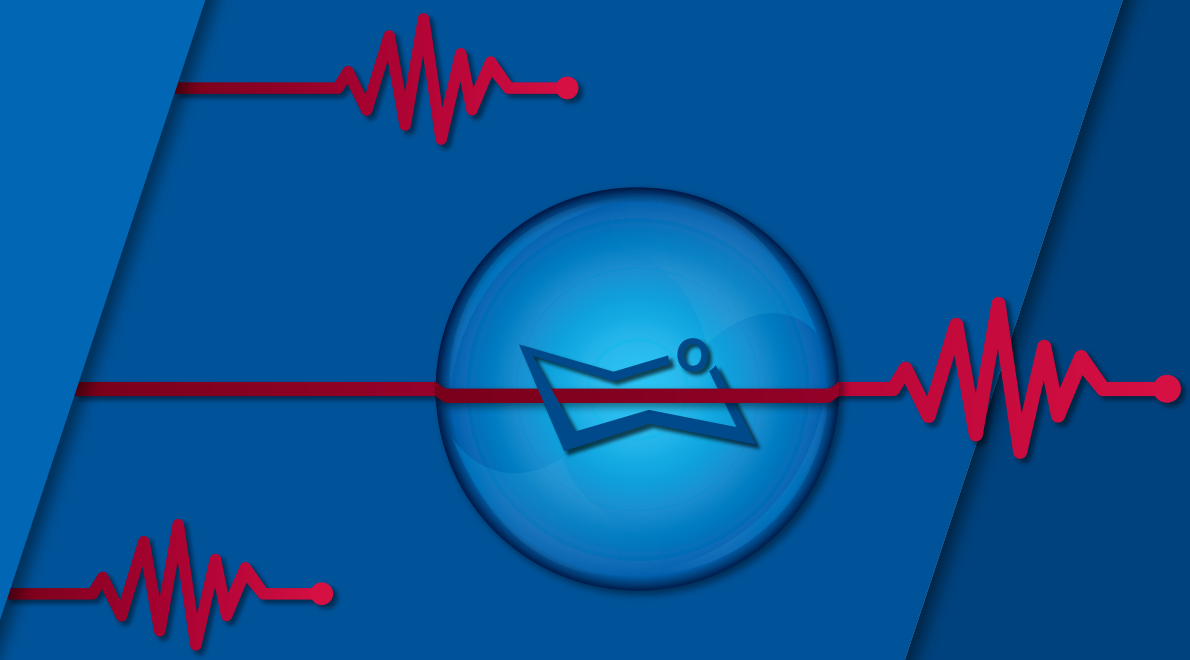
Vibrational Spectroscopy of Glycans in Helium Nanodroplets



Eike Mucha

ISBN: 978-94-028-1837-6

Printing: Ipskamp Printing, Enschede

This thesis was typeset using LaTeX, developed by Leslie Lamport, using the original TeX language designed by Donald E. Knuth. The body text is set in 11pt *Latin Modern Roman*, typographically refined by the *microtype* package. Text in figures and the book cover is set in *Scala Sans bold*, developed by Martin Majoor.

Vibrational Spectroscopy of Glycans in Helium Nanodroplets

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 van decanen
in het openbaar te verdedigen op
vrijdag 14 februari 2020
om 13:30 uur precies

door

Eike Christian Mucha

geboren op 21 februari 1991
te Peine, 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. J. Oomens

Dr. T.J. Boltje

Dr. I. Compagnon
Université de Lyon, Frankrijk

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

Vibrational Spectroscopy of Glycans in Helium Nanodroplets

Doctoral thesis

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 Council of Deans
to be defended in public on
Friday, 14 February 2020
at 13:30 hours

by

Eike Christian Mucha

born on 21 February 1991
in Peine, Germany

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. J. Oomens
	Dr. T.J. Boltje
	Dr. I. Compagnon *Université de Lyon, France*

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

"[...] Jeder von uns dagegen in der Wissenschaft weiß, dass das, was er gearbeitet hat, in 10, 20, 50 Jahren veraltet ist. Das ist das Schicksal, ja: das ist der *Sinn* der Arbeit der Wissenschaft, dem sie, in ganz spezifischem Sinne gegenüber allen anderen Kulturelementen, für die es sonst noch gilt, unterworfen und hingegeben ist: jede wissenschaftliche »Erfüllung« bedeutet neue »Fragen« und *will* überboten werden und veralten. Damit hat sich jeder abzufinden, der der Wissenschaft dienen will. Wissenschaftliche Arbeiten können gewiß dauernd, als Genußmittel, ihrer künstlerischen Qualität wegen, oder als Mittel der Schulung zur Arbeit, wichtig bleiben. Wissenschaftlich aber überholt zu werden, ist – es sei wiederholt – nicht nur unser aller Schicksal, sondern unser aller Zweck. Wir können nicht arbeiten, ohne zu hoffen, daß andere weiter kommen werden als wir."

Max Weber, Wissenschaft als Beruf, ***1919***

Abstract

A central theme among the glycosciences is the Janus-faced nature of glycans. Their tremendous structural diversity enables a myriad of biological functions ranging from energy storage to molecular recognition processes. But at the same time, this structural diversity poses a formidable challenge for glycan analysis that impedes the full development of structural glycobiology. In contrast to genomics and proteomics, glycomics lacks generic sequencing methods that allow reliable, high-throughput analyses with low sample consumption. Instead, a variety of sophisticated methods is used for glycan analysis, including mass spectrometry. A general challenge using mass spectrometry alone, however, is the unambiguous identification of isomeric glycans. Therefore, it is often coupled to orthogonal techniques, such as liquid chromatography. In the last two decades, the combination of mass spectrometry and gas-phase action spectroscopy emerged for glycan analysis. Various challenges, however, limited gas-phase spectroscopy to smaller glycans. In particular, the thermal activation of ions during the measurement in infrared multiple-photon dissociation spectroscopy leads to significant line-broadening, which limits the amount of structural information that can be obtained by this method.

This work overcomes these limitations by combining mass spectrometry and cryogenic vibrational spectroscopy using superfluid helium

nanodroplets. The unique low-temperature environment of helium droplets allows the acquisition of vibrational spectra in the absence of significant thermal contributions. A systematic study of isomeric glycans demonstrates the outstanding resolving power that provides a variety of well-resolved absorption bands that are unique to each isomer: a true spectral fingerprint. The unique optical signatures allow the resolution of even minute structural details, such as the stereochemical orientation of a single hydroxy group.

The exceptional ability of this method to resolve structural details was used to investigate an elusive rearrangement reaction, called fucose migration, which frequently leads to the detection of misleading fragment ions and erroneous sequence assignments in tandem mass spectrometry. Because fucose migration was only observed in fragment ions, it was strictly associated with the fragmentation process. In this work, cryogenic vibrational spectroscopy reveals that fucose migration is not restricted to fragment ions, and instead occurs in intact ions as well. These results generalize fucose migration to a universal issue in mass spectrometry as a whole.

In another study, the combination of cryogenic vibrational spectroscopy and first-principles theory was used to unravel the structure of glycosyl cations, the key intermediates during chemical glycosylation reactions. Various reaction pathways were postulated many years ago, but the exact structure of glycosyl cations remained obscure due to their short-lived and reactive nature in the condensed phase. In this work, the fragmentation of precursor ions was used to generate glycosyl cations in the gas phase, which provides a unique clean-room environment that stabilizes these transient intermediates. The highly resolved vibrational spectra obtained for various glycosyl cations allow an in-depth structural analysis that reveals detailed insights into the two

fundamental structural motifs that enable stereoselective glycosylation reactions: neighboring group participation and remote participation. These results facilitate the mechanistic understanding of glycosylation reactions and will eventually lead to a more rational design of building blocks that is based on structural rather than anecdotal evidence.

Samenvatting

Een centraal thema in de glyco-wetenschappen is het dubbele karakter van glycanen. Enerzijds maakt de enorme structurele diversiteit van glycanen een groot aantal biologische functies mogelijk, variërend van energieopslag tot moleculaire herkenningsprocessen. Maar tegelijkertijd vormt deze structurele diversiteit een enorme uitdaging voor de analyse van glycanen, wat de ontwikkeling van structurele glycobiologie belemmert. In tegenstelling tot genomics en proteomics, ontbreekt het glycomics aan een generieke sequentiemethode die betrouwbare, high-throughput analyses van kleine hoeveelheden glycanen mogelijk maakt. In plaats daarvan wordt een verscheidenheid aan geavanceerde methoden gebruikt, waaronder massaspectrometrie. Met uitsluitend massa-spectrometrische methoden, is een eenduidige identificatie van verschillende isomeren van glycanen echter slechts beperkt mogelijk. Daarom wordt massaspectrometrie vaak gekoppeld aan orthogonale technieken, zoals vloeistofchromatografie. In de laatste twee decennia is ook de combinatie van massaspectrometrie met gasfase-actiespectroscopie opgekomen voor glycaan analyse. Deze combinatie is tot dusver met name gebruikt voor de analyse van kleinere glycanen. In het bijzonder leidt de thermische activering van ionen gedurende infrarood multifoton dissociatie spectroscopie tot een significante lijnverbreding, hetgeen

de hoeveelheid structurele informatie beperkt die met deze methode verkregen kan worden.

Het onderzoek beschreven in dit proefschrift lost deze beperkingen op door massaspectrometrie te combineren met cryogene vibratiespectroscopie, gebruikmakend van superfluïde helium nanodruppels. De lage temperatuur in de heliumdruppels maakt de opname van vibratiespectra mogelijk in afwezigheid van significante thermische bijdragen. Een systematische studie van isomere glycanen toont het uitstekende oplossend vermogen van deze techniek. De spectra bevatten volledig opgeloste absorptiebanden die uniek zijn voor een specifiek isomeer: een echte spectrale vingerafdruk. Deze spectra maken een identificatie mogelijk van zeer subtiele structurele details, zoals de stereochemie van een enkele hydroxygroep.

Het uitzonderlijke vermogen van deze methode om structurele details te onderscheiden is gebruikt om een lastige migratie-reactie, de zogenaamde fucose migratie, te onderzoeken. Deze migratie-reactie leidt vaak tot de detectie van misleidende fragmentionen en foutieve sequentietoewijzingen in tandem massaspectrometrie. Omdat fucose migratie alleen werd waargenomen in fragmentionen, werd het enkel geassocieerd met het fragmentatieproces. Zoals beschreven in dit proefschrift, onthult cryogene vibratiespectroscopie dat fucose migratie niet beperkt is tot fragmentionen maar dat dit ook voorkomt in intacte ionen. Deze resultaten maken duidelijk dat fucose migratie een universele kwestie is in de massaspectrometrie.

In een andere studie is de combinatie van cryogene vibratiespectroscopie en ab initio theorie gebruikt om de structuur van glycosylkationen, de belangrijkste tussenproducten tijdens chemische glycosyleringsreacties, te ontrafelen. Verschillende reactiepaden zijn vele jaren geleden gepostuleerd, maar de exacte structuur van glycosylkationen bleef onduі-

delijk vanwege hun kortstondige en reactieve aard in de gecondenseerde fase. In dit werk is de fragmentatie van precursor-ionen gebruikt om glycosylkationen te genereren in de gasfase. Dit creëert een unieke "clean-room" omgeving die deze tijdelijke tussenproducten stabiliseert. De vibratiespectra verkregen voor verschillende glycosylkationen maken een diepgaande structurele analyse mogelijk. Dit onthult gedetailleerde inzichten in de twee fundamentele structurele motieven die stereoselectieve glycosyleringsreacties mogelijk maken: participatie van naburige groepen en participatie op afstand. Deze resultaten maken een mechanistisch begrip van glycosyleringsreacties mogelijk en zullen uiteindelijk leiden tot een rationeler ontwerp van de benodigde bouwstenen, gebaseerd op structureel in plaats van anekdotisch bewijs.

Contents

Chapter 1

Introduction

Before discussing the fundamentals of glycan structure, helium nanodroplets or vibrational spectroscopy, this chapter gives a global view on how these topics came together in the first place. After covering a general overview and the motivation of this thesis, a brief outline is given.

This chapter is partly based on reference [1] – Mucha, E., et al., In-depth structural analysis of glycans in the gas phase, *Chem. Sci.* **2019**, *10*, 1272.

1.1 Motivation

Glycans? Carbohydrates? Sugars? Oligosaccharides? All these terms are synonymously used to describe one of the four most important classes of biomolecules. They make up most of the organic matter on earth [2] and there is scientific consensus that glycans play key-roles in virtually every biological process. Historically, however, carbohydrates deserved little attention which mainly originated from the belief that they exclusively serve as an energy source, bio-scaffolds or cellular decorations, lacking other biological functions. In comparison to DNA and proteins, glycans can be considered as the scientific version of a 'late bloomer'. From the early days of discovering the double-helix DNA structure [3] to the sequencing of the human genome, [4] biological research mainly focused on the template-driven assembly of biomolecules from genes to RNA to proteins. [5] The biosynthesis of carbohydrates, on the other hand, cannot be directly predicted from DNA, which greatly impeded the understanding of their biological roles. Many different theories have been proposed for the biological functions of glycans, but their immense structural diversity and the lack of analytical tools to accurately determine glycan structures blurred the global view. The outstanding biological importance of carbohydrates became apparent in a milestone publication by Ajit Varki in 1993 titled "*Biological roles of oligosaccharides: all of the theories are correct*". [6] Citing over 1000 references, the author meticulously collected contributions from different fields of research and condensed them in several tables showcasing the diverse biological roles of glycans including inflammation, cell adhesion, molecular recognition and cell-cell interactions.

Today, the immense biological importance of carbohydrates is undisputed and the studies focusing on the structure and function of glycans in an organism were coined *glycomics*. Although glycomics is a

rapidly evolving field of research, the inherent structural diversity of glycans still creates a major analytical challenge and is one of the main reasons why glycomics appears to be in its infancy when compared to the major advances made in genomics and proteomics.

Both, genomics and proteomics greatly benefit from generic sequencing methods allowing for reliable high-throughput strategies with low sample consumption.[7,8] To date, there is no comparable generic sequencing method for glycans. Instead, the structural complexity of glycans is directly reflected in their analysis: prominent examples are nuclear magnetic resonance spectroscopy (NMR),[9,10] sequential mass spectrometry (MS^n),[11,12] capillary electrophoresis (CE),[13,14] ion-mobility mass spectrometry (IM-MS)[15–18] or liquid chromatography (LC)[19,20]. Although all of the aforementioned methods are very useful and established, they have certain limitations and leave gaps in the glycoanalytical workflow. For example, individual limitations originate from purity requirements, sample amounts, timescale, or the depth of structural information.

Another technique for the structural analysis of molecules is vibrational spectroscopy. Using infrared (IR) radiation, characteristic vibrational modes in the molecular fingerprint region can be used to determine the presence of functional groups, inter- and intramolecular interactions, or the molecular conformation. Classical IR spectroscopy utilizes the large analyte density in the condensed phase and measures the attenuation of light caused by the absorption of resonant photons. For detailed structural analyses, however, the condensed phase features several inherent issues. Instead of probing a uniform collection of molecules, a *molecular ensemble* comprising a dynamic mixture of different conformations or charge states is present. Additionally, the often unwanted interactions with the environment, e.g. solvent effects or

possible impurities can obscure the measurement of intrinsic molecular properties.

One way to circumvent these limitations is to transfer analytes from the condensed phase to the *clean room environment* of the gas phase inside a mass spectrometer. Mass spectrometry (MS) allows to isolate a particular species upon ionization from a complex mixture in solution and enables high sensitivity, low sample consumption and high-throughput measurements. These advantages are the main reason why MS-applications are a popular method of choice in many analytical tasks. A spectroscopic analysis of isolated ions, however, is not readily possible using classical absorption spectroscopy because the required large analyte density cannot be obtained in the gas phase where Coulomb repulsion limits the ion density to around 10^6 ions per cubic centimeter.[21] Instead, high-intensity laser radiation is used to enable *action spectroscopy*. In contrast to classical absorption spectroscopy that monitors how the intensity of the light is affected by the analyte, action spectroscopy monitors how the analyte is affected by the laser radiation. Here, the absorption of photons is indirectly measured for example by following the fragmentation yield of the parent ion. Combining the advantages of both, MS and vibrational spectroscopy, this approach allows to gain detailed structural information of isolated ions. As a result, gas-phase vibrational spectroscopy is widely used to elucidate the structure of small molecules, nucleic acids, peptides or proteins.[22] But is this method useful to analyze complex glycans?

This question was addressed by the group of J. P. Simons in the early 2000s pioneering the early work of spectroscopic glycan analysis in the gas phase. Using UV-IR double resonance spectroscopy, smaller carbohydrates were investigated in the low temperature environment of a free-jet expansion.[23–26] The obtained spectral signatures allowed

a full conformational assignment that revealed structural preferences and the importance of hydrogen bonding networks. However, this experimental technique is limited by the necessity of a covalently bound UV-chromophore, which is usually not found in naturally occurring glycans and therefore requires chemical derivatization. Furthermore, most analytes were transferred to the gas phase via evaporation from an oven or laser vaporization, which limits the method to smaller molecules.

Another widely used spectroscopic technique that does not require chromophores is *infrared multiple photon dissociation* (IRMPD). In combination with electrospray ionization (ESI) [27] to transfer ions into vacuum, it allows to study larger molecules or molecular clusters. In IRMPD, high-intensity laser radiation can lead to the sequential absorption of multiple photons and in turn to the fragmentation of the molecule. Monitoring the fragmentation yield as a function of wavelength generates an IR spectrum. With the advance of affordable, tunable high-intensity benchtop laser systems (OPO/OPA), IRMPD spectroscopy emerged as a widely-used spectroscopic technique. Over the years, this method was successfully used to analyze a wide range of molecules, including biomolecules such as amino acids, [28–30] peptides [31,32] and proteins, [33,34] as well as metabolites. [35–38] Most recently, several studies used IRMPD spectroscopy to investigate glycans and for example accomplished to differentiate isomeric species based on their spectral signatures, [39–41] assign the molecular conformation, [42–44] or specify the sulfation pattern. [45,46]

However, a common issue in IRMPD spectroscopy is spectral congestion due to peak broadening that partly arises from the thermal activation of the ions during the sequential absorption of multiple photons. In addition, the conformational flexibility of larger glycans at room temperature likely promotes the population of several coexisting

conformers with different vibrational signatures. These thermal contributions result in broad absorption features that reduce the informational content, and therefore limit IRMPD spectroscopy to smaller glycans and their fragments.

The motivation for this thesis was to overcome the limitations encountered in IRMPD spectroscopy by performing vibrational spectroscopy of glycan ions in the cryogenic environment of helium nanodroplets. In this work, the superfluid nature of helium nanodroplets is used to provide an ideal matrix that allows to study molecular ions near the absolute zero of temperature. The resulting absence of any significant thermal contributions reduces spectral line broadening, resulting in highly resolved vibrational signatures.

1.2 Outline of this Thesis

The fundamentals of glycan structure and their analysis, as well as the experimental basics of superfluid helium and vibrational spectroscopy are covered in **Chapter 2**. Details concerning the *modus operandi* of the experimental setup and the Fritz Haber Institute free-electron laser are presented in **Chapter 3**. In the past, the high resolving power of this setup was used to assign the conformation of peptides and reveal a charge-induced structure transition in proteins. For glycan analysis, this method could yield characteristic spectral fingerprints allowing for a differentiation of isomeric glycan structures. To assess the potential of cryogenic vibrational spectroscopy for glycan analysis, a systematic investigation of different types of isomeric glycans is presented in **Chapter 4**. The resulting vibrational fingerprints exhibit a wealth of well-resolved absorption bands that are diagnostic for minute structural variations. These observations show that cryogenic vibrational spec-

troscopy is a valuable addition to the structural analysis toolbox for carbohydrates.

In **Chapter 5**, this method was then applied to shed light on a phenomenon observed in tandem mass spectrometry experiments of fucosylated glycans called *fucose migration*. The obtained results reveal that fucose migration is not strictly limited to tandem mass spectrometry but can be considered as a general issue for mass spectrometry experiments as a whole.

This setup can also be used to investigate the structure of glycosyl cations that are believed to be the key intermediates in glycosylation reactions. The exact structure of these important cationic intermediates remained obscure due their transient nature in classical condensed-phase experiments. **Chapter 6** provides an in-depth structural analysis of glycosyl cations using a combination of cryogenic vibrational spectroscopy and first-principles theory.

Chapter 2

Fundamentals

The following chapter is intended to familiarize the reader with the most important fundamentals of glycans, superfluid helium and vibrational spectroscopy in order to provide a basis for the experimental setup, the obtained results and their implications in the subsequent chapters.

This chapter is partly based on reference [1] – Mucha, E., et al., In-depth Structural Analysis of Glycans in the Gas Phase, *Chem. Sci.* **2019**, *10*, 1272.

2.1 The Structure of Glycans

The term *carbohydrates* historically evolved because many compounds of this important class of biomolecules are literary hydrates of carbon following the chemical formula $(CH_2O)_n$. Today, this term has a broader definition and also includes derivatives that, for example, carry additional functional groups as shown in Figure 2.1. Smaller carbohydrates are often referred to as sugars, while larger carbohydrates or conjugates are often called glycans or (poly-/oligo-)saccharides.

The smallest units in glycans are monosaccharides. They are grouped – according to the number of carbon atoms – into tetroses, pentoses, hexoses or heptoses, containing four, five, six and seven carbon atoms, respectively. These classes are subdivided into aldoses, containing an aldehyde group, or ketoses, containing a ketone. Each class furthermore comprises a variety of possible monosaccharides that differ in the orientation of their hydroxy groups. The stereoconfiguration of the four hydroxy groups in aldohexoses, for example, lead to a total of $2^4 = 16$ isomers, such as galactose, glucose or mannose.

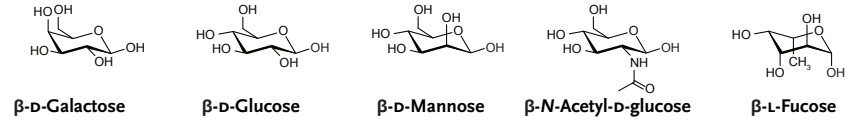


Figure 2.1: Examples for important monosaccharide building blocks. All monosaccharides are aldohexoses adopting a pyranose ring structure and are commonly found in nature.

Because monosaccharides typically lack an axis of improper rotation, they are chiral molecules. The corresponding enantiomers are labeled as L- and D-configurations. Naturally occurring carbohydrates are almost exclusively found in the D-configuration, with fucose as an important exception. Carbohydrates can be represented as an open-chain or a ring structure as shown in Figure 2.2. Typically, ring structures are

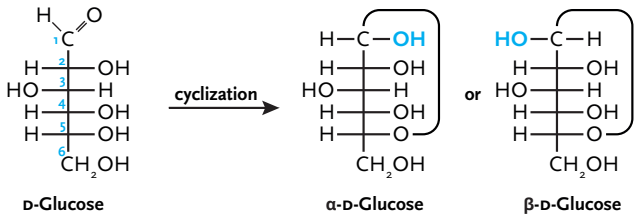


Figure 2.2: Carbohydrates can adopt open-chain or ring structures. The enthalpically favored cyclization forming pyranose structures is shown for D-glucose using the Fischer representation. The reaction with the prochiral aldehyde leads to two diastereomers denoted as α- and β-D-glucose.

enthalpically favored and are formed by a covalent bond between the hydroxy group at position 4 or 5 with the prochiral aldehyde at position 1. When a five-membered ring is formed, the structure is called furanose, while six-membered rings are called pyranose. The cyclization reaction forms a new stereocenter called the *anomeric center*. Depending on the orientation of the hydroxy group at the anomeric center relative to the oxygen atom at the reference (C5 for this example), two configurations can occur: a *cis*-relation yields the α-anomer, while a *trans*-relation yields the β-anomer. Taken together, we would categorize the glucose structure shown in Figure 2.1 as an aldohexose with a pyranose ring structure in a β-D-configuration.

In aqueous solution, α- and β-anomers can interconvert into each other in a process called mutarotation. Glucose is found to have an α:β-ratio of 36:64 in equilibrium at room temperature. From a purely sterical point of view, the high abundance of α-anomers seems surprising, because the 1,3-*syn*-diaxial repulsion generated by the axially oriented hydroxy group is energetically unfavorable. The stereoelectronic explanation for this phenomenon has come to be known as the *anomeric effect*, which stabilizes the axial substituent via hyperconjugation between the ring oxygen and the σ^*-orbital of the axial C-X bond.[47,48]

Figure 2.3: Pyranose structures can adopt various conformations. The 38 unique ways to orient pyranose rings are called puckers and were defined by Cremer and Pople.[49] These puckers contain a reference plane of at least four atoms and are grouped into five families abbreviated by a single letter: chairs (C), boats (B), envelopes (E), half-chairs (H) and skews (S).

To minimize steric repulsion between adjacent substituents, pyranose ring structures are not planar, but puckered. The 38 unique ways to arrange pyranose puckers were defined by Cremer and Pople[49] and can be grouped into five families with a single letter designation: chairs (C), boats (B), envelopes (E), half-chairs (H) and skews (S), see Figure 2.3. Every pucker contains a reference plane of at least four atoms and every atom above or below this plane are noted in superscript or subscript together with the single letter designation, e.g. ${}^{4}C_{1}$, or ${}^{1,4}B$.

Now that we have discussed the structural details of monosaccharides, we can take a closer look at how these building blocks can be linked together to create poly- or oligosaccharides. The four key elements that define the structure of glycans are the configuration, composition, connectivity and branching, as shown in Figure 2.4.

When a glycosidic bond is formed between two monosaccharides, the configuration of the anomeric carbon is no longer in dynamic equilibrium but fixed due to the condensation reaction that transforms the hemiacetal to an acetal. Because monosaccharides typically contain more than one hydroxy group, the glycosidic bond can be formed at multiple positions, referred to as connectivity. The composition is given by the identity of the individual monosaccharides. Finally, the presence of multiple hydroxy group allows the formation of branched structures.

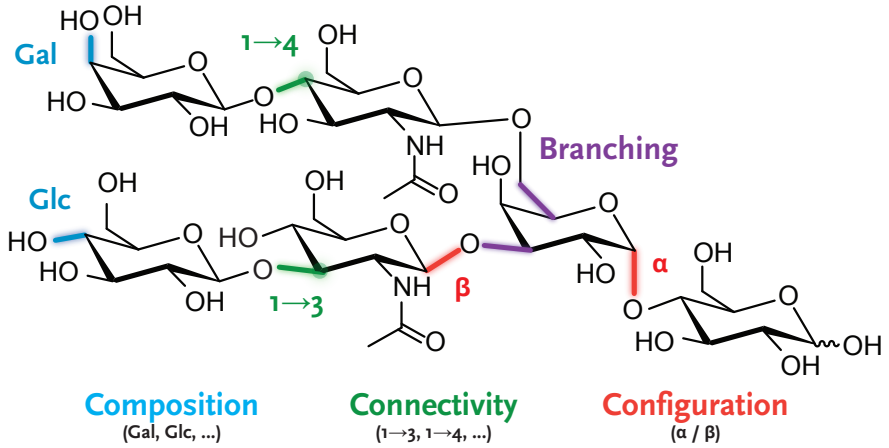


Figure 2.4: A generic glycan illustrating the structural details of glycans. The composition of a glycan is defined by the identity of the monosaccharide building blocks. The connectivity indicates which hydroxy group forms the glycosidic bond. Each glycosidic bond can adopt an α- or β-configuration. In addition, glycan structures are often characterized by branched structures.

All these elements contribute to the vast structural diversity of glycans when compared to DNA or proteins.

In a fundamental way, the structural diversity of glycans poses an immediate problem: depicting the chemical structure of a larger glycan is very accurate, but too complex to decipher, even for the trained eye. To facilitate visual recognition and simplify the complexity of glycan structures, various notations have been developed. Today, the symbol nomenclature for glycans (SNFG) [50] is widely used to depict glycan structures as shown in Figure 2.5 and will be applied throughout this thesis. It is based on symbols with different shapes and colors to depict individual monosaccharide building blocks. Glycosidic bonds are depicted as solid or dashed lines to indicate a β- or α-configuration,

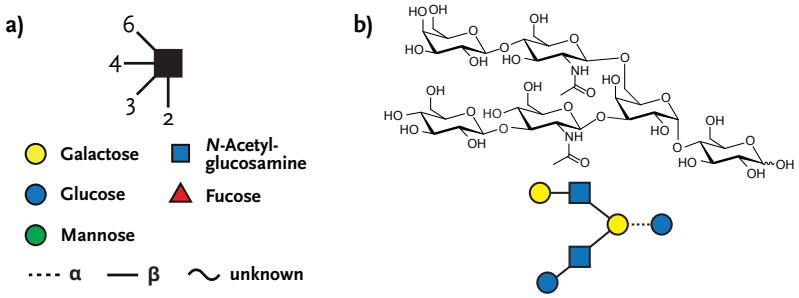


Figure 2.5: a) The symbol nomenclature for glycans (SNFG) [50] is commonly used to depict glycan structures in a simplified manner. The composition of each monosaccharide is given by a combination of different shapes and colors. A glycosidic bond is represented as a dashed or solid line indicating an α- or β-configuration. The angle of the connecting line represents the connectivity. b) The generic glycan structure previously shown in Figure 2.4 is simplified using the SNFG.

respectively. The angle of each connecting line is used to indicate the connectivity.

2.2 The Biological Roles of Glycans

The structural diversity of glycans is directly reflected in the multitude of their biological functions. On a macroscopic scale, polysaccharides provide structural scaffolds and strength such as cellulose in plants or chitin in the exoskeleton of arthropods. On a molecular scale, the biological role of glycans is so diverse, that they virtually play major roles in every biological process and many diseases.[51] In many of those processes, glycans are attached to proteins or lipids and form so-called *glycoconjugates*. In fact, glycosylation is the most common post-translational modification of proteins, and it is estimated that over 50% of all proteins in nature are glycosylated.[52] Glycoconjugates are also highly abundant on the surface of cells where they can form a dense 'sugar shell' called *glycocalyx*.[53] But what is the function of these glycans found on proteins or the surface of cells?

The historical perception of these glycans being merely decorative didn't hold very long, and research gradually unraveled their significance. In the endoplasmic reticulum, for example, protein-bound oligosaccharides induce protein folding by enzymes and aid the degradation of misfolded proteins.[54,55] Viral proteins often develop a high degree of glycosylation to evade the detection of host antibodies and suppress an immune response.[56–58] A prominent example is the envelope spike protein (Env) of the type-1 human immunodeficiency virus (HIV-1). The extensive glycan shield comprises about half of the glycoprotein's mass and effectively protects the protein surface from the immune system. In addition, the glycan shield dynamically evolves by implementing oligosaccharides synthesized by the host's glycosylation machinery.[59–61]

This molecular mimicry of the Env protein implies another crucial biological function of glycans: the specific recognition of glycans by other molecules. On the surface of cells, the interaction between glycans and

glycan-binding proteins thereby enables cell-matrix interactions,[62,63] cell adhesion,[64,65] as well as cell-signaling and communication between cells.[66–68] During the inflammation response, for example, activated endothelial cells express the glycoproteins P-selectin and E-selectin, which interact with leukocytes traversing the blood vessel. The specific interaction is based on the terminal tetrasaccharide sialyl-Lewis^x (sLe^x) and its derivatives, which are present on the surface of leukocytes. This interaction initiates a rolling motion of the leukocyte on the endothelium, which is followed by a firm adhesion that eventually enables the leukocyte to pass through the endothelium.[69,70]

Another important function of glycans is found in the classification of human blood types. The surface of red blood cells is covered with glycoproteins carrying the trisaccharides of the ABO blood group antigens. An individual can either carry the antigen A trisaccharide, the antigen B trisaccharide, both or neither, which *ipso facto* determines the blood group. In turn, the immune system develops specific antigens that target the ABO blood group antigens which are *not* present on the individuals red blood cells.[71,72]

From a pharmaceutical point of view, the wide range of biological functions render glycans as an important molecular class for the design of vaccines,[73–75] diagnostics,[76,77] and therapeutics[78–80] for many diseases. As an example, oligosaccharides of the Lewis antigen family can be used as potential biomarkers for cancer treatment. These antigens are structurally related to the ABO blood group antigens and are typically widely distributed throughout the human body, such as the Lewis b tetrasaccharide. The corresponding isomer Lewis y, on the other hand, is overexpressed on a wide range of tumors and can serve as a potential biomarker for diagnostic and drug targeting purposes.[81–83]

Overexpression of specific glycan motifs is also observed for other diseases and often considered as a mere byproduct of inflammation processes or tumor growth. For example, overexpression of the tetrasaccharide sialyl-lewis^a is often used for diagnosis and monitoring of pancreatitis and pancreatic cancer.[84–86] In a recent study of Engle and coworkers, however, expression of the sialyl-lewis^a tetrasaccharide in mice was observed to *cause* rapid and sever pancreatitis. In the presence of a mutated KRAS oncogene, which enhances cell proliferation, the tetrasaccharide also significantly accelerated the development of pancreatic cancer.[87]

2.3 Superfluid Helium

Helium is the second lightest and second most abundant element in the observable universe after hydrogen, and is found in two stable isotopes: ^{3}He and ^{4}He. Within this thesis, 'helium' refers to the most abundant isotope ^{4}He, which is of bosonic nature. The low-abundance (2 ppm) fermionic isotope ^{3}He has different physical properties and will not be further discussed.

Helium is the only element that remains liquid at the absolute zero of temperature and ambient pressure due to its high zero point energy. When cooled below the liquid/gas phase transition temperature of 4.2 K, helium does not solidify, but undergoes a phase transition at 2.17 K. This phase transition manifests itself in the discontinuity of the specific heat capacity, which divides liquid helium into the two phases Helium I and Helium II.[88–90] As shown in Figure 2.6, plotting the specific heat capacity as a function of temperature results in a curve, which resembles the Greek letter λ. The temperature at which the phase transition occurs was therefore coined the *lambda point*. At temperatures above the lambda point, Helium I is a colorless liquid with very low viscosity and density. At temperatures below the lambda point, liquid helium II exhibits remarkable characteristics. Most notable is the exceptionally high heat conductivity and the vanishingly low viscosity. In the current theoretical model, helium II is considered as a mixture of a normal fluid component, carrying a viscosity and entropy, and a superfluid component, carrying no viscosity and no entropy. The ratio between these two states is determined by the temperature: At 0 K, all helium atoms are of superfluid nature, while at 2.17 K, all Helium atoms are non-superfluid.[91,92]

The low temperature of superfluid helium, which can be maintained by evaporative cooling, as well as its weakly interacting nature inspired

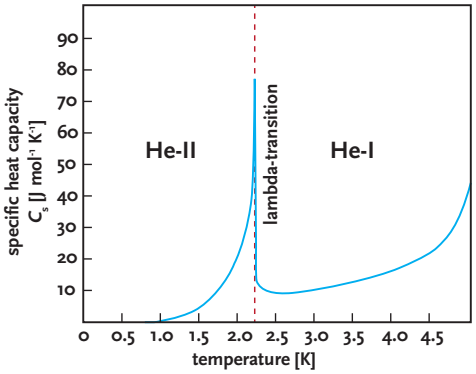


Figure 2.6: The specific heat capacity C_s of liquid helium is characterized by a unique phase transition at 2.17 K resembling the Greek letter λ, hence the name lambda point. The lambda point divides liquid helium into the two phases Helium I and Helium II. Figure adopted from Donelly et al. [90]

many researchers to utilize it as a cryogenic matrix to study atoms or molecules. In conventional matrix isolation techniques, guest molecules or atoms are trapped in an unreactive matrix, usually a solidified noble gas at low temperature. [93–95] Seeding bulk liquid helium with analytes, however, turned out to be very challenging. Because the interaction between the analytes is typically much higher than with the helium atoms, aggregation of the analytes is often unavoidable. In addition, the interaction between the analytes and the vessel material leads to the precipitation of the analyte towards the walls of the vessel. [96]

In 1990, the group of Peter Toennies demonstrated that these limitations can be overcome by employing superfluid helium nanodroplets, which are generated by a supersonic expansion of pressurized helium at low temperatures into vacuum. The beam characteristics are controlled by the stagnation pressure, the temperature of the nozzle, and the valve design. Using this technique, helium nanodroplets with an average size of 10^3 up to 10^{10} helium atoms can be generated. In

their hallmark experiment, the authors used a crossed-beam experiment and observed

"*... that it is possible to attach neutral foreign atoms to helium clusters as they fly through the vacuum*".[97]

The authors recognized the significance of their findings because it offers the possibility to utilize atoms or molecules as probes to study the inherent properties of helium nanodroplets without ionizing the cluster and thereby altering its properties. And indeed, controlled doping of helium droplets was later used to study properties of helium droplets such as superfluidity,[98] quantum dynamics[99] or quantum solvation.[100] Instead of utilizing atoms or molecules to study the properties of helium droplets, scientists realized that they could utilize helium nanodroplets to study the isolated species picked up by them using various spectroscopic techniques. This will be further discussed in Section 2.5.

2.4 Vibrational Spectroscopy

Vibrational spectroscopy is routinely used to identify unknown molecules, to gain knowledge about structural details or to elucidate the molecular conformation. In general, a molecules with N atoms has $3N - 6$ vibrational modes, $3N - 5$ in the case of a linear molecule. A prerequisite for vibrational spectroscopy is given by a selection rule: a vibrational mode will only be IR active if it changes the electric dipole moment of the molecule. If the frequency of the IR radiation matches the vibrational frequency of an IR active mode, the absorption of resonant photons can occur, leading to vibrational excitation of the mode.

Spectroscopists around the world use different physical quantities to describe electromagnetic radiation. Research fields using high energy photons agreed to use the photon energy in multiples of electron volts (eV), while other disciplines typically use the wavelength, for example in micrometers, or the frequency in GHz or THz. Vibrational spectroscopists, on the other hand, felt that they needed something more special. From all possible physical quantities, they eventually decided that *their* ideal unit of measurement for electromagnetic radiation shall be the number of wavelengths that fit into one unit length in vacuum, usually per centimeter* and written as cm^{-1}. After an inconclusive discussion about whether this reciprocal wavelength should be named "Kayser", "Rydberg", "Balmer" or "permicron", the unit was simply called the *wavenumber* with the corresponding symbol $\tilde{\nu}$.[101] Today, the wavenumber is still used in vibrational spectroscopy and the most fundamental vibrations are found in the mid-IR range between $400\,\mathrm{cm}^{-1}$ and $4000\,\mathrm{cm}^{-1}$.

*At least it's metric!

From a theoretical point of view, the bond in a simple diatomic molecule can be represented as a spring obeying Hooke's law $F = -Dx$, where the force F is directly proportional to the springs stiffness D and its displacement x. The quantum theoretical treatment of this law results in an important model in quantum mechanics: the quantum harmonic oscillator.[102] In this model, each oscillator has discrete vibrational energy levels, which are quantized as $E_v = (v + 1/2)h\nu$ with the quantum number $v = 0, 1, 2, \ldots$ and the frequency of oscillation ν multiplied by Planck's constant h. However, every scientific model is a simplified description of the more complex reality and thereby has to fail at some point. The quantum harmonic oscillator fails to describe the anharmonic nature of chemical bonds and for example cannot describe the dissociation of a molecular bond. A more sophisticated model to describe the potential energy is the Morse potential.[103] As a consequence of the anharmonic nature of the chemical bond, the energy levels are no longer equally spaced and the potential is dissociative at an infinite displacement, as shown in Figure 2.7. It is important to note, however, that the Morse potential only applies to diatomic molecules and describes vibrations that involve a change in the internuclear distance. In larger molecules, bending modes with a constant internuclear distance can occur and coupling between different vibrations is possible.

In classical absorption spectroscopy, the attenuation of light that is caused by the absorption of resonant photons is measured. Following the Beer–Lambert law,

$$A(\lambda) = \log_{10} \frac{I_1}{I_0} = -\sigma(\lambda) \cdot cd$$

the wavelength-dependent absorbance $A(\lambda)$ can be expressed using the intensity of the radiation before (I_0) and after (I_1) the propagation

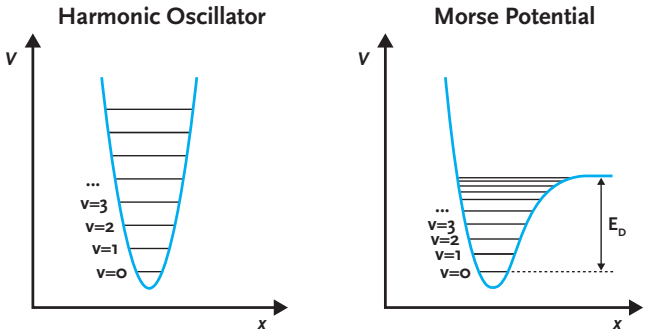


Figure 2.7: The quantum theoretical treatment of Hooke's law leads to the quantum harmonic oscillator with quantized energy levels that are equally spaced. A chemical bond, however, is better described by the Morse potential. Here, energy levels are not equally spaced and a dissociation energy is included.

through the sample. If the concentration c and the length d of the sample are known, the wavelength-dependent absorption cross-section $\sigma(\lambda)$ can be calculated. In the condensed phase, absorption spectroscopy is widely established because the high concentration of solid, liquid or dissolved analytes allow a manageable path length in the order of a few millimeters. Also gaseous samples with a sufficient vapor pressure can be measured using absorption spectroscopy. For isolated ions in the gas phase, however, the density is limited to around 10^6 ions $\cdot$ cm^{-3} due to Coulomb repulsion and the resulting space-charge limit.[21] Compared to the condensed phase, this density would translate to a solution with an analyte concentration of a few attomol per liter. Because of this low density, absorption spectroscopy cannot be applied to ionic species in the gas phase.

To overcome this limitation, a different approach called *action spectroscopy* is used. Instead of measuring how the molecules affect the intensity of radiation, action spectroscopy monitors how the radiation affects the molecules. Analogous to Beer–Lambert's law, the number

of affected molecules N_λ after irradiation at a given wavelength can be expressed as

$$N(\lambda) = N_0(1 - e^{-\sigma(\lambda)\cdot F(\lambda)}),$$

where N_0 is the initial number of molecules and $F(\lambda)$ is the wavelength dependent photon fluence. In contrast to the high analyte density* needed in absorption spectroscopy, action spectroscopy requires a high photon flux, which is typically provided by high-intensity laser radiation.

A widely used strategy in action spectroscopy is called *IR multiple photon dissociation* (IRMPD), which monitors the fragmentation of precursor ions that is induced by the sequential absorption of multiple resonant photons. It is important to note, however, that this process is not a stepwise increase of vibrational levels of a single mode (i.e. $v_0 \rightarrow v_1 \rightarrow v_2 \rightarrow \ldots$) until the dissociation threshold is exceeded. Because the vibrational energy levels are not equally spaced, such a ladder-climbing process would gradually shift the vibrational frequency to lower values and out of resonance with the monochromatic radiation, such that no more photons can be absorbed. This phenomenon is often referred to as the anharmonicity bottleneck.[104] Instead, the absorbed energy is distributed to the bath of vibrational degrees of freedom of the molecule via *intramolecular vibrational redistribution* (IVR). This process is typically very fast (sub-ns)[105,106] and allows for a sequential absorption from the same fundamental frequency, effectively heating the ion (Figure 2.8). When the internal energy exceeds the dissociation threshold, the ion can undergo fragmentation in a unimolecular reaction. Because the fragmentation yield is a function of the absorption cross-section, the ion's IR spectrum can be obtained by plotting the fragmentation yield as a function of the laser wavelength.

*strictly speaking, the large product of analyte density multiplied by the path length.

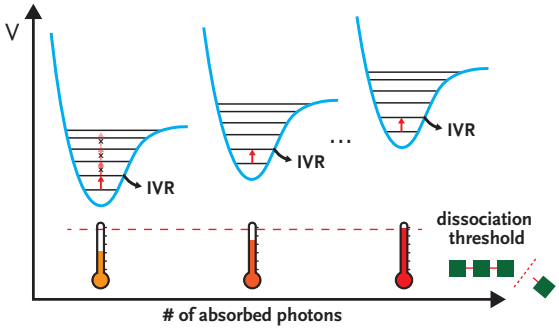


Figure 2.8: Schematic representation of the IR multiple photon dissociation (IRMPD) mechanism. The absorption of multiple resonant photons leads to an increasing internal energy of the ion. After each absorption, the absorbed energy is distributed to the bath of vibrational degrees of freedom via intramolecular vibrational redistribution (IVR). Once the dissociation threshold is exceeded, fragmentation occurs and is monitored to generate an IR spectrum.

A common issue in IRMPD spectroscopy, however, is spectral broadening, which can be attributed to two different causes: (a) the initial temperature of the ion can lead to the superposition of a thermally populated conformational ensemble with different absorption spectra that are averaged during the measurement, and (b), a quasi-continuum of vibrational states is reached at higher internal energies, which leads to the non-selective absorption of photons with large absorption cross-sections.[104] In addition, the anharmonic nature of vibrational modes can lead to cross-anharmonicities at higher internal energies, which can cause red-shifts and non-linear intensities of the measured bands. In an extreme case, an absorption band could shift out of resonance so quickly that it cannot be observed experimentally.[107]

Another spectroscopic technique that circumvents the inherent issues of the IRMPD process is called *messenger spectroscopy*.[108,109] In this method, analyte ions are cooled to cryogenic temperatures and subsequently attached to atoms or molecules (tags) via van der Waals

interactions. These tags are typically noble gas atoms or other non-interacting molecules such as helium, argon or hydrogen. Due to the weak interaction between the analyte and the tag, the absorption of a single resonant photon can lead to the dissociation of the messenger-complex, as schematically shown for a cation M^+ and Helium as a messenger atom:

$$M^+ \xrightarrow{He} M^+ \cdot He \xrightarrow{h\nu} (M^+ \cdot He)^* \xrightarrow{IVR} M^+ + He$$

The subsequent detection of the signal depletion in this single-photon process enables the acquisition of an IR spectrum that can closely resemble a linear absorption spectrum. Furthermore, the cryogenic temperature of the ions and the absence of significant thermal contributions during the measurement can lead to narrow and highly resolved spectral features. Drawbacks of this spectroscopic technique arise if the messenger-complex formation is inefficient or the messenger interacts with the analyte causing a distorted absorption spectrum.

2.5 Vibrational Spectroscopy using Helium Nanodroplets

In 1992, the first spectroscopic investigation of a molecule embedded in helium nanodroplets was reported by Goyal and coworkers.[110] Using a CO_2 laser, they measured the infrared spectrum of sulfur hexafluoride attached to helium nanodroplets and reported its narrow absorption bands. Today, helium droplets are considered as an ideal nano-matrix for spectroscopic investigations. They are optically transparent from deep ultraviolet to far infrared radiation. Due to their superfluid nature, the interactions between the analyte and the helium bath are negligibly small. This allows the assignment of molecular conformations by comparing the experimental IR spectra to simulated IR spectra from quantum chemical calculations. Vice versa, the experimental IR spectra can serve as a benchmark for quantum chemical calculations. Another advantage of using helium droplets as a matrix in optical spectroscopy is their low equilibrium temperature of 0.37 K. Maintaining this cryogenic temperature through evaporative cooling leads to the thermalization of the embedded molecule. In contrast to IRMPD spectroscopy, the low temperature drastically reduces the number of thermally populated states, which can lead to very narrow absorption lines.

Early spectroscopic works focused exclusively on smaller neutral molecules such as OCS,[111] NH_3,[112] H_2O[113] or HCN.[114] The neutral species of interest were embedded in helium nanodroplets using a pickup chamber. This chamber is often referred to as an oven because it thermally evaporates atoms or small molecules. The vaporized neutral species are then picked up by a traversing beam of helium droplets as shown in Figure 2.9a. To analyze the vibrational modes of neutrals embedded in helium droplets, photon absorption is often indirectly

detected by measuring the droplet size after irradiation. The absorption of an IR photon leads to the dissipation of the excitation energy to the helium bath, which in turn leads to the evaporation of helium atoms from the droplet. This decrease in droplet size is then used as a marker for photon absorption and is measured as a depletion spectrum. However, this technique requires small helium droplets in the order of 10^3 to 10^4 helium atoms because the evaporation of a few atoms in larger droplets cannot be detected.[115] This limits depletion spectroscopy to smaller molecules, as larger species usually require larger droplets to be efficiently picked up. Furthermore, evaporation from an oven does not allow transferring larger molecules (especially biomolecules) into vacuum due to thermal decomposition of the sample prior to evaporation.

In this work, molecular ions are embedded in helium nanodroplets for spectroscopic investigations. Spectroscopy of ions in helium droplets was enabled after the discovery that the absorption of resonant photons can lead to the release of the ion from the droplet.[116–118] Analogous to a pickup chamber, the ionic species are stored in an ion trap that is traversed by helium nanodroplets, as shown in Figure 2.9b. The doped droplets are subsequently irradiated with high-intensity laser radiation and the absorption of resonant photons can promote the release of bare, intact ions from helium droplets. The bare ions are then detected in a time of flight mass spectrometer. Using the release of ions as a marker for photon absorption is then used to generate the IR spectrum of the embedded ion by plotting the signal intensity of released ions as a function of laser wavelength.

A more detailed description of the processes involved in this kind of spectroscopy are shown in Figure 2.10. Inside the ion trap, the stored ions are thermalized via collisions with the buffer gas. When they are picked up by a helium droplet, the internal energy is dissipated to the

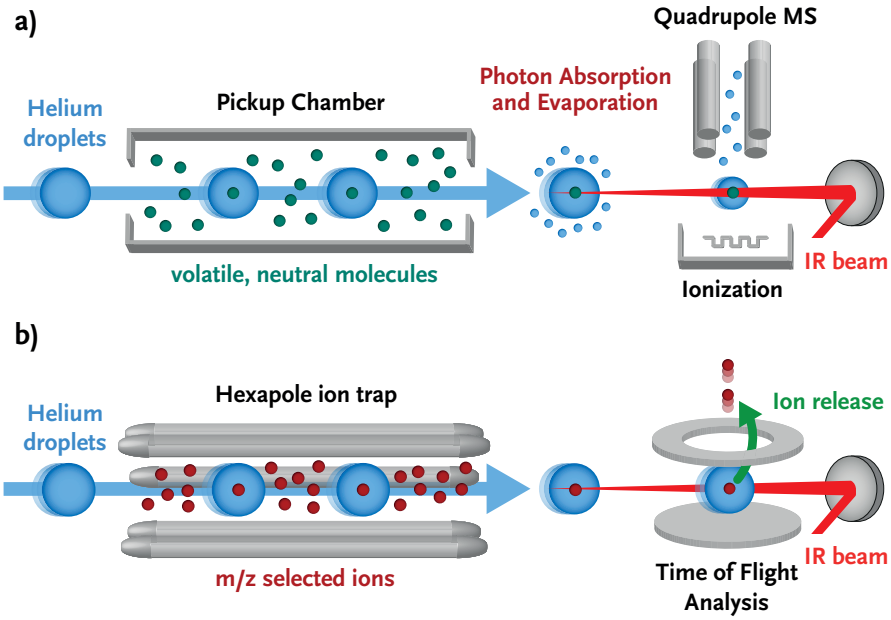


Figure 2.9: a) The pickup of neutral molecules by helium nanodroplets is typically achieved using a pickup-chamber. This chamber contains the molecules of interest and can be heated to produce a sufficient vapor pressure inside the vacuum chamber. The absorption of photons can for example be monitored by detecting the decreasing droplet size caused by evaporative cooling. b) In this work, ions are embedded in helium droplets. Analyte ions are stored inside a hexapole ion trap and can be picked up by a traversing beam of helium droplets. An IR spectrum is obtained by monitoring the number of released ions caused by the resonant absorption of multiple photons.

superfluid helium bath, which boils off helium atoms to maintain its equilibrium temperature of 0.37 K. Equivalently to neutral molecules embedded from a pickup cell, evaporative cooling of the helium droplets leads to an effective cooling of the doped ions. Because the doped droplets have a high kinetic energy, they can overcome the trapping potential of the ion trap and transport the ions downstream the instrument. Inside the detection chamber, the cryogenic ions inside the helium droplets are irradiated with infrared radiation provided by the

Fritz Haber Institute Free-electron laser (FHI FEL[119]). The details of the FHI FEL will be discussed in Section 3.5, but a brief description of the FEL's laser pulse structure will be given here: a single macro-pulse is around 10 µs long and typically contains 10^4 micro-pulses that are equally spaced by one nanosecond. Embedded ions inside the helium droplets can absorb a resonant photon, which leads to vibrational excitation of the corresponding mode. The absorbed energy is dissipated throughout the molecule to vibrational background states by intramolecular vibrational redistribution. The energy is then dissipated to the helium droplet, which leads to heating of the droplet and finally to evaporation of helium atoms and re-thermalization to the equilibrium temperature. The exact pathways and timescales of these processes are not known, but for almost all molecules comprising more than two atoms these timescales will be faster than the arrival of the next laser micro-pulse (1 ns).[105,106] This entails two important consequences: the following micro-pulse will encounter a molecule that is a), in it's vibrational ground state, and b), re-cooled to 0.37 K. Only the droplet size will be slightly smaller due to evaporative cooling. This process can repeat many times and after the successive absorption of multiple photons, the intact ion is released from the droplet and can be detected. It is possible that the observation of bare ions is the result of complete evaporation of the helium droplet. However, experimental data by Drabbels et al. suggest a non-thermal ion ejection from helium droplets.[116] Although the exact mechanism that leads to the detection of bare unsolvated ions is not fully understood and might depend on the initial droplet size, the type of ion, or the excitation process, it is used as a marker for photon absorption.

Even though spectroscopy of ions in helium droplets relies on the absorption of multiple resonant photons, the obtained spectrum is

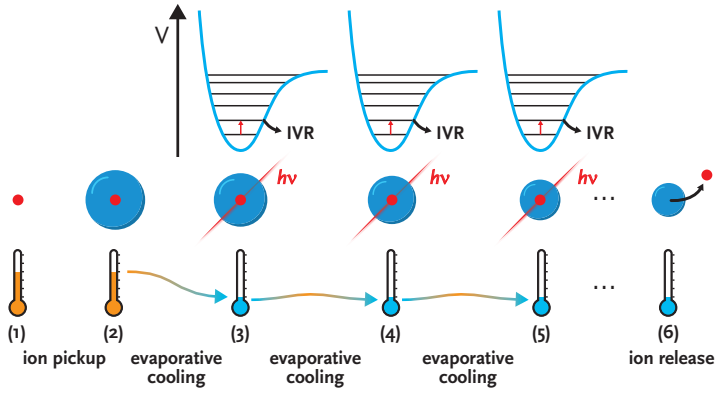


Figure 2.10: Schematic diagram of the mechanisms involved in IR spectroscopy of ions using helium droplets. The ions inside the hexapole ion trap (1) have an initial energy that is rapidly dissipated after pickup by helium droplets (2) that maintain an equilibrium temperature of 0.4 K by evaporative cooling. The absorption of a resonant photon (3) leads to vibrational excitation of the ion. The absorbed energy is quickly dissipated to the bath of vibrational degrees of freedom by intramolecular vibrational redistribution (IVR). Before the next laser micro-pulse arrives, the ion will return to its vibrational ground state by dissipating its energy to the helium droplet. Evaporation of helium atoms allows the helium droplet to re-thermalize to its equilibrium temperature. This process can repeat itself many times and after the successive absorption of multiple resonant photons (4) + (5) the ion is released from the helium droplet (6). The number of released ions is then monitored as a marker for photon absorption.

believed to correspond to the ion's vibrational ground state because of the consecutive re-thermalization events. In contrast to IRMPD spectroscopy, the absence of thermal contributions enables to acquire IR spectra that are not affected by line broadening or peak shifts due to (cross-)anharmonicities. It is important to note, however, that the release of bare ions from helium droplets has a nonlinear dependence on the photon flux. As a result, relative intensities should only be regarded as a general guide.

To date, infrared spectroscopy of ions embedded in helium nanodroplets was successfully used to investigate a variety of different

species including small organic molecules such as aniline[116] and carbonic acid derivatives,[120–122] metal oxide nanoclusters,[107] as well as biomolecules such as peptides[123] and proteins.[124]

Chapter 3

Experimental Methods

This chapter discusses the experimental methods used to investigate glycans in the cryogenic environment of superfluid helium droplets. After a brief overview of the entire instrument, the individual sections as well as the Fritz Haber Institute Free-electron Laser will be discussed separately. Part of this thesis was devoted to further develop the experimental setup. In particular, the design and construction of an optimized variable temperature ion trap and a novel detection scheme for ions after vibrational excitation are presented in sections 3.3 and 3.6.2, respectively.

3.1 Overview

The basic layout of the experimental setup is schematically shown in Figure 3.1 using the original on-axis Wiley-McLaren time-of-flight detection scheme. Ions are generated by nano-electrospray ionization and transferred into vacuum via a Z-Spray inlet-geometry. After passing two ring-electrode ion guides in differentially pumped vacuum chambers, the ions can be $^{m}/_{z}$-selected by a quadrupole mass filter. Next, a quadrupole bender determines the further path of the ions: if no potential is applied to the bender, the ions are directed to the reflectron time-of-flight mass analyzer; if the specified potentials are applied to the bender, the ions are deflected by 90° and injected into a hexapole ion trap. After a few seconds, the trap is filled, and the stored ions can be picked up by a traversing beam of helium droplets. The embedded ions are cooled to the droplet's equilibrium temperature of 0.4 K by evaporative cooling. Because the doped droplets have a velocity of around $450\,\mathrm{m\,s^{-1}}$ and a high mass, the resulting high kinetic energy allows them to overcome the trapping potential and transport the cryogenic ions to the detection region, where they are irradiated by high-intensity laser radiation produced by the Fritz Haber Institute Free-electron Laser (FHI FEL). The absorption of multiple resonant photons can lead to the release of the ion from the droplet, which is used as a messenger for photon absorption. Plotting the ion count as a function of laser wavelength yields the IR signature of the corresponding molecular ion.

In general, the experimental setup can be divided into five sections: (1) the modified *Waters Q-TOF Ultima* including the electrospray ionization source, ion guides, quadrupole mass filter and reflectron TOF, (2) the helium droplet source, (3) the variable temperature ion trap, (4) the detection region after vibrational excitation of the ions, and (5) the

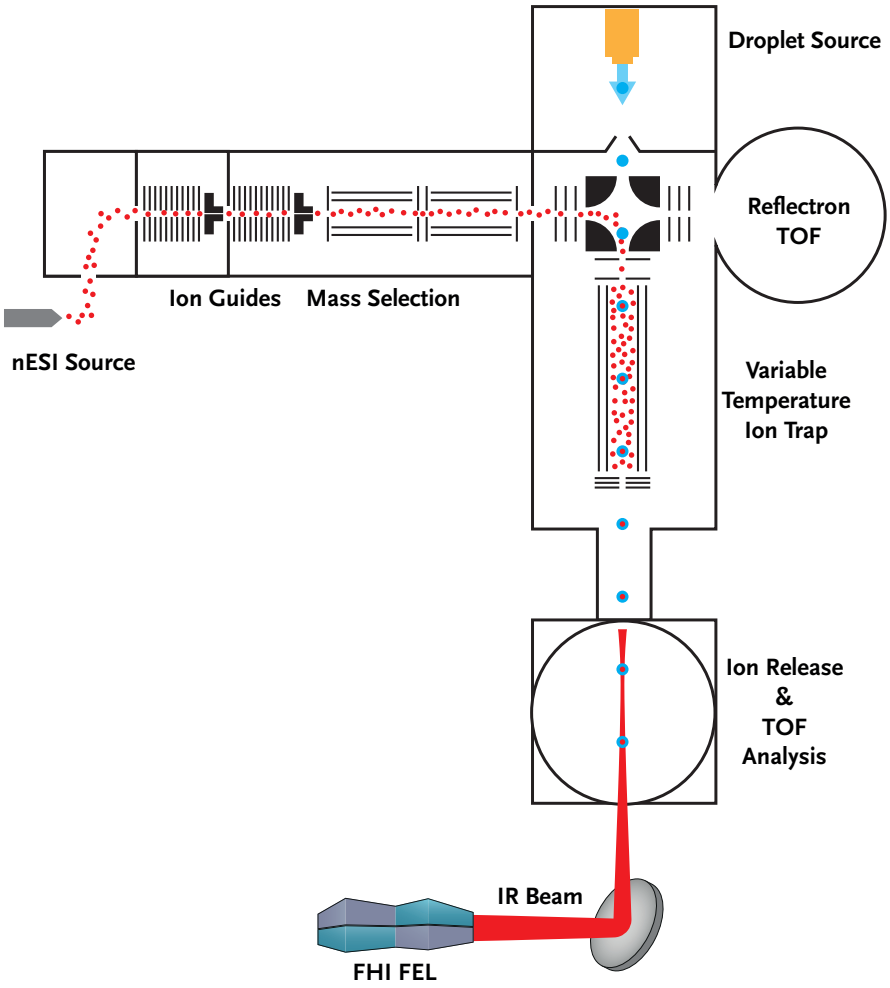


Figure 3.1: Schematic overview of the experimental setup. Ions are generated by nano-electrospray ionization and transferred into high vacuum by two ion guides. The ions can be monitored by a reflectron time-of-flight (TOF) mass spectrometer. Once the ions of interest are isolated by the quadrupole mass filter and the ion current is optimized, they are injected into the hexapole ion trap. The ion trap is traversed by a beam of helium nanodroplets that can pick up trapped ions and thermalize them to 0.4 K. The helium droplets transport the embedded ions to the detection region where they are irradiated by an IR laser beam produced by the Fritz Haber Institute Free-electron Laser (FHI FEL).[119]

FHI FEL as a laser radiation source. These individual components will be described in more detail in the following sections.

3.2 The modified Waters Q-TOF Ultima

A modified *Q-TOF Ultima* (Waters Corporation, Manchester, UK) is used for ion production, transfer of ions into vacuum, mass-to-charge selection and high-resolution mass analysis.

Ion Production

In this work, ions are produced by electrospray ionization (ESI). This technique was originally developed by Dole and coworkers in the early 1970s to produce molecular ions of high molecular weight for mass spectrometry.[125] A detailed analysis of their results, however, was not possible, because a mass spectrometer was not available in his laboratory at this time. Around 20 years later, Fenn and coworkers eventually combined electrospray ionization with a mass spectrometer and further optimized the technique.[27,126] At the time, biomolecules such as peptides and proteins were considered very large in comparison to typical gas-phase ions and could not be brought to the gas phase before the development of ESI. Today, ESI is one of the most widely used ionization techniques in mass spectrometry because it allows the transfer of ions with high molecular weight from the condensed phase to the gas phase without fragmentation. This is why John Fenn received his share of the Nobel prize in 2002 and is often referred to as 'the man who made elephants fly'.[127]

The electrospray process is schematically depicted in Figure 3.2. The analyte solution is brought to the tip of a small capillary where a high voltage is applied. The polarity of the high voltage will determine

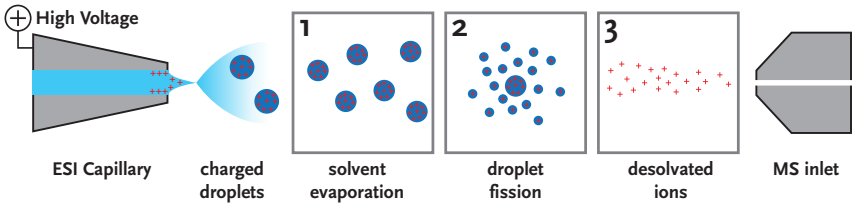


Figure 3.2: Schematic representation of the electrospray ionization process. A high voltage is applied to the capillary that leads to the accumulation of positively charged ions and the formation of a Taylor cone emitting a fine jet of charged droplets. The droplets gradually shrink due to solvent evaporation (1) until Coulomb repulsion between charges causes droplet fission (2). The smaller droplets will further evaporate solvent molecules and undergo Coulomb explosions until the bare unsolvated ions are released (3).

the polarity of the produced ions. For a positive voltage, positive ions are enriched at the surface of the liquid at the capillary tip. The polarization forces at the tip cause a distortion of the meniscus into a so-called Taylor cone.[128,129] When a critical voltage is exceeded, these forces overcome the surface tension of the liquid creating a Taylor cone that emits a fine jet of charged liquid. This jet breaks up into a plume of small charged droplets containing many analyte ions. Due to solvent evaporation, these first generation droplets continuously shrink in size while maintaining the number of charges. At some point, the increasing Coulomb repulsion between charges overcomes the surface tension of the droplet causing fission of the droplet. The maximum number of charges q a droplet of a given diameter D can carry is given by the Rayleigh-limit $q = 8\pi\sqrt{e_0\gamma D^3}$, where γ is the surface tension of the droplet.[130,131] This fission process is also known as *Coulomb explosion* and leads to the formation of smaller droplets that will further evaporate solvent molecules and undergo fission until the bare unsolvated ions are released.[132]

In this work, *nano*-ESI is used for ion generation. Instead of a stainless steel capillary, the spray emitter is a borosilicate glass capillary pulled to a tip diameter of a few µm using a micropipette puller (Model P-1000, Sutter Instrument, Novato, CA, USA). The glass capillary is subsequently coated with a palladium/platinum alloy (80:20, w/w) using a sputter coater (Sputter Coater 108auto, Cressington, UK) to make the surface electrically conductive. The small tip diameter leads to several crucial advantages: a) a very low sample consumption down to 10 nL min^{-1}, b) a more efficient ionization process that increases the sensitivity, and c) a lower voltage of around 1 kV that is required for stable operation.[133,134]

Ion Transfer into Vacuum

Different inlet geometries can be used to transfer ions from atmospheric pressure to the vacuum inside a mass spectrometer. The Q-TOF Ultima is originally equipped with a so-called Z-spray source, schematically shown in Figure 3.3a. In this double orthogonal geometry, ions pass through a skimmer and enter an inlet stage with a pressure of around 10 mbar before entering the first ion guide, which is differentially pumped to around 1 mbar. This geometry efficiently separates ions from neutral molecules, which prevents excessive contamination of the first ion guide. To aid the desolvation of ions or promote in-source fragmentation the source block can be heated.

In addition to the original Z-spray source, a capillary inlet source was custom-built (Figure 3.3b). Here, the ions follow a linear path through a stainless steel capillary of 50 mm length with an inside diameter of 254 µm (Coopers Needle Works Ltd, Birmingham, UK) directly into the vacuum stage of the mass spectrometer. Although the capillary inlet source does not provide the separation of neutral molecules, it

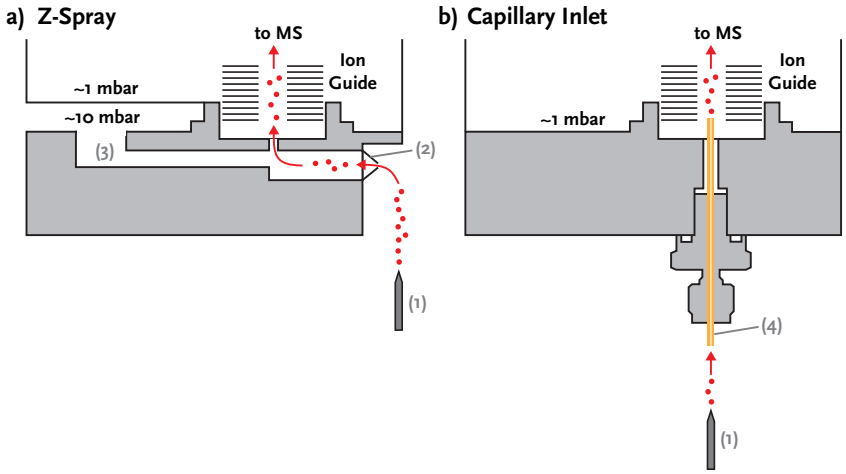


Figure 3.3: Two different inlet geometries can be used to transfer ions from atmospheric pressure to the vacuum inside the mass spectrometer. In both cases a Pd/Pt-coated borosilicate capillary (1) is used for nano-electrospray ionization. a) The commercial instrument is originally equipped with a Z-spray inlet. Here, ions first pass a skimmer (2) and enter an inlet stage (3) before traversing the first ion guide in a differentially pumped vacuum chamber. b) The capillary inlet omits an additional pumping stage and directly transfers ions from atmospheric pressure to the differentially pumped first ion guide using a 50 mm long transfer capillary (4) with an internal diameter of 254 µm.

potentially offers more gentle ion-transfer conditions than the Z-spray source.[135] As a result, the formation of non-covalent clusters can be substantially enhanced. The ion-transfer characteristics are not fully understood, but likely depend on the pressure gradient, the internal diameter of the capillary, air flow and voltage gradients that influence the collision rates and energies with the neutral buffer gas.

Multipole Radio Frequency Devices for Manipulating Ion Motion

Manipulating the motion of ions is crucial for any MS application. Under high-vacuum conditions, ions with high kinetic energy can be efficiently

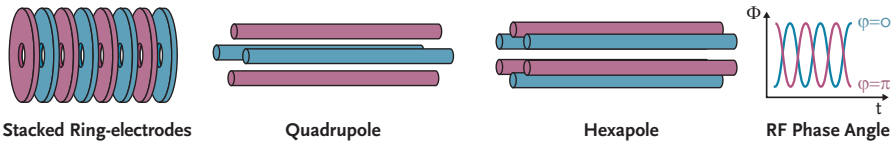


Figure 3.4: Different types of RF multipole devices used in this experimental setup. Stacked ring electrodes are used for the two ion guides that transfer ions into high vacuum as well as another ion guide used in the spectroscopy detection region. A quadrupole is used as a mass filter. A hexapole configuration is found in the collision cell as well as the ion trap. The electrodes are divided into two groups carrying a sinusoidal RF waveform that is shifted by 180° (or π rad) to one another.

transported and focused using static electromagnetic fields. Maintaining stable and directed motion of ionic beams with lower kinetic energy (in the order of a few tens of eV) in static electromagnetic fields, however, is inefficient due to particle collisions at elevated pressures, parasitic magnetic fields, electrode contamination or space charge of the ions.[136] Instead, alternating electric fields generated from radio-frequency (RF) sinusoidal waveform generators are commonly used for efficient ion transmission and focusing. These so-called multipole RF devices can additionally be used to filter ions according to their mass-to-charge ratio ($^{m}/_{z}$-ratio), or store them in ion traps over long time periods. This section describes the general operating principles of multipole RF devices, which will be discussed separately and in more detail in the following sections.

There are two types of electrode arrangements used in this experimental setup, as shown in Figure 3.4. Stacked ring-electrode arrangements are used for the two ion guides in the source region of the instrument as well as another ion guide in the spectroscopy detection region. Multipole rod electrode designs such as quadrupole or hexapole arrangements are used for mass filtering, ion transmission and ion trapping.

Multipole rod-electrode devices usually consist of $2n$ parallel metal rods of round cross-section that are arranged symmetrically around a central axis. Most commonly, the total number of rods is four (quadruple), six (hexapole) or eight (octapole). The length of the rods is generally much larger than the distance between the central axis and the nearest face of each rod. The rods are precisely machined and positioned to create an electric field that is homogeneous along the entire length of the poles. To generate the multipole potential, the rods are grouped into two pairs in a way that every second electrode is electrically connected together. Each pair carries a sum of a direct current (DC) component U and a sinusoidal waveform with an amplitude A oscillating at a radio frequency f to give the electric potential Φ_0 as

$$\Phi_0(t) = U + A\sin(2\pi f t + \phi),$$

where the phase angle ϕ is shifted by 180°(or π rad) between the two pairs. Typical oscillation frequencies are found between a few hundred kilohertz to a few megahertz with amplitudes of up to hundreds of volts peak-to-peak (V_{p-p}). In one half-wave of the RF waveform, the ions are accelerated towards one set of the rods. In the next half-wave, the polarity of the sine wave changes for each rod and the ions are now accelerated towards the other set of the rods. If the DC voltages as well as the oscillation frequency and amplitude are set appropriately, the ions will not collide with the rods and instead stay dynamically focused close to the central axis of the multipole device. However, the stable motion through a multipole device does not only depend on the different voltages that create the electric field, but also depends on the ion's m/z-ratio. This behavior is widely used to separate ions according to their m/z-ratio using the famous quadrupole mass filter.[137,138] By varying the RF voltage, the DC offset of all rods and the DC differential voltage

between the two pairs of rods, the so-called stability window can be precisely controlled. This window defines which ions will have a stable trajectory along the quadrupole mass filter and can be continuously changed: the window can be broader to allow ions of many different $^{m}/_{z}$-ratios to pass through, or it can be effectively narrowed down to a point, where only ions of a single $^{m}/_{z}$-ratio experience a stable motion through the quadrupole mass filter. In the latter case, other ions that do not share the same $^{m}/_{z}$-ratio will have unstable trajectories and collide with the quadrupole rods and will therefore not be detected. Commercial quadrupole instruments often operate at unit mass resolution, i.e. one $^{m}/_{z}$-ratio can be separated from the next integer $^{m}/_{z}$-ratio (100 and 101, or 1000 and 1001).

A different pole arrangement is found in stacked ring-electrode devices, which are commonly used as ion guides because they can efficiently transmit and focus ions from areas of high to low pressures. These devices typically consist of multiple flat ring-electrodes with a circular aperture oriented along the central axis, as shown in Figure 3.4. The RF voltages are applied with opposing phases to odd- and even-numbered electrodes. One advantage over multipole-rod devices is the possibility to control the DC offset voltage of every single electrode either individually or by applying at least two DC potentials that generate a potential gradient along the central axis by a voltage divider network. Such a DC voltage gradient can be used to drive the ions forward along the flight axis.

If the inner diameter of the aperture is constantly decreasing, diffuse ion clouds can be efficiently radially focused so they can be guided through a small orifice that serves as a conductance limit.[139–141] In addition to such an ion funnel design, stacked ring electrodes can also be used as ion traps[142,143] or ion-mobility cells.[144,145]

Ion Transmission into High Vacuum

After transfer from atmospheric pressure through the source region of the mass spectrometer, the ions traverse two differentially pumped ion guides. Using the Z-spray source, the first ion guide is operated at a pressure of around 4 mbar, whereas the capillary source allows a higher vacuum of around 1 mbar. The second ion guide does not have a dedicated pressure gauge, but the corresponding chamber is evacuated using a turbo-molecular pump, so that a pressure of around 10^{-4} mbar is expected. The collisions with neutral background gas molecules thermalize the ions as they pass through this region. Both ion guides consist of 0.5 mm thin ring electrodes with a circular aperture diameter of 5 mm. Each ion guide carries an end-cap serving as a conductance limit with an aperture diameter of 1.5 mm. The corresponding oscillators drive the ion guides at the resonance frequency of the circuit*, generating 790 kHz for the first ion guide and 1.9 MHz for the second ion guide at RF amplitudes of up 600 V_{p-p}. The potential applied to the end-cap of the first ion guide as well as the DC offset value of the second ion guide define the ion's kinetic energy of typically 34 eV.

Quadrupole Mass Filter

In this experimental setup, ions of a specific $^{m}/_{z}$-ratio can be selected using a quadrupole mass filter with a rod diameter of 9.5 mm at a length of 200 mm operating at 880 kHz, which results in a mass range of up to 2000 amu. The quadrupole is driven by an Extrel RF power supply unit (model 150-QC, Extrel, Pittsburgh, PA, USA) and an Extrel DC power supply (model 813370). Short quadrupole segments of 20 mm length are mounted at the entrance and exit of the mass filter and operate at

*The resonance frequency of the system is determined by the combined capacitance and inductance of all circuit components (oscillator, cables, and ion guide electrodes).

the same radio frequency. These segments carry a separate DC bias voltage in order to focus the ions and to create a more homogeneous electric field at both ends of the mass filter to improve ion transmission.

After passing the quadrupole mass filter, the ions are guided into the next chamber via a hexapole collision cell. This collision cell operates at 770 kHz and is intended to aid the structural identification in the commercial instrument by generating fragment ions. In this experimental setup, however, it is exclusively used to guide ions into the next chamber without significant transmission losses.

Quadrupole Bender

After passing the hexapole collision cell and entering the next vacuum chamber, the ions are focused by an Einzel lens before they enter a quadrupole bender. This bender consists of two pairs of diagonally opposing hyperbolic electrodes, as shown schematically in Figure 3.5. If an opposing DC voltage is applied to the two pairs of rods, the ion beam can be deflected by 90°. Because the pathway of the ions depends on their kinetic energy, the magnitude of the DC voltage has to match the kinetic energy of the ions. If the correct voltages are applied (in this case ±34 V), the ions are deflected and can be accumulated in the hexapole ion trap, which will be discussed later. If all electrodes are grounded, the ions traverse the bender without being deflected and pass through another Einzel lens that focuses the ion beam before it enters the time-of-flight mass analyzer.

Time-of-Flight Mass Analysis

The generated ions can be detected using a time-of-flight (TOF) mass analyzer. This device accelerates ions using pulsed high voltages and measures the time it takes the ions to travel from the acceleration region

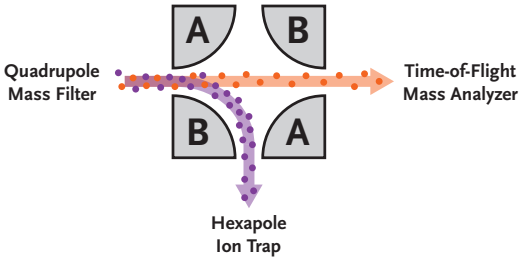


Figure 3.5: A quadrupole bender is used to determine the ion's pathway after leaving the quadrupole mass filter. If no voltage is applied to all segments of the bender, the ions will traverse in a straight path and enter the reflectron time-of-flight mass spectrometer. To deflect the ions by 90° and redirect them into the hexapole ion trap, a static potential of +34 V is applied to electrodes A and −34 V to electrodes B.

to the detector. Because the velocity of ions depends on their mass and charge, ions of different $^{m}/_{z}$-ratios will have different arrival times at the detector and can be distinguished. The flight time t is proportional to the square root of the $^{m}/_{z}$-ratio and can be expressed as

$$t = k\sqrt{\frac{m}{z}},$$

where k is an instrument-dependent constant representing the voltage settings and instrument characteristics.[146] The Waters Q-TOF Ultima is originally equipped with a Micromass reflectron TOF analyzer. The following paragraphs will describe the mode of operation and necessary voltages for positive ion-mode.

After passing the quadrupole bender and the Einzel Lens, the ions traverse another focusing element before they are guided into the entrance of the TOF region by a voltage V1 of around 11 V. Although the ESI process continuously generates ions that enter the mass spectrometer, the TOF operates in a pulsed mode, such that only a small fraction of the ion beam is analyzed per pulse. In the entrance region

of the TOF, high voltages are applied to two electrodes called *pusher* and *puller* using fast push-pull high-voltage switches (model HTS 31-03-GSM, Behlke, Krohnberg, Germany). With a typical rise-time of 50 ns, voltages of +940 V and –640 V are applied to the pusher and puller, respectively. These high-voltage pulses accelerate an approximately 5 cm-long cloud of ions towards the flight tube, where a potential of –10 kV further accelerates the ion package towards the reflectron, as shown in Figure 3.6. The reflectron carries a constant DC potential of +1.78 kV and reflects the ions back into the flight tube. To create a more homogenous electric field, multiple layers of fine wire-mesh grids are stacked and connected electrically via a voltage-divider network. Using a reflectron can drastically improve the mass resolution of the mass spectrometer because it compensates small differences in the ion's kinetic energy.[147,148] More energetic ions travel faster, but penetrate the reflectron deeper and follow a longer path to the detector. As a result, ions of the same $^{m}/_{z}$-ratio, but with different kinetic energies arrive at the detector at the same time.

The value of the entrance voltage V1 will mainly determine the operational mode of the mass analyzer, i.e. which pathway the ions travel inside the TOF-region. If ions are more accelerated along the z-axis (straight towards the detector), they will leave the reflectron and directly hit the detector in a V-shaped path, hence its name *V-mode*. This TOF mass analyzer, however, is typically operated in *W-mode*, as shown in Figure 3.6. Here, the accelerated ions pass the reflectron and enter a second, smaller reflectron called an ion mirror, carrying the same static voltage as the reflectron. This mirror reflects the ions back into the reflectron before traveling to the detector, resembling the letter W. Operating the TOF in W-mode effectively doubles the ion's flight time

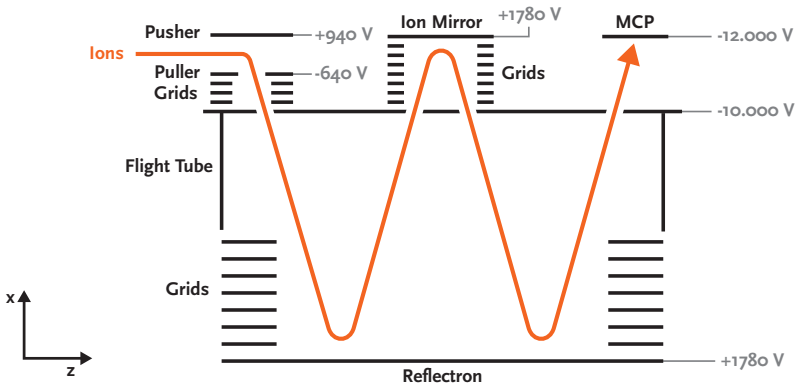


Figure 3.6: A reflectron time-of-flight mass spectrometer is used to analyze ions generated by electrospray ionization. Pulsed high voltages are applied to accelerate ions towards the flight tube. A reflectron and an ion mirror are used to redirect the ions and compensate small differences in their kinetic energy. After traveling a W-shaped pathway, the ions are detected by a micro-channel plate that is connected to a preamplifier and an analog-to-digital converter processing the signal.

and leads to a higher mass-resolution, typically around 10.000 ($^{m}/_{\Delta m}$, with Δm being measured as full-width at half maximum).

The ions are detected using a micro-channel plate that is connected to a current-to-voltage converter (preamplifier model 6954, Phillips Scientific, Mahwah, NJ, USA). The signal is then digitized using an analog-to-digital converter (PXIe 5160, National Instruments, Austin, TX, USA) operating at 10 bit resolution, 500 MHz bandwidth and a sampling rate of up to 2.5 Gs/s.

3.3 The Variable Temperature Hexapole Ion Trap

After adjusting the individual voltages to increase the signal of the $^{m}/_{z}$-selected ions of interest, the quadrupole bender is enabled and deflects the ions by 90°, injecting them into the hexapole ion trap.

The hexapole ion trap consists of six cylindrical electrodes with a diameter of 5 mm and a length of 300 mm, assembled along the central axis with an inscribed circle diameter of 9.1 mm. The sinusoidal waveform is generated by an in-house designed RF-generator operating at a frequency of 1.1 MHz with voltages of up to 250 V_{p-p}. This RF voltage is shifted by an offset DC voltage of 30 V. In order to achieve longitudinal confinement, two electrodes (called frontcap & endcap) are kept 3 V above the hexapole's DC offset voltage, as shown in Figure 3.7. Two additional electrodes after the endcap can be used to accelerate doped droplets.

The incoming ions have a kinetic energy of 34 eV and thus need to loose an excess kinetic energy of 4 eV for efficient longitudinal trapping. To achieve this, the ion trap is encased in a copper-housing that can be filled with helium buffer gas. There is no dedicated pressure gauge mounted inside the housing, but the pressure during the trap filling

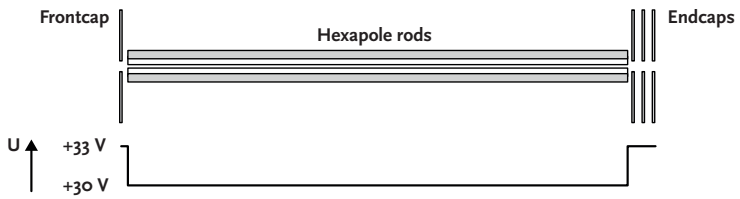


Figure 3.7: It's a trap! This sectional view shows the front- and endcap electrodes that are kept 3 V above the DC offset of the hexapole rods to enable longitudinal trapping. The three endcaps can be individually controlled to accelerate doped droplets.

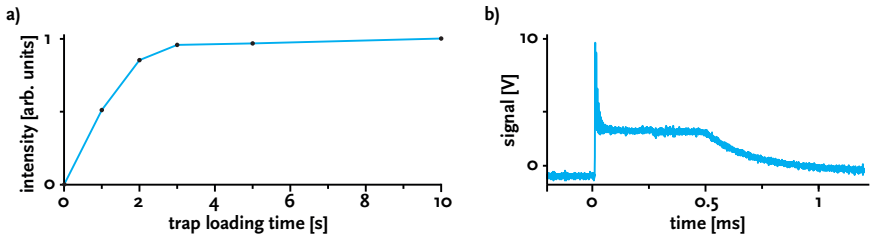


Figure 3.8: a) Typical ion currents generated by electrospray ionization allow to load the hexapole ion trap within a few seconds. b) The number of ions inside the ion trap is limited by Coulomb repulsion and the resulting space-charge limit. Using a Faraday cup detector, a maximum number of around 10^7 trapped ions was measured.

cycle is estimated to be around 10^{-4} mbar. The trap is usually filled for several seconds and the kinetic energy of the incoming ions is reduced due to collisions with the buffer gas during that time. For typical ion currents generated by the ESI process, it only takes a few seconds to fill the trap until the space-charge limit is reached, as shown in Figure 3.8a. Using a Faraday-plate detector placed directly behind the endcaps, the number of trapped ions can be determined by lowering the endcap potential, effectively guiding the ions towards the detector. The obtained signal shown in Figure 3.8b can be converted from a voltage scale to a current scale, such that integrating the signal gives the number of charges that hit the detector. Because the charge-state of the ions inside the trap is known, the number of charges can be converted into the number of ions hitting the detector. Using this detection scheme, the total number of ions inside the trap at the space-charge limit was calculated to be on the order of 10^7 ions.

Once the trap is filled and the buffer gas is pumped away, trapped ions can be picked up by a traversing beam of helium droplets. Due to their velocity and high mass, doped droplets have a high kinetic energy of a few hundred eV and can overcome the trapping potential, effectively

transporting ions towards the detection region. It is important to note that the majority of helium droplets is expected to contain only a single ion. The pickup of multiple ions is theoretically possible, but the low ion density inside the hexapole ion trap as well as Coulomb repulsion between ions render successive pickup events highly unlikely. This was confirmed by previous experiments.[149]

An important feature of this ion trap is the ability to thermalize trapped ions down to 80 K by buffer gas cooling.[150,151] This is achieved by cooling down the trap's copper housing with a flow of nitrogen gas that is externally cooled by liquid nitrogen. In turn, the helium buffer gas is cooled by collisions with the copper surface, which subsequently leads to collisional cooling of the trapped ions. In the original design of the variable temperature ion trap, the cold nitrogen gas flowed through copper tubing that was wound around the trap's copper housing and hard-soldered for good thermal contact, as shown in Figure 3.9a. This design, however, had two severe flaws: first, the copper tubing only was in direct contact with the central part of the housing. This led to temperature gradients of up to 40 K measured from the center of the trap to both ends. Second, the solder-joint became brittle after a few temperature cycles and released internally trapped rosin flux and air, creating virtual leaks in the vacuum chamber.

In the frame of this work, an optimized second-generation variable temperature ion trap was developed. To avoid an inhomogeneous temperature distribution, the cooled nitrogen gas flows through six internal cooling channels that almost span the entire length of the trap housing, as shown in Figure 3.9b. Instead of hard-soldering, the trap body was assembled by tungsten inert-gas welding (TIG), which does not require any flux. To avoid heat transfer from the base of the trap to the copper housing, thin-wall stainless steel mounts with minimal

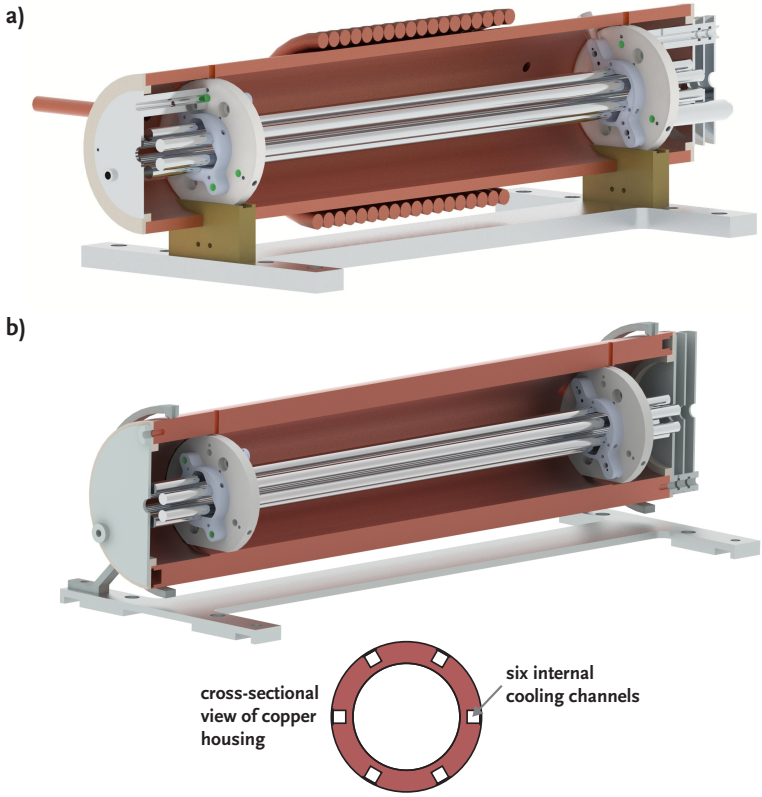


Figure 3.9: Sectional views of the variable temperature ion traps used to thermalize ions prior to pickup by helium droplets. a) The original design consisted of a copper housing that was wrapped by copper tubing. A flow of externally cooled nitrogen gas through the copper tubing was used to thermalize the trap. b) A novel variable temperature ion trap was developed during this work. The copper housing is substantially thicker and accommodates six internal cooling channels almost spanning the entire length of the trap.

contact area are used to hold the trap in place. Due to the large thermal mass of the copper housing, it takes around 90 minutes to cool down the trap below 90 K, as shown in Figure 3.10.

Another issue that is independent of the trap design arises from the fact that the IR radiation is aligned on the central axis of the ion trap. Although the laser beam is focused near the detection region and continually diverges as it approaches the ion trap, the residual photon flux inside the trap can lead to the resonant absorption of photons. In some cases the number of absorbed photons is sufficient to induce fragmentation of trapped ions. To avoid this IRMPD process, a substrate-free wire-grid polarizer (model P02, InfraSpecs, Filderstadt, Germany) was placed between the ion trap and the detection region. After traversing the detection region, this polarizer attenuates the linearly polarized radiation generated by the FHI FEL between 30–2500 cm^{-1} by a factor of 10^4. Because the polarizer consists of a fine wire-grid without any substrate, a sufficient number of doped droplets can pass from the exit of the trap to the detection region.

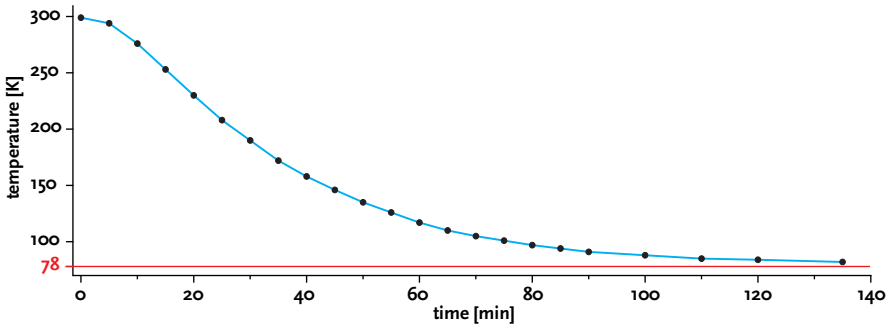


Figure 3.10: The temperature of the hexapole ion trap is measured by a thermocouple placed on the copper housing. Plotting the temperature as a function of time illustrates the cooling process. Due to the large thermal mass of the ion trap and the indirect cooling mechanism, it requires around 90 minutes to cool down the trap below 90 K.

3.4 The Helium Droplet Source

Superfluid helium nanodroplets are generated by expanding pressurized helium into vacuum at low temperatures. There are two types of expansion sources: continuous or pulsed sources. Continuous sources are widely used because they are affordable and create high beam fluxes of helium droplets. However, these sources require large vacuum pumps with high pumping speeds to maintain the required vacuum level in the source chamber. Moreover, when combined with pulsed laser detection, the fraction of droplet beam that can actually be probed is drastically reduced. As an alternative, pulsed sources can be used that provide high peak fluxes of helium droplets with reasonable pumping speeds. The beam characteristics of pulsed valves generally depend on the type of valve, its nozzle diameter and shape, the opening time, temperature and the stagnation pressure.

In this experimental setup, a pulsed, cryogenic Even–Lavie valve (EL-C-C-2013, Uzi Even & Nachum Lavie, Tel Aviv, Israel) is used to generate superfluid helium droplets. The operating principle is illustrated in Figure 3.11. The heart of the Even–Lavie valve is the plunger that has a highly polished curved sealing surface that is pressed against a Kapton gasket by a return spring. This plunger sits inside a high-performance alloy tubing and is held in place by two ruby guides. When a short pulse of high current (up to 15 A) is sent through the solenoid windings, the electromagnetic field actuates the plunger, releasing it from the Kapton gasket and allowing pressurized helium to expand through the trumpet-shaped nozzle. Although the valve is designed to operate at high repetition rates of up to 1000 Hz, it is typically matched to the 10 Hz repetition rate of the FHI FEL.

The valve-body is cooled down to 6 K using a closed-cycle helium cryocooler (RDK 408D2, Sumitomo Heavy Industries Ltd., Tokyo,

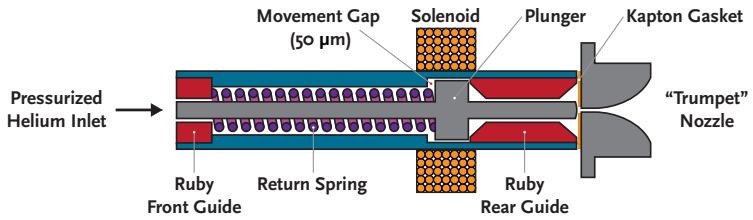


Figure 3.11: Superfluid helium nanodroplets are generated by a cryogenic Even–Lavie valve. The plunger is actuated by a solenoid and releases pressurized helium (70 bar) into vacuum through a trumpet shaped nozzle. The valve body is typically cooled to 15–25 K generating helium droplets with an average size of 10^4 to 10^6 helium atoms.

Japan) with a cooling power of 40 W at 30 K for the first stage and 1 W at 4 K for the second stage. Combining a resistive heating element with a precision diode temperature sensor allows to precisely control the temperature of the Even–Lavie valve, which is typically kept between 15–25 K. Controlling the temperature is typically used to change the size of the helium droplets. However, the expansion of helium into vacuum does not generate helium droplets with a single size, but with a size distribution that is best described by a log-normal distribution. In this experimental setup, the mean value of the size distribution can be shifted from around 10^4 helium atoms at 25 K to around 10^6 helium atoms at 15 K.[149] For experiments presented in this work, the valve was operated at a temperature of 21 K, a stagnation pressure of 70 bar and a typical opening time of around 10 µs, generating helium droplets with an average size of 10^5 helium atoms and a mean velocity of around $450\,\mathrm{m\,s^{-1}}$.

The vacuum chamber is pumped by a turbomolecular pump with a pumping speed of $2400\,\mathrm{l\,s^{-1}}$ (Turbo-V 3K-T, Agilent Technologies Italia, Italy) maintaining a pressure of $\leq 10^{-5}$ mbar during operation. The central part of the molecular beam is transmitted towards the hexapole ion trap using a skimmer (Model 50.8, Beam Dynamics, Jacksonville,

FL, USA) that has an aperture diameter of 4 mm and is placed around 15 cm away from the nozzle. The valve assembly is mounted to the vacuum chamber via a stainless steel bellow allowing the valve to be aligned using three threaded steel tensioners.

3.5 The Fritz-Haber-Institute Infrared Free-electron Laser

To date, a wide range of the electromagnetic spectrum can be covered by various commercially available benchtop laser systems. However, a single laser system often only provides a narrow spectral range with low-intensity radiation. If, on the other hand, high intensity radiation over a broad spectral range is required, a free-electron laser (FEL) is usually the only available option. The development of FELs was pioneered by John Madey and coworkers in the 1970s.[152] Unlike conventional laser systems that stimulate molecular or atomic transitions in solid, gaseous or liquid media to generate laser radiation, the lasing medium of FELs are relativistic electrons moving through periodic magnetic fields in high vacuum.

The radiation used to perform IR spectroscopy of glycans in helium droplets in this work was generated by the Fritz Haber Institute Free-electron Laser (FHI FEL).[119] Regular user operation of the FHI FEL began in 2013 after first lasing was achieved in 2012. A schematic overview of the laser and its main components is shown in Figure 3.12. The electron beam used to generate the laser beam is produced by an electron gun emitting pulses of free electrons into vacuum. The time structure of the pulsed electron gun is characterized by a macro- and a micro-repetition rate that also determines the time structure of the laser radiation. The macro-pulse repetition rate is usually set

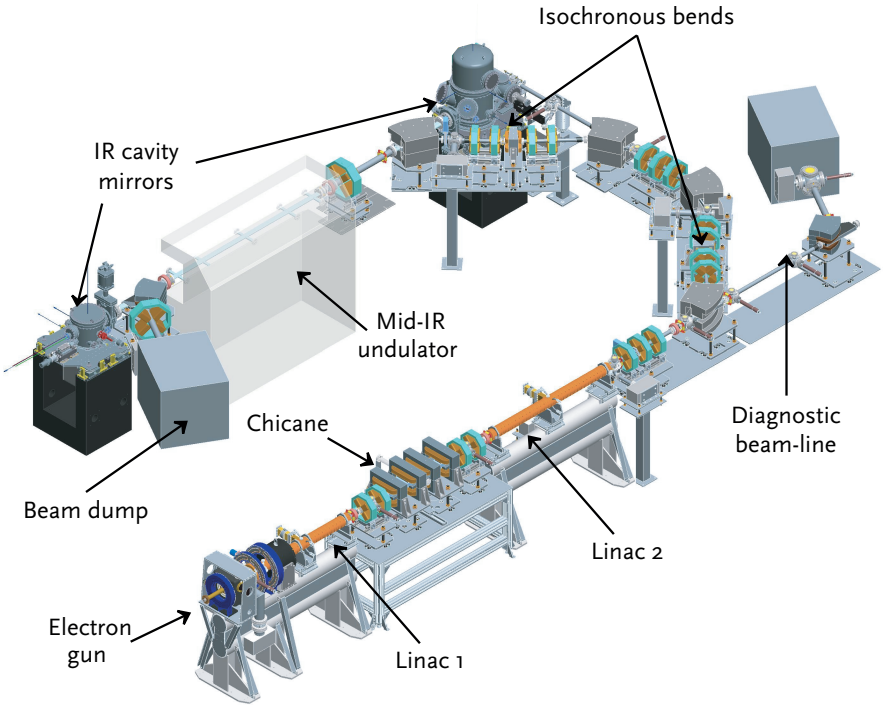


Figure 3.12: Schematic overview of the Fritz Haber Institute Free-electron Laser (FHI FEL).[119] An electron gun releases bunches of electrons into vacuum that are accelerated to relativistic velocities by two linear accelerators (LINACs). Inside the laser cavity, a periodic array of strong magnets forces the relativistic electrons on an oscillatory wiggling motion emitting monochromatic radiation. Constructive interference between the electromagnetic field of the radiation and the electrons amplifies the emission of coherent photons.

to 10 Hz, and each macro-pulse consists of 10^4 micro-pulses generated at a repetition rate of 1 GHz. After traversing a buncher cavity that compresses the bunch-length, the electrons are accelerated by two linear accelerators (LINACs). The first LINAC accelerates the electrons to a fixed energy of around 20 MeV and the second LINAC is used to vary the final electron energy between 18 and 45 MeV. Next, the electron beam is directed through a U-shaped bend by dipole and quadrupole magnets before entering the laser cavity. Inside the cavity, the 2 m long undulator contains a periodic array of oppositely poled magnets creating strong magnetic fields that force the electrons on a sinusoidal 'wiggling' motion. This deflection of electrons leads to the emission of radiation. The first electron bunch passing through the undulator emits an initial incoherent IR pulse. This optical pulse is reflected by the FEL cavity mirrors and passes through the undulator again. The length of the laser cavity (5.4 m) is set such, that the cavity round-trip time of the optical pulse is synchronized with the electron bunches entering the cavity. As a consequence, both pulses travel along the magnetic fields of the undulator, and the electromagnetic field of the radiation can interact with each electron bunch. This interaction leads to micro-bunching of the electrons and the emission of coherent radiation with high intensity gains. The laser output characteristics can be precisely controlled by 'detuning' (shortening) the optical cavity by a multiple of the laser wavelength. A small detuning, such as 1 λ, generates short IR micro-pulses with high peak power. Because the spectrum of the radiation is Fourier-limited, these short pulses have a broader spectral width. In this work, the FEL was typically operated at a detuning between 3–5 λ, which increases the micro-pulse length to a few picoseconds and decreases the spectral width to around 0.3–0.5% (FWHM).

A through hole in one of the cavity mirrors is used to transmit a fraction of the radiation to the user experiments and diagnostic elements. During a beam shift (typically one day), the electron energy is kept constant and the laser wavelength can be tuned by changing the gap between the undulator magnets. The resulting laser radiation has a macro-pulse length of 10 µs and a macro-pulse energy of up to 120 mJ, maintaining a bandwidth of around 0.5% (FWHM). Tuning the electron energy allows the FHI FEL to produce photons between 3 µm and 60 µm.

Because the characteristics of pulsed laser sources are often inconceivable by just considering a single micro- or macro-pulse of radiation, a real-world example is given here:
During this work, a total of 100 beam shifts was used to collect data and characterize the experimental setup. On average, the FHI FEL was operated 13 hours per beam shift at a wavelength of around 7 µm and macro-pulse energy of 70 mJ. This equates to a total of 10^{26} photons* with a total energy of 3.3 MJ. However, the length of a single micro-pulse is only around 3 ps. Therefore, the actual time to generate the total number of photons produced during 100 beam shifts was only around one second.

*A back-of-the-envelope calculation shows that this is approximately the same number of photons that hit the premises of the Fritz Haber Institute during a sunny summer day . . . every second.

3.6 Ion Detection Following Vibrational Excitation

When doped helium droplets and the IR laser beam overlap inside the detection region of the instrument, the absorption of multiple resonant photons can lead to the release of bare ions from helium droplets. The yield of released ions is used as a marker for photon absorption and is monitored by subsequent time-of-flight detection. At this point, the astute reader might wonder about the necessity to detect the number of released ions by means of TOF mass analysis. After all, the ions injected into the hexapole ion trap are already mass-selected by the quadrupole mass filter. However, there are important reasons that render the detection of the ion's m/z-ratio upon release from the droplet indispensable. For example, if the m/z-ratios of two different species are very close together ($\Delta\frac{m}{z} < 1$), they cannot be fully separated by the quadrupole mass filter. In this case, both species will be analyzed and detected after release from the helium droplets. However, high-resolution TOF-analysis allows to distinguish the released ions and monitor them independently to generate the respective IR spectrum. Another example that requires mass analysis of released ions is the partial fragmentation of meta-stable ions inside the hexapole ion trap prior to pickup by helium droplets. In this scenario, both parent and fragment ions are picked up by helium droplets and will be collectively analyzed. By monitoring the m/z-ratio of the released ions, the individual IR spectra can be assigned accordingly. Another fragmentation pathway of parent ions can occur during photon absorption. In this case, the photon flux during the interaction with the doped droplets is too high and likely induces an IRMPD process of ions shortly after being released from the droplet. It is beneficial to detect such fragmentation processes because

they often indicate saturation effects that lead to line broadening and shifts in relative peak intensities. To counteract these saturation effects, the photon flux can be continuously varied by shifting the focal point of the laser beam using a pneumatically controlled adaptive mirror (model AO90/70, Kugler GmbH, Germany). Lastly, experiments that involve H/D exchange require mass analysis of released ions. Although deuterated species can often be isolated by the quadrupole mass filter, the ions can undergo partial back-exchange inside the trap during the long accumulation time. Using high-resolution TOF-analysis, the correct IR spectrum can be reliably assigned by monitoring the corresponding $^{m}/_{z}$-channel.

3.6.1 On-Axis Wiley-McLaren Time-of-flight Detection

The original design of the detection region is shown in Figure 3.13 and consists of a home-built Wiley–McLaren[153] time-of-flight mass spectrometer. First, doped droplets enter the interaction region that consists of two electrodes called *repeller* and *extractor* and another electrode connected to ground. Immediately after the pulsed IR beam traverses the interaction region, high voltages are applied to the repeller and extractor to accelerate the released ions towards the detector. As the velocity of the accelerated ions depends on their mass and charge, the arrival time at the detector can be used to distinguish ions according to their $^{m}/_{z}$-ratio. Ions are detected using an electron multiplier (model 14880, ETP ion detect, Clyde, Australia) connected to a current-to-voltage amplifier (model VT120A, Ortec, Oak Ridge, TN, USA). The signal is processed using the second channel of the analog-to-digital converter described in Section 3.2.

This detection scheme is a rather simple design that allows the detection of released ions with high sensitivity. However, the short

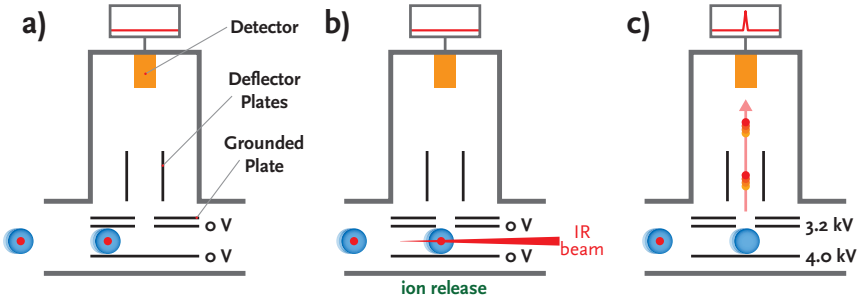


Figure 3.13: A Wiley–McLaren time-of-flight mass spectrometer was originally used to detect released ions from helium droplets as a marker for photon absorption. It consists of three electrodes that are used to accelerate released ions towards the detector. a) The doped helium droplets approach two electrodes called *repeller* and *extractor* that are kept at 0 V. After the IR laser beam traversed the interaction region (b), high voltages are applied to the extractor and repeller (c) that accelerate released ions towards the detector.

and linear flight path of around 20 cm length results in a low mass resolution of around 150 ($\frac{m}{\Delta m}$). While such a resolution is often high enough to distinguish between parent ions and their fragments, it fails to distinguish smaller differences in $^{m}/_{z}$-values such as in H/D exchange experiments. Although the mass resolution could to some extent be increased by increasing the length of the flight path, the sensitivity would suffer.

3.6.2 Off-Axis Reflectron Time-of-flight Detection

In the course of this work, a novel detection scheme was developed and constructed to improve the mass resolution without sacrificing sensitivity. The starting point of this reconstruction was a spare commercial Micromass reflectron mass spectrometer, identical to the one described in Section 3.2. However, an on-axis implementation analogous to the Wiley–McLaren design described above is not feasible for this mass spectrometer because the incoming ions are required to have a kinetic

energy of around 34 eV. Therefore, release of ions and their subsequent detection are spatially separated in the new design shown in Figure 3.14. After a brief overview, a few important aspects of the new detection scheme are described in more detail.

After exiting the hexapole ion trap, the beam of doped droplets is focused by an Einzel lens before entering a stacked ring-electrode ion guide. Inside the ion guide the doped droplets are irradiated by the IR laser beam. The resonant absorption of multiple photons can lead to the release of bare ions from the droplets that are axially confined by an RF sinusoidal waveform. Once the IR light pulse passed the ion guide, the released ions are gently accelerated towards a set of acceleration electrodes to define the ion's kinetic energy to around 34 eV. Before and after deflection by 90° using a quadrupole bender, the ions are focused by Einzel lenses. Four deflector plates are used to guide the released ions towards the high-resolution TOF mass spectrometer that allows a mass resolution of up to 10.000 ($\frac{m}{\Delta m}$).

An important detail of the new detection scheme is the custom-built ion guide that serves as the interaction region. It comprises 54 stacked ring-electrodes with an internal diameter of 5 mm, a thickness of 0.5 mm and a spacing of 2 mm. The electrodes are sandwiched between two custom-built printed circuit boards (PCBs) that each contain two RF-bias tees that mix the RF sinusoidal waveform with DC voltages, and a resistor network used as a voltage divider. The sinusoidal waveform is generated by an RF-oscillator operating at the self-resonance frequency of the system (1 MHz). Two DC voltages are used to define the offset of the sinusoidal RF waveform at the first and last electrode of the ion guide and are used to define the kinetic energy of the ions and guide them towards the exit of the ion guide. However, a constant DC gradient across the ion guide cannot be used to transfer the ions.

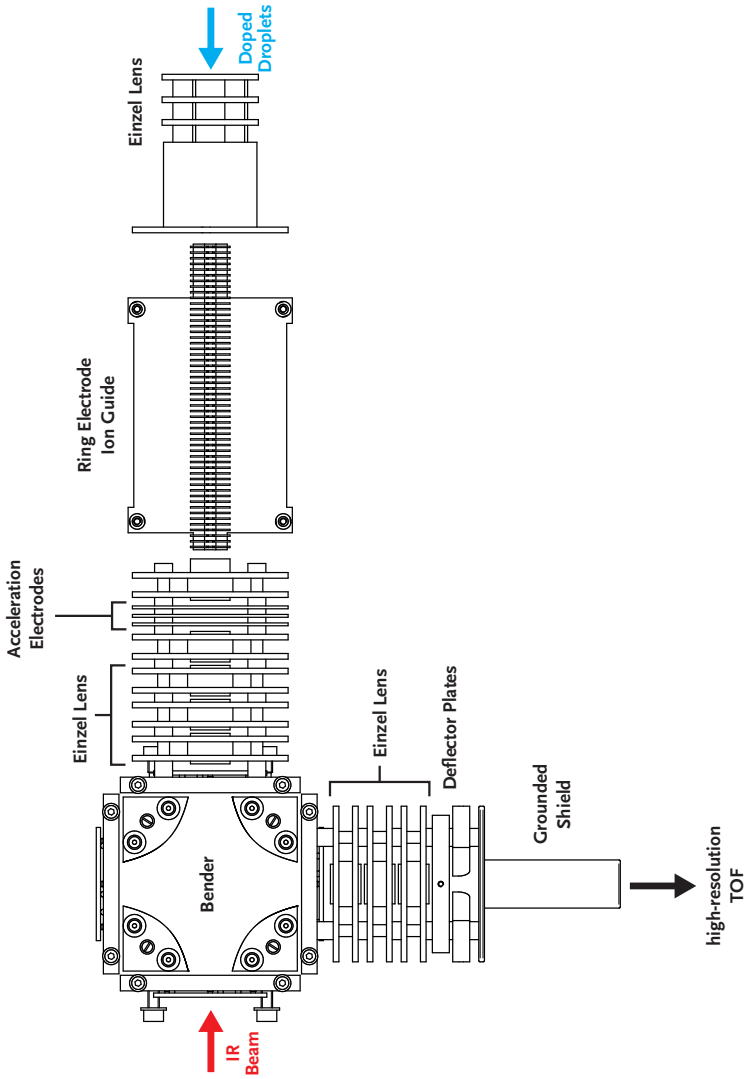


Figure 3.14: A novel detection scheme was developed during this work. It consists of an Einzel lens that focuses doped droplets into a ring-electrode ion guide where they are irradiated by IR laser beam. Immediately after the laser beam traversed the ion guide, a DC gradient guides the released ions towards a stack of acceleration electrodes that define the kinetic energy of the ions. A quadrupole bender is used to deflect the released ions by 90° and direct them towards a reflectron time-of-flight mass analyzer.

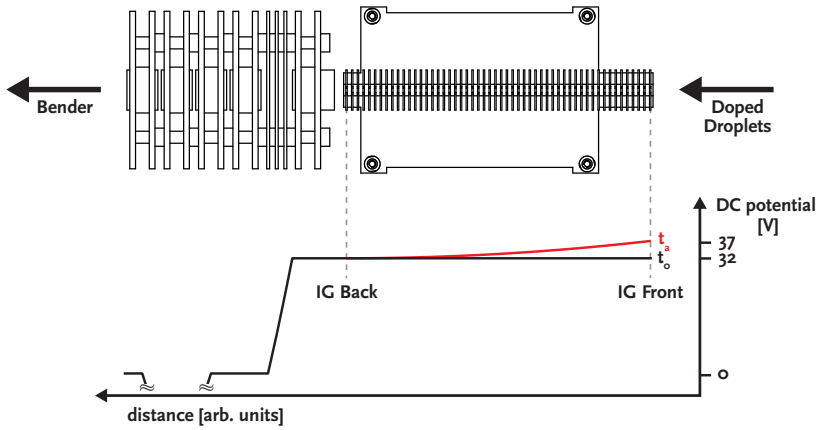


Figure 3.15: The DC potentials of the ion guide, the acceleration lenses as well as the first Einzel lens are schematically shown. While the beam of doped helium droplets enters the ion guide, all electrodes guide carry the same DC potential. Once the IR laser beam passed the ion guide, the potential at the first electrode (IG Front) is raised and creates a hyperbolic potential across the ion guide by a voltage divider network. This hyperbolic potential focuses the spatially spread cloud of released ions, which are then accelerated by three ring electrodes and focused by an Einzel lens before they enter the quadrupole bender.

Due to the initial spread of the doped droplets (around 5 cm) and the time structure of the laser pulse, the resulting accelerated ion package traveling towards the entrance region of the TOF mass spectrometer would be more than one meter long. This would result in a substantial loss of sensitivity because the pulsed operation of the TOF analyzer only allows a small portion of around 5 cm length to be analyzed at once. Therefore, the initial spread of released ions has to be compensated by compressing the ions along the flight-axis to avoid sensitivity losses.

Compression of the ion cloud is achieved by switching the DC offset voltages of the ion guide immediately after the laser pulse has traversed the ion guide (Fig. 3.15). Initially (t_0), the two DC offset voltages at the front and back of the ion guide (IG Front and IG Back, respectively)

are kept at the same potential of around 32 V. Once the IR laser pulse passed the ion guide (t_a), the DC offset voltage at the front of the ion guide is switched to 37 V. Because the individual electrodes are connected via a resistor network, the DC offset voltage will gradually decrease from IG Front to IG Back. In this case, the resistor network is chosen such that the resulting electric field describes a parabolic function along the central axis of the ion guide. This parabolic electric field pushes the released ions towards the bender and compresses the ion cloud to a point around two centimeters behind the last electrode. Efficient spatial compression, however, can only be achieved if the DC offset voltage that generates the parabolic electric field can be raised reasonably fast. Specifically, a rise-time of not more than 1 µs was targeted during development. At this point it is important to mention that fast switching of DC voltages feeding a resistor network over an RF bias tee, which tries to suppress transient DC voltages is rather unusual and cannot be accomplished by out of the box solutions. For more details regarding PCB and circuit design, a custom-built fast DC switch and the resulting switching performance of the ion guide, the interested reader is referred to Appendix A.

After the released ions are spatially compressed, they are accelerated towards a ground potential by three successive electrodes that create a homogeneous electric field gradient. After passing the Einzel lenses, the quadrupole bender and the deflector plates, ions are detected by the high-resolution Micromass TOF analyzer, which is typically operated in V-mode for maximum sensitivity. Typical TOF signals are shown in an exemplary way for the protonated pentapeptide Leu-enkephalin (LEK) in Figure 3.16a. When the IR laser wavelength is not resonant with any vibrational mode of the probed ion, no signal is detected. If, on the other hand, the IR laser wavelength matches a strong vibrational transition

of the ion, a strong signal with a high signal-to-noise ratio is detected, yielding a high dynamic range. An infrared spectrum is obtained by plotting the yield of released ions as a function of laser wavelength (Figure 3.16b). In comparison to previously published data,[123] the IR spectrum of protonated LEK exhibits partially more narrow absorption bands, which is likely attributed to an overall increase in performance of the experimental setup, the FHI FEL, and the implementation of the wire-grid polarizer that prevents activation of ions inside the hexapole ion trap.

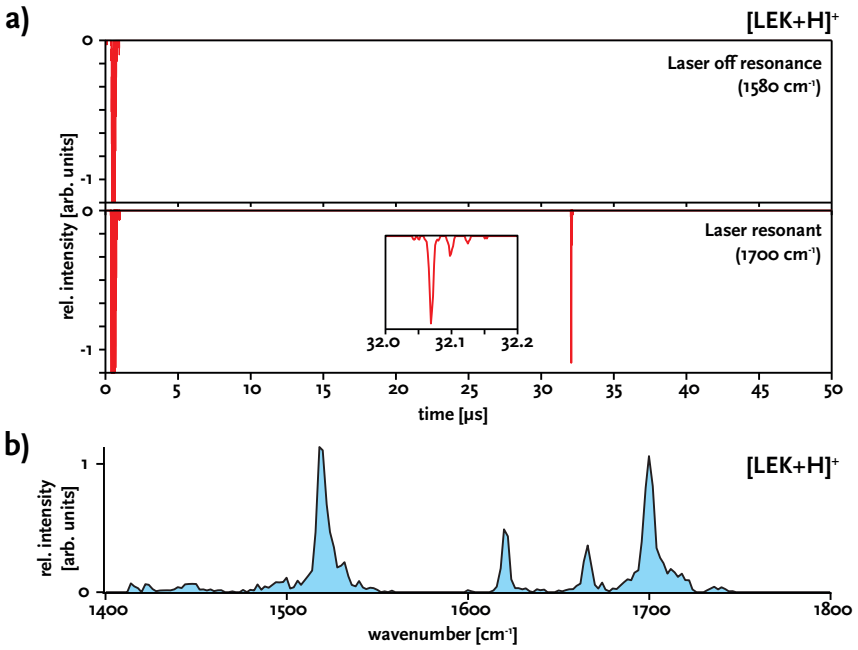


Figure 3.16: Photon absorption is monitored by detecting the number of released ions using a reflectron time-of-flight mass spectrometer, which is exemplarily shown for the protonated pentapeptide Leu-enkephalin (LEK). a) When the IR radiation is not in resonance with a molecular vibration, no ions are released from the droplets and no ion signal can be measured. If, however, the IR radiation is in resonance with a strong vibrational transition, a large signal of released ions is detected. b) Plotting the yield of released ions as a function of laser wavelength yields an IR spectrum.

Chapter 4

Fingerprinting Glycans

Cryogenic IR spectroscopy of ions was successfully used to elucidate the structure of biomolecules such as amino acids, peptides or proteins. But is this experimental approach useful to analyze complex glycans? To answer this question, sets of isomeric oligosaccharides were used to systematically assess the potential of vibrational spectroscopy using helium nanodroplets. The obtained results are presented in this chapter and compared to recently published data. Finally, a brief perspective on the future of glycan analysis is discussed.

This chapter is based on reference [154] – Mucha, E., et al., Glycan Fingerprinting via Cold-Ion Infrared Spectroscopy, *Angew. Chem. Int. Ed.* **2017**, *56*, 11248, and reference [1] – Mucha, E., et al., In-depth Structural Analysis of Glycans in the Gas Phase, *Chem. Sci.* **2019**, *10*, 1272. Figures and content adapted with permission from John Wiley and Sons, and The Royal Society of Chemistry.

4.1 Introduction

Time and time again, research unravels novel biological functions of glycans confirming their tremendous relevance in virtually every biological process. The diversity of biological functions is enabled by the enormous structural complexity and diversity of glycans. This, however, has direct consequences on the glycoanalytical workflow, especially when compared to genomics and proteomics, which both benefit from generic high-throughput sequencing methods. To date, a major hurdle in the glycosciences remains the lack of a universal sequencing method. Instead, a multitude of sophisticated methods are used for glycan analysis including nuclear magnetic resonance (NMR) spectroscopy, liquid chromatography (LC) or sequential mass spectrometry (MS^n).

NMR spectroscopy, for example, is the analytical gold standard in glycosynthesis. The identification of synthetic glycans greatly benefits from large sample amounts that enable one- and multidimensional NMR analyses, which are often supported by high resolution mass spectrometry. Each chemical reaction is planned beforehand and the analytical data can be used to confirm the outcome of the reaction. The requirements for the identification of natural samples, on the other hand, are extremely different. The sample amount of glycans in biological samples is often limited and they typically exist in complex mixtures. NMR spectroscopy is therefore generally not feasible for natural product identification, because it requires milligrams of relatively pure sample. Furthermore, and in contrast to chemical synthesis, the chemical structure of glycans from biological samples is not known *a priori*.

Naturally occurring oligosaccharides are often attached to proteins as posttranslational modifications. To fully explore the biological roles of such glycoproteins it is crucial to investigate both, the sequence of

the glycan as well as the amino acid residue it is attached to. However, most glycoanalytical workflows follow the bottom-up approach, which cleave the glycoproteins into smaller units and eventually release the glycans using enzymatic or chemical cleavage. The released glycans can then be analyzed or separated by various chromatographic techniques including hydrophilic interaction liquid chromatography (HILIC), anion-exchange chromatography (AEC), porous graphitic carbon chromatography (PGC) or capillary electrophoresis (CE). Analyte detection after chromatographic separation is typically achieved by ultraviolet (UV) absorption or fluorescence detection schemes. Glycans, however, are not UV active and therefore require sample derivatization prior to analysis by attaching UV chromophores or fluorescent labels.

Another important member of the glycoanalytical toolbox that does not require chemical derivatization is mass spectrometry (MS). The low sample requirements and the amount of information that can be obtained within a short time render MS as a widely used tool in glycan analysis. Sequential mass spectrometry (MS^n) is typically employed to maximize the amount of information extracted from a sample by generating fragment ions. Depending on the polarity, different fragmentation pathways are accessed. Fragmentation of positive glycan ions typically leads to the cleavage of glycosidic bonds, which can be used to gain information about the glycan sequence and composition. Fragmentation of negatively charged glycan ions leads to cross-ring fragments and can be used to gain information about connectivity and branching of the glycan. A major hurdle, however, remains the unambiguous identification of isomeric glycans. Monosaccharide building blocks often share the same mass and therefore cannot be distinguished by measuring their m/z-ratio. In addition, information about the configuration of the glycosidic bond cannot be easily obtained from MS techniques alone.

To overcome these limitations, mass spectrometry is often combined with LC methods as an orthogonal separation technique. This powerful combination is widely used for the identification of glycans and glycoproteins from biological samples because it offers high reproducibility and a great depth of structural information. Complete structural elucidation, however, requires time consuming data analysis and often involves enzymatic digestion of the glycan, which increases the analysis time and sample consumption.

A more recent addition to the structural analysis toolbox for glycans is ion mobility-mass spectrometry (IM-MS). A typical experimental setup comprises an ion mobility cell operated under a constant pressure of 1–15 mbar using a neutral buffer gas such as helium or nitrogen. The ions are guided through the cell under the influence of a weak electric field, and undergo collisions with the buffer gas atoms or molecules. These collision determine the ions' velocity, which is a function of their mass, size, shape and charge. For example, larger structures collide more frequently with the buffer gas than smaller structures and therefore experience longer drift times. The instrument-dependent drift time is commonly used to derive the collision cross section (CCS), which is an instrument-independent parameter resembling the ions' rotationally averaged area inside the gas phase. Several studies showcase the resolving power of IM-MS that allows to distinguish connectivity and configurational isomers of glycans. Similarly to MS^n experiments, the incremental analysis of characteristic fragments can aid the structural identification. In general, the quality of separation in ion mobility spectrometry can vary significantly when different adduct ions are investigated. For example, the coordination of different alkali metals as adduct ions affects the molecular conformation and therefore the CCS, which determines whether or not two isomers can be separated.

As a result, it is often necessary to optimize the separation power by investigating unknown structures in both ion polarities and with a variety of adduct ions. Currently, various types of ion mobility-mass spectrometers are commercially available and technical developments continue to enhance the resolving power. Because IM-MS experiments operate on a millisecond time-scale, they can also be readily coupled to LC resulting in a three-dimensional analysis of intact ions and their fragments based on LC retention time, CCS, and $^{m}/_{z}$-ratio.

Another orthogonal method that is often combined with mass spectrometry is gas-phase IR spectroscopy, as discussed in sections 2.4 and 2.5. Vibrational spectroscopy is highly sensitive towards structural details and enables the investigation of functional groups, intra- or intermolecular interactions, or the complete molecular conformation. As a consequence, gas-phase IR spectroscopy is routinely used to probe the structure of biomolecules such as amino acids, peptides, proteins, nucleic acids or metabolites. In the early 2000s, the group of John P. Simons pioneered the use of gas-phase IR spectroscopy for carbohydrate analysis using UV-IR double resonance spectroscopy. In the low temperature environment of a free-jet expansion, smaller neutral carbohydrates yielded conformer-specific IR spectra that allowed a full conformational assignment highlighting structural preferences and the importance of hydrogen-bonding networks. However, the neutral carbohydrates were mostly transferred to the gas phase by evaporation from an oven or laser vaporization, which limits this method to smaller carbohydrates that are not prone to thermal decomposition. In addition, this method necessitates a covalently attached UV-chromophore, which is usually not found in naturally occurring glycans and therefore requires chemical derivatization.

A different spectroscopical approach that does not require the chemical derivatization of samples is IR multiple photon dissociation (IRMPD). Here, the sequential absorption of multiple photons induces fragmentation of molecular ions, which is monitored as a function of the laser wavelength. Details of this method are presented in Section 2.4. Within the last decade, numerous studies used IRMPD spectroscopy to distinguish isomeric glycans or identify their structure. Recently, Compagnon et al. used the unique vibrational fingerprints obtained by IRMPD to identify sialic acid linkage isomers of human milk oligosaccharides and N-glycans. In another work, they demonstrate that carbohydrates can retain the stereochemical information of the glycosidic bond upon fragmentation, which is crucial for glycan sequencing. This anomeric memory was used to fully characterize the structure of different chito-oligosaccharides. In another recent publication, Tan et al. measured unique IR fingerprints of lithiated *N*-acetyl-D-hexosamines, and used quantum chemical calculations to identify individual molecular conformations. Martens et al. used IRMPD to identify metabolites from human body fluid samples and obtained a unique IR signature that was assigned to *N*-acetylmannosamine, a biomarker for NANS-deficiency. [35]

However, a common limitation of IRMPD spectroscopy arises from peak broadening and spectral congestion, which is partly caused by the thermal activation of ions during the sequential absorption of multiple photons. Also the conformational flexibility of larger glycans may lead to the thermal population of coexisting conformers that have different absorption spectra. The resulting spectral congestion typically limits the scope of IRMPD spectroscopy to mono- and disaccharides.

As mentioned in Section 2.5, the limitations encountered in IRMPD spectroscopy can be overcome by probing ions in the cryogenic environment of superfluid helium droplets. The experimental setup described

in chapter 3 was used to elucidate the structure of different molecular classes including carbonic acid derivatives,[120–122] metal oxide nanoclusters,[107] as well as biomolecules such as peptides[123] and proteins.[124] But can this cryogenic spectroscopy scheme provide highly resolved IR spectra of larger isomeric glycans and resolve their structural details? To answer this question, this chapter presents a systematic evaluation of the helium droplet spectroscopy method using various isomeric glycans.

4.2 Experimental Details

Samples

In this work, synthetic as well as naturally occurring glycans were investigated. The synthetic standards consist of one or more monosaccharide building blocks and an aminopentyl linker attached to the reducing end. Synthesis was performed by the group of Peter H. Seeberger (Max Planck Institute of Colloids and Interfaces, Potsdam, Germany) using automated solid-phase glycan assembly. A detailed description of the synthesizer is given in the literature.[155–157] In addition to the synthetic standards, lactose was purchased from Merck Millipore (Darmstadt, Germany), the tetrasaccharides Lewis b and Lewis y were purchased from Dextra Laboratories (Reading, UK), and the milk sugars lacto-*N*-tetraose and lacto-*N*-neotetraose were purchased from Carbosynth Limited (Berkshire, UK). All samples were used without further purification. HPLC-grade solvents were purchased from Sigma-Aldrich (now Merck MilliporeSigma). For nano-electrospray ionization, aqueous 1 mM stock solutions were diluted with water/methanol (v/v, 1/1) to yield a 50 µM analyte solution.

Experimental Setup

Vibrational spectra of glycans were recorded using the experimental setup described in Chapter 3 using the on-axis detection scheme described in Section 3.6.1. As a brief summary, glycan ions are brought into the gas phase by nano-electrospray ionization. These ions are mass-to-charge selected using a quadrupole mass filter and accumulated inside a hexapole ion trap operated at room temperature. Superfluid helium nanodroplets with an average size of 10^5 helium atoms traverse the trap and can pick up ions and rapidly cool them to the equilibrium

temperature of the droplet of 0.4 K. Inside the detection chamber, the embedded ions are investigated using laser radiation produced by the FHI FEL. The subsequent absorption of resonant photons can lead to the release of the bare ion from the droplet, which is used as a marker for photon absorption. Plotting the yield of released ions as a function of laser wavelength yields a highly reproducible IR spectrum. Each spectrum shown here consists of at least two independent scans recorded with a wavenumber step-size of $\Delta\tilde{\nu} = 2\,\mathrm{cm}^{-1}$. Although the yield of released ions scales non-linearly with laser energy, a linear correction has been performed by dividing the signal intensity by the laser energy. As a result, relative intensities can be distorted to some extent. The spectra shown here were recorded between $950\,\mathrm{cm}^{-1}$ and $1700\,\mathrm{cm}^{-1}$ with two different photon fluxes using a variable focusing mirror. A softer focus (lower photon flux) was used between $950\,\mathrm{cm}^{-1}$ and $1200\,\mathrm{cm}^{-1}$. To access the absorption bands between $1150\,\mathrm{cm}^{-1}$ and $1700\,\mathrm{cm}^{-1}$, a separate spectrum with a tighter focus (higher photon flux) was recorded.

4.3 Results and Discussion

To assess the potential of cryogenic ion spectroscopy to resolve the minute structural details present in complex carbohydrates, a series of well-defined amino-alkyl linked carbohydrates as well as natural samples were analyzed. The analyzed samples consists of monosaccharides **1**–**5**, disaccharides **6**–**7**, trisaccharides **8**–**13**, and naturally occurring tetrasaccharides **14**–**17**, as shown in Figure 4.1.

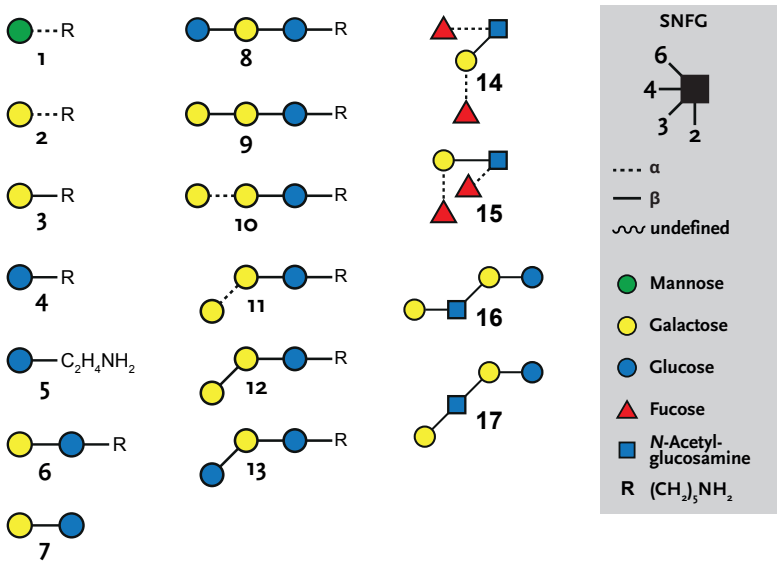


Figure 4.1: Schematic representation of the investigated molecules in this chapter using the SNFG. The sets consist of monosaccharides **1–5**, disaccharides **6**&**7**, trisaccharides **8–13**, and biologically relevant tetrasaccharides **14–17**.

Mono- and Disaccharides

Monosaccharides are the simplest form of sugars and represent the fundamental building blocks of oligosaccharides. Different monosaccharides often share the same mass and only differ in the stereochemistry at single carbon atoms. Using the cryogenic ion spectroscopy technique described in Chapter 3, IR spectra of the protonated aminoalkyl-linked monosaccharides were recorded (Figure 4.2). Because the primary amine of the linker has the largest basicity, it will most likely accept a proton upon ionization and provide a localized charge.

The IR spectrum of α-mannose (**1**) shows a variety of absorption bands. Around 1150 cm^{-1}, two narrow and well-resolved absorption bands are found. Preliminary quantum chemical calculations indicate that this region (1000–1150 cm^{-1}) is mainly governed by C−O-stretching vibrations with strong transition dipole moments. Between 1150 cm^{-1} and 1400 cm^{-1}, additional bands that likely originate from O−H-bend modes are present. Above 1400 cm^{-1}, a series of partially resolved bands are obtained. According to calculations, these bands correspond to the primary ammonium group of the linker. Interestingly, only three NH_3^+–bending modes are expected in this region. The presence of multiple, partially unresolved bands indicates that α-mannose adopts multiple coexisting conformers with different IR fingerprints in the gas phase. Moving to the isomeric α-galactose (**2**), the stereoconfiguration at C2 and C4 is inverted. The IR spectrum also shows a variety of resolved absorption bands, but the peak positions and intensities render it distinctly different to the spectrum obtained for α-mannose. Again, the presence of multiple absorption bands above 1400 cm^{-1} indicates multiple coexisting conformers. The IR spectrum of the epimeric β-galactose (**3**) is characterized by a number of highly resolved absorption bands between 950 cm^{-1} and 1400 cm^{-1}, and broader absorptions above

1400 cm^{-1}. Interestingly, some bands are as narrow as the bandwidth of the IR laser radiation (FWHM around 4 cm^{-1}). Also the IR spectrum of β-glucose (**4**) exhibits a number of highly resolved absorption bands up to 1150 cm^{-1}. Between 1150 cm^{-1} and 1500 cm^{-1} only a few weaker absorptions are present. To examine the influence of the attached aminopentyl linker, β-glucose with a shorter aminoethyl linker (**5**) was investigated. A direct comparison shows that both spectra are distinctly different. The spectrum of **5** exhibits less absorption bands than the one obtained for **4** and the peak positions do not coincide. In contrast to the vibrational spectra of monosaccharides **1**–**4**, only one strong and three weak resolved bands appear above 1400 cm^{-1}. This could indicate that the gas-phase structure of **5** mainly adopts a single conformer, which could be caused by the reduced flexibility of the shorter linker.

Upon moving to disaccharides, which consist of two monosaccharide building blocks joined by a glycosidic bond, the structural complexity increases significantly. The glycosidic bond can be formed at different hydroxy groups (connectivity isomers) with either an α- or β-configuration. The increasing system size and the torsional angle of the glycosidic bond furthermore increase the conformational complexity. Although an additional building block is present, the IR spectrum of β-lactose (**6**) shows a similar appearance to the monosaccharides described above. The variety of highly resolved bands provide a unique diagnostic pattern. To further investigate the influence of the aminoalkyl linker, a spectrum of protonated free lactose (**7**) was recorded. Here, the reducing end does not contain any linker and the anomeric center can adopt either an α- or β-configuration. In addition, the location of the charge is uncertain and different protomers may coexist. Nonetheless, the IR spectrum of **7** exhibits multiple well-resolved absorption bands between 950 cm^{-1} and 1400 cm^{-1}. There are no obvious similarities between the vibrational

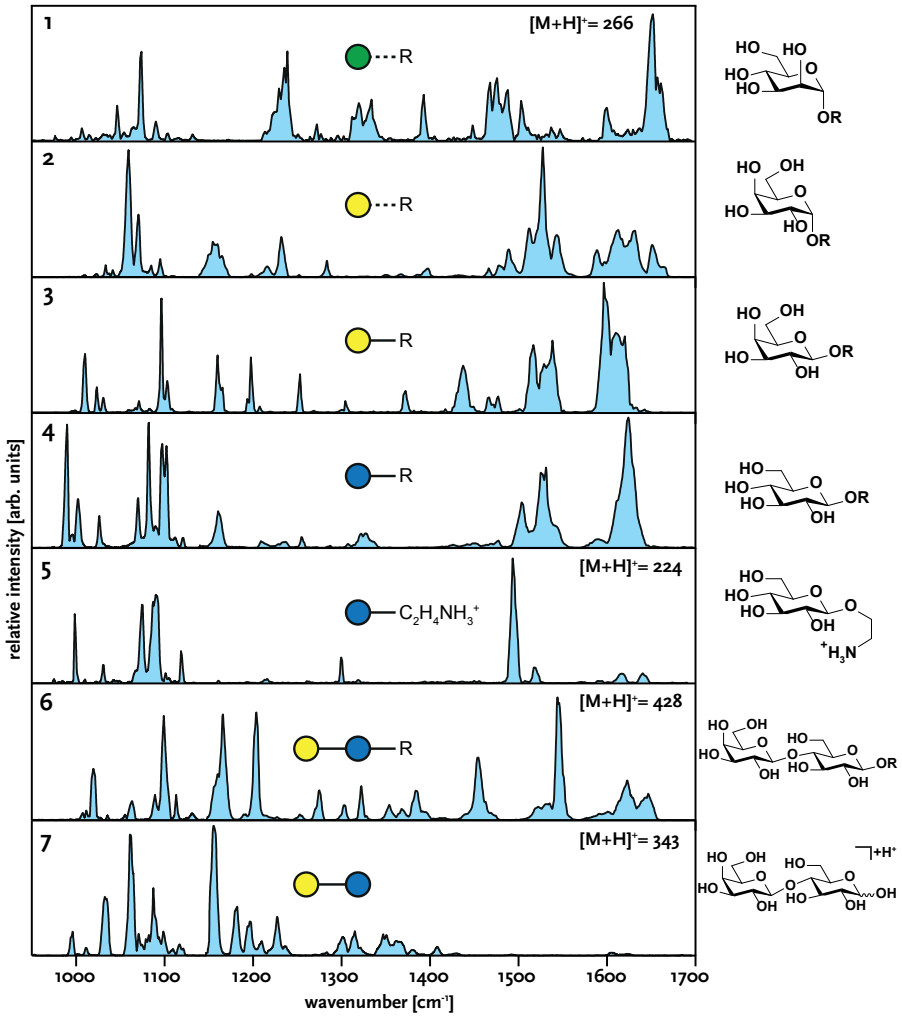


Figure 4.2: IR spectra of aminoalkyl-linked monosaccharides **1–7** as well as free (**7**) and aminoalkyl-linked lactose (**6**) investigated as $[M + H]^+$ ions.

signatures of both lactose variants, which indicates that their gas-phase structures are widely different. These structural differences most likely result from interactions between the charged linker and hydroxy groups of the disaccharide, or a distinct charge distribution in **7**. The absence of characteristic absorption bands above 1400 cm^{-1} confirms the assumption that this region is mainly governed by vibrational transitions of the charged aminoalkyl linker.

Trisaccharides

Next, a set of six isomeric trisaccharides was used to benchmark the method. The structures of trisaccharides **8**–**13** share the same reducing-end β-lactose core-motif and an aminopentyl linker, as shown in Figure 4.3. The terminal building block was systematically varied to generate isomers that differ in composition, connectivity or configuration. For example, the glycan pairs **8**/**9** and **12**/**13** are compositional isomers and only differ in the identity of the terminal building block (Glc vs. Gal). Glycan pairs **9**/**10** and **11**/**12** are configurational isomers and differ in the stereoconfiguration of the glycosidic bond. Finally, glycan pairs **8**/**13**, **9**/**12** and **10**/**11** are connectivity isomers that only differ in the position of the glycosidic bond (1→3 vs. 1→4).

These isomeric trisaccharides are extremely difficult to distinguish using established LC-MS techniques. Recent studies by Hofmann et al.[16] used ion mobility-mass spectrometry to approach this analytical challenge and measured the arrival-time distributions and collision cross sections of trisaccharides **8**–**13**. In negative ion mode, connectivity and configurational isomers were efficiently separated with baseline separation, allowing an unambiguous identification. The remarkable baseline separation also enabled a quantitative analysis of coexisting isomers inside a mixture. However, compositional isomers consistently

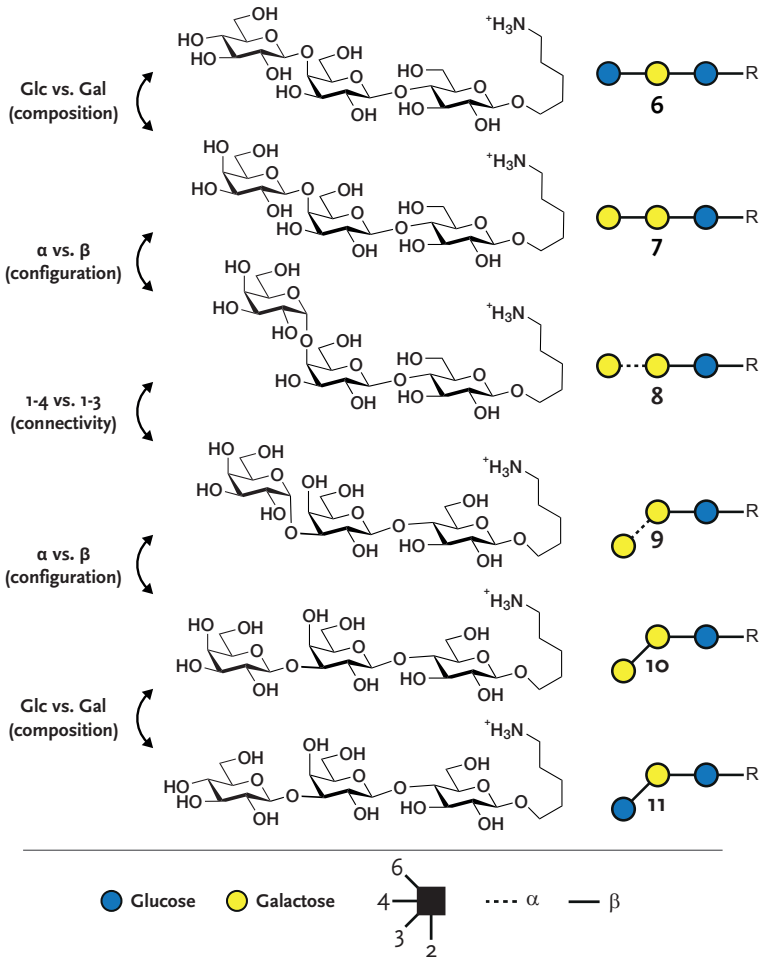


Figure 4.3: The synthetic trisaccharides **8–13** only differ in the connectivity (1→3 vs. 1→4), configuration (α vs. β), or composition (Glc vs. Gal) of the terminal building block. These isomeric oligosaccharides are extremely difficult to distinguish using established methods and therefore serve as a benchmark for this experimental approach.

showed very similar arrival time distributions and could therefore not be distinguished using IM-MS.

The IR spectra of trisaccharides **8**–**13** are shown in Figure 4.4. In general, each spectrum exhibits a large number of highly resolved absorption bands and no significant spectral congestion. The configurational isomers **9**/**10** and **11**/**12** share the same connectivity and configuration, but differ in the configuration of the terminal glycosidic bond (α/β). Distinct differences in their absorption patterns allow to readily distinguish these conformational isomers. Similar results are obtained for connectivity isomers **8**/**13**, **9**/**12** and **10**/**11**, which exhibit either a 1→3 or 1→4 terminal glycosidic bond. Again, the highly resolved optical signatures reveal differences that allow an unambiguous identification of each connectivity isomer. Most striking are the observations for compositional isomers **8**/**9** and **12**/**13**, which share the same connectivity and configuration and only differ in the identity of the terminal building block (Glc vs. Gal). In other words, the trisaccharide structures within each isomeric pair only differ in the stereochemical orientation of a *single* hydroxy group. Surprisingly, these minute structural variations lead to distinct differences in the IR signatures, especially above 1300 cm^{-1}. For example, a characteristic high-intensity transition is observed at 1450 cm^{-1} for trisaccharide **9**, but is absent in the spectrum of the corresponding compositional isomer **8**. Also trisaccharide **12** features a strong absorption band around 1310 cm^{-1} that is absent in the spectrum of the trisaccharide **13**. Taken together, each of the six trisaccharide isomers exhibits a variety of resolved absorption bands that lead to a unique IR signature and allow their unambiguous identification. Similarly to the previously discussed mono- and disaccharides, it is interesting to note that some IR spectra feature more than three absorption bands expected for the charged

aminopentyl linker. For example, trisaccharide **12** features five resolved bands, which indicates the presence of multiple coexisting conformers. Other trisaccharides such as **10** show a much cleaner spectrum with a few absorption bands that are as narrow as the bandwidth of the laser radiation.

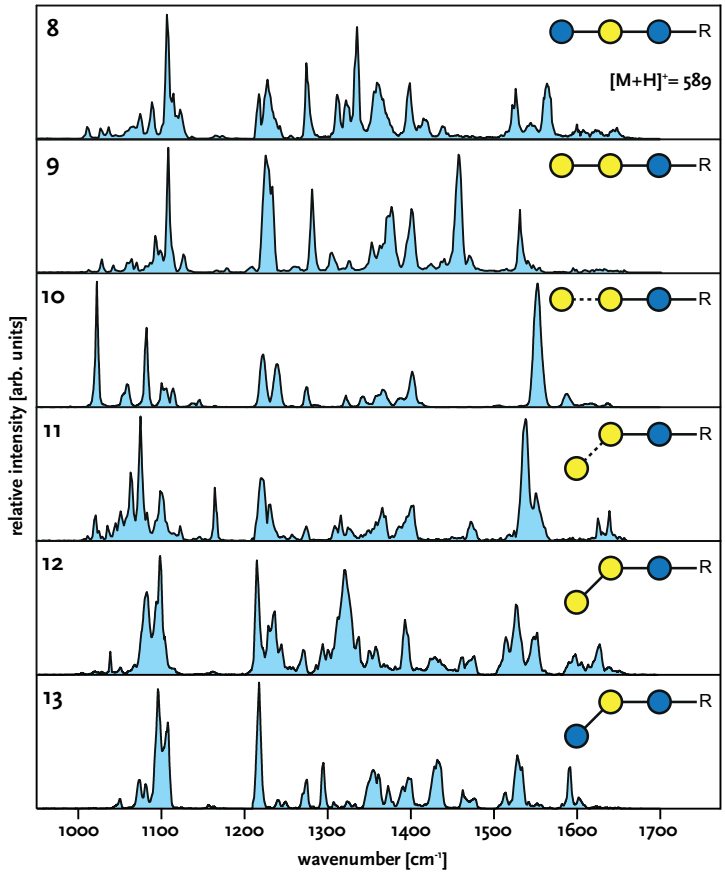


Figure 4.4: IR spectra of the isomeric trisaccharides **8–13**. Despite the marginal structural differences between these species, each trisaccharide can be readily distinguished by its unique absorption pattern that contains a variety of resolved bands.

Naturally Occurring Tetrasaccharides

To extend the scope of this method from synthetic standards to naturally occurring glycans, the blood-group antigens Lewis b (Le^b, **14**) and Lewis y (Le^y, **15**) were investigated as sodium adducts. Although it is in principle possible to generate protonated species of these samples, the high salt concentrations found in many samples typically lead to strong signals of sodium adducts. The isomeric tetrasaccharides each consist of an *N*-acetylglucosamine, a galactose and two fucose building blocks and differ in their glycosidic linkages. Similarly to free lactose **7**, the absence of a linker leads to an undefined anomeric center that can adopt either an α- or β-configuration. Although the number of expected vibrational transitions for molecules of this size is large, the corresponding IR spectra (Figure 4.5) exhibit a remarkably small number of well-resolved absorption bands that allow an unambiguous discrimination between the two isomeric species. Especially the Le^y tetrasaccharide shows an extremely clean IR spectrum with a total of eight features. The transitions around 1500 cm^{-1} and 1680 cm^{-1} likely stem from the amide group of the GlcNAc building block and are assigned as amide II and amide I bands, respectively. A single amide group is expected to give only one amide II and one amide I band. Here, however, multiple bands are observed, which indicates that multiple coexisting conformers with different amide band positions coexist in the gas phase.

As another example for naturally occurring glycans, lacto-*N*-neotetraose (**16**) and lacto-*N*-tetraose (**17**) were investigated. These isomeric tetrasaccharides belong to the group of human milk oligosaccharides and only differ in the connectivity of the terminal galactose building block. The corresponding IR spectra of the sodium adducts shown in Figure 4.5 also exhibit distinct and well-resolved absorption features

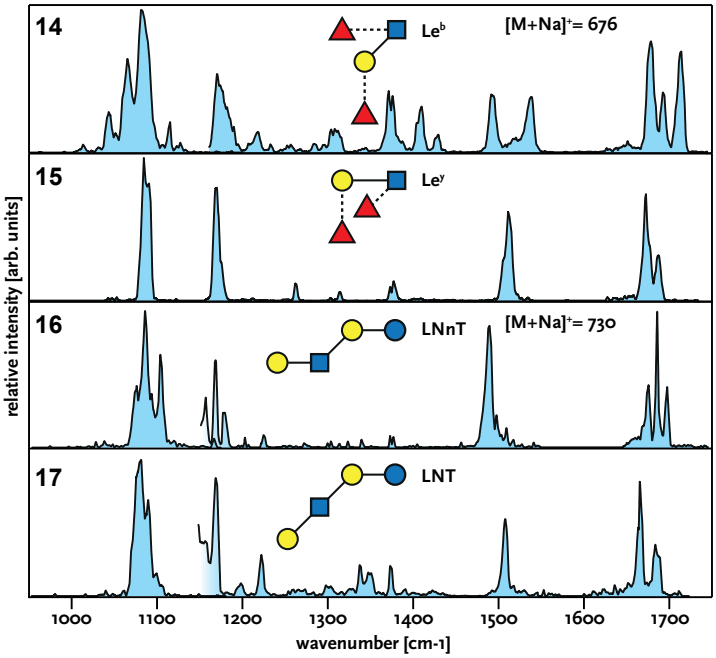


Figure 4.5: As representatives of biologically relevant glycans, the blood-group antigens Le^b (**14**) and Le^y (**15**), as well as the milk-sugar tetrasaccharides lacto-*N*-neotetraose (**16**) and lacto-*N*-tetraose (**17**) were investigated as $[M + Na]^+$ ions. The characteristic absorption patterns allow unambiguous discrimination between the corresponding isomers. The jump around 1150 cm^{-1} results from two different instrumental settings (see experimental details).

that allow their discrimination. Again, multiple bands in the amide II and amide I region indicate the presence of coexisting conformers.

4.4 What about Theory?

The highly resolved IR spectra obtained in this work allow an unambiguous identification of complex isomeric glycans. But what are the underlying structures that lead to those absorption patterns? In general, theoretical methods are widely used to compute structural candidates and calculate their corresponding IR spectra. These theoretical IR spectra are then compared to experimental results in order to extract the structural information that is encoded in the vibrational fingerprint, and identify the underlying structure.

Full of expectations about the results described in this chapter, I approached a knowledgeable theoretician and naïvely asked him to "run a few calculations" in order to identify the conformations of trisaccharides **8**–**13**. However, the highly resolved IR spectra didn't seem to spark much enthusiasm in said theoretician. Instead, the theoretician carefully tried to convey how complicated and time-consuming the demanded task really is. But why is it so difficult to identify the underlying structures of these samples?

Different computational methods have previously been employed to identify the structure of carbohydrates in the gas phase. Although only smaller systems were analyzed in these studies, a substantial computational effort was required for their theoretical treatment. Usually, the conformational space is rigorously scanned at a lower level of theory to generate a multitude of structural candidates. Selected structural candidates are then gradually refined with increasing levels of theory, while the size of the set is further reduced after each step until the most

relevant structural candidates undergo a frequency analysis to generate theoretical IR spectra. A general issue of this stepwise approach is the propagation of errors. For example, if relevant low-energy structures are not sampled during the initial screening of conformational space, or they are discarded as energetically unfavorable, the following structure optimization steps will most likely not be able to return this conformer to the pool of candidates. This is particularly challenging for complex glycans, because the initial sampling of possible conformers and their energy evaluation become increasingly difficult. To save computational time, the initial sampling often involves empirically derived force fields (FF). These FF-methods are highly valuable to simulate the behavior of carbohydrates in aqueous solution or their biological environment. However, they are parametrized for the condensed phase, often rely on crude approximations, and are potentially biased towards certain structural motifs. This bias often leads to the exclusive formation of $^4C_1$ ring puckers, which are certainly favored in aqueous solution. Isolated glycans in the gas phase, however, do not have a solvation shell and are subject to strong intramolecular hydrogen-bonding networks and charge-dipole interactions. Force fields are generally not parametrized for these boundary conditions and therefore severely underestimate non-chair ring puckers that might play key roles in isolated glycans.

These limitations advocate to omit FF-based methods and instead use first principles methods rooted in quantum mechanics such as density-functional theory (DFT). Already computationally inexpensive DFT methods promise an accuracy that far exceeds force field methods. Particularly, the novel SCAN functional has proven to reproduce energy hierarchies of mono- and disaccharides with greater accuracy than the popular hybrid functionals (B3LYP, PBE0, or M06-2X) at a fraction of the computational cost.[158] In combination with a genetic algorithm

that randomly generates possible structures, the conformational space of glycans can be efficiently screened without an inherent bias for certain structural motifs.

Another issue that needs to be addressed is the final evaluation of theoretical IR spectra. To compare theoretical and experimental IR spectra, the *Pendry R-factor* can be used to create an absolute measure and assign the best match.[159] However, this method has a limited application for complex IR spectra with narrow absorption bands. Instead, the best match is often identified on the basis of subjective estimates. The question arises, whether advanced algorithms can be employed to reliably compare experimental data with theory to avoid a subjective judgment by eye, which is prone to unconscious (or conscious) bias.

But even if the conformation that is responsible for the experimental observations is accurately reproduced by theory, another potential issue remains when calculating the vibrational transitions. Most quantum chemistry packages routinely predict IR spectra assuming that the vibrational transitions are harmonic. Complex carbohydrates, however, are believed to have anharmonic and coupled vibrational modes, which cannot be reproduced using the quantum harmonic oscillator as an approximation.

4.5 Comparison to Messenger Spectroscopy

Using superfluid helium nanodroplets is not the only possibility to investigate isolated ions at cryogenic temperatures. As described in Section 2.4, messenger spectroscopy is a viable alternative that does not require high-intensity radiation produced by a free-electron laser. Instead, commercially available benchtop laser systems (OPO/OPA)

provide sufficient pulse energies of up to 10 mJ for this spectroscopic scheme.

Shortly after publishing the results described in this chapter, the group of Thomas Rizzo (EPFL, Lausanne) published several studies on glycan analysis using a combination of IM-MS and messenger IR spectroscopy.[160–164] In their experimental setup, ion mobility-selected glycan ions are trapped and cooled to low temperatures by collisions with a cold buffer gas. The buffer gas typically consists of helium, hydrogen or nitrogen, and its condensation at cryogenic temperatures leads to the formation of messenger-'tagged' ions. The absorption of a single resonant photon transfers sufficient energy to overcome the dissociation threshold of the ion-messenger complex and induces photodissociation. Monitoring the wavelength dependent dissociation yield is then used to generate the ions' vibrational spectrum.

In 2017, Masellis et al. used this experimental setup to investigate a variety of isomeric disaccharides as well as two pentasaccharides at a temperature of 13 K using helium atoms as a messenger. In the spectral range between 3200 cm^{-1} and 3700 cm^{-1}, the large number of O−H- and N−H-stretching vibrations provide a structure-sensitive probe. The resulting absorption patterns as well as the arrival time distributions after ion mobility separation allowed an unambiguous identification of the isomeric species.[160] In subsequent studies, this approach was used to distinguish isomeric glycosaminoglycans (GAGs), and human milk oligosaccharides (HMOs) up to a size of hexasaccharides.[161,162]

A direct comparison between those results and the data presented in this chapter should be treated with caution because different spectral ranges are probed (1000–1800 cm^{-1} vs. 3200–3700 cm^{-1}), the spectroscopy schemes greatly differs (non-linear, multiple photon vs. linear, single-photon), and the internal temperature of the ions can be vastly different

(0.4 K vs. 13–69 K). In spite of that, a direct comparison shows that the vibrational signatures in messenger spectroscopy seem slightly more crowded and less resolved, which is probably caused by the ions' higher internal temperature. Nevertheless, their obtained spectral resolution is sufficient to distinguish even larger glycans based on their unique spectral signatures. In addition, a newly implemented high-resolution ion mobility cell using structures for lossless ion manipulation (SLIM) promises a powerful multidimensional approach to analyze isomeric glycans.[164]

4.6 Conclusion

In summary, cryogenic vibrational spectroscopy is a valuable addition to the structural analysis toolbox for glycans. The low temperature environment of superfluid helium nanodroplets allows to record highly resolved absorption spectra: a true spectral fingerprint that is unique for each glycan. Even minute structural differences such as the stereochemical orientation of a single hydroxy group within trisaccharides lead to spectral differences that allow an unambiguous identification. The sophisticated experimental setup used in this work involving a free-electron laser will arguably not find a commercial application in the glycosciences. However, similar results were recently reported by Rizzo et al. using a cryogenic ion trap and commercially available benchtop laser systems.

Considering this additional analytical dimension of highly resolved IR spectra, one is inclined to raise the question "*Quo vadis*, glycoanalysis?". This question specifically aims at the current lack of a universally accepted analytical technique that allows routine, high-throughput and comprehensive analysis of glycans with minimal sample requirements.

A very promising contender for such an analytical platform is the combination of mass spectrometry, ion mobility spectrometry, and cryogenic vibrational spectroscopy in a single instrument, as recently demonstrated by the group of Thomas Rizzo. The addition of ion mobility spectrometry has two crucial benefits. First, it adds another dimension of separation and allows to determine the ions' collision cross section. Second, by allowing to separate mixtures of isomeric glycans, it serves as a pre-filter and enables subsequent isomer-selective spectroscopic analyses. Such a multidimensional approach records the mass-to-charge ratio, the collision cross section, and the vibrational fingerprint as instrument-independent molecular properties. These complementary data sets can be stored in databases to provide a reference library for unknown samples, analogous to numerous LC-MS applications such as protein analysis, drug determination or pesticide residue analysis. For larger glycans, an incremental fragment-based analysis could aid the structural identification of unknown samples.

Chapter 5

Where The Fucose Are You?

Fucose is an essential deoxysugar found in a variety of biologically relevant glycans and their conjugates. A frequently observed phenomenon in tandem MS experiments of glycans is the intramolecular rearrangement of fucose units called *fucose migration*, which can lead to erroneous sequence assignments. Historically, fucose migration was only observed for fragment ions and was therefore strictly associated with the fragmentation process in tandem MS experiments. In this chapter, cryogenic ion spectroscopy is used to show that fucose migration is not limited to fragments and instead can be observed in intact glycan ions as well.

This chapter is based on reference [165] – Mucha, E., et al., Fucose Migration in Intact Protonated Glycan Ions: A Universal Phenomenon in Mass Spectrometry, *Angew. Chem. Int. Ed.* **2018**, *57*, 7440. Figures and content adapted with permission from John Wiley and Sons. Credits for the witty chapter title go to Michael T. Bowers, UC Santa Barbara.

5.1 Introduction

Fucose: A Special Building Block?

Fucose is a prevalent, but very unusual monosaccharide building block found in glycolipids and glycoproteins produced by mammalian cells. A distinct structural feature of fucose is the lack of a hydroxy group at the C6 position, which classifies it as a deoxy sugar (Fig. 5.1). In contrast to most other biologically relevant monosaccharides that are found in a D-configuration, naturally occurring fucose is solely found in the L-configuration. In oligosaccharides, fucose is almost exclusively found as a terminal or subterminal monosaccharide unit, attached during the final step of glycan processing by various fucosyltransferases in the endoplasmic reticulum or the Golgi apparatus.[166] Interestingly, fucose is only attached to other monosaccharides via α-glycosidic linkages. A structural investigation of more than 3000 mammalian oligosaccharides revealed that fucose is the fifth most abundant building block with an abundance of 7%, and the second most abundant terminal unit with an abundance of 24%.[167] Examples for fucosylated oligosaccharides and N-glycans are shown in Figure 5.1.

The Biological Roles of Fucosylated Glycans

Fucosylated glycans are frequently described using the term 'fucose-decorated'. In the authors opinion, this is highly misleading because it implies that fucosylation of glycans is merely cosmetic and not functional. However, enzymatic fucosylation is essential for a variety of biological processes. Several studies show that the lack of specific fucosyltransferases is lethal to embryonic mice, as the development of heart, blood vessels, neurons and other central structures was disturbed by severe defects.[169–172]

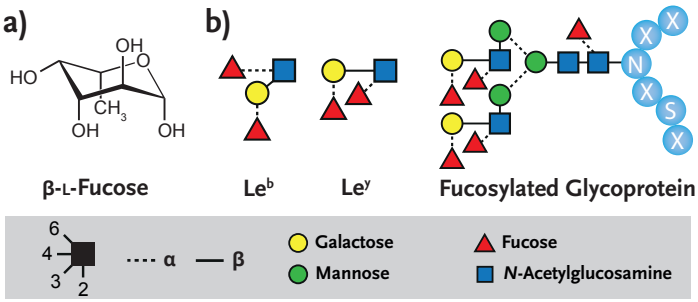


Figure 5.1: a) L-Fucose is a ubiquitous monosaccharide building block that lacks a hydroxy group at the C6 position. b) Lewis b and Lewis y tetrasaccharides as an example for biologically relevant fucosylated glycans. The highly fucosylated N-glycan was extracted from a glycoprotein found in parotid gland tissue.[168]

Fucosylated glycoproteins, for example, play key roles in neural development and growth, and are associated with the formation of long-term memory storage. Kalovidouris et al. identified the central structural motif within glycoproteins that mediate these cognitive processes: a terminal fucose unit linked to galactose via an α-(1→2) glycosidic bond.[173] In a subsequent study, they identified Synapsin I as the major glycoproteins carrying this fucose-motif in mature rat neurons and the hippocampus. Synapsin I glycoproteins are synapse-specific proteins involved in the regulation of neurotransmitter release. It was shown that the expression and turnover of these glycoproteins was regulated by fucosylation, which also prevents their degradation by proteases.[174]

Fucosylation also enables important functions during fertilization. In a variety of mammalian species including humans, fucosylated N-glycans mediate sperm binding to the egg's extracellular matrix.[175–178] Although the function of these proteins was unraveled in the late 1990s, it was not until 2011 that sialyl-Lewis x (sLe^x) was identified as the underlying structural motif in the sperm–egg recognition process.[178] As discussed in Section 2.2, the sLe^x tetrasaccharide is furthermore

a well-studied ligand for selectins and enables leukocyte adhesion on endothelial cells.

The structurally related Lewis b (Le^b) tetrasaccharide carries two fucose units and is involved in a variety of biological processes. For example, Helicobacter pylori – a bacterium found in the upper gastrointestinal tract of more than 50% of the world's population – recognizes Le^b motifs expressed by the host to attach to gastric epithelial cells. [179] For further information on the biological relevance of fucosylated glycans, the reader is referred to more comprehensive reviews in the literature. [166,180]

Fucose Migration in Tandem Mass Spectrometry

A frequently observed phenomenon for fucose-containing glycans in tandem MS experiments is an intramolecular rearrangement reaction that transfers terminal fucose units to adjacent or remote building blocks. This transfer of fucose units was coined *fucose migration* and manifests itself by the presence of misleading fragment ions, which in turn can lead to erroneous structural assignments. A typical fucose migration reaction is shown in Figure 5.2, using the tetrasaccharide lacto-*N*-fucopentaose as an example. As a result of collision induced dissociation (CID) of the protonated precursor ion, a misleading fragment ion is observed that suggests the presence of a fucose unit attached to the reducing-end lactose motif.

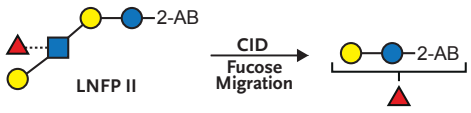


Figure 5.2: Fucose migration is an intramolecular rearrangement reaction, which is commonly observed in tandem MS experiments. This migration reaction leads to misleading fragment ions, which can lead to erroneous sequence assignments. In this example from Harvey et al., [181] the obtained fragment ion erroneously suggests that a fucose unit is glycosidically linked to the reducing-end lactose motif.

However, fucose was not the first building block for which a rearrangement reaction was observed. In 1983, McNeil investigated per-O-alkylated oligosaccharides and observed fragments that could only be explained by the loss of an internal building block, thus called *internal residue loss* (IRL).[182] Although this term can be somewhat misleading, because it focuses on the residue that is lost upon fragmentation and not the residue that is migrating across the glycan, it is used interchangeably with the terms fucose migration or the generalized *hexose migration*. In 1995, Kováčik et al. observed the migration of L-rhamnose, a 6-deoxy sugar isomeric to L-fucose, and reported fragment ions that were "difficult to rationalize".[183] Two years later, Ernst et al. first reported the migration of fucose units in sialyl-Lewis tetrasaccharides.[184] To date, numerous publications observed misleading fragment ions due to hexose migrations in various oligosaccharides.[181,185–191]

Hexose migration reactions are often compared to another gas-phase rearrangement reaction observed in the early days of proteomics, called *peptide scrambling*.[192–194] Both rearrangements reactions occur during the fragmentation of precursor ions, and both can lead to erroneous sequence assignments. However, there is one essential difference between peptide scrambling and hexose migration: whereas the mechanism that leads to peptide scrambling was unraveled using a variety of different analytical techniques,[195–198] the mechanism underlying hexose migration remains entirely unknown. Many structural prerequisites were proposed for fucose migrations to occur, such as an α-1→2 glycosidic bond of the migrating unit,[183] or the presence of a GlcNAc unit as a destination for the transfer reaction.[184] Based on these assumptions, various reaction mechanisms were proposed over the years,[181,183,184,187,188,191] but most of the initial assumptions were proven wrong by subsequent publications. Two observations, however, unite all of these publications: first, hexose

migrations were only observed in fragment ions, and second, only protonated precursor ions ($[M+H]^+$), or ammonium adducts ($[M+NH_4]^+$) lead to rearranged fragment ions. The latter observation led to the assumption that the migration reaction is proton catalyzed.

To avoid misleading fragment ions and prevent false sequence assignments caused by hexose migrations in positive ion mode, sodium or other metal adducts can be measured instead.[186] Also labeling of the reducing end,[199] or the use of free radicals[200] can prevent fucose migration. In negative ion mode, such a rearrangement reaction has not been observed.

In analogy to peptide scrambling, a variety of different methods is likely necessary to understand the principles of hexose migration and eventually understand the underlying reaction mechanism. An extensively used method that was used to unravel peptide scrambling was gas-phase vibrational spectroscopy, because it is highly sensitive to the molecular structure. In Chapter 4, vibrational spectroscopy using helium droplets was assessed as a powerful and structure sensitive method for glycan analysis. Here, it is used for a spectroscopic investigation of fucosylated glycans in order to gain further insight into this elusive migration reaction.

5.2 Experimental Details

Samples

Lewis and blood group antigens were purchased from Biozol (Eching, Germany) and used without further purification. For nano-electrospray ionization, aqueous 1 mM stock solutions were diluted with water/methanol (v/v, 1/1) to yield 100 µM analyte solutions.

Experimental Setup

Vibrational spectra of Lewis and blood group antigens were recorded using the experimental setup described in Chapter 3 using the on-axis detection scheme described in Section 3.6.2. For a brief summary of the experimental setup, see Section 4.2. Fragment ions were generated by tuning the source voltages to induce in-source fragmentation (see Appendix B). Each spectrum shown here is the result of averaging at least two independent scans recorded with a wavenumber step-size of $\Delta\tilde{\nu} = 2\,\mathrm{cm}^{-1}$. Although the yield of ions released from helium droplets scales non-linearly with laser energy, a linear correction has been performed by dividing the signal intensity by the laser energy. As a result, relative intensities should be regarded as a general guide. The spectra shown here were recorded between $1000\,\mathrm{cm}^{-1}$ and $1800\,\mathrm{cm}^{-1}$ with two different photon fluxes using a variable focusing mirror. A softer focus (lower photon flux) was used between $1000\,\mathrm{cm}^{-1}$ and $1200\,\mathrm{cm}^{-1}$. To access the absorption bands between $1200\,\mathrm{cm}^{-1}$ and $1800\,\mathrm{cm}^{-1}$, a separate spectrum with a tighter focus (higher photon flux) was recorded.

5.3 Results and Discussion

To investigate the structure and fragmentation behavior of fucosylated glycans, a set of structurally related tri- and tetrasaccharides of the Lewis and blood group antigen family was analyzed (Figure 5.3). The Lewis y (Le^y) and Lewis b (Le^b) tetrasaccharides are structural isomers and each consist of one *N*-acetylglucosamine, one galactose and two terminal fucose units. The four trisaccharides Lewis x (Le^x), Lewis a (Le^a), blood group antigen H1 (BG-H1), and blood group antigen H2 (BG-H2) only carry a single fucose unit and are substructures of the two tetrasaccharides. The anomeric center of these oligosaccharides is not defined and can therefore adopt an α- or β-configuration.

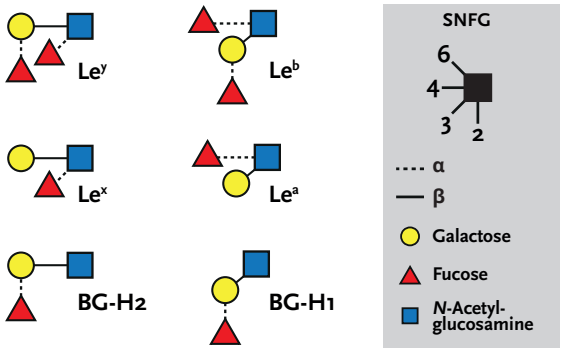


Figure 5.3: The fucosylated glycans investigated in this chapter. The trisaccharides Lewis x (Le^x) and blood group antigen H2 (BG-H2) are substructures of the Lewis y tetrasaccharide (Le^y). Analogously, the trisaccharides Lewis a (Le^a) and blood group antigen H1 (BG-H1) are substructures of the Lewis b tetrasaccharide (Le^b).

Sodium Adduct Ions

To validate the experimental approach for fragment analysis of fucosylated glycans, IR spectra of the Lewis and blood group antigens were recorded as sodium adducts. The results will serve as a refer-

ence, because sodium adduct ions effectively inhibit fucose migration reactions. IR spectra of Le^y and Le^b tetrasaccharides, measured as sodium adducts, were already presented in Chapter 4 and yield unique absorption patterns that allow their identification. Here, harsher source conditions were applied to induce in-source fragmentation of the intact precursor ions. Exemplary mass spectra using softer or harsher source conditions are shown in Appendix B.

Activation of the $[Le^y + Na]^+$ precursor ions ($m/z = 698$) yields an abundant signal that corresponds to the neutral loss of a single fucose unit at $m/z = 552$. Depending on which of the two fucose units is lost upon fragmentation, two different fragment structures are possible (Figure 5.4a). Because both trisaccharide fragments share the same mass, the mass spectrum alone cannot reveal whether one fragment structure is preferably formed over the other. Because the two trisaccharides Le^x and BG-H2 are substructures of Le^y, they are possibly formed during the fragmentation of the tetrasaccharide and can serve as structural references. To elucidate the fragment structure, IR spectra of the Le^y fragment as well as the intact trisaccharides Le^x and BG-H2 were recorded as sodium adducts (Figure 5.4b). All three IR spectra show well-resolved absorption bands in the fingerprint region between 1000 cm^{-1} and 1800 cm^{-1}. Similarly to the tetrasaccharides described in Section 4.3, the spectra show strong absorption bands around 1100 cm^{-1}, attributed to C−O-stretching vibrations, and additional features at 1520 cm^{-1} and 1680 cm^{-1} that likely stem from the characteristic amide II (N−H-bend) and amide I (C=O-stretching) vibrations of the *N*-acetylglucosamine unit. At closer inspection, the spectroscopic signatures of the two trisaccharides show enough differences to be unambiguously distinguished. The IR spectrum of the Le^y fragment, however, shares absorption bands with both trisaccharide

spectra. If both trisaccharides are formed during fragmentation of the Le^y precursor ion, the resulting spectral signature is expected to be a linear combination of the two individual trisaccharide spectra. Figure 5.4c shows a linear combination of the Le^x and BG-H2 spectra in a 1:1 ratio that matches the Le^y fragment spectrum surprisingly well. Although this approach is by no means quantitative, it suggests that the two trisaccharide structures are equally formed during fragmentation of the Le^y precursor ion.

These results are in accord with previous studies that utilized ion-mobility mass spectrometry to study Lewis and blood group antigens, and their characteristic fragment ions. Upon fragmentation of the Le^y sodium adduct, two separate peaks in the arrival time distribution are observed for the loss of a single fucose unit. Based on their distinct collision cross sections, the two substructures Le^x and BG-H2 were identified as fragments by comparison to the intact standards. The two separate peaks in the fragments arrival time distribution approximately share the same integrated area suggesting that both trisaccharide structures are about equally formed during the fragmentation process. Fragmentation of the isomeric Le^b tetrasaccharide, however, resulted only in a single peak in the arrival time distribution that could not be assigned to either of the intact trisaccharide substructures Le^a and BG-H1, because they have identical collision cross sections.[18]

In order to investigate the fragmentation behavior of the Le^b tetrasaccharide, cryogenic IR spectroscopy is used here. The sodium adduct of the Le^b precursor ion ($m/z = 698$) shows a similar fragmentation pattern as the isomeric Le^y antigen: an abundant signal ($m/z = 552$) corresponding to the loss of a neutral fucose unit is observed (Figure 5.5a). Again, two different fragment structures with the same m/z-ratio are possible depending on which fucose unit is cleaved

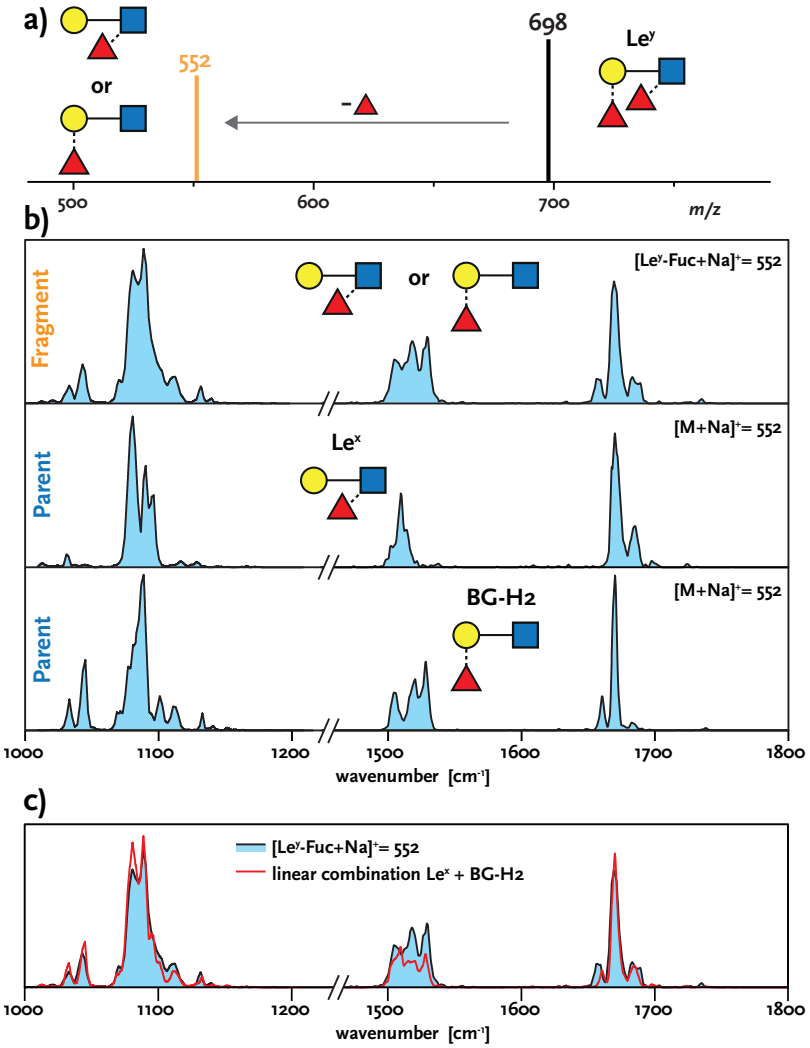


Figure 5.4: a) Schematic tandem MS spectrum of the Le^y tetrasaccharide as $[M + Na]^+$ species. The major fragment results from the neutral loss of a single fucose unit. b) IR spectra of the $[Le^y - Fuc + Na]^+$ fragment ion as well as the intact trisaccharide standards Le^x and BG-H2. The spectral region between 1200–1500 cm^{-1} only contains minor absorptions and is omitted for clarity. c) The IR spectrum of the Le^y fragment can be represented by a linear combination of the individual spectra of the trisaccharide standards (ratio 1:1). Complete IR spectra are shown in Appendix B.

upon fragmentation. The IR spectrum of this fragment as well as the two candidate structures Le^b and BG-H1 are shown in Figure 5.5b. The two intact trisaccharides show large spectral differences between 1000 cm^{-1} and 1150 cm^{-1} making them easily distinguishable, whereas the IR spectrum of the fragment ion shares absorption bands with both intact trisaccharides. The linear combination shown in Figure 5.5c matches the fragment spectrum well and again suggests that both trisaccharide structures are formed equally during fragmentation of the tetrasaccharide precursor ion.

Taken together, the fragmentation behavior of sodiated Le^y and Le^b tetrasaccharides was elucidated using cryogenic IR spectroscopy. As expected from the literature, the sodium adduct ions show no indication for fucose migration reactions. Both tetrasaccharides have two fucose units that each have approximately the same probability of being cleaved off during fragmentation. This suggests that the glycosidic linkages of the terminal fucose units have a very similar bond-dissociation energy.

Protonated Ions

In the literature, fucose migration reactions are reported for fragment ions, predominantly generated from protonated precursor ions. To identify possible rearrangement reactions, fragments of the Le^y and Le^b tetrasaccharides were investigated as protonated ions.

The fragmentation pattern of the protonated Le^y tetrasaccharide ($m/z = 676$) is very similar to its sodium adduct ion: the loss of a single neutral fucose unit ($m/z = 530$) generates an abundant fragment ion signal (Figure 5.6a). Again, two resulting fragment structures are possible depending on which of the fucose unit is cleaved upon fragmentation. The corresponding IR spectrum of this fragment is shown in Figure 5.6b and displays a number of resolved bands between

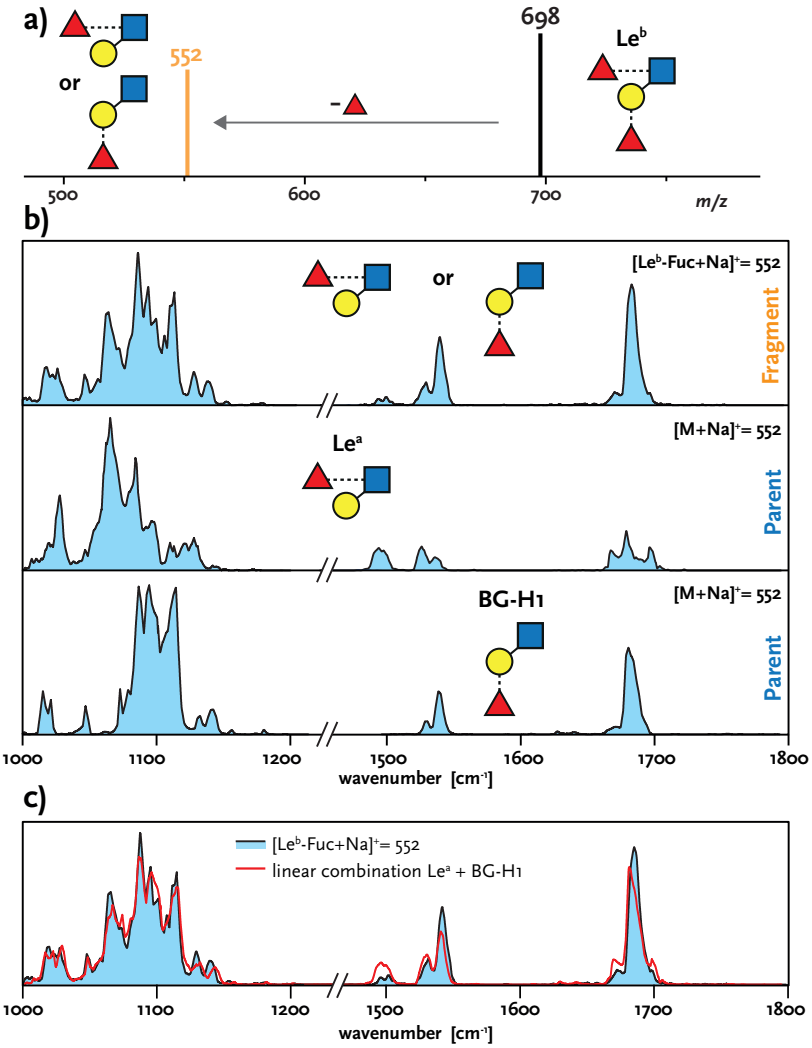


Figure 5.5: a) Schematic tandem MS spectrum of the Le^b tetrasaccharide as $[M + Na]^+$ species. The major fragment results from the neutral loss of a single fucose unit. b) IR spectra of the $[Le^b - Fuc + Na]^+$ fragment ion as well as the intact trisaccharide standards Le^a and BG-H1. The spectral region between 1200–1500 cm^{-1} only contains minor absorptions and is omitted for clarity. c) The IR spectrum of the Le^b fragment can be represented by a linear combination of the individual spectra of the trisaccharide standards (ratio 1:1). Complete IR spectra are shown in Appendix B.

1000 cm^{-1} and 1200 cm^{-1} and a single band at 1680 cm^{-1}. The two trisaccharides Le^x and BG-H2 are candidate structures that might be formed during fragmentation of the Le^y precursor ion and were investigated separately. Interestingly, the spectral signatures of both intact trisaccharides as well as the fragment ion are nearly identical. Every significant absorption band is found in each of the individual IR spectra, and only the relative intensity of some bands differs slightly between the Le^y fragment and the intact trisaccharide standards. At this point we have to ask the question: why are the three individual IR spectra almost identical?

As a first hypothesis, it may be postulated that the IR spectra are identical purely by chance. However, the results presented in Chapter 4 unambiguously show that this experimental technique is able to resolve subtle structural variations. Even changing the stereochemical orientation of a single hydroxy group in isomeric trisaccharides results in sufficient spectral differences making them easily distinguishable. Here, the trisaccharide structures are constitutional isomers, which should result in unique spectral fingerprints. To further evaluate how likely this hypothesis is, we have to take a closer look at the spectral signatures. If we imagine that only one or two absorption bands were present in each spectrum and the peak positions would happen to coincide, it would not be unusual. In this case, however, around nine resolved absorption bands share the same peak positions, most of them with comparable relative intensities. Therefore, the probability that both intact trisaccharides coincidentally yield the same IR spectrum is certainly different from zero, but negligibly small.

The second, and more likely hypothesis is the following: the three IR spectra are the same, because the *underlying chemical structures* are the same. More precisely, the Le^y fragment ion as well as both

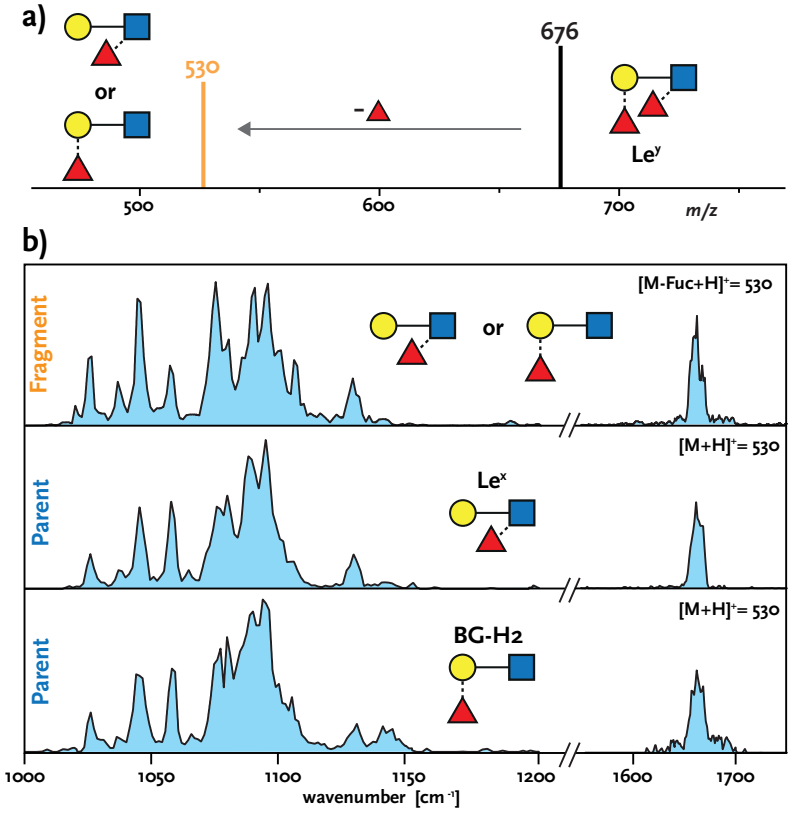


Figure 5.6: a) Schematic tandem MS spectrum of the Le^y tetrasaccharide as $[M + H]^+$ species. The major fragment results from the neutral loss of a single fucose unit. b) IR spectra of the $[Le^y - Fuc + H]^+$ fragment ion as well as the intact trisaccharide standards Le^x and BG-H2. The spectral region between 1200–1600 cm^{-1} only contains minor absorptions and is omitted for clarity. Complete IR spectra are shown in Appendix B.

intact trisaccharide ions undergo a rearrangement reaction – a fucose migration – to the same chemical structure, which leads to three identical IR signatures. Although the molecular structure of such a rearrangement product cannot be deduced from from the IR spectra without quantum chemical calculations, this hypothesis seems far more likely than pure coincidence*.

A possible explanation for the different peak intensities of individual bands, such as the band at 1040 cm^{-1}, is a different ratio of alpha and beta anomers between the intact ions and the activated fragment. The minor absorption band at 1145 cm^{-1}, which is only found in the spectrum of BG-H2, may stem from a small fraction of either a different molecular conformation or a different rearrangement product.

In previous studies, fucose migration was indirectly observed by the occurrence of misleading fragment ions in tandem MS experiments. Here, the spectroscopic signatures provide the first direct evidence for this elusive structural rearrangement. It was previously conjectured that fucose migration is caused by activation during collision-induced dissociation (CID) of precursor ions. Observing this rearrangement reaction in intact ions, however, shows that fragmentation is not a prerequisite for fucose migration. In general, it is not known whether the rearrangement reaction leads to an interconversion of one intact trisaccharide structure to the other, or whether both trisaccharides rearrange to a different, unknown structure. The energy barrier for fucose migrations can likely not be generalized and strongly depends on the glycan structure. In this work, however, soft source conditions were used to study the intact trisaccharide ions and minimize ion activation

*"*If it walks like a duck, and swims like a duck, and quacks like a duck, it probably is a duck, right?*".

during transfer into the mass spectrometer. Therefore, a low activation energy barrier can be expected for fucose migration in the Le^y series.

But what about the two structurally related trisaccharides of the Le^b series? To investigate, whether fucose migration can be observed in these samples as well, IR spectra of the Le^b fragment ion as well as the intact Le^a and BG-H1 trisaccharides were recorded (Figure 5.7b). The three IR spectra are generally rather congested and display less well-resolved absorption bands. The broad absorption pattern in the amide I region (1650 cm^{-1}) indicates that multiple conformers or chemical structures coexist for each structure. Overall, the three IR spectra show some similarities, but remain distinguishable. This suggests that fucose migration is either inhibited for these structures, or that the products after a possible rearrangement reaction are different.

Spectroscopy of Pre-cooled Ions

It is generally unknown, whether the observed fucose migration reaction occurs during electrospray ionization, during the transfer of ions from atmospheric to low pressure, or during the long trapping times in the hexapole ion trap prior to spectroscopy. To exclude chemical reactions inside the hexapole ion trap, the protonated trisaccharides Le^x and BG-H2 were pre-cooled to around 80 K inside the ion trap by buffer gas cooling before being picked up by helium droplets. In Figure 5.8, the resulting IR spectra are compared to those obtained without pre-cooling the ions. For the protonated Le^x trisaccharide, both IR spectra are virtually identical and only a minor band at 1690 cm^{-1} appears for the pre-cooled ions. This indicates, that the structures of the protonated Le^x trisaccharide do not change when the trap temperature is reduced. For BG-H2, on the other hand, the obtained spectrum using the cold trap is similar, but less resolved between 1000–1200 cm^{-1}. Between 1600–

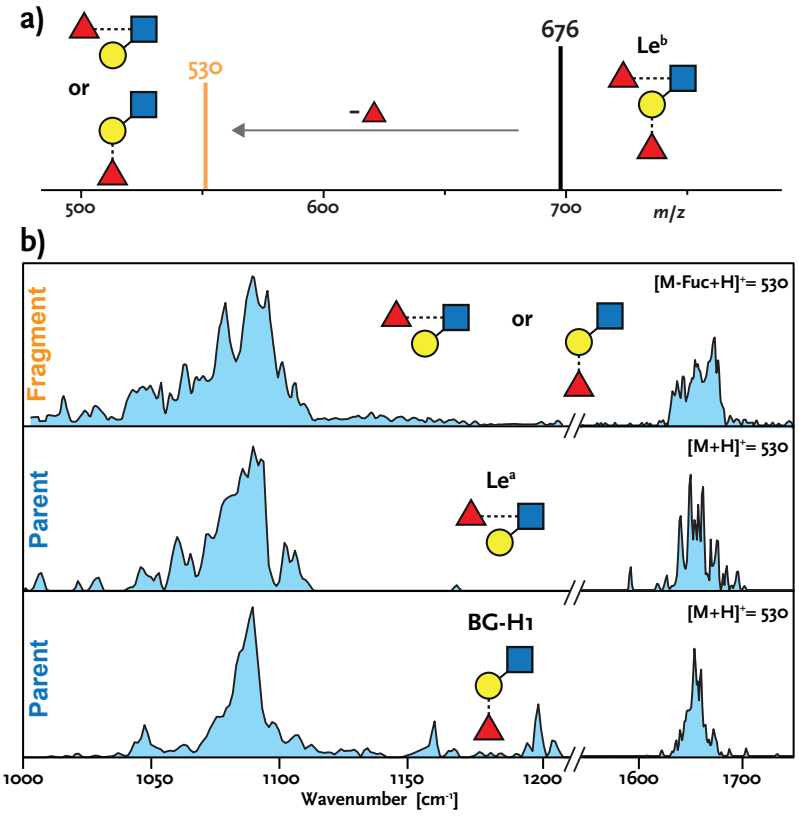


Figure 5.7: a) Schematic tandem MS spectrum of the Le^b tetrasaccharide as $[M + H]^+$ species. The major fragment results from the neutral loss of a single fucose unit. b) IR spectra of the $[Le^b - Fuc + H]^+$ fragment ion as well as the intact trisaccharide standards Le^a and BG-H1. The spectral region between 1200–1600 cm^{-1} only contains minor absorptions and is omitted for clarity. Complete IR spectra are shown in Appendix B.

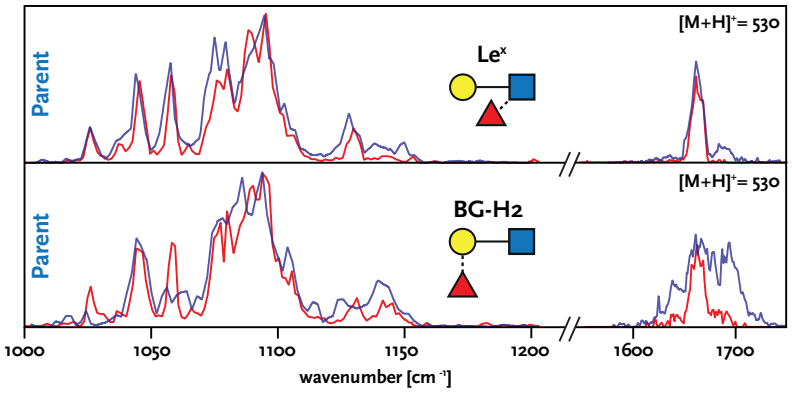


Figure 5.8: IR spectra of the intact Le^x and BG-H2 trisaccharides as $[M + H]^+$ species using a temperature of $T = 298\,K$ (red trace) or $T = 80\,K$ (blue trace) to thermalize ions inside the variable temperature ion trap prior to pick up by helium droplets.

$1730\,cm^{-1}$, a broad absorption band suggests the presence of multiple conformers, and/or chemical structures. This might indicate that fucose migration is partially inhibited for BG-H2 and a fraction of unreacted or intermediate structures coexists inside the cryogenic ion trap.

Observing a different IR signature upon pre-cooling the ion trap exclusively for protonated BG-H2 could have two causes: either BG-H2 interconverts to Le^x and reducing the trap temperature partially inhibits this conversion, or alternatively, both trisaccharides undergo a rearrangement reaction to a third unknown structure and the reaction rate is slower for BG-H2.

5.4 Conclusion

In summary, cryogenic IR spectroscopy reveals that fucose migration can occur in intact protonated glycan ions and is not necessarily a result of collision-induced dissociation (CID). Previous studies indirectly

observed fucose migration in tandem MS experiments by the presence of misleading fragment ions and coined the term 'internal residue loss'. The results obtained here suggest that this term is misleading for two reasons. First, the term focuses on the residues that are lost upon fragmentation, and not the migrating unit. Second, and more importantly, it is likely that the fragmentation itself is not the *reason* for the rearrangement reaction, but rather an *independent event* following the migration reaction. More precisely, fucose migration might occur prior to activation during collision induced dissociation and can only be observed in mass spectrometry when fragmentation occurs. Therefore, fucose migration has to be generalized to an issue that may universally occur in any type of mass spectrometry experiment. Observing fucose migration in intact protonated glycan ions furthermore suggests a potentially low activation energy barrier. In agreement with the literature, fucose migration was not observed in sodium adduct ions, which further supports the theory that the reaction is proton-catalyzed. The mechanism and final product of the rearrangement reaction remain elusive and require extensive theoretical calculations to understand the underlying processes. Recent studies by Lettow et al. further investigated the role of the mobile proton in fucose migrations and show that ammonium adducts also trigger fucose migration in intact Le^x and BG-H2 because the proton is located at the glycan, formally as an $[MH + NH_3]^+$ ion.[201]

Chapter 6

The Structure of Glycosyl Cations

Glycosylation reactions enable the synthetic assembly of oligosaccharides and build the foundation for modern glycochemistry. For the most part, these reactions are believed to proceed via a key ionic intermediate: the glycosyl cation. Although reaction pathways were postulated many years ago, the structure of glycosyl cations remained largely unknown due to the short-lived and elusive nature of these intermediates. This chapter presents an in-depth structural analysis of glycosyl cations using a combination of cryogenic vibrational spectroscopy and first-principles theory. The highly resolved infrared spectra allow detailed insights into the two fundamental motifs that enable stereoselective glycosylation reactions: neighboring group participation and remote participation.

This chapter is based on reference[202] – Mucha, E., et al., Unravelling the Structure of Glycosyl Cations via Cold-Ion Infrared Spectroscopy, *Nat. Commun.* **2018**, *9*, 4174 (neighboring group participation), and other, unpublished results (remote participation).

6.1 Introduction

Glycans are essential to life as their structural complexity enables a variety of biological functions including energy storage, structural scaffolds, cell-cell communication and molecular recognition (see Section 2.2). However, their vast structural diversity greatly impedes the full development of structural glycobiology. A major bottleneck in almost every branch of the glycosciences has been the lack of access to pure and structurally defined glycans. In contrast, proteomics and genomics heavily benefited from automated synthesis, enabling rapid progress in molecular biology.

One way to obtain glycan structures is the extraction from natural samples.[203–205] However, this strategy is often very time-consuming and usually fails to provide sufficient quantities of pure glycans to perform glycobiological studies. The presence of isomeric glycans and other impurities furthermore complicates their isolation and identification. Access to defined oligosaccharides in reasonable quantities therefore relies largely on chemical synthesis*. The central task in the chemical synthesis of oligosaccharides is seemingly simple: connect monosaccharide building blocks by forming glycosidic bonds. The realization of this task, however, remains a major challenge to date. Every addition of a monosaccharide building block requires the stereo- and regioselective formation of a glycosidic bond, which is usually aided by elaborate protecting group strategies. Permanent protecting groups, such as benzyl ethers, are employed to mask hydroxy groups of the target compound and are released at the very end of the assembly. Temporary protecting groups such as esters, on the other hand, are removed during synthesis to release

*In addition to chemical synthesis, other strategies such as enzymatic or chemoenzymatic synthesis are available, which are beyond the scope of this chapter and will not be discussed here.

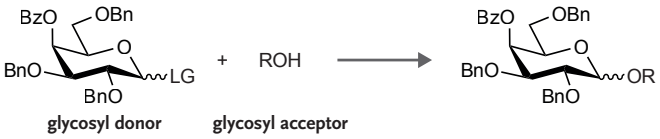


Figure 6.1: An exemplary glycosylation reaction using a protected galactose building block with a leaving group (LG) at the anomeric carbon. During oligosaccharide assembly, permanent protecting groups (e.g. Bn) are released after the last assembly step. Temporary protecting groups (e.g. Bz) are released after each coupling step to release a specific hydroxy group that acts as a nucleophile in the subsequent glycosylation reaction.

specific hydroxy groups, which are used as nucleophiles in subsequent coupling steps. As a consequence, regiocontrol during the glycosylation reaction is readily achieved. Stereocontrol, on the other hand, remains challenging to date.

The traditional solution-phase synthesis of oligosaccharides requires many purification steps in between and is very labor intensive.[206] Various strategies have been developed to accelerate and automate solution-based methods, for example by using one-pot syntheses,[207–209] convergent syntheses,[210–212] or fluorous tags that facilitate purification steps.[213–215] However, most approaches are not yet fully automated and remain time-consuming. A fully automated alternative to solution-phase synthesis is solid-phase synthesis. In analogy to the immensely successful solid-phase methods used for automated peptide and nucleic acid assembly, an automated glycan assembly platform was developed in the early 2000s.[155,216] Today, the Glyconeer 2.1™ is a commercially available glycan synthesizer that carries a library of around 20 building blocks used to assemble oligosaccharides via repeating cycles of glycosylation, capping and deprotection steps on a polystyrene resin support.[217,218] Each building block is assigned with specific reaction parameters such as reaction temperature, solvent, activator, and stoichiometry.[219] The

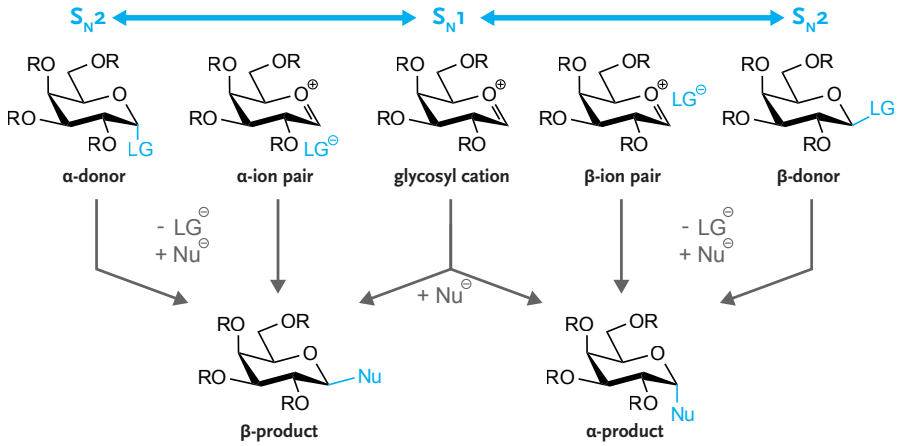


Figure 6.2: The stereochemical outcome of a glycosylation reaction strongly strongly depends on the type of reaction mechanism. S_N2 type reactions mark one end of the mechanistic continuum, which proceeds with inversion of configuration. At the other end, a dissociative S_N1 reaction can yield both diastereomeric products. The gap between S_N1 and S_N2 reactions is generally described by a mechanistic continuum of contact ion pairs.[222] LG: leaving group; Nu: nucleophile.

recently reported synthesis of a 50mer polymannoside is yet another milestone of this enabling technology.[220]

Despite these major technological developments in glycan synthesis and the various experimental breakthroughs, little is known about the actual mechanism of glycosylation reactions. Therefore, the development of novel building blocks with improved stereoselectivity or glycosylation yield is mostly based on empirical optimization.[221] A detailed understanding of the glycosylation reaction, however, is the foundation for rational building block design and improved methods.

Glycosylation reactions are usually described by two limiting mechanisms: a concerted, bimolecular S_N2 reaction through an associative mechanism, and a dissociative, unimolecular S_N1 reaction involving glycosyl cation intermediates. The gap between these two limiting

cases is often described by a mechanistic continuum of contact ion pairs between the glycosyl cation and the leaving group or promoter (Figure 6.2).[222]

The stereochemical outcome of a glycosylation reaction is directly linked to the type of reaction mechanism (S_N1 vs. S_N2). Reactions that proceed solely through an associative S_N2 reaction cause the inversion of configuration. Therefore, α-donors would only yield β-products, whereas β-donors would only yield α-products. However, many glycosylation reactions yield isomeric mixtures of α- and β-products, even if isomerically pure glycosyl donors are used. This observation is frequently rationalized by a shift towards the S_N1 side of the mechanistic continuum, where both α- and β-products can be formed.[222] Many parameters are believed to influence the stereochemical outcome of glycosylation reactions by shifting the mechanism within the S_N1/S_N2 boundaries.[223] The list of individual parameters that influence glycosylation reactions is long[222,224] and includes the nature of the donor, the acceptor, the type and configuration of the leaving group, the temperature, the type of solvent, the stoichiometry, water content, and the type of activator*.

A widely used method for the stereoselective formation of glycosidic linkages involves the participation of protecting groups. The two main strategies are *neighboring group participation* and *remote participation*, sometimes generalized as *anchimeric assistance*. Neighboring group participation (NGP), is frequently used to aid the formation of 1,2-*trans* glycosidic linkages by employing ester protecting groups at the C2 position, such as 2-*O*-acetyl or 2-*O*-benzoyl groups.[206,225] The reaction mechanism is believed to proceed via a key ionic intermediate, the glycosyl cation. Instead of an oxocarbenium ion, which would lead to

*some organic chemists would probably include the relative position of the sun, the earth and the moon to this list.

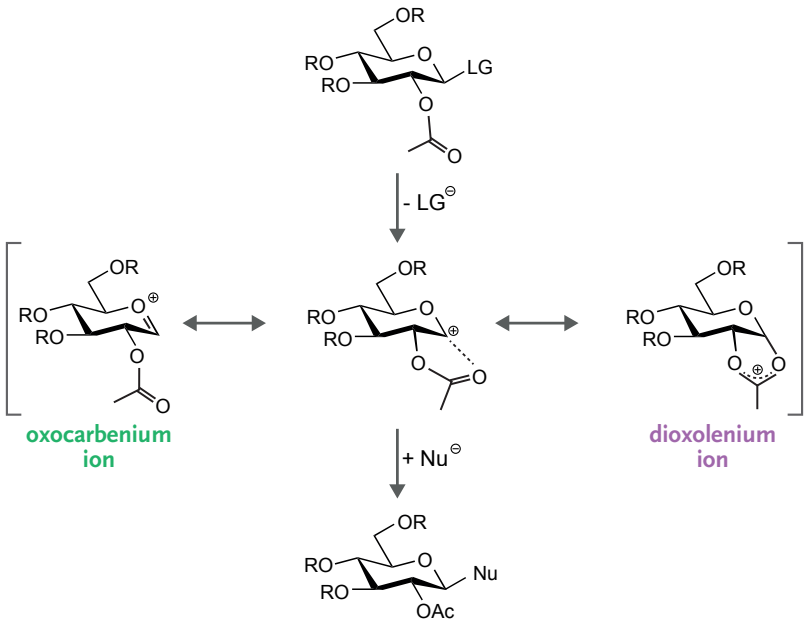


Figure 6.3: Participating protecting groups at C2 are widely used to aid the formation of 1,2-*trans* glycosidic linkages. Upon cleavage of a leaving group (LG), a glycosyl cation is formed. Instead of a non-stereoselective oxocarbenium ion, the reaction is believed to proceed via a dioxolenium ion, where the carbonyl oxygen forms a covalent bond with the anomeric carbon and promotes a nucleophilic attack from the *trans*-face. Nu: nucleophile.

the non-selective formation of α- and β-products, the reaction is believed to proceed through a dioxolenium ion, where the carbonyl oxygen of the ester forms a covalent bond with the anomeric carbon (Figure 6.3). In such an intermediate, the *cis*-face of the anomeric carbon is shielded, enabling a nucleophilic attack from the *trans*-side.[226–228]

The formation of dioxolenium ions, however, effectively shifts the positive charge and therefore the reactivity towards the former carbonyl carbon (C7) of the ester. As a consequence, another possible reaction mechanism was proposed: instead of a direct nucleophilic attack at

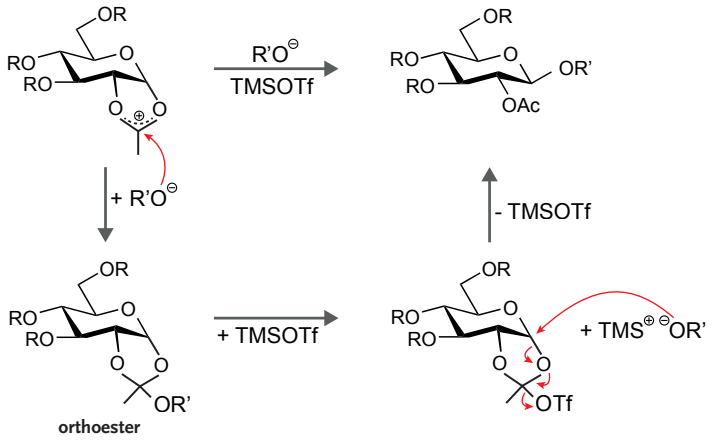


Figure 6.4: A nucleophilic attack at the dioxolenyl carbocation can lead to the formation of orthoesters, which are often isolated as side products during glycosylation reactions. Lewis or protic acids can be used to transform orthoesters to the corresponding product. [229–231]

C1, the product is formed via a nucleophilic attack at C7 that initially generates orthoesters (Figure 6.4). These orthoesters are often isolated as side products in glycosylation reactions and can be transformed to the corresponding products with the aid of protic or Lewis acids. [229–231] Another observation in favor of such a reaction mechanism is the cleavage of the C2 protecting group during glycosylation reactions. [232]

Although the stereoselective synthesis of 1,2-*trans*-glycosidic linkages using neighboring group participation may lead to side products, it is widely and successfully used in glycosynthesis. The stereoselective introduction of 1,2-*cis*-glycosidic linkages, on the other hand, is far more challenging and usually involves participating protecting groups at the C3, C4 or C6 position, [221,233–235] often referred to as remote participation. Chiral auxiliary protecting groups at C2 are also reported to yield 1,2-*cis*-glycosidic linkages with high stereoselectivity. This type of neighboring group participation, however, is not the focus of

this research and will therefore not be discussed here. The interested reader is referred to the literature.[236–239] Glycosylations using remote participation are often not as stereoselective as those using neighboring group participation, and are still subject to cumbersome empirical optimizations. In analogy to C2-participation, the reaction mechanism is believed to proceed via glycosyl cation intermediates, where the *trans*-side of the anomeric carbon is shielded by remote ester groups (Figure 6.5). This shielding was hypothesized either as a weak interaction, or a covalent bond between the carbonyl oxygen and the anomeric carbon.[206,234,240–242] A few studies report the isolation of side products such as oxazines and orthoesters, which suggest a mechanistic pathway via dioxolenium ions.[235,243–245]

Despite the widespread application of both neighboring group and remote participation, their underlying mechanisms remain poorly understood to date. Detailed structural knowledge about the key ionic intermediates – the glycosyl cations – would greatly enhance insight into the mechanistic pathways that are responsible for stereoselective glycosylation reactions. However, a major challenge that greatly impeded the experimental characterization of glycosyl cation intermediates is their short-lived and reactive nature, which prevents their isolation in the condensed phase. Recently, Martin et al. managed to overcome this limitation and stabilized glycosyl oxocarbenium ions using a hydrofluoric acid–antimony pentafluoride (HF/SbF_5) superacid.[246,247] Subsequent NMR analysis and computational chemistry was used to assign the structure of glycosyl cations from various building blocks. This admittedly 'brute-force' approach, however, does not generate glycosyl cations that occur during glycosynthesis, because the superacid medium also protonates many protecting groups. As a consequence, the investigated glycosyl cations carry a total of four to five charges that

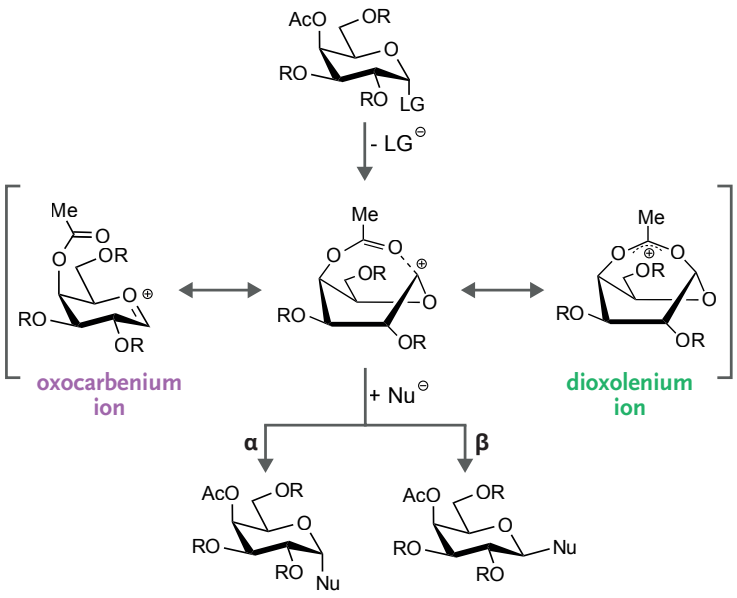


Figure 6.5: Participating protecting groups at C3, C4, or C6 are often used to aid the formation of 1,2-*cis* glycosidic bonds. Here, the remote participation of an acetyl group at C4 is exemplarily shown. The acetyl group is believed to interact with the anomeric carbon, effectively shielding the *trans*-face from a nucleophilic attack. The exact type of interaction was hypothesized as a weak interaction, or a covalent bond formation. LG: leaving group; Nu: nucleophile.

are expected to distort the cations' structure by Coulomb repulsion. The structural data can therefore not be directly translated to classical glycosylation conditions.

Another method to generate glycosyl cations is mass spectrometry. Ionization of precursor ions and subsequent collision-induced dissociation frequently leads to the formation of glycosyl cations in the clean room environment of a mass spectrometer, where the explicit absence of solvent and nucleophiles stabilizes these intermediates. The specific fragmentation behavior was previously monitored to investigate the stabilization of oxocarbenium ions by neighboring groups,[248,249] or as

an analytical tool to distinguish monosaccharide anomers[250] and disaccharide isomers.[251] Mass spectrometry alone, however, is inherently blind towards the ions' structure. The combination of mass spectrometry with structure sensitive methods, on the other hand, provides an opportunity to unravel the structure of glycosyl cations. In this chapter, cryogenic vibrational spectroscopy and first-principles theory are used to investigate glycosyl cations and identify the underlying structural motifs of neighboring group- and remote participation. Within the time frame of this work, the group of Boltje and coworkers used a very similar approach to study glycosyl cations,[252,253] which will be discussed separately in Section 6.5.

6.2 Experimental Details

Samples

Two sets of monosaccharide building blocks were used as precursors to generate glycosyl cations involved in neighboring participation, or remote participation. Synthesis was performed by the group of Peter H. Seeberger (Max Planck Institute of Colloids and Interfaces, Potsdam, Germany) and all samples were characterized by NMR and high-resolution mass spectrometry. All samples were used without further purification. HPLC-grade solvents were purchased from Sigma-Aldrich (now Merck MilliporeSigma). For nano-electrospray ionization, aqueous 1 mM stock solutions were diluted with water/methanol (v/v, 1/1) to yield a 50 µM analyte solution.

Experimental Setup

Glycosyl cations were generated by electrospray ionization of precursor ions and subsequent fragmentation using collision-induced dissociation (in-source). For a brief summary of the experimental setup, see Section 4.2. The original on-axis detection scheme described in Section 3.6.1 was used to record vibrational spectra glycosyl cations involved in neighboring group participation. The novel off-axis design described in Section 3.6.2 was used to record vibrational spectra of glycosyl cations involved in remote participation. In both cases, the trapped ions were thermalized to 80 K by buffer gas cooling before being picked up by helium droplets. Each spectrum shown here is the result of averaging at least two independent scans recorded with a wavenumber step-size of $\Delta\tilde{\nu} = 2\,\mathrm{cm}^{-1}$. Although the yield of ions released from helium droplets scales non-linearly with laser energy, a linear correction has been performed by dividing the signal intensity by the laser energy. As

a result, relative intensities should be regarded as a general guide. The spectra shown here were recorded between 1000 cm^{-1} and 1800 cm^{-1}.

Theory

Theory was performed by Mateusz Marianski (Hunter College, New York, NY, USA) and Kim Greis (Fritz Haber Institute, Berlin, Germany). Only a brief overview about the theoretical methods will be given here, as a full description is beyond the scope of this thesis. The interested reader is again referred to the original publication.[165]

Candidate low-energy structures were obtained in a stepwise procedure. For a reliable screening of conformational space, initial structures were generated using a genetic algorithm. This algorithm randomly changes torsional angles and the ring pucker to produce an initial array of structures. Atomic connectivity is not constrained during the structural search, which allows the formation of new covalent bonds. The generated structures are then optimized by an external software; in this work, FHI-aims[254,255] using dispersion-corrected PBE+vdWTS density functional theory.[256,257] Next, the genetic algorithm creates a set of child structures that each have features from two parent structures. The resulting structures are again optimized and evaluated by external DFT calculations. This process is repeated for a certain number of times, or until no change in the lowest-energy structure occurs after after a certain number of generations. In this work, the *flexible algorithm for optimization of molecules* (FAFOOM) was used to screen the conformational space of glycosyl cations.[158,258] In the following step, several low-energy structures were selected and further optimized at a higher level of theory using Gaussian 09[259] at the PBE0+D3/6-311+G(d,p) level of theory.[260,261] For each structure, a frequency analysis was performed within the harmonic approximation and the resulting vibra-

tional frequencies were scaled by a standard scaling factor of 0.965, and convoluted with a Gaussian profile. Finally, the potential energy of each conformer was computed using the Resolution-of-Identity MP2 method,[262] extrapolated to the complete basis set.[263] These energies were corrected by the free-energy contributions at 78 K, which are derived from the harmonic vibrational calculations.

6.3 Glycosyl Cations Featuring Neighboring Group Participation

In this section, the structure of three glycosyl cations carrying an acetyl group at C2 are determined by cryogenic vibrational spectroscopy and first-principles theory. The three corresponding precursor ions are shown in Figure 6.6 and are believed to exhibit neighboring group participation during glycosylation reactions. The glucose, mannose and galactose samples were custom tailored and carry a thioether as a leaving group, as well as methoxy protecting groups at C3, C4 and C6 to ease theoretical calculations by reducing the overall system size. To avoid lengthy names during the discussion, the three species are referred to as **2Ac-Glc**, **2Ac-Man** and **2Ac-Gal**.

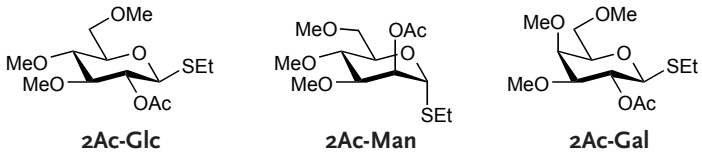


Figure 6.6: Glycosyl cations were generated from three precursor ions (ethyl 2-*O*-acetyl-3,4,6-tri-*O*-methyl-1-thio-D-hexapyranosides). The glucose, mannose and galactose precursors carry an acetyl group at C2, and a thioether at the anomeric carbon, which serves as a leaving group. The C3, C4 and C6 positions carry methoxy protecting groups.

Glycosyl cations were generated from in-source fragmentation of the respective thioglycoside precursor ions (see Appendix C). The recorded IR spectra are shown in Figure 6.7 and exhibit well-resolved absorption bands between 900 cm^{-1} and 1800 cm^{-1}. Vibrational transitions occurring in this spectral range can be divided into three regions. The first two regions are dominated by complex C−O and C−C stretching modes below 1250 cm^{-1}, and low-intensity C−OMe and C−H bending modes between 1250 cm^{-1} and 1450 cm^{-1}. Both regions, however, are expected to yield coupled and anharmonic vibrations that do not provide enough characteristic bands for an unambiguous assignment of the cation's structure. More characteristic bands are expected above 1450 cm^{-1}, where C=O stretching vibrations of the acetyl group or C−O−C vibrations of the dioxolenyl moiety are expected. The exact position of these vibrations is highly indicative of the interaction between the carbonyl group and the anomeric carbon; in oxocarbenium ions, strong absorptions above 1600 cm^{-1} indicate a free or weakly interacting carbonyl group, while dioxolenium ions are expected to show characteristic absorptions between 1500 cm^{-1} and 1600 cm^{-1}.

The theoretical treatment was performed using a genetic algorithm combined with density functional theory, as described above. For each glycosyl cation, ten individual genetic algorithm runs yielded around 200 unique structures. These structures readily divide into two groups by evaluating the distance $d_{\mathrm{O-C1}}$ between the carbonyl oxygen of the acetyl group and the anomeric carbon (C1). Oxocarbenium ions show a distance $d_{\mathrm{O-C1}}$ between 2.5–5 Å, whereas dioxolenium ions show a covalent bond with a distance $d_{\mathrm{O-C1}}$ of around 1.6 Å. A subset of the lowest-energy structures from both types of ions was further optimized using a higher level of theory. In addition, the relative free-energies for this subset of structures was further refined at the MP2

level of theory, extrapolated to the complete basis set, which show that dioxolenium ions are considerably more stable than oxocarbenium ions ($\Delta F_{harm} > 6\,\mathrm{kcal\,mol^{-1}}$). In Figure 6.7, predicted IR spectra of the lowest-energy conformers are compared to the experimental data. The corresponding glycosyl cation structures are shown schematically; detailed structures are shown in Appendix C.

The lowest-energy structure of the glucose variant (**2Ac-Glc**) is predicted to give characteristic symmetric and antisymmetric stretching modes of the dioxolenyl motif at $1500\,\mathrm{cm^{-1}}$ and $1570\,\mathrm{cm^{-1}}$, in good agreement with the experiment (Figure 6.7a). The absorptions below $1500\,\mathrm{cm^{-1}}$ are less indicative of the ions' structure, but the general agreement between experiment and theory supports the structural assignment. The covalent bond between the carbonyl oxygen and the anomeric carbon leads to a $^3S_1$ ring pucker, where the positive charge of the dioxolenyl motif is further stabilized by the axial methoxy group at C4. In addition, weak absorption bands are found above $1600\,\mathrm{cm^{-1}}$. The lower energy bands around $1615\,\mathrm{cm^{-1}}$ are assigned to [C1=O5$^+$] stretching modes of an oxocarbenium ion. The presence of a small fraction of oxocarbenium ions is further indicated by higher energy bands at $1720\,\mathrm{cm^{-1}}$ and $1780\,\mathrm{cm^{-1}}$, which likely originate from carbonyl stretching vibrations of the acetyl group.

For the epimeric mannose variant **2Ac-Man**, the two diagnostic bands measured at $1500\,\mathrm{cm^{-1}}$ and $1550\,\mathrm{cm^{-1}}$, which again originate from the dioxolenyl moiety, are well-reproduced by theory (Figure 6.7b). The satisfying agreement for the remaining bands below $1500\,\mathrm{cm^{-1}}$ further substantiate the structural assignment. This glycosyl cation structure is characterized by a dioxolenium motif that promotes the formation of a $B_{O,3}$ ring pucker, where the positive charge is stabilized by the spatially adjacent methoxy group at C6. But how sensitive is this method to the

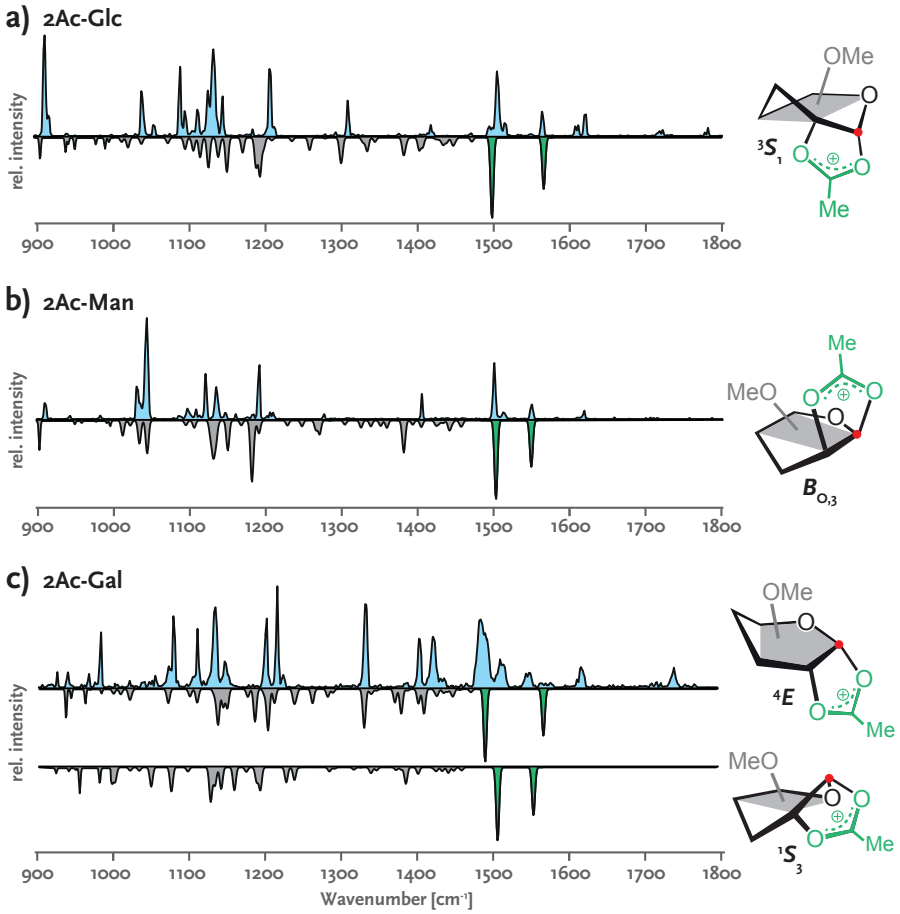


Figure 6.7: Comparison of the experimental (blue) and theoretical (gray) IR spectra of **2Ac-Glc** (a), **2Ac-Man** (b) and **2Ac-Gal** (c) glycosyl cations. The highly diagnostic spectral region above 1450 cm^{-1} reveals the identity and conformation of the respective glycosyl cation structures, shown schematically on the right. A red dot is used to highlight the anomeric carbon. Detailed structures are shown in Appendix C.

actual ring pucker? For this molecule, another low-energy candidate structure is found to have a $^3H_4$ ring pucker with distinctly different vibrations for the dioxolenyl motif, clearly dismissing this structure (see Appendix C). Above 1600 cm^{-1}, only one weak absorption indicates the minor presence of oxocarbenium ions.

The IR spectrum of the galactosyl cation **2Ac-Gal** is more congested than the two previous spectra. Additional intense bands between 1200 cm^{-1} and 1450 cm^{-1} appear, and the bands between 1450 cm^{-1} and 1600 cm^{-1} are broadened. Here, the lowest-energy structure adopts a ^4E ring pucker, where the positive charge is not stabilized by the surrounding substituents. As a consequence, the O−C1 bond distance is slightly elongated (1.58 Å) when compared to the **2Ac-Gal** and **2Ac-Man** glycosyl cation (1.55 Å, and 1.52 Å, respectively). The calculated IR spectrum explains the identity of most of the additional bands below 1450 cm^{-1}. The characteristic bands predicted for the ^4E structure at 1490 cm^{-1} and 1570 cm^{-1} agree reasonably well with the experimental data. Two additional smaller bands observed at 1510 cm^{-1} and 1550 cm^{-1} suggest the presence of another low-energy dioxolenium structure, which is characterized by a $^1S_3$ ring pucker. Therefore, two structures with different ring puckers likely coexist for the **2Ac-Gal** glycosyl cation. In addition, two bands above 1600 cm^{-1} indicate the presence of oxocarbenium ions.

6.4 Glycosyl Cations Featuring Remote Participation

After analyzing the structural details of neighboring group participation in the previous section, a detailed analysis of building blocks that are believed to exhibit remote participation is presented here. Instead of custom-tailored building blocks, the study focuses on glycosyl donors routinely used in glycosynthesis. The four structures are based on galactose equipped with acetyl and/or benzyl groups, as shown in Figure 6.8. Two building blocks carry a single acetyl group either at C4 (**4Ac**) or C6 (**6Ac**), and one building block carries two acetyl groups, both at C4 and C6 (**4,6Ac**). The last building block does not carry acetyl groups and is fully protected by benzyl groups (**Bn**).

A series of glycosylation reactions using these glycosyl donors serves as a basis to systematically investigate the effect of remote participation. The nucleophilic substitution with isopropanol was carried out at five different temperatures between –50 °C and 30 °C under well-defined reaction conditions using flow-chemistry. For each temperature, the α/β-product ratio was determined by HPLC analysis and is plotted in Figure 6.9. The glycosyl donors can be divided into two groups that either

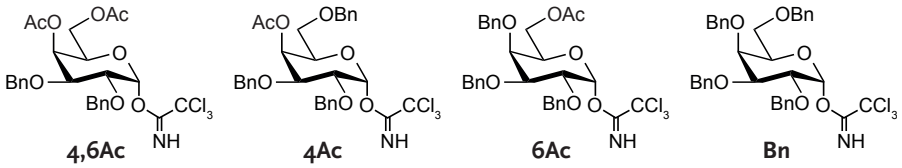


Figure 6.8: Structures of the four galactose building blocks. Two glycosyl donors carry a single acetyl group at either C4 (**4Ac**) or C6 (**6Ac**). One glycosyl donor carries an acetyl group at both C4 and C6 (**4,6Ac**), and one donor is solely equipped with benzyl protecting groups (**Bn**). All building blocks carry an α-trichloroacetimidate as a leaving group at C1.

yield mainly α-glycosides (**4Ac** and **4,6Ac**), or little α-glycosides (**6Ac** and **Bn**). If these glycosylation reactions would solely proceed via an associative S_N2 mechanism, the β-product would be formed exclusively as a result of stereochemical inversion. However, the reactions produce α-glycosides with a yield between 20% and 90%, suggesting a dissociative S_N1 mechanism. All building blocks show a consistent increase in α-selectivity with increasing temperature, which is in agreement with the general assumption that dissociative mechanisms are favored at higher temperatures due to a gain in entropy.[222,223] The difference in stereoselectivity between the two groups is consistent for all measured temperatures and suggests that different reaction intermediates are responsible for the outcome of the reaction.

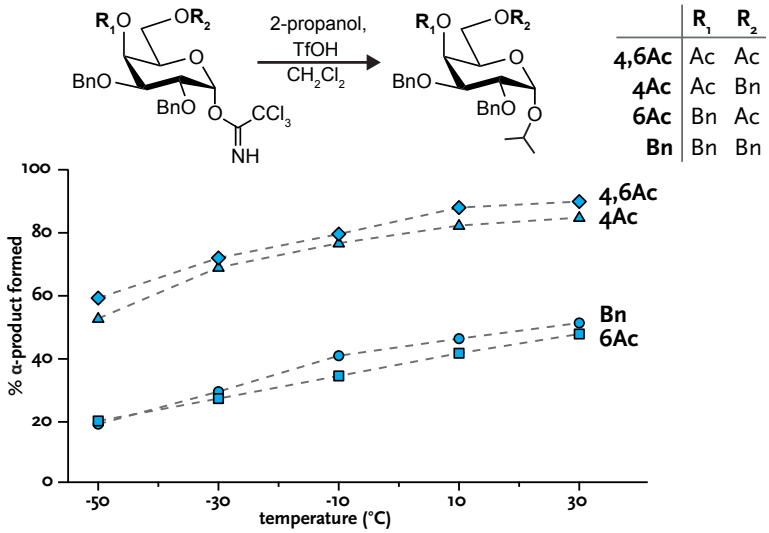


Figure 6.9: The nucleophilic substitution of the four galactose building blocks with isopropanol was used as a test reaction to determine the stereochemical outcome of the glycosylation reaction. The reactions were carried out at different temperatures between –50 °C and 30 °C and show an increasing α-selectivity with increasing temperature. Building blocks carrying an acetyl group at C4 (**4,6Ac** and **4Ac**) yield more α-products than building blocks carrying an acetyl group at C6 (6Ac) or no acetyl group at all (**Bn**).

But what are the underlying cationic structures that enable the formation of 1,2-*cis* glycosidic linkages? To answer this question, we focus on the structural investigation of glycosyl cations generated from the four galactose building blocks by in-source fragmentation (see Appendix C). The isolated glycosyl cations are characterized by combining cryogenic vibrational spectroscopy and first-principles theory. The recorded IR spectra are shown in Figures 6.10 to 6.13 and exhibit well resolved absorption bands between 1000 cm^{-1} and 1800 cm^{-1}. Here, the spectrum is divided into two main regions. Below 1200 cm^{-1} complex C−O and C−C stretching vibrations dominate the spectrum. The coupled and possibly anharmonic nature of these bands, however, is often not well-reproduced by the theoretical methods employed here. The region between 1200 cm^{-1} and 1800 cm^{-1}, on the other hand, features characteristic modes such as C=O, C−O−C, or O−C−O stretching vibrations that enable the identification of the glycosyl cation's structure.

The theoretical treatment was performed in the same way as for the C2-acetylated samples described in the previous section. Ten individual genetic algorithm runs yielded around 300 unique candidate structures for each glycosyl cation. The obtained structures again divide into two groups according to the distance d_{O-C_1} between the carbonyl oxygen of the acetyl group and the anomeric carbon (C1). Dioxolenium ions form a covalent bond to the anomeric carbon with a distance d_{O-C1} of around 1.5 Å. In other structures, the acetyl groups do not interact with the anomeric carbon and show a distance d_{O-C1} between 3–6 Å. A subset of low-energy structures from both groups was further reoptimized using a higher level of theory. For each reoptimized structure, single-point energies were calculated at the MP2 level of theory, extrapolated to the complete basis set.

The di-acetylated **4,6Ac** building block has the highest selectivity for α-glycosidic bond formation. The vibrational spectrum of the corresponding glycosyl cation is shown in Figure 6.10 and displays six well-resolved absorption bands. The lowest-energy structure **A** of this species is predicted to adopt a $^1S_5$ ring pucker featuring a covalent bond between the carbonyl oxygen of the C4-acetyl group and the anomeric carbon. The characteristic absorption bands associated with this dioxolenium motif are the symmetric and antisymmetric O−C−O-stretching modes predicted at 1463 cm^{-1} and 1566 cm^{-1}, both in very good agreement with the experiment. Two additional bands for the non-interacting C6-acetyl group are predicted at 1208 cm^{-1} and 1778 cm^{-1}, in line with the experimental data. Theory also predicts an alternative dioxolenium ion **B**, where the carbonyl oxygen of the C6-acetyl group forms a covalent bond with the anomeric carbon. This structure, however, is predicted to have a free-energy of around 5 kcal mol^{-1} higher compared to structure **A**, and the calculated spectrum agrees less well with the experimental one. Oxocarbenium ions, where the anomeric carbon does not interact with either of the acetyl groups are energetically unfavored, as they show a free-energy difference of more than 9 kcal mol^{-1} compared to **A** (see Appendix C.). In summary, the vibrational signature of the **4,6Ac** glycosyl cation provides direct evidence that the C4-acetyl group can form a covalent bond with anomeric carbon, yielding dioxolenium ions, which enable α-selectivity. Participation via the C6-acetyl group as well as oxocarbenium ions are energetically unfavored.

A similar C4 dioxolenium structure could also be responsible for the high α-selectivity of the **4Ac** glycosyl donor. The infrared spectrum of the corresponding glycosyl cation exhibits characteristic absorption bands between 1450 cm^{-1} and 1600 cm^{-1}, similar to those observed for the **4,6Ac** cation. The lowest-energy structure **C** features a covalent bond

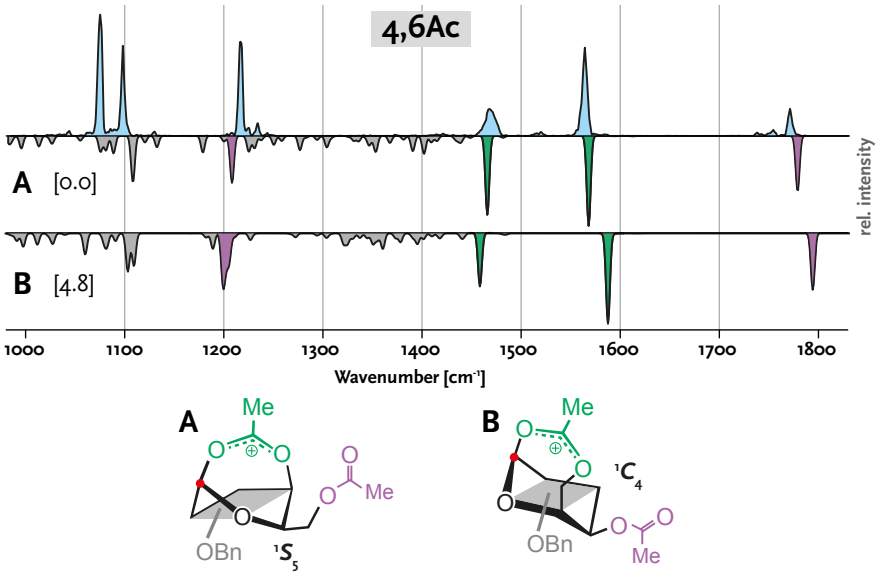


Figure 6.10: Experimental vibrational spectrum (blue) of the glycosyl cation generated from the **4,6Ac** precursor. Calculated spectra (gray, inverted) correspond to the low energy structures shown below. The relative free-energies in kcal mol^{-1} are shown in brackets. Complete structures are shown in Appendix C. The highlighted absorption bands indicate vibrations of non-interacting acetyl groups (purple) and participating acetyl groups (green). A red dot is used to highlight the anomeric carbon.

between the carbonyl oxygen at C4 and the anomeric carbon, which leads to a $^1S_5$ ring pucker. The predicted absorption bands at 1465 cm^{-1} and 1568 cm^{-1} agree well with the experiment, but the presence of multiple absorption bands in this region suggest that a structurally very similar dioxolenium ion coexists ($\Delta F = 0.7$ kcal mol^{-1}, see Appendix C). A closer inspection of the IR spectrum shows that two additional low intensity bands are found at 1220 cm^{-1} and 1780 cm^{-1}, indicating that a small fraction of glycosyl cations adopt structures with a non-interacting acetyl group. Surprisingly, theory predicts the low-energy oxonium structure **D**, which is characterized by a covalent bond between the O6-oxygen

and the anomeric carbon (1.52 Å). Observing this mode of participation is particularly surprising, because benzyl groups are generally considered as non-participating protecting groups. Recently, this mode of remote participation was proposed by Boltje et al. after observing a 1,4-anhydro-3,6-lactone as a side product during the stereoselective synthesis of 1,2-*cis* mannosides. This side product strongly suggests the presence of a glycosyl oxonium ion with remote participation of the C4 benzyl group.[239] The predicted oxonium structure **D** has a free-energy of around 4 kcal mol^{-1} higher, compared to structure **C**, and the vibrations of the carbonyl group at 1227 cm^{-1} and 1763 cm^{-1} align well with the experimentally observed bands. An alternative structure with similar band positions is oxocarbenium ion **E**. This structure, however, has a large free-energy difference of 11.9 kcal mol^{-1} compared to **C**, and a less good agreement between theory and experiment. Therefore, the spectroscopic data suggests that the **4Ac** glycosyl cation predominantly adopts an α-selective dioxolenium structure, similar to the one observed for **4,6Ac**.

The two glycosyl donors **6Ac** and **Bn** consistently produce a low to medium abundance of α-products, suggesting that the underlying reaction intermediates are structurally different from those observed for **4,6Ac** and **4Ac**. The infrared spectrum of the glycosyl cation generated from the **6Ac** precursor is shown in Figure 6.12 and displays a variety of absorption bands between 1200 cm^{-1} and 1800 cm^{-1}. The lowest free-energy structure **F** is characterized by a $^{1}C_{4}$ ring pucker and a covalent bond between the carbonyl oxygen of the C6-acetyl group and the anomeric carbon. The predicted vibrations of this dioxolenyl motif at 1455 cm^{-1} and 1590 cm^{-1}, however, do not match the experimental data particularly well. Alternatively, theory predicts remote participation of the C4-benzyl group, which results in oxonium struc-

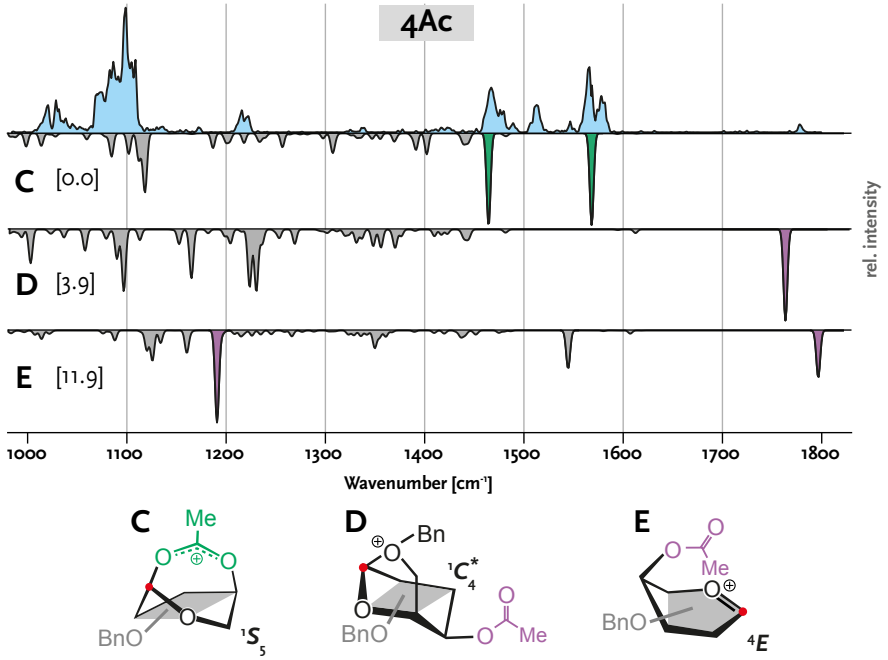


Figure 6.11: Experimental vibrational spectrum (blue) of the glycosyl cation generated from the **4Ac** precursor. Calculated spectra (gray, inverted) correspond to the low energy structures shown below. The relative free-energies in kcal mol^{-1} are shown in brackets. Complete structures are shown in Appendix C. The highlighted absorption bands indicate vibrations of non-interacting acetyl groups (purple) and participating acetyl groups (green). A red dot is used to highlight the anomeric carbon. Structure **D** adopts a distorted ring pucker, indicated by a star (*).

ture **G**. This structure is characterized by a covalent bond between the C4-oxygen and the anomeric carbon (1.51 Å) that leads to a $^{1,4}B$ ring pucker. The characteristic carbonyl vibrations of the acetyl group are calculated at 1220 cm^{-1} and 1762 cm^{-1}, which is in good agreement with the experimental data. Theory also predicts another, energetically very similar oxocarbenium structure **H**, where the *trans*-side of the anomeric carbon is shielded by the acetyl group ($d_{\mathrm{C{=}O{-}C_1}} = 2.8$ Å).

In addition to the carbonyl vibrations at 1222 cm^{-1} and 1749 cm^{-1}, the characteristic C1=O5$^+$ stretching vibration of the oxocarbenium motif is predicted at 1567 cm^{-1}. All three bands are in good agreement with the experimental data. Although both structures **G** and **H** are energetically less favored than dioxolenium ion **F**, the match between theory and experiment shows that either one, or both of these intermediates exist as the dominant structures. Combining these results with the low to medium abundance of α-products in the test reactions suggests that the formation of dioxolenium ion **F** might be kinetically suppressed by a higher activation energy barrier. Instead, the formation of energetically similar α-selective oxonium structures **G** and the non-selective oxocarbenium structures **H** might compete during the reaction. The relative energetics between these structures could also be different in the condensed phase reaction conditions. Furthermore, it is possible that the mechanistic continuum does not shift as strongly towards the S_N1-side, compared to the **4,6Ac** and **4Ac** donors. It has to be noted that the triplet of absorption bands measured between 1300 cm^{-1} and 1500 cm^{-1} is not reproduced by theory. Either another (unknown) structure is present, or the harmonic approximation cannot predict the position and/or intensity of certain absorption bands.

The low abundance of α-glycosides obtained from **6Ac** is comparable to the fully benzylated **Bn** building block. The vibrational spectrum of the corresponding glycosyl cation is shown in Figure 6.13 and does not feature any significant absorption bands above 1200 cm^{-1}. Theory predicts the low-energy structure **I**, which is characterized by a $^1C_4$ ring pucker and a covalent bond between the C6-oxygen and the anomeric carbon (1.53 Å). The predicted spectrum for this oxonium structure does not show any strong absorption bands above 1200 cm^{-1} and therefore provides a good match to the experimental data. The

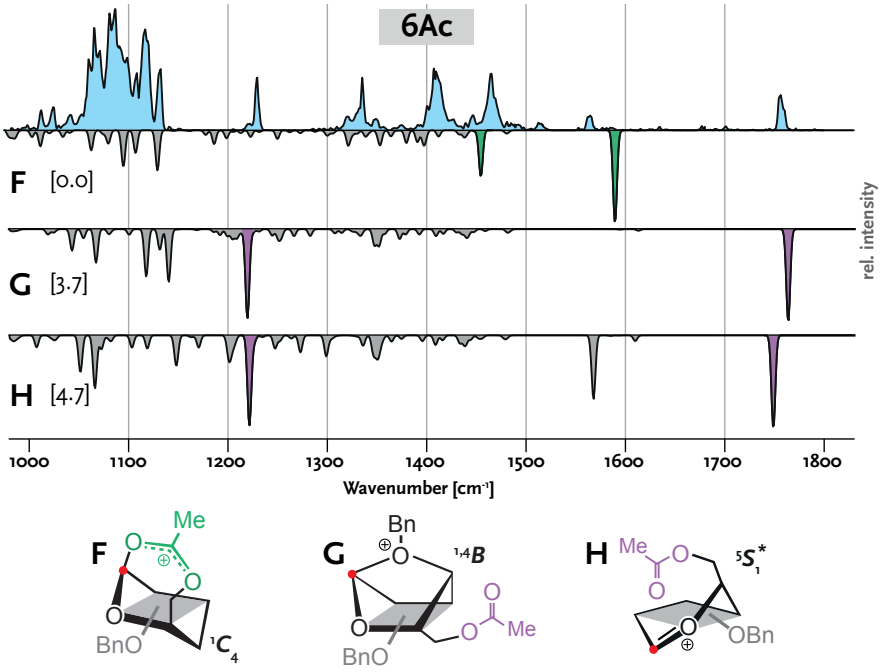


Figure 6.12: Experimental vibrational spectrum (blue) of the glycosyl cation generated from the **6Ac** precursor. Calculated spectra (gray, inverted) correspond to the low energy structures shown below. The relative free-energies in kcal mol^{-1} are shown in brackets. Complete structures are shown in Appendix C. The highlighted absorption bands indicate vibrations of non-interacting acetyl groups (purple) and participating acetyl groups (green). A red dot is used to highlight the anomeric carbon. Structure **H** adopts a distorted ring pucker, indicated by a star (*).

qualitative match for the bands below 1200 cm^{-1} is reasonable as well. Surprisingly, the gain in free-energy caused by this oxonium motif is significant: the lowest-energy oxocarbenium structure **J** is less stable by 4.6 kcal mol^{-1}. This structure adopts a ^4E ring pucker and exhibits a cation–π-interaction between the C2-benzyl group and the anomeric carbon, effectively shielding the *cis*-face from a nucleophilic attack. The importance of cation–π-interactions during glycosylation reactions

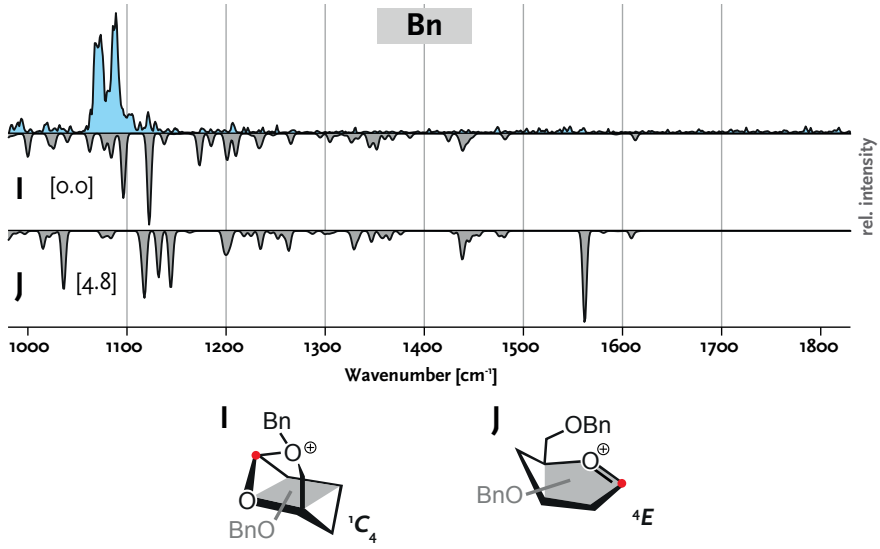


Figure 6.13: Experimental vibrational spectrum (blue) of the glycosyl cation generated from the **Bn** precursor. Calculated spectra (gray, inverted) correspond to the low energy structures shown below. The relative free-energies in kcal mol^{-1} are shown in brackets. Complete structures are shown in Appendix C. A red dot is used to highlight the anomeric carbon.

was long hypothesized and very recently demonstrated by Asensio et al., who studied model glycosyl donors equipped with aromatic platforms.[264] The stabilized oxocarbenium motif in structure **J** is predicted to give a strong absorption band at 1562 cm^{-1}, which is not found in the experimental spectrum.

Although glycosylation reactions through the low-energy oxonium structure **I** should be highly α-selective, the test reactions yield α/β-ratios of around 1:1 at best. It is possible that the mechanistic continuum does not shift as strongly towards the S_N1-side, or that the energy difference between α-selective structure **I** and β-selective structure **J** is decreased in solution. It has to be noted, however, that the strong nucleophile used in the test reactions generally leads to more S_N2-type reactions.[222,265]

In oligosaccharide synthesis, nucleophiles are usually weaker, and highly α-selective reactions were reported for this particular building block (α:β = 14:1).[221] Therefore, the α-selective oxonium structure **I** might be an important intermediate of this building block during glycan assembly.

6.5 Comparison to Recently Published Results

Within the time frame of this work, Boltje et al. used IR multiple photon dissociation (IMPRD) spectroscopy to elucidate the structure of glycosyl cations in the gas phase. Their approach is very similar to the one used in this work; glycosyl cations are generated from precursor ions by collision-induced dissociation, and subsequently investigated using vibrational spectroscopy. A marked difference is the spectroscopic analysis and in particular the temperature of the ions during the measurement. Whereas cryogenic ions are utilized in this work, IRMPD measures ions at room temperature and thermally activates them during the measurement (for more details regarding both techniques, see Sections 2.4 and 2.5). In their earlier work, Boltje et al. proved the utility of IRMPD spectroscopy by unraveling the structure of glycosyl cations from various building blocks including mannuronic acid-, mannose-, glucose- and galactose derivatives.[252] In a more recent work, they designed a mannuronic acid 6,3-lactone building block equipped with an acetyl group at C4, which enables the selective formation of 1,2-*cis* glycosidic bonds. Spectroscopic and computational analysis of the corresponding glycosyl cation shows that remote participation of the C4-acetyl group leads to the formation of energetically preferred dioxolenium ions, enabling stereoselectivity.[253]

Apart from the two fundamentally different spectroscopy schemes, also two different computational methods were employed. Instead of the DFT-based genetic algorithm used in this work, Boltje et al. employed

a molecular mechanics/dynamics (MM/MD) approach that uses DFT-optimized glycosyl cation structures for parametrization. A simulated annealing procedure was used to capture structural snapshots, which are first optimized using molecular mechanics, and then reoptimized using DFT. It would be interesting to see how both theoretical methods compare against each other, and how efficiently both methods can screen the initial conformational space.

Also a direct comparison of both spectroscopic techniques would be highly interesting. It is possible that the same glycosyl cation yields fundamentally different vibrational spectra, when measured at room temperature using IRMPD, or at sub-Kelvin temperatures using helium nanodroplets. Elevated temperatures during the measurement likely lead to the thermal population of higher-energy conformers. This effect could also be simulated in the experimental setup described in this work, by operating the hexapole ion trap at elevated temperatures and kinetically trap the conformational distribution by rapid cooling of the helium droplets.

6.6 Conclusions

In summary, cryogenic vibrational spectroscopy is a valuable method to elucidate the gas-phase structures of glycosyl cations. The low-temperature environment of superfluid helium droplets allows to record highly resolved vibrational spectra that are used to elucidate fundamental structural motifs that enable stereoselective glycosylation reactions: neighboring group participation and remote participation.

Glycosyl cations generated from three custom-tailored building blocks equipped with an acetyl group at C2 readily form a covalent bond between the carbonyl oxygen of the acetyl group and the anomeric

carbon. These interactions were identified by the characteristic vibrations of the resulting dioxolenium motif, which enables the stereoselective formation of 1,2-*trans* glycosidic linkages. As a result of this covalent bond formation, the pyranose ring is heavily distorted and each glycosyl cation adopts a different ring pucker.

The combination of cryogenic vibrational spectroscopy and first-principles theory was then transferred from custom-tailored building blocks to galactose building blocks that are routinely used in glycosynthesis. The stereoselectivity obtained by these building blocks is believed to be a result of participation by remote protecting groups. The results provide direct evidence for the remote participation of C4-acetyl groups, promoting energetically preferred dioxolenium structures with a covalent bond between the carbonyl oxygen and the anomeric carbon. Participation of the C6-acetyl group, on the other hand, is energetically less favored. Instead, the formation of oxocarbenium or oxonium structures was observed. The oxonium structure is characterized by a remote participation of the C4-benzyl group, which results in a covalent bond between the O4-oxygen and the anomeric carbon. Observing this mode of remote participation is particularly surprising, because benzyl groups are generally considered as non-participating protecting groups. In case of the fully benzylated galactosyl cation, remote participation of the C6-benzyl group leads to the formation of oxonium ions, which are thermodynamically more stable than the corresponding oxocarbenium ions. These oxonium ions might be predominant structures during glycosylation reactions that follow an S_N1-pathway, leading to the selective formation of 1,2-*cis* glycosidic linkages. It has to be noted, however, that the conditions of glycosyl cations in the gas phase cannot be directly translated to the conditions in solution during glycosynthesis. The strength of intramolecular interactions in glycosyl cations is expected

to be slightly altered in the environment of solvents with (typically) low dielectric constants. Other factors such as solvent effects, or activators are not taken into account. The combination of vibrational spectroscopy and theory nevertheless provides a valuable tool to investigate the inherent structural properties of isolated glycosyl cations, which leads to a better mechanistic understanding of glycosylation reactions. This allows a more rational development of building blocks that relies on structural rather than anecdotal evidence.

Chapter 7

Summary and Future Perspectives

This thesis shows that cryogenic vibrational spectroscopy using helium nanodroplets is a valuable method for glycan analysis. In a systematic proof-of-principle study, the high resolving power was demonstrated using a variety of isomeric glycans ranging from mono- to tetrasaccharides. A set of six custom-tailored isomeric trisaccharides served as an ultimate benchmark for this method and revealed that even minute structural variations result in distinct spectral differences that provide a true spectral fingerprint unique to every isomer. The reliable identification of connectivity-, configurational, and composition isomers shows that cryogenic vibrational spectroscopy is a valuable addition to the structural analysis toolbox for glycans. However, the current void of a universal technique that allows the routine high-throughput analysis of glycans is not filled by cryogenic vibrational spectroscopy. Recent developments that combine cryogenic ion spectroscopy with ion-mobility mass spectrometry, on the other hand, are a promising

contender for such an analytical platform and could ultimately lead to the development of commercial instruments that utilize benchtop laser systems.

Apart from identifying isomeric glycans, this method can also be used to unravel the structure and conformation of glycan ions. The superfluid nature of helium nanodroplets provides an ideal nano-matrix that allows to study molecular ions near the absolute zero of temperature, and without any significant thermal contributions during the measurement. As a consequence, spectral line broadening is drastically reduced and highly resolved vibrational signatures can be obtained. After benchmarking this method with various isomeric glycans, it was used to shed light on an elusive rearrangement reaction observed during tandem-MS experiments of fucosylated glycans, called fucose migration. This migration reaction was previously associated with the fragmentation process and frequently leads to misleading fragment ions that can result in false sequence assignments. In this work, cryogenic vibrational spectroscopy reveals that fucose migration is not limited to fragment ions, and instead can be observed in intact protonated ions as well. The two protonated glycan ions Le^x and BG-H2 exhibit identical optical signatures, which is the result of a fucose migration reaction. Both glycan ions undergo a rearrangement reaction to the same chemical structure, which leads to the two identical infrared signatures. Observing fucose migration in intact protonated glycan ions generalizes this phenomenon to a universal issue in mass spectrometry, including ion-mobility spectrometry and ion spectroscopy. Further studies by Lettow et al. probed the influence of a mobile proton in fucose migration reactions and show that ammonium adducts of Le^x and BG-H2 undergo fucose migration because the proton is formally located at the glycan. The actual mechanism underlying fucose migration reactions as well

as the final structure after the rearrangement remain largely unknown. Extensive theoretical calculations and orthogonal structure sensitive methods could eventually unravel this elusive rearrangement reaction.

In another work, cryogenic vibrational spectroscopy was used to unravel the structure of glycosyl cations, the key intermediates during chemical glycosylation reactions. Because of their short-lived and reactive nature, these intermediates cannot be studied in the condensed phase and have remained elusive until recently. Here, glycosyl cations were isolated in the gas phase and analyzed using helium droplet isolation spectroscopy. The highly resolved vibrational spectra encode a wealth of structural information that was used to elucidate the structural motifs of glycosyl cations that enable stereoselective glycosynthesis: neighboring group participation and remote participation. Custom-tailored glycosyl donors carrying a C2-acetyl group consistently show characteristic vibrations that indicate the formation of dioxolenium ions, where the carbonyl oxygen of the acetyl group forms a covalent bond with the anomeric carbon. Extensive theoretical calculations using a DFT-based genetic algorithm screened the initial conformational space and generated structural candidates, which were later refined at a higher level of theory. The comparison of experimental spectra to calculated spectra confirms the presence of dioxolenyl motifs, which heavily influence the conformation of the pyranose rings. More importantly, the bicyclic structure shields the *cis*-face of the anomeric carbon and promotes the nucleophilic attack from the *trans*-face, which enables the stereoselective formation of 1,2-*trans* glycosidic linkages.

This approach was transferred from custom-tailored glycosyl donors to galactosyl building blocks that are commonly used in glycosynthesis. Structural analysis of the corresponding glycosyl cations reveals that remote participation of C4-acetyl groups leads to the energetically

preferred formation of dioxolenium ions. Analogously to dioxolenium ions involved in neighboring group participation, the carbonyl oxygen of C4-acetyl group forms a covalent bond with the anomeric carbon. In this case, the covalent bond leads to a bicyclic structure where the *trans*-face is shielded from a nucleophilic attack and enables the stereoselective formation of 1,2-*cis* glycosidic linkages. Remote participation of C6-acetyl groups, on the other hand, is energetically less preferred and competes with the formation of oxocarbenium and oxonium ions. The formation of oxonium ions is initiated by remote participation of benzyl protecting groups and was recently hypothesized in the literature. This is particularly surprising because benzyl groups were generally considered as non-participating protecting groups. This mode of remote participation is energetically preferred for the fully benzylated galactose building block, where the oxygen of the C6-benzyl ether forms a covalent bond with the anomeric carbon. Further studies on other commonly used building blocks such as glucosyl or mannosyl donors can lead to a more general mechanistic insight into glycosylation reactions. Understanding the structural preferences of individual building blocks can promote the rational design of novel, more stereoselective building blocks by fine-tuning specific protecting groups that steer the formation of certain structural motifs in the corresponding glycosyl cations.

Another part of this work focused on the development and implementation of a novel spectroscopic detection scheme. The original on-axis Wiley–McLaren time-of-flight detector was replaced by an off-axis design, where the release of ions from helium droplets and their detection is spatially separated. The newly implemented time-of-flight mass spectrometer allows to analyze released ions with a significantly improved mass resolution. This is especially useful for H/D-exchange experiments, the detection of interfering ions with similar masses, or

the analysis of ions with complex and overlapping isotope patterns. Another modification of the experimental setup was the design and implementation of a new variable-temperature ion trap. This ion trap allows to define the temperature of trapped ions continuously between 80–350 K. This can be used to either anneal trapped ions to cryogenic temperatures prior to pickup by helium droplets, or thermally populate higher energy conformers inside the trap. In the latter case, the conformational distribution can be kinetically trapped by the rapid cooling upon pickup by helium droplets. In future experiments, this can be used to measure relative populations of different conformers at various temperatures and determine their relative energies via van't Hoff plots.

Appendix A

Circuit Design

Richard Feynman once stated that "If you think you understand quantum mechanics, you don't understand quantum mechanics". Some people would probably argue that the same holds true for radio frequency (RF) electronics engineering. To understand this analogy, we first take a look at circuit design for static or low frequency potentials. Here, passive components such as resistors, capacitors or inductors are well behaved and can often be treated as ideal-components that have a resistance R, a capacitance C or an inductance L. In RF circuits, these devices are typically treated as non-ideal components whose behavior is frequency dependent. For example, a capacitor does not simply have *a* capacitance. Instead, a real-world capacitor is described by a primary capacitance that is in series with an equivalent resistance and an equivalent inductance, and in parallel with an equivalent resistance. Neglecting these additional terms can lead to rather strange behaviors: an inductor might behave 'normally' at one frequency, but suddenly behave more like a capacitor at higher frequencies. However, before going down the rabbit hole of frequency-dependent impedance, impedance matching,

Smith charts, parasitic capacitance, Q-factors or the skin-effect, the interested reader is referred to a dedicated text book.[266]

Fortunately, a degree in electronics engineering is not required to utilize software packages such as *LTSpice* to simulate RF circuits with real components and optimize their behavior on a trial and error basis. Using this heuristic approach, the circuitry driving the ion guide inside the spectroscopy detection region was developed. It consists of 54 stacked ring-electrodes that are sandwiched between two printed circuit boards (PCBs). The ring-electrodes carry a sinusoidal RF waveform (1 MHz) and DC offset that is defined by two potentials applied over a voltage divider network, as shown in Figure A.1. During operation, the DC offset voltage IG Front is switched between two states, as previously described in Figure 3.15. Because the rise-time of the DC voltages is critical for ion detection performance, the development target was to minimize the rise-time while maintaining efficient RF filtering to prevent the RF voltages from passing to the DC power supplies. To improve the switching behavior, ceramic coupling capacitors with a capacitance of 100 pF are used to transmit the RF components to the electrodes. The voltage divider network consists of resistors with gradually increasing resistance values (5, 10, 15, 20,. . . , 130) to generate the hyperbolic electrical field along the central axis of the ion guide. Small resistor values help to reduce the RC-time constant of the system. High-performance inductors with a resonance frequency of more than 4 MHz are used to transmit the DC potential and filter the RF waveform. An inductance value of 300 µH is sufficient to attenuate the high voltage RF waveforms.

A custom-built switch is used to quickly raise the DC offset voltage (Figure A.2). It consists of a p-channel MOSFET that features a low gate capacitance, a low on-resistance and a short rise-time. The

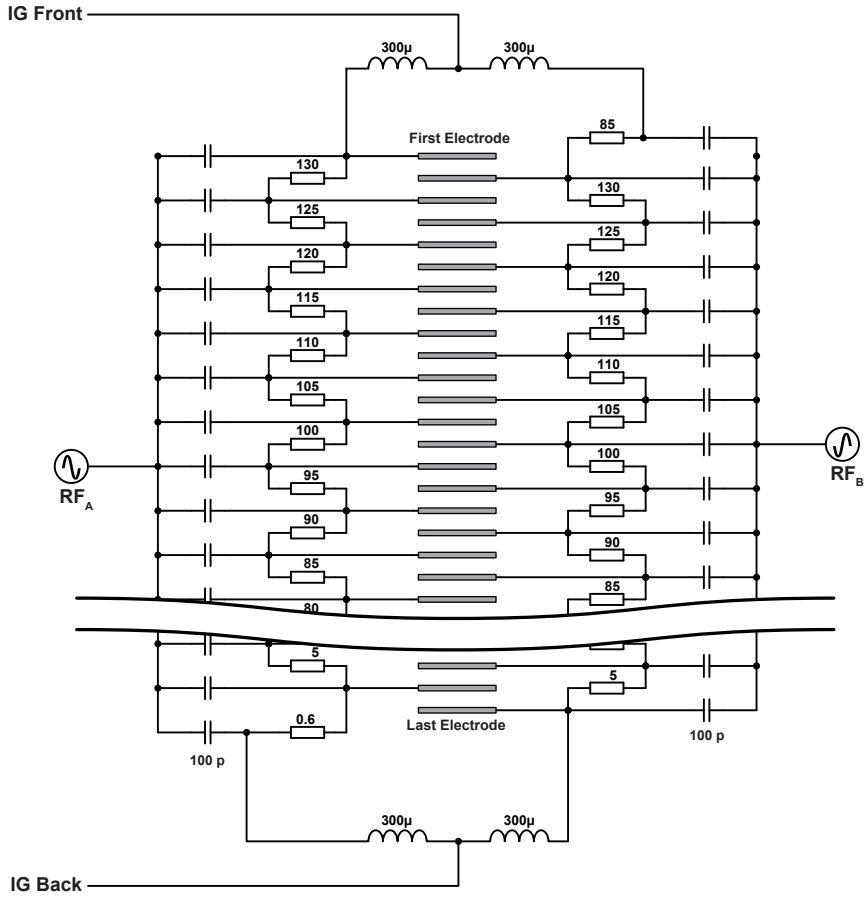


Figure A.1: Circuit design driving the ion guide used inside the spectroscopy detection region. It consists of a voltage divider network, coupling capacitors and high-performance inductors. The two DC offset voltages are controlled by the custom built switch shown in Figure A.2. The PCB was produced using vacuum-compatible RO4350B (Rogers Corporation, Chandler, AZ, USA), a proprietary composite of ceramic, PTFE and prefabulated amulite to reduce parasitic capacitance at high frequencies. Due to space restrictions, not all 54 electrodes are shown in the schematic.

gate of the MOSFET is driven by a high-voltage operational amplifier configured in a differential amplification circuit. The output of the operational amplifier has a relatively high slew-rate capable of driving the MOSFET's gate. A TTL-signal is used to trigger the operational amplifier. Two 2.2 kΩ resistors and a Zener diode are used for reverse-current protection of the DC power supplies.

The switching behavior of the custom-built switch and the ion guide are shown in Figure A.3. Shortly after the TTL signal triggers the operational amplifier, the MOSFET is enabled and raises the output voltage IG Front from +32 V to +37 V within 200 ns. The rise-time of the voltage applied to the first ion guide is around 1 µs.

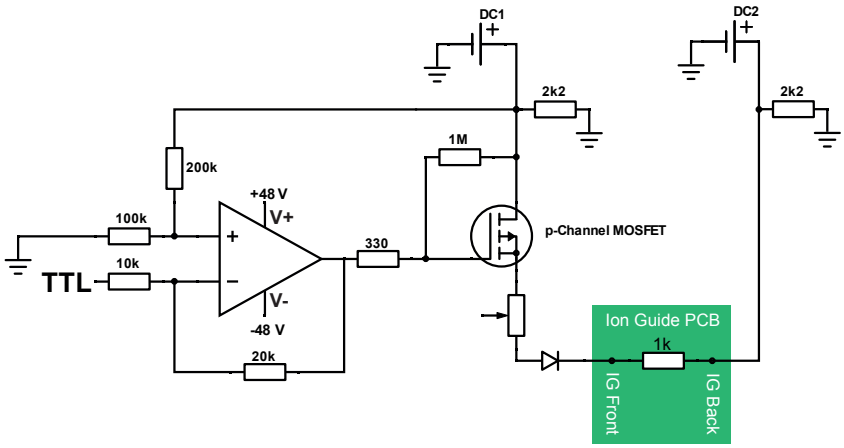


Figure A.2: A custom-built switch is used to quickly raise the DC offset voltage IG Front from 32 V to 37 V. It is triggered by a TTL signal controlling an operational amplifier, which drives the gate of a p-channel MOSFET.

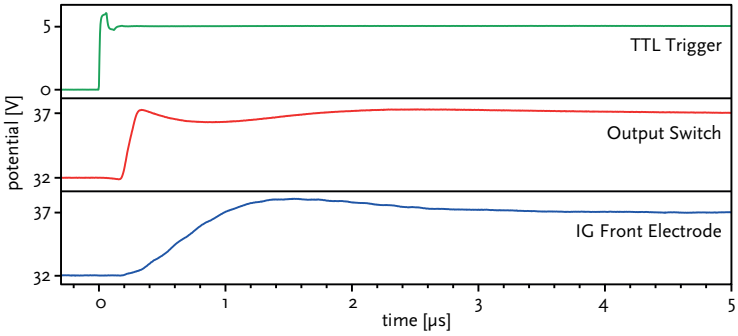


Figure A.3: Switching characteristics of the custom-built switch driving the ion guide. The TTL signal (green trace) triggers the output stage of the switch (red trace), which has a rise-time of around 200 ns. This leads to a rise-time around 1 µs measured at the first electrode of the ion guide (blue trace).

Appendix B

Lewis and Blood Group Antigens

This Appendix contains mass spectra as well as complete IR spectra of the Lewis and blood group antigens described in Chapter 5.

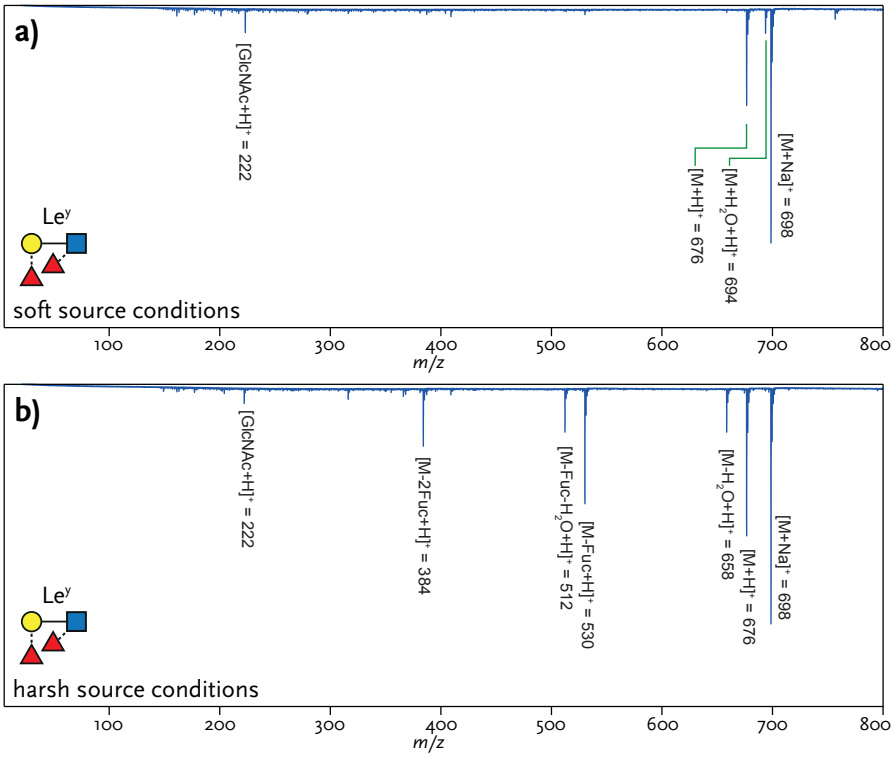


Figure B.1: Exemplary mass spectra for the Lewis y tetrasaccharide using soft (a) and harsh (b) source conditions. In-source fragmentation yields the $[M - Fuc + H]^+$ fragment. The same procedure was used for the Lewis b tetrasaccharide.

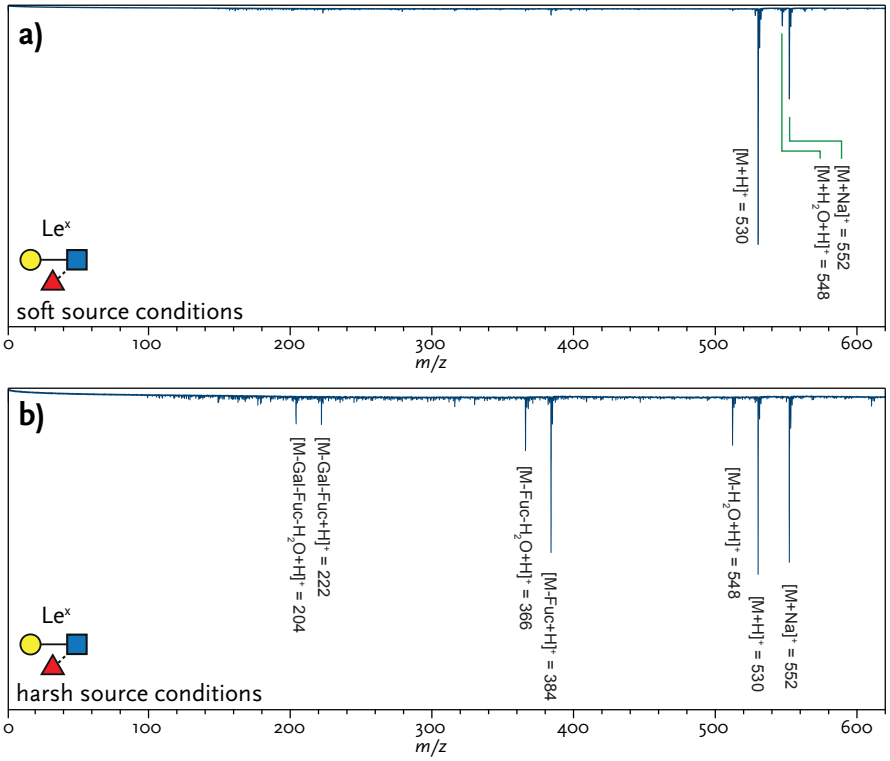


Figure B.2: Exemplary mass spectra for the Lewis y tetrasaccharide using soft (a) and harsh (b) source conditions. Under soft source conditions, only intact trisaccharide ions are observed as proton, water or sodium adducts. These conditions were used for to record IR spectra for all four trisaccharides.

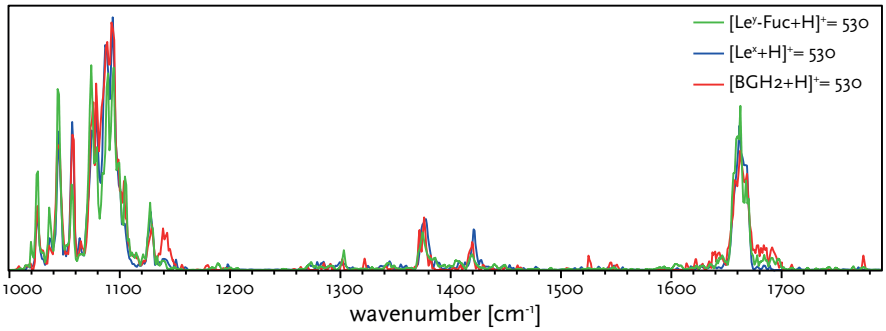


Figure B.3: Overlay of the three complete spectra obtained from the protonated Le^y fragment as well as both intact trisaccharides Le^x and BG-H2.

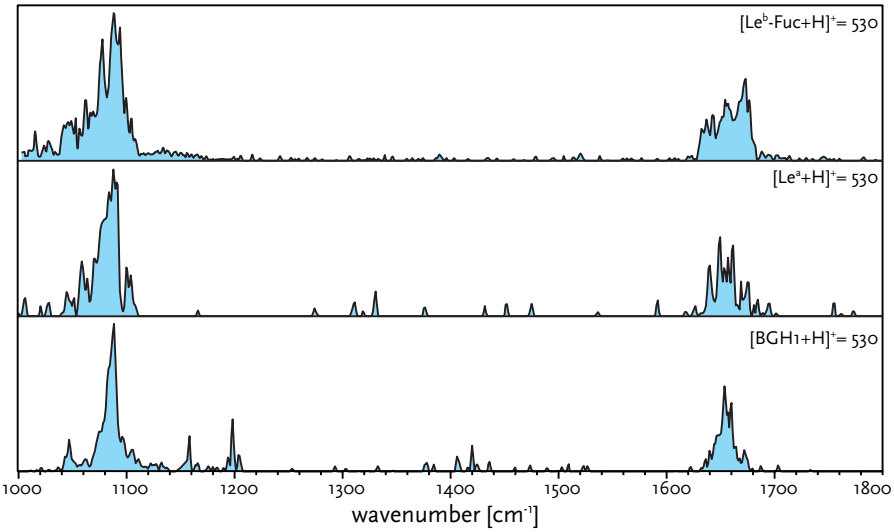


Figure B.4: Complete IR spectra of the Le^b fragment and the two intact trisaccharides Le^a and BG-H1 as protonated ions.

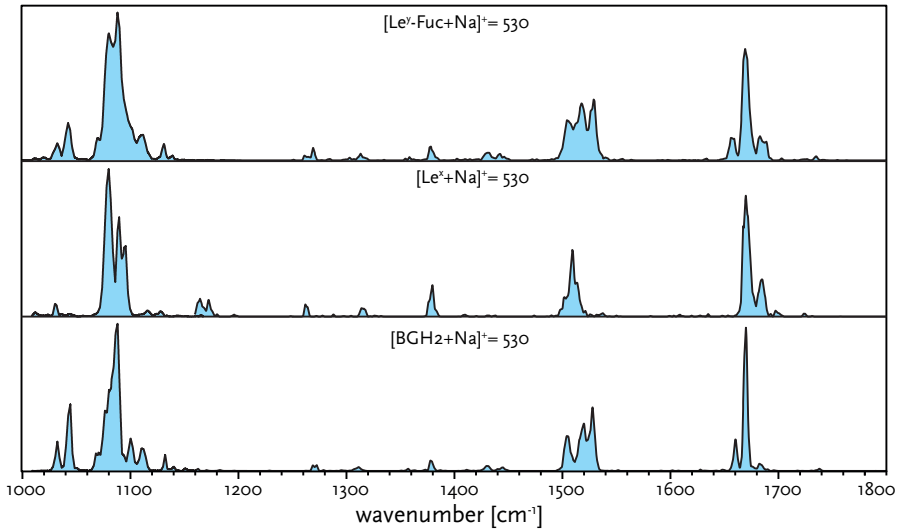


Figure B.5: Complete IR spectra of the Le^y fragment and the two intact trisaccharides Le^x and BG-H2 as sodium adducts.

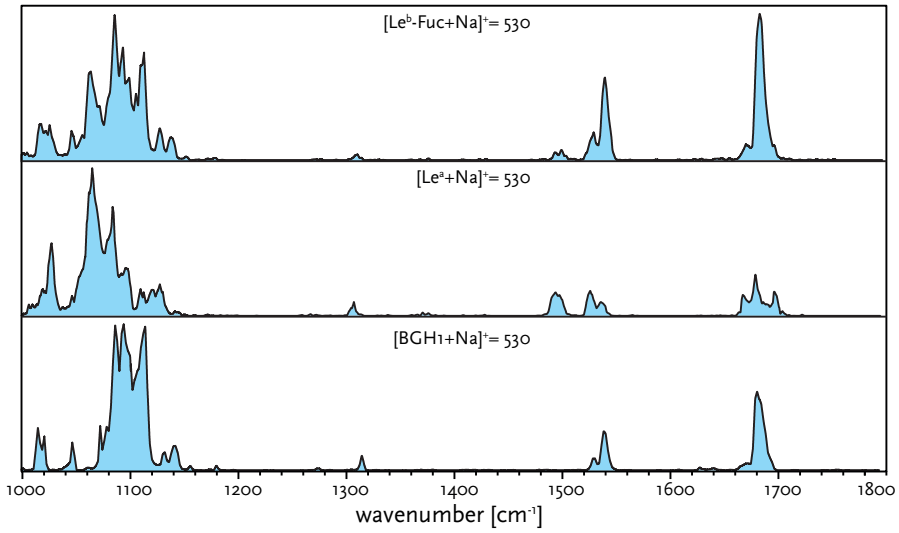


Figure B.6: Complete IR spectra of the Le^b fragment and the two intact trisaccharides Le^a and BG-H1 as sodium adducts.

Appendix C

Glycosyl Cations

This Appendix contains mass spectra, energy hierarchies as well as 3D structures of glycosyl cations described in Chapter 6. Quantum chemical calculations were performed by Mateusz Marianski (Hunter College, New York, NY, USA) and Kim Greis (Fritz Haber Institute, Berlin, Germany).

Glycosyl Cations Featuring Neighboring Group Participation

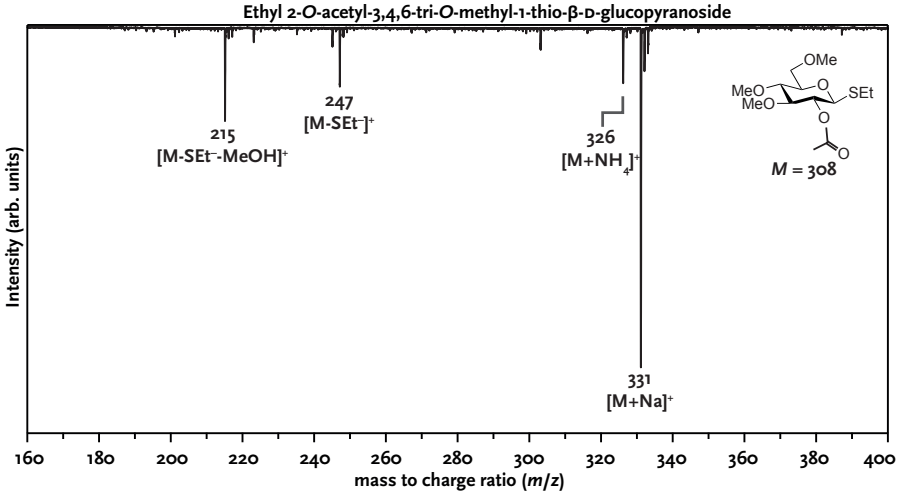


Figure C.1: Exemplary mass spectrum of the glucosyl building block carrying an acetyl group at C2. Glycosyl cations ($m/z = 247$) were generated by in-source fragmentation of the respective precursor ions.

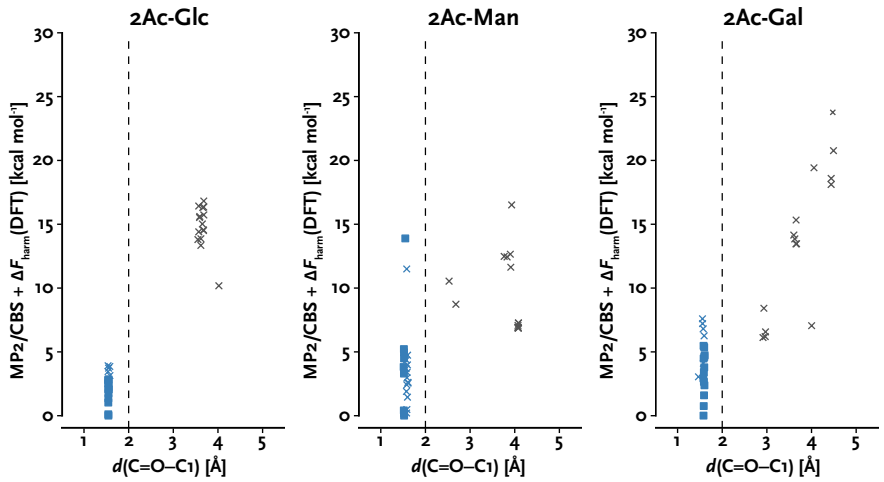


Figure C.2: Relative harmonic free-energies of glycosyl cations carrying acetyl groups at C2 as a function of the distance between the carbonyl oxygen of the acetyl group and the anomeric carbon. The glycosyl cations adopt dioxolenium (blue) and oxocarbenium (gray) structures. Geometries were obtained from a genetic algorithm and optimized at the PBE0+D3/6-311+G(d,p) level. The free-energy was calculated at the RI-MP2 level of theory, extrapolated to the complete basis set, and harmonic free-energies at the DFT level were added. Squares indicate structures that have the same ring pucker as the lowest-energy conformer.

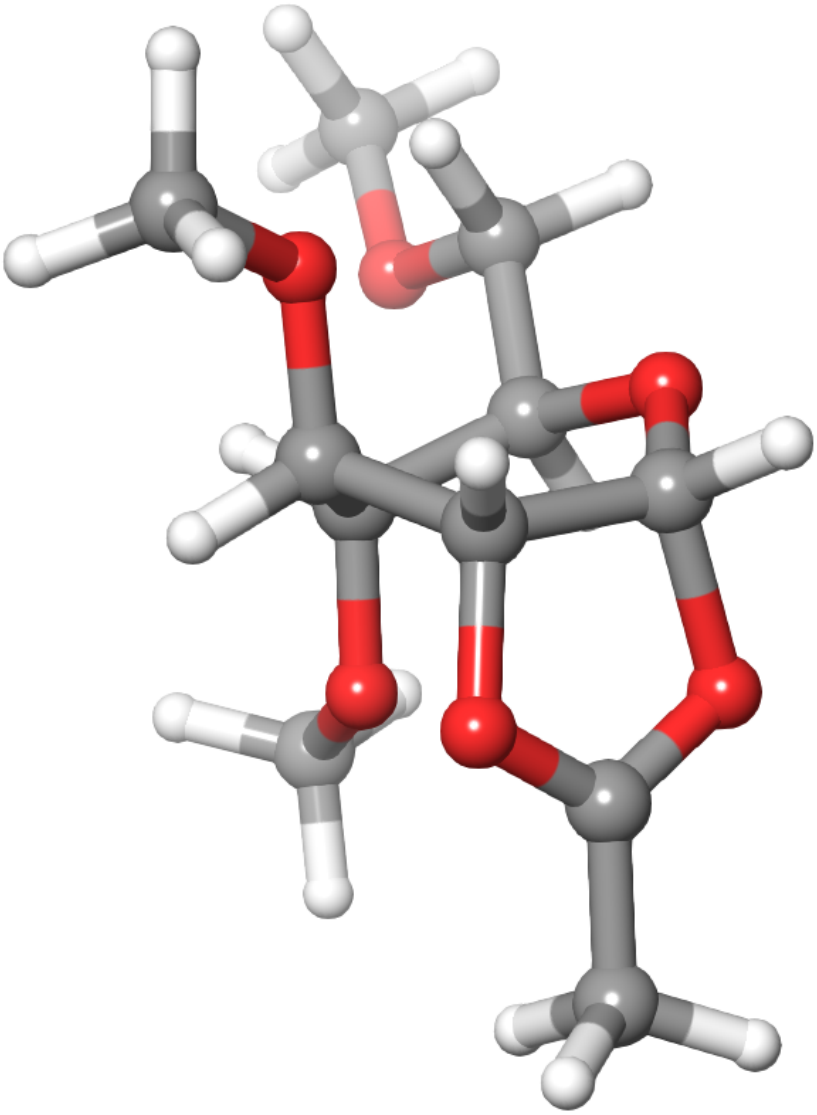


Figure C.3: The lowest-energy structure of the **2Ac-Glc** glycosyl cation adopts a $^{3}S_{1}$ ring pucker.

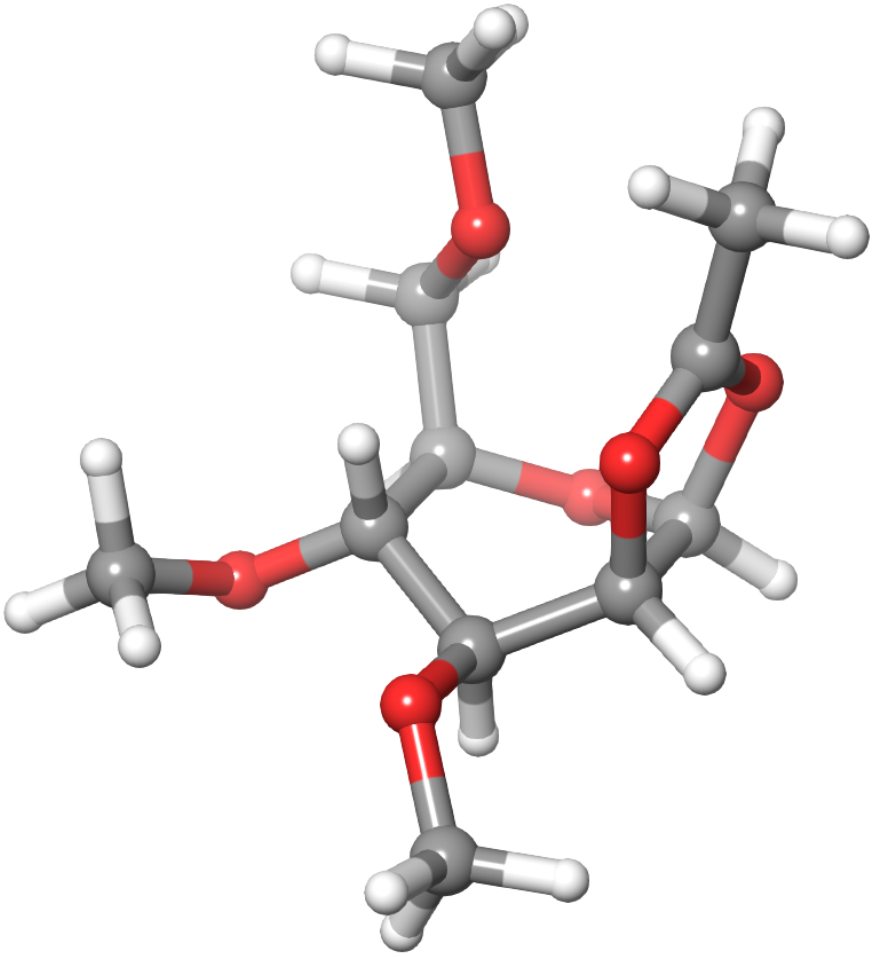


Figure C.4: The lowest-energy structure of the **2Ac-Man** glycosyl cation adopts a $B_{O,3}$ ring pucker.

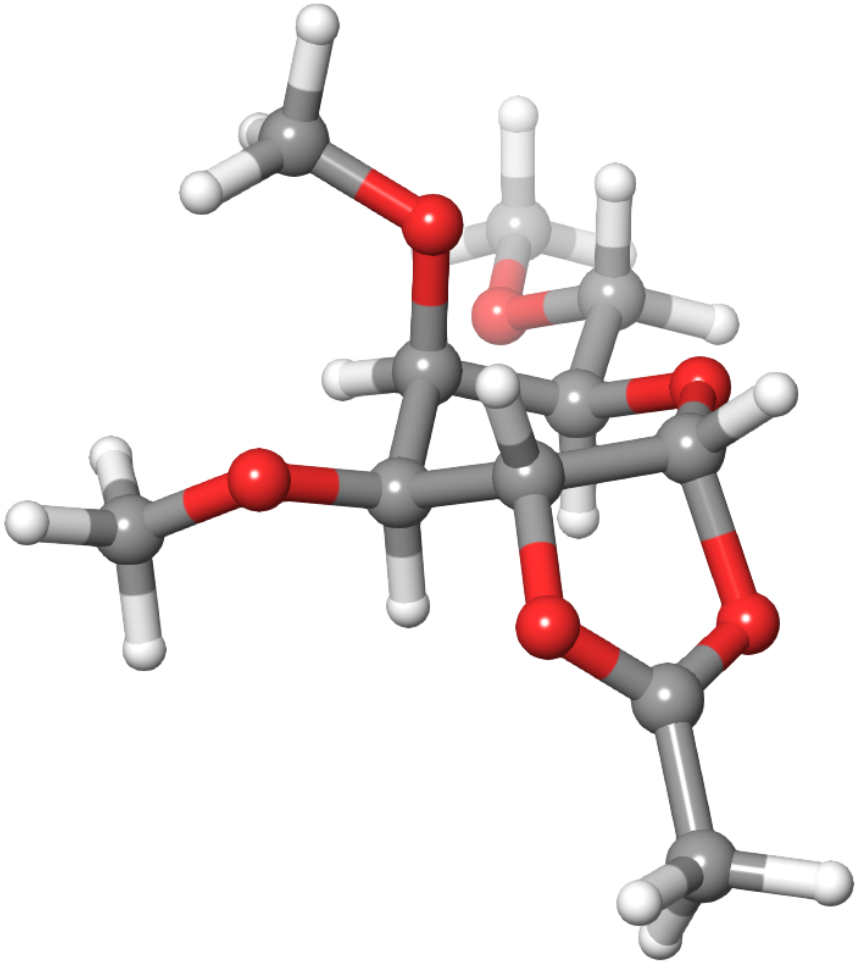


Figure C.5: The lowest-energy structure of the **2Ac-Gal** glycosyl cation adopts a ^4E ring pucker.

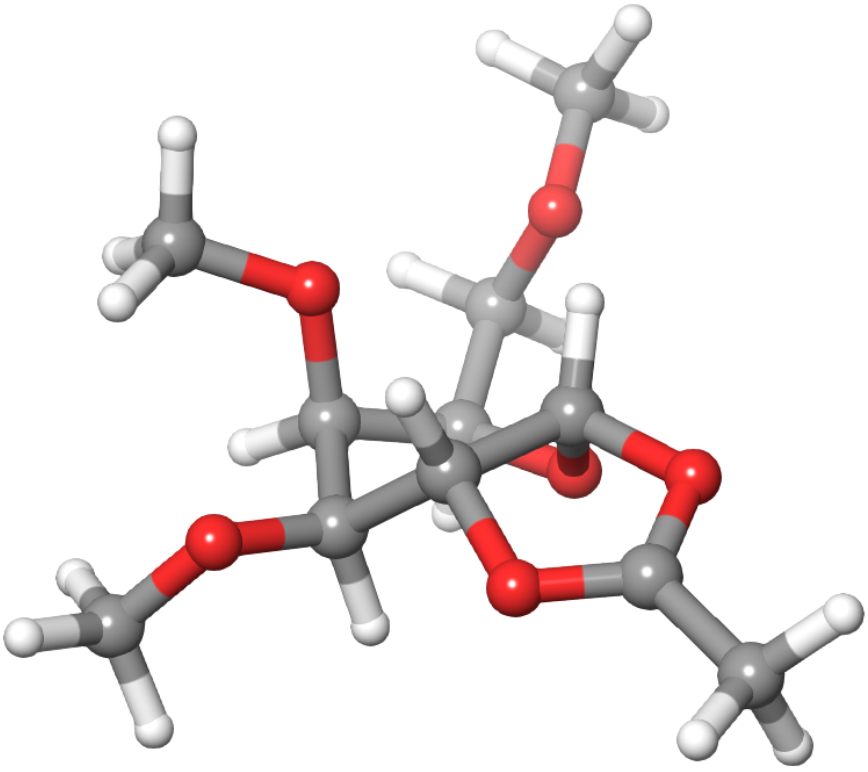


Figure C.6: Another low-energy structure of the **2Ac-Gal** glycosyl cation adopts a $^{1}S_{3}$ ring pucker and likely coexists during the measurement.

Glycosyl Cations Featuring Remote Participation

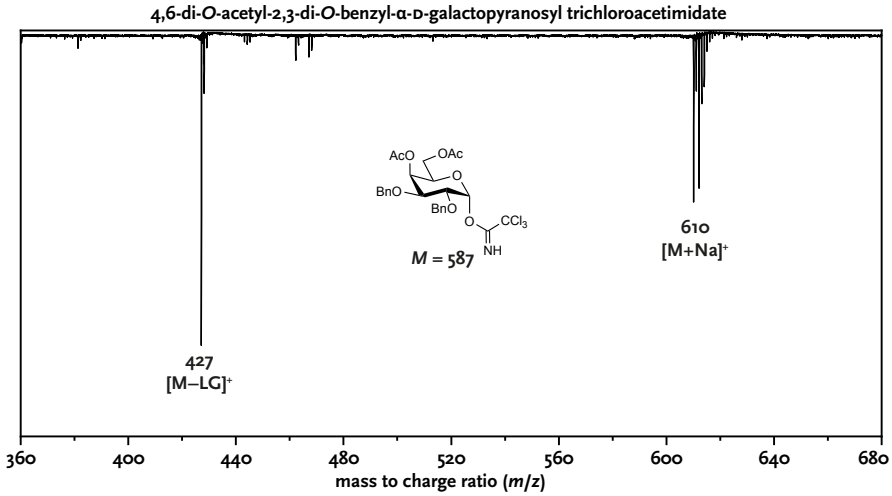


Figure C.7: Exemplary mass spectrum of the **4,6Ac** building block. Glycosyl cations ($m/z = 427$) were generated by in-source fragmentation of the respective precursor ions.

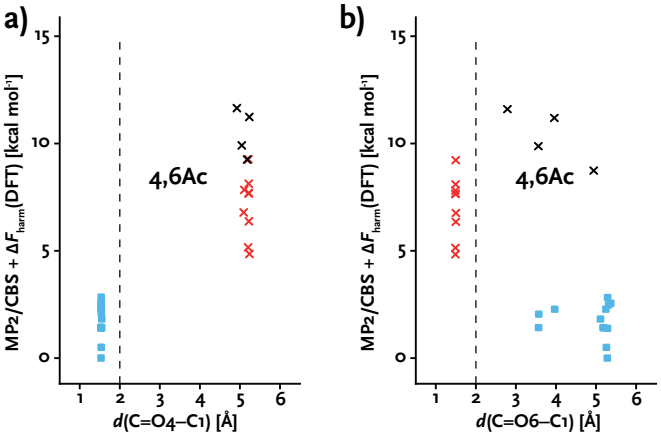


Figure C.8: Relative harmonic free-energies of glycosyl cations generated from the **4,6Ac** building block as a function of the distance between the oxygen of the C4-acetyl group (a) or the C6-acetyl group (b) and the anomeric carbon. The glycosyl cations adopt dioxolenium structures via the C4-acetyl group (blue markers) and the C6-acetyl group (red markers), or oxocarbenium structures (black markers). Square markers are used to indicate structures that have the same ring pucker as the lowest-energy structure. Geometries were obtained from a genetic algorithm and optimized at the PBE0+D3/6-311+G(d,p) level. The free-energy was calculated at the RI-MP2 level of theory, extrapolated to the complete basis set, and harmonic free-energies at the DFT level were added.

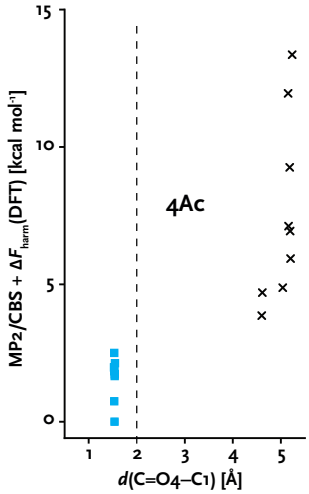


Figure C.9: Relative harmonic free-energies of glycosyl cations generated from the **4Ac** building block as a function of the distance between the oxygen of the C4-acetyl group and the anomeric carbon. The glycosyl cations adopt dioxolenium structures via the C4-acetyl group (blue markers), or oxocarbenium structures (black markers). Square markers are used to indicate structures that have the same ring pucker as the lowest-energy structure. Geometries were obtained from a genetic algorithm and optimized at the PBE0+D3/6-311+G(d,p) level. The free-energy was calculated at the RI-MP2 level of theory, extrapolated to the complete basis set, and harmonic free-energies at the DFT level were added.

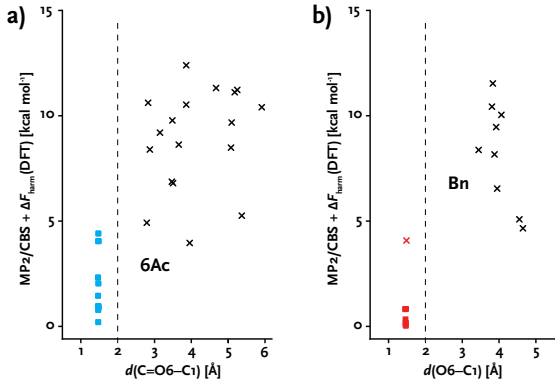


Figure C.10: Relative harmonic free-energies of glycosyl cations generated from the **6Ac** (a) and **Bn** (b) building block as a function of the distance between the oxygen of the C6-acetyl group (a) or the C6-oxygen (b) and the anomeric carbon. The **6Ac** glycosyl cations adopt dioxolenium structures via the C4-acetyl group (blue markers), or oxocarbenium structures (black markers). The **Bn** glycosyl cations adopt oxonium structures via O6 (red markers), oxocarbenium structures (black markers). Square markers are used to indicate structures that have the same ring pucker as the lowest-energy structure. Geometries were obtained from a genetic algorithm and optimized at the PBE0+D3/6-311+G(d,p) level. The free-energy was calculated at the RI-MP2 level of theory, extrapolated to the complete basis set, and harmonic free-energies at the DFT level were added.

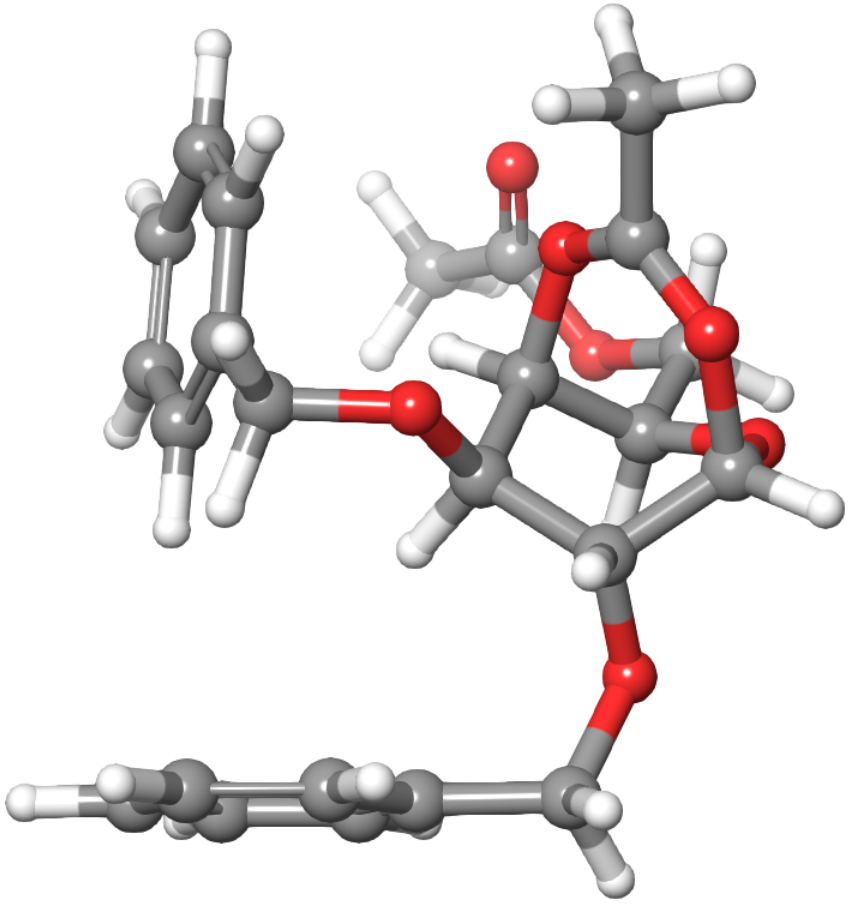


Figure C.11: Calculated structure of glycosyl cation **A** generated from the **4,6Ac** building block.

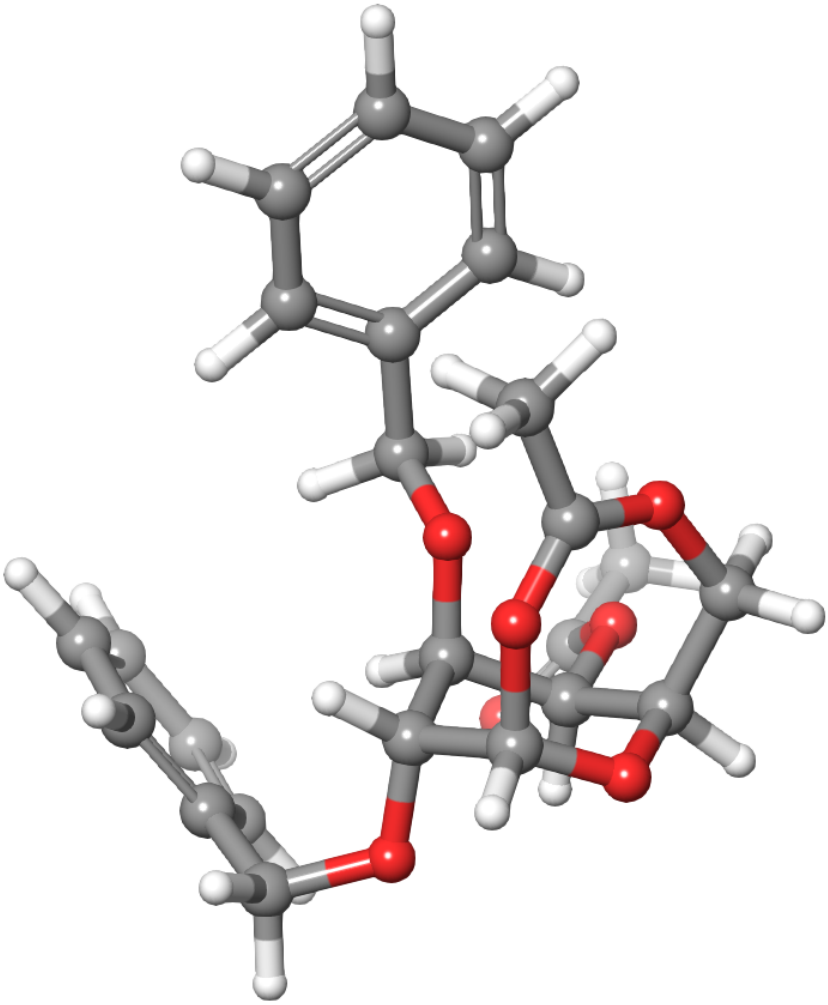


Figure C.12: Calculated structure of glycosyl cation **B** generated from the **4,6Ac** building block.

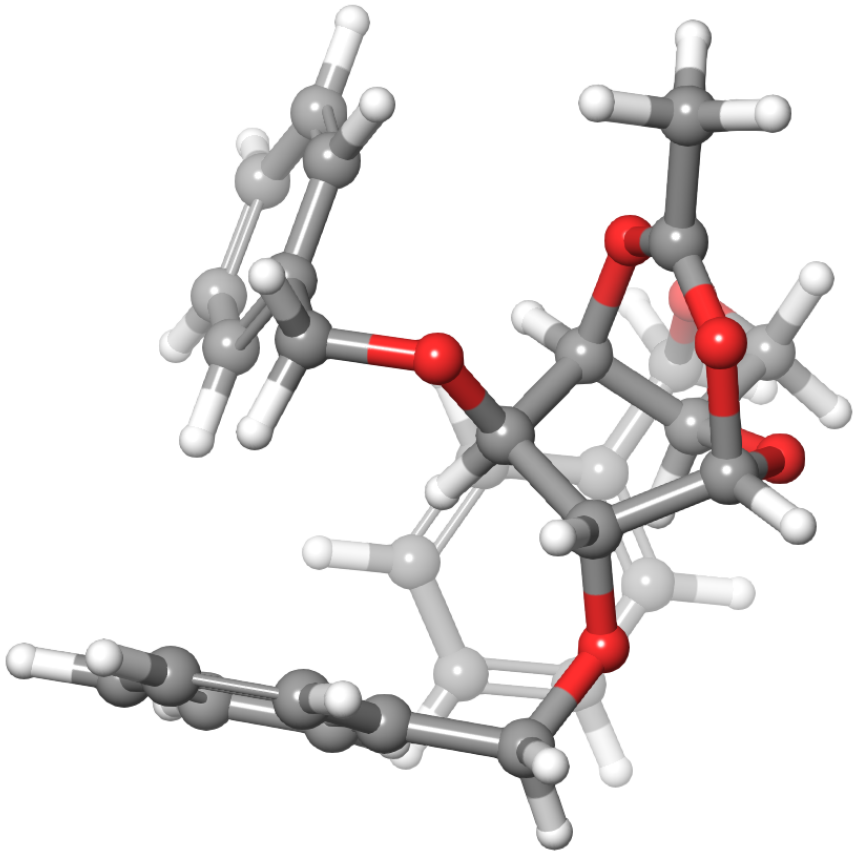


Figure C.13: Calculated structure of glycosyl cation **C** generated from the **4Ac** building block.

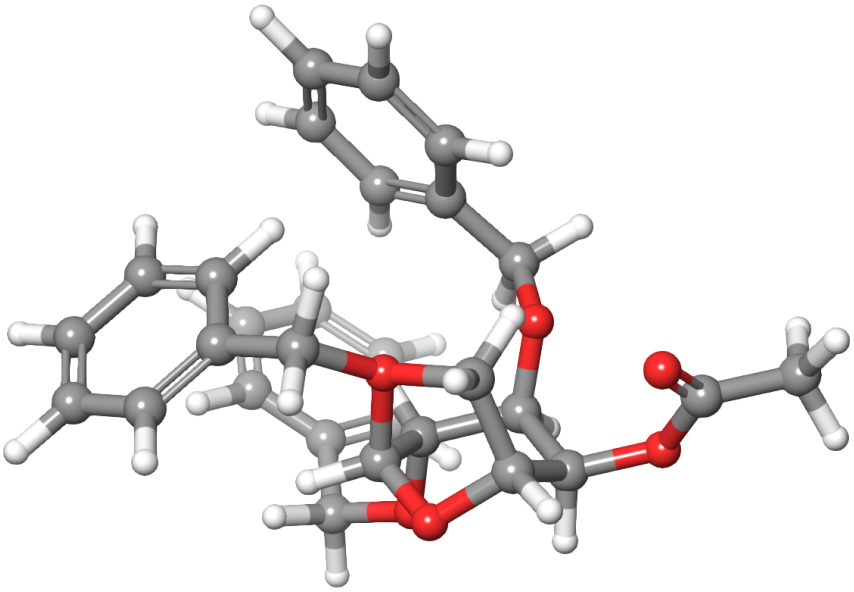


Figure C.14: Calculated structure of glycosyl cation **D** generated from the **4Ac** building block.

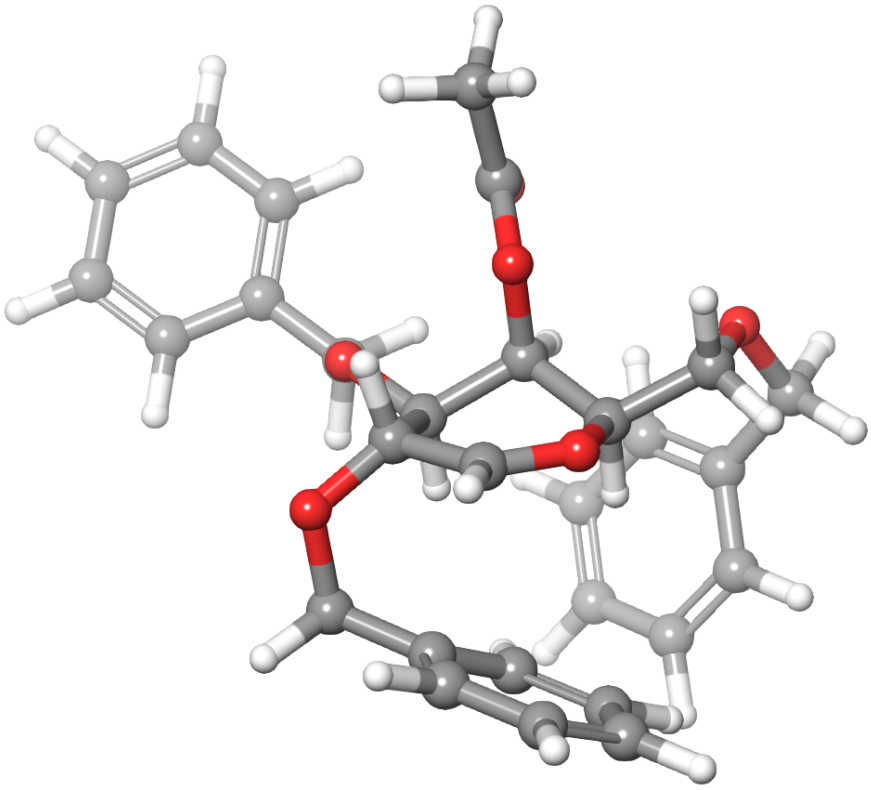


Figure C.15: Calculated structure of glycosyl cation **E** generated from the **4Ac** building block.

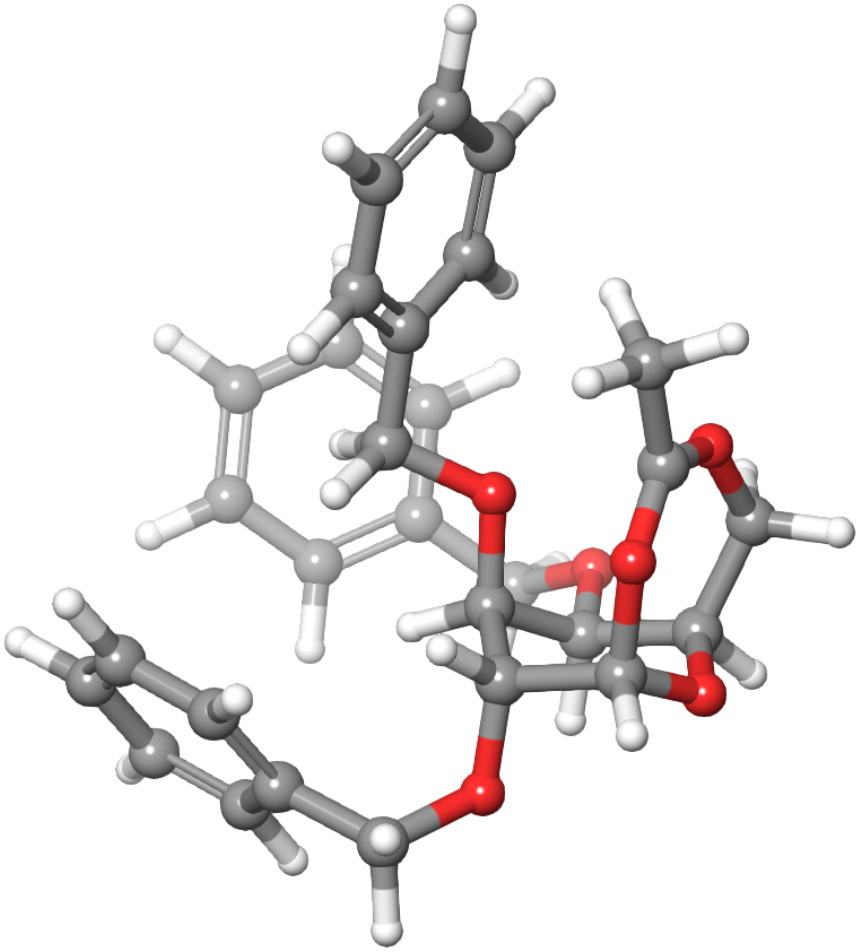


Figure C.16: Calculated structure of glycosyl cation **F** generated from the **6Ac** building block.

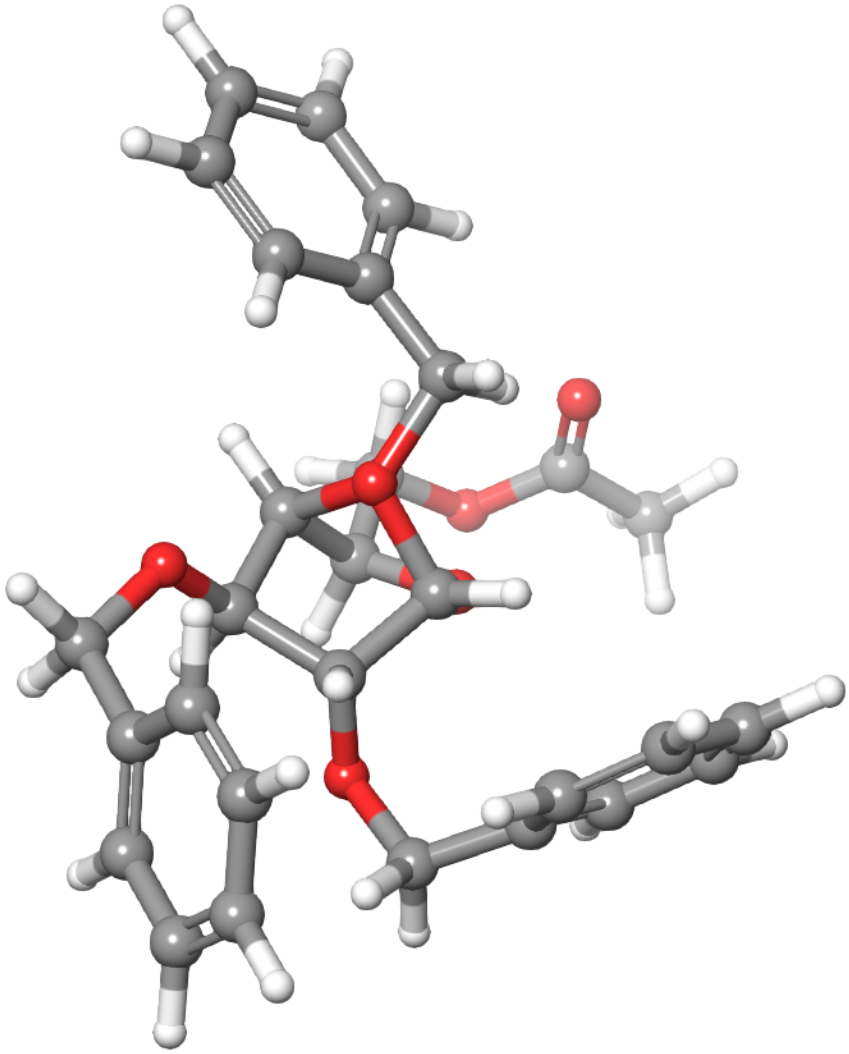


Figure C.17: Calculated structure of glycosyl cation **G** generated from the **6Ac** building block.

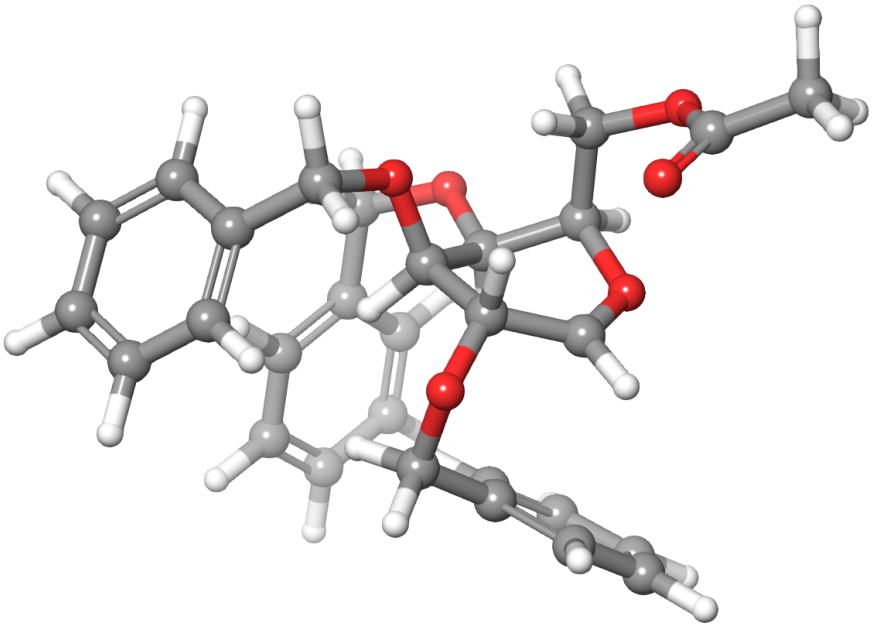


Figure C.18: Calculated structure of glycosyl cation **H** generated from the **6Ac** building block.

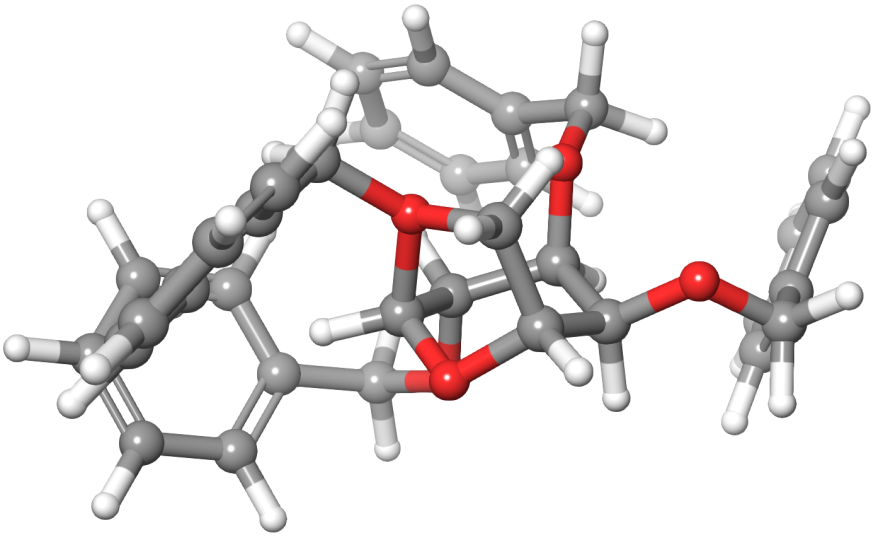


Figure C.19: Calculated structure of glycosyl cation **I** generated from the **Bn** building block.

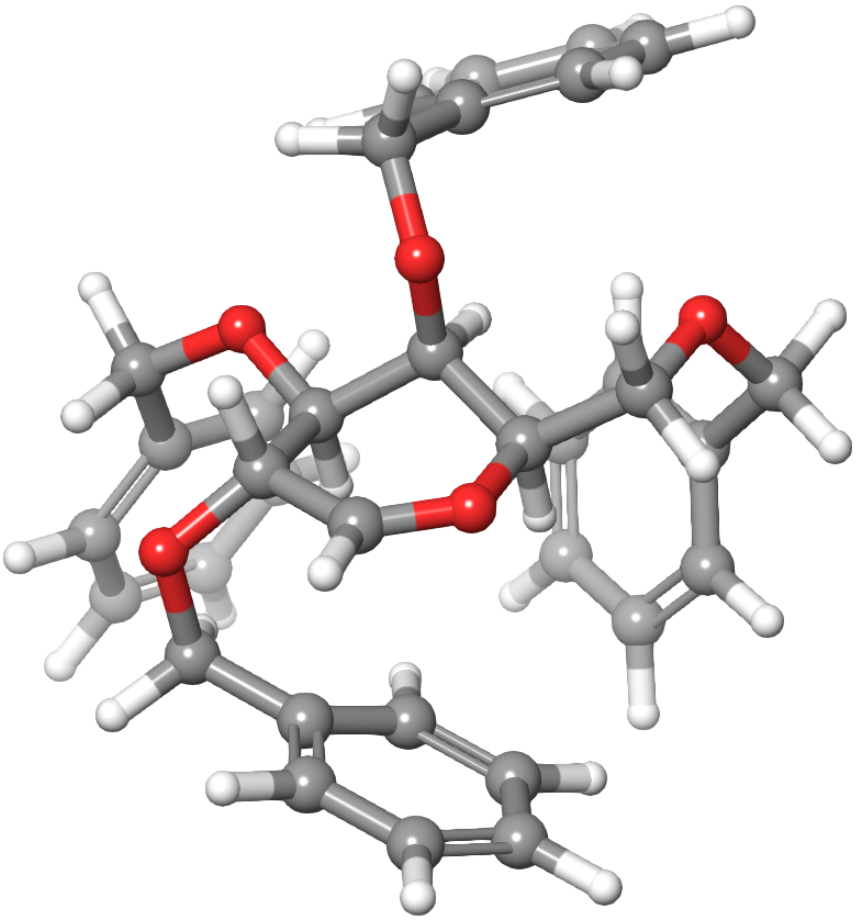


Figure C.20: Calculated structure of glycosyl cation **J** generated from the **Bn** building block.

Bibliography

[1] E. Mucha, A. Stuckmann, M. Marianski, W. B. Struwe, G. Meijer, K. Pagel, In-depth structural analysis of glycans in the gas phase, *Chem. Sci.* **2019**, *10*, 1272–1284.

[2] J. M. Berg, J. L. Tymoczko, L. Stryer, *Biochemistry*, W.H. Freeman, New York, 6th ed., **2007**.

[3] J. d. Watson, F. H. C. Crick, Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid, *Nature* **1953**, *171*, 737–738.

[4] J. C. Venter *et al.*, The Sequence of the Human Genome, *Science* **2001**, *291*, 1304–1351.

[5] J. Lucas-Lenard, Protein biosynthesis, *Annu. Rev. Biochem.* **1971**, *40*, 409–448.

[6] A. Varki, Biological roles of oligosaccharides: All of the theories are correct, *Glycobiology* **1993**, *3*, 97–130.

[7] A. Shevchenko, M. Wilm, O. Vorm, M. Mann, Mass Spectrometric Sequencing of Proteins from Silver-Stained Polyacrylamide Gels, *Anal. Chem.* **1996**, *68*, 850–858.

[8] E. S. Lander *et al.*, Initial sequencing and analysis of the human genome, *Nature* **2001**, *409*, 860–921.

[9] J. Ø. Duus, C. H. Gotfredsen, K. Bock, Carbohydrate Structural Determination by NMR Spectroscopy: Modern Methods and Limitations, *Chem. Rev.* **2000**, *100*, 4589–4614.

[10] M. D. Battistel, H. F. Azurmendi, B. Yu, D. I. Freedberg, NMR of glycans: shedding new light on old problems, *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *79*, 48–68.

[11] Y. Mechref, M. V. Novotny, C. Krishnan, Structural Characterization of Oligosaccharides Using Maldi-TOF/TOF Tandem Mass Spectrometry, *Anal. Chem.* **2003**, *75*, 4895–4903.

[12] M. Wuhrer, Glycomics using mass spectrometry, *Glycoconjugate J.* **2013**, *30*, 11–22.

[13] D. Reusch *et al.*, Comparison of methods for the analysis of therapeutic immunoglobulin G Fc-glycosylation profiles–part 1: separation-based methods, *mAbs* **2015**, *7*, 167–179.

[14] G. S. M. Lageveen-Kammeijer, N. de Haan, P. Mohaupt, S. Wagt, M. Filius, J. Nouta, D. Falck, M. Wuhrer, Highly sensitive CE-ESI-MS analysis of N-glycans from complex biological samples, *Nat. Commun.* **2019**, *10*, 2137.

[15] M. D. Plasencia, D. Isailovic, S. I. Merenbloom, Y. Mechref, M. V. Novotny, D. E. Clemmer, Resolving and assigning N-linked glycan structural isomers from ovalbumin by IMS-MS, *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1706–1715.

[16] J. Hofmann, H. S. Hahm, P. H. Seeberger, K. Pagel, Identification of carbohydrate anomers using ion mobility-mass spectrometry, *Nature* **2015**, *526*, 241–244.

[17] H. Hinneburg *et al.*, Distinguishing N-acetylneuraminic acid linkage isomers on glycopeptides by ion mobility-mass spectrometry, *Chem. Commun.* **2016**, *52*, 4381–4384.

[18] J. Hofmann, A. Stuckmann, M. Crispin, D. J. Harvey, K. Pagel, W. B. Struwe, Identification of Lewis and Blood Group Carbohydrate Epitopes by Ion Mobility-Tandem-Mass Spectrometry Fingerprinting, *Anal. Chem.* **2017**, *89*, 2318–2325.

[19] L. Royle, C. M. Radcliffe, R. A. Dwek, P. M. Rudd, Detailed structural analysis of N-glycans released from glycoproteins in SDS-PAGE gel bands using HPLC combined with exoglycosidase array digestions, *Methods Mol. Biol.* **2006**, *347*, 125–143.

[20] L. Royle *et al.*, HPLC-based analysis of serum N-glycans on a 96-well plate platform with dedicated database software, *Anal. Biochem.* **2008**, *376*, 1–12.

[21] J. Oomens, B. G. Sartakov, G. Meijer, G. von Helden, Gas-phase infrared multiple photon dissociation spectroscopy of mass-selected molecular ions, *Int. J. Mass Spectrom.* **2006**, *254*, 1–19.

[22] A. M. Rijs, J. Oomens, *Gas-Phase IR Spectroscopy and Structure of Biological Molecules*, *Vol. 364*, Springer International Publishing, Cham, **2015**.

[23] F. O. Talbot, J. P. Simons, Sugars in the gas phase: The spectroscopy and structure of jet-cooled phenyl β-D-glucopyranoside, *Phys. Chem. Chem. Phys.* **2002**, *4*, 3562–3565.

[24] R. A. Jockusch *et al.*, Probing the glycosidic linkage: UV and IR ion-dip spectroscopy of a lactoside, *J. Am. Chem. Soc.* **2004**, *126*, 5709–5714.

[25] P. Carçabal *et al.*, Hydrogen bonding and cooperativity in isolated and hydrated sugars: mannose, galactose, glucose, and lactose, *J. Am. Chem. Soc.* **2005**, *127*, 11414–11425.

[26] B. Brauer, M. Pincu, V. Buch, I. Bar, J. P. Simons, R. B. Gerber, Vibrational spectra of alpha-glucose, beta-glucose, and sucrose: anharmonic calculations and experiment, *J. Phys. Chem. A* **2011**, *115*, 5859–5872.

[27] J. Fenn, M. Mann, C. Meng, S. Wong, C. Whitehouse, Electrospray ionization for mass spectrometry of large biomolecules, *Science* **1989**, *246*, 64–71.

[28] J. Seo, S. Warnke, K. Pagel, M. T. Bowers, G. von Helden, Infrared spectrum and structure of the homochiral serine octamer-dichloride complex, *Nat. Chem.* **2017**, *9*, 1263–1268.

[29] V. Scutelnic *et al.*, The Structure of the Protonated Serine Octamer, *J. Am. Chem. Soc.* **2018**, *140*, 7554–7560.

[30] W. Hoffmann *et al.*, An Intrinsic Hydrophobicity Scale for Amino Acids and Its Application to Fluorinated Compounds, *Angew. Chem. Int. Ed.* **2019**, *58*, 8216–8220.

[31] J. Seo *et al.*, An infrared spectroscopy approach to follow β-sheet formation in peptide amyloid assemblies, *Nat. Chem.* **2017**, *9*, 39–44.

[32] W. Hoffmann *et al.*, NFGAIL Amyloid Oligomers: The Onset of Beta-Sheet Formation and the Mechanism for Fibril Formation, *J. Am. Chem. Soc.* **2018**, *140*, 244–249.

[33] J. Seo, W. Hoffmann, S. Warnke, M. T. Bowers, K. Pagel, G. von Helden, Retention of Native Protein Structures in the Absence of Solvent: A Coupled Ion Mobility and Spectroscopic Study, *Angew. Chem. Int. Ed.* **2016**, *55*, 14173–14176.

[34] S. Warnke, W. Hoffmann, J. Seo, E. de Genst, G. von Helden, K. Pagel, 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*, 638–646.

[35] J. Martens, G. Berden, R. E. van Outersterp, L. A. J. Kluijtmans, U. F. Engelke, C. D. M. van Karnebeek, R. A. Wevers, J. Oomens, Molecular identification in metabolomics using infrared ion spectroscopy, *Sci. Rep.* **2017**, *7*, 3363.

[36] J. Martens, V. Koppen, G. Berden, F. Cuyckens, J. Oomens, Combined Liquid Chromatography-Infrared Ion Spectroscopy for Identification of Regioisomeric Drug Metabolites, *Anal. Chem.* **2017**, *89*, 4359–4362.

[37] J. Martens *et al.*, Unraveling the unknown areas of the human metabolome: the role of infrared ion spectroscopy, *Journal Inh. Metab. Dis.* **2018**, *41*, 367–377.

[38] R. E. van Outersterp *et al.*, Reference-standard free metabolite identification using infrared ion spectroscopy, *Int. J. Mass Spectrom.* **2019**, *443*, 77–85.

[39] N. C. Polfer, J. J. Valle, D. T. Moore, J. Oomens, J. R. Eyler, B. Bendiak, Differentiation of isomers by wavelength-tunable infrared multiple-photon dissociation-mass spectrometry: application to glucose-containing disaccharides, *Anal. Chem.* **2006**, *78*, 670–679.

[40] Y. Tan, N. C. Polfer, Linkage and anomeric differentiation in trisaccharides by sequential fragmentation and variable wavelength infrared photodissociation, *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 359–368.

[41] B. Schindler, G. Renois-Predelus, N. Bagdadi, S. Melizi, L. Barnes, S. Chambert, A.-R. Allouche, I. Compagnon, MS/IR, a new MS-based hyphenated method for analysis of hexuronic acid epimers in glycosaminoglycans, *Glycoconj. J.* **2017**, *34*, 421–425.

[42] S. Rudić, H.-b. Xie, R. B. Gerber, J. P. Simons, Protonated sugars: Vibrational spectroscopy and conformational structure of protonated O-methyl-α-D-galactopyranoside, *Mol. Phys.* **2012**, *110*, 1609–1615.

[43] L. Barnes, B. Schindler, S. Chambert, A.-R. Allouche, I. Compagnon, 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.

[44] B. Schindler, L. Legentil, A.-R. Allouche, V. Ferrières, I. Compagnon, Spectroscopic diagnostic for the ring-size of carbohydrates in the gas phase: furanose and pyranose forms of GalNAc, *Phys. Chem. Chem. Phys.* **2019**, *21*, 12460–12467.

[45] B. Schindler *et al.*, IRMPD Spectroscopy Sheds New (Infrared) Light on the Sulfate Pattern of Carbohydrates, *J. Phys. Chem. A* **2017**, *121*, 2114–2120.

[46] G. Renois-Predelus, B. Schindler, I. Compagnon, Analysis of Sulfate Patterns in Glycosaminoglycan Oligosaccharides by MSn Coupled to Infrared Ion Spectroscopy: the Case of GalNAc4S and GalNAc6S, *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 1242–1249.

[47] P. von Ragué Schleyer, A. J. Kos, The importance of negative (anionic) hyperconjugation, *Tetrahedron* **1983**, *39*, 1141–1150.

[48] E. Juaristi, G. Cuevas, *The anomeric effect*, CRC Press, Boca Raton and London, **1995**.

[49] D. Cremer, J. A. Pople, General definition of ring puckering coordinates, *J. Am. Chem. Soc.* **1975**, *97*, 1354–1358.

[50] A. Varki *et al.*, Symbol Nomenclature for Graphical Representations of Glycans, *Glycobiology* **2015**, *25*, 1323–1324.

[51] National Research Council (U.S.) Board on Chemical Sciences and Technology, *Transforming glycoscience: A roadmap for the future*, National Academies Press, Washington, D.C., **2012**.

[52] R. Apweiler, On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database, *Biochim. Biophys. Acta Gen. Subj.* **1999**, *1473*, 4–8.

[53] S. Reitsma, D. W. Slaaf, H. Vink, M. A. M. J. van Zandvoort, M. G. A. oude Egbrink, The endothelial glycocalyx: composition, functions, and visualization, *Pflugers Arch.* **2007**, *454*, 345–359.

[54] M. Molinari, N-glycan structure dictates extension of protein folding or onset of disposal, *Nat. Chem. Biol.* **2007**, *3*, 313–320.

[55] D. N. Hebert, M. Molinari, In and out of the ER: protein folding, quality control, degradation, and related human diseases, *Physiol. Rev.* **2007**, *87*, 1377–1408.

[56] D. J. Vigerust, V. L. Shepherd, Virus glycosylation: role in virulence and immune interactions, *Trends Microbiol.* **2007**, *15*, 211–218.

[57] M. A. Wolfert, G.-J. Boons, Adaptive immune activation: glycosylation does matter, *Nat. Chem. Biol.* **2013**, *9*, 776–784.

[58] S. S. L. Yap, T. Nguyen-Khuong, P. M. Rudd, S. Alonso, Dengue Virus Glycosylation: What Do We Know?, *Front. Microbiol.* **2017**, *8*, 1415.

[59] K. J. Doores, C. Bonomelli, D. J. Harvey, S. Vasiljevic, R. A. Dwek, D. R. Burton, M. Crispin, C. N. Scanlan, Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens, *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 13800–13805.

[60] K. J. Doores, The HIV glycan shield as a target for broadly neutralizing antibodies, *FEBS J.* **2015**, *282*, 4679–4691.

[61] A.-J. Behrens, M. Crispin, Structural principles controlling HIV envelope glycosylation, *Curr. Opin. Struct. Biol.* **2017**, *44*, 125–133.

[62] D. E. Michele *et al.*, Post-translational disruption of dystro-glycan-ligand interactions in congenital muscular dystrophies, *Nature* **2002**, *418*, 417–422.

[63] P. K. Grewal, J. E. Hewitt, Glycosylation defects: a new mechanism for muscular dystrophy?, *Human Mol. Gen.* **2003**, *12*, R259–64.

[64] J. Gu, N. Taniguchi, Potential of N-glycan in cell adhesion and migration as either a positive or negative regulator, *Cell Adh. Migr.* **2008**, *2*, 243–245.

[65] Y.-Y. Zhao, M. Takahashi, J.-G. Gu, E. Miyoshi, A. Matsumoto, S. Kitazume, N. Taniguchi, Functional roles of N-glycans in cell signaling and cell adhesion in cancer, *Cancer science* **2008**, *99*, 1304–1310.

[66] C. Boscher, J. W. Dennis, I. R. Nabi, Glycosylation, galectins and cellular signaling, *Curr. Opin. Cell. Biol.* **2011**, *23*, 383–392.

[67] J. B. Konopka, N-acetylglucosamine (GlcNAc) functions in cell signaling, *Scientifica* **2012**, *2012*, 1–15.

[68] D. Freitas *et al.*, O-glycans truncation modulates gastric cancer cell signaling and transcription leading to a more aggressive phenotype, *EBioMedicine* **2019**, *40*, 349–362.

[69] M. B. Lawrence, T. A. Springer, Leukocytes roll on a selectin at physiologic flow rates: Distinction from and prerequisite for adhesion through integrins, *Cell* **1991**, *65*, 859–873.

[70] R. P. McEver, Selectins: lectins that initiate cell adhesion under flow, *Curr. Opin. Cell. Biol.* **2002**, *14*, 581–586.

[71] *Advances in Human Genetics 10*, (Eds.: H. Harris, K. Hirschhorn), Springer US, Boston, MA, **1980**.

[72] W. M. Watkins, The ABO blood group system: historical background, *Transfus. Med. (Transfusion Medicine)* **2001**, *11*, 243–265.

[73] L. Schofield, M. C. Hewitt, K. Evans, M.-A. Siomos, P. H. Seeberger, Synthetic GPI as a candidate anti-toxic vaccine in a model of malaria, *Nature* **2002**, *418*, 785–789.

[74] B. Lepenies, P. H. Seeberger, The promise of glycomics, glycan arrays and carbohydrate-based vaccines, *Immunopharmacol. Immunotoxicol.* **2010**, *32*, 196–207.

[75] R. D. Astronomo, D. R. Burton, Carbohydrate vaccines: developing sweet solutions to sticky situations?, *Nature Rev. Drug Discovery* **2010**, *9*, 308–324.

[76] D. H. Dube, C. R. Bertozzi, Glycans in cancer and inflammation–potential for therapeutics and diagnostics, *Nature Rev. Drug Discovery* **2005**, *4*, 477–488.

[77] Z. Kyselova *et al.*, Breast cancer diagnosis and prognosis through quantitative measurements of serum glycan profiles, *Clin. Chem.* **2008**, *54*, 1166–1175.

[78] Z. Shriver, S. Raguram, R. Sasisekharan, Glycomics: a pathway to a class of new and improved therapeutics, *Nature Rev. Drug Discovery* **2004**, *3*, 863–873.

[79] H. Ghazarian, B. Idoni, S. B. Oppenheimer, A glycobiology review: carbohydrates, lectins and implications in cancer therapeutics, *Acta Histochem.* **2011**, *113*, 236–247.

[80] N. Rudman, O. Gornik, G. Lauc, Altered N-glycosylation profiles as potential biomarkers and drug targets in diabetes, *FEBS Lett.* **2019**, *593*, 1598–1615.

[81] R. B. Myers, S. Srivastava, W. E. Grizzle, Lewis Y Antigen as Detected by the Monoclonal Antibody BR96 is Expressed Strongly in Prostatic Adenocarcinoma, *J. Urol.* **1995**, *153*, 1572–1574.

[82] B. W. T. Yin *et al.*, Serological and immunochemical analysis of Lewis Y (Ley) blood group antigen expression in epithelial ovarian cancer, *Int. J. Cancer* **1996**, *65*, 406–412.

[83] X. Xie, M. Boysen, O. P. Clausen, M. A. Bryne, Prognostic value of Le(y) and H antigens in oral tongue carcinomas, *Laryngoscope* **1999**, *109*, 1474–1480.

[84] C. Haglund, J. Lindgren, P. J. Roberts, S. Nordling, Gastrointestinal cancer-associated antigen CA 19-9 in histological specimens

of pancreatic tumours and pancreatitis, *Br. J. Cancer* **1986**, *53*, 189–195.

[85] F. Safi, R. Roscher, R. Bittner, B. Schenkluhn, H. P. Dopfer, H. G. Beger, High sensitivity and specificity of CA 19-9 for pancreatic carcinoma in comparison to chronic pancreatitis. Serological and immunohistochemical findings, *Pancreas* **1987**, *2*, 398–403.

[86] Y. Zhang, J. Yang, H. Li, Y. Wu, H. Zhang, W. Chen, Tumor markers CA19-9, CA242 and CEA in the diagnosis of pancreatic cancer: a meta-analysis, *Int. J. Clin. Exp. Med.* **2015**, *8*, 11683–11691.

[87] D. D. Engle *et al.*, The glycan CA19-9 promotes pancreatitis and pancreatic cancer in mice, *Science* **2019**, *364*, 1156–1162.

[88] P. Kapitza, Viscosity of Liquid Helium below the λ-Point, *Nature* **1938**, *141*, 74.

[89] J. F. Allen, A. D. Misener, Flow of Liquid Helium II, *Nature* **1938**, *141*, 75.

[90] R. J. Donnelly, C. F. Barenghi, The Observed Properties of Liquid Helium at the Saturated Vapor Pressure, *J. Phys. Chem. Ref. Data* **1998**, *27*, 1217.

[91] L. TISZA, Transport Phenomena in Helium II, *Nature* **1938**, *141*, 913.

[92] L. Landau, Theory of the Superfluidity of Helium II, *Phys. Rev.* **1941**, *60*, 356–358.

[93] E. Whittle, D. A. Dows, G. C. Pimentel, Matrix Isolation Method for the Experimental Study of Unstable Species, *J. Chem. Phys.* **1954**, *22*, 1943.

[94] M. van Thiel, E. D. Becker, G. C. Pimentel, Infrared Studies of Hydrogen Bonding of Water by the Matrix Isolation Technique, *J. Chem. Phys.* **1957**, *27*, 486–490.

[95] S. Riedel, T. Köchner, X. Wang, L. Andrews, Polyfluoride anions, a matrix-isolation and quantum-chemical investigation, *Inorg. Chem.* **2010**, *49*, 7156–7164.

[96] I. F. Silvera, Ultimate fate of a gas of atomic hydrogen in a liquid-helium chamber: Recombination and burial, *Phys. Rev. B* **1984**, *29*, 3899–3904.

[97] A. Scheidemann, J. P. Toennies, J. A. Northby, Capture of neon atoms by He4 clusters, *Phys. Rev. Lett.* **1990**, *64*, 1899–1902.

[98] J. Gspann, Atomic impact experiments with free helium-3 and helium-4 clusters, *Z. Phys. B: Condens. Matter* **1995**, *98*, 405–411.

[99] F. Stienkemeier, K. K. Lehmann, Spectroscopy and dynamics in helium nanodroplets, *J. Phys. B: At. Mol. Opt. Phys.* **2006**, *39*, R127–R166.

[100] S. Moroni, N. Blinov, P.-N. Roy, Quantum Monte Carlo study of helium clusters doped with nitrous oxide: quantum solvation and rotational dynamics, *J. Chem. Phys.* **2004**, *121*, 3577–3581.

[101] C. Candler, A Unit of Wave-number, *Nature* **1952**, *170*, 43.

[102] P. A. M. Dirac, The Fundamental Equations of Quantum Mechanics, *Proc. R. Soc. London Ser. A* **1925**, *109*, 642–653.

[103] P. M. Morse, Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels, *Phys. Rev.* **1929**, *34*, 57–64.

[104] N. C. Polfer, Infrared multiple photon dissociation spectroscopy of trapped ions, *Chem. Soc. Rev.* **2011**, *40*, 2211–2221.

[105] K. K. Lehmann, G. Scoles, B. H. Pate, Intramolecular Dynamics from Eigenstate-Resolved Infrared Spectra, *Annu. Rev. Phys. Chem.* **1994**, *45*, 241–274.

[106] D. J. Nesbitt, R. W. Field, Vibrational Energy Flow in Highly Excited Molecules: Role of Intramolecular Vibrational Redistribution, *J. Phys. Chem.* **1996**, *100*, 12735–12756.

[107] M. Marianski *et al.*, Structural Characterization of Molybdenum Oxide Nanoclusters Using Ion Mobility Spectrometry–Mass Spectrometry and Infrared Action Spectroscopy, *J. Phys. Chem. C* **2019**, *123*, 7845–7853.

[108] N. Heine, K. R. Asmis, Cryogenic ion trap vibrational spectroscopy of hydrogen-bonded clusters relevant to atmospheric chemistry, *Int. Rev. Phys. Chem.* **2014**, *34*, 1–34.

[109] C. T. Wolke *et al.*, Spectroscopic snapshots of the proton-transfer mechanism in water, *Science (New York N.Y.)* **2016**, *354*, 1131–1135.

[110] S. Goyal, D. L. Schutt, G. Scoles, Vibrational spectroscopy of sulfur hexafluoride attached to helium clusters, *Phys. Rev. Lett.* **1992**, *69*, 933–936.

[111] S. Grebenev, M. Hartmann, M. Havenith, B. Sartakov, J. P. Toennies, A. F. Vilesov, The rotational spectrum of single OCS molecules in liquid 4He droplets, *J. Chem. Phys.* **2000**, *112*, 4485.

[112] M. N. Slipchenko, K. E. Kuyanov, B. G. Sartakov, A. F. Vilesov, Infrared intensity in small ammonia and water clusters, *J. Chem. Phys.* **2006**, *124*, 241101.

[113] K. Nauta, Formation of Cyclic Water Hexamer in Liquid Helium: The Smallest Piece of Ice, *Science* **2000**, *287*, 293–295.

[114] K. Nauta, Nonequilibrium Self-Assembly of Long Chains of Polar Molecules in Superfluid Helium, *Science* **1999**, *283*, 1895–1897.

[115] D. Verma, R. M. P. Tanyag, S. M. O. O'Connell, A. F. Vilesov, Infrared spectroscopy in superfluid helium droplets, *Adv. Phys. X* **2019**, *4*, 1553569.

[116] S. Smolarek, N. B. Brauer, W. J. Buma, M. Drabbels, IR spectroscopy of molecular ions by nonthermal ion ejection from helium nanodroplets, *J. Am. Chem. Soc.* **2010**, *132*, 14086–14091.

[117] X. Zhang, N. B. Brauer, G. Berden, A. M. Rijs, M. Drabbels, Mid-infrared spectroscopy of molecular ions in helium nanodroplets, *J. Chem. Phys.* **2012**, *136*, 044305.

[118] F. Filsinger, D.-S. Ahn, G. Meijer, G. von Helden, Photoexcitation of mass/charge selected hemin$^+$, caught in helium nanodroplets, *Phys. Chem. Chem. Phys.* **2012**, *14*, 13370–13377.

[119] W. Schöllkopf, S. Gewinner, H. Junkes, A. Paarmann, G. von Helden, H. Bluem, A. M. M. Todd, *Proc. SPIE Int. Soc. Opt. Eng.*, **2015**, p. 95121L.

[120] D. A. Thomas, M. Marianski, E. Mucha, G. Meijer, M. A. Johnson, G. von Helden, Ground-State Structure of the Proton-Bound Formate Dimer by Cold-Ion Infrared Action Spectroscopy, *Angew. Chem. Int. Ed.* **2018**, *57*, 10615–10619.

[121] D. A. Thomas, E. Mucha, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, Vibrational Spectroscopy of Fluoroformate, FCO_2^-, Trapped in Helium Nanodroplets, *J. Phys. Chem. Lett.* **2018**, *9*, 2305–2310.

[122] D. A. Thomas, E. Mucha, M. Lettow, G. Meijer, M. Rossi, G. von Helden, Characterization of a trans-trans Carbonic Acid-Fluoride Complex by Infrared Action Spectroscopy in Helium Nanodroplets, *J. Am. Soc. Mass Spectrom.* **2019**, *141*, 5815–5823.

[123] A. I. González Flórez, D.-S. Ahn, S. Gewinner, W. Schöllkopf, G. von Helden, IR spectroscopy of protonated leu-enkephalin and its 18-crown-6 complex embedded in helium droplets, *Phys. Chem. Chem. Phys.* **2015**, *17*, 21902–21911.

[124] A. I. González Flórez, E. Mucha, D.-S. Ahn, S. Gewinner, W. Schöllkopf, K. Pagel, G. von Helden, Charge-Induced Un-

zipping of Isolated Proteins to a Defined Secondary Structure, *Angew. Chem. Int. Ed.* **2016**, *55*, 3295–3299.

[125] M. Dole, L. L. Mack, R. L. Hines, R. C. Mobley, L. D. Ferguson, M. B. Alice, Molecular Beams of Macroions, *J. Chem. Phys.* **1968**, *49*, 2240–2249.

[126] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, Electrospray ionization-principles and practice, *Mass Spectrom. Rev.* **1990**, *9*, 37–70.

[127] J. B. Fenn, Electrospray wings for molecular elephants (Nobel lecture), *Angew. Chem. Int. Ed.* **2003**, *42*, 3871–3894.

[128] G. I. Taylor, Disintegration of water drops in an electric field, *Proc. R. Soc. Lond. A* **1964**, *280*, 383–397.

[129] M. S. Wilm, M. Mann, Electrospray and Taylor-Cone theory, Dole's beam of macromolecules at last?, *Int. J. Mass Spectrom. Ion Processes* **1994**, *136*, 167–180.

[130] D. C. Taflin, T. L. Ward, E. J. Davis, Electrified droplet fission and the Rayleigh limit, *Langmuir* **1989**, *5*, 376–384.

[131] A. Gomez, K. Tang, Charge and fission of droplets in electrostatic sprays, *Phys. Fluids* **1994**, *6*, 404–414.

[132] P. Kebarle, U. H. Verkerk, Electrospray: from ions in solution to ions in the gas phase, what we know now, *Mass Spectrometry Reviews* **2009**, *28*, 898–917.

[133] G. A. Valaskovic, N. L. Kelleher, D. P. Little, D. J. Aaserud, F. W. McLafferty, Attomole-Sensitivity Electrospray Source for Large-Molecule Mass Spectrometry, *Anal. Chem.* **1995**, *67*, 3802–3805.

[134] M. Wilm, M. Mann, Analytical Properties of the Nanoelectrospray Ion Source, *Anal. Chem.* **1996**, *68*, 1–8.

[135] J. L. Lippens, J. B. Mangrum, W. McIntyre, B. Redick, D. Fabris, A simple heated-capillary modification improves the analysis of non-covalent complexes by Z-spray electrospray ionization, *Rapid Commun. Mass Spectrom.* **2016**, *30*, 773–783.

[136] D. Gerlich in *Advances in Chemical Physics*, *Vol. 75* (Eds.: C.-Y. Ng, M. Baer, I. Prigogine, S. A. Rice), John Wiley & Sons, Inc, Hoboken, NJ, USA, **1992**, pp. 1–176.

[137] W. Paul, H. P. Reinhard, U. von Zahn, Das elektrische Massenfilter als Massenspektrometer und Isotopentrenner, *Z. Phys.* **1958**, *152*, 143–182.

[138] P. E. Miller, M. B. Denton, The quadrupole mass filter: Basic operating concepts, *J. Chem. Educ.* **1986**, *63*, 617.

[139] S. A. Schaffer, K. Tang, G. A. Anderson, D. C. Prior, H. R. Udseth, R. D. Smith, A novel ion funnel for focusing ions at elevated pressure using electrospray ionization mass spectrometry, *Rapid Commun. Mass Spectrom.* **1997**, *11*, 1813–1817.

[140] T. Kim *et al.*, Design and Implementation of a New Electrodynamic Ion Funnel, *Anal. Chem.* **2000**, *72*, 2247–2255.

[141] R. T. Kelly, A. V. Tolmachev, J. S. Page, K. Tang, R. D. Smith, The ion funnel: theory, implementations, and applications, *Mass Spectrom. Rev.* **2010**, *29*, 294–312.

[142] A. Luca, S. Schlemmer, I. Čermák, D. Gerlich, On the combination of a linear field free trap with a time-of-flight mass spectrometer, *Rev. Sci. Instrum.* **2001**, *72*, 2900–2908.

[143] D. J. Goebbert, T. Wende, R. Bergmann, G. Meijer, K. R. Asmis, Messenger-tagging electrosprayed ions: vibrational spectroscopy of suberate dianions, *J. Phys. Chem. A* **2009**, *113*, 5874–5880.

[144] A. A. Shvartsburg, R. D. Smith, Fundamentals of traveling wave ion mobility spectrometry, *Anal. Chem.* **2008**, *80*, 9689–9699.

[145] R. Cumeras, E. Figueras, C. E. Davis, J. I. Baumbach, I. Gràcia, Review on ion mobility spectrometry. Part 1: current instrumentation, *The Analyst* **2015**, *140*, 1376–1390.

[146] C. G. Herbert, R. A. W. Johnstone, *Mass Spectrometry Basics*, CRC Press, Boca Raton, **2003**.

[147] B. A. Mamyrin, V. I. Karataev, D. V. Shmikk, V. A. Zagulin, The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution, *J. Exp. Theor. Phys.* **1973**, *37*, 45.

[148] V. M. Doroshenko, R. J. Cotter, Ideal velocity focusing in a reflectron time-of-flight mass spectrometer, *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 992–999.

[149] A. I. González Flórez, *Biomolecular Ions in Helium Nanodroplets*, PhD Thesis, Freie Universität Berlin, Berlin, **2016**.

[150] G. Savard *et al.*, A new cooling technique for heavy ions in a Penning trap, *Phys. Lett. A* **1991**, *158*, 247–252.

[151] D. Gerlich, G. Borodi, Buffer gas cooling of polyatomic ions in rf multi-electrode traps, *Faraday Discussions* **2009**, *142*, 57.

[152] J. M. J. Madey, Stimulated Emission of Bremsstrahlung in a Periodic Magnetic Field, *J. Appl. Phys.* **1971**, *42*, 1906.

[153] W. C. Wiley, I. H. McLaren, Time–of–Flight Mass Spectrometer with Improved Resolution, *Rev. Sci. Instrum.* **1955**, *26*, 1150–1157.

[154] E. Mucha *et al.*, Glycan Fingerprinting via Cold-Ion Infrared Spectroscopy, *Angew. Chem. Int. Ed.* **2017**, *56*, 11248–11251.

[155] O. J. Plante, E. R. Palmacci, P. H. Seeberger, Automated solid-phase synthesis of oligosaccharides, *Science* **2001**, *291*, 1523–1527.

[156] S. Eller, M. Collot, J. Yin, H. S. Hahm, P. H. Seeberger, Automated solid-phase synthesis of chondroitin sulfate glycosaminoglycans, *Angew. Chem. Int. Ed.* **2013**, *52*, 5858–5861.

[157] M. P. Bartetzko, F. Schuhmacher, H. S. Hahm, P. H. Seeberger, F. Pfrengle, Automated Glycan Assembly of Oligosaccharides Related to Arabinogalactan Proteins, *Org. Lett.* **2015**, *17*, 4344–4347.

[158] M. Marianski, A. Supady, T. Ingram, M. Schneider, C. Baldauf, 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*, 6157–6168.

[159] J. B. Pendry, Reliability factors for LEED calculations, *J. Phys. C: Solid State Phys.* **1980**, *13*, 937–944.

[160] C. Masellis, N. Khanal, M. Z. Kamrath, D. E. Clemmer, T. R. Rizzo, Cryogenic Vibrational Spectroscopy Provides Unique Fingerprints for Glycan Identification, *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 2217–2222.

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

[162] N. Khanal, C. Masellis, M. Z. Kamrath, D. E. Clemmer, T. R. Rizzo, Cryogenic IR spectroscopy combined with ion mobility spectrometry for the analysis of human milk oligosaccharides, *The Analyst* **2018**, *143*, 1846–1852.

[163] V. Scutelnic, T. R. Rizzo, Cryogenic Ion Spectroscopy for Identification of Monosaccharide Anomers, *J. Phys. Chem. A* **2019**, *123*, 2815–2819.

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

[165] E. Mucha *et al.*, Fucose Migration in Intact Protonated Glycan Ions: A Universal Phenomenon in Mass Spectrometry, *Angew. Chem. Int. Ed.* **2018**, *57*, 7440–7443.

[166] M. Schneider, E. Al-Shareffi, R. S. Haltiwanger, Biological functions of fucose in mammals, *Glycobiology* **2017**, *27*, 601–618.

[167] D. B. Werz, R. Ranzinger, S. Herget, A. Adibekian, C.-W. von der Lieth, P. H. Seeberger, Exploring the structural diversity of mammalian carbohydrates (glycospace) by statistical databank analysis, *ACS Chem. Biol.* **2007**, *2*, 685–691.

[168] G. R. Guile *et al.*, Identification of highly fucosylated N-linked oligosaccharides from the human parotid gland, *Eur. J. Biochem.* **1998**, *258*, 623–656.

[169] S. Shi, P. Stanley, Protein O-fucosyltransferase 1 is an essential component of Notch signaling pathways, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5234–5239.

[170] X. Wang *et al.*, Dysregulation of TGF-beta1 receptor activation leads to abnormal lung development and emphysema-like phenotype in core fucose-deficient mice, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15791–15796.

[171] J. Du, H. Takeuchi, C. Leonhard-Melief, K. R. Shroyer, M. Dlugosz, R. S. Haltiwanger, B. C. Holdener, O-fucosylation of thrombospondin type 1 repeats restricts epithelial to mesenchymal transition (EMT) and maintains epiblast pluripotency during mouse gastrulation, *Dev. Biol.* **2010**, *346*, 25–38.

[172] P. Stanley, What Have We Learned from Glycosyltransferase Knockouts in Mice?, *J. Mol. Biol.* **2016**, *428*, 3166–3182.

[173] S. A. Kalovidouris, C. I. Gama, L. W. Lee, L. C. Hsieh-Wilson, A role for fucose alpha(1-2) galactose carbohydrates in neuronal growth, *J. Am. Chem. Soc.* **2005**, *127*, 1340–1341.

[174] H. E. Murrey, C. I. Gama, S. A. Kalovidouris, W.-I. Luo, E. M. Driggers, B. Porton, L. C. Hsieh-Wilson, Protein fucosylation regulates synapsin Ia/Ib expression and neuronal morphology in primary hippocampal neurons, *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 21–26.

[175] R. Lefebvre, M. C. Lo, S. S. Suarez, Bovine sperm binding to oviductal epithelium involves fucose recognition, *Biol. Reprod.* **1997**, *56*, 1198–1204.

[176] N. Yonezawa, S. Mitsui, K. Kudo, M. Nakano, Identification of an N-glycosylated region of pig zona pellucida glycoprotein ZPB that is involved in sperm binding, *Eur. J. Biochem. (Eur. J. Biochem.)* **1997**, *248*, 86–92.

[177] D. S. Johnston, W. W. Wright, J. H. Shaper, C. H. Hokke, D. H. van den Eijnden, D. H. Joziasse, Murine sperm-zona binding, a fucosyl residue is required for a high affinity sperm-binding ligand. A second site on sperm binds a nonfucosylated, beta-galactosyl-capped oligosaccharide, *J. Biol. Chem.* **1998**, *273*, 1888–1895.

[178] P.-C. Pang *et al.*, Human sperm binding is mediated by the sialyl-Lewis(x) oligosaccharide on the zona pellucida, *Science* **2011**, *333*, 1761–1764.

[179] D. Ilver *et al.*, Helicobacter pylori adhesin binding fucosylated histo-blood group antigens revealed by retagging, *Science* **1998**, *279*, 373–377.

[180] D. J. Becker, J. B. Lowe, Fucose: biosynthesis and biological function in mammals, *Glycobiology* **2003**, *13*, 41R–53R.

[181] D. J. Harvey, T. S. Mattu, M. R. Wormald, L. Royle, R. A. Dwek, P. M. Rudd, Internal residue loss: rearrangements occurring during

the fragmentation of carbohydrates derivatized at the reducing terminus, *Anal. Chem.* **2002**, *74*, 734–740.

[182] M. McNeil, Elimination of internal glycosyl residues during chemical ionization-mass spectrometry of per-O-alkylated oligosaccharide-alditols, *Carbohydr. Res.* **1983**, *123*, 31–40.

[183] V. Kováčik, J. Hirsch, P. Kováč, W. Heerma, J. Thomas-Oates, J. Haverkamp, Oligosaccharide characterization using collision-induced dissociation fast atom bombardment mass spectrometry: Evidence for internal monosaccharide residue loss, *J. Mass Spectrom.* **1995**, *30*, 949–958.

[184] B. Ernst, D. R. Müller, W. J. Richter, False sugar sequence ions in electrospray tandem mass spectrometry of underivatized sialyl-Lewis-type oligosaccharides, *Int. J. Mass Spectrom. Ion Processes* **1997**, *160*, 283–290.

[185] L. P. Brüll, W. Heerma, J. Thomas-Oates, J. Haverkamp, V. Kovácik, P. Kovác, Loss of internal 1→6 substituted monosaccharide residues from underivatized and per-O-methylated trisaccharides, *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 43–49.

[186] L. P. Brüll, V. Kovácik, J. E. Thomas-Oates, W. Heerma, J. Haverkamp, Sodium-cationized oligosaccharides do not appear to undergo 'internal residue loss' rearrangement processes on tandem mass spectrometry, *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1520–1532.

[187] Y.-L. Ma, I. Vedernikova, H. Heuvel, M. Claeys, Internal glucose residue loss in protonated O-diglycosyl flavonoids upon low-energy collision-induced dissociation, *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 136–144.

[188] A. H. Franz, C. B. Lebrilla, Evidence for long-range glycosyl transfer reactions in the gas phase, *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 325–337.

[189] M. Wuhrer, C. A. M. Koeleman, C. H. Hokke, A. M. Deelder, Mass spectrometry of proton adducts of fucosylated N-glycans: fucose transfer between antennae gives rise to misleading fragments, *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1747–1754.

[190] M. Wuhrer, C. A. M. Koeleman, A. M. Deelder, Hexose rearrangements upon fragmentation of N-glycopeptides and reductively aminated N-glycans, *Anal. Chem.* **2009**, *81*, 4422–4432.

[191] E. S. Hecht, P. L. Loziuk, D. C. Muddiman, Xylose Migration During Tandem Mass Spectrometry of N-Linked Glycans, *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 729–732.

[192] X. J. Tang, P. Thibault, R. K. Boyd, Fragmentation reactions of multiply-protonated peptides and implications for sequencing by tandem mass spectrometry with low-energy collision-induced dissociation, *Anal. Chem.* **1993**, *65*, 2824–2834.

[193] R. W. Vachet, B. M. Bishop, B. W. Erickson, G. L. Glish, Novel Peptide Dissociation: Gas-Phase Intramolecular Rearrangement of Internal Amino Acid Residues, *J. Am. Chem. Soc.* **1997**, *119*, 5481–5488.

[194] I. S. Saminathan, X. S. Wang, Y. Guo, O. Krakovska, S. Voisin, A. C. Hopkinson, K. W. M. Siu, The extent and effects of peptide sequence scrambling via formation of macrocyclic B ions in model proteins, *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 2085–2094.

[195] N. C. Polfer, J. Oomens, S. Suhai, B. Paizs, 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*, 5887–5897.

[196] U. Erlekam, B. J. Bythell, D. Scuderi, M. van Stipdonk, B. Paizs, P. Maître, Infrared spectroscopy of fragments of protonated peptides: direct evidence for macrocyclic structures of b5 ions, *J. Am. Chem. Soc.* **2009**, *131*, 11503–11508.

[197] X. Chen, L. Yu, J. D. Steill, J. Oomens, N. C. Polfer, Effect of peptide fragment size on the propensity of cyclization in collision-induced dissociation: oligoglycine b(2)-b(8), *J. Am. Chem. Soc.* **2009**, *131*, 18272–18282.

[198] R. Chawner, S. W. Holman, S. J. Gaskell, C. E. Eyers, Peptide scrambling during collision-induced dissociation is influenced by N-terminal residue basicity, *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1927–1938.

[199] C. Nwosu, H. K. Yau, S. Becht, Assignment of Core versus Antenna Fucosylation Types in Protein N-Glycosylation via Procainamide Labeling and Tandem Mass Spectrometry, *Anal. Chem.* **2015**, *87*, 5905–5913.

[200] N. Desai, D. A. Thomas, J. Lee, J. Gao, J. L. Beauchamp, Eradicating mass spectrometric glycan rearrangement by utilizing free radicals, *Chem. Sci.* **2016**, *7*, 5390–5397.

[201] M. Lettow, E. Mucha, C. Manz, D. A. Thomas, M. Marianski, G. Meijer, G. von Helden, K. Pagel, The role of the mobile proton in fucose migration, *Anal. Bioanal. Chem.* **2019**, *411*, 4637–4645.

[202] E. Mucha, M. Marianski, F.-F. Xu, D. A. Thomas, G. Meijer, G. von Helden, P. H. Seeberger, K. Pagel, Unravelling the Structure of Glycosyl Cations via Cold-Ion Infrared Spectroscopy, *Nat. Commun.* **2018**, *9*, 4174.

[203] R. M. Albano, P. A. Mourão, Isolation, fractionation, and preliminary characterization of a novel class of sulfated glycans from the tunic of Styela plicata (Chordata Tunicata), *J. Biol. Chem.* **1986**, *261*, 758–765.

[204] Z. Zhuang, J. A. Starkey, Y. Mechref, M. V. Novotny, S. C. Jacobson, Electrophoretic analysis of N-glycans on microfluidic devices, *Anal. Chem.* **2007**, *79*, 7170–7175.

[205] W. R. Alley, B. F. Mann, V. Hruska, M. V. Novotny, Isolation and purification of glycoconjugates from complex biological sources by recycling high-performance liquid chromatography, *Anal. Chem.* **2013**, *85*, 10408–10416.

[206] X. Zhu, R. R. Schmidt, New principles for glycoside-bond formation, *Angew. Chem. Int. Ed.* **2009**, *48*, 1900–1934.

[207] Z. Zhang, I. R. Ollmann, X.-S. Ye, R. Wischnat, T. Baasov, C.-H. Wong, Programmable One-Pot Oligosaccharide Synthesis, *J. Am. Chem. Soc.* **1999**, *121*, 734–753.

[208] P. H. Seeberger, W.-C. Haase, Solid-Phase Oligosaccharide Synthesis and Combinatorial Carbohydrate Libraries, *Chem. Rev.* **2000**, *100*, 4349–4394.

[209] T. Nokami, R. Hayashi, Y. Saigusa, A. Shimizu, C.-Y. Liu, K.-K. T. Mong, J.-i. Yoshida, Automated solution-phase synthesis of oligosaccharides via iterative electrochemical assembly of thioglycosides, *Org. Lett.* **2013**, *15*, 4520–4523.

[210] D. Sames, X. T. Chen, S. J. Danishefsky, Convergent total synthesis of a tumour-associated mucin motif, *Nature* **1997**, *389*, 587–591.

[211] P. Pornsuriyasak, A. V. Demchenko, Glycosyl thioimidates in a highly convergent one-pot strategy for oligosaccharide synthesis, *Tetrahedron: Asymmetry* **2005**, *16*, 433–439.

[212] S. Meng, T. Tian, Y.-H. Wang, X.-B. Meng, Z.-J. Li, Convergent synthesis of oligosaccharides on the gram-scale using cetyl thioglycoside based on a hydrophobically assisted switching phase method, *Org. Biomol. Chem.* **2016**, *14*, 7722–7730.

[213] T. Miura, Y. Hirose, M. Ohmae, T. Inazu, Fluorous oligosaccharide synthesis using a novel fluorous protective group, *Org. Lett.* **2001**, *3*, 3947–3950.

[214] T. Miura, K. Goto, D. Hosaka, T. Inazu, Oligosaccharide synthesis on a fluorous support, *Angew. Chem. Int. Ed.* **2003**, *42*, 2047–2051.

[215] S.-L. Tang, N. L. B. Pohl, Automated Solution-Phase Synthesis of β-1,4-Mannuronate and β-1,4-Mannan, *Org. Lett.* **2015**, *17*, 2642–2645.

[216] E. R. Palmacci, M. C. Hewitt, P. H. Seeberger, "Cap-Tag"—Novel Methods for the Rapid Purification of Oligosaccharides Prepared by Automated Solid-Phase Synthesis, *Angew. Chem. Int. Ed.* **2001**, *40*, 4433.

[217] P. H. Seeberger, The logic of automated glycan assembly, *Acc. Chem. Res.* **2015**, *48*, 1450–1463.

[218] H. S. Hahm, M. K. Schlegel, M. Hurevich, S. Eller, F. Schuhmacher, J. Hofmann, K. Pagel, P. H. Seeberger, Automated glycan assembly using the Glyconeer 2.1 synthesizer, *Proceedings of the National Academy of Sciences of the United States of America* **2017**, *114*, E3385–E3389.

[219] A. Pardo-Vargas, M. Delbianco, P. H. Seeberger, Automated glycan assembly as an enabling technology, *Curr. Opin. Chem. Biol.* **2018**, *46*, 48–55.

[220] Y. Yu, A. Kononov, M. Delbianco, P. H. Seeberger, A Capping Step During Automated Glycan Assembly Enables Access to Complex Glycans in High Yield, *Chem. Eur. J.* **2018**, *24*, 6075–6078.

[221] H. S. Hahm, M. Hurevich, P. H. Seeberger, Automated assembly of oligosaccharides containing multiple cis-glycosidic linkages, *Nat. Commun.* **2016**, *7*, 12482.

[222] P. O. Adero, H. Amarasekara, P. Wen, L. Bohé, D. Crich, The Experimental Evidence in Support of Glycosylation Mechanisms at the SN1-SN2 Interface, *Chem. Rev.* **2018**, *118*, 8242–8284.

[223] S. Chatterjee, S. Moon, F. Hentschel, K. Gilmore, P. H. Seeberger, An Empirical Understanding of the Glycosylation Reaction, *J. Am. Chem. Soc.* **2018**, *140*, 11942–11953.

[224] S. S. Nigudkar, A. V. Demchenko, Stereocontrolled 1,2-cis glycosylation as the driving force of progress in synthetic carbohydrate chemistry, *Chem. Sci.* **2015**, *6*, 2687–2704.

[225] Y. Zeng, J. Ning, F. Kong, Pure α-linked products can be obtained in high yields in glycosylation with glucosyl trichloroacetimidate donors with a C2 ester capable of neighboring group participation, *Tetrahedron Lett.* **2002**, *43*, 3729–3733.

[226] D. Crich, Z. Dai, S. Gastaldi, On the Role of Neighboring Group Participation and Ortho Esters in β-Xylosylation: 13 C NMR Observation of a Bridging 2-Phenyl-1,3-dioxalenium Ion, *J. Org. Chem.* **1999**, *64*, 5224–5229.

[227] A. Bérces, G. Enright, T. Nukada, D. M. Whitfield, The conformational origin of the barrier to the formation of neighboring group assistance in glycosylation reactions: a dynamical density functional theory study, *J. Am. Chem. Soc.* **2001**, *123*, 5460–5464.

[228] K. J. Jensen, O-Glycosylations under neutral or basic conditions, *J. Chem. Soc. Perkin Trans. 1* **2002**, 2219–2233.

[229] J. Banoub, D. R. Bundle, 1,2-Orthoacetate intermediates in silver trifluoromethanesulphonate promoted Koenigs–Knorr synthesis of disaccharide glycosides, *Can. J. Chem.* **1979**, *57*, 2091–2097.

[230] T. Nukada, A. Berces, M. Z. Zgierski, D. M. Whitfield, Exploring the Mechanism of Neighboring Group Assisted Glycosylation Reactions, *J. Am. Chem. Soc.* **1998**, *120*, 13291–13295.

[231] Z. Yang, W. Lin, B. Yu, Rearrangement of sugar 1,2-orthoesters to glycosidic products: a mechanistic implication, *Carbohydr. Res.* **2000**, *329*, 879–884.

[232] T. Nukada, A. Berces, D. M. Whitfield, Acyl Transfer as a Problematic Side Reaction in Polymer-Supported Oligosaccharide Synthesis, *J. Org. Chem.* **1999**, *64*, 9030–9045.

[233] A. V. Demchenko, E. Rousson, G.-J. Boons, Stereoselective 1,2-cis-galactosylation assisted by remote neighboring group participation and solvent effects, *Tetrahedron Lett.* **1999**, *40*, 6523–6526.

[234] C. de Meo, M. N. Kamat, A. V. Demchenko, Remote Participation-Assisted Synthesis of β-Mannosides, *Eur. J. Org. Chem.* **2005**, *2005*, 706–711.

[235] B. S. Komarova, Y. E. Tsvetkov, N. E. Nifantiev, Design of α-Selective Glycopyranosyl Donors Relying on Remote Anchimeric Assistance, *Chem. Rec.* **2016**, *16*, 488–506.

[236] J.-H. Kim, H. Yang, J. Park, G.-J. Boons, A general strategy for stereoselective glycosylations, *J. Am. Chem. Soc.* **2005**, *127*, 12090–12097.

[237] T. J. Boltje, J.-H. Kim, J. Park, G.-J. Boons, Chiral-auxiliary-mediated 1,2-cis-glycosylations for the solid-supported synthesis of a biologically important branched alpha-glucan, *Nat. Chem.* **2010**, *2*, 552–557.

[238] T. J. Boltje, J.-H. Kim, J. Park, G.-J. Boons, Stereoelectronic effects determine oxacarbenium vs β-sulfonium ion mediated glycosylations, *Org. Lett.* **2011**, *13*, 284–287.

[239] H. Elferink, R. A. Mensink, P. B. White, T. J. Boltje, Stereoselective β-Mannosylation by Neighboring-Group Participation, *Angew. Chem.* **2016**, *128*, 11383–11386.

[240] Y.-P. Cheng, H.-T. Chen, C.-C. Lin, A convenient and highly stereoselective approach for α-galactosylation performed by galactopyranosyl dibenzyl phosphite with remote participating groups, *Tetrahedron Lett.* **2002**, *43*, 7721–7723.

[241] J. Kalikanda, Z. Li, Study of the stereoselectivity of 2-azido-2-deoxygalactosyl donors: remote protecting group effects and temperature dependency, *J. Org. Chem.* **2011**, *76*, 5207–5218.

[242] L. Bohé, D. Crich, A propos of glycosyl cations and the mechanism of chemical glycosylation; the current state of the art, *Carbohydr. Res.* **2015**, *403*, 48–59.

[243] X. S. Ye, C. H. Wong, Anomeric reactivity-based one-pot oligosaccharide synthesis: a rapid route to oligosaccharide libraries, *J. Org. Chem.* **2000**, *65*, 2410–2431.

[244] J. Y. Baek, B.-Y. Lee, M. G. Jo, K. S. Kim, Beta-directing effect of electron-withdrawing groups at O-3, O-4, and O-6 positions and alpha-directing effect by remote participation of 3-O-acyl and 6-O-acetyl groups of donors in mannopyranosylations, *J. Am. Chem. Soc.* **2009**, *131*, 17705–17713.

[245] K. Bock, I. Lundt, C. Pedersen, H. Pedersen, Glycosylation Reactions with Di-O-acetyl-2,6-dibromo-2,6-dideoxy-alpha-D-mannopyranosyl Bromide. A Simple Synthesis of Methyl 2,6-Dideoxy-D-arabino-hexopyranoside, *Acta Chem. Scand.* **1988**, *42b*, 640–645.

[246] A. Martin *et al.*, Catching elusive glycosyl cations in a condensed phase with HF/SbF_5 superacid, *Nat. Chem.* **2016**, *8*, 186–191.

[247] L. Lebedel *et al.*, Structural and Computational Analysis of 2-Halogeno-Glycosyl Cations in the Presence of a Superacid: An Expansive Platform, *Angew. Chem. Int. Ed.* **2019**, *58*, 13758–13762.

[248] C. Denekamp, Y. Sandlers, Formation and stability of oxocarbenium ions from glycosides, *J. Mass Spectrom.* **2005**, *40*, 1055–1063.

[249] P. K. Kancharla, C. Navuluri, D. Crich, Dissecting the influence of oxazolidinones and cyclic carbonates in sialic acid chemistry, *Angew. Chem. Int. Ed.* **2012**, *51*, 11105–11109.

[250] C. Denekamp, Y. Sandlers, Anomeric distinction and oxonium ion formation in acetylated glycosides, *J. Mass Spectrom.* **2005**, *40*, 765–771.

[251] E. R. Molina *et al.*, Characterization of Protonated Model Disaccharides from Tandem Mass Spectrometry and Chemical Dynamics Simulations, *ChemPhysChem* **2017**, *18*, 2812–2823.

[252] H. Elferink *et al.*, Direct Experimental Characterization of Glycosyl Cations by Infrared Ion Spectroscopy, *J. Am. Chem. Soc.* **2018**, *140*, 6034–6038.

[253] H. Elferink *et al.*, The Glycosylation Mechanisms of 6,3-Uronic Acid Lactones, *Angew. Chem. Int. Ed.* **2019**, *58*, 8746–8751.

[254] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, M. Scheffler, Ab initio molecular simulations with numeric atom-centered orbitals, *Comput. Phys. Commun.* **2009**, *180*, 2175–2196.

[255] V. Havu, V. Blum, P. Havu, M. Scheffler, Efficient integration for all-electron electronic structure calculation using numeric basis functions, *J. Comput. Phys.* **2009**, *228*, 8367–8379.

[256] Perdew, Burke, Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[257] A. Tkatchenko, M. Scheffler, Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data, *Phys. Rev. Lett.* **2009**, *102*, 073005.

[258] A. Supady, V. Blum, C. Baldauf, First-Principles Molecular Structure Search with a Genetic Algorithm, *J. Chem. Inf. Model.* **2015**, *55*, 2338–2348.

[259] M. J. Frisch *et al.*, *Gaussian 09 Revision D.01*.

[260] C. Adamo, V. Barone, Toward reliable density functional methods without adjustable parameters: The PBE0 model, *J. Chem. Phys.* **1999**, *110*, 6158–6170.

[261] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, 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*, 154104.

[262] F. Weigend, M. Kattannek, R. Ahlrichs, Approximated electron repulsion integrals: Cholesky decomposition versus resolution of the identity methods, *J. Chem. Phys.* **2009**, *130*, 164106.

[263] F. Neese, E. F. Valeev, Revisiting the Atomic Natural Orbital Approach for Basis Sets: Robust Systematic Basis Sets for Explicitly Correlated and Conventional Correlated ab initio Methods?, *J. Chem. Theory Comput.* **2011**, *7*, 33–43.

[264] L. Montalvillo-Jiménez, A. G. Santana, F. Corzana, G. Jiménez-Osés, J. Jiménez-Barbero, A. M. Gómez, J. L. Asensio, Impact of Aromatic Stacking on Glycoside Reactivity: Balancing CH/π and Cation/π Interactions for the Stabilization of Glycosyl-Oxocarbenium Ions, *J. Am. Chem. Soc.* **2019**, *141*, 13372–13384.

[265] S. van der Vorm, T. Hansen, H. S. Overkleeft, G. A. van der Marel, J. D. C. Codée, The influence of acceptor nucleophilicity on the glycosylation reaction mechanism, *Chem. Sci.* **2017**, *8*, 1867–1875.

[266] C. Bowick, J. Blyler, C. J. Ajluni, *RF circuit design*, Elsevier, Burlington, MA, 2nd ed., **2008**.

Acknowledgments

Numerous persons have contributed to this thesis one way or another, for which I am deeply grateful. First and foremost, I wish to thank my three supervisors Gert von Helden, Kevin Pagel and Gerard Meijer. Over the last few years, this unique constellation of supervisors tempted various people to ask me the same question from time to time: "Having *two* PhD supervisors can be very complicated, but how on earth do you handle *three* of them?". Fortunately, however, this was never an issue, and I am extremely grateful for their combined and continuous support. In particular, I would like to thank

Gert for always encouraging me to further develop the wonderful helium droplet experiment, and motivating me to step into new directions such as mechanical engineering, machining, or programming.

Kevin for his profound knowledge of the glycosciences, for always being enthusiastic about the latest results, his optimism when it comes to publishing said results, and his unique way of outlining stories to convey the key message of a manuscript.

Gerard for providing an amazing work environment in the Molecular Physics department concerning both social aspects and infrastructure, for sharing my excitement when yet another precision part was finished on our new engine lathe, and for being the director of our department,

who somehow managed to find the time to carefully proof-read this thesis.
Next, I would like to acknowledge Jos Oomens, Thomas Boltje, and Isabelle Compagnon for readily agreeing to evaluate this thesis as part of the manuscript committee.

When I started working at the Fritz Haber Institute, my knowledge concerning vibrational spectroscopy, mass spectrometry, and helium droplets was rather limited. Thanks to Isabel González Flórez, this changed after she introduced me to the experimental setup, taught me how to operate it, and supervised my very first experiments. Shortly after Isabel finished her PhD and left the FHI, Daniel Thomas joined our group as a Postdoc working on the helium droplet experiment. I am extremely grateful to Daniel for always supporting me in the lab, staying late during beam shifts, his expertise in so many different areas, and for being able to recall every single research article that remotely connects to our research. The experimental work in this thesis is often complemented by theory. This would have not been possible without the great effort by Mateusz Marianski, who burned many CPU-hours to calculate possible structures for these interesting molecules. I also had the great pleasure of supervising Maike Lettow, Carla Kirschbaum, and Kim Greis as three extremely talented students during their master theses. It was a true pleasure working with you and I couldn't be more confident in 'handing over' the helium droplet experiment to you as the next PhD-generation.

The long nights in the lab during beam shifts and the successful measurements would have not been possible without two particular things: the delicious number 27 from that Vietnamese place around the corner, and the large number of photons generated by the Fritz Haber Institute Free-electron Laser. For the latter, I would like to

thank Wieland Schöllkopf and Sandy Gewinner for providing around 10^{24} photons per day, for tuning the laser to our needs (the more power, the more better), and for staying late during beam shifts. In addition, I am deeply honored to have gained their trust to become a laser operator. Furthermore, special thanks go to Sandy for introducing me to speed-solving the Rubik's Cube – maybe I will beat your personal record some day. The FEL crew recently gained support by Marco De Pas, a true master of RF-electronics. Thank you for patiently explaining the underlying concepts of RF-engineering, and helping me to translate my visions into a circuit. When it came to mechanical engineering, Henrik Haak always took the time to review my designs, and suggest helpful modifications. I also appreciate the occasional chats in our workshop with Wolfgang Erlebach. Today, our workshop is in an extremely good shape thanks to the combined effort of Sebastian Kray and Klaus Peter Vogelgesang. Also the incredible support of Andreas Liedke regarding electricity and infrastructure around our experiment was of enormous help. Uwe Hoppe is greatly acknowledged for developing the data acquisition software in LabView, which he kept updating to our needs.

I also would like to acknowledge the other (former and current) members of our research group, namely Stephan Warnke, Jongcheol Seo, Waldemar Hoffmann, Leonhard Urner, Melanie Krahlisch, Johanna Hofmann, Rayoon Chang, Christian Manz, Márkó Grabarics, Jakob Xiao Schmidt, and Sebastian smalerz. You greatly contributed to making the time at the FHI so enjoyable. Also the non-scientific group events, such as several barbecues on the terrace, pub-nights, or the epic skiing-trip are unforgettable.

A lot of people also took great care of organizing our department. Therefore I would like to acknowledge Inga von Dölln + Zitrone, An-

drea Braker, Karin Grassow, and Manuela Misch for making sure that everything runs as smoothly as possible.

The experimental setup described in this work greatly benefited from the central services of the Fritz Haber Institute. The enormous support and technical possibilities of both the mechanical workshop and the electronics lab were invaluable to the success of this thesis, and I would like to thank the colleagues from both departments. In particular, I want to emphasize the great effort of Thorsten Nemes and Dieter Leese, who put their blood, sweat and tears into the variable temperature ion trap described in Section 3.3.

Next, I would like to explicitly thank the expertise of various anonymous reviewers, who always came up with new ingenious requests to 'improve' our work, such as measuring O−H-stretch vibrations of a molecule that does not contain any O−H groups.

Special thanks go to the wonderful friends I made during the Bachelor and Master studies: Lauryna, Lars, Tim, Daniel, and of course Marc. Thank you for making life at the university (and outside) so enjoyable.

I also wish to thank everyone who supported me beyond the scientific world. In particular, I want to mention all my friends from Groß Lafferde. We are currently quite spread across the globe and therefore do not manage to see each other too often, but every time we do meet in person, it's like we were hardly ever separated.

And then, there is one very special person who actually started her Bachelor of Chemistry at the Freie Universität just one semester after me. For some miraculous reason, we never actually met during this time. Fortunately, however, our paths finally crossed when I started my PhD and she came back from the UK to finish her master studies in our research group. Alex, I am so happy to have you by my side and I

would like to thank you for all the caring, love and support. Because I fail to put this into words, see the last page of this thesis.

Finally, I owe an enormous debt of gratitude to my family, and in particular my parents, Carmen and Thomas. They have supported me throughout my whole life, and I can always rely on them. Thank you for always being there for me.

List of Publications

1. M. Lettow, M. Grabarics, E. Mucha, D. A. Thomas, L. Polewski, J. Blaszkiewicz, J. Rademann, G. Meijer, G. von Helden, K. Pagel; IR Action Spectroscopy of Glycosaminoglycan Oligosaccharides, *Anal. Bioanal. Chem.* **2019**, *in press.*

2. D. A. Thomas, E. Mucha, M. Lettow, G. Meijer, M. Rossi, G. von Helden; Characterization of a *trans–trans* Carbonic Acid–Fluoride Complex by Infrared Action Spectroscopy in Helium Nanodroplets, *J. Am. Chem. Soc.* **2019**, *14*, 5815.

3. M. Lettow, E. Mucha, C. Manz, D. A. Thomas, M. Marianski, G. Meijer, G. von Helden, K. Pagel; The Role of the Mobile Proton in Fucose Migration, *Anal. Bioanal. Chem.* **2019**, *19*, 4637.

4. E. Mucha, A. Stuckmann, M. Marianski, W. B. Struwe, G. Meijer, K. Pagel; In-depth structural analysis of glycans in the gas phase, *Chem. Sci.* **2019**, *10*, 1272.

5. E. Mucha, M. Marianski, F.-F. Xu, D. A. Thomas, G. Meijer, G. von Helden, P. H. Seeberger, K. Pagel; Unravelling the Structure of Glycosyl Cations via Cold-Ion Infrared Spectroscopy, *Nat. Commun.* **2018**, *9*, 4147.

6. M. Marianski, J. Seo, E. Mucha, D. A. Thomas, S. Jung, R. Schlögl, G. Meijer, A. Trunschke, G. von Helden; Structural Characterization of Molybdenum Oxide Nanoclusters Using Ion Mobility Spectrometry-Mass Spectrometry and Infrared Action Spectroscopy, *J. Phys. Chem. C* **2018**, *123*, 7845.

7. D. A. Thomas, M. Marianski, E. Mucha, G. Meijer, M. A. Johnson, G. von Helden; Ground-State Structure of the Proton-Bound Formate Dimer by Cold-Ion Infrared Action Spectroscopy, *Angew. Chem. Int. Ed.* **2018**, *57*, 10615.

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

9. D. A. Thomas, E. Mucha, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden; Vibrational Spectroscopy of Fluoroformate, FCO_2^-, Trapped in Helium Nanodroplets, *J. Phys. Chem. Lett.* **2018**, *9*, 2305.

10. E. Mucha, A. I. González Flórez, M. Marianski, D. A. Thomas, W. Hoffmann, W. B. Struwe, H. S. Hahm, S. Gewinner, W. Schöllkopf, P. H. Seeberger, G. von Helden, K. Pagel; Glycan Fingerprinting via Cold-Ion Infrared Spectroscopy, *Angew. Chem. Int. Ed.* **2017**, *56*, 11248.

11. A. I. González Flórez, E. Mucha, D.-S. Ahn, S. Gewinner, W. Schöllkopf, K. Pagel, G. von Helden; Charge-Induced Unzipping of Isolated Proteins to a Defined Secondary Structure, *Angew. Chem. Int. Ed.* **2016**, *55*, 3295.

12. M. Kandziora, E. Mucha, S. Zucker, H.-U. Reissig; Syntheses of Mono- and Divalent C-Aminoglycosides Using 1,2-Oxazine Chemistry and Olefin Metathesis, *Synlett* **2015**, *26*, 367.