

Fucosylated and Sulfated Glycans

Investigated using Cryogenic
Infrared Spectroscopy



Maike Lettow

Fucosylated and Sulfated Glycans

Investigated using Cryogenic Infrared Spectroscopy

Fucosylated and Sulfated Glycans Investigated using Cryogenic Infrared Spectroscopy
PhD Thesis, Radboud University, Nijmegen, the Netherlands

© 2022 Maïke Lettow

ISBN: 978-94-6421-927-2

Printed by IPSKAMP Printing, Enschede

The work described in this thesis has been performed at the Fritz-Haber-Institut der Max-Planck-Gesellschaft in Berlin, Germany.



Fucosylated and Sulfated Glycans

Investigated using Cryogenic Infrared Spectroscopy

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,

volgens besluit van het college voor promoties

in het openbaar te verdedigen op

maandag 19 december 2022

om 12:30 uur precies

door

Maike Lettow

geboren op 17 juli 1990

te Brunsbüttel, Duitsland

Promotoren:

Prof. dr. G.J. von Helden

Prof. dr. K. Pagel (Freie Universität Berlin, Duitsland)

Prof. dr. G.J.M. Meijer

Manuscriptcommissie:

Prof. dr. B. Redlich

Prof. dr. J. Roithová

Prof. dr. M. Wuhler (Universiteit Leiden)

Fucosylated and Sulfated Glycans

Investigated using Cryogenic Infrared Spectroscopy

Dissertation

to obtain the degree of doctor

from Radboud University Nijmegen

on the authority of the Rector Magnificus prof. dr. J.H.J.M. van Krieken,

according to the decision of the Doctorate Board

to be defended in public on

Monday, December 19, 2022

at 12:30 pm

by

Maike Lettow

born on July 17, 1990

in Brunsbüttel, Germany

PhD supervisors:

Prof. dr. G.J. von Helden

Prof. dr. K. Pagel (Freie Universität Berlin, Germany)

Prof. dr. G.J.M. Meijer

Manuscript Committee:

Prof. dr. B. Redlich

Prof. dr. J. Roithová

Prof. dr. M. Wuhrer (Leiden University)

Contents

1. Introduction	1
1.1. Glycans and the Diversity in their Chemical Structure	1
1.2. Standard Techniques for the Structural Analysis of Glycans.....	8
1.3. Ion Mobility-Mass Spectrometry	12
1.4. Mass Spectrometry-based Infrared Spectroscopy.....	14
1.4.1. General Principles.....	14
1.4.2. Experiment Designs	18
1.5. Aim and Outline.....	22
2. Experimental Overview.....	25
2.1. Infrared Multiple Photon Dissociation Spectroscopy.....	25
2.2. Messenger-Tagging Spectroscopy.....	28
2.2.1. The General Design	28
2.2.2. The Cryogenic Hexapole Ion Trap	31
2.2.3. Time-Sequence Generation.....	33
2.2.4. Performance Characterization	35
2.3. Infrared Spectroscopy in Helium Nanodroplets.....	37
2.4. The Fritz Haber Institute free-electron laser	40
3. The Role of the Mobile Proton in Fucose Migration.....	43
3.1. Introduction	44
3.2. Experimental Details	47
3.3. Results and Discussion.....	48
3.3.1. Ammonium adduct ions	48
3.3.2. Alkylammonium adduct ions	52
3.3.3. Fluorescently labeled glycan ions	53
3.4. Conclusion.....	57

4. Decoding the Fucose Migration Product in Blood Group Epitopes	59
4.1. Introduction	60
4.2. Experimental Details.....	61
4.3. Results and Discussion	63
4.4. Conclusion	76
5. Infrared Action Spectroscopy of Glycosaminoglycans.....	77
5.1. Introduction	78
5.2. Experimental Details.....	80
5.3. Results and Discussion	81
5.3.1. Fondaparinux.....	81
5.3.2. Hyaluronic Acid Tetrasaccharides	83
5.4. Conclusion	85
6. Cryogenic Infrared Spectroscopy Reveals Structural Modularity in the Vibrational Fingerprints of Heparan Sulfate Diastereomers	87
6.1. Introduction	88
6.2. Experimental Details.....	89
6.3. Results and Discussion	90
6.4. Conclusion	96
7. Chondroitin Sulfate Disaccharides in the Gas Phase: Differentiation and Conformational Constraints	99
7.1. Introduction	100
7.2. Experimental Details.....	102
7.3. Results and Discussion	103
7.4. Conclusion	112

8. Predicting Structural Motifs of Heparan- and Chondroitin Sulfates using Cryogenic Infrared Spectroscopy and Random Forest	115
8.1. Introduction	116
8.2. Experimental Details	118
8.3. Results and Discussion	120
8.4. Conclusion.....	132
Summary	133
Samenvatting.....	137
Research Data Management.....	141
Appendices.....	143
References	149
Acknowledgements	I
List of Publications	V
Curriculum Vitae.....	VII

Chapter 1

Introduction

1.1. Glycans and the Diversity in their Chemical Structure

Glycans are among the most abundant biomacromolecules in our biosphere and are essential in growth, function, and fitness of multicellular organisms.^{1,2} Since the first half of the 20th century, the study of glycans has advanced when glycans were considered primarily as source of energy and structural material. Revolutionary is the understanding that biological information is not only contained in DNA (the genome), RNA (the transcriptome) and proteins (the proteome). In fact, metabolites (the metabolome), glycans (the glycome) and lipids (the lipidome) contribute to the enormous biological complexity. By definition, the glycome comprises all glycans in an organism and, indeed, may be considered one of the most diverse entities.

Glycans can exist in free form or in covalent complexes with proteins, lipids or RNA³ as glycoconjugates. In proteins, glycosylation is the most abundant form of post-translational modification. Most often *N*- or *O*-linked, the glycans are involved in structural^{4, 5} or recognition tasks, enabling a variety of specific biological functions, such as fertilization⁶⁻⁸ or immune response.^{9, 10} The glycosylation patterns on proteins can be highly dynamic, which is obvious in an evolutionary sense, but also accounts for pathological variations.^{11, 12} It is often the glycosylation pattern that introduces the distinct diversity in glycoproteins.

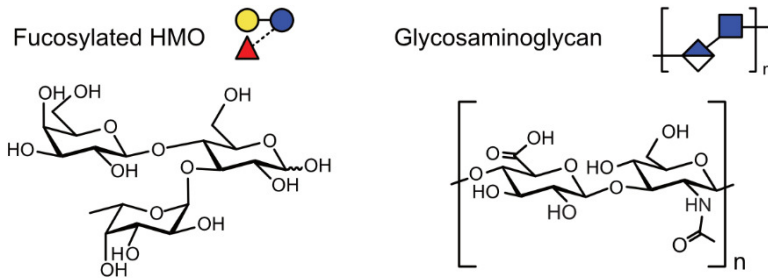
Glycosylation has been the objective of research for novel therapeutic targets in tumor and cancer therapy. During ovarian tumor growth and proliferation, sulfated glycosaminoglycans (GAGs) bound to proteins on the surface of cells facilitate signal transduction and therewith act as co-receptor.¹³ For early detection of hepatocellular carcinoma, glycosylation with fucose (fucosylation) at core positions at the protein is approved as biomarker distinguishing it from e.g. chronic hepatitis.¹⁴

The COVID-19 pandemic accelerated basic research targeting viruses to find drugs and vaccines to treat and prevent billions of infections. Viruses exist in very diverse morphologies, yet some are largely glycosylated to form a glycosylation shield on their surface, which can play a role in the effectiveness of immune evasion.¹⁵ The highly glycosylated HIV-1 virus is considered evasion strong, which can hinder the development of broadly neutralizing antibodies. In contrast, following an infection with or immunization against the SARS and MERS viruses, neutralizing antibodies are readily produced. The overall structure, density and repetitive glycosylation with mannose is found to determine the strength of the glycosylation shield, yet, the immunogenic response of viruses is orchestrated by a multitude of factors.¹⁶

As human milk oligosaccharides (HMO), free glycans have an important role in the development of newborns. They serve as a metabolic substrate for desired intestinal bacteria and therewith lay the basis for a healthy intestinal microbiota composition.^{17, 18} Acting as antimicrobials, HMOs can also lower the risk for viral, bacterial, and protozoan parasite infections. Breast feeding surely provides infants with the best nutrition, nevertheless, the supplement of free HMOs, example given in Figure 1, in commercial infant formula is nowadays extremely well researched and available due to advancements in enzymatic synthesis.^{19, 20}

The terms sugars, glycans, carbohydrates, or saccharides are generally interchangeably used. The diversity in this large family of biomolecules originates already in the smallest building blocks: the monosaccharides. In mammals, ten monosaccharides are considered the most abundant: In a study of more than 3000 mammalian glycans, *N*-acetyl-D-glucosamine was found with 31.8% the most abundant.²¹ In Figure 1, these monosaccharides are shown in a simplified form, using the symbol nomenclature for glycans (SNFG).²² Chemists make use of this short form to ease the reading of larger biopolymers.

A. Example Structures of Common Glycan Classes



B. Symbol Nomenclature for Glycans – SNFG

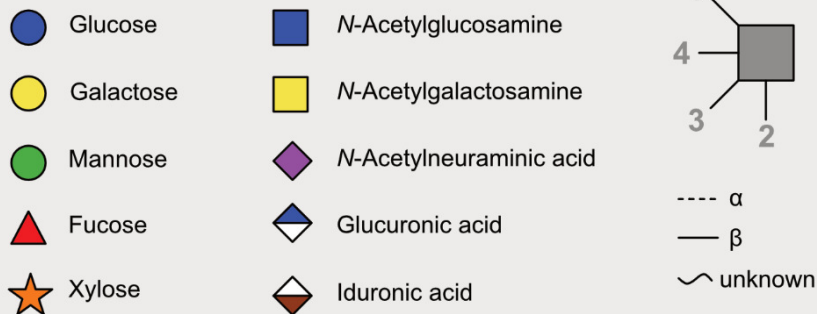


Figure 1: Representation of complex glycan structures in mammals. A. Examples of human milk oligosaccharides (HMOs) and glycosaminoglycans are given. B. The symbol nomenclature for glycans (SNFG)²² offers simplified representations of glycan structures. Symbols are shown for some abundant monosaccharides.

Present in an open chain or heterocyclic five- or six-membered ring form, monosaccharides are grouped by the number of carbon atoms into trioses, tetroses, pentoses, hexoses, and heptoses. These groups are further divided by the functional groups present (in the open chain form), such as an aldehyde in aldoses or a ketone in ketoses. The carbon atoms are labelled numerically beginning with the carbon atom clockwise next to the heterocyclic ring oxygen, see first chemical structure in Figure 2. The orientation of the hydroxy groups in the chain defines the composition, depicted in Figure 2. Already for the aldohexoses, including glucose and galactose, sixteen isomers are chemically possible. The L- or D-configuration are the two enantiomers (image and mirror image) for a monosaccharide, of which the latter is more prominent in nature.

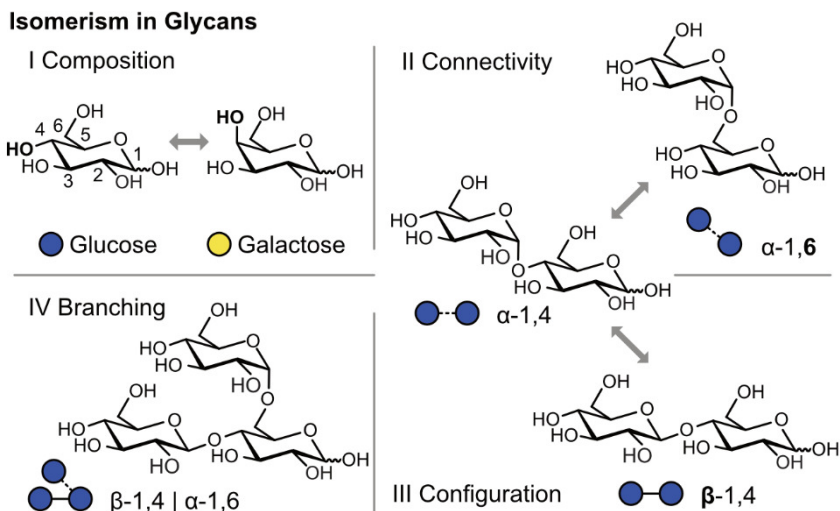


Figure 2: Isomerism in glycans occurs on different levels: glycan composition, connectivity, configuration, and branching.

Via glycosidic bond formation, monosaccharides are polymerized to form disaccharides and further oligo- to polysaccharides. Upon bond formation, another stereocenter is formed, the anomeric center, which defines the configuration of the glycan, see Figure 2. The reducing end (or core; opposed to non-reducing end or terminal) of a glycan chain is at the anomeric C1 position which is not involved in a glycosidic bond but can be otherwise covalently bound, such as in glycoconjugates. Monosaccharides provide different linkage positions so that different connectivity and branching are possible, as depicted in Figure 2. The diversity in glycans—from their composition, connectivity, configuration, and branching—exceeds that of both proteins and DNA. Further attachments such as sulfation, methylation, and phosphorylation are not even considered, yet, important in their functioning. Overall, many aspects in the chemical space of glycans shape their diversity and in the present thesis, two aspects shall be the topic of discussion and introduced in the following: fucosylation in glycans and sulfated glycosaminoglycans.

Fucose is an unusual monosaccharide and abundant in mammals in free glycans, glycoproteins, and glycolipids. A distinct feature is the lack of a hydroxy group

at the C6 position, Figure 3. Further characteristic of the deoxy sugar is the naturally occurring L-configuration, whereas most other biologically relevant monosaccharides are found in D-configuration. Fucose, a 6-deoxy-L-galactose, is almost exclusively the ultimate or penultimate monosaccharide in glycan chains and always in α -configuration. It is attached at a terminal site in *N*-glycans, *O*-glycans or glycolipids, at the core of complex *N*-glycans or also in some cases directly as *O*-glycan to serine or threonine. Thirteen to-date identified fucosyltransferases are involved in the transfer of fucose from guanosine 5'-diphospho- β -L-fucose to the acceptor site at the glycan chain. The enzymes are located in the Golgi apparatus and in case of *O*-fructosyltransferases, in the endoplasmic reticulum.^{21, 23, 24}

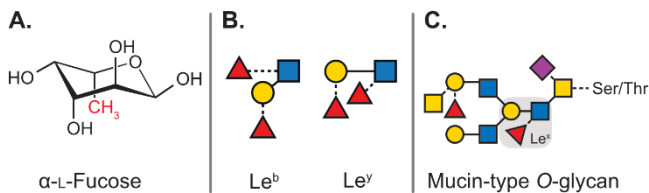


Figure 3: The unusual monosaccharide fucose. A. Most characteristic for the chemical structure is the lack of a hydroxy group at the C6 position. B. The two isomers Lewis b (Le^b) and Lewis y (Le^y) are important blood group antigens. C. In *O*-glycans, which are most abundant on mucins, fucosylation plays an important role in recognition processes, such as fertilization²⁵, and blood group antigens, here Lewis x (Le^x), can be found at terminal positions.

Glycosaminoglycans (GAGs) are a large class of glycans that share common and distinct chemical features that make them diverse and well-recognizable. Probably most characteristic is their long, unbranched and often highly acidic glycan core. The smallest building block are disaccharides with an alternating *N*-acetyl hexosamine (occasionally deacetylated) and a hexuronic acid or galactose. Based on connectivity and monosaccharides, four families are typically defined: hyaluronan, chondroitin- and dermatan sulfate, heparan sulfate and heparin, and keratan sulfate. The hexuronic acids present in GAGs are D-glucuronic acid and its C5 epimer L-iduronic acid. Only keratan sulfate lacks the hexuronic acid and instead possess a galactose monosaccharide. In heparan sulfate and heparin, monosaccharides are exclusively 1,4-linked, whereas in hyaluronan and chondroitin- and dermatan sulfate the hexuronic

acid is 1,3-linked to the *N*-acetyl hexosamine. In keratan sulfate, the *N*-acetyl hexosamine is 1,3-linked to the galactose monosaccharide. The four families are shown in their SNFG representation in Figure 4.

GAGs are produced in different parts of cells. Heparan sulfate and heparin and chondroitin- and dermatan sulfate are synthesized in the Golgi apparatus and are part of large glycoproteins, termed proteoglycans. Sharing a common linker, they are *O*-linked to serine. The family of keratan sulfates is complex and can have different *N*- and *O*-linked conjugation motifs. Hyaluronans are free glycans and are directly synthesized in the cell membrane.

Diversity in glycans and specific biological functions can generally arise from unusual covalent modifications. With the exception of hyaluronan, all GAG families are sulfated with different naturally occurring sulfation motifs, see Figure 4.^{26, 27} Heparan sulfate and heparin are most extensively sulfated and heparin, a highly sulfated form of heparan sulfate, is certainly among the highest negatively charged biopolymers. Heavily sulfated domains can have more than two sulfates per disaccharide on average. In heparan sulfate, the sulfation pattern in a single glycan chain is organized within domains to form heterogeneous sites. In keratan sulfates, the acidic character only stems from the sulfate functional groups.

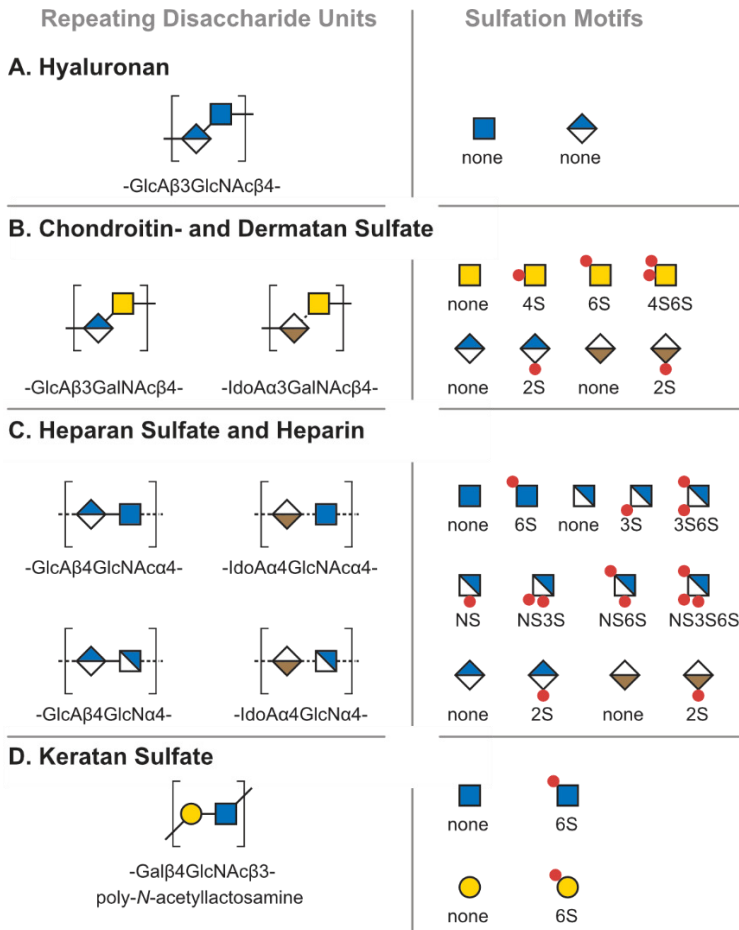


Figure 4: Structural features in glycosaminoglycans: the basic building blocks and sulfation motifs. A. Hyaluronans are the only non-sulfated GAGs. B. Chondroitin and dermatan sulfate are closely related. Chondroitin sulfate contains only D-glucuronic acid whereas dermatan sulfate contains both epimers. C. Heparan sulfate and heparin are the most highly charged GAGs. D. Keratan sulfate contains galactose instead of hexuronic acid.

In proteoglycans, heparan sulfate, heparin, chondroitin, and dermatan sulfate often participate in GAG-protein interactions that mediate biological functions, such as the regulation of fibroblast growth factor signaling, in which the specific sulfation motifs on heparan sulfate (co-receptor) is found essential,^{28, 29} however

this is debated in recent literature.^{30, 31} The bioactive motif in large GAG chains is generally shorter and ranges from a few disaccharides to a dodecasaccharide. The GAG-protein interaction can follow different mechanisms involving discrete bioactive sequences or electrostatic interactions.³² In the latter case, the interactions from negatively charged groups in the GAGs and basic amino-acid residues in the protein ligands are less specific. Nonetheless, the elucidation of the bioactive subsequence to deduce structure-activity relationships is for many interactions the key to understanding the biological functions.

In summary, glycans play an important role in all sources of life. The diversity in their chemical structures—from isomerism, unusual glycans or covalent modifications—influences molecular recognition and may encode biological functions. The analysis and knowledge of certain structural features in glycans is, therefore, a premise to basic research. Elucidating a glycan's code challenges the field of instrumental analysis, and the toolbox of techniques shall be introduced in the following sections.

1.2. Standard Techniques for the Structural Analysis of Glycans

Analytical chemists strive to find the *gold standard* for the structural analysis of a certain class of molecules. In glycan analysis, the utilized techniques most often originate from genomics and proteomics. Contrary to proteins, the structure of glycans from mammalian source is not directly encoded in the genome. Glycan synthesis is non-template derived and underlies a higher dynamic, hindering the simple transfer of techniques. Over the past years, the analyses of the glycome has defined its own measures and techniques, which shall be briefly reviewed here. The focus in the present thesis lies on the analysis of glycans, therefore, selected literature for the analysis of intact glycoconjugates shall only be cited here.³³⁻³⁶

Glycans from natural sources often require diverse isolation and enrichment strategies prior to their structural analysis.³⁷⁻⁴⁰ In case of glycoconjugates, the glycans are often enzymatically or chemically released. Large glycosaminoglycan (GAG) chains can further be reduced in size by chemical depolymerization or lyase digestion to determine the bulk properties of the GAG population.^{41, 42} For the separation of glycans, commonly employed techniques include several

types of liquid or, less frequently electrophoretic chromatography, such as reversed-phase high-performance liquid chromatography (RP-HPLC),⁴³ capillary electrophoresis (CE),⁴⁴ hydrophilic interaction liquid chromatography (HILIC),⁴⁵ porous graphitized carbon (PGC)⁴⁶ chromatography or strong anion exchange (SAX)⁴⁷ chromatography. For the detection of the analytes, mass spectrometry (MS) analysis and ultraviolet or fluorescence detection are used, of which the latter demands additional derivatization steps. Derivatization steps can not only serve the detectability of glycans but also e.g. enhance their ionization in the MS detection and aid their sequencing.^{40, 48, 49}

Nuclear magnetic resonance (NMR) spectroscopy gives information on the primary structure of glycans and many glycans from natural sources have been identified this way. Especially from the signals of the anomeric carbons, the configuration of monosaccharides can be directly deduced. In combination with databases, a number of structural features are distinguishable, but most often NMR is combined with data from MS. Elaborate pulse sequences and NMR spectroscopy of active nuclei of low natural abundances require large sample amounts, long measurement times and large expenses for instrumentation and consumables. Still, in the early years of glycan analysis, NMR was one of the most important techniques.^{37, 50, 51}

Mass spectrometry (MS) experiments increased in popularity in glycan analysis due to its information content at high sensitivity and speed. Indeed, excellent reviews have been published throughout the last years.^{33, 34, 52-56} A mass spectrum directly yields the mass-to-charge ratio (m/z) of ionized molecules over a large range of biomolecules, from small metabolites to large viruses.⁵⁷⁻⁵⁹ Moreover, the mass of a small ionized glycan—with sufficient accuracy—can provide information on the present monosaccharides in terms of hexoses versus *N*-acetyl hexosamines and deoxyhexoses. The broad applicability of MS certainly arises from different ionization modes (cations or anions, the latter is useful for GAGs) and the use of coordinating ions (adduct formation).

The ionization of the sample is the beginning of every MS analysis. In the past decade, electrospray ionization (ESI)^{60, 61} has become the dominant ionization method over matrix-assisted laser desorption/ionization (MALDI)^{62, 63} for glycans. Small to large glycans are especially gently transferred from solution to

the gas phase. In the ESI process, a high voltage, typically +/- 0.5 to 3 kV, is applied to a capillary containing the analyte solution. A counter electrode is placed at a short distance from the capillary's tip. Here, the strong electric field leads to a distortion of the liquid surface's shape to form a cone (Taylor cone). When the electric field exceeds a critical voltage, the liquid Taylor cone breaks up into droplets that move towards the counter electrode. From these charged droplets that contain the analyte, solvent evaporates until the Coulombic repulsion inside the droplet is too large. Moreover, the fission of the droplet leads to second generation droplets. After several iterations of evaporation and fission, the droplets are so small that the bare ion is eventually released into (ion evaporation model) or obtained in (charge residue model) the gas phase. ESI capillaries with a smaller opening of the tip, termed nano ESI (nESI), are a variation of increasing popularity due to the reduced sample consumption, increased sensitivity, and enhanced ionization efficiency.⁶⁴⁻⁶⁶

In tandem MS (or MS/MS) and MSⁿ, ionized analytes are activated in the gas phase by collisions,⁶⁷⁻⁶⁹ electrons⁷⁰⁻⁷² or photons^{73, 74} to produce targeted fragmentation. Each method allows for specific fragmentation mechanisms and can yield complementary information. The fragment ions are detected and are ideally highly specific to the precursor ion to deduce its chemical structure. Collision-induced dissociation (CID) is probably most widely used in glycan analysis. Ions are accelerated in an electric field and undergo several inelastic collisions with neutral gas molecules. With the sequential activation, energy is redistributed throughout the ion and the weakest bonds are typically cleaved. In positive ion mode, glycosidic bond cleavage is most dominant but also labile functional groups, such as sulfates, are dissociated easily. For anions, cross-ring cleavages are often present, which can indicate branching and connectivity. A limitation of tandem MS is that many glycan isomers with identical mass exhibit highly similar fragmentation spectra, especially in positive ion mode. This circumstance highlights the need for an additional separation dimension, such as in LC-MS/MS.

Rearrangement reactions in targeted fragmentation experiments can be used to produce highly characteristic fragment ions but they have always challenged the interpretation of results.⁷⁵ In the early days of proteomics, a gas-phase

rearrangement reaction was observed which led to erroneous sequence assignments. Peptide scrambling in protonated or multiply protonated ions involves head-to-tail macrocyclization and reopening prior to CID. As a consequence, the original sequence information is lost.⁷⁶⁻⁸⁰

A rearrangement reaction frequently observed in CID of glycan cations is the intramolecular transfer of a terminal fucose monosaccharide, termed fucose migration or internal residue loss (IRL).⁸¹ The rearrangement leads to the presence of misleading fragment ions, which in turn can produce faulty structural assignments. This rearrangement reaction is regarded a universal phenomenon and is frequently observed for *O*-glycans but also for fucosylated glycans from other glycan classes. During migration, a fucose residue is transferred from the nonreducing end of a glycan to an adjacent or remote site within the same molecule. In fragmentation experiments using CID, fucose migration may be accompanied by internal residue loss (IRL). Unexpected *m/z* fragments indirectly indicate the occurrence of the reaction. Besides fucose, rearrangement reactions have been observed for xylose,^{82, 83} rhamnose,^{83, 84} and glucuronic acid,⁸³ with distinct masses, but also mannose⁸⁵ is able to migrate and more challenging to detect. Several studies based on tandem MS have been published investigating the reaction mechanism and driving forces, the kinetics and energetics, and, of utmost concern, the destination of the migrating monosaccharide. Both short-⁸⁶ and long-range⁸⁷ migration reactions are possible. For both cases, the close proximity in space of the migrating residue and the destination within the chain are important, yet difficult to predict. The rearrangement reaction has been observed for 1,2-, 1,3-, 1,4-, and 1,6-linked migrating monosaccharides. Considering the time scale of the reaction, experiments using MALDI-TOF/TOF-MS show that the reaction for the investigated ions is faster than microseconds.⁸⁸ The abundance of ion signals resulting from IRL is collision energy-dependent. When plotting the relative abundance of an ion against the collision energy, the ion from a simple loss of a terminal fucose and the ion from IRL with migration of fucose show the same

curve shape, indicating a similar mechanism.⁸⁶ Yet, the exact mechanism of fucose migration remains unresolved to date.*

Activation of ions in mass spectrometers is not only associated to targeted fragmentation experiments but can also occur unintended in other parts of the instrument. Radio frequency (RF) fields in regions with higher pressures, e.g. in ion traps or ion guides in transfer stages, are known to induce ion heating.⁸⁹ Furthermore, also in the source region, collisions with residual gas molecules, that cause unintended activation of the analytes, can frequently occur.⁹⁰ Subsequently, a major challenge in the investigation of sulfated GAGs is the neutral loss of SO₃ without explicit activation.^{42, 91} The decomposition hinders the analysis of the sulfate pattern and can lead to ambiguous results. In positive ion mode, counter ions such as quaternary ammonium and phosphonium salts have shown to stabilize sulfate groups.⁹² In negative ion mode, the extend of sulfate loss is considerably decreased for charged sulfate groups.⁹³ Using coordinating cations, such as metals, the charge density can be further reduced in highly sulfated and highly charged GAGs to stabilize the ions in the gas phase.⁹⁴

To summarize, in recent years, a variety of approaches evolved for the identification of complex glycans. Mass spectrometry (MS) is certainly one of the most widely used tools in glycan analysis. However, structure-sensitive methods are needed to enable unambiguous sequencing or to fully understand gas-phase rearrangement reactions. Therefore, ion mobility-mass spectrometry (IM-MS) and MS-based infrared (IR) spectroscopy shall be introduced as orthogonal techniques to advance the MS-based toolbox.

1.3. Ion Mobility-Mass Spectrometry

Ion mobility spectrometry (IMS) is a versatile gas-phase separation technique routinely applied in combination with mass spectrometry (IM-MS) for the separation and structural analysis of glycans.⁹⁵⁻⁹⁸ Different measurement

* Paragraph adapted from: Grabarics, M.; [Lettow, M.](#); Kirschbaum, C.; Greis, K.; Manz, C.; Pagel, K., Mass Spectrometry-Based Techniques to Elucidate the Sugar Code. *Chem. Rev.* **2022**, 122, 7840-7908.

principles have been developed and utilized for the analysis of glycans, e.g., travelling wave IM-MS or trapped IM-MS. In the present thesis, drift tube IM-MS (DTIM-MS) shall be addressed, which is the oldest IM-MS technique. The basic idea of DTIM-MS is that ionized molecules of varying size, charge, and shape drift with different velocities through a buffer gas under the influence of an electric field. Upon their motion through the electric field, the ions collide with the particles of the neutral and inert buffer gas, e.g., nitrogen or helium. Compact ions undergo less collisions and travel faster through the drift cell than extended ions. While MS alone is sensitive to the weight of an ion, IMS is sensitive to the shape. In Figure 5, a drift cell is schematically drawn with molecules of identical m/z -ratios but different shapes. MS experiments are typically performed in microseconds and IMS experiments are on the millisecond range, perfectly matching orthogonally.

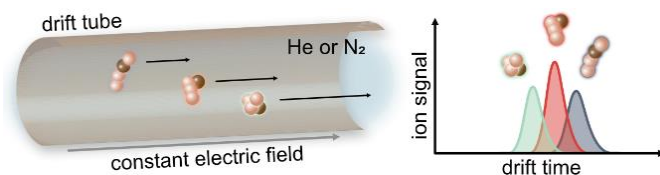


Figure 5: Scheme of three molecules of identical m/z -ratio that drift through a buffer gas-filled drift cell under the influence of a constant electric field. Their drift time through the drift cell is recorded and plotted as arrival time distribution.

In recent years, standardized protocols to perform measurements and to report results have been developed to further aid the comparability across laboratories.^{99, 100} For the optimization of measurement conditions and to increase the separation power of an instrument, a practical approach has been developed that transfers the widely-used plate-height model from chromatography to IM-MS.^{101, 102} These efforts further strengthen the position of IM-MS in the analytical chemistry community.

In DTIM-MS, ions are directed through a stationary gas by applying a constant and homogeneous electric field E so that ions with higher mobilities undergo less collisions and travel faster. Therefore, DTIM-MS is a time-dispersive measurement principle. When operated in the “low-field limit”, the mobility K

of an ion is directly proportional to the velocity v_d , which derives from the length of the drift tube L and the observable drift time t_d .

$$K = \frac{v_d}{E} = \frac{L}{E \cdot t_d} \quad (1)$$

The ion's mobility depends on the pressure p and the temperature T . For comparison across laboratories, the reduced mobility K_0 is, therefore, normalized for standard pressure p_0 and temperature T_0 .

$$K_0 = K \cdot \frac{p}{p_0} \cdot \frac{T_0}{T} \quad (2)$$

Generally, the mobility of an ion is further converted to the rotationally-averaged collision cross section (CCS) Ω of an ion. The Mason-Shamp equation^{103, 104} relates the ion's reduced mobility to the ion's CCS:

$$\Omega = \frac{3ze}{16N} \sqrt{\frac{2\pi}{\mu k_B T}} \frac{1}{K_0} \quad (3)$$

where z is the ionic charge state, e the elementary charge, N the buffer gas number density, μ the reduced mass of the ion-neutral collision partners and k_B the Boltzmann constant. An advantage of DTIM-MS is that the CCS (^DTCCS) can be determined experimentally without further calibration.

Hence, CCS values are comparable, physical quantities that can be readily stored in databases¹⁰⁵ and be compared to calculated CCS, commonly using the trajectory method today.¹⁰⁶⁻¹⁰⁸ This extends the objective of IM-MS experiments from separating glycans to identifying and sequencing glycans based on their gas-phase shape.¹⁰⁹⁻¹¹¹

1.4. Mass Spectrometry-based Infrared Spectroscopy*

1.4.1. General Principles

Mass spectrometry-based infrared (IR) spectroscopy is directly implemented within the mass spectrometer and yields a structure-sensitive fingerprint of a glycan, complementing the information from MS.¹¹²⁻¹¹⁷ In the following

*Adapted from: Grabarics, M.; Lettow, M.; Kirschbaum, C.; Greis, K.; Manz, C.; Pagel, K., Mass Spectrometry-Based Techniques to Elucidate the Sugar Code. *Chem. Rev.* **2022**, 122, 7840-7908.

sections, the general principles are introduced and compared to classical absorption spectroscopy. Furthermore, the types of instruments are described.

IR spectroscopy is a powerful tool used to directly identify unknown molecules from their IR spectrum and deduce information on their structure, such as functional groups, intra- and intermolecular interactions, as well as molecular conformations. Electromagnetic radiation in the IR range can excite molecules if the frequency of the incident radiation is in resonance with IR active vibrational transitions. IR radiation expands from the edge of the visible spectrum to the microwave range and can be further divided into the higher energy, near IR ($>4000\text{ cm}^{-1}$), mid-IR ($4000\text{--}400\text{ cm}^{-1}$) and lower energy, far IR ($400\text{--}10\text{ cm}^{-1}$) regime. Most stretching and bending vibrations are found in the mid-IR range. Classical IR spectroscopy techniques are based on direct absorption spectroscopy, which measures the attenuation of light after passing through a solid, liquid, or gaseous sample. The absorbance as a function of the frequency is derived from the Lambert-Beer law. The concept of direct absorption spectroscopy is depicted in Figure 6.

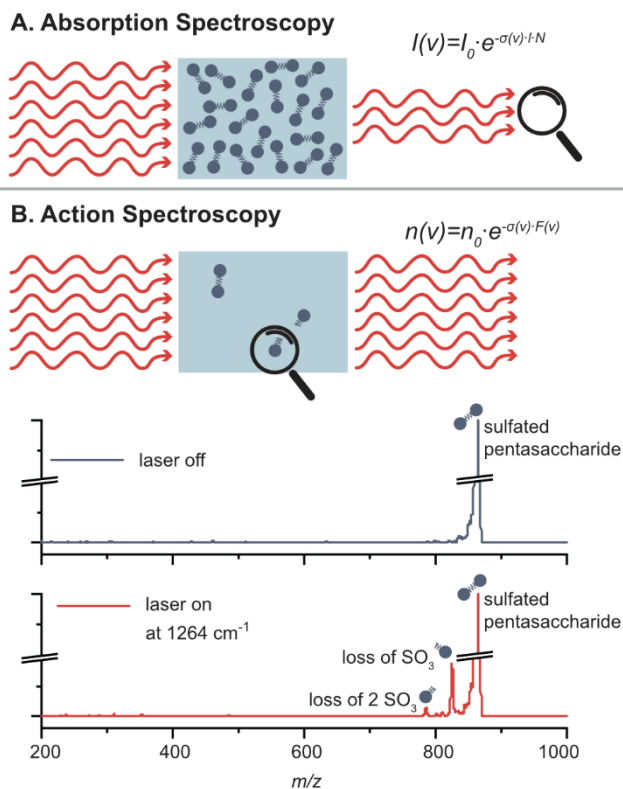


Figure 6: Direct absorption vs. action spectroscopy. A. Concept of direct absorption spectroscopy. The Lambert-Beer law relates the intensity of transmitted light I at a specific frequency ν to the intensity of the incident light I_0 , the absorption cross section σ , the path length l , and the particle density N . B. Concept of action spectroscopy (upper panel). The equation is derived from the Lambert-Beer law and relates the number of unaffected ions n at a specific frequency ν to the number of precursor ions n_0 , the absorption cross section $\sigma(\nu)$, and the photon fluence $F(\nu)$. Infrared multiple photon dissociation time-of-flight mass spectra of a m/z -selected highly sulfated pentasaccharide without (middle panel) and with (lower panel) resonant IR irradiation. Two fragments with a sequential loss of neutral SO_3 are observed upon multiple photon dissociation at 1264 cm^{-1} .

For m/z -selected ions in the high vacuum of a mass spectrometer, the sample density is usually limited to 10^6 singly charged ions per cubic centimeter due to Coulomb repulsion and the resulting space-charge limit.¹¹⁸ The attenuation of light after passing through the low-density sample is so low that it is very hard

or in most cases even impossible to measure. Therefore, vibrational spectroscopy of ions in the gas phase is typically performed as action spectroscopy, as addressed in recent reviews.^{117, 119} Here, the absorption of photons is measured indirectly by following an action, e.g., the dissociation of the intact ion or electron detachment as in gas-phase ultraviolet spectroscopy. A schematic concept of action spectroscopy can be found in Figure 6.

In action spectroscopy, the number of unaffected ions n at a specific frequency ν can be expressed as a function of the number of precursor ion n_0 , the absorption cross section $\sigma(\nu)$, and the photon fluence $F(\nu)$ in the following equation which is derived from the Lambert-Beer law:¹¹⁷

$$n(\nu) = n_0 \cdot e^{-\sigma(\nu)F(\nu)} \quad (4)$$

The challenge is that IR light is usually much lower in energy than the threshold to dissociation; therefore, the analytes must have a low barrier to dissociate, or the laser system must be of high enough intensity to enable multiple photon absorption. With the advance of tunable and powerful laser systems such as IR free-electron lasers (FELs)¹²⁰⁻¹²² and benchtop laser systems such as optical parametric oscillators and optical parametric amplifiers (OPO/OPA), the interest in gas-phase IR spectroscopy of m/z -selected ions for their structural analyses started growing. Early works by the Simons group^{123, 124} investigated small, neutral glycans in a UV-IR double-resonance experiment from a free jet expansion. Limited to the presence of UV chromophores and the harsh ionization method, the method was only applicable to small glycans. Combined with ESI sources, the field of IR spectroscopy finally opened up for the investigation of larger glycan ions.¹²⁵⁻¹²⁷

Theoretical calculations serve to correlate and explain experimental results from IR spectroscopy. To model glycan ions, empirical force field and semiempirical and first-principle methods are typically used.^{114, 128} The conformational space and structural features of glycans, with ring puckering, glycosidic bond geometry, charge migration processes, and hydrogen bonding, renders the sampling of possible chemical structures challenging. For harmonic^{129, 130} and anharmonic¹³¹ frequency calculations, the sampled structures need to be reoptimized at a sufficiently high first-principle level of theory, which is costly,

especially for anharmonic calculations. Yet, on the long run, computational resources and the accuracy of calculations could match the needs of larger glycans.

1.4.2. Experiment Designs

Infrared Multiple Photon Dissociation

One of the most widely used types of IR action spectroscopy is infrared multiple photon dissociation (IRMPD) spectroscopy, in which weakly bound clusters or ions with energetically low fragmentation barriers are fragmented upon irradiation.^{117, 119} An IR spectrum is recorded by measuring the fragmentation yield as a function of the wavelength. The IRMPD mechanism is a nonlinear process involving the sequential absorption of a large number of IR photons (typically tens to hundreds) in which the energy of a single photon is distributed throughout the ion via internal vibrational distribution (IVR); i.e., the original vibrational mode relaxes via anharmonic coupling to vibrational background states. The sequential absorption of single photons, hence, takes place via the same fundamental vibrational level; yet the internal energy of the ion rises with vibrational excitation until the dissociation threshold is reached and the ion fragments. With the IVR-coupled excitation process, the vibrational excitation is randomized throughout the ion so that the fragmentation occurs statistically, and usually the weakest bonds dissociate. For the above reason, IRMPD fragmentation patterns closely resemble those obtained by CID. Inherent to the process, red-shifting of absorption bands and spectral congestion can leave an imprint on the IRMPD spectrum. Additionally, glycans typically populate a large number of coexisting conformers at room temperature due to their conformational flexibility, which further increases spectral congestion and limits IRMPD spectroscopy to mostly mono- and disaccharides.^{125, 132-134}

Various designs of IRMPD instruments are published, yet in the most basic approach an interaction region with electrodynamic ion optics is sufficient to radially define the ion cloud and achieve an efficient overlap of the ions with the laser. Ion storage proves beneficial for slow dissociation processes so that the ions can be irradiated for a longer time. Online coupling of IRMPD with

LC-MS workflows is challenging because the time required to record an IRMPD spectrum usually exceeds 10 min, which is not compatible with the time scale of LC peaks. However, efforts are underway to generate IRMPD spectra from chromatographically separated glycans.¹³⁵

Messenger-Tagging Spectroscopy

Another approach that does not require multiple photon excitation steps is messenger-tagging spectroscopy which follows the detachment of a weakly bound and weakly interacting messenger atom or molecule upon resonant irradiation, schematically shown in Figure 7. The typical tag for the investigation of glycans is nitrogen,¹²⁶ and for other molecules the use of helium,¹³⁶ argon,¹³⁷ and hydrogen¹³⁸ has been explored. Contrary to IRMPD spectroscopy, the IVR-mediated detachment mechanism of the tag in messenger-tagging spectroscopy is a nearly linear process overcoming red-shifting and reducing spectral congestion. Furthermore, the power of the tunable IR laser required for messenger-tagging spectroscopy is significantly lower and thus renders the approach more compatible with benchtop lasers. Coincidentally, this limits the numbers of glycan conformers present, therewith further reducing spectral broadening. Several IR spectroscopic studies on glycans, small building blocks up to oligosaccharides, using messenger-tagging spectroscopy have been published.^{139, 140}

Already in the 1980s, messenger-tagging spectroscopy was developed as a concept using a supersonic jet expansion followed by electron ionization or corona discharge. With the low temperatures in the supersonic jet, ions or clusters formed weakly bound complexes with coexpanded hydrogen. The detachment of the hydrogen tag upon irradiation was detected in a quadrupole mass analyzer. Since the mid-1990s, the development of cryogenic ion traps^{141, 142} enabled the investigation of more fragile (bio-)molecules produced from soft ionization sources. Here, ions are transferred to the cryogenic ion trap, cooled by collisions, and eventually tagged with buffer gas at temperatures between 3 and 70 K. Various designs for cryogenic ion traps are developed to serve different purposes: such as temperatures below 3 K, long storage times, spatial spread of the ion cloud, or space focusing of ions with the ejection from the trap. First, multipole ion traps were developed,¹⁴³⁻¹⁴⁵ and later planar multipole

ion traps¹⁴⁶⁻¹⁴⁸ and ring ion guide traps¹⁴⁹ followed. It was recently published¹³⁶ that a wire quadrupole ion trap, with a linear quadrupole geometry in which each of the rods is approximated by six copper wires, reaches temperatures below 3 K and is attached to a commercial mass spectrometer.

Infrared Spectroscopy in Helium Nanodroplets

IR action spectroscopy in helium nanodroplets is currently applied in basic research only, yet the spectral quality which is achievable for glycan ions sets new benchmarks in the field.^{96, 114} The superfluid helium nanodroplets are produced from a precooled reservoir of helium with evaporative cooling and formation of nanodroplets of a defined size distribution (typically 10^5 helium atoms). Before irradiation, the trapped ions are picked up by the traversing helium nanodroplets and cooled to the equilibrium temperature of the droplet of 0.37 K. Upon irradiation with a resonant photon, the ion is vibrationally excited and quickly cooled again to its ground state by evaporation of helium from the shell of the droplet. After several iterations, the ion is eventually released from the droplet and detected background-free. With this approach, a unique resolving power and distinct spectral fingerprints have been recorded for various classes of glycans and glycoconjugates.^{35, 127, 150-153}

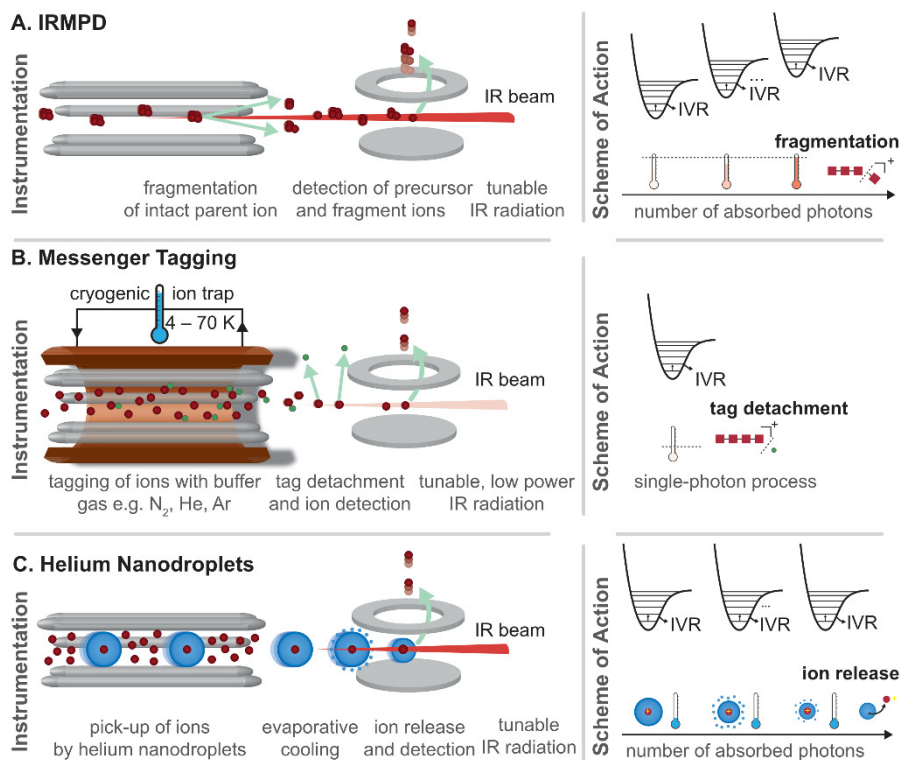


Figure 7: Comparison of types of IR action spectroscopy. A. Instrumentation and scheme of action for infrared multiple photon dissociation (IRMPD) spectroscopy. With resonant irradiation, multiple IR photons excite the intact precursor ion until the fragmentation threshold is reached. Fragment and precursor ions are detected. B. Instrumentation and scheme of action for messenger (tagging) spectroscopy. Ions are tagged with buffer gas atoms or molecules, e.g., N₂, He, or Ar, in a cryogenic ion trap. With resonant irradiation typically with a single photon, the tag is detached, and the bare ion is detected. C. Instrumentation and scheme of action for cryogenic spectroscopy in helium nanodroplets. The ions are picked up by helium nanodroplets in an ion trap and cooled to 0.4 K. With resonant irradiation, the ion is excited and immediately cooled again by evaporative cooling. After several iterations, the ion is released from the nanodroplet and detected.

1.5. Aim and Outline

As already stated before, glycans are omnipresent in nature and make up the largest amount of the biomass found on Earth. The field of glycosciences has only recently emerged more strongly. The understanding that glycans play central roles in biological processes has eventually shifted the focus from the genome and proteome to the glycome. Here, terminal sequences, unusual glycans (such as fucose), and covalent modifications (such as sulfation) are known to play a central role. The recent technological developments in IR spectroscopy hold great potential for the identification of complex glycans in the gas phase.

The first aim of this doctoral work is to understand the rearrangement reaction, termed fucose migration, which is frequently observed in tandem MS experiments of fucosylated glycans using cryogenic IR spectroscopy in the gas phase. The work is built on the previous observation that the reaction takes place in intact glycan ions and is promoted without explicit activation of the ions. To date, only assumptions have been made on the final product of the rearrangement reaction in the prominent fucosylated epitopes Lewis x and blood group H type 2. Furthermore, several mechanisms have been proposed but a common understanding has not been achieved. The aim is to give an in-depth insight into the triggers of the reaction, the final product and eventually the mechanism. From the exact chemical structure of an ion, a deeper insight into the reaction can prove beneficial for the development of MS strategies.

The second aim of this doctoral work is to establish MS-based IR spectroscopy as an addition to the MS-based toolbox for the sequencing of glycosaminoglycans. The linear, acidic and often highly sulfated glycans still call for more advanced analytical tools that can unambiguously identify sulfation positions and epimerization. To date, standard techniques, such as LC-MS/MS and IM-MS, reach their limits and novel methods are being explored.

An overview of the gas-phase IR spectroscopy experiments performed in this thesis, namely IRMPD spectroscopy and spectroscopy in helium nanodroplets, is given in **Chapter 2**. Furthermore, the operational mode of a new apparatus for messenger-tagging IR spectroscopy of mobility-selected ions is described

and initial experiments presented. Messenger-tagging IR spectroscopy for glycan ions will presumably be more broadly accessible to analytical chemists due to the simpler design (compared to helium nanodroplet spectroscopy) but also feature a high resolving power which is needed for larger glycan ions (compared to IRMPD).

In the past, it has been shown that fucose migration can occur in intact glycan ions to which MS alone is blind. In **Chapter 3**, the prototypical glycans Lewis x and blood group H type 2 are studied with adduct ions and additional functional groups of varying proton affinities to deduce the role of a mobile proton in the reaction. This chapter is based on the publication "*The Role of the Mobile Proton in Fucose Migration*" published in the journal Analytical Bioanalytical Chemistry. Using a combination of three MS-based techniques (IM-MS, radical-directed dissociation (RDD) and cryogenic IR spectroscopy) combined with computational methods, the final structure and the mechanism of the migration reaction are finally decoded in **Chapter 4**. This chapter is based on an early version of "*Decoding the Fucose Migration Product in Blood Group Epitopes*".

The high resolving power of cryogenic IR spectroscopy was previously used to identify conformation in peptides, charge-induced structural changes in proteins, small isomeric glycans, reaction intermediates in glycosylation reactions and glycolipids. To assess the potential of cryogenic IR spectroscopy for GAG analysis, the IR spectra of a sulfated GAG pentasaccharide and tetrasaccharide are discussed in **Chapter 5**, which is based on "*Infrared Action Spectroscopy of Glycosaminoglycans*" published in the journal Analytical Bioanalytical Chemistry. Subsequently, systematic investigations of epimerization and sulfation are presented in **Chapter 6** and **Chapter 7**, respectively. Chapter 6 is based on "*Cryogenic Infrared Spectroscopy Reveals Structural Modularity in the Vibrational Fingerprints of Heparan Sulfate Diastereomers*" published in Analytical Chemistry and Chapter 7 is adapted from "*Chondroitin Sulfate Disaccharides in the Gas Phase: Differentiation and Conformational Constraints*" published in Journal of Physical Chemistry A. Finally, in **Chapter 8**, cryogenic IR spectroscopy is combined with random forest modelling to predict structural features in GAGs that almost fully characterize the underlying chemical structure, namely the class and sulfation positions. This chapter is in parts based on an early version of

“Predicting Structural Motifs of Heparan- and Chondroitin Sulfates using Cryogenic Infrared Spectroscopy and Random Forest”.

Chapter 2

Experimental Overview

2.1. Infrared Multiple Photon Dissociation Spectroscopy

Infrared multiple photon dissociation (IRMPD) is the most widely applied type of infrared (IR) spectroscopy in the gas-phase environment of a mass spectrometer. The original instrumental setup for IRMPD spectroscopy, utilized in this thesis, is described in the following section. The apparatus is redesigned in the course of this thesis, which is the topic of the following section. The setup is connected to the Fritz Haber Institute free-electron laser (FHI FEL), which is briefly introduced at the end of this chapter. Figure 8 shows a schematic overview of the instrument, which can be divided into four parts: (A) mass spectrometry: with nano-electrospray ionization (nESI) source and quadrupole mass filter, (B) drift tube ion mobility spectrometry (DTIMS), (C) interaction region for IRMPD and (D) the IR light source.

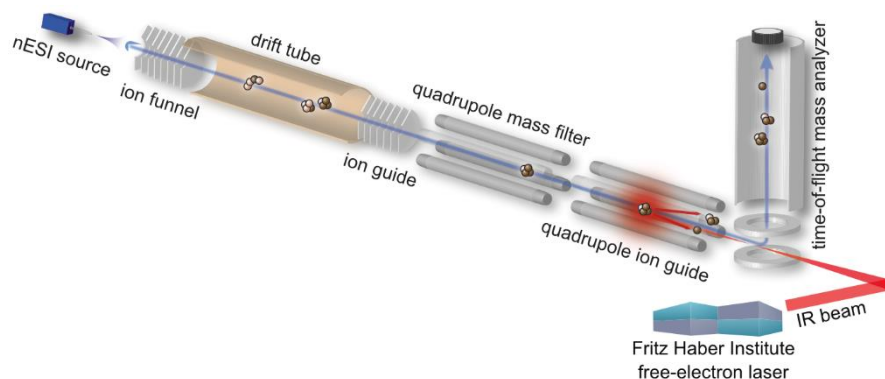


Figure 8: Scheme of the home-built setup for infrared multiple photon dissociation spectroscopy combined with drift tube ion mobility spectrometry which was a user station at the free-electron laser of the Fritz Haber Institute.

With the nESI source of the home-built instrument, ions were produced by applying needle voltages of approx. 0.5 kV for the investigated cations. The nESI glass capillaries which are used on all mass spectrometers in this thesis are

produced using a micropipette puller (Flaming/Brown P-1000, Sutter Instrument, USA) and coated with Pd/Pt alloy with a sputter coater (108auto, Cressington Scientific Instruments, UK). The tip of the glass capillaries is cut by hand using tweezers. The ions are transferred to the entrance funnel via a capillary which is slightly off-axis to the first, hourglass-shaped ion funnel. The charged analytes are collected and radially focused in the ring-electrode, radio frequency ion funnel, or entrance funnel. The ions are trapped in the second half of the entrance funnel before being injected to the drift tube, typically at 10–20 Hz. The drift region consists of conductive glass tubes and is filled with helium or nitrogen buffer gas as drift gas at typically 2–3 mbar for nitrogen and 3–6 mbar for helium. Ion mobility spectrometry measurements are performed within the low-field regime with field strengths of 5–18 V cm⁻¹. For CCS calculation, the temperature in the drift region is measured with a Pt100 resistance thermometer and the buffer gas pressure is monitored with a high-precision absolute pressure transducer (type 627, MKS Instruments, USA). A pressure difference of 0.6 mbar is maintained between the drift region and the entrance funnel to avoid contamination. At the end of the drift region, the ions are collected and radially confined in a second ring-electrode, radio frequency ion funnel, or exit funnel. The ions are transferred to high vacuum via two ring-electrode ion guides which are differently pumped. Via electrostatic deflection in an ion gate, ions of a particular arrival time are selected for further investigation. Before entering the interaction region, a quadrupole mass filter is used to further select the ions of interest by their m/z -ratio. The interaction region is a second quadrupole in which the IR beam from the FHI FEL overlaps the radially confined ion cloud. The focus point of the IR beam is adapted dynamically using a pneumatically controlled adaptive mirror (AO90/70, Kugler, Germany) to control the number of photons per area unit sufficient for the dissociation process. For ATD measurements, the ions are detected using an off-axis electron multiplier (AF632, ETP Ion Detect, Adaptas, USA). For IRMPD spectroscopy measurements, the ions are deflected perpendicular to a home-built Wiley-McLaren-type of time-of-flight mass analyzer equipped with an electron multiplier (AF882, ETP Ion Detect, Adaptas, USA).

In spectroscopy experiments, the abundances of the precursor ion N_p and fragment ions $\sum N_f$ are recorded as a function of the laser's wavenumber in steps of 2 cm^{-1} . The IMS part of the instrument is operated at 20 Hz and the FHI FEL at 10 Hz. Monitoring the intensities of the non-irradiated ions with every scan step gives a background spectrum and helps to determine the stability of the ion source. After two IR spectra are successfully recorded, the laser power and wavenumber are calibrated by measuring the power and the exact wavenumber at ten wavenumbers over the operated wavenumber range. The laser power is fitted using a quartic polynomial function. To obtain the IR spectrum from the absorption cross section $\sigma(\nu)$ in Equation (4), Section 1.4.1, the total number of ions N_0 is replaced with the sum of N_p and $\sum N_f$ and the equation is rearranged to:

$$\sigma(\nu) \propto -\ln\left(\frac{N_p(\nu)}{N_p(\nu) + \sum N_f(\nu)}\right) \cdot \frac{1}{F(\nu)} \quad (5)$$

The ion signal is divided by the laser power as a first order approximation. For each analyte, at least two independent scans with high reproducibility are measured and averaged. The ion signal intensity is normalized to [0,1]. The relative intensity of each partial spectrum, when using different laser focus settings, is scaled on the basis of the intensity of spectral lines found in the overlapping measurement regions.

The drift region consists of segmented, conductive tubes of glass (Photonis Scientific, USA) and was extended from 805.5 mm to 1612 mm in length in the frame of this thesis. Each new segment is commercially manufactured to a homogenous electric resistance in the $\text{G}\Omega$ range over a length of approx. 200 mm. Additional resistors on the outside of each segment and on the connection between two segments modulate the resistance so that the electric field throughout the old and new drift region and among each segment is nearly homogenous. The electric resistance over the total length of the drift region is $185 \text{ M}\Omega$, hence each segment has an electric resistance of $23 \text{ M}\Omega$. The time period that a measurement is being made over one measurement cycle, or the IMS duty cycle, is now twice as high. The new IMS resolving power $R = t_d/\Delta t$ is experimentally determined approx. 55 using the tetrabutylammonium iodide trimer $[(\text{TBA})_3\text{I}_2]^+$ with m/z 980 and $\text{D}^1\text{CCS}_{\text{He}}$ of 270 \AA^2 (Appendix).

2.2. Messenger-Tagging Spectroscopy

2.2.1. The General Design

Messenger-tagging spectroscopy addresses the shortcomings of IRMPD, i.e. heating of the ion with subsequent red-shifting and broadening, by altering the type of action. Here, the action is the dissociation of a messenger tag from the ion which generally requires much less energy. The design and the operation of the setup for messenger-tagging spectroscopy (Figure 9) is the topic of the following sections. It is an extension to the DTIM-MS part and a redesign of the spectroscopy part of the setup described in Section 2.1. In the frame of this thesis, the apparatus was built and operated for the first beam time shifts at the FHI FEL. The development of the idea and design of the apparatus is the work of several people over the last years.

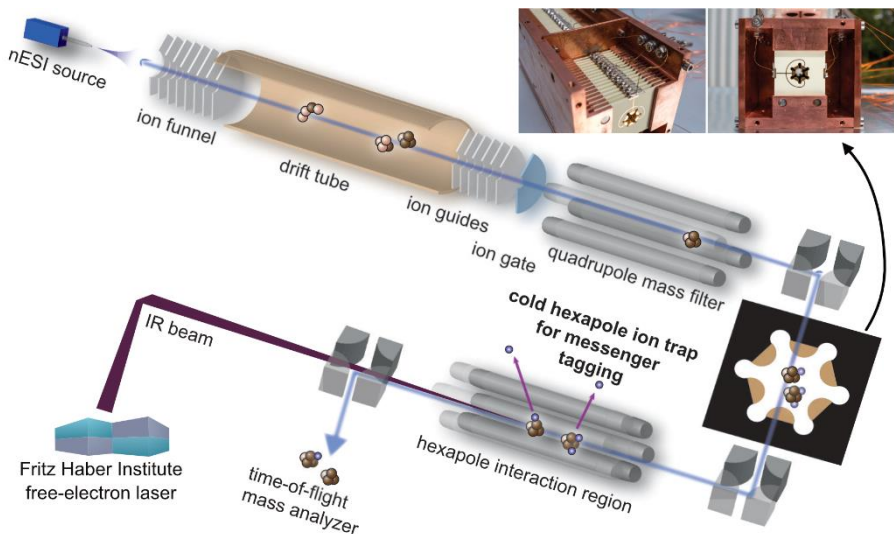


Figure 9: Scheme of the setup for messenger-tagging spectroscopy combined with DTIM-MS.

The front end of the instrumental setup is described in Section 2.1 and remains unchanged up to the quadrupole mass filter. In Figure 10, a more detailed technical drawing of the setup is depicted in top view. Furthermore, a detailed scheme of all DC and RF voltages is given in the Appendix. The former interaction region is now used to guide the ions to the first quadrupole bend.

For IMS experiments, the ions travel linearly and are detected using a high energy dynode (HED) electron multiplier (DM283, ETP ion detect, Adaptas, USA).

For messenger-tagging spectroscopy, the ions are perpendicularly deflected into the cryogenic hexapole ion trap. The novel ion trap contains two compartments: tagging and ejection and is described in further detail in the following section. For diagnostic reasons, a second electron multiplier (AF632, ETP Ion Detect, Adaptas, USA) is attached to the quadrupole bender to monitor the ions on-axis coming from the cryogenic ion trap. In the hexapole interaction region, the messenger-tagged ions overlap the IR beam from the FHI FEL for infrared photon dissociation (IRPD). The focus point of the IR beam is adapted using a pneumatically controlled adaptive mirror (model AO90/70, Kugler GmbH, Germany). Every second ion package is irradiated with IR light. In a commercial, high-resolution time-of-flight mass analyzer (Micromass, UK) the ions are detected off-axis to the interaction region. For data acquisition, a rapid analog-to-digital converter (PXIe 5160, National Instruments, USA) with 10-bit resolution, 500 MHz bandwidth and a sampling rate of up to 2.5 Gs/s, is employed to digitize the signal.

A LabVIEW controlled National Instruments platform provides several analog and digital in- and outputs to generate control voltages for e.g. data acquisition or the drift voltage power supply. An analog-to-digital converter from a multifunction module (PXIe 6363, National Instruments, USA) is used to digitize the ATD signal with 16-bit resolution, 2 MHz bandwidth and a sampling rate of 2 MS/s.

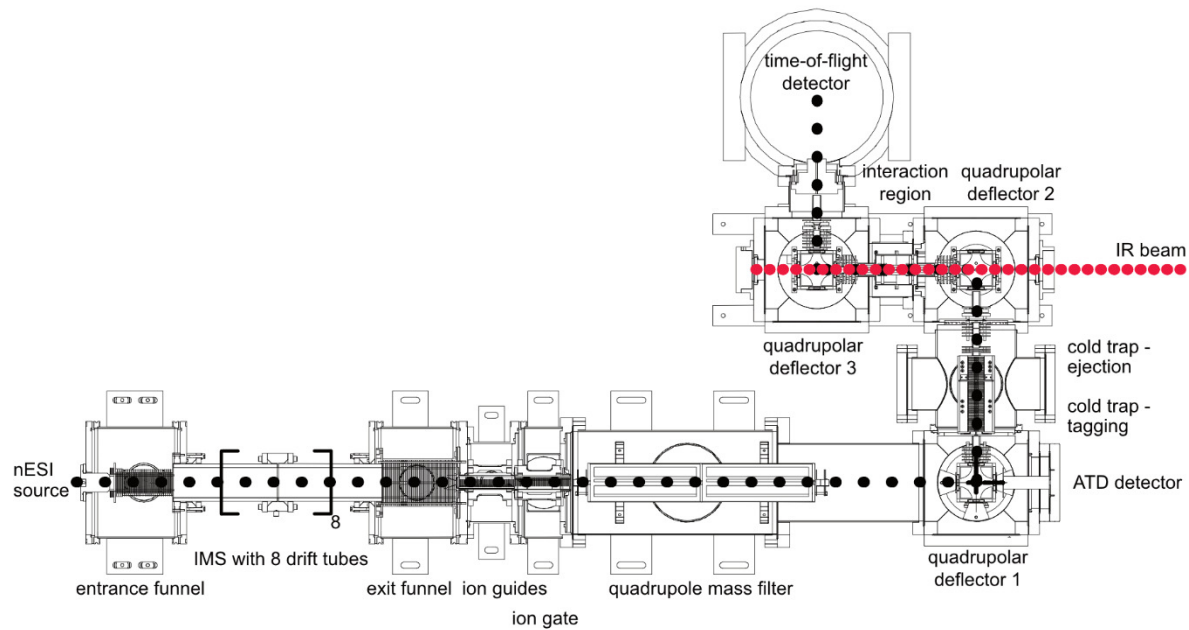


Figure 10: Simplified technical drawing of the instrumental setup for messenger-tagging spectroscopy combined with drift tube ion mobility-mass spectrometry

2.2.2. The Cryogenic Hexapole Ion Trap

The heart of the IR spectroscopy experiment is the new cryogenic ion trap that was assembled and tested in the framework of this thesis. As discussed in Section 1.4.2, several designs of ion traps for IR spectroscopy have been published. This novel ion trap is designed with two compartments to allow for an independent operational mode of the IMS part and IR spectroscopy part to improve in speed and stability. In the following section, the design, the mode of action and applied voltages are described.

To briefly explain the mode of action, the drift time- and m/z -selected ions from the IMS part enter the first compartment of the cryogenic ion trap, the tagging compartment. The ions are immediately cooled with buffer gas to the temperature of the copper housing. They lose an excess of kinetic energy and are efficiently trapped for up to several seconds. The buffer gas is a mixture of nitrogen in helium. Nitrogen freezes to the ions as messenger tag whereas helium stays gaseous as carrier gas. A package of messenger tagged ions is transferred to the ejection compartment by lowering the trapping potential of the ion gate. As soon as the ion gate closes, the ion package is ejected from the cryogenic ion trap and accelerated and steered towards the interaction region.

The 36 segments of the ion trap are made of printed circuit boards (PCBs) with an opening in the center in the shape of a hexapole, see schematic drawing in Figure 9. The PCBs are mounted in a copper housing. The copper housing has apertures at the front, the end and in the middle at the ion gate with diameters of 3 mm, 60 mm, and 6 mm, respectively. A turbomolecular pump is connected to the chamber so that the two compartments of the trap and the further flight path of the ions are differently pumped. Hence, cycles of pulsing and pumping of gas are not needed, as in most other designs, see Chapter 1. In the trapping compartment, a pressure of typically 10^{-3} mbar is achieved, whereas the ejection compartment is maintained at a pressure of approx. 10^{-5} mbar and the further flight path is at a pressure of approx. 10^{-6} mbar, see Figure 11. As a result, the tagged ions can be efficiently tagged, then gently ejected and accelerated, respectively, towards the interaction region.

DC potentials are connected to the front and rear of each tagging and ejection region with chains of resistors in between. Figure 11 describes the applied DC and RF potentials in the different stages of operation: trapping and tagging, transfer and ejection. In the tagging compartment, a shallow confining potential of 3 V is used to trap the ions. A small ramp rising from front to end of 1 V has experimentally shown to stabilize the ions during trapping. To spatially compress the ion package, the DC offset voltages across the ejection compartment are designed with a parabolic gradient. An RF voltage with a frequency of 1.32 MHz is applied to radially confine and stabilize the ions in both compartments.

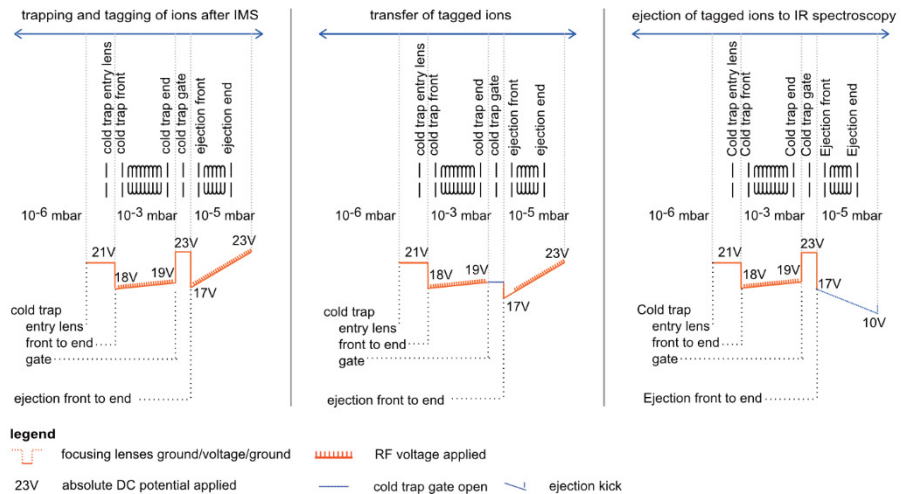


Figure 11: DC and RF voltages applied in the cryogenic ion trap in different operational stages: trapping and tagging, transfer and ejection. The pressures in the two compartments of the ion trap and in the chamber range from 10⁻³ to 10⁻⁶ mbar, of which the pressures inside the trap are an estimation.

The ion trap is cooled using a closed-cycle helium cryocooler (Sumitomo RDK-408D2, SHI, Japan) and a compressor unit (Sumitomo CSW-71D, SHI, Japan). The trap is directly mounted to the adapter stage or second stage of the cryocooler and electrically isolated with a thin foil of Kapton. Three precision diode temperature sensors are attached to the first stage of the cryocooler, the adapter stage and the bottom of the ion trap. A 25 W resistive heating element

controls the temperature from to 26 K at minimum to 150 K, which is ideal for nitrogen tagging experiments. A defined buffer gas mixture is prepared directly prior to the use. To avoid heating, the buffer gas is precooled at the interface of the cryogenic ion trap and the adapter stage before it flows into the tagging compartment.

The design of the cryogenic ion trap with its two compartments (I) tagging and (II) ejection is unique. The missing necessity of pumping the ion trap before IR spectroscopy accelerates the experiment on the orders of seconds in one duty cycle. Furthermore, the continuous flow of ions into the cryogenic ion trap allows a more stable operation of IR spectroscopy experiments and flexibility in timing of the IMS part. The IMS part can be operated up to repetition rates 10-times higher compared to a synchronous operational mode of IMS and spectroscopy.

2.2.3. Time-Sequence Generation

The timing of voltage switches that control ion optics such as ion gates is crucial for the operation of the instrumental setup. A detailed scheme of the voltages applied is given in Appendix. Two 8-channel delay generators (PXI 6602 and PXIe 6612, National Instruments, USA) were installed to generate trigger pulses which are controlled using LabVIEW.

In the front end of the instrument, the ions are injected to the drift tube at repetition rates up to 200 Hz with a length of typically 150 μs controlled by an internal clock from LabVIEW. A fraction of an ATD with lengths of 50 to 200 μs is selected, the ATD slice. The ATD slice is referenced to the IMS gate and several milliseconds later depending on the ions' mobilities. Figure 12 shows the pulse sequence of the IMS part and a representative drift time plot with ATD slicing of the monomer of human islet amyloid polypeptide (hIAPP) in charge state 4+ with two conformations of different mobilities.

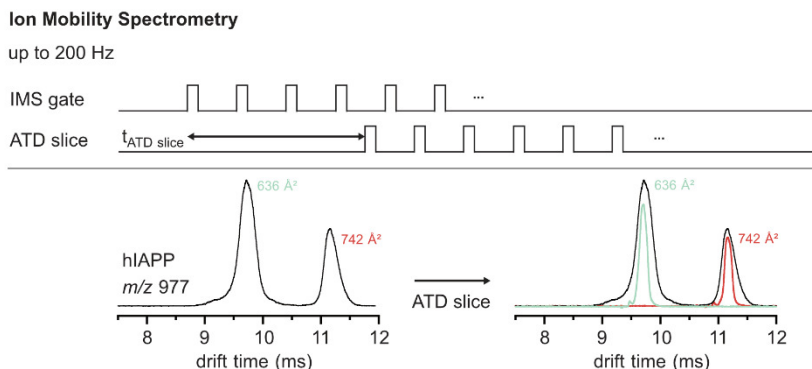


Figure 12: Scheme of trigger pulses for ion mobility spectrometry (IMS). The slicing of arrival time distributions (ATD) for further investigation is shown in a representative ATD plot of human islet amyloid polypeptide (hIAPP) as [hIAPP+4H]⁴⁺ with m/z 977. The cation has two conformational shapes with baseline-separated mobilities.

The trigger pulse from the FHI FEL with a frequency of 10 Hz is used to control the spectroscopy part and synchronize laser and ions. It is defined with a fixed lead of 700 μs prior to the arrival of the laser beam.

To overlap the ions and the laser in the interaction region, a series of timings are experimentally defined and depicted in Figure 13. The trigger pulse of the cryogenic trap gate, termed $t_{\text{ct gate}}$, to transfer the ions from tagging to ejection has a typical length of 400 μs , which can vary for different ions. The ions are ejected from the ejection compartment with a delay, termed $\Delta_{\text{ct gate-ejection}}$, after closing the gate. $\Delta_{\text{ct gate-ejection}}$ is experimentally determined 0 μs which indicates an instability of the ions in the ejection compartment. The ejection trigger pulse has a fixed length of 150 μs . The trigger pulse to control the pusher and puller of the time-of-flight mass analyzer has a fixed length of 150 μs . The time t_{TOF} is the time that the ions need from the ejection compartment to the time-of-flight mass analyzer which depends on their m/z -ratios. Due to the length of the ion path from the cryogenic ion trap to the time-of-flight mass analyzer, the spread of the ion cloud does not allow to detect ions with a m/z -difference of more than approx. 20 with the same efficiency. The ions overlap the light from the FEL at approx. half of t_{TOF} . This time is determined experimentally for each ion and termed t_{light} .

Messenger Tagging Spectroscopy

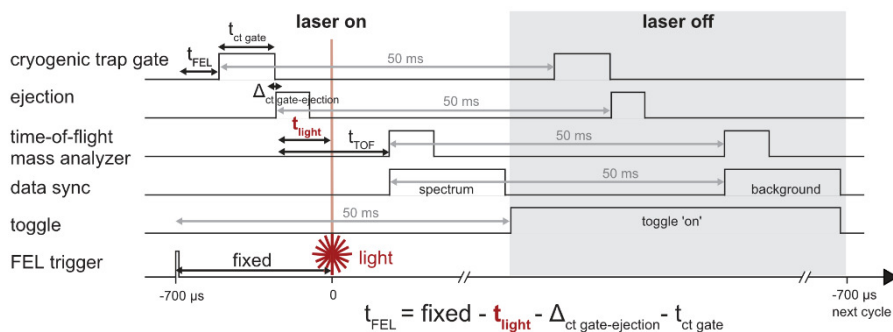


Figure 13: Scheme of trigger pulses for messenger-tagging spectroscopy. The trigger pulse of the Fritz Haber Institute free-electron laser (FHI FEL) is used to synchronize the ions and the laser beam in the instrumental setup.

The frequency for the ions is doubled to 20 Hz to record a background spectrum, i.e. without irradiation of the ions, laser off, in each cycle. Therefore, all trigger pulses, except the reference, induce a second train of pulses with a delay of 50 ms and the two trigger pulses are combined in a logic box before they reach certain voltage switches. Grey arrows are used to show the dependencies in time for the train of pulses. The state of the toggle trigger pulse is used to acquire the data on two independent channels of the analog-to-digital converter: the IR spectrum and the background spectrum.

2.2.4. Performance Characterization

The ideal tagging condition for a specific ion is mainly affected by three factors: temperature, gas mixture, and trapping time. The trapping time can be neglected in this instrumental setup because of the unique situation that ions constantly or at higher frequencies arrive in the tagging compartment. The temperature and gas mixture are determined for every ion.

The tagging behavior was evaluated for the protonated peptide leucine-enkephalin $[\text{LEK}+\text{H}]^+$ (m/z 556), the sodium adduct of the trisaccharide lewis x $[\text{Le}^x+\text{Na}]^+$ (m/z 552) and its protonated equivalent $[\text{Le}^x+\text{H}]^+$ (m/z 530). Interestingly, despite extensive tuning of the tagging conditions, the three ions showed different tagging behaviors. $[\text{LEK}+\text{H}]^+$ is predominantly present as

singly tagged ion, whereas $[\text{Le}^x+\text{Na}]^+$ shows efficient tagging for higher order or multiply tagging. $[\text{Le}^x+\text{H}]^+$ shows least tagging efficiency and is mostly present as untagged ion. The three cations are similar in their m/z -ratios, yet differ in their physical-chemical properties, most importantly their charge solvation and polarity. Considering the differences in tagging behavior as well as the limitations to detect ions of different m/z -ratio in the time-of-flight mass analyzer (see previous section), the measurement routine is adapted to three different cases. (1) For predominantly singly tagged ions, the intensity of the bare, untagged ion is recorded. (2) For predominantly multiply tagged ions, the depletion of the highest multiply tagged ion is recorded and (3) for predominantly non-tagged ions, the depletion of the fraction of singly tagged ions is recorded. The three cases are depicted in the Appendix.

In the first tagging spectroscopy experiments, the ion yields of the bare ion and tagged ions with irradiation N_{b-on} , N_{tag-on} and without irradiation N_{b-off} , $N_{tag-off}$ are recorded as a function of the laser's wavenumber in steps of 2 cm^{-1} . With the newly implemented control software in LabVIEW, the exact wavenumber is recorded directly with every scan step and the laser power is measured every 30 scan steps. The laser power is fit using a quartic polynomial function. To calculate the IR spectrum, Equation (4), Section 1.4.1, was rearranged, in a first order approximation, to:

$$\sigma(\nu) \propto -\ln\left(\frac{N_{tag-on}(\nu)}{N_{b-on}(\nu)+N_{tag-on}(\nu)}\right) \cdot \frac{1}{F(\nu)} \quad (6)$$

The number of unaffected ions n is defined as N_{tag-on} and the total number of ions is the sum of N_{b-on} and N_{tag-on} . To account for variations in laser fluence, the ion signal is divided by the laser power F . For predominantly multiply tagged ions and for predominantly non-tagged ions, tagging cases (2) and (3), the IR spectrum is obtained from:

$$\sigma(\nu) \propto -\ln\left(\frac{N_{tag-on}(\nu)}{N_{tag-off}(\nu)}\right) \cdot \frac{1}{F(\nu)} \quad (7)$$

Here, the ion yield $N_{tag-off}$ is the total number of ions. For each analyte, at least two independent scans are measured and averaged. The ion signal intensity is normalized to $[0,1]$.

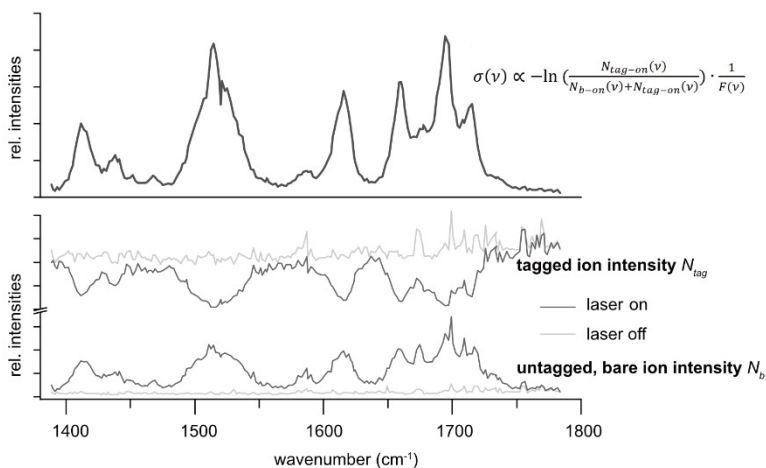


Figure 14: Upper panel: IRPD spectrum of leucine-enkephalin [LEK+H]⁺. Lower panel: Recorded ion yield for the tagged ion [LEK+H+N₂]⁺ and the bare ion [LEK+H]⁺ with irradiation (laser on, dark grey) and without irradiation (laser off, light grey).

Initial experiments were performed with [LEK+H]⁺ ($m/z \approx 556$). The peptide has been studied using different types of gas-phase IR spectroscopy techniques.^{154, 155} The IRPD spectrum with nitrogen tagging at 40 K with 10% nitrogen in helium is shown in Figure 14.

2.3. Infrared Spectroscopy in Helium Nanodroplets

The following section describes the home-built setup for IR spectroscopy in helium nanodroplets with the basic layout and operational mode. The setup is a user station at the Fritz Haber Institute free-electron laser (FHI FEL) which is also briefly introduced at the end of this chapter.

Figure 15 shows a scheme of the instrumental setup which consists of five parts: (A) the front end with the ionization source, transfer to high vacuum and mass filter, (B) the helium droplet source, (C) the variable temperature ion trap, (D) the interaction and detection region and (E) the IR light source.

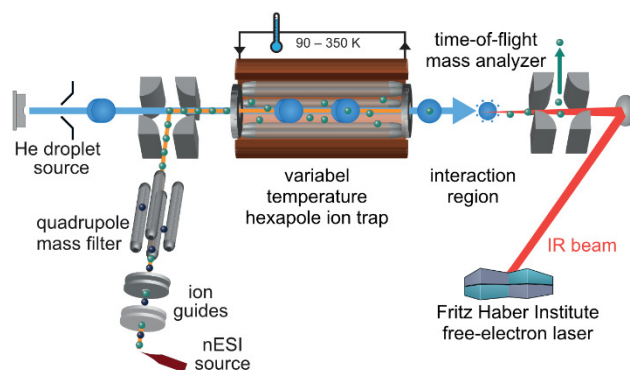


Figure 15: Scheme of the home-built setup for IR spectroscopy in helium nanodroplets which is attached to the Fritz Haber Institute free-electron laser.

The front end of the apparatus is a commercial and modified quadrupole time-of-flight mass spectrometer (Q-TOF Ultima, Waters Corporation, UK). The analyte is ionized using a nano-electrospray ionization (nESI) source by applying voltages to the nESI needle containing the analyte solution between -0.5 and -1.0 kV for anions and about 1 kV for cations. The Pd/Pt-coated needles are prepared in-house as described in Section 2.1. Molecular ions form from charged droplets and are doubly deflected into the apparatus in a Z-type pathway. The ions are transferred to high vacuum through two differential pumping stages using ring electrode ion guides. In the quadrupole mass filter, the ions are selected by their m/z ratios up to m/z ratios of 2000. In the first quadrupole bender, the ions travel linearly towards the first commercial, high-resolution time-of-flight mass analyzer (Micromass, UK) operated in W-mode, not depicted in Figure 15 and solely used for diagnostic reasons.

For IR spectroscopy, the ions are orthogonally deflected towards the variable temperature hexapole ion trap. The ion trap is filled with helium buffer gas for several seconds in which the ions undergo collisions to lose an excess of kinetic energy for efficient trapping and to thermalize with the housing of the trap. The ion trap has a length of 300 mm with an inner diameter of 9.1 mm. At the space-charge limit, the total number of ions inside the trap were calculated to be approx. 10^7 ions. The housing of the ion trap is made of copper and contains six internal cooling channels through which nitrogen gas flows, that is externally

cooled by liquid nitrogen. The housing is stepwise cooled to a maximum of 90 K and heated to 350 K using a thermocouple to measure the temperature and a heater to regulate the temperature on the copper housing.

A beam of helium nanodroplets traverse the ion trap to capture single ions immediately after the helium buffer gas is pumped from the trap. The nanodroplets of an average size of 10^5 helium atoms are produced by expanding high purity helium gas into high vacuum using an Even-Lavie valve (EL-C-C-2013, Uzi Even & Nachum Lavie, Israel) with a nozzle temperature of 21 K and an opening time of 17 μ s. The valve body is cooled using a closed-cycle helium cryocooler (RDK 408D2, Sumitomo Heavy Industries Ltd., Japan). The ion-doped helium nanodroplets have a high kinetic energy and easily overcome the shallow confining potential of the ion trap towards the interaction region of the instrument.

In the interaction region of the instrument, the ion-doped helium nanodroplets overlap the IR beam from the Fritz Haber Institute free-electron laser and at a resonant transition, ions are eventually released from the droplet. The mechanism of ion release from helium nanodroplets is discussed in Chapter 1. The focus point of the IR beam is shifted along the interaction region using a pneumatically controlled adaptive mirror (model AO90/70, Kugler GmbH, Germany). Therewith, the photon density, i.e. the number of photons per area unit, is controlled. The interaction region consists of a stacked ring electrode ion guide in which a spatial compression of the formerly large ion cloud is achieved by applying DC offset voltages across the ion guide with a parabolic gradient. In a second quadrupole bender, the ions are directed towards the off-axis detection region. The released ions are detected in a second, commercial, high-resolution time-of-flight mass analyzer (Micromass, UK) operated in V-mode. Early experiments in this thesis were performed using an on-axis home-built Wiley-McLaren-type of time-of-flight mass analyzer with a lower resolution.

An IR spectrum is recorded as the integrated ion intensity as a function of the laser's wavelength in steps of 2 cm^{-1} . The laser power is measured every 30 scan steps and fit using a quartic polynomial function. The ion signal is divided by the laser power as a first order approximation. The exact wavenumber of the

laser is recorded with every step. The ion signal intensity is normalized to [0,1]. When using different focus settings in certain wavenumber ranges, i.e. higher or lower photon flux per unit area, the different wavenumber ranges are normalized relative to each other on the basis of the intensity of spectral lines found in the overlapping measurement regions. For each analyte, at least two independent scans are measured and averaged.

2.4. The Fritz Haber Institute free-electron laser

All instrumental setups for IR spectroscopy are attached as user stations to the Fritz Haber Institute free-electron laser (FHI FEL). In free-electron lasers, photons are produced from relativistic electrons as gain medium passing through a periodic magnetic structure, the undulator, in high vacuum. The kinetic energy of the electrons and the periodic magnetic structure are tunable so that free-electron lasers are tunable over a wide range of wavelengths. The FHI FEL is schematically shown in Figure 16 and has been described in several publications. Regular user operation of the FHI FEL started in 2013 and has been continuously running since then.

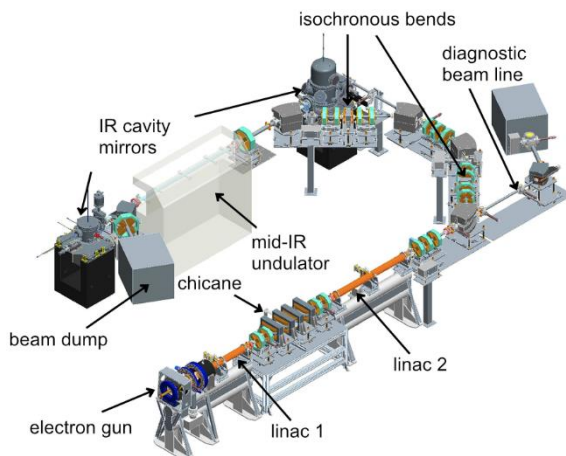


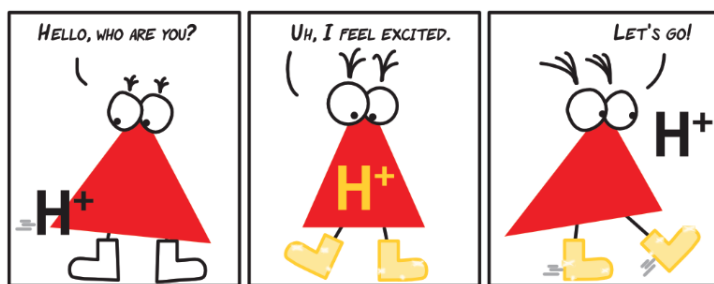
Figure 16: Schematic overview of the Fritz Haber Institute free-electron laser (FHI FEL). Figure adapted from literature.¹²⁰

An electron gun produces an electron beam at 1 GHz micro-pulse repetition rate. The electron bunches are compressed in a buncher cavity and subsequently accelerated by two consecutive linear accelerators (LINACs). The latter defines the kinetic energy of the electrons between 15 and 50 MeV and 36 MeV for the experiments in this thesis. A chicane between the two LINACs can be used to additionally compress the length of the electron beam down to picoseconds. The electron beam is bent into the laser cavity. The 2 m long mid-IR undulator with NdFeB permanent magnets inside the laser cavity forces the electrons to a sinusoidal transverse motion, called wiggling motion. The undulator gap, i.e. the distance between the permanent magnets, defines the strength of the magnetic field and therewith the electrons' trajectories through the undulator. With the wiggling motion of the first electrons, incoherent, monochromatic radiation is emitted. The wavelength of the radiation is determined by the undulator period, gap and electron energy. The radiation is reflected by the gold-plated laser cavity mirrors at the ends. The cavity mirrors are at a distance equal to a multipole of the generated wavelength. After reflection, the emitted light then overlaps the next incoming electron beam, so that a process called micro-bunching of the electrons is achieved which modulates the spatial distribution of the electrons along their propagation direction evenly. This has the effect that coherent radiation is emitted. After several roundtrips, the gain in radiation is strongly increased and the IR beam is outcoupled through a hole in one of the mirrors in the cavity. A shortening of the laser cavity leads to longer micro-pulses with smaller band width but also less power. The IR beam of the FHI FEL has a pulse structure of macro-pulses at 10 Hz with a length of 10 μ s consisting of micro-pulses at a repetition rate of 1 GHz with a length of a few picoseconds and a spectral width of 0.3 to 0.5% (FWHM) of the respective wavelength. Macro-pulse energies of 200 mJ are achieved at maximum.

Chapter 3

The Role of the Mobile Proton in Fucose Migration*

Fucose migration reactions represent a substantial challenge in the analysis of fucosylated glycan structures by mass spectrometry. In addition to the well-established observation of transposed fucose residues in glycan-dissociation product ions, recent experiments show that the rearrangement can also occur in intact glycan ions. These results suggest a low-energy barrier for migration of the fucose residue and broaden the relevance of fucose migration to include other types of mass spectrometry experiments, including ion mobility-mass spectrometry and ion spectroscopy. In this work, cold-ion infrared spectroscopy was utilized to provide further insight into glycan scrambling in intact glycan ions. Our results show that the mobility of the proton is a prerequisite for the migration reaction. For the prototypical fucosylated glycans Lewis x and blood group antigen H-2, the formation of adduct ions or the addition of functional groups with variable proton affinity yields significant differences in the infrared spectra. These changes correlate well with the promotion or inhibition of fucose migration through the presence or absence of a mobile proton.



* Adapted from: [Lettow, M.; Mucha, E.; Manz, C.; Thomas, D. A.; Marianski, M.; Meijer, G.; von Helden, G.; Pagel, K., The role of the mobile proton in fucose migration. *Anal. Bioanal. Chem.* **2019**, *411* \(19\), 4637-4645.](#)

3.1. Introduction

Glycans are among the most abundant biomacromolecules in living organisms and are essential in a variety of biological functions, such as cell-signaling and molecular recognition.^{2, 156} The glycosylation patterns of proteins, lipids, or other organic molecules are in many cases highly dynamic, which is obvious in an evolutionary sense,¹⁵⁷ but also accounts for physiological as well as pathological variations.¹⁵⁸ Fucose is a deoxy hexose found in a wide range of biologically relevant glycans and is easily distinguishable from many other hexoses due to the lack of the hydroxy group at the C-6 position. In mammalian glycans, fucose is the fifth most abundant monosaccharide (7.2%), and nearly every fourth terminal end (23.8%) carries a fucose.²¹ Among human milk oligosaccharides, up to 80% are fucosylated, depending on the individual's secretor status.¹⁹ Glycans attached to proteins are usually fucosylated in the Golgi apparatus or in the endoplasmic reticulum during the final step of glycan processing by different fucosyltransferases.²³ Altered fucosylation can account for pathological variations, for example, the overexpression of fucose in cancer cells. The fucosylated Lewis y (Le^y) antigen is highly expressed on a wide range of tumors, and tumor growth can be inhibited by suppressing Le^y expression.¹⁵⁸⁻¹⁶⁰

The inherent structural diversity of glycans and the limited availability of standards have made glycan structural characterization a major analytical challenge. Tandem mass spectrometry (MS) is a powerful tool in the analysis of glycan structures^{33, 161-163} and is often combined with chromatography¹⁶⁴ and/or ion mobility spectrometry.^{97, 109} In general, tandem MS can provide the mass of the intact ion and, additionally, the masses of its fragment ions in a single experiment. Different types of fragmentation techniques yield distinct types of fragment ions that provide information on the original glycan sequence. Information-rich fragment spectra contain glycosidic fragments and/or cross-ring fragments, but can also show rearrangement products and internal fragments. An effective application of MS in glycan analysis therefore requires comprehensive knowledge of potential gas-phase reactions and fragmentation pathways.

Suitable techniques to shed light on the mechanism of a gas-phase reaction in MS are infrared (IR) spectroscopy,^{80, 155} ion mobility-mass spectrometry (IM-MS),^{78, 165} hydrogen-deuterium exchange (HDX),⁷⁹ stable isotope labeling¹⁶⁶ or computational modeling.⁷⁷ A very prominent rearrangement reaction in biomolecules that has been investigated extensively in the last years is peptide scrambling. Peptide scrambling in protonated or multiply protonated b-type peptide ions involves a head-to-tail macrocyclization and subsequent reopening reaction prior to fragmentation in tandem MS with collision-induced dissociation (CID). Consequently, the original sequence information is lost during scrambling, which may result in erroneous assignments. The phenomenon was first reported in 1997.⁷⁶ Since the beginning of the 2000s, the underlying mechanism and the structure of the macrocyclic intermediate have been successfully proven using mainly a combination of IR spectroscopy and computational modeling.¹⁶⁷ It has been found that the incidence of this gas-phase rearrangement is negligible in shotgun proteomics but of major interest in understanding the fragmentation pathways in CID.¹⁶⁸⁻¹⁷⁰

For more than 20 years, researchers have also been studying glycan rearrangement reactions in tandem MS experiments.^{83, 86} Most of these studies are based on the observation of unusual fragment masses that arise from a rearrangement of fucose and occasionally other monosaccharide units such as xylose⁸² during the CID process. This phenomenon is also often referred to as internal residue loss (IRL) and has been studied extensively by probing the influence of different adduct ions^{171, 172} or an aglycon⁸⁴ and different derivatization strategies.^{87, 88, 173} Based on these findings, various mechanisms have been proposed,^{84, 86} all of which involve a mobile proton that is able to migrate to the reactive site. However, despite these efforts, the underlying mechanism is not fully understood to date.

Recently, a glycan sequence scrambling reaction in the intact trisaccharide ions Lewis x (Le^x) and blood group antigen H-2 (BG-H2), which are substructures of the tetrasaccharide Lewis y (Le^y), was reported.¹⁷⁴ The observed fucose migration involves the cleavage of a glycosidic bond and the subsequent or concerted formation of a new glycosidic bond (Figure 17). The reaction takes place in the presence of a proton and is inhibited in its absence, namely in the

sodium adducts. Interestingly, this reaction was found to occur under soft conditions without substantial collisional activation, which indicates a rather low-energy barrier for the rearrangement. Based on these results, the conclusion was that fucose residues could spontaneously migrate from the terminal end of a glycan chain to adjacent or remote intramolecular sites of the glycan without CID. The previously reported occurrence of unexpected fragment masses in IRL is therefore likely only a result, but not the cause of fucose migration.

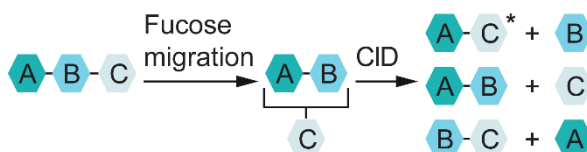


Figure 17: Glycan rearrangement reactions consist of two independent, consecutive steps: fucose migration and then collision-induced dissociation (CID) with internal residue loss (IRL) products and glycosidic fragments. The energy barrier for a fucose migration is low and the product of a fucose migration reaction is stable. The IRL fragment is labeled with an asterisk. A, B, and C represent monosaccharides. Glycosidic linkages are indicated with bars.

Here, we build on the previous findings and investigate the role of the mobile proton in fucose migration in the intact ion pair BG-H2 and Le^x using cryogenic IR spectroscopy. In order to increase or diminish the mobility of protons in the investigated ions, functional groups and adduct ions with competitive proton affinity have been utilized. Our results clearly show that a mobile proton at a specific site of the molecule is required for fucose migration.

3.2. Experimental Details

Sample preparation

Lewis and blood group antigens were purchased from Dextra Laboratories (UK) and Biozol (Germany) and used without further purification. Solvents (HPLC grade) and labeling reagents were purchased from Sigma-Aldrich (USA) and used without further purification. Aqueous glycan stock solutions (1 mM and 100 μ M for labeled glycans) were further diluted prior use with water/methanol (*v/v*, 50/50) to yield 20–500 μ M analyte solutions. To promote the adduct formation in the measurements of $[M+NH_4]^+$, $[M+NMe_3H]^+$ and $[M+NEt_3H]^+$, ammonium acetate, trimethylammonium chloride, and triethylammonium acetate, respectively, were added in concentrations of 1–5 mM.

Synthesis of labeled glycans

For glycan labeling, the Lewis and blood group antigens were dissolved in water to yield 100 mM stock solutions. The stock solutions were divided into 10 μ L (1 μ mol) aliquots and freeze-dried. The glycans were labeled with 2-aminobenzoic acid (AA) and 4-amino-N-(2-diethylaminoethyl) benzamide (procainamide) via reductive amination using standard protocols.¹⁷⁵ Removal of the excess label was performed using paper chromatography.¹⁷⁶ The labeled glycans were further purified using HyperSep Hypercarb SPE cartridges (ThermoFisher Scientific, USA) according to manufacturer's instructions. The labeled glycans were then freeze-dried and dissolved to aqueous 100 μ M stock solutions.

Cryogenic IR spectroscopy

The experimental setup is comprehensively described in Section 2.3.

3.3. Results and Discussion

3.3.1. Ammonium adduct ions

First, trisaccharides of the Lewis y-series (Le^y -series) were investigated as ammonium adducts $[M+NH_4]^+$ (m/z 547). The Le^y -series consists of the tetrasaccharide Le^y and the two trisaccharides BG-H2 and Le^x that are substructures of Le^y (Figure 18).

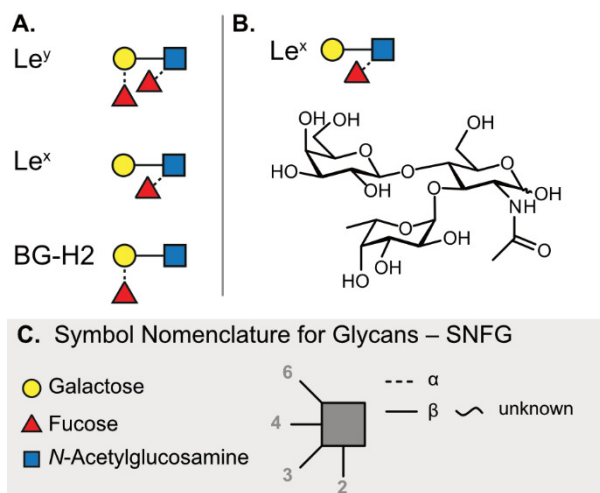


Figure 18: A. The investigated Fucosylated glycans. The trisaccharide Le^x and BG-H2 are substructures of the tetrasaccharide Le^y . B. The trisaccharide Le^x is exemplarily translated into its skeletal structure. C. The symbol nomenclature for glycans (SNFG)²² offers simplified representations. Symbols are shown for the relevant monosaccharides in this chapter.

The IR spectra of the trisaccharide standards BG-H2 and Le^x are shown in Figure 19 and were recorded using the instrumental setup described above. The instrument settings were tuned to minimize activation of the parent ion. As expected, both IR spectra show well-resolved features in the range of 1000 cm^{-1} to 1700 cm^{-1} . Generally, vibrations in glycans around 1100 cm^{-1} are associated with numerous C–O and coupled C–O–C–O-stretch vibrations of the (hemi-)acetal. Features above 1500 cm^{-1} and around 1650 cm^{-1} canonically stem from the amide II (N–H bend) and amide I (C=O stretch) vibrations in N -acetylglucosamine monosaccharides. Weaker vibrations between 1200 and

1500 cm^{-1} are O–H-bending modes. With a charged site at the amide group, the amide vibrational modes can be strongly shifted, e.g. with a proton at the oxygen, the amide I vibration is likely red-shifted. A confident assignment and differentiation of the amide I and II features can therefore not be made.

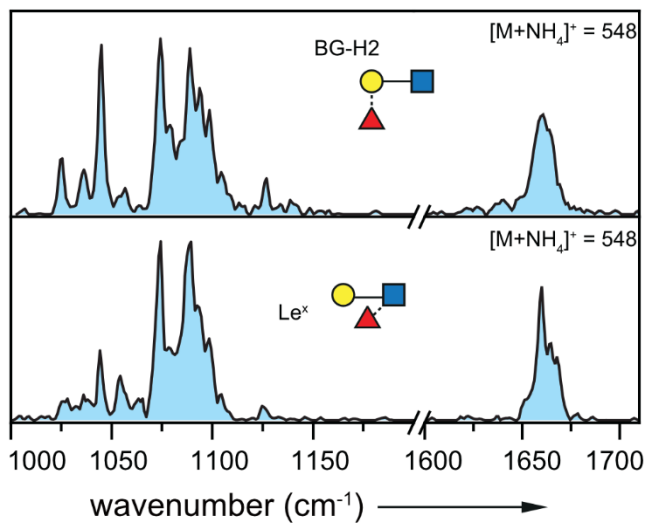


Figure 19: IR spectra of the parent ions BG-H2 (upper panel) and Le^x (lower panel) investigated as $[\text{M}+\text{NH}_4]^+$ species ($m/z \approx 547$) in the range of 1000–1200 cm^{-1} and 1600–1710 cm^{-1} . The two IR spectra are nearly identical.

The IR spectrum of BG-H2 as $[\text{M}+\text{NH}_4]^+$ (Figure 19, upper panel) shows nine well-resolved bands in the range from 1000 to 1200 cm^{-1} . The strongest absorptions are found at 1045, 1075 and 1089 cm^{-1} . Between 1200 and 1600 cm^{-1} , weaker absorption bands, which are not resolvable with the utilized instrumental setup, are observed. A broad absorption is found in the higher wavenumber range of the spectrum around 1660 cm^{-1} . A comparison of the IR spectrum of BG-H2 with that of Le^x as $[\text{M}+\text{NH}_4]^+$ ion (Figure 19, lower panel) reveals many similarities. The main absorptions in the IR spectrum of Le^x are found between 1000 and 1200 cm^{-1} and around 1660 cm^{-1} . In the lower wavenumber range, the strongest absorptions are at 1075 and 1089 cm^{-1} and in the higher wavenumber range, a broad absorption is found around 1660 cm^{-1} .

In general, all features in the IR spectrum of Le^x are also present in the spectrum of BG-H2, however, with varying intensity, which likely results from differing ratios of alpha and beta anomers in the gas phase. The nearly identical IR spectra of the intact ammonium adduct ions therefore imply that the trisaccharides BG-H2 and Le^x rearrange to the same chemical structure, not only as protonated ions as observed previously¹⁷⁴ but also as ammonium adduct ions. This is in good agreement with previously reported CID experiments in which IRL reactions have been found for ammonium adducts.¹⁷²

Observing fucose migration in protonated ions and ammonium adducts at almost identical instrument conditions indicates a comparable energy barrier, migration mechanism and possibly even structure in the two types of ions. In order to test this hypothesis, the optical signatures of protonated and ammonium adduct ions of the Ley series are compared. Remarkably, very similar IR signatures are observed for the isomers BG-H2 and Le^x , regardless whether they are present as $[\text{M}+\text{NH}_4]^+$ or $[\text{M}+\text{H}]^+$ species. The stacked overlay in Figure 20 reveals that the main features are qualitatively present in all four IR spectra (indicated by vertical dashed lines). This is a rather surprising observation, as different adduct ions typically change the gas-phase conformation of the molecule substantially and therefore lead to distinct IR traces. A possible explanation for the striking similarities observed here is that the present species are in fact not strict $[\text{M}+\text{NH}_4]^+$ adducts, but rather neutral ammonia adducts of a protonated glycan $[\text{NH}_3+\text{MH}]^+$. The proton directly transfers to the glycan while the ammonia is coordinated as the neutral species, which does not seem to change the gas-phase conformation of the glycan. With only weak interactions of the neutral ammonia molecule, the observed absorptions largely correspond to those of the protonated glycan ion, which in turn is able to undergo fucose migration.

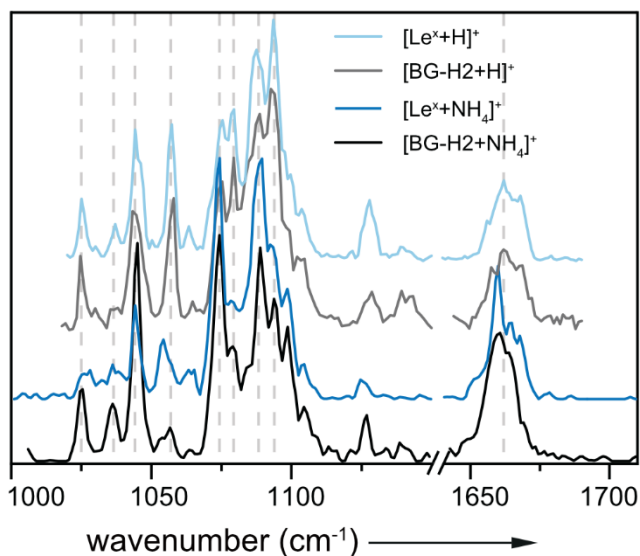


Figure 20: Stacked overlay of the IR spectra of the parent ions BG-H2 and Le^x investigated as $[\text{M}+\text{H}]^+$ and $[\text{M}+\text{NH}_4]^+$ species in the range of 1000–1710 cm^{-1} (break between 1150 and 1640 cm^{-1}). The spectra of the protonated species are reproduced from Mucha et al.¹⁷⁴ and shown in the range of 1020–1690 cm^{-1} . The four IR spectra are highly similar.

Ammonia loss can be an indicator for the $[\text{NH}_3+\text{MH}]^+$ species. Recorded as $[\text{M}+\text{H}]^+$ ions, the ammonia loss can either be produced by varying the cone voltage or also after ejection from the helium nanodroplet. The latter has been recorded over all scans. A fraction of $[\text{M}+\text{H}]^+$ ions are detectable in each scan indicating loosely bound ammonia. It should be pointed out that the $[\text{M}+\text{H}]^+$ ion signal is recorded with each scan but it is not part of the resulting IR spectrum of the $[\text{M}+\text{NH}_4]^+$ species.

Gas-phase proton transfer is one of the most commonly observed processes in ion-molecule chemistry and studies have provided a substantial database of thermochemical parameters. For singly charged species, the protonation propensity is predominantly governed by the gas-phase proton affinity (PA) of the present functional groups. The gas-phase basicity (GB) and the PA are directly derived from the reaction of a neutral molecule with a proton to form the protonated species, with the GB corresponding to the Gibbs free energy change and the PA as the respective enthalpy change.¹⁷⁷ In the trisaccharides

BG-H2 and Le^x , the proton is likely located at the C=O of the amide (PA 230 kcal mol⁻¹ for C=O compared to 220 kcal mol⁻¹ for the N–H group).¹⁷⁸ Even though the actual proton affinity can deviate depending on the local environment, the PA of free ammonia is considerably lower (204 kcal mol⁻¹)¹⁷⁷ which supports the formation of a $[\text{NH}_3+\text{MH}]^+$ -type ion and the presence of a mobile proton.

3.3.2. Alkylammonium adduct ions

To further test the impact of proton mobility, the trisaccharide Le^x was investigated with two alkylammonium adduct ions as $[\text{Le}^x+\text{NMe}_3\text{H}]^+$ and $[\text{Le}^x+\text{NEt}_3\text{H}]^+$ (Figure 21).

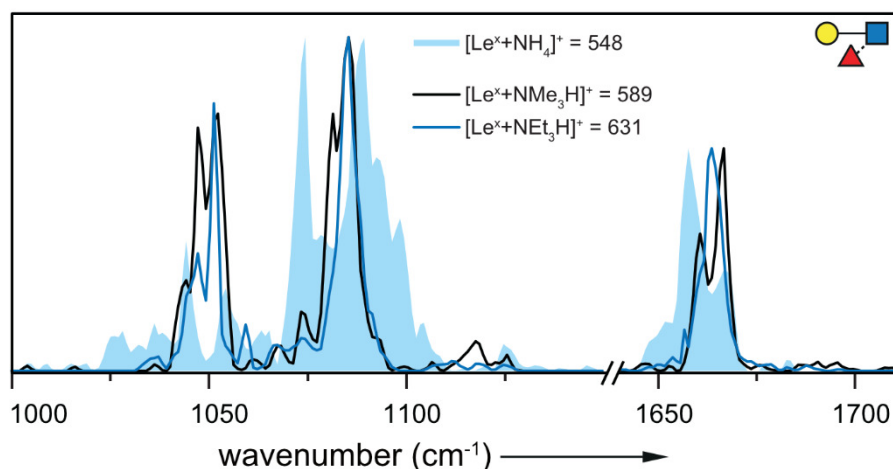


Figure 21: Overlay of the IR spectra of $[\text{Le}^x+\text{NMe}_3\text{H}]^+$ (black line), $[\text{Le}^x+\text{NEt}_3\text{H}]^+$ (blue line) and $[\text{Le}^x+\text{NH}_4]^+$ (blue filled) in the range of 1000–1710 cm⁻¹ (break between 1150 and 1640 cm⁻¹). The spectra of the alkylammonium adducts are very similar and differ from that of the ammonium adduct of Le^x .

The spectra of BG-H2 with different alkylammonium adduct ions could not be recorded as a result of unstable electrospray conditions. As shown in Figure 22, trimethylammonia (PA 227 kcal mol⁻¹) and triethylammonia (PA 235 kcal mol⁻¹)¹⁷⁷ have substantially higher gas-phase PAs than ammonia. A proton transfer from the alkylammonium ions to the glycan as observed for ammonium ions is therefore energetically unlikely. The spectra of $[\text{Le}^x+\text{NMe}_3\text{H}]^+$ and

$[\text{Le}^x + \text{NEt}_3\text{H}]^+$ shown in Figure 21 confirm this hypothesis. The two traces are almost identical, but differ considerably from the spectrum of $[\text{Le}^x + \text{NH}_4]^+$. The spectra of the alkylammonium adducts show two well-resolved bands at 1050 cm^{-1} and 1085 cm^{-1} , of which the first is a split band. In the higher wavenumber range, a peak is found, which in comparison to the ammonium adduct is slightly blue-shifted. The only noticeable difference in the two alkylammonium adduct spectra are a side peak at 1081 cm^{-1} in the spectrum of $[\text{Le}^x + \text{NMe}_3\text{H}]^+$ and a subtle shift of the feature at higher wavenumbers. However, the tremendous difference to the spectrum of the ammonium adduct is apparent. Alkylammonium adduct ions demobilize the proton or at least offer two possible protonation sites. Therefore, the gas-phase structures of the alkylammonium adducts of Le^x are rather similar while those of the ammonium adduct closely resembles that of the protonated species.

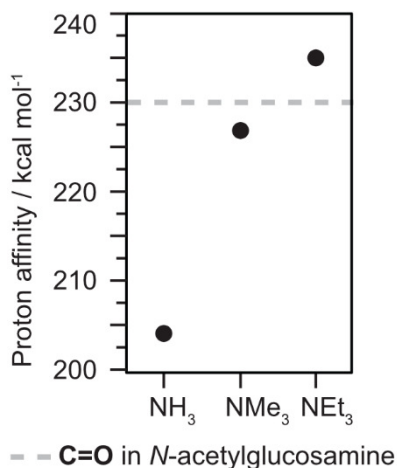


Figure 22: Proton affinity of different adduct ions in kcal mol^{-1} compared to the proton affinity of the C=O group in a representative *N*-acetylglucosamine (gray line).^{177, 178} Due to the local environment, the PA of the C=O group in Le^x can be slightly different.

3.3.3. Fluorescently labeled glycan ions

Glycans are challenging to study with common analytical techniques, especially at low analyte concentrations. A modification of the sugars with fluorescent labels facilitates their analysis and quantification by liquid chromatography

using fluorescence detection. Since glycans themselves do not fluoresce, they are typically modified with a fluorophore at their reducing end by reductive amination. Popular chromophores in LC are 2-aminobenzamide (2-AB), anthranilic acid (AA) or 4-amino-*N*-(2-diethylaminoethyl) benzamide (procainamide).^{40, 179} Labeling often goes in-line with improved protonation efficiency due to the presence of one or more apparent basic sites at the label. As a result, fluorescence labeling is also beneficial for hyphenated LC-MS techniques.⁵⁶ However, to date relatively little is known about the impact of fluorescence labeling on the previously described fucose migration reaction. In order to evaluate this, the trisaccharides of the Le^x-series were modified with the common AA (m/z 651) and procainamide-labels (m/z 749) (Figure 23).

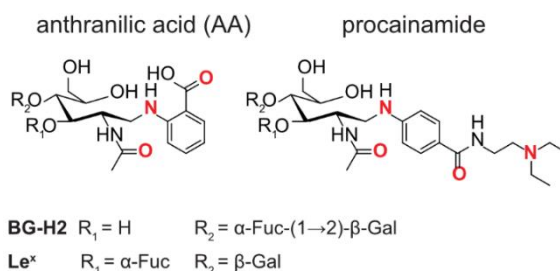


Figure 23: Chemical structures of the investigated, labeled trisaccharides. Bold red atoms indicate possible protonation sites.

IR spectra of the protonated ions $[M+H]^+$ were recorded in the range from 900 to 1800 cm^{-1} and 1000 to 1800 cm^{-1} , respectively, and the obtained traces are shown in Figure 24. In the range from 1000 to 1200 cm^{-1} , the dominant vibrations arise from the typical C–O and coupled C–O–C–O-stretch vibrations of the (hemi-)acetal. In stark contrast to the spectra of the unmodified trisaccharides, all four spectra show a large number of vibrations in the range from 1250 to 1500 cm^{-1} , which are predominantly associated with functional groups of the labels. In the last part of the spectra, from 1500 to 1800 cm^{-1} , the vibrations derive either from the amide vibrations in the *N*-acetylglucosamine or from the label.

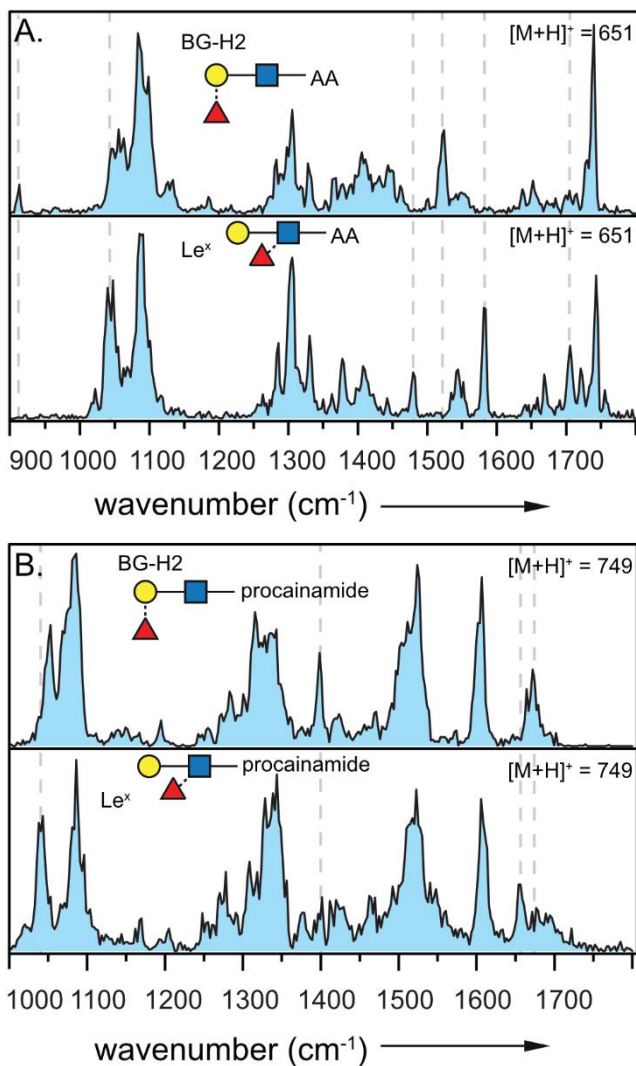


Figure 24: A. IR spectra of the parent ions Le^x-AA (upper panel) and BG-H2-AA (lower panel) investigated as [M+H]⁺ species (*m/z* 651) in the range of 900–1800 cm⁻¹. The two IR spectra are distinguishable from each other, and the gray dashed lines indicate major differences. B. IR spectra of Le^x-procainamide (upper panel) and BG-H2-procainamide (lower panel) investigated as [M+H]⁺ species (*m/z* 749) in the range of 1000–1800 cm⁻¹. For both isomeric sets, the two IR spectra are clearly distinguishable from each other, and the gray dashed lines indicate major differences.

In AA-labeled glycans (Figure 24), vibrations above 1700 cm^{-1} stem from the C=O stretch in the carboxylic acid at the aromatic ring. The IR spectra of the AA-labeled glycans show well-resolved bands but also broad features. The latter may result from a larger set of accessed low-energy conformers and/or protonation isomers (protomers).¹⁸⁰ In the IR spectrum of BG-H2-AA, the main features are found at 1087, 1305, 1523, and 1740 cm^{-1} . In the IR spectrum of the isomeric trisaccharide Le^x-AA, the main absorptions are at 1044 and 1087 cm^{-1} in the lower wavenumber range and at 1305 and 1582 cm^{-1} and at 1705 and 1743 cm^{-1} in the higher wavenumber range. The IR fingerprints of the two isomeric trisaccharides differ. Especially above 1450 cm^{-1} , the vibration bands are diagnostic. The main differences in the IR spectra are marked with vertical gray dashed lines.

In procainamide-labeled glycans (Figure 24), a second amide group and a tertiary amine contribute to the IR spectrum. The spectra of BG-H2-procainamide and Le^x-procainamide show less well-resolved features than the spectra of the AA-labeled derivatives. In both spectra, strong amide vibrations are found from 1500 to 1700 cm^{-1} . The spectral signatures or fingerprints of the two isomeric trisaccharides are clearly distinguishable from each other.

Based on the observed spectra, both labels—anthranilic acid and procainamide—seem to alter the proton mobility in the Le^y series, which leads to an inhibition of fucose migration. Based on the data, it is not possible to exclude a rearrangement reaction; however, the observed chemical structures are clearly different.

For procainamide, these results are consistent with previous CID experiments, which show that this label reduces or inhibits IRL reactions in fucosylated glycans.¹⁸¹ Glycans labeled with anthranilic acid, on the other hand, have been reported to produce unknown mass-to-charge signals from IRL reactions in tandem MS experiments.⁸⁵ A possible explanation for this different behavior might be the competition between protonation sites. In AA-labeled glycans, the proton has at least two possible sites with comparable affinity: the amide group on the *N*-acetylglucosamine and the secondary, aromatic amine. The former would facilitate a migration reaction, while the latter would inhibit it. In the procainamide label, the basic tail with a tertiary amine probably has the highest

proton affinity while the secondary, aromatic amine and the aromatic amide have compatible affinities to the *N*-acetylglucosamine. It is therefore conceivable that in both cases more than one protomer is populated at room temperature. A migration, however, is—if at all—only conceivable in case of the AA label. It can therefore be concluded that an additional functional group with comparable or higher proton affinity inhibits or at least alters the fucose migration reaction in the isomeric trisaccharides of the *Le^x*-series.

3.4. Conclusion

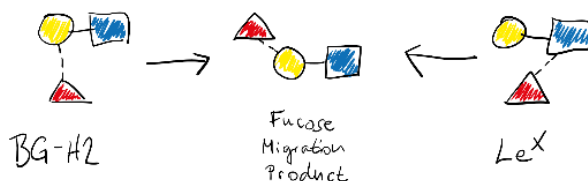
Cryogenic IR spectroscopy was used to study the impact of adduct formation and proton mobility on fucose migration reactions in the glycan isomers BG-H2 and *Le^x*. In ammonium adducts, the proton is located at the glycan in the form of a $[\text{NH}_3+\text{MH}]^+$ -type ion of the individual isomers. These ions are in many ways similar to $[\text{MH}]^+$ ions and show fucose migration to the same, yet not fully unraveled, chemical structure. In the alkylammonium adducts of *Le^x*, on the other hand, the proton is immobilized at the adduct, which leads to a distinct glycan structure in which fucose migration is likely not occurring. A similar behavior was observed when different functional groups with competitive proton affinity are introduced. In these structures, fucose migration to one chemical structure is inhibited or at least altered, since multiple competing protonation sites in the glycan reduce proton mobility. These results agree well with our previous results in which fucose migration is inhibited in the sodium adducts.

In summary, our results suggest that a proton needs to be localized on the trisaccharide and presumably in close proximity to the rearranging monosaccharide for the fucose migration reaction to take place. This observation validates what has already been found in work on IRL. The proton mobility is influenced in the presence of adduct ions or functional groups with competitive proton affinity.

Chapter 4

Decoding the Fucose Migration Product in Blood Group Epitopes*

Fucose, a signaling carbohydrate that is typically attached in the final step of glycan processing, is involved in a range of processes such as the regulation of selectin-dependent leukocyte adhesion or pathogen-receptor interactions. Mass-spectrometric techniques, commonly used to determine the structure of glycans, frequently show fucose-containing chimeric fragments that obfuscate the analysis. The rearrangement reaction leading to these fragments—often referred to as fucose migration—has been known for more than 25 years, but the chemical identity of the rearrangement products and the mechanism remain unknown. In this work, we combine ion-mobility spectrometry, radical-directed dissociation mass spectrometry, and cryogenic-ion IR spectroscopy with high-throughput density functional theory calculations to deduce the structure of the rearrangement products in the model trisaccharides Lewis x and blood group H type 2. The structurally unconstrained search yields the fucose moiety attached to the galactose with an α 1,6 glycosidic bond as the most thermodynamically stable ion. The computed CCSs and IR spectra of the model agree well with the experimental values, which further supports the α 1,6-linked structure as the rearrangement product. Finally, we discuss two possible rearrangement mechanisms which would lead to the formation of such a product.



* Based on an early version of: [Lettow, M.](#); Mucha, E.; Lambeth, T. R.; Greis, K.; Yaman, M.; Manz, C.; Hoffmann, W.; Meijer, G.; Julian, R. R.; von Helden, G.; Marianski, M.; Pagel K., Decoding the Fucose Migration Product in Blood Group Epitopes. *Manuscript in preparation*.

4.1. Introduction

Glycans are responsible for a range of key-biological functions, such as nutrient storage, molecular signaling, and cell-cell communications. The widespread use of glycans is enabled by their large structural heterogeneity, owed to the non-trivial assembly rules of isomeric monosaccharides, and their large conformational flexibility. This structural complexity poses a challenge for analytical methods aimed at determining glycan molecular structure. Colluded with a lack of a simple amplification mechanism, this renders glycomics lagging behind genomics and proteomics.¹⁸² Today, the structural analysis of glycans is typically performed by sophisticated tandem mass spectrometry (MS) techniques. The tandem MS spectrum of a glycan usually contains glycosidic and cross-ring fragments, which can be used to reconstruct the topology of the precursor.³³ The assignment, however, is often hampered by the appearance of non-native chimeric fragments, which are caused by an internal rearrangement of particular monosaccharide building blocks. Fucose, a C6-deoxy L-monosaccharide involved in the extracellular communication, such as the regulation of selectin-dependent leukocyte adhesion,¹⁸³ and the viral docking to the host cell,¹⁸⁴ is particularly prone to migration along the glycan sequence.^{88, 172, 185-187} This phenomenon has been known for more than 25 year, however, the mechanism as well as structural triggers of the rearrangement remain unknown.^{81, 82, 174} Finally, the migration is not limited to fucose, but has also been observed for other biologically relevant saccharides such as rhamnose,⁸³ mannose,⁸⁵ and recently, xylose.⁸²

In recent years, the capabilities of MS for glycan analysis have been significantly enhanced by the introduction of orthogonal gas-phase techniques. Ion mobility-mass spectrometry (IM-MS) adds a structural dimension to the mass-to-charge ratio in form of a rotationally-averaged collision cross section^{109, 111} and action infrared spectroscopy probes the spatial structure of the ion in form of an IR spectrum.^{126, 127, 188} In the past, we have applied both techniques to identify the glycosidic bond regio- and stereochemistry in glycans¹²⁷ and glycoconjugates,³⁵ and unraveled the structure of reaction intermediates in glycosylation reactions.^{150, 152, 189} Radical-mediated dissection of glycans in tandem MS has shown to be structure sensitive to discriminate isomeric glycans.^{190, 191} In radical-

directed dissociation (RDD) MS, a radical, generated upon UV photodissociation, migrates along the glycan core to produce diverse dissociation products, such as cross-ring cleavages for cations, increasing its sensitivity compared to collision-induced dissociation.¹⁹²

Recently, we elucidated the migration of fucose in the Lewis x (Le^x) and blood group antigen H, type 2 (BG-H2) trisaccharides. They share a D-galactopyranose β 1,4D-N-acetylglucosaminepyranose ($\text{Gal}\beta$ 1,4GlcNAc) core, but differ by the attachment of the fucosyl moiety (Figure 25). In Le^x , the fucose is attached via an α 1,3 glycosidic bond to the reducing GlcNAc, whereas in BG-H2 the fucose is linked by an α 1,2 bond to the galactose. Cryogenic IR spectroscopy revealed identical mid-IR fingerprints for the protonated $[\text{Le}^x+\text{H}]^+$ and $[\text{BG-H2}+\text{H}]^+$ trisaccharide ions, which indicates that fucose migration can occur at ambient conditions and does not require collision-induced dissociation.¹⁷⁴ In a subsequent work, we further showed that the rearrangement is promoted by a mobile proton, whereas lithium, sodium and adducts with an immobilized proton inhibit migration, see Chapter 3. However, the final structure of the rearrangement product and the triggers that initiate migration remain elusive to date. In this work, we combine structural data from IM-MS, cryogenic IR spectroscopy and RDD with high-throughput density functional theory (DFT) modeling to elucidate the molecular structure of the rearrangement product and discuss possible mechanisms leading to its formation.

4.2. Experimental Details

Mass spectrometry-based experiments

The three utilized instruments have been described in great detail in previous publications,¹⁹²⁻¹⁹⁴ and therefore only relevant details are discussed here. In the utilized IM-MS experiment, molecular ions are generated by nano-electrospray ionization and transferred into the instrument via a capillary inlet. Before being injected to an 80.55 cm long drift tube for mobility separation, the ions are accumulated and stored in an hourglass ion funnel. This section can be used to collisionally activate the ions below their dissociation threshold (sub-CID, harsh conditions) by applying an electric field of up to 190 V cm^{-1} over a distance of

approx. 1 cm. Under soft conditions, the electric field strength is only 30 V cm^{-1} across the respective distance. With a helium gas pressure of approx. 4.1 mbar, the ions quickly lose an excess of kinetic energy before entering the drift tube. After the separation and subsequent m/z -selection, their arrival time distributions (ATD) are recorded in the high vacuum stages of the mass spectrometer. From the ATDs, drift times are extracted from which mobilities are calculated. The mobilities are subsequently converted into $^{D^+}\text{CCS}_{\text{He}}$ using the Mason-Schamp equation.¹⁹⁵

RDD experiments were performed using a modified Thermo Fisher LTQ linear ion trap. Prior to analysis, the molecules are derivatised with a para-iodoaniline linker to provide a photocleavable C–I bond using established protocols.¹⁹⁰ To technically achieve that, the trisaccharides were modified by an additional galactose monosaccharide at the reducing end, such that the Le^x and BG-H2 motives are preserved (Figure 26). In a multistep MS workflow, the parent ions are generated by nano-electrospray ionization, collisionally activated below their dissociation threshold (sub-CID, harsh conditions), m/z -selected and irradiated using fourth harmonic (266 nm) pulses from a Continuum Minilite II Nd:YAG laser. The generated radical ions of interest are isolated and activated via CID with a relative energy of 20 to yield the respective mass spectrum.

In cryogenic IR spectroscopy, the ions are generated by nano-electrospray ionization. To avoid excessive activation, soft source conditions are used to transfer the intact ions to the gas phase.¹⁷⁴ Next, the m/z -selected species are accumulated in a hexapole ion trap, which can be maintained under room temperature (warm trap) or cooled to 90 K (cold trap). These ions are then captured by superfluid helium nanodroplets (approx. size of 10^5 atoms) traversing the trap. Inside the droplet, which acts as a transparent thermostat, the ion is rapidly cooled down to 0.4 K. Downstream the apparatus, these droplets are irradiated by the Fritz Haber Institute IR free-electron laser (FHI-FEL). The resonant frequency of the laser pulse and a molecular vibration causes the droplet evaporation, leading to the ion release from the droplet, which is recorded using the time-of-flight (TOF) spectrometer. The ion count recorded at the TOF as a function of the frequency yields the IR spectrum of the ion in cryogenic conditions.

Computational modelling*

The here presented DFT-based search protocol has been previously successfully applied to the structural assignment of glycoconjugates³⁵ and amino acids complexes.¹⁹⁶ The search protocol, which has been implemented in the glypy python package,¹⁹⁷ is outlined below. First, the conformational spaces of the amide-protonated Le^x and BG-H2 ions, with both the α and β -anomers and cis (-c-) and trans orientation (-t-) of the amide bond, were sampled using replica-exchange molecular dynamics that led to eight conformational spaces. Next, two distinct structures, which differed by the orientation of at least one glycosidic bond, were selected to initialize a structural search using CREST, a part of the xtb package.¹⁹⁸ The GFN2-xTB energy function was chosen, because it outperforms other empirical and semiempirical methods when tested against mono- and disaccharide benchmark sets.^{128, 199} The conformational ensembles were then merged and clustered to select a set of 750 conformers, which were reoptimized using the PBE0 hybrid density functional and a small basis set.²⁰⁰ For each ion, the resulting ensembles were clustered again using tighter geometrical criteria. We collected the 121 conformers within a 6 kcal mol⁻¹ energy window above the respective lowest energy structure, which were then reoptimized at the PBE0+D3 level of theory in a larger basis set. Finally, each of these eight sets was appended with structures generated from a third independent CREST search and reoptimized at the same level of theory, which resulted in additional 185 conformers. The relaxed geometries were used for the harmonic frequency analysis to derive their free energies, derivation of the anharmonic spectra for selected ions, and prediction of their CCS using the projection approximation method implemented in the sigma code (hereafter referred as ^PACCs).²⁰¹

4.3. Results and Discussion

The arrival-time distributions (ATDs) of the $[\text{Le}^x+\text{H}]^+$ and $[\text{BG-H2}+\text{H}]^+$ ions, which were recorded at soft and harsh (sub-CID) conditions to mimic the source conditions in commercial mass spectrometers, are shown in the Figure 25. The $[\text{Le}^x +\text{H}]^+$ ions show highly similar ATD profiles under both conditions

* Calculations were performed by Kim Greis, Murat Yaman, and Mateusz Marianski.

with a collision cross section in helium buffer gas $^{DT}CCS_{He}$ of 143 \AA^2 . The ATD of the $[\text{BG-H2+H}]^+$ ions, on the other hand, changes upon activation. Under soft conditions, the ions yield a $^{DT}CCS_{He}$ of 146 \AA^2 , which decreases to 143 \AA^2 at harsh conditions. A structural compaction upon heating is not observed for glycan ions of that size, and the subtle drop in size instead is likely caused by structural isomerism yielding ions of a similar size as $[\text{Le}^x+\text{H}]^+$.

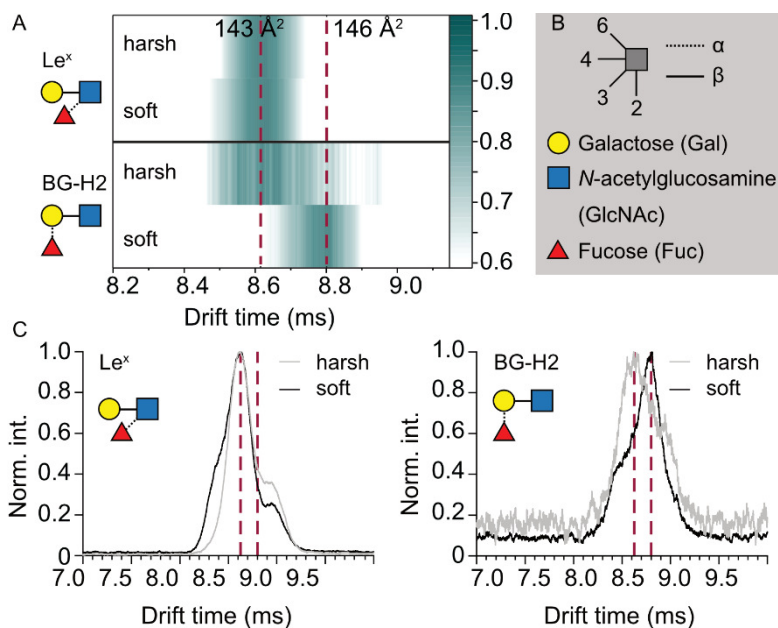


Figure 25: IM-MS data at soft and harsh (sub-CID) conditions of Le^x and BG-H2. A. Arrival-time distributions of activated (harsh conditions) and non-activated (soft conditions) intact precursor standards of $[\text{Le}^x+\text{H}]^+$ and $[\text{BG-H2+H}]^+$ in heatmap representation. Upon activation, $^{DT}CCS_{He}$ of the $[\text{BG-H2+H}]^+$ ion decreases from 146 \AA^2 to 143 \AA^2 , but $^{DT}CCS_{He}$ of $[\text{Le}^x+\text{H}]^+$ of 143 \AA^2 remains unchanged. B. Symbol nomenclature for glycans (SNFG) eases the representation of complex glycans. C. Arrival-time distributions of $[\text{Le}^x+\text{H}]^+$ and $[\text{BG-H2+H}]^+$ from (A) shown in classical plots for comparison.

To shed light on the potential rearrangement, RDD spectra of Le^x - or BG-H2-based tetrasaccharides were recorded without (soft) and with sub-CID (harsh) activation (Figure 26). The BG-H2-based structure generally shows more pronounced differences in fragment abundance between soft and harsh

activation conditions than the Le^x-based tetrasaccharide. In the respective RDD spectrum, the glycosidic fragment at m/z 512 increases in intensity at sub-CID conditions, whereas the glycosidic fragment at m/z 443 decreases, pointing towards a rearrangement within the core Gal-GlcNAc unit. In contrast, the ratios of the most intense ion fragments in the RDD spectrum of the Le^x-based tetrasaccharide do not change significantly. Furthermore, in all RDD spectra, fucose migration to the linker or the additional galactose can be observed (m/z 404, denoted with an asterisk), which highlights the general ability of fucose to undergo migration reactions to different sites. While some of these sites could be easily identified by their masses, migration along the sequence might yield fragments with unchanged m/z value so that such rearrangement is camouflaged.

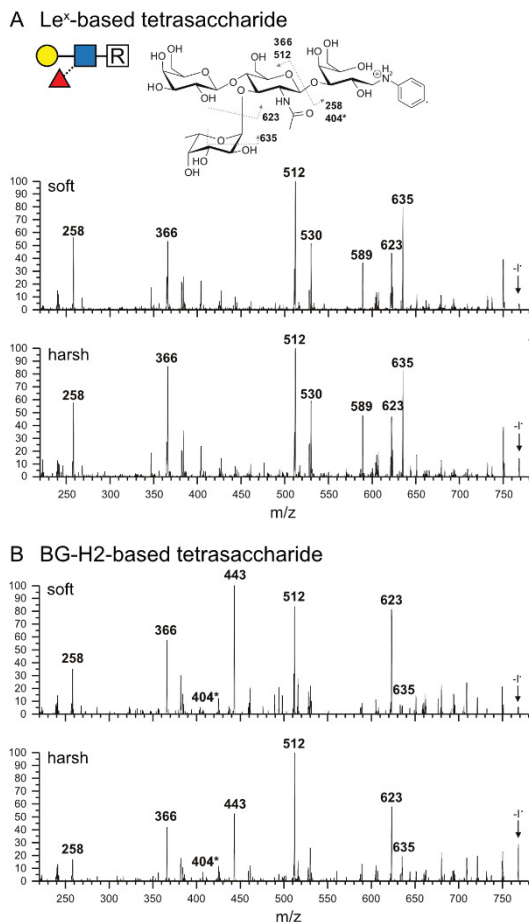


Figure 26: Radical fragmentation spectra afforded by photodissociation of the precursor m/z 895 and RDD of the generated radical glycan m/z 768 (indicated with an arrow) of the isomeric A. Le^x-based tetrasaccharide and B. BG-H2-based tetrasaccharide. The spectra of the BG-H2-based tetrasaccharide show more pronounced differences between soft and harsh conditions compared to the spectra of the Le^x-based tetrasaccharide.

To corroborate these results, we recorded the cryogenic-IR spectra (using cold trap) of the two modified tetrasaccharides (Figure 27). The vibrational spectrum of the Le^x-based tetrasaccharide appears as sub-spectrum of that of the BG-H2-based tetrasaccharide, which is readily visible in their amide region. The former

shows only one band in the amide I region (main contribution from $\nu(\text{C}=\text{O})$ of the GlcNAc), which, in addition to another blue-shifted band, is mirrored in the spectrum of the latter. With only one $\text{C}=\text{O}$ oscillator in these molecules, the two bands in the IR spectrum of the BG-H2-based tetrasaccharide, strongly shifted by approx. 50 cm^{-1} , can be associated to the presence of two isomeric structures.

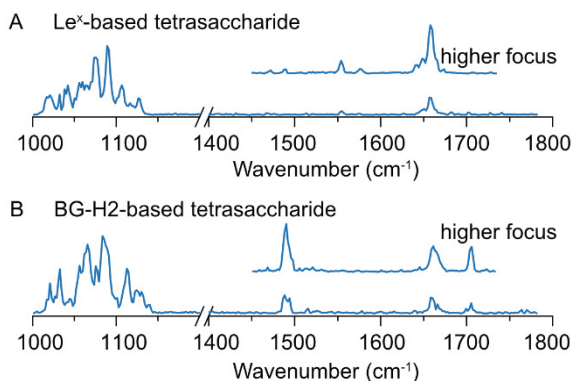


Figure 27: IR spectra of the intact A. Le^x - and B. BG-H2-based tetrasaccharides. The IR spectrum of the Le^x -based tetrasaccharide features a subset of vibrational bands from the IR spectrum of the BG-H2-based tetrasaccharide. As an inset, IR spectra recorded with a higher laser focus (i.e. higher photon density) are given with each spectrum to potentially resolve weaker vibrational modes.

Finally, we recorded the IR spectra of the two trisaccharides at different trap temperatures. We hypothesize that accumulation of the ions in the cold trap should slow the rearrangement process enough to record the IR spectra of the precursor ions. The IR spectra of the $[\text{Le}^x+\text{H}]^+$ ions obtained in a warm or cold trap show little difference, while the IR signatures of the $[\text{BG-H2}+\text{H}]^+$ ions are clearly sensitive to the trap's temperature (Figure 28). In a warm trap, the $[\text{BG-H2}+\text{H}]^+$ spectrum is similar to that of $[\text{Le}^x+\text{H}]^+$. Quenching of the isomerization kinetics by cooling of the trap leads to several additional features in the IR spectra. Both spectra of $[\text{Le}^x+\text{H}]^+$ exhibit a complex feature in the region of C–C and C–O stretching modes between $1025\text{--}1175\text{ cm}^{-1}$, and a narrow band in the amide I region at 1660 cm^{-1} . By contrast, the spectrum of the $[\text{BG-H2}+\text{H}]^+$ ions accumulated in the cold trap lacks the band at 1060 cm^{-1} and instead displays additional peaks at 1100 cm^{-1} and 1140 cm^{-1} as well as a

broad feature in the amide I region between 1640 cm^{-1} and 1700 cm^{-1} . Especially the amide I region is sensitive to the chemical environment of the *N*-acetyl group in the GlcNAc monosaccharide. At room temperature, the ions are expected to be present in more conformations than at 90K. Upon contact with the helium droplets, the ensemble is flash frozen, leading to kinetic trapping of individual conformers. Usually, this leads to crowded and more complex spectra at higher trap temperature. However, in the particular case of $[\text{BG-H2+H}]^+$, richer features in the amide I region are observed for the cold-trap settings. This is likely the result of the trapping of multiple chemical structures—precursor, product structures, and potential reaction intermediates—rather than different conformers of the same ion.

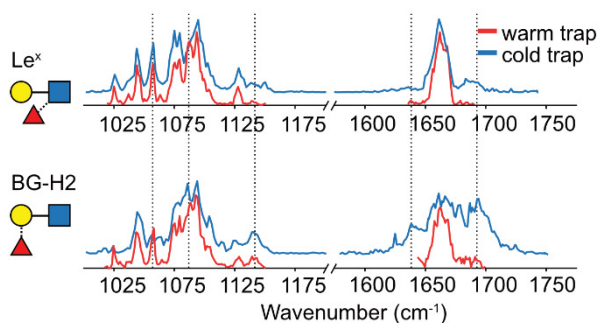


Figure 28: IR spectra of the intact $[\text{Le}^x+\text{H}]^+$ and $[\text{BG-H2+H}]^+$ standards recorded using warm and cold traps. The spectrum of $[\text{Le}^x+\text{H}]^+$ remains unchanged, but the spectrum of $[\text{BG-H2+H}]^+$ changes (guided by dashed lines) when the temperature of the trap is changed.

In summary, the three MS-based experiments can be interpreted as the fucose rearrangement from $[\text{BG-H2+H}]^+$ to the $[\text{Le}^x+\text{H}]^+$ ion. The spectra of $[\text{BG-H2+H}]^+$ show strong dependence on the activation condition, while the $[\text{Le}^x+\text{H}]^+$ appears to be insensitive. However, at this point we cannot exclude the alternative scenario, in which both glycans rearrange to a third common—yet unknown—structure. In this case, the fucose of the $[\text{Le}^x+\text{H}]^+$ ion is more labile and undergoes the rearrangement already during the ionization process, while the $[\text{BG-H2+H}]^+$ ion is stable enough to be captured in the cold trap.

To determine which of these two migration mechanisms—the interconversion between the two precursors, or the rearrangement towards a third unknown

structure—occurs, we applied methods of computational chemistry to identify the lowest-free energy structure that matches the experimental $^{\text{DT}}\text{CCS}_{\text{He}}$ and IR spectrum. The resulting free-energy hierarchy and the spectra of the two most stable anomers are shown in Figure 29.

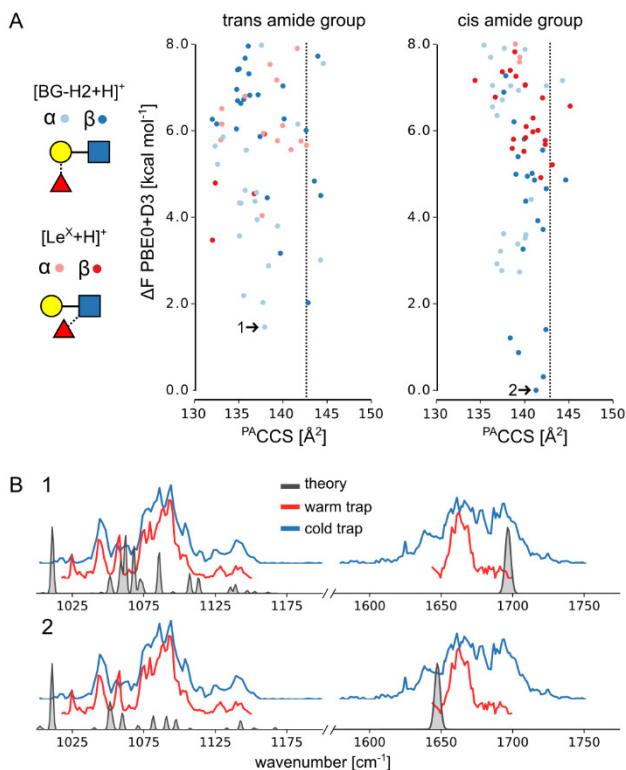


Figure 29: A. The $^{\text{PA}}\text{CCS}$ vs. relative free energy (ΔF) of conformers of the $[\text{Le}^x+\text{H}]^+$ and $[\text{BG-H2+H}]^+$ at PBE0+D3/6-311+G(d,p) level of theory. The dashed line indicates the experimental CCS, and the free energies are plotted relative to the lowest-free energy conformer of $[\beta\text{-c-BG-H2+H}]^+$. B. The comparison of the experimental and vibrational spectrum of the most stable anomers of $[\text{BG-H2+H}]^+$.

The theory predicts that the two most stable anomers are $[\beta\text{-c-BG-H2+H}]^+$ ($\Delta F=0.0$ kcal mol⁻¹), and $[\alpha\text{-t-BG-H2+H}]^+$ ($\Delta F=1.5$ kcal mol⁻¹), whereas the lowest-free energy structures of $[\alpha\text{-t-Le}^x+\text{H}]^+$ and $[\beta\text{-t-Le}^x+\text{H}]^+$ are 4.1 and 3.5 kcal mol⁻¹ less stable, respectively. The $^{\text{PA}}\text{CCS}$ of the $[\beta\text{-t-Le}^x+\text{H}]^+$ of 133 Å²

is substantially smaller than the experimental value of 143 Å². Furthermore, it is 9 Å² lower than the ^{PA}CCS of the [β-c-BG-H2+H]⁺ (142 Å²), which is larger than the contraction of 3 Å² that occurs upon activation of the [BG-H2+H]⁺ ion. Furthermore, the simulated IR spectra of the [Le^x+H]⁺ ion show a poor agreement with the experimental spectrum, as the theory does not predict the correct shape of the band in the 1075–1125 cm⁻¹ region. It also misplaces the position of the protonated amide band, predicting it at 1650 and 1700 cm⁻¹ for the β- and α-anomers, respectively. The simulated spectrum of the [β-BG-H2+H]⁺ ion (Figure 29), on the other hand, resembles to some degree the amide I features observed in the cold-trap spectrum. Specifically, the protonated amide bands of the β- and α-anomers at 1640 and 1700 cm⁻¹, respectively, match the additional shoulders in the broad feature visible in the amide region. Thus, in addition to the rearrangement product, also the [BG-H2+H]⁺ ions are observed in the IR spectrum recorded using the cold trap. However, the [Le^x+H]⁺ ion is not the product of this rearrangement and there is abundant evidence supporting this claim: 1) the α- and β-anomers of [Le^x+H]⁺ are thermodynamically less stable than the [BG-H2+H]⁺ isomers; 2) the [Le^x+H]⁺ ions are too compact and the predicted compaction of CCS is much larger than the observed 3 Å² and 3) the predicted IR spectrum of [Le^x+H]⁺ matches poorly to the experimental spectrum. Based on these grounds, we conclude that the rearrangement must proceed towards a third, unknown structure.

To reveal the identity of the rearrangement product, we have repeated the structural search for all conceivable constitutional isomers of the trisaccharide. We have investigated each type of linkage between the fucose and the Galβ1,4GlcNAc core in both anomeric forms, in both configurations of the glycosidic bond, and both orientations of the amide group, which in total equals to 56 types of structures. This involved 4999 additional optimizations in the smaller basis set, followed by 786 additional geometry relaxations and frequency calculations in the larger basis set. The relative free energies of the most stable conformers, along with the difference between ^{PA}CCS and ^{DT}CCS_{He}, are summarized in Table 1.

The expanded structural search confirms that neither $[\text{Le}^x+\text{H}]^+$ nor $[\text{BG-H2}+\text{H}]^+$ constitutes the thermodynamic global minimum of the trisaccharide structural space. Instead, the global minimum features the fucose moiety attached by a $\alpha 1,6$ glycosidic bond to the galactose. This trans-amide bond α -anomer ($[\alpha\text{-t-}\alpha 1,6+\text{H}]^+$) is more stable by $5.7 \text{ kcal mol}^{-1}$ than the most stable conformer of the BG-H2/ Le^x pair, and its $^{\text{PACCS}}$ matches the experimental value well. Two other structures, which have fucose connected by either the $\alpha 1,1\alpha$ or $\alpha 1,1\beta$ linkages to the reducing GlcNAc, are less stable by 1.0 and $2.3 \text{ kcal mol}^{-1}$ respectively, and all other structures are at least 4 kcal mol^{-1} less stable than the global minimum. Most notably, the fucose attached to the *N*-acetyl group of the GlcNAc, which has recently been reported as the potential fucose rearrangement product,²⁰² is at least $10.9 \text{ kcal mol}^{-1}$ less stable than the $[\alpha\text{-}\alpha 1,6+\text{H}]^+$ ion, and has a $^{\text{PACCS}}$ 11 \AA^2 smaller than the experimental value. In case of the β -anomer, there are six chemically different ions within a free-energy range of $1.0 \text{ kcal mol}^{-1}$. These include the $[\beta\text{-c-BG-H2}+\text{H}]^+$, the β -anomer of the fucose linked via the $\alpha 1,6\text{Gal}$ bond, an ion with $\text{GlcNAc}\beta 1,1\beta\text{Fuc}$ bond, and three other ions that have the β -glycosidic bond and cis-amide bond orientation. Other rearrangement products are at least $3.0 \text{ kcal mol}^{-1}$ less stable than the most stable β -anomer.

Table 1: The relative free energy in kcal mol⁻¹ at 300 K, with respect to the lowest-free energy α -anomer of Fuc α 1,6Gal β 1,4GlcNAc α and the difference between calculated ^{PA}CCS and experimental ^{DT}CCS_{Hlc} of 143 Å² of the lowest-free energy conformer of each possible fucose position around the Gal β 1,4GlcNAc core. Cells with the data for the Le^x and BG-H2 are in bold letters. The lowest energy conformer is highlighted in gray.

Bond configuration			α -anomer				β -anomer			
			α		β		α		β	
Linkage	Amide	ΔF	ΔCCS	ΔF	ΔCCS	ΔF	ΔCCS	ΔF	ΔCCS	
GlcNAc	1,1	trans	1.0	1	2.3	2	6.0	0	4.6	0
	1,1	cis	7.6	-12	7.7	-15	6.6	-4	6.6	-9
	1,2	trans	10.9	-11	14.9	-11	14.0	-10	17.5	-13
	1,2	cis	13.3	6	18.8	-5	15.1	3	17.7	0
	1,3	trans	8.3	-5	8.7	-8	7.7	-7	7.7	-4
	1,3	cis	11.8	-3	10.1	-4	9.1	-1	8.2	-2
	1,6	trans	5.9	-3	5.9	-8	7.6	-7	5.4	-7
	1,6	cis	9.3	1	7.1	-3	8.1	3	4.9	-3
	Gal	1,2	trans	5.7	-5	6.3	-5	6.3	0	7.4
1,2		cis	7.0	-3	7.4	-5	4.2	-1	5.0	-3
1,3		trans	8.2	0	4.2	-10	8.6	3	8.9	2
1,3		cis	9.4	1	8.9	5	7.2	2	6.8	6
1,4		trans	5.8	2	4.4	-3	7.7	3	7.3	-1
1,4		cis	10.1	-12	6.5	0	9.1	0	4.5	0
1,6		trans	0.0	0	4.1	1	4.4	0	7.0	-4
1,6		cis	7.0	0	7.9	-10	6.2	-5	8.7	2

Among these potential candidates, the RDD data excludes the 1,1 rearrangement product as potential candidate for the Le^x and BG-H2-based tetrasaccharides. In these, this position is blocked by the label but still rearrangement is observed when induced with sub-CID conditions. Therefore, the 1,1 rearrangement product shall also be excluded for the trisaccharides. This leaves only one possible candidate for the rearrangement product in Le^x and BG-H2 trisaccharides, which has the fucose linked with the α 1,6 glycosidic bond to the terminal galactose. To validate this prediction, we have compared the calculated spectra of several different low-energy ions with the experimental IR spectrum (Figure 30). The computed spectrum of the $[\alpha\text{-t-}\alpha\text{1,6+H}]^+$ provides an excellent match in the amide I region, which explains the narrowing of the band that happens in combination with the warm trap. In addition, the 1000–1200 cm^{-1} region qualitatively agrees surprisingly well with the experimental data. The spectrum shows several bands at 1040, 1075, 1100 and 1130 cm^{-1} , which are also present in the experimental IR spectrum, whereas none of the spectra of other low-energy structures provide a comparable match. Therefore, we conclude that the unknown rearrangement product observed in the three MS-based experiments has the fucose connected with the α 1,6 glycosidic bond to the terminal galactose. This assignment is validated by the three independent measurements: the thermodynamic stability of the ion, its CCS, and the IR fingerprint. The small discrepancies in the IR spectrum, such as the missing band at 1060 cm^{-1} and a small shoulder at 1700 cm^{-1} , are likely caused by the presence of small amounts of the β -anomer which features more chemically different ions of similar stability.

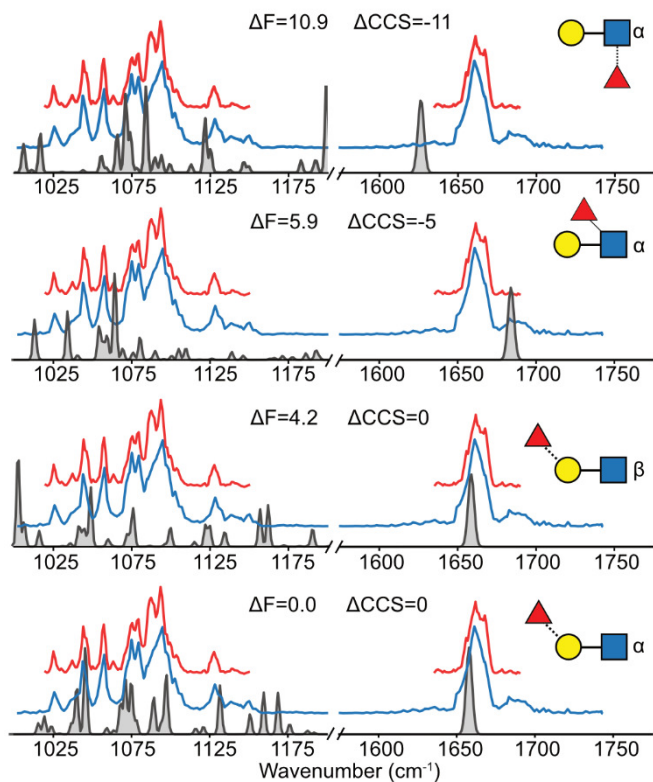


Figure 30: The comparison of the experimental and anharmonic theoretical spectra for few low-energy ions of different chemical structures.

Finally, the rearrangement mechanism that leads to the formation of the ion has to be considered. The rearrangement product retains the α -configuration of the glycosidic bond, which can be achieved by two subsequent $\text{S}_{\text{N}}2$ -like reactions, proceeding through a less stable intermediate (Figure 31). This proposed mechanism does not suffer from the entropy penalty associated with, for example, the ring opening and closing mechanism proposed in the literature.⁸⁶ During the first step of the rearrangement of $[\text{BG-H}_2+\text{H}]^+$, the protonated amide acts as a proton donor, which protonates the $\text{Fuc}\alpha 1,2\text{Gal}$ glycosidic bond through a chain of hydroxy groups. After the rotation around the $\text{Gal}\beta 1,4\text{GlcNAc}$ bond, the anomeric carbon of the fucose moiety is attacked

from the β -side by the carbonyl oxygen in the amide group. The resulting intermediate has a similar structure to that reported recently by the group of Boons.²⁰² Next, another rotation around the Gal β 1,4GlcNAc bond prearranges the ion for the second Sn2 reaction: a nucleophilic attack of the 6-O of the galactose to the α -side of the fucose anomeric carbon, which recovers the initial α -configuration. In the proposed mechanism, the first fucose migration to the amide group constitutes the rate-determining step. In $[\text{Le}^x+\text{H}]^+$, the fucose moiety is already in the proximity of the amide group. As a result, this first step of the rearrangement is accelerated, making the ion more prone to the rearrangement—so prone in fact that none of the three experiments records the true spectrum of the Le^x trisaccharide.

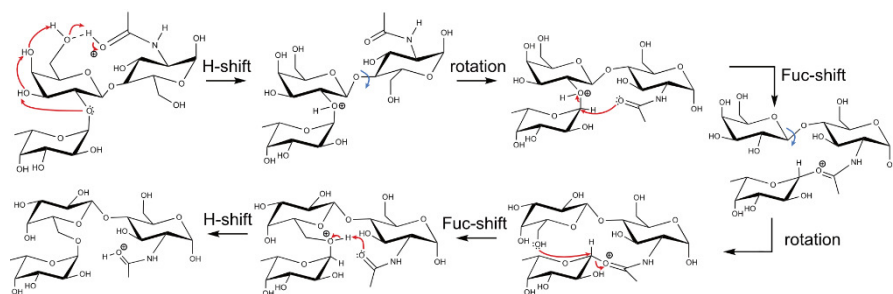


Figure 31: The proposed closed-ring reaction mechanism of the $[\alpha\text{-t-BG-H2}]^+$ leading to the formation of the $[\alpha\text{-t-}\alpha 1,6+\text{H}]^+$.

Alternatively, the rearrangement might proceed via a modified mechanism proposed by Harvey et. al,⁸⁶ which involves a protonation-induced fucose-ring opening step (Figure 32). In the $[\text{BG-H2}+\text{H}]^+$, the ring opening would form a carbocation on the anomeric carbon of the fucose group, which constitutes a good target for the nucleophilic attack by the adjacent 3-O group. After the attack, the resulting intermediate ion would feature three fused rings. The fucose transfer to 3-O could be completed by a stereospecific nucleophilic attack by the 5-O, which, by ring closing, would lead to the formation of the transition $[\text{Le}^x+\text{H}]^+$ ion. Next, the rotation around the Gal β 1,4GlcNAc would position the 6-O on the galactose for the second step of this rearrangement, which leads to the transfer of the fucose moiety to the 6-O on the galactose. In this mechanism, the $[\text{Le}^x+\text{H}]^+$ acts as intermediate ion, after the rate-limiting

first rearrangement step, which would explain why the true spectrum of the Le^x ion was not recorded.

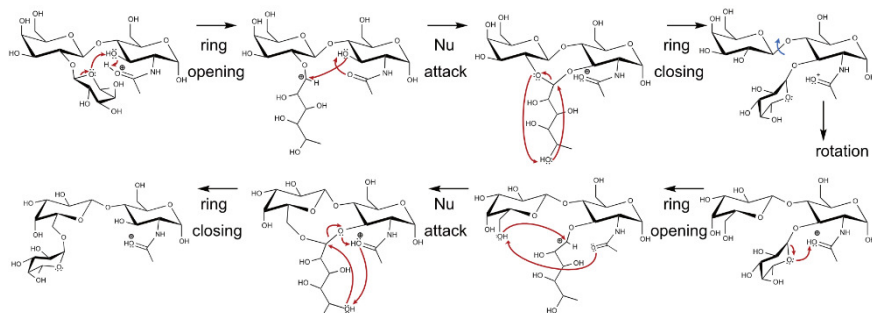


Figure 32: The proposed open-ring reaction mechanism of the $[\alpha\text{-t-BG-H2}]^+$ leading to the formation of the $[\alpha\text{-t-}\alpha 1,6\text{+H}]^+$.

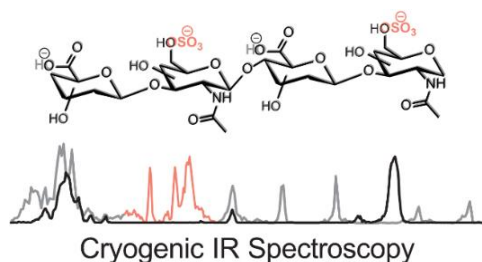
4.4. Conclusion

In summary, using three MS-based methods—IM-MS, RDD-MS and cryogenic-IR spectroscopy—we investigated the structure and stability of the intact, protonated ions of the human blood group epitopes Le^x and BG-H2 and identified the final rearrangement product of fucose migration. First, our data confirm that CID is not a prerequisite for fucose migration. Second, the large-scale DFT calculations likely identified the structure of the rearrangement product. Surprisingly, this product is neither Le^x nor BG-H2. Instead it is a third structure which is common to both of them and carries the fucose moiety attached via an $\alpha 1,6$ glycosidic bond to the terminal galactose. Regardless of the activation conditions, the CCS and spectral fingerprint of this structure match well with that recorded for rearranged Le^x . This indicates that the rearrangement barrier leading to this structure is particularly low, so low that none of the experiments probes the true Le^x structure. $[\text{BG-H2+H}]^+$ on the other hand, requires more activation energy and can therefore be trapped intact in the mass spectrometer. This difference in migration barrier highlights that structural motifs which appear similar at first glance may behave very different in tandem MS experiments—which has broad implications for the MS-based sequencing of carbohydrates.

Chapter 5

Infrared Action Spectroscopy of Glycosaminoglycans*

Glycosaminoglycans (GAGs) are a physio- and pharmacologically highly relevant class of complex saccharides, possessing a linear sequence and strongly acidic character. Their repetitive linear core makes them seem structurally simple at first glance, yet differences in sulfation and epimerization lead to an enormous structural diversity with only a few GAGs having been successfully characterized to date. Recent infrared action spectroscopic experiments on sulfated mono- and disaccharide ions show great promise. Here, we assess the potential of two types of gas-phase action spectroscopy approaches in the range from 1000 to 1800 cm^{-1} for the structural analysis of complex GAG oligosaccharides. Synthetic tetra- and pentasaccharides were chosen as model compounds for this benchmark study. Utilizing infrared multiple photon dissociation action spectroscopy at room temperature, diagnostic bands are largely unresolved. In contrast, cryogenic infrared action spectroscopy of ions trapped in helium nanodroplets yields resolved infrared spectra with diagnostic features for monosaccharide composition and sulfation pattern. The analysis of GAGs could therefore significantly benefit from expanding the conventional MS-based toolkit with gas-phase cryogenic IR spectroscopy.



* Adapted from: [Lettow, M.](#); Grabarics, M.; Mucha, E.; Thomas, D. A.; Polewski, L.; Freyse, J.; Rademann, J.; Meijer, G.; von Helden, G.; Pagel, K., IR Action Spectroscopy of Glycosaminoglycan Oligosaccharides. *Anal. Bioanal. Chem.* **2020**, *412*, 533-537. Maike Lettow and Márkó Grabarics contributed equally.

5.1. Introduction

Living systems encode information and function in the sequence of biopolymers. Determining the primary structure of nucleic acids and proteins has played a central role in the progress in life sciences, with an arsenal of sensitive and automated sequencing strategies currently available. The high-throughput *de novo* sequencing of glycans, however, is still one of the biggest challenges in bioanalytics.^{114, 116} An important group of complex carbohydrates are glycosaminoglycans (GAGs). These are strongly acidic polysaccharides with a linear sequence of repeating disaccharide units.²⁰³ Based on the structure of these disaccharides, five families of GAGs are distinguished. With the exception of naturally occurring hyaluronan, the backbone is diversely sulfated, giving rise to a multitude of isomeric sequences.

GAGs are ubiquitous in the extracellular matrix and on cell surfaces.²⁰⁴ Both as glycoconjugates and in unconjugated form, they mediate various physio- and pathophysiological processes, such as haemostasis, inflammation, tumorigenesis or target-cell recognition in infections.^{12, 205} Heparin—arguably the best known and pharmaceutically most relevant natural GAG—is a widely used and potent anticoagulant.²⁰⁶ Due to its unfavorable pharmacokinetic properties, natural heparin is being increasingly substituted by low molecular weight heparins (LMWHs) and synthetic GAG analogues. The best example of the latter approach is fondaparinux (Arixtra®), a heparin-related fully synthetic pentasaccharide, approved by the EMA and the FDA (Figure 33).

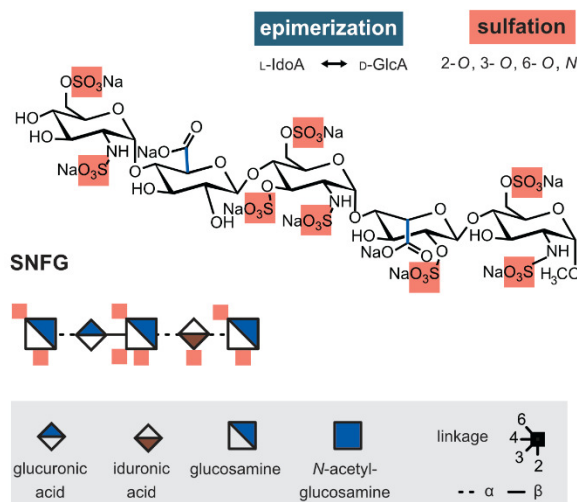


Figure 33: The synthetic anticoagulant fondaparinux (as sodium salt) presented as chemical structure and in the symbol nomenclature for glycans (SNFG, legend given in grey box).²² Highlighted in red are the positions of the sulfate groups.

Improved quality control in pharmaceutical analysis and more broadly the pursuit to elucidate the glycode, both urge the development of efficient sequencing methods. However, the analysis of GAGs has proven to be extremely challenging. Despite the relatively simple backbone, they exhibit immense complexity, which arises from two aspects of their structure: size polydispersity and sequence microheterogeneity. As polydisperse systems, their degree of polymerization is not well-defined, making efficient separation methods essential. Sequence microheterogeneity, on the other hand, is the result of uronic acid epimerization and the extent and position of sulfation. Multidimensional workflows combining chromatographic or electrophoretic separation with mass spectrometry (MS) have been traditionally employed to tackle the aforementioned difficulties.^{207, 208} Here, the high charge density and the instantaneous, often unintended loss of sulfates during ion manipulation cause additional problems. In recent years, significant progress has been made in using novel electron-based ion activation methods (EDD,²⁰⁹ EID,²¹⁰ NETD)²¹¹ and ultraviolet photodissociation (UVPD)⁷³ to improve the capabilities of MS to overcome the unique challenges of GAG analysis. Another

promising direction is to employ ion mobility spectrometry (IMS) to separate isomers or isomeric fragments prior to MS analysis.^{212, 213} A complementary method to MS-based approaches is Fourier transform infrared (FTIR) spectroscopy, which has been employed for the spectroscopic profiling of intact biological samples.^{214, 215} However, despite the aforementioned successful attempts, GAGs are still largely underexplored.

The potential of gas-phase infrared (IR) spectroscopy on mass-selected ions to tackle the complexity of GAGs has been demonstrated using infrared multiple photon dissociation (IRMPD) spectroscopy^{129, 216} and messenger-tagging spectroscopy.²¹⁷ Generally, spectra measured at low temperatures and without excessive ion heating proved to be better resolved. Here, we employ IRMPD spectroscopy and cryogenic IR spectroscopy in helium nanodroplets in a comparative manner to benchmark the analytical performance of each method. The aforementioned spectroscopic studies focused on disaccharides. Here, we extend the molecular space and report gas-phase IR spectra for synthetic tetra- and pentasaccharides with a different degree of sulfation. Such well-defined structures are well-suited for proof-of-principle analyses, exhibiting every aspect of sequence microheterogeneity, including uronic acid stereochemistry. In addition, the investigated molecules are similar in size to the smallest GAG oligosaccharides that carry biological information.²⁷ Finally, they provide residue overlap that is essential for sequencing longer GAG chains based on their characteristic fragments.

5.2. Experimental Details

Sample preparation

Fondaparinux-sodium salt (Arixtra®) was purchased from Sigma-Aldrich (St Louis, USA). The sulfated oligohyaluronans were synthesized as described previously.²¹⁸ All samples were used without further purification. Solvents (HPLC grade) were purchased from Sigma-Aldrich (St Louis, USA). Aqueous glycan stock solutions with a concentration of 1 mM were further diluted prior use with water/methanol (v/v, 50/50) to yield 50–500 μ M analyte solutions.

Infrared multiple photon dissociation spectroscopy

The experimental setup is comprehensively described in Section 2.1.

Cryogenic infrared spectroscopy in helium nanodroplets

The experimental setup is comprehensively described in Section 2.3.

5.3. Results and Discussion

5.3.1. Fondaparinux

First, the commercially available fondaparinux-sodium salt (1727 Da, Figure 33) was investigated as doubly protonated species [fondaparinux+2H]²⁺ (m/z 864) and as adduct with two additional sodium ions [fondaparinux+2Na]²⁺ (m/z 886). Using IRMPD, IR spectra for both species were recorded (Figure 34). Activation of both precursor ions leads to the loss of one of the eight possible sulfate groups as neutral SO₃. The level of activation, i.e., the number of photons exciting the molecule, is tuned such that ideally only one loss channel is populated. Both spectra show three regions of absorptions. Most specific for the investigated ions are the positions of bands above 1600 cm⁻¹, a region that is typically attributed to the stretching vibrations of carbonyl and carboxylate functional groups. The [fondaparinux+2H]²⁺ ion exhibits a strong absorption centered at 1755 cm⁻¹, which indicates that two protons are located at the carboxyl functional groups, making them neutral. In comparison, the sodiated analogue [fondaparinux+2Na]²⁺ shows a strong stretching vibration of the charged carboxylates at 1630 cm⁻¹. Between 1200 and 1450 cm⁻¹, mainly the antisymmetric SO₃⁻ stretching modes from multiple sulfate groups are observable,¹²⁹ with minor contributions of weak C–H bending vibrations. In the lower wavenumber range, combined C–O and C–C stretching vibrations from the glycan core and also the symmetric SO₃⁻ stretching modes are typically found. Overall, the spectra are rather congested, especially in the lower wavenumber range.

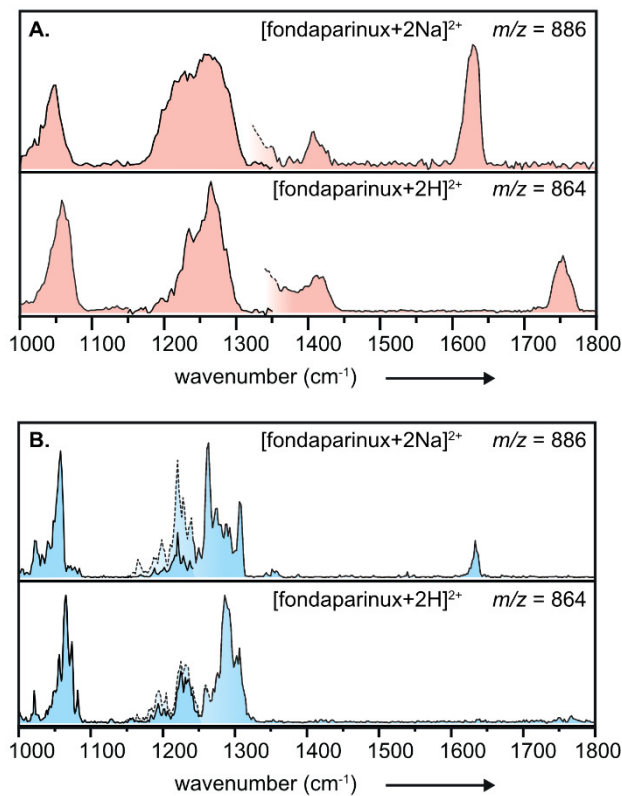


Figure 34: A. IRMPD spectroscopy of fondaparinux-sodium salt (1727 Da) investigated as adduct with two additional sodium ions $[\text{fondaparinux}+2\text{Na}]^{2+}$ (upper panel) and as doubly protonated species $[\text{fondaparinux}+2\text{H}]^{2+}$ (lower panel). B. Cryogenic IR spectroscopy in helium nanodroplets of the aforementioned ions. Dashed lines indicate an overlap of measurements using different laser focus.

To qualitatively benchmark the gain in spectral quality at low ion temperature, the IR spectra for both species $[\text{fondaparinux}+2\text{Na}]^{2+}$ and $[\text{fondaparinux}+2\text{H}]^{2+}$ were recorded using cryogenic IR spectroscopy (Figure 34). Generally, the position of the absorption bands obtained in IRMPD is reproduced. Yet, the spectral resolution in the cryogenic IR spectra is much higher even in the lower wavenumber range. A limitation is that the intensity of bands in the higher wavenumber range is much lower and as a consequence the carbonyl vibration above 1750 cm⁻¹ cannot be sufficiently resolved.

5.3.2. Hyaluronic Acid Tetrasaccharides

As cryogenic IR spectroscopy yields spectra of improved resolution, the ability of the method to provide discrete spectral features potentially diagnostic to the sulfation pattern was tested. To do so, a model system of the hyaluronic acid (HA) family was studied. Sulfated hyaluronic acid (SHA) derivatives are in focus of current research to systematically tailor GAG-protein interactions (Figure 35).²¹⁸ These newly developed potential drug candidates serve as well-defined standards in this study. With their low degree of sulfation, high stability in a mass spectrometry experiment and acidic nature, negative ion mode is well-suited for these molecules.

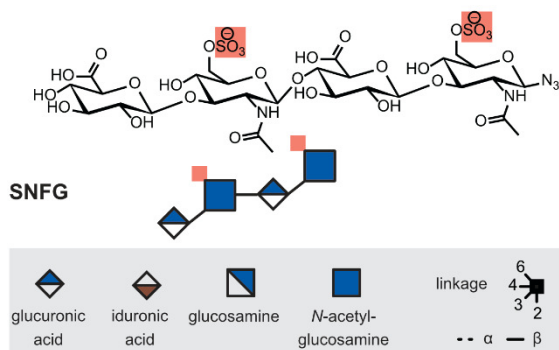


Figure 35: Chemical structure and SNFG representation²² of the investigated synthetic, sulfated (highlighted in red) hyaluronic acid (SHA) derivative.

The IR spectra for the non-sulfated HA tetrasaccharide as $[\text{HA-2H}]^{2-}$ ($m/z \approx 400$) and the doubly sulfated 2SHA derivative as $[\text{2SHA-2H}]^{2-}$ ($m/z \approx 480$) were recorded using cryogenic IR spectroscopy (Figure 36). The strongest absorption in the spectrum of the non-sulfated $[\text{HA-2H}]^{2-}$ is the carboxylate antisymmetric stretching vibration at 1645 cm^{-1} . The presence of this band confirms the deprotonation of the carboxyl functional group. The amide I band (the $\text{C}=\text{O}$ stretching vibration of secondary amides) arising from the GlcNAc moiety is most likely overlapped by the strong vibration of the carboxylate anion. The corresponding amide II and III vibrations are assigned to the features observed at 1585 cm^{-1} and 1372 cm^{-1} , respectively. Absorptions between 1000 and 1200 cm^{-1} mainly stem from the glycan core.

In order to elucidate the impact of sulfation, the spectrum of the doubly sulfated, but otherwise chemically identical glycan $[2\text{SHA}-2\text{H}]^{2-}$, was recorded for comparison. As expected from the relative proton affinities of the carboxylate and sulfate moieties, the carboxyl functional groups are neutral, which is confirmed by the $\text{C}=\text{O}$ stretching vibration at 1770 cm^{-1} . The signal at 1685 cm^{-1} is assigned to the amide I vibrations. The amide II and III vibrations can be assigned qualitatively to the bands at 1547 and 1372 cm^{-1} . The vibration at 1455 cm^{-1} is in a spectral region in which $\text{O}-\text{H}$ bending modes in carboxyl groups are typically found. Most importantly, the sulfates show multiple well-resolved bands between 1200 and 1350 cm^{-1} (Figure 36, highlighted with red squares).

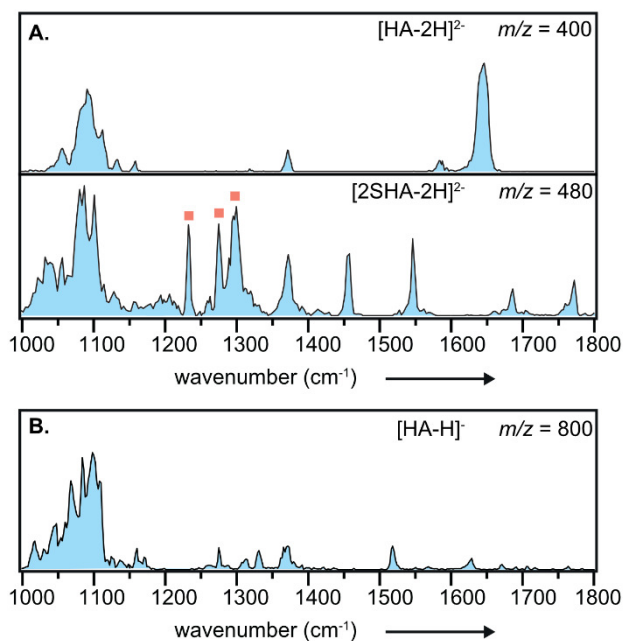


Figure 36: A. Cryogenic IR spectroscopy in helium nanodroplets of the non-sulfated hyaluronic acid (HA) as $[\text{HA}-2\text{H}]^{2-}$ (upper panel) and the 2SHA derivative as $[2\text{SHA}-2\text{H}]^{2-}$ (lower panel). Absorption bands corresponding to sulfate groups are highlighted with red squares. B. Cryogenic IR spectroscopy in helium nanodroplets of the non-sulfated HA as $[\text{HA}-\text{H}]^{-}$.

Finding the optimal charge state is crucial to obtaining good spectral quality. With two carboxylate functional groups, the singly deprotonated ion of the tetrasaccharide HA yields a congested spectrum (Figure 36). Two chemically almost identical deprotonation sites aid the formation of two deprotomers, each with a multitude of individual conformers. In addition, the dense hydrogen bonding network within the molecule can promote charge migration, which further increases the number of potentially observable species.²¹⁹

5.4. Conclusion

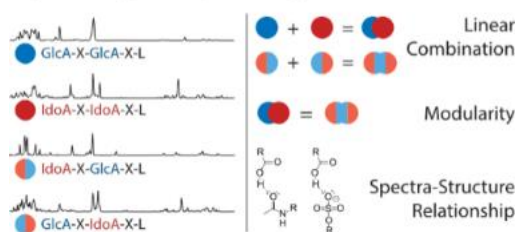
In summary, we here demonstrate the potential of IR action spectroscopy in the range from 1000 to 1800 cm^{-1} for the structural characterization of highly complex GAG oligosaccharides. Well-resolved IR spectra of oligosaccharides up to pentasaccharides were obtained using cryogenic IR spectroscopy in helium nanodroplets. Signals arising from sulfate groups appear in a spectral range in which typically no other diagnostic vibrations occur. As a result, vibrational patterns with high informational content are highly resolved. The optimal charge state depends on the functional groups present and is crucial for the spectral quality. Besides being sensitive to minute structural differences, cryogenic IR spectroscopy on mass-selected ions has the potential to be implemented in existing MS-based workflows. Current MS-based databases²²⁰ could be extended with IR fingerprints of intact and fragment ions of GAGs. As such, gas-phase IR spectroscopy could serve as a key analytical technique for the characterization of GAG oligosaccharides in the future.

Chapter 6

Cryogenic Infrared Spectroscopy Reveals Structural Modularity in the Vibrational Fingerprints of Heparan Sulfate Diastereomers*

Heparan sulfate and heparin are highly acidic polysaccharides with a linear sequence, consisting of alternating glucosamine and hexuronic acid building blocks. The identity of hexuronic acid units shows a variability along their sequence, as D-glucuronic acid and its C5 epimer, L-iduronic acid, can both occur. The resulting backbone diversity represents a major challenge for an unambiguous structural assignment by mass spectrometry-based techniques. Here, we employ cryogenic infrared spectroscopy on mass-selected ions to overcome this challenge and distinguish isomeric heparan sulfate tetrasaccharides that differ only in the configuration of their hexuronic acid building blocks. High-resolution infrared spectra of a systematic set of synthetic heparan sulfate stereoisomers were recorded in the fingerprint region from 1000 to 1800 cm^{-1} . The experiments reveal a characteristic combination of spectral features for each of the four diastereomers studied and imply structural modularity in the vibrational fingerprints. Strong spectrum-structure correlations were found and rationalized by state-of-the-art quantum chemical

Glycosaminoglycan IR Spectroscopy



calculations. The findings demonstrate the potential of cryogenic infrared spectroscopy to extend the mass spectrometry-based toolkit for the sequencing of heparan sulfate and structurally related biomolecules.

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

6.1. Introduction

Heparan sulfate and heparin are structurally closely related, sulfated representatives of glycosaminoglycans (GAGs), a class of acidic polysaccharides with a linear sequence of repeating disaccharide units.^{203, 221} Being present in the extracellular matrix and at the surface of cells, they regulate various biological processes, ranging from hemostasis to inflammation and tumor metastasis.^{26, 206, 222} The backbone of heparan sulfate and heparin consists of alternating α -D-glucosamine (GlcN, often acetylated as GlcNAc) and hexuronic acid building blocks. The identity of the hexuronic acid units varies in heparan sulfate and heparin chains: β -D-glucuronic acid (GlcA) and its C5 epimer, α -L-iduronic acid (IdoA), are both present. This backbone diversity is unique among GAGs and represents a major challenge for structural characterization.

Recent developments in electron-based ion dissociation methods enabled tandem mass spectrometry to overcome many difficulties in GAG analysis stemming from the presence of multiple sulfate groups^{71, 72, 211, 223} and from the potential variability of hexuronic acid stereochemistry.²²⁴⁻²²⁶ Despite these improvements, an unambiguous distinction between GlcA and IdoA in isomeric heparan sulfate sequences is still extremely challenging and, in certain cases, may prove to be impossible by mass spectrometry alone.²²⁷

A promising strategy to unravel hexuronic acid stereochemistry is to employ ion mobility-mass spectrometry (IM-MS). IM-MS has been successfully applied for the analysis of intact GAG ions^{212, 228, 229} as well as of their fragments to facilitate the sequencing of larger structures in a bottom-up approach.¹¹⁰ Besides improving the speed and confidence of analyte identification, IM-MS often enables the separation and differentiation of heparan sulfate ions that differ only in the stereochemistry of hexuronic acid building blocks. In certain cases, however, the mobilities of such diastereomers are too similar, impeding an unambiguous structural assignment by IM-MS.

Previous studies revealed that mass-selected GAG ions exhibit characteristic spectral features in the mid-infrared¹²⁹ and OH-stretching region.^{217, 230} Cryogenic temperatures in either messenger-tagging infrared (IR) spectroscopy²¹⁷ or helium nanodroplet IR spectroscopy, see Chapter 5, are

crucial in the identification of GAG ions larger than monosaccharides. Both cryogenic IR spectroscopy approaches yield spectra that are reproducible and comparable across laboratories and instruments.²³¹ Building on these findings, we here employ cryogenic IR spectroscopy in helium nanodroplets to study the influence of hexuronic acid stereochemistry on the vibrational fingerprints of heparan sulfate oligosaccharide ions from 1000 to 1800 cm^{-1} .

6.2. Experimental Details

Sample Preparation

The tetrasaccharides were chemically synthesized as described previously.²³² Solvents (HPLC grade) were purchased from Sigma-Aldrich (St Louis, USA). The tetrasaccharides were dissolved prior to use in water/methanol (v/v, 50/50) to yield 50 μM analyte solutions.

Cryogenic Infrared Spectroscopy

The utilized experimental setup is comprehensively described in Section 2.3 and has been described in detail previously.^{154, 233}

Computational Methods*

The conformational space of the disaccharide building blocks GlcA-GlcNAc6S, IdoA-GlcNAc6S, GlcA-GlcNAc6S-L and IdoA-GlcNAc6S-L as singly deprotonated anions were sampled using the evolutionary algorithm FAFOOM²³⁴ utilizing the external software FHI-aims²³⁵ for local DFT optimization of each sampled structure. Mutation of all rotatable bonds and ring puckers was allowed during conformational search. Each generated structure was optimized at PBE+vdWTS level of theory^{236, 237} with light basis set settings. That method yielded excellent results for glycans in past studies.^{128, 150, 189} For each building block ten genetic algorithms were performed. Low-energy structures with distinct structural motifs were reoptimized in Gaussian 16²³⁸ at PBE0-D3/6-31G(d) level of theory²⁰⁰ for all atoms, except sulfur, for which the 6-311+G(2df,2pd) basis set was employed.¹²⁹ The 6-31G(d) basis set is too small to model anions accurately and solely using that basis set leads to

* Calculations were performed by Kim Greis

unreasonable structures. It was found that reasonable structures can be efficiently modelled if the large basis set is solely used for sulfur while maintaining a small basis set for the other atoms. The harmonic frequencies and the relative free energies at 90 K of the building blocks were computed at the same level of theory. The lowest-energy structures for each building block were reoptimized and the harmonic frequencies computed at PBE0-D3/ma-def2-TZVP level of theory.

6.3. Results and Discussion

For the experiments, a set of well-defined, aminoalkyl-linked synthetic heparan sulfate tetrasaccharides was chosen. The four tetrasaccharides represent a complete set of all possible permutations with repetitions of GlcA and IdoA; the rest of the molecule is identical.²³² The chemical structures are shown in Figure 37. In the four tetrasaccharides, the configurations at two stereogenic C5 atoms are varied, which makes them diastereomers and in pairs of epimers with only one alteration. These moderately sulfated tetrasaccharides were investigated as doubly deprotonated ions [molecule-2H]²⁻ with m/z 510, which makes them relatively stable in the gas-phase environment of a mass spectrometer. The tetrasaccharides carry two 6-*O* sulfate functional groups (6S) at which the charges are localized. Thus, charge migration and population of multiple deprotomers can be limited which decreases the spectral complexity,^{129, 219} which is also discussed in Chapter 5.

Cryogenic Infrared Spectroscopy Reveals Structural Modularity in the Vibrational Fingerprints of Heparan Sulfate Diastereomers

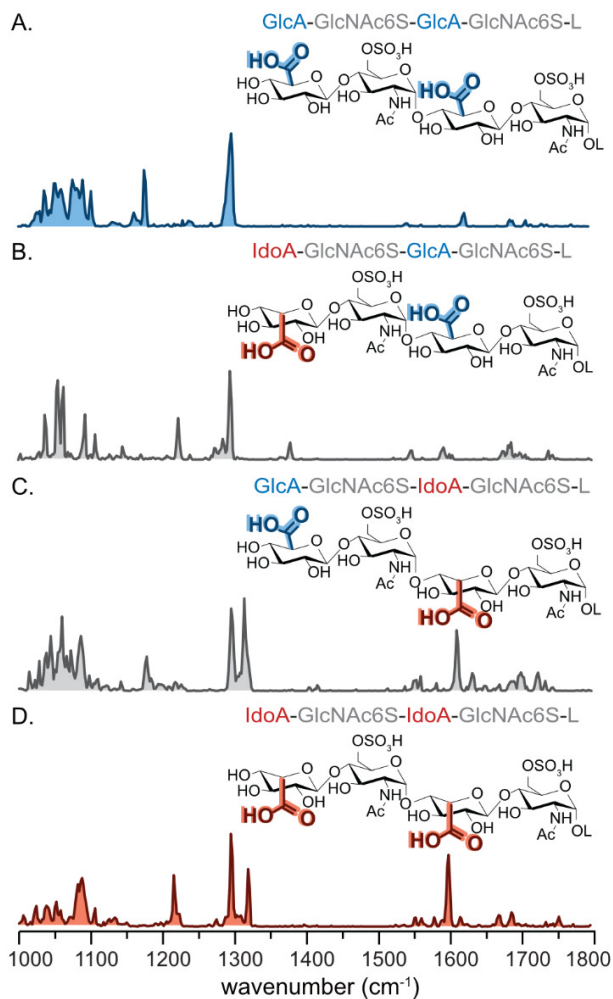


Figure 37: Cryogenic IR spectroscopy in helium nanodroplets from 1000 to 1800 cm^{-1} of heparan sulfate tetrasaccharides: (A) GlcA-GlcNAc6S-GlcA-GlcNAc6S-L, (B) IdoA-GlcNAc6S-GlcA-GlcNAc6S-L, (C) GlcA-GlcNAc6S-IdoA-GlcNAc6S-L, (D) IdoA-GlcNAc6S-IdoA-GlcNAc6S-L, investigated as [molecule-2H] $^{2-}$ anions with m/z 510. Abbreviated are β -D-glucuronic acid GlcA, α -L-iduronic acid IdoA, 6-O-sulfo-N-acetyl- α -D-glucosamine GlcNAc6S, and aminoalkyl-linker L ((CH_2) $_5$ NH $_2$) from synthesis.

Cryogenic IR spectra in the fingerprint mid-IR range from 1000 to 1800 cm^{-1} were recorded and are shown in Figure 37. All spectra consist of well-resolved lines, and the four diastereomers can be unambiguously distinguished from each other based on their IR signatures. Assignments of spectral lines within wavenumber ranges are based on theoretical calculations of disaccharides and previously published calculations of sulfated monosaccharides.¹²⁹ For all four anions, the spectral features observed from 1000 to 1150 cm^{-1} can be assigned to combined C–O and C–C stretching vibrations of the glycan core. Additionally, the symmetric SO_3^- stretching modes of the sulfate functional groups are typically found around or below 1050 cm^{-1} and usually overlap. Spectral lines corresponding to the combined, antisymmetric SO_3^- stretching modes are observed between 1150 and 1350 cm^{-1} . Smaller features from 1200 to 1500 cm^{-1} correspond to multiple C–H and O–H bending modes. For the GlcNAc moieties, the amide I and II vibrations, the stretching vibration of the C=O, and the combined bending modes of the N–H, respectively, are expected to be active between 1500 and 1700 cm^{-1} . Above 1700 cm^{-1} , spectral lines correspond to the C=O stretching vibration in the neutral carboxyl functional groups of GlcA and IdoA.

The region between 1000 and 1150 cm^{-1} in the spectra in Figure 37 is congested for all four species. At higher wavenumbers, however, characteristic bands can be observed, allowing for specific assignments. In the spectra in Figure 37 A., C., a band at around 1175 cm^{-1} is observed. In the spectra in Figure 37 B., D., this band seems to be replaced by a band at approx. 1220 cm^{-1} . It is thus reasonable to assign the band at 1175 cm^{-1} to stem from the sulfate group in GlcA-GlcNAc6S and the band at approx. 1220 cm^{-1} from the sulfate group in IdoA-GlcNAc6S, both at the nonreducing end. Moving to higher wavenumbers, in all four spectra, a band is observed at 1295 cm^{-1} . To the higher wavenumber side in the spectra in Figure 37 C., D., a band at 1330 cm^{-1} is observed, which appears to be connected to the presence of the sulfate group in IdoA-GlcNAc6S-L at the reducing end of the glycan. At higher wavenumbers around 1600 cm^{-1} , an intense band is observed in the spectra in Figure 37C., D., which again coincides with the disaccharide IdoA-GlcNAc6S-L at the reducing end. The absorption bands in the region of the

antisymmetric SO_3^- stretching modes are always present in two IR spectra, in an IR spectrum of a homogeneous diastereomer (either two GlcA or two IdoA) and in an IR spectrum of a heterogeneous diastereomer (one GlcA and one IdoA), with the exception of the vibration at 1295 cm^{-1} .

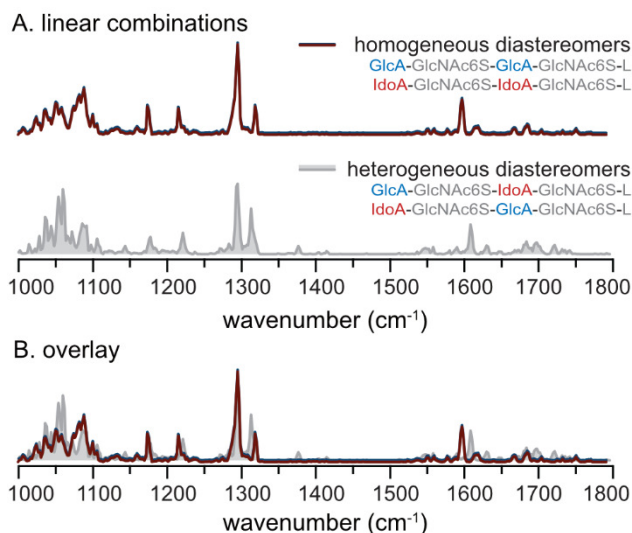


Figure 38: A. Linear combinations of the cryogenic IR spectra in helium nanodroplets from 1000 to 1800 cm^{-1} of the homogeneous diastereomers (upper panel) GlcA-GlcNAc6S-GlcA-GlcNAc6S-L and IdoA-GlcNAc6S-IdoA-GlcNAc6S-L and of the heterogeneous diastereomers (lower panel) IdoA-GlcNAc6S-GlcA-GlcNAc6S-L and GlcA-GlcNAc6S-IdoA-GlcNAc6S-L. B. Overlay of the linear combinations. Abbreviated are β -D-glucuronic acid GlcA, α -L-iduronic acid IdoA, 6-O-sulfo-N-acetyl- α -D-glucosamine GlcNAc6S and aminoalkyl-linker L ($(\text{CH}_2)_5\text{NH}_2$).

The difference in the absorption pattern in the region of the antisymmetric SO_3^- stretching region likely reflects a hexuronic acid-dependent change in the chemical environment of the sulfate groups. With the variation in the hexuronic acid stereochemistry, the hydrogen-bonding network in the tetrasaccharides seems to change. Furthermore, the local conformation in the disaccharide modules appears discrete, i.e., without strong hydrogen-bonding networks to the rest of the chain.

To test whether the absorptions derive from chemically discrete modules of disaccharides in the tetrasaccharide, the linear combinations of the IR spectra

of the homogeneous diastereomers and of the IR spectra of the heterogeneous diastereomers were constructed and are shown in Figure 38 A. Assuming discrete increments of disaccharides with discrete IR signatures, both linear combinations represent a 1:1:1:1 combination of the IR signatures of all possible disaccharide modules, GlcA-GlcNAc6S, IdoA-GlcNAc6S, GlcA-GlcNAc6S-L, IdoA-GlcNAc6S-L, at the nonreducing end or at the reducing end adjacent to the aminoalkyl-linker L. In the region of the antisymmetric SO_3^- stretching modes and also around 1600 cm^{-1} , the two linear combinations reveal very similar absorption patterns, which becomes especially apparent in the overlay of the linear combinations in Figure 38 B. The IR signatures between 1000 and 1150 cm^{-1} and above 1600 cm^{-1} are more congested and, even though qualitatively matching, did not line up perfectly.

To investigate the structural motifs that lead to the differences in the IR signatures, the conformational space of the four disaccharides was explored using the evolutionary algorithm FAFOOM²³⁴ with local density functional theory optimization in FHI-aims.²³⁵ The sampled structures were screened for differences in the interaction between the carboxyl, amide, sulfate, and potentially present aminoalkyl-linker with relevant distances less than 2 \AA . For each building block, low-energy structures with these distinct interaction motifs were reoptimized, and vibrational frequencies were computed in Gaussian 16²³⁸ at the PBE0-D3/6-31G(d)^{200, 239} level of theory for all atoms except sulfur, for which the basis set 6-311+G(2df,2pd) was employed. Low-energy structures below an arbitrary threshold of 15 kJ mol^{-1} relative to the lowest-energy structure of each building block are listed with their type of interaction and their relative free energy in Table 2. The lowest-energy structure of each building block was reoptimized and the frequencies computed at PBE0-D3/ma-def2-TZVP^{240, 241} level of theory for all atoms.

Cryogenic Infrared Spectroscopy Reveals Structural Modularity in the Vibrational Fingerprints of Heparan Sulfate Diastereomers

Table 2: Low-energy structures of the four investigated disaccharides and their structural motifs in the hydrogen-bonding pattern of the carboxyl, amide, sulfate and linker, the type of interaction and computed relative free energies at 90 K (ΔF). The free energies have been computed at PBE0-D3/6-31G(d) level of theory for all atoms except for sulfur, for which 6-311+G(2df,2pd) has been employed, using Gaussian 16. Type A, B, Y, Z interactions are specified in Figure 39. Type 0 is a non-specific interaction in which the carboxyl, amide and sulfate (and linker) only interact with hydroxy groups. Type 1 is any other, specific interaction that is not described in Figure 39. Abbreviated are β -D-glucuronic acid GlcA, α -L-iduronic acid IdoA, 6-O-sulfo-N-acetyl- α -D-glucosamine GlcNAc6S and aminoalkyl-linker L ((CH₂)₅NH₂).

Disaccharide	Structure	Structural motif in the hydrogen-bonding pattern	Type of interaction	ΔF_{90K} in kJ mol ⁻¹
GlcA-GlcNAc6S	1	COOH \cdots O=C(amide)	A	0
	2	non-specific interaction	0	11.8
IdoA-GlcNAc6S	1	COOH \cdots O ₃ SO	B	0
	2	non-specific interaction	0	1.1
	3	COOH \cdots O ₃ SO	B	2.9
	4	non-specific interaction	0	11.7
GlcA-GlcNAc6S-L	1	C=O(amide) \cdots HOCO \cdots H ₂ N	Y	0
	2	COOH \cdots NH ₂ \cdots O=C(amide)	1	4.0
	3	COOH \cdots O ₃ SO	B	10.3
IdoA-GlcNAc6S-L	1	NH(amide) \cdots NH ₂ \cdots OCOH	Z	0
	2	non-specific interaction	0	1.0
	3	NH ₂ \cdots O ₃ SO	1	3.9
	4	NH(amide) \cdots NH ₂	(Z)	5.5
	5	non-specific interaction	0	11.6

The lowest-energy structures for the four disaccharides reveal four different structural motifs, which are depicted in Figure 39 for the nonreducing end disaccharides GlcA-GlcNAc6S and IdoA-GlcNAc6S and for the disaccharides with aminoalkyl-linker GlcA-GlcNAc6S-L and IdoA-GlcNAc6S-L. The native disaccharide GlcA-GlcNAc6S predominantly forms a Type A pattern in hydrogen bonding in which the amide carbonyl oxygen acts as hydrogen bond acceptor for the neutral carboxyl OH.

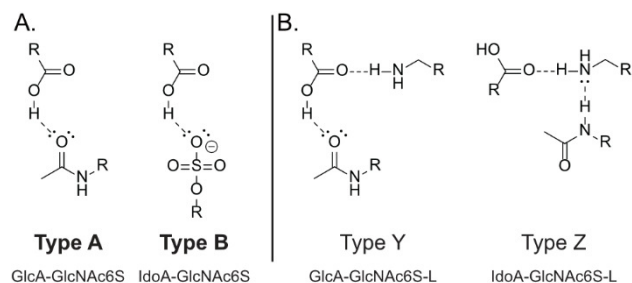


Figure 39: Structural motifs in the lowest-energy structures for (A) GlcA-GlcNAc6S (left), IdoA-GlcNAc6S (right) and (B) GlcA-GlcNAc6S-L (left), IdoA-GlcNAc6S-L (right). Abbreviated are β -D-glucuronic acid GlcA, α -L-iduronic acid IdoA, 6-O-sulfo-N-acetyl- α -D-glucosamine GlcNAc6S and aminoalkyl-linker L ((CH₂)₅NH₂).

The epimer IdoA-GlcNAc6S shows preference for a Type B hydrogen bonding pattern in which the charged sulfate group forms an ionic hydrogen bond with the neutral carboxyl. The structural motif in the disaccharides with aminoalkyl-linker is dominated by an interaction between three of the functional groups in question: the amide carbonyl from GlcNAc, the neutral carboxyl group from the hexuronic acid and the amino group from the linker.

6.4. Conclusion

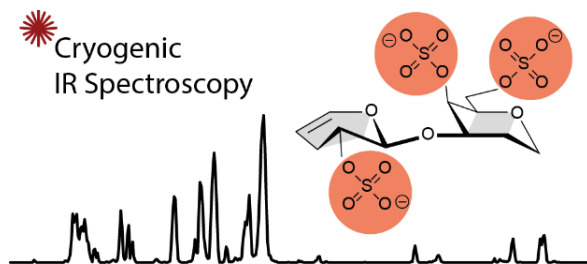
In conclusion, we show here that cryogenic IR spectroscopy is capable of unambiguously distinguishing isomeric heparan sulfate oligosaccharides that differ only in the configuration of hexuronic acid building blocks. In a wider context, the findings show that IR spectroscopy of mass-selected ions displays great potential for the analysis of heparan sulfate and heparin. Reference spectra from further small building blocks such as those reported here facilitates their identification. Cryogenic IR spectroscopy may be combined with offline

chromatographic or electrophoretic separation, ion mobility spectrometry, or tandem MS techniques to significantly improve the fragment-based sequencing¹¹⁰ of longer GAG chains in the future. Furthermore, our study reveals a strong spectra-structure correlation for this set of ions, which results from key intramolecular interactions that are unique for the individual stereoisomers and can be directly linked to specific structural elements. It is not clear whether these unique structural motifs are preserved in larger GAG ions, which is an interesting question for further investigations. The differences in the detailed molecular structure of the ions will help to unravel differences in the currently poorly understood fragmentation behavior of GAGs in MS experiments.

Chapter 7

Chondroitin Sulfate Disaccharides in the Gas Phase: Differentiation and Conformational Constraints*

Glycosaminoglycans (GAGs) are a family of complex carbohydrates vital to all mammalian organisms and involved in numerous biological processes. Chondroitin and dermatan sulfate, an important class of GAGs, are linear macromolecules consisting of disaccharide building blocks of *N*-acetylgalactosamine and two different uronic acids. The varying degree and the site of sulfation render their characterization challenging. Here, we combine mass spectrometry with cryogenic infrared spectroscopy in the wavenumber range from 1000 to 1800 cm^{-1} . Fingerprint spectra were recorded for a comprehensive set of disaccharides bearing all known motifs of sulfation. In addition, state-of-the-art quantum chemical calculations were performed to aid the understanding of the differences in the experimental fingerprint spectra. The results show that the degree and position of charged sulfate groups define the size of the conformational landscape in the gas phase. The detailed understanding of cryogenic infrared spectroscopy for acidic and often highly sulfated glycans may pave the way to utilize the technique in fragment-based sequencing approaches.



* Adapted from: [Lettow, M.](#); Greis, K.; Grabarics, M.; Horlebein, J.; Miller, R. L.; Meijer, G.; von Helden, G.; Pagel, K., Chondroitin Sulfate Disaccharides in the Gas Phase: Differentiation and Conformational Constraints. *J. Phys. Chem. A* **2021**, *125*, 4373-4379.

7.1. Introduction

Carbohydrates or glycans are biological macromolecules that make up a significant amount of organic matter on earth and fulfill essential functional and structural roles in living organisms.¹ Glycosaminoglycans (GAGs) are a major family of glycans, which participate in biological processes, such as hemostasis, inflammation, cell migration, proliferation, and differentiation.^{26, 32, 242} GAGs are long, acidic, and often sulfated, and within this family, an important class is chondroitin sulfate and dermatan sulfate (CS/DS). CS/DS are usually a part of larger glycoproteins, so-called proteoglycans, and often have chain lengths of over 50 disaccharide units.^{208, 243, 244} A characteristic for the glycan backbone of CS/DS are repeating disaccharide units of two different uronic acids (UA) linked to *N*-acetylgalactosamine (GalNAc) (Figure 40). The UA in CS is glucuronic acid (GlcA), whereas the backbone of DS contains iduronic acid (IdoA) monosaccharides. Disaccharide fragments from larger chains are the product of bacterial chondroitinase digestion, which introduces a double bond between C4 and C5 in the UA (then abbreviated Δ UA). Varying *O*-sulfations at C2 of the uronic acid and at C4 and C6 of the GalNAc further increase the complexity of CS/DS chains significantly.²⁷ The most informative level of GAG analysis resides in its disaccharide building blocks; therefore, disaccharide analysis is essential for any oligosaccharide or full-length GAG sequence.

Recent developments in mass spectrometry-based techniques, such as electron-based dissociation^{211, 212, 245} and ion mobility-mass spectrometry,^{246, 247} have improved the sensitivity and integrity in GAG analysis toward varying uronic acid stereochemistry and in sulfation.^{52, 248} Nevertheless, the distinction of certain sulfation motifs, especially *O*-sulfations at C4 and C6 in CS/DS, is still exceptionally challenging by mass spectrometry-based techniques.^{247, 249-253}

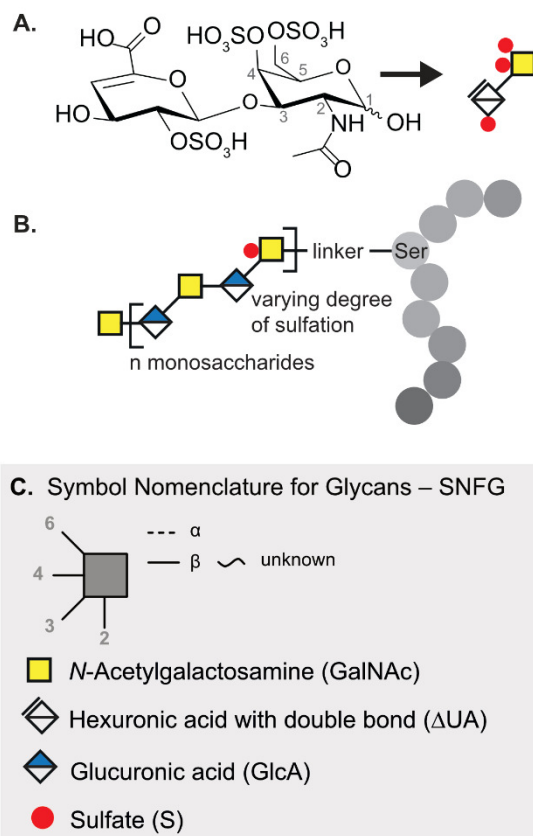


Figure 40: A. Investigated chondroitin sulfate disaccharide **1** in chemical representation followed by its representation in the symbol nomenclature for glycans (SNFG)²². Both α and β anomers are present in the sample. B. Representative structure of the proteoglycan bikunin, which carries one site for an *O*-linked chondroitin sulfate chain.^{8, 208, 254, 255} The protein chain is depicted with gray circles and the three-letter code is used to highlight serine. C. The SNFG offers simplified representations. Symbols are shown for the relevant monosaccharides in this chapter.

A novel approach for GAG structural characterization is infrared (IR) spectroscopy of ions in the gas phase.¹¹⁵ GAGs and their fragments ranging from monosaccharides to pentasaccharides (Chapter 5 and Chapter 6) have shown to exhibit a wealth of characteristic spectral features.^{129, 216, 217} Cryogenic temperatures have enhanced the spectral quality significantly in the midinfrared (Chapter 5) and the OH-stretching region.²¹⁷ Here, we assess the capability of

cryogenic IR spectroscopy for the differentiation and structural characterization of uronic acids of CS/DS disaccharides with all known motifs of sulfation.

7.2. Experimental Details

Sample Preparation

Glycosaminoglycans were purchased from Iduron (Alderley Edge, U.K.) and used without further purification. For each disaccharide, a stock solution of 1 mM in H₂O was further diluted prior to use to yield a 50 μM analyte solution in H₂O/MeOH (*v/v*, 1/1).

Ion Mobility-Mass Spectrometry

Drift tube (DT) ion mobility-mass spectrometry measurements were performed on a modified Synapt G2-S HDMS instrument (Waters, Manchester, U.K.) containing a drift tube instead of the commercial traveling wave cell²⁵⁶ and equipped with a nanoelectrospray ionization source. Ion mobilities were determined employing the stepped-field approach, and collision cross sections (CCS) were derived from the respective mobilities using the Mason-Schamp equation.¹⁹⁵

Cryogenic Infrared Spectroscopy

The experimental setup is comprehensively described in Section 2.3 and in recent publications.^{127, 151, 233}

Computational Methods*

The conformational space of the disaccharides was sampled using the evolutionary algorithm FAFOOM²³⁴ utilizing external software FHI-aims²³⁵ for local DFT optimization of each sampled structure at the PBE + vdW^{TS237} level of theory and *light* basis set settings. Mutation of all rotatable bonds and ring puckers was allowed during the conformational search. The algorithm allows functional groups to interact. In total approx. 250 structures were sampled for each disaccharide using the described approach. The methodology has previously yielded excellent accuracy for small glycans.^{128, 189}

* Calculations were performed by Kim Greis.

7.3. Results and Discussion

Chondroitin sulfate disaccharides ranging from non-sulfated species up to triply sulfated species were chosen to cover all known sulfate combinations. In Figure 40, the chemical structure of the triply sulfated disaccharide **1** is depicted. To assure molecular stability in the gas phase and limit charge migration, the disaccharides were investigated as deprotonated anionic species with a charge state equal to the number of sulfates at which the charges are localized,¹²⁹ also see Chapter 5. The non-sulfated disaccharide was investigated with a charge state equal to the number of carboxyl groups, i.e., as singly charged species.

Cryogenic IR spectra in the mid-IR range were recorded at a minimum from 1000 to a maximum of 1800 cm^{-1} (Figure 41 and Figure 42). The assignment of spectral ranges to vibrational features is based on previous theoretical calculations of sulfated monosaccharides¹²⁹ and tetrasaccharides (Chapter 6). In brief, the spectral range below 1100 cm^{-1} is most characteristic for symmetric stretching of the charged sulfate $\nu_s(\text{SO}_3^-)$ and overlaps the spectral range from 1000 to 1150 cm^{-1} in which $\nu(\text{C}-\text{O})$ and $\nu(\text{C}-\text{C})$ are found. From 1150 to 1350 cm^{-1} , strong antisymmetric stretching vibrations of the charged sulfate $\nu_a(\text{SO}_3^-)$ typically dominate the IR signature in sulfated glycosaminoglycans. Multiple minor vibrational features between 1200 and 1500 cm^{-1} correspond to C–H and O–H bending modes. Furthermore, above 1300 up to 1700 cm^{-1} , the amide vibrations I to III are found in *N*-acetylated monosaccharides, of which amide I between 1600 and 1700 cm^{-1} is usually the most intense. Minor contributions of $\nu(\text{C}=\text{C})$ in the range of amide I are characteristic for GAGs derived from lyase digestion. Above 1700 cm^{-1} , $\nu(\text{C}=\text{O})$ indicates the absence of a charged site at the carboxyl functional group. Carboxylate anions on the other hand yield an intense vibrational feature around 1650 cm^{-1} , corresponding to the antisymmetric stretch $\nu_a(\text{COO}^-)$, whereas the potentially weaker $\nu_s(\text{COO}^-)$ is typically found between 1300 and 1400 cm^{-1} .

The IR signature of the triply sulfated disaccharide **1** exhibits a unique absorption pattern with the strongest feature centered at 1292 cm^{-1} in the spectral range of $\nu_a(\text{SO}_3^-)$. Single bands are, in some cases, close to the bandwidth of the free-electron laser (FWHM approx. 5 cm^{-1} , i.e., 0.3–0.5% of the respective wavelength). The IR signatures of the isomeric disaccharides **2** to

4 show a variety of well-resolved bands and are unambiguously distinguishable from each other. Especially, the IR signatures of disaccharides **2** and **3** with *O*-sulfation at C2 exhibit surprisingly narrow spectral features for ions of this size.

Although smaller in molecular size compared to the disaccharides with higher sulfation, the singly sulfated disaccharides **5** to **7** show a more complex IR signature, as shown in Figure 42. Especially the range up to 1400 cm^{-1} is more congested and shows less deviation between weak- and high-intensity signals. In disaccharide **6** carrying *O*-sulfation at C4, the range between 1650 and 1800 cm^{-1} shows a multitude of vibrational features for the two strongest potential oscillators in that range, i.e., C=O in the carboxyl and the amide groups. The non-sulfated disaccharide **8** was included to assess the impact of the backbone, and the IR spectrum clearly reveals that this ion carries the charge at the carboxylate.

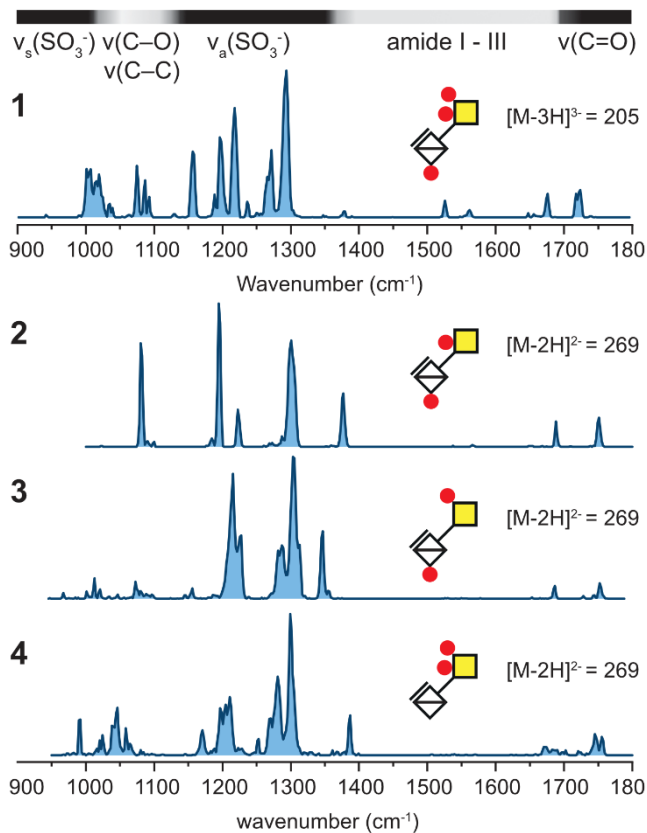


Figure 41: Cryogenic IR spectra of triply sulfated disaccharide **1** investigated as a $[\text{M}-3\text{H}]^3-$ anion with a mass-to-charge ratio (m/z) of 205 and doubly sulfated disaccharides **2-4** investigated as $[\text{M}-2\text{H}]^2-$ isomeric anions with m/z of 269. On top of the first spectrum, the main vibrational features for these ions are qualitatively assigned in a horizontal bar.

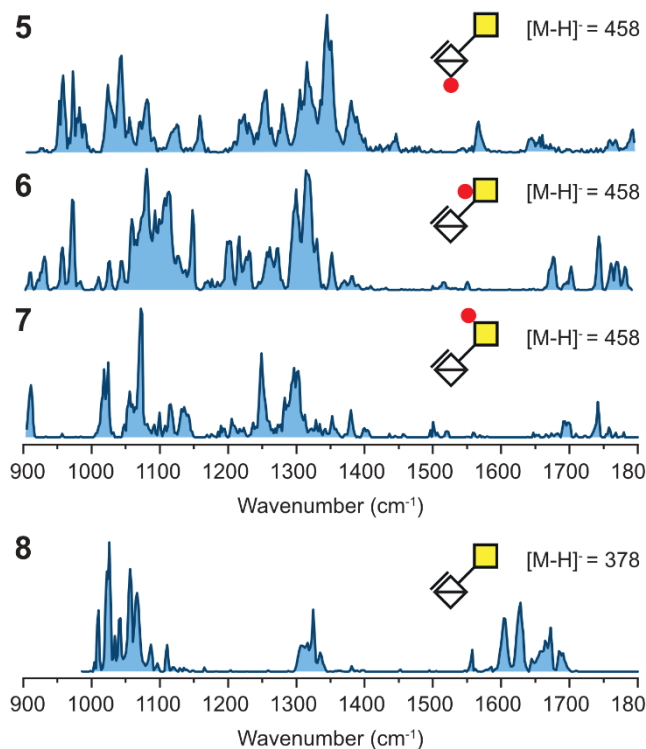


Figure 42: Cryogenic IR spectra of sulfated disaccharides **5–7** investigated as $[M-H]^-$ isomeric anions with m/z of 458 and non-sulfated disaccharide **8** investigated as a $[M-H]^-$ anion with m/z of 378.

The size of the ensemble of conformations present during the experiment correlates to the congestion of the IR spectrum and the number of vibrational features in a range in which a limited number of functional groups absorb. In our experimental setup, the ensemble of conformations is mainly defined by the temperature of the coolable ion trap (90 K). When the molecules are picked up by the superfluid helium nanodroplets (0.4 K), the rate of cooling is so fast that ions are kinetically trapped in local conformational minima.¹⁹⁴ The IR spectra of the higher sulfated disaccharides are remarkably well resolved, whereas the IR spectra of the singly sulfated disaccharides are highly congested. Our hypothesis is that the higher sulfated and also higher charged disaccharides

exhibit only a limited number of low-energy conformers, whilst the singly sulfated and also singly charged disaccharides exhibit a larger ensemble of low-energy conformers. To test this hypothesis, quantum chemical calculations were performed and the conformational space of the sulfated disaccharide β anomers was extensively sampled to yield approx. 250 conformers for each disaccharide. The β anomer was chosen under the premise that the stereocenter at C1 has only a minor influence on the overall conformation of the disaccharide. For each ion, all sampled conformers with relative energies (ΔE_{PBE}) up to 50 kJ mol⁻¹ were selected and characterized using two criteria: (A) the intramolecular distances of charged sulfates toward the hydrogens of the carboxyl and amide groups and (B) the dihedral angles $\Psi(\text{C1-O-C3-C4})$ and $\Phi(\text{C2-C1-O-C3})$ (Figure 43).

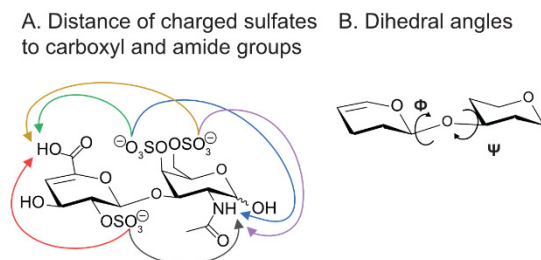


Figure 43: I. The intramolecular distances of charged sulfates toward the hydrogens of the carboxyl and amide groups is the first criterion to qualitatively assess the conformational diversity. II. The dihedral angles $\Psi(\text{C1-O-C3-C4})$ and $\Phi(\text{C2-C1-O-C3})$ at the glycosidic bond are the second criterion.

To qualitatively assess the conformational diversity, the diagrams relating relative energies to structural parameters and Ramachandran-type plots for glycosidic dihedral angles give insights into the orientation of selected functional groups and the glycan backbone, respectively. Key interactions in these molecules are hydrogen bonds of the charged sulfates toward the hydrogens of the carboxyl and amide groups. For this reason, the x-axis in the energy diagram shows the distances of all charged sulfates toward either the secondary amide of *N*-acetyl or the neutral carboxyl functional groups and on the y-axis of the relative energies. The different distances in a single conformation are, thus, given from left to right at the same relative energy. The

Ramachandran-type plots for glycosidic dihedral angles, analogous to Ramachandran plots of peptides and proteins, show the dihedral angles $\Psi(\text{C1-O-C3-C4})$ and $\Phi(\text{C2-C1-O-C3})$ on y- and x-axes, respectively, with dots for single conformers, which are colored according to the relative energy.

The conformers of disaccharide **1**, Figure 44, with three charged sulfate groups show trends in the orientation of the sulfate groups. More importantly, one conformer is with 15 kJ mol^{-1} significantly lower in energy than the second- and third-lowest-energy conformers. The hydrogen-bonding network is dominated by the interaction between the *O*-sulfate group at C4 and the neutral carboxyl functional group but also by the *O*-sulfate group at C2 and the secondary amide. The *O*-sulfate group at C6 has in all conformers a distance larger than 5 \AA toward the amide and carboxyl to minimize Coulombic repulsion with the sulfate groups in their proximity. The Ramachandran-type plot shows a very narrow distribution, indicating similarities in the glycan backbone of the low-energy conformers. Conformers of high similarity are more likely to relax into the same local minimum or the global minimum in the cryogenic ion trap.

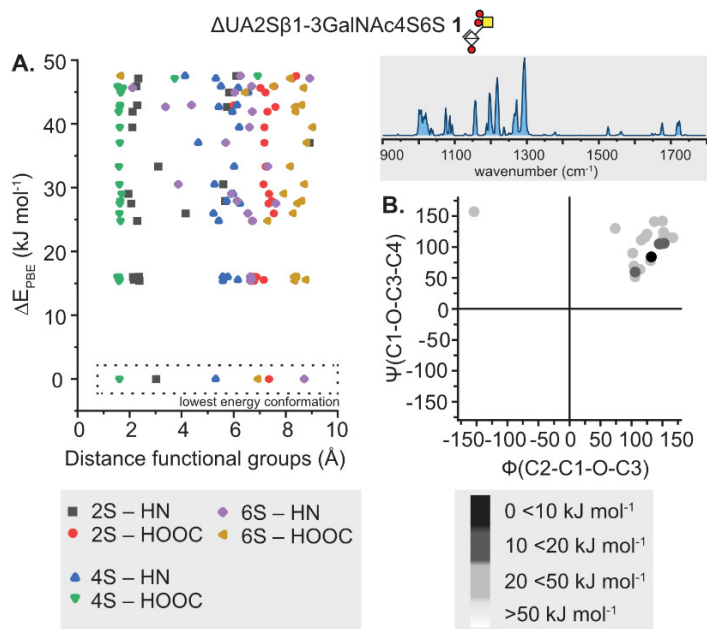


Figure 44: Results of quantum chemical calculations for the triply sulfated disaccharide **1**. For comparison, a scaled version of the IR spectrum from Figure 41 is reproduced. A. Calculation of the intramolecular distance of selected functional groups in Å in conformers with relative energies $\Delta E_{\text{PBE}} < 50$ kJ mol⁻¹. B. Calculation of the dihedral angles $\Psi(\text{C1-O-C3-C4})$ and $\Phi(\text{C2-C1-O-C3})$ in the degree at the glycosidic bond with respect to the relative energies of the conformers. The results are presented in Ramachandran-type plots for glycosidic linkages.

Disaccharide **2** with two charged sulfate groups likewise reveals trends in the energy diagram, yet, far more conformers are within 20 kJ mol⁻¹ of the lowest-energy conformer (Figure 45). The distribution in the Ramachandran-type plot is narrow. In the singly sulfated disaccharide **6**, on the contrary, the O-sulfate at C4 does not form a hydrogen bond with the neutral carboxyl (Figure 45). The Ramachandran-type plot shows that two clusters of conformers with large differences in their glycan backbone are present. This can indicate that at least two but probably several local minima are populated that are conformationally constrained to reach the same or global minimum structure.

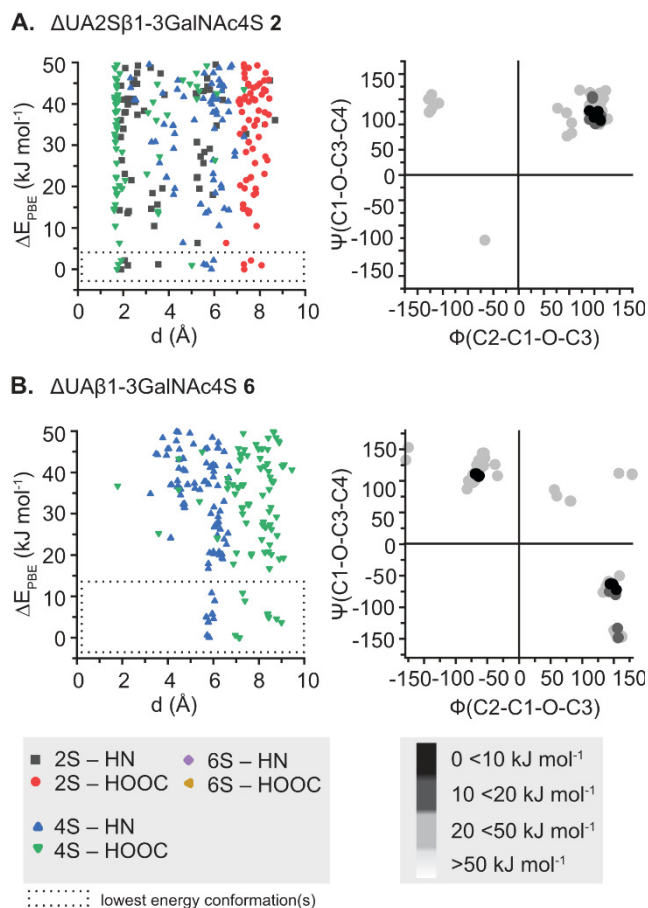


Figure 45: **Left panel:** calculation of the intramolecular distance of selected functional groups in Å in conformers with relative energies $\Delta E_{\text{PBE}} < 50 \text{ kJ mol}^{-1}$. **Right panel:** calculation of the dihedral angles $\Psi(\text{C1-O-C3-C4})$ and $\Phi(\text{C2-C1-O-C3})$ in the degree at the glycosidic bond with respect to the relative energies of the conformers. The results are presented in Ramachandran-type plots for glycosidic linkages. Results for the doubly sulfated disaccharide **2** (A) and singly sulfated disaccharide **6** (B). Generally, an increase in conformational heterogeneity is observed with a decreasing extent of sulfation and charge.

The plots of the disaccharides **3–5** and **7** (Appendix) further support the observed trend of a growing conformational heterogeneity in chondroitin

sulfate ions with a decreasing degree of charged groups, which are sulfate groups in this case.

For disaccharides **1**, **2**, and **6**, selected low-energy conformers were reoptimized at a higher level of theory. Theoretical infrared spectra were computed for the lowest-energy conformers and compared to the experimental spectra. The match between computed and experimental spectra is satisfying, yet, anharmonicities challenge the calculations for glycosaminoglycans and extend the computational effort.¹³¹

In addition to gas-phase IR spectroscopy, drift tube ion mobility-mass spectrometry (DTIM-MS) experiments were performed to compare the potential of the method for disaccharide differentiation. The rotationally averaged collision cross section (CCS) reflects the shape of an ion and is an instrument-independent physical property. CCS values of the disaccharides, measured in nitrogen and helium, are summarized in Table 3. The most challenging isomeric disaccharides **6** and **7** with *O*-sulfate either at C4 or at C6 are not distinguishable based on their CCS. The results indicate that differentiating certain isomeric disaccharides in sets **2–4** and **5–7** is possible but may require IMS instrumentation with resolving power on the order of 100 when they are present in a mixture and measurement uncertainties below 0.5% when present alone.

Table 3: Collision cross sections (CCSs) of the investigated chondroitin sulfate disaccharides.

	Molecule	z	m/z^a	$^{DT}CCS_{N_2}^b$	$^{DT}CCS_{He}^b$
1	$\Delta U A 2 S \beta 1-3 G a l N A c 4 S 6 S$	-3	205	317	160
2	$\Delta U A 2 S \beta 1-3 G a l N A c 4 S$	-2	269	241	138
3	$\Delta U A 2 S \beta 1-3 G a l N A c 6 S$	-2	269	244	140
4	$\Delta U A \beta 1-3 G a l N A c 4 S 6 S$	-2	269	239	135
5	$\Delta U A 2 S \beta 1-3 G a l N A c$	-1	458	190	127
6	$\Delta U A \beta 1-3 G a l N A c 4 S$	-1	458	189	122
7	$\Delta U A \beta 1-3 G a l N A c 6 S$	-1	458	189	122
8	$\Delta U A \beta 1-3 G a l N A c$	-1	378	178	112

^a monoisotopic mass-to-charge (m/z) ratio. ^b in \AA^2 , DT: drift tube.

7.4. Conclusion

In summary, we show here that cryogenic IR spectroscopy has the potential to unambiguously differentiate all known sulfation patterns in uronic acids of chondroitin sulfate and dermatan sulfate disaccharides, including the distinction of 4-*O* and 6-*O* sulfation, a long-standing challenge in GAG analysis. Identification of disaccharide isomers based purely on their ion mobility-derived CCS values is partially ambiguous and requires outstanding resolving power. Furthermore, we show that for chondroitin sulfate ions with a low charge state and a low degree of sulfation, large ensembles of low-energy conformers are present. Whereas for higher charged and higher sulfated chondroitin sulfate ions, only a few low-energy conformers are present. This conformational constraint in highly charged gas-phase ions has also been described for other biomolecules such as peptides and proteins.²⁵⁷ Yet,

additional experiments are necessary to evaluate whether this effect purely results from Coulombic repulsion or also from the nature of the *O*-sulfate groups, which, in contrast to more localized carboxylates, potentially induce further constraints with their steric repulsion²⁵⁸ and conformational flexibility.

Chapter 8

Predicting Structural Motifs of Heparan- and Chondroitin Sulfates using Cryogenic Infrared Spectroscopy and Random Forest*†

Machine learning algorithms for extracting patterns in data sets have become important tools in diverse areas of chemistry. Especially in infrared (IR) spectroscopy, random forest algorithms have been successfully used for classification tasks and regression analysis in small to large data sets. With the unprecedented resolution of cryogenic IR spectroscopy for glycosaminoglycans (GAGs), the question was raised whether random forest modelling could aid to predict structural motifs in complex GAG samples from their IR signature. A proof-of-concept study is presented to combine cryogenic IR spectroscopy with random forest to deduce the sulfation pattern and GAG family from the IR spectrum. With a comparably small training set of 22 cryogenic IR spectra, a hexasaccharide ion could be correctly classified for the GAG family and the presence of *N*-, 2-*O*-, 4-*O*-, and 6-*O* sulfation. Current similarity, score-based techniques that operate on databases holding a rich collection of GAG spectra may thus be superseded by the predictive power of machine learning.

* Based on an early version of: Riedel, J.; [Lettow, M.](#); Grabarics, M.; Götze, M.; Miller, R. L.; Boons, G.-J., Meijer, G.; von Helden, G.; Szekeres, G. P.; Pagel, K., Predicting Structural Motifs of Heparan- and Chondroitin Sulfates using Cryogenic Infrared Spectroscopy and Random Forest. *Manuscript in preparation*.

Maik Lettow carried out the cryogenic infrared spectroscopy experiments and Jerome Riedel performed the computational work which is only briefly summarized in the present thesis.

† Partly, the introduction is adapted from: Grabarics, M.; [Lettow, M.](#); Kirschbaum, C.; Greis, K.; Manz, C.; Pagel, K., Mass Spectrometry-Based Techniques to Elucidate the Sugar Code. *Chem. Rev.* **2022**, 122, 7840-7908.

8.1. Introduction

Glycosaminoglycans (GAGs) are a large class of linear, often highly acidic polysaccharides found in nature. Depending on the disaccharide building block, four families are generally differentiated, namely heparin and heparan sulfate (Hp/HS), chondroitin and dermatan sulfate (CS/DS), keratan sulfate (KS) and hyaluronan (HA). For their analysis, tandem mass spectrometry (MS/MS) using mostly electron-mediated fragmentation⁵³ and more recently ion mobility-mass spectrometry (IM-MS)¹¹⁰ are being employed nowadays. Yet, the sequencing of GAGs is still extremely challenging and pure synthetic or naturally-derived standards are rare.

As a relatively new field augmenting the GAG sequencing toolbox, GAG analysis by gas-phase ion spectroscopy in the infrared (IR) range emerged only throughout the past decade.^{96, 115} In the following paragraph, the research from different groups is reviewed to strengthen that the field is at a stage in which large data analysis tools become more relevant.

The Dugourd group published the first experimental gas-phase spectra of sulfated disaccharide anions in the UV range from 220 to 290 nm.²⁵⁹ The exact gas-phase molecular structure of the set was further investigated using density functional theory (DFT). The experimental spectra were found to be in reasonable agreement with the calculated spectra and revealed that UV spectroscopy is sensitive to the modes of the sulfate groups present.²⁶⁰ The first gas-phase IR spectrum using infrared multiple photon dissociation (IRMPD) spectroscopy in an FT-ICR cell from 3400 to 3700 cm^{-1} of D-glucuronic and L-iduronic acid monosaccharides was published by Polfer and co-workers.²⁶¹ They investigated rubidium adducts of monosaccharides and showed that the method is sensitive to the epimerization at C5. Similar results were obtained for the same monosaccharides investigated as anions.²³⁰ Compagnon and co-workers¹²⁵ used IRMPD spectroscopy from 3200 to 3700 cm^{-1} to record the spectrum of a protonated glucosamine with 6-*O* sulfation. In their publication, they describe the potential of gas-phase spectroscopy coupled to MS as a tool to differentiate sulfated from phosphorylated saccharides based on their vibrational modes. Harmonic and anharmonic (VPT2 and finite temperature molecular dynamics) frequency simulations were used to predict the wavenumber of

certain OH stretches in the sulfate and phosphate functional groups. Glucosamine with 6-O sulfation was later used as a model to test the accuracy of a large number of quantum chemical methods.¹³¹ Methods which are satisfying at higher wavenumbers are shown to fail from 500 and 1700 cm^{-1} as the anharmonic nature of the sulfates challenges the calculations. The best match between experimental and theoretical spectra from 500 to 1700 cm^{-1} was obtained using the computationally most demanding method. DFT was further applied to anionic monosaccharides and compared to their IRMPD experimental spectra.¹²⁹

Going larger, IR spectra of isomeric CS and HS disaccharides as positively charged sodium adducts were published in the range from 3200 to 3700 cm^{-1} using messenger-tagging spectroscopy coupled to IM separation.²¹⁷ At cryogenic temperatures, the resolving power is sufficient to differentiate the five isomers only based on their IR fingerprints. Using IRMPD spectroscopy, four isomeric CS and HS disaccharides were investigated from 2700 to 3700 cm^{-1} and two of the set also from 550 to 1850 cm^{-1} .¹²⁹ The authors explored the use of different charge states, species, and ionization modes to limit the numbers of conformers present in the gas phase. In another study, MS^n in a 3D ion trap analyzer was combined with IRMPD spectroscopy from 2800 to 3700 cm^{-1} for Y- and B-type fragment anions of CS disaccharide isomers with 4-O or 6-O sulfation.²¹⁶ Generally, these two isomers cannot be differentiated by MS alone and are challenging in IMS. The resolving power of IRMPD reached its limitations for GAG disaccharides, yet here, the two isomers could still be distinguished from their spectroscopic fingerprints.

With the advent of machine learning, algorithms for pattern recognition in IR spectra are now frequently used to circumvent the need for comprehensive databases to identify unknown analytes. Random forest (RF)^{262, 263} is a supervised machine learning technique which has been successfully applied in vibrational spectroscopy, e.g., to identify diseases in serum samples or to follow pathological processes in cells.²⁶⁴⁻²⁶⁶ Random forest classifiers train an ensemble of classification and regression trees (CART) to create a network of decision rules. CARTs answer the classification by probing spectral features with respect to their expected value inferred from model training. In random forest

classifiers, the prediction probability is simply obtained as the ensemble average of each CART prediction.

Here, we combine cryogenic IR spectroscopy in the gas phase and random forest to predict structural motifs—namely the GAG family (HS versus CS), *N*-sulfation (*N*-S), *O*-sulfation at 2-*O* (2-*O*-S), 4-*O* (4-*O*-S), and 6-*O* (6-*O*-S)—that almost fully define the chemical structure of GAGs. We present a workflow to train random forest classifiers that can be readily extended for further classifications, such as *O*-sulfation at C3. A machine learning approach for IR spectroscopy should not be used as a black box tool. The in-depth understanding of vibrational features in the IR spectra is still mandatory for an accurate interpretation. Therefore, a detailed discussion of spectral bands is included before the data is fed into the algorithm.

8.2. Experimental Details

Materials and Methods

A library of GAG standards is assembled to cover sulfation variation, epimerization, backbone diversity, degree of polymerization, and chain length. The spectrum library includes sixteen disaccharides (CS **1-8**, HS **9-16**), six tetrasaccharides (**17-22**), and one hexasaccharide (**23**). The disaccharides were purchased from Iduron (Manchester, UK). The tetrasaccharides and hexasaccharide were chemically synthesized as previously published.^{232, 267} Solvents (HPLC grade) were purchased from Sigma-Aldrich (St Louis, USA). All glycans were dissolved prior to use in water/methanol (v/v, 50/50) to yield 50 μM analyte solutions. The gas-phase IR spectra of the library were recorded on the custom-built helium nanodroplet instrument in the range of 1000–1800 cm^{-1} . A description of the experimental setup can be found in Section 2.3.

Computational Details

The experimental IR spectra were extracted from time-of-flight intensity values measured as a function of the laser wavelength. The ion signal was divided by the laser power as a first order approximation. Linear interpolation with 2 cm^{-1} steps was employed to align spectra on a commonly shared wavenumber axis before finding the largest shared wavenumber range among spectra (1010–

1786 cm^{-1}). In addition to the spectroscopic features, information about the charge, degree of sulfation, and, when tetrasaccharides were part of the training set, the presence of a linker was added. The wavenumber features were binned by integration of the IR trace in windows of fixed width (15 cm^{-1}) for the purpose of feature reduction. After spectral binning, bins were normalized to [0,1] to minimize the influence of absolute intensity values as latent variables in growing of the decision trees, because in helium nanodroplet spectroscopy intensities are not linear due to the absorption of multiple photons before the ions are ejected. This can result in over- or underestimating intensity values of vibrational modes.¹⁵⁴ For model training, feature selection was performed using an evolutionary optimization algorithm as implemented in the *GAFeatureSelectionCV* class provided by *sklearn-genetic-opt*.^{268, 269} The algorithm uses selection, reproduction, and mutation stages to choose the locally optimal features. In each generation, only the fittest candidates are used for reproduction. Classification was performed using a random forest classifier trained by a Leave-One-Out cross-validation approach. Computationally, the *RandomForestClassifier* class provided by *scikit_learn* was used.²⁷⁰

8.3. Results and Discussion

















Cryogenic Infrared Spectroscopy

The set of di-, tetra- and hexasaccharide GAGs form a diverse spectral library, from which structural motifs could be inferred. The full library of glycosaminoglycans is shown in Table 4 and a legend to the symbol nomenclature for glycans (SNFG) is given in Figure 46. The CS disaccharide spectra of **1–8** are presented and discussed in Chapter 7. Whereas, the HS tetrasaccharide spectra of **17–20** are shown and part of the discussion in Chapter 6.

Table 4: Assembling a library of GAG standards. The library contains of 16 disaccharides, six tetrasaccharides and one hexasaccharide. The GAGs can consist of glucuronic acid (GlcA), iduronic acid (IdoA), uronic acid with double bond (Δ UA), *N*-acetylglucosamine (GlcNAc), *N*-acetylgalactosamine (GalNAc) or glucosamine (GlcN) and can carry *N*- (NS), 2-*O* (2S), 4-*O* (4S), and 6-*O* (6S) sulfation. The m/z -ratio of the ions are given and to ease the reading of their chemical structures, the symbol nomenclature for glycans (SNFG) is depicted for each GAG (for the legend, see Figure 46). The tetrasaccharides carry an aminopentyl linker (L1), and the hexasaccharide a para-methoxyphenyl linker (L2).

Label	Structure	SNFG	m/z
1	Δ UA2S β 1,3GalNAc4S6S		205
2	Δ UA2S β 1,3GalNAc4S		269
3	Δ UA2S β 1,3GalNAc6S		269
4	Δ UA β 1,3GalNAc4S6S		269
5	Δ UA2S β 1,3GalNAc		458
6	Δ UA β 1,3GalNAc4S		458
7	Δ UA β 1,3GalNAc6S		458

Predicting Structural Motifs of Heparan- and Chondroitin Sulfates using Cryogenic
Infrared Spectroscopy and Random Forest

8	Δ UA β 1,3GalNAc		378
9	Δ UA2S β 1,4GlcNS6S		191
10	Δ UA2S β 1,4GlcNAc6S		269
11	Δ UA2S β 1,4GlcNS		247
12	Δ UA β 1,4GlcNS6S		247
13	Δ UA β 1,4GlcNAc6S		458
14	Δ UA2S β 1,4GlcNAc		458
15	Δ UA β 1,4GlcNS		416
16	Δ UA β 1,4GlcNAc		378
17	GlcA β 1,4GlcNAc6S α 1,4GlcA β 1,4GlcNAc6S α 1,4-L1		510
18	IdoA β 1,4GlcNAc6S α 1,4GlcA β 1,4GlcNAc6S α 1,4-L1		510
19	GlcA β 1,4GlcNAc6S α 1,4IdoA β 1,4GlcNAc6S α 1,4-L1		510
20	IdoA β 1,4GlcNAc6S α 1,4IdoA β 1,4GlcNAc6S α 1,4-L1		510
21	GlcA β 1,4GlcNS6S α 1,4IdoA β 1,4GlcNS6S α 1,4-L1		273
22	IdoA β 1,4GlcNS6S α 1,4GlcA β 1,4GlcNS6S α 1,4-L1		273
23	GlcNAc6S α 1,4GlcA β 1,4GlcNAc6S α 1,4GlcA β 1,4GlcNAc6S α 1,4GlcA β 1,4-L2		499

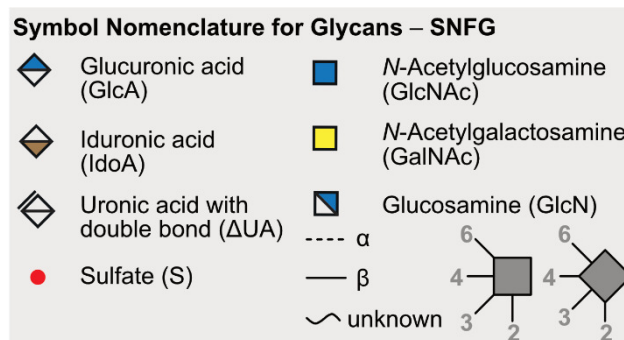


Figure 46: Legend of the symbol nomenclature for glycans (SNFG).²² SNFG is used for a concise representation of the investigated GAGs.

The cryogenic IR spectra of the heparan sulfate disaccharides **9-16** were measured in the range from 1000 to 1800 cm^{-1} and are plotted in Figure 47. Selected vibrational features (summarized in Table 5) were assigned based on previous experimental and theoretical results of glycan cations and anions.^{129, 271, 272} In sulfated GAGs, the 1000 to 1150 cm^{-1} range is populated by $\nu(\text{C}-\text{O})$ and $\nu(\text{C}-\text{C})$ modes, which overlap with the symmetric stretching of charged sulfates. The strong antisymmetric SO_3^- stretching vibrations appear between 1150 to 1350 cm^{-1} . These features serve as a unique fingerprint of the sulfation pattern in glycans. Further minor bands between 1200-1500 cm^{-1} can be assigned to C–H and O–H bending modes. The distinct amide I-III modes between 1300 and 1700 cm^{-1} (of which amide III is generally very weak in IR spectra) stem from *N*-acetylation in GAGs. The amide I, with major contribution from $\nu(\text{C}=\text{O})$, is typically the most intense amide region in the spectra, and it is highly sensitive to the intramolecular hydrogen bonding pattern associated to the conformation of the glycan. With strong hydrogen bonding at the carbonyl functional group, the C=O bond is weakened and the amide I is red-shifted. The amide II region is mainly populated by the N–H in-plane bending vibration with decreasing contributions from the C–N stretching motion and other minor components. The weak stretching mode $\nu(\text{C}=\text{C})$ around 1600 cm^{-1} is characteristic for GAGs derived from lyase digestion.

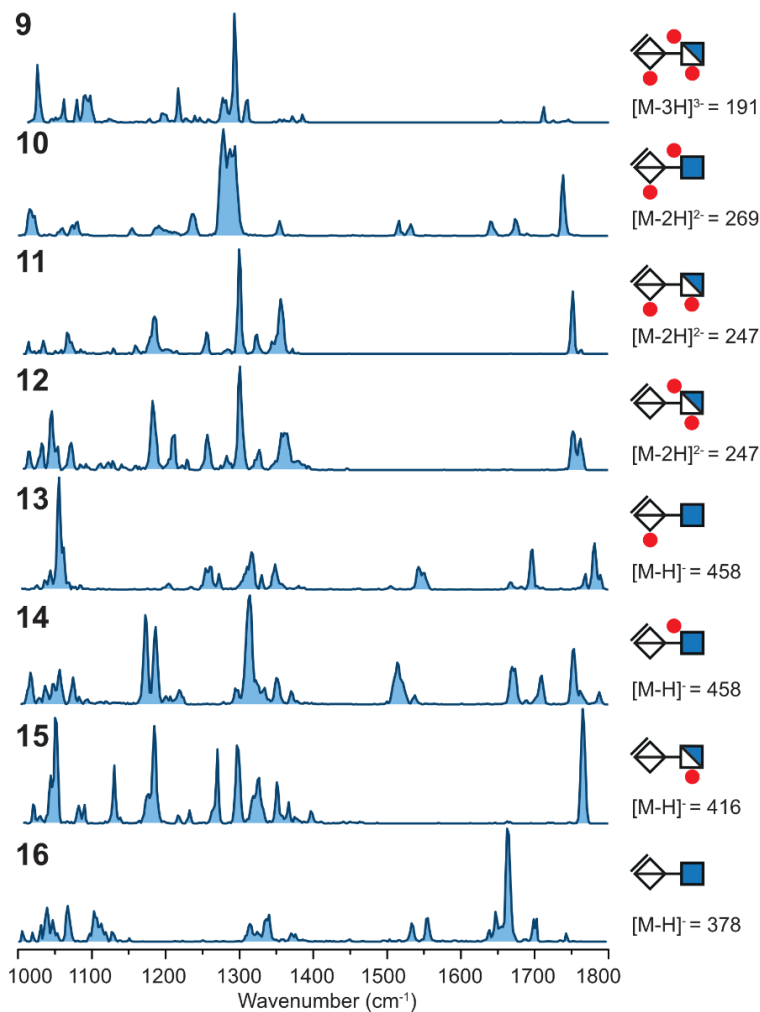


Figure 47: Cryogenic IR spectra of heparan sulfate disaccharides with varying and naturally occurring sulfation patterns. Top to bottom: the triply sulfated disaccharide **9** is investigated as $[M-3H]^{3-}$ anion, the doubly sulfated disaccharides **10–12** are investigated as $[M-2H]^{2-}$ anions, the singly sulfated **13–15** and the non-sulfated **16** disaccharides are investigated as $[M-H]^{-}$ anions. Their m/z ratios are given next to the respective IR spectrum.

Above 1700 cm^{-1} , the C=O stretching vibration comes from the neutral carbonyl functional group in the hexuronic acid moiety. In non-sulfated GAGs, the charged carboxylate functional group yields a strong peak around 1650 cm^{-1} , corresponding to the antisymmetric COO^- stretching.

All HS disaccharides **9–16** show unique IR signatures. *N*-acetylation (and therewith lacking *N*-sulfation) is easy to determine due to the presence of the amide I (1600–1700 cm^{-1}) and II (1500–1600 cm^{-1}) vibrations in the IR fingerprints of disaccharides **10**, **13**, **14**, and **16**. A silent amide region can indicate that HS is either *N*-sulfated or has a free primary amine, while the *O*-sulfation motif in these ions (2-*O* or 6-*O* sulfation) are hard to differentiate based on the IR spectra. This can later be circumvented by analyzing the 1294–1300 cm^{-1} range, where a single strong mode at $\approx 1300 \text{ cm}^{-1}$ (or somewhat red-shifted in case of high sulfation) can indicate *N*-sulfation. The number of conformers captured in the cryogenic ion trap can vary even for ions of similar chemical structure (Chapter 7). A double band in the IR spectrum of disaccharide **10** with a shift in the amide I band of $\approx 35 \text{ cm}^{-1}$ can indicate that two conformers (or anomers) were present in the trap. The number of conformers in the cryogenic ion trap for disaccharides **9**, **11**, and **15** appears to be rather low, because the neutral carboxylic acid only yields a single $\nu(\text{C}=\text{O})$ mode.

In the other seven HS disaccharides, this band shows an apparent tendency: with increasing sulfation, the $\nu(\text{C}=\text{O})$ stretching vibration red-shifts from the 1753–1781 cm^{-1} region (singly sulfated) to the 1739–1762 cm^{-1} region in the doubly-sulfated HS ions, and eventually to 1712 cm^{-1} in the triply-sulfated HS ion. This red-shift is most probably a contribution by the electron-withdrawing characteristics of sulfates, leading to a reduction in electron density around the carboxylic moiety with increasing sulfation. In the spectrum of **11**, **12**, and **14**, the $\nu(\text{C}=\text{O})$ band appears at 1753 and 1763 cm^{-1} as a double peak, while the region below 1400 cm^{-1} is more populated with (often weak) modes compared to the other spectra. This leads to two conclusions: these three ions can exist in a more extended conformational space than the other ions, while they also share a highly similar chemical environment around the carboxylic moiety.

Predicting Structural Motifs of Heparan- and Chondroitin Sulfates using Cryogenic
Infrared Spectroscopy and Random Forest

Table 5: Tentative assignment of selected vibrational modes in Figure 47. ν_a designates the (antisymmetric) stretching mode.

Frequencies (cm^{-1})	9	10	11	12	13	14	15	16
$\nu_a(\text{SO}_3^-)$	1217, 1279, 1294, 1311	1238, 1278, 1287, 1294, 1355*	1185, 1256, 1300, 1324, 1356	1183, 1212, 1257, 1283, 1301, 1327, 1361	1260, 1272, 1316, 1330, 1349*	1173, 1186, 1218, 1296, 1314, 1323, 1334, 1351*, 1370*	1185, 1217, 1233, 1271, 1297, 1326, 1351, 1367, 1397	—
amide II	—	1517, 1533	—	—	1543, 1550	1515, 1538	—	1534, 1555
amide I	—	1641, 1675	—	—	1697	1670, 1710**	—	1647, 1701
$\nu(\text{C}=\text{O})$	1713	1739	—	1753, 1762	1769, 1782, 1790	1754, 1763, 1788	1766	—
$\nu_a(\text{COO}^-)$	—	—	—	—	—	—	—	1664

*could also be amide III vibrational mode.

**could also be a $\nu(\text{C}=\text{O})$ mode of the neutral carboxylic acid.

The $\nu(\text{C}=\text{O})$ region in the spectrum of **13** contains several additional bands, with unique band positions compared to the previous case pointing towards a different chemical environment for the carboxylic moiety. This can result from the fact that the only charged sulfate in this case is located on the same

monosaccharide unit as the carboxylic group itself. The three charged sulfates in disaccharide **9** show intense $\nu_a(\text{SO}_3^-)$ vibrational features in the IR spectrum, which can overpower the $\nu(\text{C}=\text{O})$ and account for relatively low intensity. It is important to note that the relative intensities only serve as qualitative measures in helium nanodroplet spectroscopy.¹⁵⁴ Interestingly, the most intense band in the IR spectrum of disaccharide **13**, which carries 2-*O* sulfation, is at 1055 cm^{-1} . The $\nu(\text{C}-\text{O})$ and $\nu(\text{C}-\text{C})$ of the glycan core, and $\nu_s(\text{SO}_3^-)$ in this region are difficult to differentiate, therefore this strong mode remains unassigned.

The cryogenic IR spectra of HS tetrasaccharides **21** and **22** were measured in the range from 1000 to 1800 cm^{-1} and are plotted in Figure 48. The two diastereomers are differentiable from their IR signature and selected vibrational modes are summarized in Table 6.

The residual aminopentyl linker present from the chemical synthesis has additional contributions in the 1200–1500 cm^{-1} region corresponding to C–H deformation, while the scissor bends of neutral, primary amines are typically found around 1600 cm^{-1} and are of very low intensity.^{273, 274} This highly flexible linker can form interactions of different strength with several different functional groups in the molecules, which may distribute their vibrational contribution over an extended wavenumber range, thus resulting in an overall weak NH_2 signal.

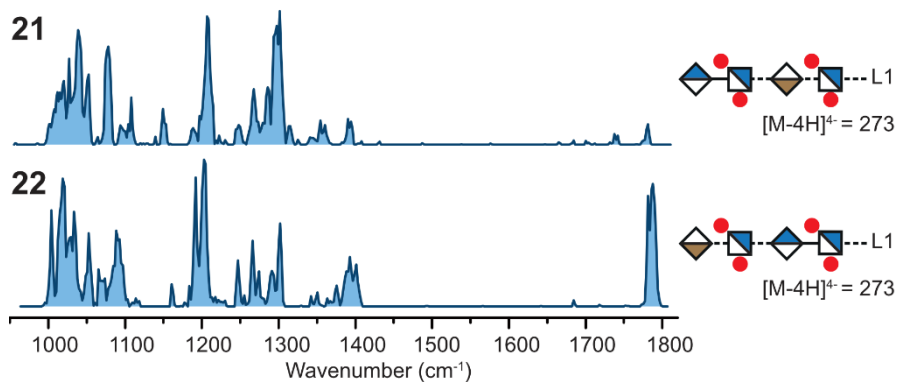


Figure 48: Cryogenic IR spectra of heparan sulfate tetrasaccharides from 1000 to 1800 cm^{-1} . The tetrasaccharides are *O*-sulfated at C6 and *N*-sulfated at every glucosamine. They are diastereomers and carry an aminoalkyl-linker (L1) from synthesis.

To date, the pentasaccharide Fondaparinux was the longest GAG investigated using IR spectroscopy, see Chapter 5. Here, the cryogenic IR spectrum of the hexasaccharide **23** was recorded and is shown in Figure 49. The heparan sulfate hexasaccharide carries three negatively charged sulfate groups and a residual para-methoxyphenyl linker from chemical synthesis. The band positions of selected vibrational features are summarized in Table 6.

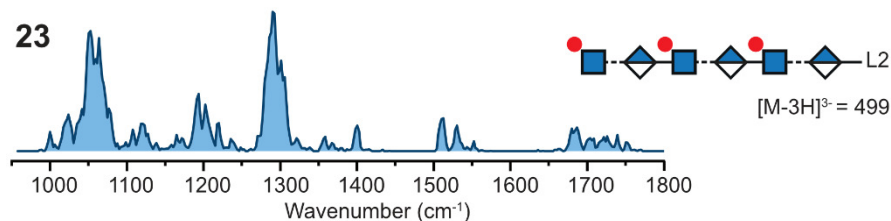


Figure 49: Cryogenic IR spectra of a heparan sulfate hexasaccharide from 1000 to 1800 cm^{-1} . The hexasaccharide is *O*-sulfated at C6 at every glucosamine and carries a para-methoxyphenyl linker (L2) from their chemical synthesis.

Table 6: Tentative assignment of selected vibrational modes in Figure 48 and Figure 49. ν_a designates the (antisymmetric) stretching mode.

Frequencies (cm^{-1})	21	22	23
$\nu_a(\text{SO}_3^-)$	1150, 1207, 1267, 1286, 1296, 1301, 1357, 1393	1160, 1191, 1203, 1247, 1266, 1274, 1292, 1302, 1392, 1400	1192, 1202, 1290, 1301, 1400*
amide II	—	—	1511, 1530
amide I	—	—	1684, 1704**
$\nu(\text{C}=\text{O})$	1739, 1780	1781, 1787	1725, 1740

*could also be amide III vibrational mode.

**could also be a $\nu(\text{C}=\text{O})$ mode of the neutral carboxylic acid.

Random Forest

To optimize the feature-to-sample ratio for the present study with a comparably small data set, spectral binning and subsequent feature selection are applied to reduce the feature space, i.e. the intensity values measured at each wavenumber step. In IR spectroscopy, neighbouring IR channels are typically strongly correlated, thus, requiring additional feature engineering beyond spectral binning.

Predicting Structural Motifs of Heparan- and Chondroitin Sulfates using Cryogenic Infrared Spectroscopy and Random Forest

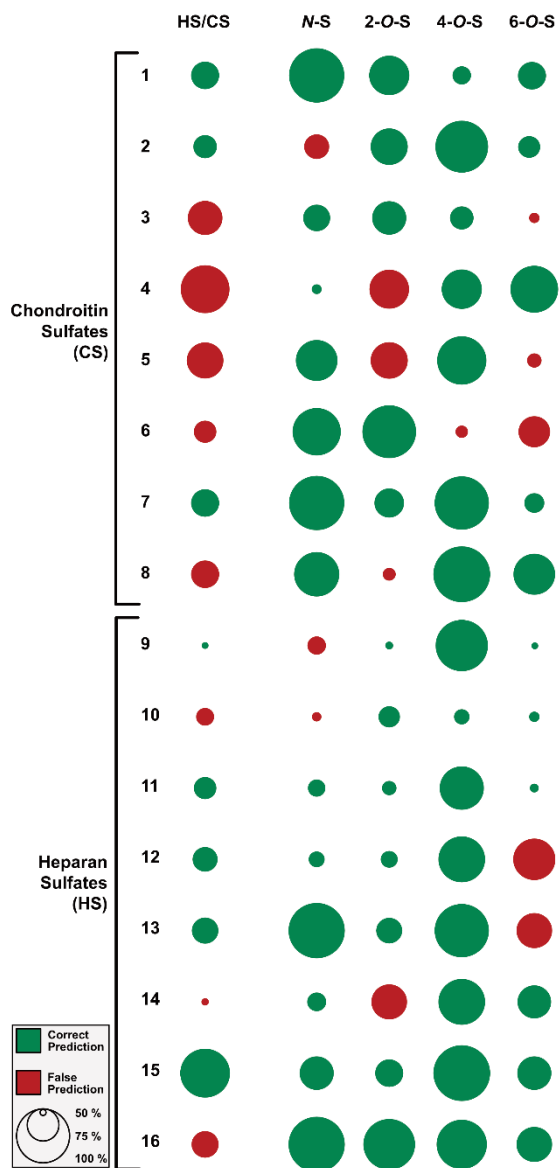


Figure 50: Prediction outcome and confidence for the disaccharide classifications in the complete combinatorial space excluding one sample from the base set. The features included here are HS/CS, and the presence or absence of N-, 2-O-, 4-O-, and 6-O-sulfation (N-S, 2-O-S, 4-O-S, and 6-O-S, respectively). High prediction score was achieved in classification tasks for 4-O-S and N-S with 94% and 81%, respectively. Lower prediction score was achieved for 2-O-S and 6-O-S and HS versus CS classification, with 73%, 63% and 50%, respectively. The prediction confidence is given by the size of the dot and is generally lower for false predictions.

In the initial step, the training set contains 15 (of total 16) disaccharides to evaluate the classification outcome in predicting the structure (HS or CS), and the presence or absence of N-, 2-O-, 4-O-, and 6-O-sulfation for the remaining disaccharide, see Figure 48. In total, the models predicted the correct structural motifs in 73% of the cases and the prediction confidence was generally lower for false predictions. In agreement with the in-depth analysis of the IR spectra, the prediction score (or prediction accuracy) for N-sulfation is with 81% generally high due to the absence of amide vibrations. Only the classification of 4-O-S achieved a higher prediction score with 94%. All disaccharides were correctly classified except for disaccharide **6** which was falsely labelled as non-4-O sulfated molecule. Lower prediction scores were observed in predicting 2-O-S and 6-O-S with 73% and 63%, respectively. The decision rules for the distinction of HS and CS, which can have different structural motifs in the set, are more complex and here, the highest number of features were selected for classification. Still, the prediction score of only 50% shows that the classification for HS/CS remains inconclusive. Especially for the CS disaccharides, the distinction failed more often. Disaccharides **3–6** and **8** were falsely labelled as HS disaccharides while disaccharides **10**, **14** and **16** were incorrectly assigned as CS disaccharides.

In the next step, the IR spectra of the tetrasaccharides **17–22** were systematically added to the training set to test the influence of the training set size on the prediction score. With increasing training set size, the prediction score strongly improved and scored 97% with 21 analytes in the training set. The tetrasaccharides are not only different in their GAG chain length but also carry a linker from their chemical synthesis (aminoalkyl-linker, L_1), introducing additional vibrational contributions to regions that were previously not affected by the presence of a linker, therefore challenging the robustness of the algorithm. With a complete training set of 21 di- and tetrasaccharides, the remaining tetrasaccharide was used for the prediction of structural motifs. The results are shown in Figure 51. With one exception, i.e. 6-O sulfation in tetrasaccharide **19**, all structural motifs were predicted correctly across the excluded samples. Tetrasaccharide **19** is structurally related to samples **17**, **18** and **20**, see Chapter 6, which are exclusively sulfated at 6-O. Especially in the

wavenumber region above 1450 cm^{-1} , the IR spectra of **19** and **20** share selected vibrational features and reveal significant differences to the IR spectra of **17** and **18**. These vibrational features were associated to the terminal motif IdoA-GlcNAc6S-L1 present only in **19** and **20**. For tetrasaccharide **19** (and in contrast to tetrasaccharide **20**), many wavenumber bins of the random forest classifier for 6-O sulfation are located in the higher wavenumber region, most likely introducing ambiguity to the classification.

Finally, a hexasaccharide with a different linker from chemical synthesis (para-methoxyphenyl linker, L2) was tested against the training set including 16 disaccharides and 6 tetrasaccharides. The results are shown in Figure 51, last row. With a high confidence, all structural motifs in the triply sulfated hexasaccharide are predicted correctly only from the IR signature.

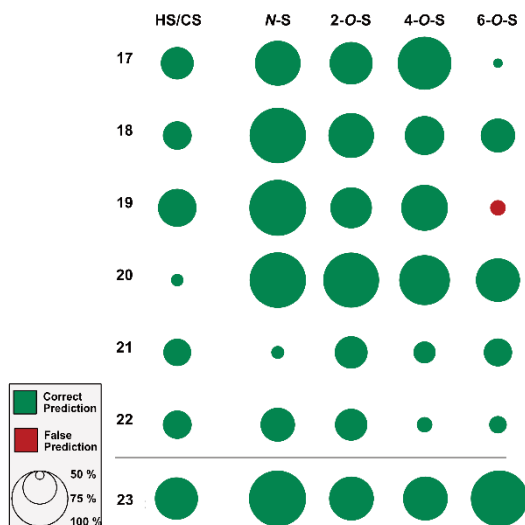


Figure 51: Prediction outcome and confidence for the tetrasaccharide and hexasaccharide classifications. The hexasaccharide **23** is only used for classification and is not part of the training set for the tetrasaccharides. In each classification task, the respective sample is excluded from the training set. The training set proves sufficient to make high-confidence predictions about structural motifs in unknown GAGs.

8.4. Conclusion

Here, the gas-phase cryogenic IR spectra of a comprehensive set of GAGs were recorded and used in random forest classifications for predicting the class (HS/CS) and sulfate positions (2-*O*-, 4-*O*-, 6-*O*-, *N*-sulfation). The library of IR spectra is sufficiently complex for a proof-of-concept study to combine cryogenic IR spectroscopy and random forest modelling, yet small enough to justify the measurement time at the FHI FEL, until tunable table-top systems become more available in the mid-IR range.²⁷⁵ The chemical features included nearly fully characterize the underlying chemical structure, with exception to modifications among domains and acetylation (i.e. GlcNAc versus GlcN). While some structural motifs can be directly identified from the IR fingerprint, machine learning tools could automate the identification. It was shown that already a small random forest training set with 21 samples allows to predict GAG structural motifs with high confidence in tetra- and hexasaccharides. The results demonstrate that random forest modeling can circumvent the need of large databases to match IR spectra of GAGs for structural identification and automate the interpretation of complex signatures.

Summary

Unusual monosaccharides (fucose), covalent modifications of glycans (sulfation) and terminal sequences play important biological roles in physiology and pathology of living organisms. Furthermore, in an evolutionary sense, uncommon structures are often the result of selection pressures and can be the source to a deeper understanding of the evolution of glycosylation.¹⁵⁷ At the same time, fucosylated glycans and sulfated glycans still challenge standard mass spectrometry (MS)-based analytical workflows in glycan analysis. MS emerged throughout the last decade as the most widely used analytical technique in glycan analysis. As a stand-alone technique, it is limited in glycan analysis due to the presence of isomers. Isomerism in glycans arises from their composition, connectivity, configuration, and branching. Therefore, MS is often coupled to orthogonal techniques such as liquid chromatography (LC) and ion mobility spectrometry (IM-MS). Most recently, the combination of cryogenic IR spectroscopy in the gas phase with MS proved beneficial for the identification of smaller glycans. At low measurement temperatures, the IR spectrum of small glycans provides a unique fingerprint to the underlying chemical structure and conformation.

In this thesis, cryogenic IR spectroscopy as an addition to the MS-based analytical toolbox was used to shed light on the migration of fucose residues in MS experiments. This elusive rearrangement reaction is not restricted to tandem MS workflows but is recently found to occur in intact ions without extensive activation. Here, the role of the proton in fucose migration reactions was investigated for the two glycan epitopes Lewis x and blood group H type 2. A systematic study of adduct ions and functional groups with competing proton affinities demonstrated that the proton can be selectively mobilized and demobilized. Planning MS-based experiments of fucosylated glycan cations certainly needs an effective strategy to circumvent the presence of a mobile proton in order to avoid erroneous sequence assignments.

In a multidimensional approach, IR spectroscopy, IM-MS, RDD and computational modelling were combined to decode the rearrangement product

and the reaction mechanism. The trisaccharides Lewis x and blood group H type 2 were found to migrate to a third chemical structure, in which the fucose moiety is most likely 1,6-linked to galactose. The barrier is much higher for blood group H type 2 compared to Lewis x and it is feasible that the latter is never detected in its original chemical structure in the mass spectrometer. These results generalize fucose migration to a universal issue in any mass spectrometer to which even various orthogonal MS-based techniques can be blind.

In the second part of this thesis, cryogenic IR spectroscopy in combination with computational modelling was employed for the structural analysis of sulfated glycosaminoglycans (GAGs). Diversity in the chemical structure of linear and acidic GAGs arises from the GAG class, sulfation, epimerization and acetylation. Using messenger tagging IR spectroscopy, sulfated mono- and disaccharides have been characterized successfully recently. In the present thesis, the prominent anticoagulant pentasaccharide fondaparinux which carries eight sulfate functional groups was investigated using cryogenic IR spectroscopy in helium nanodroplets as a proof-of-concept. The spectroscopic fingerprint features unique absorption bands in the mid-IR range for the sulfate functional groups. With this knowledge, a systematic set of all naturally occurring sulfation variations in chondroitin and dermatan sulfate (CS/DS) further demonstrated the capabilities of cryogenic IR spectroscopy for their differentiation. Moreover, from their IR fingerprints in combination with computational modelling, conformational diversity arising from sulfation and charge density distribution could be derived. In a different study, the IR fingerprints of four heparan sulfate (HS) diastereomers revealed a modularity in their chemical structure which was explained, using computational modelling, from their unique hydrogen bonding patterns. The knowledge of the preferred hydrogen bonding pattern could aid e.g. the development for labelling strategies in IM-MS. The results show that the high resolution in the optical fingerprints of GAGs allows to unambiguously resolve their diversity arising from GAG class, sulfation and epimerization. The results exemplify the importance of gas-phase cryogenic IR spectroscopy to enhance future analytical workflows for GAG sequencing. A fully MS-based workflow could involve the ionization of an intact GAG chain and combine tandem MS with IM-MS and cryogenic IR

spectroscopy of respective fragments to unambiguously characterize a GAG chain in a single MS experiment.

In the last part, cryogenic IR spectroscopy was combined with random forest modelling to extract vibrational features that are characteristic to structural features in GAGs. The selected structural features included the GAG class and sulfation and therefore, almost fully characterize the underlying chemical structure. In a proof-of-concept study, a prediction score of >97% could be achieved for HS tetra- and hexasaccharides based on a training set of only 21 spectra. Especially for certain marker motifs, such as 3-O-sulfation in cancer cells, this workflow could prove beneficial. With machine learning algorithms, the need for comprehensive spectral databases could be circumvented for the identification of unknowns. Overall, the results show that MS-based IR spectroscopy certainly has the potential to leave the framework of academic basic research and add as a valuable addition to the MS-based analytical toolbox.

Samenvatting

Weinig voorkomende monosachariden (fucose), covalente modificaties van glycanen (sulfering) en terminale sequenties spelen belangrijke rollen in de fysiologie en pathologie van levende organismen. Weinig voorkomende structuren zijn in evolutionaire zin vaak het resultaat van selectiedruk en kunnen derhalve een dieper inzicht leveren in de evolutie van glycosylering. Gefucosyleerde glycanen en gesulfoneerde glycanen vormen echter nog steeds een uitdaging voor standaard workflows in glycaananalyse. Massaspectrometrie (MS) heeft zich in het laatste decennium ontwikkeld tot de meest gebruikte techniek voor glycaananalyse, maar is beperkt door de aanwezigheid van isomeren. Isomeren van glycanen zijn het gevolg van hun samenstelling, connectiviteit, configuratie en vertakking. MS wordt daarom vaak gekoppeld aan complementaire technieken zoals vloeistofchromatografie (LC) en ion-mobiliteitsspectrometrie (IM-MS). Gedurende de laatste jaren is de combinatie van cryogene infrarood (IR)-spectroscopie in de gasfase met MS van grote waarde gebleken voor de identificatie van kleinere glycanen. Bij lage meettemperaturen geeft het IR spectrum van kleine glycanen een unieke vingerafdruk van de onderliggende chemische structuur en conformatie.

In dit proefschrift is cryogene IR-spectroscopie in combinatie met MS-gebaseerde analytische technieken gebruikt om licht te werpen op de migratie van fucose in MS-experimenten. Deze ongrijpbare migratiereactie is niet beperkt tot tandem MS workflows, maar is recentelijk ook waargenomen in intacte ionen zonder uitgebreide activering. De rol van het proton in fucose-migratiereacties is onderzocht voor de twee glycaanepitopen Lewis x en bloedgroep H type 2. In een systematische studie van adductie-ionen en functionele groepen met concurrerende protonaffiniteiten is aangetoond dat het proton selectief gemobiliseerd en gedemobiliseerd kan worden. Het meten van gefucosyleerde glycaan-kationen met MS vereist een effectieve strategie om de aanwezigheid van een mobiel proton te omzeilen om foutieve sequentie-toewijzingen te voorkomen.

In een multidimensionele benadering zijn IR spectroscopie, IM-MS, radical-directed dissociation (RDD) MS en computationele modellering gecombineerd

om het migratieproduct en het reactiemechanisme te ontcijferen. De trisachariden Lewis x en bloedgroep H type 2 blijken te migreren naar een chemische structuur, waarin fucose hoogstwaarschijnlijk 1,6-gekoppeld is aan galactose. De barrière is veel hoger voor bloedgroep H type 2 dan voor Lewis x en het is goed mogelijk dat de laatste nooit in zijn oorspronkelijke chemische structuur gedetecteerd is in de massaspectrometer. Uit deze resultaten blijkt dat fucose-migratie een universeel probleem is in elke massaspectrometer en dat ook het gebruik van verschillende complementaire MS-gebaseerde technieken dit probleem niet geheel kan oplossen.

In het tweede deel van dit proefschrift is cryogene IR spectroscopie in combinatie met computationele modellering gebruikt voor de structurele analyse van gesulfoneerde glycosaminoglycanen (GAG's). De verscheidenheid in de chemische structuur van lineaire zure GAG's komt voort uit de GAG klasse, sulfatie, epimerisatie en acetylatie. Met behulp van messenger tagging IR spectroscopie zijn recentelijk met succes gesulfoneerde mono- en disachariden gekarakteriseerd. In dit proefschrift is het anticoagulant pentasaccharide fondaparinux, dat acht sulfaatgroepen bevat, onderzocht met behulp van cryogene IR spectroscopie in helium nanodruppels om het werkingsprincipe van de meting aan te tonen. De spectroscopische vingerafdruk toont unieke absorptiebanden in het midden-IR bereik voor de sulfaatgroepen. Het meten van een systematische set van alle natuurlijk voorkomende sulfatievariëaties in chondroïtine- en dermatan-sulfaat (CS/DS) heeft de differentiatie mogelijkheden met behulp van cryogene IR spectroscopie verder aangetoond. Uit de IR-vingerafdruk in combinatie met computationele modellering kan bovendien conformationele diversiteit als gevolg van sulfatie en ladingsdichtheidsverdeling worden afgeleid. In een andere studie onthullen de IR-vingerafdrukken van vier heparansulfaat (HS) diastereomeren een modulariteit in hun chemische structuur die verklaard is met behulp van computationele modellering door hun unieke waterstofbrugpatronen. De kennis van het geprefereerde waterstofbindingspatroon zou bijvoorbeeld kunnen helpen bij de ontwikkeling van labelingstrategieën in IM-MS. De resultaten laten zien dat de hoge resolutie in de optische vingerafdrukken van GAG's het mogelijk maakt om eenduidig de diversiteit op te lossen die

voortkomt uit GAG klasse, sulfatie en epimerisatie. De resultaten illustreren het belang van gas-fase cryogene IR spectroscopie om toekomstige analytische workflows voor GAG sequencing te verbeteren. Een volledig op MS gebaseerde workflow zou de ionisatie van een intacte GAG-keten kunnen omvatten en tandem MS met IM-MS en cryogene IR-spectroscopie van de respectieve fragmenten kunnen combineren om een GAG-keten eenduidig te karakteriseren in één enkel MS-experiment.

In het laatste deel van het proefschrift is cryogene IR-spectroscopie gecombineerd met random forest modellering om vibratie patronen die kenmerkend zijn voor structurele eigenschappen in GAG's aan te tonen. De geselecteerde structurele eigenschappen omvatten de GAG-klasse en sulfatie en karakteriseren derhalve bijna volledig de onderliggende chemische structuur. In een proof-of-concept studie is een voorspellingscore van >97% bereikt voor HS tetra- en hexasachariden op basis van een trainingsset van slechts 21 spectra. Vooral voor bepaalde markermotieven, zoals 3-O-sulfatie in kankercellen, zou deze workflow nuttig kunnen blijken. Met algoritmen voor machine learning zou de noodzaak voor het gebruik van uitgebreide spectrale databanken voor de identificatie van onbekende GAG's kunnen worden omzeild. Concluderend kan gesteld worden dat de resultaten zoals beschreven in dit proefschrift aantonen dat IR-spectroscopie op basis van MS zeker het potentieel heeft om het stadium van het academisch basisonderzoek te verlaten en een waardevolle aanvulling vormt op MS gebaseerde analytische technieken.

Research Data Management

This thesis has been carried out in accordance with the research data management policy of the High Field Magnet Laboratory & FELIX Laboratory (HFML-FELIX), Radboud University, the Netherlands.

Data sets for the results presented in this thesis are deposited in Zenodo, part of the CERN's Data Centre, as a data sharing collection with a unique digital object identifier (DOI) 10.5281/zenodo.6881155. It can be accessed via <https://doi.org/10.5281/zenodo.6881155>.

Data sets available are:

- Experimental IR spectral data as xy-files.
- Experimental IM-MS data as xy-files.

Data sets exist for the Chapters 3, 4, 5, 6, 7 and 8.

Appendices

Appendix to Chapter 2

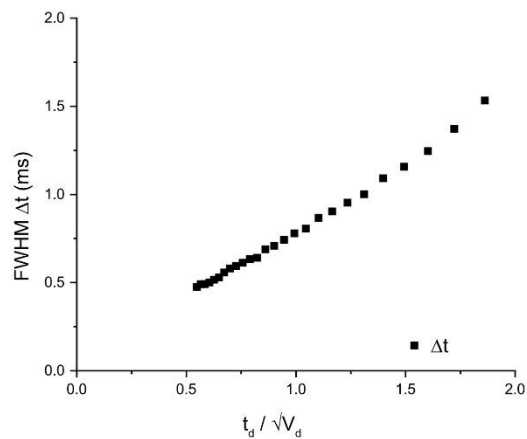


Figure 52: ATD peak width Δt as a function of $t_d / \sqrt{V_d}$.

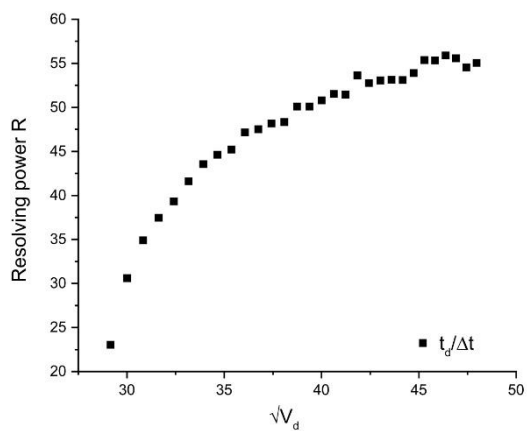


Figure 53: IMS resolving power R as a function of $\sqrt{V_d}$.

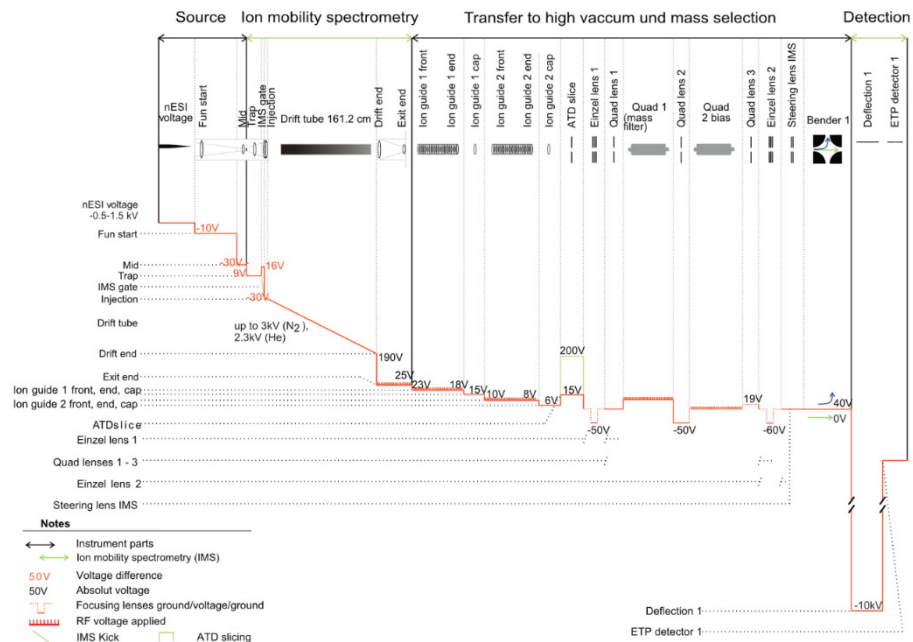


Figure 54: Voltage diagram for the DTIM-MS part of the experiment described in Section 2.2.

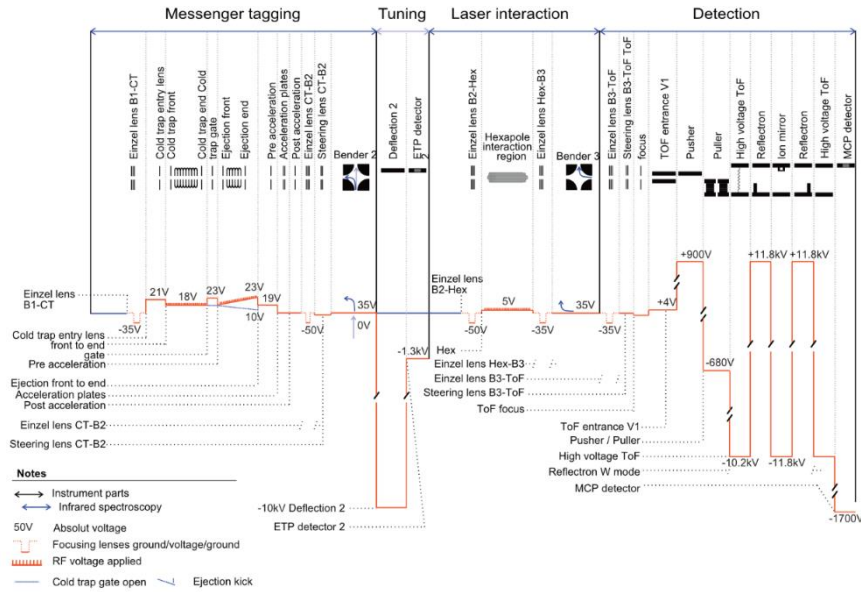


Figure 55: Voltage diagram for the spectroscopy part of the experiment described in Section 2.2.

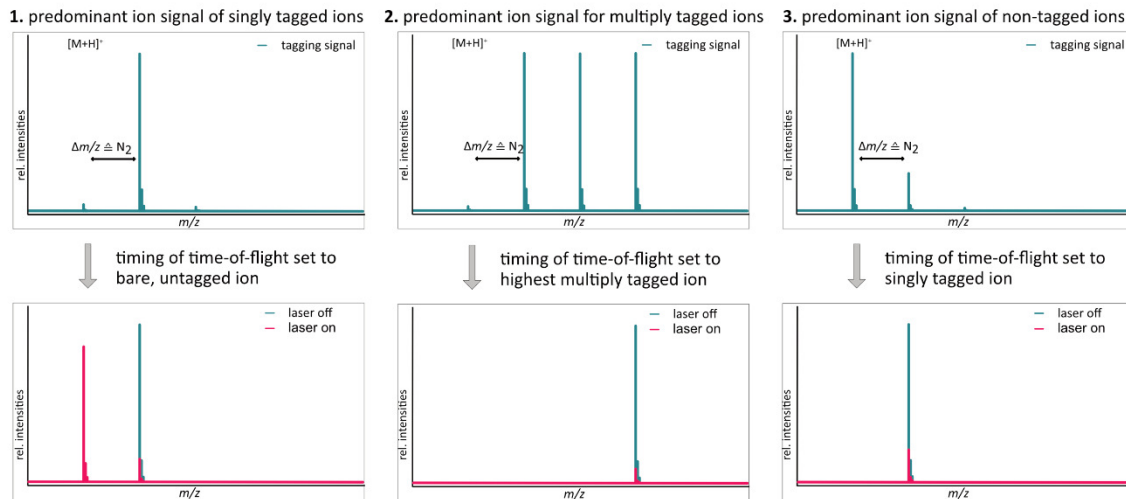
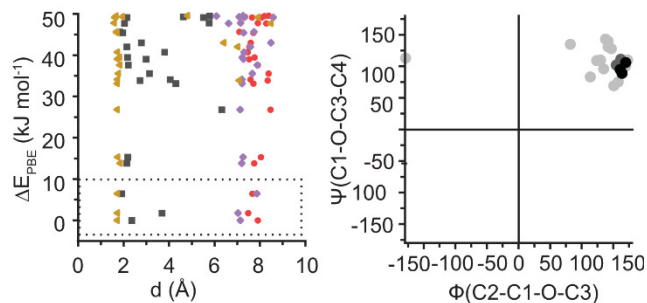
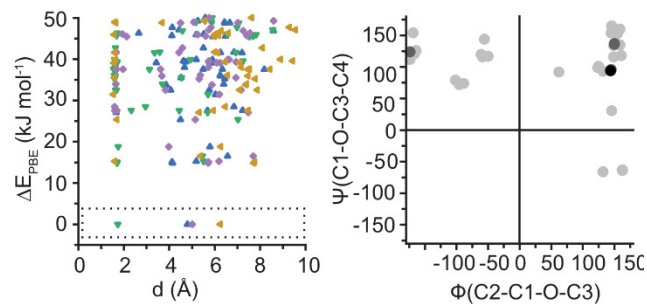


Figure 56: Mode of operation of the experiment described in Section 2.2.

Appendix to Chapter 7

A. $\Delta\text{UA}2\text{S}\beta\text{1-3GalNAc}6\text{S } 3$ B. $\Delta\text{UA}\beta\text{1-3GalNAc}4\text{S}6\text{S } 4$ 

■ 2S – HN 2S ● 6S – HN 6S
 ● – HOOC ● – HOOC

▲ 4S – HN 4S
 ▼ – HOOC

⋯⋯⋯⋯ lowest energy conformation(s)

■ 0 < 10 kJ mol⁻¹
 ■ 10 < 20 kJ mol⁻¹
 ■ 20 < 50 kJ mol⁻¹
 ■ > 50 kJ mol⁻¹

Figure 57: **Left panel:** calculation of the intramolecular distance of selected functional groups in Å in conformers with relative energies $\Delta E_{\text{PBE}} < 50$ kJ mol⁻¹. **Right panel:** calculation of the dihedral angles $\Psi(\text{C1-O-C3-C4})$ and $\Phi(\text{C2-C1-O-C3})$ in the degree at the glycosidic bond with respect to the relative energies of the conformers. The results are presented in Ramachandran-type plots for glycosidic linkages. Results for disaccharide **3** (A) and disaccharide **4** (B).

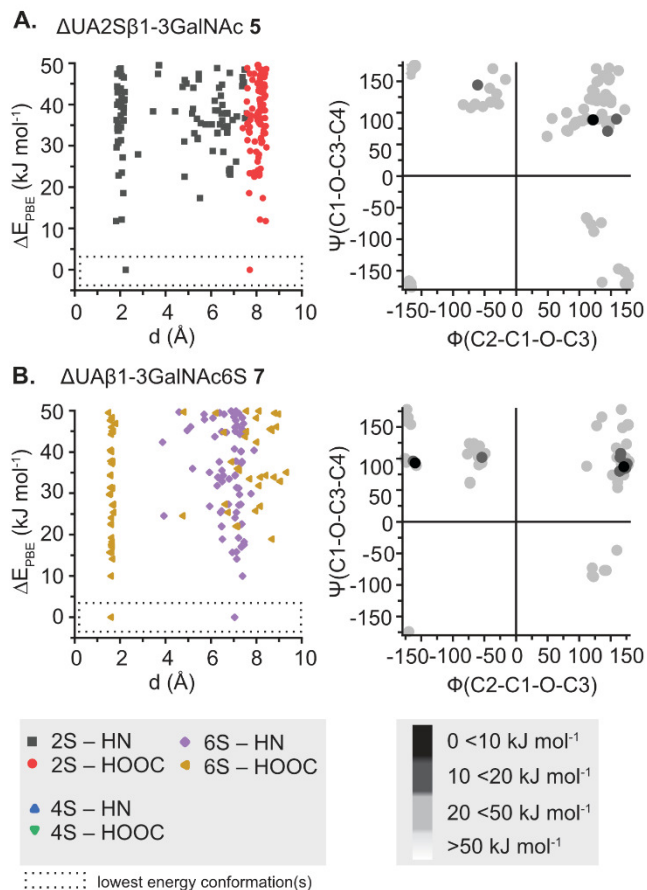


Figure 58: **Left panel:** calculation of the intramolecular distance of selected functional groups in Å in conformers with relative energies $\Delta E_{\text{PBE}} < 50 \text{ kJ mol}^{-1}$. **Right panel:** calculation of the dihedral angles $\Psi(\text{C1-O-C3-C4})$ and $\Phi(\text{C2-C1-O-C3})$ in the degree at the glycosidic bond with respect to the relative energies of the conformers. The results are presented in Ramachandran-type plots for glycosidic linkages. Results for disaccharide **5** (A) and disaccharide **7** (B).

References

1. Varki, A., Biological roles of glycans. *Glycobiology* **2017**, *27* (1), 3-49.
2. Dwek, R. A., Glycobiology: Toward Understanding the Function of Sugars. *Chem. Rev.* **1996**, *96* (2), 683-720.
3. Flynn, R. A.; Pedram, K.; Malaker, S. A.; Batista, P. J.; Smith, B. A. H.; Johnson, A. G.; George, B. M.; Majzoub, K.; Villalta, P. W.; Carette, J. E.; Bertozzi, C. R., Small RNAs are modified with N-glycans and displayed on the surface of living cells. *Cell* **2021**, *184* (12), 3109-3124.e22.
4. Petrescu, A.-J.; Wormald, M. R.; Dwek, R. A., Structural aspects of glycomes with a focus on N-glycosylation and glycoprotein folding. *Curr. Opin. Struct. Biol.* **2006**, *16* (5), 600-607.
5. Shental-Bechor, D.; Levy, Y., Folding of glycoproteins: toward understanding the biophysics of the glycosylation code. *Curr. Opin. Struct. Biol.* **2009**, *19* (5), 524-533.
6. Tulsiani, D. R. P.; Yoshida-Komiya, H.; Araki, Y., Mammalian Fertilization: A Carbohydrate-Mediated Event1. *Biol. Reprod.* **1997**, *57* (3), 487-494.
7. Snell, W. J.; White, J. M., The Molecules of Mammalian Fertilization. *Cell* **1996**, *85* (5), 629-637.
8. Lepedda, A. J.; De Muro, P.; Capobianco, G.; Formato, M., Role of the small proteoglycan bikunin in human reproduction. *Hormones* **2020**, *19* (2), 123-133.
9. Marth, J. D.; Grewal, P. K., Mammalian glycosylation in immunity. *Nat. Rev. Immunol.* **2008**, *8* (11), 874-887.
10. Rudd, P. M.; Elliott, T.; Cresswell, P.; Wilson, I. A.; Dwek, R. A., Glycosylation and the Immune System. *Science* **2001**, *291* (5512), 2370-2376.
11. Reily, C.; Stewart, T. J.; Renfrow, M. B.; Novak, J., Glycosylation in health and disease. *Nat. Rev. Nephrol.* **2019**, *15* (6), 346-366.
12. Kamhi, E.; Joo, E. J.; Dordick, J. S.; Linhardt, R. J., Glycosaminoglycans in infectious disease. *Biol. Rev.* **2013**, *88* (4), 928-943.

13. Fuster, M. M.; Esko, J. D., The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat. Rev. Cancer* **2005**, *5* (7), 526-542.
14. Pinho, S. S.; Reis, C. A., Glycosylation in cancer: mechanisms and clinical implications. *Nat. Rev. Cancer* **2015**, *15* (9), 540-555.
15. Shajahan, A.; Supekar, N. T.; Gleinich, A. S.; Azadi, P., Deducing the N- and O-glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology* **2020**, *30* (12), 981-988.
16. Watanabe, Y.; Berndsen, Z. T.; Raghwani, J.; Seabright, G. E.; Allen, J. D.; Pybus, O. G.; McLellan, J. S.; Wilson, I. A.; Bowden, T. A.; Ward, A. B.; Crispin, M., Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. *Nat. Commun.* **2020**, *11* (1), 2688.
17. Kostopoulos, I.; Elzinga, J.; Ottman, N.; Klievink, J. T.; Blijenberg, B.; Aalvink, S.; Boeren, S.; Mank, M.; Knol, J.; de Vos, W. M.; Belzer, C., Akkermansia muciniphila uses human milk oligosaccharides to thrive in the early life conditions in vitro. *Sci. Rep.* **2020**, *10* (1), 14330.
18. Gotoh, A.; Katoh, T.; Sakanaka, M.; Ling, Y.; Yamada, C.; Asakuma, S.; Urashima, T.; Tomabeche, Y.; Katayama-Ikegami, A.; Kurihara, S.; Yamamoto, K.; Harata, G.; He, F.; Hirose, J.; Kitaoka, M.; Okuda, S.; Katayama, T., Sharing of human milk oligosaccharides degradants within bifidobacterial communities in faecal cultures supplemented with Bifidobacterium bifidum. *Sci. Rep.* **2018**, *8* (1), 13958.
19. Bode, L., Human milk oligosaccharides: every baby needs a sugar mama. *Glycobiology* **2012**, *22* (9), 1147-62.
20. Grabarics, M.; Csernak, O.; Balogh, R.; Beni, S., Analytical characterization of human milk oligosaccharides - potential applications in pharmaceutical analysis. *J. Pharm. Biomed. Anal.* **2017**, *146*, 168-178.
21. Werz, D. B.; Ranzinger, R.; Herget, S.; Adibekian, A.; von der Lieth, C.-W.; Seeberger, P. H., Exploring the Structural Diversity of Mammalian Carbohydrates (“Glycospace”) by Statistical Databank Analysis. *ACS Chem. Biol.* **2007**, *2* (10), 685-691.
22. 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. M., Many human milk oligosaccharides are sialylated. *J. Biol. Chem.* **2005**, *280* (12), 11730-11736.

- 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.
23. Schneider, M.; Al-Shareffi, E.; Haltiwanger, R. S., Biological functions of fucose in mammals. *Glycobiology* **2017**, *27* (7), 601-618.
24. Becker, D. J.; Lowe, J. B., Fucose: biosynthesis and biological function in mammals. *Glycobiology* **2003**, *13* (7), 41R-53R.
25. Pang, P.-C.; Chiu, P. C. N.; Lee, C.-L.; Chang, L.-Y.; Panico, M.; Morris, H. R.; Haslam, S. M.; Khoo, K.-H.; Clark, G. F.; Yeung, W. S. B.; Dell, A., Human Sperm Binding Is Mediated by the Sialyl-Lewis^x Oligosaccharide on the Zona Pellucida. *Science* **2011**, *333* (6050), 1761-1764.
26. Soares da Costa, D.; Reis, R. L.; Pashkuleva, I., Sulfation of Glycosaminoglycans and Its Implications in Human Health and Disorders. *Annu. Rev. Biomed. Eng.* **2017**, *19* (1), 1-26.
27. Gama, C. I.; Tully, S. E.; Sotogaku, N.; Clark, P. M.; Rawat, M.; Vaidehi, N.; Goddard, W. A.; Nishi, A.; Hsieh-Wilson, L. C., Sulfation patterns of glycosaminoglycans encode molecular recognition and activity. *Nat. Chem. Biol.* **2006**, *2* (9), 467-473.
28. Townley, R. A.; Bülow, H. E., Deciphering functional glycosaminoglycan motifs in development. *Curr. Opin. Struct. Biol.* **2018**, *50*, 144-154.
29. Chan, W. K.; Howe, K.; Clegg, J. M.; Guimond, S. E.; Price, D. J.; Turnbull, J. E.; Pratt, T., 2-O Heparan Sulfate Sulfation by Hs2st Is Required for Erk/Mapk Signalling Activation at the Mid-Gestational Mouse Telencephalic Midline. *PLOS ONE* **2015**, *10* (6), e0130147.
30. Rudd, T. R.; Uniewicz, K. A.; Ori, A.; Guimond, S. E.; Skidmore, M. A.; Gaudesi, D.; Xu, R.; Turnbull, J. E.; Guerrini, M.; Torri, G.; Siligardi, G.; Wilkinson, M. C.; Fernig, D. G.; Yates, E. A., Comparable stabilisation, structural changes and activities can be induced in FGF by a variety of HS and non-GAG analogues: implications for sequence-activity relationships. *Org. Biomol. Chem.* **2010**, *8* (23), 5390-5397.

31. Meneghetti, M. C. Z.; Hughes, A. J.; Rudd, T. R.; Nader, H. B.; Powell, A. K.; Yates, E. A.; Lima, M. A., Heparan sulfate and heparin interactions with proteins. *J. R. Soc. Interface* **2015**, *12* (110), 0589-0589.
32. Xu, D.; Esko, J. D., Demystifying Heparan Sulfate–Protein Interactions. *Annu. Rev. Biochem.* **2014**, *83* (1), 129-157.
33. Ruhaak, L. R.; Xu, G.; Li, Q.; Goonatilake, E.; Lebrilla, C. B., Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses. *Chem. Rev.* **2018**, *118* (17), 7886-7930.
34. Wuhler, M., Glycomics using mass spectrometry. *Glycoconj. J.* **2013**, *30* (1), 11-22.
35. Kirschbaum, C.; Greis, K.; Mucha, E.; Kain, L.; Deng, S.; Zappe, A.; Gewinner, S.; Schöllkopf, W.; von Helden, G.; Meijer, G.; Savage, P. B.; Marianski, M.; Teyton, L.; Pagel, K., Unravelling the structural complexity of glycolipids with cryogenic infrared spectroscopy. *Nat. Commun.* **2021**, *12* (1), 1201.
36. Riley, N. M.; Hebert, A. S.; Westphall, M. S.; Coon, J. J., Capturing site-specific heterogeneity with large-scale N-glycoproteome analysis. *Nat. Commun.* **2019**, *10* (1), 1311.
37. Pabst, M.; Altmann, F., Glycan analysis by modern instrumental methods. *Proteomics* **2011**, *11* (4), 631-643.
38. Riley, N. M.; Bertozzi, C. R.; Pitteri, S. J., A Pragmatic Guide to Enrichment Strategies for Mass Spectrometry–Based Glycoproteomics. *Mol. Cell Proteomics* **2021**, *20*, 100029.
39. Ahn, Y. H.; Kim, J. Y.; Yoo, J. S., Quantitative mass spectrometric analysis of glycoproteins combined with enrichment methods. *Mass Spectrom. Rev.* **2015**, *34* (2), 148-165.
40. Ruhaak, L. R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A. M.; Wuhler, M., Glycan labeling strategies and their use in identification and quantification. *Anal. Bioanal. Chem.* **2010**, *397* (8), 3457-3481.
41. Zaia, J., On-line separations combined with MS for analysis of glycosaminoglycans. *Mass Spectrom. Rev.* **2009**, *28* (2), 254-272.
42. Zaia, J., Glycosaminoglycan glycomics using mass spectrometry. *Mol. Cell Proteomics* **2013**, *12* (4), 885-92.

43. Moran, A. B.; Gardner, R. A.; Wuhrer, M.; Lageveen-Kammeijer, G. S. M.; Spencer, D. I. R., Sialic Acid Derivatization of Fluorescently Labeled N-Glycans Allows Linkage Differentiation by Reversed-Phase Liquid Chromatography–Fluorescence Detection–Mass Spectrometry. *Anal. Chem.* **2022**, *94* (18), 6639-6648.
44. Lu, G.; Crihfield, C. L.; Gattu, S.; Veltri, L. M.; Holland, L. A., Capillary Electrophoresis Separations of Glycans. *Chem. Rev.* **2018**, *118* (17), 7867-7885.
45. Zauner, G.; Deelder, A. M.; Wuhrer, M., Recent advances in hydrophilic interaction liquid chromatography (HILIC) for structural glycomics. *Electrophoresis* **2011**, *32* (24), 3456-3466.
46. Ashwood, C.; Pratt, B.; MacLean, B. X.; Gundry, R. L.; Packer, N. H., Standardization of PGC-LC-MS-based glycomics for sample specific glycotyping. *Analyst* **2019**, *144* (11), 3601-3612.
47. Lei, M.; Novotny, M. V.; Mechref, Y., Sequential enrichment of sulfated glycans by strong anion-exchange chromatography prior to mass spectrometric measurements. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (3), 348-357.
48. Harvey, D. J., Derivatization of carbohydrates for analysis by chromatography; electrophoresis and mass spectrometry. *J. Chromatogr. B* **2011**, *879* (17), 1196-1225.
49. Liu, H.; Liang, Q.; Sharp, J. S., Peracylation coupled with tandem mass spectrometry for structural sequencing of sulfated glycosaminoglycan mixtures without depolymerization. *J. Am. Soc. Mass Spectrom.* **2020**.
50. Bubb, W. A., NMR spectroscopy in the study of carbohydrates: Characterizing the structural complexity. *Concepts Magn. Reson. A: Bridg. Educ. Res.* **2003**, *19A* (1), 1-19.
51. Duus, J. Ø.; Gotfredsen, C. H.; Bock, K., Carbohydrate Structural Determination by NMR Spectroscopy: Modern Methods and Limitations. *Chem. Rev.* **2000**, *100* (12), 4589-4614.
52. Kailemia, M. J.; Ruhaak, L. R.; Lebrilla, C. B.; Amster, I. J., Oligosaccharide Analysis by Mass Spectrometry: A Review of Recent Developments. *Anal. Chem.* **2014**, *86* (1), 196-212.

53. Pepi, L. E.; Sanderson, P.; Stickney, M.; Amster, I. J., Developments in Mass Spectrometry for Glycosaminoglycan Analysis: A Review. *Mol. Cell Proteomics* **2021**, *20*, 100025.
54. Morelle, W.; Canis, K.; Chirat, F.; Faid, V.; Michalski, J.-C., The use of mass spectrometry for the proteomic analysis of glycosylation. *Proteomics* **2006**, *6* (14), 3993-4015.
55. Zaia, J., Mass Spectrometry and Glycomics. *OMICS J. Integr. Biol.* **2010**, *14* (4), 401-418.
56. Han, L.; Costello, C. E., Mass spectrometry of glycans. *Biochem. (Mosc.)* **2013**, *78* (7), 710-720.
57. Snijder, J.; Rose, R. J.; Veessler, D.; Johnson, J. E.; Heck, A. J. R., Studying 18 MDa Virus Assemblies with Native Mass Spectrometry. *Angew. Chem. Int. Ed.* **2013**, *52* (14), 4020-4023.
58. Keifer, D. Z.; Motwani, T.; Teschke, C. M.; Jarrold, M. F., Measurement of the accurate mass of a 50 MDa infectious virus. *Rapid Commun. Mass Spectrom.* **2016**, *30* (17), 1957-1962.
59. Alseekh, S.; Aharoni, A.; Brotman, Y.; Contrepois, K.; D'Auria, J.; Ewald, J.; C. Ewald, J.; Fraser, P. D.; Giavalisco, P.; Hall, R. D.; Heinemann, M.; Link, H.; Luo, J.; Neumann, S.; Nielsen, J.; Perez de Souza, L.; Saito, K.; Sauer, U.; Schroeder, F. C.; Schuster, S.; Siuzdak, G.; Skirycz, A.; Sumner, L. W.; Snyder, M. P.; Tang, H.; Tohge, T.; Wang, Y.; Wen, W.; Wu, S.; Xu, G.; Zamboni, N.; Fernie, A. R., Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices. *Nat. Methods* **2021**, *18* (7), 747-756.
60. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray ionization—principles and practice. *Mass Spectrom. Rev.* **1990**, *9* (1), 37-70.
61. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989**, *246* (4926), 64-71.
62. Harvey, D. J., Analysis of carbohydrates and glycoconjugates by matrix-assisted laser desorption/ionization mass spectrometry: An update for 2015–2016. *Mass Spectrom. Rev.* **2021**, *40* (4), 408-565.

63. Harvey, D. J., Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates and glycoconjugates. *Int. J. Mass Spectrom.* **2003**, *226* (1), 1-35.
64. Covey, T. R.; Thomson, B. A.; Schneider, B. B., Atmospheric pressure ion sources. *Mass Spectrom. Rev.* **2009**, *28* (6), 870-897.
65. Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S., Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **2013**, *85* (1), 2-9.
66. Kebarle, P.; Verkerk, U. H., Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.* **2009**, *28* (6), 898-917.
67. Kailemia, M. J.; Li, L.; Ly, M.; Linhardt, R. J.; Amster, I. J., Complete Mass Spectral Characterization of a Synthetic Ultralow-Molecular-Weight Heparin Using Collision-Induced Dissociation. *Anal. Chem.* **2012**, *84* (13), 5475-5478.
68. Harvey, D. J.; Struwe, W. B., Structural Studies of Fucosylated N-Glycans by Ion Mobility Mass Spectrometry and Collision-Induced Fragmentation of Negative Ions. *J. Am. Soc. Mass Spectrom.* **2018**.
69. Harvey, D. J., Negative Ion Mass Spectrometry for the Analysis of N-Linked Glycans. *Mass Spectrom. Rev.* **2020**, *39* (5-6), 586-679.
70. Adamson, J. T.; Håkansson, K., Electron Capture Dissociation of Oligosaccharides Ionized with Alkali, Alkaline Earth, and Transition Metals. *Anal. Chem.* **2007**, *79* (7), 2901-2910.
71. Huang, Y.; Yu, X.; Mao, Y.; Costello, C. E.; Zaia, J.; Lin, C., De Novo Sequencing of Heparan Sulfate Oligosaccharides by Electron-Activated Dissociation. *Anal. Chem.* **2013**, *85*, 11979-11986.
72. Wolff, J. J.; Leach, F. E.; Laremore, T. N.; Kaplan, D. A.; Easterling, M. L.; Linhardt, R. J.; Amster, I. J., Negative Electron Transfer Dissociation of Glycosaminoglycans. *Anal. Chem.* **2010**, *82*, 3460-3466.
73. Klein, D. R.; Leach, F. E.; Amster, I. J.; Brodbelt, J. S., Structural Characterization of Glycosaminoglycan Carbohydrates Using Ultraviolet Photodissociation. *Anal. Chem.* **2019**, *91* (9), 6019-6026.

74. Ko, B. J.; Brodbelt, J. S., 193 nm Ultraviolet Photodissociation of Deprotonated Sialylated Oligosaccharides. *Anal. Chem.* **2011**, *83* (21), 8192-8200.
75. McLafferty, F. W., Mass Spectrometric Analysis. Molecular Rearrangements. *Anal. Chem.* **1959**, *31* (1), 82-87.
76. Vachet, R. W.; Bishop, B. M.; Erickson, B. W.; Glish, G. L., Novel Peptide Dissociation: Gas-Phase Intramolecular Rearrangement of Internal Amino Acid Residues. *J. Am. Chem. Soc.* **1997**, *119* (24), 5481-5488.
77. Harrison, A. G.; Young, A. B.; Bleiholder, C.; Suhai, S.; Paizs, B., Scrambling of Sequence Information in Collision-Induced Dissociation of Peptides. *J. Am. Chem. Soc.* **2006**, *128* (32), 10364-10365.
78. Garcia, I. R.; Giles, K.; Bateman, R. H.; Gaskell, S. J., Studies of peptide a- and b-type fragment ions using stable isotope labeling and integrated ion mobility/tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (12), 1781-1787.
79. Chen, X.; Yu, L.; Steill, J. D.; Oomens, J.; Polfer, N. C., Effect of Peptide Fragment Size on the Propensity of Cyclization in Collision-Induced Dissociation: Oligoglycine b₂-b₈. *J. Am. Chem. Soc.* **2009**, *131* (51), 18272-18282.
80. Erlekam, U.; Bythell, B. J.; Scuderi, D.; Van Stipdonk, M.; Paizs, B.; Maître, P., Infrared Spectroscopy of Fragments of Protonated Peptides: Direct Evidence for Macrocyclic Structures of b₅ Ions. *J. Am. Chem. Soc.* **2009**, *131* (32), 11503-11508.
81. Wührer, M.; Deelder, A. M.; van der Burgt, Y. E., Mass spectrometric glycan rearrangements. *Mass Spectrom. Rev.* **2011**, *30* (4), 664-80.
82. Hecht, E. S.; Loziuk, P. L.; Muddiman, D. C., Xylose Migration During Tandem Mass Spectrometry of N-Linked Glycans. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (4), 729-732.
83. Kováčik, V.; Hirsch, J.; Kováč, P.; Heerma, W.; Thomas-Oates, J.; Haverkamp, J., Oligosaccharide characterization using collision-induced dissociation fast atom bombardment mass spectrometry: Evidence for internal monosaccharide residue loss. *J. Mass Spectrom.* **1995**, *30* (7), 949-958.

84. Ma, Y.-L.; Vedernikova, I.; Heuvel, H.; Claeys, M., Internal glucose residue loss in protonated O-diglycosyl flavonoids upon low-energy collision-induced dissociation. *J. Am. Soc. Mass Spectrom.* **2000**, *11* (2), 136-144.
85. Wuhrer, M.; Koeleman, C. A.; Deelder, A. M., Hexose rearrangements upon fragmentation of N-glycopeptides and reductively aminated N-glycans. *Anal. Chem.* **2009**, *81* (11), 4422-32.
86. Harvey, D. J.; Mattu, T. S.; Wormald, M. R.; Royle, L.; Dwek, R. A.; Rudd, P. M., "Internal residue loss": rearrangements occurring during the fragmentation of carbohydrates derivatized at the reducing terminus. *Anal. Chem.* **2002**, *74* (4), 734-40.
87. Franz, A. H.; Lebrilla, C. B., Evidence for long-range glycosyl transfer reactions in the gas phase. *J. Am. Soc. Mass Spectrom.* **2002**, *13* (4), 325-37.
88. Wuhrer, M.; Koeleman, C. A.; Hokke, C. H.; Deelder, A. M., Mass spectrometry of proton adducts of fucosylated N-glycans: fucose transfer between antennae gives rise to misleading fragments. *Rapid Commun. Mass Spectrom.* **2006**, *20* (11), 1747-54.
89. Tolmachev, A. V.; Vilkov, A. N.; Bogdanov, B.; Páša-Tolić, L.; Masselon, C. D.; Smith, R. D., Collisional activation of ions in RF ion traps and ion guides: The effective ion temperature treatment. *J. Am. Soc. Mass Spectrom.* **2004**, *15* (11), 1616-1628.
90. Xu, Y.-F.; Lu, W.; Rabinowitz, J. D., Avoiding Misannotation of In-Source Fragmentation Products as Cellular Metabolites in Liquid Chromatography–Mass Spectrometry-Based Metabolomics. *Anal. Chem.* **2015**, *87* (4), 2273-2281.
91. Zappe, A.; Miller, R. L.; Struwe, W. B.; Pagel, K., State-of-the-art glycosaminoglycan characterization. *Mass Spectrom. Rev.* **2021**, *n/a* (n/a).
92. Gunay, N. S.; Tadano-Aritomi, K.; Toida, T.; Ishizuka, I.; Linhardt, R. J., Evaluation of Counterions for Electrospray Ionization Mass Spectral Analysis of A Highly Sulfated Carbohydrate, Sucrose Octasulfate. *Anal. Chem.* **2003**, *75* (13), 3226-3231.
93. Shi, X.; Huang, Y.; Mao, Y.; Naimy, H.; Zaia, J., Tandem Mass Spectrometry of Heparan Sulfate Negative Ions: Sulfate Loss Patterns and Chemical Modification Methods for Improvement of Product Ion Profiles. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (9), 1498-1511.

94. Naggar, E. F.; Costello, C. E.; Zaia, J., Competing fragmentation processes in tandem mass spectra of heparin-like glycosaminoglycans. *J. Am. Soc. Mass Spectrom.* **2004**, *15* (11), 1534-1544.
95. Hofmann, J.; Pagel, K., Glycan Analysis by Ion Mobility–Mass Spectrometry. *Angew. Chem. Int. Ed.* **2017**, *56* (29), 8342-8349.
96. Manz, C.; Pagel, K., Glycan analysis by ion mobility-mass spectrometry and gas-phase spectroscopy. *Curr. Opin. Chem. Biol.* **2018**, *42* (Supplement C), 16-24.
97. Gray, C. J.; Thomas, B.; Upton, R.; Migas, L. G.; Eysers, C. E.; Barran, P. E.; Flitsch, S. L., Applications of ion mobility mass spectrometry for high throughput, high resolution glycan analysis. *Biochim. Biophys. Acta - Gen. Subj.* **2016**, *1860* (8), 1688-1709.
98. Mookherjee, A.; Guttman, M., Bridging the structural gap of glycoproteomics with ion mobility spectrometry. *Curr. Opin. Chem. Biol.* **2018**, *42*, 86-92.
99. Gabelica, V.; Marklund, E., Fundamentals of ion mobility spectrometry. *Curr. Opin. Chem. Biol.* **2018**, *42*, 51-59.
100. 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.
101. Grabarics, M.; Lettow, M.; Kirk, A. T.; von Helden, G.; Causon, T. J.; Pagel, K., Plate-height model of ion mobility-mass spectrometry. *Analyst* **2020**, *145* (19), 6313-6333.
102. Grabarics, M.; Lettow, M.; Kirk, A. T.; von Helden, G.; Causon, T. J.; Pagel, K., Plate-height model of ion mobility-mass spectrometry: Part 2—Peak-to-peak resolution and peak capacity. *J. Sep. Sci.* **2021**, *44* (14), 2798-2813.
103. Mason, E. A.; Schamp, H. W., Mobility of gaseous ions in weak electric fields. *Ann. Phys.* **1958**, *4* (3), 233-270.
104. Revercomb, H. E.; Mason, E. A., Theory of plasma chromatography/gaseous electrophoresis. Review. *Anal. Chem.* **1975**, *47* (7), 970-983.

105. Struwe, W. B.; Pagel, K.; Benesch, J. L. P.; Harvey, D. J.; Campbell, M. P., GlycoMob: an ion mobility-mass spectrometry collision cross section database for glycomics. *Glycoconj. J.* **2016**, *33* (3), 399-404.
106. Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F., Structural Information from Ion Mobility Measurements: Effects of the Long-Range Potential. *J. Phys. Chem.* **1996**, *100* (40), 16082-16086.
107. Ieritano, C.; Crouse, J.; Campbell, J. L.; Hopkins, W. S., A parallelized molecular collision cross section package with optimized accuracy and efficiency. *Analyst* **2019**, *144* (5), 1660-1670.
108. Zanotto, L.; Heerdt, G.; Souza, P. C. T.; Araujo, G.; Skaf, M. S., High performance collision cross section calculation—HPCCS. *J. Comput. Chem.* **2018**, *39* (21), 1675-1681.
109. Hofmann, J.; Hahm, H. S.; Seeberger, P. H.; Pagel, K., Identification of carbohydrate anomers using ion mobility-mass spectrometry. *Nature* **2015**, *526* (7572), 241-4.
110. Miller, R. L.; Guimond, S. E.; Schwörer, R.; Zubkova, O. V.; Tyler, P. C.; Xu, Y.; Liu, J.; Chopra, P.; Boons, G.-J.; Grabarics, M.; Manz, C.; Hofmann, J.; Karlsson, N. G.; Turnbull, J. E.; Struwe, W. B.; Pagel, K., Shotgun ion mobility mass spectrometry sequencing of heparan sulfate saccharides. *Nat. Comm.* **2020**, *11* (1), 1481.
111. Both, P.; Green, A. P.; Gray, C. J.; Šardžik, R.; Voglmeir, J.; Fontana, C.; Austeri, M.; Rejzek, M.; Richardson, D.; Field, R. A.; Widmalm, G.; Flitsch, S. L.; Eyers, C. E., Discrimination of epimeric glycans and glycopeptides using IM-MS and its potential for carbohydrate sequencing. *Nat. Chem.* **2014**, *6* (1), 65-74.
112. Boyarkin, O. V., Cold ion spectroscopy for structural identifications of biomolecules. *Int. Rev. Phys. Chem.* **2018**, *37* (3-4), 559-606.
113. Greis, K.; Kirschbaum, C.; von Helden, G.; Pagel, K., Gas-phase infrared spectroscopy of glycans and glycoconjugates. *Curr. Opin. Struct. Biol.* **2022**, *72*, 194-202.
114. 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**.

115. Compagnon, I.; Schindler, B.; Renois-Predelus, G.; Daniel, R., Lasers and ion mobility: new additions to the glycosaminoglycanomics toolkit. *Curr. Opin. Struct. Biol.* **2018**, *50*, 171-180.
116. Gray, C. J.; Migas, L. G.; Barran, P. E.; Pagel, K.; Seeberger, P. H.; Eyers, C. E.; Boons, G.-J.; Pohl, N. L. B.; Compagnon, I.; Widmalm, G.; Flitsch, S. L., Advancing Solutions to the Carbohydrate Sequencing Challenge. *J. Am. Chem. Soc.* **2019**.
117. Rijs, A. M.; Oomens, J., IR Spectroscopic Techniques to Study Isolated Biomolecules. In *Gas-Phase IR Spectroscopy and Structure of Biological Molecules*, Rijs, A. M.; Oomens, J., Eds. Springer International Publishing: Cham, 2015; pp 1-42.
118. Oomens, J.; Sartakov, B. G.; Meijer, G.; von Helden, G., Gas-phase infrared multiple photon dissociation spectroscopy of mass-selected molecular ions. *Int. J. Mass Spectrom.* **2006**, *254* (1-2), 1-19.
119. Polfer, N. C.; Oomens, J., Vibrational spectroscopy of bare and solvated ionic complexes of biological relevance. *Mass Spectrom. Rev.* **2009**, *28* (3), 468-94.
120. Schöllkopf, W.; Gewinner, S.; Junkes, H.; Paarmann, A.; von Helden, G.; Bluem, H.; Todd, A. M. M., The new IR and THz FEL facility at the Fritz Haber Institute in Berlin. In *Proc. of SPIE - Advances in X-ray Free-Electron Lasers Instrumentation III*, Biedron, S. G., Ed. Prague, Czech Republic, 2015; Vol. 9512.
121. Lucas, B.; Grégoire, G.; Lemaire, J.; Maître, P.; Glotin, F.; Schermann, J. P.; Desfrancois, C., Infrared multiphoton dissociation spectroscopy of protonated N-acetyl-alanine and alanyl-histidine. *Int. J. Mass Spectrom.* **2005**, *243* (2), 105-113.
122. Oomens, J.; van Rooij, A. J. A.; Meijer, G.; von Helden, G., Gas-Phase Infrared Photodissociation Spectroscopy of Cationic Polyaromatic Hydrocarbons. *Astrophys. J.* **2000**, *542* (1), 404-410.
123. Talbot, F. O.; Simons, J. P., Sugars in the gas phase: the spectroscopy and structure of jet-cooled phenyl β -D-glucopyranoside. *Phys. Chem. Chem. Phys.* **2002**, *4* (15), 3562-3565.
124. Jockusch, R. A.; Kroemer, R. T.; Talbot, F. O.; Snoek, L. C.; Carcabal, P.; Simons, J. P.; Havenith, M.; Bakker, J. M.; Compagnon, I.; Meijer, G.; von

- Helden, G., Probing the glycosidic linkage: UV and IR ion-dip spectroscopy of a lactoside. *J. Am. Chem. Soc.* **2004**, *126* (18), 5709-14.
125. Schindler, B.; Joshi, J.; Allouche, A. R.; Simon, D.; Chambert, S.; Brites, V.; Gaigeot, M. P.; Compagnon, I., Distinguishing isobaric phosphated and sulfated carbohydrates by coupling of mass spectrometry with gas phase vibrational spectroscopy. *Phys. Chem. Chem. Phys.* **2014**, *16* (40), 22131-8.
126. Masellis, C.; Khanal, N.; Kamrath, M. Z.; Clemmer, D. E.; Rizzo, T. R., Cryogenic Vibrational Spectroscopy Provides Unique Fingerprints for Glycan Identification. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (10), 2217-2222.
127. Mucha, E.; Flórez, A. I. G.; Marianski, M.; Thomas, D. A.; Hoffmann, W.; Struwe, W. B.; Hahm, H. S.; Gewinner, S.; Schöllkopf, W.; Seeberger, P. H.; Helden, G. v.; Pagel, K., Glycan Fingerprinting via Cold-Ion Infrared Spectroscopy. *Angew. Chem. Int. Ed.* **2017**, *56* (37), 11248-11251.
128. Marianski, M.; Supady, A.; Ingram, T.; Schneider, M.; Baldauf, C., Assessing the Accuracy of Across-the-Scale Methods for Predicting Carbohydrate Conformational Energies for the Examples of Glucose and α -Maltose. *J. Chem. Theory Comput.* **2016**, *12* (12), 6157-6168.
129. 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.
130. Elferink, H.; Severijnen, M. E.; Martens, J.; Mensink, R. A.; Berden, G.; Oomens, J.; Rutjes, F. P. J. T.; Rijs, A. M.; Boltje, T. J., Direct Experimental Characterization of Glycosyl Cations by Infrared Ion Spectroscopy. *J. Am. Chem. Soc.* **2018**.
131. Barnes, L.; Schindler, B.; Allouche, A. R.; Simon, D.; Chambert, S.; Oomens, J.; Compagnon, I., Anharmonic simulations of the vibrational spectrum of sulfated compounds: application to the glycosaminoglycan fragment glucosamine 6-sulfate. *Phys. Chem. Chem. Phys.* **2015**, *17* (39), 25705-13.
132. Barnes, L.; Schindler, B.; Chambert, S.; Allouche, A.-R.; Compagnon, I., Conformational preferences of protonated N-acetylated hexosamines probed by InfraRed Multiple Photon Dissociation (IRMPD) spectroscopy and ab initio calculations. *Int. J. Mass Spectrom.* **2017**, *421*, 116-123.

133. Tan, Y.; Polfer, N. C., Linkage and anomeric differentiation in trisaccharides by sequential fragmentation and variable-wavelength infrared photodissociation. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (2), 359-68.
134. 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.
135. Schindler, B.; Laloy-Borgna, G.; Barnes, L.; Allouche, A. R.; Bouju, E.; Dugas, V.; Demesmay, C.; Compagnon, I., Online Separation and Identification of Isomers Using Infrared Multiple Photon Dissociation Ion Spectroscopy Coupled to Liquid Chromatography: Application to the Analysis of Disaccharides Regio-Isomers and Monosaccharide Anomers. *Anal. Chem.* **2018**, *90* (20), 11741-11745.
136. Roithova, J.; Gray, A.; Andris, E.; Jasik, J.; Gerlich, D., Helium Tagging Infrared Photodissociation Spectroscopy of Reactive Ions. *Acc. Chem. Res.* **2016**, *49* (2), 223-30.
137. Mosley, J. D.; Young, J. W.; Agarwal, J.; Schaefer III, H. F.; Schleyer, P. v. R.; Duncan, M. A., Structural Isomerization of the Gas-Phase 2-Norbornyl Cation Revealed with Infrared Spectroscopy and Computational Chemistry. *Angew. Chem. Int. Ed.* **2014**, *53* (23), 5888-5891.
138. Kamrath, M. Z.; Garand, E.; Jordan, P. A.; Leavitt, C. M.; Wolk, A. B.; Van Stipdonk, M. J.; Miller, S. J.; Johnson, M. A., Vibrational Characterization of Simple Peptides Using Cryogenic Infrared Photodissociation of H₂-Tagged, Mass-Selected Ions. *J. Am. Chem. Soc.* **2011**, *133* (16), 6440-6448.
139. Khanal, N.; Masellis, C.; Kamrath, M. Z.; Clemmer, D. E.; Rizzo, T. R., Cryogenic IR spectroscopy Combined with Ion Mobility Spectrometry for the Analysis of Human Milk Oligosaccharides. *Analyst* **2018**.
140. 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.
141. Gerlich, D., Ion-neutral collisions in a 22-pole trap at very low energies. *Phys. Scr.* **1995**, *T59*, 256-263.

142. Gerlich, D., Inhomogeneous RF Fields: A Versatile Tool for the Study of Processes with Slow Ions. In *Adv. Chem. Phys.*, 2007; pp 1-176.
143. Wang, Y. S.; Tsai, C. H.; Lee, Y. T.; Chang, H. C.; Jiang, J. C.; Asvany, O.; Schlemmer, S.; Gerlich, D., Investigations of Protonated and Deprotonated Water Clusters Using a Low-Temperature 22-Pole Ion Trap. *J. Phys. Chem. A* **2003**, *107* (21), 4217-4225.
144. Boyarkin, O. V.; Mercier, S. R.; Kamariotis, A.; Rizzo, T. R., Electronic spectroscopy of cold, protonated tryptophan and tyrosine. *J. Am. Chem. Soc.* **2006**, *128* (9), 2816-7.
145. Asmis, K. R.; Brümmer, M.; Kaposta, C.; Santambrogio, G.; von Helden, G.; Meijer, G.; Rademann, K.; Wöste, L., Mass-selected infrared photodissociation spectroscopy of V4O10+. *Phys. Chem. Chem. Phys.* **2002**, *4* (7), 1101-1104.
146. Lorenz, U. J.; Rizzo, T. R., Planar multipole ion trap/time-of-flight mass spectrometer. *Anal. Chem.* **2011**, *83* (20), 7895-901.
147. Bansal, P.; Yatsyna, V.; AbiKhodr, A. H.; Warnke, S.; Ben Faleh, A.; Yalovenko, N.; Wysocki, V. H.; Rizzo, T. R., Using SLIM-Based IMS-IMS Together with Cryogenic Infrared Spectroscopy for Glycan Analysis. *Anal. Chem.* **2020**, *92* (13), 9079-9085.
148. Yalovenko, N.; Yatsyna, V.; Bansal, P.; AbiKhodr, A. H.; Rizzo, T. R., Analyzing glycans cleaved from a biotherapeutic protein using ultrahigh-resolution ion mobility spectrometry together with cryogenic ion spectroscopy. *Analyst* **2020**, *145* (20), 6493-6499.
149. Heine, N.; Asmis, K. R., Cryogenic ion trap vibrational spectroscopy of hydrogen-bonded clusters relevant to atmospheric chemistry. *Int. Rev. Phys. Chem.* **2015**, *34* (1), 1-34.
150. Mucha, E.; Marianski, M.; Xu, F. F.; Thomas, D. A.; Meijer, G.; von Helden, G.; Seeberger, P. H.; Pagel, K., Unravelling the structure of glycosyl cations via cold-ion infrared spectroscopy. *Nat. Commun.* **2018**, *9* (1), 4174.
151. Greis, K.; Kirschbaum, C.; Lechnitz, S.; Gewinner, S.; Schöllkopf, W.; von Helden, G.; Meijer, G.; Seeberger, P. H.; Pagel, K., Direct Experimental Characterization of the Ferrier Glycosyl Cation in the Gas Phase. *Org. Lett.* **2020**, *22* (22), 8916-8919.

152. Greis, K.; Kirschbaum, C.; Fittolani, G.; Mucha, E.; Chang, R.; von Helden, G.; Meijer, G.; Delbianco, M.; Seeberger, P. H.; Pagel, K., Neighboring Group Participation of Benzoyl Protecting Groups in C3- and C6-Fluorinated Glucose. *Eur. J. Org. Chem.* **2022**, 2022 (15), e202200255.
153. Greis, K.; Mucha, E.; Lettow, M.; Thomas, D. A.; Kirschbaum, C.; Moon, S.; Pardo-Vargas, A.; von Helden, G.; Meijer, G.; Gilmore, K.; Seeberger, P. H.; Pagel, K., The Impact of Leaving Group Anomericity on the Structure of Glycosyl Cations of Protected Galactosides. *ChemPhysChem* **2020**, 21 (17), 1905-1907.
154. Gonzalez Florez, A. I.; Ahn, D. S.; Gewinner, S.; Schollkopf, W.; von Helden, G., IR spectroscopy of protonated leu-enkephalin and its 18-crown-6 complex embedded in helium droplets. *Phys. Chem. Chem. Phys.* **2015**, 17 (34), 21902-11.
155. Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B., Infrared Spectroscopy and Theoretical Studies on Gas-Phase Protonated Leu-enkephalin and Its Fragments: Direct Experimental Evidence for the Mobile Proton. *J. Am. Chem. Soc.* **2007**, 129 (18), 5887-5897.
156. Varki, A., Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **1993**, 3 (2), 97-130.
157. Varki, A., Nothing in Glycobiology Makes Sense, except in the Light of Evolution. *Cell* **2006**, 126 (5), 841-845.
158. Myers, R. B.; Srivastava, S.; Grizzle, W. E., Lewis Y Antigen as Detected by the Monoclonal Antibody BR96 is Expressed Strongly in Prostatic Adenocarcinoma. *J. Urol.* **1995**, 153 (5), 1572-1574.
159. Zhang, Z.; Sun, P.; Liu, J.; Fu, L.; Yan, J.; Liu, Y.; Yu, L.; Wang, X.; Yan, Q., Suppression of FUT1/FUT4 expression by siRNA inhibits tumor growth. *BBA - Mol. Cell Res.* **2008**, 1783 (2), 287-296.
160. Yin, B. W. T.; Finstad, C. L.; Kitamura, K.; Federici, M. G.; Welshinger, M.; Kudryashov, V.; Hoskins, W. J.; Welt, S.; Lloyd, K. O., Serological and immunochemical analysis of Lewis Y (Ley) blood group antigen expression in epithelial ovarian cancer. *Int. J. Cancer* **1996**, 65 (4), 406-412.
161. Harvey, D. J., Identification of protein-bound carbohydrates by mass spectrometry. *Proteomics* **2001**, 1 (2), 311-328.

162. Leymarie, N.; Zaia, J., Effective Use of Mass Spectrometry for Glycan and Glycopeptide Structural Analysis. *Anal. Chem.* **2012**, *84* (7), 3040-3048.
163. Sheeley, D. M.; Reinhold, V. N., Structural Characterization of Carbohydrate Sequence, Linkage, and Branching in a Quadrupole Ion Trap Mass Spectrometer: Neutral Oligosaccharides and N-Linked Glycans. *Anal. Chem.* **1998**, *70* (14), 3053-3059.
164. Wuhrer, M.; Deelder, A. M.; Hokke, C. H., Protein glycosylation analysis by liquid chromatography–mass spectrometry. *J. Chromatogr. B* **2005**, *825* (2), 124-133.
165. Jia, C.; Wu, Z.; Lietz, C. B.; Liang, Z.; Cui, Q.; Li, L., Gas-Phase Ion Isomer Analysis Reveals the Mechanism of Peptide Sequence Scrambling. *Anal. Chem.* **2014**, *86* (6), 2917-2924.
166. Bythell, B. J.; Molesworth, S.; Osburn, S.; Cooper, T.; Paizs, B.; Van Stipdonk, M., Structure and reactivity of a n and a* n peptide fragments investigated using isotope labeling, tandem mass spectrometry, and density functional theory calculations. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (12), 1788-1798.
167. Harrison, A. G., To b or not to b: The ongoing saga of peptide b ions. *Mass Spectrom. Rev.* **2009**, *28* (4), 640-654.
168. Goloborodko, A. A.; Gorshkov, M. V.; Good, D. M.; Zubarev, R. A., Sequence Scrambling in Shotgun Proteomics is Negligible. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (7), 1121-1124.
169. Saminathan, I. S.; Wang, X. S.; Guo, Y.; Krakovska, O.; Voisin, S.; Hopkinson, A. C.; Siu, K. W. M., The Extent and Effects of Peptide Sequence Scrambling Via Formation of Macrocyclic b Ions in Model Proteins. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (12), 2085-2094.
170. Patrick, A. L.; Polfer, N. C., Peptide Fragmentation Products in Mass Spectrometry Probed by Infrared Spectroscopy. In *Gas-Phase IR Spectroscopy and Structure of Biological Molecules*, Rijs, A. M.; Oomens, J., Eds. Springer International Publishing: Cham, 2015; pp 153-181.
171. Brüll, L. P.; Kováčik, V.; Thomas-Oates, J. E.; Heerma, W.; Haverkamp, J., Sodium-cationized oligosaccharides do not appear to undergo ‘internal residue loss’ rearrangement processes on tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **1998**, *12* (20), 1520-1532.

172. Ernst, B.; Müller, D. R.; Richter, W. J., False sugar sequence ions in electrospray tandem mass spectrometry of underivatized sialyl-Lewis-type oligosaccharides. *Int. J. Mass Spectrom. Ion Processes* **1997**, *160* (1-3), 283-290.
173. Brüll, L. P.; Heerma, W.; Thomas-Oates, J.; Haverkamp, J.; Kováčik, V.; Kovác, P., Loss of internal 1 → 6 substituted monosaccharide residues from underivatized and per-O-methylated trisaccharides. *J. Am. Soc. Mass Spectrom* **1997**, *8* (1), 43-49.
174. Mucha, E.; Lettow, M.; Marianski, M.; Thomas, D. A.; Struwe, W. B.; Harvey, D. J.; Meijer, G.; Seeberger, P. H.; von Helden, G.; Pagel, K., Fucose Migration in Intact Protonated Glycan Ions: A Universal Phenomenon in Mass Spectrometry. *Angew. Chem. Int. Ed.* **2018**, *57* (25), 7440-7443.
175. Pabst, M.; Kolarich, D.; Pörtl, G.; Dalik, T.; Lubec, G.; Hofinger, A.; Altmann, F., Comparison of fluorescent labels for oligosaccharides and introduction of a new postlabeling purification method. *Anal. Biochem.* **2009**, *384* (2), 263-273.
176. Royle, L.; Dwek, R. A.; Rudd, P. M., Determining the Structure of Oligosaccharides N- and O-Linked to Glycoproteins. *Curr. Protoc. Protein Sci.* **2006**, *43* (1), 12.6.1-12.6.45.
177. Hunter, E. P. L.; Lias, S. G., Evaluated Gas Phase Basicities and Proton Affinities of Molecules: An Update. *J. Phys. Chem. Ref. Data* **1998**, *27* (3), 413-656.
178. Re, S.; Watabe, S.; Nishima, W.; Muneyuki, E.; Yamaguchi, Y.; MacKerell, A. D.; Sugita, Y., Characterization of Conformational Ensembles of Protonated N-glycans in the Gas-Phase. *Sci. Rep.* **2018**, *8* (1), 1644.
179. O'Flaherty, R.; Trbojević-Akmačić, I.; Greville, G.; Rudd, P. M.; Lauc, G., The sweet spot for biologics: recent advances in characterization of biotherapeutic glycoproteins. *Expert Rev. Proteomics* **2018**, *15* (1), 13-29.
180. Warnke, S.; Seo, J.; Boschmans, J.; Sobott, F.; Scrivens, J. H.; Bleiholder, C.; Bowers, M. T.; Gewinner, S.; Schöllkopf, W.; Pagel, K.; von Helden, G., Protomers of Benzocaine: Solvent and Permittivity Dependence. *J. Am. Chem. Soc.* **2015**, *137* (12), 4236-4242.

181. Nwosu, C.; Yau, H. K.; Becht, S., Assignment of Core versus Antenna Fucosylation Types in Protein N-Glycosylation via Procainamide Labeling and Tandem Mass Spectrometry. *Anal. Chem.* **2015**, *87* (12), 5905-13.
182. Agre, P.; Bertozzi, C.; Bissell, M.; Campbell, K. P.; Cummings, R. D.; Desai, U. R.; Estes, M.; Flotte, T.; Fogleman, G.; Gage, F.; Ginsburg, D.; Gordon, J. I.; Hart, G.; Hascall, V.; Kiessling, L.; Kornfeld, S.; Lowe, J.; Magnani, J.; Mahal, L. K.; Medzhitov, R.; Roberts, R. J.; Sackstein, R.; Sarkar, R.; Schnaar, R.; Schwartz, N.; Varki, A.; Walt, D.; Weissman, I., Training the next generation of biomedical investigators in glycosciences. *J. Clin. Invest.* **2016**, *126* (2), 405-408.
183. Li, J.; Hsu, H.-C.; Mountz, J. D.; Allen, J. G., Unmasking Fucosylation: from Cell Adhesion to Immune System Regulation and Diseases. *Cell Chem. Biol.* **2018**, *25* (5), 499-512.
184. Watanabe, Y.; Bowden, T. A.; Wilson, I. A.; Crispin, M., Exploitation of glycosylation in enveloped virus pathobiology. *Biochim. Biophys. Acta - Gen. Subj.* **2019**, *1863* (10), 1480-1497.
185. Chen, X.; Flynn, G. C., Analysis of N-glycans from recombinant immunoglobulin G by on-line reversed-phase high-performance liquid chromatography/mass spectrometry. *Anal. Biochem.* **2007**, *370* (2), 147-161.
186. Broberg, A., High-performance liquid chromatography/electrospray ionization ion-trap mass spectrometry for analysis of oligosaccharides derivatized by reductive amination and N,N-dimethylation. *Carbohydr. Res.* **2007**, *342* (11), 1462-1469.
187. Desai, N.; Thomas, D. A.; Lee, J.; Gao, J.; Beauchamp, J. L., Eradicating mass spectrometric glycan rearrangement by utilizing free radicals. *Chem. Sci.* **2016**, *7* (8), 5390-5397.
188. 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.
189. Marianski, M.; Mucha, E.; Greis, K.; Moon, S.; Pardo, A.; Kirschbaum, C.; Thomas, D. A.; Meijer, G.; von Helden, G.; Gilmore, K.; Seeberger, P. H.; Pagel, K., Remote Participation during Glycosylation Reactions of

- Galactose Building Blocks: Direct Evidence from Cryogenic Vibrational Spectroscopy. *Angew. Chem. Int. Ed.* **2020**, *59* (15), 6166-6171.
190. Zhang, X.; Julian, R. R., Radical mediated dissection of oligosaccharides. *Int. J. Mass Spectrom.* **2014**, *372*, 22-28.
191. Gao, J.; Thomas, D. A.; Sohn, C. H.; Beauchamp, J. L., Biomimetic Reagents for the Selective Free Radical and Acid–Base Chemistry of Glycans: Application to Glycan Structure Determination by Mass Spectrometry. *J. Am. Chem. Soc.* **2013**, *135* (29), 10684-10692.
192. Riggs, D. L.; Hofmann, J.; Hahm, H. S.; Seeberger, P. H.; Pagel, K.; Julian, R. R., Glycan Isomer Identification Using Ultraviolet Photodissociation Initiated Radical Chemistry. *Anal. Chem.* **2018**, *90* (19), 11581-11588.
193. Warnke, S.; von Helden, G.; Pagel, K., Analyzing the higher order structure of proteins with conformer-selective ultraviolet photodissociation. *Proteomics* **2015**, *15* (16), 2804-2812.
194. Thomas, D. A.; Chang, R.; Mucha, E.; Lettow, M.; Greis, K.; Gewinner, S.; Schöllkopf, W.; Meijer, G.; von Helden, G., Probing the conformational landscape and thermochemistry of DNA dinucleotide anions via helium nanodroplet infrared action spectroscopy. *Phys. Chem. Chem. Phys.* **2020**, *22* (33), 18400-18413.
195. Gabelica, V.; Shvartsburg, A. A.; Afonso, C.; Barran, P.; Benesch, J. L. P.; Bleiholder, C.; Bowers, M. T.; Bilbao, A.; Bush, M. F.; Campbell, J. L.; Campuzano, I. D. G.; Causon, T.; Clowers, B. H.; Creaser, C. S.; De Pauw, E.; Far, J.; Fernandez-Lima, F.; Fjeldsted, J. C.; Giles, K.; Groessl, M.; Hogan Jr, C. J.; Hann, S.; Kim, H. I.; Kurulugama, R. T.; May, J. C.; McLean, J. A.; Pagel, K.; Richardson, K.; Ridgeway, M. E.; Rosu, F.; Sobott, F.; Thalassinou, K.; Valentine, S. J.; Wyttenbach, T., Recommendations for reporting ion mobility Mass Spectrometry measurements. *Mass Spectrom. Rev.* **2019**, *38* (3), 291-320.
196. Scutelnic, V.; Perez, M. A. S.; Marianski, M.; Warnke, S.; Gregor, A.; Rothlisberger, U.; Bowers, M. T.; Baldauf, C.; von Helden, G.; Rizzo, T. R.; Seo, J., The Structure of the Protonated Serine Octamer. *J. Am. Chem. Soc.* **2018**, *140* (24), 7554-7560.
197. Rahim, S.; Yaman, M.; Marianski, M., glyP: An automated tool for structural analysis of glycans. *Manuscript in preparation*.

198. Pracht, P.; Bohle, F.; Grimme, S., Automated exploration of the low-energy chemical space with fast quantum chemical methods. *Phys. Chem. Chem. Phys.* **2020**, *22* (14), 7169-7192.
199. Bannwarth, C.; Ehlert, S.; Grimme, S., GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15* (3), 1652-1671.
200. Adamo, C.; Barone, V., Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110* (13), 6158-6170.
201. von Helden, G.; Hsu, M. T.; Gotts, N.; Bowers, M. T., Carbon cluster cations with up to 84 atoms: structures, formation mechanism, and reactivity. *J. Phys. Chem.* **1993**, *97* (31), 8182-8192.
202. Sastre Torano, J.; Gagarinov, I. A.; Vos, G. M.; Broszeit, F.; Srivastava, A. D.; Palmer, M.; Langridge, J. I.; Aizpurua-Olaizola, O.; Somovilla, V. J.; Boons, G. J., Ion-Mobility Spectrometry Can Assign Exact Fucosyl Positions in Glycans and Prevent Misinterpretation of Mass-Spectrometry Data After Gas-Phase Rearrangement. *Angew. Chem. Int. Ed.* **2019**, *58* (49), 17616-17620.
203. Gandhi, N. S.; Mancera, R. L., The Structure of Glycosaminoglycans and their Interactions with Proteins. *Chem. Biol. Drug Des.* **2008**, *72* (6), 455-482.
204. Lindahl, U.; Couchman, J.; Kimata, K.; Esko, J. D., Proteoglycans and Sulfated Glycosaminoglycans. In *Essentials of Glycobiology* 3ed.; 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): Cold Spring Harbor Laboratory Press: 2017.
205. Sasisekharan, R.; Shriver, Z.; Venkataraman, G.; Narayanasami, U., Roles of heparan-sulphate glycosaminoglycans in cancer. *Nat. Rev. Cancer* **2002**, *2* (7), 521-528.
206. Capila, I.; Linhardt, R. J., Heparin–Protein Interactions. *Angew. Chem. Int. Ed.* **2002**, *41* (3), 390-412.

207. Jones, C. J.; Beni, S.; Limtiaco, J. F. K.; Langeslay, D. J.; Larive, C. K., Heparin Characterization: Challenges and Solutions. *Ann. Rev. Anal. Chem.* **2011**, *4* (1), 439-465.
208. Ly, M.; Leach, F. E.; Laremore, T. N.; Toida, T.; Amster, I. J.; Linhardt, R. J., The proteoglycan bikunin has a defined sequence. *Nat. Chem. Biol.* **2011**, *7* (11), 827-833.
209. Wolff, J. J.; Amster, I. J.; Chi, L.; Linhardt, R. J., Electron detachment dissociation of glycosaminoglycan tetrasaccharides. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (2), 234-244.
210. Wolff, J. J.; Laremore, T. N.; Aslam, H.; Linhardt, R. J.; Amster, I. J., Electron-induced dissociation of glycosaminoglycan tetrasaccharides. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (10), 1449-1458.
211. Leach, F. E.; Riley, N. M.; Westphall, M. S.; Coon, J. J.; Amster, I. J., Negative Electron Transfer Dissociation Sequencing of Increasingly Sulfated Glycosaminoglycan Oligosaccharides on an Orbitrap Mass Spectrometer. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (9), 1844-1854.
212. Wei, J.; Wu, J.; Tang, Y.; Ridgeway, M. E.; Park, M. A.; Costello, C. E.; Zaia, J.; Lin, C., Characterization and Quantification of Highly Sulfated Glycosaminoglycan Isomers by Gated-Trapped Ion Mobility Spectrometry Negative Electron Transfer Dissociation MS/MS. *Anal. Chem.* **2019**, *91* (4), 2994-3001.
213. Seo, Y.; Andaya, A.; Leary, J. A., Preparation, Separation, and Conformational Analysis of Differentially Sulfated Heparin Octasaccharide Isomers Using Ion Mobility Mass Spectrometry. *Anal. Chem.* **2012**, *84* (5), 2416-2423.
214. Brézillon, S.; Untereiner, V.; Lovergne, L.; Tadeo, I.; Noguera, R.; Maquart, F.-X.; Wegrowski, Y.; Sockalingum, G. D., Glycosaminoglycan profiling in different cell types using infrared spectroscopy and imaging. *Anal. Bioanal. Chem.* **2014**, *406* (24), 5795-5803.
215. Mohamed, H. T.; Untereiner, V.; Sockalingum, G. D.; Brézillon, S., Implementation of infrared and Raman modalities for glycosaminoglycan characterization in complex systems. *Glycoconj. J.* **2017**, *34* (3), 309-323.
216. Renois-Predelus, G.; Schindler, B.; Compagnon, I., Analysis of Sulfate Patterns in Glycosaminoglycan Oligosaccharides by MS(n) Coupled to

- Infrared Ion Spectroscopy: the Case of GalNAc4S and GalNAc6S. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (6), 1242-1249.
217. 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.
218. Köhling, S.; Blaszkiewicz, J.; Ruiz-Gómez, G.; Fernández-Bachiller, M. I.; Lemmnitzer, K.; Panitz, N.; Beck-Sickinger, A. G.; Schiller, J.; Pisabarro, M. T.; Rademann, J., Syntheses of defined sulfated oligohyaluronans reveal structural effects, diversity and thermodynamics of GAG–protein binding. *Chem. Sci.* **2019**, *10* (3), 866-878.
219. Struwe, W. B.; Baldauf, C.; Hofmann, J.; Rudd, P. M.; Pagel, K., Ion mobility separation of deprotonated oligosaccharide isomers - evidence for gas-phase charge migration. *Chem. Comm.* **2016**, *52* (83), 12353-12356.
220. Hogan, J. D.; Klein, J. A.; Wu, J.; Chopra, P.; Boons, G.-J.; Carvalho, L.; Lin, C.; Zaia, J., Software for Peak Finding and Elemental Composition Assignment for Glycosaminoglycan Tandem Mass Spectra. *Mol. Cell Proteomics* **2018**, *17* (7), 1448-1456.
221. Rabenstein, D. L., Heparin and heparan sulfate: structure and function. *Nat. Prod. Rep.* **2002**, *19*, 312-331.
222. Kjellén, L.; Lindahl, U., Specificity of glycosaminoglycan–protein interactions. *Curr. Opin. Struc. Biol.* **2018**, *50*, 101-108.
223. Alocci, D.; Ghraichy, M.; Barletta, E.; Gastaldello, A.; Mariethoz, J.; Lisacek, F., Understanding the glycome: an interactive view of glycosylation from glyco compositions to glycoepitopes. *Glycobiology* **2018**, *28* (6), 349-362.
224. Wolff, J. J.; Chi, L.; Linhardt, R. J.; Amster, I. J., Distinguishing Glucuronic from Iduronic Acid in Glycosaminoglycan Tetrasaccharides by Using Electron Detachment Dissociation. *Anal. Chem.* **2007**, *79*, 2015-2022.
225. Oh, H. B.; Leach, F. E.; Arungundram, S.; Al-Mafraji, K.; Venot, A.; Boons, G.-J.; Amster, I. J., Multivariate Analysis of Electron Detachment Dissociation and Infrared Multiphoton Dissociation Mass Spectra of Heparan Sulfate Tetrasaccharides Differing Only in Hexuronic acid Stereochemistry. *J. Am. Soc. Mass Spectrom.* **2011**, *22* (3), 582-590.

226. Agyekum, I.; Zong, C.; Boons, G.-J.; Amster, I. J., Single Stage Tandem Mass Spectrometry Assignment of the C-5 Uronic Acid Stereochemistry in Heparan Sulfate Tetrasaccharides using Electron Detachment Dissociation. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (9), 1741-1750.
227. Wu, J.; Wei, J.; Chopra, P.; Boons, G.-J.; Lin, C.; Zaia, J., Sequencing Heparan Sulfate Using HILIC LC-NETD-MS/MS. *Anal. Chem.* **2019**, *91* (18), 11738-11746.
228. Kailemia, M. J.; Park, M.; Kaplan, D. A.; Venot, A.; Boons, G.-J.; Li, L.; Linhardt, R. J.; Amster, I. J., High-Field Asymmetric-Waveform Ion Mobility Spectrometry and Electron Detachment Dissociation of Isobaric Mixtures of Glycosaminoglycans. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 258-268.
229. Miller, R. L.; Wei, W.; Schwörer, R.; Zubkova, O. V.; Tyler, P. C.; Turnbull, J. E.; Leary, J. A., Composition, Sequencing and Ion Mobility Mass Spectrometry of Heparan Sulfate-like Octasaccharide Isomers Differing in Glucuronic and Iduronic Acid Content. *Eur. J. Mass Spectrom.* **2015**, *21* (3), 245-254.
230. Schindler, B.; Renois-Predelus, G.; Bagdadi, N.; Melizi, S.; Barnes, L.; Chambert, S.; Allouche, A. R.; Compagnon, I., MS/IR, a new MS-based hyphenated method for analysis of hexuronic acid epimers in glycosaminoglycans. *Glycoconj. J.* **2017**, *34* (3), 421-425.
231. Thomas, D. A.; Marianski, M.; Mucha, E.; Meijer, G.; Johnson, M. A.; von Helden, G., Ground-State Structure of the Proton-Bound Formate Dimer by Cold-Ion Infrared Action Spectroscopy. *Angew. Chem. Int. Ed.* **2018**, *57* (33), 10615-10619.
232. Arungundram, S.; Al-Mafraji, K.; Asong, J.; Leach, F. E.; Amster, I. J.; Venot, A.; Turnbull, J. E.; Boons, G.-J., Modular Synthesis of Heparan Sulfate Oligosaccharides for Structure–Activity Relationship Studies. *J. Am. Chem. Soc.* **2009**, *131* (47), 17394-17405.
233. Thomas, D. A.; Mucha, E.; Lettow, M.; Meijer, G.; Rossi, M.; von Helden, G., Characterization of a trans-trans Carbonic Acid-Fluoride Complex by Infrared Action Spectroscopy in Helium Nanodroplets. *J. Am. Chem. Soc.* **2019**, *141* (14), 5815-5823.

234. Supady, A.; Blum, V.; Baldauf, C., First-Principles Molecular Structure Search with a Genetic Algorithm. *J. Chem. Inf. Model.* **2015**, *55* (11), 2338-48.
235. Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M., Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **2009**, *180* (11), 2175-2196.
236. Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865-3868.
237. Tkatchenko, A.; Scheffler, M., Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.* **2009**, *102* (7), 073005.
238. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsujii, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16 (Revision A.03)*, Guassian Inc.: Wallingford, CT, 2016.
239. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H., A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104.
240. Weigend, F.; Ahlrichs, R., Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297-305.
241. Zheng, J.; Xu, X.; Truhlar, D. G., Minimally augmented Karlsruhe basis sets. *Theor. Chem. Acc.* **2010**, *128* (3), 295-305.

242. Hwang, H.-Y.; Olson, S. K.; Esko, J. D.; Robert Horvitz, H., *Caenorhabditis elegans* early embryogenesis and vulval morphogenesis require chondroitin biosynthesis. *Nature* **2003**, *423* (6938), 439-443.
243. van Kuppevelt, T. H.; Oosterhof, A.; Versteeg, E. M. M.; Podhumljak, E.; van de Westerlo, E. M. A.; Daamen, W. F., Sequencing of glycosaminoglycans with potential to interrogate sequence-specific interactions. *Sci. Rep.* **2017**, *7* (1), 14785.
244. Yu, Y.; Duan, J.; Leach, F. E.; Toida, T.; Higashi, K.; Zhang, H.; Zhang, F.; Amster, I. J.; Linhardt, R. J., Sequencing the Dermatan Sulfate Chain of Decorin. *J. Am. Chem. Soc.* **2017**, *139* (46), 16986-16995.
245. Leach, F. E.; Ly, M.; Laremore, T. N.; Wolff, J. J.; Perlow, J.; Linhardt, R. J.; Amster, I. J., Hexuronic Acid Stereochemistry Determination in Chondroitin Sulfate Glycosaminoglycan Oligosaccharides by Electron Detachment Dissociation. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (9), 1488-1497.
246. Lemmitzer, K.; Riemer, T.; Groessl, M.; Süß, R.; Knochenmuss, R.; Schiller, J., Comparison of ion mobility-mass spectrometry and pulsed-field gradient nuclear magnetic resonance spectroscopy for the differentiation of chondroitin sulfate isomers. *Anal. Methods* **2016**, *8* (48), 8483-8491.
247. Poyer, S.; Lopin-Bon, C.; Jacquinet, J.-C.; Salpin, J.-Y.; Daniel, R., Isomer separation and effect of the degree of polymerization on the gas-phase structure of chondroitin sulfate oligosaccharides analyzed by ion mobility and tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2017**, *31* (23), 2003-2010.
248. Solakyildirim, K., Recent advances in glycosaminoglycan analysis by various mass spectrometry techniques. *Anal. Bioanal. Chem.* **2019**, *411* (17), 3731-3741.
249. Zaia, J.; McClellan, J. E.; Costello, C. E., Tandem Mass Spectrometric Determination of the 4S/6S Sulfation Sequence in Chondroitin Sulfate Oligosaccharides. *Anal. Chem.* **2001**, *73* (24), 6030-6039.
250. Zaia, J.; Li, X.-Q.; Chan, S.-Y.; Costello, C. E., Tandem mass spectrometric strategies for determination of sulfation positions and uronic acid epimerization in chondroitin sulfate oligosaccharides. *J. Am. Soc. Mass Spectrom.* **2003**, *14* (11), 1270-1281.

251. Bielik, A. M.; Zaia, J., Multistage tandem mass spectrometry of chondroitin sulfate and dermatan sulfate. *Int. J. Mass Spectrom.* **2011**, *305* (2), 131-137.
252. Zamfir, A. D.; Flangea, C.; Sisu, E.; Serb, A. F.; Dinca, N.; Bruckner, P.; Seidler, D. G., Analysis of novel over- and under-sulfated glycosaminoglycan sequences by enzyme cleavage and multiple stage MS. *Proteomics* **2009**, *9* (13), 3435-3444.
253. Persson, A.; Vorontsov, E.; Larson, G.; Nilsson, J., Glycosaminoglycan Domain Mapping of Cellular Chondroitin/Dermatan Sulfates. *Sci. Rep.* **2020**, *10* (1), 3506.
254. Lord, M. S.; Day, A. J.; Youssef, P.; Zhuo, L.; Watanabe, H.; Caterson, B.; Whitelock, J. M., Sulfation of the bikunin chondroitin sulfate chain determines heavy chain-hyaluronan complex formation. *J. Biol. Chem.* **2013**, *288* (32), 22930-22941.
255. Chi, L.; Wolff, J. J.; Laremore, T. N.; Restaino, O. F.; Xie, J.; Schiraldi, C.; Toida, T.; Amster, I. J.; Linhardt, R. J., Structural Analysis of Bikunin Glycosaminoglycan. *J. Am. Chem. Soc.* **2008**, *130* (8), 2617-2625.
256. Allen, S. J.; Giles, K.; Gilbert, T.; Bush, M. F., Ion mobility mass spectrometry of peptide, protein, and protein complex ions using a radio-frequency confining drift cell. *Analyst* **2016**, *141* (3), 884-891.
257. Warnke, S.; Hoffmann, W.; Seo, J.; De Genst, E.; von Helden, G.; Pagel, K., From Compact to String—The Role of Secondary and Tertiary Structure in Charge-Induced Unzipping of Gas-Phase Proteins. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (4), 638-646.
258. Queiroz, I. N.; Wang, X.; Glushka, J. N.; Santos, G. R.; Valente, A. P.; Prestegard, J. H.; Woods, R. J.; Mourão, P. A.; Pomin, V. H., Impact of sulfation pattern on the conformation and dynamics of sulfated fucan oligosaccharides as revealed by NMR and MD. *Glycobiology* **2014**, *25* (5), 535-547.
259. Racaud, A.; Antoine, R.; Joly, L.; Mesplet, N.; Dugourd, P.; Lemoine, J., Wavelength-tunable ultraviolet photodissociation (UVPD) of heparin-derived disaccharides in a linear ion trap. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (9), 1645-51.

260. Racaud, A.; Allouche, A. R.; Antoine, R.; Lemoine, J.; Dugourd, P., UV electronic excitations in acidic sugars. *Theochem* **2010**, *960* (1-3), 51-56.
261. Cagmat, E. B.; Szczepanski, J.; Pearson, W. L.; Powell, D. H.; Eyler, J. R.; Polfer, N. C., Vibrational signatures of metal-chelated monosaccharide epimers: gas-phase infrared spectroscopy of Rb⁺-tagged glucuronic and iduronic acid. *Phys. Chem. Chem. Phys.* **2010**, *12* (14), 3474-9.
262. Breiman, L., Random Forests. *Mach. Learn.* **2001**, *45* (1), 5-32.
263. Meza Ramirez, C. A.; Greenop, M.; Ashton, L.; Rehman, I. u., Applications of machine learning in spectroscopy. *Appl. Spectrosc. Rev.* **2021**, *56* (8-10), 733-763.
264. Menze, B. H.; Petrich, W.; Hamprecht, F. A., Multivariate feature selection and hierarchical classification for infrared spectroscopy: serum-based detection of bovine spongiform encephalopathy. *Anal. Bioanal. Chem.* **2007**, *387* (5), 1801-1807.
265. Smith, B. R.; Ashton, K. M.; Brodbelt, A.; Dawson, T.; Jenkinson, M. D.; Hunt, N. T.; Palmer, D. S.; Baker, M. J., Combining random forest and 2D correlation analysis to identify serum spectral signatures for neuro-oncology. *Analyst* **2016**, *141* (12), 3668-3678.
266. Živanović, V.; Seifert, S.; Drescher, D.; Schrade, P.; Werner, S.; Guttman, P.; Szekeres, G. P.; Bachmann, S.; Schneider, G.; Arenz, C.; Kneipp, J., Optical Nanosensing of Lipid Accumulation due to Enzyme Inhibition in Live Cells. *ACS Nano* **2019**, *13* (8), 9363-9375.
267. Schworer, R.; Zubkova, O. V.; Turnbull, J. E.; Tyler, P. C., Synthesis of a targeted library of heparan sulfate hexa- to dodecasaccharides as inhibitors of beta-secretase: potential therapeutics for Alzheimer's disease. *Chem. Eur. J.* **2013**, *19* (21), 6817-23.
268. sklearn-genetic-opt sklearn-genetic-opt. <https://sklearn-genetic-opt.readthedocs.io> (accessed 02/05/2022).
269. Fortin, F.-A. D. R., F.-M.; Gardner, M.-A. G.; Parizeau, M.; Gagné, C., DEAP: Evolutionary algorithms made easy. *J. Mach. Learn. Res.* **2012**, *13*, 2171-2175.
270. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.;

- Duchesnay, É., Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12* (null), 2825–2830.
271. Hansen, T.; Elferink, H.; van Hengst, J. M. A.; Houthuijs, K. J.; Remmerswaal, W. A.; Kromm, A.; Berden, G.; van der Vorm, S.; Rijs, A. M.; Overkleeft, H. S.; Filippov, D. V.; Rutjes, F. P. J. T.; van der Marel, G. A.; Martens, J.; Oomens, J.; Codée, J. D. C.; Boltje, T. J., Characterization of glycosyl dioxolenium ions and their role in glycosylation reactions. *Nat. Commun.* **2020**, *11* (1), 2664.
272. Martens, J.; Berden, G.; van Outersterp, R. E.; Kluijtmans, L. A. J.; Engelke, U. F.; van Karnebeek, C. D. M.; Wevers, R. A.; Oomens, J., Molecular identification in metabolomics using infrared ion spectroscopy. *Sci. Rep.* **2017**, *7* (1), 3363.
273. Bush, M. F.; Forbes, M. W.; Jockusch, R. A.; Oomens, J.; Polfer, N. C.; Saykally, R. J.; Williams, E. R., Infrared Spectroscopy of Cationized Lysine and ϵ -N-methyllysine in the Gas Phase: Effects of Alkali-Metal Ion Size and Proton Affinity on Zwitterion Stability. *J. Phys. Chem. A* **2007**, *111* (32), 7753-7760.
274. Bush, M. F.; Oomens, J.; Williams, E. R., Proton Affinity and Zwitterion Stability: New Results from Infrared Spectroscopy and Theory of Cationized Lysine and Analogues in the Gas Phase. *J. Phys. Chem. A* **2009**, *113* (2), 431-438.
275. van Outersterp, R. E.; Martens, J.; Peremans, A.; Lamard, L.; Cuyckens, F.; Oomens, J.; Berden, G., Evaluation of table-top lasers for routine infrared ion spectroscopy in the analytical laboratory. *Analyst* **2021**, *146* (23), 7218-7229.

List of Publications

1. Riedel, J.; Lettow, M.; Grabarics, M.; Götze, M.; Miller, R. L.; Boons, G.-J., Meijer, G.; von Helden, G.; Szekeres, G. P.; Pagel, K., Predicting Structural Motifs of Heparan- and Chondroitin Sulfates using Cryogenic Infrared Spectroscopy and Random Forest. *Manuscript in preparation*.
2. Lettow, M.; Mucha, E.; Lambeth, T. R.; Greis, K.; Yaman, M.; Manz, C.; Hoffmann, W.; Meijer, G.; Julian, R. R.; von Helden, G.; Marianski, M.; Pagel, K., Decoding the Fucose Migration Product in Blood Group Epitopes. *Manuscript in preparation*.
3. Lettow, M.; Greis, K.; Grabarics, M.; Horlebein, J.; Miller, R. L.; Meijer, G.; von Helden, G.; Pagel, K., Chondroitin Sulfate Disaccharides in the Gas Phase: Differentiation and Conformational Constraints. *J. Phys. Chem. A* **2021**, *125*, 4373-4379.
4. Kirschbaum, C.; Greis, K.; Lettow, M.; Gewinner, S.; Schöllkopf, W.; Meijer, G.; von Helden, G.; Pagel, K., Non-Covalent Double Bond Sensors for Gas-Phase Infrared Spectroscopy of Unsaturated Fatty Acids. *Anal. Bioanal. Chem.* **2021**, *413*, 3643-3653.
5. Grabarics, M.; Lettow, M.; Kirschbaum, C.; Greis, K.; Manz, C.; Pagel, K., Mass Spectrometry-Based Techniques to Elucidate the Sugar Code. *Chem. Rev.* **2022**, *122*, 7840-7908.
6. Grabarics, M.; Lettow, M.; Kirk, A. T.; von Helden, G.; Causon, T. J.; Pagel, K., Plate-Height Model of Ion Mobility-Mass Spectrometry: Part 2—Peak-to-Peak Resolution and Peak Capacity. *J. Sep. Sci.* **2021**, *44*, 2798-2813.
7. Thomas, D. A.; Chang, R.; Mucha, E.; Lettow, M.; Greis, K.; Gewinner, S.; Schöllkopf, W.; Meijer, G.; von Helden, G., Probing the Conformational Landscape and Thermochemistry of DNA Dinucleotide Anions Via Helium Nanodroplet Infrared Action Spectroscopy. *Phys. Chem. Chem. Phys.* **2020**, *22*, 18400-18413.
8. Lettow, M.; Grabarics, M.; Greis, K.; Mucha, E.; Thomas, D. A.; Chopra, P.; Boons, G. J.; Karlsson, R.; Turnbull, J. E.; Meijer, G.; Miller, R. L.; von Helden, G.; Pagel, K., Cryogenic Infrared Spectroscopy Reveals Structural Modularity in the Vibrational Fingerprints of Heparan Sulfate Diastereomers. *Anal. Chem.* **2020**, *92*, 10228-10232.

9. Lettow, M.; Grabarics, M.; Mucha, E.; Thomas, D. A.; Polewski, L.; Freyse, J.; Rademann, J.; Meijer, G.; von Helden, G.; Pagel, K., IR Action Spectroscopy of Glycosaminoglycan Oligosaccharides. *Anal. Bioanal. Chem.* **2020**, *412*, 533-537.
10. Greis, K.; Mucha, E.; Lettow, M.; Thomas, D. A.; Kirschbaum, C.; Moon, S.; Pardo-Vargas, A.; von Helden, G.; Meijer, G.; Gilmore, K.; Seeberger, P. H.; Pagel, K., The Impact of Leaving Group Anomerism on the Structure of Glycosyl Cations of Protected Galactosides. *Chem. Phys. Chem.* **2020**, *21* (17), 1905-1907.
11. Grabarics, M.; Lettow, M.; Kirk, A. T.; von Helden, G.; Causon, T. J.; Pagel, K., Plate-Height Model of Ion Mobility-Mass Spectrometry. *Analyst* **2020**, *145*, 6313-6333.
12. Thomas, D. A.; Mucha, E.; Lettow, M.; Meijer, G.; Rossi, M.; von Helden, G., Characterization of a Trans-Trans Carbonic Acid-Fluoride Complex by Infrared Action Spectroscopy in Helium Nanodroplets. *J. Am. Chem. Soc.* **2019**, *141*, 5815-5823.
13. Lettow, M.; Mucha, E.; Manz, C.; Thomas, D. A.; Marianski, M.; Meijer, G.; von Helden, G.; Pagel, K., The Role of the Mobile Proton in Fucose Migration. *Anal. Bioanal. Chem.* **2019**, *411*, 4637-4645.
14. Schwaar, T.; Lettow, M.; Remmler, D.; Börner, H. G.; Weller, M. G., Efficient Screening of Combinatorial Peptide Libraries by Spatially Ordered Beads Immobilized on Conventional Glass Slides. *High-Throughput* **2019**, *8*, 11.
15. E. Mucha, M. Lettow, M. Marianski, D. A. Thomas, W. B. Struwe, D. J. Harvey, G. Meijer, P. H. Seeberger, G. v. Helden, K. Pagel, *Angew. Chem. Int. Ed.* **2018**, *57*, 7440-7443.