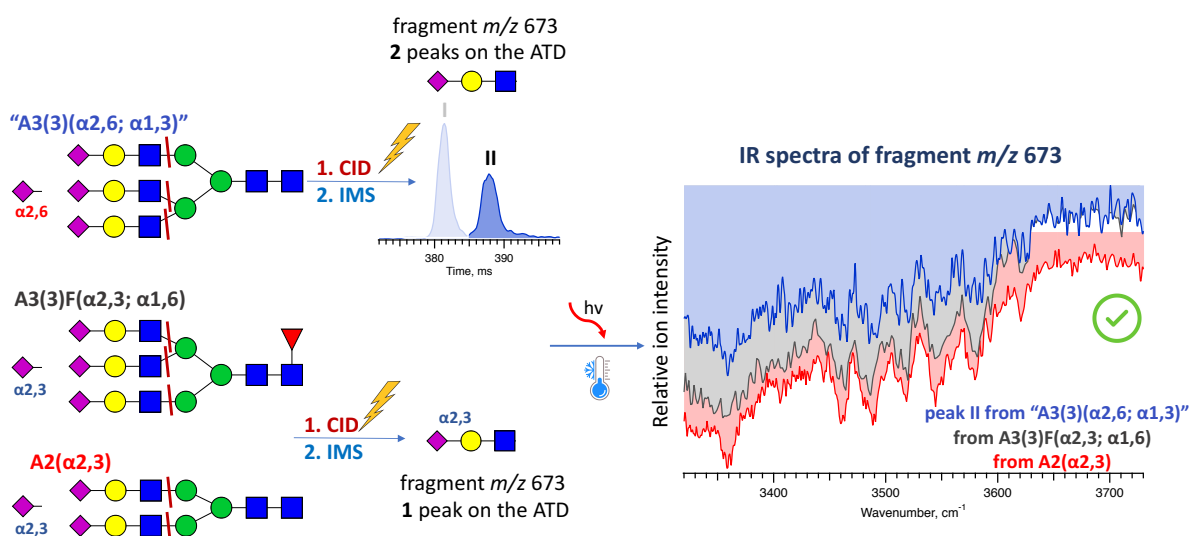


Figure 5.12. IR spectra of the  $C_3(\alpha2,3)$  fragments obtained by CID of the  $A3(3)(\alpha2,6; \alpha1,3)$ , (peak II in the ATD – blue), the  $A3(3)F(\alpha2,3; \alpha1,6)$  (grey) and  $A2(\alpha2,3)$  (red).

We experienced a similar issue with samples of pure positional and linkage isomers of glycans from different suppliers. The standard methods they use cannot detect such subtle structural differences. Adding the IMS-IR dimension would definitely improve the quality of the final glycan products.

## Tetrasaccharide fragments

The fragments generated from A3(3)F( $\alpha$ 2,3;  $\alpha$ 1,6) with  $m/z = 835$  correspond to two tetrasaccharide structures, differing in the linkage between the N-acetylglucosamine and the mannose monosaccharides (Figure 5.13). Breaking the Man-Man bond on the  $\alpha$ 1,3 branch yields a tetrasaccharide fragment with a  $\beta$ 1-2 GlcNAc–Man linkage (**C<sub>4</sub>( $\alpha$ 2,3;  $\beta$ 1-2)**); breaking the Man-Man bond on the  $\alpha$ 1,6 branch and the GlcNAc–Man bond with  $\beta$ 1-6 or  $\beta$ 1-2 linkages yields a tetrasaccharide fragment with a  $\beta$ 1-2 GlcNAc–Man linkage (**C<sub>4</sub>( $\alpha$ 2,3;  $\beta$ 1-2)**) or a  $\beta$ 1-6 GlcNAc–Man linkage (**C<sub>4</sub>( $\alpha$ 2,3;  $\beta$ 1-6)**), respectively. The fragments of A2( $\alpha$ 2,3) with  $m/z = 835$  correspond to one tetrasaccharide fragment with a  $\beta$ 1-2 GlcNAc–Man linkage (**C<sub>4</sub>( $\alpha$ 2,3;  $\beta$ 1-2)**).

The **C<sub>4</sub>( $\alpha$ 2,3;  $\beta$ 1-2)** standard generated from A2( $\alpha$ 2,3) allows to identify the **C<sub>4</sub>( $\alpha$ 2,3;  $\beta$ 1-6)** fragment from A3(3)F( $\alpha$ 2,3;  $\alpha$ 1,6). After 2 cycles separation of the fragments generated from A3(3)F( $\alpha$ 2,3;  $\alpha$ 1,6) and A2( $\alpha$ 2,3), we observed two peaks in both cases. The IR fingerprints of the separated peaks are depicted in Figure 5.13. Peaks I and II from A3(3)F( $\alpha$ 2,3;  $\alpha$ 1,6) represent a good match with peaks I and II from A2( $\alpha$ 2,3), respectively. These results demonstrate that under the applied CID conditions, we only perform a single bond fragmentation of A3(3)F( $\alpha$ 2,3;  $\alpha$ 1,6), generating the fragment **C<sub>4</sub>( $\alpha$ 2,3;  $\beta$ 1-2)**.

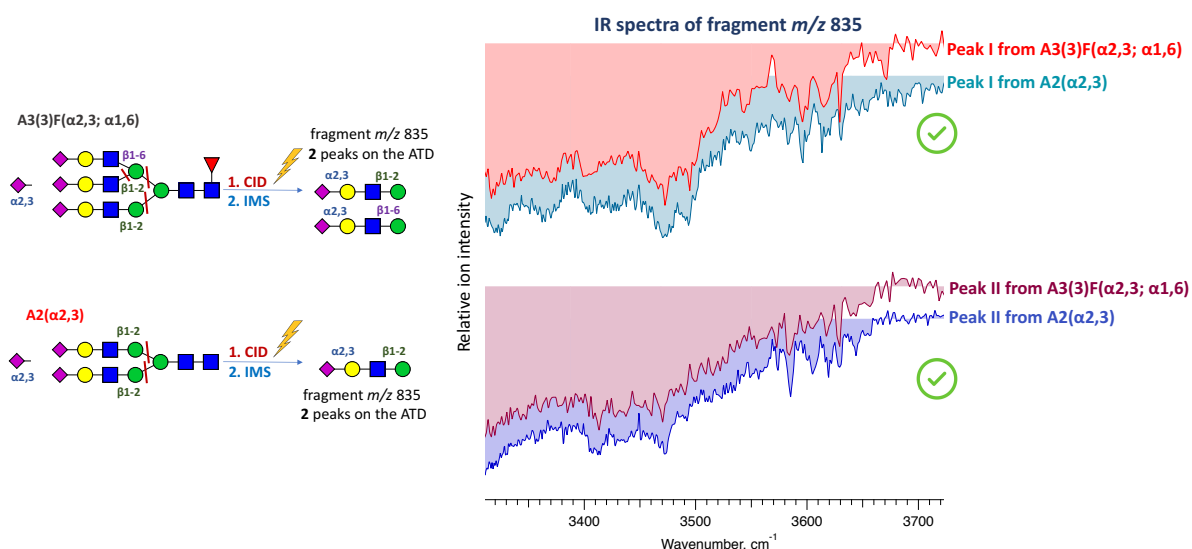


Figure 5.13. IR spectra of peaks I and II of the C<sub>4</sub> fragments obtained by CID of A3(3)F( $\alpha$ 2,3;  $\alpha$ 1,6) and A2( $\alpha$ 2,3).

A similar result was obtained for the fragments with  $m/z = 835$  generated from A3(3)( $\alpha$ 2,6;  $\alpha$ 1,3) and A2( $\alpha$ 2,6). After 2 cycles separation of the fragment, we observed two peaks in both cases. The IR fingerprints of the separated peaks are depicted in Figure 5.14. Peaks I and II

from  $A3(3)(\alpha2,6; \alpha1,3)$  represent a good match with peaks I and II from  $A2(\alpha2,3)$ , respectively. These results demonstrate that under the applied CID conditions, we only perform a single bond fragmentation of  $A3(3)(\alpha2,6; \alpha1,3)$ , generating the fragment  $C_4(\alpha2,6; \beta1-2)$ .

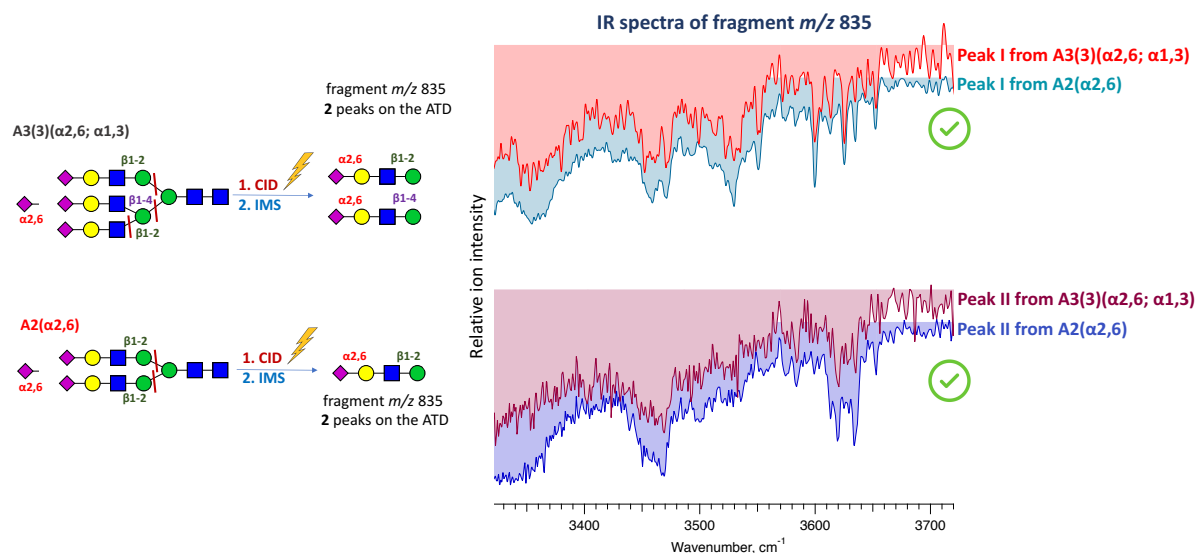


Figure 5.14. IR spectra of peaks I and II of the  $C_4$  fragments obtained by CID of  $A3(3)F(\alpha2,6; \alpha1,3)$  and  $A2(\alpha2,6)$ .

The spectra of the trisaccharide fragments –  $C_3(\alpha2,3)$ ,  $C_3(\alpha2,6)$  and of the tetrasaccharide fragments –  $C_4(\alpha2,3; \beta1-2)$ ,  $C_4(\alpha2,6; \beta1-2)$  were added to our database. Future work on optimization of the CID conditions that yield the  $C_4(\alpha2,3; \beta1-6)$  and  $C_4(\alpha2,6; \beta1-4)$  diagnostic fragments will continue to complete the database. This will allow for full identification of the isomer content of the glycan  $A3(3)F$  extracted from the serum samples.

## Conclusions

In this work, we demonstrated that ultra-high resolution IMS represents an effective tool for selecting potential biomarkers of diseases associated with changes in the glycocontent at the isomer level. The application of this method to a set of serum samples from patients with breast cancer revealed a significant difference between the ATDs of the triply-sialylated glycan A3(3)F from healthy and sick individuals. A future study of this potential biomarker on a large group of patients will allow for more complete conclusions. It is important to note that the same procedure can be applied to any glycan released from different biological fluids, including urine, sputum, blood serum and plasma.

Based on the high reproducibility of IR fingerprints, the assignment of glycan biomarkers can be done automatically using a spectroscopic database. Furthermore, we showed that adding a CID dimension to IMS-IR provides a way for isomer identification without the need for glycan standards that are not yet contained in our database and are impossible to obtain commercially. Thus, our four-dimensional IMS-MS-IR-CID approach is proven to be an important addition to existing techniques for discovering new glycan biomarkers.

# References

1. Siegel, R. L.; Miller, K. D.; Fuchs, H. E.; Jemal, A., Cancer statistics, 2022. *CA Cancer J Clin* **2022**, *72* (1), 7-33.
2. Glavey, S. V.; Huynh, D.; Reagan, M. R.; Manier, S.; Moschetta, M.; Kawano, Y.; Roccaro, A. M.; Ghobrial, I. M.; Joshi, L.; O'Dwyer, M. E., The cancer glycome: Carbohydrates as mediators of metastasis. *Blood Rev* **2015**, *29* (4), 269-279.
3. Dube, D. H.; Bertozzi, C. R., Glycans in cancer and inflammation--potential for therapeutics and diagnostics. *Nat Rev Drug Discov* **2005**, *4* (6), 477-88.
4. Adamczyk, B.; Tharmalingam, T.; Rudd, P. M., Glycans as cancer biomarkers. *Bba-Gen Subjects* **2012**, *1820* (9), 1347-1353.
5. Reis, C. A.; Osorio, H.; Silva, L.; Gomes, C.; David, L., Alterations in glycosylation as biomarkers for cancer detection. *J Clin Pathol* **2010**, *63* (4), 322-9.
6. Ruhaak, L. R.; Miyamoto, S.; Lebrilla, C. B., Developments in the identification of glycan biomarkers for the detection of cancer. *Mol Cell Proteomics* **2013**, *12* (4), 846-55.
7. de Leoz, M. L.; Young, L. J.; An, H. J.; Kronewitter, S. R.; Kim, J.; Miyamoto, S.; Borowsky, A. D.; Chew, H. K.; Lebrilla, C. B., High-mannose glycans are elevated during breast cancer progression. *Mol Cell Proteomics* **2011**, *10* (1), M110 002717.
8. Kyselova, Z.; Mechref, Y.; Kang, P.; Goetz, J. A.; Dobrolecki, L. E.; Sledge, G. W.; Schnaper, L.; Hickey, R. J.; Malkas, L. H.; Novotny, M. V., Breast cancer diagnosis and prognosis through quantitative measurements of serum glycan profiles. *Clin Chem* **2008**, *54* (7), 1166-75.
9. Saldova, R.; Reuben, J. M.; Abd Hamid, U. M.; Rudd, P. M.; Cristofanilli, M., Levels of specific serum N-glycans identify breast cancer patients with higher circulating tumor cell counts. *Ann Oncol* **2011**, *22* (5), 1113-1119.
10. Scott, E.; Munkley, J., Glycans as Biomarkers in Prostate Cancer. *Int J Mol Sci* **2019**, *20* (6).
11. Saldova, R.; Royle, L.; Radcliffe, C. M.; Abd Hamid, U. M.; Evans, R.; Arnold, J. N.; Banks, R. E.; Hutson, R.; Harvey, D. J.; Antrobus, R.; Petrescu, S. M.; Dwek, R. A.; Rudd, P. M., Ovarian cancer is associated with changes in glycosylation in both acute-phase proteins and IgG. *Glycobiology* **2007**, *17* (12), 1344-56.
12. Sethi, M. K.; Hancock, W. S.; Fanayan, S., Identifying N-Glycan Biomarkers in Colorectal Cancer by Mass Spectrometry. *Accounts Chem Res* **2016**, *49* (10), 2099-2106.

13. Liu, X. E.; Desmyter, L.; Gao, C. F.; Laroy, W.; Dewaele, S.; Vanhooren, V.; Wang, L.; Zhuang, H.; Callewaert, N.; Libert, C.; Contreras, R.; Chen, C., N-glycomic changes in hepatocellular carcinoma patients with liver cirrhosis induced by hepatitis B virus. *Hepatology* **2007**, *46* (5), 1426-35.
14. Arnold, J. N.; Saldova, R.; Galligan, M. C.; Murphy, T. B.; Mimura-Kimura, Y.; Telford, J. E.; Godwin, A. K.; Rudd, P. M., Novel Glycan Biomarkers for the Detection of Lung Cancer. *J Proteome Res* **2011**, *10* (4), 1755-1764.
15. Varki, A., Sialic acids in human health and disease. *Trends Mol Med* **2008**, *14* (8), 351-60.
16. Zhang, Z.; Wuhler, M.; Holst, S., Serum sialylation changes in cancer. *Glycoconj J* **2018**, *35* (2), 139-160.
17. Sethi, M. K.; Thaysen-Andersen, M.; Smith, J. T.; Baker, M. S.; Packer, N. H.; Hancock, W. S.; Fanayan, S., Comparative N-glycan profiling of colorectal cancer cell lines reveals unique bisecting GlcNAc and alpha-2,3-linked sialic acid determinants are associated with membrane proteins of the more metastatic/aggressive cell lines. *J Proteome Res* **2014**, *13* (1), 277-88.
18. Lomax-Browne, H. J.; Robertson, C.; Antonopoulos, A.; Leathem, A. J. C.; Haslam, S. M.; Dell, A.; Dwek, M. V., Serum IgA1 shows increased levels of alpha2,6-linked sialic acid in breast cancer. *Interface Focus* **2019**, *9* (2), 20180079.
19. Schultz, M. J.; Swindall, A. F.; Bellis, S. L., Regulation of the metastatic cell phenotype by sialylated glycans. *Cancer Metastasis Rev* **2012**, *31* (3-4), 501-18.
20. Bansal, P.; Ben Faleh, A.; Warnke, S.; Rizzo, T. R., Identification of N-glycan positional isomers by combining IMS and vibrational fingerprinting of structurally determinant CID fragments. *Analyst* **2022**, *147* (4), 704-711.
21. Nishikaze, T.; Tsumoto, H.; Sekiya, S.; Iwamoto, S.; Miura, Y.; Tanaka, K., Differentiation of Sialyl Linkage Isomers by One-Pot Sialic Acid Derivatization for Mass Spectrometry-Based Glycan Profiling. *Anal Chem* **2017**, *89* (4), 2353-2360.
22. Li, H.; Gao, W.; Feng, X.; Liu, B. F.; Liu, X., MALDI-MS analysis of sialylated N-glycan linkage isomers using solid-phase two step derivatization method. *Anal Chim Acta* **2016**, *924*, 77-85.





# Chapter 6. Relative quantification of sialylated N-glycans

## 6.1. Introduction

Follicle-stimulating hormone (FSH, follitropin) belongs to the gonadotropin family of hormones which are involved in the regulation of essential reproductive processes. FSH is a heterodimer that consists of a common  $\alpha$ -subunit noncovalently linked to a hormone-specific  $\beta$ -subunit for all members of this family. In humans, the subunits each contain two glycosylation sites. Glycans are one of the main components of FSH, representing ~30 % of the total mass of the hormone. They are involved in a number of processes, including the assembly of subunits, intracellular sorting and metabolic clearance. FSH glycosylation exhibits both macro and micro-heterogeneity. The macro-heterogeneity relates to the absence or presence of glycosylation of the  $\beta$ -subunit (the  $\alpha$ -subunit being always glycosylated on both sites). The micro-heterogeneity comes from the wide variety of oligosaccharide structures, including isoforms.<sup>1,2</sup>

Follitropin is widely used in the clinical treatment of both men and women. Between 88–99% of FSH oligosaccharides are sialylated, and several studies have shown that the sialylation has a significant impact on the pharmacokinetics and biological effects of the medicine.<sup>3,4</sup> The attached glycans can be singly, doubly, triply and quadruply sialylated (Figure 6.1). Different cell lines and production processes can lead to a difference in the glycosylation profiles, including the sialic acid pattern. For example, a first generation of FSH (FSH alpha) is more acidic and contains more tetra-antennary sialylated structures than the second generation (FSH beta).<sup>5</sup> Follitropin beta induces a greater estradiol response in patients, whereas follitropin alpha is associated with a higher pregnancy rate.<sup>6</sup> Thus, it is critical for the pharmaceutical industry to monitor the degree of sialylation of the recombinant FSH.

The content of sialylated glycans is generally detected by capillary electrophoresis and anion exchange chromatography. The separation is based on the different charge states of glycans and depends on the number of sialic acids in the molecule. While both techniques have been proven to be powerful tools for the study of FSH sialylation, they are time-consuming and require a derivatization step. In this work, we demonstrate our approach for the relative

quantification of sialylated glycans based on ion mobility spectrometry that allows to perform an analysis of glycans in their native form and on the millisecond time scale.

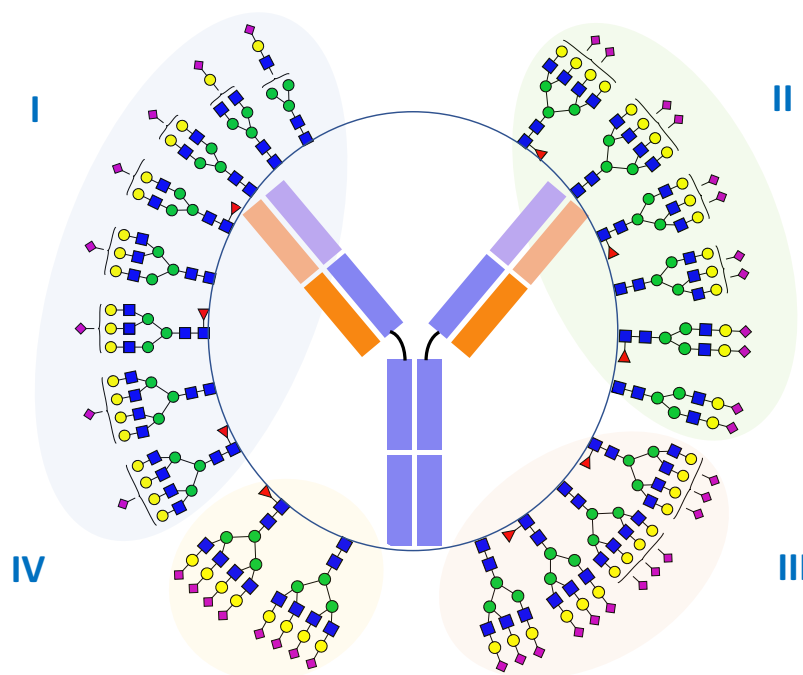


Figure 6.1. Four groups of sialylated N-glycans studied in this work.

## 6.2. Releasing sialylated N-glycans from r-hFSH

To release the sialylated N-glycans from r-hFSH, 250  $\mu\text{g}$  of the lyophilized hormone was first reconstituted in 100  $\mu\text{l}$  of water. The sample was then denatured by incubation with 10  $\mu\text{l}$  of ammonium carbonate (2M) and 10  $\mu\text{l}$  of DTT (1 M) for 1h at 60  $^{\circ}\text{C}$ . Subsequently, the sample was cooled down to ambient temperature and incubated with 10  $\mu\text{l}$  of iodoacetamide (1 M) in a dark room for 1h. Finally, the reaction mixture was digested with 5  $\mu\text{l}$  of PNGaseF enzyme (BioConcept, Allschwil, Switzerland) at 50  $^{\circ}\text{C}$  for 2 h and left overnight at 37  $^{\circ}\text{C}$ . After the reaction, the glycan mixture was cleaned from the protein and peptides with C18 cartridges and desalted using a dialysis device Float-A-Lyzer G2 (100-500 Da, SpectraPor, USA). The mixture of released N-linked glycans was lyophilized to dry powder and reconstituted in a 50:50 water/methanol solution. To confirm the complete release of all four glycan groups, the mixture was analyzed by UPLC-MS. Figure 6.2 displays a 60 min LC run on the example of some of the most abundant singly, doubly, triply and quadruply sialylated glycans released from the r-hFSH samples.

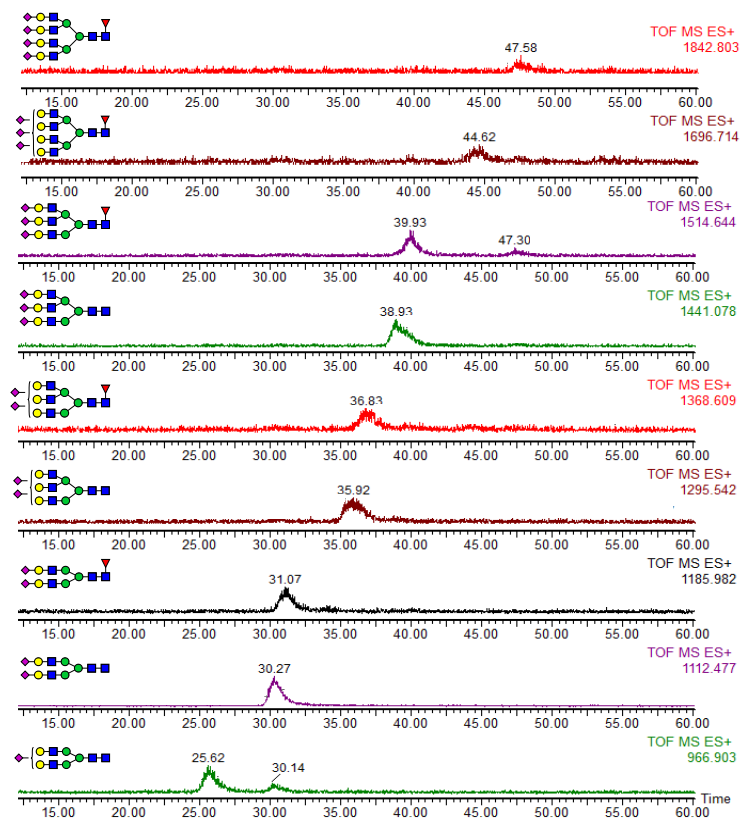


Figure 6.2. LC of the most abundant singly, doubly, triply and quadruply sialylated glycans released from the r-hFSH samples.

### 6.3. Ion mobility for relative quantification of glycans

For the IMS measurements, the lyophilized glycans were reconstituted in a 50:50 water/methanol solution up to a final concentration of 5  $\mu$ M. In this work, we aimed to establish the ratio between four families of sialylated glycans. For this reason, the ATDs were recorded without isomer separation. A crucial factor in observing such large glycan structures was a long time (5 s) in the accumulation region. We demonstrate our approach on the example of doubly sialylated glycans. Figure 6.3A illustrates their ATD profiles. In this mode, doubly sialylated glycans arrived within a 65–75 ms range, overlapping with each other. For the sake of clarity, the ATDs of A2F, A3(2), A3(2)F, A4(2) and A4(2)F were shifted along the x-axis in Figure 6.3B. We then integrated the area of each mobility peak of the group and summed them up. This procedure was performed for singly, triply and quadruply sialylated groups as well.

The intensities of the mobility peaks depend on many parameters, including the TW speed and height, the accumulation time, the drift gas pressure and the nESI conditions. In order to evaluate the reproducibility of our measurements for the relative quantification of sialylated

glycans, we chose one set of parameters and applied it to all glycans in the studied mixture. Our method favored singly and doubly charged states. Singly sialylated glycans were observed in a singly deprotonated charge state, whereas doubly, triply and quadruply sialylated groups were observed in a doubly deprotonated charge state.

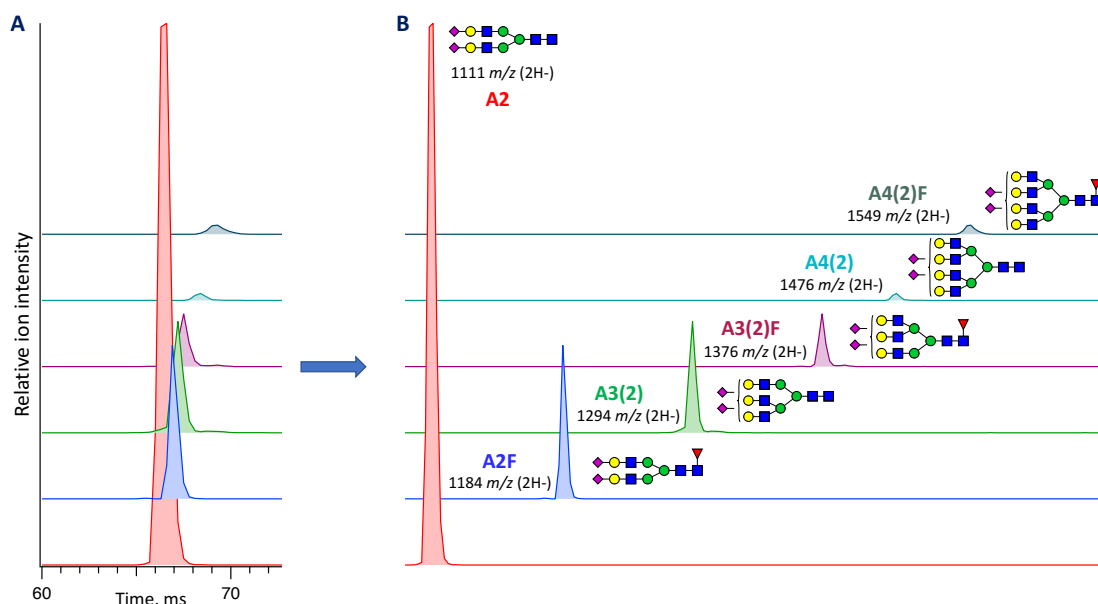


Figure 6.3. A) ATDs of doubly sialylated glycans released from the r-hFSH; B) Peak areas in 6.3.A shifted along the X-axis.

### 6.3.1 Relative quantification with nESI needles

To evaluate the reproducibility of our measurements, we performed experiments using nESI needles (Figure 6.4). We pulled the needles from borosilicate capillaries (Warner Instruments; OD = 1.2 mm, ID = 0.69 mm) on a micropipette puller (Sutter Instrument Co., model P-2000). The same glycan mixture was sprayed using three different needles produced from capillaries randomly taken from the same batch.

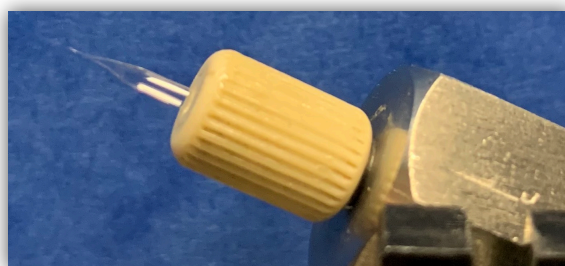


Figure 6.4. nESI needles used in this work.

Figure 6.5 illustrates the calculated content (%) of singly, doubly, triply and quadruply sialylated glycans after three independent measurements (experiments I, II, III) with the nESI needles (see **Appendix I** for details). The calculations were performed using the following expression:

$$\% = \frac{\sum \text{integrated areas of ATD peaks of glycans from the same group}}{\text{total ion signal from four groups of glycans}} * 100$$

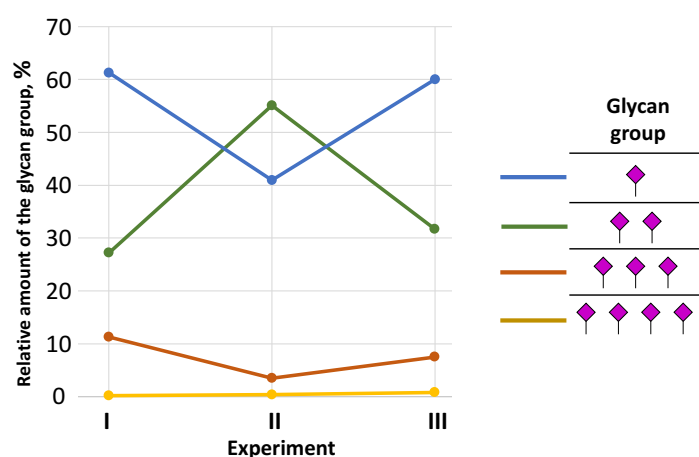


Figure 6.5. Relative amount of singly (blue), doubly (green), triply (orange) and quadruply (yellow) sialylated glycans obtained in three parallel experiments (I, II, III) for the same sample using different nESI needles.

These results show a significant inconsistency between our measurements. Relying on the reproducibility of the capillary manufacturing process and the puller program, the reason for this difference is likely to come from the needle-cutting step, even if done carefully. It is not an automated process but done by hand. The way needles are cut affects the inner diameter and as a result, the diameter of the droplets. The latter has a strong influence on the electrospray ionization process, leading to a different distribution of the sialylated glycan groups. In order to perform reproducible measurements, it was critical to eliminate the inconsistency introduced by human error.

### 6.3.2 Relative quantification with ESI source

In order to improve the reproducibility of the measurements, we installed a new ESI source with a syringe pump on the IMS-CIT II instrument (Figure 6.6).



Figure 6.6. ESI source with a syringe pump used in this work.

The measurements were performed at 7  $\mu\text{L}/\text{min}$  injection flow rate, and the source was kept at 70  $^{\circ}\text{C}$ . The syringe (100  $\mu\text{l}$ ) was filled with the same mixture of glycans that was used in the previous part 6.4. Prior to each experiment, the syringe and the connecting line were flushed with a mixture of solvents (50/50 MeOH/H<sub>2</sub>O) until the total signal in the ATD dropped to the baseline. Figure 6.7A illustrates the calculated content of sialylated glycans after three independent measurements with a syringe pump in one day, and Figure 6.7B shows measurements for three different days.

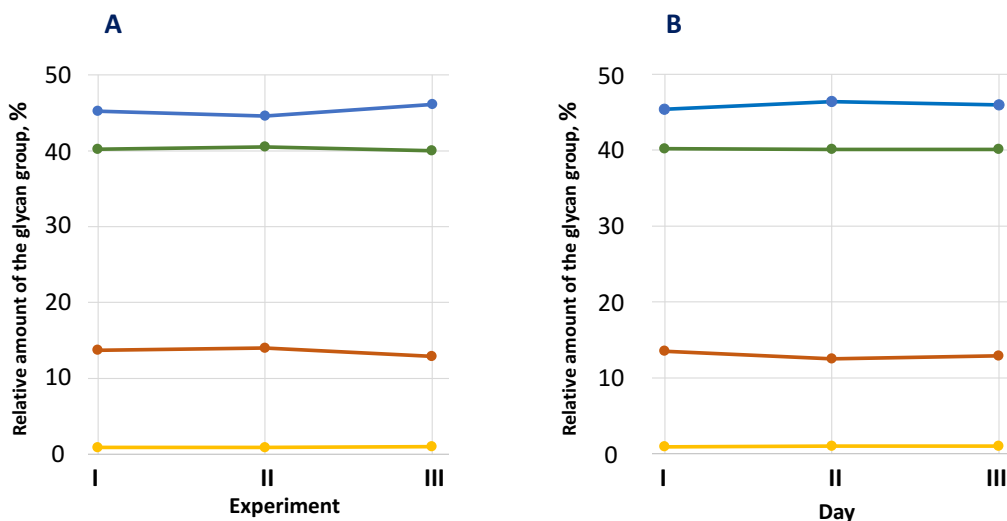


Figure 6.7. A) Relative amount of singly (blue), doubly (green), triply (orange) and quadruply (yellow) sialylated glycans obtained in three parallel experiments (I, II, III) for the same sample using a nESI source with a pumped syringe in one day; B) in different days.

Figure 6.7 demonstrates the degree of reproducibility between measurements performed with a syringe pump on one day and between different days. The values for the different days

were taken as the average of three measurements on each day. Table 6.1. shows the content (%) of four glycan groups in the mixture with confidence intervals ( $\Delta$ ) given by:

$$\Delta = \frac{\sigma * 1.960}{\sqrt{N}}, \sigma = \sqrt{\frac{\sum(x_i - \mu)^2}{N}}$$

where  $\sigma$  – standard deviation,  $N$  – number of samples,  $x_i$  – each value of the group,  $\mu$  – average value of the group, 1.960 – coefficient for 95% confidence level.

Table 6.1. Calculated content (%) of sialylated glycans for Figure 6.7B

Glycan group	Number of sialic acids	%
A <sub>1</sub>	1	45.9 ± 0.5
A <sub>2</sub>	2	40.1 ± 0.1
A <sub>3</sub>	3	13.0 ± 0.5
A <sub>4</sub>	4	1.0 ± 0.1

We subsequently performed experiments with different samples. Each sample was taken from the same batch and independently treated using the same glycan releasing procedure described in part 6.2. The ATDs of released glycans from five samples were measured one after another. We repeated this cycle two more times to obtain three independent measurements.

Figure 6.8 shows the reproducibility between five different samples (I–V) from the same batch.

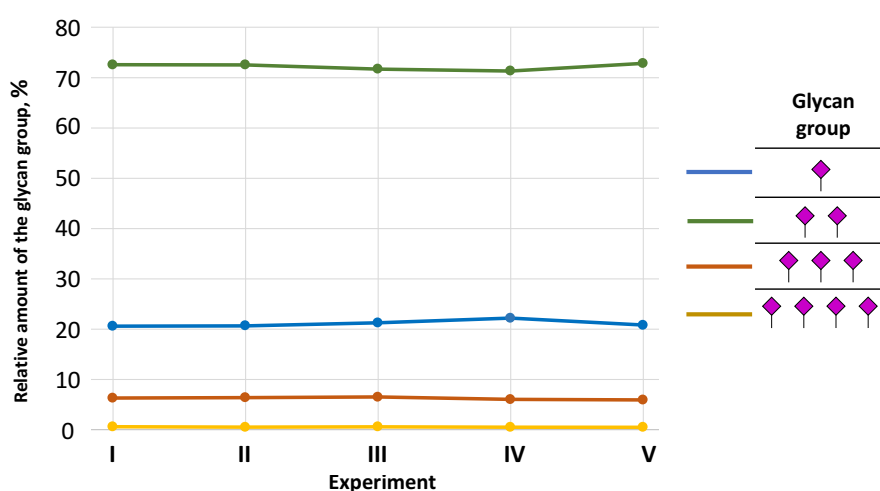


Figure 6.8. Relative amounts of singly (blue), doubly (green), triply (orange) and quadruply (yellow) sialylated glycans obtained in 5 samples from the same batch using a nESI source with a pumped syringe.



Table 6.2. Calculated content (%) of sialylated glycans for Figure 6.8

Glycan group	Number of sialic acids	%
A <sub>1</sub>	1	21.1 ± 0.5
A <sub>2</sub>	2	72.2 ± 0.5
A <sub>3</sub>	3	6.2 ± 0.2
A <sub>4</sub>	4	0.5 ± 0.1

These results show a good reproducibility both for measurements and for the sample preparation procedure. The next step in this research would be to monitor the difference in glycosylation profiles of FSH from different batches. The corresponding samples will be provided by the same manufacturer, we have been collaborating with in this project.

## Conclusions

In this work, ion mobility has been demonstrated to be a fast and reliable method for the relative quantification of sialylated glycans. It is important to note that with this method, glycans can be analyzed in their native form and directly from a complex mixture. We demonstrated that the inner diameter of nESI needles has a strong influence on the electrospray process and cannot be used for the quantification of glycans since the needle-cutting step introduces inconsistency in measurements. In contrast, the capillary of an ESI source has a fixed inner diameter. We showed that the replacement of the nESI needles by an ESI source with a syringe pump allows obtaining reproducible results between measurements performed on different days with the same set of parameters. Moreover, good reproducibility between independently treated samples from the same batch confirmed the accuracy of our sample preparation procedure. Thus, ion mobility is a promising technique that can be used to control the quality of recombinant glycoproteins, including FSH, by monitoring their glycosylation pattern.

# References

1. Ulloa-Aguirre, A.; Timossi, C.; Barrios-de-Tomasi, J.; Maldonado, A.; Nayudu, P., Impact of carbohydrate heterogeneity in function of follicle-stimulating hormone: studies derived from in vitro and in vivo models. *Biol Reprod* **2003**, *69* (2), 379-89.
2. Bousfield, G. R.; Harvey, D. J., Follicle-Stimulating Hormone Glycobiology. *Endocrinology* **2019**, *160* (6), 1515-1535.
3. Dias, J. A.; Ulloa-Aguirre, A., New Human Follitropin Preparations: How Glycan Structural Differences May Affect Biochemical and Biological Function and Clinical Effect. *Front Endocrinol (Lausanne)* **2021**, *12*, 636038.
4. Bousfield, G. R.; May, J. V.; Davis, J. S.; Dias, J. A.; Kumar, T. R., In Vivo and In Vitro Impact of Carbohydrate Variation on Human Follicle-Stimulating Hormone Function. *Front Endocrinol (Lausanne)* **2018**, *9*, 216.
5. Olijve, W.; de Boer, W.; Mulders, J. W.; van Wezenbeek, P. M., Molecular biology and biochemistry of human recombinant follicle stimulating hormone (Puregon). *Mol Hum Reprod* **1996**, *2* (5), 371-82.
6. Orvieto, R.; Nahum, R.; Rabinson, J.; Ashkenazi, J.; Anteby, E. Y.; Meltzer, S., Follitropin-alpha (Gonal-F) versus follitropin-beta (Puregon) in controlled ovarian hyperstimulation for in vitro fertilization: is there any difference? *Fertil Steril* **2009**, *91* (4 Suppl), 1522-5.



## Chapter 7. Summary and future perspectives

Traditional analytical methods alone cannot unravel glycans' structural complexity and heterogeneity. However, over the last two decades, orthogonal approaches have opened new ways to solve this problem. Despite the significant progress, glycan analysis remains challenging, particularly for isomeric structures. In this work, we presented a novel, state-of-the-art technique that introduces an additional metric for glycan identification based on IR fingerprinting. The development and gradual improvement of our method were accomplished in three generations of homebuilt instruments.

The first-generation instrument incorporates tandem mass spectrometry and cryogenic messenger-tagging spectroscopy, which provides unique vibrational fingerprints of biomolecules. The extreme sensitivity of cryogenic IR to the slightest structural differences is a prerequisite for applying this method to the study of glycans. The preliminary data showed that even glycans consisting of nine monosaccharide units exhibit sufficiently distinct spectral fingerprints that can serve as unique identifiers. Based on these results, we proposed a new approach for glycan identification using a spectroscopic database, without high-level quantum-chemical calculations. While this method can deal with structures that are already present in the database, we needed to find a solution for the identification of unknown glycans. By combining selective enzymatic cleavage with cryogenic infrared spectroscopy, we have developed a mechanism to identify unknown glycans and expand our database. The IR spectra of enzymatically cleaved glycans demonstrated no difference regarding the cleavage procedure and were in very good agreement with those of reference standards. These results served as a starting point for constructing a spectroscopic glycan database.

The second-generation instrument provides an additional dimension for analysis by implementing an ultra-high resolution ion mobility module based on SLIM technology. This allows to separate isomeric glycan structures and obtain isomer-specific IR fingerprints. We showed that even structures as close as positional isomers could be easily distinguished. Moreover, the robustness and speed of the experiment were significantly improved due to recent progress in solid-state IR laser sources. This has made IR spectroscopy competitive with other techniques for routine analysis.

The third-generation instrument is an upgraded version of the previous one for more accurate, sensitive and rapid analysis of glycans. It provides a higher IMS resolving power and one more dimension by implementing on-board CID fragmentation. We demonstrated our approach on the example of two positional isomers of the glycan G1F. The observed isomers differ only by 0.2% in their CCS, which makes it challenging to distinguish them by IMS-MS alone. Moreover, we implemented a chemoenzymatic approach for the selective synthesis of one of the isomers, since it is not possible to obtain pure isomers of G1F commercially. We then employed the synthesized standard for assigning the mobility-separated peaks based on their unique IR fingerprints. These results were used to monitor the difference in G1F isomers extracted from IgGs that have been produced in two different cell lines. We thereby demonstrated the potential of our approach to control the glycosylation in mAbs.

Our IMS-MS-IR approach was then used for the identification and relative quantification of sialylated glycans released from different biological samples. We have shown it to be an effective tool for discovering new glycan biomarkers for different diseases and controlling the quality of recombinant glycoproteins at the glycan isomer level. It is important to note that this method allows for the analysis of glycans in their native form and directly from a complex mixture, which eliminates additional steps in the sample preparation procedure. Furthermore, we developed a method for the identification of sialyl-linkage and positional isomers using C-fragments in the CID experiment. This eliminates the necessity of glycan standards that are difficult or impossible to obtain. Table 7.1 summarizes the main advantages and limitations of the techniques (MS, IMS, IR) combined in our approach.

Table 7.1. Summary of the main advantages and limitations of the techniques (MS, IMS, IR) combined in our approach

Advantages	Limitations
<b>Mass spectrometry</b>	
<ul style="list-style-type: none"> <li>- High sensitivity</li> <li>- Wide dynamic range</li> <li>- Does not require a derivatization step for detection</li> </ul>	<ul style="list-style-type: none"> <li>- Ion suppression in the electrospray when working with complex mixture</li> <li>- Requires a pre-separation step (LC, CE)</li> <li>- Quantification requires an internal standard from a similar class to the studied analytes</li> </ul>

<b>Advantages</b>	<b>Limitations</b>
<b>SLIM ion mobility</b>	
<ul style="list-style-type: none"> <li>- Ultra-high resolution separation</li> <li>- Fast separation on the timescale of milliseconds</li> <li>- Easy to fabricate and has a small size</li> </ul>	<ul style="list-style-type: none"> <li>- Limited peak capacity</li> </ul>
<b>Cryogenic messenger-tagging spectroscopy</b>	
<ul style="list-style-type: none"> <li>- High scanning speed</li> <li>- Orthogonality to IMS-MS</li> <li>- Provides unique vibrational fingerprints that are sensitive to the slightest structural differences</li> <li>- Highly reproducible IR spectra</li> </ul>	<ul style="list-style-type: none"> <li>- Tag distribution can affect the IR spectrum</li> <li>- Requires low temperatures</li> </ul>
<b>Additionally, for all of the above</b>	
<ul style="list-style-type: none"> <li>- Compatibility with fragmentation techniques that reduces the need of standards for identification</li> </ul>	

The continuation of this work will include the following steps:

1. Method improvement: expanding the spectroscopic database of glycans, implementing the multiplexing approach for high throughput analysis and applying machine learning for the automatic spectra assignment.
2. Collaboration with hospitals to access the donor libraries for the glycan biomarker search.
3. Collaboration with biopharmaceutical companies to monitor the glycosylation pattern of their products.
4. Expanding the classes of biomolecules to metabolites and glycolipids; applying our method to issues which traditional analytical methods for biomolecular identification cannot solve.



# Appendix I

Calculated content (%) of sialylated glycans from Figure 6.5

Glycan group	Number of sialic acids	Experiment I	Experiment II	Experiment III
A <sub>1</sub>	1	27.2	55.1	31.7
A <sub>2</sub>	2	61.3	41.0	60.0
A <sub>3</sub>	3	11.3	3.5	7.5
A <sub>4</sub>	4	0.2	0.4	0.8

Calculated content (%) of sialylated glycans from Figure 6.7A

Glycan group	Number of sialic acids	Experiment I	Experiment II	Experiment III
A <sub>1</sub>	1	45.2	44.6	46.1
A <sub>2</sub>	2	40.2	40.5	40.0
A <sub>3</sub>	3	13.7	14.0	12.9
A <sub>4</sub>	4	0.9	0.9	1.0

Calculated content (%) of sialylated glycans from Figure 6.7B

Glycan group	Number of sialic acids	Day I	Day II	Day III
A <sub>1</sub>	1	45.4	46.4	46.0
A <sub>2</sub>	2	40.2	40.1	40.1
A <sub>3</sub>	3	13.5	12.5	12.9
A <sub>4</sub>	4	0.9	1.0	1.0

Calculated content (%) of sialylated glycans from Figure 6.8.

Glycan group	Number of sialic acids	I	II	III	IV	V
A <sub>1</sub>	1	20.5	20.6	21.2	22.2	20.8
A <sub>2</sub>	2	72.6	72.5	71.7	71.3	72.8
A <sub>3</sub>	3	6.3	6.4	6.5	6.0	5.9
A <sub>4</sub>	4	0.6	0.5	0.6	0.5	0.5





# Acknowledgements

From the bottom of my heart, I would like to thank Prof. Thomas Rizzo, the person who made this work possible. Tom, thank you that four years ago you believed in me and gave me the opportunity to join the LCPM group. I really appreciate all the independency and resources you have provided me during this time. You helped me to develop my critical thinking and presenting skills. This allowed me to feel confident, even giving talks at big international conferences. Thank you for the guidance, support and being fair in every situation. Thanks for the great 12-hour trip to Austria and back and for sharing many stories on the way; for the excellent barbeque and hospitality that you and Karen offered us in your house.

I am grateful to the PhD exam committee: Prof. Jacques-E. Moser, Prof. Sandrine Gerber, Prof. Anouk Rijs and Prof. Brian Clowers for accepting my invitation and evaluating this thesis.

Another person I would like to thank is Angeles Alarcon who freed me from all administrative work and just let me do science. Thank you for always being positive, kind and patient even when I once again messed up my invoices in Catalysis.

Thanks to Chiara for showing me around and answering endless questions during my interview. A big thank goes to Val who introduced me to the world of lasers and spent quite some time showing the OPO alignment.

To be honest, during the first month of my PhD, just looking at the machines made my blood run cold. I am deeply grateful to Robert who became a good friend of mine for all his technical help and moral support. He was the first person to introduce me to the experimental work on Machine II. His exceptional sense of humor did not let me give up even when it was tough and it seemed to be hopeless. We spent so much time aligning the laser, fixing the machine and just laughing. Later we were very lucky to have Eduardo who completed our “dream team”. I really appreciated our discussions about life and science, which both always ended with energy barrier diagrams. Thank you for believing in me and offering your help.

I would like to thank Priyanka for being the best travel buddy. We had an unforgettable summer school in the CCRC and ASMS conference in Minneapolis. Thank you for instilling in me a love for Indian food and sharing my love for the TV show Friends. A big thank goes to Ali with whom we had very nice discussions, cheered up each other and went to many conferences together. I would also like to thank my colleagues Natalia and Lei for the fruitful discussions and thoughtful suggestions during the group meetings.

I am thankful to David Hacker from the Protein facility at EPFL for providing IgG samples, David Kidwell from MS vision for his help with the Q-TOF and the LCBM group (Kristina, Alex, Eduard, Maxime and Harsh) for letting me use their equipment and borrow chemicals.

I am extremely grateful to Stephan and Ahmed for the great work we have done together and all the help they provided with Machine IV. I learned so much from them that helped me to develop my problem-solving skills and critical thinking. Thanks, Ahmed, for the support during tough times and for the great atmosphere you and Sarah always have in your house. I

am also thankful to Vasyl for his constant help with Machine III, for always being very patient and willing to explain the same thing again and again.

A warm thought goes to my LCPM friend Bo-Jung. We did paragliding, hiking, cooking and many other things together. Thank you for the outstanding cakes and delicious hot chocolate, but most importantly, for your friendship and endless optimism that charges everyone around you.

I was extremely lucky to have very close friends of mine from Russia in the same laboratory. With Erik and Vika, we spent much time together and supported each other, especially during lunch breaks. In the first year, we travelled almost every weekend to different countries. It was crazy but we had so much fun!

No words to explain how grateful I am to have these three girls in my life: Nastia, Lera, and Kristina. We could write a book of thousands of pages telling all our stories, adventures and jokes. Most importantly, we lived through all of them together and were a great support for each other. This year we celebrated the 10-year anniversary of our friendship and I believe that this is just the very beginning because real friendship never ends!

I am also thankful to all new friends I made here at EPFL (Acia, Andrey, Johann, Eugene Shulga, Eugene Upenik, Ben, the DiableRabattable music band, ...), for many parties, trips and outdoor activities we had together. Thanks to all the LCPM members from Rainer's and Andreas's groups for hiking, skiing and having fondue together.


*Огромное спасибо моей любимой семье: маме Галине, папе Ивану и сестре Кате за их невероятную поддержку на протяжении всей моей жизни. Без их безграничной любви и заботы я бы не смогла пройти этот путь. Они моя опора и я знаю, что могу на них положиться в любой ситуации. Спасибо что вы есть у меня и всегда ждете дома!*

The most special thank goes to my dear Raphael, my best friend, my husband and the strictest reviewer of this thesis. Even when the whole world is shaking, you are my stability and protection whatever happens. Thank you for your love and support that make me stronger and confident to climb new heights but only with you by my side. Thanks to the Märki family for the warm welcome and for being the best reason to improve my French.

And, finally, you, dear reader, thank you for your interest in my work. Hopefully, in this thesis, you could find something useful and inspiring for your own research.

# Irina Dyukova Märki

 7 October 1994

 irinamarki94@gmail.com

 +41 76 395 10 94

 Rue de la Blancherie 28  
1022 Chavannes

 Irina Dyukova



## EDUCATION

- 2018 – present
- PhD-candidate in the Laboratory of Molecular Physical Chemistry
  - École Polytechnique Fédérale de Lausanne
  - Combining ion mobility spectrometry and cryogenic ion spectroscopy with enzymatic digestion/synthesis for the study of glycans
  - Reference: Prof. T. Rizzo ([thomas.rizzo@epfl.ch](mailto:thomas.rizzo@epfl.ch))
- August 2019
- Summer school. Complex Carbohydrate Research Center, University of Georgia, USA
- 2012 – 2017
- Master in chemistry, GPA 4.82 / 5.00
  - Novosibirsk State University (Novosibirsk, Russia)

## RESEARCH METHODS AND SKILLS

- Mass spectrometry
- Liquid/gas chromatography
- Ion mobility separation
- Infrared spectroscopy
- Enzymatic cleavage/synthesis
- Analytical biochemistry
- Cryogenic systems
- Vacuum systems
- Inorganic synthesis

## LANGUAGES

Russian – native, English – C1, French – B1.

## RESEARCH EXPERIENCE

- 2016 – 2018
- Engineer in the Laboratory of Complex Compounds.
  - Nicolaev Institute of Inorganic Chemistry (Novosibirsk, Russia)
  - Synthesis and characterization of complex compounds with polynitrogen ligands

- Reference: Prof. L. Lavrenova (ludm@niic.nsc.ru)
- 2016 – 2017
- Intern at the Well Services group
  - Novosibirsk Technological Center Schlumberger (Novosibirsk, Russia)
  - Improvement of the efficiency of hydraulic fracturing
  - Reference: Dr. A. Alexeev (aalekseev4@slb.com)
- 2015 – 2016
- Engineer in the Analytical Laboratory
  - Boreskov Institute of Catalysis (Novosibirsk, Russia)
  - Design of MEMS-based columns for micro gas chromatography

### **TEACHING EXPERIENCE**

- 2018 – 2022
- Teaching assistant for Experimental Physical Chemistry (1<sup>st</sup> year master students)
  - Teaching assistant for General Chemistry (1<sup>st</sup> year bachelor students)
  - École Polytechnique Fédérale de Lausanne
- 2017 – 2018
- Teacher for Analytical Chemistry (2<sup>nd</sup> year bachelor students)
  - Novosibirsk state university

### **TALKS AT CONFERENCES**

1. The 70th ASMS Conference on Mass Spectrometry 2022: “Isomer separation and relative quantification of sialylated N-linked glycans”. Glycomics section. Minneapolis, USA.
2. Swiss Chemical Society Fall Meeting 2021: “A new approach for identifying positional isomers of glycans cleaved from monoclonal antibodies”. Section of Analytical Sciences. University of Zurich.
3. Swiss Chemical Society Fall Meeting 2019: “Determination of the primary structure of glycans by enzymatic cleavage and cryogenic IR-spectroscopy”. Runner-up in the Section of Analytical Sciences. University of Zurich.

### **POSTERS AT CONFERENCES**

1. Isolated Biomolecules and Biomolecular Interactions 2022: “A new approach for identifying positional isomers of glycans cleaved from monoclonal antibodies”. University of Innsbruck.

2. The 69th ASMS Conference on Mass Spectrometry 2021: “A new approach for identifying positional isomers of glycans cleaved from monoclonal antibodies”. Carbohydrates section. Philadelphia, USA.
3. The 68th ASMS Conference on Mass Spectrometry 2020: “Combining cryogenic infrared spectroscopy with selective enzymatic cleavage for determining glycan primary structure”. Carbohydrates section. Online.

### LIST OF PUBLICATIONS

1. **I. Dyukova**, A. Ben Faleh, S. Warnke, N. Yalovenko, V. Yatsyna, P. Bansal, and T. R. Rizzo. A new approach for identifying positional isomers of glycans cleaved from monoclonal antibodies. *CHIMIA*, **2022**, 76, 363.
2. A. Ben Faleh, S. Warnke, P. Bansak, R.P. Pellegrinelli, **I. Dyukova**, and T.R. Rizzo. Identification of mobility-resolved N-glycan isomers. *Anal. Chem.* **2022**, 94, 10101–10108.
3. **I. Dyukova**, A. Ben Faleh, S. Warnke, N. Yalovenko, V. Yatsyna, P. Bansal, and T. R. Rizzo. A new approach for identifying positional isomers of glycans cleaved from monoclonal antibodies. *Analyst*, **2021**, 146, 4789-4795.
4. **I. Dyukova**, E. Carrascosa, R. P. Pellegrinelli, and T. R. Rizzo. Combining cryogenic infrared spectroscopy with selective enzymatic cleavage for determining glycan primary structure. *Anal. Chem.* **2020**, 92, 1658–1662.
5. L.G. Lavrenova, **I.I. Dyukova**, E.V. Korotaev, L.A. Sheludyakova, V.A. Varnek. Spin Crossover in New Iron(II) Complexes with 2,6-Bis(benzimidazole-2-yl)pyridine. *Russian Journal of Coordination Chemistry*, **2020**, 65, 30-35.
6. **I.I. Dyukova**, L.G. Lavrenova, T.A. Kuz'menkoV. Yu. Komarov, T.S. Sukhikh, E.V. Vorontsova. Coordination compounds of Co(II) and Cu(II) halides with 2,4-dimethyl-1,2,4-triazolo[1,5-a]benzimidazole. *Inorganica Chimica Acta*, **2019**, 486, 406-411.
7. **I.I. Dyukova**, T.A. Kuz'menkoV. Yu. Komarov, T.S. Sukhikh, E.V. Vorontsova, L.G. Lavrenova. Coordination Compounds of Cobalt(II), Nickel(II), and Copper(II) Halides with 2-Methyl-1,2,4-Triazolo[1,5-a]benzimidazole. *Russian Journal of Coordination Chemistry*, **2018**, 44, 755-764.

