

Structural Analysis of Glycosyl Cations and Other Intermediates Using Cryogenic Infrared Spectroscopy

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy
of Freie Universität Berlin

by

Kim Greis

2023

This work was performed from November 2019 to April 2023 at Freie Universität Berlin and the Fritz-Haber-Institut der Max-Planck-Gesellschaft under the supervision of Prof. Kevin Pagel and Prof. Gert von Helden.

First Reviewer: Prof. Dr. Kevin Pagel, Freie Universität Berlin

Second Reviewer: Prof. Dr. Gert von Helden, Fritz-Haber-Institut der Max-Planck-Gesellschaft, Berlin and Radboud Universiteit, Nijmegen

Date of Defense: July 12, 2023

Statutory Declaration

by Kim Greis

I assert that I have written this thesis independently and entirely by myself using only the indicated sources and support. All quotations and citations are acknowledged. I also declare that I have not submitted the dissertation in this or any other form to another institution as a dissertation.

Berlin, May 10, 2023

KIM GREIS

Abstract

Knowing the structure of reactive intermediates can yield unprecedented insight into organic reaction mechanisms. In particular for glycosyl cations – the reactive intermediates in glycosylations – the stereoselectivity of the reaction could be predicted by knowing the structure of the intermediate. The structure reveals whether an acyl protecting group of the monosaccharide unit interacts with the positively charged anomeric carbon so that it would shield one side from nucleophilic attack and thus steer the stereoselectivity of the reaction. These postulated approaches have been termed neighboring-group and remote participation. However, the short lifetime of reactive intermediates impedes their structural characterization in solution. Hence, for glycosyl cations, the structure remained elusive until very recently. These intermediates are not intrinsically unstable, but well-defined minima on the potential energy surface. Therefore, the ionic intermediates can be generated inside the vacuum of a mass spectrometer, free from nucleophiles or solvent molecules. In this environment, the isolated intermediates are stable and can subsequently be characterized using spectrometric or spectroscopic techniques. Recent advances in instrumentation allow coupling mass spectrometers with infrared lasers for infrared ion spectroscopy. Thus, highly-resolved infrared spectra of the analyte ions can be obtained by using cryogenic infrared spectroscopy in helium nanodroplets. To assign the obtained spectrum to a structure, it can be compared to harmonic frequencies of promising candidate structures calculated using density functional theory. This workflow was successfully used to determine the structure of several glycosyl cations, based on which, a new selective building block for 1,2-*cis* galactosylations was developed and its stereoselectivity was rationalized. Furthermore, it was determined that *c*-fragments of RNA dinucleotides are identical to the intermediate of RNA autohydrolysis. Finally, potentially antiaromatic carbocations were investigated.

Zusammenfassung

Die Kenntnis der Struktur reaktiver Zwischenprodukte kann neue Einblicke in organische Reaktionsmechanismen liefern. Für Glykosylkationen – die Zwischenprodukte der Glykosylierungsreaktion – könnte die Stereoselektivität der Reaktion vorhergesagt werden, wenn die Struktur des Intermediats bekannt ist. Die Struktur zeigt, ob eine Acyl-Schutzgruppe mit dem positiv geladenen anomeren Kohlenstoff des Monosaccharids interagiert, um eine Seite vor nukleophilen Angriffen zu schützen. Diese Ansätze werden als Nachbargruppen- und Fernpartizipation bezeichnet. Die kurze Lebensdauer reaktiver Zwischenprodukte erschwert ihre strukturelle Charakterisierung in Lösung. Daher blieb die Struktur von Glykosylkationen bis vor Kurzem unklar. Die Intermediate sind jedoch nicht intrinsisch instabil, sondern definierte Minima auf der Potenzialhyperfläche. Daher können sie im Vakuum eines Massenspektrometers erzeugt werden, das frei von nukleophilen Verbindungen und Lösungsmittelmolekülen ist. In dieser Umgebung sind die isolierten Intermediate stabil und können anschließend mittels spektrometrischen und spektroskopischen Methoden charakterisiert werden. Fortschritte in der Entwicklung von Massenspektrometern ermöglichen ihre Kopplung mit Infrarotlasern, um Ionen mit Infrarotspektroskopie zu untersuchen. Daher können hochaufgelöste Infrarotspektren der Analytionen durch kryogene Infrarotspektroskopie in Heliumtröpfchen erhalten werden. Um das gemessene Spektrum einer Struktur zuzuordnen, kann es mit harmonischen Frequenzen von vielversprechenden Strukturen, die mit Dichtefunktionaltheorie gerechnet wurden, verglichen werden. Diese Methode wurde in dieser Arbeit erfolgreich eingesetzt, um die Struktur mehrerer Glykosylkationen zu bestimmen. Basierend auf den so gewonnenen Daten wurde ein neuer selektiver Baustein für 1,2-*cis* Galaktosylierungen entwickelt und dessen Stereoselektivität wurde rationalisiert. Darüber hinaus wurde festgestellt, dass *c*-Fragmente von RNA-Dinukleotiden identisch mit dem Intermediat der RNA-Autohydrolyse sind. Schließlich wurden potenziell antiaromatische Carbokationen untersucht.

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Outline of the Thesis	4
2	Fundamentals	7
2.1	Glycosylation Reaction	7
2.2	Infrared Spectroscopy of Ions in the Gas Phase	12
2.2.1	Mass Spectrometry	12
2.2.2	Infrared Spectroscopy	15
2.2.3	Infrared Action Spectroscopy	16
2.3	Computational Methods	19
2.3.1	Electronic Energies: Wave Function-Based Methods	19
2.3.2	Electronic Energies: Density Functional Theory	21
2.3.3	Empirical Dispersion Correction	22
2.3.4	Basis Sets	23
2.3.5	Optimization and Frequency Calculations	24
2.3.6	Conformational Sampling	26
3	Experimental Methods	29
3.1	Infrared Spectroscopy in Helium Nanodroplets	29
3.2	Fritz Haber Institute Free-Electron Laser	32
3.3	Computational Methods	33
4	Influence of the Electron Density of Acyl Protecting Groups on the Selectivity of Galactosylations	37
4.1	Author Contributions	38
4.2	Project Hypothesis and Summary	38
	Publication	41
5	Neighboring-Group Participation of Benzoyl Protecting Groups in Fluorinated Glucose	51
5.1	Author Contributions	52
5.2	Project Hypothesis and Summary	52
	Publication	55

6	The Ferrier Glycosyl Cation in the Gas Phase	61
6.1	Author Contributions	62
6.2	Project Hypothesis and Summary	62
	Publication	65
7	The Fate of a Septanosyl Ferrier Cation in the Gas Phase	69
7.1	Author Contributions	70
7.2	Project Hypothesis and Summary	70
	Publication	73
8	The Key Intermediate of RNA Autohydrolysis	85
8.1	Author Contributions	86
8.2	Project Hypothesis and Summary	86
	Publication	89
9	Cyclopentadienyl and Fluorenyl Cations	97
9.1	Introduction	98
9.2	Experimental Details	100
9.3	Results and Discussion	101
9.4	Conclusion	106
10	Summary and Outlook	107
	References	109
	Appendix	129
	Appendix A	129
	Appendix B	283
	Appendix C	315
	Appendix D	359
	Appendix E	409
	Appendix F	431
	Acknowledgments	IX
	List of Publications	XI

Abbreviations

Ac	acetyl
AGA	automated glycan assembly
Bn	benzyl
Bz	benzoyl
CC	coupled cluster
CID	collision-induced dissociation
CREST	Conformer–Rotamer Ensemble Sampling Tool
DFT	density functional theory
DLPNO	domain-based local pair natural orbital
ESI	electrospray ionization
FAFOOM	Flexible Algorithm for Optimization of Molecules
FEL	free-electron laser
FHI FEL	Fritz Haber Institute free-electron laser
FHWM	full width at half maximum
Fmoc	fluorenylmethoxycarbonyl
GGA	generalized gradient approximation
HF	Hartree-Fock
IR	infrared
IRMPD	infrared multiple photon dissociation
IVR	intramolecular vibrational energy redistribution
LDA	local density approximation
LINAC	linear accelerator
LRP	long-range participation
MP	Møller-Plesset
MS	mass spectrometry
MS/MS	tandem mass spectrometry
<i>m/z</i>	mass-to-charge ratio
nESI	nano-electrospray ionization

NGP neighboring-group participation

NMR nuclear magnetic resonance

PES potential energy surface

Piv pivaloyl

RI resolution of identity

RNA ribonucleic acid

TFA trifluoroacetyl

TOF time-of-flight

TS transition state

ZPVE zero-point vibrational energy

1 | Introduction

Knowing the structure of intermediates is of tremendous importance to fully understand the mechanism of a chemical reaction. However, in some cases, the structural characterization of intermediates is impeded by their short lifetime. A prominent example of such an intermediate is the glycosyl cation – the intermediate of the glycosylation reaction. Despite this challenge, a few techniques have been developed in the past decades that allow us to elucidate their structure. In this chapter, the importance of the intermediate on the mechanism as well as the limitations of previously used techniques for their characterization will be discussed. Cryogenic infrared spectroscopy will be introduced as a method that can overcome the limitations of previous techniques. Eventually the outline of this thesis is presented.

1.1 | Motivation

Determining the mechanism of a chemical reaction can be a cumbersome task. Chemical intuition often allows chemists to draw a reasonable path between reactants and products, potentially involving several intermediates and transition states (TSs). The potential energy surface (PES) is a multidimensional function that describes how the energy of molecular systems varies with respect to the atomic coordinates. Only structures that are minima on the PES have a definite lifetime. TSs are not minima on the PES but saddle points, whereas intermediates are local minima. It is theoretically possible to isolate intermediates but this endeavor is, in practice, limited by their often short lifetime.^[1] Computational methods,^[2,3] spectroscopy,^[4-8] and kinetic studies^[9] provided valuable information on reactive intermediates in the past. Directly characterizing the structure of a reactive intermediate, however, remains a challenging task. The information that could be obtained from such studies would yield tremendous mechanistic insight, leading to a better understanding of chemical reactivity.^[10]

A prominent example for a poorly understood reaction is the glycosylation reaction with its elusive intermediate, the glycosyl cation (Figure 1.1a). It was first reported by Emil Fischer more than a hundred years ago.^[11] The glycosylation reaction allows the synthesis of oligosaccharides from monosaccharide building blocks. Oligosaccharides are carbohydrates, which constitute one of the main classes of biomolecules, next to proteins, lipids, and nucleic

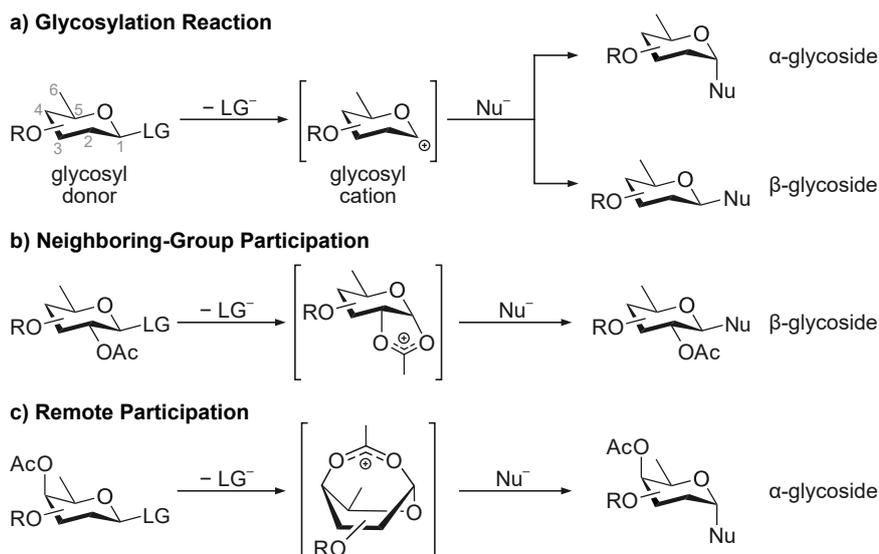


Figure 1.1: (a) Schematic glycosylation reaction. Cleavage of the leaving group (LG) leads to the formation of a glycosyl cation. With non-participating protecting groups (such as R = Bn), a mixture of α - and β -glycosides is obtained. (b) Neighboring-group participation of an acyl group at the C2-position is the gold-standard method for stereoselective 1,2-*trans* glycosylations (equaling to β -glucosylations, β -galactosylations, or α -mannosylations). (c) Remote participation is an approach in which remote acyl groups (for example located at the C4-position) interact with the positively charged anomeric carbon leading, for example, to an increase in α -selectivity in galactosylations.

acids. Carbohydrates play a role in several pivotal processes in living organisms, such as the immune response,^[12,13] fertilization,^[14–16] or protein folding.^[17,18] The carbohydrates that are synthesized *in vitro*, on the other hand, are used for pharmaceuticals,^[19–24] antigens for vaccines,^[25–28] or biomaterials.^[29–31] However, the synthesis of carbohydrates is a tedious process, as it is challenging to control the regio- and stereoselectivity.^[32,33]

This sheer complexity of carbohydrates can be illustrated best by comparing it to another class of biomolecules, *e.g.*, proteins. Out of three distinct amino acids, six different tripeptides can be built. In strong contrast, 768 different trisaccharides could be built from three distinct hexose monosaccharide building blocks.^[34] This immense number results from the sequence of the three building blocks, the regioselectivity, as there are four nucleophilic hydroxyl groups of each monosaccharide unit that could be potentially attached to the next building block, and the stereoconfiguration of three anomeric carbon atoms.

Several methods have been reported to control the sequence of monosaccharide building blocks in oligosaccharide synthesis, such as the programmable one-pot chemical synthesis,^[35–38] the preactivation method,^[39–44] and solid-phase synthesis.^[45–48] Synthesizers – instruments allowing to synthesize large oligosaccharides in an automated fashion – have been developed and improved in the past two decades. The use of such synthesizers is commonly referred to as automated glycan assembly (AGA).^[49,50] The regioselectivity in glycosylation reactions

can be controlled by using orthogonal protecting group strategies.^[51–53] The control of the stereoconfiguration of the glycosidic linkage is, however, far more challenging. A gold-standard method for obtaining 1,2-*trans*¹ glycosidic linkages is neighboring-group participation (NGP) of acyl groups located at the C2-position of a glycosyl donor.^[54]

In NGP, it is assumed that the intermediate – the glycosyl cation – plays an important role in steering the stereoselectivity of the reaction. This intermediate is formed after cleavage of a leaving group at the anomeric carbon. In glycosyl cations generated from a glucosyl donor, for example, the carbonyl oxygen from an acyl group at the C2-position would form a covalent bond with the anomeric carbon from below (the 1,2-*cis*- or α -side) so that the attack of a nucleophile (*i.e.*, a glycosyl acceptor) could only proceed from above (the 1,2-*trans*- or β -side) (Figure 1.1b). As a consequence, 1,2-*trans* glycosides are obtained as the only product.

To synthesize 1,2-*cis* glycosides, there is no gold-standard method available yet. Several promising methods such as benzylidene-directed glycosylations,^[55–57] the use of chiral auxiliaries,^[58–61] or participation of remote acyl groups^[62] have been reported to increase the content of 1,2-*cis* glycosides. However, these methods can either not be readily implemented in AGA workflows or the selectivity is not sufficient. The first results obtained as part of this thesis deal with the question how the yield of 1,2-*cis* glycosides by remote participation can be improved. In this method, which is also sometimes termed long-range participation (LRP), it is assumed that acyl groups located on one of the remote C3-, C4-, or C6-carbon atoms could form a covalent bond with the anomeric carbon in the charged glycosyl cation intermediate. Then, the 1,2-*trans* side would be shielded and attack of a nucleophile could only proceed from the 1,2-*cis* side. Specifically for galactose building blocks with an acyl group at the C4-position, a clear and consistent increase in the 1,2-*cis* selectivity has been observed in several studies.^[63–67]

To get direct insight into the mechanism of the glycosylation reaction and determine how the interaction of protecting groups with the anomeric carbon influences the stereoselectivity, it is highly important to determine the structure of glycosyl cations. Due to their short lifetimes (ca. 1 ps in aqueous solution at room temperature),^[1,68–70] it was for a long time not possible to get structural insight into this important intermediate. Therefore, it was only characterized by computational^[3,71–74] and kinetic methods.^[1,75,76] Besides, indirect information on the mechanism was gained by structural analysis of stable sideproducts.^[65,77,78] Recently, it was possible to stabilize the glycosyl cations by super acids, which allowed recording nuclear magnetic resonance (NMR) spectra *in situ*.^[79–82] Furthermore, in certain cases, it was also possible to record low-temperature (–80 to –20 °C) NMR spectra of dissolved glycosyl cations *in situ*.^[59,61,83,84]

¹The 1,2-*cis/trans* and α/β nomenclatures can both be used to describe the stereoconfiguration of glycosidic linkages. In the former, the configuration of C1 is defined directly relative to that of the C2-atom, whereas in the latter the configuration of C1 is defined relative to that of C5 in hexoses. If the configuration of C1 and C5 is not identical, it is an α -glycosidic linkage and if it is identical, then it is a β -glycosidic linkage. Both nomenclatures will be used in this work. For glucosides and galactosides, 1,2-*cis* equals to α and 1,2-*trans* to β , whereas it is the other way around in mannosides.

While these results yielded tremendous insight into the mechanism of the glycosylation reaction, it also comes with drawbacks. In the case of super acids, for example, the glycosyl cations are fully protonated, which distorts their structure and also inhibits the capacity of acyl groups to stabilize the positive charge at the anomeric carbon. Recording low-temperature NMR spectra, on the other hand, yields direct structural data on the *real* glycosyl cation. However, this method could so far only be applied to a limited number of glycosyl cations, with sufficiently long lifetimes for an NMR experiment.

In the groups of Kevin Pagel and Thomas Boltje a different approach was introduced to structurally characterize glycosyl cations, which involves a combination of mass spectrometry (MS) and infrared (IR) spectroscopy.^[85,86] Here, glycosyl donors are transferred to the gas phase using electrospray ionization (ESI). Subsequently, using in-source fragmentation or collision-induced dissociation (CID), the leaving group of the protonated or sodiated glycosyl donors is cleaved, leading to glycosyl cations. The ions can be isolated by their mass-to-charge ratio (m/z) using a quadrupole mass filter. With IR ion spectroscopy, the vibrational modes of the glycosyl cations can be probed and linked to computationally simulated structures. Thomas Boltje *et al.* are probing the ions at room temperature, whereas Kevin Pagel *et al.* are investigating the ions in superfluid helium droplets (0.4 K), close to the absolute zero temperature. Several studies from both groups were published in the past five years and show correlations between the gas-phase intermediate and the stereoselective outcome of the glycosylation reaction.^[87–93] Hence, with this method, it was possible to confirm the structure of intermediates postulated many decades earlier.

The aim of this thesis is to exploit the advantages of IR ion spectroscopy in superfluid helium droplets to get insight into the mechanism of the glycosylation reaction and to rationalize the selectivity of glycosyl donors that are used in glycan synthesis. Eventually, the gained knowledge will lead to a deeper understanding of the glycosylation reaction and potentially to the development of stereoselective building blocks for 1,2-*cis* glycosylations. Furthermore, the ability of the method to probe other intermediates, as the intermediate of ribonucleic acid (RNA) autohydrolysis, fluorenyl cations, and cyclopentadienyl cations will be investigated.

1.2 | Outline of the Thesis

This work deals with the structural characterization of several intermediates, which are relevant for understanding organic reaction mechanisms, such as glycosyl cations. An in-depth understanding of these structures can be generated through the combination of cryogenic IR spectroscopy and computational methods. In **Chapter 2** the fundamentals of the glycosylation reaction, mass spectrometry, infrared ion spectroscopy, and computational methods are introduced. Details on how the experiment was performed and on the used computational methods are provided in **Chapter 3**. In **Chapter 4** it is shown how the stereoselectivity of glycosylation

reactions can be tuned by implementing electron-donating and -withdrawing substituents in remote acyl protecting groups. The structures and relative stabilities of the resulting glycosyl cations are studied. In **Chapter 5**, the influence of fluorine substituents in glycosyl cations with differential protecting groups are studied. Then, the structure of the Ferrier glycosyl cation is experimentally characterized in **Chapter 6**. In **Chapter 7** the influence of a larger sugar ring size on the structure of the resulting septanosyl Ferrier cation is investigated. **Chapter 8** explores RNA autohydrolysis and its intermediate in the gas phase. In **Chapter 9**, cryogenic IR spectroscopy is used to investigate small carbocations that may exhibit antiaromatic properties. Finally, this thesis is concluded and an outlook is given in **Chapter 10**.

2 | Fundamentals

In this chapter, the background knowledge required for a deeper understanding of this thesis will be described. The first section will be dedicated to the glycosylation reaction and its intermediate, the glycosyl cation, which is at the center of this work. The second and third sections will deal with the methods that are used in this work to investigate reactive intermediates: cryogenic infrared spectroscopy and computational methods.

2.1 | Glycosylation Reaction

Controlling the stereoselectivity of the glycosylation reaction is one of the most challenging aspects in organic chemistry. Although the first glycosylation reaction was already reported in 1893 by Emil Fischer,^[11] the mechanism of this reaction is still not fully understood today.^[40,94,95] In this reaction, a monosaccharide building block carrying a leaving group (glycosyl donor or electrophile) at the anomeric carbon (C1) is attacked by a nucleophile (acceptor). This reaction leads to an anomeric mixture of α - and β -glycosides in most cases (Figure 1.1a).^[45,88,96,97]

To ensure that the hydroxyl groups of a glycosyl donor do not act as nucleophiles during a glycosylation reaction, all hydroxyl groups are usually protected (for example with benzyl (Bn) protecting groups). Furthermore, to build oligosaccharides, the nucleophile needs to be a glycoside as well. Regioselectivity is achieved when all hydroxyl groups are protected, except for the one that is supposed to attack the glycosyl donor. To attach a third glycosyl donor to the scaffold, one hydroxyl group on the scaffold needs to be selectively deprotected. In order to perform this step, orthogonal protecting groups need to be used.^[51,53,98-100] This strategy involves at least two distinct protecting groups that can be cleaved using different conditions. For example, the protecting groups Bn and fluorenylmethoxycarbonyl (Fmoc) (structures in Figure 2.1) can be removed by hydrogenolysis and under basic conditions, respectively.^[53,101] With this strategy, after coupling one glycosyl donor to a nucleophile, this donor can be selectively deprotected, yielding a hydroxyl group that acts as the nucleophile for the next glycosyl donor. This concept is especially useful in solid-phase synthesis of sugars.^[45-47] Here, the glycosyl donors are linked to a nucleophilic linker (nucleophile attached to a solid phase), all unreacted nucleophiles are capped, and unreacted donors are washed away. Then, the terminal monosaccharide on the

2.1. Glycosylation Reaction

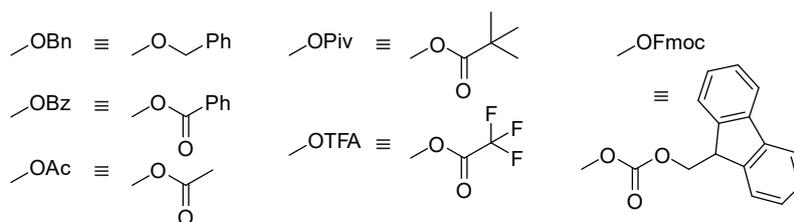


Figure 2.1: Structures of protecting groups commonly used in glycosylations, which are relevant for this work.

solid support is selectively deprotected and reacts with the next glycosyl donor. This cycle can be repeated numerous times and can lead to polysaccharides of 100 repeating units.^[48] AGA is the fully automated version of this process. Further commonly used methods to synthesize oligosaccharides are the preactivation method^[39–43] and the programmable one-pot chemical synthesis.^[35–38]

While the regioselectivity is well-controlled in glycosylation reactions, control of the stereoselectivity is more challenging. This drawback arises from the complexity of the mechanism of the glycosylation reaction that impedes in-depth understanding. Several factors, such as the temperature, the solvent, as well as the reactivity of the glycosyl donors and acceptors impact the stereoselectivity.^[96,97,102–109] It is assumed that the reaction is governed by an S_N1 – S_N2 continuum (Figure 2.2).^[94] Depending on several factors, such as those previously mentioned, the reaction is more inclined towards one side of the continuum. For example, strong nucleophiles tend to react *via* an S_N2 mechanism, whereas weak nucleophiles prefer to react *via* an S_N1 mechanism.^[102] However, each glycosylation reaction is unique and it is the interplay of many different factors that will ultimately steer its preference towards one or multiple distinct mechanistic pathways.^[94]

For glycosylation reactions located on the S_N2 side, one might postulate that nucleophiles simply attack the anomeric carbon leading to the cleavage of the leaving group and stereo-inversion (Walden inversion) of the anomeric carbon. However, the mechanism is more complex. Crich et al. emphasized the important role of triflates during the glycosylation reaction.^[40,55,110,111] Triflic acid is used as an activator in many glycosylation reactions. As such, it can replace the leaving group leading to the formation of covalent intermediates or contact-ion pairs.^[112] It is not understood in great detail, which type of intermediates are formed and what their influence on the stereoselectivity of the reaction is. In certain cases, it was possible to study triflate intermediates *via* NMR spectroscopy.^[113–115] Based on the reported results, triflates seem to be mainly present as covalent intermediates and react with nucleophiles according to an S_N2 mechanism. These reports show that considering triflate intermediates is of tremendous importance for understanding the S_N2 side of the glycosylation reaction.

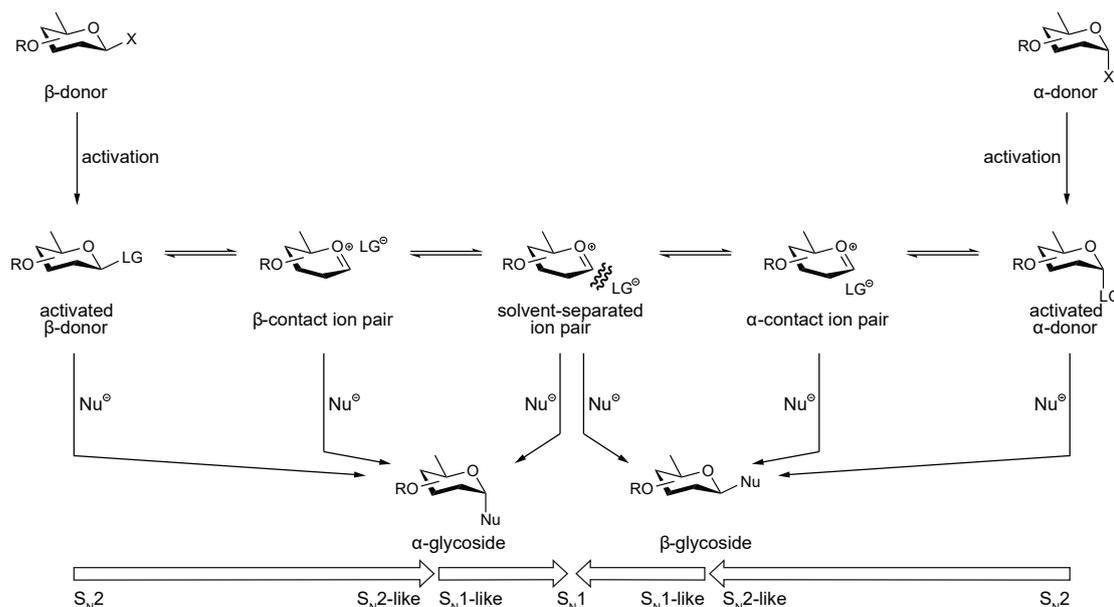


Figure 2.2: Schematic depiction of the glycosylation reaction governed by an S_N1 – S_N2 continuum. Adapted with permission from Adero et al.^[94] Copyright 2018 American Chemical Society.

On the S_N1 side, a carbocationic intermediate – the glycosyl cation – is generated.^[80,116,117] This ion is either directly generated after cleavage of the leaving group or after cleavage of a triflate that might temporarily be connected to the anomeric carbon. Here, it is important that a solvent-separated ion pair is formed, where formally no interactions between the glycosyl cation and the leaving group occur anymore. The charge at the anomeric carbon is stabilized by a mesomeric +M effect of the neighboring oxygen atom, leading to a delocalization of the charge between the two atoms.^[118,119] In this case, the glycosyl cation can also be termed oxocarbenium ion. A nucleophile could attack the oxocarbenium ion from both sides. Statistically, a 1:1 mixture of 1,2-*trans* and 1,2-*cis* glycosides could be expected as products. However, empirical evidence suggests that the mechanism is more complex and that there are multiple factors, such as the conformation of the oxocarbenium ion, its substituents, and the solvent, pushing the selectivity to one side or the other.^[63,96,120]

For many decades, neighboring-group participation, which is an effect that requires S_N1 -type glycosylations, has been used to yield 1,2-*trans* glycosidic bonds.^[78,121–123] Here, a participating acyl protecting group, such as acetyl (Ac) or benzoyl (Bz), at the C2-position is postulated to interact with the positively charged anomeric carbon after cleavage of the leaving group. During this interaction, the 1,2-*cis* side of the glycosyl cation would be shielded, so that nucleophiles could only attack from the 1,2-*trans* side, leading to complete stereoselectivity in many cases (Figure 1.1b). This method is well-established and the mechanism generally accepted across the community.^[100,124] Contrary to participating protecting groups, non-participating protecting

groups, such as Bn or methyl (Me), are believed to have less influence on the stereoselectivity of a glycosylation reaction.^[63]

In contrast to the well-established method to generate 1,2-*trans* glycosides, a gold-standard method to generate 1,2-*cis* glycosides is not yet available. Several methods have been reported that increase selectivity for 1,2-*cis* glycosides, such as 4,6-*O*-benzylidene-directed glycosylations, chiral auxiliaries, or the participation of remote acyl groups (remote participation, Figure 1.1c).^[56,58,63,125,126]

The benzylidene group is an acetal protecting group that protects two neighboring hydroxyl groups of a sugar at the same time. Upon protection, 4,6-*O*-benzylidene sugars are the major product.^[127] Benzylidene-directed glycosylations are mainly known for consistently yielding high amounts of 1,2-*cis* (β -) mannosides. Here, it is postulated that during activation, highly reactive α -triflates are formed that react with nucleophiles in an S_N2 fashion to yield β -mannosides (Figure 2.3a).^[57,111] With the same rationale one would also obtain β -glucosides and -galactosides, which are 1,2-*trans* glycosides. However, with these sugars, the selectivity is not as high as for mannose.^[128] Furthermore, these linkages can be straightforwardly generated *via* NGP as well. Using similar types of protecting groups, such as 4,6-*O*-phenylborate, a high selectivity for β -mannosylations was observed.^[128] The related di-*tert*-butylsilylene protecting group can be used for 1,2-*cis* (α -) selective galactosylations.^[129,130] A major drawback of this method is that the glycosyl donors cannot be readily used in automated workflows, because the acetal protecting group would yield two nucleophilic hydroxyl groups upon deprotection.

Chiral auxiliaries are substituents at the C2-position on a glycosyl donor that carry a chiral carbon atom. Attached to this carbon atom are a bulky group (*e.g.*, phenyl), which prevents free rotation of the auxiliary, and a substituent carrying a nucleophilic atom, such as a sulfur atom^[59] or a carbonyl oxygen.^[58] In the pendant glycosyl cation, the bulky substituent forces the nucleophilic group to interact with the positively charged anomeric carbon either from above or below the mean plane of the sugar, depending on the chirality of the carbon atom in the auxiliary. As a consequence, the 1,2-*cis* or the 1,2-*trans* side is blocked, so that a nucleophile would attack from the other side (Figure 2.3b). Several studies have shown that this concept can be used to selectively yield 1,2-*cis* glycosides for glucosides, galactosides, and mannosides.^[60,61,131–134] The use of chiral auxiliaries is laborious, potentially leading to an overall decrease in yield. Furthermore, their removal is in certain cases not orthogonal to the often used Bn groups and their steric demand might hamper subsequent glycosylations, impeding their use in AGA workflows.^[63,135]

A further promising and versatile approach to generate 1,2-*cis* glycosides is remote participation. The principle of this approach is similar to that of NGP. An S_N1 mechanism is assumed in which a glycosyl cation is formed. Then the positively charged anomeric carbon is stabilized by a remote acyl group located at the C3-, C4-, or C6-carbon atom of the glycosyl donor. In galactose, for example, an acyl group at the C4-position could shield the 1,2-*trans* side of the

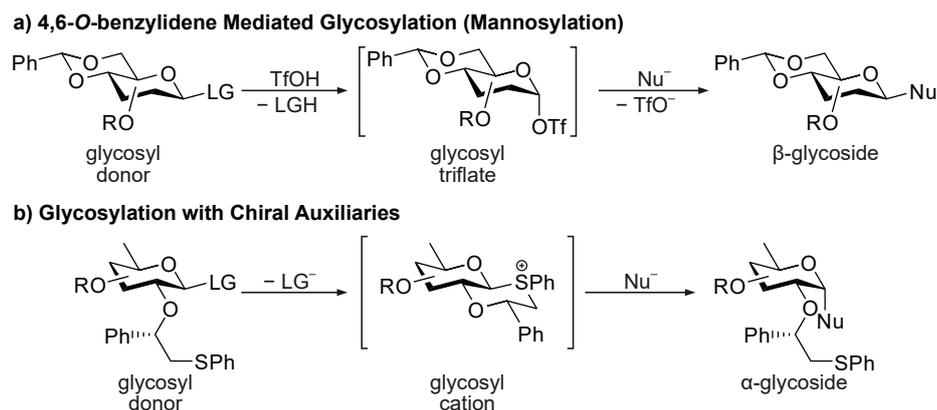


Figure 2.3: (a) Schematic depiction of a 4,6-*O*-benzylidene mediated glycosylation. A leaving group is replaced by an α -triflate that is highly stereoselective. This methodology efficiently generates β -mannosides. (b) Glycosylation reactions using a chiral auxiliary at C2. In the glycosyl cation, the auxiliary interacts with the anomeric carbon so that attack from a nucleophile could only proceed from the α -side.

glycosyl cation. Subsequently, the attack of a nucleophile from the 1,2-*cis* side would be favored (Figure 1.1c). Especially for galactose, a clear and consistent trend has been observed that an acyl group at the C4-position enhances the 1,2-*cis* selectivity.^[63–67] Generally, however, the role of acyl groups at both the C3- and the C4-positions is not fully understood.^[63,88] Furthermore, evidence suggests that an acetyl group at the C6-position would even have adverse effects on the 1,2-*cis* selectivity.^[64] Although this technique is very promising, especially for 1,2-*cis* galactosylations, the stereoselectivity has so far not been high enough to readily implement it in AGA workflows. Here, total stereoselectivity is required to ensure a high yield of the desired product after several coupling steps.

Due to the short lifetime of glycosyl cations of picoseconds,^[1] it was for a long time not possible to get direct evidence on the mechanism. Nonetheless, tremendous mechanistic evidence on the glycosylation reaction was gained by determining kinetics^[57,76,136] or characterizing stable sideproducts.^[65,77,78] The latter confirms that neighboring and remote acyl groups can interact with the positively charged anomeric carbon in a glycosylation reaction. In one study, a sideproduct diagnostic for the participation of a C4-Bn group was characterized.^[61] It was also possible to study both triflate intermediates^[113–115] and glycosyl cations stabilized by chiral auxiliaries^[59,61,137] using NMR spectroscopy at low temperatures (ca. $-20\text{ }^{\circ}\text{C}$). Furthermore, glycosyl cations can be stabilized in super acids and subsequently be probed by NMR spectroscopy.^[79–82] Here, evidence for NGP could be found, whereas remote participation did not seem to occur. However, in super acids the nucleophilicity of carbonyl groups and thus their ability to stabilize the positively charged anomeric carbon is decreased, as all carbonyl groups are fully protonated in this medium. In the last two years, two studies were published on

systems, in which it was possible to study galactosyl and mannosyl cations exhibiting C3-acyl^[84] and C4-acyl^[83] remote participation, respectively, using NMR spectroscopy at ca. $-80\text{ }^{\circ}\text{C}$.

Recently, a different approach was established to characterize the structure of glycosyl cations. It was determined that glycosyl donors can be ionized by ESI coupled to a mass spectrometer, leading to the formation of protonated and sodiated glycosyl donors. By in-source fragmentation or CID of the ionized donors, the leaving group can be cleaved, leading to the glycosyl cation intermediate. *In vacuo*, where solvent molecules and nucleophiles are absent, glycosyl cations remain stable and can be directly probed with IR ion spectroscopy.^[85] This technique allows to probe the vibrational modes of *m/z*-selected ions. The acquired IR spectrum can be compared to computed spectra to assign the structure. With this method, the structure of glucosyl, galactosyl, and mannosyl cations exhibiting NGP^[85,86] and remote participation of C3- and C4-acetyl groups has been elucidated.^[87-89,138] Furthermore, it was determined that a 2,2-dimethyl-2-(*ortho*-nitrophenyl)acetyl (DMNPA) protecting group at C6 can stabilize the glycosyl cation.^[91] In uronic acids, an interaction between the positively charged anomeric carbon and the carbonyl oxygen of the methylester at C5 was observed.^[90]

Despite tremendous mechanistic investigations in the past decades, the glycosylation reaction is still far from being completely understood. The past studies made clear that it is an extremely complex reaction and that it is unlikely that there is one single mechanism that describes all glycosylations. The use of unconventional methods, such as ion spectroscopy or NMR in super acids, has proven beneficial for a deeper understanding of the mechanism.

2.2 | Infrared Spectroscopy of Ions in the Gas Phase

Recording IR spectra of ions in the gas phase requires in-depth knowledge on both MS and IR spectroscopy. Therefore, relevant aspects of both techniques will be introduced, before IR action spectroscopy and its variations are discussed.

2.2.1 | Mass Spectrometry

A mass spectrometer allows to analyze ions according to their mass-to-charge ratio. Several developments in the past decades helped MS ascend to be one of the most important analytical techniques. These improvements include broad applicability,^[139,140] versatility,^[141-145] availability of commercial instruments, as well as ease of use.^[146,147] A few of its many applications are identification and quantification of compounds,^[148,149] determining the structure and sequence of biomolecules,^[150-155] or to generate and isolate ions that are otherwise not stable.^[156-159] In its simplest form, a mass spectrometer consists of an ion source, a mass analyzer, and a detector. Most importantly, the technique relies on manipulating and detecting charged particles, *i.e.*, ions. Therefore, every sample needs to be ionized before it can be measured with MS.

ESI is one of the most commonly used techniques to generate and transfer ions into the gas phase.^[160] It was developed by John B. Fenn in the late 1980s, which was awarded with the Nobel Prize in 2002.^[161,162] ESI is a gentle ionization technique that does not necessarily require a large sample volume or high purity. MS stands in strong contrast to NMR spectroscopy, where purity is an important requirement, as well as sample amounts that are at least two to three orders of magnitude higher.^[163]

The analyte solution is introduced into the mass spectrometer through a needle, to which a voltage of a few kV is applied.^[161] This leads to the formation of highly charged droplets containing solvent and analyte molecules. Subsequently, bare ions are generated by a mechanism that is dependent on the sample. Small ions favor the ion evaporation model, whereas larger, globular ions favor the charge residue model.^[164] The generated ions, which are usually protonated, sodiated, or deprotonated, are guided into the mass spectrometer. Depending on the system, they can carry multiple charges. Furthermore, molecules of a wide spectrum of sizes can be ionized with this method, ranging from a few Daltons up to the range of Megadaltons, as observed for intact viruses.^[165] The molecular ions investigated for this thesis are comparably small and only carry one charge. A drawback of ESI is that non-polar molecules do not ionize very well.^[166]

A modification of ESI that was used in this thesis is nano-electrospray ionization (nESI). Here, a disposable metal-coated glass capillary is used, with the emitter tip only having a few μm in diameter. The sample consumption can be reduced from $\mu\text{L min}^{-1}$ to nL min^{-1} , and lower needle voltages (0.8 – 1.0 kV) are required for ionization. Furthermore, as the capillaries are disposable, the contamination from previous samples can be kept at a minimum level.^[167]

The generated ions are guided into vacuum, where they can be separated according to their m/z with a mass analyzer. One commonly used mass analyzer that combines fast acquisition times, a broad mass range, and a high resolution is the time-of-flight (TOF) mass analyzer.^[168] In TOFs, ions with the charge z are accelerated in an electric field with a potential difference U . Thereby, all ions obtain an equal amount of potential energy. Due to the acceleration of the ions in the electric field, the potential energy is converted to kinetic energy. The velocity of an ion is equal to the length of the flight path l in the TOF cavity over the measured time of flight t required to traverse this path. Generally, the flight time is proportional to the square-root of m/z :

$$t = \frac{l}{\sqrt{2U}} \sqrt{\frac{m}{z}}. \quad (2.1)$$

This equation shows that m/z can be calculated based on values that are known from the experiment.

The resolution of a TOF mass analyzer can be significantly enhanced by using a reflectron. A reflectron consists of a set of grids, to which a voltage is applied, leading to the reflection of ions. A reflectron increases the time of flight of the ions, however, this is not the main source of

increase in resolution. Depending on the initial potential energy of the ions, they penetrate the reflectron more or less deep. This innovation reduces the inhomogeneities in the energy distribution of the ions and ultimately leads to more confined ion packets.^[169] Eventually, the ions are detected using a microchannel plate detector. This detector amplifies the signal of an ion by secondary emission of multiple electrons that can subsequently be measured. The obtained signal is plotted against the measured time of flight, which can be converted to m/z after calibration, leading to a mass spectrum.

Depending on the scientific question to be answered, mass spectrometers can be heavily modified. In the following paragraphs, the processes and parts of a mass spectrometer that are important for obtaining the results of this thesis will be introduced.

Although ESI is a soft ionization method, fragmentation processes are desirable under certain circumstances. For example, MS alone cannot discriminate isomeric ions, but their fragment ions can contain diagnostic information that allows differentiating the parent ions.^[170–172] Therefore, various techniques have been introduced that allow the fragmentation of ions. A popular technique is CID. Here, ions are guided into a collision cell that is filled with an inert collision gas, such as helium, nitrogen, or argon. Collisions between the analyte ion and the collision gas leads to the cleavage of the weakest bond.^[173,174] CID does not necessarily need to occur in a dedicated collision cell, but can also occur in the source region of an instrument, which is a process called in-source fragmentation.^[175,176] Before entering the high vacuum of a mass spectrometer, ions have to pass through multiple differential pressure regions. In these regions collisions with residual air and solvent molecules might occur and can be increased by accelerating the ions. As in CID, this leads to fragmentation of the weakest bond. This technique has been used to generate intermediates, such as glycosyl cations, in the gas phase.^[85,87]

For certain experiments, it is desirable that only ions with a specific m/z are probed. However, especially after using a fragmentation method, ions of multiple m/z may populate the mass spectrum. Furthermore, certain samples do not contain a pure analyte, also leading to heavily populated mass spectra. In these cases, a quadrupole mass filter can be used to select a specific m/z -value.^[177,178] A quadrupole is composed of four cylindrical rods arranged in square shape with negative and positive polarities, while opposing rods have the same polarity. With the application of a combination of DC and RF voltages to the rods, an electrical field is generated in which the resulting trajectory of the ions can be described by the Mathieu/Hill equations.^[179] The voltages can be tuned such that only ions with a distinct m/z -value have a stable trajectory and can pass the quadrupole. Similar to quadrupole mass filters are quadrupole benders. Here, hyperbolically-shaped rods are used to generate a field in which an ion beam can be bent by 90°. Furthermore, various linear ion traps using a similar design are used in physicochemical research labs, such as hexapoles, octopoles, or even 22-pole ion traps. They can be used to trap ions with a distinct potential depending on the number of rods.^[180,181]

2.2.2 | Infrared Spectroscopy

The wavelength of IR light lies between 0.75–1000 μm . Conventionally, the range of an IR spectrum is indicated in cm^{-1} , leading to a range from 10 to 33333 cm^{-1} . When molecules interact with resonant mid-IR photons (400–4000 cm^{-1}), their vibrational modes are excited.^[182] Hence, IR spectroscopy is a form of vibrational spectroscopy. However, only vibrations leading to a change in the dipole moment of the molecule can be excited by IR light. Except for symmetric molecules, this limitation does not have big implications. Most molecules have $3N - 6$ vibrational degrees of freedom, with N being the number of atoms. Three degrees of freedom are occupied by rotation and translation each. The only exception are linear molecules that only exhibit two rotational degrees of freedom and therefore have $3N - 5$ vibrational degrees of freedom.^[183]

In quantum mechanics, the model of the harmonic oscillator was used to describe the vibrational behavior of a chemical bond, assuming that it is similar to a spring. This model, which uses a quadratic function to fit the potential of a chemical bond, is a good approximation in the ground state of a molecule. However, it most prominently fails to describe the dissociation of a chemical bond. The dissociation is more correctly described by the Morse potential. The model incorporates quantized vibrational states that are equidistant in the case of the harmonic oscillator. However, the energy between vibrational states becomes smaller with rising excitation, a phenomenon known as anharmonicity bottleneck.^[184] As a consequence, a molecule that is singly vibrationally excited by a photon cannot be excited to its second vibrational state by a photon of the same energy.

The vibrational frequency of a covalent bond is dependent on the bond order, the involved elements, and the chemical environment. Therefore, functional groups exhibit vibrations in a specific wavenumber region and almost every molecule has a distinct IR pattern. As a consequence, isomeric molecules that cannot be differentiated by MS may be distinguishable by IR spectroscopy. In the 2600–4000 cm^{-1} region, C–H, N–H, and O–H stretching can be observed. However, this region is not accessible with the setup used in this thesis. Therefore, the 1000–1800 cm^{-1} region (for Chapter 9: 600–1700 cm^{-1}) was measured. Within 600–1400 cm^{-1} C–C and C–O stretches, C–H bends, as well as symmetric and antisymmetric stretches of the O=P–O[−] moiety in phosphates can be observed. Although this region has provided diagnostic information for certain ions investigated in this work, the diagnostic content of the functional group region (1400–1800 cm^{-1}) is often higher. Here, diagnostic symmetric and antisymmetric dioxolenium (O–C=O⁺), carbon-carbon double bond (C=C), oxocarbenium (C=O⁺), and carbonyl (C=O) stretches can be localized.

In a typical IR experiment, the attenuation of IR light with an initial intensity I_0 and the wavelength λ passing through a sample is measured (Figure 2.4a). The transmission of light traversing through a sample with the concentration (or number density) n , path length L and the

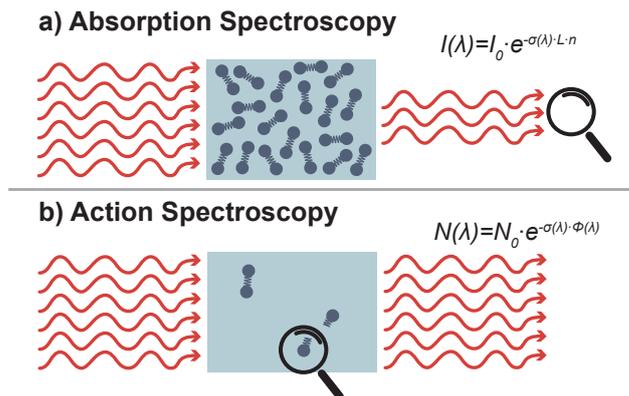


Figure 2.4: Comparison of the working principles of absorption and action spectroscopy. (a) In absorption spectroscopy, the light that is transmitted after interaction with the sample is measured. (b) In action spectroscopy, the light interacts with the analyte ion, inducing an action event, such as the fragmentation of a bond, which can be subsequently detected. Hence, the influence of the light on the analyte ion is measured. This event can be used to probe the IR spectrum of an ion. Adapted with permission from Grabarics et al.^[189] Creative Commons Attribution 4.0 International License.

wavelength-dependent absorption cross section $\sigma(\lambda)$ can be described with the Beer-Lambert law^[185]:

$$I(\lambda) = I_0 \cdot e^{-\sigma(\lambda) \cdot L \cdot n}. \quad (2.2)$$

All the light that is not transmitted is absorbed. Absorption IR experiments can be performed on any state of matter, as long as the density of the analyte is high enough. However, for ions in the gas phase, the density is limited by Coulomb-charge repulsion, leading to a limitation of ca. $10^6 - 10^7$ ions per cm^3 .^[186,187] Hence, it would not be possible to measure an absorption inside a mass spectrometer. To overcome this limitation, a few methods have been developed to record IR spectra that do not rely on detecting the absorption of IR photons by the sample but on the effect of IR photons on the sample (Figure 2.4b).^[188] The umbrella term for these methods is *infrared action spectroscopy*.

2.2.3 | Infrared Action Spectroscopy

Due to the previously discussed limitations, ions cannot be probed by using IR absorption spectroscopy in the gas phase. Nevertheless, IR photons still have an effect on irradiated ions. Several methods use IR action spectroscopy to measure that effect in order to record IR spectra of gaseous ions. First, the Lambert-Beer law needs to be rewritten for this process to:

$$N(\lambda) = N_0 \cdot e^{-\sigma(\lambda) \cdot \Phi(\lambda)}. \quad (2.3)$$

This equation shows the number of unaffected ions $N(\lambda)$ after irradiation with IR photons and the photon fluence $\Phi(\lambda)$. The effect of radiation on gaseous ions can be leveraged by multiple techniques. The most widely used technique is infrared multiple photon dissociation (IRMPD) spectroscopy (Figure 2.5a).^[184,190–192] Here, the absorption of multiple resonant photons ultimately leads to fragmentation of the analyte ion. Plotting the fragmentation yield against the IR wavenumber leads to an IR spectrum.

On a microscopic scale, a distinct resonant IR photon excites a vibration to its first excited vibrational state. Due to the anharmonicity bottleneck, a second resonant IR photon of the same wavenumber cannot excite the ion further to its second excited vibrational state. Instead, a fast process called intramolecular vibrational energy redistribution (IVR) occurs. In this process, the vibrational energy is redistributed over the whole molecular ion into its vibrational background states within picoseconds.^[193] The exact amount of time is dependent on several factors, such as the size of the molecule. Subsequently, although the molecule has an elevated thermal energy, it returns to its vibrational ground state. For the laser parameters that are commonly used for IR action spectroscopy, it is assumed that IVR is faster than the absorption of a second photon. After multiple absorption and IVR cycles, the internal energy of the molecular ion is so high that fragmentation occurs. This event can be monitored by detecting the difference in ion signal with mass analyzers, such as a TOF mass analyzer, and subsequent ion detection. Monitoring the photofragmentation yield as a function of the IR wavenumber yields an IR spectrum.

Although IRMPD spectroscopy has yielded tremendous insight into the structure of biomolecules,^[194–196] fragmentation mechanisms,^[197–199] or analytical chemistry,^[200,201] it comes with inherent drawbacks. In a typical IRMPD experiment, the analyte ions are probed at room temperature. Subsequent thermal activation of the ions and anharmonic coupling lead to line broadening and shifts to lower wavenumbers of the absorption bands.^[189,202] Especially in the case of flexible biomolecules, multiple conformers might be simultaneously populated at room temperature. Two conformers would have a similar but not exactly the same IR pattern and therefore the IR bands would overall get broader. This phenomenon would also lead to an overall decrease in intensity. Due to these effects, IRMPD spectra can be congested, especially for biomolecular ions.^[203–205]

The drawbacks of IRMPD spectroscopy are mitigated in cryogenic vibrational spectroscopy. It is used in several research groups in the form of messenger-tagging spectroscopy (Figure 2.5b).^[188,206] Here, the analyte ion is stored in a cryogenic ion trap (4–70 K). At these low temperatures, the analyte ion can form van der Waals complexes with a neutral messenger that is seeded into the helium buffer gas, such as H₂, N₂, or Ar.^[207,208] In certain cases, such as for analytes containing a metal ion, helium can also be used as a tag.^[209] Ideally, the tag is only weakly bound to the ion and does not influence its structure. The absorption of a single resonant photon can induce the dissociation of the tag.^[210] Hence, a linear IR spectrum of the tagged ion can be obtained by monitoring the dissociation yield of the ion-neutral complex. As the ions

2.2. Infrared Spectroscopy of Ions in the Gas Phase

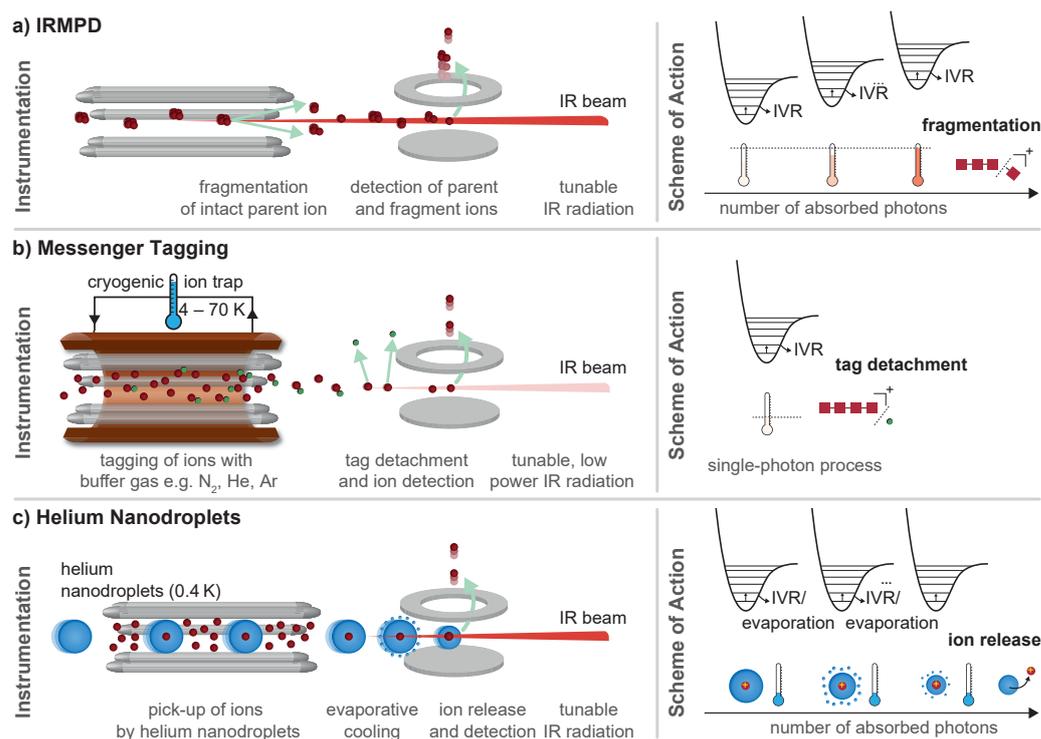


Figure 2.5: Comparison of three different types of IR spectroscopy of ions. (a) In IRMPD spectroscopy, the absorption of multiple resonant IR photons by the analyte ion leads to its fragmentation. With the absorption of every photon, the internal energy of the ion rises, ultimately leading to its fragmentation. (b) In messenger-tagging IR spectroscopy, weakly bound van-der-Waals complexes of the analyte ion with a messenger tag (e.g., H_2 , N_2 , or Ar) are formed. The detachment of this tag ideally occurs upon the absorption of a single photon. Therefore, ion heating is reduced. (c) In helium nanodroplet IR spectroscopy, the analyte ions are embedded in superfluid helium nanodroplets (0.4 K). After the absorption of multiple resonant IR photons, the ions are released from the droplets. Here, the helium droplet suppresses heating of the ion and keeps its temperature at 0.4 K. The photofragmentation or -release in the three methods can be monitored to yield an IR spectrum. Adapted with permission from Grabarics et al.^[189] Creative Commons Attribution 4.0 International License.

are probed at low temperatures and their dissociation occurs after the absorption of only one photon, no excessive ion heating occurs. Therefore, the ion is probed close to its ground state. However, the influence of the tag on the structure of the ion is not always clear.^[211,212]

In this thesis, a different approach was chosen to generate cryogenic IR spectra. Here, the analyte ions are probed using cryogenic vibrational spectroscopy in helium nanodroplets (Figure 2.5c). Helium is an element with unique properties, as it is the only element that remains liquid at 0 K under atmospheric pressure. It liquefies at 4.2 K at ambient pressure and in the case of 4He undergoes a phase transition at 2.18 K from He I to superfluid He II.^[213] The change in phase can be determined by the high heat conductivity and extremely low viscosity of He II.^[214,215] Superfluid helium nanodroplets with an internal equilibrium temperature of

0.37 K can be generated by the release of pressurized ^4He into vacuum at low temperatures (ca. 20 K).^[216–218]

Helium droplets can encapsulate neutral and charged particles *via* collisions. By traversing a filled ion trap, superfluid helium droplets collide with ions, which will then be picked up by the droplet.^[217,219,220] Subsequently, the encapsulated ion is rapidly cooled down to the intrinsic temperature of the droplets by evaporative cooling ($10^{10} - 10^{11} \text{ K s}^{-1}$) enabled by the high heat capacity of superfluid helium.^[221] Furthermore, superfluid He is transparent for IR light and its low viscosity allows the ions to freely undergo rotation and vibration inside the droplet.^[216,222] When the ion is vibrationally excited, similar to IRMPD, relaxation occurs *via* IVR. However, the energy is not stored within the molecular ion, but is dissipated to the helium matrix, which subsequently shrinks to maintain its equilibrium temperature. This process is repeated multiple times until the ion is released from the droplet, an event that can be monitored by a TOF mass analyzer. The exact mechanism of ion release from the droplet is not fully understood, although a non-thermal ejection mechanism seems to be most likely.^[223] Eventually, the signal of released ions is plotted against the IR wavenumber to yield an IR spectrum.^[224,225]

2.3 | Computational Methods

Although this thesis is experimental in nature, in most cases the interpretation of the experimental results requires results from computational chemistry. Therefore, this section aims to give a brief overview of computational chemistry and the methods used to generate the results for this thesis. These results include sampling of the conformational space of flexible molecules, computing accurate electronic energies using density functional theory (DFT), localizing and optimizing minima and in some cases saddle points along the PES of a system, and generating harmonic frequencies.

Nowadays, several quantum chemistry software packages are accessible to non-theoretical chemists. However, it is necessary to respect a few guidelines to generate meaningful results.^[226] The choice of the software package is mainly dependent on the scientific problem that needs to be solved but also simply on which software is available on the institutional high-performance computing resources. Details on the methods and software that were used in this work can be found in section 3.3.

2.3.1 | Electronic Energies: Wave Function-Based Methods

Ab initio quantum chemical methods allow to compute the intrinsic energy of a defined molecular system using the Schrödinger equation. Using these methods, the electronic energy can be calculated without any empirical knowledge on the investigated system. Generally, the Schrödinger equation is time-dependent. However, the investigated properties for this thesis do

not include time-dependent reaction dynamics or photochemical processes, which is why the time-independent Schrödinger equation,

$$H(r)\Psi(r) = E(r)\Psi(r), \quad (2.4)$$

can be used instead.^[227,228] Here, H is the Hamiltonian, Ψ is the wave function, E the energy, while r denotes the position vector of all electrons. The Hamiltonian is a mathematical operator that describes the relationship between the wave function and the electronic energy. It can be split into the kinetic and potential energy operators. A fundamental assumption that needs to be made in order to render quantum chemical treatment of molecules computationally feasible is the Born-Oppenheimer approximation.^[229] As the electrons move much faster than the nuclei, it is assumed that the velocities of both are independent from each other. This approximation allows separating the kinetic energy operators of the nuclei and electrons. As a consequence, the electronic Schrödinger equation can be solved, while the nuclei are treated as point charges that are fixed in space. The electronic Schrödinger equation can only be solved analytically for systems with no more than one electron, such as H, H₂⁺, or He⁺. For larger systems, the Schrödinger equation needs to be approximated numerically.

One of the big challenges in quantum chemistry is to account for the interactions between electrons. The Hartree-Fock (HF) theory^[230,231] was the first widely-used method to solve the Schrödinger equation using electronic interactions. The fermionic behavior of electrons and the resulting antisymmetry of their wave function can be described by arranging the orbitals in a Slater determinant. The HF orbitals are the orbitals with the lowest energy and can be determined using the variational principle. The HF wave function is the best single-determinant wave function. However, the HF theory does not include electron correlation. Proper treatment of such interactions would, for example, require a wave function composed of multiple determinants, which is computationally much more demanding than HF. It is important to note that using more determinants would systematically improve the obtained results.

Although electron correlation methods are less prominent in this thesis, two of these methods should be briefly mentioned for the sake of completeness. The first is based on Møller-Plesset (MP) perturbation theory.^[232] In the widely-used MP2 method, the correlation energy is described as perturbation of the HF energy. Perturbation theory also allows to study higher degrees of electron correlation, leading to an unfavorable increase in computational cost. MP2 is considered the most economical method to calculate electron correlation.^[233] Another systematic improvement of HF is derived from coupled cluster (CC) theory.^[234] Here, all interactions conceivable between a given number of electrons are considered. Furthermore, all possible excited determinants are considered. Due to computational limitations, truncated coupled cluster methods are commonly employed, which restrict the calculation to singly-, doubly-, and triply excited determinants. This method is called CCSDT. The computation of

triply excited states is in many cases hardly feasible, which is why the the triples contribution can be approximated as perturbation to the CCSD energy, leading to CCSD(T), a widely used electronic structure method considered as a gold-standard.

In computational approaches, the cost of computation has to be constantly weighed against the required accuracy, which generally correlates with the sophistication of the method. The computational cost scales exponentially with the number of basis functions N used to describe each electron. For HF, MP2, and CCSD(T), the computational cost increases with N^4 , N^5 , and N^7 , respectively.^[235] While pure HF calculations are outdated nowadays, some systems are too large for proper treatment using the MP2 or CCSD(T) methods. To mitigate this high computational cost, approximations have been introduced that lead to a more favorable scaling. Two prominent examples are the resolution of identity (RI)^[236] and the domain-based local pair natural orbital (DLPNO)^[237] approximations. For RI-MP2, for example, the computational cost is decreased by one order of magnitude, while maintaining similar accuracy as MP2. DLPNO-CCSD(T), on the other hand, has shown to calculate 99.9% of the CCSD(T) correlation energy, with an almost linear scaling behavior.^[238] Alternatively, one could also use the less expensive density functional theory methods.

2.3.2 | Electronic Energies: Density Functional Theory

Another class of electronic structure calculation methods is density functional theory, which relies on a different philosophy than applying the Schrödinger equation to the wave function of the investigated system. Based on the Hohenberg-Kohn theorem,^[239] the ground state electron energy of a system can be completely determined by its electron density through a functional expression. Contrary to wave function-based methods, where each electron is described by four variables, three spatial- and one spin coordinate, the whole electron density of a system is only dependent of three spatial variables, regardless of the number of electrons. However, the functional that connects the electronic energy and the electron density is unknown.

In the beginning of DFT development, functionals were generated that would connect all components of the electronic energy to the electron density. Their performance was very poor compared to wave function-based methods. A significant improvement was made by Kohn and Sham who established that the kinetic energy of the electrons can be calculated from an auxiliary set of orbitals used for representing the electron density.^[240] Now, the only unknown term is the exchange-correlation energy, which is a small part (ca. 1%) of the total electronic energy of a system. However, an accurate calculation of this 1% is important, as it describes the interaction between electrons; thus, it should not be neglected. Although the introduction of orbitals in DFT increased the dependence of the electron density from only three variables to $3N$ variables, it is often considered as the advent of modern DFT.^[241]

In the following decades, several functionals were developed to model the exchange-correlation energy. Although systematic improvement as for wave function-based methods is not possible, most functionals can be assigned to five main categories of functionals, which are ranked according to their accuracy. These categories are often compared to steps of the DFT Jacob's Ladder (Figure 2.6a),^[241] where methods closer to the top are generally more accurate. The first functionals made use of the local density approximation (LDA),^[240] in which the density is like a uniform electron gas. The obtained results are often very inaccurate for molecular systems, which is why a non-uniform electron gas is considered in the generalized gradient approximation (GGA).^[242] Here, the first derivative of the electron density is introduced as a variable. Although significantly more accurate functionals are available, the use of GGA functionals, such as the PBE functional, is still established in the community for certain applications. Pure DFT functionals can be further improved by including second derivatives, leading to the so-called meta-GGA methods.^[243,244]

So far, most functionals performed exceptionally well in modelling the correlation energy that is neglected in HF. However, their performance at calculating the exchange energy is underwhelming compared to HF. Therefore, hybrid-DFT functionals were introduced, in which the exchange energy is composed of a DFT and a HF term.^[245] The use of hybrid-DFT methods is well-established in the community to get accurate electronic energies and frequencies of systems that are too big to be treated with post-HF methods. Furthermore, compared to pure DFT methods, hybrid-DFT methods can yield a high enough accuracy to compute reaction mechanism and TSs. Some frequently used hybrid-DFT methods are B3LYP^[245-247] and PBE0.^[248] Currently the most accurate DFT methods are double hybrid-DFT functionals.^[249,250] Here, the correlation energy is composed of a DFT and, for example, an MP2 term, leading to the currently most accurate electronic energies that can be computed with DFT (Figure 2.6b). Modern DFT methods present a good tradeoff between computational cost and accuracy, compared to wave function-based methods. Most double-hybrid DFT methods clearly outperform MP2, with a similar scaling behavior.^[251]

2.3.3 | Empirical Dispersion Correction

Some electron correlation effects, such as dispersion, are not properly treated using DFT methods. These forces play a huge role for determining the most stable conformation of a molecule and for determining how multiple molecules interact with each other. Here, intra- and intermolecular interactions between functional groups and the non-polar residues play an important role. In this work, many probed molecular ions carry electron-rich phenyl groups, which are heavily influenced by dispersion interactions. In order to account for dispersion in DFT, several corrections have been developed, such as the D3 and D4 corrections developed in the research group of Stefan Grimme^[252,253] or Scheffler's and Tkatchenko's vdW^{TS} correction.^[254]

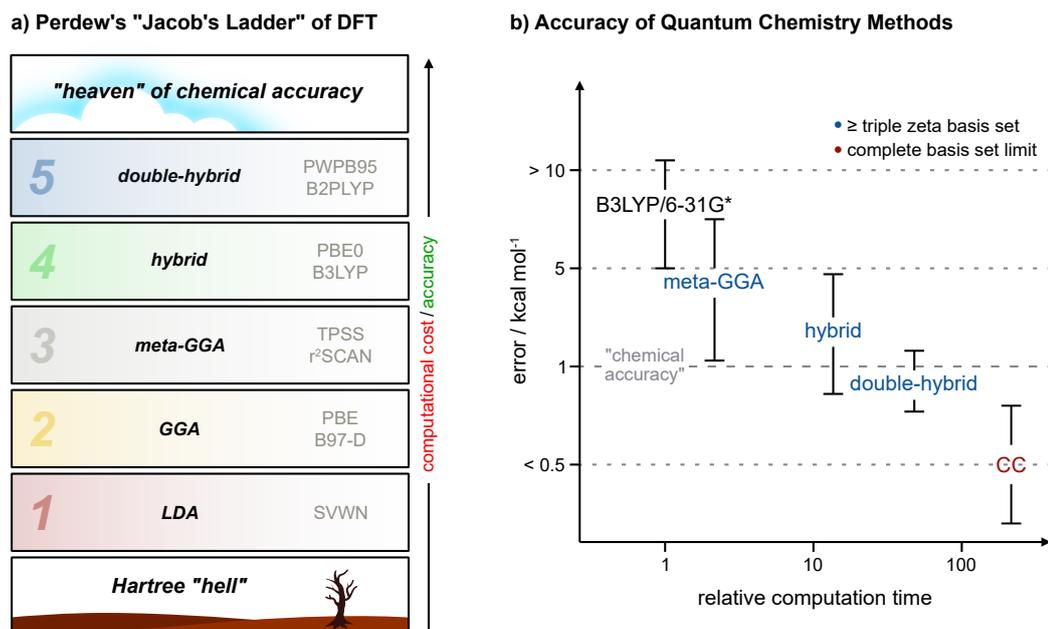


Figure 2.6: (a) Categorization of density functionals according to Perdew's "Jacob's ladder". Each successive rung represents a category of density functionals that is typically more accurate than the previous one, but comes at the cost of greater computational expense. (b) Accuracy of quantum chemical methods as a function of computational time on a logarithmic scale. CC methods provide an accuracy below 1 kcal mol^{-1} ($\approx 4.2 \text{ kJ mol}^{-1}$), which is defined as "chemical accuracy". Hybrid DFT methods can reach chemical accuracy if an appropriate basis set is used, while double-hybrid DFT methods can reach the accuracy of CC methods. Using hybrid DFT methods with a small basis set, such as B3LYP/6-31G*, can lead to large errors. Adapted with permission from Bursch et al.^[226] Creative Commons Attribution-NonCommercial 4.0 International License.

In their most simple form, an empirical r^{-6} -dependent term is added to the DFT energy. These methods are nowadays established among quantum chemists and it is encouraged to use them. The increase in computational cost depends on the chosen dispersion correction, but is either negligible or at least significantly lower than the cost of the DFT calculation itself.^[226]

2.3.4 | Basis Sets

Besides choosing a wave function method or a functional, computational chemists also need to choose a basis set for their calculation. A basis set is a set of basis functions that approximate atomic orbitals. Molecular orbitals are linear combinations of atomic orbitals. The number and shape of basis functions to describe an atomic orbital directly reflects in the accuracy of the basis set. However, the computational cost of the calculation increases with the number of basis functions chosen. A viable approach to mitigate a high computational cost is to use an elevated number of basis functions to describe the valence electrons, which play a more important role in describing the interactions between atoms. The core electrons, on the other

hand, are described with fewer basis functions. Furthermore, polarization and diffuse functions can be added to a basis set to improve its accuracy.^[255,256]

A large number of basis sets have been developed. Based on conditions, such as availability for a large number of elements from the periodic table, systematic improvability by adding more functions, and its universal availability for all types of quantum chemical calculations, computational chemists can choose the most suitable basis set for their endeavour. Basis sets from two established *families* have been used extensively in this thesis. Pople style basis sets have been developed in the 1970s.^[257] However, many computational chemists now prefer to use basis sets that can be systematically improved towards the complete basis set limit.^[255,256] Nevertheless, all glycans in this thesis are modelled with the Pople basis set 6-311+G(d,p), as this basis set provided accurate results for very similar systems in the past. Ahlrichs style basis sets, such as def2-TZVPP,^[258,259] allow for better control and systematic improvement. They are available for a large fraction of the periodic table (including the sixth period) and are used for both DFT and wave function-based methods.^[260] Ahlrichs-type basis sets were used in this thesis for computing RNA fragments and small carbocations.

2.3.5 | Optimization and Frequency Calculations

Ultimately the aim of using computational methods in this thesis is to generate accurate structures and use their harmonic frequencies to match them with the experimental IR spectra. As optimization and frequency calculations were heavily used in this work, this section will provide a brief introduction on how these processes work.

An optimization is an iterative procedure with which stationary points on a potential energy surface can be located. A PES of a molecule is a multidimensional ($3N_{at}$ dimensions, with N_{at} being the number of atoms) hypersurface that describes the energy as a function of the position of atoms in a molecule, but also chemical reactions. Each system has a unique PES and the exact shape of the PES is also dependent on the quantum chemical method that is used. As the human brain cannot easily process more than three dimensions, parts of a PES are often visualized in two or three dimensions. Minima on the PES correspond to stable structures with a definite lifetime. First-order saddle points, which in 2D- or 3D-representations appear as maxima, correspond to transition states, which connect two minima. They do not have a lifetime.

A geometry optimization is usually performed on a structure that was either built using chemical intuition or a structure yielded from a structure sampling tool (more about the latter can be read in the next section). Structures generated by chemical intuition do not usually correspond to stationary points. Structures generated from a sampling tool were potentially already optimized with a certain method, which is, however, different to the one that will be used for the calculation of accurate electronic energies and harmonic frequencies. Normally,

all structures are optimized at the same level of theory that is used to generate energies and harmonic frequencies. However, multi-level workflows are possible, for example to perform high-level single-point energy calculations on a structure optimized at a lower level of theory.

In most cases, optimizations are performed to locate minima on the PES. However, an optimization usually does not locate the global minimum, but only the local minimum closest to the input geometry. During an optimization, the first derivatives (gradient) of the energy with respect to the atomic coordinates are calculated. With this information, the atoms will be displaced in the direction that yields the largest energetic stabilization. Then the step will be repeated until the energy and the gradient lie below predefined thresholds.

In certain situations, it can be helpful to calculate the second derivatives, also known as the Hessian matrix. For example, the optimization of transition states (saddle points) requires second derivatives. Otherwise, if the atomic coordinates are optimized according to the greatest decrease in energy, the system will quickly fall into a close local minimum. In contrast to minima, TSs have exactly one negative second derivative (imaginary mode), which corresponds to the trajectory that connects two minima. Therefore, an initial guess for a TS must already incorporate the imaginary frequency describing the reaction trajectory. Such a guess can be generated by chemical intuition or by scanning a distinct coordinate from the reactant or product geometries. The geometry of the energetic maximum (saddle point) of such a scan would correspond the closest to the TS.

To determine whether a structure is a minimum or a saddle point, a frequency calculation needs to be performed. A minimum only has real frequencies, while saddle points also have one imaginary frequency. The squares of the eigenvalues of the Hessian matrix are proportional to the harmonic frequencies, while the eigenvectors indicate the direction of the vibrational mode. Besides identifying a geometry as a stationary point, a frequency calculation also allows us to extract vibrational spectra, zero-point vibrational energy (ZPVE), and thermodynamic quantities. It is important to note that for the calculation of harmonic frequencies, a parabolic function is used to approximate the motion of nuclei around the minimum. This approximation works comparably well for minima, as the experimental frequencies align well with their computed counterparts, after the latter are corrected by an empirical scaling factor to account for anharmonicities.

Scaling factors for harmonic frequencies are dependent on the method and the basis set.^[261] They can be empirically determined by minimizing the deviations between experimentally resolved bands and the computed frequencies.^[262,263] Some vibrations exhibit a strong anharmonic behavior and are, therefore, inaccurately described by the harmonic approximation. Methods are available to compute anharmonic frequencies, but they come at a great computational cost. As a consequence, anharmonic frequencies were not used in this thesis, except for Chapters 8 and 9. To illustrate the excessive computational cost of anharmonic frequencies, the GVPT2 method^[264–266] implemented in Gaussian 16^[267] was used on a medium-sized system with 33

atoms and 170 electrons with the method PBE0+D3/def2-TZVP.^[268] A harmonic frequency calculation on that system took 1.5 h, whereas the anharmonic frequency calculation took 337.6 h on an Intel® Xeon® Gold 6130 CPU @ 2.10 GHz. The reaction ran over 32 cores, which means that on one single core the calculation would have taken ca. 10 600 h, which corresponds to 1.2 years.

2.3.6 | Conformational Sampling

Geometries of molecules built by chemical intuition are likely not close to the global minimum in the case of flexible biomolecules but would rather fall into a local minimum upon optimization. To mitigate human bias, software packages have been developed for conformational search. Some of these algorithms can screen a large part of the PES and can potentially yield structures close to the global minimum under appropriate parametrization.

In this thesis two tools were used to sample flexible molecular ions. For all pyranose-based sugars the tool Flexible Algorithm for Optimization of Molecules (FAFOOM) was used.^[269] FAFOOM is a genetic algorithm using biological concepts to find the lowest-energy structure. FAFOOM generates an initial population by randomly varying torsion angles and pyranose ring puckers. Each structure is sent to an external software for optimization at the desired level of theory, yielding minimum structures and their respective electronic energies. These electronic energies are used to determine the fitness of the molecules. The fittest structures are chosen as parent structures from which features are taken to generate the next generation of structures. Then, the fitness is re-assessed and a new pair of parent structures is chosen. Along the way, random mutations can occur, leading to changes in torsion angles or ring pucker. The algorithm terminates if after a certain number of generations the fitness does not improve anymore or if a predefined threshold of sampled structures is reached.

For all other flexible structures, the Conformer–Rotamer Ensemble Sampling Tool (CREST) was used.^[270] CREST is very user-friendly, as it only requires an input geometry and a command line to automatically output potentially hundreds of low-energy conformers by combining molecular dynamics, genetic structure crossing, and meta-dynamics.^[270,271] CREST employs both empirical force field and semi-empirical methods. The latter are based on *ab initio* methods, such as DFT or HF, but some parameters are obtained by fitting to experimental data or high-level calculations.^[272] This method was used for Chapter 7 and Chapter 8, as FAFOOM does not include parameters to systematically sample sugar rings that are not six-membered.

It is important to note that less accurate methods are often used to generate conformers to speed up the conformational search. In many cases, force field methods are used. Here, the motions and interactions of atoms are described by classical mechanics, while electrons are not quantum mechanically treated. Empirical parameters are used for the calculation of electronic energies and the optimization of structures. The empirical methods are often not reliable for

ions in the gas phase, especially in the case of sugars. DFT methods, on the other hand, using the general gradient approximation yield accurate results for sugars.^[273] Therefore, FAFOOM was combined with DFT methods for sampling pyranose-based glycans in this work. Although this workflow comes with an increase in computational cost, compared to semi-empirical methods, the cost is reasonable for the herein-treated systems. For larger systems, the use of semi-empirical or empirical methods is recommended, as implemented in CREST.^[189] More details on the computational workflow can be found in Chapter 3.

3 | Experimental Methods

This chapter describes the experimental methods and instruments used for the structural investigation of reactive intermediates. The chapter starts with a description of the custom-built instrument that combines mass spectrometry and infrared spectroscopy in helium droplets to record infrared spectra of reactive intermediates. Infrared action spectroscopy in helium droplets requires a powerful laser source, such as the free-electron laser (FEL), which will be described after the helium droplet instrument. The last section of this chapter gives an overview of the computational methods used to model IR spectra of reactive intermediates. Further details on the experimental methods can be found in Chapters 4–9.

3.1 | Infrared Spectroscopy in Helium Nanodroplets

All experimental data for this thesis have been collected using a custom-built instrument that combines MS and IR spectroscopy in helium nanodroplets (Figure 3.1). This section serves as an overview of the experimental setup. Further details can be found in previous doctoral theses.^[274,275] This instrument is based on a Waters Ultima quadropole time-of-flight mass spectrometer (Waters Corporation, Manchester, UK) that was subsequently heavily modified. The original instrument is equipped with a Z-spray nESI source. The nESI source requires using disposable glass needles that are coated by a thin metallic layer. The needles are prepared in-house from borosilicate capillaries pulled to a tip of a few μm using a P-1000 micropipette puller (Sutter Instrument, Novato, USA). Subsequently, the needles are coated with Pd/Pt (80:20, w:w) with a 108auto sputter coater (Cressington Scientific Instruments, Watford, UK).

Precursors for glycosyl cations, RNA fragments, or carbocations were either synthesized by collaborators or purchased. Details on the precursors can be found in the respective chapters. The samples were dissolved in acetonitrile and water (9:1, v:v), yielding 100–250 μM solutions. Further details can be found in Chapters 4–9. Ions were generated by applying ± 0.8 –1.1 kV to the needle containing the diluted sample. The ions are guided through differential pressure stages into the high vacuum ($\leq 10^{-5}$ mbar during operation) using ion optics. First, the ions are guided through two ring-electrode ion guides. As the pressure in this region is comparably high (1–10 mbar), acceleration of ions in the first ion guide leads to more collisions with residual

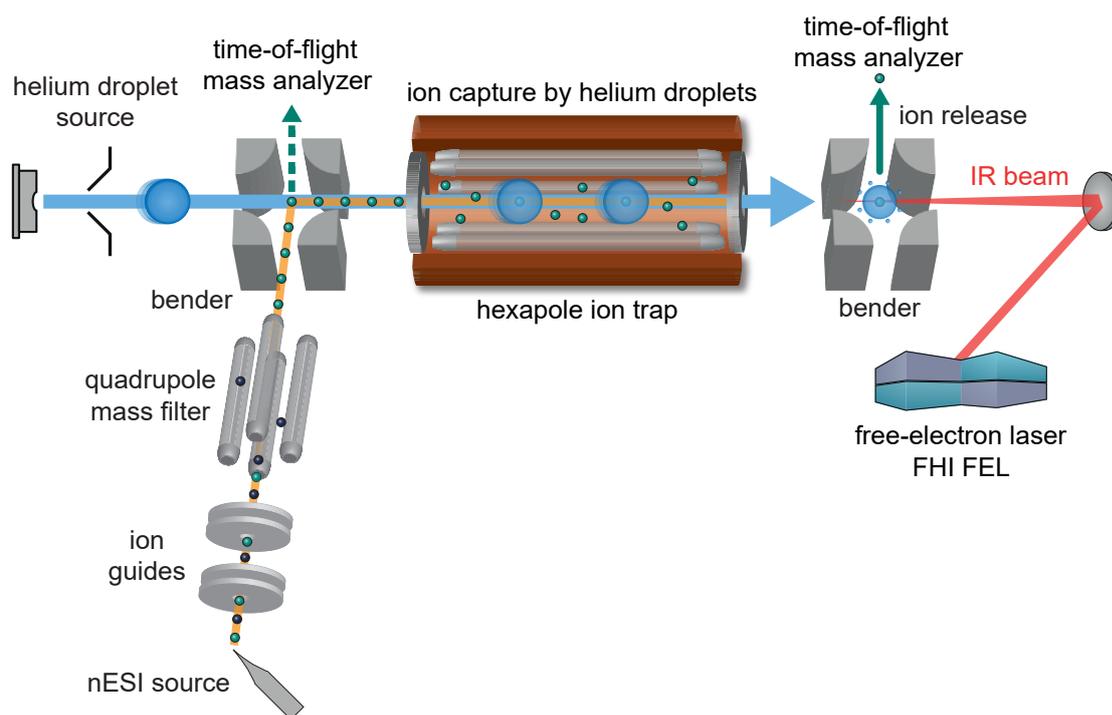


Figure 3.1: Schematic overview of the helium droplet instrument. The precursor is ionized and fragmented using nESI and in-source fragmentation. The ions of interest are m/z -selected in a quadrupole mass filter and then guided into a hexapole ion trap. Here, the ions are picked up by superfluid helium nanodroplets and guided to the detection region, where the doped droplets overlap with the Fritz Haber Institute free-electron laser (FHI FEL) IR beam. After absorption of multiple resonant photons, the ions are released from the droplet and detected using a TOF mass analyzer.

solvent and air molecules, leading to in-source fragmentation. This fragmentation method has been used in this thesis to generate reactive intermediates and is similar to CID. After passing through the ion guides, the generated ions enter a quadrupole mass filter. All ions are guided through the filter to record mass spectra, whereas the ions of interest are m/z -selected to record IR spectra. Then the ions pass a collision cell, which only serves for ion transfer.

After the collision cell the customized part of the instrument starts, as the ions enter a quadrupole bender that can be switched off to let the ions pass through to a TOF mass analyzer. This mass analyzer also originates from a Water Ultima quadrupole time-of-flight mass spectrometer and serves for recording mass spectra and generally monitoring and optimizing the ion signal. If a potential is applied to the rods of the quadrupole bender (DC voltage of ± 33 V), the ions are deflected by 90° into a hexapole ion trap. When recording IR spectra, the ion trap is filled for 2.5 s before measuring a data point. Upon entering the hexapole ion trap, which is surrounded by a copper housing, the ions are thermalized by collisions with helium buffer gas. The temperature of the ion trap can be modulated by adjusting the temperature of nitrogen gas that can optionally flow through a circuit in the copper housing. Before entering

the copper housing, the temperature of the flowing nitrogen can either be increased up to ca. 400 K with a heater or be cooled to ca. 90 K with liquid nitrogen. In this thesis, all ions were probed at ca. 90 K. Especially in the case of flexible biomolecules, higher temperatures lead to the population of many conformers, leading to broader IR signatures and a lower overall signal.^[276] Slow sympathetic cooling with the buffer gas allows the ions to populate the global minimum to a higher degree. After cooling the ions the buffer gas is pumped out of the trap. Radial confinement of the ions is achieved by an RF potential on the hexapole rods. The ions are longitudinally trapped by endcap electrodes to which a DC potential 3–5 V above the offset voltage of the hexapole is applied.

The hexapole ion trap is traversed by superfluid helium nanodroplets that are generated by a pulsed Even-Lavie valve (10 Hz). Expansion of pressurized helium (70–80 bar) through a cryogenic nozzle (21 K) into the vacuum leads to the formation of a beam of superfluid helium droplets with an average size of 10^5 atoms and a velocity of ca. 450 m s^{-1} . After passing through a skimmer, the helium droplet beam passes through the ion trap, picks up ions, and guides them out of the trap. Due to the high kinetic energy of the helium nanodroplets, ions inside the droplet can overcome the trapping potential. Subsequently, the beam of doped helium nanodroplets enters the detection region, consisting of a third ion guide followed by a second quadrupole bender. Here, the encapsulated ions overlap with a pulsed IR beam, generated by the Fritz Haber Institute free-electron laser (FHI FEL).^[277] The FHI FEL was operated with a repetition rate of 10 Hz, macropulse energies between 50–130 mJ, and a full width at half maximum (FWHM) of the IR beam of ca. 0.3–0.6 % of the central wavelength. Every macropulse is composed of about 10000 micropulses at a repetition rate of 1 GHz. The absorption of resonant IR photons leads to vibrational excitation of the ion inside the droplet. The vibrational energy is dissipated to the helium matrix that subsequently shrinks through evaporation. After the absorption of multiple resonant IR photons, the ions are released from the droplets, before entering the quadrupole bender. Here, the released ions are deflected by 90° into a second TOF mass analyzer, which is of the same type as the first TOF mass analyzer. For every data point, the ion signal is averaged over 25 FEL macropulses. Plotting the ion signal as a function of the IR wavenumber leads to an IR spectrum. The spectra in this work were measured in the $1000\text{--}1800 \text{ cm}^{-1}$ range, except for the spectra shown in Chapter 9 that were measured in the $600\text{--}1700 \text{ cm}^{-1}$ range. Each spectrum consists of an average of at least two measured spectra, measured with a step size of 2 cm^{-1} . Note that measuring a spectrum with a range of 800 cm^{-1} in that mode takes approximately 60 min, assuming that both the ion signal and the FEL are stable.

Each spectrum undergoes a first-order correction by dividing the ion signal by the energy of the IR pulse and is then normalized using a customized script for the software Wolfram Mathematica 11.3 (Wolfram Research, Champaign, IL, USA). Then the spectra are plotted using OriginPro 2018 (OriginLab Corporation, Northampton, MA, USA) and Affinity Designer 1.9.2.1035 (Serif Europe, West Bridgford, UK).

3.2 | Fritz Haber Institute Free-Electron Laser

In this thesis, the signal of ions released from helium droplets, as a function of the IR wavenumber of the photons that interact with the embedded ions, is used to generate IR spectra. To release the ions from helium droplets, a high-energy source for IR radiation is necessary. Currently, FELs are the most suitable source for such high-intensity radiation. It is worth noting that FELs cannot be readily bought and can only be used in certain facilities. Commonly, interested users apply for FEL beamtime at the facilities FELIX^[278] in Nijmegen (Netherlands) or CLIO^[279] in Orsay (France). Another FEL, the FHI FEL, is located at the Fritz Haber Institute of the Max Planck Society in Berlin (Figure 3.2).^[277] In operation, an FEL emits a high amount of X-ray radiation. Therefore, it is shielded behind thick walls of concrete. This section is intended to give a brief overview of the working principle of FELs and how it was operated to generate the results of this work.

A gridded electron gun that includes a heated cathode made out of doped tungsten produces electrons with a repetition rate of 1 GHz. Using a subharmonic buncher, the electron bunches are compressed into shorter bunches that are compatible with the following linear accelerators (LINACs). In the following LINAC, the electrons are accelerated to 20 MeV, corresponding to 99.97% of the speed of light. In a second LINAC the kinetic energy of the electrons can be increased or decreased to values between 15–50 MeV. In this work, electrons with energies of 36–37 MeV have been used, except for Chapter 9, where also electrons with 32 MeV were used. With this setup, electron macrobunches with a duration of 10 μ s and a repetition rate of 10 Hz are generated. Every macrobunch consists of about 10^4 microbunches. Passing two isochronous bends, the electron beam is deflected twice by 90° before entering the optical cavity.

The optical cavity has a length of 5.4 m that is delimited by two gold-plated copper mirrors. In the cavity, the electron beam interacts with a magnetic field generated by an undulator. The FHI FEL undulator consists of two periodic series of permanent NdFeB magnets above and below the cavity axis. In the magnetic field of the undulator, the negatively charged electrons experience the Lorentz force. Thereby, the electrons deviate from their linear trajectory to undergo sinusoidal motions perpendicular to the magnetic field and emit radiation in the forward direction. While the electrons are deliberately steered into an electron-beam dump, the emitted radiation gets reflected by the cavity mirrors and interacts with another incoming electron bunch. Thereby, the IR radiation will be amplified and it obtains the pulsed nature of the electron beam. The obtained IR light is coherent and monochromatic to a high degree with a bandwidth of 0.3–0.6 % (FWHM) of the central wavelength.

The wavelength of the emitted IR light can mainly be controlled by two parameters, which are the electron energy and the strength of the magnetic field. The former is a value that needs to be fixed at the beginning of every FEL beamtime. Therefore, in operation, only the strength of the magnetic field can be adjusted to manipulate the IR wavelength. The strength of the magnetic

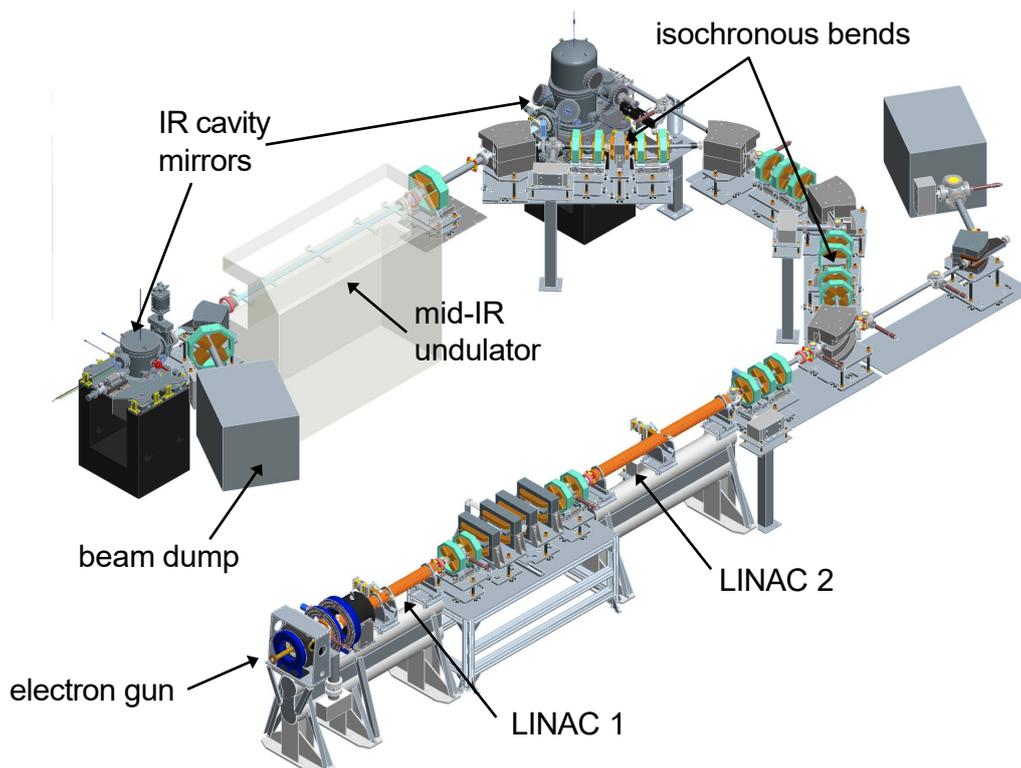


Figure 3.2: Schematic representation of the FHI FEL. Adapted with permission from W. Schöllkopf.

field can be manipulated by adjusting the gap between the upper and lower array of magnets in the undulator. A smaller gap leads to a stronger magnetic field and longer wavelengths. With an electron energy of 36–37 MeV, typically wavelengths of 5–12.5 μm (800–2000 cm^{-1}) can be obtained, whereas with 32 MeV wavelengths of 7.2–16.6 μm (600–1400 cm^{-1}) are achievable.

3.3 | Computational Methods

Except for Chapter 9, every computed IR spectrum shown in this work originates from sampled geometries. In Chapter 9, the structures are small and rigid, rendering conformational sampling obsolete. The first structures were built by chemical intuition using either the program GaussView 6.0.16^[280] or Avogadro 1.2.0.^[281] Subsequently the sampling tools FAFOOM^[269] and CREST (Version 2.9)^[270] were used to generate conformers of a given structure. The use of FAFOOM is established for pyranose-based sugars, for which it also performs well.^[273] Therefore, it has been applied in Chapters 4–6 with the settings defined in Table 3.1. However, FAFOOM is not a standalone program but is relying on an external software for structure optimization. Every structure generated by FAFOOM was optimized at the DFT level of theory PBE+vdW^{TS/light}^[254,282] in FHI-aims (version 171221)^[283] for Chapters 4 and 6, whereas the

Table 3.1: Parameters used for the conformational search in FAFOOM. *Distance_cutoff_1* defines the lowest distance possible between non-bonded atoms, whereas *Distance_cutoff_2* defines the maximum distance between bonded atoms. *Rmsd_cutoff_uniq* defines the value of the root-mean-square deviation between two structures that must be overcome to classify them as structurally distinct. *Popsiz*e defines the size of the population that is necessary, before the genetic algorithm starts. *Prob_for_crossing*, *Prob_for_mut_pyranosering*, and *Prob_for_mut_torsion* define the probability for a crossing or a mutation event to occur. Therefore, a random number between 0 and 1 will be generated, and if it is above the defined threshold, the event will occur. Furthermore, the maximum number of mutations to occur per GA run can be defined by *Max_mutations_torsion* and *Max_mutations_pyranosering*. The structures that undergo crossing are selected by the mode *Roulette wheel*, where the fittest structures are chosen, based on computed electronic energies. The sum of the fitness of two structures that are crossed needs to be above a defined threshold *Fitness_sum_limit*.

	Parameter	Value
Molecule	Distance_cutoff_1	1.2
	Distance_cutoff_2	2.15
	Rmsd_cutoff_uniq	0.25
GA settings	Popsiz	10
	Prob_for_crossing	0.95
	Prob_for_mut_pyranosering	0.6
	Prob_for_mut_torsion	0.8
	Fitness_sum_limit	1.2
	Selection	Roulette wheel
	Max_mutations_torsion	3
	Max_mutations_pyranosering	1

structures in Chapter 5 were optimized with the PBE/def2-SVP^[259] method in ORCA 4.1.1.^[284] Although the combination of FAFOOM with FHI-aims is established,^[85,87] structure optimizations for rather large systems containing fluorine often failed (see Chapter 5). Therefore, the program was for this study interfaced with ORCA instead. In Chapters 7 and 8, seven-membered sugar rings and RNA fragments, containing five-membered ribose, needed to be sampled. However, the possible conformations of such rings are not implemented in the code of FAFOOM. Therefore, CREST was used to sample these structures. Here, the structures were optimized with the semi-empirical GFN2-xTB^[285] method, whereas some structures in Chapter 7 were also sampled using the empirical GFN-FF force field.^[286] Both methods are implemented in the xtb package (version 6.3.0).^[287] With semi-empirical methods in CREST sometimes erroneous structures are obtained, as some unreasonable bonds can be formed. With GFN-FF, the initial connectivity of the system remains intact, although the performance is generally weaker than that of GFN2-xTB.^[287]

A subset of structures with distinct structural types (e.g., oxocarbenium- or dioxolenium-type structures) were selected from the sampled geometries and reoptimized at the hybrid DFT level of theory PBE0+D3/6-311+G(d,p)^[248,252,257] implemented in Gaussian 16, Revision

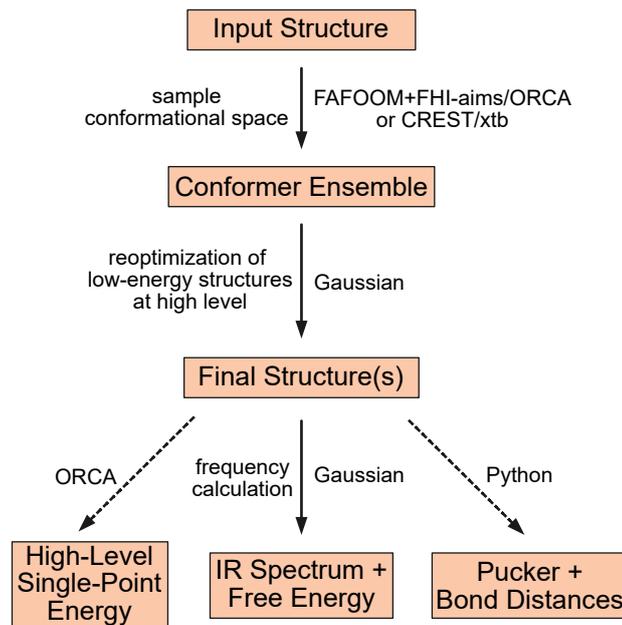


Figure 3.3: Computational workflow that was employed in this thesis to probe the conformational space of reactive intermediates and to generate their IR spectra. Steps that are indicated with dashed lines are optional and were not employed for all investigated intermediates.

A.03,^[267] which has previously yielded reliable results for glycosyl cations.^[85,87] The same method has been employed to compute harmonic frequencies, which were subsequently scaled by an empirical factor of 0.965. For Chapter 8, the basis set was adapted to def2-TZVPP, as the Ahlrichs-type basis set yielded more accurate frequencies for RNA anions than the Pople-type basis set. In Chapter 9, the CAM-B3LYP+D3/def2-TZVPP^[288] method was used, which in this case, also reproduced the experimental frequencies better than the previously used methods. For certain ions in Chapters 8 and 9, also anharmonic frequency calculations were performed using the GVPT2 method.^[264–266] Energies and frequencies are extracted from the output files of the calculations using a custom Python script. The employed workflow is visualized in Figure 3.3.

The structures are ranked using their relative free energies at 90 K, which is the temperature of the ion trap in the experiment. In Chapters 4 and 7, the DFT-optimized geometries were used for single-point energy calculations at the DLPNO-CCSD(T)/def2-TZVPP^[237] level of theory in ORCA 5.0.3.^[289] Here, the final energy used for the structural ranking was the electronic energy at the DLPNO-CCSD(T) level in addition to a zero-point vibrational and free energy correction originating from the PBE0+D3 frequency calculation.

Chapters 4, 7, and 8 also contain TS calculations. Candidates for transition states were generated using the `opt=modredundant` function in Gaussian 16. This function allows to increase or decrease the distance between two atoms in a stepwise manner, optimize the geometry with the positions of the two atoms fixed, and ultimately obtain the electronic energy

of the system as a function of distance that is scanned. Maxima (saddle points) in the resulting energy diagram are potential candidates for TSs. When this structure is optimized as a TS, it should have one imaginary frequency corresponding to the reaction trajectory. In a subsequent intrinsic reaction coordinate calculation, the two minima that are connected by the TS can be determined.^[290]

4 | Influence of the Electron Density of Acyl Protecting Groups on the Selectivity of Galactosylations

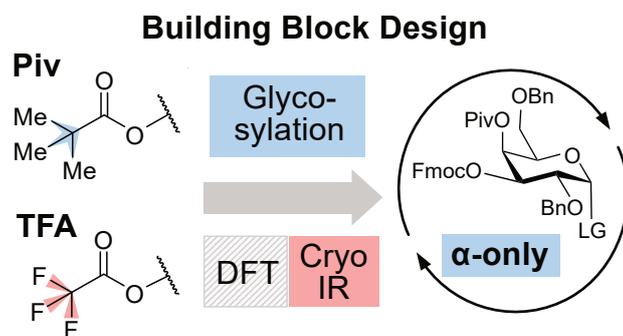
This work is published^[291] and available online:

DOI: <https://doi.org/10.1021/jacs.2c05859>

Details on the related Supporting Information can be found in Appendix A.

K. Greis,[†] S. Lechnitz,[†] C. Kirschbaum, C.-W. Chang, M.-H. Lin, G. Meijer, G. von Helden, P. H. Seeberger,* K. Pagel*; The Influence of the Electron Density in Acyl Protecting Groups on the Selectivity of Galactose Formation. *J. Am. Chem. Soc.* **2022**, *144*, 20258–20566.

[†]equal contributions



Submitted: 2 June 2022; **Published:** 27 October 2022

Copyright © 2022 The Authors. Published by American Chemical Society

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

4.1 | Author Contributions

The project was conceived by **Kim Greis**, Sabrina Lechnitz, Peter Seeberger, and Kevin Pagel. **Kim Greis** (main) and Sabrina Lechnitz (support) wrote the first draft of the manuscript. **Kim Greis** (main) and Carla Kirschbaum (support) performed the mass spectrometry and cryogenic infrared spectroscopy experiments. The cryogenic infrared spectroscopy experiments were supervised by Gerard Meijer and Gert von Helden. **Kim Greis** performed all quantum chemical calculations. Sabrina Lechnitz synthesized the galactose building blocks and performed the automated glycan assembly. Sabrina Lechnitz (main), Chun-Wei Chang, and Mei-Huei Lin (both support) performed the glycosylation reactions. **Kim Greis** and Sabrina Lechnitz analyzed the data.

4.2 | Project Hypothesis and Summary

As detailed in Chapter 2, the synthesis of 1,2-*cis* glycosidic linkages is one of the major challenges in glycochemistry. These linkages can be found in many biological systems, such as bacterial lipopolysaccharide antigens^[292,293] or blood group epitopes.^[294] The work presented in this chapter mainly focuses on how the synthesis of 1,2-*cis* glycosidic linkages can be optimized. Furthermore, the mechanism leading to these sugars is investigated using cryogenic IR spectroscopy and computational methods.

It is known from previous studies that an acyl group at the C4-position in galactose building blocks enhances the 1,2-*cis* (or α -) selectivity.^[63-67] While this increase in selectivity is consistent, it is not high enough to be readily implemented in AGA workflows, which requires total selectivity. The main hypothesis of this work was that remote participation could be enhanced and maybe even deactivated depending on how the acyl group is functionalized. Electron-donating substituents would increase the electron density of the carbonyl oxygen in the acyl group, increasing its ability to engage in remote participation, whereas electron-withdrawing substituents would decrease the efficiency of remote participation or deactivate it completely. In a previous study, glycosylation reactions of acetylated galactosyl donors were compared to the structure of glycosyl cations derived from cryogenic IR spectroscopy combined with DFT calculations.^[87] This study demonstrated that remote participation of C4-acetyl groups can be directly observed for galactosyl cations. A similar workflow was employed in this study, where galactosyl donors are protected with electron-donating pivaloyl (Piv) and electron-withdrawing trifluoroacetyl (TFA) groups, instead of Ac groups.

In a large set of glycosylation reactions involving seven different galactosyl donors and four nucleophiles (see Figure 1 of the related published work), weak nucleophiles generally produced a higher amount of 1,2-*cis* glycosides than strong nucleophiles. This observation is in line with previous studies.^[88] As previously determined^[87] for acetyl groups at the C6-position,

Piv or TFA groups at this position do not positively affect the 1,2-*cis* selectivity of such building blocks. For the C4-functionalized galactosyl donor, a clear increase in 1,2-*cis* selectivity can be observed for the pivaloylated building block compared to its trifluoroacetylated counterpart, affirming the initial hypothesis. According to this hypothesis the 1,2-*cis* selectivity of the C4-acetylated building block would need to lie between that of its functionalized counterparts. However, its selectivity is even lower than that of the C4-trifluoroacetylated building block. This result suggests that the mechanism of the reaction cannot be explained by remote participation alone. A building block carrying pivaloyl groups at both the C4- and the C6-positions is less selective than its C4-pivaloylated counterpart, whereas the opposite trend was observed in a previous study for acetylated building blocks.^[87] This observation could be attributed to the increased steric hindrance caused by the bulky Piv group, compared to its smaller Ac counterpart. Intrigued by these observations, the structures of the glycosyl cations generated by these donors were studied using cryogenic IR spectroscopy and DFT calculations.

Electrospray ionization followed by in-source fragmentation of the pivaloylated and trifluoroacetylated galactosyl donors only afforded glycosyl cations in the case of pivaloylated precursors. Trifluoroacetylated donors ionize well and can be detected as proton or sodium adducts, but do not fragment in the same way as the pivaloylated donors. It is likely that a participating protecting group is necessary to ease the cleavage of the leaving group leading to the formation of glycosyl cations. This behavior would corroborate the initial hypothesis that the ability of TFA groups to engage in remote participation is decreased. Hence, only the experimental IR spectra of the pivaloylated galactosyl cations were probed (see Figures 2 and 3 of the related publication). These IR spectra show a set of diagnostic absorption bands in the 1400–1800 cm^{-1} region. In combination with DFT-computed harmonic frequencies, it is shown that in the case of the galactosyl cations carrying a pivaloyl group at the C4-position, a large fraction adopt dioxolenium-type structures, in which the positively charged anomeric carbon is stabilized by remote participation of the C4-protecting group (see Scheme 1 of the related publication). However, the identity of the second fraction as well as the structures adopted by the galactosyl cation carrying a pivaloyl group only at the C6-position could not be revealed based on other structural types that are commonly adopted by glycosyl cations, *i.e.*, oxocarbenium or oxonium structures. Instead, ring-opening of the pyranose-ring occurs after attack of the C4- or the C6-carbonyl oxygen of the pivaloyl group at the C5-atom, as previously reported for acetylated glycosyl cations.^[88] This ring-opening reaction leads to the formation of a five-membered dioxolenium-ring and an aldehyde functional group. This gas-phase rearrangement reaction has not been observed in the condensed phase yet. Calculations of the transition states leading to rearrangement and remote participation show that both are thermodynamically favored in the case of C4-pivaloylated galactosyl cations. The barrier leading to rearrangement is comparably high (138 kJ mol^{-1} vs. 4 kJ mol^{-1} for remote participation). However, it is known from previous studies that processes with a similar barrier height occur

during in-source fragmentation.^[268,295] As the barrier for remote participation is rather low, this process is believed to occur in the condensed phase as well, which would explain the high 1,2-*cis* selectivity of the C4-pivaloylated building block. The experimental IR signature of the related C6-pivaloylated galactosyl cation shows that it is mainly forming rearranged structures, while remote participation cannot be observed. Similar to the C4-pivaloylated galactosyl cation, the C4,C6-dipivaloylated counterpart is showing both rearranged and dioxolenium-type structures, indicative for remote participation.

Although it was not possible to study the structure of galactosyl cations of TFA precursors experimentally, their structures and stabilities can be studied using DFT calculations. These calculations reveal that oxocarbenium structures, in which no participation occurs, are either more stable or of similar stability than the dioxolenium structures for these glycosyl cations. Furthermore, the bond length between the carbonyl oxygen and the anomeric carbon is decreased in dioxolenium structures of trifluoroacetylated galactosyl cations, compared to pivaloylated structures, indicating that remote participation would be weaker in this case. Yet, the galactosyl building block carrying a C4-TFA group exhibits an 1,2-*cis* selectivity that is higher than that of its C4-acetylated counterpart, a building block for which remote participation occurs.^[87] To explain this discrepancy, a second mechanism needs to be invoked, which involves triflate intermediates. This mechanism has been discussed in Chapter 2. As electron-withdrawing substituents favor the formation of triflate intermediates,^[296] it is likely that such intermediates are generated from trifluoroacetylated donors. Here, β -triflates would be preferentially formed, leading to a comparably high α -selectivity.

Based on these results, a modified galactose building block with a pivaloyl group at the C4-position was developed, which can be readily implemented in AGA workflows. In a proof-of-principle experiment, an α (1,3)-galactose trisaccharide was generated with complete α -selectivity at high yield (69%) (see Figure 4 of the related publication). Hence, it was possible to develop an α -selective galactose building block based on rational considerations, and its stereoselectivity was explained using a combination of cryogenic IR spectroscopy and DFT calculations.

The Influence of the Electron Density in Acyl Protecting Groups on the Selectivity of Galactose Formation

Kim Greis, Sabrina Lechnitz, Carla Kirschbaum, Chun-Wei Chang, Mei-Huei Lin, Gerard Meijer, Gert von Helden, Peter H. Seeberger,* and Kevin Pagel*



Cite This: *J. Am. Chem. Soc.* 2022, 144, 20258–20266



Read Online

ACCESS |

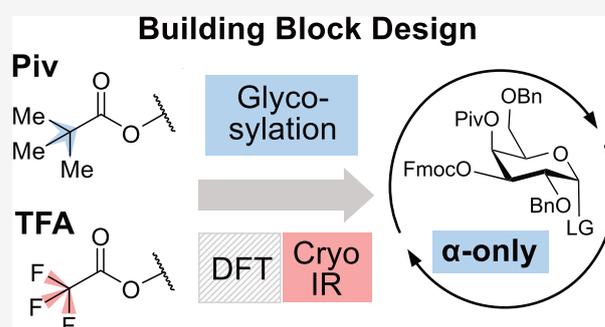
Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: The stereoselective formation of 1,2-*cis*-glycosidic bonds is a major bottleneck in the synthesis of carbohydrates. We here investigate how the electron density in acyl protecting groups influences the stereoselectivity by fine-tuning the efficiency of remote participation. Electron-rich C4-pivaloylated galactose building blocks show an unprecedented α -selectivity. The trifluoroacetylated counterpart with electron-withdrawing groups, on the other hand, exhibits a lower selectivity. Cryogenic infrared spectroscopy in helium nanodroplets and density functional theory calculations revealed the existence of dioxolenium-type intermediates for this reaction, which suggests that remote participation of the pivaloyl protecting group is the origin of the high α -selectivity of the pivaloylated building blocks.

According to these findings, an α -selective galactose building block for glycosynthesis is developed based on rational considerations and is subsequently employed in automated glycan assembly exhibiting complete stereoselectivity. Based on the obtained selectivities in the glycosylation reactions and the results from infrared spectroscopy and density functional theory, we suggest a mechanism by which these reactions could proceed.



INTRODUCTION

The chemical synthesis of carbohydrates requires stereochemical control during glycoside formation. While neighboring-group participation is key to synthesizing 1,2-*trans* glycosides, methods to generate 1,2-*cis* glycosides are less reliable. Many biologically important oligosaccharides contain 1,2-*cis* linkages, such as the blood group systems¹ or bacterial lipopolysaccharide antigens.^{2,3} Participation of remote acyl groups,^{4–6} chiral auxiliaries,⁷ or 4,6-benzylidene^{8,9} protecting groups helps to increase the ratio of 1,2-*cis* glycosides. Previous studies on galactose building blocks suggest that participating acetyl protecting groups at the C4 position lead to *cis*-selectivity (defined as α -selectivity for galactose).^{10,11} The remote acetyl protecting group is shielding the positive charge of the anomeric carbon by forming a temporary covalent bond that prevents nucleophiles from attacking the 1,2-*trans*-side, leading to 1,2-*cis*-selectivity. However, the ability of acetyl groups to remotely participate is limited, as the selectivity differs dramatically depending on the strength of the nucleophile. This is problematic because efficient glycan synthesis requires full stereocontrol. Total stereoselectivity is particularly important in sequential synthetic methods such as automated glycan assembly (AGA)¹² to avoid the formation of complex mixtures of stereoisomers, which leads to a drastic drop in overall yield.

Besides high yields and an excellent stereoselectivity, differential protecting groups are a requirement for implementation in AGA. Therefore, strategies involving 4,6-*O*-*tert*-butylsilylene (DTBS) protecting groups, showing full α -selectivity in galactosylations,^{13,14} cannot be employed, as this protecting group would yield two nucleophilic OH groups after deprotection. Moreover, AGA requires an excess amount of promoters (NIS and TfOH); however, DTBS is labile toward such acidic conditions. While the position of the acyl protecting group and the influence of nucleophile strength have been investigated before,^{10,15} the effect of electron density on acyl protecting groups in galactosylations has been ignored.

Generally, the mechanism of glycosylation reactions is not entirely understood to date.^{16,17} It is generally accepted that the mechanism is governed by an S_N1 – S_N2 continuum¹⁸ that can be shifted toward one side by adjusting various parameters. When it comes to the formation of α -selective linkages in galactose building blocks, a consistent increase in selectivity has been observed for C4-acylated building blocks.^{4,10,11,15,19,20}

Received: June 2, 2022

Published: October 27, 2022



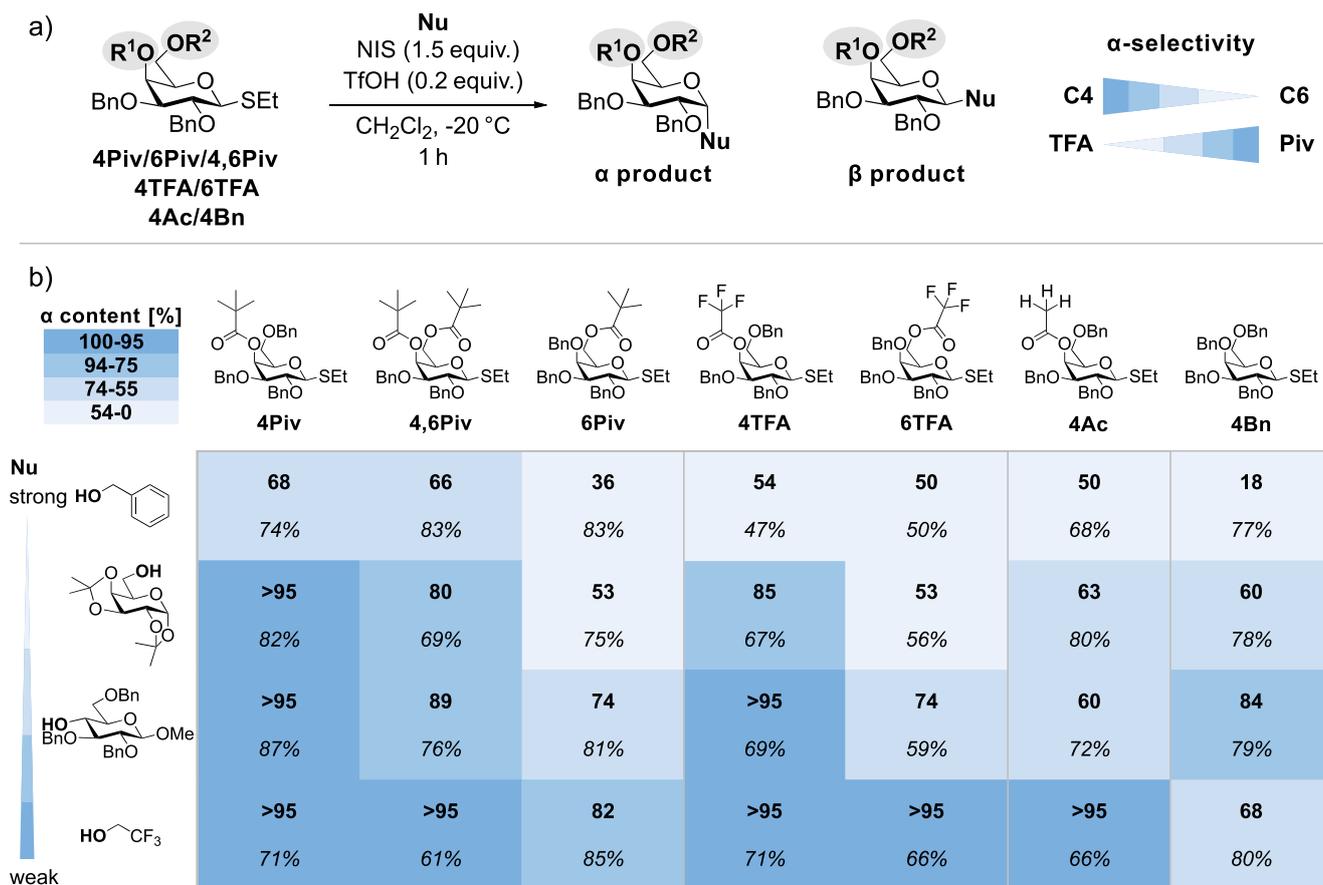


Figure 1. (a) Glycosylation conditions of galactose building blocks (17.5 mM) carrying either Piv, TFA, Ac, or Bn protecting groups at the C4 and C6 positions ($R^{1/2}$). (b) Stereochemical outcome (α content in relation to β -content and yield) of the glycosylation reactions of the respective building blocks with different nucleophiles (Nu) of decreasing strength. Piv protecting groups at the C4 position lead to increased α -selectivity, while the selectivity is reduced for building blocks carrying the less electron rich TFA, Ac, or Bn protecting groups at that position. Protecting groups at the C6 position do not increase the α -selectivity. Results for 4,6TFA cannot be shown, as this building block is rapidly decomposing.

Strong evidence suggests that this selectivity is aided by remote participation of the C4-acyl group.²¹ On the other hand, it has been reported that the formation of β -triflates^{22–24} can lead to α -selectivity upon the attack of a nucleophile. Evidence for remote participation has been provided indirectly by bridged side products extracted from glycosylation reactions^{25,26} or directly by low-temperature NMR experiments in organic solvents^{19,27} and gas-phase infrared spectroscopy.^{10,15,28,29} It should be noted that the intermediate showing remote participation in solution can only be observed under very limited circumstances, as the lifetime of the glycosyl cation is usually shorter than the relaxation time in NMR experiments.²⁷ Furthermore, glycosyl cations with remote acetyl groups were stabilized in super acids.³⁰ Here, remote participation was not observed. However, all carbonyl groups are protonated, which drastically reduces their nucleophilicity. Hence, they are unable to engage in remote participation.

Here, we systematically investigate how electron-donating and electron-withdrawing substituents in acyl protecting groups influence the stereoselectivity of galactosylations. Custom-tailored galactosyl building blocks were investigated carrying pivaloyl (trimethylacetyl, Piv) or trifluoroacetyl (TFA) protecting groups at C4, C6, or both positions, while the remaining hydroxyl groups are benzylated. The building blocks (4/6/4,6Piv and 4/6/4,6TFA) were assessed in glycosylation test reactions to determine their selectivity with

four distinct nucleophiles. Their selectivities were compared to acetylated and benzylated building blocks 4Ac and 4Bn. In parallel, the glycosyl cation intermediates of the 4/6/4,6Piv building blocks were structurally characterized using cryogenic gas-phase infrared (IR) spectroscopy in helium nanodroplets and density functional theory (DFT).^{31,32} This approach allows investigating the intermediate of S_N1 -like glycosylation reactions. Finally, the most promising building block, 4Piv, was used in automated glycan assembly¹² to synthesize an $\alpha(1,3)$ -D-trigalactopyranoside.

METHODS

The instrumental setup for gas-phase IR spectroscopy in helium nanodroplets has been described previously^{33,34} (see SI and Figure S1). Briefly, glycosyl cations are generated by nanoelectrospray ionization and subsequent in-source fragmentation of thioglycoside galactose building blocks. The mass-to-charge ratio of the generated ions can be monitored by a time-of-flight mass spectrometer. A quadrupole mass filter allows for mass-to-charge selection of the ions of interest that are then guided into a hexapole ion trap, where the ions are cooled to ca. 90 K by collisions with the helium buffer gas. A beam of superfluid helium nanodroplets (0.37 K) is generated by a pulsed Even–Lavie valve.³⁵ The beam is guided through the ion trap, where the droplets pick up the ions and lead them to a detection region, where the beam of doped droplets overlaps with an IR beam generated by the tunable Fritz Haber Institute free-electron laser³⁶ (FHI FEL). The interaction with resonant IR photons

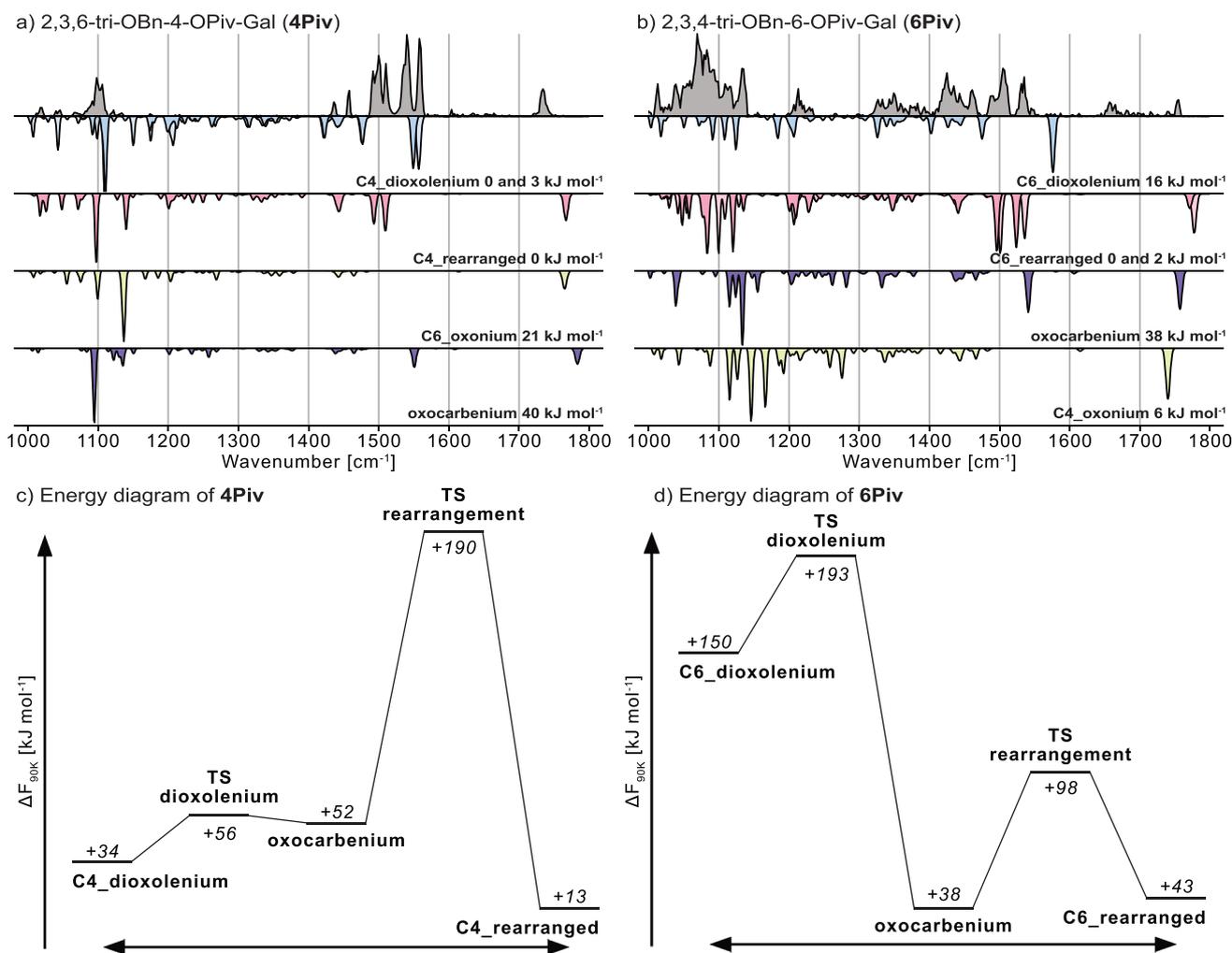


Figure 2. Cryogenic infrared spectra of (a) **4Piv** and (b) **6Piv** galactosyl cations (gray). Computed infrared spectra are shown in the inverted traces for structures showing remote acyl participation (dioxolenium, blue), rearrangement (red), remote benzyl participation (oxonium, yellow), and no participation (oxocarbenium, purple). For **4Piv** the positive charge at the anomeric carbon is mainly stabilized by remote participation of the C4-pivaloyl protecting group. However, further signals can be observed in the experimental spectrum that can be linked to an isoenergetic rearranged structure. The rearranged structure is the dominant motif in the experimental spectrum of the **6Piv** galactosyl cation. Energy diagrams of (c) **4Piv** and (d) **6Piv** show the barriers for remote participation and rearrangement (note that the minimum structures in the diagram do not necessarily correspond to the global minimum). The barrier for remote participation in **4Piv** (C4_dioxolenium) is surprisingly small.

(1000–1800 cm^{-1}) leads to the release of the probed glycosyl cations, which are subsequently detected by a second time-of-flight mass spectrometer. The ion count is plotted against the wavenumber to yield an IR spectrum. By comparison with computed harmonic frequencies, the structure of the probed ion can be determined. This approach and others based on infrared multiple photon dissociation (IRMPD) spectroscopy have successfully been applied to probe the structure of glycosyl cations exhibiting remote and neighboring group participation.^{10,15,28,29,34,37–41}

For structural assignment, the experimental IR spectra are compared with theoretical spectra derived from computed structures. A genetic algorithm⁴² was employed to sample the conformational space of glycosyl cations at the PBE+vdW^{TS43,44} level of theory using *light* basis set settings, implemented in FHI-aims.⁴⁵ The geometries of a subset of low-energy structures were reoptimized and their frequencies computed at the PBE0+D3/6-311+G(d,p)^{46,47} level of theory in Gaussian 16.⁴⁸ All calculated IR spectra are normalized and scaled by an empirical factor of 0.965.^{10,34} The reoptimized geometries were used to compute accurate single-point energies at the DLPNO-CCSD(T)/Def2-TZVPP^{49,50} level of theory in ORCA.⁵¹ Pyranose ring puckers are assigned according to Cremer–Pople coordinates.⁵² The free energy at 90 K is used as a relevant parameter

to rank the reoptimized structures. Detailed information on the computed structures, such as energetics, ring puckers, or *xyz*-coordinates, can be found in the SI.

RESULTS AND DISCUSSION

Glycosylation Reactions. Six galactose building blocks carrying pivaloyl or trifluoroacetyl protecting groups at C4, C6, or both positions were synthesized (see SI). Furthermore, two other galactose building blocks, known from previous studies, that are fully benzylated or carry an acetyl group at the C4 position were synthesized. The building blocks were employed in glycosylation reactions with four distinct nucleophiles of different strengths (Figure 1). Generally, weak nucleophiles lead to a higher α -selectivity, which decreases with increasing strength of the nucleophile, in agreement with previous reports.¹⁵ Glycosyl alcohols are weak nucleophiles,⁵³ and hence the observed trend is desirable for the synthesis of α -glycosidic bonds. Furthermore, the α -selectivity is higher for building blocks with an acyl protecting group at C4 than for those with the protecting group at C6. Interestingly, for **4,6Piv**,

43

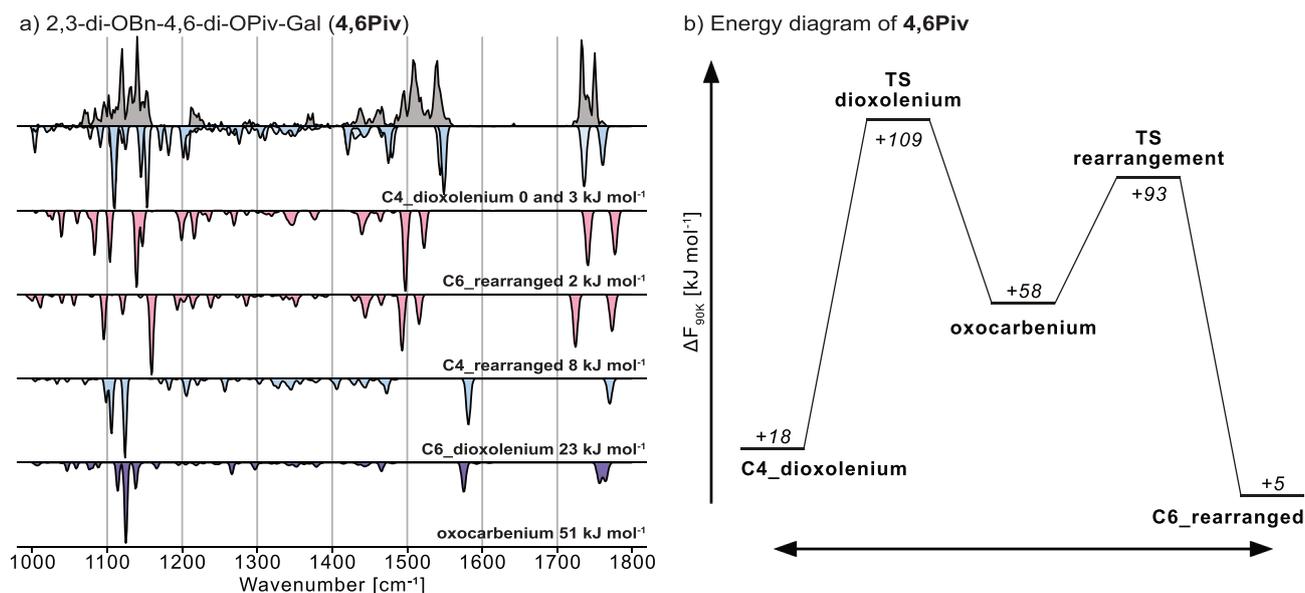


Figure 3. (a) Cryogenic infrared spectrum of the (a) **4,6Piv** galactosyl cation (gray). Computed infrared spectra are shown in the inverted traces for structures showing remote acyl participation (dioxolenium, blue), rearrangement (red), and no participation (oxocarbenium, purple). The positive charge at the anomeric carbon is mainly stabilized by remote participation of the C4-pivaloyl protecting group. However, further signals in the experimental spectrum can be linked to an isoenergetic rearranged structure. (b) Energy diagram of **4,6Piv** showing the barrier for remote participation and rearrangement (note that the minimum structures in the diagram do not necessarily correspond to the global minimum).

the α -selectivity is lower than for **4Piv**, although an inverse trend has been reported for similar acetyl building blocks.¹⁰

Our research groups^{4,10} and others^{11,15,29} have found strong evidence suggesting that remote participation of the C4 protecting group is the origin of the increased α -selectivity of C4-acylated galactose building blocks. For C6-acyl groups, such an effect is not observed. Other groups reported strong evidence that the formation of β -triflates contributes to α -selectivity in glycosylations.^{22–24} The central question of this work is how α -selectivity can be modulated by alterations in the electron density of the acyl protecting groups. For building blocks carrying the acyl group at the C4 position, the electron-rich **4Piv** provides high α -selectivity. The electron-withdrawing **4TFA**, on the other hand, results in significantly lower α -selectivity for the strong nucleophile benzyl alcohol and a sugar nucleophile carrying a free OH group at C6. This result implies that an increase in the electron density on the carbonyl oxygen of the acyl group more likely leads to the formation of a covalent bond with the positively charged anomeric carbon and with that a better shielding of the β -side. However, counterintuitively, the α -selectivity for the **4TFA** building block is higher than expected. There are two possible explanations for this unexpected behavior. Either the electron-withdrawing groups do not inhibit remote participation, but rather weaken it (leading to an equilibrium, where both structures with and without remote participation are present), or a second mechanism, based on α -selective β -triflates could play a role here because their formation is favored due to the longer lifetime of the oxocarbenium species without remote participation.

To elucidate which mechanism is more likely, we performed the same set of test reactions on a **4Ac** building block. Evidence for remote participation on this and similar building blocks has previously been reported.^{4,10,11,15} Solely based on the electron density, this building block would exhibit an α -selectivity that is higher than that of **4TFA** but lower than that

of **4Piv**. Interestingly, its selectivity is lower than that of **4TFA**. This finding suggests that in the case of **4TFA** remote participation does not play a role, but rather the formation of β -triflates. This finding is corroborated by a previous study on glucosyl donors, where it was found that dioxolenium ions are the intermediate of donors carrying electron-rich protecting groups, while triflates are the major intermediates when electron-withdrawing groups are used.⁵⁴

The decreased selectivity for **4,6Piv** compared to **4Piv** and **4TFA** can likely be attributed to the steric effects because of the bulky Piv group. Remote participation of the C4-pivaloyl group is less efficient in this building block, as the C6-pivaloyl is partially blocking its trajectory for an intramolecular attack.

In contrast to the C4-acyl variant, a participating protecting group at C6 seems to have no (**6TFA**) or adverse effects (**6Piv**) on the α -selectivity. Adverse effects of C6-acetyl groups on the α -selectivity have been previously reported.^{10,11} With strong nucleophiles, **6Piv** predominantly forms β -products, whereas **6TFA** is not stereoselective. For weaker nucleophiles, the α -selectivity increases, which might be due to counterions or the formation of β -triflates as previously reported.^{18,53,55}

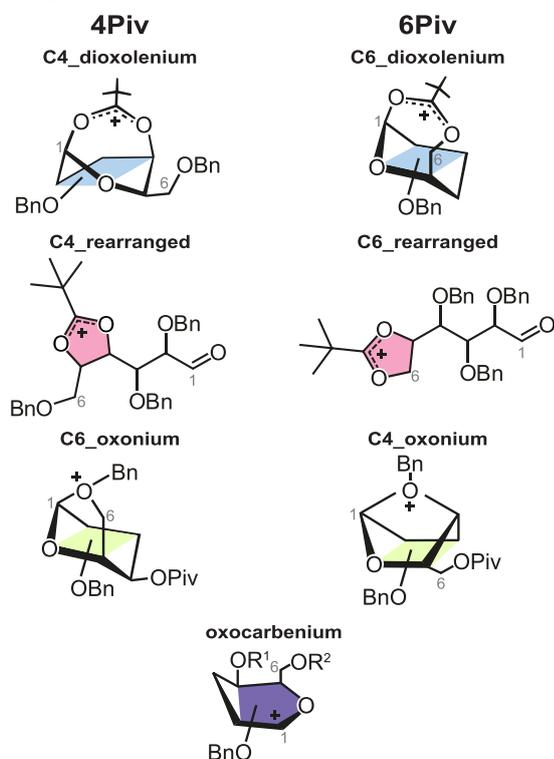
As a reference, glycosylation reactions were performed on a fully benzylated galactose building block (**4Bn**). This building block generally exhibits a decreased α -selectivity compared to its C4-acylated counterparts, indicating the importance of a C4-acyl group in achieving high α -selectivity in galactosylations. Intriguingly, the glycosylation reaction with a sugar nucleophile carrying a free OH group at C4 shows a surprisingly high α -selectivity of 84%. Further, it is important to highlight the high yield of the coupling reactions of **4Piv** with sugars, as this is a crucial requirement for AGA.

Cryogenic Infrared Spectroscopy and Density Functional Theory Investigations of Glycosyl Cations. In parallel to the test reactions, the intermediates of the glycosylations—the glycosyl cations—were structurally characterized by cryogenic IR spectroscopy and DFT calculations.

Thioglycoside precursors were subjected to in-source fragmentation after nano-electrospray ionization. Surprisingly, only in the case of pivaloylated building blocks this approach leads to the desired glycosyl cation intermediates. Trifluoroacetylated molecules on the other hand did not fragment sufficiently or decomposed by losing TFA (Figures S2 and S3). Therefore, only galactosyl cations of **4Piv**, **6Piv**, and **4,6Piv** were subjected to cryogenic IR spectroscopy (Figures 2a,b and 3a). The glycosyl cations of **4Ac** and **4Bn** were already probed with the same method in a previous publication.¹⁰ The experimental spectra can be divided into two main regions: (1) the fingerprint region (1000–1400 cm⁻¹), which is predominantly populated by C–O and C–C stretching modes as well as C–H bending vibrations. Due to the complex nature of carbohydrates, this region is usually very challenging to model.^{56,57} (2) The functional group region (1400–1800 cm⁻¹) contains most of the diagnostic vibrations of the investigated ions, such as symmetric and antisymmetric dioxolenium $\nu(\text{O}-\text{C}-\text{O}^+)$ and carbonyl stretches $\nu(\text{C}=\text{O})$. To determine the structure of the probed glycosyl cations, the IR spectra are compared to harmonic frequencies of sampled structures. The sampling mainly yielded dioxolenium structures, which exhibit remote participation of the C4- or the C6-acyl protecting group and oxocarbenium structures (Scheme 1), where no participation occurs at the anomeric carbon (C1). Furthermore, oxonium structures that feature participation of the C4- or C6-benzyl protecting groups at C1 were generated.

For **4Piv**, C4-dioxolenium structures are the lowest in energy and match the experimentally resolved signals at 1090–1110, 1540, and 1558 cm⁻¹ well (Figure 2a). The presence of

Scheme 1. Structures That Can Be Adopted by (Left) 4Piv and (Right) 6Piv Galactosyl Cations^a



^aOxocarbenium structures can be adopted by both cations. Except for oxonium structures, all five structures can be adopted by **4,6Piv**.

two absorption bands diagnostic for antisymmetric dioxolenium stretches is likely due to the presence of two conformers carrying this structural motif. However, the signals at 1492–1510 cm⁻¹ cannot be explained with the sampled structures, and also the carbonyl stretch at 1734 cm⁻¹ shows that another type of structures must be present. In a previous study,¹⁵ it was suggested that acyl groups may attack the C5 atom in glycosyl cations, leading to ring opening and an aldehyde as a product. Therefore, these rearranged ions have been added to the list of structural motifs and were sampled as well (Scheme 1). As the rearranged ions feature a five-membered dioxolenium moiety (compared to the seven-membered dioxolenium moiety observed for remote participation), they are expected to show diagnostic absorption bands in the functional group region.⁵⁸ Indeed, the C4-rearranged structure is isoenergetic to the lowest-energy C4-dioxolenium structure, and its dioxolenium and carbonyl stretches match the remaining experimentally resolved absorption bands. The observations indicate that the spectra observed for **4Piv** are resulting from a mixture of C4-dioxolenium ions and rearrangement products present in the hexapole ion trap after ionization.

The C4-dioxolenium structure is in line with the α -selectivity observed in the glycosylation reactions. In contrast, our results indicate that the C4-rearrangement product is unique to the gas-phase conditions, as none of the expected side products is observed in the test reactions. The literature on the presence of rearranged structures in the condensed phase is generally scarce. Ring opening occurring to a minor degree after the glycosylation reaction of a glucosyl donor carrying three trichloroacetimidate groups has been reported.⁵⁹ Based on our results the rearranged structure does not seem to play a dominant role in the here reported glycosylation reactions. Other structural motifs including C6-oxonium and oxocarbenium ions were sampled, and their harmonic frequencies are compared to the experimental infrared spectrum. Contrary to the dioxolenium and rearranged structures, their computed spectra do not match with the experiment. Based on this result and their higher relative free energy of 21 and 40 kJ mol⁻¹, respectively, their presence in the ion trap can be ruled out.

For **6Piv**, the computed harmonic frequencies of the sampled C6-dioxolenium structure do not match the experimental spectrum (Figure 2b). Furthermore, the corresponding C6-rearranged structure is stabilized by –16 kJ mol⁻¹, and its frequencies match the experimentally resolved absorption bands at 1421–1461, 1506, and 1533 cm⁻¹ well. The oxonium structure is surprisingly low in energy (+6 kJ mol⁻¹), but can, like the oxocarbenium structure (+38 kJ mol⁻¹), be ruled out due to its poor spectral match. Hence, C6-acyl participation is unlikely to exist for Piv groups, in line with the poor α -selectivity of these building blocks.

These findings are corroborated by computed transition states that are connecting dioxolenium, oxocarbenium, and rearranged structures for **4Piv** and **6Piv** glycosyl cations displayed in the energy diagrams in Figure 2c,d. The geometries that are connected by the transition states do not necessarily correspond to the global minima that we previously sampled. For **4Piv**, the diagram shows that the surface is shallow except for the transition state leading from the oxocarbenium to the rearranged structure. The barrier for remote participation is surprisingly small (+4 kJ mol⁻¹), and therefore remote participation is very likely occurring for this species. Hence, the high kinetic barrier that was postulated⁶⁰

for this type of interaction does at least for the gas phase not exist. The relative barrier of +138 kJ mol⁻¹ for rearrangement can according to previous studies^{61,62} be overcome using in-source fragmentation, leading to the thermodynamically stable rearranged ion. Once the energy in the ion source is high enough to overcome the transition state, thermodynamically stable species can coexist in the ion trap.

For **6Piv**, the formation of the rearranged product is favored both kinetically and thermodynamically. Furthermore, in previous studies on similar acetylated building blocks, the rearrangement was only observed for C6-acetylated galactosyl cations, whereas it was not reported for its C4-acetylated counterparts.^{10,15} Hence, the results suggest that increasing the electron density within the acyl protecting group enhances remote participation (in both the gas and the condensed phase), but also facilitates a gas-phase rearrangement of the ions. However, the latter does not have an implication on condensed-phase reactivity of the precursors.

For **4,6Piv**, the experimental IR signature (Figure 3a) is similar to that of **4Piv**. The absorption band at 1540 cm⁻¹ is diagnostic for C4-dioxolenium structures, whereas the absorption bands at 1495 and 1508 cm⁻¹ are diagnostic for the five-membered dioxolenium motif in rearranged structures. Although the predicted frequencies for C6- and C4-rearranged **4,6Piv** are similar, the C6-rearranged structure matches slightly better, especially in the carbonyl stretch region, and is also lower in energy than the C4-rearranged analog (+2 vs +8 kJ mol⁻¹). The harmonic frequencies of computed low-energy C6-dioxolenium and oxocarbenium ions do not match the experimental data, and their relative free energies are significantly higher than those of the C4-dioxolenium and rearranged structures. Hence, similarly to **4Piv**, this result suggests that the formation of C4-dioxolenium intermediates with remote participation of the C4-pivaloyl group contributes to the α -selectivity of **4,6Piv** that can be observed in condensed-phase glycosylation reactions. Transition states connecting dioxolenium, oxocarbenium, and rearranged structures and subsequent energy diagrams were also computed for **4,6Piv** (Figure 3b). Here, both the transition states and the products are similar in energy, explaining their coexistence in the experiment. Furthermore, the barrier of C4-dioxolenium ion formation (remote participation) from oxocarbenium ions is significantly higher for **4,6Piv** than for **4Piv** (difference of +47 kJ mol⁻¹). This finding highlights that the steric demand of two pivaloyl groups on one glycosyl cation is decreasing the efficiency of remote participation, likely being the cause for the decreased α -selectivity of **4,6Piv** compared to **4Piv** in glycosylation reactions.

Although it was not possible to generate glycosyl cations out of the TFA protected building blocks for cryogenic IR spectroscopy, it is still possible to compute their structures and energetics to rationalize the observed reactivity in glycosylation reactions. The energetics shown in Tables S3–S5 (**4/6/4,6Piv**) and Tables S6–S8 (**4/6/4,6TFA**) show that remote participation of the C4-pivaloyl leading to dioxolenium structures is favored by 40–51 kJ mol⁻¹ over oxocarbenium structures in which no participation takes place. Structures with remote participation of C4-TFA can be generated, but their relative energetics are similar (2–4 kJ mol⁻¹) to oxocarbenium structures. Interestingly, for **4TFA** C6-oxonium structures are stabilized by –24 kJ mol⁻¹ compared to low-energy C4-dioxolenium structures. Such a structure was previously reported for a fully benzylated galactosyl cation,

without a clear implication on the condensed phase reactivity.¹⁰ Furthermore, the calculations show that the C–O bond between the acyl protecting group and the anomeric carbon is, in comparison to **Piv**, significantly weakened when remote participation of TFA occurs (1.61 vs 1.52 Å). These results, as well as the energy diagram shown in Figure S7a, indicate that remote participation of the C4-trifluoroacetyl group is thermodynamically unfavored, while the energy of the transition state leading to a C4-dioxolenium structure is not particularly high. Furthermore, if remote participation takes place, the effect is weaker than for the C4-pivaloylated counterparts, indicating that it would be less efficient and therefore lead to a decreased α -selectivity. Yet, even though the α -selectivity of **4TFA** is clearly lower than that of **4Piv**, it is higher than one would expect without remote participation and higher than for **4Ac**, a precursor for which remote participation has previously been reported.^{10,15}

The gas-phase conditions under which we study the glycosyl cations are not identical to those in the condensed phase during glycosylation reactions, yet there are clear correlations that are worth pointing out. In this study, it was found that C4-pivaloylated glycosyl cations are stabilized by remote participation in the gas phase. If that intermediate is attacked by a nucleophile, the α -product would preferentially be formed. Based on previous studies, there is a consistent trend in the condensed phase that C4-acylated species are more α -selective than their non- or differently acetylated counterparts.^{4,10,11,15} Furthermore, the bridged dioxolenium intermediate was linked to the α -selectivity observed for these building blocks by condensed phase studies in organic solvents using low-temperature NMR spectroscopy.¹⁹ Because of those findings, we are convinced that remote participation is at least contributing to the selectivity of C4-acylated building blocks.

The glycosylation reaction and its selectivity are governed by an S_N1–S_N2 continuum, and the herein presented selectivities are illustrating this continuum. The selectivity of the S_N1 side is dominated by the structure of the glycosyl cation, whereas the S_N2 side is dominated by the structure of the glycosyl triflates.^{18,60} In the condensed phase, the lifetime of the glycosyl cation is very short, leading to the quick formation of a thermodynamically stable intermediate that is potentially stereoselective.²⁴ The exact mechanism of the glycosylation reaction is currently unclear. Based on the current knowledge it is likely that there are at least two pathways that are contributing to the selectivity observed in glycosylation reactions, depending on various parameters, such as the donor and acceptor reactivities, temperature, solvent, or promoters.^{23,63}

This and other studies showed that remote participation of C6-acyl groups does not occur.^{10,11,15} Except for weak nucleophiles, C6-acylated building blocks are not α -selective. For building blocks carrying a C4-acyl group, it is established that remote participation occurs. The fundamental question of this manuscript is how the electron density in acyl protecting groups influences the stereochemical outcome of a glycosylation reaction. Here, the electron density increases as **4TFA** < **4Ac** < **4Piv**, while the α -selectivity increases as **4Ac** < **4TFA** < **4Piv**. As a consequence, it is not the α -selectivity that increases with increasing electron density on the C4-acyl protecting group, but rather the strength of remote participation. Also, based on these findings, remote participation alone can explain the high α -selectivity of the electron-rich **4Piv**, but not that of the less selective electron-deficient

4TFA building block. Here, the longer lifetime of oxocarbenium-type intermediates without remote participation could favor the formation of β -triflates, leading to an increased α -selectivity.⁵⁴ For 4Ac, on the other hand, remote participation was previously established, but is not as selective as for 4Piv due to the decreased electron density.

Automated Glycan Assembly. The combination of the results on the nature and position of the acyl groups on the α -selectivity led to the design of the 4Piv building block 1 (Figure 4), which can be readily implemented in AGA

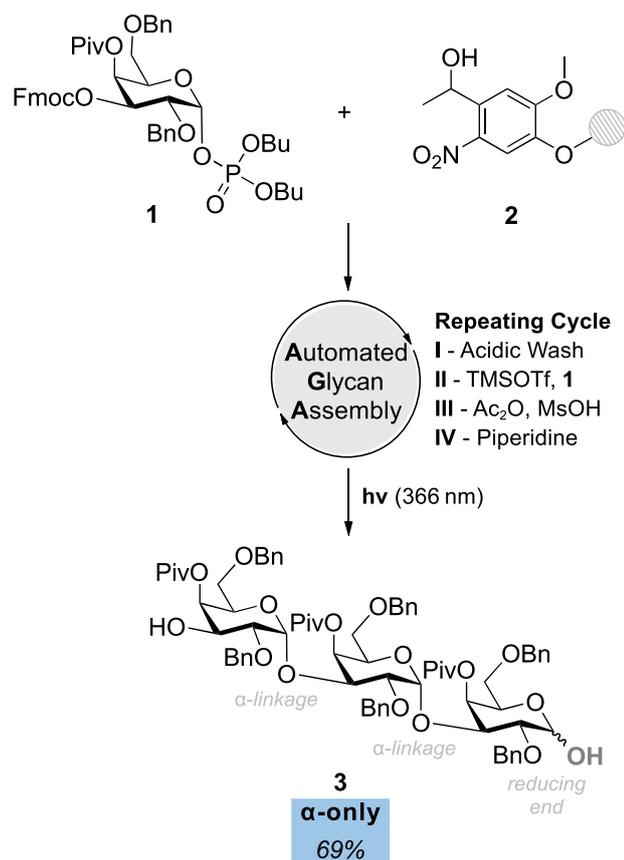


Figure 4. Automated glycan assembly employing the 4Piv galactosyl building block 1 leads to an $\alpha(1,3)$ -D-trigalactopyranoside 3 with total α -selectivity (determined after the AGA assembly, but before isolation) and a yield of 69% in 6 h using solid support 2 and a coupling cycle consisting of module I, acidic wash; II, glycosylation; III, capping; and IV, Fmoc-deprotection, followed by photocleavage (for more detailed conditions see S1).

workflows due to differential protecting groups. Temporary fluorenylmethoxycarbonyl (Fmoc) protection at the C3 position ensured regioselective extension, while the more reactive phosphate leaving group at C1 ensured high yields. Employed in AGA (Figure 4 and S1), 1 was used to assemble the $\alpha(1,3)$ -galactose trisaccharide 3 at high yield and with full α -selectivity.

CONCLUSIONS

This study shows that electron-donating substituents on participating acyl protecting groups increase the efficiency of remote participation, leading to a higher α -selectivity in glycosylation reactions, as shown for pivaloyl groups. Computational results suggest that electron-withdrawing

substituents, such as trifluoroacetyl groups, on the other hand, deactivate remote participation, possibly leading to a decrease in selectivity of the reaction. However, the 4TFA building block is more α -selective than expected, which can be attributed to a favored formation of β -triflates. The presented data confirm that the C4 position plays a more important role in inducing selectivity than the C6 position. In the gas phase, remote participation of the C4-pivaloyl group can be observed, suggesting a role of that effect in the high α -selectivity for the 4Piv building block. Furthermore, the computed barrier for remote participation is very low, and therefore it can be assumed that it is a fast process. The increased electron density in pivaloyl groups also leads to an increased rearrangement of glycosyl cations in the gas phase, for which no influence on the reactivity in solution was observed. The mechanistic insights were used to tailor a 4Piv building block that was successfully employed in AGA to synthesize an $\alpha(1,3)$ -trigalactopyranoside with total α -selectivity. In summary, our results show how α -selective building blocks can be developed by rational design and thus provide guiding points on how to fine-tune the selectivity and efficiency of glycosylations.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.2c05859>.

A detailed description of the experiment, mass spectra, computed energetics, spectra, and structures (PDF) xyz-coordinates (XYZ)

AUTHOR INFORMATION

Corresponding Authors

Kevin Pagel – Institute of Chemistry and Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany; Fritz Haber Institute of the Max Planck Society, 14195 Berlin, Germany; orcid.org/0000-0001-8054-4718; Email: kevin.pagel@fu-berlin.de

Peter H. Seeberger – Institute of Chemistry and Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany; Max Planck Institute of Colloids and Interfaces, 14476 Potsdam, Germany; Email: peter.seeberger@mpikg.mpg.de

Authors

Kim Greis – Institute of Chemistry and Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany; Fritz Haber Institute of the Max Planck Society, 14195 Berlin, Germany; orcid.org/0000-0002-9107-2282

Sabrina Lechnitz – Institute of Chemistry and Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany; Max Planck Institute of Colloids and Interfaces, 14476 Potsdam, Germany

Carla Kirschbaum – Institute of Chemistry and Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany; Fritz Haber Institute of the Max Planck Society, 14195 Berlin, Germany; orcid.org/0000-0003-3192-0785

Chun-Wei Chang – Institute of Chemistry and Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany; orcid.org/0000-0002-0253-3463

Mei-Huei Lin – Institute of Chemistry and Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany; Max Planck Institute of Colloids and Interfaces, 14476 Potsdam, Germany; orcid.org/0000-0002-0842-1964

Gerard Meijer – Fritz Haber Institute of the Max Planck Society, 14195 Berlin, Germany; orcid.org/0000-0001-9669-8340

Gert von Helden – Fritz Haber Institute of the Max Planck Society, 14195 Berlin, Germany; orcid.org/0000-0001-7611-8740

Complete contact information is available at:
<https://pubs.acs.org/10.1021/jacs.2c05859>

Author Contributions

K.G. and S.L. contributed equally.

Funding

Open access funded by Max Planck Society.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank Dr. Wieland Schöllkopf and Sandy Gewinner for operating the FHI FEL. K.G. is grateful to the Fonds National de la Recherche (FNR), Luxembourg, for funding the project GlycoCat (13549747). S.L. acknowledges the CRC 1449 funded by the DFG. C.K. is grateful for financial support by Fonds der Chemischen Industrie. K.P. acknowledges generous funding by the European Research Council, ERC-2019-CoG-863934-GlycoSpec.

REFERENCES

- (1) Kappler, K.; Hennet, T. Emergence and significance of carbohydrate-specific antibodies. *Genes Immun.* **2020**, *21* (4), 224–239.
- (2) Zou, X.; Qin, C.; Pereira, C. L.; Tian, G.; Hu, J.; Seeberger, P. H.; Yin, J. Synergistic Glycosylation as Key to the Chemical Synthesis of an Outer Core Octasaccharide of *Helicobacter pylori*. *Chem. Eur. J.* **2018**, *24* (12), 2868–2872.
- (3) Paramonov, N.; Bailey, D.; Rangarajan, M.; Hashim, A.; Kelly, G.; Curtis, M. A.; Hounsell, E. F. Structural analysis of the polysaccharide from the lipopolysaccharide of *Porphyromonas gingivalis* strain W50. *Eur. J. Biochem.* **2001**, *268* (17), 4698–707.
- (4) Hahm, H. S.; Hurevich, M.; Seeberger, P. H. Automated assembly of oligosaccharides containing multiple cis-glycosidic linkages. *Nat. Commun.* **2016**, *7*, 12482.
- (5) Zhu, Y.; Delbianco, M.; Seeberger, P. H. Automated Assembly of Starch and Glycogen Polysaccharides. *J. Am. Chem. Soc.* **2021**, *143* (26), 9758–9768.
- (6) Lin, M. H.; Chang, C. W.; Chiang, T. Y.; Dhurandhare, V. M.; Wang, C. C. Thiocarbonyl as a Switchable Relay-Auxiliary Group in Carbohydrate Synthesis. *Org. Lett.* **2021**, *23* (19), 7313–7318.
- (7) Boltje, T. J.; Kim, J. H.; Park, J.; Boons, G. J. Chiral-auxiliary-mediated 1,2-cis-glycosylations for the solid-supported synthesis of a biologically important branched alpha-glucan. *Nat. Chem.* **2010**, *2* (7), 552–7.
- (8) Crich, D.; Sun, S. Direct Formation of β -Mannopyranosides and Other Hindered Glycosides from Thioglycosides. *J. Am. Chem. Soc.* **1998**, *120* (2), 435–436.
- (9) Crich, D.; Chandrasekera, N. S. Mechanism of 4,6-O-benzylidene-directed beta-mannosylation as determined by alpha-deuterium kinetic isotope effects. *Angew. Chem. Int. Ed.* **2004**, *43* (40), 5386–9.
- (10) Marianski, M.; Mucha, E.; Greis, K.; Moon, S.; Pardo, A.; Kirschbaum, C.; Thomas, D. A.; Meijer, G.; von Helden, G.; Gilmore, K.; Seeberger, P. H.; Pagel, K. Remote Participation during Glycosylation Reactions of Galactose Building Blocks: Direct Evidence from Cryogenic Vibrational Spectroscopy. *Angew. Chem. Int. Ed.* **2020**, *59* (15), 6166–6171.
- (11) Li, Z.; Zhu, L.; Kalikanda, J. Development of a highly α -selective galactopyranosyl donor based on a rational design. *Tetrahedron Lett.* **2011**, *52* (43), 5629–5632.
- (12) Plante, O. J.; Palmacci, E. R.; Seeberger, P. H. Automated solid-phase synthesis of oligosaccharides. *Science* **2001**, *291* (5508), 1523–7.
- (13) Imamura, A.; Ando, H.; Korogi, S.; Tanabe, G.; Muraoka, O.; Ishida, H.; Kiso, M. Di-tert-butylsilylene (DTBS) group-directed α -selective galactosylation unaffected by C-2 participating functionalities. *Tetrahedron Lett.* **2003**, *44* (35), 6725–6728.
- (14) Imamura, A.; Kimura, A.; Ando, H.; Ishida, H.; Kiso, M. Extended applications of di-tert-butylsilylene-directed alpha-predominant galactosylation compatible with C2-participating groups toward the assembly of various glycosides. *Chem.—Eur. J.* **2006**, *12* (34), 8862–70.
- (15) Hansen, T.; Elferink, H.; van Hengst, J. M. A.; Houthuijs, K. J.; Remmerswaal, W. A.; Kromm, A.; Berden, G.; van der Vorm, S.; Rijs, A. M.; Overkleeft, H. S.; Filippov, D. V.; Rutjes, F.; van der Marel, G. A.; Martens, J.; Oomens, J.; Codee, J. D. C.; Boltje, T. J. Characterization of glycosyl dioxolenium ions and their role in glycosylation reactions. *Nat. Commun.* **2020**, *11* (1), 2664.
- (16) Bohe, L.; Crich, D. A propos of glycosyl cations and the mechanism of chemical glycosylation; the current state of the art. *Carbohydr. Res.* **2015**, *403*, 48–59.
- (17) Crich, D. Mechanism of a chemical glycosylation reaction. *Acc. Chem. Res.* **2010**, *43* (8), 1144–53.
- (18) Adero, P. O.; Amarasekara, H.; Wen, P.; Bohe, L.; Crich, D. The Experimental Evidence in Support of Glycosylation Mechanisms at the SN1-SN2 Interface. *Chem. Rev.* **2018**, *118* (17), 8242–8284.
- (19) Upadhyaya, K.; Subedi, Y. P.; Crich, D. Direct Experimental Characterization of a Bridged Bicyclic Glycosyl Dioxocarbenium Ion by (1) H and (13) C NMR Spectroscopy: Importance of Conformation on Participation by Distal Esters. *Angew. Chem. Int. Ed.* **2021**, *60* (48), 25397–25403.
- (20) Kalikanda, J.; Li, Z. Study of the stereoselectivity of 2-azido-2-deoxygalactosyl donors: remote protecting group effects and temperature dependency. *J. Org. Chem.* **2011**, *76* (13), 5207–18.
- (21) Demchenko, A. V.; Rousson, E.; Boons, G.-J. Stereoselective 1,2-cis-galactosylation assisted by remote neighboring group participation and solvent effects. *Tetrahedron Lett.* **1999**, *40* (36), 6523–6526.
- (22) Santana, A. G.; Montalvillo-Jimenez, L.; Diaz-Casado, L.; Corzana, F.; Merino, P.; Canada, F. J.; Jimenez-Oses, G.; Jimenez-Barbero, J.; Gomez, A. M.; Asensio, J. L. Dissecting the Essential Role of Anomeric beta-Triflates in Glycosylation Reactions. *J. Am. Chem. Soc.* **2020**, *142* (28), 12501–12514.
- (23) van der Vorm, S.; Hansen, T.; Overkleeft, H. S.; van der Marel, G. A.; Codee, J. D. C. The influence of acceptor nucleophilicity on the glycosylation reaction mechanism. *Chem. Sci.* **2017**, *8* (3), 1867–1875.
- (24) Franconetti, A.; Arda, A.; Asensio, J. L.; Bleriot, Y.; Thibaudeau, S.; Jimenez-Barbero, J. Glycosyl Oxocarbenium Ions: Structure, Conformation, Reactivity, and Interactions. *Acc. Chem. Res.* **2021**, *54* (11), 2552–2564.
- (25) Ma, Y.; Lian, G.; Li, Y.; Yu, B. Identification of 3,6-di-O-acetyl-1,2,4-O-orthoacetyl-alpha-D-glucopyranose as a direct evidence for the 4-O-acyl group participation in glycosylation. *Chem. Commun.* **2011**, *47* (26), 7515–7.
- (26) Yao, D.; Liu, Y.; Yan, S.; Li, Y.; Hu, C.; Ding, N. Evidence of robust participation by an equatorial 4-O group in glycosylation on a 2-azido-2-deoxy-glucopyranosyl donor. *Chem. Commun.* **2017**, *53* (20), 2986–2989.
- (27) de Kleijne, F. F. J.; Elferink, H.; Moons, S. J.; White, P. B.; Boltje, T. J. Characterization of Mannosyl Dioxanion Ions in Solution Using Chemical Exchange Saturation Transfer NMR Spectroscopy. *Angew. Chem. Int. Ed.* **2022**, *61* (6), No. e202109874.
- (28) Elferink, H.; Remmerswaal, W. A.; Houthuijs, K. J.; Jansen, O.; Hansen, T.; Rijs, A. M.; Berden, G.; Martens, J.; Oomens, J.; Codee, J.

- D. C.; Boltje, T. J. Competing C-4 and C-5-Acyl Stabilization of Uronic Acid Glycosyl Cations *Chem. Eur. J.* **2022**, e202201724.
- (29) Elferink, H.; Mensink, R. A.; Castelijns, W. W. A.; Jansen, O.; Bruekers, J. P. J.; Martens, J.; Oomens, J.; Rijs, A. M.; Boltje, T. J. The Glycosylation Mechanisms of 6,3-Uronic Acid Lactones. *Angew. Chem. Int. Ed.* **2019**, *58* (26), 8746–8751.
- (30) Lebedel, L.; Ardá, A.; Martin, A.; Désiré, J.; Mingot, A.; Aufiero, M.; Aiguabella Font, N.; Gilmour, R.; Jiménez-Barbero, J.; Blériot, Y.; Thibaudeau, S. Structural and Computational Analysis of 2-Halogeno-Glycosyl Cations in the Presence of a Superacid: An Expansive Platform. *Angew. Chem. Int. Ed.* **2019**, *58* (39), 13758–13762.
- (31) González Flórez, A. I.; Mucha, E.; Ahn, D. S.; Gewinner, S.; Schöllkopf, W.; Pagel, K.; von Helden, G. Charge-Induced Unzipping of Isolated Proteins to a Defined Secondary Structure. *Angew. Chem. Int. Ed.* **2016**, *55* (10), 3295–9.
- (32) Greis, K.; Kirschbaum, C.; von Helden, G.; Pagel, K. Gas-phase infrared spectroscopy of glycans and glycoconjugates. *Curr. Opin. Struct. Biol.* **2022**, *72*, 194–202.
- (33) Thomas, D. A.; Chang, R.; Mucha, E.; Lettow, M.; Greis, K.; Gewinner, S.; Schöllkopf, W.; Meijer, G.; von Helden, G. Probing the conformational landscape and thermochemistry of DNA dinucleotide anions via helium nanodroplet infrared action spectroscopy. *Phys. Chem. Chem. Phys.* **2020**, *22* (33), 18400–18413.
- (34) Greis, K.; Kirschbaum, C.; Lechnitz, S.; Gewinner, S.; Schöllkopf, W.; von Helden, G.; Meijer, G.; Seeberger, P. H.; Pagel, K. Direct Experimental Characterization of the Ferrier Glycosyl Cation in the Gas Phase. *Org. Lett.* **2020**, *22* (22), 8916–8919.
- (35) Even, U. The Even-Lavie valve as a source for high intensity supersonic beam. *EPJ. Technol. Instrum* **2015**, *2* (1), 17.
- (36) Schöllkopf, W.; Gewinner, S.; Junkes, H.; Paarmann, A.; von Helden, G.; Bluem, H. P.; Todd, A. M. M. The new IR and THz FEL facility at the Fritz Haber Institute in Berlin. *Proc. SPIE Int. Soc. Opt. Eng.* **2015**, 9512, 95121L.
- (37) Elferink, H.; Severijnen, M. E.; Martens, J.; Mensink, R. A.; Berden, G.; Oomens, J.; Rutjes, F.; Rijs, A. M.; Boltje, T. J. Direct Experimental Characterization of Glycosyl Cations by Infrared Ion Spectroscopy. *J. Am. Chem. Soc.* **2018**, *140* (19), 6034–6038.
- (38) Mucha, E.; Marianski, M.; Xu, F.-F.; Thomas, D. A.; Meijer, G.; von Helden, G.; Seeberger, P. H.; Pagel, K. Unravelling the structure of glycosyl cations via cold-ion infrared spectroscopy. *Nat. Commun.* **2018**, *9* (1), 4174.
- (39) Greis, K.; Mucha, E.; Lettow, M.; Thomas, D. A.; Kirschbaum, C.; Moon, S.; Pardo-Vargas, A.; von Helden, G.; Meijer, G.; Gilmore, K.; Seeberger, P. H.; Pagel, K. The Impact of Leaving Group Anomericity on the Structure of Glycosyl Cations of Protected Galactosides. *ChemPhysChem* **2020**, *21* (17), 1905–1907.
- (40) Greis, K.; Kirschbaum, C.; Fittolani, G.; Mucha, E.; Chang, R.; von Helden, G.; Meijer, G.; Delbianco, M.; Seeberger, P. H.; Pagel, K. Neighboring Group Participation of Benzoyl Protecting Groups in C3- and C6-Fluorinated Glucose. *Eur. J. Org. Chem.* **2022**, 2022 (15), No. e202200255.
- (41) Remmerswaal, W. A.; Houthuijs, K. J.; van de Ven, R.; Elferink, H.; Hansen, T.; Berden, G.; Overkleeft, H. S.; van der Marel, G. A.; Rutjes, F.; Filippov, D. V.; Boltje, T. J.; Martens, J.; Oomens, J.; Codee, J. D. C. Stabilization of Glycosyl Dioxolenium Ions by “Dual Participation” of the 2,2-Dimethyl-2-(ortho-nitrophenyl)acetyl (DMNPA) Protection Group for 1,2-cis-Glycosylation. *J. Org. Chem.* **2022**, *87* (14), 9139–9147.
- (42) Supady, A.; Blum, V.; Baldauf, C. First-Principles Molecular Structure Search with a Genetic Algorithm. *J. Chem. Inf. Model* **2015**, *55* (11), 2338–48.
- (43) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868.
- (44) Tkatchenko, A.; Scheffler, M. Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.* **2009**, *102* (7), 073005.
- (45) Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **2009**, *180* (11), 2175–2196.
- (46) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110* (13), 6158–6170.
- (47) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56* (5), 2257–2261.
- (48) Frisch, M. J.; et al. *Gaussian 16*, Revision A.03; Gaussian, Inc.: Wallingford, CT, 2016.
- (49) Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *J. Chem. Phys.* **2013**, *139* (13), 134101.
- (50) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297–305.
- (51) Neese, F. Software update: The ORCA program system—Version 5.0. *WIREs Comput. Mol. Sci.* **2022**, *12* (5), No. e1606.
- (52) Cremer, D.; Pople, J. A. General definition of ring puckering coordinates. *J. Am. Chem. Soc.* **1975**, *97* (6), 1354–1358.
- (53) van der Vorm, S.; Hansen, T.; van Hengst, J. M. A.; Overkleeft, H. S.; van der Marel, G. A.; Codee, J. D. C. Acceptor reactivity in glycosylation reactions. *Chem. Soc. Rev.* **2019**, *48* (17), 4688–4706.
- (54) Zeng, Y.; Wang, Z.; Whitfield, D.; Huang, X. Installation of electron-donating protective groups, a strategy for glycosylating unreactive thioglycosyl acceptors using the preactivation-based glycosylation method. *J. Org. Chem.* **2008**, *73* (20), 7952–62.
- (55) Andreana, P. R.; Crich, D. Guidelines for O-Glycoside Formation from First Principles. *ACS Cent. Sci.* **2021**, *7* (9), 1454–1462.
- (56) Mucha, E.; Stuckmann, A.; Marianski, M.; Struwe, W. B.; Meijer, G.; Pagel, K. In-depth structural analysis of glycans in the gas phase. *Chem. Sci.* **2019**, *10* (5), 1272–1284.
- (57) Grabarics, M.; Lettow, M.; Kirschbaum, C.; Greis, K.; Manz, C.; Pagel, K. Mass Spectrometry-Based Techniques to Elucidate the Sugar Code. *Chem. Rev.* **2022**, *122* (8), 7840–7908.
- (58) Kirschbaum, C.; Greis, K.; Polewski, L.; Gewinner, S.; Schöllkopf, W.; Meijer, G.; von Helden, G.; Pagel, K. Unveiling Glycerolipid Fragmentation by Cryogenic Infrared Spectroscopy. *J. Am. Chem. Soc.* **2021**, *143* (36), 14827–14834.
- (59) Xu, K.; Man, Q.; Zhang, Y.; Guo, J.; Liu, Y.; Fu, Z.; Zhu, Y.; Li, Y.; Zheng, M.; Ding, N. Investigation of the remote acyl group participation in glycosylation from conformational perspectives by using trichloroacetimidate as the acetyl surrogate. *Org. Chem. Front* **2020**, *7* (13), 1606–1615.
- (60) Hettikankanamalage, A. A.; Lassfolk, R.; Ekholm, F. S.; Leino, R.; Crich, D. Mechanisms of Stereodirecting Participation and Ester Migration from Near and Far in Glycosylation and Related Reactions. *Chem. Rev.* **2020**, *120* (15), 7104–7151.
- (61) Greis, K.; Kirschbaum, C.; Taccone, M. I.; Götze, M.; Gewinner, S.; Schöllkopf, W.; Meijer, G.; von Helden, G.; Pagel, K. Studying the Key Intermediate of RNA Autohydrolysis by Cryogenic Gas-Phase Infrared Spectroscopy. *Angew. Chem. Int. Ed.* **2022**, *61* (19), No. e202115481.
- (62) Kirschbaum, C.; Greis, K.; Gewinner, S.; Schöllkopf, W.; Meijer, G.; von Helden, G.; Pagel, K. Cryogenic infrared spectroscopy provides mechanistic insight into the fragmentation of phospholipid silver adducts. *Anal. Bioanal. Chem.* **2022**, *414* (18), 5275–5285.
- (63) Chatterjee, S.; Moon, S.; Hentschel, F.; Gilmore, K.; Seeberger, P. H. An Empirical Understanding of the Glycosylation Reaction. *J. Am. Chem. Soc.* **2018**, *140* (38), 11942–11953.

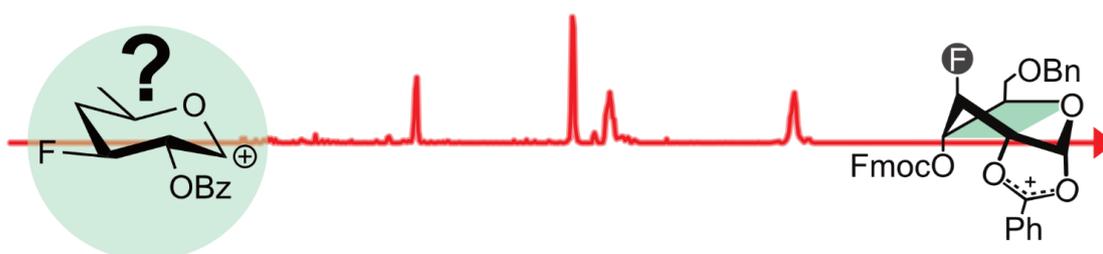
5 | Neighboring-Group Participation of Benzoyl Protecting Groups in Fluorinated Glucose

This work is published^[297] and available online:

DOI: <https://doi.org/10.1002/ejoc.202200255>

Details on the related Supporting Information can be found in Appendix B.

K. Greis, C. Kirschbaum, G. Fittolani, E. Mucha, R. Chang, G. Meijer, G. von Helden, M. Delbianco, P.H. Seeberger, K. Pagel*; Neighboring Group Participation of Benzoyl Protecting Groups in C3- and C6-Fluorinated Glucose. *Eur. J. Org. Chem.* **2022**, e202200255.



Submitted: 2 March 2022; **Published:** 24 March 2022

Copyright © 2022 The Authors. Published by Wiley-VCH GmbH

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

5.1 | Author Contributions

The project was conceived by **Kim Greis** and Kevin Pagel. **Kim Greis** wrote the first draft of the manuscript. **Kim Greis** (main), Carla Kirschbaum, Eike Mucha, and Rayoon Chang (all three support) performed the mass spectrometry and cryogenic infrared spectroscopy experiments. The cryogenic infrared spectroscopy experiments were supervised by Gerard Meijer and Gert von Helden. **Kim Greis** performed all quantum chemical calculations. The glucose building blocks were provided by Giulio Fittolani, Martina Delbianco, and Peter Seeberger. **Kim Greis** analyzed the data.

5.2 | Project Hypothesis and Summary

Previous studies on glycosyl cations mainly focused on model compounds to study the effect of one participating acetyl group.^[85–88] In oligosaccharide synthesis, the monosaccharides are usually differentially protected, allowing for selective deprotection during sequential coupling of monosaccharide units. In this project, the effect of differential protecting groups and the replacement of a protecting group by fluorine on the structure of glucosyl cations was studied by combining cryogenic IR spectroscopy and DFT calculations. Fluorinated sugars play a role for pharmaceuticals^[20,21] and in material science,^[31,298] as fluorination significantly alters the lipophilicity^[299–301] and stability^[19,22,302] of glycans. In total, five glucose building blocks carrying Bz, Fmoc, and Bn were investigated (see Scheme 2 of the related published work). In this work, fluorine is either introduced at the C3- or the C6-position.

The thioether leaving groups of the studied building blocks can be readily cleaved by in-source fragmentation after electrospray ionization, leading to the formation of glucosyl cations. In all cases, the glucosyl cations adopt dioxolenium structures by neighboring-group participation of the C2-benzoyl group. DFT-computed harmonic frequencies show that only this structural motif has matching absorption bands to those that are experimentally resolved (see Figures 1 and 2 of the related publication). The harmonic frequencies of other structural motifs, such as C4- or C6-remote participation of the Fmoc group, C6-remote participation of the benzyl group, or no participation do not match the experiment (see Scheme 1 of the related published work). Further, DFT-calculated energetics confirm that neighboring-group participation of the benzoyl group is favored by 52–95 kJ mol⁻¹. The observed vibrational pattern in the 1400–1600 cm⁻¹ region suggests that this type of participation is further enhanced by a positive mesomeric effect of the phenyl moiety in the benzoyl group. The strong signal at ca. 1520 cm⁻¹ can be associated to the C–C stretching mode connecting the phenyl and the COO⁺ moieties of the benzoyl group, indicating a high double-bond character of this C–C bond (see Scheme 3 of the related published work).

Fluorination does not change the overall structural motif that can be observed in glucosyl cations, but it impacts the conformation. The glucosyl cation **Glc1**, which carries an Fmoc group at the C4-position and a Bn group at the C6-position, adopts an ${}^O,{}^3B$ pucker, whereas its C3- and C6-fluorinated counterparts rather adopt 3S_1 puckers. In systems with a decreased steric demand, glucosyl cations prefer to adopt 3S_1 ring puckers.^[85] Although one would only expect to observe one carbonyl band in the spectra of these ions, a second one can be observed in the C6-fluorinated system. This band is indicative for the presence of an isoenergetic conformer of the same structural motif adopting an ${}^O,{}^3B$ pucker in the ion trap.

For the glucosyl cation **Glc2**, with an Fmoc group at the C6-position and a Bn group at the C4-position, a 3S_1 pucker is most likely. Here, the steric demand is generally smaller, as the bulky Fmoc group is located at the more flexible C6-arm. Due to the increased flexibility for this system, two conformers populate the ion trap, as two carbonyl bands can be observed in its spectrum. Interestingly, the lowest-energy structure of its C3-fluorinated counterpart does not exhibit a 3S_1 pucker but a 5H_4 pucker. However, the bands in the diagnostic functional group region ($1400\text{--}1800\text{ cm}^{-1}$) match better for a second structure, slightly destabilized by 3 kJ mol^{-1} , which adopts an OS_2 pucker.

This project has shown that it is possible to assign the structure of comparably large, differentially protected glycosyl cations and study the effect of fluorination and permutation of protecting groups on the structure using cryogenic IR spectroscopy and DFT calculations. In previous publications,^[30,303,304] it has been determined that the precursors for the investigated glycosyl cations are β -selective in glycosylation reactions. The origin of this selectivity is the structure of the reactive intermediate, which exhibits neighboring-group participation of the C2-benzoyl group, as determined in this work. Here, the α -side is efficiently shielded from nucleophilic attack.

Neighboring Group Participation of Benzoyl Protecting Groups in C3- and C6-Fluorinated Glucose

Kim Greis,^[a, b] Carla Kirschbaum,^[a, b] Giulio Fittolani,^[a, c] Eike Mucha,^[b] Rayoon Chang,^[a, b] Gert von Helden,^[b] Gerard Meijer,^[b] Martina Delbianco,^[c] Peter H. Seeberger,^[a, c] and Kevin Pagel^{*[a, b]}

Fluorination is a potent method to modulate chemical properties of glycans. Here, we study how C3- and C6-fluorination of glucosyl building blocks influence the structure of the intermediate of the glycosylation reaction, the glycosyl cation. Using a combination of gas-phase infrared spectroscopy and first-principles theory, glycosyl cations generated from fluorinated and non-fluorinated monosaccharides are structurally characterized. The results indicate that neighboring group participation of the C2-benzoyl protecting group is the dominant structural

motif for all building blocks, correlating with the β -selectivity observed in glycosylation reactions. The infrared signatures indicate that participation of the benzoyl group is enhanced by resonance effects. Participation of remote acyl groups such as Fmoc or benzyl on the other hand is unfavored. The introduction of the less bulky fluorine leads to a change in the conformation of the ring pucker, whereas the structure of the active dioxolenium site remains unchanged.

Introduction

Beyond the various roles of glycans in biological processes,^[1] they exhibit a great pharmaceutical potential. Fractionated heparin is used as anti-coagulating agent since the 1940s. Glycans used in biomedical applications are often extracted from natural sources. This approach not only limits the number of available compounds to those occurring in nature, but also requires elaborate separation workflows to produce pure and well-defined molecules.^[2] Furthermore, the short lifetimes of glycan-based pharmaceuticals and their absorption properties, such as low lipophilicity, in the human body are impeding their usage.^[3] An efficient method to modulate glycan properties is the incorporation of fluorine. Fluorinated glycans are more stable,^[4] exhibit an increased lipophilicity^[5] and are more potent against certain pathogens than their non-fluorinated counterparts.^[6] Moreover, site-selective introduction of fluorine

impacts material properties of carbohydrates as demonstrated for cellulose.^[7]

Well-defined fluorinated glycans can be synthesized by automated glycan assembly (AGA)^[8] using fluorinated monosaccharide building blocks. AGA allows to control sequence, branching, and length, up to 100-mers.^[9] A major challenge in the glycosylation reactions is the stereoselective formation of α - and β -glycosidic linkages. However, the underlying reaction mechanism is still not fully understood today, thus rendering the prediction of the stereochemical outcome of a reaction difficult. Generally, it is believed that the reaction is governed by a mechanistic continuum between S_N1 and S_N2 , dependent on various parameters such as the nature of acceptor and donor, temperature, solvent, counter ions, or leaving groups.^[10] Recently, a correlation between the stereoselectivity of the S_N1 side of the continuum and the structure of the positively charged intermediate that is formed during the reaction, the glycosyl cation, has been determined.^[11] To selectively generate 1,2-*trans* linkages, participating acyl protecting groups such as benzoyl or acetyl at the C2 position are commonly used.^[12] For glucose, it has been postulated that these neighboring protecting groups (PGs) shield the α -side in glycosyl cations, forcing nucleophiles to attack from the β -side.

Due to their short lifetimes, it is generally difficult to directly characterize glycosyl cations experimentally. They can be stabilized by super acids and subsequently be probed via NMR spectroscopy. However, the super acids fully protonate the glycosyl cation, leading to a distortion of its structure and properties.^[13] Recently, it was shown that bare glycosyl cations can be isolated in the "clean-room" environment of a mass spectrometer and subsequently characterized by gas-phase infrared spectroscopy. First experiments demonstrated that acetyl groups in model building blocks show neighboring group participation (I, Scheme 1)^[11a,b,14] and remote participation (II),^[15] in which the carbonyl oxygen forms a covalent bond with

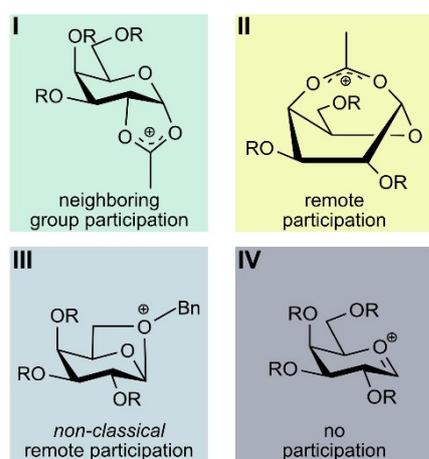
[a] K. Greis, C. Kirschbaum, G. Fittolani, R. Chang, Prof. Dr. P. H. Seeberger, Prof. Dr. K. Pagel
Institute of Chemistry and Biochemistry
Freie Universität Berlin
Arnimallee 22, 14195 Berlin, Germany
E-mail: kevin.pagel@fu-berlin.de
<https://www.bcp.fu-berlin.de/chemie/pagel>

[b] K. Greis, C. Kirschbaum, Dr. E. Mucha, R. Chang, Prof. Dr. G. von Helden, Prof. Dr. G. Meijer, Prof. Dr. K. Pagel
Fritz Haber Institute of the Max Planck Society
Faradayweg 4–6, 14195 Berlin, Germany

[c] G. Fittolani, Dr. M. Delbianco, Prof. Dr. P. H. Seeberger
Max Planck Institute of Colloids and Interfaces
Am Mühlenberg 1, 14476 Potsdam, Germany

 Supporting information for this article is available on the WWW under <https://doi.org/10.1002/ejoc.202200255>

 © 2022 The Authors. European Journal of Organic Chemistry published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

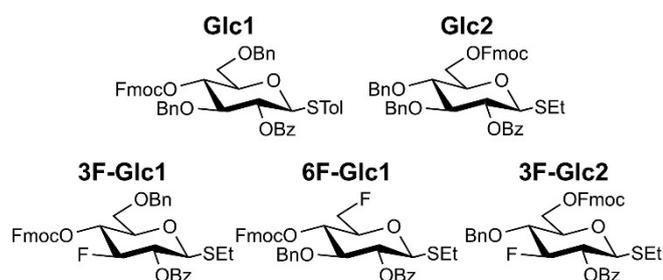


Scheme 1. Modes of participation in glycosyl cations.

the anomeric carbon to yield a bicyclic dioxolenium intermediate. The studies also revealed that the gas-phase structures of the investigated glycosyl cations correlate with the experimental stereoselectivity observed in solution-phase studies of their precursors.

Interestingly, despite being formally known as non-participating PGs, benzyl ether oxygens can also stabilize the positive charge at the anomeric carbon, resulting in the formation of oxonium ions (III).^[15b]

Here, we combine cryogenic infrared spectroscopy with density functional theory (DFT) to probe glycosyl cations of functionalized glucose building blocks that are commonly used in glycan synthesis. The C2 position is always benzoylated (Bz), while the other hydroxy groups are either protected with fluorenylmethoxycarbonyl (Fmoc) or benzyl (Bn) groups. In selected building blocks fluorine is introduced at the C3 or C6 position to study its impact on the structure of the glycosyl cation (Scheme 2). Further, the gas-phase structures are correlated to the experimentally observed β -stereoselectivity.

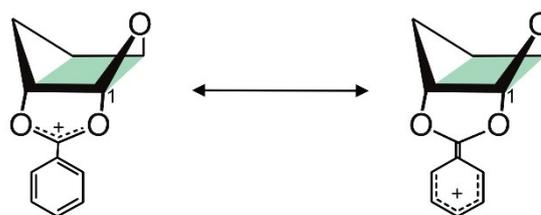


Scheme 2. Differentially protected monosaccharide building blocks used in this study to generate glycosyl cations, which are subsequently probed by cryogenic infrared spectroscopy.

Results and Discussion

First, the IR signature of the non-fluorinated glycosyl cation **Glc1** is shown (Figure 1a). The functional group region (1450–1800 cm^{-1}) shows five resolved absorption bands that clearly match the computed spectrum of the lowest-energy structure I (**Glc1**), with an ${}^{\text{O},3}\text{B}$ ring pucker, exhibiting neighboring group participation (NGP) of the C2-benzoyl group with a covalent bond (1.51 Å) between the carbonyl oxygen and the anomeric carbon. The signals at 1466 and 1500 cm^{-1} originate from the symmetric and antisymmetric dioxolenium stretches $\nu(\text{O}-\text{C}-\text{O})$ of the participating Bz PG, while the signal at 1759 cm^{-1} stems from a carbonyl stretch $\nu(\text{C}=\text{O})$ within the non-participating Fmoc PG. Interestingly, the vibrations at 1519 and 1600 cm^{-1} are due to $\nu(\text{C}=\text{C})$ stretches connected to resonance stabilization of the positive charge by the phenyl ring of the Bz PG in the dioxolenium motif (Scheme 3). The strong absorption at 1519 cm^{-1} is caused by the vibration of the formed $\text{C}=\text{C}$ double bond, while the weak absorption at 1600 cm^{-1} can be attributed to the $\nu(\text{C}=\text{C})$ stretches within the phenyl ring. The increased partial double bond character is also visible in the length of the $\text{C}-\text{C}$ bond that decreases from 1.47 to 1.43 Å compared to the lowest-energy oxocarbenium structure where the PGs do not participate. Thus, the charge of the glycosyl cation is not only delocalized within the dioxolenium motif, but also within the phenyl ring, leading to further stabilization. A similar stabilization by resonance effects in cations was previously reported for 4-aminobenzoic acid in gas-phase IR experiments.^[16]

The fingerprint region (1000–1450 cm^{-1}) contains a unique signature for each species, however, it is rather difficult to derive a structural assignment solely based on this region. Computational methods often fail to accurately model the fingerprint region in more complex systems, also due to anharmonicities.^[17] The vibrations observed herein are mainly originating from $\text{C}-\text{C}$ and $\text{C}-\text{O}$ stretching vibrations (1000–1350 cm^{-1}) as well as $\text{C}-\text{H}$ bends (1350–1450 cm^{-1}). The spectral signature corresponds the best to the lowest-energy structure I (**Glc1**). Other structural motifs, such as remote participation of the Fmoc PG II (**Glc1**) (+61 kJ mol^{-1}), remote benzyl ether participation III (**Glc1**) (+57 kJ mol^{-1}) or oxocarbenium structures IV (**Glc1**) (+80 kJ mol^{-1}), can be clearly ruled out due to two reasons: 1) their free energies at 90 K are significantly higher than those of structures exhibiting NGP; 2)



Scheme 3. Resonance stabilization of the positive charge by the phenyl ring in benzoyl neighboring group participation. Glycosyl cations with this mode of participation are further stabilized by increased delocalization of the positive charge.

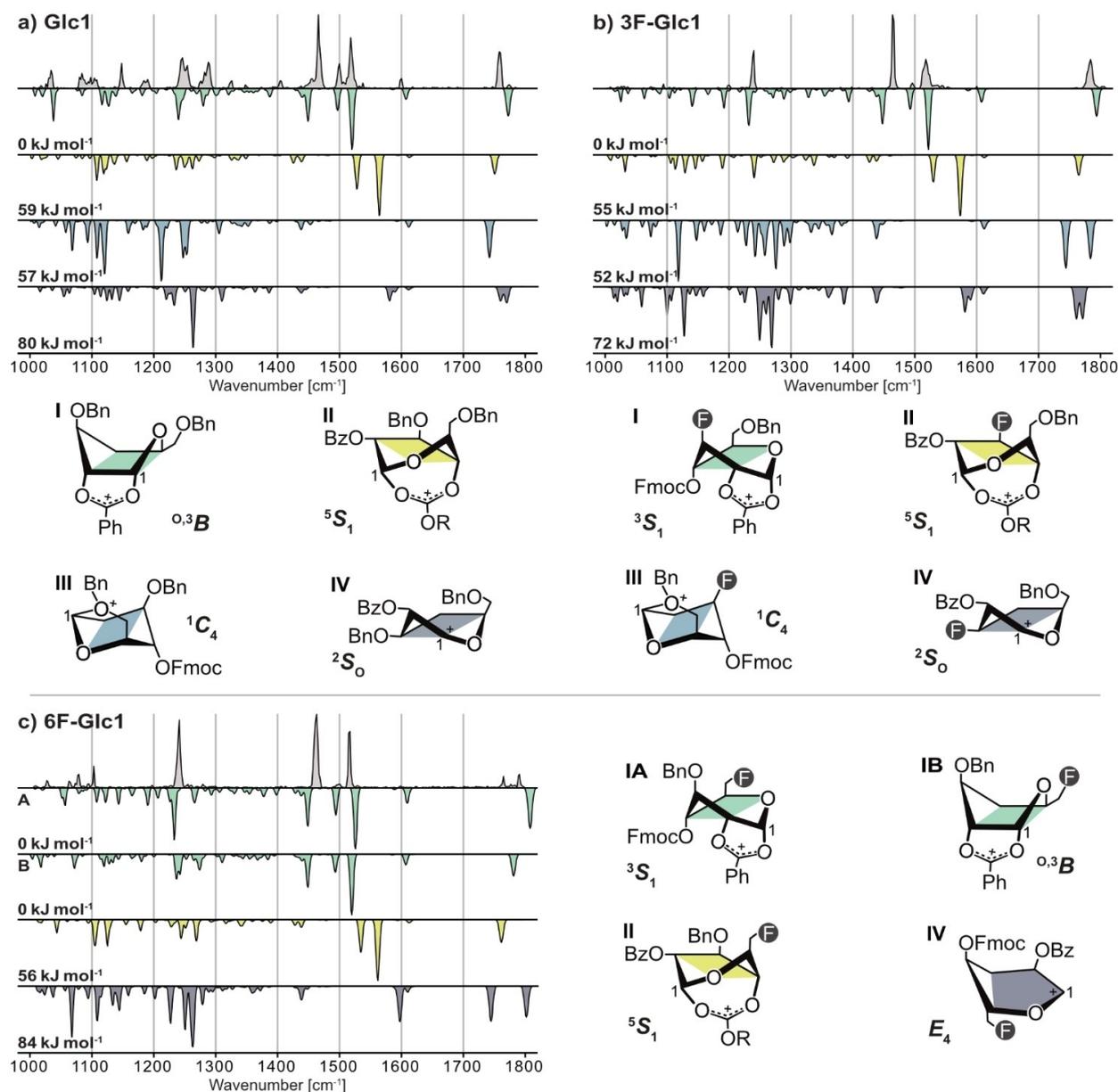


Figure 1. Infrared spectra of (a) **Glc1**, (b) **3F-Glc1**, and (c) **6F-Glc1** glycosyl cations generated from β -thiotolyl (a) and β -thioethyl (b,c) precursors. Experimental IR spectra are shown as light gray traces. Computed spectra of lowest-energy dioxolenium structures, exhibiting neighboring group (green) and remote participation (yellow), oxonium (blue), and oxocarbenium structures (dark gray) are shown as inverted traces in respective colors. Relative free energies at 90 K are indicated. The lowest-energy structures are shown in a simplified representation below the spectra, with their ring pucker annotated. For clarity, some protecting groups have been omitted and R used as abbreviation for fluorenylmethyl. 3D-representation of the structures and xyz-coordinates can be found in the SI.

their computed infrared spectra do not agree with the experimental spectrum (Figure 1a).

The IR spectra of the C3- and C6-fluorinated glycosyl cations **3F-Glc1** and **6F-Glc1** are shown in Figure 1b and Figure 1c. Compared to **Glc1**, the spectral signature of the fluorinated counterparts is less crowded in the fingerprint region. Here, mainly one intense absorption band can be observed at 1234 cm^{-1} associated with a $\nu(\text{C}=\text{O})$ stretch within the Fmoc PG. Otherwise, the spectral signature resembles that of **Glc1**. As a consequence, the glycosyl cations **3F-Glc1** and **6F-Glc1** mainly

adopt dioxolenium-type structures I exhibiting benzoyl NGP. Although all three experimental spectra share some similarities, the absorption bands differ in shape and exact position. Thus, each spectrum is a unique pattern for the probed glycosyl cation. Further evidence for a C2-dioxolenium motif is provided by the computed spectra of structures exhibiting benzoyl NGP that also possess the lowest free energy of all sampled structures. In both cases, a 3S_1 pucker is adopted with a bond distance of 1.50 \AA between the carbonyl oxygen of the Bz PG and the anomeric carbon. The vibrations associated with the

dioxolenium motif and the +M effect within the benzoyl group clearly correspond to the experimental signature. The carbonyl absorption band in **I** (**3F-Glc1**) corresponds to the experiment, while the experimental spectrum of **6F-Glc1** exhibits two carbonyl bands, which is diagnostic for a second low-energy conformer (**IB**) simultaneously present in the ion trap. Like for **Glc1**, other structural motifs can be excluded based on their computed spectral signatures and unfavorable free energies.

Although the substitution of a benzyl group by fluorine changes the ring pucker from 0_3B in **Glc1** to 3S_1 in **3/6F-Glc1**, it does not have an influence on the participation of the neighboring benzoyl group. The changes in ring pucker could be attributed to a decreased steric hindrance of fluorine compared to the bulkier benzyl PG. In all three cases, the α -side of the glycosyl cation is efficiently shielded, leading to β -stereoselectivity. This selectivity was observed in the AGA of deoxyfluorinated $\beta(1,4)$ hexaglycoside analogues (employing building blocks **Glc1**, **3F-Glc1**, and **6F-Glc1**, see the Supporting Information).^[7b,18]

In a second set of glycosyl cations, **Glc2** and **3F-Glc2**, the C4 and C6 PGs are permuted, compared to **Glc1** analogues. The IR spectra are shown in Figure 2. Generally, the spectral signature is slightly more congested than the corresponding **Glc1** species, which is attributed to the population of multiple low-energy conformers enabled by the increased flexibility of the Fmoc PG now located at the C6 position. In the functional

group region, the spectra look similar to those previously shown, being diagnostic for C2-dioxolenium structures exhibiting NGP. For **Glc2**, the lowest-energy structure **IA** exhibits benzoyl NGP, with a 3S_1 pucker and a bond distance of 1.51 Å between the carbonyl oxygen of the benzoyl group and the anomeric carbon. A second low-energy conformer **IB** (+5 kJ mol⁻¹) was sampled, in which the Fmoc, the C4-Bn and the participating Bz PG are stacked. The differently orientated Fmoc PG leads to a shift of the position of the carbonyl band. The population of these two low-energy conformers might explain the presence of two carbonyl bands and the wealth of absorption bands in the fingerprint region in the experimental spectrum. For **3F-Glc2**, the lowest-energy conformer **IA** exhibits a 5H_4 pucker, however, its IR signature matches the experiment slightly less well than that of a second low-energy structure **IB** (+3 kJ mol⁻¹) with a 0S_2 pucker and a 1.50 Å bond distance. Again, other structural motifs are unlikely, considering their spectral signature and energetics. Here, fluorine has an influence on the ring pucker, but not on the overall structural motif, strongly correlated to the experimental β -stereoselectivity. Formation of β -linkages was observed in the AGA of deoxyfluorinated glucosides (employing building blocks **Glc2** and **3F-Glc2**).^[19]

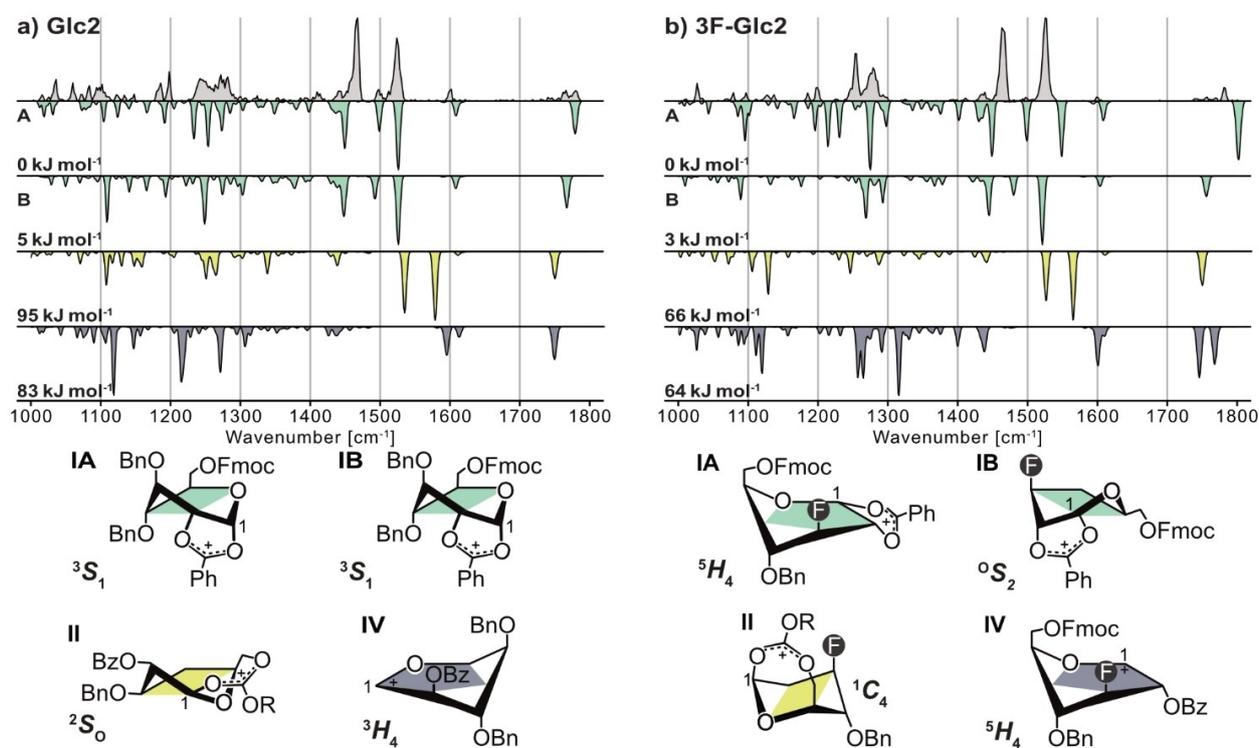


Figure 2. Infrared spectra of (a) **Glc2** and (b) **3F-Glc2** glycosyl cations generated from β -thioethyl precursors. Experimental IR spectra are shown as light gray traces. Computed spectra of lowest-energy dioxolenium structures, exhibiting neighboring group (green) and remote participation (yellow), and oxocarbenium structures (dark gray) are shown as inverted traces in respective colors. Relative free energies at 90 K are indicated. The lowest-energy structures are shown in a simplified representation below the spectra, with their ring pucker annotated (for **Glc2**, **IA** and **IB** the differences in structures are too subtle to represent them in the simplified representation, therefore, the reader is referred to the 3D-structure in Figure S12). For clarity, some protecting groups have been omitted and R used as abbreviation for fluorenylmethyl. 3D-representation of the structures and xyz-coordinates can be found in the SI.

Conclusion

To conclude, we have shown that it is possible to generate and probe glycosyl cations and their fluorinated analogues from precursors readily used in glycan synthesis. In each case, the underlying structural motif can be clearly identified as neighboring group participation of C2-benzoyl protecting groups. Interestingly, participation of the Bz protecting groups is connected to resonance effects involving the phenyl ring, which can be directly monitored due to vibrations associated with the delocalized electrons. The permutation of the protecting groups as well as their substitution by the less bulky fluorine leads to a change in the conformation of the ring pucker. However, the structure of the active dioxolenium site remains unchanged and the stereoselectivity observed for these building blocks in glycosylation reactions is therefore not affected. Further experiments are needed to explore the effects of a C2- and C4-fluorination, which are expected to have a much more significant impact on the structure of the reactive glycosylation intermediate.

Experimental Section

Cryogenic infrared spectroscopy

A detailed description of the experimental setup can be found in the SI (Figure S1) and in previous publications.^[20] Briefly, thioglycoside precursors were transferred into the gas phase *via* nano-electrospray ionization (nESI). The leaving group is cleaved by in-source fragmentation leading to glycosyl cations. Mass spectra can be found in the SI (Figures S2–S6). The ions of interest are mass-to-charge selected by a quadrupole mass filter and accumulated in a hexapole ion trap, which is cooled to approximately 90 K by liquid nitrogen. Superfluid helium nanodroplets (0.4 K) are generated by an Even-Lavie valve and traverse the ion trap, picking up ions, and guide them to a detection region, where the embedded ions are excited by IR photons generated by the free-electron laser of the Fritz Haber Institute (FHI FEL^[21]). Upon absorption of resonant photons, ions are eventually released from the droplets and afterwards detected by a time-of-flight detector. Monitoring the ion signal as a function of the IR photon wavenumber leads to a high-resolution IR signature of the probed ion.

Computational methods

To model the IR spectra of the probed ions, candidate structures were sampled using the genetic algorithm (GA) FAFOOM.^[22] The GA allows sampling flexible bonds and ring puckers and sends each sampled geometry to an external software (ORCA 4.1.1)^[23] for DFT optimization at the PBE/def2-SVP^[24] level of theory. This conformational search mainly yielded dioxolenium-type structures **I**, in which the benzoyl group shields the anomeric carbon from the α -side, and oxocarbenium-type structures **IV**, in which no participation takes place. Furthermore, the algorithm also generated structures in which either the remote Fmoc (C4 and C6) or Bn PGs (C6 only) interact with the anomeric carbon (dioxolenium **II** and oxonium structures **III**). A subset of structures of each type was reoptimized and their harmonic frequencies computed at the PBE0+D3/6-311+G(d,p)^[25] level of theory using Gaussian 16.^[26] Each computed IR spectrum was normalized and scaled by 0.965. Ring puckers were assigned according to Cremer-Pople coordinates.^[27] The employed

DFT functionals were chosen because they showed chemical accuracy in a benchmark study on carbohydrates.^[28] Details of the reoptimized structures, such as energetics, ring puckers, and coordinates can be found in the Supporting Information.

Acknowledgements

The authors gratefully acknowledge the expertise of Dr. Wieland Schöllkopf and Sandy Gewinner for running the FHI FEL. K.G. thanks the Fonds National de la Recherche (FNR), Luxembourg, for funding the project GlycoCat (13549747). C.K. is grateful for financial support by Fonds der Chemischen Industrie. R.C. and K.P. thank the Deutsche Forschungsgemeinschaft (DFG) for support under project number 387284271-SFB 1349. K.P. acknowledges generous funding by the European Research Council, ERC-2019-CoG-863934-GlycoSpec. M.D., G.F., and P.H.S. thank the MPG-FhG Cooperation Project Glyco3Dysplay and the German Federal Ministry of Education and Research (BMBF, grant number 13XP5114) for financial support. P.H.S. thanks the Max Planck Society for generous financial support. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords: Carbohydrates · Fluorine · Glycosylation · IR Spectroscopy · Mass spectrometry

- [1] A. Varki, *Glycobiology* **2017**, *27*, 3.
- [2] P. H. Seeberger, R. D. Cummings, in: *Essentials of Glycobiology*, 3rd ed. (Eds.: rd, A. Varki, R. D. Cummings, J. D. Esko, P. Stanley, G. W. Hart, M. Aebi, A. G. Darvill, T. Kinoshita, N. H. Packer, J. H. Prestegard, R. L. Schnaar, P. H. Seeberger), Cold Spring Harbor (NY), **2015**, pp. 729.
- [3] a) R. Hevey, *Chem. Eur. J.* **2021**, *27*, 2240; b) B. Linclau, A. Ardá, N. C. Reichardt, M. Sollogoub, L. Unione, S. P. Vincent, J. Jiménez-Barbero, *Chem. Soc. Rev.* **2020**, *49*, 3863.
- [4] a) A. Geissner, L. Baumann, T. J. Morley, A. K. O. Wong, L. Sim, J. R. Rich, P. P. L. So, E. M. Dullaghan, E. Lessard, U. Iqbal, M. Moreno, W. W. Wakarchuk, S. G. Withers, *ACS Cent. Sci.* **2021**, *7*, 345; b) A. Axer, R. P. Jumde, S. Adam, A. Faust, M. Schäfers, M. Fobker, J. Koehnke, A. K. H. Hirsch, R. Gilmour, *Chem. Sci.* **2021**, *12*, 1286; c) H. J. Lo, L. Krasnova, S. Dey, T. Cheng, H. Liu, T. I. Tsai, K. B. Wu, C. Y. Wu, C. H. Wong, *J. Am. Chem. Soc.* **2019**, *141*, 6484.
- [5] a) J. St-Gelais, E. Côté, D. Lainé, P. A. Johnson, D. Giguère, *Chem. Eur. J.* **2020**, *26*, 13499; b) J. St-Gelais, M. Bouchard, V. Denavit, D. Giguère, *J. Org. Chem.* **2019**, *84*, 8509; c) D. Lainé, O. Lessard, J. St-Gelais, D. Giguère, *Chem. Eur. J.* **2021**, *27*, 3799.
- [6] J. Vaugenot, A. El Harras, O. Tasseau, R. Marchal, L. Legentil, B. Le Guennic, T. Benvegno, V. Ferrières, *Org. Biomol. Chem.* **2020**, *18*, 1462.
- [7] a) M. Delbianco, P. H. Seeberger, *Mater. Horiz.* **2020**, *7*, 963; b) Y. Yu, T. Tyrikos-Ergas, Y. Zhu, G. Fittolani, V. Bordoni, A. Singhal, R. J. Fair, A.

- Grafmuller, P. H. Seeberger, M. Delbianco, *Angew. Chem. Int. Ed.* **2019**, *58*, 13127.
- [8] O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Science* **2001**, *291*, 1523.
- [9] A. A. Joseph, A. Pardo-Vargas, P. H. Seeberger, *J. Am. Chem. Soc.* **2020**, *142*, 8561.
- [10] a) P. O. Adero, H. Amarasekara, P. Wen, L. Bohe, D. Crich, *Chem. Rev.* **2018**, *118*, 8242; b) S. Chatterjee, S. Moon, F. Hentschel, K. Gilmore, P. H. Seeberger, *J. Am. Chem. Soc.* **2018**, *140*, 11942.
- [11] a) E. Mucha, M. Marianski, F.-F. Xu, D. A. Thomas, G. Meijer, G. von Helden, P. H. Seeberger, K. Pagel, *Nat. Commun.* **2018**, *9*, 4174; b) H. Elferink, M. E. Severijnen, J. Martens, R. A. Mensink, G. Berden, J. Oomens, F. Rutjes, A. M. Rijs, T. J. Boltje, *J. Am. Chem. Soc.* **2018**, *140*, 6034; c) A. A. Hettikankanamalage, R. Lassfolk, F. S. Ekholm, R. Leino, D. Crich, *Chem. Rev.* **2020**, *120*, 7104.
- [12] H. S. Hahm, M. Hurevich, P. H. Seeberger, *Nat. Commun.* **2016**, *7*, 12482.
- [13] L. Lebedel, A. Ardá, A. Martin, J. Désiré, A. Mingot, M. Aufiero, N. Aiguabella Font, R. Gilmore, J. Jiménez-Barbero, Y. Blériot, S. Thibaud-deau, *Angew. Chem. Int. Ed.* **2019**, *58*, 13758.
- [14] K. Greis, C. Kirschbaum, S. Lechnitz, S. Gewinner, W. Schöllkopf, G. von Helden, G. Meijer, P. H. Seeberger, K. Pagel, *Org. Lett.* **2020**, *22*, 8916.
- [15] a) H. Elferink, R. A. Mensink, W. W. A. Castelijns, O. Jansen, J. P. J. Bruekers, J. Martens, J. Oomens, A. M. Rijs, T. J. Boltje, *Angew. Chem. Int. Ed.* **2019**, *58*, 8746; b) M. Marianski, E. Mucha, K. Greis, S. Moon, A. Pardo, C. Kirschbaum, D. A. Thomas, G. Meijer, G. von Helden, K. Gilmore, P. H. Seeberger, K. Pagel, *Angew. Chem. Int. Ed.* **2020**, *59*, 6166; c) T. Hansen, H. Elferink, J. M. A. van Hengst, K. J. Houthuijs, W. A. Remmerswaal, A. Kromm, G. Berden, S. van der Vorm, A. M. Rijs, H. S. Overkleeft, D. V. Filippov, F. Rutjes, G. A. van der Marel, J. Martens, J. Oomens, J. D. C. Codee, T. J. Boltje, *Nat. Commun.* **2020**, *11*, 2664; d) K. Greis, E. Mucha, M. Lettow, D. A. Thomas, C. Kirschbaum, S. Moon, A. Pardo-Vargas, G. von Helden, G. Meijer, K. Gilmore, P. H. Seeberger, K. Pagel, *ChemPhysChem* **2020**, *21*, 1905.
- [16] a) J. Seo, S. Warnke, S. Gewinner, W. Schöllkopf, M. T. Bowers, K. Pagel, G. von Helden, *Phys. Chem. Chem. Phys.* **2016**, *18*, 25474; b) T. Khuu, N. Yang, M. A. Johnson, *Int. J. Mass Spectrom.* **2020**, 457.
- [17] a) B. Brauer, M. Pincu, V. Buch, I. Bar, J. P. Simons, R. B. Gerber, *J. Phys. Chem. A* **2011**, *115*, 5859; b) E. Mucha, A. Stuckmann, M. Marianski, W. B. Struwe, G. Meijer, K. Pagel, *Chem. Sci.* **2019**, *10*, 1272; c) M. Grabarics, M. Lettow, C. Kirschbaum, K. Greis, C. Manz, K. Pagel, *Chem. Rev.* **2021**.
- [18] G. Fittolani, E. Shanina, M. Guberman, P. H. Seeberger, C. Rademacher, M. Delbianco, *Angew. Chem. Int. Ed.* **2021**, *60*, 13302.
- [19] a) M. Delbianco, A. Kononov, A. Poveda, Y. Yu, T. Diercks, J. Jiménez-Barbero, P. H. Seeberger, *J. Am. Chem. Soc.* **2018**, *140*, 5421; b) S. Gim, G. Fittolani, Y. Yu, Y. Zhu, P. H. Seeberger, Y. Ogawa, M. Delbianco, *Chem. Eur. J.* **2021**, *27*, 13139; c) A. Poveda, G. Fittolani, P. H. Seeberger, M. Delbianco, J. Jiménez-Barbero, *Front. Mol. Biosci.* **2021**, *8*, 784318.
- [20] a) D. A. Thomas, E. Mucha, M. Lettow, G. Meijer, M. Rossi, G. von Helden, *J. Am. Chem. Soc.* **2019**, *141*, 5815; b) D. A. Thomas, R. Chang, E. Mucha, M. Lettow, K. Greis, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, *Phys. Chem. Chem. Phys.* **2020**, *22*, 18400; c) M. Lettow, M. Grabarics, K. Greis, E. Mucha, D. A. Thomas, P. Chopra, G. J. Boons, R. Karlsson, J. E. Turnbull, G. Meijer, R. L. Miller, G. von Helden, K. Pagel, *Anal. Chem.* **2020**, *92*, 10228.
- [21] W. Schöllkopf, S. Gewinner, H. Junkes, A. Paarmann, G. von Helden, H. P. Bluem, A. M. M. Todd, *Proc. SPIE-Int. Soc. Opt. Eng.* **2015**, 9512, 95121 L.
- [22] A. Supady, V. Blum, C. Baldauf, *J. Chem. Inf. Model.* **2015**, *55*, 2338.
- [23] F. Neese, *WIREs Comput. Mol. Sci.* **2012**, *2*, 73.
- [24] a) J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865; b) F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- [25] a) C. Adamo, V. Barone, *J. Chem. Phys.* **1999**, *110*, 6158; b) S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
- [26] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Wallingford, CT, **2016**.
- [27] D. Cremer, J. A. Pople, *J. Am. Chem. Soc.* **1975**, *97*, 1354.
- [28] M. Marianski, A. Supady, T. Ingram, M. Schneider, C. Baldauf, *J. Chem. Theory Comput.* **2016**, *12*, 6157.

Manuscript received: March 2, 2022
Revised manuscript received: March 23, 2022
Accepted manuscript online: March 24, 2022

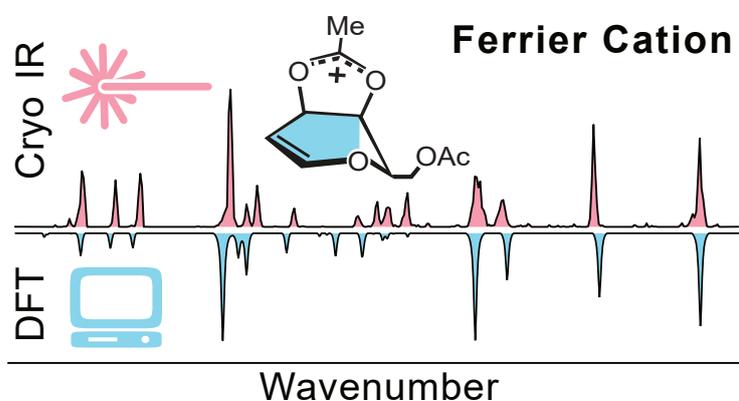
6 | The Ferrier Glycosyl Cation in the Gas Phase

This work is published^[305] and available online:

DOI: <https://doi.org/10.1021/acs.orglett.0c03301>

Details on the related Supporting Information can be found in Appendix C.

K. Greis, C. Kirschbaum, S. Lechnitz, S. Gewinner, W. Schöllkopf, G. von Helden, G. Meijer, P. H. Seeberger, K. Pagel*; Direct Experimental Characterization of the Ferrier Glycosyl Cation in the Gas Phase. *Org. Lett.* **2020**, *22*, 8916-8919.



Submitted: 1 October 2020; **Published:** 5 November 2020

Copyright © 2020 American Chemical Society

6.1 | Author Contributions

The project was conceived by **Kim Greis** and Kevin Pagel. **Kim Greis** wrote the first draft of the manuscript. **Kim Greis** (main) and Carla Kirschbaum (support) performed the mass spectrometry and cryogenic infrared spectroscopy experiments. The cryogenic infrared spectroscopy experiments were supervised by Gerard Meijer and Gert von Helden. Sandy Gewinner and Wieland Schöllkopf operated the free-electron laser. **Kim Greis** performed all quantum chemical calculations. Sabrina Lechnitz and Peter Seeberger provided the glycal building blocks. **Kim Greis** analyzed the data.

6.2 | Project Hypothesis and Summary

The Ferrier rearrangement reaction is a glycosylation reaction that involves glycals, which are monosaccharides featuring a double bond between the C1- and the C2-positions instead of hydroxyl groups.^[306–308] Upon activation, the C3-acetoxy group of fully acetylated glycals is cleaved leading to the formation of the Ferrier cation, a special type of glycosyl cations. Two structures have been postulated for this ion: either the positive charge generated at C3 is delocalized along the O–C1–C2–C3 axis or it is stabilized by neighboring-group participation of the C4-acetyl group. Nucleophiles commonly attack the C1-position, leading to a shift of the double bond to the 2,3 position (see Scheme 1 of the related published work). This reaction is an important step in the synthesis of certain pharmaceuticals^[309,310] and natural products.^[311] A fully protonated Ferrier cation was characterized by NMR spectroscopy in super acids, where only the structure with a delocalized double bond within the pyranose ring could be identified.^[82]

In this work, the structure of bare Ferrier cations is characterized experimentally for the first time using cryogenic IR spectroscopy in helium droplets. Ferrier cations are generated from fully acetylated glycals by electrospray ionization followed by in-source fragmentation, leading to cleavage of the C3-acetoxy group (see Figure 1 of the related publication). Ferrier cations from acetylated glucal and galactal were subsequently characterized. In combination with DFT-generated harmonic frequencies, the experimental spectra of both Ferrier cations can be clearly assigned to dioxolenium-type structures, in which the positive charge at the C3-position is stabilized by neighboring-group participation of the C4-acetyl group (see Figures 2 and 3 of the related publication). The 1500–1800 cm^{-1} region of the Ferrier cations is dominated by four absorption bands, which are originating from symmetric and antisymmetric dioxolenium, carbon-double bond, and carbonyl stretching modes. For oxocarbenium-type structures, in which the charge is delocalized within the pyranose ring, no matching absorption bands can be found. Furthermore, oxocarbenium-type structures are destabilized compared to dioxolenium-type structures by ca. 29–30 kJ mol^{-1} .

This work clearly shows that bare Ferrier cations only adopt dioxolenium-type structures originating from neighboring-group participation of the C4-acetyl group at the C3-position in the gas phase. Furthermore, similar to the previous chapters, this work is another example how cryogenic IR spectroscopy can be combined with DFT calculations to investigate the structure of reactive intermediates in organic reactions.

7 | The Fate of a Septanosyl Ferrier Cation in the Gas Phase

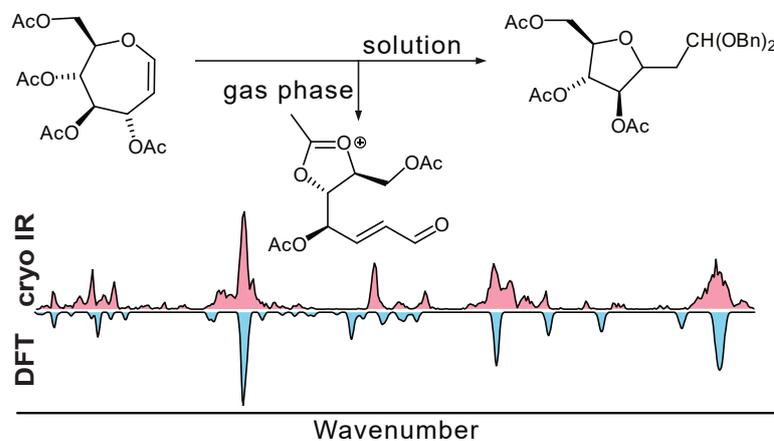
This work is published^[312] and available online:

DOI: <https://doi.org/10.1021/acs.joc.3c00079>

Details on the related Supporting Information can be found in Appendix D.

K. Greis,[†] C. E. Griesbach,[†] C. Kirschbaum, G. Meijer, G. von Helden, K. Pagel,^{*} M. W. Peczuł^{*}; Characterization and Fate of a Septanosyl Ferrier Cation in the Gas and Solution Phases. *J. Org. Chem.* **2023**, *88*, 5543–5553.

[†]equal contributions



Submitted: 12 January 2023; **Published:** 24 April 2023

Copyright © 2023 The Authors. Published by American Chemical Society

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

7.1 | Author Contributions

The project was conceived by **Kim Greis**, Caleb Griesbach, Kevin Pagel, and Mark Peczuh. **Kim Greis**, Caleb Griesbach, and Mark Peczuh wrote the first draft of the manuscript. **Kim Greis** (main) and Carla Kirschbaum (support) performed the mass spectrometry and cryogenic infrared spectroscopy experiments. The cryogenic infrared spectroscopy experiments were supervised by Gerard Meijer and Gert von Helden. **Kim Greis** (main) and Caleb Griesbach (support) performed the quantum chemical calculations. Caleb Griesbach synthesized the septanose building blocks and performed all glycosylation reactions. **Kim Greis** and Caleb Griesbach analyzed the data.

7.2 | Project Hypothesis and Summary

Building on the work of the previous chapter, the aim of this work was to characterize the intermediate of a Ferrier rearrangement reaction featuring a septanose instead of a pyranose ring. In parallel to this study, the collaborators who provided the precursors for the septanosyl Ferrier cations (carbohydrate-based oxepines),^[313] conducted condensed-phase reactions on the same system to get further information on the intermediate. Previously, it has been reported that the reactivity of the double bond of carbohydrate-based oxepines is similar to that of glycals^[314] and that these oxepines can undergo Ferrier rearrangement (see Figure 1 of the related published work).^[315]

Like fully acetylated pyranose-based glycals, septanose-based glycals can fragment by cleavage of an acetoxy group at the C3-position after electrospray ionization followed by in-source fragmentation. For septanose-based glycals derived from D-glucose and D-mannose, which only differ in the stereoconfiguration of the C3-atom, the generated fragment should be identical, as the C3-acetoxy group is cleaved to yield the Ferrier cation. The identity of both ions is confirmed by cryogenic IR spectroscopy (see Figure 3 of the related publication).

To identify the structure of the generated septanosyl Ferrier cations, several structural motifs were considered. In principle any of the three acetyl groups at the C4-, C5-, or the C7-atom could stabilize the positive charge at the C3- or at the C1-position. As the double bond is formally delocalized, the charge could be either at the initial C3-position or be shifted to the C1-position. Based on sampled lowest-energy structures for each structural motif, the dioxolenium-type structure featuring neighboring-group participation of the C4-acetyl group at the C3-position is overall the most stable. This motif is the same as experimentally confirmed in the previous chapter for pyranosyl Ferrier cations. However, its harmonic frequencies do not match the experiment (see Figure 4 of the related publication). Hence, the ions populating the ion trap must adopt a different structure.

Inspired by previous publications, including the work reported in Chapter 4,^[88,291] the rearrangement previously observed for pyranose-based glycosyl cations was considered to occur in septanosyl Ferrier cations. This rearrangement would proceed by attack of the carbonyl oxygen of the C4-, C5-, or C7-acetyl groups at the C6-atom, leading to ring opening and the formation of a positively charged dioxolenium- and a neutral enal moiety. Attack of the C5- or C7-acetyl group would lead to a five-membered dioxolenium moiety, whereas a six-membered ring would be formed upon the attack of the C4-acetyl group. The lowest-energy structures sampled for the C5- and C7-rearranged ions are stabilized by ca. 32–33 kJ mol⁻¹, compared to the structure exhibiting neighboring-group participation. The C4-rearranged ion would only be stabilized by 9 kJ mol⁻¹. Furthermore, the harmonic frequencies of the C5-rearranged structure match the experimental spectrum the best (see Figure 5 of the related publication). However, it is also apparent that the experimental signature is more complex and that therefore the ion trap might be populated by multiple structures. The additional absorption bands in the experimental spectrum are likely originating from the C7-rearranged structure which is only slightly destabilized compared to the C5-rearranged structure (ca. 1 kJ mol⁻¹). The harmonic frequencies of oxocarbenium-type structures, in which the charge is stabilized by delocalization within the septanose ring and long-range interactions of acetyl groups, match the experimental spectrum less well and their electronic energies are significantly higher than the most stable rearranged structure (ca. 48–59 kJ mol⁻¹).

In the condensed phase, this rearrangement did not occur. However, for a palladium-mediated *O*-glycosylation reaction, a ring-contraction of the septanose moiety to a furanose-moiety can be observed, if no precautions to exclude water from the reaction are taken (see Scheme 1 of the related publication). Under the exclusion of water, a typical Ferrier rearrangement occurs, leading to attack of a benzyl alcohol at the C1-position. This ring contraction was not observed in a *C*-glycosylation reaction using allyltrimethylsilane as nucleophile (see Scheme 2 of the related publication).

Both the gas-phase and the condensed-phase work show the intrinsic instability of seven-membered rings. Although it is known that five- and six-membered rings are more stable than their seven-membered counterparts, specific examples for ring-contraction in carbohydrates are less known. The gas-phase work shows that cryogenic IR spectroscopy is suitable to determine the structure of intermediates of organic reactions. Even though the observed rearrangement is exclusively observed the gas phase, it helps understanding the intrinsic reactivity of this system.

Characterization and Fate of a Septanosyl Ferrier Cation in the Gas and Solution Phases

Kim Greis,^{||} Caleb E. Griesbach,^{||} Carla Kirschbaum, Gerard Meijer, Gert von Helden, Kevin Pagel,^{*} and Mark W. Peczuł^{*}



Cite This: *J. Org. Chem.* 2023, 88, 5543–5553



Read Online

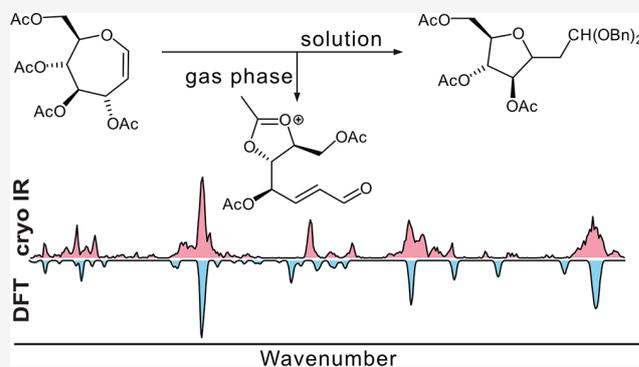
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Ferrier reactions follow a mechanistic pathway whereby Lewis acid activation of a cyclic enol ether facilitates departure of an allylic leaving group to form a glycosyl Ferrier cation. Attack on the Ferrier cation provides a new acetal linkage concurrent with the transposition of the alkene moiety. The idiosyncratic outcomes of Ferrier reactions of seven-membered ring carbohydrate-based oxepines prompted an investigation of its corresponding septanosyl Ferrier cation. Experiments that characterized the ion, including gas-phase cryogenic IR spectroscopy matched with density functional theory-calculated spectra of candidate cation structures, as well as product analysis from solution-phase Ferrier reactions, are reported here. Results from both approaches revealed an inclination of the seven-membered ring cation to contract to five-membered ring structures. Gas-phase IR spectra matched best to calculated spectra of structures in which five-membered dioxolenium formation opened the oxepine ring. In the solution phase, an attack on the ion by water led to an acyclic enal that cyclized to a C-methylene-aldehydo arabinofuranoside species. Attack by allyl trimethylsilane, on the other hand, was diastereoselective and yielded a C-allyl septanoside.



INTRODUCTION

Glycols have proven to be valuable starting materials for the synthesis of numerous oligosaccharides, glycosylated natural products, and even small-molecule targets.^{1–3} The advantages of these compounds come from their inherently rich stereochemistry, the unique reactivity of their enol ether units, and the ability to modulate their reactivity by varying the protecting groups attached to the oxygens.^{2,4} The archetypal reaction of glycols is the functionalization of the double bond with an electrophilic oxygen species (e.g., DMDO) followed by nucleophilic attack on the newly formed 1,2-anhydro sugar (not shown) to form a glycosidic bond, as depicted for the conversion of D-glucal **1** to methyl β-glucoside **2** (Figure 1a).^{5,6} Ring-expanded glycols, informally referred to as carbohydrate-based oxepines (i.e., **3** in Figure 1a), react in a similar fashion. Glucose-based oxepine **3**, for example, was converted to methyl β-D-glycero-D-guloseptanoside **4** and α-D-glycero-D-idoseptanoside **5** under conditions that were nearly identical to those used for glycols. The diastereomeric mixture of glycosides in the case of the seven-membered ring system mostly reflected the low selectivity of epoxidation of oxepine **3**.⁷ Nonetheless, the common reactivity pattern of glycols and oxepines in terms of the direct addition across the enol ether double bond is apparent.

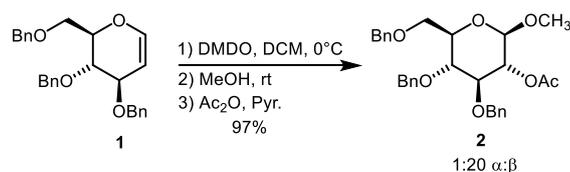
The Ferrier rearrangement is another reaction typical of glycols (i.e., conversion of **6** to **7** in Figure 1b).⁸ Here, the nucleophilic attack is concomitant with the migration of the double bond as a leaving group is ejected from C3. For the Ferrier reaction, formation of a glycosyl cation under the action of a Lewis acid is essential and nucleophilic attack occurs therefore under S_N1 conditions. Hence, the stereoselectivity depends on the nature of the glycosyl cation and preferred pathways for additions to it.^{9–11} We recently reported that, under conditions established for glycols, oxepine **8**¹² could be converted to hexafluoroisopropyl 2,3-dideoxy-β-D-arabino-hex-2-enoseptanose **9** by a Ferrier rearrangement.¹³ Even though the yield was modest, the reaction reinforced the similarity in reactivity of oxepines to glycols. To our surprise, the C3-epimeric oxepine **10**,¹² derived from D-mannose, was unreactive under the conditions that afforded the Ferrier product from **8**. The low-energy conformations of **8** and **10** are largely the same—both favor ⁴H₆ conformations with minor

Received: January 12, 2023

Published: April 24, 2023



a) Direct Functionalization



b) Ferrier Reaction

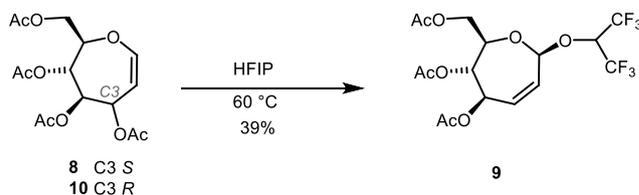
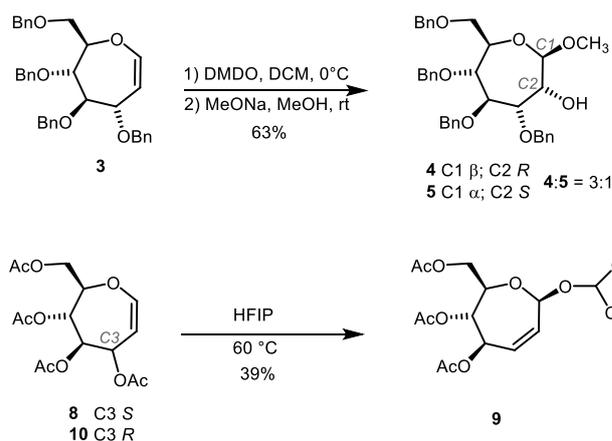
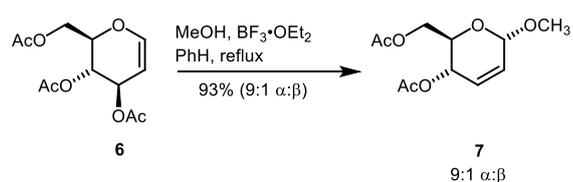


Figure 1. (a) Direct functionalization of D-glucal **1** and D-glucose-based oxepine **3** via epoxidation and methanolysis; (b) Ferrier reactions of D-glucal **6** and D-glucose-based oxepine **8**; D-mannose-based oxepine **10** does not react under the HFIP Ferrier conditions.

populations of twist-half and chair structures. The consequence is that the C3-acetyl group of **8** is pseudoaxial but pseudoequatorial for **10**. We invoked the vinylogous anomeric effect as part of the explanation of this differential reactivity of the oxepines. Furthermore, the preference for the β -anomeric configuration of septanoside **9** was initially unexpected, considering that Ferrier rearrangements with D-glycals have often favored the α -anomer. We speculated that the β -selectivity likely arose either *via* a preference for selective β -attack into a cationic intermediate or *via* anomerization to the thermodynamic product. The latter rationale was reinforced by the susceptibility of the hexafluoroisopropyl group to anomeric stabilization compared to less electron-withdrawing aglycons.¹⁴

The results from our initial investigation into the Ferrier reactivity of carbohydrate-based oxepines **8** and **10** challenged us to consider in greater detail the cationic intermediate—henceforth referred to as the septanosyl Ferrier cation. This intermediate is generated after the cleavage of the C3 protecting group, with the positive charge formally localized at the C3 atom. The mechanism of glycosylation depends on several factors—especially, the structure of the reactant and the reaction conditions. Depending on the conditions, the mechanism will fall somewhere along an S_N1 – S_N2 continuum.^{15,16} Traditionally, Ferrier rearrangements were considered to proceed through an allyl oxocarbenium ion intermediate.^{2,17} However, a recent report using cryogenic vibrational spectroscopy in the gas phase revealed that, in isolation, the Ferrier cations generated from acetylated D-glucal and D-galactal exist as dioxolenium ions stabilized by neighboring-group participation (NGP) from an acetyl group at the C4 position on the ring.¹⁸ In another study, fully protonated Ferrier cations stabilized by superacids were measured by NMR spectroscopy.¹⁹ Due to their reduced nucleophilicity in this medium, the acetyl groups do not engage in NGP. We reasoned that septanosyl Ferrier cations prepared from oxepines **8** and **10** might similarly be subjected to NGP or long-range participation (LRP, sometimes termed remote participation) which could influence the outcomes of reactions involving them.²⁰ Herein, we report a two-pronged approach to investigate the septanosyl Ferrier cation generated from oxepine **8** or **10** in the gas phase using cryogenic vibrational spectroscopy and in the solution phase by Ferrier reactions followed by product characterization. Our investigation reveals

74

a preference for α -attack and subsequent anomerization of the product. In addition, we observe the preference of the septanosyl Ferrier cation to ring-contrast to a thermodynamic product in the gas and condensed phases.

Previous reports from our group chronicle a proclivity toward complex reactivity by cationic septanosyl intermediates. Notable among these instances were intramolecular reactions that formed bicyclic products.^{11,21} We also reported a ring contraction in which methyl septanoside **11** was converted to substituted C-methylene-aldehydo arabinofuranoside **12** (Figure 2).²² This unexpected product arose under conditions

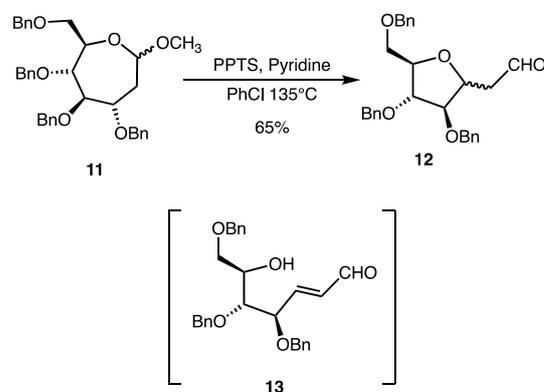


Figure 2. Previously reported ring contraction of methyl 2-deoxyseptanoside **11** to C-methylene-aldehydo arabinofuranoside **12** under acidic conditions *via* a putative enal intermediate **13**.

aimed at performing a regioselective, acid-mediated elimination of methanol across the C1–C2 bond of **11** to deliver oxepine **3**. An α,β -unsaturated aldehyde, **13**, was invoked as a likely intermediate in the transformation. Enal **13** underwent *oxa*-Michael addition to form a unique C-methylene-aldehydo arabinofuranoside **12**. During the investigation of the septanosyl Ferrier cation reported here, we observed a similar ring contraction, as detailed in the Results and Discussion. Taken together, the two examples of ring contractions highlight a hierarchy of thermodynamic stabilities where seven-membered rings are less stable compared to five- and six-membered rings. This hierarchy is not unique to systems where ring contractions are a thermodynamic sink.²³ In fact,

dynamic equilibria between minor seven- and major five-membered ring products can be observed in aqueous media.^{24,25}

RESULTS AND DISCUSSION

Characterization of a Septanosyl Ferrier Cation in the Gas Phase. Nano-electrospray ionization (nESI) of per-*O*-acetyl oxepines **8** and **10**, derived from glucose and mannose, respectively, yielded three main signals at m/z 285, 367, and 711 (Figure S1). These signals correspond to $[M - OAc]^+$, $[M + Na]^+$, and $[2M + Na]^+$ ions. The $[M - OAc]^+$ ion most likely arises from cleavage of the C3-acetoxy group, leading to a Ferrier-like carbocation (Figure 3a). Traditionally, such ions

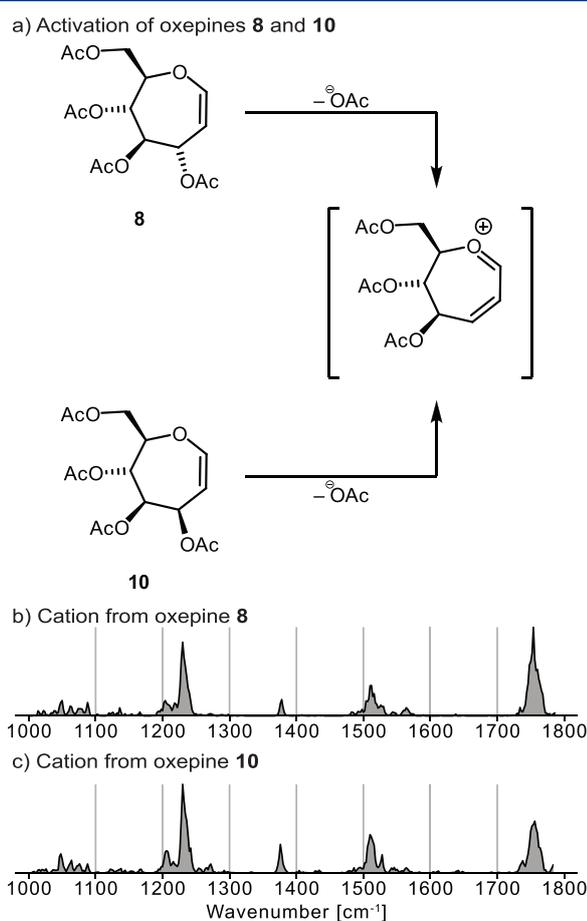


Figure 3. (a) Cleavage of the C3-acetoxy group from oxepines **8** and **10** lead to the same septanosyl Ferrier cation. Cryogenic infrared spectroscopy in helium nanodroplets of septanosyl Ferrier cations $[M - OAc]^+$ generated from (b) glucose- and (c) mannose-derived per-*O*-acetyl oxepines **8** and **10**, respectively, reveals the identity of both cations.

can be stabilized by resonance and/or by the participation of one of the remaining acetyl groups. Based on the mass spectrum alone, oxepines **8** and **10** cannot be differentiated. Furthermore, the septanosyl Ferrier cations generated from both precursors should be identical as they only differ in the absolute configuration of the group at C3, which is cleaved upon activation.

To find out if the septanosyl Ferrier cations generated from oxepines **8** and **10** are identical and what their structure is, they were investigated using cryogenic infrared (IR) spectroscopy

(Figure 3b,c). Recently, this technique was used to investigate the structure of Ferrier cations generated from glycol precursors¹⁸ and to probe intramolecular interactions in various glycosyl cations.^{26–29} The IR spectra displayed in Figure 3b reveal that both ions are identical as their IR signatures are essentially superimposable. Hence, as anticipated, cleavage of the C3-acetoxy group gives rise to the same cation. Generally, the vibrations observed in the fingerprint region around 1000–1300 cm^{-1} can be assigned to C–C and C–O stretches, whereas the absorption bands observed in 1300–1450 cm^{-1} originate from C–H bends. As has been observed in related systems,¹⁸ the functional group region contains symmetric and antisymmetric dioxolenium (COO^+), oxocarbenium ($C=O^+$), and C=C-stretches in the 1450–1700 cm^{-1} range, while carbonyl ($C=O$) stretches are commonly found around 1700–1800 cm^{-1} .

To get insight into the structure of the Ferrier-like ion, the experimental spectrum was compared to computed spectra derived from harmonic frequency calculations for several possible structural motifs. Structural motifs that were considered made use of the C4-, C5-, or C6-acetyl groups to stabilize the positive charge of the oxocarbenium ion *via* NGP or LRP. Because the charge of the Ferrier-like cation is formally delocalized along the four-atom O–C3 unit of the septanose ring, the acetyl groups could participate at both C1 and C3 positions. It has been determined that such structures—where the positive charge is stabilized by Ferrier glycosyl cations based on pyranose sugars.¹⁸ Geometries for each structural motif were built and their conformational spaces were sampled. For each motif, a subset of low-energy structures was selected for reoptimization and computation of harmonic frequencies at a higher level of theory PBE0+D3/6-311+G(d,p).^{30–33} For the lowest-energy structure of each motif, more accurate single-point energies were obtained at the DLPNO-CCSD(T)/Def2-TZVPP^{34–36} level of theory. Overall, similar to pyranose-based Ferrier cations, the overall lowest-energy structure is a cation in which the charge is stabilized by the NGP of the C4-acetyl group (I). Hence, this lowest-energy structure serves as a reference. Computed IR spectra of the lowest-energy structure for each of the six structural motifs are depicted in Figure 4.

Based on the lowest-energy structures for each structural motif, the stability decreases in the following order C4_C3_NGP (I) > C5_C3_LRP (II) > C7_C1_LRP (III) > C5_C1_LRP (IV) > C4_C1_LRP (V) > C7_C3_LRP (VI) (Figure 4). This ranking indicates that the relative stability of the respective structural motif is dependent on the ring size of the newly formed ring after participation, which are five- (I), six- (II), seven- (III, IV, V), and eight-membered rings (VI). Relative to I, the other interactions are destabilized by 15–62 $kJ mol^{-1}$. Generally, and similar to previous studies,^{18,28} NGP is always favored over LRP.

Previously, the identity of the two septanosyl Ferrier cations was confirmed based on their experimental IR spectra. To assign a computed structure to the experimental spectrum, it has been rerecorded with a higher power of the free-electron laser, leading to a better-resolved spectrum (Figure 4). The experimental spectrum is significantly crowded in comparison to the computed spectra, suggesting that the ensemble of previously mass-to-charge selected ions was composed of more than one conformer. It is possible that there was more than one structural motif present in the ion trap. While some

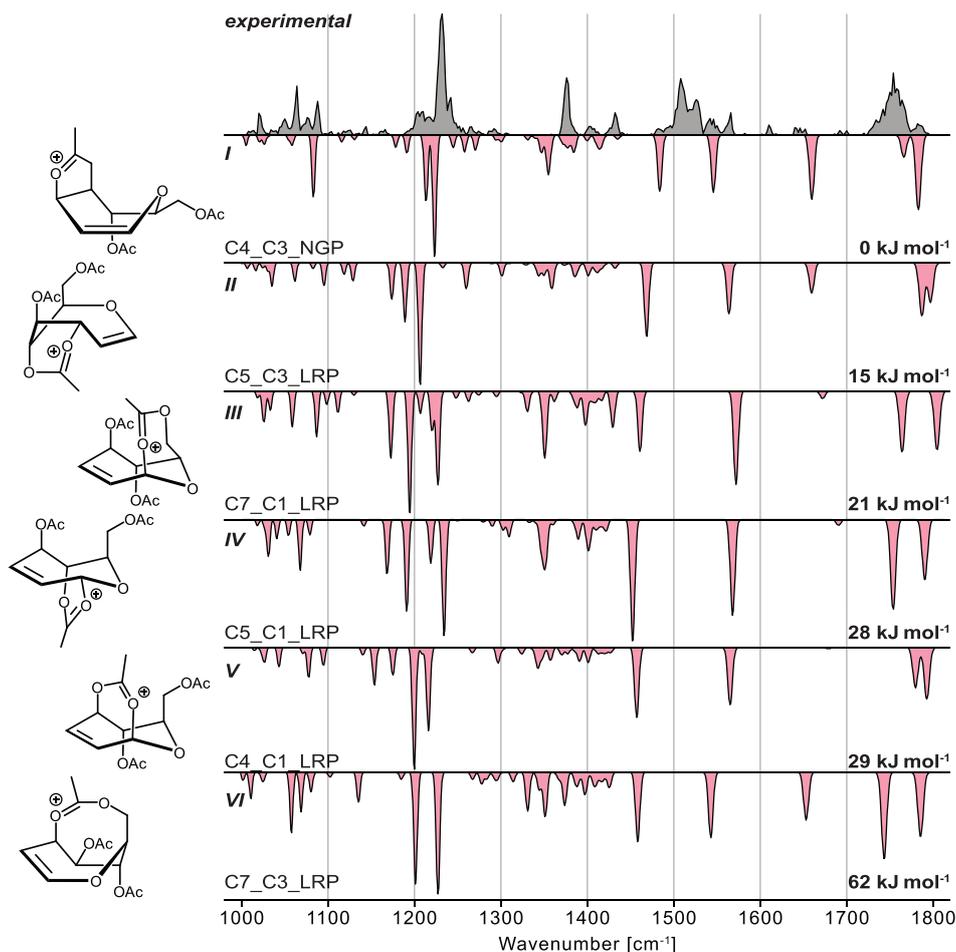


Figure 4. Experimental infrared spectrum (gray) of the septanosyl Ferrier cation $[M - OAc]^+$ compared to computed spectra (red, inverted traces) of structures exhibiting (I) NGP of the C4-acetyl group at the C3 position, (II) LRP of the C5-acetyl group at the C3 position, (III) LRP of the C7-acetyl group at the C1 position, (IV) LRP of the C5-acetyl group at the C1 position, (V) LRP of the C4-acetyl group at the C1 position, and (VI) LRP of the C7-acetyl group at the C3 position. The relative free energy at 90 K as well as schematic depictions of each structure are indicated.

harmonic frequencies of the intermediate exhibiting NGP (I) have matching absorption bands to the experimental spectrum, they are low in intensity. Hence, such a structure may be present in the ion trap but only to a lesser extent. The computed IR spectra of the other species match even less well with the experimental spectrum than the spectrum of I. Thus, based on their poor match and their unfavorable energetics, these structures can be discarded. Furthermore, the lowest-energy structure—if present among these—is only partially populating the ion trap.

Due to the unsatisfactory structural match of I with the experimental spectrum, other structural motifs were considered as well. Recent publications reported that rearrangement occurs for certain pyranose-based glycosyl cations in the gas phase.^{29,37} There, an acetyl group attacks the C5 carbon atom of a pyranose, leading to the opening of the pyranose ring and the formation of a five-membered dioxolenium moiety and an aldehyde group. Such dioxolenium ions have previously been stabilized in super acids, where they rearranged to oxonium ions.³⁸ In our system, rearrangement could potentially arise from the attack of each of the acetyl groups at C4, C5, and C7 onto the C6 atom of the seven-membered ring. Mechanistically, such an attack would proceed *via* an S_N2 mechanism, hence leading to inversion of the stereoconfiguration at C6.

76

Therefore, the configuration at C4/C5/C6 of the rearranged septanosyl Ferrier cations would be (R,S,S). Although mechanistically less likely, the C6-epimers of the rearranged ions were considered as well. Additionally, species that did not employ the participation of an acetyl group—oxocarbenium structures—were investigated. Similar to the previously considered structures, the conformational space of the new structural motifs was sampled, a subset of low-energy structures was reoptimized, and harmonic frequencies were computed at a higher level of theory. Computed IR spectra of the lowest-energy structures in comparison to the experimental spectrum are depicted in Figures 5 and S2. It is apparent that the rearranged species are significantly lower in energy by 9–33 kJ mol⁻¹ than the one stabilized by NGP (I). The oxocarbenium structures are higher in energy relative to I. The computed energetics of the rearranged structure formed by the C5-(VII) or the C7-acetyl group (VIII) are very similar; however, the computed spectrum of the C5_rearranged structure matches the experimental spectrum slightly better. Here, mainly the carbonyl stretches of the free acetyl groups at 1761 and 1756 cm⁻¹, the symmetric and antisymmetric dioxolenium stretches at 1511 and 1569 cm⁻¹, and the C–O stretch at 1230 cm⁻¹ match exceptionally well. C4_rearranged structures (IX) or the lowest-energy oxocarbenium-type

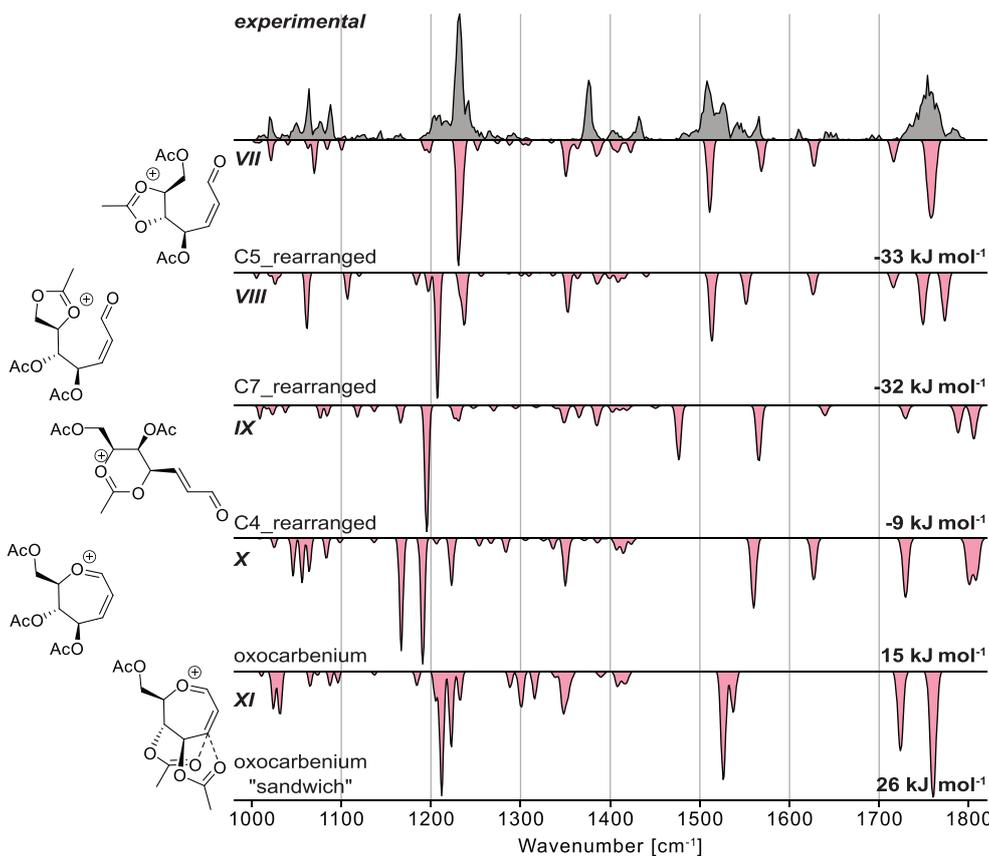


Figure 5. Experimental infrared spectrum (gray) of the septanosyl Ferrier cation $[M - OAc]^+$ compared to computed spectra (red, inverted traces) of structures exhibiting rearrangement by the attack of the (VII) C5-, (VIII) C7-, and (IX) C4-acetyl groups at the C6 position leading to ring opening, (X) oxocarbenium structures that are stabilized by one acetyl group interacting *via* long-range, or (XI) oxocarbenium-type structures stabilized by long-range interaction of two acetyl groups. The relative free energy at 90 K as well as schematic depictions of each structure are indicated.

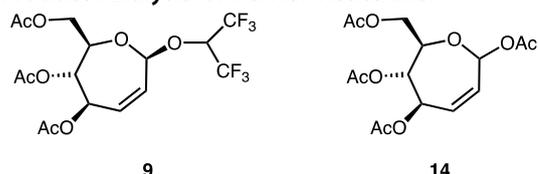
structures (X and XI) generally match less well. However, in one oxocarbenium structure, the charge center is “sandwiched” by two acetyl groups (XI) (Figure S5), leading to a strong change in IR absorption. The gaps in the experimental spectrum that cannot be filled by VII are, based on the energetics, most likely filled by structure VIII, while matching absorption bands can also be observed for the higher-energy structures I and XI.

Transition states leading to the rearranged structures (Figure S3a) indicate a comparably low barrier for rearrangement of ca. 68–84 kJ mol^{-1} . This value is in line with previously computed transition states for rearrangement of pyranose-based glycosyl cations (35–138 kJ mol^{-1}).²⁹ The required energy for the rearrangement is transferred to the ion during the ion-source fragmentation process. The C5- and C7-rearranged species VII and VIII cannot directly interconvert into each other. However, the (*R,S,S*) (C4/C5/C6) diastereomer of the C5-rearranged ion can convert to the (*R,S,R*) diastereomer of the C7-rearranged ion *via* $\text{S}_{\text{N}}2$ attack of the C7-acetyl group at C6. The same reaction can proceed for the (*R,S,S*) diastereomer of the C7-rearranged ion. The barrier of this reaction is computed to be only 61–62 kJ mol^{-1} (Figure S3b). However, the formation of the diastereomers is thermodynamically not favored.

Overall, our assessment of the experimental and computational data in the gas phase is that the most abundant species observed is the ring-opened ion VII formed by the attack of

the C5-acetyl group at the C6 position. However, there are peaks in the experimental region that are broader than predicted by theory and some peaks are not reproduced at all by ion VII. Therefore, it is likely that a fraction of the probed ions adopt other structures, such as C7_rearranged (VIII). Importantly, the fact that the seven-membered ring system followed trajectories on the potential energy surface to form the five-membered rings was significant. The propensity for seven-membered ring cations to decompose into the more stable five-membered rings cannot only be observed in the gas phase but also in solution-phase experiments *via* a different mechanism for ring opening (*vide infra*).

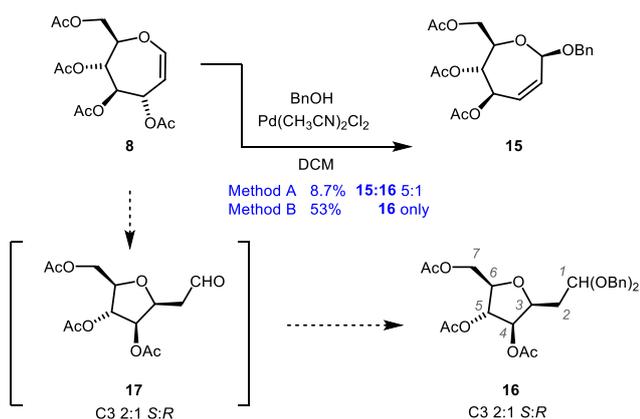
Characterization of the Septanosyl Cation in Solution via Product Analysis of Ferrier Reactions.



Previous Ferrier reactions of oxepine 8 using alcoholic nucleophiles were of mixed success. Initial reactions of 8 with benzyl alcohol in the presence of Lewis acids (i.e., FeCl_3 and $\text{BF}_3 \cdot \text{OEt}_2$) gave intractable product mixtures. On the other hand, HFIP septanoside 9 and septanose acetate 14 were prepared in modest yields (39 and 26%, respectively)¹³ under less common conditions.^{39,40} We then turned to a palladium-

mediated rearrangement introduced by Galan and Sau.² Care was taken to dry donor **8** and benzyl alcohol by azeotropic distillation before dissolving in anhydrous dichloromethane. Addition of vacuum-desiccated Pd(MeCN)₂Cl₂ to the solution while under a nitrogen atmosphere (Method A) resulted in the disappearance of the starting material by TLC and the appearance of two new spots. One fraction isolated by chromatography in low yield (8.7%) proved to be a 5:1 mixture of Ferrier product **15** and C-methylene-acetal arabinofuranoside **16** (Scheme 1 and Figure 6a). The similarity

Scheme 1. Pd(II)-Mediated Ferrier Reaction and Ring Contraction



of the olefinic ¹H signals for H2 and H3 in the mixture to those of earlier Ferrier product **9** helped to assign the structure of **15**. Irradiation of H2 in a selective TOCSY experiment was used to set the signals of the seven-membered ring. The β-configuration of the anomeric center was assigned based on its similarity to the C1 chemical shift of **9** and was reinforced by an NOE cross peak between H1 and H6 (see the Supporting Information). Structural assignment of **16** was done retroactively based on additional experiments (*vide infra*). The other product fraction from chromatography was C-methylene-aldehyde arabinofuranoside **17**. Compound **17** proved to be unstable in our hands; we were, however, able to conduct an explicit experiment to isolate it (12% BRSM)⁴¹ and collect NMR spectra used in its structural assignment. Analysis of the data revealed that **17** was isolated as a 2:1 mixture of stereoisomers at the C3 position (Scheme 1). Observation of this species suggested both the likely structure of **16** and an experiment to prepare it.

The appearance of the C-methylene-aldehyde compound **17** was reminiscent of compound **12** (Figure 2) that had arisen under reaction conditions where an oxocarbenium ion was a plausible intermediate. In that previous case, adventitious water was implicated in formation of the ring-contracted compound. Running the palladium-mediated Ferrier reaction under conditions where measures to remove water were not taken (Method B) consequently resulted in the isolation of compound **16** as the sole product in a 53% yield (BRSM). In fact, it was the analysis of NMR spectra of the sample of **16** (Figure 6c) obtained under these conditions that enabled its assignment as a product in the earlier anhydrous reaction. Presumably, attack on the septanosyl Ferrier cation by water forms a lactol that tautomerizes to the unsaturated aldehyde; *oxa*-Michael addition by the C6 hydroxyl then leads to C-methylene aldehyde species **17**, followed by acetalization to

78

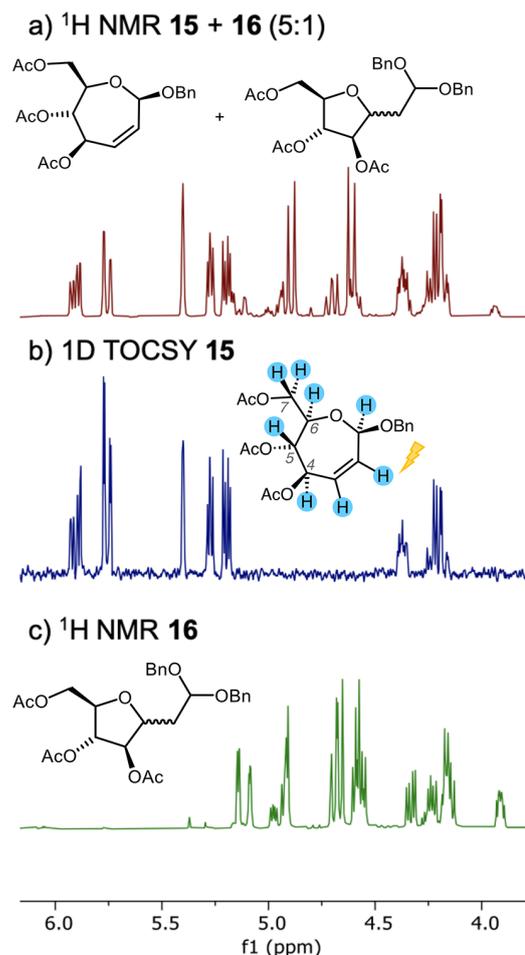
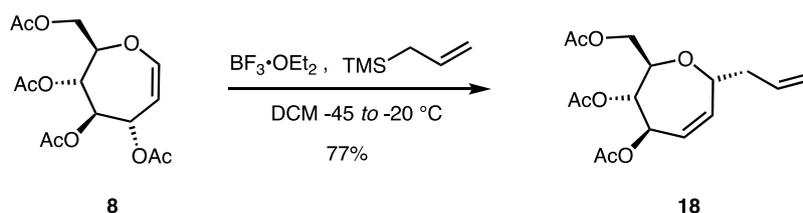
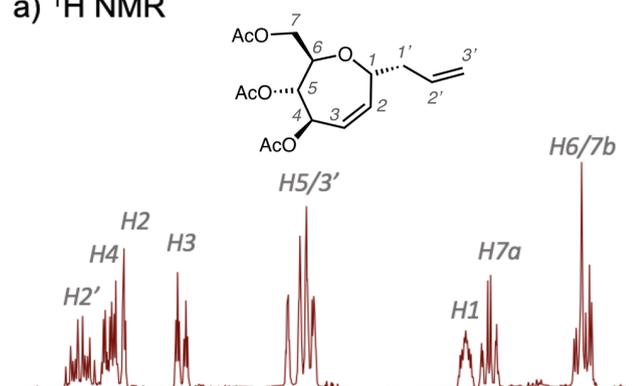


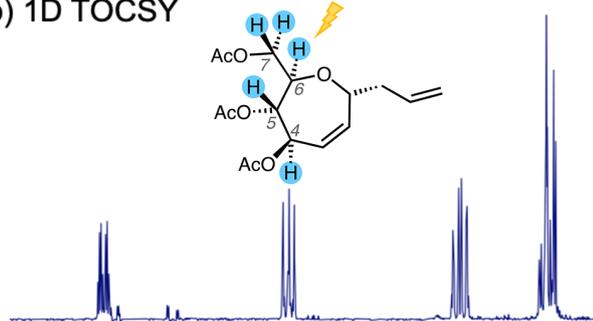
Figure 6. (a) ¹H NMR spectrum of an isolated fraction of the anhydrous Ferrier rearrangement containing **15** and **16** (5:1); (b) selective 1D TOCSY spectrum of **15** arising from irradiation of the H2 chemical shift region (δ 5.76 ppm) on a sample of **15** + **16** (5:1); (c) ¹H NMR spectrum of major product **16** arising from the “wet” Ferrier rearrangement.

provide **16** as a 2:1 mixture of C3 diastereomers (2:1 S/R. The major isomer is shown in Scheme 1). See the Supporting Information for additional spectroscopic details on the structure of **16**.

Additional evidence of the septanosyl Ferrier cation in the solution phase was inferred by characterizing the product of a kinetic trap experiment. Allylation was used because the stereocenter formed in the reaction reflects the facial selectivity of attack and is unable to equilibrate to a thermodynamic product.^{11,19} The reaction was performed on oxepine **8** (Scheme 2), where reagents were added at -45 °C. The reaction was allowed to warm to -20 °C and held at that temperature for 1 h. A single product was isolated from the reaction mixture, in a 77% yield, whose NMR spectra proved to be consistent with allyl C-septanoside **18**. In deuteriochloroform, the ¹³C{¹H} NMR spectrum of the product showed one set of signals, indicating that a single diastereomer was the product of the reaction. The ¹H NMR spectrum, however, suffered from overlapping signals that prevented the analysis of ³J_{H,H} coupling constants and H,H NOEs to characterize which stereoisomer had formed. When the solvent was changed to acetone-*d*₆ (Figure 7a), several of the signals became

Scheme 2. Allylation Reaction on Oxepine 8 to Form α -C-Allyl Septanoside 18a) ^1H NMR

b) 1D TOCSY



c) 1D NOE

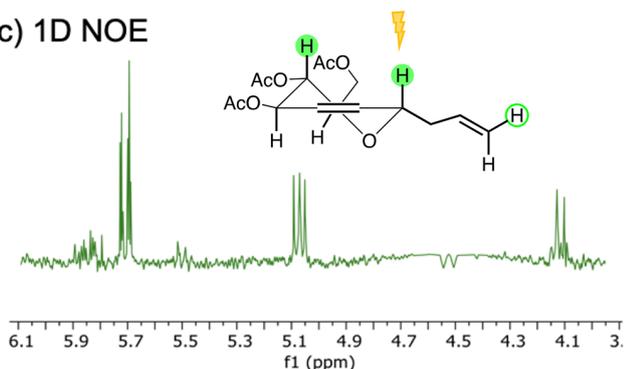


Figure 7. (a) Detail of the ^1H NMR spectrum of α -C-allyl septanoside **18** with signals assigned according to the structure shown. (b) 1D TOCSY spectrum of **18** arising from irradiation of the H6/H7b chemical shift region (δ 4.02 ppm); the structure highlights protons assigned *via* spin diffusion. (c) 1D NOE spectrum of **18** arising from irradiation on H1.⁴² NOE assignments based on the $^5\text{H}_0$ conformer of α -C-allyl septanoside **18**.

sufficiently resolved to enable analysis of NOEs. Both 1D and 2D ^1H – ^1H NOESY NMR spectra showed NOEs between H1 (δ 4.44 ppm) and a signal (δ 5.04 ppm) that corresponded to H5 and/or a vinylic signal from the allyl group (H3'). To tease out if H1 was in close proximity to both or just one of these

protons, a selective 1D TOCSY experiment was implemented. Irradiation of the chemical shift region shared by signals for H6 and H7 (δ 4.02 ppm) identified the spin system corresponding to H4 through H7 and enabled the assignment of H5 (Figure 7b). In light of that assignment, and with regard to the likely preference for the $^5\text{H}_0$ conformation¹³ of the compound, it was clear that it was H1 that overlapped with H5 in the NOE experiments (Figure 7c). Based on this NOE and the known configuration of C5, we assigned the product as the α -anomer, compound **18**.

The α -configured C-glycoside **18** formed in the kinetic trap experiment was of the opposite configuration of HFIP glycoside **9** that we had observed previously, as well as benzyl septanoside **15**. This result suggested that, in O-glycosylation reactions, the product equilibrates to the thermodynamically favored anomer. Furthermore, the selective formation of the α -anomer was consistent with a similar kinetic trap experiment where D-glucal was used as the starting material and the α -configured C-allyl glucoside was isolated as a product. Concomitant with the results of the kinetic trap experiment, $^{13}\text{C}\{^1\text{H}\}$ NMR spectra were collected in superacid media in conjunction with density functional theory calculations and used to characterize a protonated oxocarbenium ion exhibiting a β -face that was significantly hindered by the C6 acetoxy group.¹⁹ Looking at the calculated structures of oxocarbenium ion conformers, the C4_C3_NGP I and the oxocarbenium “sandwich” species XI should be quite similar; an attack on either of them should favor the α -face because it minimizes transannular interactions between the ring and the nucleophile (Figure S6).^{9,10} Particularly, C4_C3_NGP I—the lowest-energy cation—has a β -face hindered by the participating acetate and acetoxy group. The α -face of cation I is unhindered, and nucleophilic attack is expected to be highly stereoselective; however, other ions have similar profiles. Upon generation of an oxocarbenium ion, stabilization of the electrophilic C1 and C3 positions by the C4 and C7 acetates would effectively block its β -face. Also, the α -product probably adopts a half-chair conformation that projects substituents in a quasi-equatorial arrangement.

CONCLUSIONS

The propensity for ring opening of the septanosyl Ferrier cation is the common theme that emerged from the gas- and solution-phase experiments reported here. IR spectra of the gas-phase ion match with computed spectra where an acetyl group at either C5 or C7 attacks at C6, rupturing the septanosyl ring to form a five-membered dioxolenium ion (i.e., cations VII and VIII in Figure 5) with a pendant enal moiety. These dioxolenium enals were also calculated to be among the most stable structures of the studied cations. Previously, a similar rearrangement was reported in the gas phase for pyranose-based glycosyl cations.^{29,37} In the solution phase, Ferrier product benzyl septanoside **15** was only obtained in low yield

and it was accompanied by the ring-opened C-methylene dibenzyl acetal **16**. Ring opening was facilitated by the attack of adventitious water onto the septanosyl Ferrier cation. The subsequent septanose lactol then relaxed to the acyclic enal followed by an *oxa*-Michael attack, delivering C-methylene-aldehyde species **17**. Under conditions where measures to remove water were abandoned, ring-contracted product **16** predominated. Critically, the septanosyl Ferrier cation could be efficiently trapped under kinetic conditions. Using allyl trimethylsilane as a nucleophile C-allyl septanoside **18** was obtained with good yield and diastereoselectivity. In total, the results suggest that Ferrier reactions of oxepines **8** and **10** with alcoholic nucleophiles will be vexed by low yields, but they should be amenable to formation of other C-glycosides.⁴³ Finally, the results from the gas and solution phases both point to the formation of a stable five-membered ring from a less stable seven-membered ring. While the relative stabilities of these rings (in combination with six-membered rings) are considered to be well established,^{24,25} specific examples demonstrating them are rare.

EXPERIMENTAL SECTION

General Methods. Commercially available reagents were used without further purification with the exception of benzyl alcohol which was checked for benzaldehyde before use in experiments and distilled over potassium hydroxide when necessary. Solvents for anhydrous reactions were dried over calcium hydride and distilled. Solid reagents were dried in a vacuum desiccator in the presence of phosphorous pentoxide as a desiccant prior to use. Compounds not purchased were synthesized in accordance with the literature precedent and matched reported spectra. Structural assignments were made with additional information from gCOSY, gHSQC, and gHMBC experiments.

Mass Spectrometry and Cryogenic Infrared Spectroscopy. Acetylated oxepines derived from glucose- and mannose-based oxepines were dissolved in a mixture of acetonitrile and water (9:1, V/V) to yield 100 μM solutions. The oxepines were ionized by nano-electrospray ionization (nESI) on a custom-built mass spectrometer that allows for infrared ion spectroscopy in helium nanodroplets, previously described in detail.^{44–46} For nESI, Pd/Pt-coated glass capillaries prepared in-house were used. Septanosyl Ferrier cations are generated by in-source fragmentation of protonated or sodiated oxepines. The ion beam is focused by two ring-electrode ion guides, and the ions of interest are mass-to-charge-selected in a quadrupole mass filter. Subsequently, the ions are guided to a quadrupole bender, where the ions either pass through to a time-of-flight detector to monitor the ion signal and record mass spectra or are bent into a hexapole ion trap. Here, the ions are thermalized by buffer gas cooling to the temperature of the ion trap (90 K) achieved by cooling with liquid nitrogen.

A beam of superfluid helium nanodroplets (0.37 K) is generated by a pulsed Even-Lavie valve (nozzle temperature of 21 K). The helium nanodroplets pass through the hexapole ion trap, picking up the ions and guiding them to a detection region where the beam of doped helium nanodroplets overlaps with an infrared (IR) beam of the Fritz Haber Institute free-electron laser (FHI FEL).⁴⁷ Infrared radiation leads to the excitation of resonant vibrational modes of the analyte ions. By relaxation, the energy is dissipated to the helium matrix that subsequently evaporates. The helium matrix acts as a cryostat that keeps the ions at 0.4 K. After the absorption of multiple IR photons, the ion is released from the droplet and detected by a time-of-flight detector. Monitoring the ion yield as a function of the IR wavenumber leads to an IR spectrum. The ions were probed in the 1000–1800 cm^{-1} range.

Computational Methods. To assign a structure to the intermediate ion characterized by infrared ion spectroscopy, the conformational space of potential candidates was sampled using the

software CREST⁴⁸ (version 2.9) with the semiempirical method GFN2-xTB,⁴⁹ the empirical method GFN-FF⁵⁰ (using xtb version 6.3.0) and Schrödinger Maestro^{51,52} (version 2021-3). As the C3-acetyl group in the oxepines is cleaved, several structural motifs are conceivable (displayed in Figures 4 and 5). The conformational spaces of non-rearranged dioxolenium-type structures exhibiting long-range or NGP were sampled using CREST with GFN2-xTB, while the other structures were sampled using Maestro and CREST with GFN2-xTB/GFN-FF. Sampling these other structures in CREST with GFN2-xTB is nontrivial as these structures often tend to rearrange or form erroneous bonds during sampling.

Oxocarbenium and rearranged dioxolenium ions were loaded into Schrödinger Maestro.^{51,52} A Monte Carlo search using the OPLSe forcefield in vacuum was performed to sample the conformational space for each ion. Newly found conformers within 63 kJ mol^{-1} were tested by an rmsd statistic. Conformers with an rmsd > 0.5 Å from all previously generated conformers were considered unique. These were then optimized in Maestro at a PBE0+D3/6-31G(d) level of theory and again tested for uniqueness by an rmsd statistic.

All geometries generated by the CREST sampling were optimized at the PBE0+D3/6-31G(d)^{30–33} level of theory implemented in Gaussian 16.⁵³ All unique structures optimized at the PBE0+D3/6-31G(d) level of theory below 21 kJ mol^{-1} , relative to the lowest-energy structure of one structural type, were reoptimized, and harmonic frequencies were computed at the PBE0+D3/6-311+G(d,p) level theory in Gaussian 16. The relative free energy at 90 K ($\Delta F_{90\text{K}}$ according to the temperature of the ion trap) from the harmonic frequency calculation was used to rank all final structures (Table S1 and Figures S4 and S5). All harmonic infrared spectra were scaled by an empirical scaling factor of 0.965. For the lowest-energy structure of each motif, single-point energy calculations at the DLPNO-CCSD(T)/Def2-TZVPP^{54–56} level of theory were performed in ORCA 5.0.3⁵⁴ (Table S2). The xyz coordinates of the reoptimized geometries can be found in the Supporting Information.

Transition states were located using relaxed scans of the reaction coordinate in Gaussian 16. The saddle points were optimized as transition states, and the harmonic frequencies were computed at the PBE0+D3/6-311+G(d,p) level of theory. The existence of one imaginary frequency corresponding to the reaction coordinate confirms the existence of the transition state. The transition states were linked to minima using intrinsic reaction coordinate calculations (Figure S3). Single-point energies of all optimized structures along the reaction trajectory were computed at the DLPNO-CCSD(T)/Def2-TZVPP level of theory using ORCA.

Reactions of Oxepine 8. Method A. (Ferrier reaction under anhydrous conditions) To a 10 mL round-bottom flask were added oxepine **8** (120.9 mg, 0.351 mmol) and benzyl alcohol (40.0 μL , 0.385 mmol, 1.1 equiv). The contents were dried by azeotropic distillation with toluene (3 \times 2 mL) and then dissolved in dry DCM (2 mL) under N_2 at rt. Then, bis(acetonitrile)dichloropalladium(II) (9.1 mg, 0.035 mmol, 0.1 equiv), dried in a vacuum desiccator prior to use, was added as a solid in one portion. After 5 h, the reaction was quenched with aq. NaHCO_3 (1 mL). The mixture was diluted with DCM (10 mL) and sequentially washed with saturated NaHCO_3 (1 \times 10 mL), water (1 \times 10 mL), and brine (1 \times 10 mL). The organic layer was dried with Na_2SO_4 and filtered, and the solvent was removed under reduced pressure. The residue was purified by column chromatography (40% EtOAc/hexane) to afford 12.6 mg (8.7%) of a yellow syrup which was a mixture of **15** and **16** (5:1).

Method B. (Ferrier reaction without rigorous exclusion of water) To a 10 mL round-bottom flask, **1** (24.0 mg, 0.0726 mmol) and benzyl alcohol (8.0 μL , 0.077 mmol, 1.1 equiv) were added. The catalyst bis(acetonitrile)dichloropalladium(II) was added as a solution in dichloromethane (0.40 mL, 4.42 mg mL^{-1} , 0.1 equiv). The solution was stirred open to the atmosphere for 5 h. Then, saturated aq. NaHCO_3 (1 mL) was added to the reaction (quench), and this mixture was diluted with dichloromethane (10 mL). The solution was next washed with saturated NaHCO_3 (1 \times 10 mL), water (1 \times 10 mL), and brine (1 \times 10 mL). The organic layer was dried with Na_2SO_4 and filtered, and the solvent was removed under reduced

pressure. The condensed crude was purified by column chromatography (40% EtOAc/hexane) to afford **16** (13.6 mg, R/S 1:2, 53% BRSM) as a yellow syrup in a mixture of C3 diastereomers.

Benzyl 4,5,7-tri-O-acetyl-2,3-dideoxy-β-D-arabino-2-enoseptanoside (15). Synthesized using **Method A**. R_f 0.61 (40% EtOAc/Hex). Spectroscopic data for compound **15**: ^1H NMR (400 MHz, CDCl_3): δ ppm 7.41–7.24 (m, 5H), 5.88 (ddd, $J = 12.1, 5.9, 1.5$ Hz, 1H), 5.73 (dd, $J = 12.1, 2.2$ Hz, 1H), 5.38 (dd, $J = 1.9, 1.9$ Hz, 1H), 5.25 (dd, $J = 5.9, 4.2$ Hz, 1H), 5.21–5.04 (m, 1H), 4.87 (d, $J = 12.0$ Hz, 1H), 4.59 (d, $J = 11.9$ Hz, 1H), 4.35 (ddd, $J = 9.4, 5.8, 2.9$ Hz, 1H), 4.21 (dd, $J = 11.9, 5.7$ Hz, 1H), 4.15 (dd, $J = 12.1, 2.9$ Hz, 1H), 2.06 (s, 3H), 2.05 (s, 3H), 2.01 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, CDCl_3): δ ppm 170.9, 169.9, 169.4, 137.5, 132.3, 128.7, 128.7, 127.8, 127.1, 126.8, 98.2, 73.2, 70.9, 69.8, 69.0, 65.5, 64.4, 21.0, 20.94, 20.88; TOF HRMS (ESI) m/z calcd for $\text{C}_{20}\text{H}_{25}\text{O}_8$ [$\text{M} + \text{H}$] $^+$ 393.1549; found, 393.1519.

2-(2,3,5-tri-O-acetyl-D-arabino-pentofuranosyl)-1,1-dibenzylxyethane (16). Synthesized using **Method B**. R_f 0.61 (40% EtOAc/Hex); major C3 isomer: ^1H NMR (400 MHz, CDCl_3): δ ppm 7.41–7.25 (m, 10H), 5.15 (dd, $J = 3.5, 1.0$ Hz, 1H), 4.95–4.90 (m, 2H), 4.68 (dd, $J = 11.6, 9.1$ Hz, 2H), 4.57 (dd, $J = 11.7, 4.5$ Hz, 2H), 4.33 (dd, $J = 11.6, 4.9$ Hz, 1H), 4.20–4.11 (m, 2H), 3.91 (ddd, $J = 6.4, 4.8, 3.4$ Hz, 1H), 2.08 (s, 3H), 2.08 (s, 3H), 2.07–2.05 (m, 2H), 2.10–2.00 (m, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, CDCl_3): δ 170.9, 169.8, 169.8, 138.1, 128.6, 128.0, 127.97, 127.94, 100.0, 81.4, 79.1, 77.6, 77.3, 68.3, 67.8, 63.9, 33.0, 20.97, 20.96, 20.8; minor C3 isomer: ^1H NMR (400 MHz, CDCl_3): δ 7.41–7.25 (m, 10H), 5.09 (m, 2H), 4.98 (dd, $J = 7.4, 4.1$ Hz, 1H), 4.68 (dd, $J = 11.6, 9.1$ Hz, 2H), 4.57 (dd, $J = 11.7, 4.5$ Hz, 2H), 4.29–4.25 (m, 2H), 4.20–4.11 (m, 2H), 2.08 (s, 3H), 2.08 (s, 3H), 2.07–2.05 (m, 2H), 2.06 (s, 3H). $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, CDCl_3): δ ppm 170.9, 169.8, 169.8, 138.1, 128.6, 127.97, 127.94, 127.9, 99.9, 81.0, 80.7, 80.1, 78.9, 68.2, 67.7, 63.6, 36.9, 20.97, 20.95, 20.8; TOF HRMS (ESI) m/z calcd for $\text{C}_{27}\text{H}_{32}\text{NaO}_9$ [$\text{M} + \text{Na}$] $^+$ 523.1944; found, 523.1927.

2-(2,3,5-tri-O-acetyl-D-arabino-pentofuranosyl)-acetaldehyde (17). Synthesized using **Method A** using 69.3 mg (0.201 mmol) of **8**. Isolated as a colorless syrup (17.2 mg, 12% BRSM). R_f 0.24 (40% EtOAc/Hex); major isomer (S): ^1H NMR (400 MHz, CDCl_3): δ 9.79 (dd, $J = 1.8, 1.0$ Hz, 1H), 5.29 (dd, $J = 3.8, 1.2$ Hz, 1H), 4.99 (dd, $J = 3.5, 1.2$ Hz, 1H), 4.60–4.53 (m, 1H), 4.36 (dd, $J = 11.6, 4.7$ Hz, 1H), 4.17 (dd, $J = 11.6, 6.4$ Hz, 1H), 4.02 (ddd, $J = 6.5, 4.7, 3.6$ Hz, 1H), 2.89–2.81 (m, 1H), 2.78 (dd, $J = 6.8, 1.9$ Hz, 1H), 2.11 (s, 3H), 2.10 (s, 3H), 2.09 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, CDCl_3): δ 198.8, 170.8, 169.8, 169.7, 81.6, 78.8, 77.5, 75.5, 63.7, 42.9, 20.94, 20.90, 20.8. Minor isomer (R): ^1H NMR (400 MHz, CDCl_3): δ 9.77 (dd, $J = 2.0, 1.5$ Hz, 1H), 5.15 (dd, $J = 2.7, 2.7$ Hz, 1H), 5.05 (dd, $J = 3.9, 2.4$ Hz, 1H), 4.60–4.53 (m, 1H), 4.29–4.21 (m, 3H), 2.74 (dd, $J = 5.8, 1.0$ Hz, 1H), 2.70 (dd, $J = 5.7, 1.0$ Hz, 1H), 2.12 (s, 3H), 2.11 (s, 3H), 2.10 (s, 3H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, CDCl_3): δ 199.8, 170.8, 169.8, 169.6, 81.2, 80.8, 78.5, 78.3, 63.2, 46.6, 20.94, 20.90, 20.8; TOF HRMS (ESI) m/z calcd for $\text{C}_{13}\text{H}_{19}\text{O}_8$ [$\text{M} + \text{H}$] $^+$ 303.1080; found, 303.1083.

3-(4,5,7-tri-O-acetyl-2,3-dideoxy-α-D-arabino-2-enoseptanosyl)-1-propene (18). To a 10 mL round-bottom flask was added oxepine **8** (23.9 mg, 0.0694 mmol), and the sample was dried by azeotropic distillation from toluene (3 × 2 mL). The contents were then dissolved in dry DCM (2 mL) and cooled to –45 °C. Allyltrimethylsilane (12.0 μL, 0.0755 mmol) was added to the solution, followed by the slow addition of $\text{BF}_3 \cdot \text{OEt}_2$ (8.6 μL, 0.067 mmol). The mixture was immediately switched to a bath at –20 °C and stirred at that temperature for 1 h. Then, the reaction was quenched with aq. NaHCO_3 (1 mL). The mixture was then diluted with DCM (10 mL) which was sequentially washed with saturated NaHCO_3 (1 × 10 mL), water (1 × 10 mL), and brine (1 × 10 mL). The organic layer was dried with Na_2SO_4 and filtered, and the solvent was removed under reduced pressure to give compound **18** (17.5 mg, 77%) as a colorless oil. R_f 0.63 (40% EtOAc/hexanes); ^1H NMR (400 MHz, acetone- d_6): δ ppm 5.87 (dddd, $J = 17.2, 13.9, 10.2, 6.9$ Hz, 1H), 5.78 (ddd, $J = 11.7, 9.2, 2.6$ Hz, 1H), 5.73 (ddd, $J = 11.8, 2.2, 2.2$ Hz, 1H), 5.51 (ddd, $J = 11.8, 2.5, 2.5$ Hz, 1H), 5.12 (dddd, $J = 17.2,$

2.0, 1.5, 1.5 Hz, 1H), 5.10–5.02 (m, 2H), 4.49 (dddd, $J = 8.0, 4.9, 2.4, 2.4$ Hz, 1H), 4.42 (ddd, $J = 8.8, 8.8, 1.9$ Hz, 1H), 4.11–4.03 (m, 2H), 2.32 (dddd, $J = 12.0, 5.6, 2.6, 1.2$ Hz, 2H), 2.03 (s, 3H), 2.02 (s, 6H); $^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz): δ ppm 170.7, 170.1, 170.0, 135.6, 134.8, 127.9, 117.5, 75.3, 73.2, 72.4, 71.3, 62.8, 40.2, 20.8, 20.7, 20.7; TOF HRMS (ESI) m/z calcd for $\text{C}_{16}\text{H}_{26}\text{NO}_7$ [$\text{M} + \text{NH}_4$] $^+$ 344.1709; found, 344.1695.

■ ASSOCIATED CONTENT

Data Availability Statement

The data underlying this study are available in the published article and its [Supporting Information](#).

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.joc.3c00079>.

Mass spectra, computed energetics, 3D structures, and NMR spectra (PDF)
xyz coordinates (XYZ)

■ AUTHOR INFORMATION

Corresponding Authors

Kevin Pagel – Fritz Haber Institute of the Max Planck Society, 14195 Berlin, Germany; Institute of Chemistry and Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany; orcid.org/0000-0001-8054-4718; Email: kevin.pagel@fu-berlin.de

Mark W. Peczuh – Department of Chemistry, University of Connecticut, Storrs, Connecticut 06269, United States; Email: mark.peczuh@uconn.edu

Authors

Kim Greis – Fritz Haber Institute of the Max Planck Society, 14195 Berlin, Germany; Institute of Chemistry and Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany; orcid.org/0000-0002-9107-2282

Caleb E. Griesbach – Department of Chemistry, University of Connecticut, Storrs, Connecticut 06269, United States

Carla Kirschbaum – Fritz Haber Institute of the Max Planck Society, 14195 Berlin, Germany; Institute of Chemistry and Biochemistry, Freie Universität Berlin, 14195 Berlin, Germany; orcid.org/0000-0003-3192-0785

Gerard Meijer – Fritz Haber Institute of the Max Planck Society, 14195 Berlin, Germany; orcid.org/0000-0001-9669-8340

Gert von Helden – Fritz Haber Institute of the Max Planck Society, 14195 Berlin, Germany; orcid.org/0000-0001-7611-8740

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.joc.3c00079>

Author Contributions

† K.G. and C.E.G. contributed equally.

Funding

Open access funded by Max Planck Society.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Dr. Wieland Schöllkopf and Sandy Gewinner for operating the FHI FEL. K.G. is grateful to the Fonds National de la Recherche (FNR), Luxembourg, for funding the project GlycoCat (13549747). C.K. thanks the

Fonds der Chemischen Industrie for financial support. K.P. acknowledges generous funding by the European Research Council, ERC-2019-CoG-863934-GlycoSpec. This work was partially supported by NSF CHE-1506567 to M.W.P.

REFERENCES

- (1) Bennett, C. S.; Galan, M. C. Methods for 2-Deoxyglycoside Synthesis. *Chem. Rev.* **2018**, *118*, 7931–7985.
- (2) Sau, A.; Galan, M. C. Palladium-Catalyzed α -Stereoselective O-Glycosylation of O(3)-Acylated Glycols. *Org. Lett.* **2017**, *19*, 2857–2860.
- (3) Sau, A.; Palo-Nieto, C.; Galan, M. C. Substrate-Controlled Direct α -Stereoselective Synthesis of Deoxyglycosides from Glycols Using $B(C_6F_5)_3$ as Catalyst. *J. Org. Chem.* **2019**, *84*, 2415–2424.
- (4) Sau, A.; Williams, R.; Palo-Nieto, C.; Franconetti, A.; Medina, S.; Galan, M. C. Palladium-Catalyzed Direct Stereoselective Synthesis of Deoxyglycosides from Glycols. *Angew. Chem., Int. Ed.* **2017**, *56*, 3640–3644.
- (5) Halcomb, R. L.; Danishefsky, S. J. On the Direct Epoxidation of Glycols: Application of a Reiterative Strategy for the Synthesis of β -Linked Oligosaccharides. *J. Am. Chem. Soc.* **1989**, *111*, 6661–6666.
- (6) Nomenclature designating the configuration of anomeric centers (α/β) follows IUPAC rules. See McNaught, A. D. Nomenclature of Carbohydrates (IUPAC Recommendations 1996). *Pure Appl. Chem.* **1996**, *68*, 1919–2008.
- (7) Markad, S. D.; Xia, S.; Snyder, N. L.; Surana, B.; Morton, M. D.; Hadad, C. M.; Peczu, M. W. Stereoselectivity in the Epoxidation of Carbohydrate-Based Oxepines. *J. Org. Chem.* **2008**, *73*, 6341–6354.
- (8) Ferrier, R. J.; Prasad, N. Unsaturated Carbohydrates. Part IX. Synthesis of 2,3-dideoxy- α -D-erythro-hex-2-enopyranosides from tri-O-acetyl-D-glucal. *J. Chem. Soc. C* **1969**, 570–575.
- (9) Ayala, L.; Lucero, C. G.; Romero, J. A. C.; Tabacco, S. A.; Woerpel, K. A.; Woerpel, K. A. Stereochemistry of Nucleophilic Substitution Reactions Depending upon Substituent: Evidence for Electrostatic Stabilization of Pseudoaxial Conformers of Oxocarbenium Ions by Heteroatom Substituents. *J. Am. Chem. Soc.* **2003**, *125*, 15521–15528.
- (10) Beaver, M. G.; Buscagan, T. M.; Lavinda, O.; Woerpel, K. A. Stereoelectronic Model To Explain Highly Stereoselective Reactions of Seven-Membered-Ring Oxocarbenium-Ion Intermediates. *Angew. Chem., Int. Ed.* **2016**, *55*, 1816–1819.
- (11) Luu, K. B.; Woerpel, K. A. Involvement of an Oxonium Ion Intermediate in Controlling the Diastereoselectivity of Nucleophilic Substitution Reactions of Septanoses. *Org. Lett.* **2023**, *25*, 152–157.
- (12) Vannam, R.; Peczu, M. W. A Practical and Scalable Synthesis of Carbohydrate Based Oxepines. *Org. Biomol. Chem.* **2016**, *14*, 3989–3996.
- (13) Griesbach, C. E.; Peczu, M. W. Characterization of the Low Energy Conformations and Differential Reactivities of D-Glucose and D-Mannose Based Oxepines. *Org. Biomol. Chem.* **2021**, *19*, 10635–10646.
- (14) Demkiw, K. M.; Hu, C. T.; Woerpel, K. A. Hyperconjugative Interactions of the Carbon–Halogen Bond That Influence the Geometry of Cyclic α -Haloacetals. *J. Org. Chem.* **2022**, *87*, 5315–5327.
- (15) Andreana, P. R.; Crich, D. Guidelines for O-Glycoside Formation from First Principles. *ACS Cent. Sci.* **2021**, *7*, 1454–1462.
- (16) Adero, P. O.; Amarasekara, H.; Wen, P.; Bohé, L.; Crich, D. The Experimental Evidence in Support of Glycosylation Mechanisms at the S_N1 – S_N2 Interface. *Chem. Rev.* **2018**, *118*, 8242–8284.
- (17) Gómez, A. M.; Lobo, F.; Uriel, C.; López, J. C. Recent Developments in the Ferrier Rearrangement. *Eur. J. Org. Chem.* **2013**, *2013*, 7221–7262.
- (18) Greis, K.; Kirschbaum, C.; Lechnitz, S.; Gewinner, S.; Schöllkopf, W.; von Helden, G.; Meijer, G.; Seeberger, P. H.; Pagel, K. Direct Experimental Characterization of the Ferrier Glycosyl Cation in the Gas Phase. *Org. Lett.* **2020**, *22*, 8916–8919.
- (19) Bhumra, N.; Lebedel, L.; Yamashita, H.; Shimizu, Y.; Abada, Z.; Ardá, A.; Désiré, J.; Michelet, B.; Martin-Mingot, A.; Abou-Hassan, A.; Takumi, M.; Marrot, J.; Jiménez-Barbero, J.; Nagaki, A.; Blériot, Y.; Thibaudeau, S. Insight into the Ferrier Rearrangement by Combining Flash Chemistry and Superacids. *Angew. Chem., Int. Ed.* **2021**, *60*, 2036–2041.
- (20) Demchenko, A. V.; Rousson, E.; Boons, G. Stereoselective 1,2-Cis-Galactosylation Assisted by Remote Neighboring Group Participation and Solvent Effects. *Tetrahedron Lett.* **1999**, *40*, 6523–6526.
- (21) Vannam, R.; Pote, A. R.; Peczu, M. W. Formation and Selective Rupture of 1,4-Anhydroseptanoses. *Tetrahedron* **2017**, *73*, 418–425.
- (22) Castro, S.; Peczu, M. W. Sequential Cyclization–Elimination Route to Carbohydrate-Based Oxepines. *J. Org. Chem.* **2005**, *70*, 3312–3315.
- (23) Redlich, H. Ring Contractions in Carbohydrates. *Angew. Chem., Int. Ed. Engl.* **1994**, *33*, 1345–1347.
- (24) Pakulski, Z. Seven Membered Ring Sugars. *Pol. J. Chem.* **1996**, *70*, 667–707.
- (25) Pakulski, Z.; Zamojski, A. 6-Deoxy-Heptoses. *Pol. J. Chem.* **1995**, *69*, 509–528.
- (26) Mucha, E.; Marianski, M.; Xu, F. F.; Thomas, D. A.; Meijer, G.; von Helden, G.; Seeberger, P. H.; Pagel, K. Unravelling the Structure of Glycosyl Cations via Cold-Ion Infrared Spectroscopy. *Nat. Commun.* **2018**, *9*, 4174–4175.
- (27) Marianski, M.; Mucha, E.; Greis, K.; Moon, S.; Pardo, A.; Kirschbaum, C.; Thomas, D. A.; Meijer, G.; Helden, G.; Gilmore, K.; Seeberger, P. H.; Pagel, K. Remote Participation during Glycosylation Reactions of Galactose Building Blocks: Direct Evidence from Cryogenic Vibrational Spectroscopy. *Angew. Chem., Int. Ed.* **2020**, *59*, 6166–6171.
- (28) Greis, K.; Kirschbaum, C.; Fittolani, G.; Mucha, E.; Chang, R.; Helden, G.; Meijer, G.; Delbianco, M.; Seeberger, P. H.; Pagel, K. Neighboring Group Participation of Benzoyl Protecting Groups in C3- and C6-Fluorinated Glucose. *Eur. J. Org. Chem.* **2022**, *2022*, No. e202200255.
- (29) Greis, K.; Lechnitz, S.; Kirschbaum, C.; Chang, C.-W.; Lin, M.-H.; Meijer, G.; von Helden, G.; Seeberger, P. H.; Pagel, K. The Influence of the Electron Density in Acyl Protecting Groups on the Selectivity of Galactose Formation. *J. Am. Chem. Soc.* **2022**, *144*, 20258–20266.
- (30) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (31) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- (32) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H–Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (33) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261.
- (34) Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. Natural Triple Excitations in Local Coupled Cluster Calculations with Pair Natural Orbitals. *J. Chem. Phys.* **2013**, *139*, 134101.
- (35) Purvis, G. D.; Bartlett, R. J. A Full Coupled-Cluster Singles and Doubles Model: The Inclusion of Disconnected Triples. *J. Chem. Phys.* **1982**, *76*, 1910–1918.
- (36) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- (37) Hansen, T.; Elferink, H.; van Hengst, J. M. A.; Houthuijs, K. J.; Remmerswaal, W. A.; Kromm, A.; Berden, G.; van der Vorm, S.; Rijs, A. M.; Overkleef, H. S.; Filippov, D. V.; Rutjes, F. P. J. T.; van der Marel, G. A.; Martens, J.; Oomens, J.; Codée, J. D. C.; Boltje, T. J.

Characterization of Glycosyl Dioxolenium Ions and Their Role in Glycosylation Reactions. *Nat. Commun.* **2020**, *11*, 2664–2669.

(38) Childs, R. F.; Kang, G. J.; Wark, T. A.; Frampton, C. S. Structural Studies on Dioxan-2-Ylium Ions; Intramolecular Attack of an Oxygen Atom on a Carbenium Ion Centre. *Can. J. Chem.* **1994**, *72*, 2084–2093.

(39) Inaba, K.; Matsumura, S.; Yoshikawa, S. Novel C-3 Epimerization of Glycals by Metal Chlorides. *Chem. Lett.* **1991**, *20*, 485–488.

(40) Di Salvo, A.; David, M.; Crousse, B.; Bonnet-Delpon, D. Self-Promoted Nucleophilic Addition of Hexafluoro-2-Propanol to Vinyl Ethers. *Adv. Synth. Catal.* **2006**, *348*, 118–124.

(41) The experiment followed the anhydrous method using 69.3 mg (0.201 mmol) of **8**.

(42) NMR spectra shown in this figure had their noise reduced, and the NOE irradiated signal for H1 (4.5 ppm) was suppressed for clarity. See the Supporting Information for raw spectra.

(43) Pal, K. B.; Lee, J.; Das, M.; Liu, X. W. Palladium(II)-Catalyzed Stereoselective Synthesis of: C-Glycosides from Glycals with Diaryliodonium Salts. *Org. Biomol. Chem.* **2020**, *18*, 2242–2251.

(44) González Flórez, A. I.; Mucha, E.; Ahn, D.-S.; Gewinner, S.; Schöllkopf, W.; Pagel, K.; von Helden, G. Charge-Induced Unzipping of Isolated Proteins to a Defined Secondary Structure. *Angew. Chem., Int. Ed.* **2016**, *55*, 3295–3299.

(45) Thomas, D. A.; Chang, R.; Mucha, E.; Lettow, M.; Greis, K.; Gewinner, S.; Schöllkopf, W.; Meijer, G.; Von Helden, G. Probing the Conformational Landscape and Thermochemistry of DNA Dinucleotide Anions: Via Helium Nanodroplet Infrared Action Spectroscopy. *Phys. Chem. Chem. Phys.* **2020**, *22*, 18400–18413.

(46) Greis, K.; Kirschbaum, C.; Taccone, M. I.; Götz, M.; Gewinner, S.; Schöllkopf, W.; Meijer, G.; von Helden, G.; Pagel, K. Studying the Key Intermediate of RNA Autohydrolysis by Cryogenic Gas-Phase Infrared Spectroscopy. *Angew. Chem., Int. Ed.* **2022**, *61*, No. e202115481.

(47) Schöllkopf, W.; Gewinner, S.; Junkes, H.; Paarmann, A.; von Helden, G.; Bluem, H.; Todd, A. M. M. The New IR and THz FEL Facility at the Fritz Haber Institute in Berlin. In *Advances in X-ray Free-Electron Lasers Instrumentation III*; Biedron, S. G., Ed.; SPIE, 2015; Vol. 9512, p 95121L. DOI: DOI: 10.1117/12.2182284.

(48) Pracht, P.; Bohle, F.; Grimme, S. Automated Exploration of the Low-Energy Chemical Space with Fast Quantum Chemical Methods. *Phys. Chem. Chem. Phys.* **2020**, *22*, 7169–7192.

(49) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-XTB - An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.

(50) Spicher, S.; Grimme, S. Robust Atomistic Modeling of Materials, Organometallic, and Biochemical Systems. *Angew. Chem., Int. Ed.* **2020**, *132*, 15795–15803.

(51) *Schrödinger Release 2021-3: MacroModel*, Schrödinger, LLC, New York, NY, 2021.

(52) Bochevarov, A. D.; Harder, E.; Hughes, T. F.; Greenwood, J. R.; Braden, D. A.; Philipp, D. M.; Rinaldo, D.; Halls, M. D.; Zhang, J.; Friesner, R. A. Jaguar: A High-Performance Quantum Chemistry Software Program with Strengths in Life and Materials Sciences. *Int. J. Quantum Chem.* **2013**, *113*, 2110–2142.

(53) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.

Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16, Revision A.03*; Gaussian, Inc.: Wallingford CT, 2016.

(54) Neese, F. Software Update: The ORCA Program System—Version 5.0. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2022**, *12*, 1–15.

Recommended by ACS

3-Substituted-1-(Trimethylsilylmethyl)cyclobutyl Cations: Stereochemistry of Solvent Capture of β -Trimethylsilyl Carbocations

Xavier Creary.

JANUARY 26, 2023
THE JOURNAL OF ORGANIC CHEMISTRY

READ 

Synthesis of the Tetrasaccharide Glycone Part of Tetrocarcin A

Jihun Kang and Young Ho Rhee

FEBRUARY 22, 2023
THE JOURNAL OF ORGANIC CHEMISTRY

READ 

Experimental and Theoretical Investigation of the Synchronicity of Ambident Silyloxypyrone-Based (5 + 2) Cycloadditions

Adam J. Youman, T. Andrew Mitchell, *et al.*

APRIL 14, 2023
THE JOURNAL OF ORGANIC CHEMISTRY

READ 

Taming Keteniminium Reactivity by Steering Reaction Pathways: Computational Predictions and Experimental Validations

Mark A. Maskeri, K. N. Houk, *et al.*

DECEMBER 16, 2022
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

READ 

Get More Suggestions >

8 | The Key Intermediate of RNA Autohydrolysis

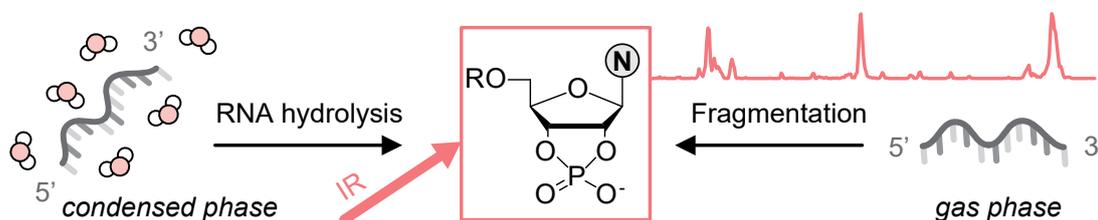
This work is published^[268] and available online:

DOI: <https://doi.org/10.1002/anie.202115481>

Details on the related Supporting Information can be found in Appendix E.

K. Greis,[†] C. Kirschbaum,[†] M. I. Taccone, M. Götze, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, K. Pagel*; Studying the Key Intermediate of RNA Autohydrolysis by Cryogenic Gas-Phase Infrared Spectroscopy. *Angew. Chem. Int. Ed.* **2022**, *61*, e202115481.

[†]equal contributions



Submitted: 14 November 2021; **Published:** 1 March 2022

Copyright © 2022 The Authors. Published by Wiley-VCH GmbH

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

8.1 | Author Contributions

The project was conceived by **Kim Greis**, Carla Kirschbaum, and Kevin Pagel. **Kim Greis** (main) and Carla Kirschbaum (support) wrote the first draft of the manuscript. **Kim Greis**, Carla Kirschbaum (both main), and Martín Taccone (support) performed the mass spectrometry and cryogenic infrared spectroscopy experiments. The cryogenic infrared spectroscopy experiments were supervised by Gerard Meijer and Gert von Helden. Sandy Gewinner and Wieland Schöllkopf operated the free-electron laser. **Kim Greis** (main) and Carla Kirschbaum (support) performed the quantum chemical calculations. **Kim Greis**, Carla Kirschbaum (both main), and Michael Götze (support) analyzed the data. Carla Kirschbaum designed the frontispiece artwork.

8.2 | Project Hypothesis and Summary

The idea of the following project was to study the structure of the intermediate of RNA autohydrolysis. In this reaction, the phosphate backbone of RNA is attacked by the 2'-OH group of ribose, leading to the formation of an intermediate with a five-membered cyclic phosphate and to the cleavage of the 3'-RNA strand. Hydrolysis leads to opening of the phosphate ring and the formation of nucleoside phosphates (see Figure 1 of the related published work).^[316] Due to this reaction, RNA is intrinsically unstable and decomposes even without the presence of degrading enzymes, acids, or bases. In the presence of the latter the decomposition rate is accelerated up to a million times.^[317] DNA, on the other hand, does not feature a hydroxyl group at the 2'-position and can remain stable for thousands of years.^[318]

Similarly, the fragment ions obtained upon tandem mass spectrometry (MS/MS) of RNA and DNA differ. A prominent fragment of RNA oligonucleotides are *c*-fragments, where cleavage occurs between the phosphorus and 5'-oxygen atom (see Figure 2a of the related published work).^[319] This fragment cannot be observed in MS/MS for analogous DNA nucleotides.^[320] Interestingly, in negative ion mode, the *m/z* of *c*-ions is identical to that of the intermediate of base-catalyzed RNA autohydrolysis. In this work, cryogenic IR spectroscopy was used in combination with DFT calculations to investigate the structure of these *c*-fragments.

Upon activation, RNA dinucleotides provide a variety of fragments, among them *c*-ions. Previously, it was postulated that the structure of this ion might correspond to the intermediate of RNA autohydrolysis^[321-323] or it might adopt a linear structure.^[150,324] In the case of dinucleotides, a 3',5'-cyclic phosphate would be a further candidate structure for this intermediate. The *c*-ions were generated for each nucleobase and probed with cryogenic IR spectroscopy (see Figure 2b of the related published work). DFT calculations show that linear fragments are not minima on the potential energy surface of *c*-fragments, whereas five-membered 2',3'-cyclic phosphates are stabilized by 5–27 kJ mol⁻¹ compared to six-membered 3',5'-cyclic phosphates. Although both the harmonic and the anharmonic frequencies for 2',3'-cyclic phosphates match

the experimental spectrum slightly better than those of 3',5'-cyclic phosphates, the assignment solely based on computed spectra is not fully convincing (see Figure 2d of the related published work). This assignment would mainly rely on the absence of vibrational bands in the experimental spectrum, which are predicted for 3',5'-cyclic phosphates, but cannot be observed experimentally. Therefore, an experimental approach based on standards was designed to differentiate between these two structural types, as 2',3'- and 3',5'-cyclic phosphates of adenosine are commercially available. They readily yield the same m/z as c -ions in negative ion mode. Based on their experimental IR spectra, the structure of c -ions can be unambiguously assigned to 2',3'-cyclic phosphates (see Figure 2e of the related published work). Hence, autohydrolysis of dinucleotides can be induced in the gas phase, leading to the same type of intermediates as observed in the condensed phase.

As the relevant absorption bands of the phosphate moiety are similar in all probed c -fragments, it can be assumed that the c -ions of other nucleobases also form 2',3'-cyclic phosphates. Only in the case of the c -ion containing cytosine, the spectrum is considerably more complex. Therefore, this fragment ion spectrum was compared to those generated by commercially available standards of cyclic cytidine phosphates. The spectra of the standards confirmed that the six-membered 3',5'-cyclic phosphate is not present in the c -fragment ion spectrum (see Figure 3a of the related published work). However, the spectra of the standards are significantly less complex than that of the fragment ion, especially in the functional group region, where the vibrations of the nucleobase can be observed. Initially, the standards were ionized using soft source conditions, whereas activating source conditions are necessary to generate the fragments. Therefore, in a subsequent experiment, the standards were also ionized using activating source conditions (see Figure 3b of the related published work). This experiment showed that activating ionization conditions lead to a more complex spectrum for the standards as well. This change in spectrum can only be explained by an isomerization reaction. The nucleobase cytosine is known to undergo a tautomerization reaction from the amino-oxo to the imino-oxo tautomer (see Figure 3c of the related published work), which potentially happens under physiological conditions and leads to mutations.^[325]

DFT calculations confirm that for c -ions containing cytosine, the tautomer is only destabilized by a mere 3 kJ mol^{-1} , indicating that a copopulation of both isomers in the ion trap is likely. The anharmonic frequencies of the tautomerized structures match with the additional absorption bands in the experimental spectrum, which do not match with those of the non-tautomerized structures in the functional group region (see Figure 3d of the related published work). Hence, the broad feature in the $1600\text{--}1800 \text{ cm}^{-1}$ region can be explained by partially overlapping vibrations originating from the tautomerized and the non-tautomerized cytosine moiety. To get further insight into the tautomerization process, the reaction barrier was computed and found to be 156 kJ mol^{-1} , which is in agreement with reaction barriers of processes previously detected on the same instrument under activating source conditions.^[295] The barrier

of this process, which essentially involves the shift of a proton from the amino group to the imine function in the cytosine heterocycle, could be lowered to 47 kJ mol^{-1} , if a water molecule acts as a proton shuttle. Although water molecules are likely present in the source region of the instrument, it is not clear if a water molecule catalyzes this reaction.

The experimental results show that the structure of *c*-fragments of RNA dinucleotides is identical to the 2',3'-cyclic phosphate intermediate of RNA autohydrolysis for all four RNA nucleobases. Furthermore, in combination with DFT calculations, it was elucidated that the cytosine moiety undergoes partial tautomerization under the conditions that are necessary to generate *c*-fragments. In cyclic standards, this process can be avoided, as these generate the same ions under soft ionization conditions. Generally, this work shows that the observed reactions are intrinsic to the investigated molecules and are not dependent on a specific environment, as they are both known from condensed-phase reactions but also occur in the vacuum of a mass spectrometer.



Analytical Methods Hot Paper

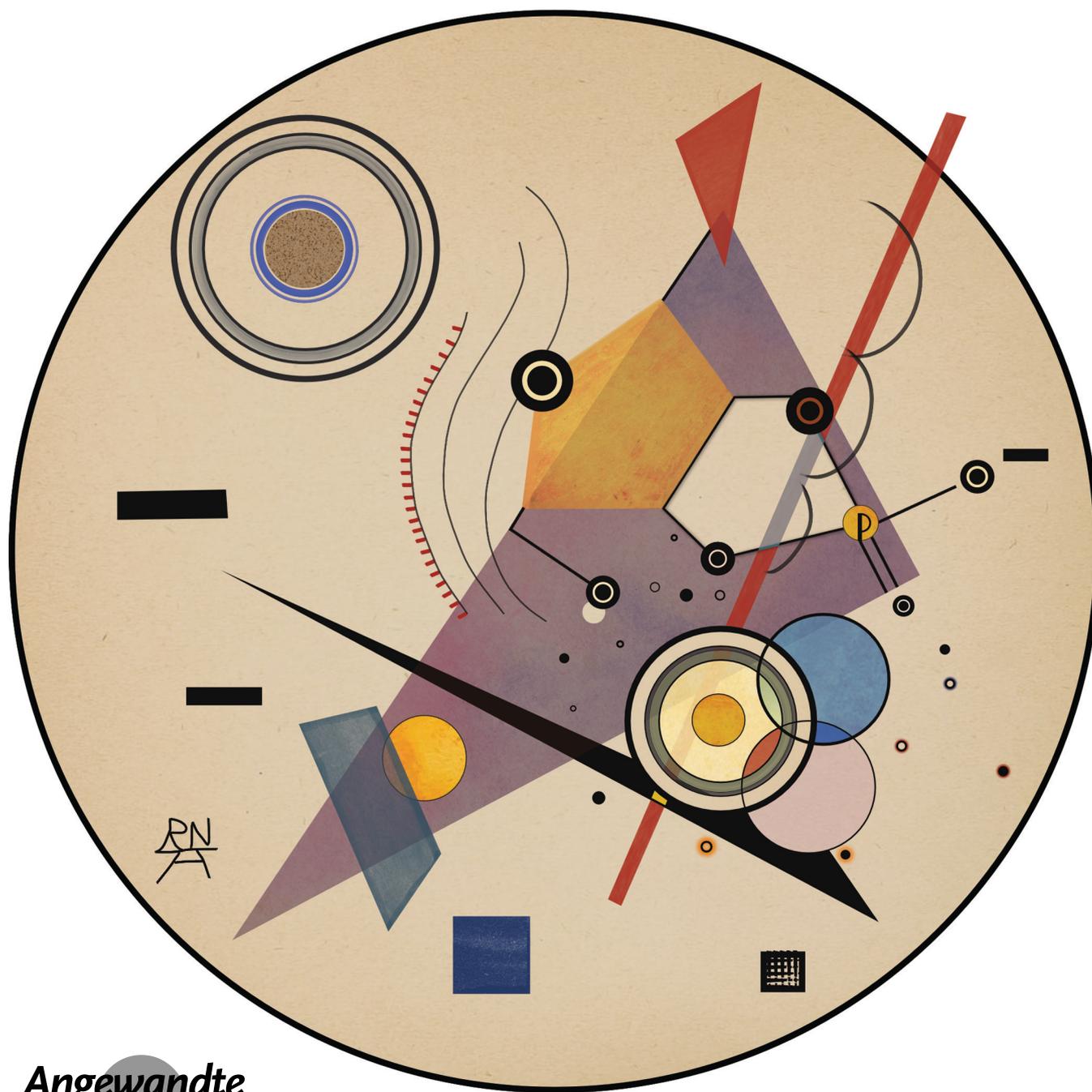
How to cite: *Angew. Chem. Int. Ed.* **2022**, *61*, e202115481

International Edition: doi.org/10.1002/anie.202115481

German Edition: doi.org/10.1002/ange.202115481

Studying the Key Intermediate of RNA Autohydrolysis by Cryogenic Gas-Phase Infrared Spectroscopy

Kim Greis⁺, Carla Kirschbaum⁺, Martín I. Taccone, Michael Götze, Sandy Gewinner, Wieland Schöllkopf, Gerard Meijer, Gert von Helden, and Kevin Pagel*

Angewandte
International Edition
Chemie

Abstract: Over the course of the COVID-19 pandemic, mRNA-based vaccines have gained tremendous importance. The development and analysis of modified RNA molecules benefit from advanced mass spectrometry and require sufficient understanding of fragmentation processes. Analogous to the degradation of RNA in solution by autohydrolysis, backbone cleavage of RNA strands was equally observed in the gas phase; however, the fragmentation mechanism remained elusive. In this work, autohydrolysis-like intermediates were generated from isolated RNA dinucleotides in the gas phase and investigated using cryogenic infrared spectroscopy in helium nanodroplets. Data from both experiment and density functional theory provide evidence for the formation of a five-membered cyclic phosphate intermediate and rule out linear or six-membered structures. Furthermore, the experiments show that another prominent condensed-phase reaction of RNA nucleotides can be induced in the gas phase: the tautomerization of cytosine. Both observed reactions are therefore highly universal and intrinsic properties of the investigated molecules.

The development of mRNA vaccines has recently experienced an unparalleled boost in the course of the worldwide COVID-19 pandemic.^[1] Their sudden breakthrough largely rests on the fact that, contrary to conventional vaccines, mRNA vaccines can be rapidly developed, produced on a large scale and adapted to different pathogens.^[2] The development of RNA-based pharmaceuticals requires advanced analytical methods including mass spectrometry (MS) and a profound understanding of fragmentation mechanisms for reliable identification of artificial RNA structures.^[3]

DNA is significantly more stable than RNA. This stability difference is mainly attributed to the different sugars incorporated in the backbone. Ribose in RNA contains a hydroxyl group at the C2' position, which destabilizes the phosphodiester bonds. This 2'-OH group can intramolecularly attack the phosphate group at C3' of the same nucleotide, leading to autohydrolysis of RNA even

in the absence of degrading enzymes. The deoxyribose in DNA on the other hand, does not contain this OH group and therefore does not undergo autohydrolysis. As a result, DNA fragments may remain stable for hundreds, sometimes even thousands of years.^[4]

The mechanism of in vitro RNA-autohydrolysis has been studied extensively.^[5] The reaction is initiated by a nucleophilic attack of the 2'-OH group on the 3'-phosphate, proceeding via a phosphorane to yield a cyclic 2',3'-phosphate—the key intermediate of RNA hydrolysis (Figure 1a). The intermediate is subsequently hydrolyzed in aqueous media to yield a mixture of 2'- and 3'-phosphates. RNA autohydrolysis is accelerated up to a million fold in basic or acidic media compared to spontaneous hydrolysis at neutral pH.^[6]

In accordance with its reactivity in solution, the 2'-OH group also induces a substantially different behavior of RNA in the gas phase. Tandem mass spectrometry (MS/MS) probes the fragmentation of ions in the gas phase and can be used to deduce structural information of nucleic acids.^[7] The major difference in fragmentation of DNA and RNA is the facile formation of *c*-ions from RNA precursors, which are not observed in DNA fragmentation.^[8] *c*-Ions arise from cleavage between the 5'-O and the phosphorous atom (Figure S4)^[9] and their formation relies on the 2'-OH group providing a mobile proton.^[10] *c*-Ions generated from RNA dinucleotides are of identical *m/z* as the key intermediate of base-catalyzed RNA autohydrolysis in solution. However, there is disagreement about the fragmentation mechanism and exact structure of *c*-ions. In the literature, mechanisms proceeding via a linear structure^[10] or direct formation of 2',3'-cyclic structures^[11] as in the condensed phase have been proposed. In the case of dinucleotides, isomeric 3',5'-cyclic nucleoside monophosphates are equally conceivable fragmentation structures (Figure 1b).

Gas-phase infrared (IR) ion spectroscopy is a powerful tool for the direct structural analysis of intact molecular ions or short-lived fragments. Infrared multiple photon dissociation (IRMPD) spectroscopy has been successfully used to study RNA building blocks,^[12] nucleoside triphosphate ions^[13] and 3',5'-cAMP.^[14] More recently, intact DNA dinucleotide anions were analysed by cryogenic IR spectroscopy in helium droplets, a technique offering an increased spectral resolution due to the suppression of ion heating.^[15] However, to date no tandem MS fragments of nucleotides have been spectroscopically studied with either method.

Here we investigate *c*-ions obtained by dissociation of deprotonated RNA dinucleotides ApA, GpG, UpG, and CpG (Figure 1c) using cryogenic IR spectroscopy in helium droplets. The utilized experimental setup has been described previously (see Supporting Information).^[16] Briefly, the ions of interest are generated by nano-electrospray ionization in negative ion mode and can be fragmented in the ion source by collisions with residual gas molecules in a process identical to collision-induced dissociation (CID). The ions are mass-to-charge selected in a quadrupole mass filter and accumulated in a hexapole ion trap (90 K). The thermalized, trapped ions are subsequently picked up by superfluid helium nanodroplets, generated by an Even-Lavie valve.^[17]

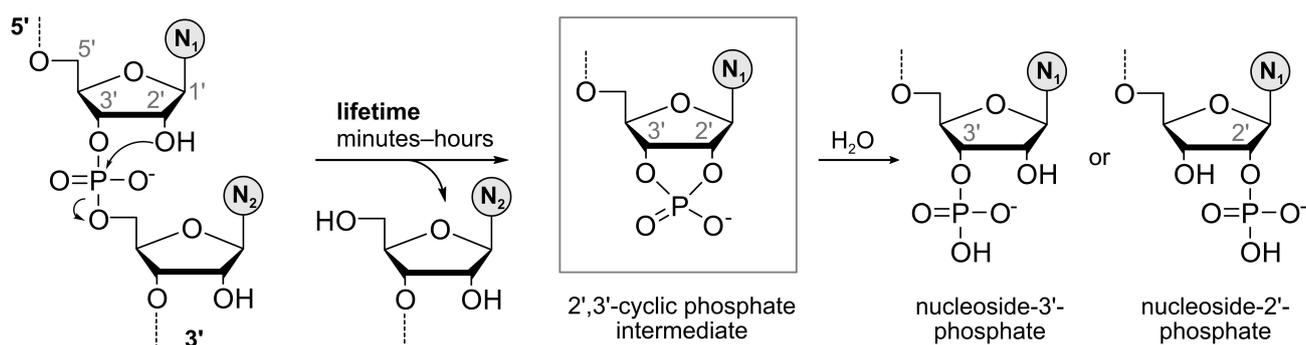
[*] K. Greis,[†] C. Kirschbaum,[†] Dr. M. Götze, Prof. K. Pagel
Institut für Chemie und Biochemie, Freie Universität Berlin
Arnimallee 22, 14195 Berlin (Germany)
E-mail: kevin.pagel@fu-berlin.de

K. Greis,[†] C. Kirschbaum,[†] Dr. M. I. Taccone, Dr. M. Götze,
S. Gewinner, Dr. W. Schöllkopf, Prof. G. Meijer, Prof. G. von Helden,
Prof. K. Pagel
Fritz-Haber-Institut der Max-Planck-Gesellschaft
Faradayweg 4–6, 14195 Berlin (Germany)

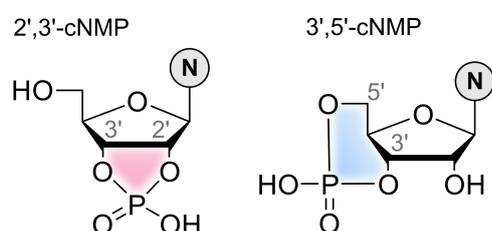
[†] These authors contributed equally to this work.

© 2022 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

a. Autohydrolysis of RNA



b. Cyclic Nucleoside Monophosphates



c. Nucleobases

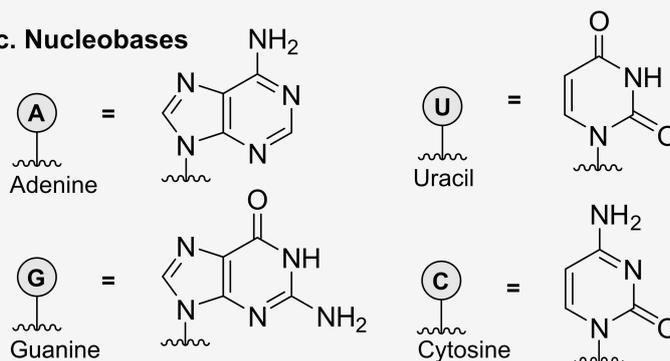


Figure 1. Autohydrolysis reduces the stability of RNA. a) Schematic autohydrolysis of an RNA strand with a proposed 2',3'-cyclic phosphate intermediate. An intramolecular nucleophilic attack of the 2' OH group is crucial to initiate cleavage of the phosphodiester. In DNA, the absence of a 2' OH group prevents cleavage. b) Structures of cyclic nucleoside monophosphate isomers. c) Nomenclature and structures of RNA nucleobases (N).

Inside the helium droplets, ions are rapidly cooled to 0.37 K and guided to a detection region where the helium droplet beam overlaps with an IR beam of the Fritz Haber Institute free-electron laser (FHI FEL).^[18] Photon absorption excites vibrational modes, leading to evaporation of the helium shell and subsequent release of the ions. The ion yield is measured in a time-of-flight mass spectrometer and plotted against the photon wavenumber to obtain an IR spectrum. The experimentally obtained IR spectra are compared to computed vibrational frequencies. For this, the conformational space of 2',3'- and 3',5'-cyclic nucleoside monophosphates was explored for each nucleobase using CREST^[19] with the semiempirical method GFN2-xTB.^[20] Unique conformations were optimized and harmonic frequencies computed at the PBE0+D3/def2-TZVPP^[21] level of theory in Gaussian 16^[22] and scaled by an empirical factor of 0.965. Anharmonic frequencies were calculated using the GVPT2^[23] method at the PBE0+D3/def2-TZVP level of theory. Previously reported linear *c*-fragments were ruled out from consideration because they are not minima on the potential energy surface of the ions at the employed level of theory, and their optimization leads to 2',3'-cyclic nucleoside monophosphates.

In the condensed phase, autohydrolysis can occur in the absence of enzymes or external reaction partners (Figure 1).

In the gas phase, dissociation of deprotonated RNA dinucleotides was induced by in-source fragmentation (Figure S2), leading to *c*-fragments with identical *m/z* as the intermediate of RNA autohydrolysis observed in solution. Once transferred to the high vacuum region of the instrument, no further ion activation occurs. The parent ion and its fragments were detected using time-of-flight MS (Figures 2a and S3). Subsequently, cryogenic IR spectra of the generated *c*-fragments were recorded for each nucleobase to determine their structure (Figure 2b). All IR signatures are similar in the fingerprint region (ca. 1000–1400 cm⁻¹) but exhibit significant differences in the functional group region (1400–1800 cm⁻¹). The vibrations in the fingerprint region are mostly originating from the ribose and phosphate moieties in the backbone. The two prominent bands at 1082 and 1326 cm⁻¹ can be assigned to symmetric and antisymmetric stretches within the phosphate group. The nucleobases mainly absorb in the functional group region. Furthermore, the IR signature of the *c*-fragment containing cytosine (C) is more prominent than its counterparts.

For dinucleotides, two different bicyclic structures are conceivable for *c*-fragments, as the phosphate backbone can be either attacked by the 2'- or the 5'-OH group leading to a five- or six-membered phosphodiester ring, respectively (Figure 2c). The experimental spectra were compared to

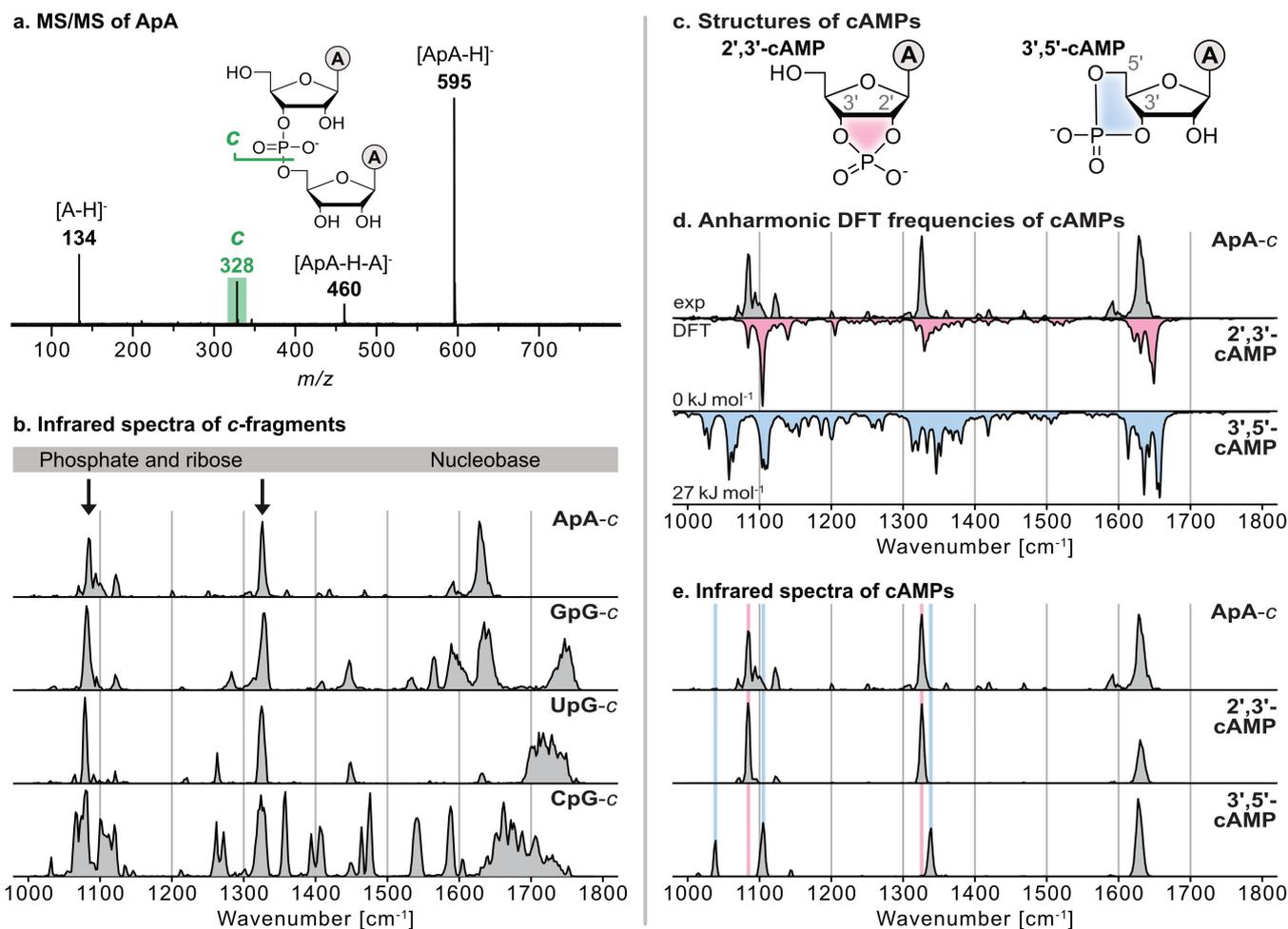


Figure 2. Fragmentation of RNA leads to 2',3'-cyclic phosphates. a) Tandem mass spectra of the ApA RNA dinucleotide in negative ion mode under in-source fragmentation conditions. The resulting *c*-fragments exhibit the same m/z -ratio as the cyclic phosphate intermediate occurring in RNA autohydrolysis. b) Cryogenic infrared spectra of the ApA, GpG, UpG and CpG *c*-fragments. Bands between 1000 and 1400 cm^{-1} are mostly originating from the ribose and phosphate moieties; vibrations between 1400 and 1800 cm^{-1} are originating from functional groups in the nucleobases. Bands highlighted with an arrow are independent of the nucleobase and present in each spectrum. c) Simplified structures of 2',3'- and 3',5'-cyclic adenosine monophosphates. d) The experimental spectrum of the *c*-fragment of deprotonated ApA (gray trace) is compared to the computed anharmonic vibrational spectra of the lowest-energy structures of 2',3'-cAMP (red) and 3',5'-cAMP (blue). Free energies of the computed structures at 90 K are indicated. e) The experimental spectrum of the ApA *c*-fragment compared to those of deprotonated 2',3'- and 3',5'-cAMP. Key features of the spectra are highlighted with red and blue lines. The absorption bands are clearly diagnostic to distinguish between penta- and hexacyclic phosphates.

computed energetics and frequencies of representative structures for both types of intermediates. First, from a thermodynamic point of view, the formation of a five-membered cyclic phosphate is favored over the formation of a six-membered phosphate by 5–27 kJ mol^{-1} in free energy ($\Delta F_{90\text{K}}$), depending on the nucleobase (Tables S1–S4). Moreover, the computed harmonic frequencies of the 2',3'-cNMPs agree well with the experimental spectra (Figure S6). The frequencies of the 3',5'-cNMPs on the other hand match less well, especially because of additional absorptions between 1020–1060 cm^{-1} in the computed spectrum, which cannot be observed in the experiment. In addition, it is noteworthy that some absorptions in the experimental spectrum are significantly broader than predicted by the harmonic frequency calculation. Anharmonic frequency calculations show that the band broadening observed for the *c*-fragments

of ApA, GpG and UpG predominantly arises from anharmonicities (Figures 2d and S8). For *c*-fragments of CpG, on the other hand, the spectrum is considerably more complex and cannot be explained by anharmonicities alone.

The above results suggest an intermediate containing a five-membered phosphate moiety. To further substantiate this finding, the IR signatures of commercially available 2',3'- and 3',5'-cyclic adenosine monophosphate anions were recorded and compared to the IR spectrum of the *c*-fragment of the ApA dinucleotide (Figures 2e and S10). The spectra confirm that the *c*-fragment corresponds to 2',3'-cAMP. Even though some weak absorption bands in the spectrum of the *c*-fragment of ApA are poorly resolved in case of 2',3'-cAMP due to day-to-day fluctuations in laser power, the main absorption bands coincide perfectly. The vibrations of 3',5'-cAMP, on the other hand, are significantly

shifted and contain an additional band at 1038 cm^{-1} , which was observed previously using IRMPD spectroscopy.^[14]

Due to the relatively straightforward assignment of *c*-fragments from ApA based on cyclic standards, we applied the same methodology to obtain further information on the crowded IR spectrum of CpG fragments. The spectra of 2',3'- and 3',5'-cyclic cytidine monophosphate anions generated under non-activating conditions suggest that the *c*-fragment of CpG is also a five-membered cyclic phosphate. The absorption bands of 2',3'-cCMP anions match the spectrum of the *c*-fragment of CpG (Figures 3a and S11). However, some features in the fragment ion spectrum are much more intense than for 2',3'-cCMP and the spectra generally do not match as well as in the case of adenosine. To elucidate the origin of the unusual complexity in the fragment ion spectrum, we subjected the cyclic cytidine anions to activating source conditions and recorded their infrared signatures (Figure 3b). Interestingly, the spectrum of activated 2',3'-cCMP corresponds much better to the *c*-fragment ion spectrum than its non-activated counterpart. Certain bands are significantly more pronounced and the broad feature at $1600\text{--}1800\text{ cm}^{-1}$ is now reproduced. For the 3',5'-cCMP anion, some features gained in intensity too, yet the signature does not match that of the fragment ion spectrum.

The most significant changes in the spectra at activating conditions occur in the functional group region (1400--

1800 cm^{-1}). This implies that the underlying process involves the nucleobase cytosine, rather than the backbone of the molecule. Cytosine is known to undergo tautomerization in the condensed phase, which leads to the transformation of the amino-oxo (C) into the imino-oxo (C*) form (Figure 3c). In vivo, this tautomerization leads to errors in base pairing (CG \rightarrow C*A) which in turn can cause mutations in the next replication cycle.^[24] For protonated nucleobases in the gas phase, tautomerization has been observed previously using IRMPD and differential mobility spectrometry.^[25] For the physiologically more relevant deprotonated ions, however, this process has not been studied to date.

Density functional theory (DFT) calculations reveal that the C* tautomer is destabilized by a mere 3 kJ mol^{-1} in case of deprotonated 2',3'- and 3',5'-cyclic cytidine monophosphate (Figure 3c). However, for the 2',3'-analog, a substantial activation energy of $+156\text{ kJ mol}^{-1}$ has been computed. This barrier can be lowered to $+47\text{ kJ mol}^{-1}$ when catalyzed by a water molecule, which explains the relatively low activation energy required to induce tautomerization in the presence of water in the source region of the instrument (see Figure S12 and Table S5). The harmonic and anharmonic frequencies of tautomerized cyclic cytidine monophosphate further complement those of the non-tautomerized form (Figures 3d, S13 and S14), which explains the wealth of spectral features observed for the *c*-fragment ions and the activated cCMPs.

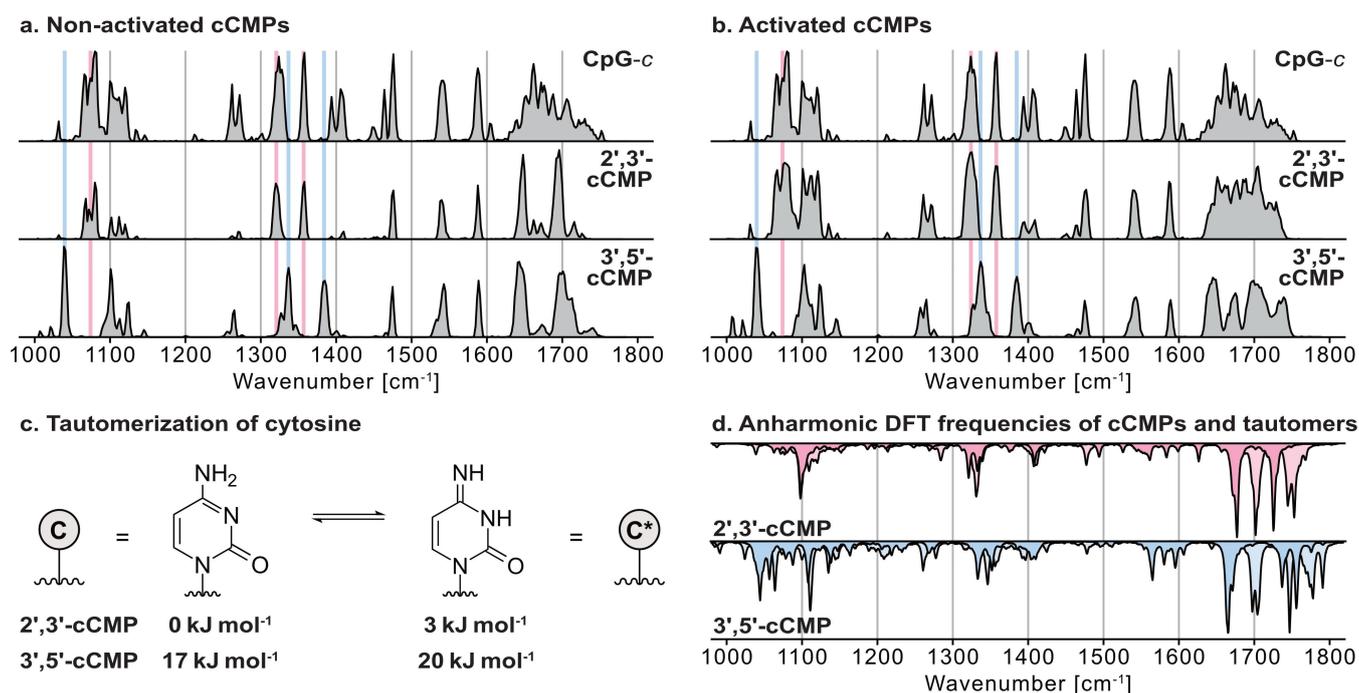


Figure 3. Tautomerization of cytosine in the gas phase. The experimental infrared spectrum of the CpG *c*-fragment, formed under activating source conditions compared to those of deprotonated 2',3'- and 3',5'-cCMP generated at a) non-activating and b) activating source conditions. Key absorption features of 2',3'-cCMP and 3',5'-cCMP spectra are highlighted in red and blue, respectively. The absorption bands are diagnostic and reveal that the fragment ion possesses a pentacyclic structure. Some absorption bands, especially above 1400 cm^{-1} , are more complex and can only be reproduced under activating conditions which induce c) the tautomerization of the cytosine moiety. According to calculations, the free energy (90 K) difference between both tautomers is only 3 kJ mol^{-1} for both the 2',3' and the 3',5' isomer. d) Anharmonic frequencies confirm that the coexistence of tautomerized and non-tautomerized cCMPs in the ion trap leads to significant broadening of the absorptions in the $1600\text{--}1800\text{ cm}^{-1}$ region.

In summary, we show here that *c*-fragments in RNA tandem mass spectrometry adopt a structure, which is identical to that of the reactive intermediate occurring during RNA autohydrolysis. Comparison between the infrared signatures of *c*-fragments and those of synthetic cyclic nucleotides in combination with DFT calculations unambiguously identified the characteristic 2',3'-cyclic structure of the intermediate. This implies that the intramolecular cyclisation reaction of RNA nucleotides is a strongly favored process with a high intrinsic driving force. The results support the previously proposed fragmentation mechanism involving a nucleophilic attack of the 2'-OH group at the phosphate rather than alternative mechanisms proceeding via linear intermediates. Further, our experiments show that tautomerization of cytosine can be followed in the gas phase as well. This tautomerization spontaneously occurs in the source region under activating conditions and leads to more complex infrared signatures. In a broader context, our results highlight that both processes, the intramolecular cyclisation of RNA nucleotides as well as the tautomerization of cytidine can occur in full isolation. This implies that both reactions are intrinsic to the individual molecules and independent of their environment.

Acknowledgements

K.G. thanks the Fonds National de la Recherche, Luxembourg, for funding the project GlycoCat (13549747). C.K. is grateful for financial support by the Fonds der Chemischen Industrie. M.I.T. gratefully acknowledges the support of the Alexander von Humboldt Foundation. M.G. and K.P. acknowledge generous funding by the European Research Council, ERC-2019-CoG-863934-GlycoSpec. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

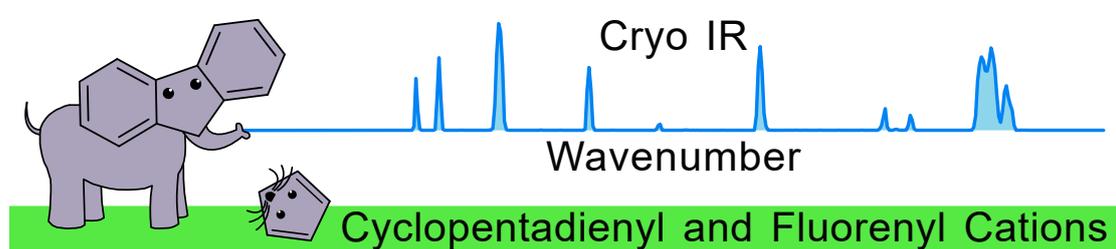
Keywords: Autohydrolysis · Fragmentation · Infrared Spectroscopy · Mass Spectrometry · RNA

- [1] J. W. Park, P. N. P. Lagniton, Y. Liu, R. H. Xu, *Int. J. Biol. Sci.* **2021**, *17*, 1446–1460.
- [2] a) N. Pardi, M. J. Hogan, F. W. Porter, D. Weissman, *Nat. Rev. Drug Discovery* **2018**, *17*, 261–279; b) C. Zhang, G. Maruggi, H. Shan, J. Li, *Front. Immunol.* **2019**, *10*, 594; c) N. A. C. Jackson, K. E. Kester, D. Casimiro, S. Gurunathan, F. DeRosa, *NPJ Vaccines* **2020**, *5*, 11.
- [3] a) M. Taucher, K. Breuker, *Angew. Chem. Int. Ed.* **2012**, *51*, 11289–11292; *Angew. Chem.* **2012**, *124*, 11451–11454; b) C. Wetzel, P. A. Limbach, *Analyst* **2016**, *141*, 16–23; c) K. Thüring, K. Schmid, P. Keller, M. Helm, *Methods* **2016**, *107*, 48–56; d) R. Lauman, B. A. Garcia, *Mol. BioSyst.* **2020**, *16*, 305–315.
- [4] S. Pääbo, H. Poinar, D. Serre, V. Jaenicke-Despres, J. Hebler, N. Rohland, M. Kuch, J. Krause, L. Vigilant, M. Hofreiter, *Annu. Rev. Genet.* **2004**, *38*, 645–679.
- [5] H. Lönnberg, *Org. Biomol. Chem.* **2011**, *9*, 1687–1703.
- [6] G. M. Emilsson, S. Nakamura, A. Roth, R. R. Breaker, *RNA* **2003**, *9*, 907–918.
- [7] S. Schürch, *Mass Spectrom. Rev.* **2016**, *35*, 483–523.
- [8] T. Y. Huang, A. Kharlamova, J. Liu, S. A. McLuckey, *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1832–1840.
- [9] S. A. McLuckey, G. J. Berkel, G. L. Glish, *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60–70.
- [10] a) S. Schürch, E. Bernal-Méndez, C. J. Leumann, *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 936–945; b) J. M. Tromp, S. Schürch, *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1262–1268.
- [11] a) T. E. Andersen, F. Kirpekar, K. F. Haselmann, *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1353–1368; b) C. Riml, H. Glasner, M. T. Rodgers, R. Micura, K. Breuker, *Nucleic Acids Res.* **2015**, *43*, 5171–5181; c) E. Fuchs, C. Falschlunger, R. Micura, K. Breuker, *Nucleic Acids Res.* **2019**, *47*, 7223–7234.
- [12] a) Y. W. Nei, K. T. Crampton, G. Berden, J. Oomens, M. T. Rodgers, *J. Phys. Chem. A* **2013**, *117*, 10634–10649; b) R. R. Wu, B. Yang, G. Berden, J. Oomens, M. T. Rodgers, *J. Phys. Chem. B* **2015**, *119*, 2795–2805.
- [13] R. E. van Outersterp, J. Martens, G. Berden, J. D. Steill, J. Oomens, A. M. Rijs, *Phys. Chem. Chem. Phys.* **2018**, *20*, 28319–28330.
- [14] B. Chiavarino, M. E. Crestoni, S. Fornarini, F. Lanucara, J. Lemaire, P. Maitre, D. Scuderi, *Int. J. Mass Spectrom.* **2008**, *270*, 111–117.
- [15] D. A. Thomas, R. Chang, E. Mucha, M. Lettow, K. Greis, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, *Phys. Chem. Chem. Phys.* **2020**, *22*, 18400–18413.
- [16] a) 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, *Angew. Chem. Int. Ed.* **2017**, *56*, 11248–11251; *Angew. Chem.* **2017**, *129*, 11400–11404; b) K. Greis, C. Kirschbaum, S. Lechnitz, S. Gewinner, W. Schöllkopf, G. von Helden, G. Meijer, P. H. Seeberger, K. Pagel, *Org. Lett.* **2020**, *22*, 8916–8919; c) C. Kirschbaum, K. Greis, L. Polewski, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, K. Pagel, *J. Am. Chem. Soc.* **2021**, *143*, 14827–14834.
- [17] U. Even, *EPL Tech. Instrum.* **2015**, *2*, 17.
- [18] W. Schöllkopf, S. Gewinner, H. Junkes, A. Paarmann, G. von Helden, H. P. Bluem, A. M. M. Todd, *Proc. SPIE-Int. Soc. Opt. Eng.* **2015**, *9512*, 95121 L.
- [19] P. Pracht, F. Bohle, S. Grimme, *Phys. Chem. Chem. Phys.* **2020**, *22*, 7169–7192.
- [20] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.
- [21] a) J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868; b) C. Adamo, V. Barone, *J. Chem. Phys.* **1999**, *110*, 6158–6170; c) S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104; d) F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- [22] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G.

- Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Wallingford, CT, **2016**.
- [23] a) V. Barone, *J. Chem. Phys.* **2005**, *122*, 014108; b) V. Barone, J. Bloino, C. A. Guido, F. Lipparini, *Chem. Phys. Lett.* **2010**, *496*, 157–161; c) J. Bloino, V. Barone, *J. Chem. Phys.* **2012**, *136*, 124108.
- [24] J. Florián, J. Leszczyński, *J. Am. Chem. Soc.* **1996**, *118*, 3010–3017.
- [25] a) J. Y. Salpin, S. Guillaumont, J. Tortajada, L. MacAleese, J. Lemaire, P. Maitre, *ChemPhysChem* **2007**, *8*, 2235–2244; b) A. Anwar, J. Psutka, S. W. C. Walker, T. Dieckmann, J. S. Janizewski, J. Larry Campbell, W. Scott Hopkins, *Int. J. Mass Spectrom.* **2018**, *429*, 174–181; c) J. M. Bakker, J.-Y. Salpin, P. Maitre, *Int. J. Mass Spectrom.* **2009**, *283*, 214–221.

Manuscript received: November 14, 2021
Accepted manuscript online: March 1, 2022
Version of record online: March 24, 2022

9 | Cyclopentadienyl and Fluorenyl Cations



The notion of (anti-)aromaticity is a successful concept in chemistry to explain structure and stability of (poly)cyclic hydrocarbons. Cyclopentadienyl and fluorenyl cations are among the most studied classical antiaromatic systems. In this work, these carbocations are investigated by high-resolution gas-phase IR spectroscopy in helium droplets. Bare cyclopentadienyl and fluorenyl cations are generated in the gas phase by nESI. After m/z selection, ions are captured in ultracold helium nanodroplets and probed by IR spectroscopy using a widely-tunable FEL in the $600\text{--}1700\text{ cm}^{-1}$ range. The highly resolved cryogenic IR spectra confirm in combination with DFT computations that all cations are present in their singlet states. Theory predicts a triplet ground state for the pentaphenylcyclopentadienyl cation, which is clearly not observed in the experiment, demanding further investigation.

The results in this chapter are not published and the text is based on a manuscript that is currently in preparation [326] – K. Greis et al., 2023, *in preparation*

9.1 | Introduction

The distribution of electrons in molecules and concepts such as electron localization and delocalization are fundamental in chemistry and are used to describe the structure, stability, and reactivity of molecules. For cyclic and polycyclic organic carbon-containing molecules, the $4n+2$ π electron Hückel rule has been particularly successful in explaining enhanced stabilities. Contrary to aromatic $4n+2$ π electron systems, $4n$ π electron systems are particularly unstable and were termed *antiaromatic* by Breslow.^[327] A prime example is the cyclopentadienyl cation (Figure 9.1), which should be antiaromatic in its singlet state. However, according to Baird's rule,^[328] the situation is reversed for triplet states where $4n$ π electron systems are expected to behave aromatic.

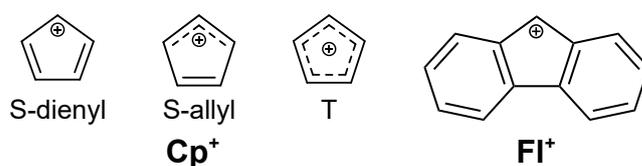


Figure 9.1: Schematic structures of the cyclopentadienyl cation Cp^+ (C_5H_5^+ ; singlet dienyl, singlet allyl, and triplet structures from left to the second from the right) and the 9-fluorenyl cation Fl^+ ($\text{C}_{13}\text{H}_9^+$; right).

Indeed, for C_5H_5^+ the triplet seems to be the ground state, however, the singlet state is only slightly higher in energy.^[329–331] Most condensed-phase experiments on cyclopentadienyl are performed on substituted species C_5X_5^+ , where X equals to phenyl (Ph) groups or halogen atoms. Early electron paramagnetic resonance experiments indicate that the singlet and triplet state of C_5Ph_5^+ are very close in energy, with the singlet being the ground state. However, even at liquid nitrogen temperature, the triplet is populated.^[332,333] In the case of chlorine substituted species, a triplet ground state is favored.^[334]

Triplet C_5X_5^+ ions are expected to be of D_{5h} symmetry, or of D_5 symmetry when bulky substituents are involved. Due to orbital degeneracy, the singlet state is susceptible to the (pseudo) Jahn-Teller effect, causing a distortion to $C_{2(v)}$ symmetry with dienyl and allylic forms as stationary states, one being a minimum and one a transition state. For C_5H_5^+ , those two structures are calculated to be extremely close in energy and which of them is predicted to be the minimum or transition state depends on the level of theory. In addition, the low lying transition state can cause bond pseudorotation, potentially leading to an average symmetric structure.^[330] Gas-phase photoelectron spectroscopy studies confirm the singlet triplet interval.^[329]

The related and potentially antiaromatic 9-fluorenyl cation (Figure 9.1) has also intrigued researchers since many decades. However, due to its high reactivity and short lifetime it is challenging to analyze and early experiments to stabilize unsubstituted fluorenyl cations in superacidic media failed due to polymerization.^[5] Measurements of rate constants for the

formation of several 9-fluorenyl cations and related species led to the conclusion that the description of the 9-fluorenyl cation as antiaromatic is misleading.^[335]

Fluorenyl cations substituted at the C9 position (the top position of the five membered ring) on the other hand are easier to generate and characterize. It is possible to stabilize 9-fluorenyl cations with alkyl, phenyl, hydroxy, and chloro groups in super acids and analyze their structure with ¹H- and ¹³C-NMR spectroscopy.^[5] Furthermore, tetrachloroaluminate crystals of a fluorenyl cation with a hydroxy group at the C9 position and methyl and mesityl substituents on the annulated benzyl rings are formed. Although the compound degrades in chlorinated NMR solvents within one day and as a solid under inert conditions within weeks, its NMR spectra and X-ray structure have been measured.^[336] In a recent study, a fluorenyl cation has been stabilized with diaminomethyl substituents, exhibiting a lifetime of minutes in moderately protic solvents.^[337] Those studies suggest that the fluorenyl cation is rather non-aromatic. While the five-membered ring itself might be described as antiaromatic, it is highly stabilized through the annulated benzene ring and potential substituents at the C9 position.

The unsubstituted 9-fluorenyl cation can be generated as a short-lived intermediate in ultrafast UV-VIS spectroscopy experiments. The ion was shown to have lifetimes of picoseconds in methanol^[338,339] and microseconds in certain zeolites,^[340] limiting detailed investigations. An alternative strategy is to study the 9-fluorenyl cation in isolation, either in the gas phase or in a non-reactive environment. Upon ionization of fluorene in the gas phase using UV light, the 9-fluorenyl cation has been identified as a fragmentation product, and some broad IR absorption bands have been tentatively assigned to it.^[341] In another study, *m/z* selected C₁₃H₉⁺ from the ionization and fragmentation of fluorene has been deposited in a neon matrix to record its electronic absorption spectrum.^[342] Further, an IR spectrum of the 9-fluorenyl cation has been measured by matrix-isolation spectroscopy in low-density water ice, after the photolysis of diazofluorene, followed by protonation. The accompanying calculations are in very good agreement with the experimental data and the authors conclude that the presence of the water ice matrix introduces only negligible shifts.^[343]

An ideal matrix for performing spectroscopic experiments is superfluid helium. Recently a technique was introduced in which *m/z* selected molecular ions are implanted into superfluid helium droplets that have an equilibrium temperature of 0.4 K. This technique allows for the spectroscopic investigation of ions at ultralow temperature, almost free of interactions with the surroundings and free of interactions with counterions. This technique has been previously applied to characterize various ions, including FH₂CO₃⁻,^[344] protein ions,^[345] and the reactive intermediate of the glycosylation reaction.^[87,305]

In this chapter, derivatized cyclopentadienyl and 9-fluorenyl cations were investigated using vibrational spectroscopy under ultracold conditions. The investigated ions are generated using nESI and then resonant IR excitation of *m/z*-selected ions in helium nanodroplets is performed.

The resulting spectra consist of narrow bands and allow for the assignment of spin state and structure by comparison with DFT-computed harmonic frequencies.

9.2 | Experimental Details

Sample Preparation

Bromopentaphenylcyclopentadiene (Sigma-Aldrich, 90%), tetraphenylcyclopentadienone (Sigma-Aldrich, 98%), 9-fluorenyl methacrylate (Sigma-Aldrich, 97%), 9-bromo-9-phenylfluorene (Sigma-Aldrich, 97%), and 9-fluorenone (Sigma-Aldrich, 98%) were purchased and were dissolved in a mixture of acetonitrile and water (9:1, v:v) to yield 200 μm solutions.

Cryogenic Infrared Spectroscopy

The experimental setup has been previously described in detail (see Chapter 3).^[87,276,305,344,345] The desired cations are generated by nESI and subsequent in-source fragmentation of precursor molecules carrying an appropriate leaving group or by protonation of precursors with a carbonyl group (Figures F1 and F2). The pentaphenylcyclopentadienyl cation **Ph₅Cp⁺** (C₃₅H₂₅⁺) and the hydrotetraphenylcyclopentadienyl cation **HOPh₄Cp⁺** (C₂₉H₂₁O⁺) were generated by fragmentation of bromopentaphenylcyclopentadiene and protonation of tetraphenylcyclopentadienone, respectively (Figure F1). The 9-fluorenyl cation **Fl⁺** (C₁₃H₉⁺), the 9-phenyl-9-fluorenyl cation **PhFl⁺** (C₁₉H₁₃⁺), and the 9-hydroxy-9-fluorenyl cation **HOFl⁺** (C₁₃H₉O⁺) were generated by fragmentation of 9-fluorenyl methacrylate, 9-bromo-9-phenylfluorene, and protonation of 9-fluorenone, respectively (Figure F2). Interestingly, no parent ion signals were visible for bromo precursors. Hence, bromide leaving groups are cleaved easily under the employed ionization conditions.

After their generation, the ions are m/z selected by a quadrupole mass filter and injected into a hexapole ion trap that is cooled to 90 K. Helium nanodroplets are generated by a cryogenic pulsed valve and directed through the ion trap where ion pick-up takes place. Due to their high kinetic energy, the ion-doped droplets can escape the longitudinal trapping potential (ca. 3 V) and travel further downstream where they are overlapped with the IR beam of the FHI FEL.^[277] When the IR wavelength of the laser is resonant with an IR-active vibrational mode of the ion, absorption of photons can take place, leading to helium evaporation and the release of the ions. Subsequently, these ions are detected in a TOF MS. This process requires the absorption of many photons. Nonetheless, due to the fast relaxation of the energy (<1 ns) and the long FEL macropulse (ca. 10 μs), each absorption event will occur from a cold (0.4 K) ion in its vibrational ground state and narrow absorption lines are expected. The infrared spectrum is then obtained by plotting the m/z selected ion signal as a function of the IR wavenumber. Spectra in the 600–1700 cm^{-1} range are recorded.

Computational Methods

Structures and harmonic frequencies of singlet and triplet states of the ions are obtained using the CAM-B3LYP hybrid exchange–correlation functional,^[288] the GD3BJ dispersion correction,^[252,346] and the def2-TZVPP basis set.^[259] Harmonic frequencies are scaled by an empirical factor of 0.965. The energetics of the singlet and triplet states of the pentaphenylcyclopentadienyl cation **Ph₅Cp⁺**, the hydroxytetraphenylcyclopentadienyl cation **HOPh₄Cp⁺**, the 9-fluorenyl cation **Fl⁺**, the 9-phenyl-9-fluorenyl cation **PhFl⁺**, and the 9-hydroxy-9-fluorenyl cation **HOFl⁺** are displayed in Table F1. All geometries were optimized using tight optimization settings in Gaussian 16^[267] at the respective level of theory.

In Tables F2–F6, the experimentally resolved absorption bands are compared to computed harmonic frequencies at the CAM-B3LYP+D3/def2-TZVPP^[252,259,288] level of theory of the singlet state, scaled by an empirical factor of 0.965. For **Fl⁺** only, the data is also compared to experimental results from matrix isolation spectroscopy,^[343] and harmonic frequencies computed at the PBE0+D3/def2-TZVPP^[248] and B3LYP+D3/def2-TZVPP^[245,246] levels of theory. For **Fl⁺** and **HOFl⁺** only, anharmonic frequencies were computed at the CAM-B3LYP+D3/def2-TZVPP level of theory using the GVPT2 method.^[264–266] Experimental and computed IR spectra are also compared in Figures 9.2, 9.4, and F3. Energetics, tables with experimental and computed IR frequencies, and 3D-structures can be found in Appendix F.

9.3 | Results and Discussion

In the top trace of Figure 9.2a, the helium droplet IR spectrum of **Ph₅Cp⁺** is shown. Sharp resonances can be observed with widths that are limited by the bandwidth of the FHI FEL (ca. 0.4% of the respective wavenumber). The experimental spectrum can be compared to results from computations. Three stationary states are found, whose 3D-structures and coordinates are shown in Figure F7. A triplet structure with *D*₅ symmetry is lowest in energy. Two singlet structures were computed, one dienyl and one allyl structure. Both have *C*₂ symmetry and an ¹A electronic state. Interestingly, those two structures are extremely close in energy. At the chosen level of theory, the dienyl structure is lower in energy by less than 0.1 kJ mol⁻¹. Such a difference is smaller than the expected error in quantum chemical calculations on molecules of that size. Compared to the triplet structure, both structures are 14 kJ mol⁻¹ higher in energy. The differences in zero-point energies of all three molecules are within 0.1 kJ mol⁻¹. Similar results are obtained when other functionals and methods are employed (see Table F1).

For all three species, harmonic IR spectra are calculated. The triplet and the dienyl singlet structures are found to be minima on their potential energy surfaces. The allyl structure, however, shows a small imaginary frequency of -10 cm^{-1} . The calculated IR spectra of the three species are compared to the experimental spectrum in Figure 9.2a. The best agreement

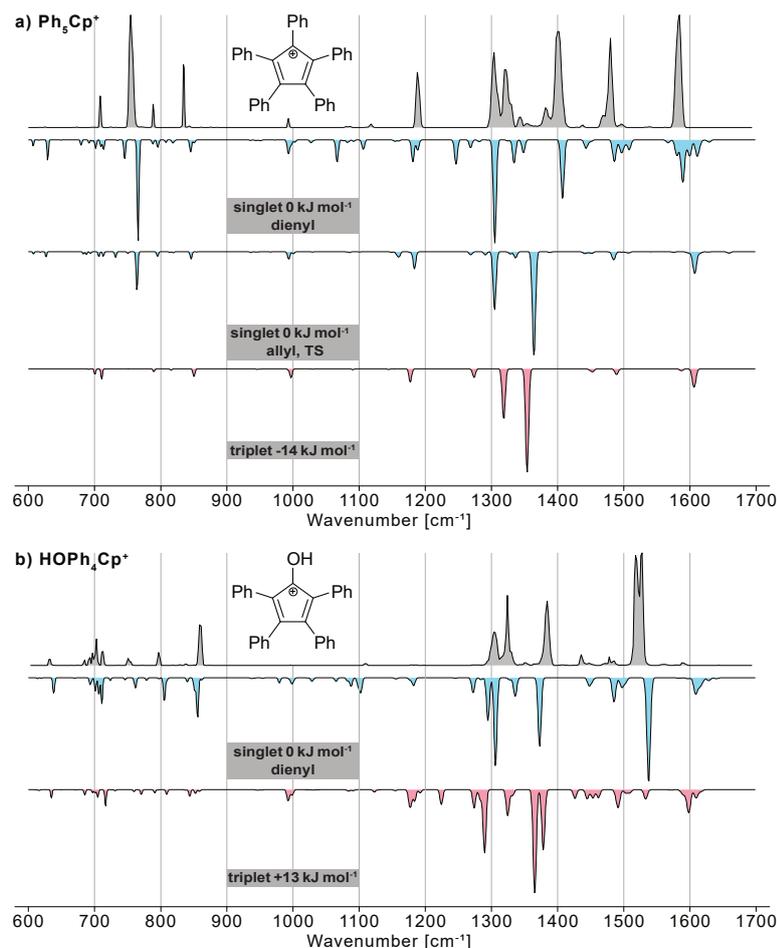


Figure 9.2: IR spectra of (a) the pentaphenylcyclopentadienyl cation Ph_5Cp^+ and (b) the tetraphenylhydroxycyclopentadienyl cation HOPh_4Cp^+ . Experimental cryogenic gas-phase IR spectra are shown as gray traces. Computed spectra for the singlet (blue) and triplet states (red) are shown in the inverted traces. The structure and relative energies computed at the CAM-B3LYP+D3/def2-TZVPP level of theory are shown.

between experimental and calculated spectra is obtained for the singlet dienyl structure. Nearly all experimentally observed bands have their counterpart in this calculated spectrum, and band positions match very well (Table F2). There are some deviations in relative intensity that could also be caused by the measurement technique, which is known to deliver reliable band positions but less accurate relative intensities due to the non-linear multiphoton absorption.^[347] Almost all calculated bands with significant IR intensity correspond to normal modes that are dominated by in-plane motions of the cyclopentadienyl ring. Particularly telling is the calculated band of the dienyl structure at 1407 cm^{-1} , which is dominated by in-plane motion parallel to the C_2 symmetry axis of the single bonded C-atom, which matches well with the experiment. On the other hand, the band of the allyl structure at 1364 cm^{-1} , which is dominated by in-plane sideways motion (perpendicular to the C_2 symmetry axis) of the central atom in the delocalized

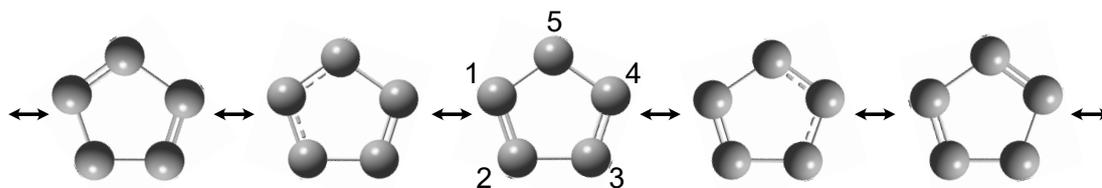


Figure 9.3: Structures and isomerization pathway of the singlet state cyclopentadienyl cation between the dienyl and the allyl form. The central dienyl structure can be converted into symmetry equivalent dienyl structures *via* allyl transition states leading to slight changes in C–C bond lengths.

bond, is not observed experimentally. However, there are also bands predicted with modest intensity for the dienyl structure at 1066 cm^{-1} and 1246 cm^{-1} that are not reproduced by the experiment. The mode predicted at 1066 cm^{-1} might have a very faint experimental counterpart at 1082 cm^{-1} , however, no intensity is observed near 1246 cm^{-1} . Nonetheless, together with the overall agreement the data strongly suggests that the singlet dienyl is the structure probed at the experimental conditions. Surprisingly, the ion does not seem to be in a triplet state, although this state is calculated to be significantly lower in energy.

A slight distortion along the -10 cm^{-1} imaginary mode of the allyl structure followed by reoptimization yields the dienyl structure. Therefore, at the present level of theory, the allyl structure is a transition state connecting two symmetry-equivalent dienyl structures as displayed in Figure 9.3. The full cycle is shown in Figure F6.

A total of five symmetry-equivalent dienyl and allyl structures can be constructed that can interconvert.^[330] The lowest vibrational modes are at 12 cm^{-1} for the dienyl and at -10 cm^{-1} (an imaginary transition state mode) for the allyl structure. Both are of B symmetry and carry by far the most IR intensity (compared to the other modes below 100 cm^{-1}), indicative of the involvement of motion of charge centers in the respective normal modes. A visual inspection of those two modes shows that significant atomic displacement stems from in-plane sideways motion of the atom, which is connected by single bonds for the dienyl structure. For the allyl structure the atom in the center of the delocalized bond, opposing the double bond, exhibits the most atomic displacement. This atomic motion is schematically indicated in Figure 9.3 and corresponds to the motion along the reaction coordinate for isomerization. As the calculated barrier is extremely small (0.1 kJ mol^{-1}), rapid interconversion leading to bond pseudorotation is expected to occur. As the zero-point vibrational energies are possibly above this barrier, this process could even occur at 0.4 K in a helium droplet. However, it is at the same time difficult to draw firm conclusions from the calculations, as the computed energy differences are within the expected error. Further, the vibrations for this ion are only calculated at the harmonic approximation even though anharmonic effects may not be negligible.

Figure 9.2b displays the experimental and computed infrared spectra of HOPh_4Cp^+ . For the molecule in its singlet state, calculations exclusively yield a dienyl type structure while no stable

point for an allyl type structure was obtained. Contrary to Ph_5Cp^+ , the triplet state is calculated to be 11 kJ mol^{-1} higher in energy, which is likely the result of the mesomeric π -donation effect of the OH group. Comparison between the calculated and experimental spectra shows a good agreement for the dienyl singlet structure. In the experiment the feature around 1520 cm^{-1} , originating from a $\text{C}=\text{O}^+$ -stretching motion, appears to consist of two partially resolved bands, while only one is predicted by harmonic frequencies (Table F3). Such a splitting might be the result of anharmonic coupling.

Figure 9.4 shows the cryogenic infrared spectra of Fl^+ , PhFl^+ and HOFl^+ . As in the case of the cyclopentadienyl cation derivatives, narrow absorption bands are observed. Although the ions only differ in the substituent at the C9 position, the spectral signatures are significantly different. The positions of the absorption bands of Fl^+ agree with those previously recorded using matrix isolation spectroscopy in water ice in the $900\text{--}1650 \text{ cm}^{-1}$ region (see Table F4).^[343]

The experimental infrared spectra are compared to the respective computed harmonic spectra for singlet and triplet states (Figure 9.4, Tables F4–F6). For all three ions, the IR signatures of the singlet and triplet ions differ significantly, and the calculated singlet spectra agree very well with the respective experimental spectra. It can thus be safely stated that the ground states of Fl^+ , PhFl^+ and HOFl^+ are singlet electronic states. This result is in line with theory, which predicts triplet states that are significantly higher in energy than the singlet states for all three ions. The singlet-triplet gap increases with increasing mesomeric π -donation effect of the substituent.

The Fl^+ cation is planar and exhibits a C_{2v} -symmetry. All high intensity modes of the calculated spectrum of the singlet state of Fl^+ are in-plane bond stretching and bending modes, except for the vibrational bands at 700 and 770 cm^{-1} , which correspond to out-of-plane C–H rocking modes. The band calculated at 984 cm^{-1} is a breathing mode of the two benzene rings. Higher in wavenumber, the modes at 1072 cm^{-1} and 1153 cm^{-1} have almost exclusive C–H bend character. The weak mode at 1234 cm^{-1} , the stronger mode at 1330 cm^{-1} , as well as the mode calculated at 1468 cm^{-1} have C–H bend character with some C–C stretch and bend motion mixed in, while the bands above 1500 cm^{-1} have mainly C=C stretch character. Experimental as well as calculated frequencies can be found in Table F4.

The PhFl^+ cation, on the other hand, is C_2 -symmetric, as the phenyl substituent at the C9 position is, despite the sp^2 -character of the C9-atom, bent out of plane. The steric repulsion between the hydrogen atoms at the *ortho*-position of the phenyl substituent and the hydrogen atoms of the fluorenyl moiety contributes to its rotation. Similar to the Fl^+ cation, the absorption bands in the $600\text{--}1000 \text{ cm}^{-1}$ are mainly originating from C–H out-of-plane rocking modes. The absorption bands in the $1000\text{--}1500 \text{ cm}^{-1}$ region are originating from bending and stretching modes of C–H and C–C moieties, while C=C stretches are visible in the $1500\text{--}1600 \text{ cm}^{-1}$ region. Interestingly, the absorption band at 1508 cm^{-1} is originating from the C=C stretch of the bond

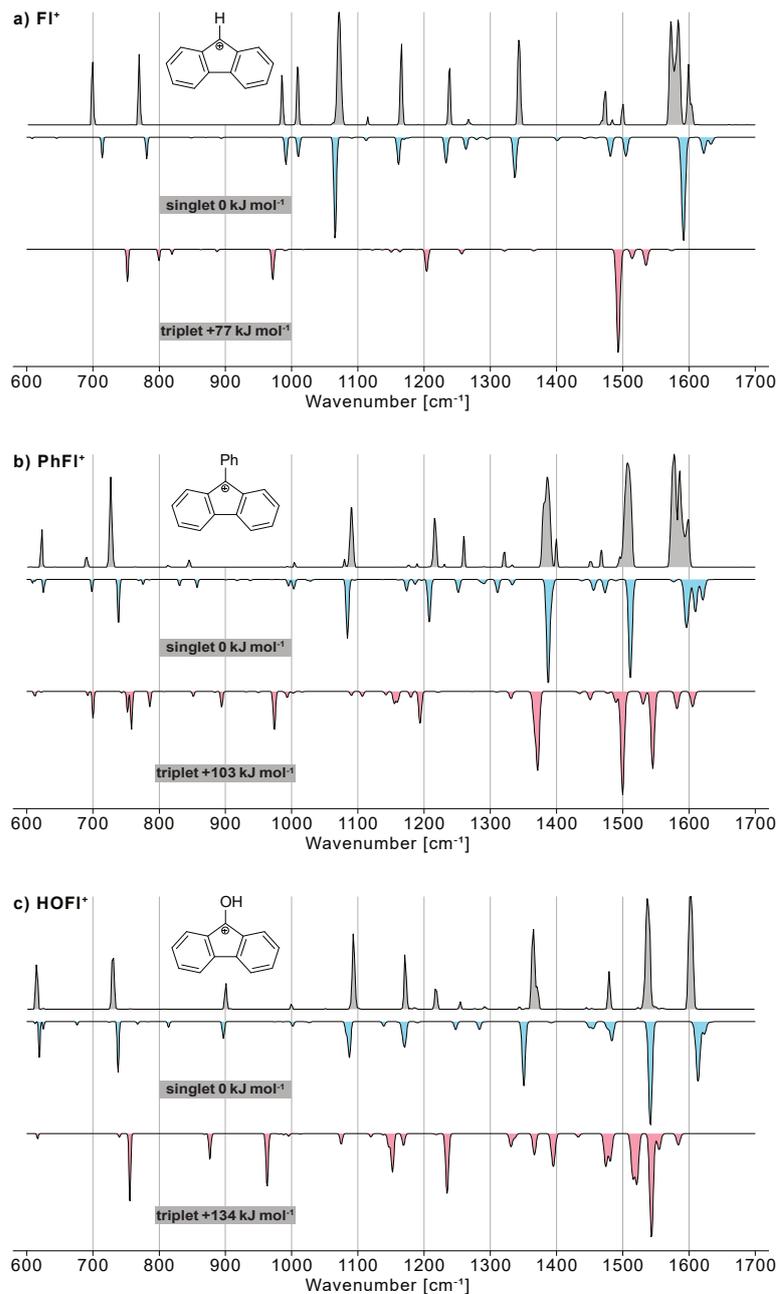


Figure 9.4: IR spectra of (a) the 9-fluorenyl cation **FI⁺**, b) the 9-phenylfluorenyl cation **PhFI⁺**, and (c) the 9-hydroxyfluorenyl cation **HOFI⁺**. Experimental cryogenic gas-phase infrared spectra are shown as gray traces. Computed spectra for the singlet (blue) and triplet states (red) are shown in the inverted traces. The structure and relative energies computed at the CAM-B3LYP+D3/def2-TZVPP level of theory are indicated.

that connects the fluorenyl and phenyl groups. The high double bond character is indicating that the **PhFI⁺** cation is stabilized by a positive mesomeric effect of the phenyl group.

The **HOFI**⁺ cation is planar and C_S -symmetric, as the protonation of the precursor is breaking the C_{2v} -symmetry of the fluorenone precursor. Interestingly, its experimental IR spectrum displayed in Figure 9.4c is sharing many similarities to the spectrum of **PhFI**⁺. As a consequence, the origin of the absorption bands is largely the same, except for the absorption bands at 615 and 1538 cm^{-1} , which are originating from an out-of-plane O–H rocking mode and a C–O stretching mode with partial double bond character, respectively.

9.4 | Conclusion

The results show that the fluorenyl cation and its derivatives are probed as singlet "dienylic" structures. Such a structure can naively be regarded as two aromatic benzene moieties, connected by a single bond and a bridging carbon atom, which also acts as charge carrier. Both an allyl-type structure or a triplet electronic state would disrupt conjugation and aromaticity in the two benzene moieties. For the cyclopentadienyl cation and its derivatives, the situation is not so clear. For this ion, good arguments can be brought forward why a triplet state would be lowest in energy. And indeed, at least for C_5H_5^+ , experiment and theory agree on a triplet ground state.^[329–331] The situation on the singlet surface is less clear. An allyl, dienyl or average symmetric structure seem possible. For the investigated **Ph₅Cp**⁺ and **HOPh₄Cp**⁺ molecules, the experimental data point towards singlet dienyl structures. For **HOPh₄Cp**⁺, this agrees with theory, however, for **Ph₅Cp**⁺, it does not. For this ion, computations suggest a triplet ground state, and for the singlet surface, one would possibly expect an average symmetric structure. The reason for this discrepancy is not clear.

In conclusion, this chapter presents a method by which bare pentaphenylcyclopentadienyl and 1-hydroxy-2,3,4,5-tetraphenylcyclopentadienyl cations as well as bare 9-fluorenyl, 9-phenyl-9-fluorenyl, and 9-hydroxy-9-fluorenyl cations can be generated and isolated in the gas phase. Their subsequent analysis by cryogenic IR spectroscopy in helium nanodroplets yields IR spectra containing narrow and highly resolved absorption bands, offering a large amount of structural information. Comparison with DFT calculations shows that in all cases, intact ions are probed in their singlet state. This result is surprising for the pentaphenylcyclopentadienyl cation, as theory predicts a triplet ground state, which is not observed in the experiment.

10 | Summary and Outlook

Mass spectrometry was combined with infrared spectroscopy in helium droplets to get structural insight into ionic intermediates from organic reactions. Electrospray ionization followed by fragmentation of distinct precursors allows for generating reactive intermediates, such as glycosyl cations. Their lifetime in solution is often too short to get meaningful structural data. With mass spectrometry-based techniques, the generated ions can be m/z selected and their vibrational modes can be probed using infrared action spectroscopy in helium nanodroplets. The obtained infrared spectrum can be directly linked to a 3D-structure by matching the data with computed harmonic frequencies. With this technique, the structures of various glycosyl cations (Chapters 4–7), the intermediate of RNA autohydrolysis (Chapter 8), and small carbocations (Chapter 9) were elucidated.

In some cases, the structure of glycosyl cations can be directly correlated with the observed stereoselectivity in glycosylation reactions. It is known that an acyl group at the C2 position of sugars leads to high 1,2-*trans* selectivity, due to neighboring-group participation. Despite its decades-long use, experimental proof of the intermediate's structure in neighboring-group participation remained elusive until recently.^[85,86] The results presented in Chapter 5 show that neighboring-group participation of C2-benzoyl groups can be directly observed in the IR spectrum of the intermediate. The probed intermediates also carried other groups, *i.e.*, Fmoc, Bn, and fluorine. Although the ring pucker changes with the replacement of a Bn group by fluorine, it does not impede the occurrence of neighboring-group participation. In Chapter 4, it is shown that acyl groups at the C4 position can form a covalent bond to the anomeric carbon, increasing the 1,2-*cis* selectivity of such building blocks. The obtained results illustrate that the substituent of acyl groups plays a major role. The electron-donating substituent pivaloyl increases the 1,2-*cis* selectivity in galactosylations compared to electron-withdrawing or neutral substituents, such as trifluoroacetyl or acetyl. Based on these results, a differentially-protected building block was developed that shows total 1,2-*cis* selectivity in galactosylations, when employed in automated glycan assembly.

A special type of glycosyl cations are Ferrier cations. They are the intermediates of the Ferrier rearrangement reaction, involving unsaturated pyranose-based monosaccharides. The intermediates are formed after cleavage of the acetoxy group at the C3 position. Out of the

two structures that were postulated for this intermediate, the results in Chapter 6 allowed to clearly determine that the intermediate is stabilized by neighboring-group participation of the C4-acetyl group at the C3 position in the gas phase. Evidence for a structure in which the positive charge is delocalized over the O–C1–C2–C3 unit was not found. The workflow was then applied to a septanose-based monosaccharide analog in Chapter 7. One might assume that the intermediate would be similar to that of the classical Ferrier rearrangement. However, after precursor activation, the generated ionic intermediate undergoes rearrangement, where the carbonyl oxygen of the C5-acetyl group attacks the C6-carbon atom, leading to ring opening, and the formation of a more stable five-membered dioxolenium ring. These rearrangements have also been previously observed in the gas phase for certain pyranose-based glycosyl cations.^[88] An acyl group at the C6 position, electron-rich acyl groups, and a larger ring size seem to favor this rearrangement, which has not yet been observed in the condensed phase. However, condensed-phase ring contractions in carbohydrates, leading to a different product, have been observed for pyranose- and septanose-based systems.^[312,348]

The experimental and computational workflow used in this thesis can also be extended to the intermediate of RNA autohydrolysis (Chapter 8). It was determined that *c*-fragments of deprotonated RNA dinucleotides correspond to the autohydrolysis intermediate, independent of the nucleobase. It was further determined that tautomerization of the nucleobase cytosine occurs to a certain degree upon activation. Both the autohydrolysis of RNA and the tautomerization of cytosine are processes that occur under physiological conditions. The results show that these processes can also happen in isolation in the gas phase, indicating that they are intrinsic to the investigated molecules. In Chapter 9, methods are presented to generate potentially antiaromatic fluorenyl and cyclopentadienyl cations. The IR spectra reveal that the intermediates are intact when probed and solely populate the singlet state. This result is particularly surprising for the pentaphenylcyclopentadienyl cation where the triplet state is calculated to be more stable than the singlet state.

Although very valuable results can be obtained using cryogenic infrared spectroscopy, it is currently a niche method and commercial setups are not available yet. In certain cases, such as IRMPD spectroscopy, it is sufficient to interface commercial instruments with a powerful infrared laser to perform the experiment. For cryogenic infrared spectroscopy, on the other hand, a significant body of custom-built hardware and knowledge is necessary for a successful experiment. Promising efforts^[349] are underway to commercialize setups employing messenger-tagging spectroscopy for bioanalytical purposes, such as the analysis of isomeric glycans. To make the technique accessible to a broad range of researchers, the commercial setups need to be equipped with user-friendly hardware and software. In the meantime, researchers that do not have direct access to such instruments can implement their ideas through collaborations or free-electron laser user facilities.

To conclude, the results show that the presented workflow combining mass spectrometry, cryogenic vibrational spectroscopy, and computational methods can be efficiently used to determine the structure of reactive intermediates. In some cases, the stereoselectivity of a glycosylation reaction can be predicted by knowing the structure of the underlying intermediate. Therefore, the herein-presented methodology could be used to probe additional glycosyl cations to generate an in-depth understanding of their reactivity. To better correlate the gas-phase data with condensed phase reactions, the influence of solvent molecules on the structure of the intermediate needs to be studied. The interactions of solvent molecules and glycosyl cations could be investigated in the gas phase after microsolvation using cryogenic infrared spectroscopy. In a next step, ion-molecule reactions could be added to the setup, allowing not only to characterize the intermediate, but to directly perform a glycosylation reaction in the gas phase. This toolbox would allow to selectively probe the structure and reactivity of glycosyl cations and could subsequently be extended to other intermediates from organic and inorganic chemistry. Thereby, a general understanding of the intrinsic reactivity of intermediates could be obtained.

Bibliography

- [1] T. L. Amyes, W. P. Jencks, *J. Am. Chem. Soc.* **1989**, *111*, 7888–7900.
- [2] F. Bravo, A. Viso, E. Alcázar, P. Molas, C. Bo, S. Castellón, *J. Org. Chem.* **2002**, *68*, 686–691.
- [3] Y. Fu, L. Bernasconi, P. Liu, *J. Am. Chem. Soc.* **2021**, *143*, 1577–1589.
- [4] G. A. Olah, M. B. Comisarow, *J. Am. Chem. Soc.* **1968**, *90*, 5033–5034.
- [5] G. A. Olah, G. K. S. Prakash, G. Liang, P. W. Westerman, K. Kunde, J. Chandrasekhar, P. V. R. Schleyer, *J. Am. Chem. Soc.* **1980**, *102*, 4485–4492.
- [6] W. Sander, *Angew. Chem. Int. Ed.* **1990**, *29*, 344–354.
- [7] H. P. Reisenauer, J. Romanski, G. Mloston, P. R. Schreiner, *Eur. J. Org. Chem.* **2006**, *2006*, 4813–4818.
- [8] J. Roithová, *Chem. Soc. Rev.* **2012**, *41*, 547–559.
- [9] R. Huisgen, *Angew. Chem. Int. Ed.* **1970**, *9*, 751–762.
- [10] *Reactive Intermediate Chemistry*, (Eds.: R. A. Moss, M. S. Platz, M. Jones), John Wiley & Sons, Inc., **2003**.
- [11] E. Fischer, *Ber. Dtsch. Chem. Ges.* **1893**, *26*, 2400.
- [12] P. M. Rudd, T. Elliott, P. Cresswell, I. A. Wilson, R. A. Dwek, *Science* **2001**, *291*, 2370–2376.
- [13] J. D. Marth, P. K. Grewal, *Nat. Rev. Immunol.* **2008**, *8*, 874–887.
- [14] W. J. Snell, J. M. White, *Cell* **1996**, *85*, 629–637.
- [15] D. R. Tulsiani, H. Yoshida-Komiya, Y. Araki, *Biol. Reprod.* **1997**, *57*, 487–494.
- [16] P. M. Wassarman, L. Jovine, E. S. Litscher, *Nat. Cell Biol.* **2001**, *3*, E59–E64.
- [17] A Petrescu, M Wormald, R Dwek, *Curr. Opin. Struct. Biol.* **2006**, *16*, 600–607.
- [18] D. Shental-Bechor, Y. Levy, *Curr. Opin. Struct. Biol.* **2009**, *19*, 524–533.
- [19] A. Geissner, L. Baumann, T. J. Morley, A. K. O. Wong, L. Sim, J. R. Rich, P. P. L. So, E. M. Dullaghan, E. Lessard, U. Iqbal, M. Moreno, W. W. Wakarchuk, S. G. Withers, *ACS Cent. Sci.* **2021**, *7*, 345–354.

- [20] B. Linclau, A. Ardá, N.-C. Reichardt, M. Sollogoub, L. Unione, S. P. Vincent, J. Jiménez-Barbero, *Chem. Soc. Rev.* **2020**, *49*, 3863–3888.
- [21] R. Hevey, *Chem. Eur. J.* **2021**, *27*, 2240–2253.
- [22] H.-J. Lo, L. Krasnova, S. Dey, T. Cheng, H. Liu, T.-I. Tsai, K. B. Wu, C.-Y. Wu, C.-H. Wong, *J. Am. Chem. Soc.* **2019**, *141*, 6484–6488.
- [23] R. Hevey, *Pharmaceuticals* **2019**, *12*, 55.
- [24] A. Tamburrini, C. Colombo, A. Bernardi, *Med. Res. Rev.* **2020**, *40*, 495–531.
- [25] J. Sianturi, P. Priegue, J. Hu, J. Yin, P. H. Seeberger, *Angew. Chem. Int. Ed.* **2022**, *61*, e202209556.
- [26] C. Qin, L. Li, G. Tian, M. Ding, S. Zhu, W. Song, J. Hu, P. H. Seeberger, J. Yin, *J. Am. Chem. Soc.* **2022**, *144*, 21068–21079.
- [27] S. G. Parameswarappa, C. L. Pereira, P. H. Seeberger, *Beilstein J. Org. Chem.* **2020**, *16*, 1693–1699.
- [28] P. H. Seeberger, *Chem. Rev.* **2021**, *121*, 3598–3626.
- [29] G. Fittolani, S. Djalali, M. A. Chaube, T. Tyrikos-Ergas, M. C. S. D. Colle, A. Grafmüller, P. H. Seeberger, M. Delbianco, *Org. Biomol. Chem.* **2022**, *20*, 8228–8235.
- [30] S. Gim, G. Fittolani, Y. Yu, Y. Zhu, P. H. Seeberger, Y. Ogawa, M. Delbianco, *Chem. Eur. J.* **2021**, *27*, 13139–13143.
- [31] M. Delbianco, P. H. Seeberger, *Mater. Horiz.* **2020**, *7*, 963–969.
- [32] S. Moon, K. Gilmore, R. W. Hicklin, P. H. Seeberger in *Flow Chemistry in Organic Synthesis*, Georg Thieme Verlag KG, **2018**.
- [33] M. Panza, S. G. Pistorio, K. J. Stine, A. V. Demchenko, *Chem. Rev.* **2018**, *118*, 8105–8150.
- [34] R. A. Laine, *Glycobiology* **1994**, *4*, 759–767.
- [35] C.-W. Cheng, C.-Y. Wu, W.-L. Hsu, C.-H. Wong, *Biochemistry* **2019**, *59*, 3078–3088.
- [36] S. S. Kulkarni, C.-C. Wang, N. M. Sabbavarapu, A. R. Podilapu, P.-H. Liao, S.-C. Hung, *Chem. Rev.* **2018**, *118*, 8025–8104.
- [37] Z. Zhang, I. R. Ollmann, X.-S. Ye, R. Wischnat, T. Baasov, C.-H. Wong, *J. Am. Chem. Soc.* **1999**, *121*, 734–753.
- [38] C.-W. Cheng, Y. Zhou, W.-H. Pan, S. Dey, C.-Y. Wu, W.-L. Hsu, C.-H. Wong, *Nat. Commun.* **2018**, *9*, 5202.
- [39] D. Crich, S. Sun, *J. Org. Chem.* **1996**, *61*, 4506–4507.
- [40] D. Crich, *Acc. Chem. Res.* **2010**, *43*, 1144–1153.

- [41] L. Yang, Q. Qin, X.-S. Ye, *Asian J. Org. Chem.* **2013**, *2*, 30–49.
- [42] X. Huang, L. Huang, H. Wang, X.-S. Ye, *Angew. Chem. Int. Ed.* **2004**, *43*, 5221–5224.
- [43] Y.-S. Lu, Q. Li, L.-H. Zhang, X.-S. Ye, *Org. Lett.* **2008**, *10*, 3445–3448.
- [44] Y. Wu, Y. Qiu, Y. Feng, J. F. Stoddart, *ACS Cent. Sci.* **2022**, *8*, 1369–1372.
- [45] O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Science* **2001**, *291*, 1523–1527.
- [46] P. H. Seeberger, *Acc. Chem. Res.* **2015**, *48*, 1450–1463.
- [47] M. Guberman, P. H. Seeberger, *J. Am. Chem. Soc.* **2019**, *141*, 5581–5592.
- [48] A. A. Joseph, A. Pardo-Vargas, P. H. Seeberger, *J. Am. Chem. Soc.* **2020**, *142*, 8561–8564.
- [49] H. S. Hahm, M. K. Schlegel, M. Hurevich, S. Eller, F. Schuhmacher, J. Hofmann, K. Pagel, P. H. Seeberger, *Proc. Natl. Acad. Sci.* **2017**, *114*, 3385–3389.
- [50] W. Yao, D.-C. Xiong, Y. Yang, C. Geng, Z. Cong, F. Li, B.-H. Li, X. Qin, L.-N. Wang, W.-Y. Xue, N. Yu, H. Zhang, X. Wu, M. Liu, X.-S. Ye, *Nat. Synth.* **2022**, *1*, 854–863.
- [51] K. Ágoston, H. Streicher, P. Fügedi, *Tetrahedron: Asymmetry* **2016**, *27*, 707–728.
- [52] J. D. Codée, A. Ali, H. S. Overkleeft, G. A. van der Marel, *C. R. Chim.* **2011**, *14*, 178–193.
- [53] B. Ghosh, S. S. Kulkarni, *Chem. Asian J.* **2020**, *15*, 450–462.
- [54] X. Zhu, R. Schmidt, *Angew. Chem. Int. Ed.* **2009**, *48*, 1900–1934.
- [55] D. Crich, S. Sun, *Tetrahedron* **1998**, *54*, 8321–8348.
- [56] D. Crich, N. S. Chandrasekera, *Angew. Chem. Int. Ed.* **2004**, *43*, 5386–5389.
- [57] M. Huang, G. E. Garrett, N. Birlirakis, L. Bohé, D. A. Pratt, D. Crich, *Nat. Chem.* **2012**, *4*, 663–667.
- [58] J.-H. Kim, H. Yang, G.-J. Boons, *Angew. Chem. Int. Ed.* **2005**, *44*, 947–949.
- [59] J.-H. Kim, H. Yang, J. Park, G.-J. Boons, *J. Am. Chem. Soc.* **2005**, *127*, 12090–12097.
- [60] T. J. Boltje, J.-H. Kim, J. Park, G.-J. Boons, *Nat. Chem.* **2010**, *2*, 552–557.
- [61] H. Elferink, R. A. Mensink, P. B. White, T. J. Boltje, *Angew. Chem. Int. Ed.* **2016**, *55*, 11217–11220.
- [62] A. V. Demchenko, E. Rousson, G.-J. Boons, *Tetrahedron Lett.* **1999**, *40*, 6523–6526.
- [63] H. S. Hahm, M. Hurevich, P. H. Seeberger, *Nat. Commun.* **2016**, *7*, 12482.
- [64] Z. Li, L. Zhu, J. Kalikanda, *Tetrahedron Lett.* **2011**, *52*, 5629–5632.
- [65] Y. Ma, G. Lian, Y. Li, B. Yu, *Chem. Commun.* **2011**, *47*, 7515–7517.
- [66] D. Yao, Y. Liu, S. Yan, Y. Li, C. Hu, N. Ding, *Chem. Commun.* **2017**, *53*, 2986–2989.
- [67] J. Kalikanda, Z. Li, *J. Org. Chem.* **2011**, *76*, 5207–5218.

- [68] A. Franconetti, A. Ardá, J. L. Asensio, Y. Blériot, S. Thibaudeau, J. Jiménez-Barbero, *Acc. Chem. Res.* **2021**, *54*, 2552–2564.
- [69] L. Bohé, D. Crich, *C. R. Chim.* **2011**, *14*, 3–16.
- [70] J. J. R. Pernía, I. Tuñón, I. H. Williams, *J. Phys. Chem. B* **2010**, *114*, 5769–5774.
- [71] B. A. Horenstein, *J. Am. Chem. Soc.* **1997**, *119*, 1101–1107.
- [72] T. Nukada, A. Berces, M. Z. Zgierski, D. M. Whitfield, *J. Am. Chem. Soc.* **1998**, *120*, 13291–13295.
- [73] P. Merino, I. Delso, S. Pereira, S. Orta, M. Pedrón, T. Tejero, *Org. Biomol. Chem.* **2021**, *19*, 2350–2365.
- [74] M. V. Panova, M. G. Medvedev, A. V. Orlova, L. O. Kononov, *ChemPhysChem* **2021**, *23*, e202100788.
- [75] B. A. Horenstein, M. Bruner, *J. Am. Chem. Soc.* **1998**, *120*, 1357–1362.
- [76] M. Huang, P. Retailleau, L. Bohé, D. Crich, *J. Am. Chem. Soc.* **2012**, *134*, 14746–14749.
- [77] S. Hanessian, J. Banoub, *Carbohydr. Res.* **1975**, *44*, C14–C17.
- [78] D. Crich, Z. Dai, S. Gastaldi, *J. Org. Chem.* **1999**, *64*, 5224–5229.
- [79] A. Martin, A. Arda, J. Désiré, A. Martin-Mingot, N. Probst, P. Sinaÿ, J. Jiménez-Barbero, S. Thibaudeau, Y. Blériot, *Nat. Chem.* **2015**, *8*, 186–191.
- [80] T. Hansen, L. Lebedel, W. A. Remmerswaal, S. van der Vorm, D. P. A. Wander, M. Somers, H. S. Overkleeft, D. V. Filippov, J. Désiré, A. Mingot, Y. Bleriot, G. A. van der Marel, S. Thibaudeau, J. D. C. Codée, *ACS Cent. Sci.* **2019**, *5*, 781–788.
- [81] L. Lebedel, A. Ardá, A. Martin, J. Désiré, A. Mingot, M. Aufiero, N. Aiguabella Font, R. Gilmour, J. Jiménez-Barbero, Y. Blériot, S. Thibaudeau, *Angew. Chem. Int. Ed.* **2019**, *58*, 13758–13762.
- [82] N. Bhuma, L. Lebedel, H. Yamashita, Y. Shimizu, Z. Abada, A. Ardá, J. Désiré, B. Michelet, A. Martin-Mingot, A. Abou-Hassan, M. Takumi, J. Marrot, J. Jiménez-Barbero, A. Nagaki, Y. Blériot, S. Thibaudeau, *Angew. Chem. Int. Ed.* **2020**, *60*, 2036–2041.
- [83] K. Upadhyaya, Y. P. Subedi, D. Crich, *Angew. Chem. Int. Ed.* **2021**, *60*, 25397–25403.
- [84] F. F. J. de Kleijne, H. Elferink, S. J. Moons, P. B. White, T. J. Boltje, *Angew. Chem. Int. Ed.* **2022**, *61*, e202109874.
- [85] E. Mucha, M. Marianski, F.-F. Xu, D. A. Thomas, G. Meijer, G. von Helden, P. H. Seeberger, K. Pagel, *Nat. Commun.* **2018**, *9*, 4174.
- [86] H. Elferink, M. E. Severijnen, J. Martens, R. A. Mensink, G. Berden, J. Oomens, F. P. J. T. Rutjes, A. M. Rijs, T. J. Boltje, *J. Am. Chem. Soc.* **2018**, *140*, 6034–6038.

- [87] M. Marianski, E. Mucha, K. Greis, S. Moon, A. Pardo, C. Kirschbaum, D. A. Thomas, G. Meijer, G. von Helden, K. Gilmore, P. H. Seeberger, K. Pagel, *Angew. Chem. Int. Ed.* **2020**, *59*, 6166–6171.
- [88] T. Hansen, H. Elferink, J. M. A. van Hengst, K. J. Houthuijs, W. A. Remmerswaal, A. Kromm, G. Berden, S. van der Vorm, A. M. Rijs, H. S. Overkleeft, D. V. Filippov, F. P. J. T. Rutjes, G. A. van der Marel, J. Martens, J. Oomens, J. D. C. Codée, T. J. Boltje, *Nat. Commun.* **2020**, *11*, 2664.
- [89] H. Elferink, R. A. Mensink, W. W. A. Castelijns, O. Jansen, J. P. J. Bruekers, J. Martens, J. Oomens, A. M. Rijs, T. J. Boltje, *Angew. Chem. Int. Ed.* **2019**, *58*, 8746–8751.
- [90] H. Elferink, W. A. Remmerswaal, K. J. Houthuijs, O. Jansen, T. Hansen, A. M. Rijs, G. Berden, J. Martens, J. Oomens, J. D. C. Codée, T. J. Boltje, *Chem. Eur. J.* **2022**, *28*, e202201724.
- [91] W. A. Remmerswaal, K. J. Houthuijs, R. van de Ven, H. Elferink, T. Hansen, G. Berden, H. S. Overkleeft, G. A. van der Marel, F. P. J. T. Rutjes, D. V. Filippov, T. J. Boltje, J. Martens, J. Oomens, J. D. C. Codée, *J. Org. Chem.* **2022**, *87*, 9139–9147.
- [92] F. ter Braak, H. Elferink, K. J. Houthuijs, J. Oomens, J. Martens, T. J. Boltje, *Acc. Chem. Res.* **2022**, *55*, 1669–1679.
- [93] K. Greis, C. Kirschbaum, G. von Helden, K. Pagel, *Curr. Opin. Struct. Biol.* **2022**, *72*, 194–202.
- [94] P. O. Adero, H. Amarasekara, P. Wen, L. Bohé, D. Crich, *Chem. Rev.* **2018**, *118*, 8242–8284.
- [95] Y. Blériot in *Comprehensive Glycoscience*, Elsevier, **2021**, pp. 49–82.
- [96] S. Chatterjee, S. Moon, F. Hentschel, K. Gilmore, P. H. Seeberger, *J. Am. Chem. Soc.* **2018**, *140*, 11942–11953.
- [97] C.-W. Chang, C.-H. Wu, M.-H. Lin, P.-H. Liao, C.-C. Chang, H.-H. Chuang, S.-C. Lin, S. Lam, V. P. Verma, C.-P. Hsu, C.-C. Wang, *Angew. Chem. Int. Ed.* **2019**, *58*, 16775–16779.
- [98] C.-H. Wong, X.-S. Ye, Z. Zhang, *J. Am. Chem. Soc.* **1998**, *120*, 7137–7138.
- [99] C.-C. Wang, J.-C. Lee, S.-Y. Luo, S. S. Kulkarni, Y.-W. Huang, C.-C. Lee, K.-L. Chang, S.-C. Hung, *Nature* **2007**, *446*, 896–899.
- [100] J. Guo, X.-S. Ye, *Molecules* **2010**, *15*, 7235–7265.
- [101] K. R. Love, P. H. Seeberger, *Angew. Chem. Int. Ed.* **2004**, *43*, 602–605.
- [102] S. van der Vorm, T. Hansen, H. S. Overkleeft, G. A. v. d. Marel, J. D. C. Codée, *Chem. Sci.* **2017**, *8*, 1867–1875.
- [103] S. van der Vorm, T. Hansen, J. M. A. v. Hengst, H. S. Overkleeft, G. A. v. d. Marel, J. D. C. Codée, *Chem. Soc. Rev.* **2019**, *48*, 4688–4706.

- [104] C.-W. Chang, M.-H. Lin, C.-H. Wu, T.-Y. Chiang, C.-C. Wang, *J. Org. Chem.* **2020**, *85*, 15945–15963.
- [105] C.-W. Chang, M.-H. Lin, C.-C. Wang, *Chem. Eur. J.* **2020**, *27*, 2556–2568.
- [106] C.-W. Chang, M.-H. Lin, C.-K. Chan, K.-Y. Su, C.-H. Wu, W.-C. Lo, S. Lam, Y.-T. Cheng, P.-H. Liao, C.-H. Wong, C.-C. Wang, *Angew. Chem. Int. Ed.* **2021**, *60*, 12413–12423.
- [107] S. Moon, S. Chatterjee, P. H. Seeberger, K. Gilmore, *Chem. Sci.* **2021**, *12*, 2931–2939.
- [108] O. T. Tuck, E. T. Sletten, J. Danglad-Flores, P. H. Seeberger, *Angew. Chem. Int. Ed.* **2022**, *61*, e202115433.
- [109] J. M. A. van Hengst, R. J. C. Hellemons, W. A. Remmerswaal, K. N. A. van de Vrande, T. Hansen, S. van der Vorm, H. S. Overkleeft, G. A. van der Marel, J. D. C. Codée, *Chem. Sci.* **2023**, *14*, 1532–1542.
- [110] D. Crich, *J. Am. Chem. Soc.* **2020**, *143*, 17–34.
- [111] D. Crich, S. Sun, *J. Am. Chem. Soc.* **1997**, *119*, 11217–11223.
- [112] T. Hosoya, T. Takano, P. Kosma, T. Rosenau, *J. Org. Chem.* **2014**, *79*, 7889–7894.
- [113] C. S. Callam, R. R. Gadikota, D. M. Krein, T. L. Lowary, *J. Am. Chem. Soc.* **2003**, *125*, 13112–13119.
- [114] D. Hou, H. A. Taha, T. L. Lowary, *J. Am. Chem. Soc.* **2009**, *131*, 12937–12948.
- [115] A. G. Santana, L. Montalvillo-Jiménez, L. Díaz-Casado, F. Corzana, P. Merino, F. J. Cañada, G. Jiménez-Osés, J. Jiménez-Barbero, A. M. Gómez, J. L. Asensio, *J. Am. Chem. Soc.* **2020**, *142*, 12501–12514.
- [116] C. G. Lucero, K. A. Woerpel, *J. Org. Chem.* **2006**, *71*, 2641–2647.
- [117] M. T. Walvoort, J. Dinkelaar, L. J. van den Bos, G. Lodder, H. S. Overkleeft, J. D. Codée, G. A. van der Marel, *Carbohydr. Res.* **2010**, *345*, 1252–1263.
- [118] C. H. Larsen, B. H. Ridgway, J. T. Shaw, K. A. Woerpel, *J. Am. Chem. Soc.* **1999**, *121*, 12208–12209.
- [119] E. R. van Rijssel, P. van Delft, G. Lodder, H. S. Overkleeft, G. A. van der Marel, D. V. Filippov, J. D. C. Codée, *Angew. Chem. Int. Ed.* **2014**, *53*, 10381–10385.
- [120] J. A. C. Romero, S. A. Tabacco, K. A. Woerpel, *J. Am. Chem. Soc.* **1999**, *122*, 168–169.
- [121] R. J. Williams, N. W. McGill, J. M. White, S. J. Williams, *J. Carbohydr. Chem.* **2010**, *29*, 236–263.
- [122] W. Koenigs, E. Knorr, *Ber. Dtsch. Chem. Ges.* **1901**, *34*, 957–981.
- [123] H. Paulsen, C.-P. Herold, *Chem. Ber.* **1970**, *103*, 2450–2462.

- [124] A. A. Hettikankanamalage, R. Lassfolk, F. S. Ekholm, R. Leino, D. Crich, *Chem. Rev.* **2020**, *120*, 7104–7151.
- [125] A. Demchenko, *Curr. Org. Chem.* **2003**, *7*, 35–79.
- [126] S. S. Nigudkar, A. V. Demchenko, *Chem. Sci.* **2015**, *6*, 2687–2704.
- [127] A. V. Demchenko, P. Pornsuriyasak, C. D. Meo, *J. Chem. Educ.* **2006**, *83*, 782.
- [128] D. Crich, M. de la Mora, A. U. Vinod, *J. Org. Chem.* **2003**, *68*, 8142–8148.
- [129] A. Imamura, H. Ando, S. Korogi, G. Tanabe, O. Muraoka, H. Ishida, M. Kiso, *Tetrahedron Lett.* **2003**, *44*, 6725–6728.
- [130] A. Imamura, A. Kimura, H. Ando, H. Ishida, M. Kiso, *Chem. Eur. J.* **2006**, *12*, 8862–8870.
- [131] T. Fang, Y. Gu, W. Huang, G.-J. Boons, *J. Am. Chem. Soc.* **2016**, *138*, 3002–3011.
- [132] R. A. Mensink, T. J. Boltje, *Chem. Eur. J.* **2017**, *23*, 17637–17653.
- [133] J.-H. Kim, H. Yang, V. Khot, D. Whitfield, G.-J. Boons, *Eur. J. Org. Chem.* **2006**, *2006*, 5007–5028.
- [134] T. J. Boltje, J.-H. Kim, J. Park, G.-J. Boons, *Org. Lett.* **2010**, *13*, 284–287.
- [135] F. Pfrengle, P. H. Seeberger, Orthogonally Protected Building Blocks for Automated Glycan Assembly, **2018**.
- [136] P. O. Adero, T. Furukawa, M. Huang, D. Mukherjee, P. Retailleau, L. Bohé, D. Crich, *J. Am. Chem. Soc.* **2015**, *137*, 10336–10345.
- [137] R. A. Mensink, H. Elferink, P. B. White, N. Pers, F. P. J. T. Rutjes, T. J. Boltje, *Eur. J. Org. Chem.* **2016**, *2016*, 4656–4667.
- [138] K. Greis, E. Mucha, M. Lettow, D. A. Thomas, C. Kirschbaum, S. Moon, A. Pardo-Vargas, G. von Helden, G. Meijer, K. Gilmore, P. H. Seeberger, K. Pagel, *ChemPhysChem* **2020**, *21*, 1905–1907.
- [139] R. D. Smith, J. A. Loo, C. G. Edmonds, C. J. Barinaga, H. R. Udseth, *Anal. Chem.* **1990**, *62*, 882–899.
- [140] F. W. McLafferty, *Anal. Chem.* **1956**, *28*, 306–316.
- [141] A. Makarov, M. Scigelova, *J. Chrom. A* **2010**, *1217*, 3938–3945.
- [142] D. Collins, M. Lee, *Anal. Bioanal. Chem.* **2001**, *372*, 66–73.
- [143] A. B. Kanu, P. Dwivedi, M. Tam, L. Matz, H. H. Hill, *J. Mass Spectrom.* **2008**, *43*, 1–22.
- [144] A. P. Cismesia, M. R. Bell, L. F. Tesler, M. Alves, N. C. Polfer, *Analyst* **2018**, *143*, 1615–1623.
- [145] J. S. Ho, A. Gharbi, B. Schindler, O. Yeni, R. Brédy, L. Legentil, V. Ferrières, L. L. Kiessling, I. Compagnon, *J. Am. Chem. Soc.* **2021**, *143*, 10509–10513.

- [146] M. L. Kovarik, J. K. Robinson, *J. Chem. Educ.* **2019**, *96*, 905–911.
- [147] S. Cheng, Z. Xie, Q. Hu, Y. Qian, X. Ma, *J. Chem. Educ.* **2023**, *100*, 1270–1276.
- [148] D. P. Demarque, A. E. M. Crotti, R. Vessecchi, J. L. C. Lopes, N. P. Lopes, *Nat. Prod. Rep.* **2016**, *33*, 432–455.
- [149] K. D. Bruycker, A. Welle, S. Hirth, S. J. Blanksby, C. Barner-Kowollik, *Nat. Rev. Chem.* **2020**, *4*, 257–268.
- [150] S. Schürch, E. Bernal-Méndez, C. J. Leumann, *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 936–945.
- [151] M. Taucher, K. Breuker, *Angew. Chem. Int. Ed.* **2012**, *51*, 11289–11292.
- [152] D. F. Hunt, J. R. Yates, J. Shabanowitz, S. Winston, C. R. Hauer, *Proc. Natl. Acad. Sci.* **1986**, *83*, 6233–6237.
- [153] B. Domon, R. Aebersold, *Science* **2006**, *312*, 212–217.
- [154] N. Viseux, E. de Hoffmann, B. Domon, *Anal. Chem.* **1998**, *70*, 4951–4959.
- [155] S. J. North, P. G. Hitchen, S. M. Haslam, A. Dell, *Curr. Opin. Struct. Biol.* **2009**, *19*, 498–506.
- [156] M. Mayer, K. R. Asmis, *J. Phys. Chem. A* **2021**, *125*, 2801–2815.
- [157] J. Zelenka, R. Cibulka, J. Roithová, *Angew. Chem. Int. Ed.* **2019**, *58*, 15412–15420.
- [158] A. Kumar, S. Mondal, S. Banerjee, *J. Am. Chem. Soc.* **2021**, *143*, 2459–2463.
- [159] A. Kumar, S. Mondal, M. Mofidfar, R. N. Zare, S. Banerjee, *J. Am. Chem. Soc.* **2022**, *144*, 7573–7577.
- [160] G. R. D. Prabhu, E. R. Williams, M. Wilm, P. L. Urban, *Nat. Rev. Methods Primers* **2023**, *3*, 23.
- [161] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, *Science* **1989**, *246*, 64–71.
- [162] J. B. Fenn, *Angew. Chem. Int. Ed.* **2003**, *42*, 3871–3894.
- [163] C. Kirschbaum, K. Greis, E. Mucha, L. Kain, S. Deng, A. Zappe, S. Gewinner, W. Schöllkopf, G. von Helden, G. Meijer, P. B. Savage, M. Marianski, L. Teyton, K. Pagel, *Nat. Commun.* **2021**, *12*, 1201.
- [164] L. Konermann, E. Ahadi, A. D. Rodriguez, S. Vahidi, *Anal. Chem.* **2012**, *85*, 2–9.
- [165] T. P. Wörner, T. M. Shamorkina, J. Snijder, A. J. R. Heck, *Anal. Chem.* **2020**, *93*, 620–640.
- [166] T. Covey in *ACS Symposium Series*, American Chemical Society, **1996**, pp. 21–59.
- [167] R. Juraschek, T. Dülcks, M. Karas, *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 300–308.
- [168] U. Boesl, *Mass Spectrom. Rev.* **2016**, *36*, 86–109.

- [169] B. A. Mamyryn, V. I. Karataev, D. V. Shmikk, V. A. Zagulin, *Sov. Phys.-JETP* **1973**, *37*, 45–48.
- [170] O. M. Saad, J. A. Leary, *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1274–1286.
- [171] J. P. O'Brien, J. S. Brodbelt, *Anal. Chem.* **2013**, *85*, 10399–10407.
- [172] S. Crotti, M. Menicatti, M. Pallecchi, G. Bartolucci, *Mass Spectrom. Rev.* **2021**, e21757.
- [173] L. Sleno, D. A. Volmer, *J. Mass Spectrom.* **2004**, *39*, 1091–1112.
- [174] A. R. Johnson, E. E. Carlson, *Anal. Chem.* **2015**, *87*, 10668–10678.
- [175] K. Sannes-Lowery, R. H. Griffey, G. H. Kruppa, J. P. Speir, S. A. Hofstadler, *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1957–1961.
- [176] D. R. Carl, R. M. Moision, P. B. Armentrout, *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2312–2317.
- [177] R. A. Yost, C. G. Enke, *J. Am. Chem. Soc.* **1978**, *100*, 2274–2275.
- [178] P. E. Miller, M. B. Denton, *J. Chem. Educ.* **1986**, *63*, 617.
- [179] G. W. Hill, *Acta Math.* **1886**, *8*, 1–36.
- [180] D. J. Douglas, A. J. Frank, D. Mao, *Mass Spectrom. Rev.* **2004**, *24*, 1–29.
- [181] D. Nolting, R. Malek, A. Makarov, *Mass Spectrom. Rev.* **2017**, *38*, 150–168.
- [182] J. Haas, B. Mizaikoff, *Annu. Rev. Anal. Chem.* **2016**, *9*, 45–68.
- [183] B. H. Stuart, *Infrared Spectroscopy: Fundamentals and Applications*, Wiley, **2004**.
- [184] J. Oomens, B. G. Sartakov, G. Meijer, G. von Helden, *Int. J. Mass Spectrom.* **2006**, *254*, 1–19.
- [185] D. F. Swinehart, *J. Chem. Educ.* **1962**, *39*, 333.
- [186] *Laser Photodissociation and Spectroscopy of Mass-separated Biomolecular Ions*, (Eds.: N. C. Polfer, P. Dugourd), Springer International Publishing, **2013**.
- [187] A. M. Rijs, J. Oomens in *Top. Curr. Chem.* Springer International Publishing, **2014**, pp. 1–42.
- [188] M. Okumura, L. I. Yeh, J. D. Myers, Y. T. Lee, *J. Chem. Phys.* **1986**, *85*, 2328–2329.
- [189] M. Grabarics, M. Lettow, C. Kirschbaum, K. Greis, C. Manz, K. Pagel, *Chem. Rev.* **2022**, *122*, 7840–7908.
- [190] R. L. Woodin, D. S. Bomse, J. L. Beauchamp, *J. Am. Chem. Soc.* **1978**, *100*, 3248–3250.
- [191] T. Oka, *Phys. Rev. Lett.* **1980**, *45*, 531–534.
- [192] P. Maitre, D. Scuderi, D. Corinti, B. Chiavarino, M. E. Crestoni, S. Fornarini, *Chem. Rev.* **2019**, *120*, 3261–3295.

- [193] D. Boyall, K. L. Reid, *Chem. Soc. Rev.* **1997**, *26*, 223.
- [194] J. Oomens, J. D. Steill, B. Redlich, *J. Am. Chem. Soc.* **2009**, *131*, 4310–4319.
- [195] P. Nieto, A. Günther, G. Berden, J. Oomens, O. Dopfer, *J. Phys. Chem. A* **2016**, *120*, 8297–8308.
- [196] W. Hoffmann, M. Marianski, S. Warnke, J. Seo, C. Baldauf, G. von Helden, K. Pagel, *Phys. Chem. Chem. Phys.* **2016**, *18*, 19950–19954.
- [197] S. Osburn, J. D. Steill, J. Oomens, R. A. J. O'Hair, M. van Stipdonk, V. Ryzhov, *Chem. Eur. J.* **2010**, *17*, 873–879.
- [198] N. C. Polfer, J. Oomens, S. Suhai, B. Paizs, *J. Am. Chem. Soc.* **2005**, *127*, 17154–17155.
- [199] M. Lesslie, Y. Yang, A. J. Canty, E. Piacentino, F. Berthias, P. Maitre, V. Ryzhov, R. A. J. O'Hair, *Chem. Commun.* **2018**, *54*, 346–349.
- [200] O. Yeni, B. Schindler, B. Moge, I. Compagnon, *Analyst* **2022**, *147*, 312–317.
- [201] R. E. van Outersterp, J. Oosterhout, C. R. Gebhardt, G. Berden, U. F. H. Engelke, R. A. Wevers, F. Cuyckens, J. Oomens, J. Martens, *Anal. Chem.* **2023**, *95*, 3406–3413.
- [202] N. C. Polfer, *Chem. Soc. Rev.* **2011**, *40*, 2211.
- [203] S. Becher, G. Berden, J. Martens, J. Oomens, S. Heiles, *J. Am. Soc. Mass Spectrom.* **2021**, *32*, 2874–2884.
- [204] N. C. Polfer, J. J. Valle, D. T. Moore, J. Oomens, J. R. Eyler, B. Bendiak, *Anal. Chem.* **2005**, *78*, 670–679.
- [205] A. Filippi, C. Fraschetti, F. Rondino, S. Piccirillo, V. Steinmetz, L. Guidoni, M. Speranza, *Int. J. Mass Spectrom.* **2013**, *354-355*, 54–61.
- [206] W. H. Robertson, E. G. Diken, E. A. Price, J.-W. Shin, M. A. Johnson, *Science* **2003**, *299*, 1367–1372.
- [207] E. H. Perez, T. Schleif, J. P. Messinger, A. G. R. Buxó, O. C. Moss, K. Greis, M. A. Johnson, *J. Am. Soc. Mass Spectrom.* **2022**, *33*, 1914–1920.
- [208] P. Bansal, V. Yatsyna, A. H. AbiKhodr, S. Warnke, A. B. Faleh, N. Yalovenko, V. H. Wysocki, T. R. Rizzo, *Anal. Chem.* **2020**, *92*, 9079–9085.
- [209] G. L. Tripodi, M. M. J. Dekker, J. Roithová, L. Que, *Angew. Chem. Int. Ed.* **2021**, *60*, 7126–7131.
- [210] J. S. Prell, J. T. O'Brien, E. R. Williams, *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 800–809.
- [211] H. Schwarz, K. R. Asmis, *Chem. Eur. J.* **2019**, *25*, 2112–2126.
- [212] A. Tsybizova, E. Paenurk, V. Gorbachev, P. Chen, *J. Phys. Chem. A* **2020**, *124*, 8519–8528.
- [213] J. P. Toennies, A. F. Vilesov, *Annu. Rev. Phys. Chem.* **1998**, *49*, 1–41.

- [214] L. Landau, *Phys. Rev.* **1941**, *60*, 356–358.
- [215] P. L. Kapitza, *Phys. Rev.* **1941**, *60*, 354–355.
- [216] M. Hartmann, R. E. Miller, J. P. Toennies, A. Vilesov, *Phys. Rev. Lett.* **1995**, *75*, 1566–1569.
- [217] A. Lindinger, J. P. Toennies, A. F. Vilesov, *J. Chem Phys.* **1999**, *110*, 1429–1436.
- [218] U. Even, *EPJ Tech. Instrum.* **2015**, *2*, 17.
- [219] A. Scheidemann, J. P. Toennies, J. A. Northby, *Phys. Rev. Lett.* **1990**, *64*, 1899–1902.
- [220] F. Bierau, P. Kupser, G. Meijer, G. von Helden, *Phys. Rev. Lett.* **2010**, *105*, 133402.
- [221] L. G. Mendoza-Luna, N. M. K. Shiltagh, M. J. Watkins, N. Bonifaci, F. Aitken, K. von Haefen, *J. Phys. Chem. Lett.* **2016**, *7*, 4666–4670.
- [222] S. Goyal, D. L. Schutt, G. Scoles, *Phys. Rev. Lett.* **1992**, *69*, 933–936.
- [223] S. Smolarek, N. B. Brauer, W. J. Buma, M. Drabbels, *J. Am. Chem. Soc.* **2010**, *132*, 14086–14091.
- [224] X. Zhang, N. B. Brauer, G. Berden, A. M. Rijs, M. Drabbels, *J. Chem. Phys.* **2012**, *136*, 044305.
- [225] A. I. González Flórez, D.-S. Ahn, S. Gewinner, W. Schöllkopf, G. von Helden, *Phys. Chem. Chem. Phys.* **2015**, *17*, 21902–21911.
- [226] M. Bursch, J.-M. Mewes, A. Hansen, S. Grimme, *Angew. Chem. Int. Ed.* **2022**, *61*, e202205735.
- [227] E. Schrödinger, *Ann. Phys.* **1926**, *384*, 361–376.
- [228] F. Jensen, *Introduction to Computational Chemistry*, Third Edition, Wiley, **2017**, pp. 1–19.
- [229] M. Born, R. Oppenheimer, *Ann. Phys.* **1927**, *389*, 457–484.
- [230] D. R. Hartree, *Math. Proc. Camb. Phil. Soc.* **1928**, *24*, 89–110.
- [231] V. Fock, *Z. Phys.* **1930**, *61*, 126–148.
- [232] C. Møller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 618–622.
- [233] F. Jensen, *Introduction to Computational Chemistry*, Third Edition, Wiley, **2017**, pp. 124–187.
- [234] G. D. Purvis, R. J. Bartlett, *J. Chem. Phys.* **1982**, *76*, 1910–1918.
- [235] D. L. Strout, G. E. Scuseria, *J. Chem. Phys.* **1995**, *102*, 8448–8452.
- [236] F. Weigend, M. Häser, H. Patzelt, R. Ahlrichs, *Chem. Phys. Lett.* **1998**, *294*, 143–152.
- [237] C. Riplinger, B. Sandhoefer, A. Hansen, F. Neese, *J. Chem. Phys.* **2013**, *139*, 134101.
- [238] D. G. Liakos, F. Neese, *J. Chem. Theory Comput.* **2015**, *11*, 4054–4063.
- [239] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864–B871.

- [240] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *140*, A1133–A1138.
- [241] J. P. Perdew, K. Schmidt, *AIP Conference Proceedings* **2001**, *577*, 1.
- [242] J. P. Perdew, *Phys. Rev. Lett.* **1985**, *55*, 1665–1668.
- [243] E. I. Proynov, S. Sirois, D. R. Salahub, *Int. J. Quantum Chem.* **1997**, *64*, 427–446.
- [244] J. P. Perdew, S. Kurth, A. Zupan, P. Blaha, *Phys. Rev. Lett.* **1999**, *82*, 2544–2547.
- [245] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- [246] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.
- [247] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- [248] C. Adamo, V. Barone, *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- [249] Y. Zhao, B. J. Lynch, D. G. Truhlar, *J. Phys. Chem. A* **2004**, *108*, 4786–4791.
- [250] L. Goerigk, S. Grimme, *J. Chem. Theory Comput.* **2010**, *7*, 291–309.
- [251] L. Goerigk, S. Grimme, *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670.
- [252] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
- [253] E. Caldeweyher, C. Bannwarth, S. Grimme, *J. Chem. Phys.* **2017**, *147*, 034112.
- [254] A. Tkatchenko, M. Scheffler, *Phys. Rev. Lett.* **2009**, *102*, 073005.
- [255] J. G. Hill, *Int. J. Quantum Chem.* **2012**, *113*, 21–34.
- [256] F. Jensen, *WIREs Comput. Mol. Sci.* **2012**, *3*, 273–295.
- [257] W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **1972**, *56*, 2257–2261.
- [258] A. Schäfer, H. Horn, R. Ahlrichs, *J. Chem. Phys.* **1992**, *97*, 2571–2577.
- [259] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- [260] F. Jensen, *Introduction to Computational Chemistry*, Third Edition, Wiley, **2017**, pp. 188–232.
- [261] J. P. Merrick, D. Moran, L. Radom, *J. Phys. Chem. A* **2007**, *111*, 11683–11700.
- [262] M. D. Halls, J. Velkovski, H. B. Schlegel, *Theor. Chem. Acc.* **2001**, *105*, 413–421.
- [263] P. Sinha, S. E. Boesch, C. Gu, R. A. Wheeler, A. K. Wilson, *J. Phys. Chem. A* **2004**, *108*, 9213–9217.
- [264] V. Barone, *J. Chem. Phys.* **2005**, *122*, 014108.
- [265] V. Barone, J. Bloino, C. A. Guido, F. Lipparini, *Chem. Phys. Lett.* **2010**, *496*, 157–161.
- [266] J. Bloino, V. Barone, *J. Chem. Phys.* **2012**, *136*, 124108.

- [267] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Gaussian 16 Revision A.03, Gaussian Inc. Wallingford CT, **2016**.
- [268] K. Greis, C. Kirschbaum, M. I. Taccone, M. Götze, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, K. Pagel, *Angew. Chem. Int. Ed.* **2022**, *61*, e202115481.
- [269] A. Supady, V. Blum, C. Baldauf, *J. Chem. Inf. Model.* **2015**, *55*, 2338–2348.
- [270] P. Pracht, F. Bohle, S. Grimme, *Phys. Chem. Chem. Phys.* **2020**, *22*, 7169–7192.
- [271] S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 2847–2862.
- [272] W. Thiel, *WIREs Comput. Mol. Sci.* **2013**, *4*, 145–157.
- [273] M. Marianski, A. Supady, T. Ingram, M. Schneider, C. Baldauf, *J. Chem. Theory Comput.* **2016**, *12*, 6157–6168.
- [274] A. I. González Flórez, PhD thesis, Freie Universität Berlin, **2016**.
- [275] E. Mucha, PhD thesis, Radboud Universiteit Nijmegen, **2020**.
- [276] D. A. Thomas, R. Chang, E. Mucha, M. Lettow, K. Greis, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, *Phys. Chem. Chem. Phys.* **2020**, *22*, 18400–18413.
- [277] W. Schöllkopf, S. Gewinner, H. Junkes, A. Paarmann, G. von Helden, H. Bluem, A. M. M. Todd, *Proc. SPIE* **2015**, (Ed.: S. G. Biedron), 95121L.
- [278] D. Oepts, A. van der Meer, P. van Amersfoort, *Infrared Phys. Technol.* **1995**, *36*, 297–308.
- [279] R. Prazeres, J. Berset, R. Chaput, F. Glotin, D. Jaroszynski, J. Ortega, *Nucl. Instr. and Meth. B* **1994**, *89*, 54–59.
- [280] R. Dennington, T. A. Keith, J. M. Millam, GaussView Version 6, Semicem Inc. Shawnee Mission KS, **2019**.
- [281] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, G. R. Hutchison, *J. Cheminform.* **2012**, *4*, 17.
- [282] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

- [283] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, M. Scheffler, *Comput. Phys. Commun.* **2009**, *180*, 2175–2196.
- [284] F. Neese, *WIREs Comput. Mol. Sci.* **2012**, *2*, 73–78.
- [285] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.
- [286] S. Spicher, S. Grimme, *Angew. Chem. Int. Ed.* **2020**, *59*, 15665–15673.
- [287] C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, S. Grimme, *WIREs Comput. Mol. Sci.* **2020**, *11*, e1493.
- [288] T. Yanai, D. P. Tew, N. C. Handy, *Chem. Phys. Lett.* **2004**, *393*, 51–57.
- [289] F. Neese, *WIREs Comput. Mol. Sci.* **2022**, *12*, e1606.
- [290] K. Fukui, *Acc. Chem. Res.* **1981**, *14*, 363–368.
- [291] K. Greis, S. Lechnitz, C. Kirschbaum, C.-W. Chang, M.-H. Lin, G. Meijer, G. von Helden, P. H. Seeberger, K. Pagel, *J. Am. Chem. Soc.* **2022**, *144*, 20258–20266.
- [292] X. Zou, C. Qin, C. L. Pereira, G. Tian, J. Hu, P. H. Seeberger, J. Yin, *Chem. Eur. J.* **2018**, *24*, 2868–2872.
- [293] N. Paramonov, D. Bailey, M. Rangarajan, A. Hashim, G. Kelly, M. A. Curtis, E. F. Hounsell, *Eur. J. Biochem.* **2001**, *268*, 4698–4707.
- [294] K. Kappler, T. Hennet, *Genes Immun.* **2020**, *21*, 224–239.
- [295] C. Kirschbaum, K. Greis, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, K. Pagel, *Anal. Bioanal. Chem.* **2022**, *414*, 5275–5285.
- [296] Y. Zeng, Z. Wang, D. Whitfield, X. Huang, *J. Org. Chem.* **2008**, *73*, 7952–7962.
- [297] K. Greis, C. Kirschbaum, G. Fittolani, E. Mucha, R. Chang, G. von Helden, G. Meijer, M. Delbianco, P. H. Seeberger, K. Pagel, *Eur. J. Org. Chem.* **2022**, *2022*, e202200255.
- [298] Y. Yu, T. Tyrikos-Ergas, Y. Zhu, G. Fittolani, V. Bordoni, A. Singhal, R. J. Fair, A. Grafmüller, P. H. Seeberger, M. Delbianco, *Angew. Chem. Int. Ed.* **2019**, *58*, 13127–13132.
- [299] J. St-Gelais, M. Bouchard, V. Denavit, D. Giguère, *J. Org. Chem.* **2019**, *84*, 8509–8522.
- [300] J. St-Gelais, Côté, D. Lainé, P. A. Johnson, D. Giguère, *Chem. Eur. J.* **2020**, *26*, 13499–13506.
- [301] D. Lainé, O. Lessard, J. St-Gelais, D. Giguère, *Chem. Eur. J.* **2021**, *27*, 3799–3805.
- [302] A. Axer, R. P. Jumde, S. Adam, A. Faust, M. Schäfers, M. Fobker, J. Koehnke, A. K. H. Hirsch, R. Gilmour, *Chem. Sci.* **2021**, *12*, 1286–1294.
- [303] M. Delbianco, A. Kononov, A. Poveda, Y. Yu, T. Diercks, J. Jiménez-Barbero, P. H. Seeberger, *J. Am. Chem. Soc.* **2018**, *140*, 5421–5426.

- [304] A. Poveda, G. Fittolani, P. H. Seeberger, M. Delbianco, J. Jiménez-Barbero, *Front. Mol. Biosci.* **2021**, *8*, 784318.
- [305] K. Greis, C. Kirschbaum, S. Lechnitz, S. Gewinner, W. Schöllkopf, G. von Helden, G. Meijer, P. H. Seeberger, K. Pagel, *Org. Lett.* **2020**, *22*, 8916–8919.
- [306] E. Fischer, *Ber. Dtsch. Chem. Ges.* **1914**, *47*, 196–210.
- [307] R. J. Ferrier, W. G. Overend, A. E. Ryan, *J. Chem. Soc.* **1962**, 3667.
- [308] R. J. Ferrier, O. A. Zubkov, *Org. React.* **2003**, 569–736.
- [309] C. S. Bennett, M. C. Galan, *Chem. Rev.* **2018**, *118*, 7931–7985.
- [310] É. Bokor, S. Kun, D. Goyard, M. Tóth, J.-P. Praly, S. Vidal, L. Somsák, *Chem. Rev.* **2017**, *117*, 1687–1764.
- [311] B. Fraser-Reid, J. Lopez, *Curr. Org. Chem.* **2009**, *13*, 532–553.
- [312] K. Greis, C. E. Griesbach, C. Kirschbaum, G. Meijer, G. von Helden, K. Pagel, M. W. Peczu, *J. Org. Chem.* **2023**, *88*, 5543–5553.
- [313] R. Vannam, M. W. Peczu, *Org. Biomol. Chem.* **2016**, *14*, 3989–3996.
- [314] S. D. Markad, S. Xia, N. L. Snyder, B. Surana, M. D. Morton, C. M. Hadad, M. W. Peczu, *J. Org. Chem.* **2008**, *73*, 6341–6354.
- [315] C. E. Griesbach, M. W. Peczu, *Org. Biomol. Chem.* **2021**, *19*, 10635–10646.
- [316] H. Lönnberg, *Org. Biomol. Chem.* **2011**, *9*, 1687.
- [317] G. M. Emilsson, S. Nakamura, A. Roth, R. R. Breaker, *RNA* **2003**, *9*, 907–918.
- [318] S. Pääbo, H. Poinar, D. Serre, V. Jaenicke-Després, J. Hebler, N. Rohland, M. Kuch, J. Krause, L. Vigilant, M. Hofreiter, *Annu. Rev. Genet.* **2004**, *38*, 645–679.
- [319] S. A. McLuckey, G. J. V. Berkel, G. L. Glish, *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60–70.
- [320] T. yi Huang, A. Kharlamova, J. Liu, S. A. McLuckey, *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1832–1840.
- [321] T. E. Andersen, F. Kirpekar, K. F. Haselmann, *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1353–1368.
- [322] C. Riml, H. Glasner, M. T. Rodgers, R. Micura, K. Breuker, *Nucleic Acids Res.* **2015**, *43*, 5171–5181.
- [323] E. Fuchs, C. Falschlunger, R. Micura, K. Breuker, *Nucleic Acids Res.* **2019**, *47*, 7223–7234.
- [324] J. M. Tromp, S. Schürch, *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1262–1268.
- [325] J. Florián, J. Leszczyński, *J. Am. Chem. Soc.* **1996**, *118*, 3010–3017.

- [326] K. Greis, C. Kirschbaum, M. I. Taccone, A. Y. Torres-Boy, K. Ober, G. Meijer, K. Pagel, G. von Helden, **2023**, *in preparation*.
- [327] R. Breslow, *Acc. Chem. Res.* **1973**, *6*, 393–398.
- [328] N. C. Baird, *J. Am. Chem. Soc.* **1972**, *94*, 4941–4948.
- [329] H. J. Wörner, F. Merkt, *J. Chem. Phys.* **2007**, *127*, 034303.
- [330] W. Zou, M. Filatov, D. Cremer, *Int. J. Mass Spectrom.* **2012**, *112*, 3277–3288.
- [331] H. Jiao, P. von Ragué Schleyer, Y. Mo, M. A. McAllister, T. T. Tidwell, *J. Am. Chem. Soc.* **1997**, *119*, 7075–7083.
- [332] W. A. Y. R. Breslow, H. W. Chang, *J. Am. Chem. Soc.* **1963**, *85*, 2033–2034.
- [333] R. Breslow, H. W. Chang, R. Hill, E. Wasserman, *J. Am. Chem. Soc.* **1967**, *89*, 1112–1119.
- [334] M. Saunders, R. Berger, A. Jaffe, J. M. McBride, J. O'Neill, R. Breslow, J. M. Hoffmann, C. Perchonock, E. Wasserman, et al., *J. Am. Chem. Soc.* **1973**, *95*, 3017–3018.
- [335] T. L. Amyes, J. P. Richard, M. Novak, *J. Am. Chem. Soc.* **1992**, *114*, 8032–8041.
- [336] D. Duvinage, S. Mebs, J. Beckmann, *Chem. Eur. J.* **2021**, *27*, 8105–8109.
- [337] C. Abdellaoui, V. Hermanns, M. Reinfelds, M. Scheurer, A. Dreuw, A. Heckel, J. Wachtveitl, *Phys. Chem. Chem. Phys.* **2022**, *24*, 5294–5300.
- [338] J. Wang, J. Kubicki, E. F. Hilinski, S. L. Mecklenburg, T. L. Gustafson, M. S. Platz, *J. Am. Chem. Soc.* **2007**, *129*, 13683–13690.
- [339] S. L. Mecklenburg, E. F. Hilinski, *J. Am. Chem. Soc.* **1989**, *111*, 5471–5472.
- [340] M. A. O'Neill, F. L. Cozens, N. P. Schepp, *Tetrahedron* **2000**, *56*, 6969–6977.
- [341] J. Oomens, G. Meijer, G. von Helden, *J. Phys. Chem. A* **2001**, *105*, 8302–8309.
- [342] J. Fulara, A. Chakraborty, J. P. Maier, *Angew. Chem. Int. Ed.* **2016**, *55*, 3424–3427.
- [343] P. Costa, I. Trosien, M. Fernandez-Oliva, E. Sanchez-Garcia, W. Sander, *Angew. Chem. Int. Ed.* **2015**, *54*, 2656–2660.
- [344] D. A. Thomas, E. Mucha, M. Lettow, G. Meijer, M. Rossi, G. von Helden, *J. Am. Chem. Soc.* **2019**, *141*, 5815–5823.
- [345] A. I. González Flórez, E. Mucha, D.-S. Ahn, S. Gewinner, W. Schöllkopf, K. Pagel, G. von Helden, *Angew. Chem. Int. Ed.* **2016**, *55*, 3295–3299.
- [346] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- [347] E. Mucha, D. Thomas, M. Lettow, G. Meijer, K. Pagel, G. von Helden in *Top. Appl. Phys.* Springer International Publishing, **2022**, pp. 241–280.

- [348] K. Xu, Q. Man, Y. Zhang, J. Guo, Y. Liu, Z. Fu, Y. Zhu, Y. Li, M. Zheng, N. Ding, *Org. Chem. Front.* **2020**, *7*, 1606–1615.
- [349] S. Warnke, A. B. Faleh, T. R. Rizzo, *ACS Meas. Sci. Au* **2021**, *1*, 157–164.

Appendix A

Influence of the Electron Density of Acyl Protecting Groups on the Selectivity of Galactosylations

This appendix contains the supporting information of the publication related to Chapter 4. It contains mass spectra, energetics, energy hierarchies, energy diagrams, and 3D-structures of pivaloylated and trifluoroacetylated galactosyl cations. Furthermore, the synthesis of the building blocks and the glycosylation reactions are described in this document and it contains the data for characterization of the synthesized compounds. The publication, the supporting information, and xyz-coordinates of the computed structures can be found online:

<https://doi.org/10.1021/jacs.2c05859>.^[291]

Supplementary Information

The Influence of the Electron Density in Acyl Protecting Groups on the Selectivity of Galactose Formation

Kim Greis,^{a,b†} Sabrina Lechnitz,^{a,c†} Carla Kirschbaum,^{a,b} Chun-Wei Chang,^a Mei-Huei Lin,^a
Gerard Meijer,^b Gert von Helden,^b Peter H. Seeberger,^{*a,c} Kevin Pagel^{*a,b}

(a) Institut für Chemie und Biochemie, Freie Universität Berlin, Arnimallee 22, 14195 Berlin, Germany

(b) Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

(c) Max-Planck-Institut für Kolloid- und Grenzflächenforschung, Am Mühlenberg 1, 14476 Potsdam, Germany

† K.G. and S.L. contributed equally.

Correspondence to: Prof. Dr. Kevin Pagel
kevin.pagel@fu-berlin.de

Prof. Dr. Peter Seeberger
peter.seeberger@mpikg.mpg.de

1	Mass Spectrometry and Infrared Spectroscopy.....	3
1.1	Experimental Setup	3
1.2	Mass Spectra	5
2	Computational Methods	7
2.1	Method Description	7
2.2	Energetics	9
2.3	Energy Hierarchies	15
2.4	Energy Diagrams	18
2.5	3D Structures	19
2.6	xyz-Coordinates of reoptimized structures	24
3	General Information	25
4	Materials and Conditions for Automated Synthesis.....	25
4.1	Materials and Measurements	25
4.2	Preparation of Stock Solutions	26
4.3	Modules for Automated Synthesis	26
4.4	Post-automated Synthesis Manipulations, Analysis and Purification	27
4.5	General Procedure for Glycosylations	28
4.6	Determination of Alpha/Beta Ratios and Purification	29
5	Synthesis of building blocks.....	30
6	Determination of alpha/beta-ratios	68
7	Automated Glycan Assembly of Building Block 1	147
8	Literature.....	151

1 Mass Spectrometry and Infrared Spectroscopy

1.1 Experimental Setup

The precursors were dissolved in a 9:1 (V:V) mixture of acetonitrile and water to yield 0.1 mM solutions. Pd/Pt coated glass capillaries (Sputter Coater HR 208, *Cressington*) for nano electrospray ionization (nESI) are pulled to a tip with an inner diameter of 1–2 μm using a micropipette puller (Model P-1000, *Sutter Instrument*).

Glycosyl cations were generated and probed using a custom-built helium droplet instrument (Figure S1). Glycosyl cations are formed after nESI (*Z*-spray) with a voltage of 1 kV to the tip of the capillary of the precursors, followed by in-source fragmentation of the generated ions. Commonly, nESI of the precursor leads to sodiated and protonated ions, however, labile leaving groups, such as S*Et*, can be cleaved by in-source fragmentation.

After passing through two ring-electrode ion guides, the ions of interest are mass-to-charge selected by a quadrupole mass filter. Then, the ions enter a quadrupole bender. If no voltage is applied, the ions directly pass through the bender to get to a time-of-flight detector to record mass spectra (Figures S2 and S3) and to monitor the ion signal. If a voltage is applied to the quadrupole bender, the ions are bent and enter a hexapole ion trap that is cooled to 90 K by liquid nitrogen in this experiment. The ions of interest are subsequently accumulated in the ion trap and thermalized by collisions with helium buffer gas.

Expansion of pressurized helium into the vacuum by a pulsed Even-Lavie valve leads to the formation of a beam of superfluid helium nanodroplets (0.4 K) that is traversing the ion trap, picking up ions, rapidly cool them, and guiding them to the detection region. Here, an infrared (IR) beam generated by the Fritz Haber Institute free-electron laser (FHI FEL¹) overlaps with the ion beam. Upon the absorption of resonant photons, vibrational modes of the molecular ions are excited. The ions dissipate the energy to the helium matrix to get back to their ground state. After the absorption of multiple photons, the probed ions are release from the helium nanodroplets and detected by a time-of-flight detector. The ion yield can be plotted as a function of the IR wavenumber, leading to an IR spectrum (Figures 2 and 3). Due to the multiphoton absorption process, the intensities in the obtained IR spectrum do not scale linearly. As a first-order correction, the ion signal is divided by the energy of the IR macropulse.

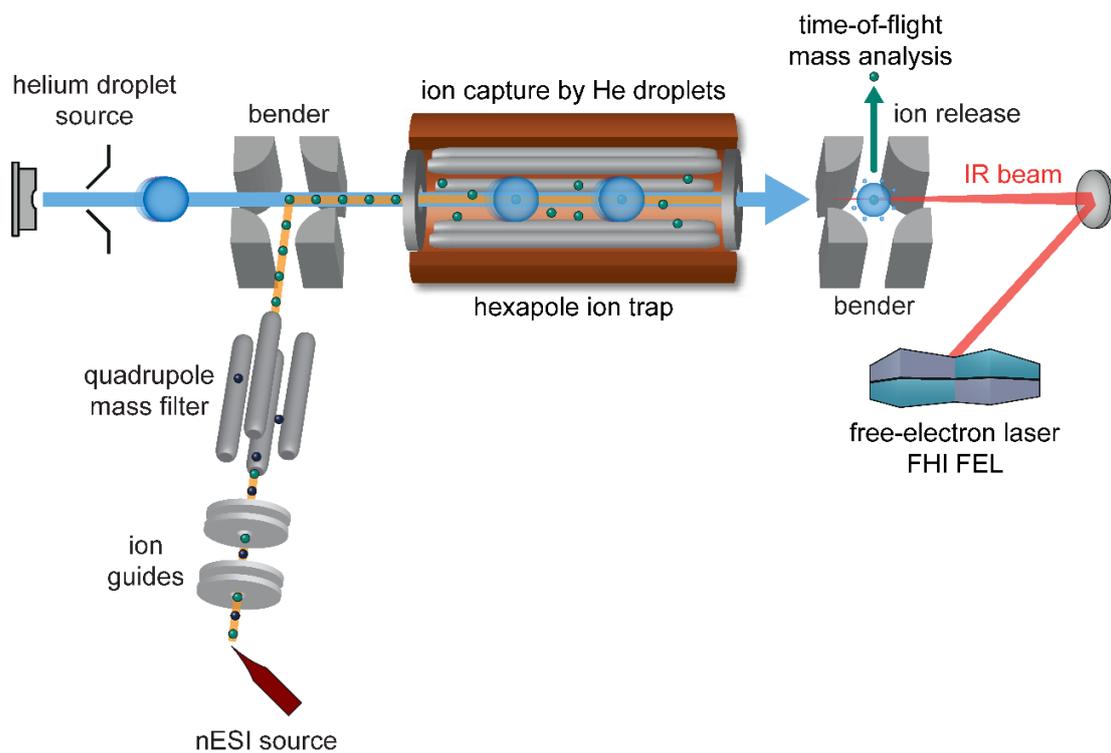


Figure S1. Schematic drawing of the custom-built helium droplet instrument combining mass spectrometry and infrared spectroscopy to probe mass-to-charge selected ions.

1.2 Mass Spectra

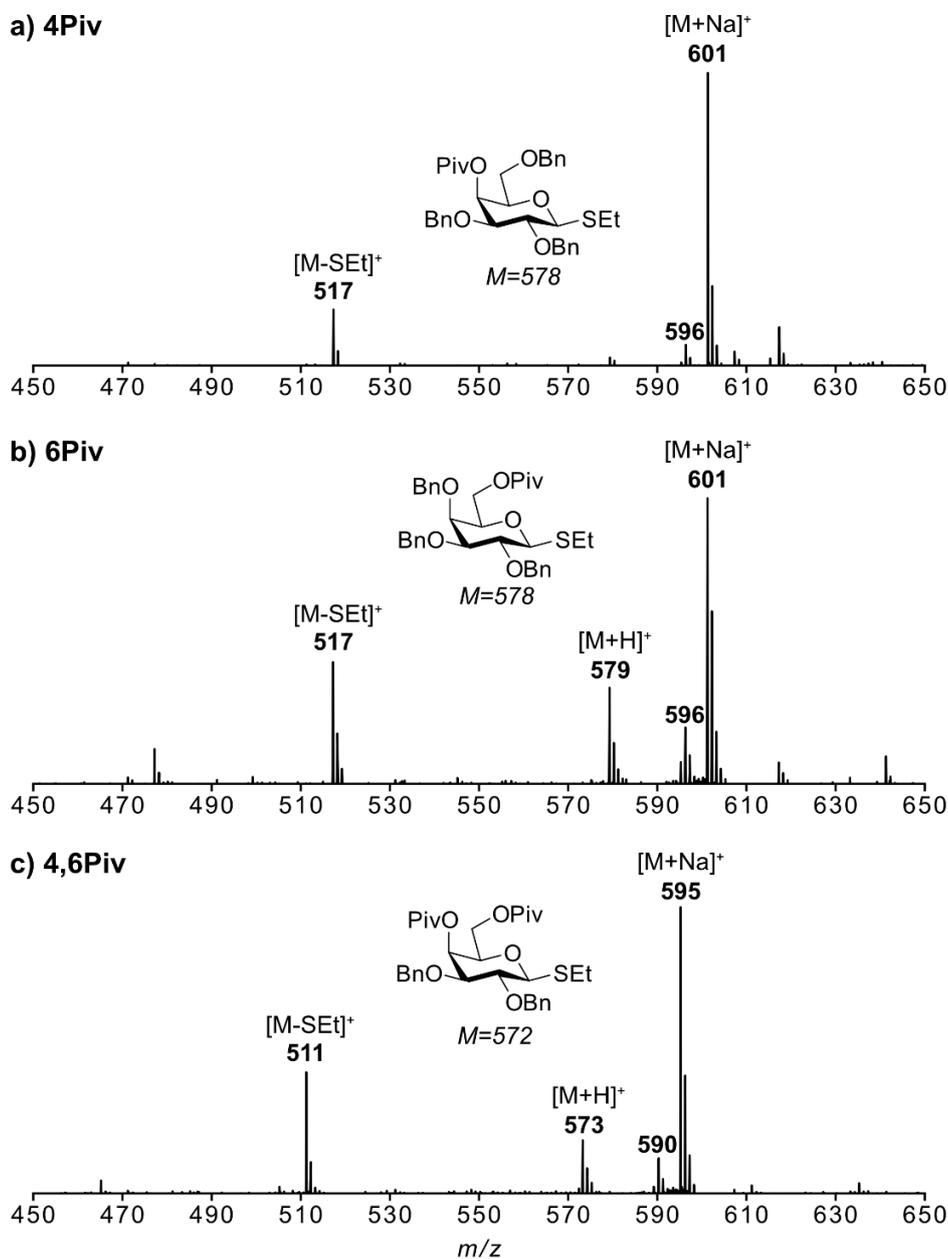


Figure S2. Mass spectra of β -thioethyl (SEt) precursors of (a) 2,3,6-tri-*O*-benzyl-4-*O*-pivaloyl-D-galactopyranosyl (**4Piv**), (b) 2,3,4-tri-*O*-benzyl-6-*O*-pivaloyl-D-galactopyranosyl (**6Piv**), and (c) 2,3-di-*O*-benzyl-4,6-di-*O*-pivaloyl-D-galactopyranosyl cations (**4,6Piv**) recorded on the helium droplet instrument. In-source fragmentation of precursor ions $[M+H]^+$ ($m/z = 579$ and 573), $[M+NH_4]^+$ ($m/z = 596$ and 590), and $[M+Na]^+$ ($m/z = 601$ and 595) leads to galactosyl cations ($m/z = 517$ and 511).

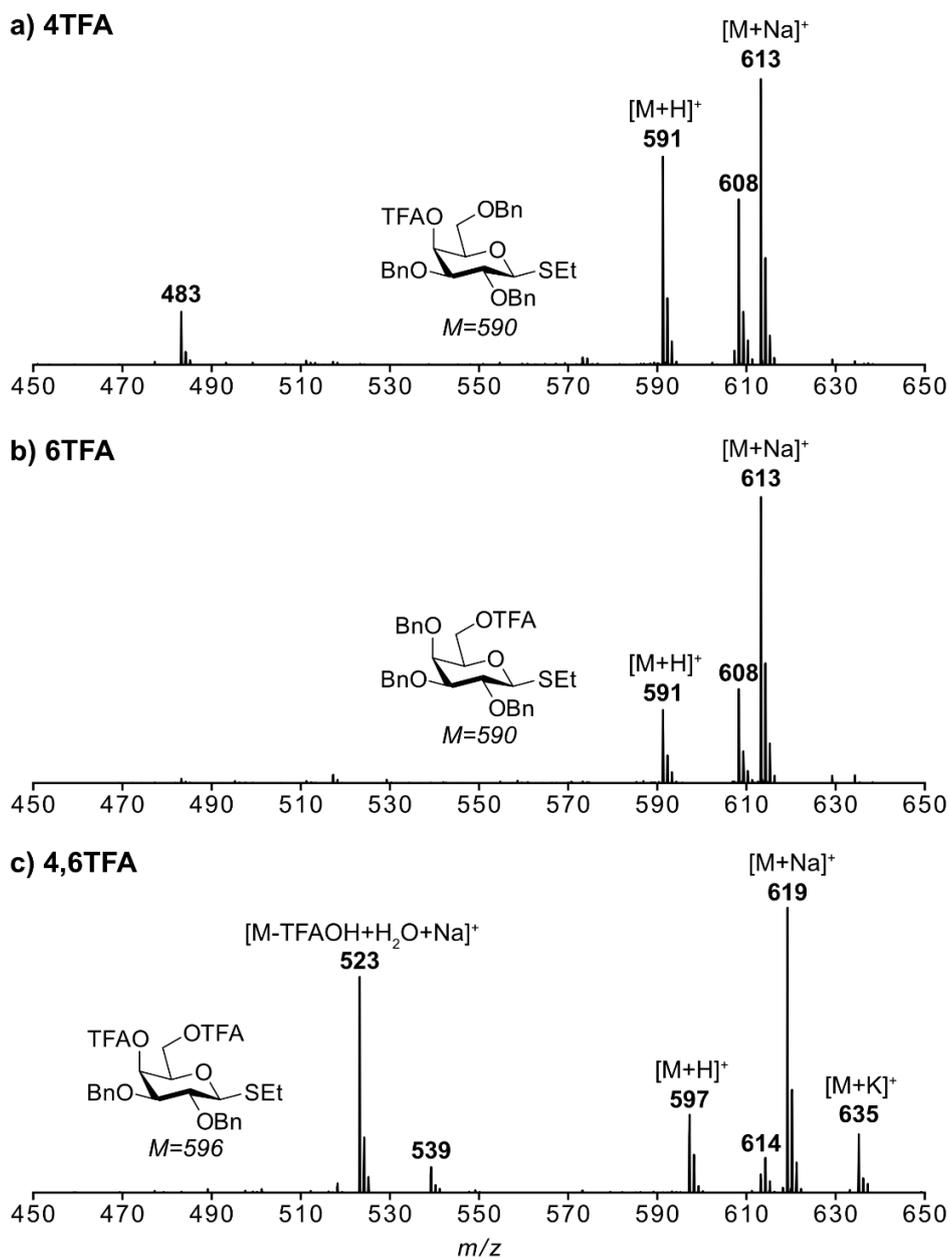


Figure S3. Mass spectra of (a) 2,3,6-tri-*O*-benzyl-4-*O*-trifluoroacetyl-D-galactopyranoside (**4TFA**), (b) 2,3,4-tri-*O*-benzyl-6-*O*-trifluoroacetyl-D-galactopyranoside (**6TFA**), and (c) 2,3-di-*O*-benzyl-4,6-di-*O*-trifluoroacetyl-D-galactopyranoside (**4,6TFA**) carrying a β -thioethyl (SEt) leaving group, recorded on the helium droplet instrument. In-source fragmentation conditions do not lead to fragmentation of the precursor ions $[M+H]^+$ (m/z = 591 and 597), $[M+NH_4]^+$ (m/z = 608 and 614), and $[M+Na]^+$ (m/z = 613 and 619). The **4,6TFA** building block partially hydrolyses, leading to the sodiated and potassiated molecular ions at m/z = 523 and 539.

2 Computational Methods

2.1 Method Description

The genetic algorithm (GA) FAFOOM² was used to sample the conformational space of intact **4Piv**, **6Piv**, **4,6Piv**, **4TFA**, **6TFA**, **4,6TFA** galactosyl cations. Furthermore, for **4Piv**, **6Piv**, and **4,6Piv** also the conformational space of rearranged galactosyl cations was sampled. With the GA all rotatable bonds and pyranose puckers can be sampled. After its generation, each structure is sent to the external software FHI-aims³ (version 171221) for local DFT geometry optimization at the dispersion corrected PBE+vdW^{TS4-5} level of theory using *light* basis set settings for all atoms. For each galactosyl cation and their rearranged counterparts multiple separate GA runs were carried using the settings in Table S1. The number of generated structures is shown in Table S2 and the energy hierarchies are represented in Figures S4-S6. The GA sampling of the intact galactosyl cations yielded structures with five different modes of participation: dioxolenium-type structures exhibiting remote participation either from the C4- or the C6-acyl protecting group, oxonium-type structures exhibiting *non-classical* remote participation either from the C4- or the C6-benzyl protecting group, and oxocarbenium-type structures in which no participation occurs (Figures S8-S13).

Table S1. GA parameters used in initial search of intact galactosyl cations. For rearranged galactosyl cations the function to sample the pyranose pucker was switched off.

	Parameter	Value
Molecule	Distance_cutoff_1	1.2
	Distance_cutoff_2	2.15
	Rmsd_cutoff_uniq	0.25
GA settings	Popsize	10
	Prob_for_crossing	0.95
	Prob_for_mut_pyranosering	0.6
	Prob_for_mut_torsion	0.8
	Fitness_sum_limit	1.2
	Selection	Roulette wheel
	Max_mutations_torsion	3
	Max_mutations_pyranosering	1

From all the structures generated by the GA, several distinct low-energy structures for each mode of participation, specified in Table S2, were selected for reoptimization and harmonic frequency calculation at the PBE0+D3/6-311+G(d,p)⁶⁻⁸ level of theory in Gaussian 16, Revision A.03⁹ using default settings. The energies including the zero-point vibrational energy (ZPVE) and free energies at 90 K (the temperature in the hexapole ion trap) of the reoptimized structures are shown in Tables S3-S8 and free energy hierarchies in

Figures S4-S6. The computed IR spectra, generated from harmonic frequency calculations, were normalized and scaled by an empirical factor of 0.965.

In a last step, the structures reoptimized at the PBE0+D3/6-311+G(d,p) level of theory were used for single-point energy calculations at the DLPNO-CCSD(T)/Def2-TZVPP¹⁰⁻¹¹ level of theory as implemented in ORCA 5.0.3.¹² The relative single-point energies are reported in Tables S3-S8 as such and with the ZPVE plus free energy correction at 90 K derived from previous calculations at the PBE0 level of theory.

Table S2. Number of generated structures during GA sampling and reoptimized structures.

Galactosyl Cations	#(GA Structures)	#(Reoptimized Structures)						
		dioxolenium		rearranged		oxonium		oxocarbenium
		C4	C6	C4	C6	C4	C6	
4Piv intact	599	10	–	–	–	–	5	5
4Piv rearranged	188	–	–	5	–	–	–	–
6Piv intact	548	–	10	–	–	5	–	5
6Piv rearranged	205	–	–	–	5	–	–	–
4,6Piv intact	673	10	5	–	–	–	–	5
4,6Piv rearranged	176	–	–	5	–	–	–	–
4,6Piv rearranged	162	–	–	–	5	–	–	–
4TFA intact	561	11	–	–	–	–	5	5
6TFA intact	533	–	3	–	–	5	–	5
4,6TFA intact	493	10	3	–	–	–	–	6

2.2 Energetics

Table S3. List of structures of the **4Piv** glycosyl cation reoptimized at the PBE0+D3/6-311+G(d,p) level of theory. Ring puckers, bond distances between the carbonyl oxygen of the C4-pivaloyl group and the anomeric carbon (C1), energies (ΔE , including zero-point-vibrational energy) and free energies (ΔF) at 90 K are assigned to each structure. Furthermore, single-point energies of these reoptimized structures at the DLPNO-CCSD(T)/Def2-TZVPP level of theory are indicated as such (ΔE , not including ZPVE) and with ZPVE plus free energy correction at 90 K (ΔF) derived from the PBE0 calculation. The infrared spectra of the structures labelled with an asterisk are represented in the manuscript. In the last six rows, the energetics of transition states connected to minima via intrinsic reaction coordinate (IRC) calculations are shown.

ID	Ring Pucker	d(C4=O—C1) [Å]	ΔE (PBE0+D3) [kJ mol ⁻¹]	ΔF (PBE0+D3) [kJ mol ⁻¹]	ΔE (DLPNO-CCSD(T)) [kJ mol ⁻¹]	ΔF (DLPNO-CCSD(T)) [kJ mol ⁻¹]
C4_dioxolenium/conf_00 (*)	1S5	1.52	0.00	0.00	0.00	0.05
C4_dioxolenium/conf_01	1S5	1.52	3.28	2.78	3.19	2.10
C4_dioxolenium/conf_02	1S5	1.52	3.07	1.80	4.29	2.25
C4_dioxolenium/conf_03	1S5	1.53	9.09	8.64	8.20	6.97
C4_dioxolenium/conf_04	1S5	1.52	9.36	7.69	9.08	6.87
C4_dioxolenium/conf_05	1S5	1.53	9.67	8.20	9.52	7.15
C4_dioxolenium/conf_06 (*)	1S5	1.51	5.76	4.22	5.64	2.56
C4_dioxolenium/conf_07	1S5	1.52	9.34	7.20	9.13	6.44
C4_dioxolenium/conf_08	1S5	1.54	11.19	8.40	11.93	7.62
C4_dioxolenium/conf_09	1S5	1.51	16.65	12.89	17.12	10.90
C4_rearranged/conf_00 (*)	–	5.93	12.71	8.84	13.44	0.00
C4_rearranged/conf_01	–	4.85	21.67	18.61	16.68	4.55
C4_rearranged/conf_02	–	4.93	27.51	23.14	26.06	12.83
C4_rearranged/conf_03	–	5.20	25.09	21.55	19.44	6.23
C4_rearranged/conf_04	–	4.96	31.85	27.36	27.70	13.52
C6_oxonium/conf_00 (*)	1C4	5.21	29.77	27.66	24.54	20.57
C6_oxonium/conf_01	1C4	5.19	31.26	30.20	25.46	22.45
C6_oxonium/conf_02	1C4	5.20	32.25	31.01	27.04	24.44
C6_oxonium/conf_03	1C4	5.19	31.27	30.22	25.46	22.47
C6_oxonium/conf_04	1C4	5.20	32.25	30.99	27.05	24.43
oxocarbenium/conf_00 (*)	4H3	5.20	40.13	37.80	50.01	40.34
oxocarbenium/conf_01	4H3	5.24	49.24	47.67	59.44	50.64
oxocarbenium/conf_02	E3	5.14	52.17	50.34	61.53	51.00
oxocarbenium/conf_03	4H3	5.16	49.25	44.04	62.48	48.75
oxocarbenium/conf_04	3H4	5.06	54.93	51.64	62.43	51.88
oxocarbenium/conf_IRC1	4H3	2.62	51.98	49.68	60.69	52.21
TS1	4E	2.33	52.65	52.16	62.72	55.80
C4_dioxolenium/conf_IRC1	1,4B	1.55	34.07	33.17	35.67	33.94
oxocarbenium/conf_IRC2	4H3	2.62	51.98	49.68	60.69	52.21
TS2	–	2.91	177.76	176.58	210.86	189.59
C4_rearranged/conf_IRC2	–	3.68	24.54	21.89	23.33	13.33

Table S4. List of structures of the **6Piv** glycosyl cation reoptimized at the PBE0+D3/6-311+G(d,p) level of theory. Ring puckers, bond distances between the carbonyl oxygen of the C6-pivaloyl group and the anomeric carbon (C1), energies (ΔE , including zero-point-vibrational energy) and free energies (ΔF) at 90 K are assigned to each structure. Furthermore, single-point energies of these reoptimized structures at the DLPNO-CCSD(T)/Def2-TZVPP level of theory are indicated as such (ΔE , not including ZPVE) and with ZPVE plus free energy correction at 90 K (ΔF) derived from the PBE0 calculation. The infrared spectra of the structures labelled with an asterisk are represented in the manuscript. In the last six rows, the energetics of transition states connected to minima via intrinsic reaction coordinate (IRC) calculations are shown.

ID	Ring Pucker	d(C4=O—C1) [Å]	ΔE (PBE0+D3) [kJ mol ⁻¹]	ΔF (PBE0+D3) [kJ mol ⁻¹]	ΔE (DLPNO-CCSD(T)) [kJ mol ⁻¹]	ΔF (DLPNO-CCSD(T)) [kJ mol ⁻¹]
C6_dioxolenium/conf_00	1C4	1.50	2.79	1.30	6.58	17.30
C6_dioxolenium/conf_01	1C4	1.50	2.87	3.73	4.51	17.70
C6_dioxolenium/conf_02 (*)	1C4	1.50	0.00	0.00	3.41	16.20
C6_dioxolenium/conf_03	1C4	1.49	8.16	8.35	10.56	22.87
C6_dioxolenium/conf_04	1C4	1.49	6.69	5.82	8.96	20.61
C6_dioxolenium/conf_05	1C4	1.51	9.50	8.31	12.13	23.69
C6_dioxolenium/conf_06	1C4	1.49	6.69	5.82	8.96	20.62
C6_dioxolenium/conf_07	1C4	1.49	6.60	5.88	8.75	20.35
C6_dioxolenium/conf_08	1C4	1.50	11.02	9.60	12.81	22.99
C6_dioxolenium/conf_09	1C4	1.50	3.44	3.39	6.19	19.92
C6_rearranged/conf_00	–	4.96	6.49	2.74	5.78	6.46
C6_rearranged/conf_01 (*)	–	4.68	6.56	2.11	2.89	2.05
C6_rearranged/conf_02 (*)	–	4.81	5.81	2.19	0.00	0.00
C6_rearranged/conf_03	–	5.30	10.80	5.66	8.47	5.85
C6_rearranged/conf_04	–	3.94	11.28	8.28	6.47	8.27
C4_oxonium/conf_00 (*)	1,4B	4.20	29.72	28.74	5.78	6.46
C4_oxonium/conf_01	1,4B	3.87	30.26	28.84	27.38	35.35
C4_oxonium/conf_02	1,4B	4.91	48.12	46.29	47.67	53.50
C4_oxonium/conf_03	1,4B	4.76	51.54	49.57	50.96	56.74
C4_oxonium/conf_04	1,4B	3.84	50.14	46.17	47.37	50.62
oxocarbenium/conf_00 (*)	4E	5.04	23.40	21.99	32.72	38.14
oxocarbenium/conf_01	4H3	4.07	33.59	28.72	45.83	45.63
oxocarbenium/conf_02	4H3	3.96	36.33	33.29	46.21	48.24
oxocarbenium/conf_03	3E	3.56	37.68	34.01	46.97	48.65
oxocarbenium/conf_04	3E	3.56	37.68	34.02	46.97	48.66
oxocarbenium/conf_IRC1	4E	4.97	23.52	22.08	32.72	38.01
TS1	E5	1.96	171.70	172.64	185.11	192.54
C6_dioxolenium/conf_IRC1	OS2	1.51	137.11	136.18	138.16	149.67
oxocarbenium/conf_IRC2	4E	5.04	23.29	22.00	33.41	38.83
TS2	–	4.81	81.38	80.53	98.36	98.23
C6_rearranged/conf_IRC2	–	4.82	41.13	38.54	40.86	42.98

Table S5. List of structures of the **4,6Piv** glycosyl cation reoptimized at the PBE0+D3/6-311+G(d,p) level of theory. Ring puckers, bond distances between the carbonyl oxygen of the C6-pivaloyl group and the anomeric carbon (C1), energies (ΔE , including zero-point-vibrational energy) and free energies (ΔF) at 90 K are assigned to each structure. Furthermore, single-point energies of these reoptimized structures at the DLPNO-CCSD(T)/Def2-TZVPP level of theory are indicated as such (ΔE , not including ZPVE) and with ZPVE plus free energy correction at 90 K (ΔF) derived from the PBE0 calculation. The infrared spectra of the structures labelled with an asterisk are represented in the manuscript. In the last six rows, the energetics of transition states connected to minima via intrinsic reaction coordinate (IRC) calculations are shown.

ID	Ring Pucker	d(C4=O—C1) [Å]	ΔE (PBE0+D3) [kJ mol ⁻¹]	ΔF (PBE0+D3) [kJ mol ⁻¹]	ΔE (DLPNO-CCSD(T)) [kJ mol ⁻¹]	ΔF (DLPNO-CCSD(T)) [kJ mol ⁻¹]
C4_dioxolenium/conf_00	1S5	1.52	0.00	0.00	0.00	0.14
C4_dioxolenium/conf_01 (*)	1S5	1.51	0.77	0.08	1.08	0.00
C4_dioxolenium/conf_02	1S5	1.51	4.30	5.62	4.57	6.15
C4_dioxolenium/conf_03	1S5	1.52	6.59	6.16	5.53	4.72
C4_dioxolenium/conf_04	1S5	1.52	5.10	5.95	6.72	7.46
C4_dioxolenium/conf_05 (*)	1S5	1.51	4.08	5.29	0.61	3.49
C4_dioxolenium/conf_06	1S5	1.52	4.78	5.76	2.91	5.17
C4_dioxolenium/conf_07	1S5	1.51	4.07	5.29	0.61	3.49
C4_dioxolenium/conf_08	1S3	1.51	18.71	18.07	21.40	20.51
C4_dioxolenium/conf_09	1S5	1.51	7.72	7.57	7.25	7.11
C6_dioxolenium/conf_00 (*)	1C4	5.21	21.42	21.83	23.13	23.37
C6_dioxolenium/conf_01	1C4	5.07	24.77	24.99	26.72	26.63
C6_dioxolenium/conf_02	1C4	5.22	23.66	23.17	26.08	25.98
C6_dioxolenium/conf_03	1C4	5.10	25.43	25.31	27.40	27.56
C6_dioxolenium/conf_04	BO,3	5.21	39.12	40.85	45.04	46.68
C4_rearranged/conf_00 (*)	–	4.34	18.22	17.29	15.93	8.07
C4_rearranged/conf_01	–	4.88	21.91	18.57	19.68	8.14
C4_rearranged/conf_02	–	4.92	23.34	19.40	22.89	10.45
C4_rearranged/conf_03	–	5.27	22.61	19.77	24.09	11.52
C4_rearranged/conf_04	–	5.28	25.41	22.42	25.10	13.18
C6_rearranged/conf_00 (*)	–	5.94	16.43	14.66	11.52	2.36
C6_rearranged/conf_01	–	4.56	15.64	13.01	13.44	3.03
C6_rearranged/conf_02	–	5.17	18.71	17.09	13.59	5.38
C6_rearranged/conf_03	–	5.34	23.84	21.50	19.48	9.94
C6_rearranged/conf_04	–	4.78	25.03	24.21	23.05	14.77
oxocarbenium/conf_00	4H3	5.21	57.76	55.21	71.09	57.95
oxocarbenium/conf_01	4H3	5.19	62.51	59.42	75.62	61.48
oxocarbenium/conf_02	4E	5.18	54.95	52.26	65.56	54.82
oxocarbenium/conf_03 (*)	3E	5.06	52.39	49.48	60.51	50.59
oxocarbenium/conf_04	4H3	5.17	59.40	56.25	71.51	57.66
oxocarbenium/conf_IRC1	4H3	5.21	57.76	55.21	71.11	57.98
TS1	4E	3.73	107.18	104.65	123.75	109.01
C4_dioxolenium/conf_IRC1	1,4B	1.52	13.92	14.87	17.66	17.94
oxocarbenium/conf_IRC2	4H3	5.16	60.75	57.85	73.74	60.24
TS2	–	5.34	91.32	89.40	109.91	93.14
C6_rearranged/conf_IRC2	–	5.32	18.84	15.51	17.02	5.17

Table S6. List of structures of the **4TFA** glycosyl cation reoptimized at the PBE0+D3/6-311+G(d,p) level of theory. Ring puckers, bond distances between the carbonyl oxygen of the C4-pivaloyl group and the anomeric carbon (C1), energies (ΔE , including zero-point-vibrational energy) and free energies (ΔF) at 90 K are assigned to each structure. Furthermore, single-point energies of these reoptimized structures at the DLPNO-CCSD(T)/Def2-TZVPP level of theory are indicated as such (ΔE , not including ZPVE) and with ZPVE plus free energy correction at 90 K (ΔF) derived from the PBE0 calculation.

ID	Ring Pucker	d(C4=O—C1) [Å]	ΔE (PBE0+D3) [kJ mol ⁻¹]	ΔF (PBE0+D3) [kJ mol ⁻¹]	ΔE (DLPNO-CCSD(T)) [kJ mol ⁻¹]	ΔF (DLPNO-CCSD(T)) [kJ mol ⁻¹]
C4_dioxolenium/conf_00	1S5	1.61	18.26	18.17	25.39	24.45
C4_dioxolenium/conf_01	1S5	1.60	26.09	26.40	33.13	32.16
C4_dioxolenium/conf_02	1S5	1.61	26.85	28.97	35.29	35.83
C4_dioxolenium/conf_03	1S5	1.59	23.64	22.54	30.85	28.76
C4_dioxolenium/conf_04	1S5	1.62	28.14	26.56	35.79	32.41
C4_dioxolenium/conf_05	1S5	1.59	31.73	30.83	37.88	35.00
C4_dioxolenium/conf_06	1S5	1.57	37.18	36.04	43.58	40.84
C4_dioxolenium/conf_07	1,4B	1.60	41.85	41.22	50.53	48.38
C4_dioxolenium/conf_08	1S5	1.59	39.80	37.76	47.30	43.28
C4_dioxolenium/conf_09	1S5	1.61	40.70	39.03	49.98	45.51
C4_dioxolenium/conf_10	1S5	1.95	23.48	21.90	36.95	31.40
C6_oxonium/conf_00	1C4	5.26	0.00	0.00	0.00	0.00
C6_oxonium/conf_01	1C4	5.10	7.03	5.60	7.06	5.64
C6_oxonium/conf_02	BO,3	4.99	14.34	12.27	15.12	12.34
C6_oxonium/conf_03	1C4	5.16	9.83	6.80	11.84	6.58
C6_oxonium/conf_04	1C4	5.09	12.66	10.85	12.23	9.48
oxocarbenium/conf_00	4E	5.24	15.85	14.57	33.58	26.58
oxocarbenium/conf_01	E3	5.20	17.94	16.88	37.46	29.12
oxocarbenium/conf_02	3H4	5.15	30.05	27.67	42.83	35.03
oxocarbenium/conf_03	4H3	5.22	24.73	22.55	43.66	35.00
oxocarbenium/conf_04	5H4	5.20	30.12	28.06	42.76	35.39
oxocarbenium/conf_IRC1	4H3	2.75	32.44	31.55	47.53	42.15
TS1	1,4B	1.71	51.89	52.52	63.52	61.01
C4_dioxolenium/conf_IRC1	1,4B	1.73	52.31	50.80	63.37	59.14
oxocarbenium/conf_IRC2	4H3	2.75	32.44	31.56	47.54	42.17
TS2	–	2.99	175.75	174.83	214.89	193.42
C4_rearranged/conf_IRC2	–	3.62	60.57	58.60	66.06	56.89

Table S7. List of structures of the **6TFA** glycosyl cation reoptimized at the PBE0+D3/6-311+G(d,p) level of theory. Ring puckers, bond distances between the carbonyl oxygen of the C4-pivaloyl group and the anomeric carbon (C1), energies (ΔE , including zero-point-vibrational energy) and free energies (ΔF) at 90 K are assigned to each structure. Furthermore, single-point energies of these reoptimized structures at the DLPNO-CCSD(T)/Def2-TZVPP level of theory are indicated as such (ΔE , not including ZPVE) and with ZPVE plus free energy correction at 90 K (ΔF) derived from the PBE0 calculation.

ID	Ring Pucker	d(C4=O—C1) [Å]	ΔE (PBE0+D3) [kJ mol ⁻¹]	ΔF (PBE0+D3) [kJ mol ⁻¹]	ΔE (DLPNO-CCSD(T)) [kJ mol ⁻¹]	ΔF (DLPNO-CCSD(T)) [kJ mol ⁻¹]
C6_dioxolenium/conf_00	1C4	1.53	10.22	12.21	5.12	11.77
C6_dioxolenium/conf_01	1C4	1.53	17.19	18.73	11.59	16.37
C6_dioxolenium/conf_02	1C4	1.53	35.29	33.10	29.99	29.96
C4_oxonium/conf_00	1,4B	5.35	42.08	36.48	34.48	25.62
C4_oxonium/conf_01	1,4B	5.23	50.67	49.18	41.62	38.59
C4_oxonium/conf_02	1,4B	3.91	48.98	44.00	37.28	31.12
C4_oxonium/conf_03	1,4B	5.15	47.40	43.90	35.53	31.27
C4_oxonium/conf_04	1,4B	5.99	70.53	67.87	60.97	57.00
oxocarbenium/conf_00	4E	5.11	0.00	0.00	0.00	0.00
oxocarbenium/conf_01	5H4	4.80	13.64	12.50	13.14	10.93
oxocarbenium/conf_02	3E	3.62	17.18	14.95	21.17	16.99
oxocarbenium/conf_03	3E	3.65	10.47	8.41	10.62	6.93
oxocarbenium/conf_04	3E	3.77	12.32	11.20	15.80	13.25

Table S8. List of structures of the **4,6TFA** glycosyl cation reoptimized at the PBE0+D3/6-311+G(d,p) level of theory. Ring puckers, bond distances between the carbonyl oxygen of the C4-pivaloyl group and the anomeric carbon (C1), energies (ΔE , including zero-point-vibrational energy) and free energies (ΔF) at 90 K are assigned to each structure. Furthermore, single-point energies of these reoptimized structures at the DLPNO-CCSD(T)/Def2-TZVPP level of theory are indicated as such (ΔE , not including ZPVE) and with ZPVE plus free energy correction at 90 K (ΔF) derived from the PBE0 calculation.

ID	Ring Pucker	d(C4=O—C1) [Å]	ΔE (PBE0+D3) [kJ mol ⁻¹]	ΔF (PBE0+D3) [kJ mol ⁻¹]	ΔE (DLPNO-CCSD(T)) [kJ mol ⁻¹]	ΔF (DLPNO-CCSD(T)) [kJ mol ⁻¹]
C4_dioxolenium/conf_00	1S5	1.59	1.82	1.00	0.01	0.01
C4_dioxolenium/conf_01	1S5	1.60	1.07	0.87	1.54	2.09
C4_dioxolenium/conf_02	1S5	1.59	0.00	0.00	1.06	1.70
C4_dioxolenium/conf_03	1S5	2.04	5.15	3.49	12.90	8.92
C4_dioxolenium/conf_04	1S5	2.09	5.34	3.16	12.61	8.07
C4_dioxolenium/conf_05	1S5	2.01	6.17	4.19	12.87	8.54
C4_dioxolenium/conf_06	1S5	1.61	6.00	3.98	4.30	2.71
C4_dioxolenium/conf_07	1S5	1.58	5.04	4.62	3.51	4.05
C4_dioxolenium/conf_08	1S5	2.01	6.17	4.20	12.87	8.54
C4_dioxolenium/conf_09	1S5	1.61	8.66	8.58	6.59	7.23
C6_dioxolenium/conf_00	1C4	5.25	8.25	8.61	12.19	14.45
C6_dioxolenium/conf_01	1C4	5.32	13.84	13.34	14.57	15.55
C6_dioxolenium/conf_02	1C4	5.29	22.38	20.84	24.71	24.03
oxocarbenium/conf_00	OS2	5.09	4.07	1.61	8.91	3.60
oxocarbenium/conf_01	OS2	5.09	4.07	1.61	8.90	3.60
oxocarbenium/conf_02	OS2	5.08	5.09	1.67	9.16	1.98
oxocarbenium/conf_03	OS2	5.09	8.04	5.25	14.76	8.62
oxocarbenium/conf_04	OS2	5.32	12.20	11.33	15.80	11.37
oxocarbenium/conf_05	3E	5.17	8.74	5.40	15.66	8.68
oxocarbenium/conf_IRC1	4H3	5.25	20.48	16.98	35.74	23.96
TS1	4H3	3.88	62.79	59.98	79.34	67.14
C4_dioxolenium/conf_IRC1	1S5	1.59	1.82	1.00	0.00	0.00
oxocarbenium/conf_IRC2	4H3	5.22	18.56	14.87	31.03	19.56
TS2	–	5.38	84.39	81.13	101.26	85.30
C6_rearranged/conf_IRC2	–	5.29	53.83	49.35	48.60	37.34

2.3 Energy Hierarchies

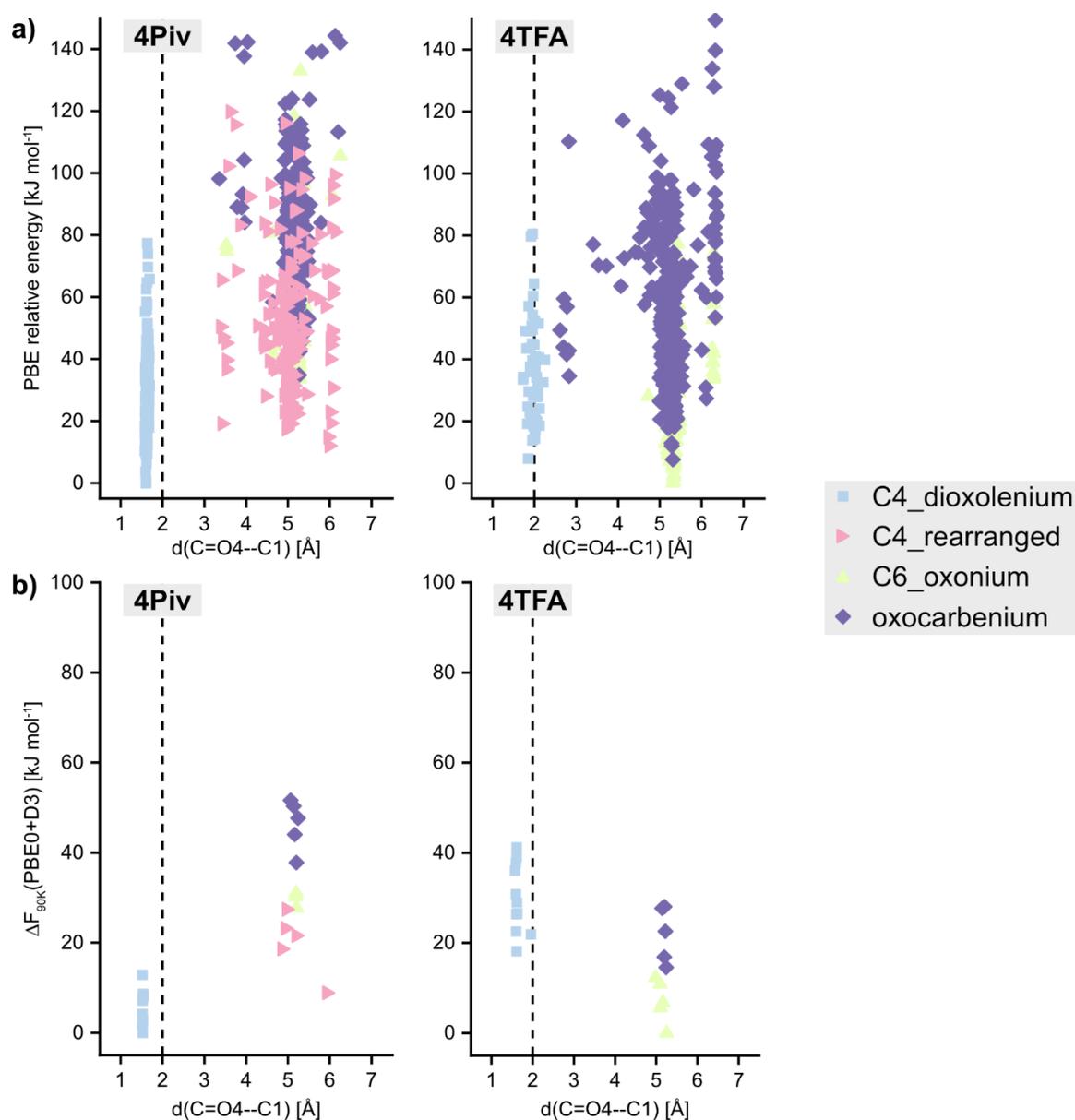


Figure S4. Energy hierarchies of (a) sampled and (b) reoptimized **4Piv** and **4TFA** galactosyl cations as a function of the distance between the carbonyl oxygen of the C4-acyl group and the anomeric carbon (C1). Blue squares indicate dioxolenium-type, red triangles rearranged, yellow triangles oxonium-type, and purple diamond oxocarbenium-type structures. The energetics of sampled and reoptimized galactosyl cations were computed at the PBE+vdW^{TS}/*light* and PBE0+D3/6-311+G(d,p) levels of theory respectively.

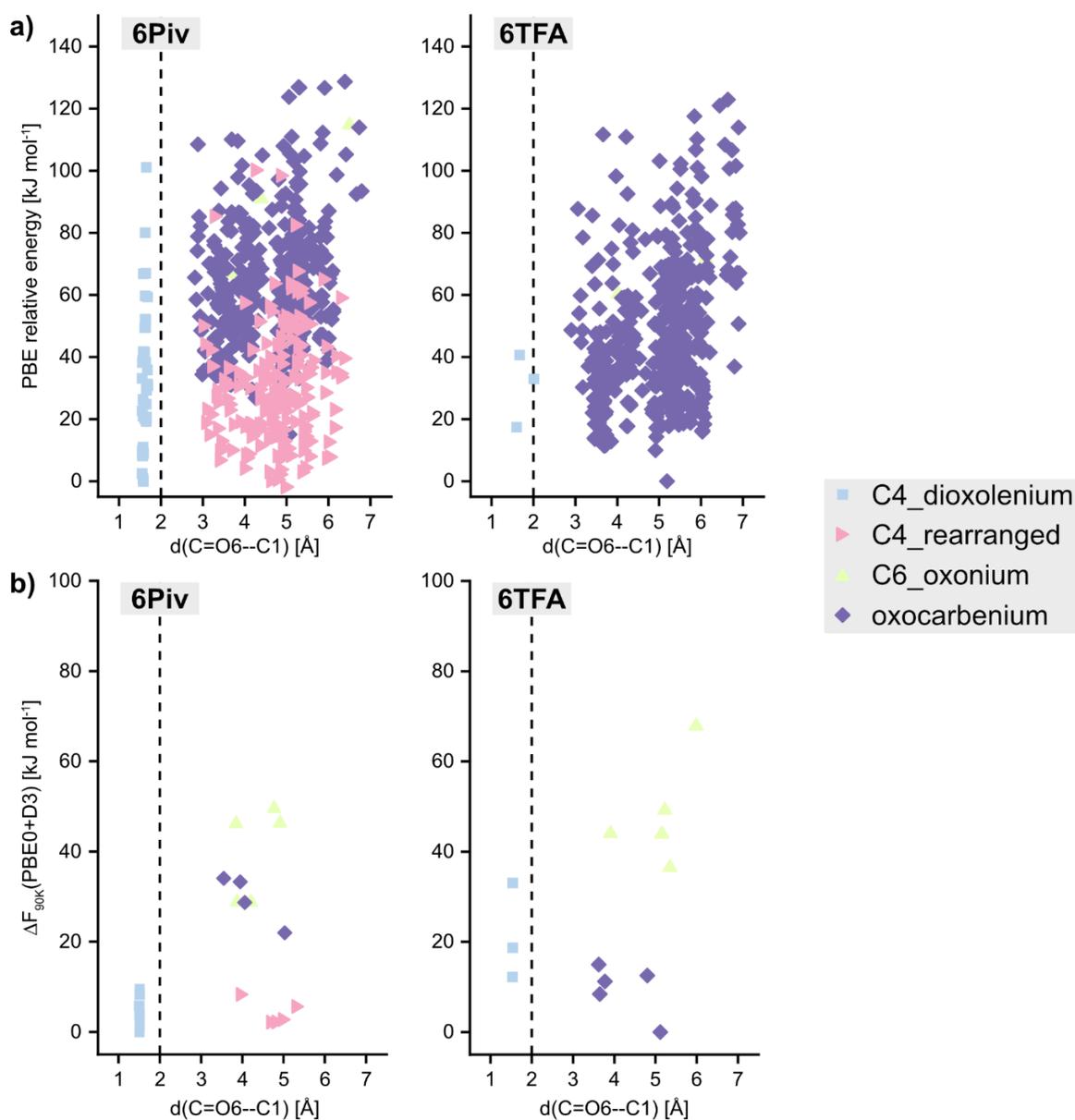


Figure S5. Energy hierarchies of (a) sampled and (b) reoptimized **6Piv** and **6TFA** galactosyl cations as a function of the distance between the carbonyl oxygen of the C6-acyl group and the anomeric carbon (C1). Blue squares indicate dioxolenium-type, red triangles rearranged, yellow triangles oxonium-type, and purple diamond oxocarbenium-type structures. The energetics of sampled and reoptimized galactosyl cations were computed at the PBE+vdW^{TS}/*light* and PBE0+D3/6-311+G(d,p) levels of theory respectively.

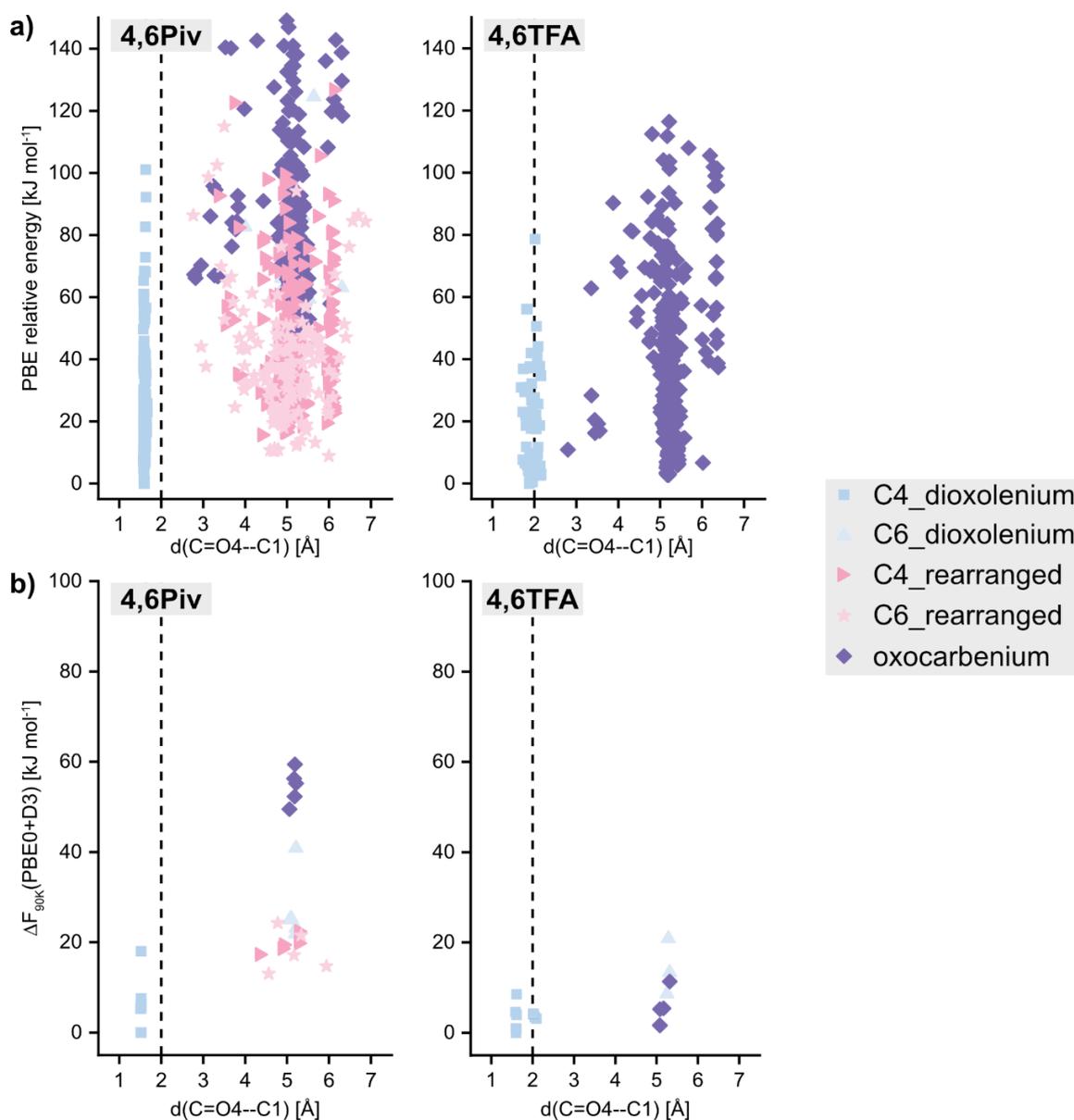


Figure S6. Energy hierarchies of (a) sampled and (b) reoptimized **4,6Piv** and **4,6TFA** galactosyl cations as a function of the distance between the carbonyl oxygen of the C4-acyl group and the anomeric carbon (C1). Blue squares indicate C4_dioxolenium-type, blue triangles C6_dioxolenium-type, red triangles C4_rearranged, red stars C6_rearranged, and purple diamond oxocarbenium-type structures. The energetics of sampled and reoptimized galactosyl cations were computed at the PBE+vdW^{TS}/light and PBE0+D3/6-311+G(d,p) levels of theory respectively.

2.4 Energy Diagrams

Transition states (TS) that connect oxocarbenium-type structures with rearranged and dioxolenium-type structures have been computed for **4,6Piv**, **4Piv**, **6Piv**, **4,6TFA**, and **4TFA** glycosyl cations. To find the transition states, relaxed scans of the bonds that form were performed in Gaussian 16. The saddle point of the obtained surfaces was optimized as a transition state and its existence confirmed by a single imaginary frequency. The transition states were connected to the structures in Tables S3-S6 and S8 by intrinsic reaction coordinate (IRC) calculations. The structures in most cases do not correspond to the lowest energy structures for each type of structure. As the conformational space of the probed ions is vast, the existence of multiple transition states that might be more or less favorable cannot be excluded. The structures generated by the IRC calculation can be transformed into their lowest-energy conformers by simple rotation of bonds, which usually requires only a low activation energy. Furthermore, we consider the single-point energy computed at DLPNO-CCSD(T)/Def2-TZVPP level of theory to construct energy diagrams including the ZPVE and free energy from the DFT calculation at PBE0 level of theory (Figures 2, 3 and S7). The energetics indicate that the rearrangement gets favored at higher temperatures, likely due to an entropic contribution. However, the exact temperature of the ions under in-source fragmentation conditions is not known.

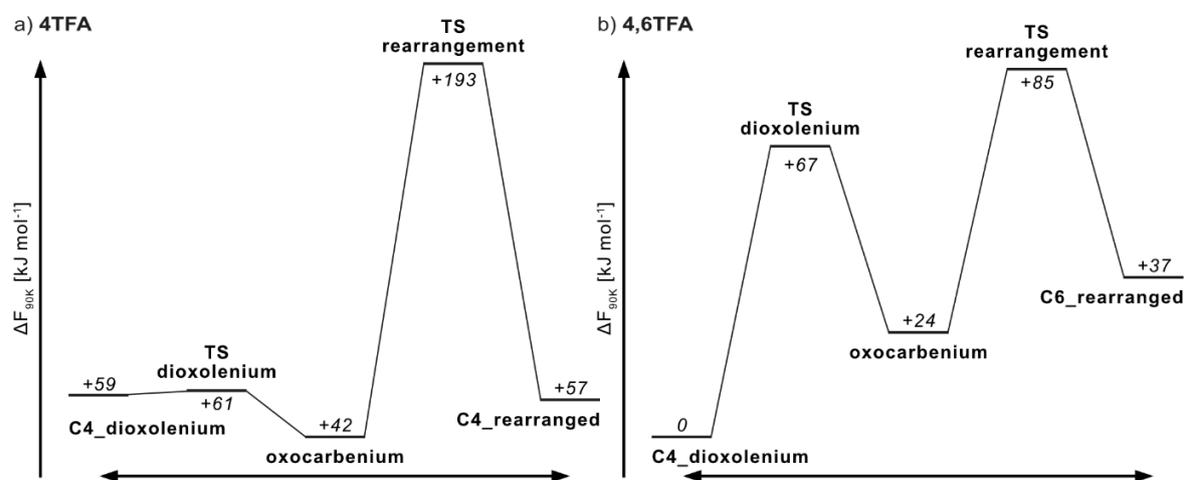


Figure S7. Energy diagrams of (a) **4TFA** and (b) **4,6TFA** galactosyl cations. Surfaces of formation of dioxolenium-type and rearranged structures from oxocarbenium-type structures are shown. As these galactosyl cations are not formed, the rearrangement and participation processes are purely hypothetical. Generally, the energetics indicate that these cations are less prone to engage in remote participation or rearrangement.

2.5 3D Structures

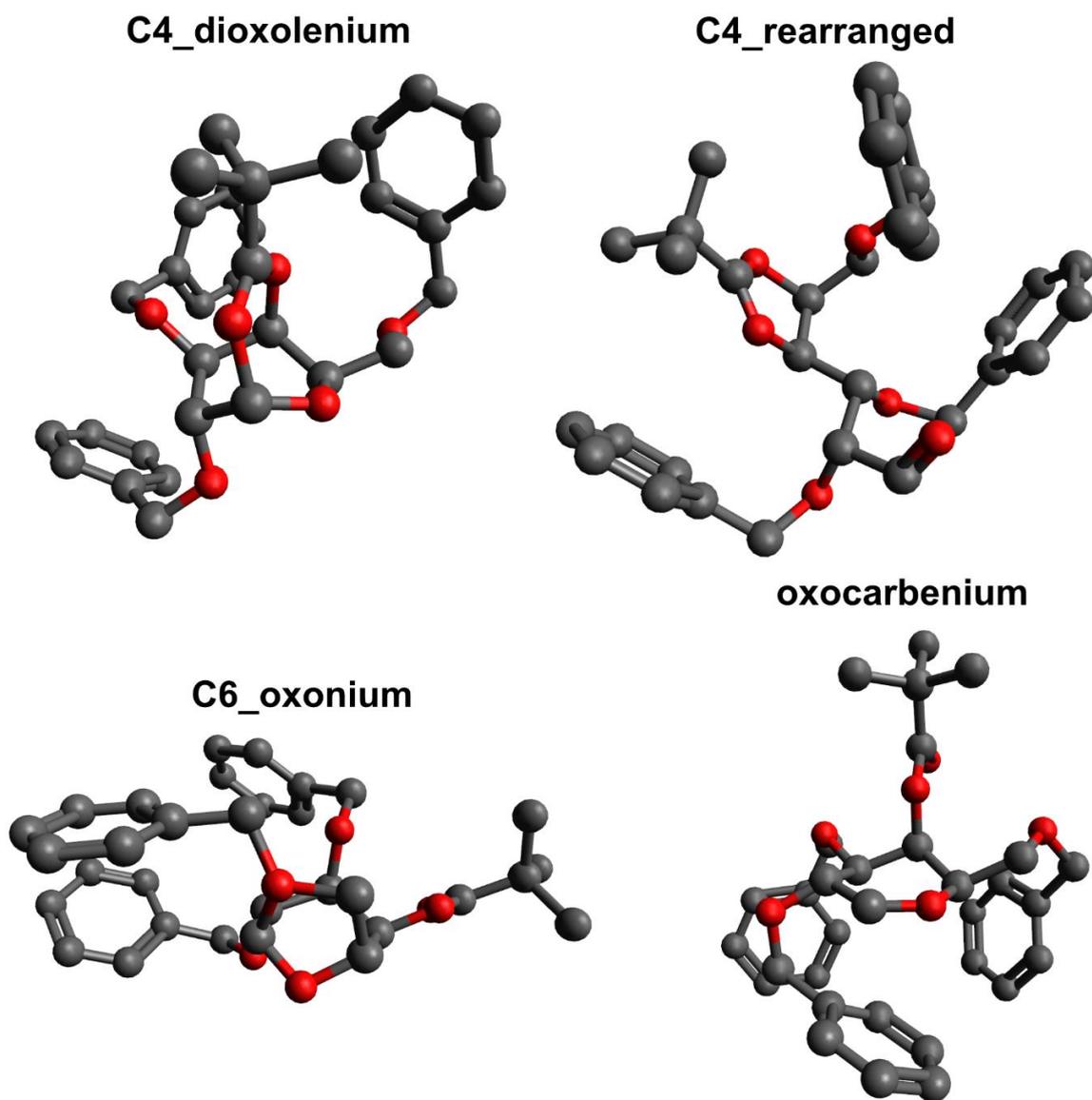


Figure S8. Reoptimized lowest-energy structures for intact and rearranged **4Piv** glycosyl cations for each structural motif. Hydrogen atoms are omitted for clarity.

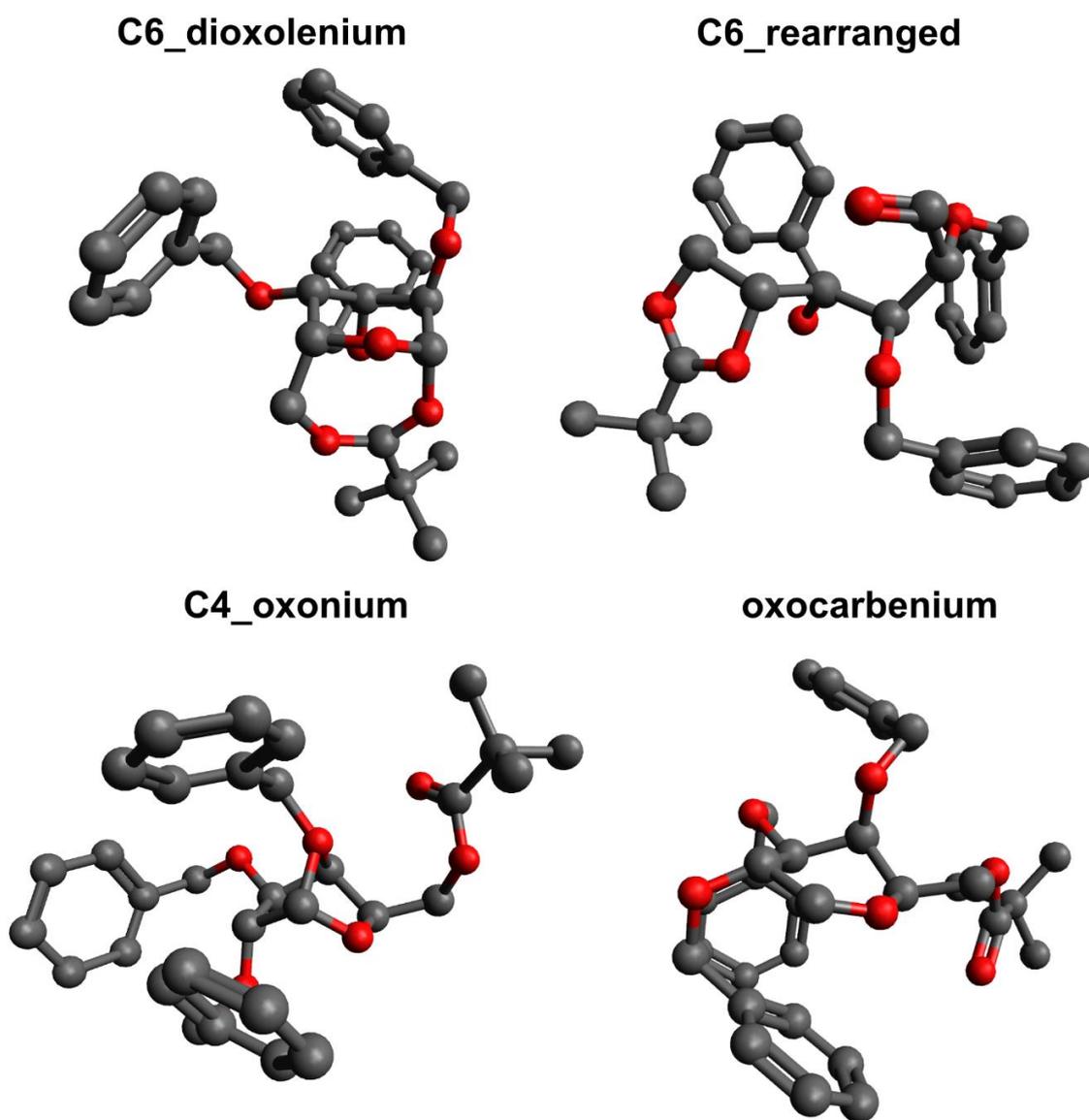


Figure S9. Reoptimized lowest-energy structures for intact and rearranged **6Piv** glycosyl cations for each structural motif. Hydrogen atoms are omitted for clarity.

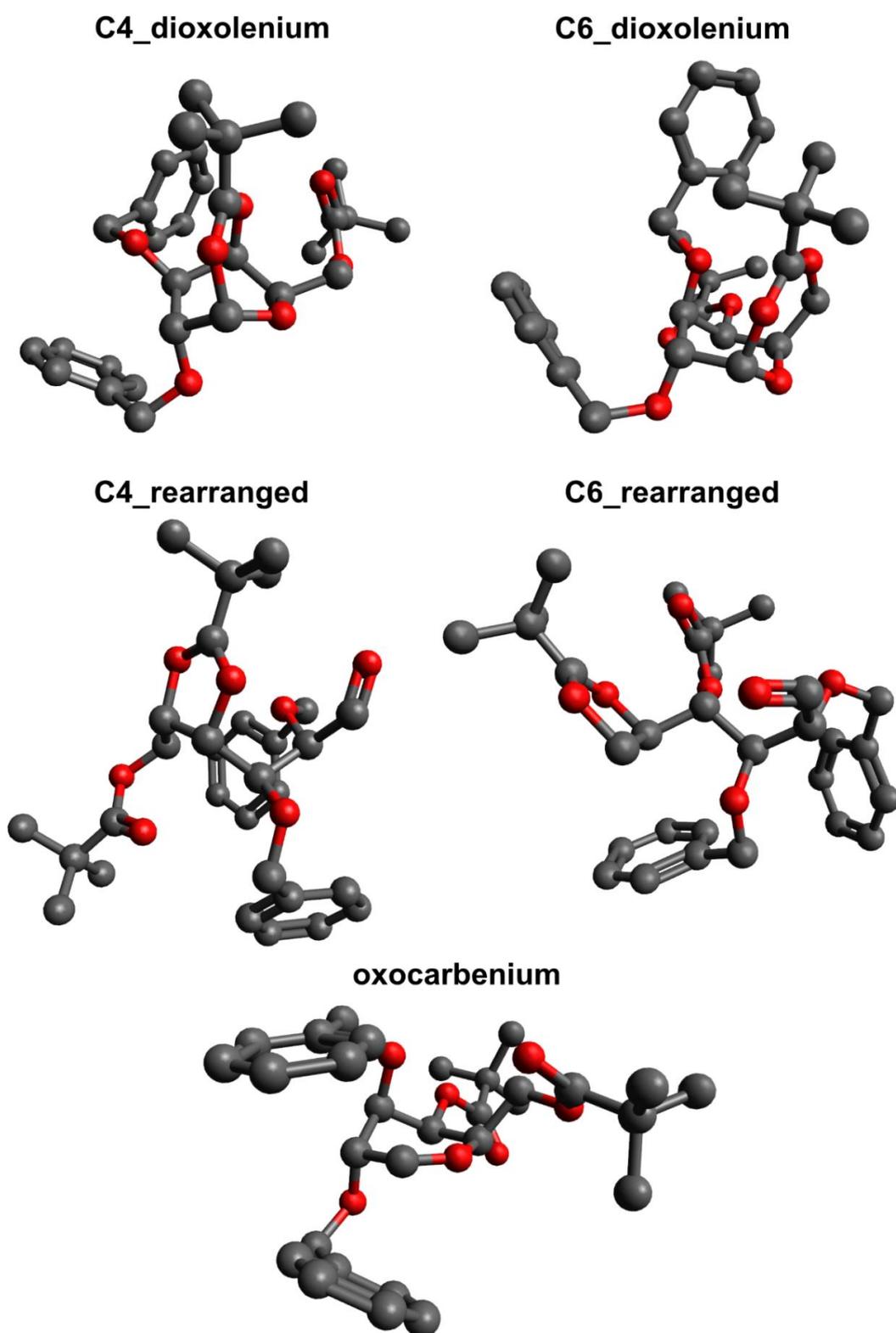


Figure S10. Reoptimized lowest-energy structures for intact and rearranged **4,6Piv** glycosyl cations for each structural motif. Hydrogen atoms are omitted for clarity.

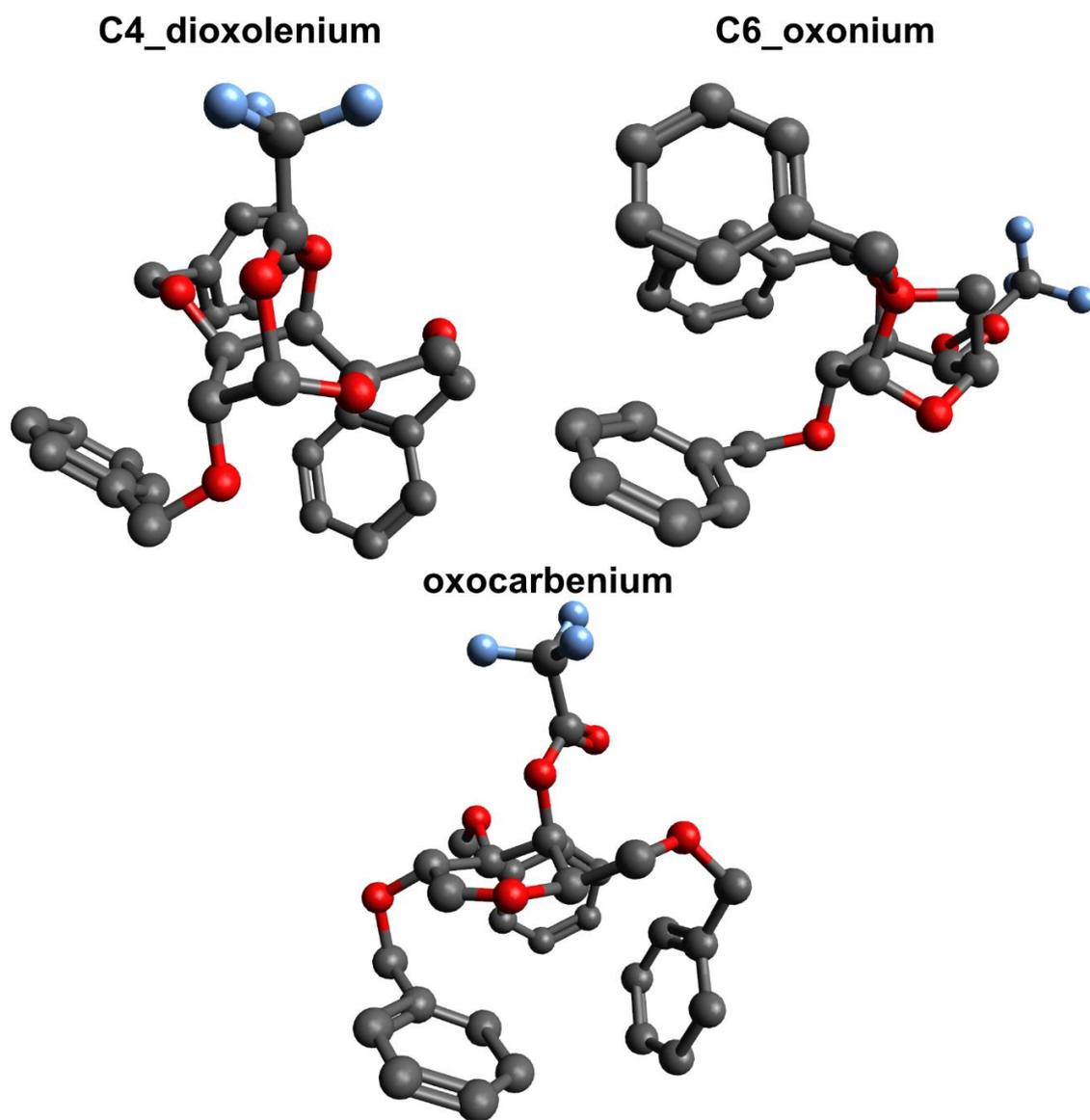


Figure S11. Reoptimized lowest-energy structures for intact **4TFA** glycosyl cations for each structural motif. Hydrogen atoms are omitted for clarity.

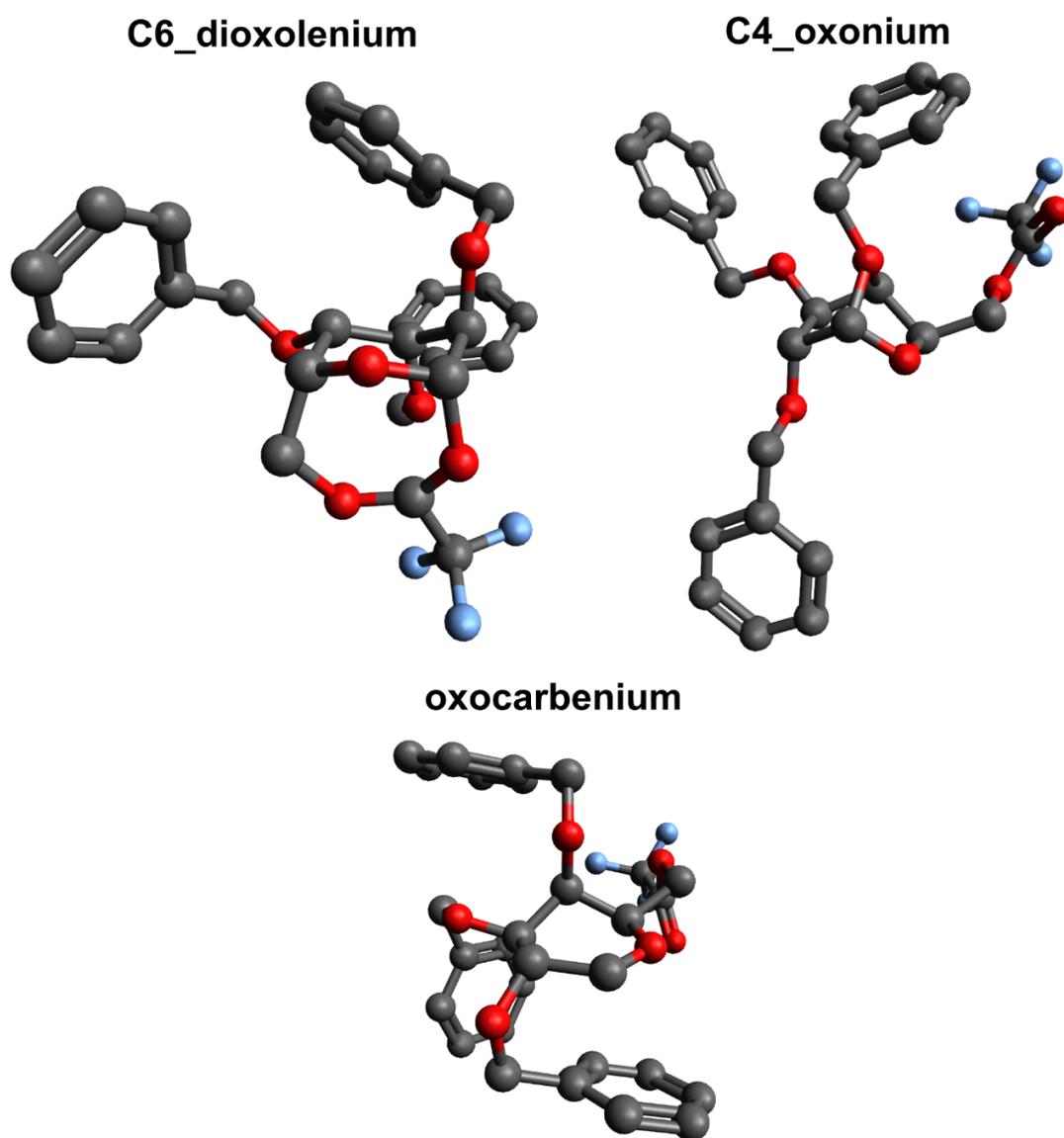


Figure S12. Reoptimized lowest-energy structures for intact **6TFA** glycosyl cations for each structural motif. Hydrogen atoms are omitted for clarity.

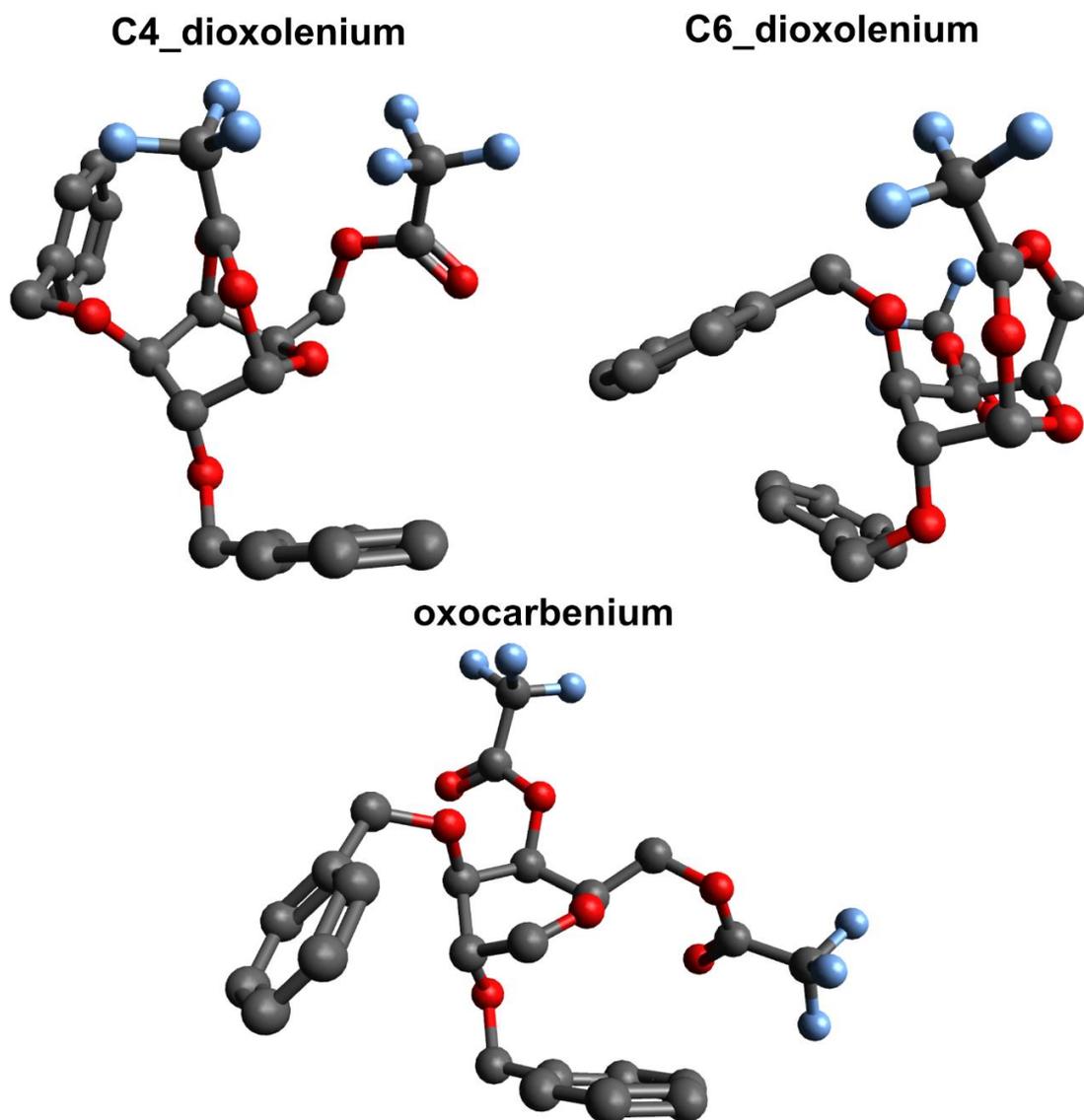


Figure S13. Reoptimized lowest-energy structures for intact 4,6TFA glycosyl cations for each structural motif. Hydrogen atoms are omitted for clarity.

2.6 xyz-Coordinates of reoptimized structures

xyz-Coordinates of all reoptimized geometries can be found in a separate document “coordinates.xyz”.

3 General Information

All **chemicals** were reagent grade and used as supplied unless otherwise noted. All **solvents** for chemical reactions were commercially purchased in p.a. quality. If stated, they were dried in a Solvent Dispensing System (J.C. Meyer). For HPLC and MS spectrometry, solvents with corresponding quality were used. Water was used from a Milli Q-station from Millipore. **The automated syntheses** were performed on a home-built synthesizer developed at the Max Planck Institute of Colloids and Interfaces.

Reaction completion, identity, and purity of all compounds were determined by low resolution mass spectrometry (**ESI-LRMS**) or analytical thin-layer chromatography (**TLC**). TLC was performed on Merck silica gel 60 F₂₅₄ plates (0.25 mm). Compounds were visualized by UV irradiation (254 nm) or stained (5% sulfuric acid in ethanol or Hanessian's Stain: 235 mL of distilled water, 12 g of ammonium molybdate, 0.5 g of ceric ammonium molybdate, and 15 mL sulfuric acid). **Flash column chromatography** was performed on Kieselgel 60 with 230-400 mesh (Sigma-Aldrich, St. Louis, USA). Analysis and purification by normal and reverse phase **HPLC** and ESI-LRMS was performed by using an Agilent 1200 series. ¹H, ¹³C, COSY and HSQC **NMR spectra** were recorded in parts per million (δ) relative to the resonance of the solvent on a Varian 400-MR (400 MHz), Varian 600-MR (600 MHz), or Bruker Biospin AVANCE700 (700 MHz) spectrometer. Assignments were supported by COSY and HSQC experiments. High resolution mass spectra (**HRMS**) were obtained using 6210 ESI-TOF mass spectrometer (Agilent) and **MALDI-TOF** autoflexTM (Bruker) instruments.

4 Materials and Conditions for Automated Synthesis

4.1 Materials and Measurements

Solvents used for dissolving all building blocks and making of various solutions were taken from Solvent Dispensing System (J.C. Meyer). Wash solvents were HPLC grade. Prior to automated synthesis, the building blocks were weighed and co-evaporated three times with anhydrous toluene and dried for at least one hour under high vacuum prior to use. All solutions were freshly prepared in oven-dried, argon-flushed glassware and kept under argon during the automation process. Isolated product yields were calculated on the basis of resin loading. Functionalized resin **2** was synthesized as previously reported¹³ and resin loading (0.40 mmol/g) was determined following a published protocol.¹⁴ Resin was placed in the reaction vessel and was swollen in

dichloromethane for 20 min at room temperature before starting the first module. During this time, all reagent lines involved in the synthesis were washed and primed.

4.2 Preparation of Stock Solutions

Building Block Solution: Glycosyl phosphate building block (0.07 mmol, 4.7 equiv. per cycle) was dissolved in 1 mL (per cycle) of anhydrous CH₂Cl₂.

Activator Solution/Acidic Wash Solution: TMSOTf (0.9 mL, 0.62 mmol) was added to 40 mL of anhydrous CH₂Cl₂.

Pre-Capping Solution: Pyridine (10 mL) was added to 90 mL of DMF.

Capping Solution: Methanesulfonic acid (1.2 mL, 18.5 mmol), acetic anhydride (6 mL, 63.5 mmol) were added to 50 mL of anhydrous CH₂Cl₂.

Fmoc Deprotection Solution: Piperidine (20 mL) was added to 80 mL anhydrous DMF.

4.3 Modules for Automated Synthesis

Initiation: The resin **2** is loaded in the reaction vessel and washed with DMF, THF, and CH₂Cl₂ (3 x 3 mL for 15 s, respectively). The resin is then swollen in 2 mL CH₂Cl₂ for 20 minutes while the temperature of the reaction vessel is cooled to the lowest temperature required throughout the synthesis. During this time, all reagent lines needed for the synthesis are washed and primed.

Module I - Acidic Washing: Once the temperature of the reaction vessel has adjusted to the desired temperature of the subsequent glycosylation, 1 mL of the **Acidic Wash Solution** is delivered to the reaction vessel. After bubbling for three minutes, the solution is drained. Finally, the resin is washed with 3 mL CH₂Cl₂ for 25 s and drained.

Module II – Glycosylation (for glycosyl phosphate): Upon draining the CH₂Cl₂ in the reaction vessel, **Building Block Solution** (1 mL) containing the appropriate building block is delivered from the building block storing component to the reaction vessel. After the temperature again reaches the desired temperature (T₁), **Activator Solution** (1 mL) is delivered to the reaction vessel from the respective activator storing component to the reaction vessel. The glycosylation mixture is incubated for the selected duration (t₁) at the desired T₁, then the reaction temperature is ramped to T₂. Once T₂ is reached, it is maintained and the reaction mixture is incubated for an additional time (t₂). Once the incubation time is finished, the reaction mixture is

drained and the resin is washed with CH₂Cl₂ (2 x 2 mL for 25 s). The temperature of the reaction vessel is increased to 25 °C for the next module.

Module III - Capping: The resin is washed with DMF (2 x 2 mL for 25 s). Then **Pre-capping Solution** (2 mL) is delivered at 25 °C. After 1 min, the reaction solution is drained and the resin is washed with CH₂Cl₂ (3 x 3 mL for 25 s). Upon washing, **Capping Solution** (4 mL) is delivered and the temperature is maintained at 25 °C. The resin and the reagents are incubated for 20 min. The solution is then drained from the reactor vessel and the resin is washed with CH₂Cl₂ (3 x 3 mL for 25 s).

Module IV - Fmoc Deprotection: The resin is first washed with DMF (3 x 2 mL for 25 s), and then **Fmoc Deprotection Solution** (2 mL) is delivered to the reaction vessel at 25 °C. After 5 min, the reaction solution is drained and the resin is washed with DMF (3 x 3 mL for 25 s) and CH₂Cl₂ (5 x 2 mL for 25 s). Then, the temperature of the reaction vessel is decreased to T₁ for the next module.

4.4 Post-automated Synthesis Manipulations, Analysis and Purification

Cleavage from Solid Support (Method A)

After automated synthesis, the resin was removed from the reaction vessel, suspended in CH₂Cl₂ (20 mL), and photocleaved in a continuous-flow photoreactor. A Vapourtec E-Series easy-MedChem, equipped with a UV-150 Photochemical reactor having a UV-150 Medium-Pressure Mercury Lamp (arc length 27.9 cm, 450 W) surrounded by a long-pass UV filter (Pyrex, 50% transmittance at 305 nm) was used. A Pump 11 Elite Series (Harvard Apparatus syringe pump at a flow rate of 0.8 mL/min) was used to pump the mixture through a FEP tubing (i.d. 3.0 inch, volume: 12 mL) at 20 °C. The reactor was washed with 20 mL CH₂Cl₂ at a flow rate of 2.0 mL/min. The output solution was filtered to remove the resin and the solvent was evaporated *in vacuo*. Crude was then analyzed by MALDI.

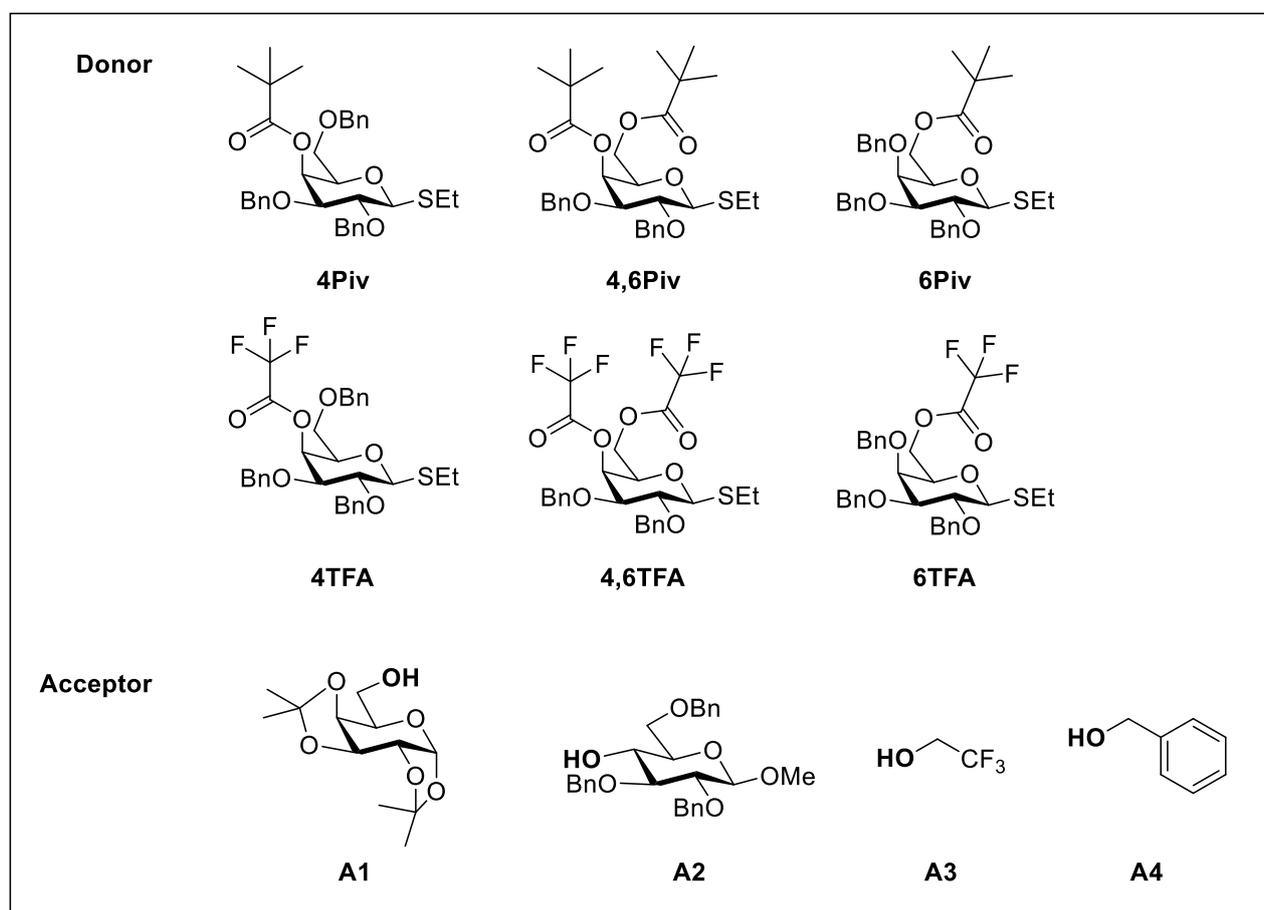
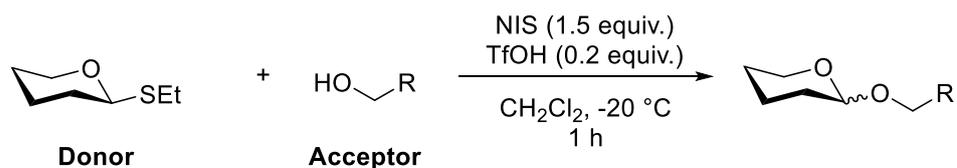
Analytical NP-HPLC of Crude Material (Method B-1a)

Analytical NP-HPLC was conducted on an Agilent 1200 Series system. A YMC-Diol-300-NP column (150 mm x 4.60 mm I.D.) was used at a flow rate of 1.00 mL/min with hexane/EtOAc as eluent (20% EtOAc in hexane for 5 min, 20 → 55% EtOAc in hexane over 35 min, 55 → 100% EtOAc in hexane over 35 min, 100% EtOAc for 10 min).

Preparative NP-HPLC of Crude Material (Method B-1b)

Preparative NP-HPLC was conducted on an Agilent 1200 Series system. A YMC-Diol-300-NP column (150 mm x 20 mm I.D.) was used at a flow rate of 15.00 mL/min with hexane/EtOAc as eluent (20% EtOAc in hexane for 5 min, 20 → 55% EtOAc in hexane over 35 min, 100% EtOAc for 10 min).

4.5 General Procedure for Glycosylations



Donor (35 μmol, 1.0 equiv.) and the acceptor **A1** or **A2** (35 μmol, 1.0 equiv.) were co-evaporated with anhydrous toluene (3 x 2 mL) and kept under high vacuum for one hour. Anhydrous CH₂Cl₂ (2 mL) was added and if acceptor **A1** or **A2** were not used, acceptor **A3** or **A4** (175 μmol, 5 equiv.) was added and the mixture

was stirred over activated molecular sieves (3 Å-AW) for 30 minutes at room temperature. The solution was cooled to -20 °C and NIS (52 μmol, 1.5 equiv.) was added followed by TfOH (60 μL of a 1% solution in CH₂Cl₂, 7 μmol, 0.2 equiv.) and the mixture was stirred for 1 h at -20 °C. The reaction mixture was quenched with pyridine, diluted with CH₂Cl₂, filtered and was then washed with 10% Na₂S₂O₃ (10 mL). The aqueous phase was extracted with CH₂Cl₂ (3 x 10 mL), dried over Na₂SO₄ and concentrated. The residue was purified by HPLC using **Method B-2**.

4.6 Determination of Alpha/Beta Ratios and Purification

Alpha/beta ratios were determined using HPLC and ¹H/¹³C/HSQC NMR spectroscopy.

Analytical NP-HPLC for alpha/beta ratio determination (Method B-2a)

Analytical NP-HPLC was conducted on an Agilent 1200 Series system. A YMC-Diol-300-NP column (150 mm x 4.600 mm I.D.) was used with a flow rate of 1.00 mL/min and hexanes/EtOAc as eluent (16 min linear gradient 2 to 25% EtOAc in hexanes, 2 min linear gradient 25 to 70% EtOAc in hexanes, 2 min isocratic 70% EtOAc).

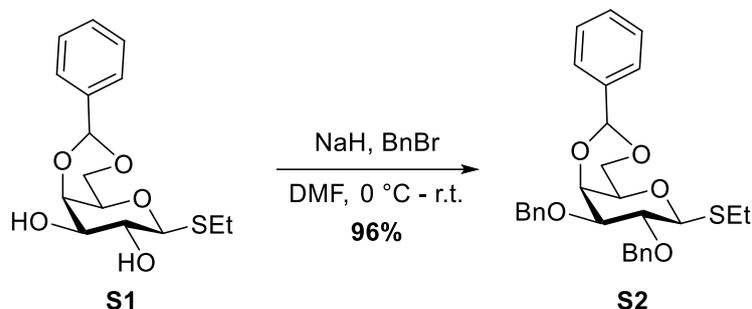
Preparative NP-HPLC of Crude Material (Method B-2b)

Preparative NP-HPLC was conducted on an Agilent 1200 Series system. A YMC-Diol-300-NP column (150 mm x 20 mm I.D.) was used at a flow rate of 15.00 mL/min with hexane/EtOAc as eluent (16 min linear gradient 2 to 25% EtOAc in hexanes, 2 min linear gradient 25 to 70% EtOAc in hexanes, 2 min isocratic 70% EtOAc).

5 Synthesis of building blocks

4Ac and **Bn** galactose building blocks were synthesized according to previously published procedures.¹⁵⁻¹⁶

Ethyl 2,3-bis-*O*-benzyl-4,6-*O*-[(*S*)-phenylmethylene]-1-thio- β -D-galactopyranoside¹⁷ (**S2**)

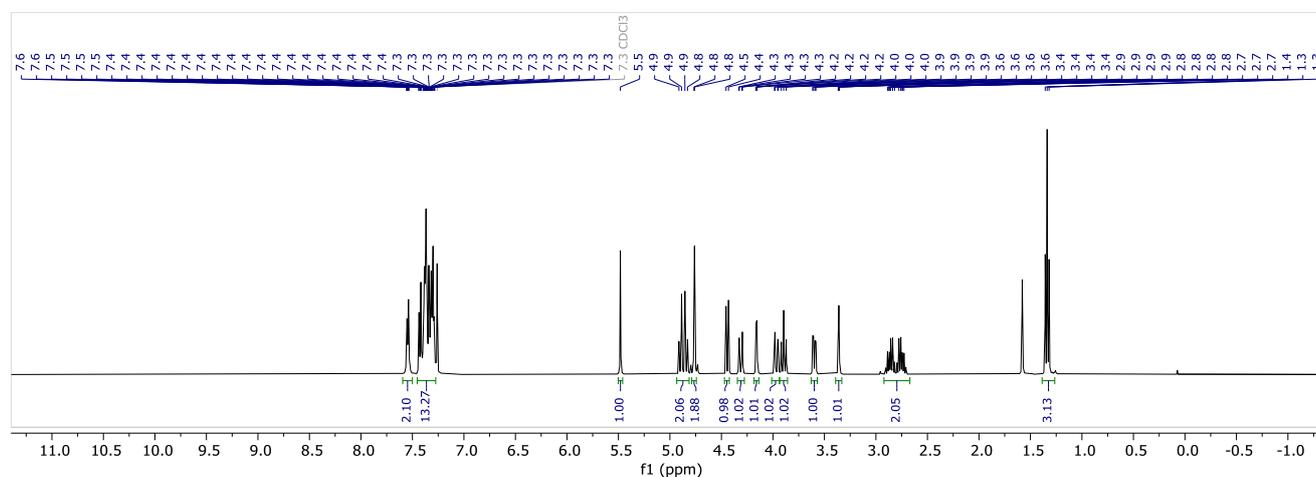


Galactose building block **S1** (2.5 g, 8.00 mmol, 1.0 eq.) was dissolved in anhydrous DMF (15 mL). The stirred solution was cooled to 0 °C and sodium hydride (1.5 g, 37.61 mmol; 60% dispersion in mineral oil, 4.7 eq.) was added in small portions. After 30 min, benzyl bromide (2.85 mL, 24.01 mmol, 3.0 eq.) was added dropwise. The reaction mixture was allowed to warm up to room temperature and was stirred overnight. Methanol (10 mL) was added, the reaction mixture was stirred for 10 min and afterwards diluted with EtOAc (50 mL). The organic layer was washed with water (2 x 30 mL). The aqueous phase was extracted with ethyl acetate (2 x 40 mL). The combined organic phase was washed with water (30 mL) followed by brine (30 mL), dried over Na₂SO₄ and concentrated. Product **S2** (3.8 g, 7.71 mmol, **96%**) was obtained as a colorless syrup after purification by column chromatography (SiO₂, Hex/EtOAc = 9:1).

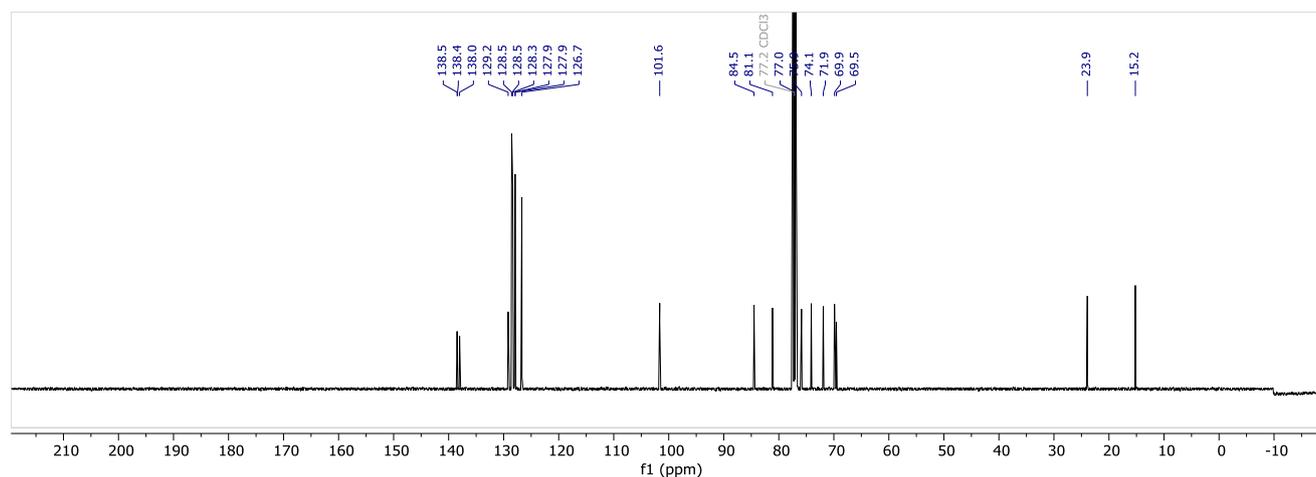
¹H NMR (400 MHz, CDCl₃) δ 7.58 – 7.51 (m, 2H), 7.46 – 7.28 (m, 13H), 5.48 (s, 1H), 4.93 – 4.83 (m, 2H), 4.76 (d, *J* = 2.0 Hz, 2H), 4.44 (d, *J* = 9.6 Hz, 1H), 4.31 (dd, *J* = 12.4, 1.6 Hz, 1H), 4.16 (dd, *J* = 3.6, 1.1 Hz, 1H), 3.97 (dd, *J* = 12.4, 1.8 Hz, 1H), 3.90 (t, *J* = 9.4 Hz, 1H), 3.60 (dd, *J* = 9.2, 3.5 Hz, 1H), 3.36 (q, *J* = 1.5 Hz, 1H), 2.92 – 2.70 (m, 2H), 1.34 (t, *J* = 7.5 Hz, 3H) ppm.

¹³C NMR (101 MHz, CDCl₃) δ 138.5, 138.4, 138.0, 129.2, 128.5, 128.5, 128.3, 127.9, 127.9, 126.7, 101.6, 84.5, 81.1, 77.0, 75.9, 74.1, 71.9, 69.9, 69.5, 23.9, 15.2 ppm.

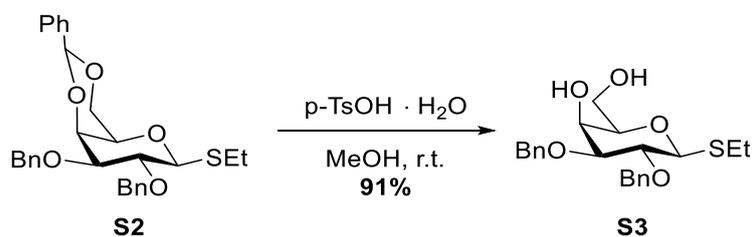
^1H NMR (400 MHz, CDCl_3) of **S2**:



^{13}C NMR (101 MHz, CDCl_3) of **S2**:

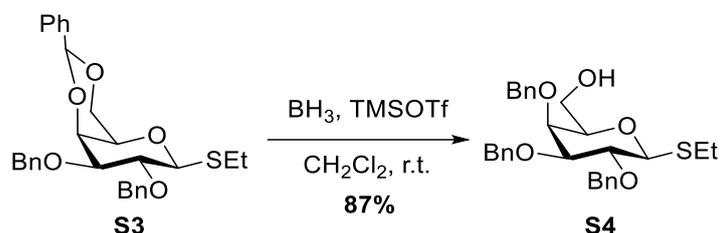


Ethyl 2,3-bis-*O*-benzyl-1-thio- β -D-galactopyranoside¹⁸ (**S3**)



A mixture of **S2** (750 mg, 1.52 mmol, 1.0 equiv.) and *p*-TsOH·H₂O (290 mg, 1.52 mmol, 1.0 equiv.) in MeOH (30 mL) was stirred at room temperature for 5 h. The mixture was diluted with ethyl acetate (100 mL) and was washed with saturated aqueous NaHCO₃ (2 x 50 mL) and brine (50 mL). The organic layer was dried over

Ethyl 2,3,4-tris-*O*-benzyl-1-thio- β -D-galactopyranoside¹⁹ (**S4**)



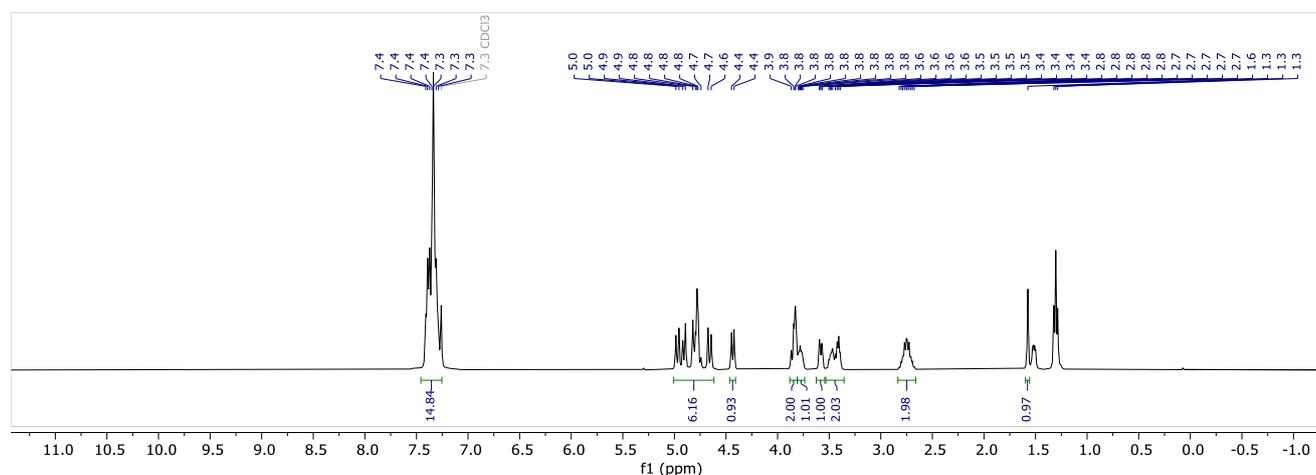
To a solution of **S3** (750 mg, 1.52 mmol, 1.0 equiv.) in anhydrous CH_2Cl_2 (10 mL) was added BH_3 (1 M solution in THF, 7.6 mL, 7.61 mmol, 5.0 equiv.) and TMSOTf (41 μL , 0.23 mmol, 0.15 equiv.). The mixture was stirred under argon atmosphere at room temperature for 4 h. Et_3N (1 mL) was added followed by MeOH until the evolution of H_2 ceased. The mixture was concentrated and coevaporated with MeOH (3 x 30 mL). Product **S4** (650 mg, 1.31 mmol, 87%) was obtained as a white solid after purification by column chromatography (Hex/EtOAc 3:1).

$R_f = 0.19$ (Hex/EtOAc 3:1).

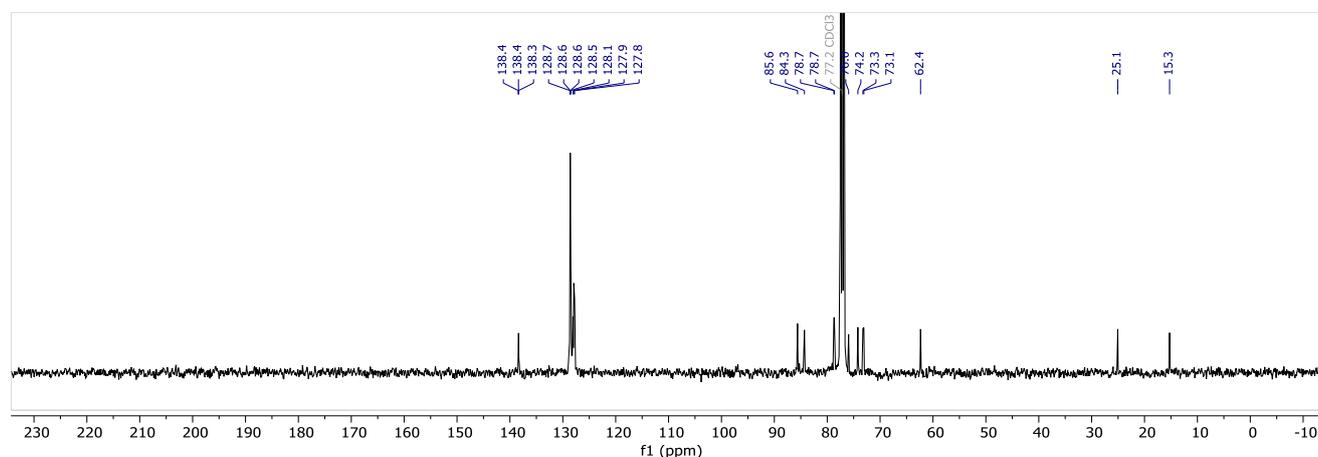
$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.43 – 7.26 (m, 15H), 5.01 – 4.60 (m, 6H), 4.43 (d, $J = 9.6$ Hz, 1H), 3.89 – 3.80 (m, 2H), 3.81 – 3.74 (m, 1H), 3.58 (dd, $J = 9.2, 2.8$ Hz, 1H), 3.52 – 3.37 (m, 2H), 2.84 – 2.65 (m, 2H), 1.57 (s, 1H), 1.30 (t, $J = 7.4$ Hz, 3H) ppm.

$^{13}\text{C NMR}$ (101 MHz, CDCl_3) δ 138.42, 138.37, 138.35, 128.65, 128.59, 128.55, 128.49, 128.08, 127.93, 127.77, 85.61, 84.31, 78.74, 78.66, 75.96, 74.21, 73.25, 73.08, 62.35, 25.06, 15.26 ppm.

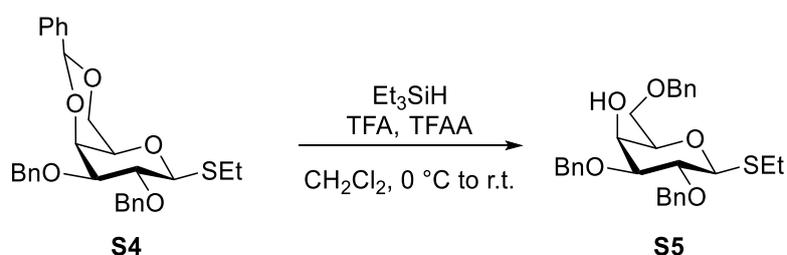
$^1\text{H NMR}$ (400 MHz, CDCl_3) of **S4**:



^{13}C NMR (101 MHz, CDCl_3) of **S4**:



Ethyl 2,3,6-tris-*O*-benzyl-1-thio- β -D-galactopyranoside²⁰ (S5)



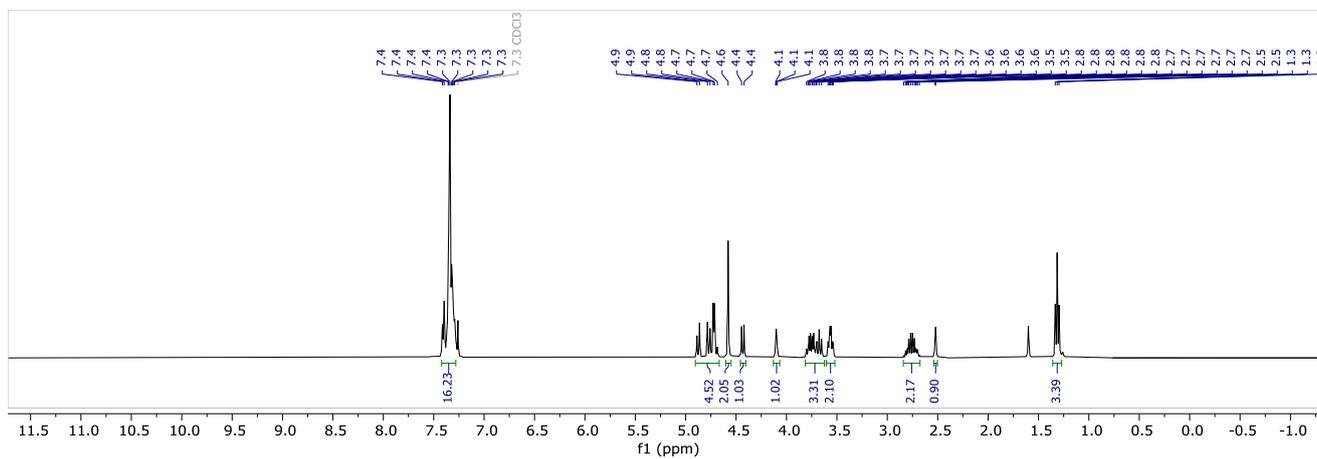
Compound **S4** (1.00 g, 2.03 mmol, 1.0 equiv.) was co-evaporated with anhydrous toluene (2 x 3 mL), and dissolved in anhydrous CH_2Cl_2 (10 mL). Triethylsilane (1.93 mL, 12.18 mmol, 6.0 equiv.) and trifluoroacetic anhydride (0.29 mL, 2.03 mmol, 1.00 equiv.) were added and the solution was cooled to 0°C . Trifluoroacetic acid (0.93 mL, 12.18 mmol, 6.00 equiv.) was added dropwise. The mixture was allowed to warm up to room temperature and was stirred for 5 h. The solution was diluted with CH_2Cl_2 and quenched with saturated aqueous NaHCO_3 (20 mL). The aqueous phase was extracted with CH_2Cl_2 (2 x 30 mL) and the combined organic phase was washed with water (30 mL), dried over Na_2SO_4 , filtered and concentrated. Product **S5** (800 mg, 1.62 mmol, 80%) was obtained as a colorless syrup after purification by column chromatography (SiO_2 , Hex/EtOAc = 9:1 to 7:3).

R_f = 0.48 (Hex/EtOAc 7:3).

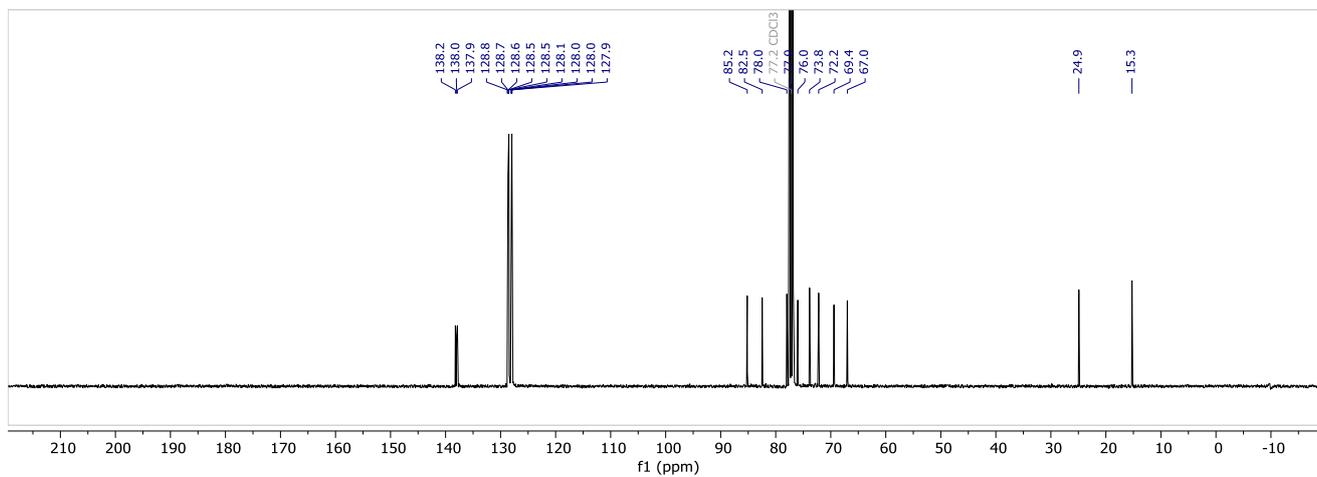
^1H NMR (400 MHz, CDCl_3) δ 7.43 – 7.28 (m, 15H), 4.90 – 4.67 (m, 4H), 4.58 (s, 2H), 4.43 (d, J = 9.7 Hz, 1H), 4.10 (s, 1H), 3.83 – 3.63 (m, 3H), 3.61 – 3.51 (m, 2H), 2.84 – 2.68 (m, 2H), 2.52 (s, 1H), 1.31 (t, J = 7.5 Hz, 3H) ppm.

^{13}C NMR (101 MHz, CDCl_3) δ 138.22, 138.00, 137.85, 128.79, 128.67, 128.58, 128.51, 128.50, 128.11, 128.00, 127.95, 127.93, 85.18, 82.46, 77.98, 76.96, 75.98, 73.85, 72.20, 69.41, 66.98, 24.91, 15.26 ppm.

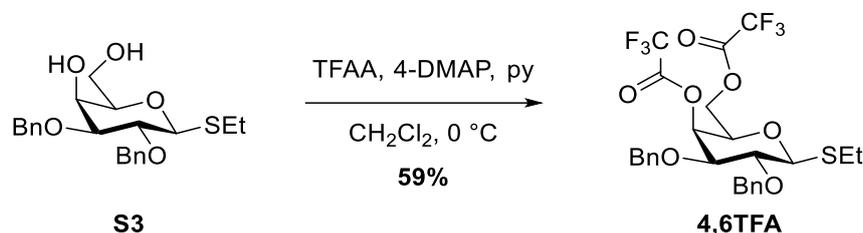
^1H NMR (400 MHz, CDCl_3) of **S5**:



^{13}C NMR (101 MHz, CDCl_3) of **S5**:



Ethyl 2,3-bis-*O*-benzyl-4,6-bis(trifluoroacetate)-1-thio- β -D-galactopyranoside (**4,6TFA**)



A stirred solution of **S3** (150 mg, 0.37 mmol, 1.0 equiv.) in anhydrous CH_2Cl_2 (3 mL) was cooled to 0 °C and anhydrous pyridine (60 μL , 0.74 mmol, 2.0 equiv.), trifluoroacetic anhydride (105 μL , 0.74 mmol, 2.0 equiv.) and 4-DMAP (5 mg, 0.04 mmol, 0.10 equiv.) were added. The solution was stirred at 0 °C for 30 minutes before it was quenched with water (10 mL). The mixture was extracted with CH_2Cl_2 (3 x 10 mL) and the combined organic phases were washed with 1 N HCl (10 mL), dried over Na_2SO_4 , filtered and concentrated. Product **4,6TFA** (130 mg, 0.22 mmol, 59%) was obtained as a colorless oil after purification by column chromatography (SiO_2 , Hex/EtOAc = 9:1 to 7:3).

R_f = 0.90 (Hex/EtOAc 3:1).

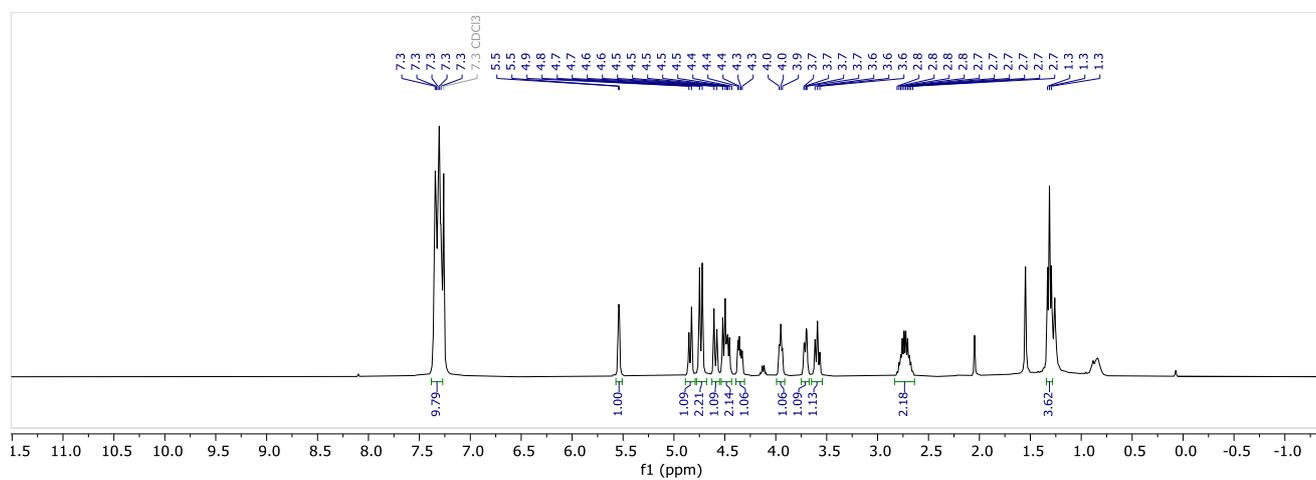
$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.37 – 7.28 (m, 10H), 5.54 (d, J = 3.2 Hz, 1H), 4.88 – 4.80 (m, 1H), 4.77 – 4.70 (m, 2H), 4.63 – 4.56 (m, 1H), 4.54 – 4.43 (m, 2H), 4.35 (dd, J = 11.4, 5.6 Hz, 1H), 3.95 (t, J = 6.4 Hz, 1H), 3.75 – 3.67 (m, 1H), 3.59 (t, J = 9.4 Hz, 1H), 2.73 (qq, J = 13.7, 7.4 Hz, 2H), 1.31 (t, J = 7.5 Hz, 3H) ppm.

$^{13}\text{C NMR}$ (101 MHz, CDCl_3) δ 137.69, 136.97, 128.66, 128.58, 128.54, 128.28, 128.16, 85.76, 79.86, 77.07, 76.18, 73.00, 72.98, 71.51, 64.93, 25.12, 15.14 ppm.

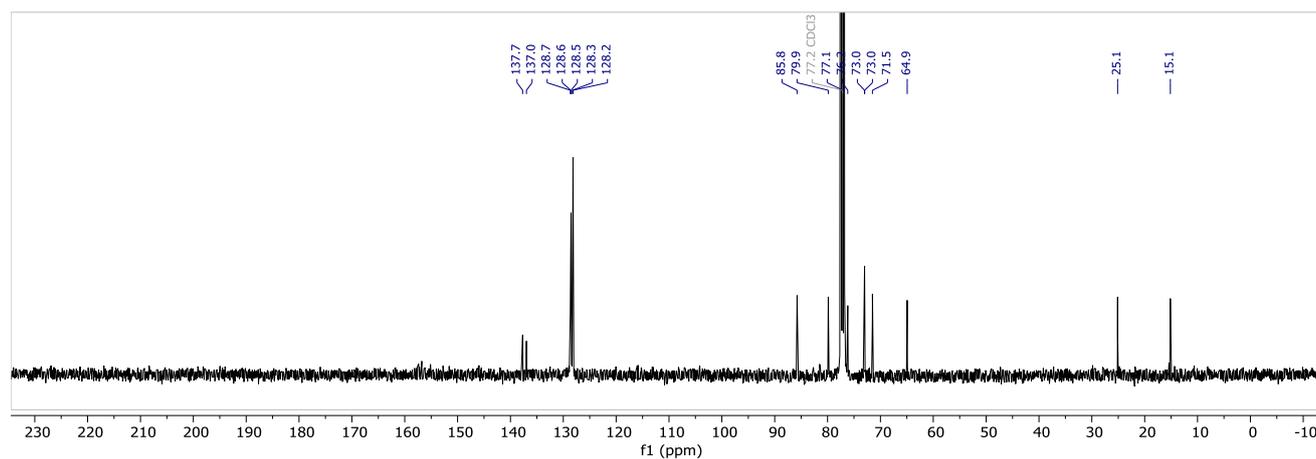
$^{19}\text{F NMR}$ (376 MHz, CDCl_3) δ -74.60 (s), -74.86 (s) ppm.

HRMS (QToF): Calcd for $\text{C}_{26}\text{H}_{26}\text{F}_6\text{O}_7\text{SNa}$ $[\text{M} + \text{Na}]^+$ 619.1196; found 619.1201.

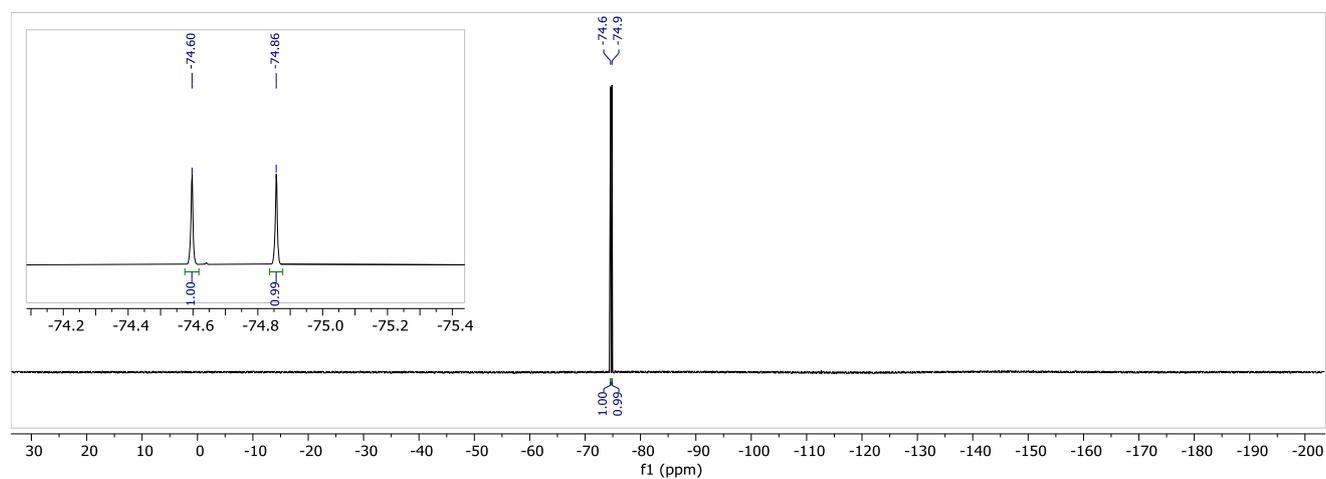
^1H NMR (400 MHz, CDCl_3) of **4,6TFA**:



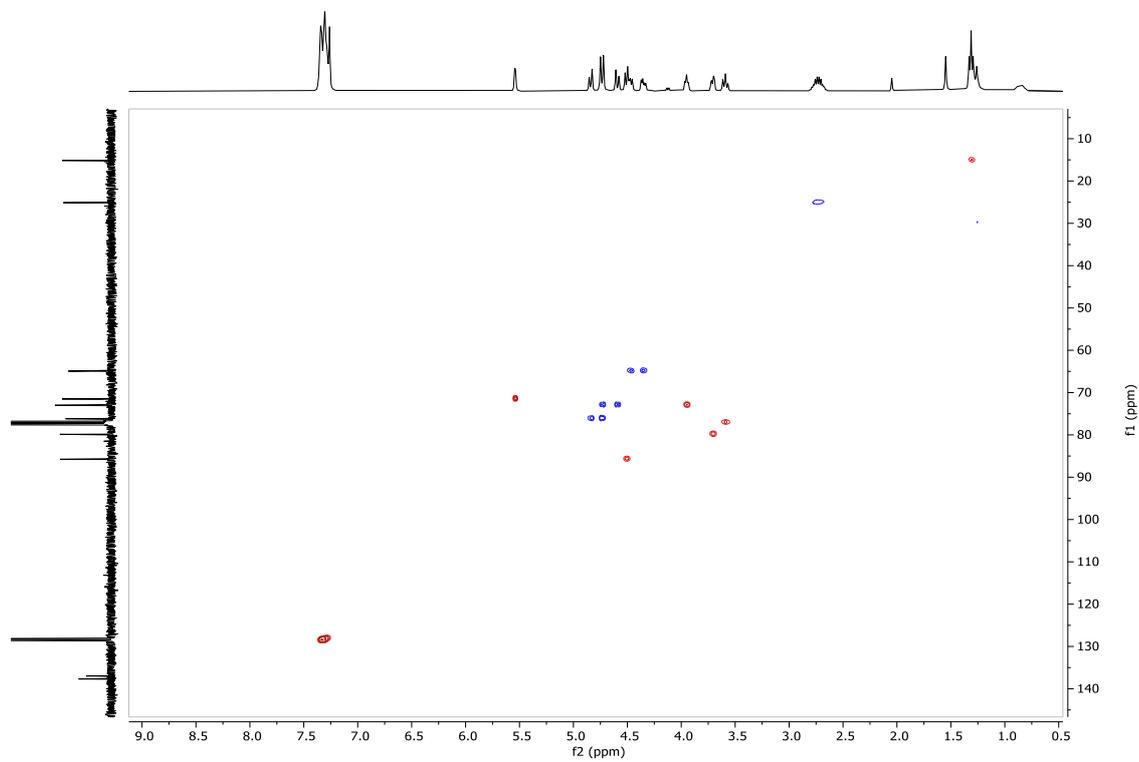
^{13}C NMR (101 MHz, CDCl_3) of **4,6TFA**:



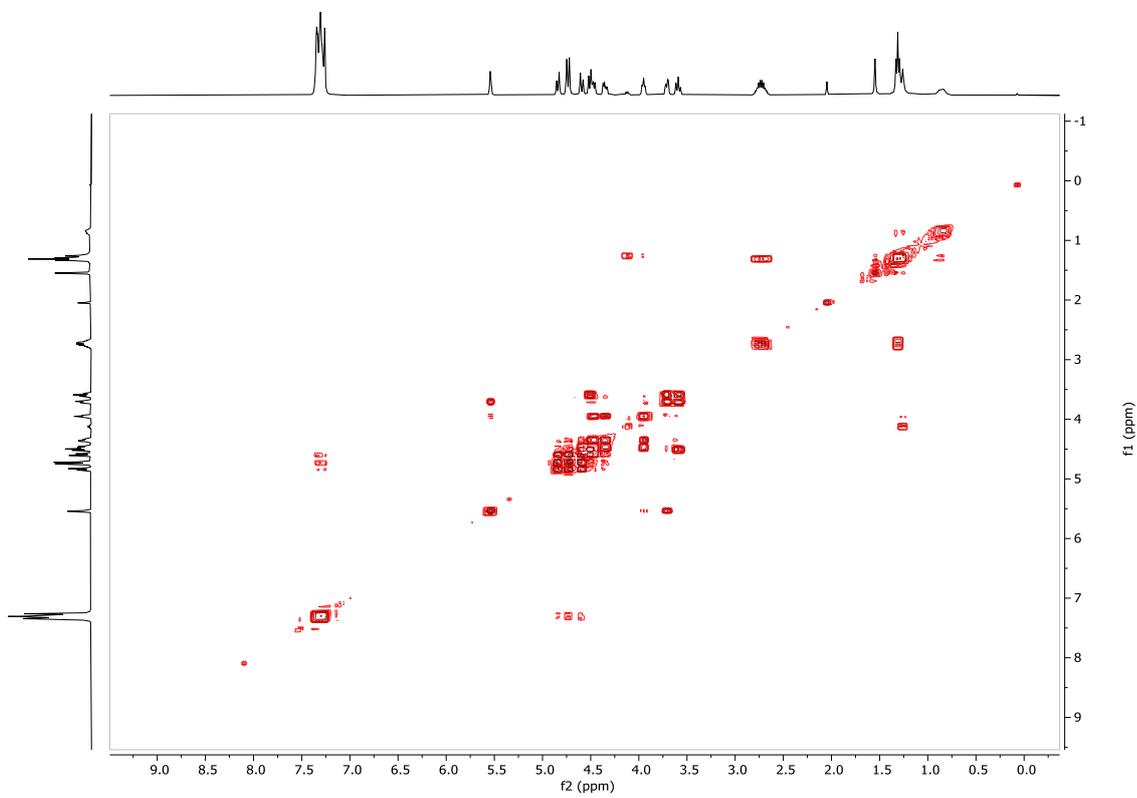
^{19}F NMR (376 MHz, CDCl_3) of **4,6TFA**:



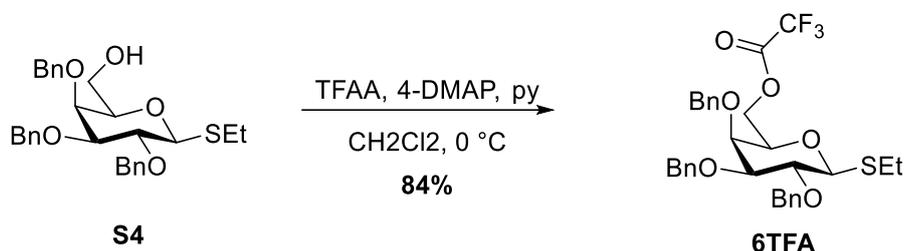
$^{13}\text{C}, ^1\text{H}$ HSQC of **4,6TFA**:



$^1\text{H}, ^1\text{H}$ COSY of **4,6TFA**:



Ethyl 2,3,4-tris-*O*-benzyl-6-(trifluoroacetate)-1-thio- β -D-galactopyranoside (**6TFA**)



A stirred solution of **S4** (100 mg, 0.20 mmol, 1.0 equiv.) in anhydrous CH_2Cl_2 (3 mL) was cooled to 0 °C and anhydrous pyridine (20 μL , 0.24 mmol, 1.2 equiv.), trifluoroacetic anhydride (34 μL , 0.24 mmol, 1.2 equiv.) and 4-DMAP (2 mg, 0.02 mmol, 0.1 equiv.) were added. The solution was stirred at 0 °C for 30 minutes before it was quenched with water (10 mL). The mixture was extracted with CH_2Cl_2 (3 x 10 mL) and the combined organic phases were washed with 1 N HCl (10 mL), dried over Na_2SO_4 , filtered and concentrated. Product **6TFA** (100 mg, 0.17 mmol, 84%) was obtained as a colorless solid after purification by column chromatography (SiO_2 , Hex/EtOAc = 9:1 to 7:3).

$R_f = 0.79$ (Hex/EtOAc 3:1).

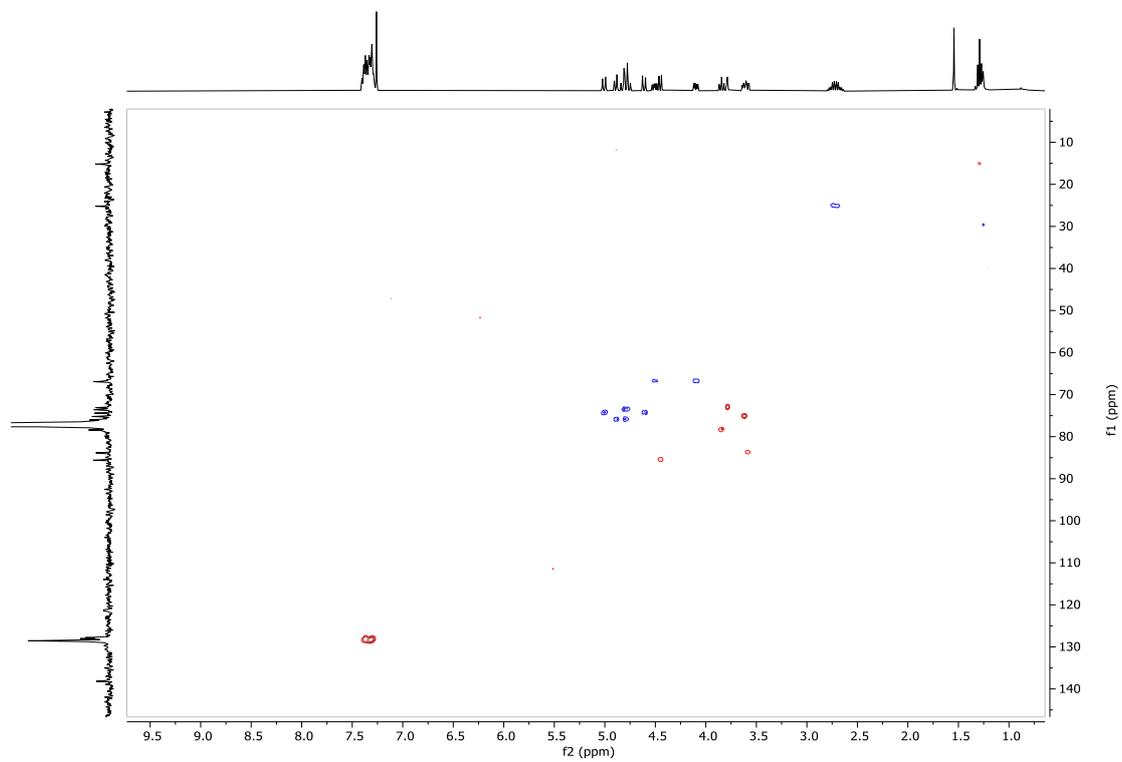
$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.43 – 7.26 (m, 15H), 5.03 – 4.58 (m, 6H), 4.45 (d, $J = 9.6$ Hz, 1H), 4.55 – 4.06 (m, 2H), 3.84 (t, $J = 9.4$ Hz, 1H), 3.81 – 3.77 (m, 1H), 3.66 – 3.55 (m, 2H), 2.82 – 2.62 (m, 2H), 1.29 (t, $J = 7.4$ Hz, 3H) ppm.

$^{13}\text{C NMR}$ (101 MHz, CDCl_3) δ 138.19, 128.70, 128.53, 128.15, 128.04, 127.81, 85.61, 83.87, 78.42, 76.02, 75.22, 74.41, 73.59, 73.11, 66.92, 25.20, 15.18 ppm.

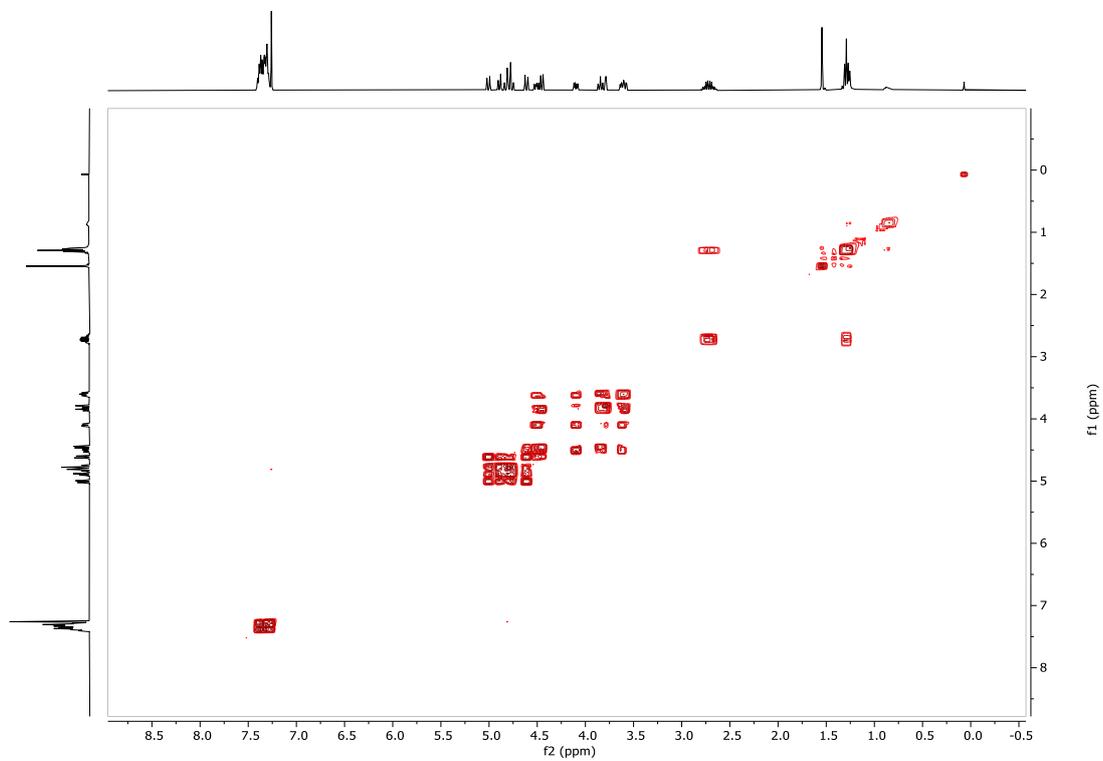
$^{19}\text{F NMR}$ (376 MHz, CDCl_3) δ -74.94 (s) ppm.

HRMS (QToF): Calcd for $\text{C}_{31}\text{H}_{33}\text{F}_3\text{O}_6\text{SNa}$ $[\text{M} + \text{Na}]^+$ 613.1842; found 613.1844.

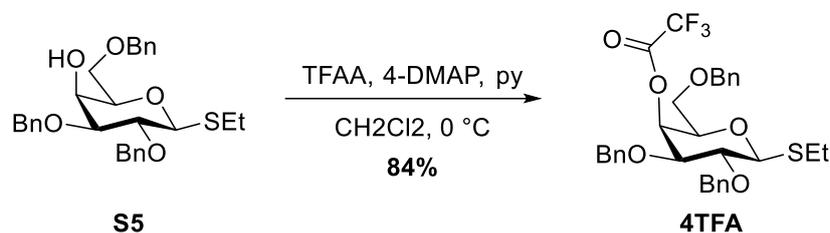
$^{13}\text{C}, ^1\text{H}$ HSQC of **6TFA**:



$^1\text{H}, ^1\text{H}$ COSY of **6TFA**:



Ethyl 2,3,6-tris-*O*-benzyl-4-(trifluoroacetate)-1-thio- β -D-galactopyranoside (**4TFA**)



A stirred solution of **S5** (130 mg, 0.30 mmol, 1.0 equiv.) in anhydrous CH_2Cl_2 (3 mL) was cooled to 0 °C and anhydrous pyridine (29 μL , 0.36 mmol, 1.2 equiv.), trifluoroacetic anhydride (51 μL , 0.36 mmol, 1.2 equiv.) and 4-DMAP (4 mg, 0.03 mmol, 0.1 equiv.) were added. The solution was stirred at 0 °C for 30 min before it was quenched with water (10 mL). The mixture was extracted with CH_2Cl_2 (3 x 10 mL) and the combined organic phases were washed with 1 N HCl (10 mL), dried over Na_2SO_4 , filtered and concentrated. Product **4TFA** (100 mg, 0.17 mmol, 84%) was obtained as a colorless oil after purification by column chromatography (SiO_2 , Hex/EtOAc = 9:1 to 7:3).

R_f = 0.69 (Hex/EtOAc 3:1).

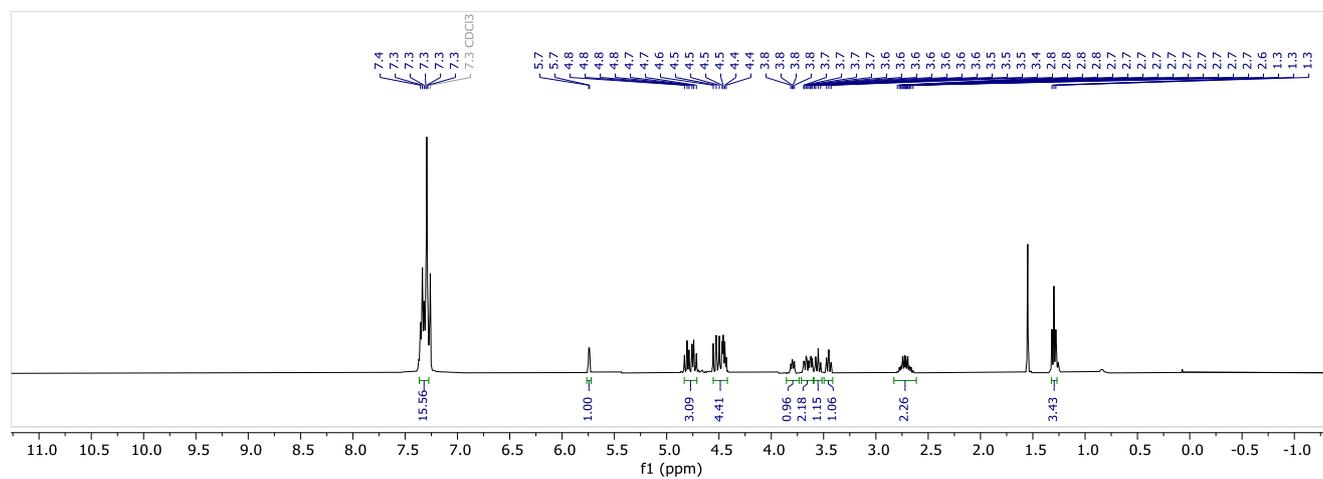
$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.36 – 7.27 (m, 15H), 5.74 (d, J = 3.1 Hz, 1H), 4.85 – 4.71 (m, 3H), 4.57 – 4.42 (m, 4H), 3.80 (dd, J = 8.3, 5.7 Hz, 1H), 3.71 – 3.60 (m, 2H), 3.55 (t, J = 9.4 Hz, 1H), 3.45 (t, J = 8.8 Hz, 1H), 2.82 – 2.62 (m, 2H), 1.30 (t, J = 7.5 Hz, 3H) ppm.

$^{13}\text{C NMR}$ (101 MHz, CDCl_3) δ 137.94, 137.37, 137.26, 128.69, 128.56, 128.48, 128.24, 128.14, 128.05, 128.03, 85.46, 80.42, 77.32, 76.07, 74.73, 73.98, 72.47, 71.52, 67.14, 24.83, 15.15 ppm.

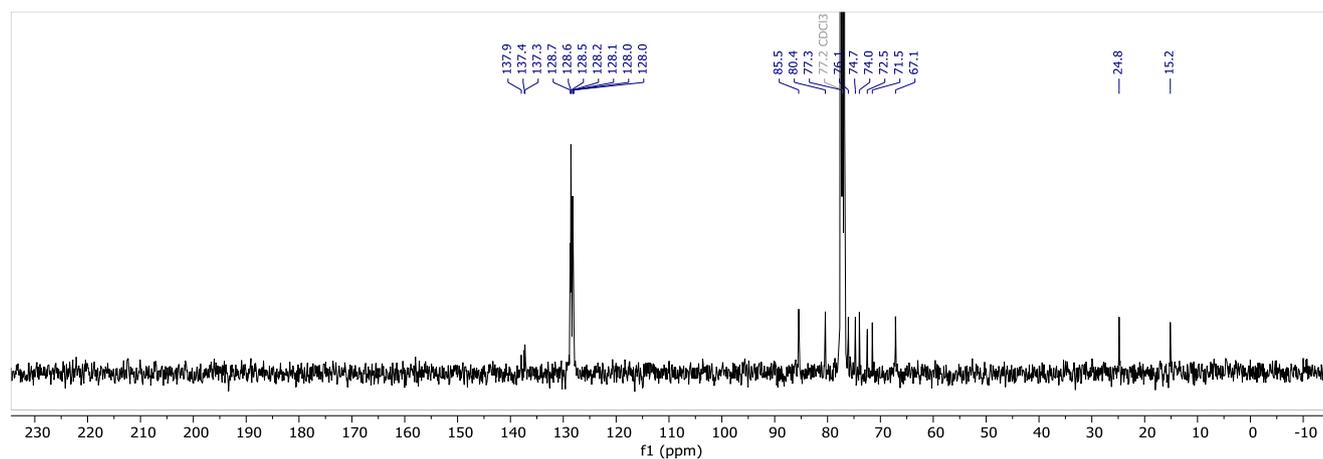
$^{19}\text{F NMR}$ (376 MHz, CDCl_3) δ -74.64 (s) ppm.

HRMS (QToF): Calcd for $\text{C}_{31}\text{H}_{33}\text{F}_3\text{O}_6\text{SNa}$ [$\text{M} + \text{Na}$] $^+$ 613.1842; found 613.1852.

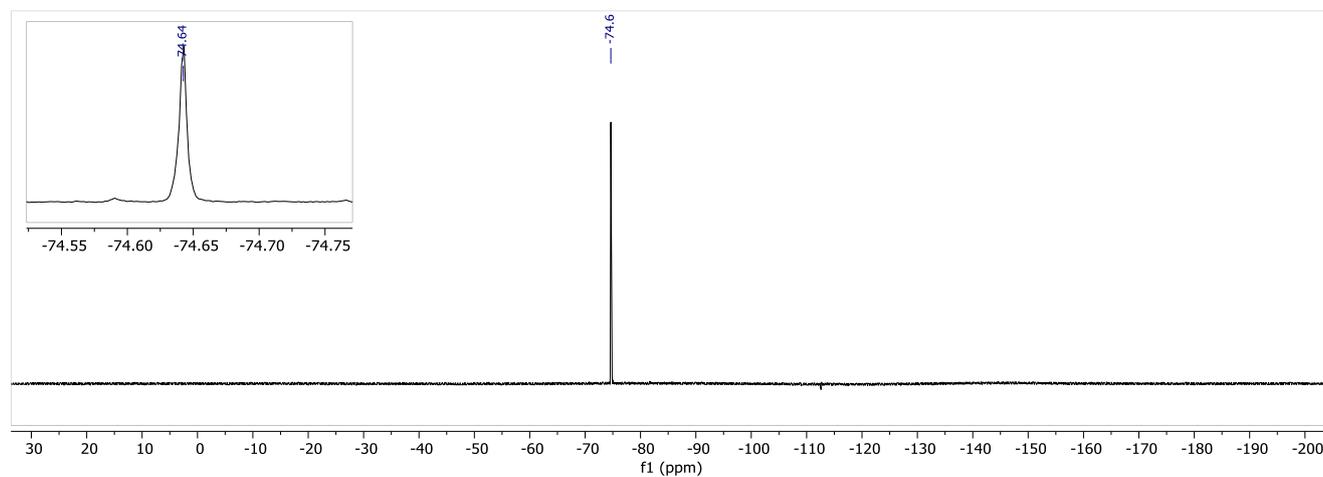
^1H NMR (400 MHz, CDCl_3) of **4TFA**:



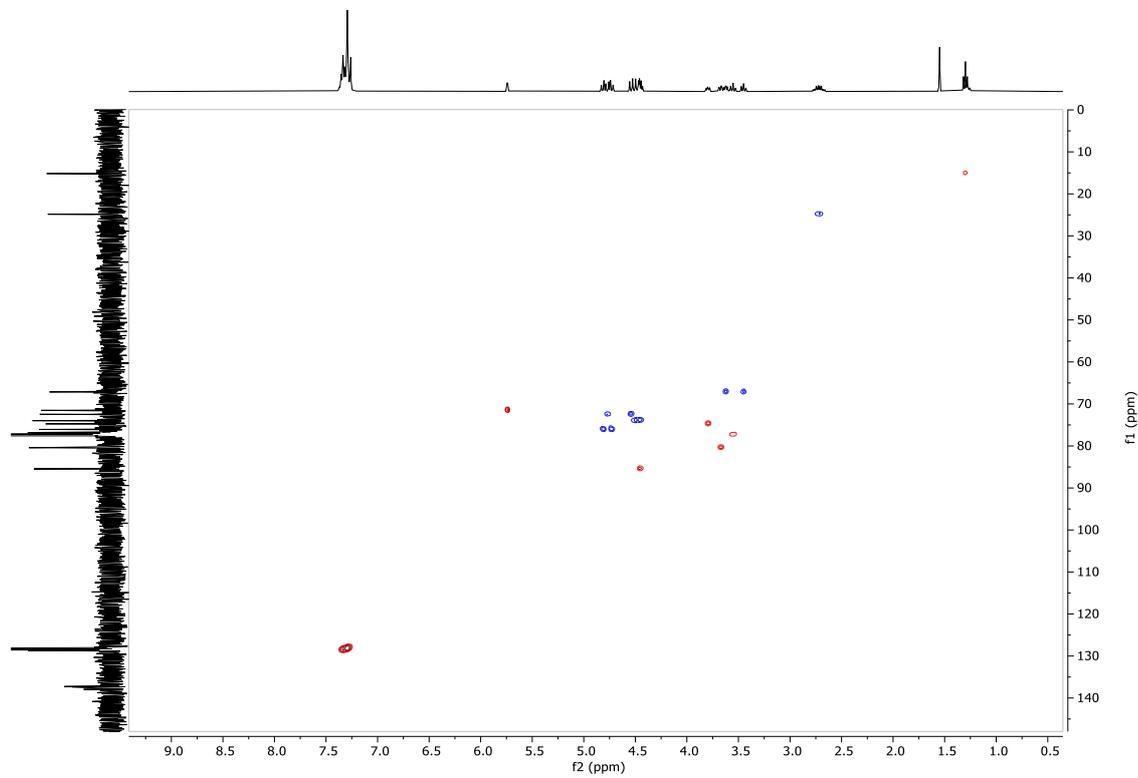
^{13}C NMR (101 MHz, CDCl_3) of **4TFA**:



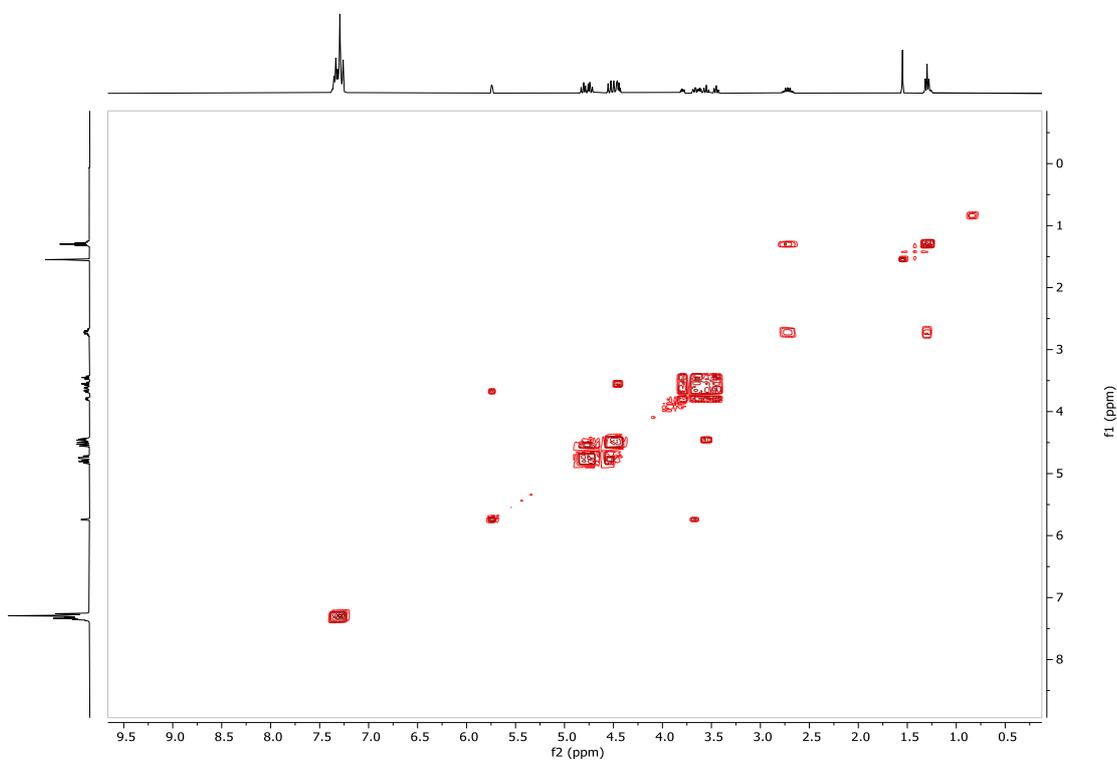
^{19}F NMR (376 MHz, CDCl_3) of **4TFA**:



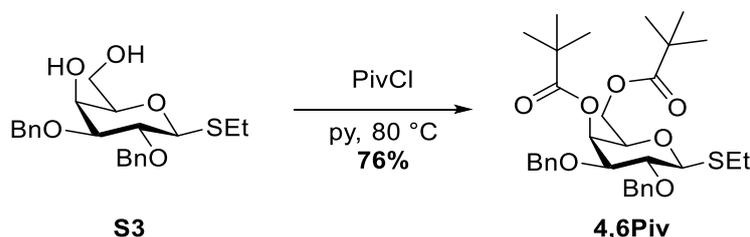
$^{13}\text{C}, ^1\text{H}$ HSQC of 4TFA:



$^1\text{H}, ^1\text{H}$ COSY of 4TFA:



Ethyl 2,3-bis-*O*-benzyl-4,6-bis(2,2-dimethylpropanoate)-1-thio- β -D-galactopyranoside (**4,6Piv**)



To a solution of **S3** (140 mg, 0.35 mmol, 1.0 equiv.) in anhydrous pyridine (3 mL) pivaloyl chloride (1.5 mL, 12.13 mmol, 35.0 equiv.) was added. The mixture was stirred at 80 °C for 2 hours. The volatiles were evaporated. Product **4,6Piv** (150 mg, 0.26 mmol, 76%) was obtained as a colorless oil after purification by column chromatography (SiO₂, Hex/EtOAc = 9:1).

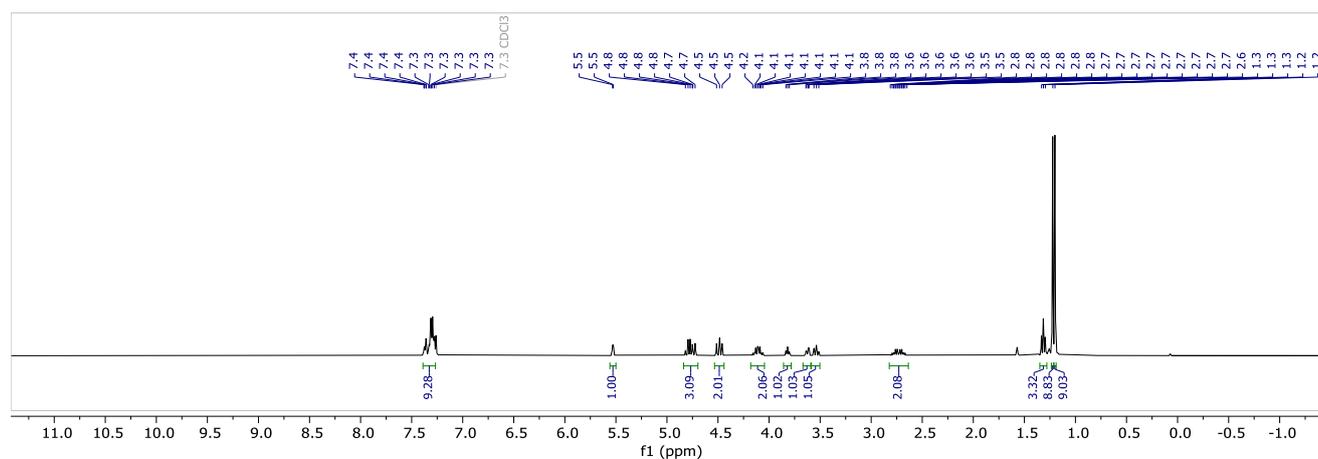
R_f = 0.47 (Hex/EtOAc 9:1).

¹H NMR (400 MHz, CDCl₃) δ 7.39 – 7.27 (m, 10H), 5.53 (d, J = 3.2 Hz, 1H), 4.84 – 4.70 (m, 3H), 4.53 – 4.44 (m, 2H), 4.11 (qd, J = 11.3, 6.8 Hz, 2H), 3.82 (t, J = 6.7 Hz, 1H), 3.62 (dd, J = 9.3, 3.2 Hz, 1H), 3.54 (t, J = 9.4 Hz, 1H), 2.84 – 2.63 (m, 2H), 1.31 (t, J = 7.4 Hz, 3H), 1.22 (s, 9H), 1.20 (s, 9H) ppm.

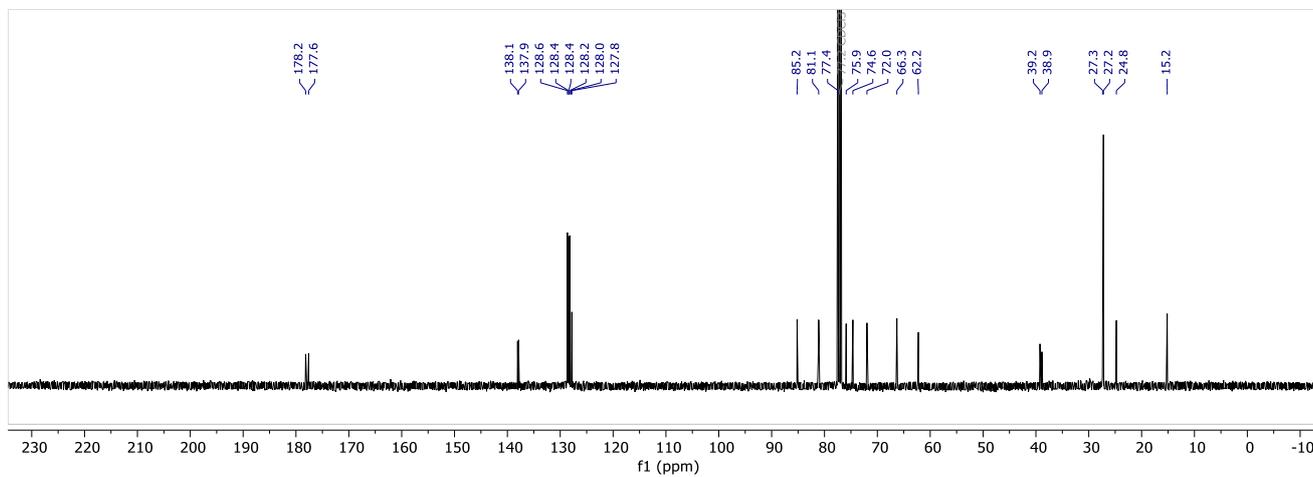
¹³C NMR (101 MHz, CDCl₃) δ 178.17, 177.60, 138.07, 137.89, 128.64, 128.43, 128.39, 128.20, 127.95, 127.81, 85.19, 81.10, 77.43, 75.90, 74.63, 71.96, 66.31, 62.24, 39.23, 38.87, 27.33, 27.23, 24.78, 15.16 ppm.

HRMS (QToF): Calcd for C₃₂H₄₄O₇SNa [M + Na]⁺ 595.2700; found 595.2701.

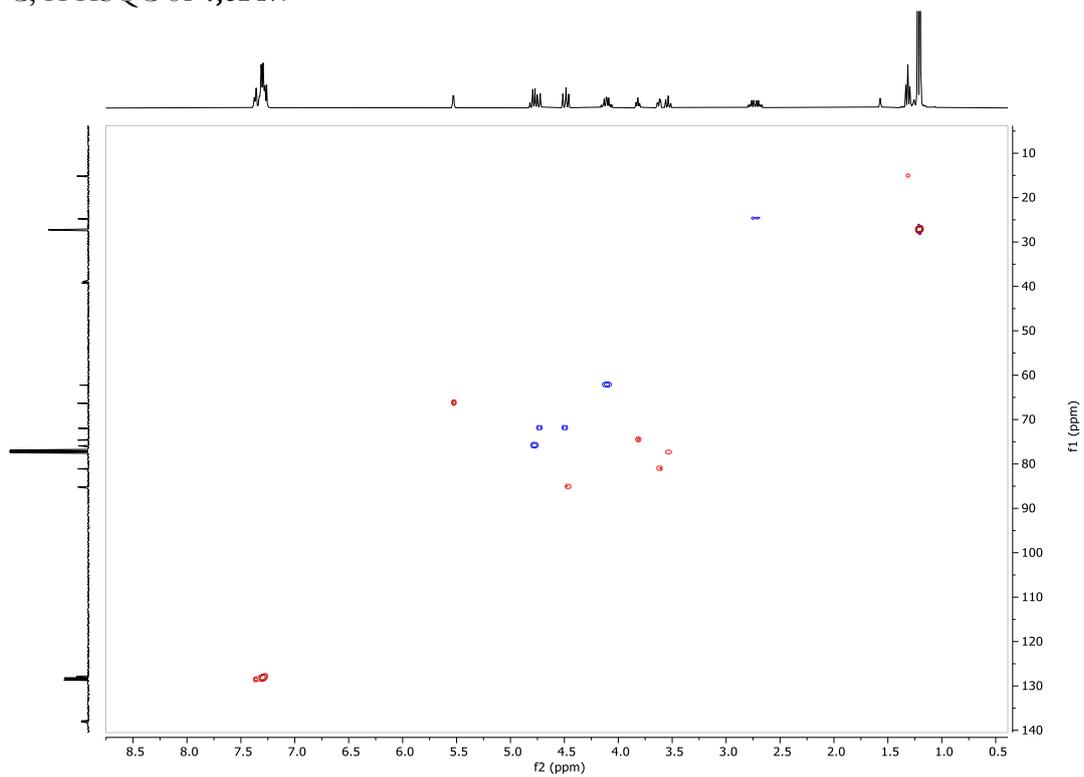
¹H NMR (400 MHz, CDCl₃) of **4,6Piv**:



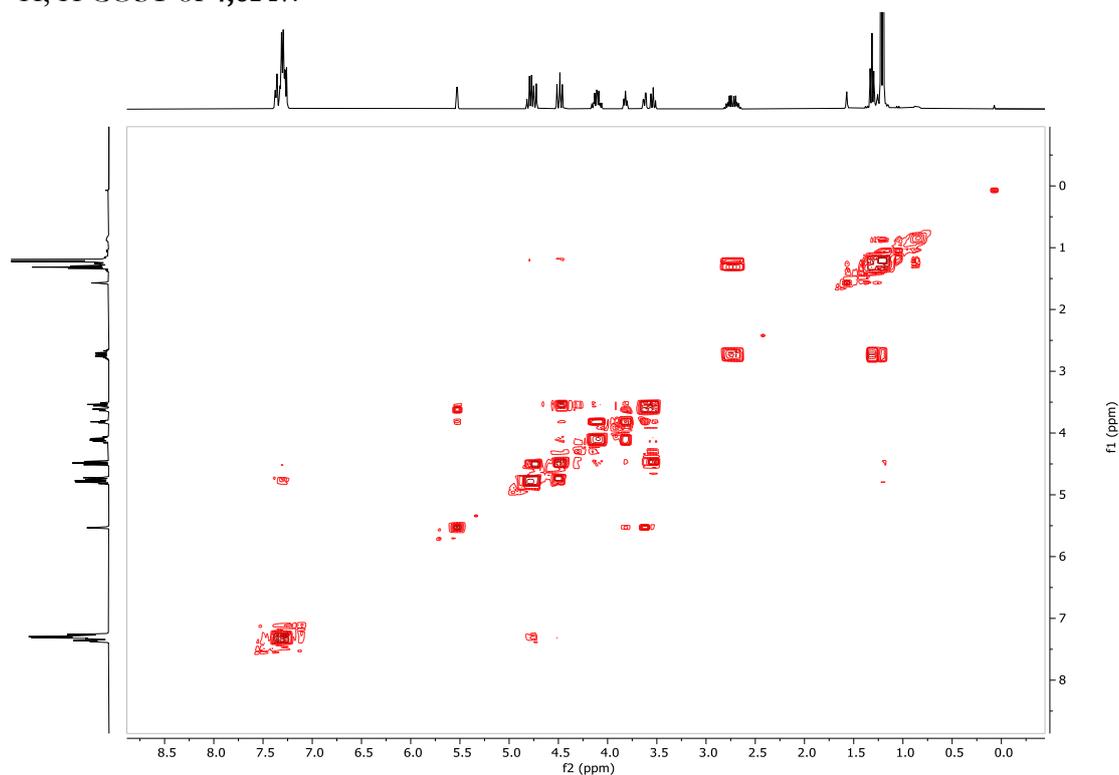
^{13}C NMR (101 MHz, CDCl_3) of **4,6Piv**:



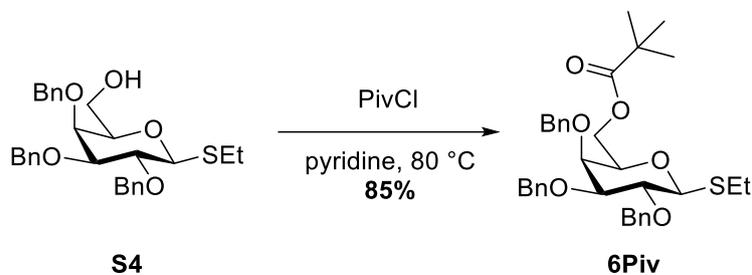
$^{13}\text{C}, ^1\text{H}$ HSQC of **4,6Piv**:



$^1\text{H},^1\text{H}$ COSY of **4,6Piv**:



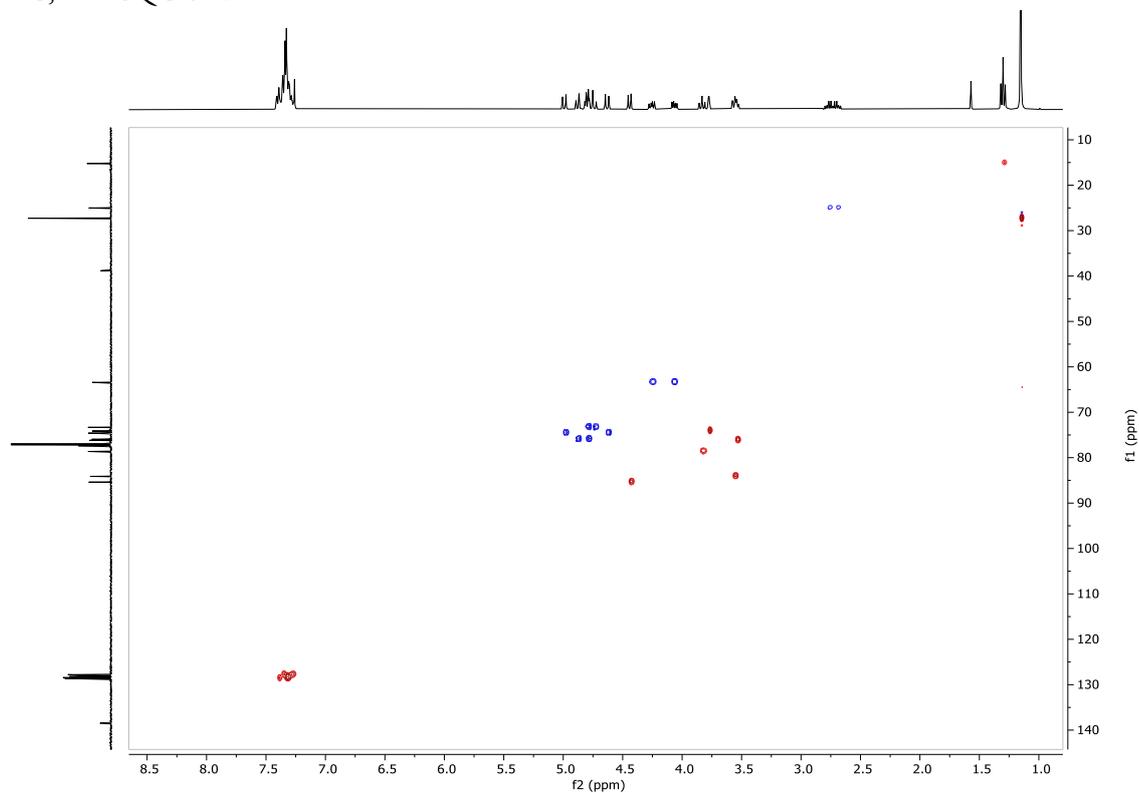
Ethyl 2,3,4-tris-*O*-benzyl-6-(2,2-dimethylpropanoate)-1-thio- β -D-galactopyranoside (**6Piv**)



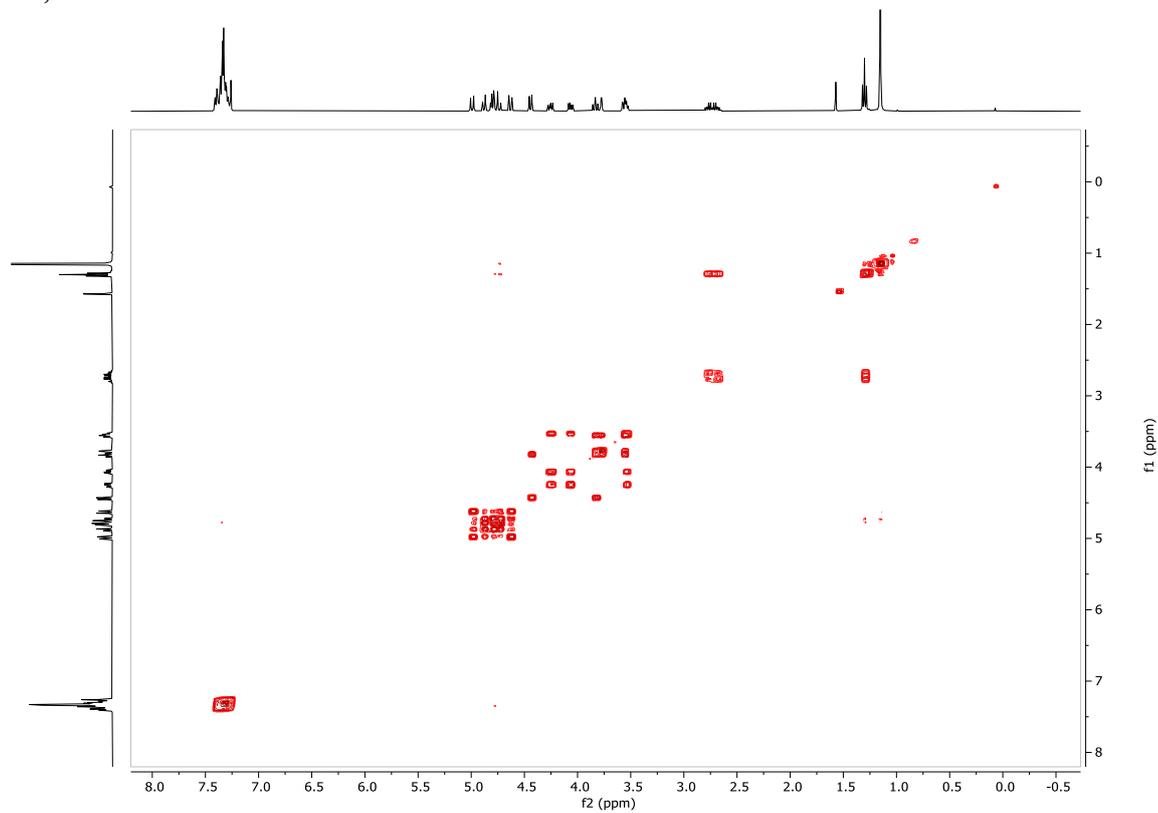
To a solution of **S4** (150 mg, 0.30 mmol, 1.0 equiv.) in anhydrous pyridine (3 mL) pivaloyl chloride (0.75 mL, 6.06 mmol, 20.0 equiv.) was added. The mixture was stirred at 80 °C for 2 hours. The volatiles were evaporated. Product **6Piv** (150 mg, 0.26 mmol, 85%) was obtained as a colorless oil after purification by column chromatography (SiO_2 , Hex/EtOAc = 9:1).

R_f = 0.53 (Hex/EtOAc 3:1).

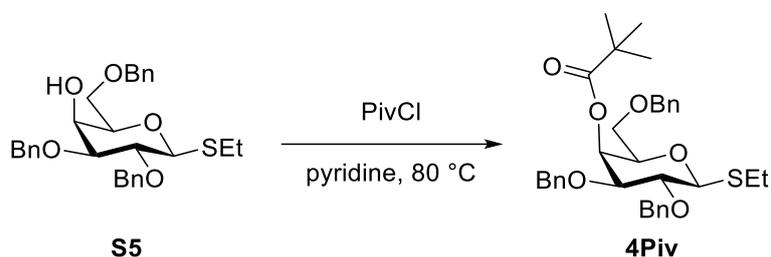
$^{13}\text{C}, ^1\text{H}$ HSQC of **6Piv**:



$^1\text{H}, ^1\text{H}$ COSY of **6Piv**:



Ethyl 2,3,6-tris-*O*-benzyl-4-(2,2-dimethylpropanoate)-1-thio- β -D-galactopyranoside (**4Piv**)



To a solution of **S5** (100 mg, 0.20 mmol, 1.0 equiv.) in anhydrous pyridine (1.5 mL) pivaloyl chloride (0.50 mL, 6.06 mmol, 20.0 equiv.) was added. The mixture was stirred at 80 °C for 2 hours. The volatiles were evaporated. Product **4Piv** (110 mg, 0.19 mmol, 94%) was obtained as a colorless oil after purification by column chromatography (SiO₂, Hex/EtOAc = 9:1).

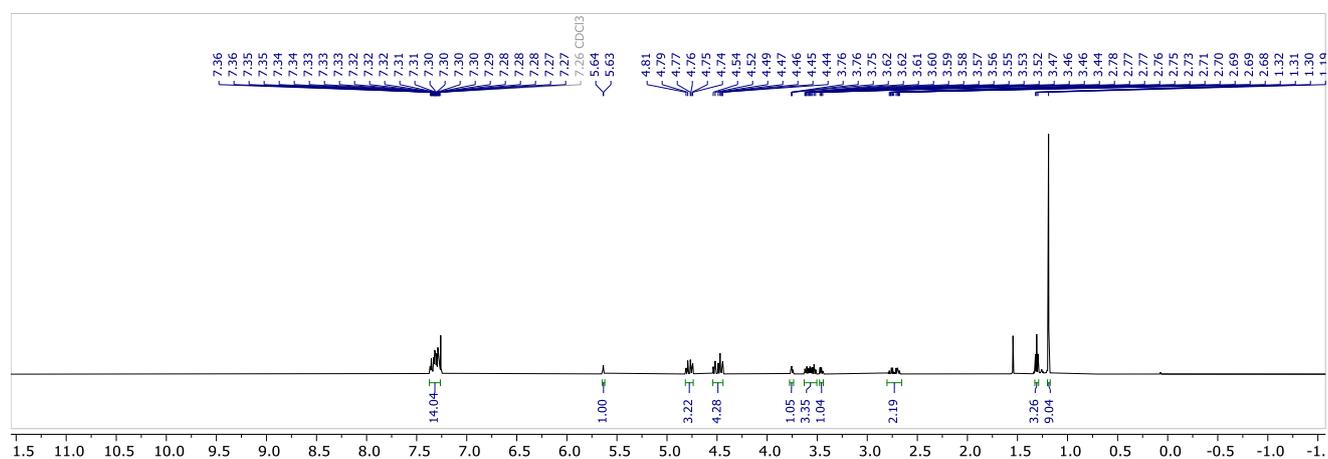
R_f = 0.49 (Hex/EtOAc 3:1).

¹H NMR (600 MHz, CDCl₃) δ 7.37 – 7.26 (m, 14H), 5.64 (d, J = 2.3 Hz, 1H), 4.84 – 4.71 (m, 3H), 4.58 – 4.39 (m, 4H), 3.75 (d, J = 1.0 Hz, 1H), 3.64 – 3.50 (m, 3H), 3.46 (dd, J = 9.5, 6.9 Hz, 1H), 2.81 – 2.65 (m, 2H), 1.31 (t, J = 7.4 Hz, 3H), 1.19 (s, 9H) ppm.

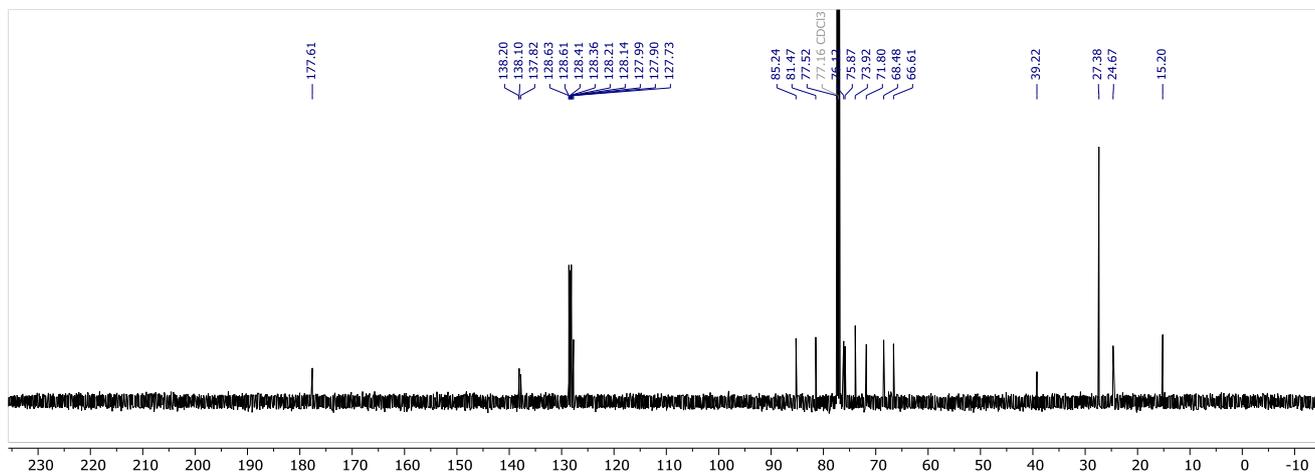
¹³C NMR (151 MHz, CDCl₃) δ 177.61, 138.20, 138.10, 137.82, 128.63, 128.61, 128.41, 128.36, 128.21, 128.14, 127.99, 127.90, 127.73, 85.24, 81.47, 77.52, 76.12, 75.87, 73.92, 71.80, 68.48, 66.61, 39.22, 27.38, 24.67, 15.20 ppm.

HRMS (QToF): Calcd for C₃₄H₄₂O₆SNa [M + Na]⁺ 601.2594; found 601.2614.

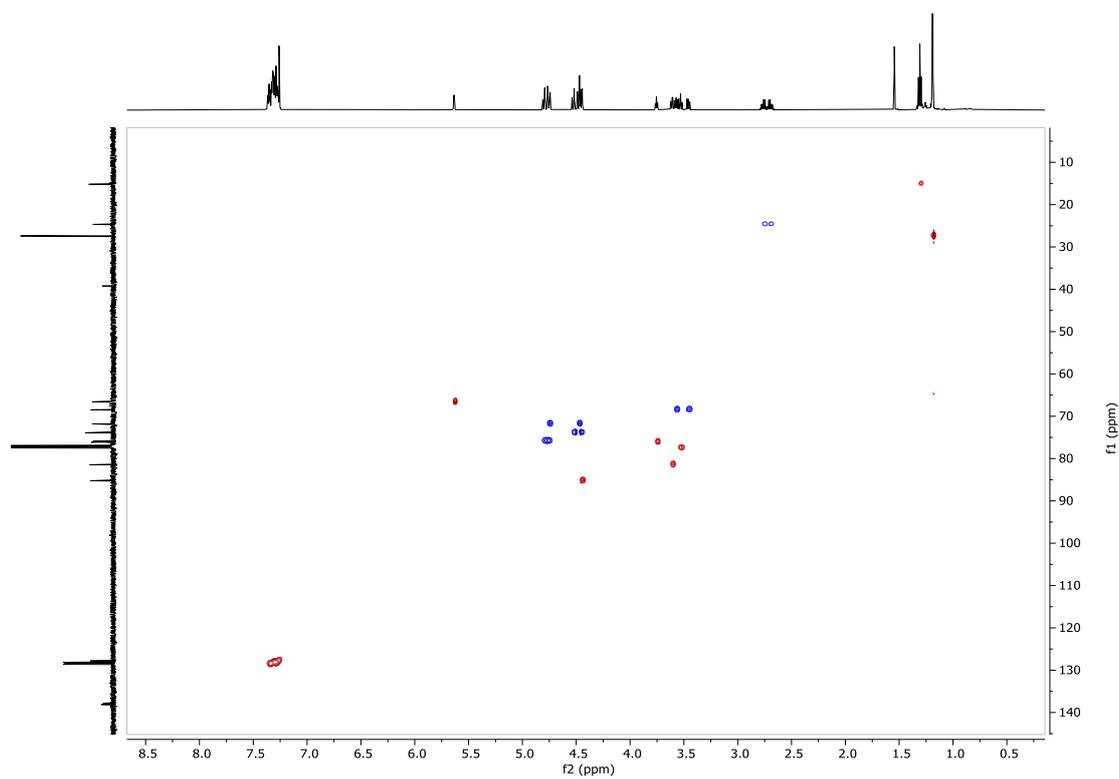
¹H NMR (400 MHz, CDCl₃) of **4Piv**:



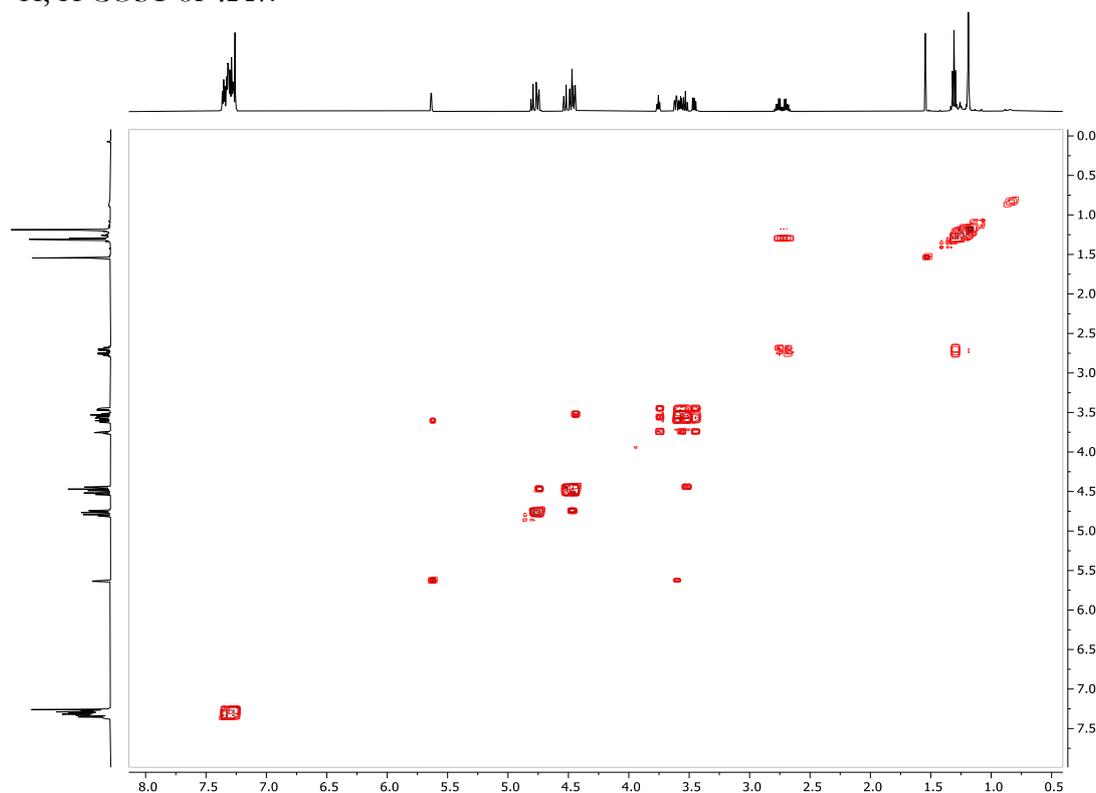
^{13}C NMR (101 MHz, CDCl_3) of **4Piv**:



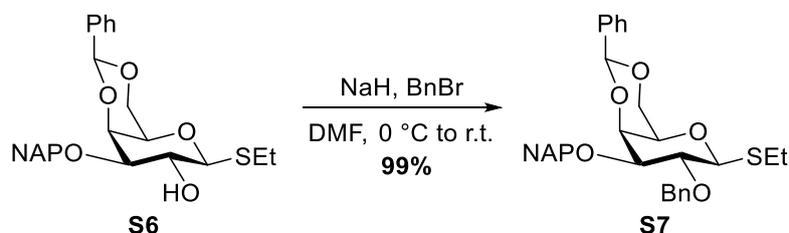
$^{13}\text{C}, ^1\text{H}$ HSQC of **4Piv**:



$^1\text{H},^1\text{H}$ COSY of **4Piv**:



Ethyl 2-*O*-benzyl-3-*O*-(2-naphthalenylmethyl)-4,6-*O*-[(*S*)-phenylmethylene]-1-thio- β -D-galactopyranoside²¹ (**S7**)



Compound **S6**²¹ (7.4 g, 16.4 mmol, 1.0 equiv.) was dissolved in anhydrous DMF (75 mL). The stirred solution was cooled to 0 °C and sodium hydride (1.5 g, 24.6 mmol; 60% dispersion in mineral oil, 2.3 equiv.) was added in small portions. After 30 min, benzyl bromide (1.3 mL, 37.7 mmol, 1.5 equiv.) was added dropwise. The reaction mixture was allowed to warm up to room temperature and was stirred overnight. Methanol (10 mL) was added, the reaction mixture was stirred for 10 min and afterwards diluted with ethyl acetate (100 mL). The organic layer was washed with water (2 x 100 mL). The aqueous phase was extracted with ethyl acetate (2 x 100 mL). The combined organic phase was washed with water (100 mL), dried over Na_2SO_4 and concentrated.

Product **S7** (8.8 g, 16.2 mmol, **99%**) was obtained as a colorless solid after purification by column chromatography (SiO₂, Hex/EtOAc = 3:1).

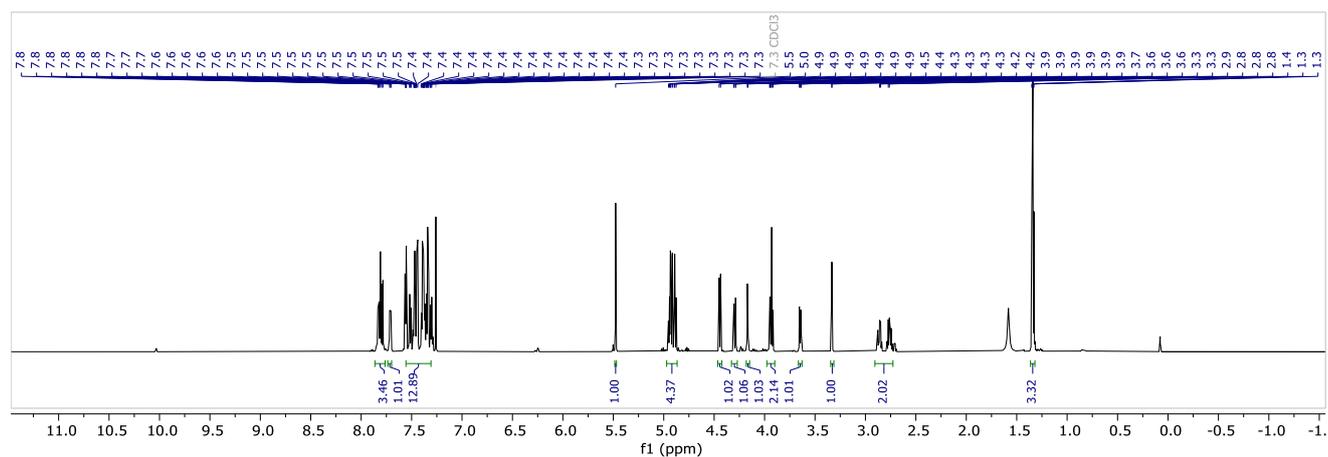
$R_f = 0.22$ (Hex/EtOAc 3:1).

¹H NMR (700 MHz, CDCl₃) δ 7.87 – 7.78 (m, 3H), 7.74 – 7.69 (m, 1H), 7.57 – 7.29 (m, 13H), 5.48 (s, 1H), 4.96 – 4.87 (m, 4H), 4.44 (d, $J = 9.7$ Hz, 1H), 4.30 (dd, $J = 12.3, 1.8$ Hz, 1H), 4.17 (d, $J = 3.7$ Hz, 1H), 3.96 – 3.90 (m, 2H), 3.65 (dd, $J = 9.2, 3.5$ Hz, 1H), 3.35 – 3.32 (m, 1H), 2.90 – 2.73 (m, 2H), 1.34 (t, $J = 7.5$ Hz, 3H) ppm.

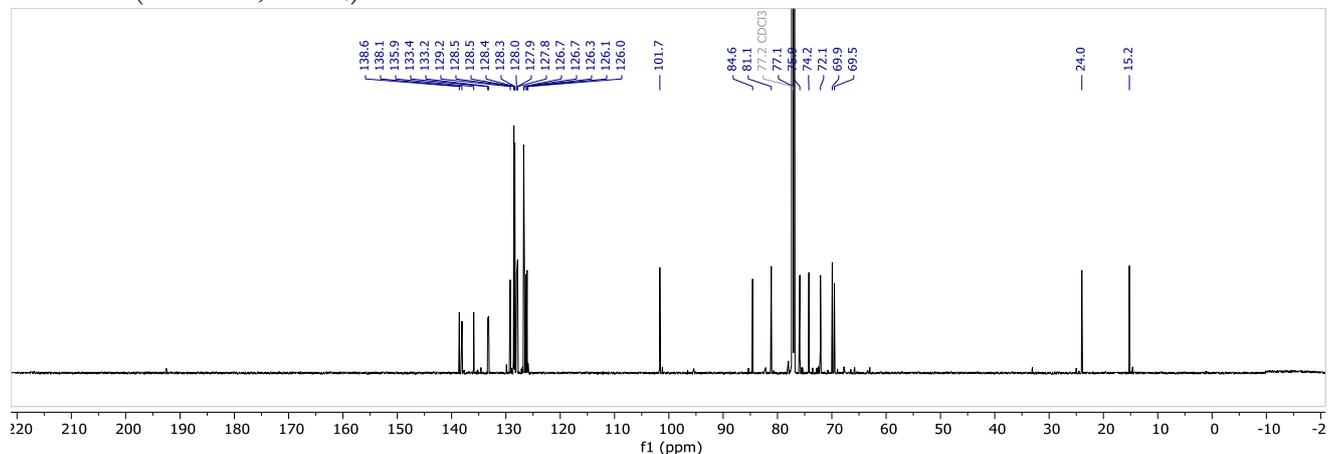
¹³C NMR (176 MHz, CDCl₃) δ 138.6, 138.1, 135.9, 133.4, 133.2, 129.2, 128.5, 128.5, 128.4, 128.3, 128.0, 127.9, 127.8, 126.7, 126.7, 126.3, 126.1, 126.0, 101.7, 84.6, 81.1, 77.1, 75.9, 74.2, 72.1, 69.9, 69.5, 24.0, 15.2 ppm.

HRMS (QToF): Calcd for C₃₃H₃₄O₅Na [M + Na]⁺ 565.2019; found 565.2018.

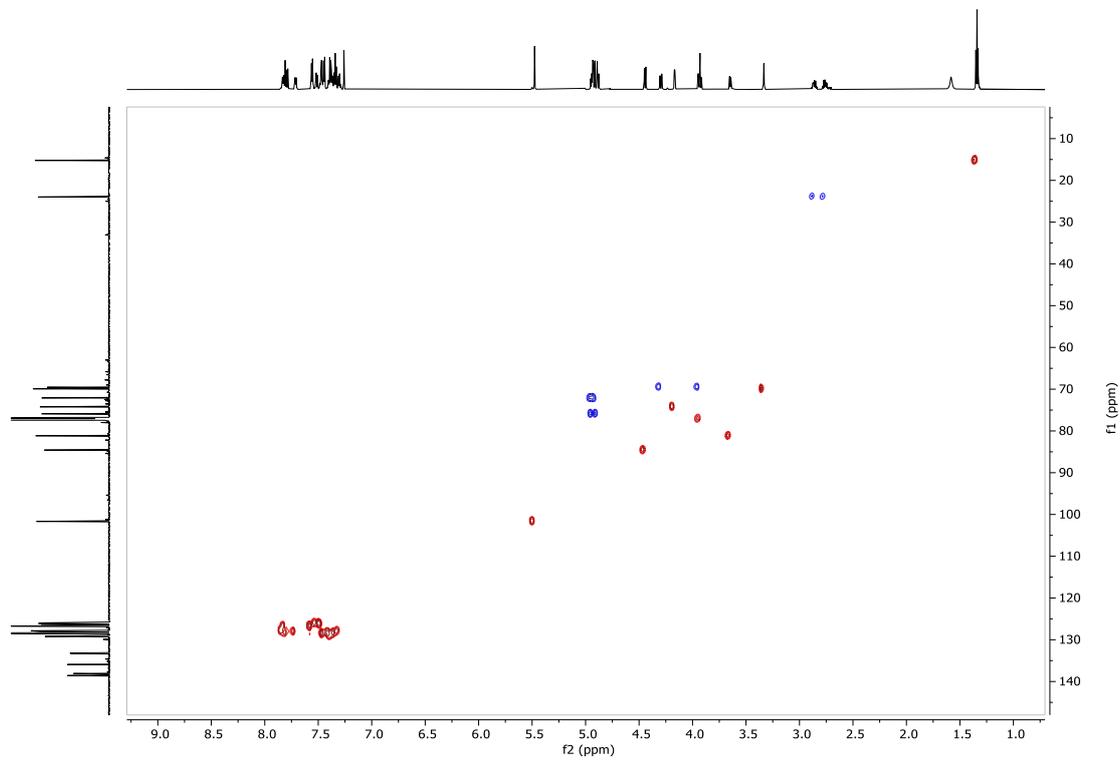
¹H NMR (400 MHz, CDCl₃) of **S7**:



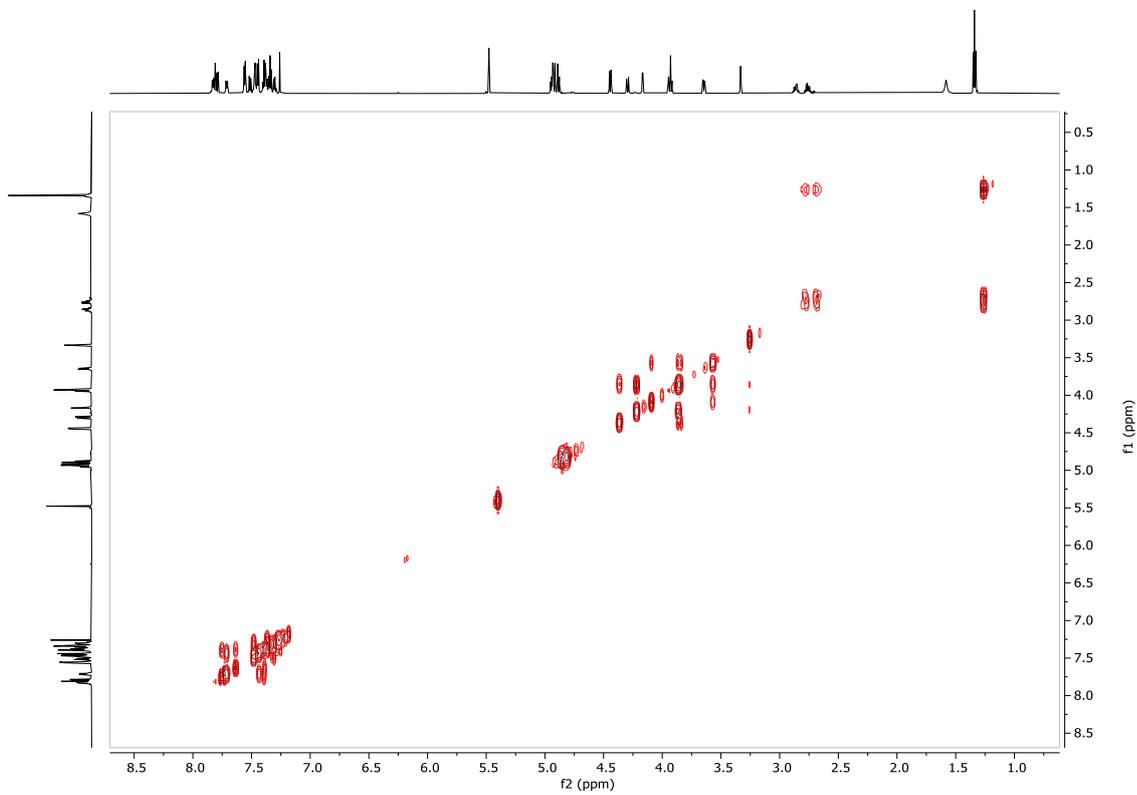
¹³C NMR (101 MHz, CDCl₃) of **S7**:



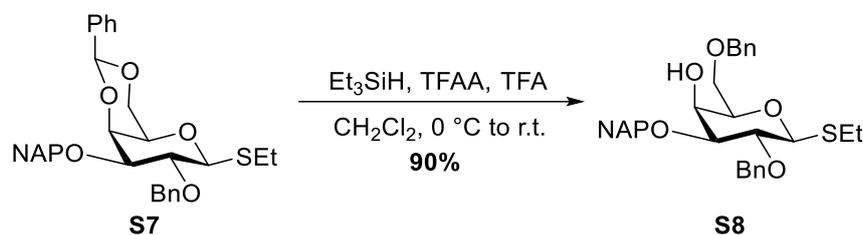
$^{13}\text{C}, ^1\text{H}$ HSQC of **S7**:



$^1\text{H}, ^1\text{H}$ COSY of **S7**:



Ethyl 2,6-bis-*O*-benzyl-3-*O*-(2-naphthalenylmethyl)-1-thio- β -D-galactopyranoside²² (**S8**)



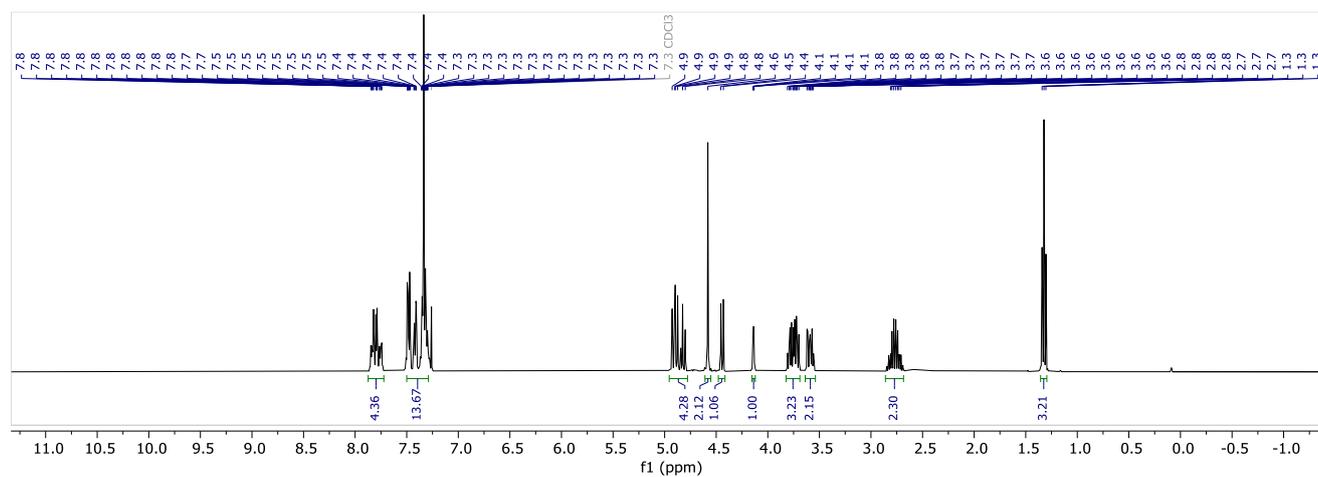
Compound **S7** (2.0 g, 3.7 mmol, 1.0 equiv.) was co-evaporated with anhydrous toluene (2 x 3 mL), and dissolved in anhydrous CH_2Cl_2 (20 mL). Triethylsilane (3.5 mL, 22.1 mmol, 6.0 equiv.) and trifluoroacetic anhydride (0.52 mL, 3.7 mmol, 1.0 equiv.) were added and the solution was cooled to $0\text{ }^\circ\text{C}$. Trifluoroacetic acid (1.7 mL, 22.1 mmol, 6.0 equiv.) was added dropwise. The mixture was allowed to warm up to room temperature and was stirred for 5 h. The solution was diluted with CH_2Cl_2 and quenched with saturated aqueous NaHCO_3 (40 mL). The aqueous phase was extracted with CH_2Cl_2 (2 x 60 mL) and the combined organic phase was washed with water (60 mL), dried over Na_2SO_4 , filtered and concentrated. Product **S8** (1.8 g, 3.3 mmol, 90%) was obtained as a colorless syrup after purification by column chromatography (SiO_2 , Hex/EtOAc = 3:1 to 1:1).

$R_f = 0.39$ (Hex/EtOAc 1:1).

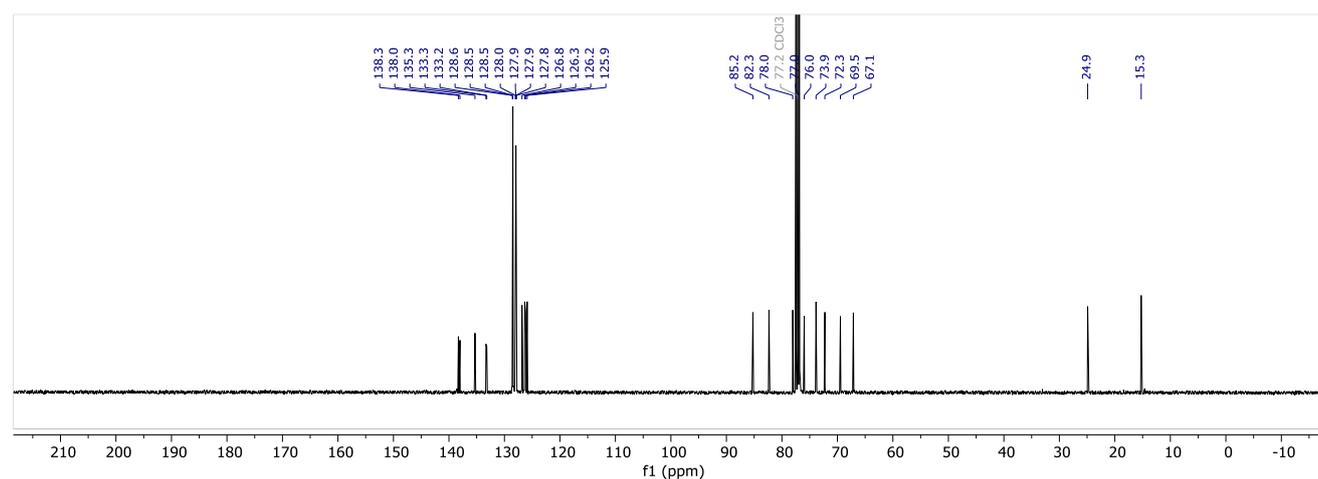
$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.88 – 7.71 (m, 4H), 7.53 – 7.27 (m, 13H), 4.95 – 4.77 (m, 4H), 4.58 (s, 2H), 4.44 (d, $J = 9.8$ Hz, 1H), 4.14 (dd, $J = 3.3, 1.1$ Hz, 1H), 3.83 – 3.68 (m, 3H), 3.64 – 3.52 (m, 2H), 2.87 – 2.66 (m, 2H), 1.32 (t, $J = 7.5$ Hz, 3H) ppm.

$^{13}\text{C NMR}$ (101 MHz, CDCl_3) δ 138.3, 138.0, 135.3, 133.3, 133.2, 128.6, 128.5, 128.5, 128.0, 127.9, 127.9, 127.8, 126.8, 126.3, 126.2, 125.9, 85.2, 82.3, 78.0, 77.0, 76.0, 73.9, 72.3, 69.5, 67.1, 24.9, 15.3 ppm.

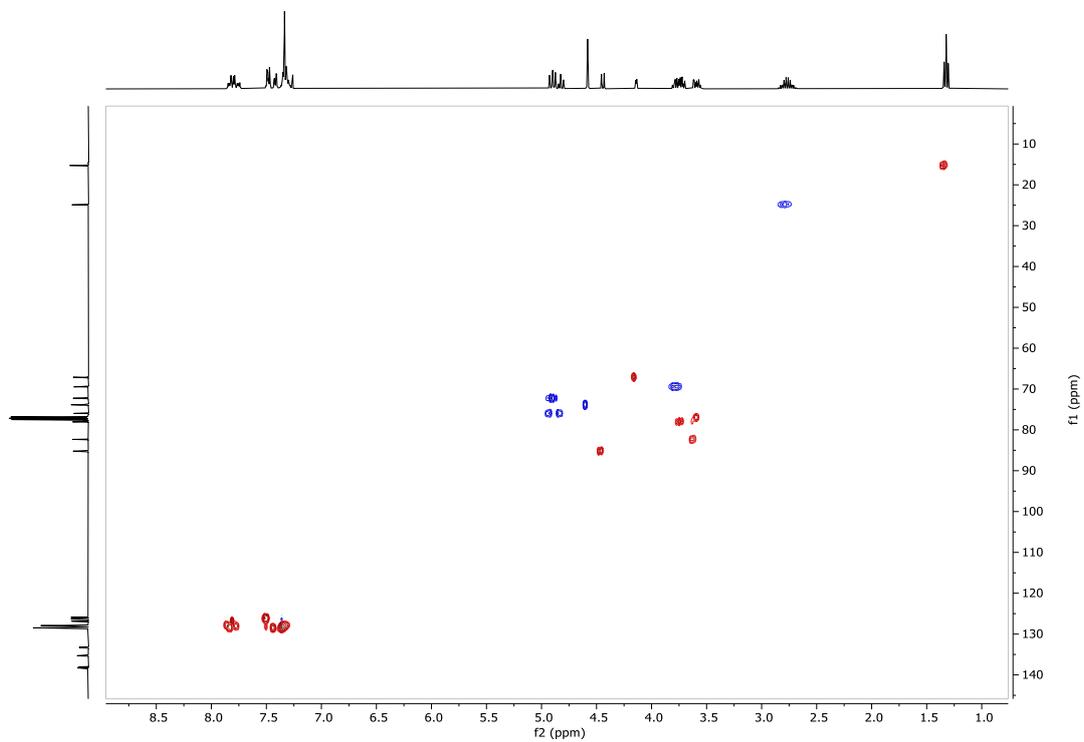
^1H NMR (400 MHz, CDCl_3) of **S8**:



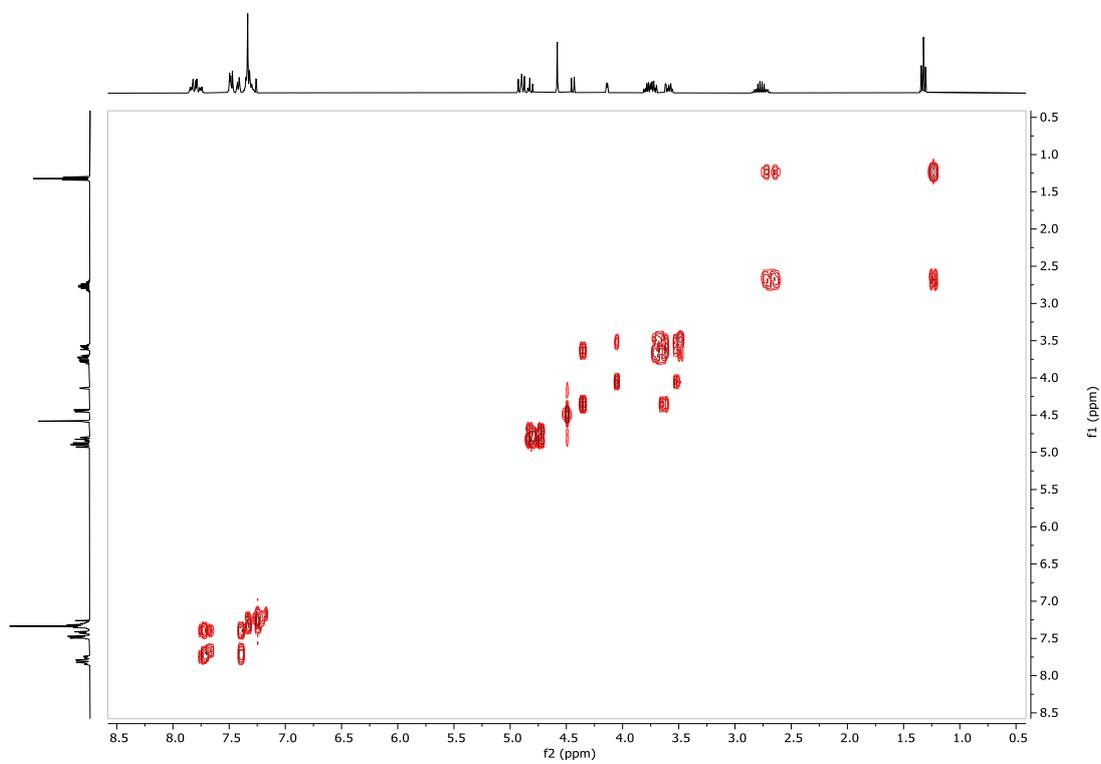
^{13}C NMR (101 MHz, CDCl_3) of **S8**:



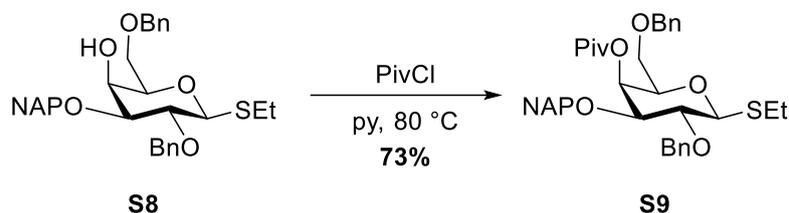
$^{13}\text{C}, ^1\text{H}$ HSQC of **S8**:



$^1\text{H}, ^1\text{H}$ COSY of **S8**:



Ethyl 2,6-bis-*O*-benzyl-3-*O*-(2-naphthalenylmethyl)-4-(2,2-dimethylpropanoate)-1-thio- β -D-galactopyranoside (S9**)**



To a solution of **S8** (1.3 g, 2.4 mmol, 1.0 equiv.) in anhydrous pyridine (15 mL) pivaloyl chloride (1.5 mL, 12.0 mmol, 5.0 equiv.) was added. The mixture was stirred at 80 °C for 2 hours. The volatiles were evaporated. Product **S9** (1.1 g, 1.7 mmol, **73%**) was obtained as a colorless oil after purification by column chromatography (SiO₂, Hex/EtOAc = 9:1 to 3:1).

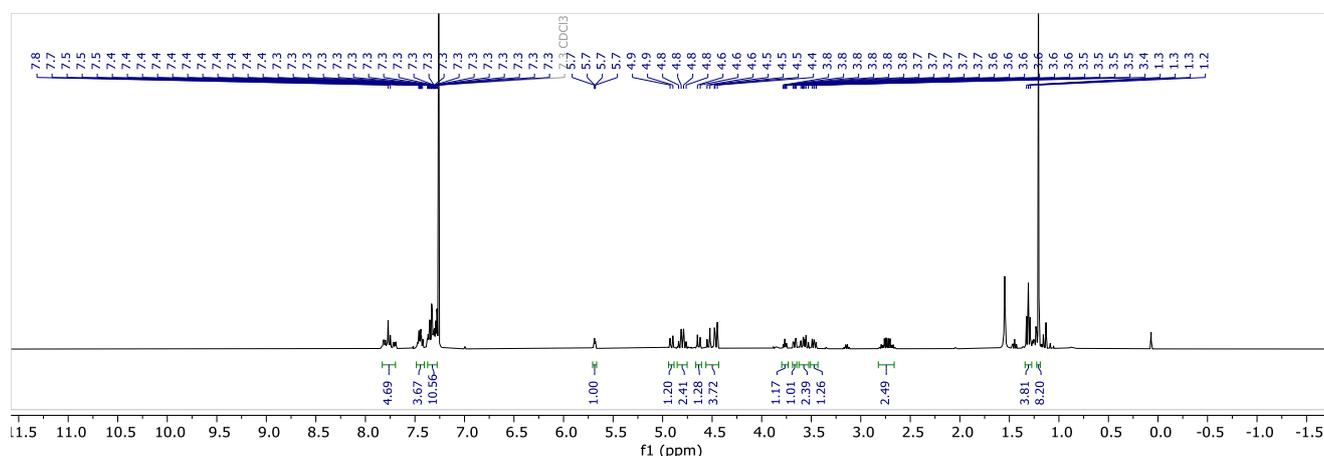
R_f = 0.43 (Hex/EtOAc 3:1).

¹H NMR (400 MHz, CDCl₃) δ 7.84 – 7.68 (m, 4H), 7.50 – 7.41 (m, 3H), 7.39 – 7.27 (m, 10H), 5.68 (dd, J = 3.3, 1.1 Hz, 1H), 4.91 (d, J = 11.2 Hz, 1H), 4.80 (q, J = 10.4 Hz, 2H), 4.63 (d, J = 11.2 Hz, 1H), 4.57 – 4.44 (m, 4H), 3.77 (ddd, J = 7.2, 5.8, 1.1 Hz, 1H), 3.67 (dd, J = 9.1, 3.3 Hz, 1H), 3.61 – 3.52 (m, 2H), 3.47 (dd, J = 9.5, 7.1 Hz, 1H), 1.31 (t, J = 7.5 Hz, 3H), 1.21 (s, 9H) ppm.

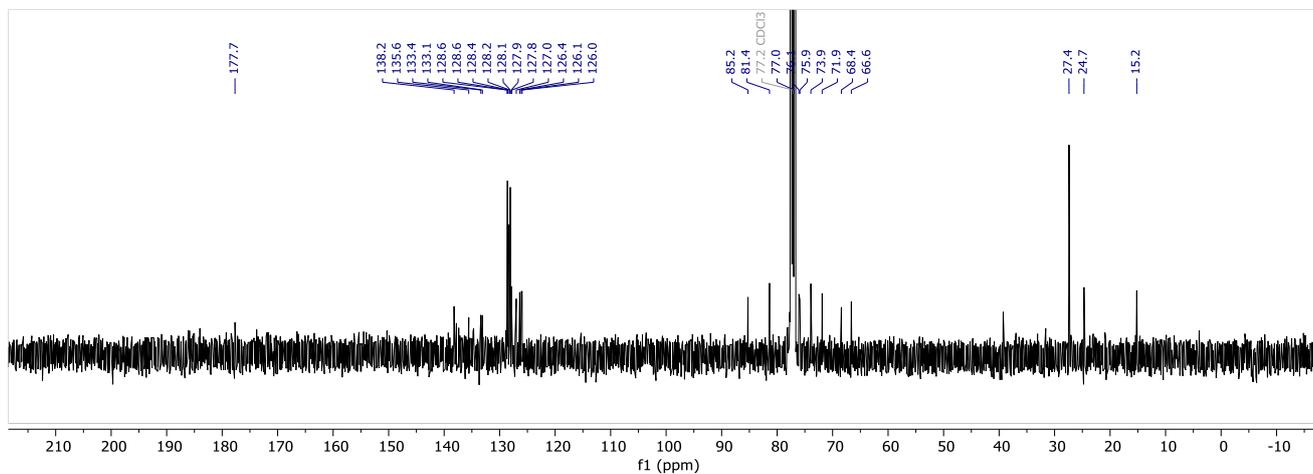
¹³C NMR (101 MHz, CDCl₃) δ 177.7, 138.2, 135.6, 133.4, 133.1, 128.6, 128.6, 128.4, 128.2, 128.1, 127.9, 127.8, 127.0, 126.4, 126.1, 126.0, 85.2, 81.4, 77.0, 76.1, 75.9, 73.9, 71.9, 68.4, 66.6, 27.4, 24.7, 15.2 ppm.

HRMS (QToF): Calcd for C₃₈H₄₄O₆SNa [M + Na]⁺ 651.2751; found 651.2740.

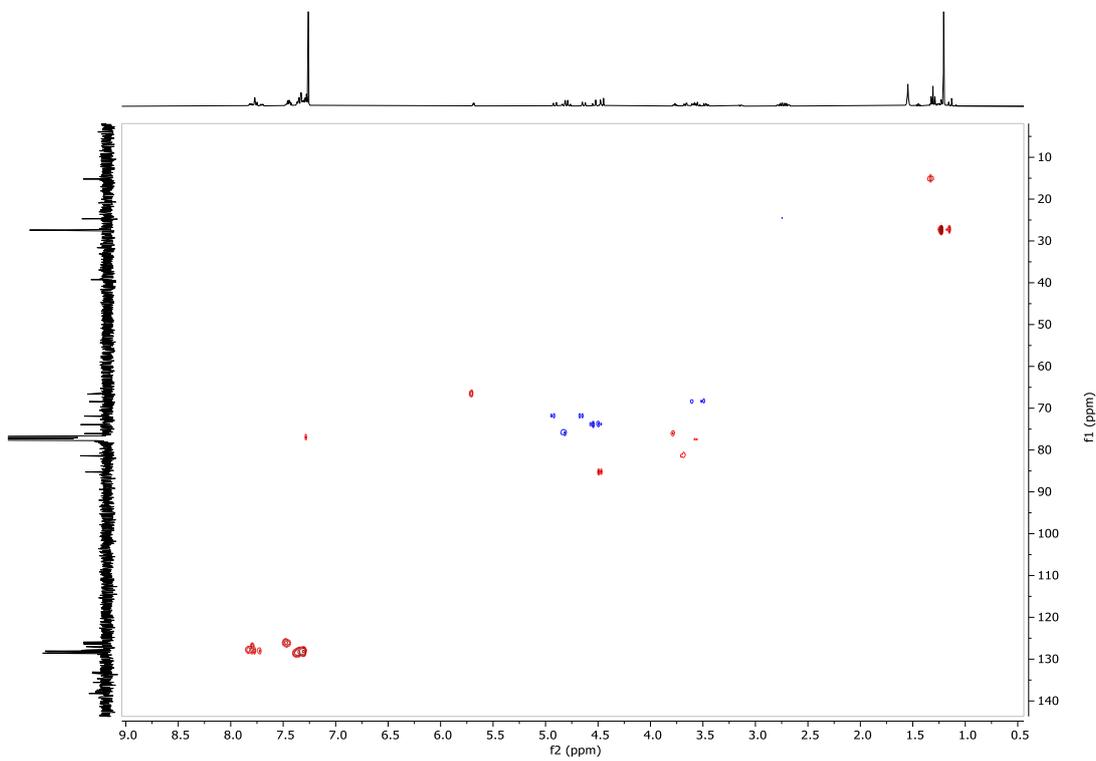
¹H NMR (400 MHz, CDCl₃) of **S9**:



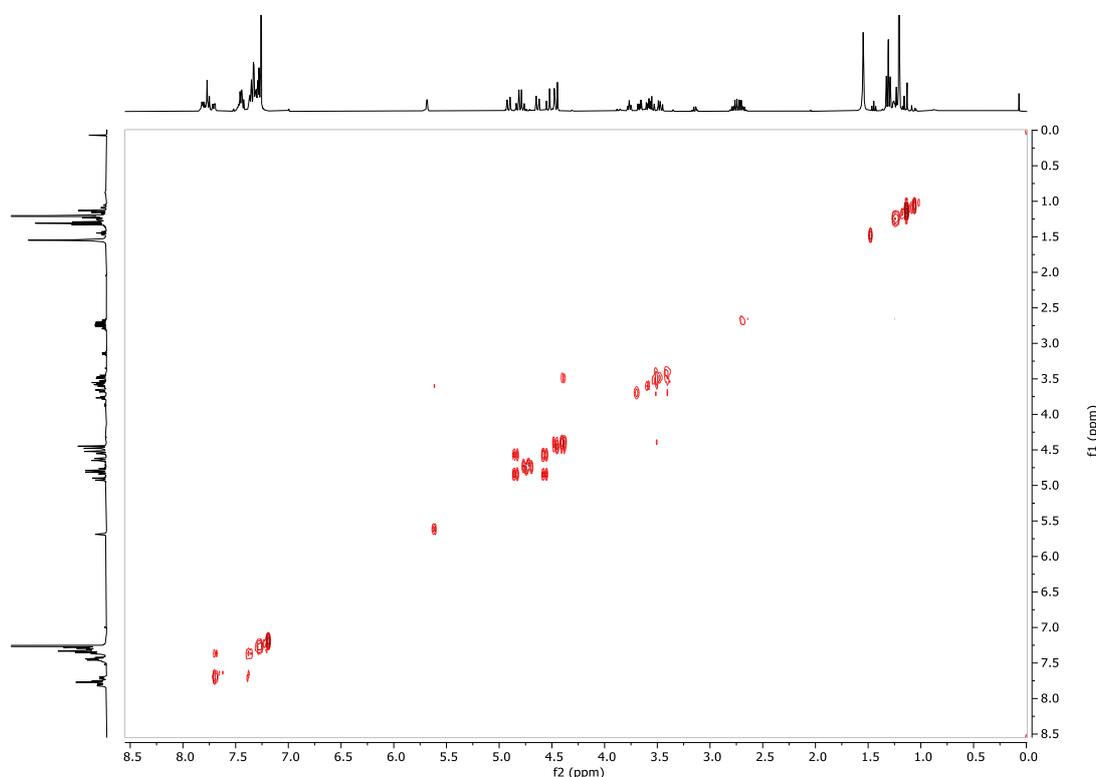
^{13}C NMR (101 MHz, CDCl_3) of **S9**:



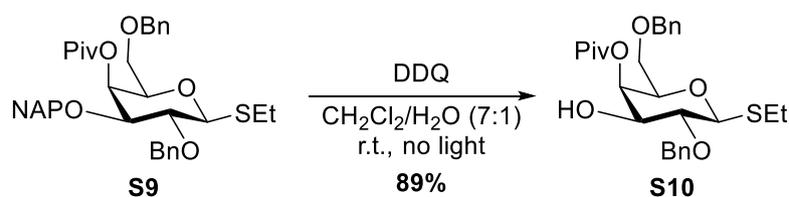
$^{13}\text{C}, ^1\text{H}$ HSQC of **S9**:



$^1\text{H},^1\text{H}$ COSY of **S9**:



Ethyl 2,6-bis-*O*-benzyl-4-(2,2-dimethylpropanoate)-1-thio- β -D-galactopyranoside (**S10**)



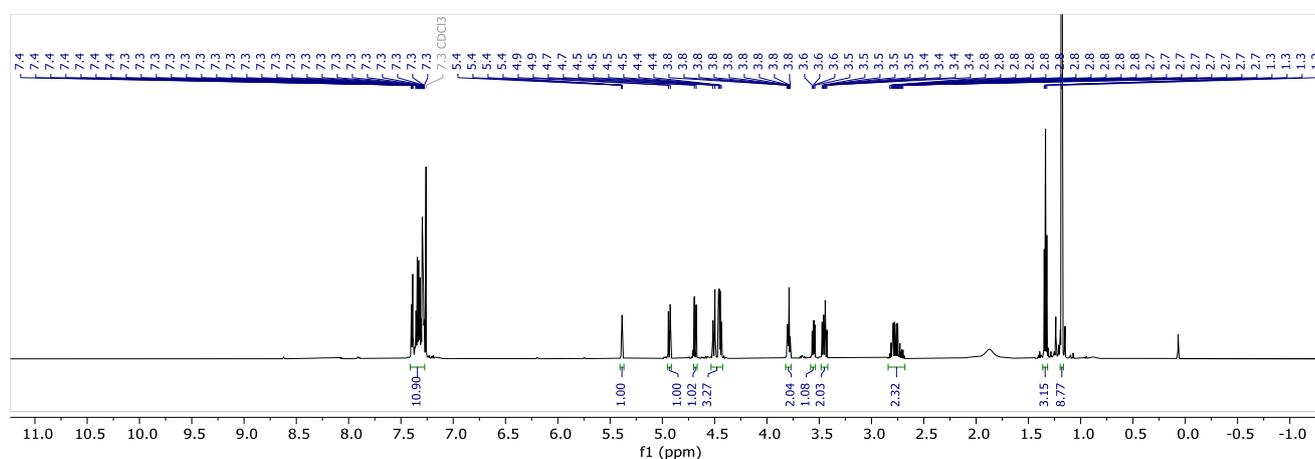
To a well stirred emulsion of **S9** (1.3 g, 2.1 mmol, 1.0 equiv.) in CH_2Cl_2 /water (7:1, 24 mL), was added DDQ (517 mg, 2.3 mmol, 1.1 equiv.) and the suspension was stirred at room temperature for 1.5 h protected from light. The mixture was diluted with CH_2Cl_2 , washed with 10% $\text{Na}_2\text{S}_2\text{O}_3$ and saturated aqueous NaHCO_3 solution. The organic layer was dried over Na_2SO_4 , filtered, concentrated and the residue was purified by column chromatography (SiO_2 , Hex/EtOAc = 9:1 to 3:1) to obtain the title compound **S10** (900 mg, 1.8 mmol, 89%) as a colorless solid.

$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.41 – 7.27 (m, 10H), 5.39 (dd, $J = 3.4, 1.2$ Hz, 1H), 4.93 (d, $J = 11.0$ Hz, 1H), 4.69 (d, $J = 11.0$ Hz, 1H), 4.53 – 4.43 (m, 3H), 3.82 – 3.76 (m, 2H), 3.55 (dd, $J = 9.6, 6.2$ Hz, 1H), 3.49 – 3.41 (m, 2H), 2.85 – 2.67 (m, 2H), 1.34 (t, $J = 7.5$ Hz, 3H), 1.18 (s, 9H) ppm.

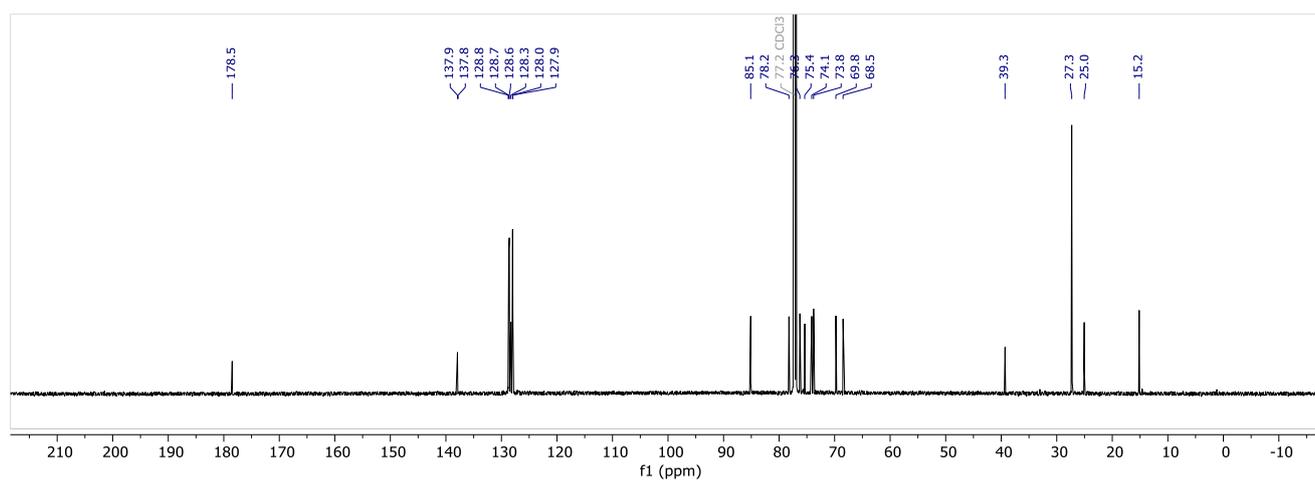
$^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 178.5, 137.9, 137.8, 128.8, 128.7, 128.6, 128.3, 128.0, 127.9, 85.1, 78.2, 76.3, 75.4, 74.1, 73.8, 69.8, 68.5, 39.3, 27.3, 25.0, 15.2 ppm.

HRMS (QToF): Calcd for $\text{C}_{27}\text{H}_{36}\text{O}_6\text{Sn}$ [$\text{M} + \text{Na}$] $^+$ 511.2125; found 511.2217.

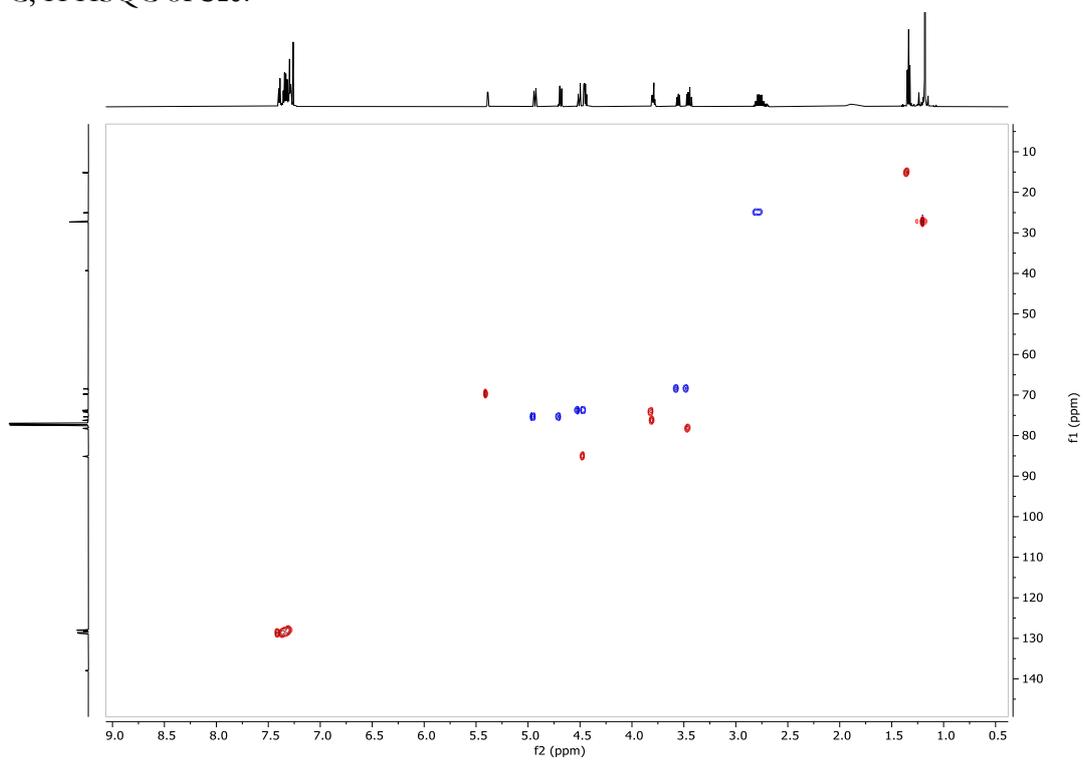
$^1\text{H NMR}$ (400 MHz, CDCl_3) of **S10**:



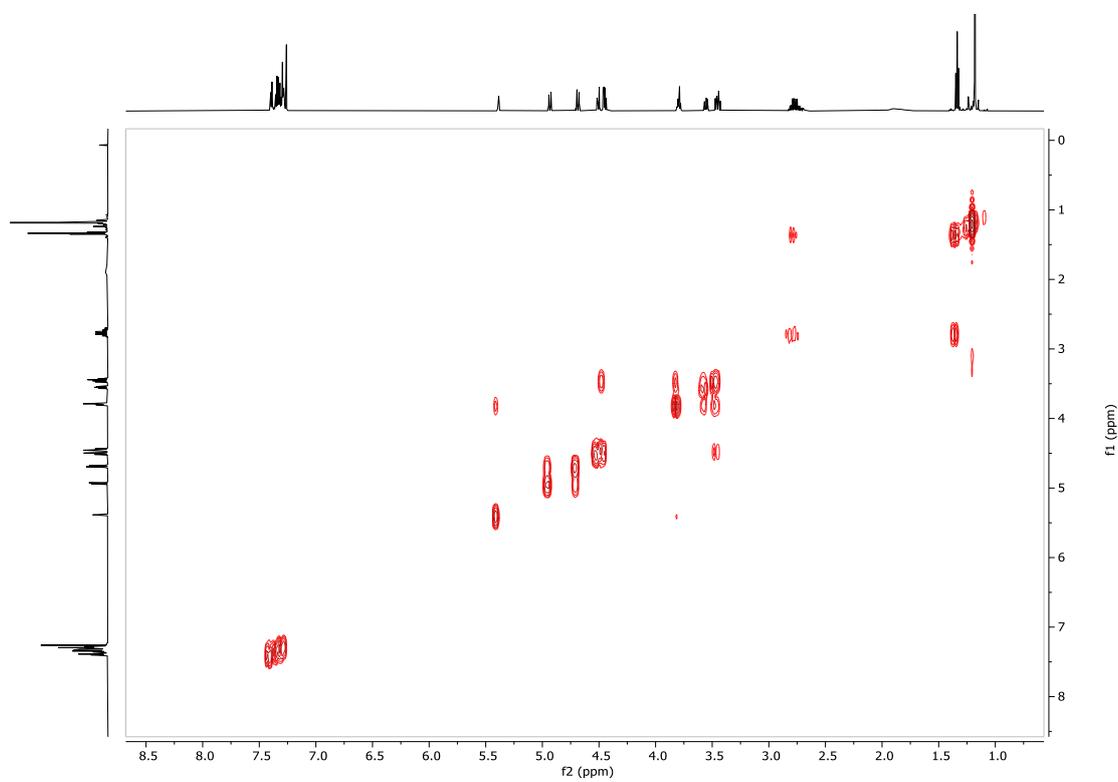
$^{13}\text{C NMR}$ (101 MHz, CDCl_3) of **S10**:



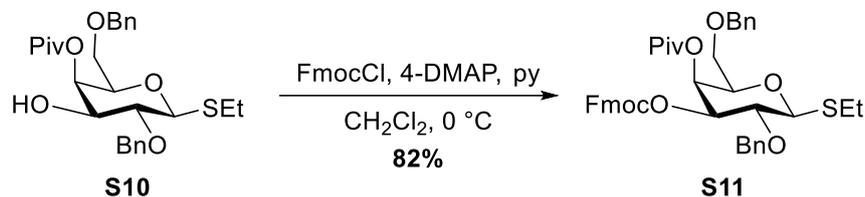
$^{13}\text{C}, ^1\text{H}$ HSQC of **S10**:



$^1\text{H}, ^1\text{H}$ COSY of **S10**:



Ethyl 2,6-bis-*O*-benzyl-4-(2,2-dimethylpropanoate)-4-*O*-fluorenylmethoxycarbonyl-1-thio- β -D-galactopyranoside (S11**)**



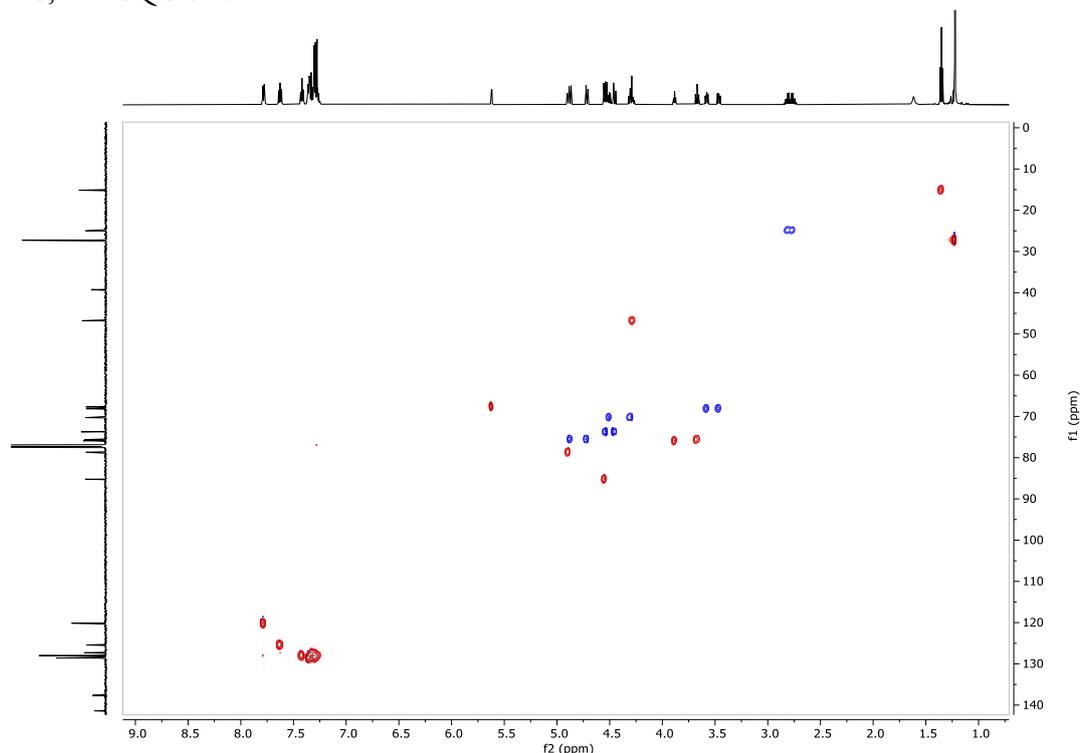
S10 was dissolved in anhydrous CH_2Cl_2 (20 mL) and anhydrous pyridine (0.41 mL, 5.12 mmol, 5.0 equiv.) was added followed by FmocCl (0.4 g, 1.54 mmol, 1.5 equiv.) at 0 °C. 4-DMAP (6 mg, 0.05 mmol, 0.05 equiv.) was added and the reaction mixture was stirred for one hour at 0 °C. Aqueous citric acid solution (10 mL) was added and the mixture was allowed to warm up to room temperature. The aqueous phase was extracted with CH_2Cl_2 and the combined organic phase was dried over Na_2SO_4 , filtered and concentrated. The title compound **S11** (600 mg, 0.84 mmol, 82%) was obtained as a white solid after purification by column chromatography (SiO_2 , Hex/EtOAc = 9:1 to 3:1).

^1H NMR (600 MHz, CDCl_3) δ 7.78 (ddt, $J = 7.5, 2.0, 0.9$ Hz, 2H), 7.63 (ddd, $J = 9.2, 7.4, 0.9$ Hz, 2H), 7.42 (tt, $J = 7.5, 1.5$ Hz, 2H), 7.38 – 7.26 (m, 12H), 5.62 (dd, $J = 3.4, 1.1$ Hz, 1H), 4.92 – 4.85 (m, 2H), 4.72 (d, $J = 10.8$ Hz, 1H), 4.57 – 4.43 (m, 4H), 4.34 – 4.25 (m, 2H), 3.88 (td, $J = 6.4, 1.2$ Hz, 1H), 3.67 (t, $J = 9.6$ Hz, 1H), 3.58 (dd, $J = 9.6, 6.2$ Hz, 1H), 3.46 (dd, $J = 9.6, 6.5$ Hz, 1H), 2.86 – 2.70 (m, 2H), 1.35 (t, $J = 7.4$ Hz, 3H), 1.22 (s, 9H) ppm.

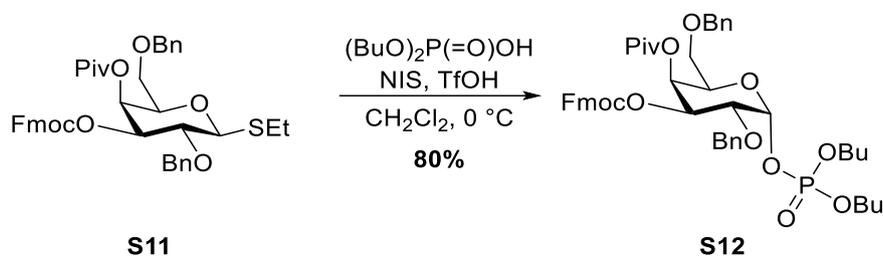
^{13}C NMR (151 MHz, CDCl_3) δ 177.5, 154.3, 143.8, 143.4, 141.4, 137.7, 137.6, 128.6, 128.5, 128.5, 128.0, 128.0, 127.9, 127.3, 127.3, 125.4, 125.3, 120.1, 85.2, 78.7, 75.9, 75.6, 75.5, 73.7, 70.3, 68.1, 67.7, 46.8, 39.3, 27.3, 25.0, 15.1 ppm.

HRMS (QToF): Calcd for $\text{C}_{42}\text{H}_{46}\text{O}_8\text{SNa}$ $[\text{M} + \text{Na}]^+$ 733.2806; found 733.2787.

$^{13}\text{C}, ^1\text{H}$ HSQC of **S11**:



Dibutoxyphosphoryloxy 2,6-bis-*O*-benzyl-4-(2,2-dimethylpropanoate)-4-*O*-fluorenylmethoxycarbonyl- α -D-galactopyranoside (**S12**)



To a solution of **S11** (300 mg, 0.4 mmol, 1.0 equiv.) and dibutyl phosphate (168 μL , 0.8 mmol, 2.0 equiv.) in anhydrous CH_2Cl_2 (4 mL) was added NIS (171 mg, 0.8 mmol, 1.5 equiv.) and TfOH (11 μL , 0.1 mmol, 0.3 equiv.) at 0 $^\circ\text{C}$. The reaction was stirred for 1 h. To the reaction mixture was added 10% sodium thiosulfate solution. The bilayer mixture was extracted with CH_2Cl_2 , dried with Na_2SO_4 , filtered, concentrated and purified by flash column chromatography (SiO_2 , Hex/EtOAc = 3:1 to 2:3) to give title compound **S12** (290 mg, 0.3 mmol, **80%**) as a colorless foam.

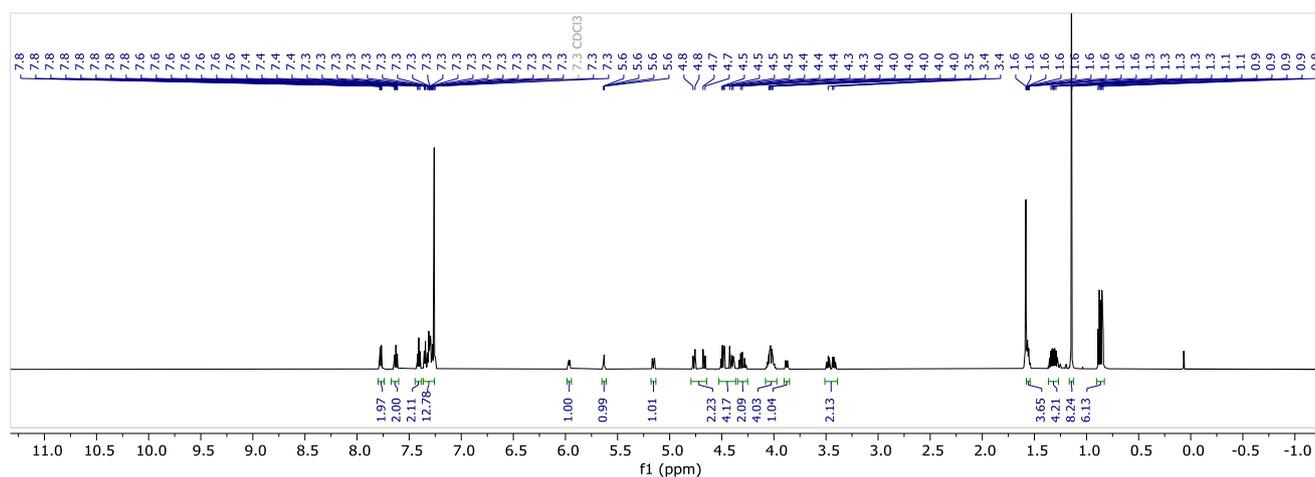
$R_f = 0.16$ (Hex/EtOAc 3:1).

¹H NMR (600 MHz, CDCl₃) δ 7.77 (ddt, *J* = 7.6, 1.8, 0.9 Hz, 2H), 7.63 (tt, *J* = 7.5, 1.0 Hz, 2H), 7.41 (t, *J* = 7.1 Hz, 2H), 7.37 – 7.24 (m, 12H), 5.96 (dd, *J* = 7.3, 3.4 Hz, 1H), 5.63 (dd, *J* = 3.4, 1.5 Hz, 1H), 5.16 (dd, *J* = 10.4, 3.3 Hz, 1H), 4.81 – 4.62 (m, 2H), 4.56 – 4.36 (m, 4H), 4.35 – 4.24 (m, 2H), 4.09 – 3.97 (m, 4H), 3.88 (dt, *J* = 10.4, 2.9 Hz, 1H), 3.54 – 3.39 (m, 2H), 1.61 – 1.53 (m, 4H), 1.40 – 1.23 (m, 4H), 1.14 (s, 9H), 0.86 (dt, *J* = 15.0, 7.4 Hz, 6H) ppm.

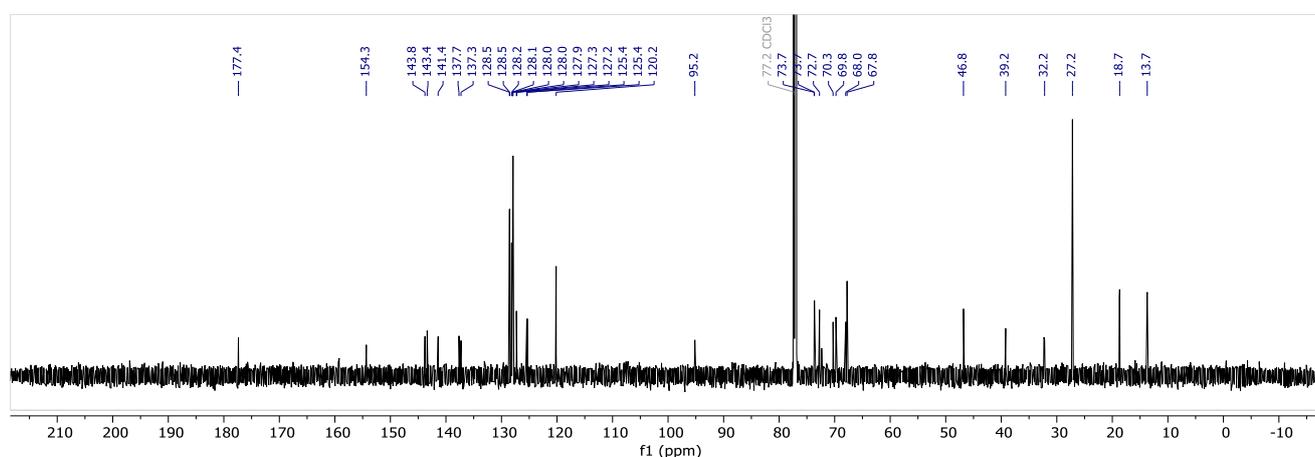
¹³C NMR (151 MHz, CDCl₃) δ 177.4, 154.3, 143.8, 143.4, 141.4, 137.7, 137.3, 128.5, 128.5, 128.2, 128.1, 128.0, 128.0, 127.9, 127.3, 127.2, 125.4, 125.4, 120.2, 95.2, 73.7, 73.7, 72.7, 70.3, 69.8, 68.0, 67.8, 46.8, 39.2, 32.2, 27.2, 18.7, 13.7 ppm.

HRMS (QToF): Calcd for C₄₈H₅₉O₁₂PNa [M + Na]⁺ 881.3636; found 881.3759.

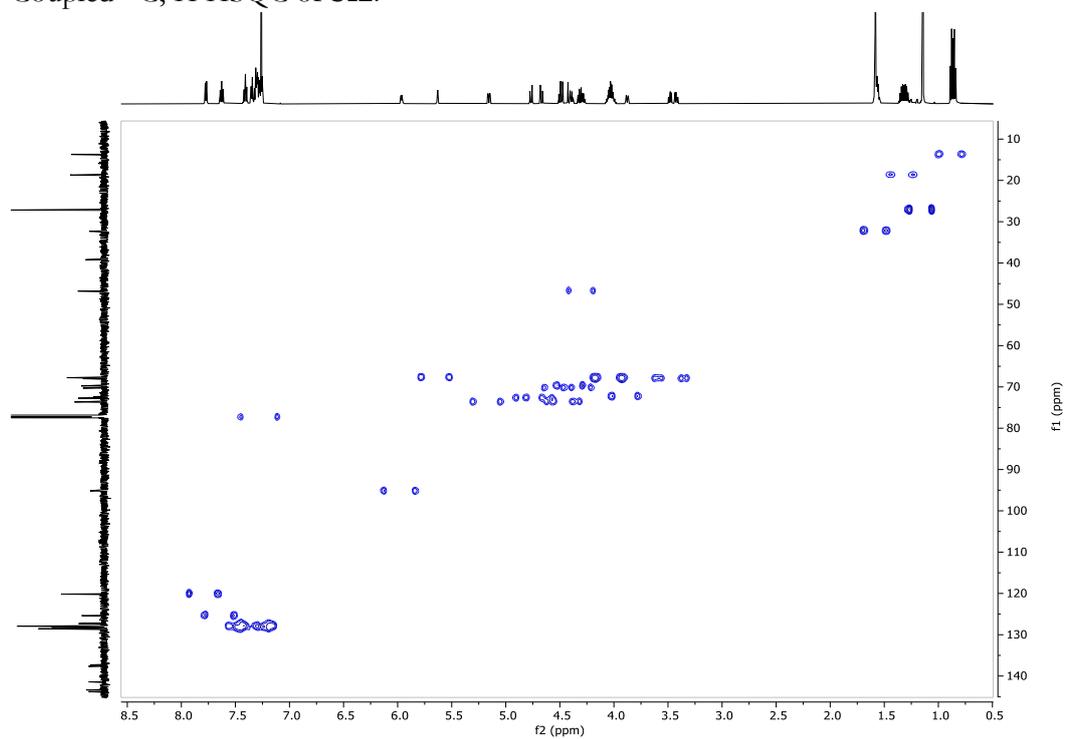
¹H NMR (400 MHz, CDCl₃) of **S12**:



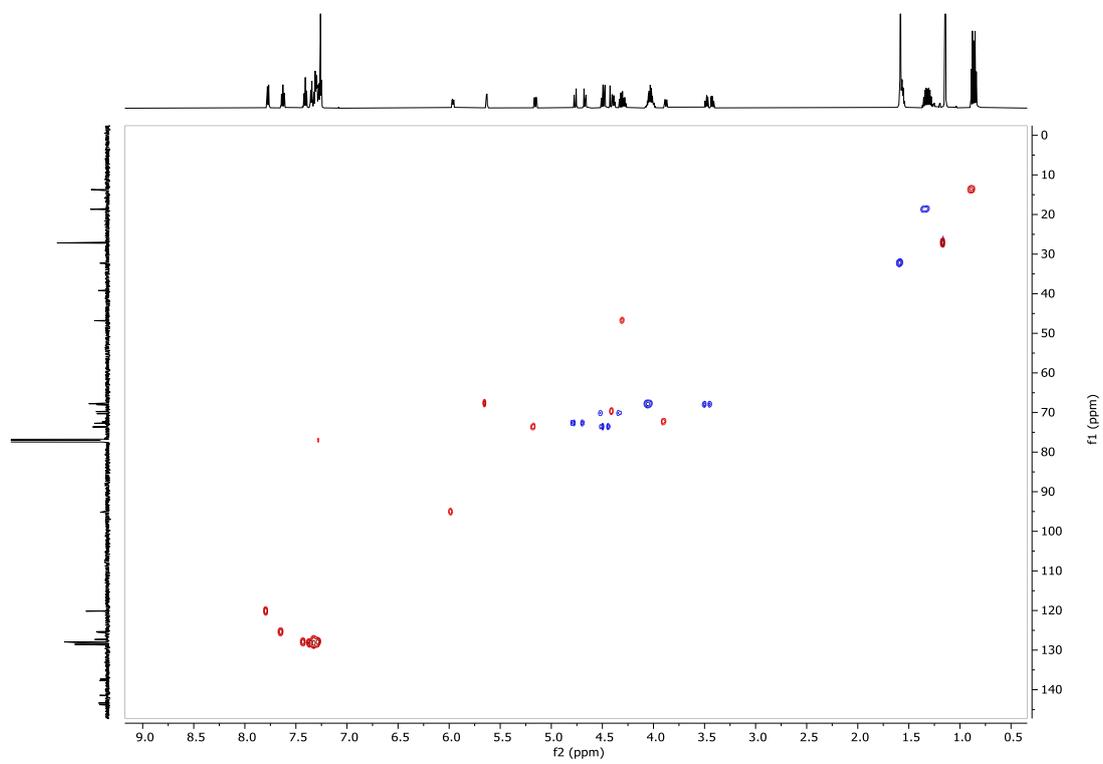
¹³C NMR (101 MHz, CDCl₃) of **S12**:



Coupled $^{13}\text{C}, ^1\text{H}$ HSQC of **S12**:

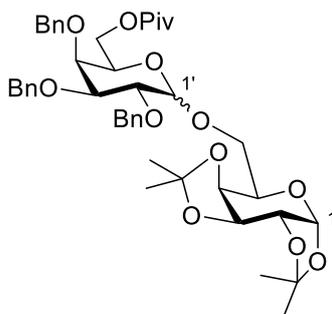


$^{13}\text{C}, ^1\text{H}$ HSQC of **S12**:



6 Determination of alpha/beta-ratios

1,2:3,4-Bis-*O*-(1-methylethylidene)-6-*O*-[2,3,4-tris-*O*-benzyl-6-(2,2-dimethylpropanoate)- α/β -D-galactopyranosyl]- α -D-galactopyranoside (**6Piv-A1**)



The title compound was prepared according to general procedure for glycosylations. Product **6Piv-A1** (20 mg, 26 μ mol, 75%, α/β 53:47) was obtained as a colorless oil after purification using **Method-2b** ($t_R(\alpha)$ = 9.8 min, $t_R(\beta)$ = 9.1 min).

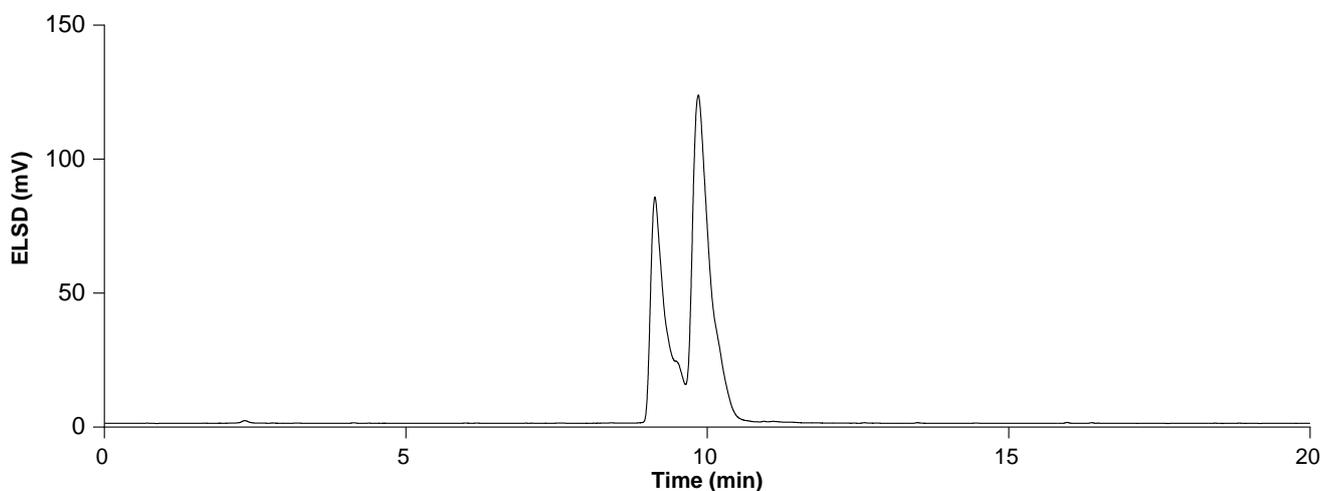
Data of the anomeric mixture:

$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.47 – 7.27 (m, 30H), 5.56 (d, J = 5.0 Hz, 1H, $\text{H}_1(\beta)$), 5.50 (d, J = 5.0 Hz, 1H, $\text{H}_1(\alpha)$), 5.05 (d, J = 11.1 Hz, 1H), 5.00 (d, J = 3.7 Hz, 1H, $\text{H}_{1'}(\alpha)$), 4.97 (dd, J = 11.3, 5.1 Hz, 2H), 4.87 (dd, J = 18.6, 11.9 Hz, 2H), 4.79 – 4.70 (m, 5H), 4.62 (d, J = 11.5 Hz, 1H), 4.60 – 4.56 (m, 3H), 4.41 (d, J = 7.7 Hz, 1H, $\text{H}_{1'}(\beta)$), 4.32 – 4.28 (m, 3H), 4.28 – 4.24 (m, 1H), 4.23 – 4.16 (m, 2H), 4.13 – 4.00 (m, 8H), 3.97 (dd, J = 10.0, 2.8 Hz, 1H), 3.88 – 3.82 (m, 2H), 3.76 (d, J = 6.6 Hz, 1H), 3.73 – 3.66 (m, 3H), 3.55 – 3.49 (m, 2H), 1.51 (s, 3H), 1.48 (s, 3H), 1.43 (s, 4H), 1.32 (s, 3H), 1.31 (s, 6H), 1.30 (s, 3H), 1.17 (s, 9H), 1.15 (s, 9H) ppm.

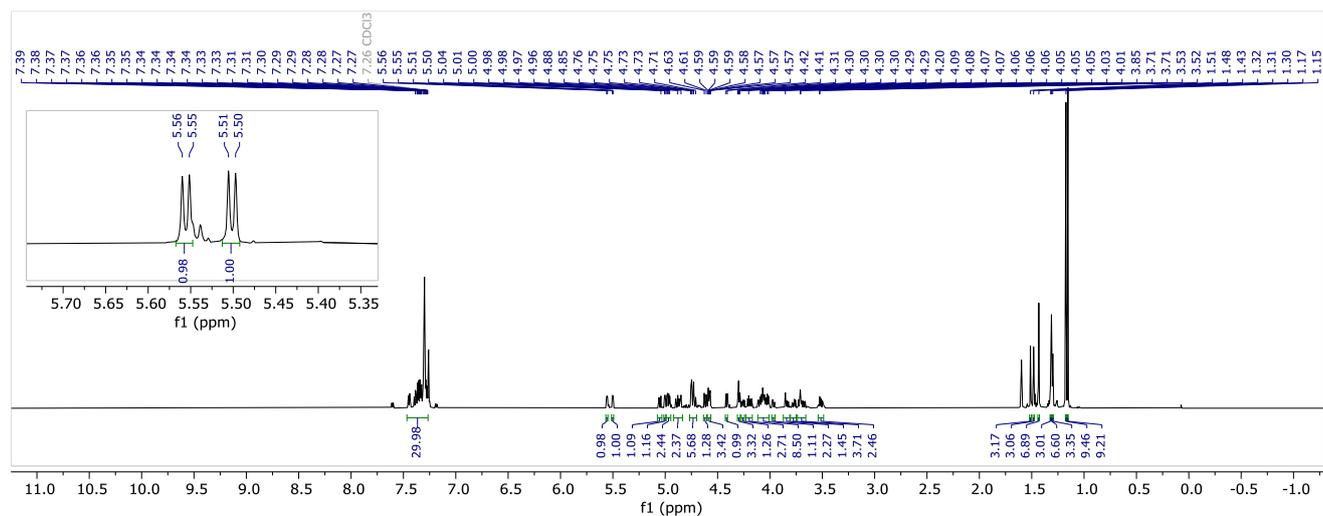
$^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 178.2, 139.1, 139.0, 138.8, 138.7, 138.6, 138.5, 137.3, 128.7, 128.6, 128.5, 128.5, 128.4, 128.4, 128.3, 128.3, 127.9, 127.8, 127.8, 127.8, 127.7, 127.7, 127.7, 127.6, 127.5, 109.5, 109.4, 108.7, 108.6, 104.8 ($\text{C}_{1'}(\beta)$), 97.5 ($\text{C}_{1'}(\alpha)$), 96.5 ($\text{C}_1(\beta)$), 96.4 ($\text{C}_1(\alpha)$), 82.0, 79.2, 79.0, 76.5, 75.5, 74.9, 74.9, 74.7, 74.0, 73.7, 73.5, 72.9, 72.1, 71.6, 71.1, 70.9, 70.8, 70.8, 70.7, 69.9, 69.7, 68.7, 67.8, 67.6, 66.6, 66.2, 63.3, 62.7, 38.8, 27.4, 27.3, 26.3, 26.2, 26.1, 25.2, 25.1, 24.6, 24.6 ppm.

HRMS (QToF): Calcd for $\text{C}_{44}\text{H}_{56}\text{O}_{12}\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 799.3664; found 799.3670.

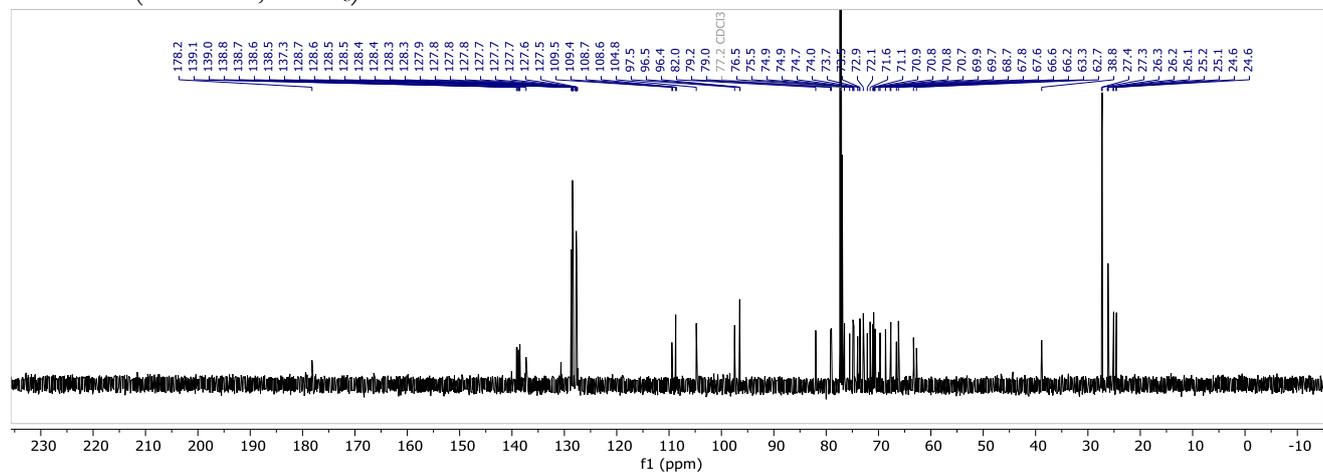
NP-HPLC of **P1-SL101** (ELSD trace, $t_R(\alpha) = 9.8$ min, $t_R(\beta) = 9.1$ min):



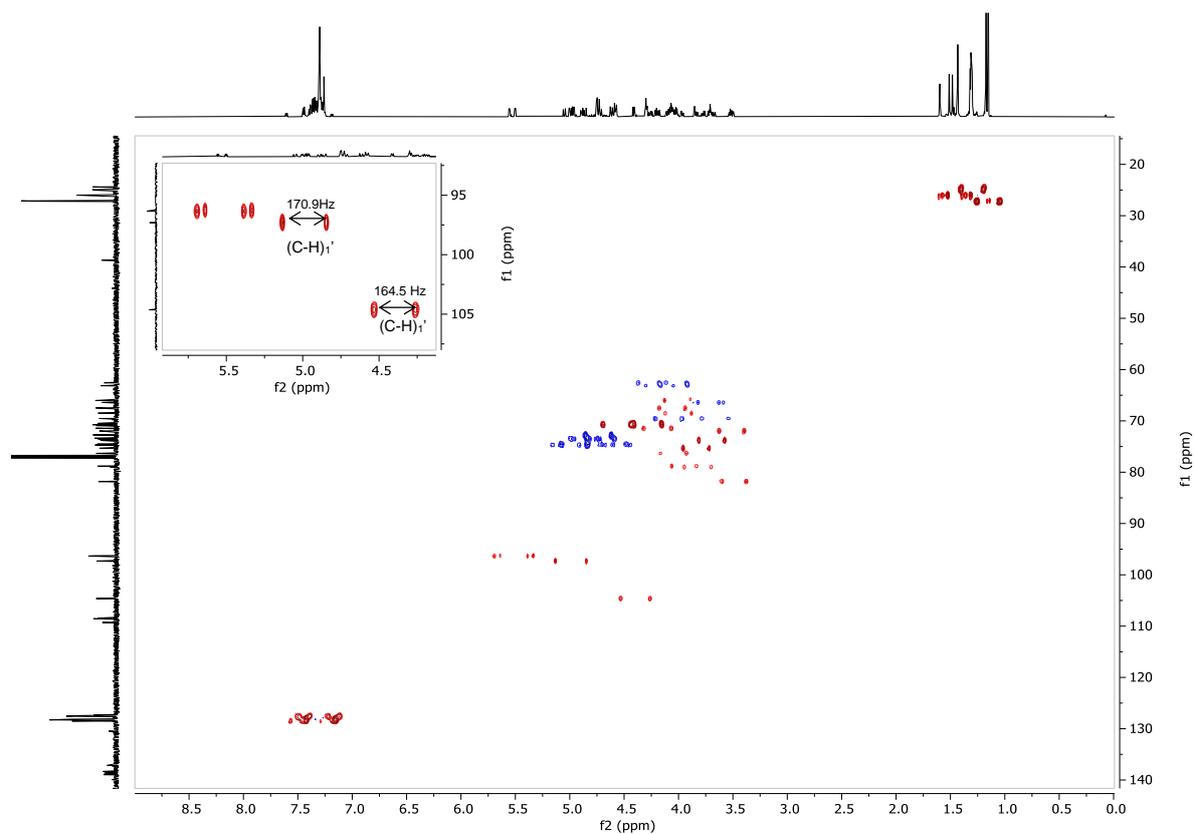
^1H NMR (600 MHz, CDCl_3) of **P1-SL101**:



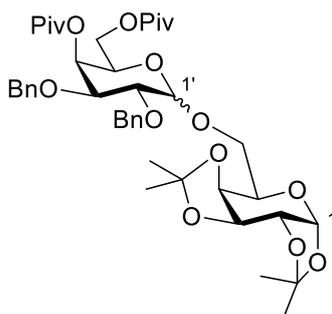
^{13}C NMR (151 MHz, CDCl_3) of **P1-SL101**:



Coupled $^{13}\text{C}, ^1\text{H}$ HSQC of **P1-SL101**:



1,2:3,4-Bis-*O*-(1-methylethylidene)-6-*O*-[2,3-bis-*O*-benzyl-4,6-bis-(2,2-dimethylpropanoate)- α/β -D-galactopyranosyl]- α -D-galactopyranoside (4,6Piv-A1)



The title compound was prepared according to general procedure for glycosylations. Product **4,6Piv-A1** (19 mg, 24 μmol , 69%, α/β 80:20) was obtained as a colorless oil after purification using **Method-2b** ($t_{\text{R}}(\alpha) = 7.6$ min, $t_{\text{R}}(\beta) = 8.6$ min).

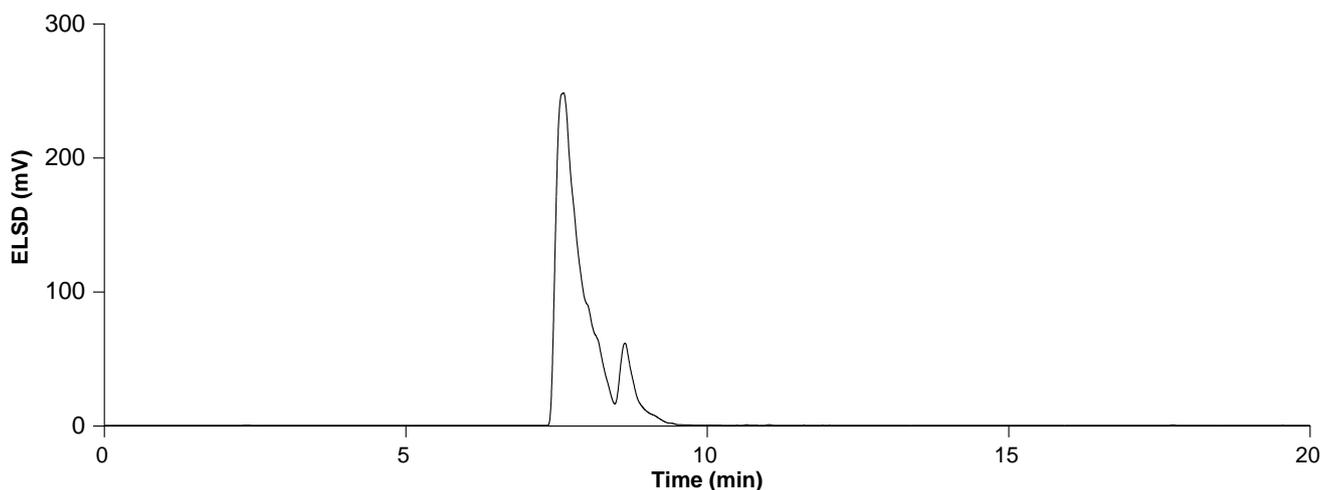
Data of the major isomer (α):

$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.35 – 7.26 (m, 10H), 5.56 (dd, $J = 3.5, 1.4$ Hz, 1H), 5.51 (d, $J = 5.0$ Hz, 1H, H_1), 4.97 (d, $J = 3.5$ Hz, 1H, H_1'), 4.80 – 4.67 (m, 3H), 4.59 (ddd, $J = 7.7, 5.3, 2.4$ Hz, 1H), 4.54 (d, $J = 10.8$ Hz, 1H), 4.34 – 4.26 (m, 3H), 4.08 (dd, $J = 11.0, 7.1$ Hz, 1H), 4.05 – 3.97 (m, 3H), 3.82 – 3.72 (m, 3H), 1.52 (s, 3H), 1.44 (s, 3H), 1.33 – 1.32 (m, 6H), 1.21 (s, 9H), 1.15 (s, 9H) ppm.

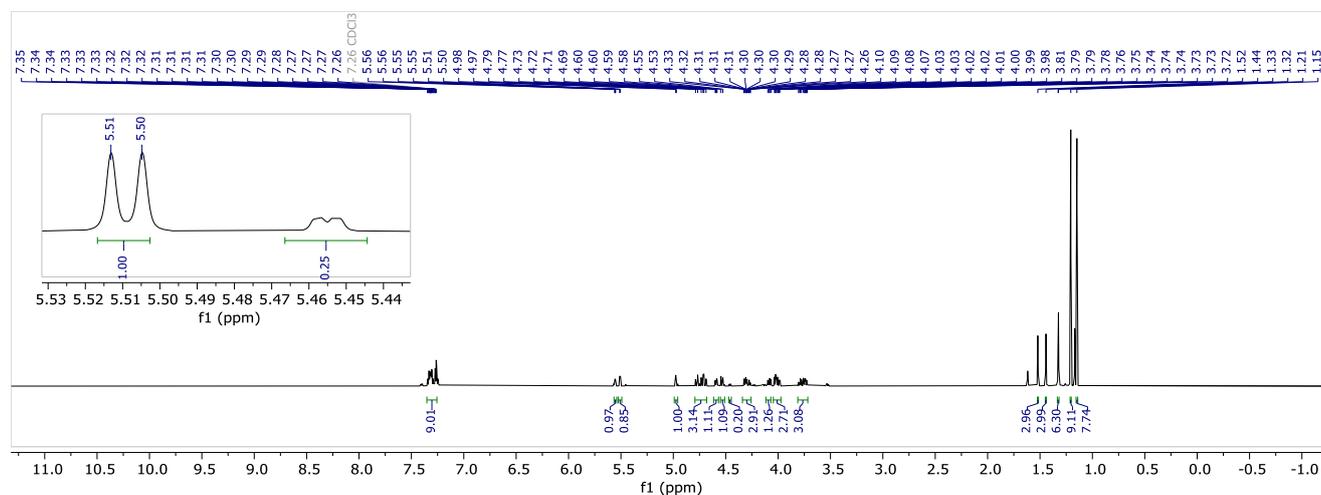
$^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 178.1, 176.9, 138.5, 138.4, 128.4, 128.3, 128.0, 128.0, 127.8, 127.5, 109.4, 108.7, 98.1 (C_1'), 96.4 (C_1), 76.1, 74.9, 73.2, 72.0, 71.1, 70.8, 70.7, 67.3, 67.1, 67.0, 66.5, 62.2, 39.2, 38.9, 27.3, 27.3, 26.3, 26.2, 25.1, 24.7 ppm.

HRMS (QToF): Calcd for $\text{C}_{42}\text{H}_{58}\text{O}_{13}\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 793.3796; found 793.3786.

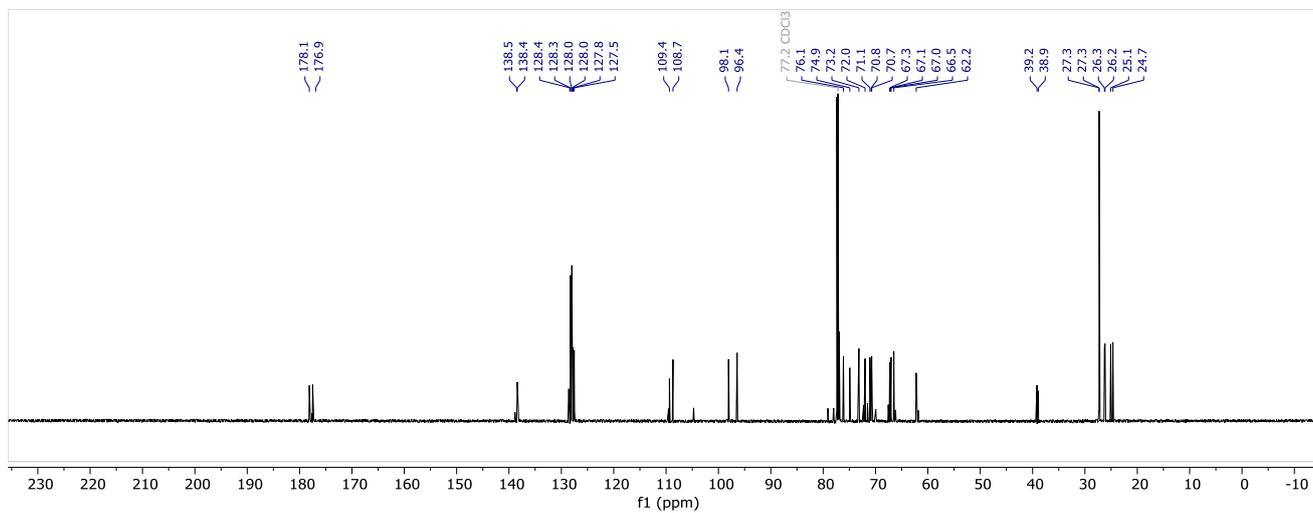
NP-HPLC of **P1-SL103** (ELSD trace, $t_R(\alpha) = 7.6$ min, $t_R(\beta) = 8.6$ min):



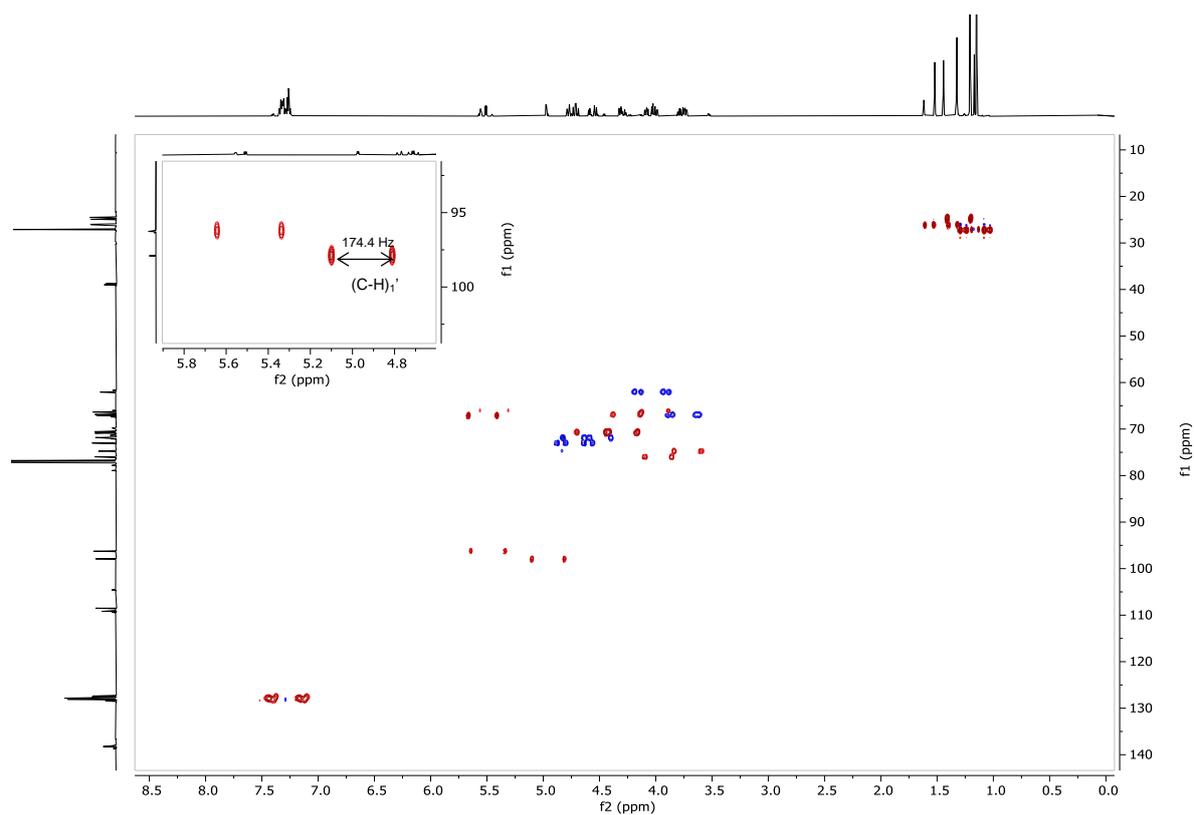
$^1\text{H NMR}$ (600 MHz, CDCl_3) of **4,6Piv-A1**:



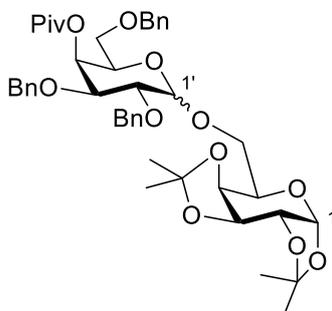
^{13}C NMR (151 MHz, CDCl_3) of **4,6Piv-A1**:



Coupled ^{13}C , ^1H HSQC of **4,6Piv-A1**:



1,2:3,4-Bis-*O*-(1-methylethylidene)-6-*O*-[2,3,6-tris-*O*-benzyl-4-(2,2-dimethylpropanoate)- α/β -D-galactopyranosyl]- α -D-galactopyranoside (4Piv-A1)



The title compound was prepared according to general procedure for glycosylations. Product **4Piv-A1** (22 mg, 28 μ mol, 82%, α/β 96:4) was obtained as a colorless oil after purification using **Method-2b** ($t_R(\alpha) = 7.9$ min).

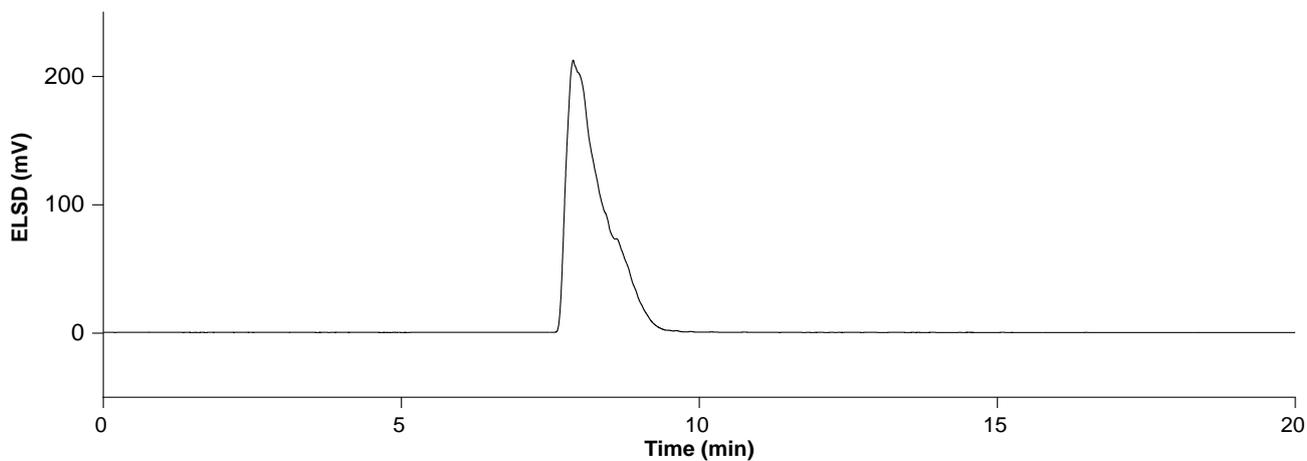
Data of the major isomer (α):

$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.35 – 7.24 (m, 15H), 5.63 (dd, $J = 3.4, 1.3$ Hz, 1H), 5.51 (d, $J = 5.1$ Hz, 1H, **H_I**), 4.98 (d, $J = 3.7$ Hz, 1H, **H_{I'}**), 4.79 – 4.67 (m, 3H), 4.58 (dd, $J = 7.9, 2.4$ Hz, 1H), 4.52 (t, $J = 10.7$ Hz, 2H), 4.48 – 4.43 (m, 1H), 4.34 (dd, $J = 7.9, 1.9$ Hz, 1H), 4.30 (dd, $J = 5.0, 2.4$ Hz, 1H), 4.23 (td, $J = 6.5, 1.3$ Hz, 1H), 4.04 (td, $J = 6.7, 1.9$ Hz, 1H), 3.98 (dd, $J = 10.0, 3.3$ Hz, 1H), 3.85 – 3.80 (m, 1H), 3.80 – 3.70 (m, 2H), 3.46 (ddd, $J = 42.0, 9.6, 6.5$ Hz, 2H), 1.53 (s, 3H), 1.43 (s, 3H), 1.33 (s, 3H), 1.32 (s, 3H), 1.12 (s, 9H) ppm.

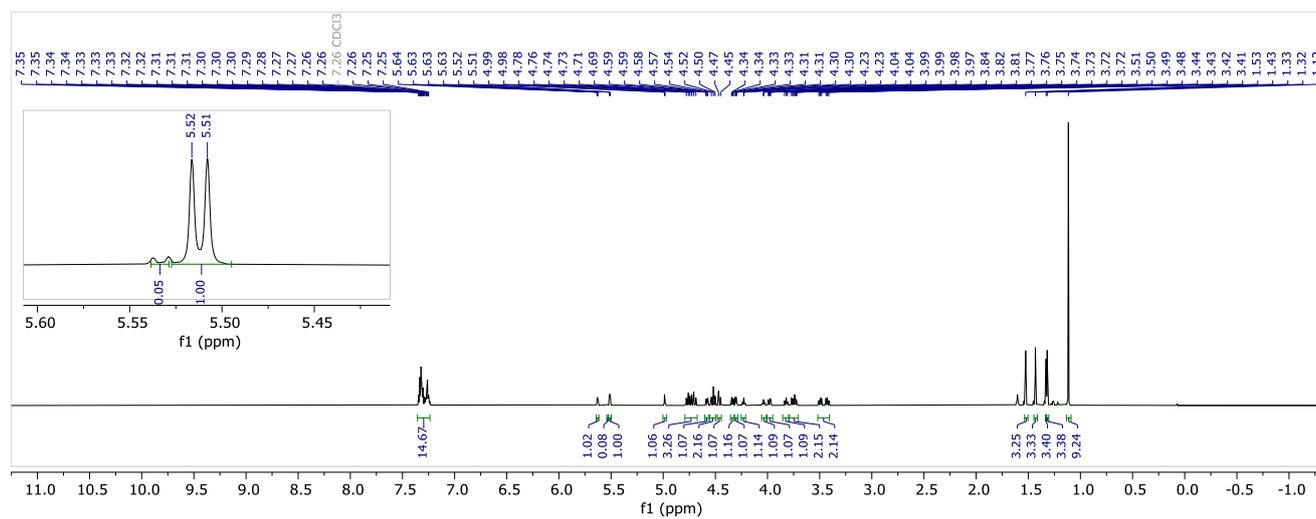
$^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 177.5, 138.7, 138.6, 138.0, 128.5, 128.4, 128.2, 128.1, 128.0, 128.0, 127.8, 127.7, 127.5, 109.3, 108.7, 98.1 (**C_{I'}**), 96.4 (**C_I**), 76.5, 75.1, 73.6, 73.1, 71.8, 71.0, 70.8, 70.8, 68.5, 67.9, 67.6, 66.8, 66.5, 66.2, 39.1, 29.8, 27.3, 26.3, 26.2, 25.1, 24.8 ppm.

HRMS (QToF): Calcd for $\text{C}_{44}\text{H}_{56}\text{O}_{12}\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 799.3664; found 799.3693.

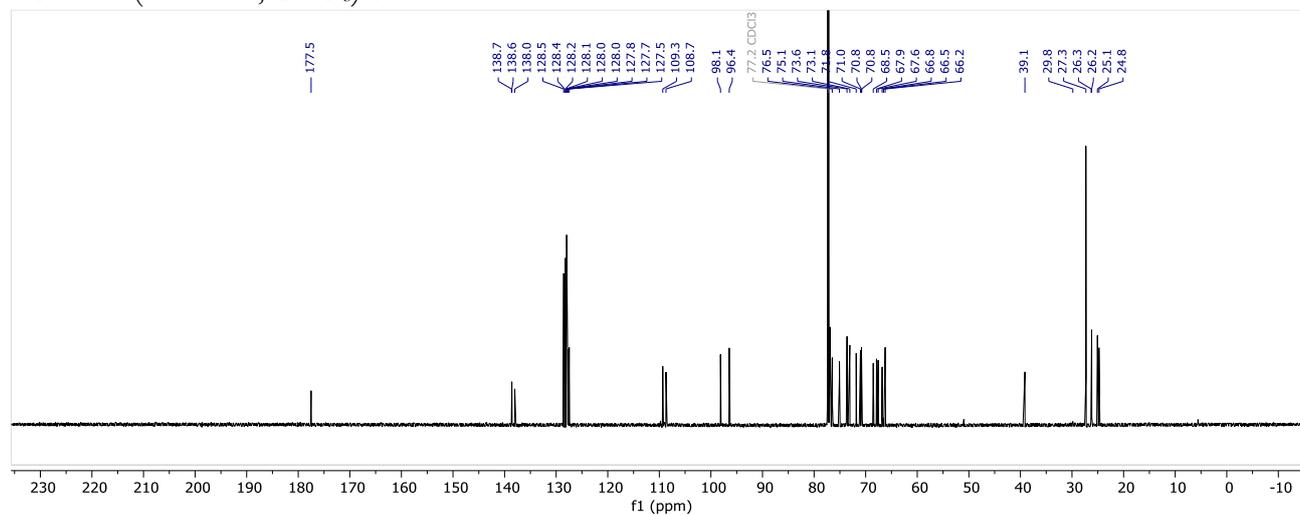
NP-HPLC of **4Piv-A1** (ELSD trace, $t_R(\alpha) = 7.9$ min):



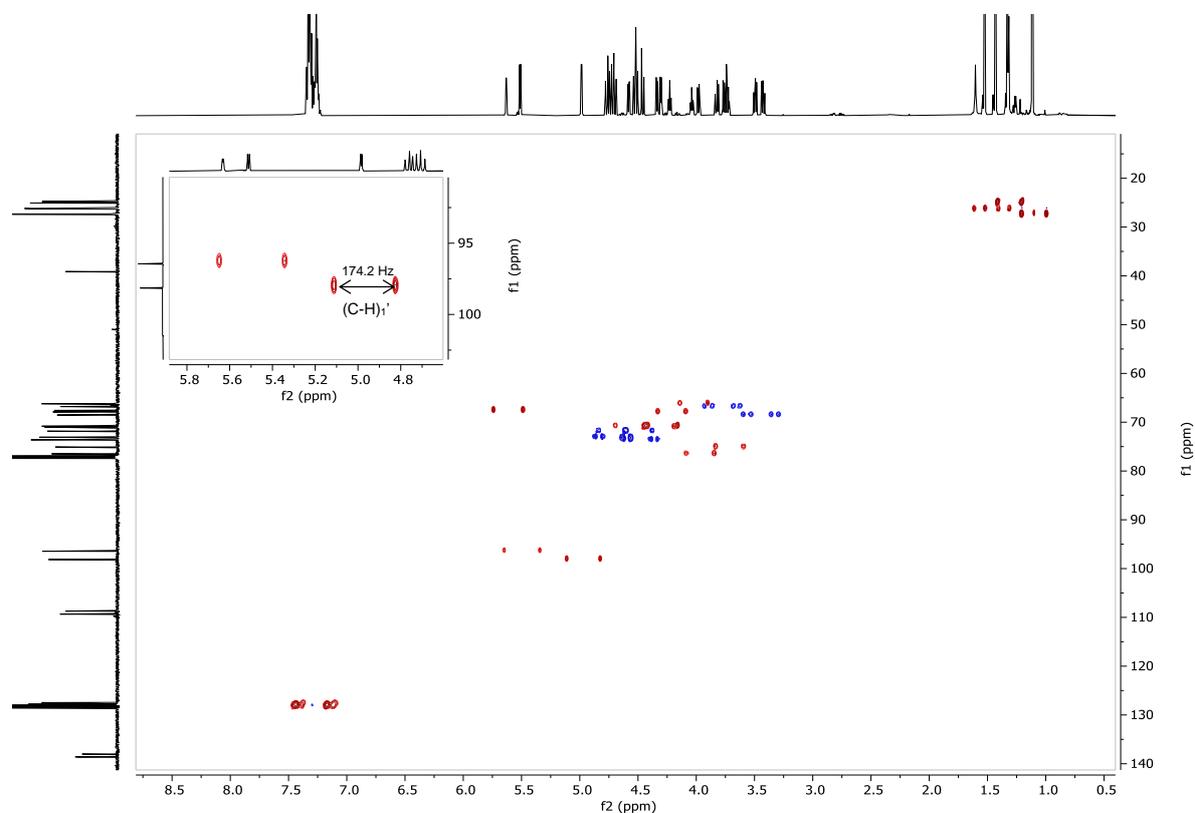
^1H NMR (600 MHz, CDCl_3) of **4Piv-A1**:



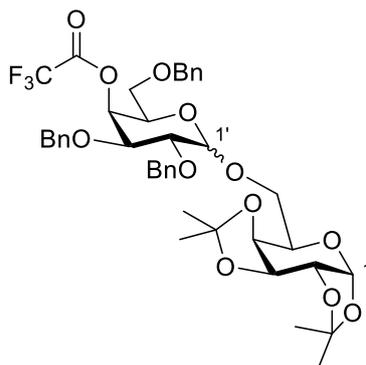
^{13}C NMR (151 MHz, CDCl_3) of **4Piv-A1**:



Coupled $^{13}\text{C}, ^1\text{H}$ HSQC of **4Piv-A1**:



1,2:3,4-Bis-*O*-(1-methylethylidene)-6-*O*-[2,3,6-tris-*O*-benzyl-4-(trifluoroacetate)- α/β -D-galactopyranosyl]- α -D-galactopyranoside (4TFA-A1**)**



The title compound was prepared according to general procedure for glycosylations. Product **4TFA-A1** (18 mg, 23 μmol , 67%, α/β 85:15) was obtained as a colorless oil after purification using **Method-2b** ($t_R(\alpha) = 6.0$ min, $t_R(\beta) = 9.1$ min).

Data of the major isomer (α):

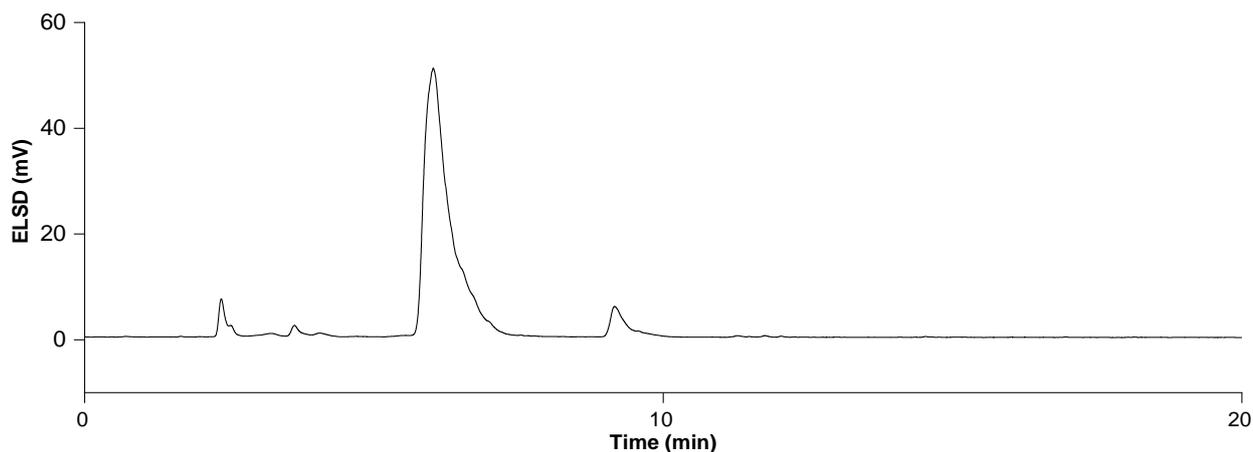
$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.37 – 7.26 (m, 15H), 5.74 (dd, $J = 3.4, 1.3$ Hz, 1H), 5.49 (d, $J = 5.0$ Hz, 1H, H_1), 4.94 (d, $J = 3.7$ Hz, 1H, H_1'), 4.75 (dt, $J = 11.6, 3.4$ Hz, 2H), 4.67 (d, $J = 12.0$ Hz, 1H), 4.61 – 4.54 (m, 2H), 4.53 – 4.43 (m, 2H), 4.33 – 4.25 (m, 3H), 4.06 – 3.98 (m, 2H), 3.80 – 3.68 (m, 3H), 3.54 (dd, $J = 9.3, 5.6$ Hz, 1H), 3.41 (dd, $J = 9.2, 8.3$ Hz, 1H), 1.51 (s, 3H), 1.43 (s, 3H), 1.33 (s, 3H), 1.30 (s, 3H) ppm.

$^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 138.4, 138.0, 137.6, 128.6, 128.4, 128.1, 128.0, 127.9, 127.8, 127.8, 109.5, 108.8, 98.0 (C_1'), 96.5 (C_1), 75.7, 75.2, 73.8, 73.4, 72.9, 72.6, 71.1, 70.8, 70.7, 67.4, 67.3, 66.6, 66.4, 29.9, 26.3, 26.2, 25.1, 24.7 ppm.

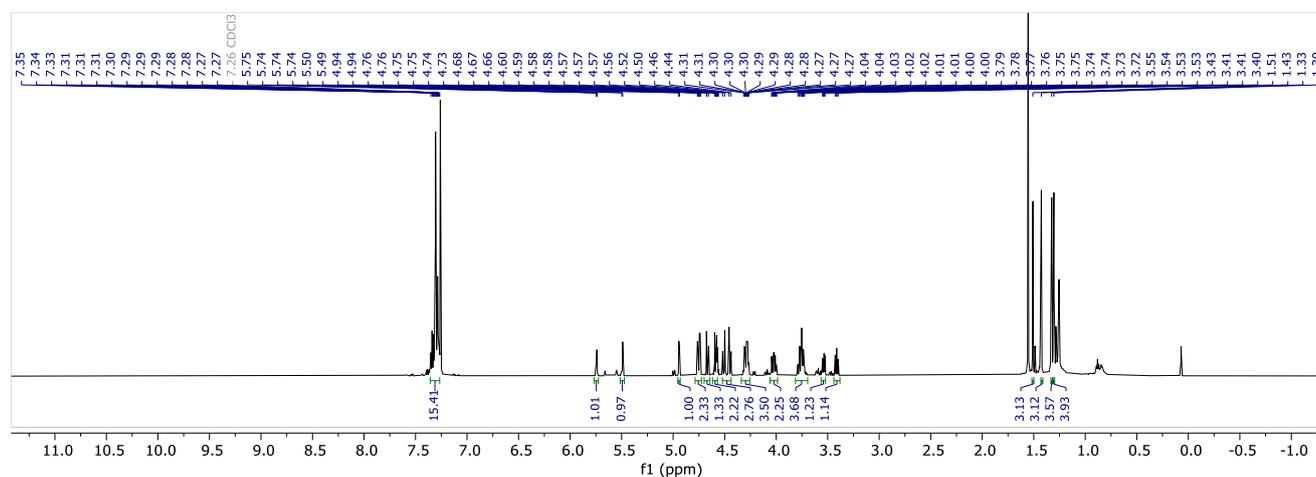
$^{19}\text{F NMR}$ (564 MHz, CDCl_3) δ -74.85 (s) ppm.

HRMS (QToF): Calcd for $\text{C}_{41}\text{H}_{47}\text{F}_3\text{O}_{12}\text{Na}$ $[\text{M} + \text{Na}]^+$ 811.2912; found 811.2944.

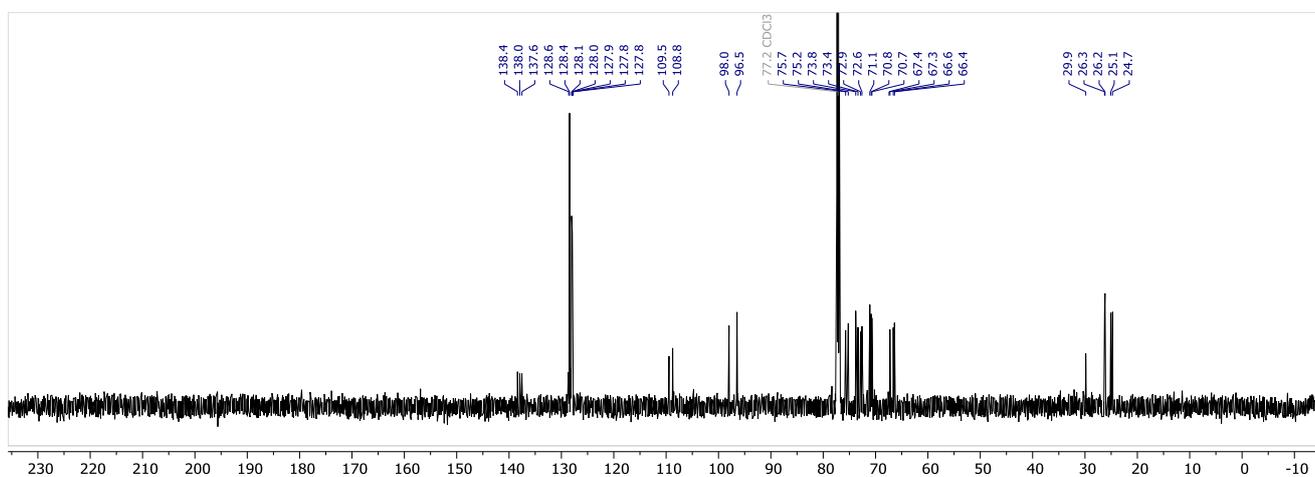
NP-HPLC of **4TFA-A1** (ELSD trace, $t_R(\alpha) = 6.0$ min, $t_R(\beta) = 9.1$ min):



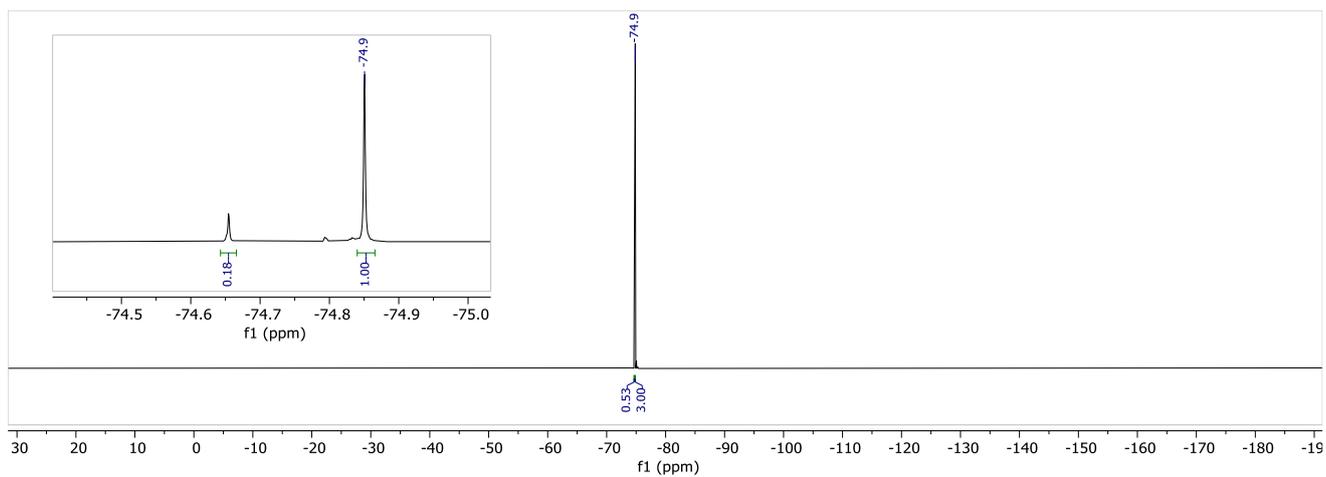
$^1\text{H NMR}$ (600 MHz, CDCl_3) of **4TFA-A1**:



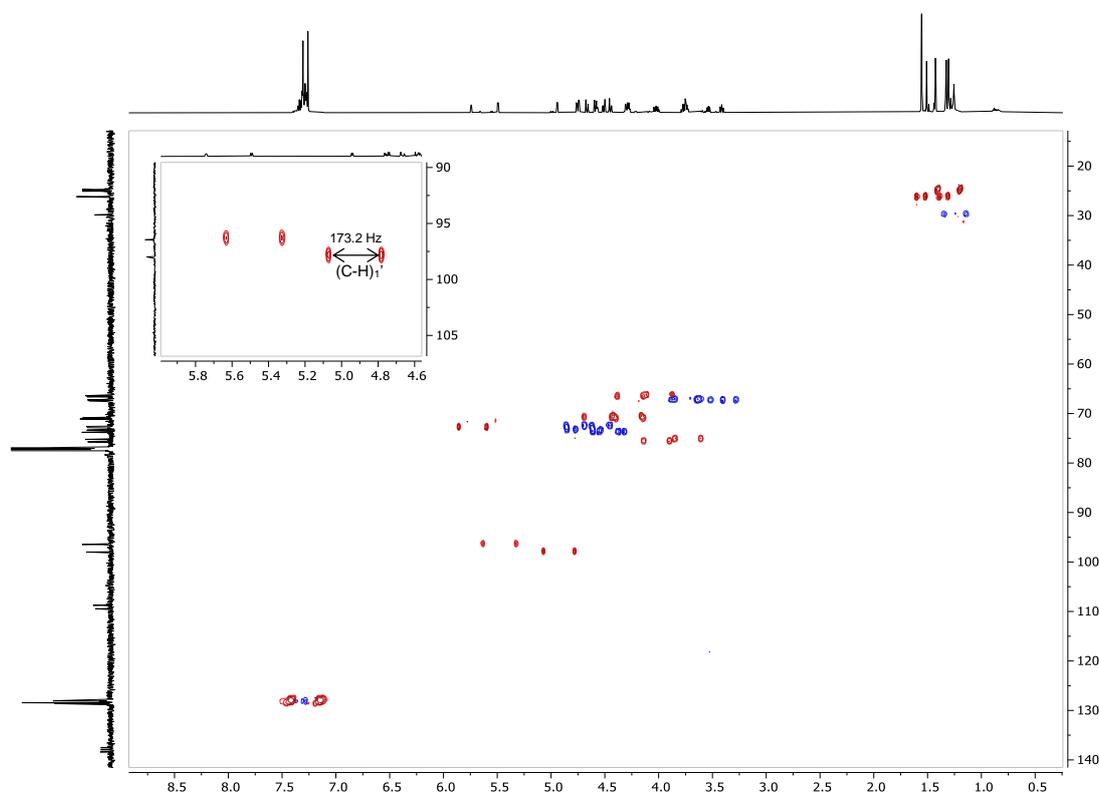
^{13}C NMR (151 MHz, CDCl_3) of **4TFA-A1**:



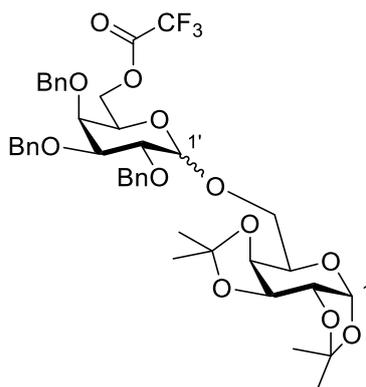
^{19}F NMR (376 MHz, CDCl_3) of **4TFA-A1**:



Coupled $^{13}\text{C}, ^1\text{H}$ HSQC of **4TFA-A1**:



1,2:3,4-Bis-*O*-(1-methylethylidene)-6-*O*-[2,3,4-tris-*O*-benzyl-6-(trifluoroacetate)- α/β -D-galactopyranosyl]- α -D-galactopyranoside (**6TFA-A1**)



The title compound was prepared according to general procedure for glycosylations. Product **6TFA-A1** (15 mg, 19 μmol , 56%, α/β 53:47) was obtained as a colorless oil after purification using **Method-2b** ($t_{\text{R}}(\alpha) = 8.3$ min, $t_{\text{R}}(\beta) = 9.0$ min).

Data of the anomeric mixture:

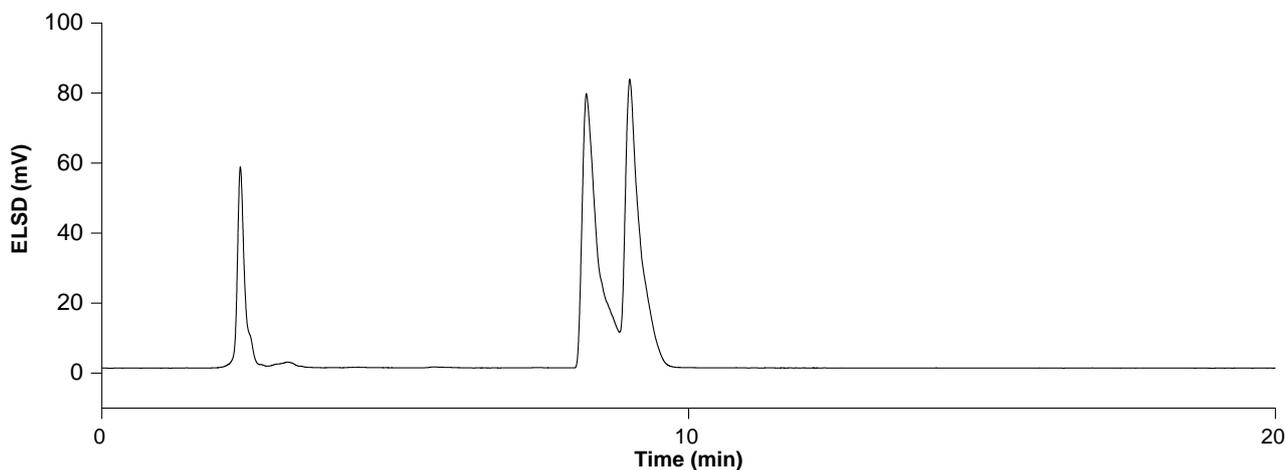
$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.46 – 6.98 (m, 30H), 5.56 (d, $J = 5.0$ Hz, 1H, $\text{H}_1(\beta)$), 5.46 (d, $J = 5.0$ Hz, 1H, $\text{H}_1(\alpha)$), 5.05 (d, $J = 11.0$ Hz, 1H, $\text{H}_1'(\alpha)$), 4.98 (dd, $J = 11.6, 7.3$ Hz, 2H), 4.93 (d, $J = 3.6$ Hz, 1H), 4.88 (dd, $J = 14.2, 11.7$ Hz, 2H), 4.81 – 4.71 (m, 5H), 4.63 – 4.55 (m, 4H), 4.51 – 4.42 (m, 3H, $\text{H}_1'(\beta)$), 4.29 (ddd, $J = 15.8, 5.0, 2.4$ Hz, 2H), 4.24 – 4.19 (m, 2H), 4.18 – 4.03 (m, 5H), 3.98 (dd, $J = 12.7, 9.8$ Hz, 2H), 3.88 – 3.82 (m, 2H), 3.78 – 3.66 (m, 4H), 3.61 (t, $J = 6.2$ Hz, 1H), 3.54 (dd, $J = 9.7, 2.9$ Hz, 1H), 1.48 (s, 6H), 1.44 (s, 3H), 1.42 (s, 3H), 1.31 (s, 3H), 1.31 (s, 3H), 1.30 (s, 3H), 1.30 (s, 4H) ppm.

$^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 139.0, 138.8, 138.6, 138.5, 138.2, 138.0, 128.7, 128.6, 128.6, 128.5, 128.3, 128.1, 128.1, 128.0, 127.9, 127.8, 127.8, 127.8, 127.7, 127.6, 109.5, 109.5, 108.8, 108.6, 104.8 ($\text{C}_1'(\beta)$), 97.7 ($\text{C}_1'(\alpha)$), 96.5 ($\text{C}_1(\beta)$), 96.4 ($\text{C}_1(\alpha)$), 81.7, 79.0, 78.8, 76.4, 74.9, 74.7, 74.5, 74.5, 73.9, 73.8, 73.3, 73.0, 71.6, 71.5, 71.3, 70.9, 70.9, 70.6, 69.8, 67.8, 67.7, 67.6, 66.9, 66.5, 66.4, 26.2, 26.2, 26.1, 26.0, 25.2, 25.0, 24.6, 24.5 ppm.

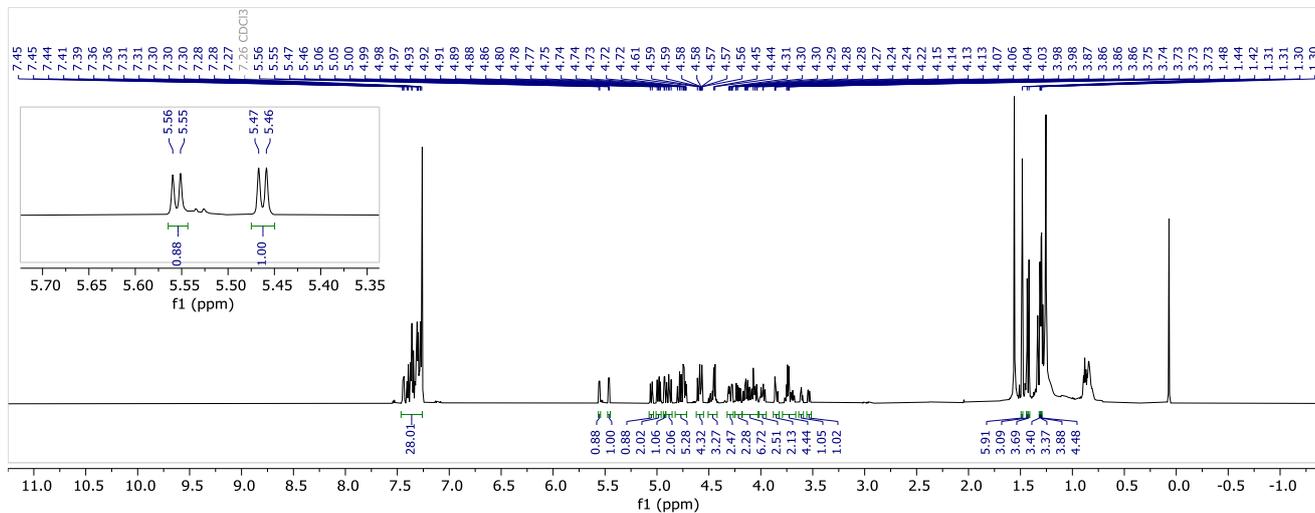
$^{19}\text{F NMR}$ (564 MHz, CDCl_3) δ -74.9 (s), -74.9 (s) ppm.

HRMS (QToF): Calcd for $\text{C}_{41}\text{H}_{47}\text{F}_3\text{O}_{12}\text{Na}$ $[\text{M} + \text{Na}]^+$ 811.2939; found 811.2912.

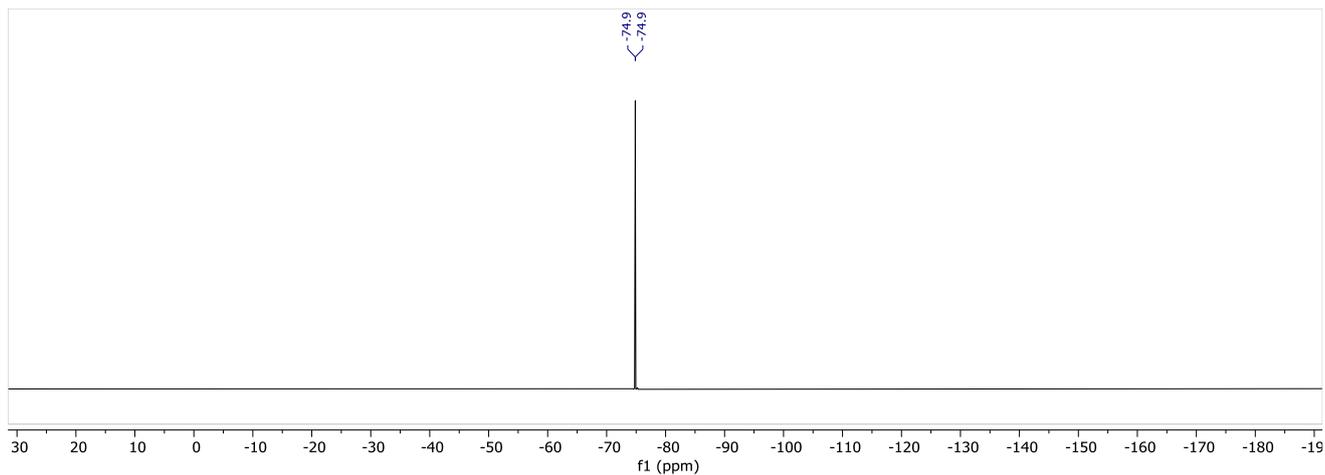
NP-HPLC of **6TFA-A1** (ELSD trace, $t_R(\alpha) = 8.3$ min, $t_R(\beta) = 9.0$ min):



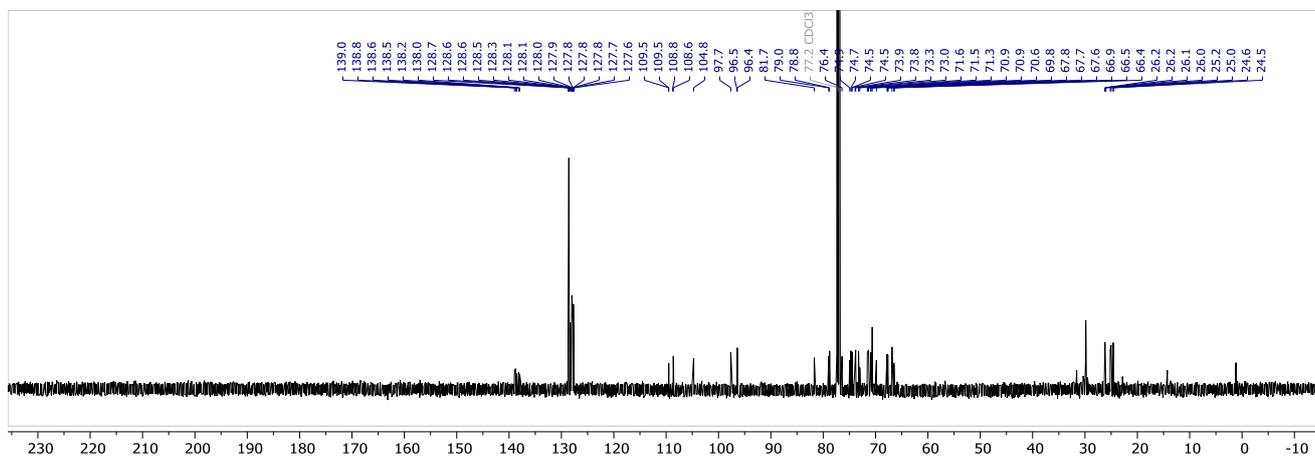
¹H NMR (600 MHz, CDCl₃) of **6TFA-A1**:



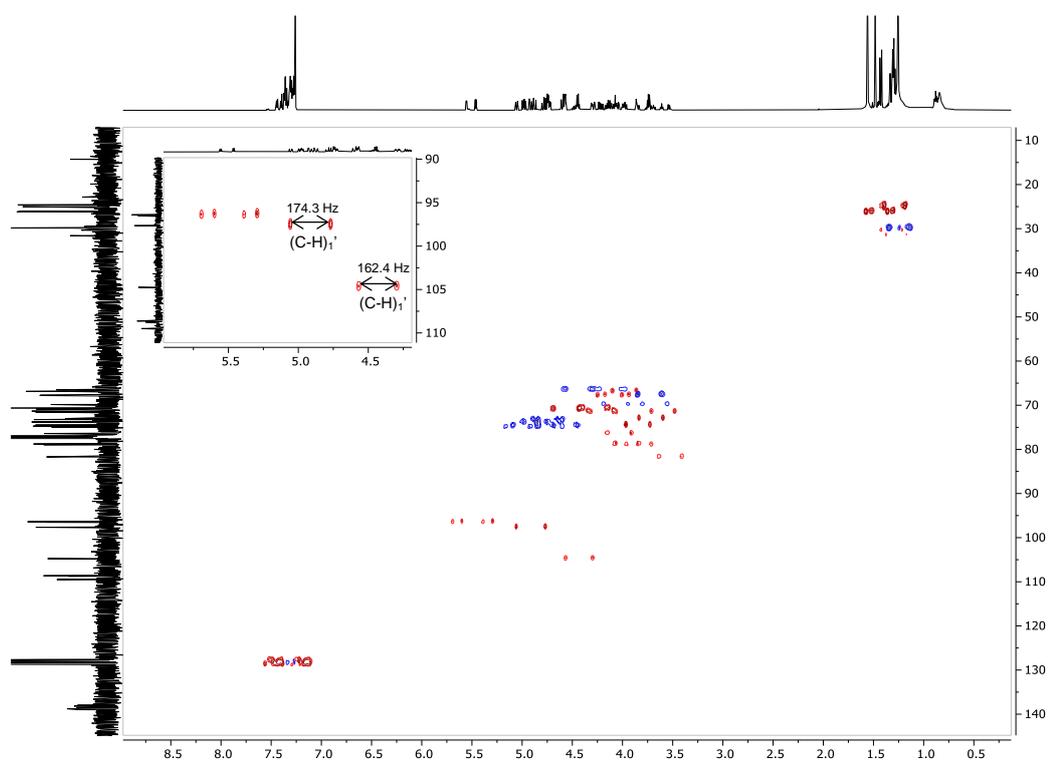
¹⁹F NMR (376 MHz, CDCl₃) of **6TFA-A1**:



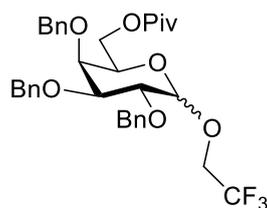
¹³C NMR (151 MHz, CDCl₃) of **6TFA-A1**:



Coupled $^{13}\text{C}, ^1\text{H}$ HSQC of **6TFA-A1**:



2,2,2-Trifluoroethyl 2,3,4-tris-*O*-benzyl-6-(2,2-dimethylpropanoate)- α/β -D-galactopyranoside (6Piv-A3)

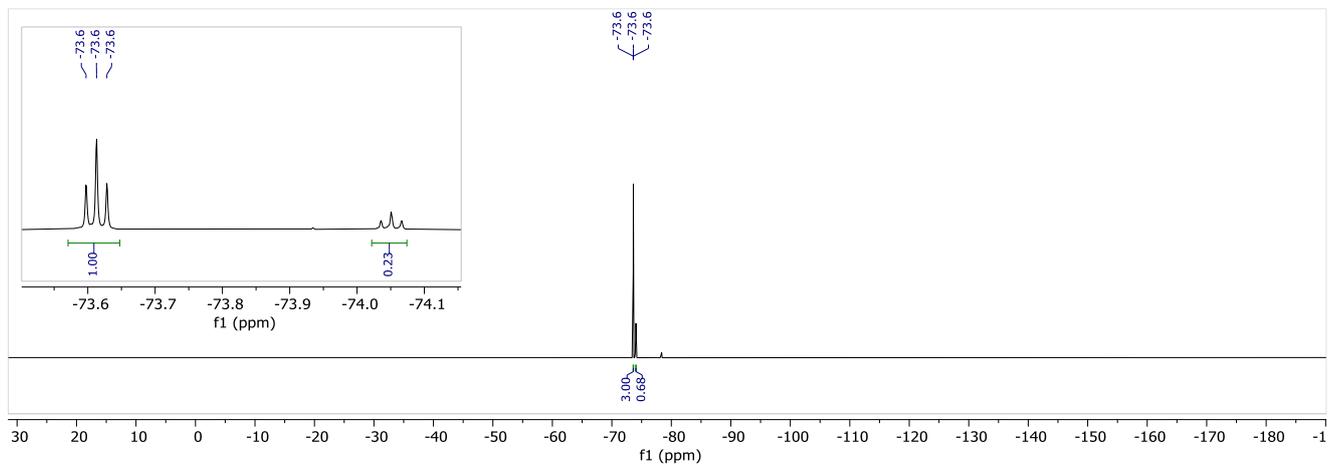


The title compound was prepared according to general procedure for glycosylations. Product **6Piv-A3** (18 mg, 29 μmol , 85%, α/β 82:18) was obtained as a colorless oil after purification using **Method-2b** ($t_{\text{R}}(\alpha) = 6.7$ min).

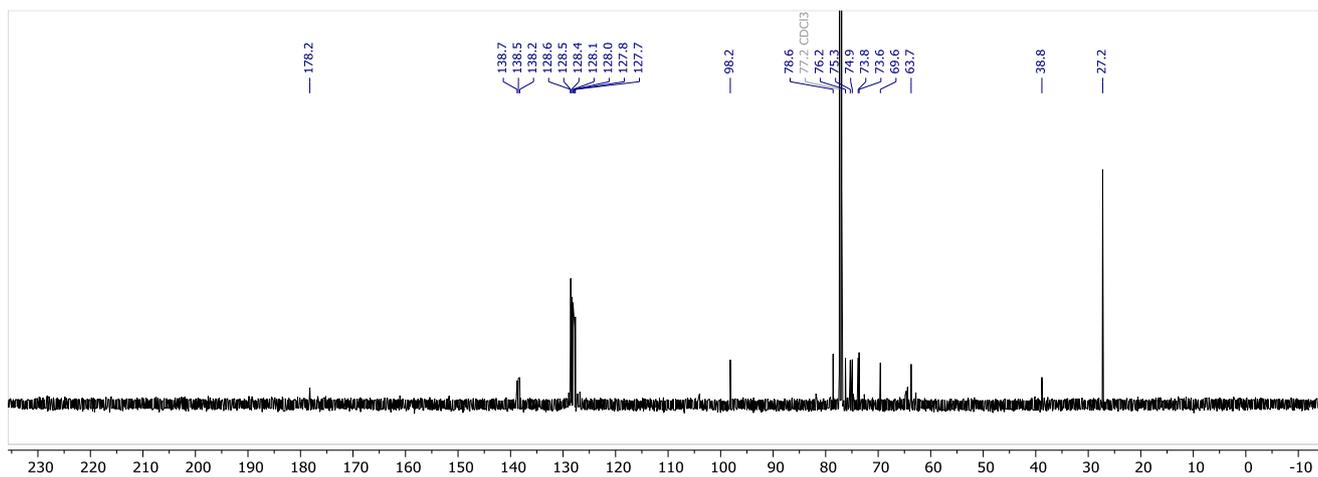
Data of the major isomer (α):

^1H NMR (600 MHz, CDCl_3) δ 7.42 – 7.27 (m, 15H), 4.99 (d, $J = 11.2$ Hz, 1H), 4.93 – 4.81 (m, 3H), 4.75 (d, $J = 11.6$ Hz, 1H, H_1), 4.66 (d, $J = 12.2$ Hz, 1H), 4.59 (d, $J = 11.3$ Hz, 1H), 4.17 (dd, $J = 11.4, 7.5$ Hz, 1H), 4.11 – 4.01 (m, 2H), 3.98 – 3.81 (m, 5H), 1.17 (s, 9H) ppm.

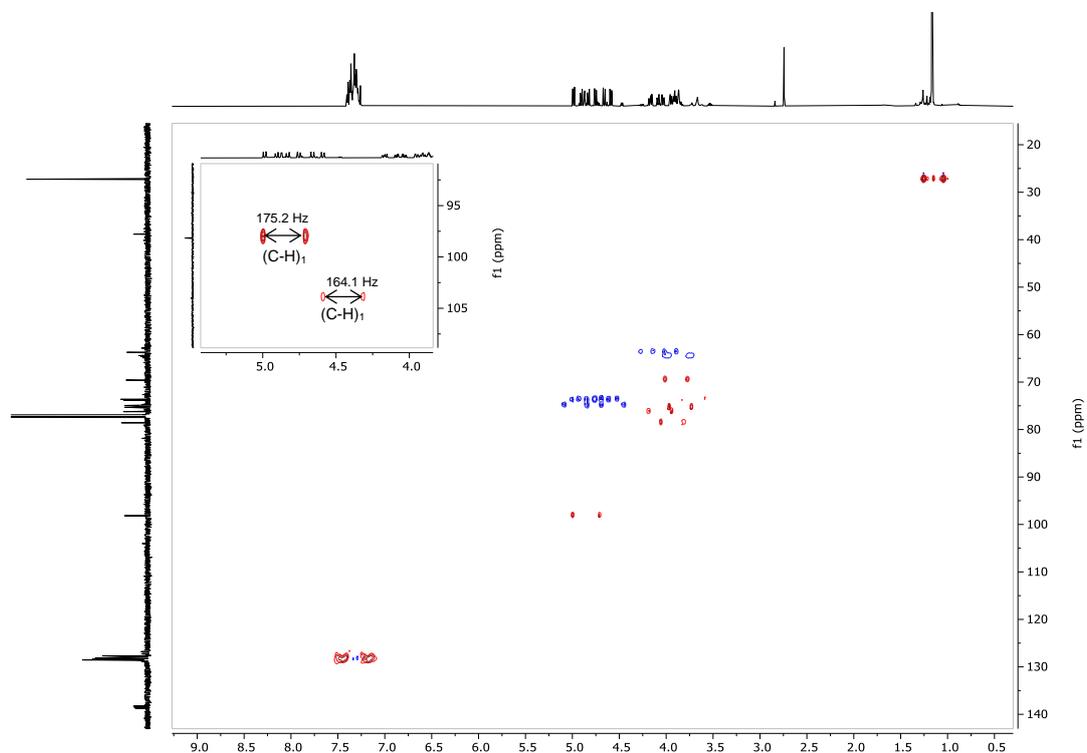
^{19}F NMR (376 MHz, CDCl_3) of **6Piv-A3**:



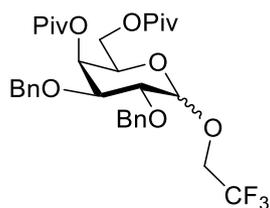
^{13}C NMR (151 MHz, CDCl_3) of **6Piv-A3**:



Coupled ^{13}C , ^1H HSQC of **6Piv-A3**:



2,2,2-Trifluoroethyl 2,3-bis-*O*-benzyl-4,6-bis-(2,2-dimethylpropanoate)- α/β -D-galactopyranoside (4,6Piv-A3)



The title compound was prepared according to general procedure for glycosylations. Product **4,6Piv-A3** (13 mg, 21 μmol , 61%, α/β 96:4) was obtained as a colorless oil after purification using **Method-2b** ($t_{\text{R}}(\alpha) = 4.5$ min).

Data of the major isomer (α):

^1H NMR (600 MHz, CDCl_3) δ 7.33 – 7.27 (m, 10H), 5.56 (dd, $J = 3.5, 1.4$ Hz, 1H), 4.88 (s, 1H, H_1), 4.82 (d, $J = 12.2$ Hz, 1H), 4.72 (d, $J = 10.8$ Hz, 1H), 4.64 (d, $J = 12.2$ Hz, 1H), 4.55 (d, $J = 10.7$ Hz, 1H), 4.16 (td, $J =$

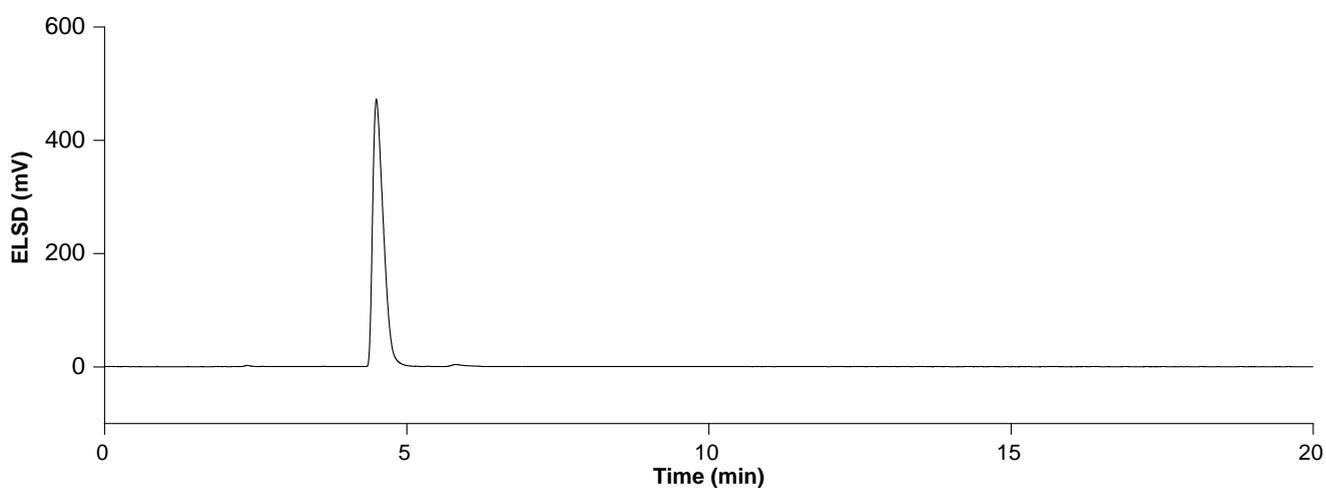
6.5, 1.4 Hz, 1H), 4.10 – 4.02 (m, 2H), 3.99 (dd, $J = 10.1, 3.4$ Hz, 1H), 3.93 (q, $J = 8.6$ Hz, 2H), 3.76 (dd, $J = 10.0, 3.7$ Hz, 1H), 1.19 (s, 9H), 1.16 (s, 9H) ppm.

^{13}C NMR (151 MHz, CDCl_3) δ 178.1, 177.4, 138.2, 138.1, 128.5, 128.3, 128.2, 128.1, 128.0, 127.7, 98.6 (C_1), 75.8, 74.4, 73.7, 72.2, 68.0, 67.1, 62.5, 39.2, 38.9, 27.3, 27.2 ppm.

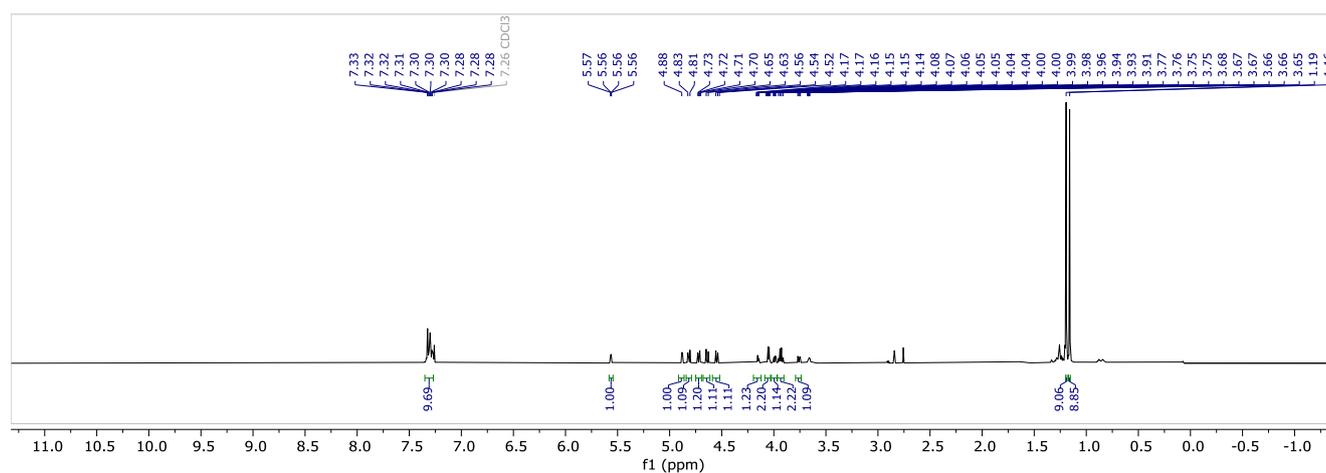
^{19}F NMR (564 MHz, CDCl_3) δ -73.63 (t, $J = 8.7$ Hz) ppm.

HRMS (QToF): Calcd for $\text{C}_{32}\text{H}_{41}\text{F}_3\text{O}_8\text{Na}$ $[\text{M} + \text{Na}]^+$ 633.2646; found 633.2651.

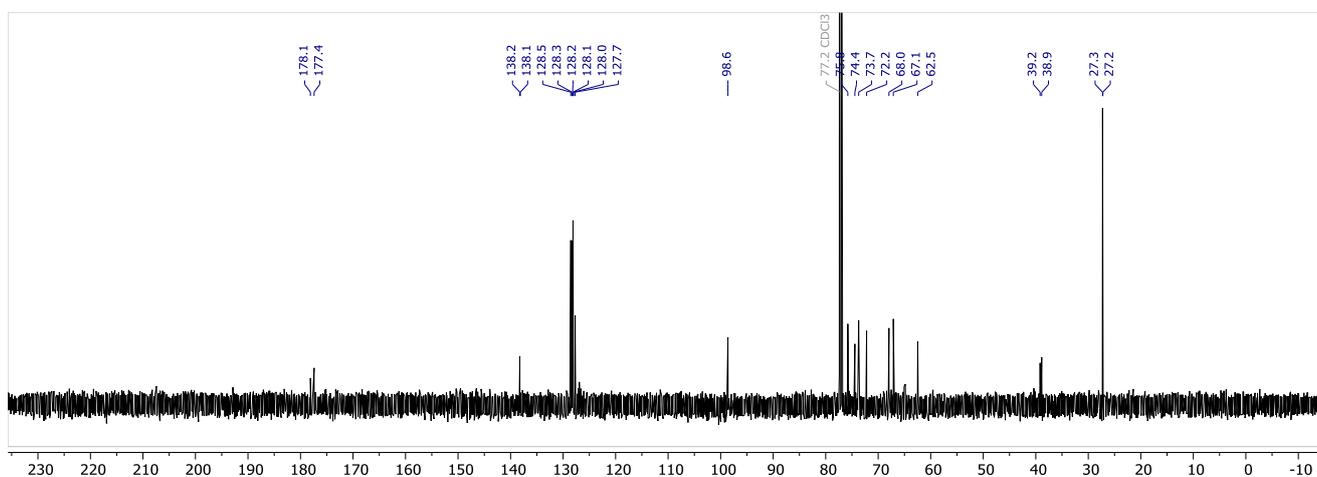
NP-HPLC of **4,6Piv-A3** (ELSD trace, $t_R(\alpha) = 4.5$ min):



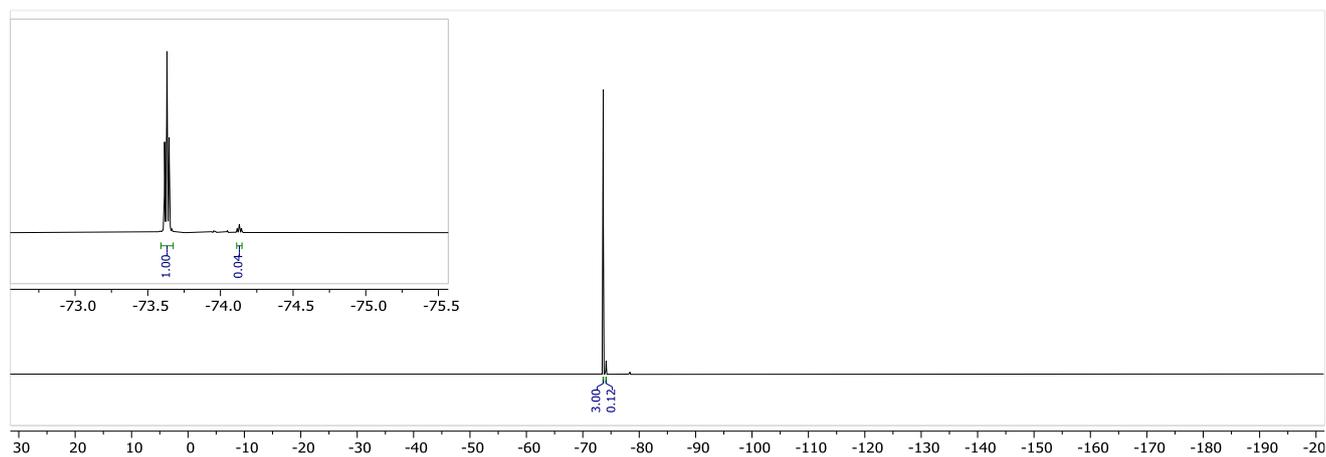
^1H NMR (600 MHz, CDCl_3) of **4,6Piv-A3**:



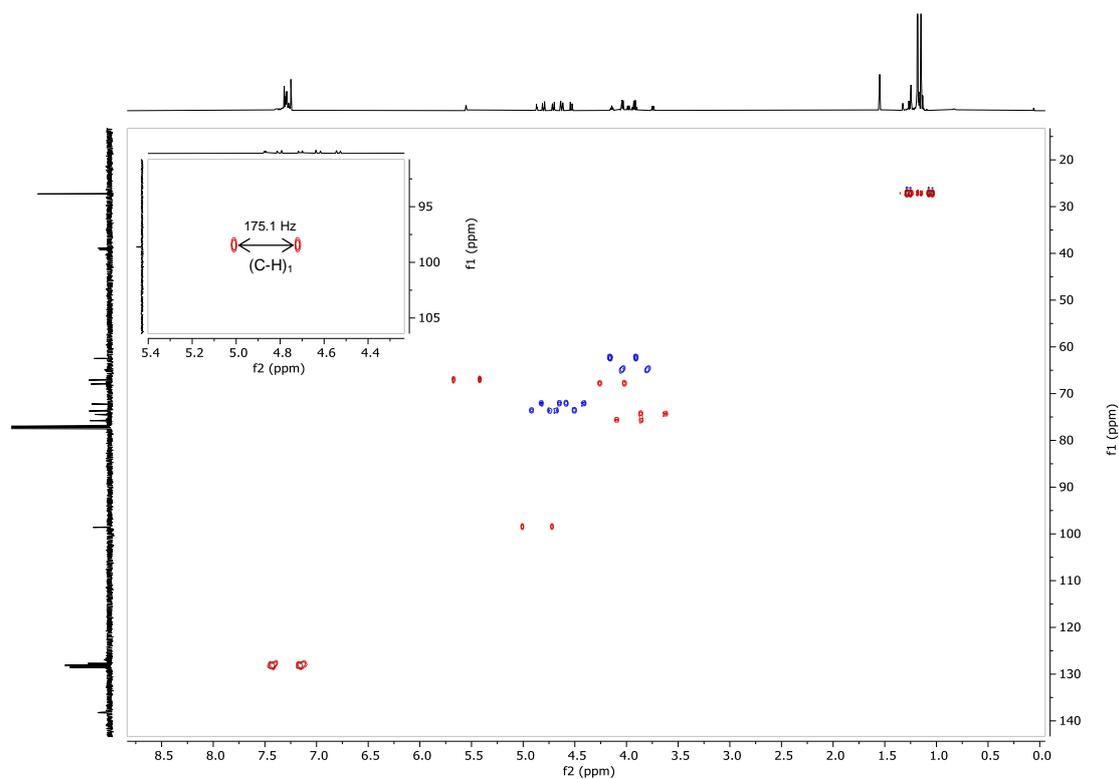
^{13}C NMR (151 MHz, CDCl_3) of **4,6Piv-A3**:



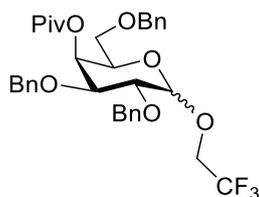
^{19}F NMR (376 MHz, CDCl_3) of **4,6Piv-A3**:



Coupled ^{13}C , ^1H HSQC of **4,6Piv-A3**:



2,2,2-Trifluoroethyl 2,3,6-tris-*O*-benzyl-4-(2,2-dimethylpropanoate)- α/β -D-galactopyranoside (4Piv-A3)



The title compound was prepared according to general procedure for glycosylations. Product **4Piv-A3** (15 mg, 24 μmol , 71%, α/β 97:3) was obtained as a colorless oil after purification using **Method-2b** ($t_{\text{R}}(\alpha) = 5.5$ min).

Data of the major isomer (α):

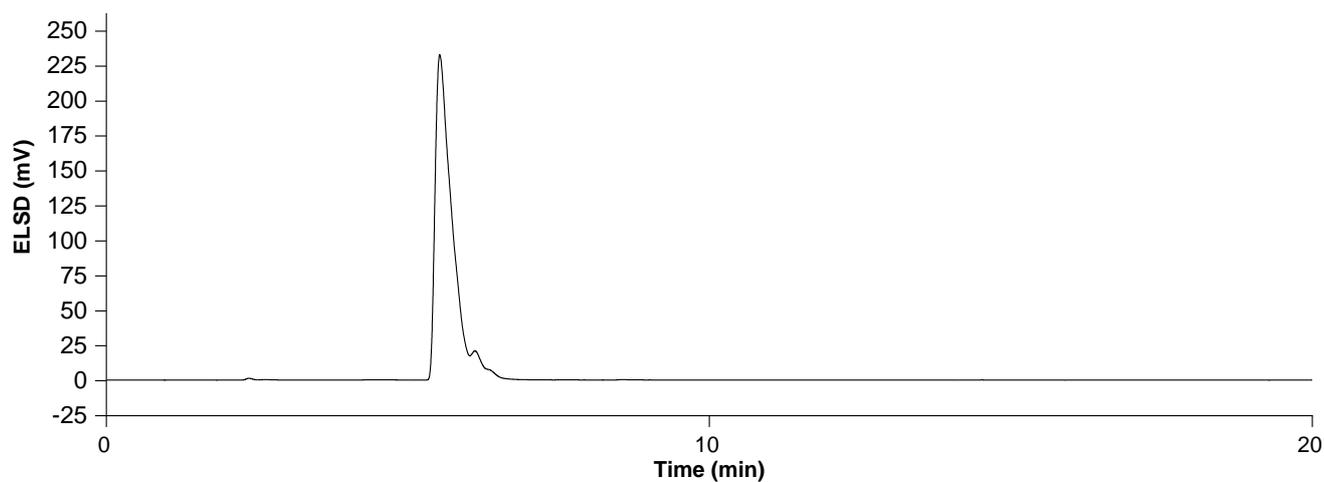
^1H NMR (600 MHz, CDCl_3) δ 7.37 – 7.27 (m, 15H), 5.62 (dd, $J = 3.4, 1.3$ Hz, 1H), 4.90 (d, $J = 3.7$ Hz, 1H, **H**₁), 4.81 (d, $J = 12.2$ Hz, 1H), 4.73 (d, $J = 10.8$ Hz, 1H), 4.64 (d, $J = 12.2$ Hz, 1H), 4.55 – 4.50 (m, 2H), 4.45 (d, $J = 11.8$ Hz, 1H), 4.11 (td, $J = 6.3, 1.3$ Hz, 1H), 4.00 – 3.91 (m, 3H), 3.76 (dd, $J = 10.0, 3.7$ Hz, 1H), 3.50 – 3.41 (m, 2H), 1.13 (s, 9H) ppm.

^{13}C NMR (151 MHz, CDCl_3) δ 177.4, 138.4, 138.3, 137.8, 128.6, 128.5, 128.3, 128.1, 128.1, 128.0, 127.9, 127.6, 98.6 (C_1), 76.1, 74.6, 73.8, 73.6, 72.0, 68.8, 68.6, 67.4, 39.1, 27.3 ppm.

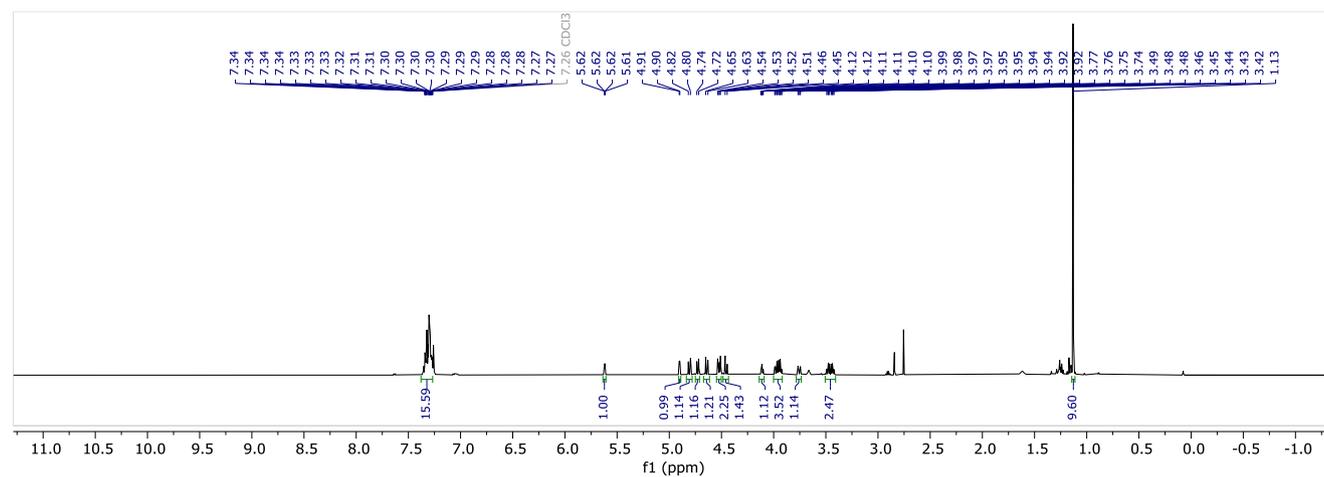
^{19}F NMR (564 MHz, CDCl_3) δ -73.54 (t, $J = 8.7$ Hz).

HRMS (QToF): Calcd for $\text{C}_{34}\text{H}_{39}\text{F}_3\text{O}_7\text{Na}$ $[\text{M} + \text{Na}]^+$ 639.2540; found 639.2535.

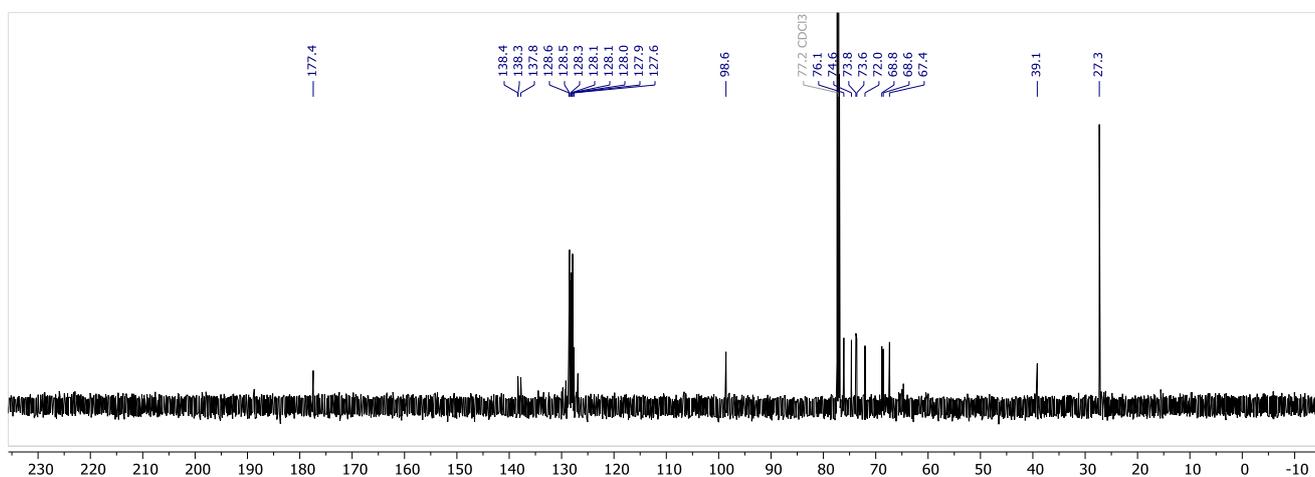
NP-HPLC of **4Piv-A3** (ELSD trace, $t_R(\alpha) = 5.5$ min):



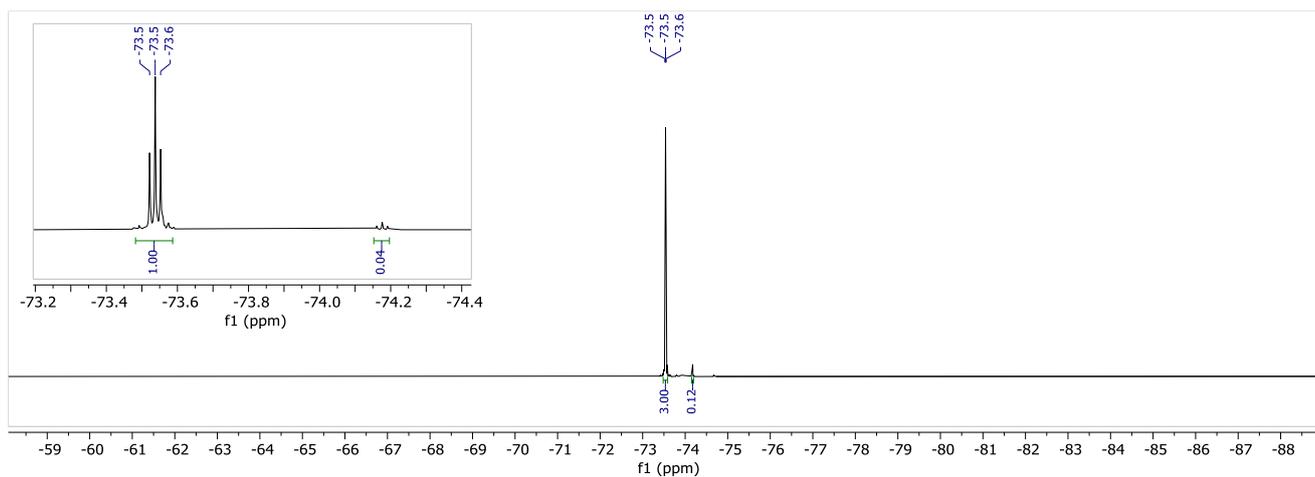
^1H NMR (600 MHz, CDCl_3) of **4Piv-A3**:



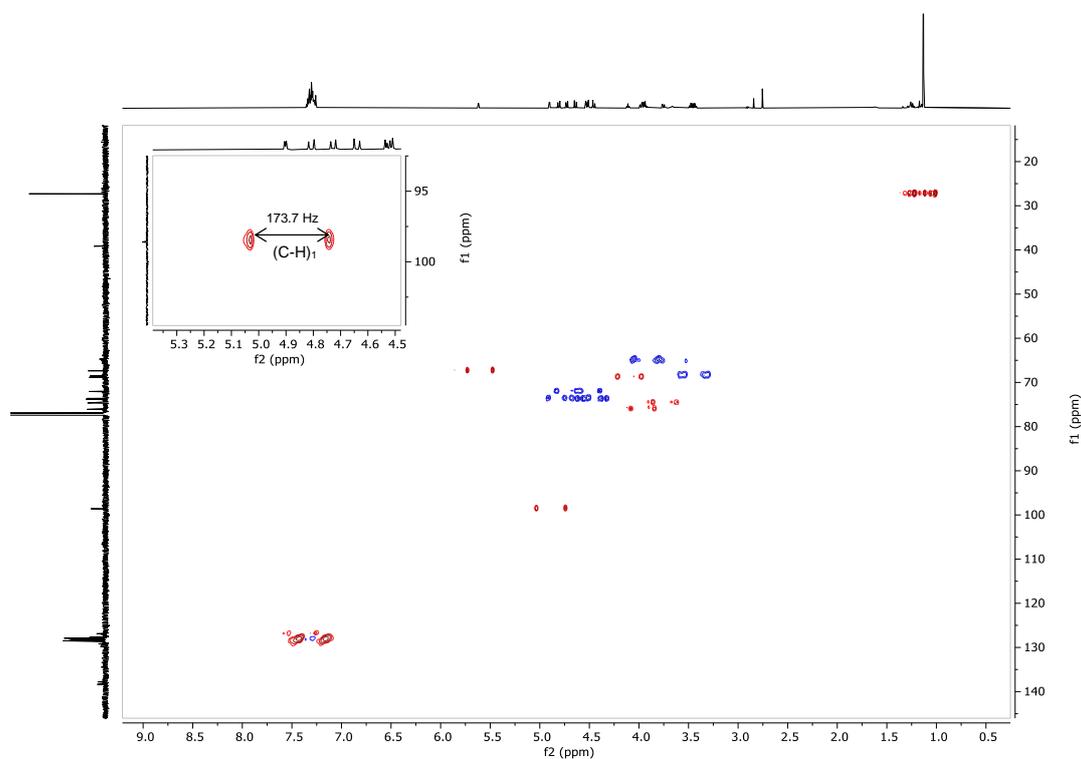
^{13}C NMR (151 MHz, CDCl_3) of **4Piv-A3**:



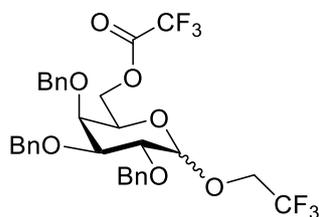
^{19}F NMR (376 MHz, CDCl_3) of **4Piv-A3**:



Coupled $^{13}\text{C}, ^1\text{H}$ HSQC of **4Piv-A3**:



2,2-Trifluoroethyl 2,3,4-tris-*O*-benzyl-6-(trifluoroacetate)- α/β -D-galactopyranoside (6TFA-A3)



The title compound was prepared according to general procedure for glycosylations. Product **6TFA-A3** (14 mg, 22 μmol , 66%, α/β 96:4) was obtained as a colorless oil.

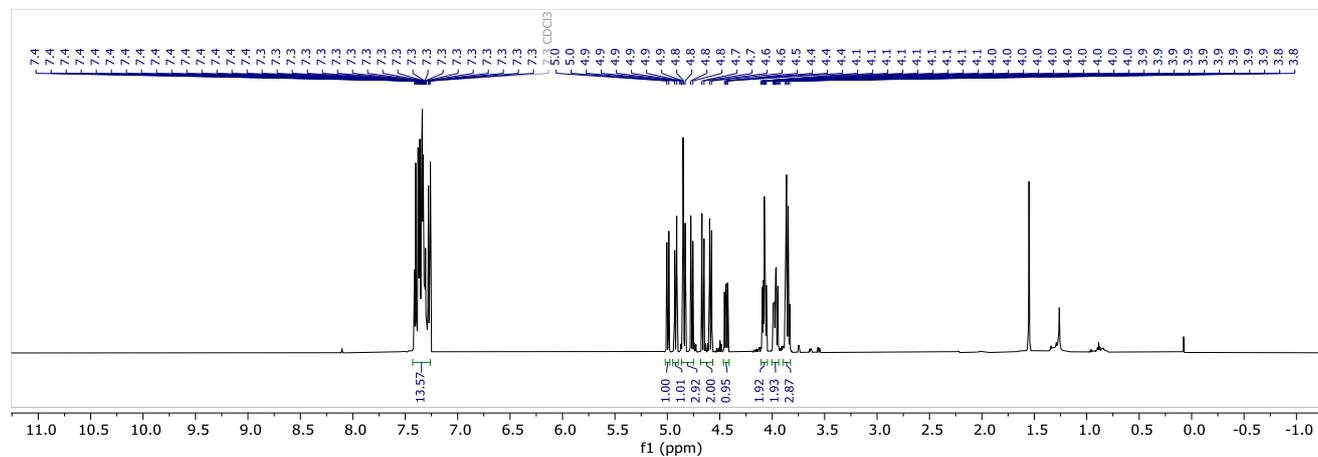
Data of the major isomer (α):

$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.45 – 7.26 (m, 15H), 5.00 (d, $J = 11.6$ Hz, 1H), 4.92 (d, $J = 11.6$ Hz, 1H), 4.86 – 4.75 (m, 3H, H_1), 4.69 – 4.57 (m, 2H), 4.44 (dd, $J = 11.3, 7.9$ Hz, 1H), 4.07 (ddd, $J = 11.4, 6.9, 3.9$ Hz, 2H), 4.01 – 3.92 (m, 2H), 3.90 – 3.83 (m, 3H) ppm.

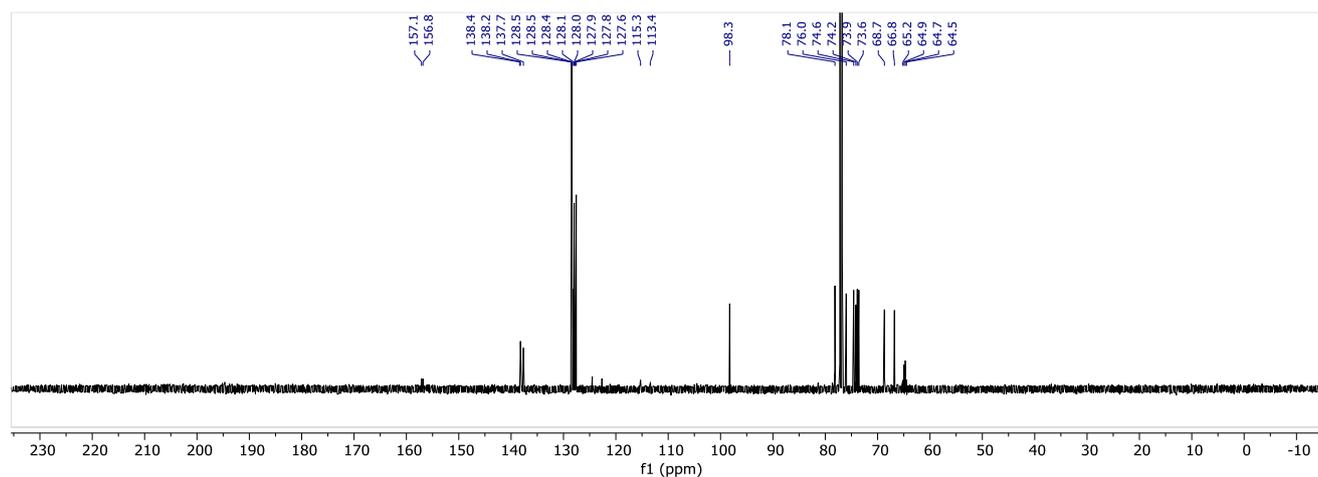
$^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 157.1, 156.8, 138.4, 138.2, 137.7, 128.5, 128.5, 128.4, 128.1, 128.0, 127.9, 127.8, 127.6, 115.3, 113.4, 98.3 (C_1), 78.1, 76.0, 74.6, 74.2, 73.9, 73.6, 68.7, 66.8, 64.8 (q, $J = 35.0$ Hz) ppm.

^{19}F NMR (564 MHz, CDCl_3) δ -73.87 (t, $J = 8.7$ Hz), -75.11 ppm.

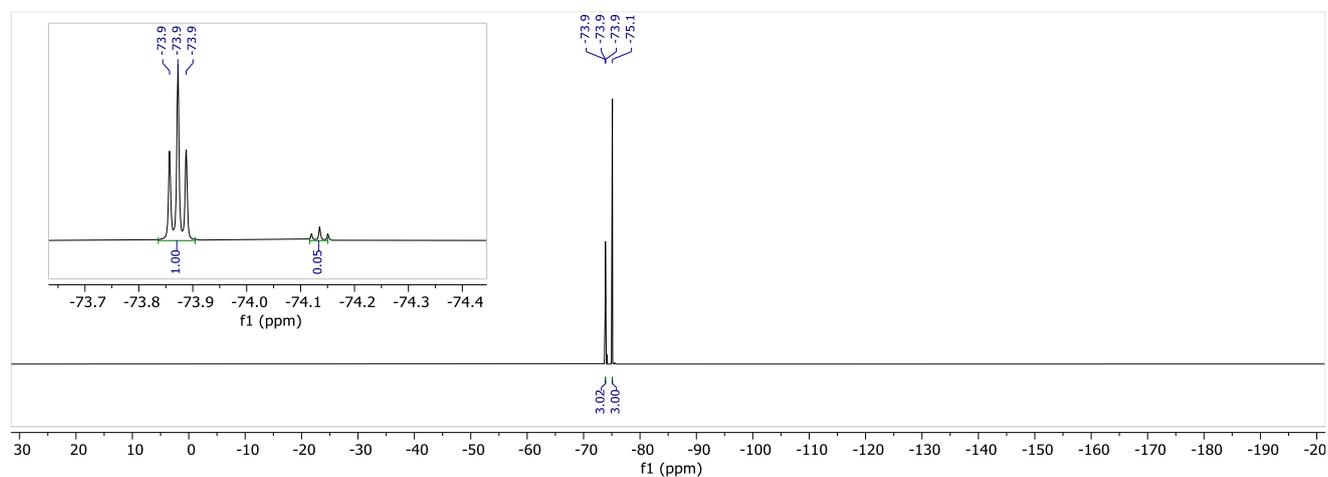
^1H NMR (600 MHz, CDCl_3) of **6TFA-A3**:



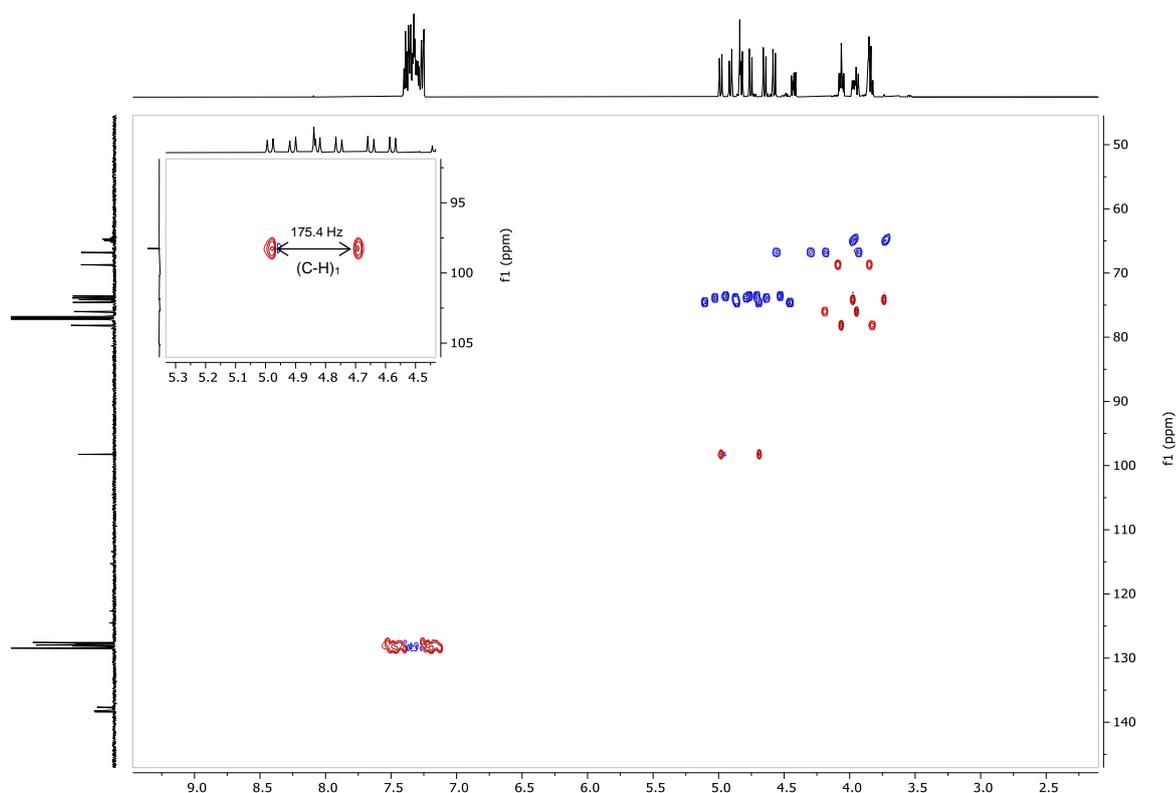
^{13}C NMR (151 MHz, CDCl_3) of **6TFA-A3**:



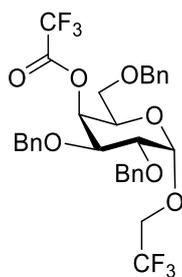
^{19}F NMR (376 MHz, CDCl_3) of **6TFA-A3**:



Coupled $^{13}\text{C}, ^1\text{H}$ HSQC of **6TFA-A3**:



2,2,2-Trifluoroethyl 2,3,6-tris-*O*-benzyl-4-(trifluoroacetate)- α -D-galactopyranoside (4TFA-A3)



The title compound was prepared according to general procedure for glycosylations. Product **4TFA-A3** (15 mg, 24 μmol , 71%, α -only) was obtained as a colorless oil.

Data of the major isomer (α):

^1H NMR (600 MHz, CDCl_3) δ 7.38 – 7.26 (m, 15H), 5.73 (dd, $J = 3.2, 1.3$ Hz, 1H), 4.85 (d, $J = 3.7$ Hz, 1H, **H**₁), 4.83 – 4.72 (m, 2H), 4.61 (dd, $J = 11.5, 2.2$ Hz, 2H), 4.53 – 4.41 (m, 2H), 4.14 (ddd, $J = 7.3, 5.7, 1.3$ Hz,

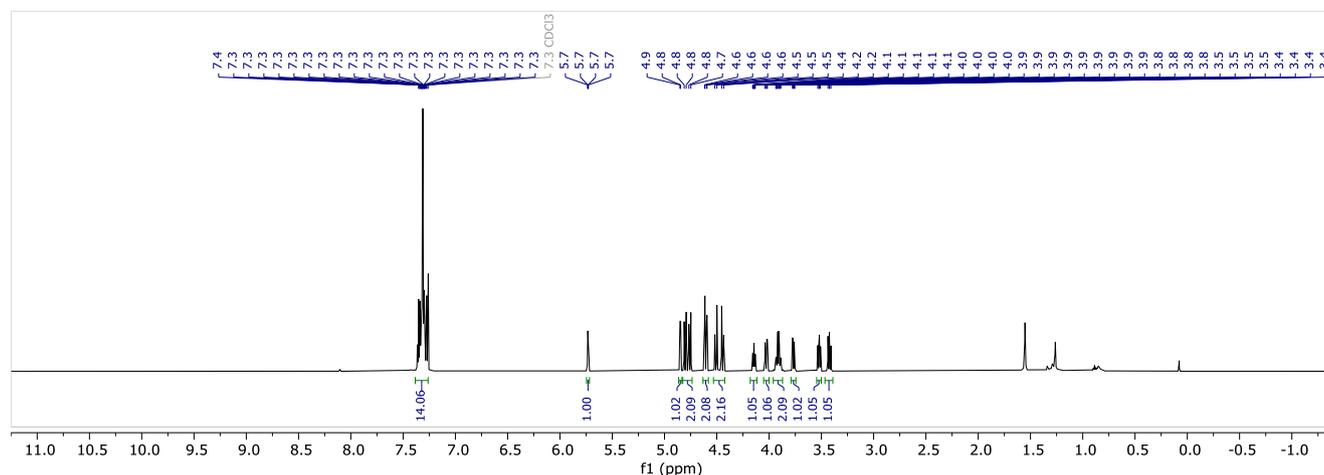
1H), 4.03 (dd, $J = 10.0, 3.3$ Hz, 1H), 3.91 (qd, $J = 8.6, 2.8$ Hz, 2H), 3.77 (dd, $J = 10.0, 3.7$ Hz, 1H), 3.52 (dd, $J = 9.4, 5.8$ Hz, 1H), 3.42 (dd, $J = 9.3, 7.9$ Hz, 1H) ppm.

^{13}C NMR (151 MHz, CDCl_3) δ 156.92 (q, $J = 42.4$ Hz), 138.1, 137.7, 137.3, 128.7, 128.6, 128.5, 128.2, 128.2, 128.1, 127.9, 127.9, 113.8, 98.7 (C_1), 75.3, 74.8, 73.9, 73.9, 72.8, 72.5, 67.5, 67.3, 65.3 (q, $J = 35.0$ Hz) ppm.

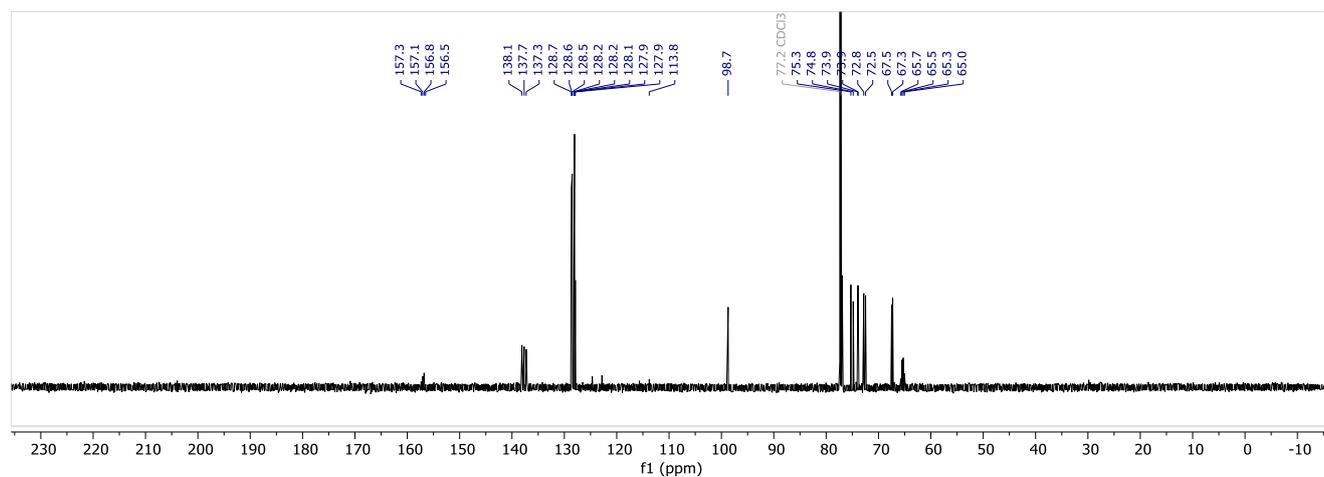
^{19}F NMR (564 MHz, CDCl_3) δ -73.69 (t, $J = 8.7$ Hz), -74.82 ppm.

HRMS (QToF): Calcd for $\text{C}_{31}\text{H}_{30}\text{F}_6\text{O}_7\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 651.1788; found 651.1815.

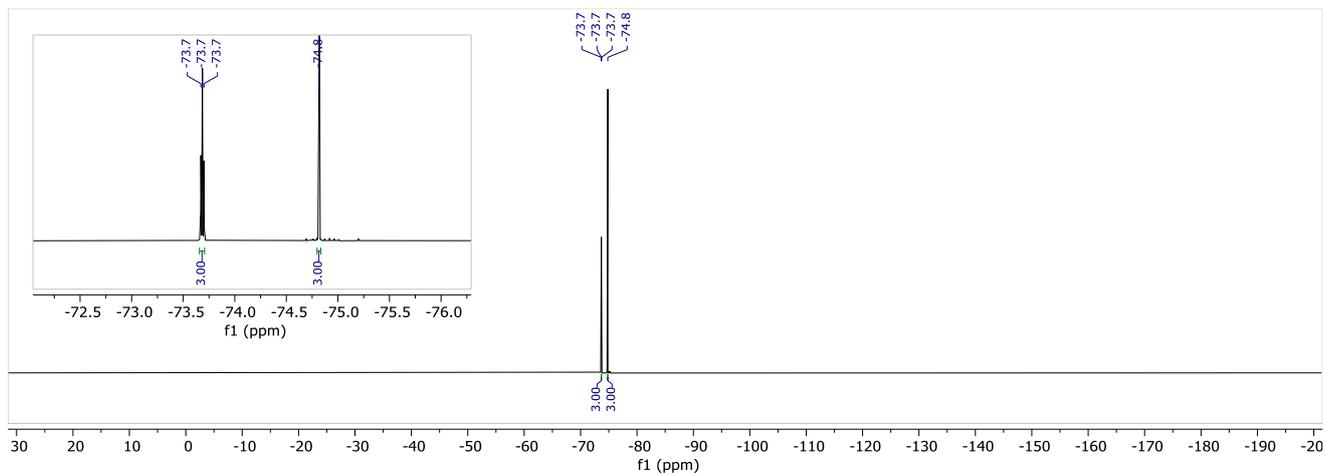
^1H NMR (600 MHz, CDCl_3) of **4TFA-A3**:



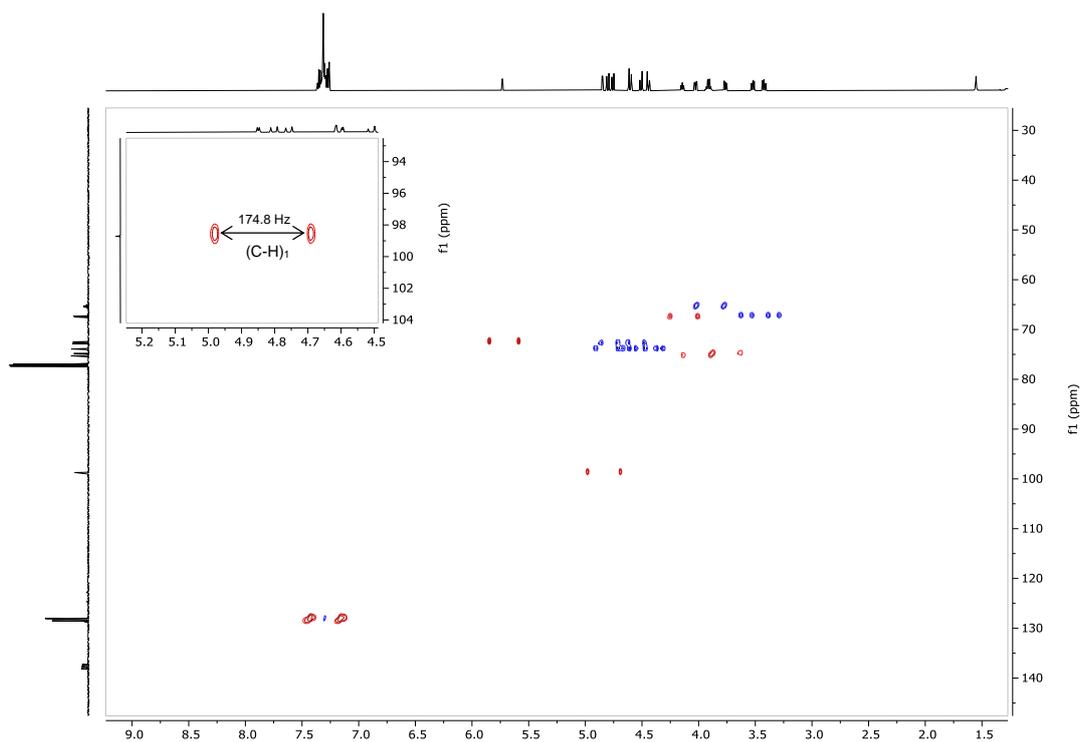
^{13}C NMR (151 MHz, CDCl_3) of **4TFA-A3**:



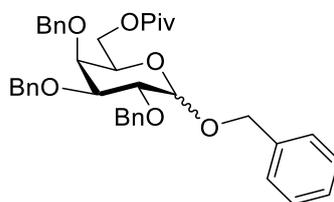
^{19}F NMR (376 MHz, CDCl_3) of **4TFA-A3**:



Coupled ^{13}C , ^1H HSQC of **4TFA-A3**:



Benzyl 2,3,4-tris-*O*-benzyl-6-(2,2-dimethylpropanoate)- α/β -D-galactopyranoside (**6Piv-A4**)



The title compound was prepared according to general procedure for glycosylations. Product **6Piv-A4** (18 mg, 29 μ mol, 83%, α/β 36:64) was obtained as a colorless oil after purification using **Method-2b** (t_R (α/β) = 6.7 min).

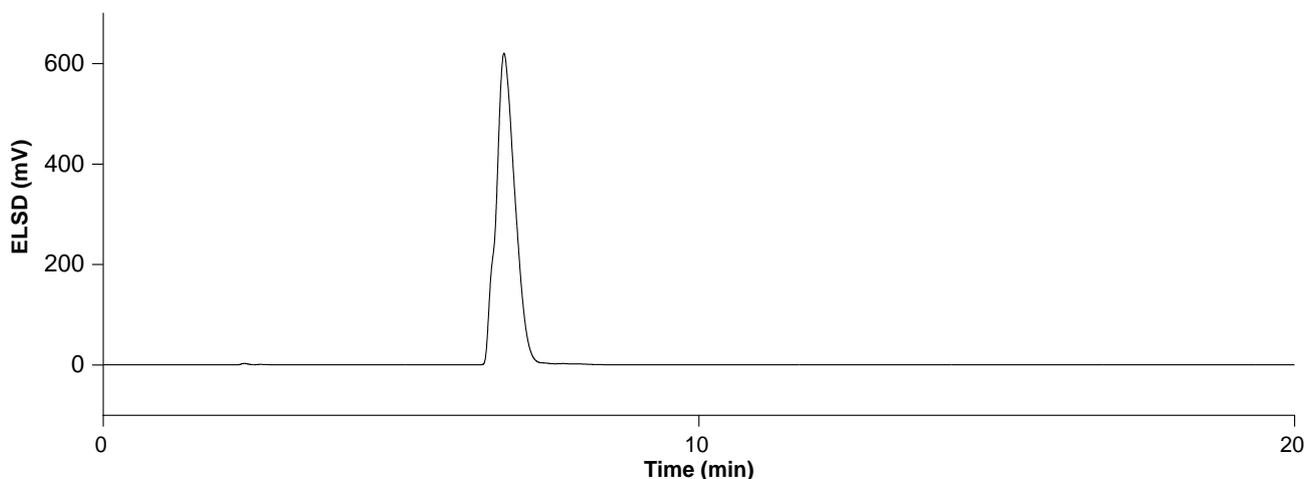
Data of the anomeric mixture:

$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.44 – 7.27 (m, 26H), 5.03 – 4.88 (m, 4H, **$\text{H}_1(\alpha)$**), 4.85 – 4.69 (m, 5H), 4.68 – 4.52 (m, 4H), 4.45 (d, $J = 7.7$ Hz, 1H, **$\text{H}_1(\beta)$**), 4.31 (dd, $J = 11.1, 7.0$ Hz, 1H), 4.18 (dd, $J = 11.2, 7.4$ Hz, 1H), 4.13 – 3.95 (m, 3H), 3.91 (dd, $J = 9.7, 7.7$ Hz, 1H), 3.88 – 3.85 (m, 1H), 3.74 (d, $J = 2.8$ Hz, 1H), 3.57 – 3.49 (m, 2H), 1.19 (s, 4H), 1.19 (s, 9H) ppm.

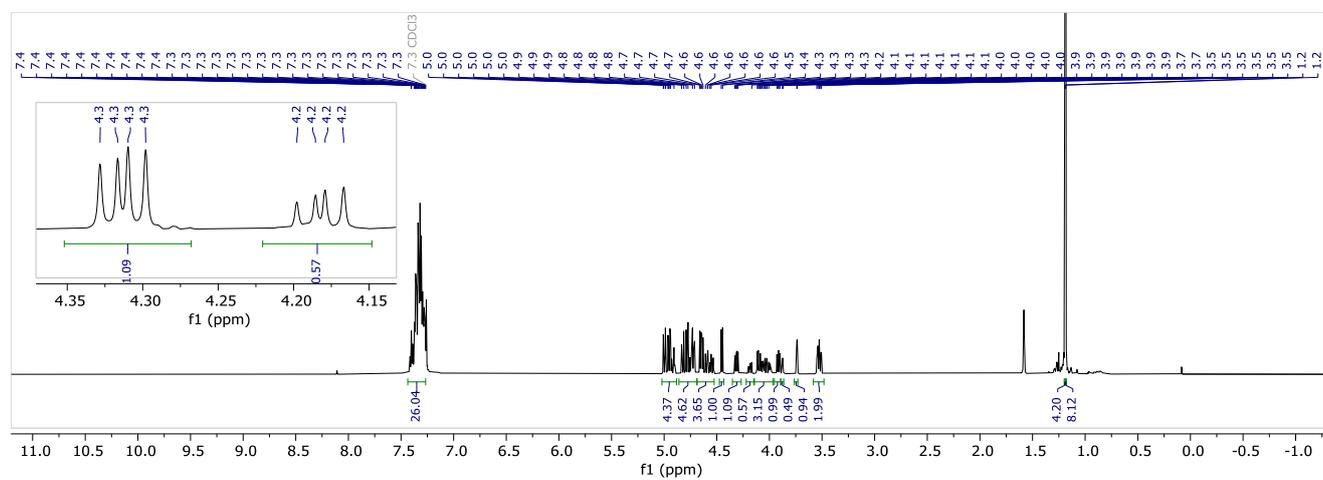
$^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 178.2, 138.9, 138.8, 138.6, 138.6, 138.5, 138.4, 137.6, 137.2, 128.5, 128.5, 128.5, 128.4, 128.4, 128.3, 128.3, 128.1, 128.0, 127.9, 127.8, 127.7, 102.7 **$\text{C}_1(\beta)$** , 95.7 **$\text{C}_1(\alpha)$** , 82.3, 79.6, 79.2, 76.5, 75.5, 75.3, 74.9, 74.7, 73.9, 73.7, 73.6, 73.3, 72.3, 70.9, 68.9, 68.7, 63.8, 63.2, 38.9, 27.3 ppm.

HRMS (QToF): Calcd for $\text{C}_{39}\text{H}_{44}\text{O}_7\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 647.2979; found 647.2974.

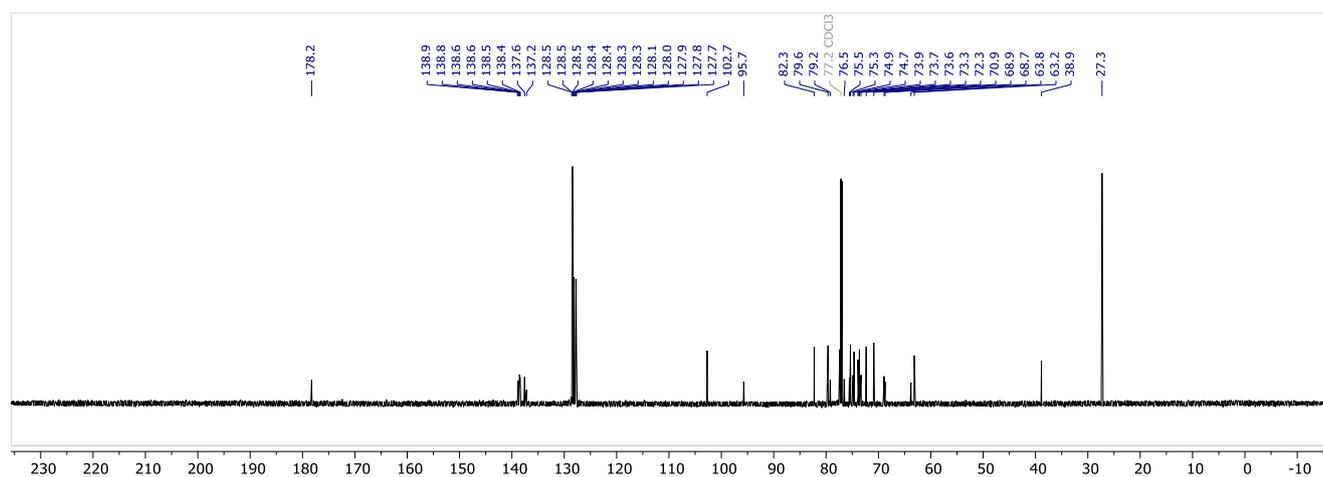
NP-HPLC of **6Piv-A4** (ELSD trace, t_R (α/β) = 6.7 min):



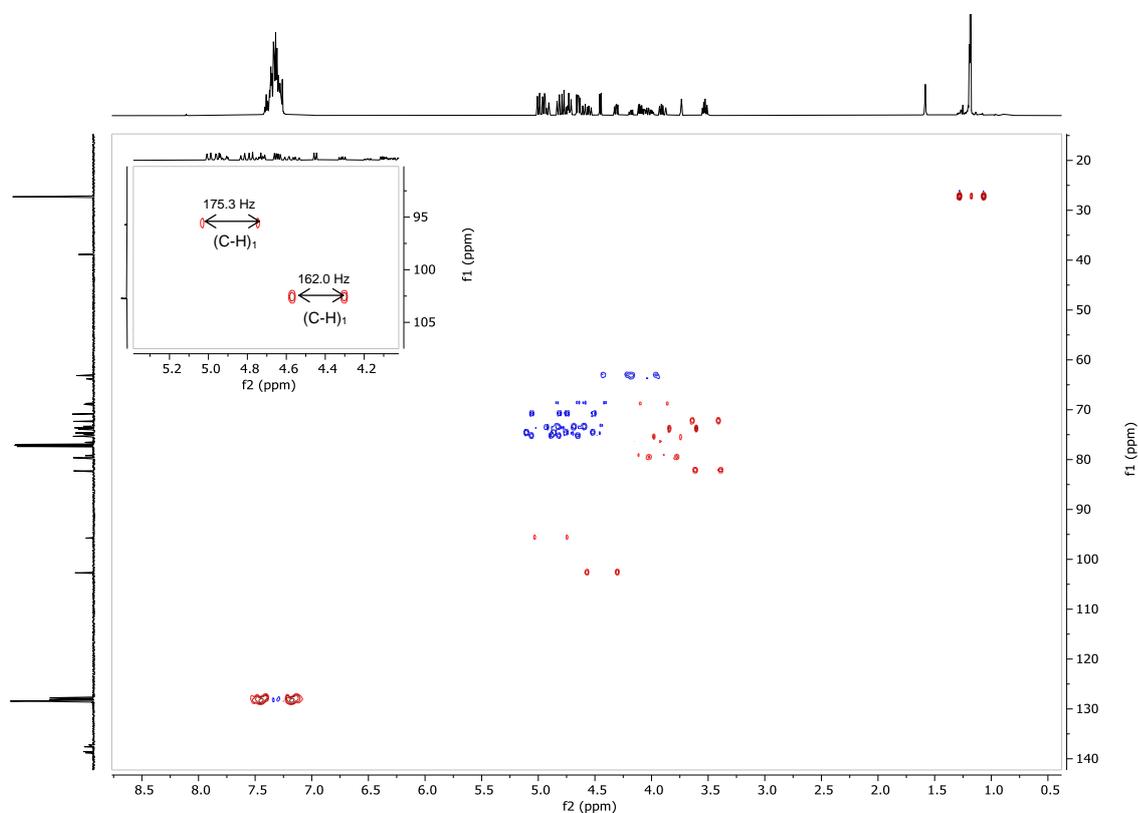
^1H NMR (600 MHz, CDCl_3) of **6Piv-A4**:



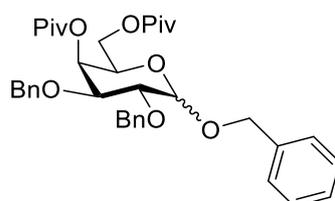
^{13}C NMR (151 MHz, CDCl_3) of **6Piv-A4**:



Coupled $^{13}\text{C}, ^1\text{H}$ HSQC of **6Piv-A4**:



Benzyl 2,3-bis-*O*-benzyl-4,6-bis-(2,2-dimethylpropanoate)- α/β -D-galactopyranoside (4,6Piv-A4)



The title compound was prepared according to general procedure for glycosylations. Product **4,6Piv-A4** (18 mg, 29 μmol , 83%, α/β **66:34**) was obtained as a colorless oil after purification using **Method-2b** ($t_{\text{R}}(\alpha) = 5.0$ min, $t_{\text{R}}(\beta) = 5.6$ min).

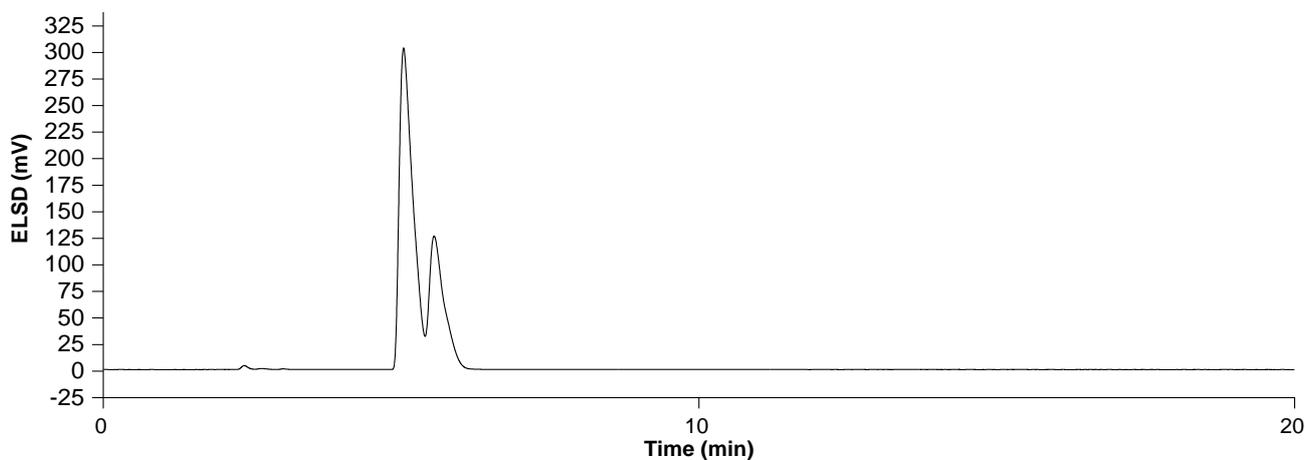
Data of the anomeric mixture:

$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.46 – 7.27 (m, 22H), 5.59 (d, $J = 2.9$ Hz, 1H), 5.50 (d, $J = 3.4$ Hz, 1H), 5.01 – 4.87 (m, 2H, $\text{H}_1(\alpha)$), 4.83 – 4.67 (m, 5H), 4.64 – 4.49 (m, 4H, $\text{H}_1(\beta)$), 4.29 – 4.12 (m, 2H), 4.12 – 4.04 (m, 3H), 3.83 (t, $J = 6.8$ Hz, 1H), 3.75 (dd, $J = 10.0, 3.7$ Hz, 1H), 3.66 – 3.55 (m, 1H), 1.25 (s, 4H), 1.25 (s, 9H), 1.23 (s, 4H), 1.18 (d, $J = 0.8$ Hz, 9H) ppm.

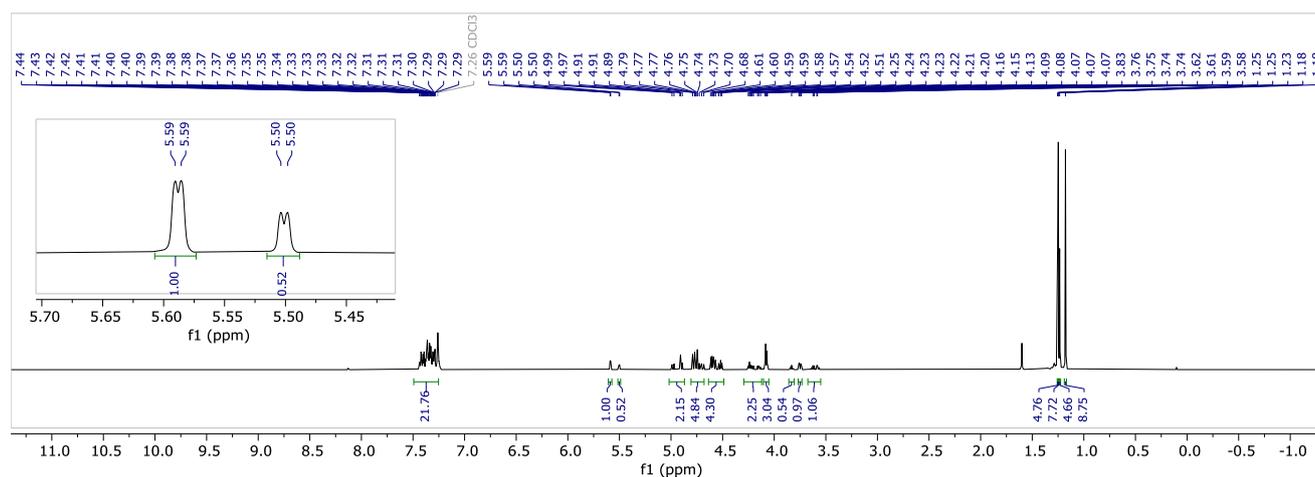
^{13}C NMR (151 MHz, CDCl_3) δ 178.2, 178.1, 177.7, 177.5, 138.6, 138.4, 138.4, 138.1, 137.4, 137.0, 128.6, 128.6, 128.4, 128.4, 128.3, 128.3, 128.3, 128.1, 128.1, 128.0, 127.8, 127.7, 127.6, 102.5 $\text{C}_1(\beta)$, 96.1 $\text{C}_1(\alpha)$, 79.4, 78.7, 76.4, 75.4, 74.9, 73.5, 72.2, 72.2, 71.2, 71.1, 69.1, 67.4, 67.3, 66.2, 62.6, 62.1, 39.2, 39.1, 38.9, 27.3, 27.3, 27.3 ppm.

HRMS (QToF): Calcd for $\text{C}_{37}\text{H}_{46}\text{O}_8\text{Na}$ $[\text{M} + \text{Na}]^+$ 641.3085; found 641.3082.

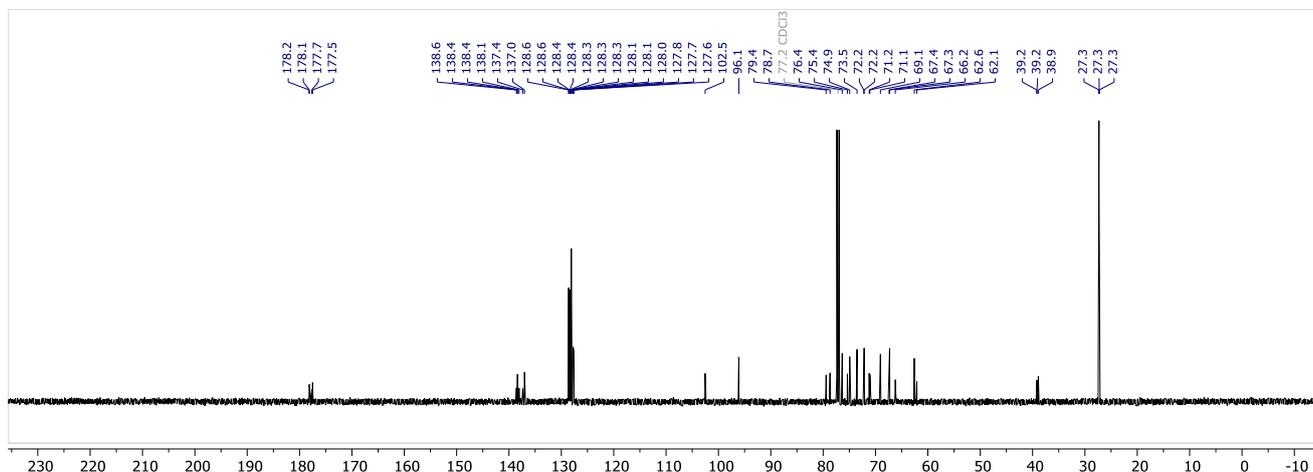
NP-HPLC of **4,6Piv-A4** (ELSD trace, $t_R(\alpha) = 5.0$ min, $t_R(\beta) = 5.6$ min):



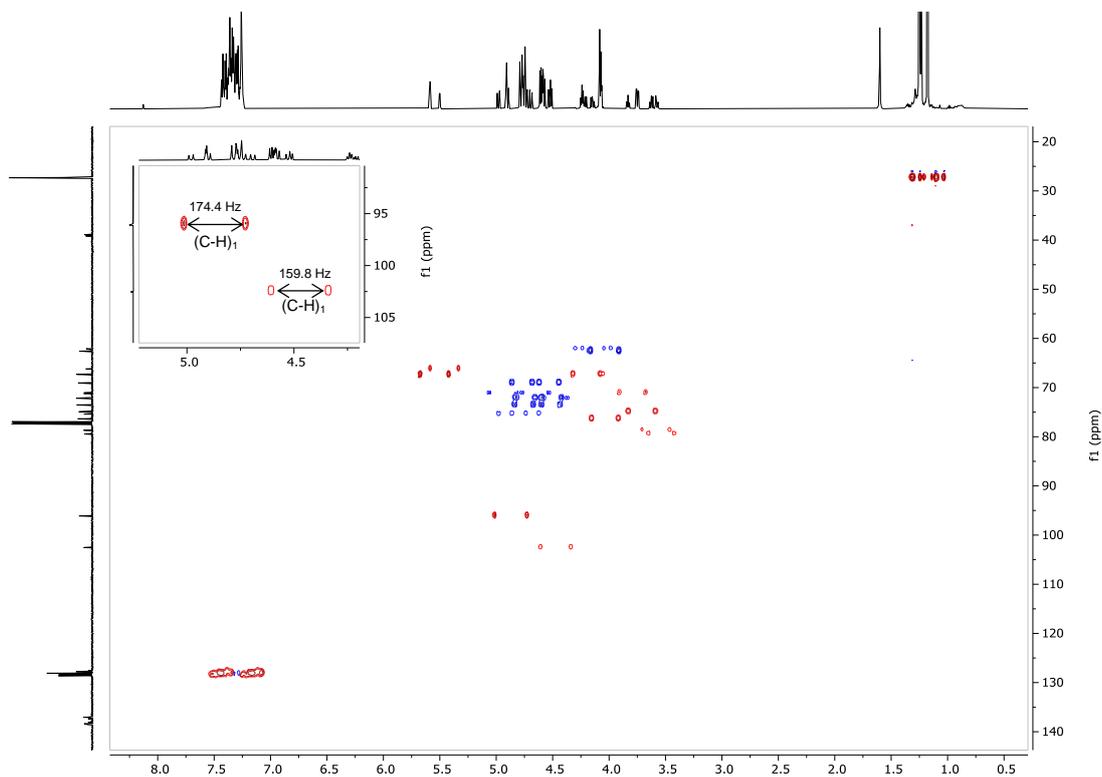
^1H NMR (600 MHz, CDCl_3) of **4,6Piv-A4**:



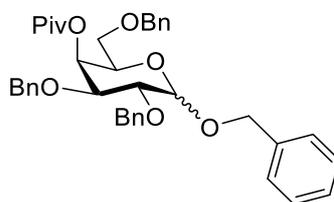
^{13}C NMR (151 MHz, CDCl_3) of **4,6Piv-A4**:



Coupled $^{13}\text{C}, ^1\text{H}$ HSQC of **4,6Piv-A4**:



Benzyl 2,3,6-tris-*O*-benzyl-4-(2,2-dimethylpropanoate)- α/β -D-galactopyranoside (**4Piv-A4**)



The title compound was prepared according to general procedure for glycosylations. Product **4Piv-A4** (16 mg, 26 μ mol, 74%, α/β **68:32**) was obtained as a colorless oil after purification using **Method-2b** ($t_R(\alpha)$ = 6.8 min, $t_R(\beta)$ = 8.1 min).

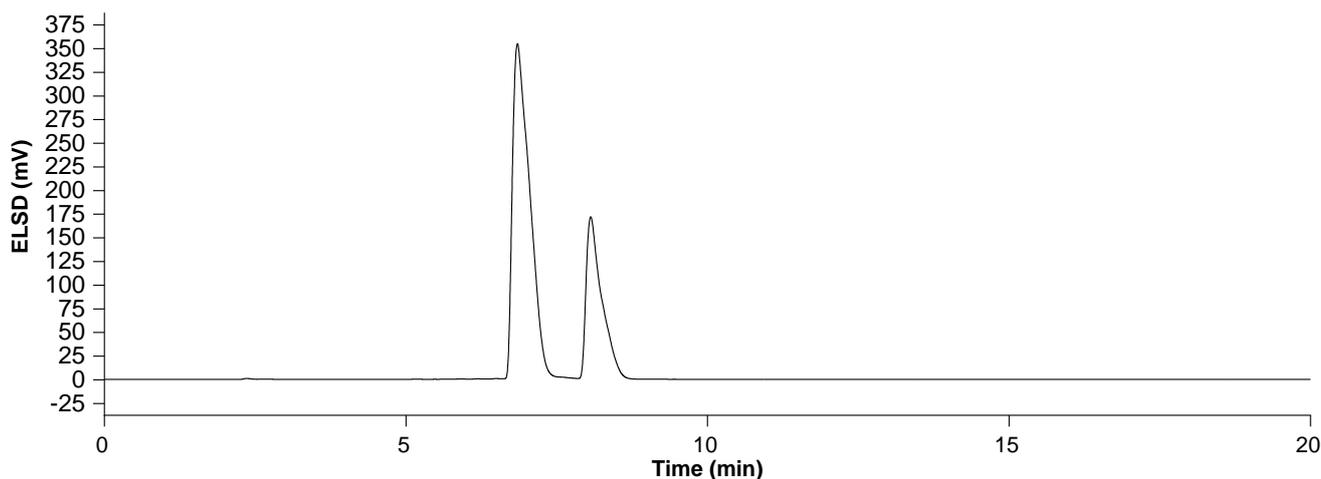
Data of the major isomer (α):

$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.43 – 7.25 (m, 15H), 5.61 (d, J = 3.0 Hz, 1H), 4.99 – 4.83 (m, 2H, H_1), 4.77 – 4.66 (m, 4H), 4.63 – 4.42 (m, 6H), 4.16 (t, J = 6.4 Hz, 1H), 4.03 (dd, J = 10.0, 3.3 Hz, 1H), 3.77 – 3.69 (m, 1H), 3.64 – 3.50 (m, 2H), 3.43 (dd, J = 6.3, 1.2 Hz, 2H), 1.12 (s, 9H) ppm.

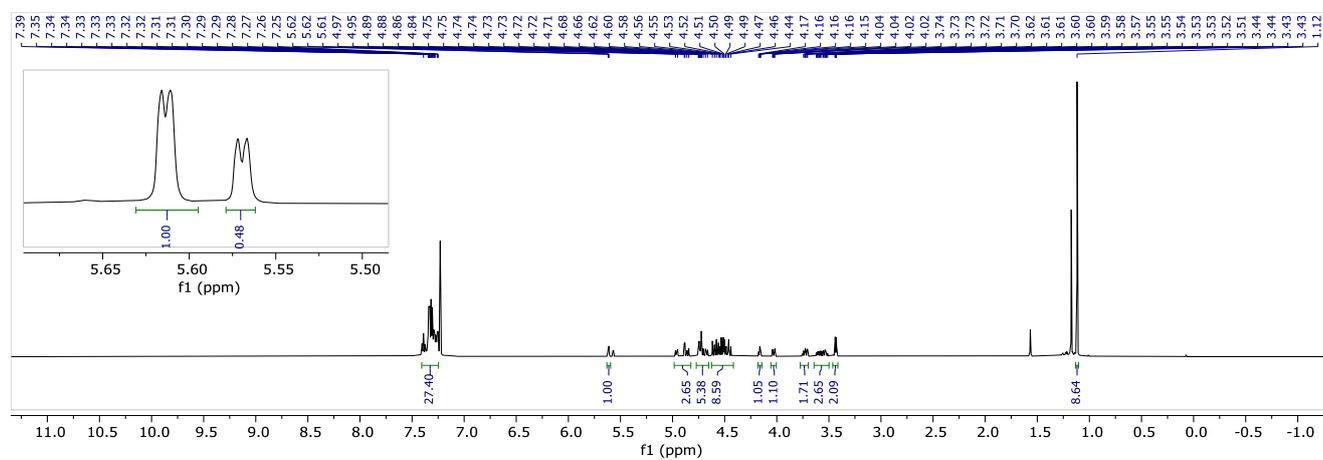
$^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 177.5, 138.5, 138.0, 137.3, 128.6, 128.5, 128.5, 128.3, 128.3, 128.1, 127.9, 127.7, 127.6, 96.4 (C_1), 76.7, 75.1, 73.8, 73.5, 72.0, 69.3, 68.8, 68.3, 67.7, 39.1, 27.3 ppm.

HRMS (QToF): Calcd for $\text{C}_{39}\text{H}_{44}\text{O}_7\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 647.2979; found 647.2978.

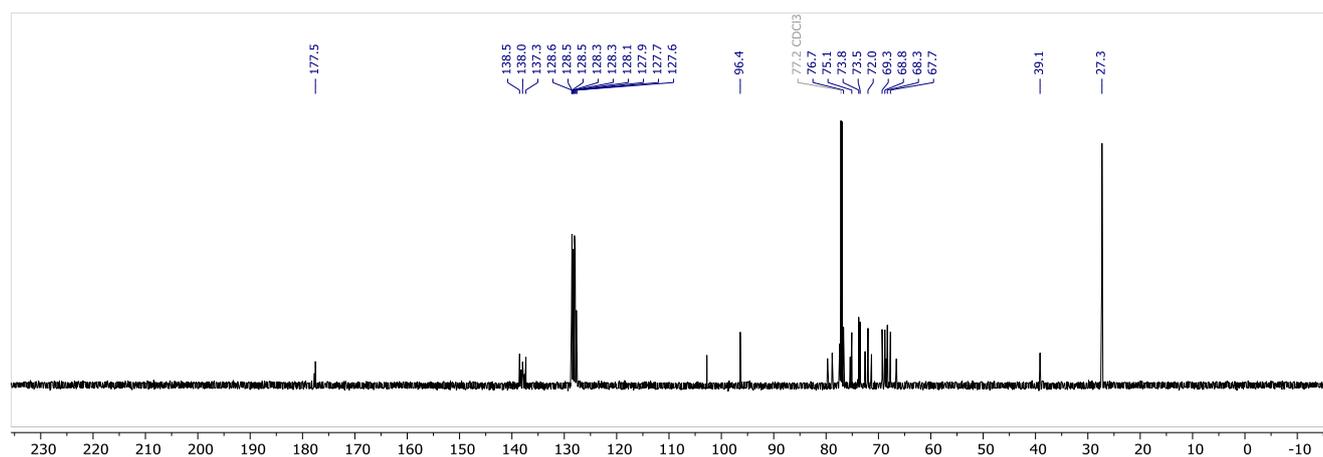
NP-HPLC of **4Piv-A4** (ELSD trace, $t_R(\alpha)$ = 6.8 min, $t_R(\beta)$ = 8.1 min):



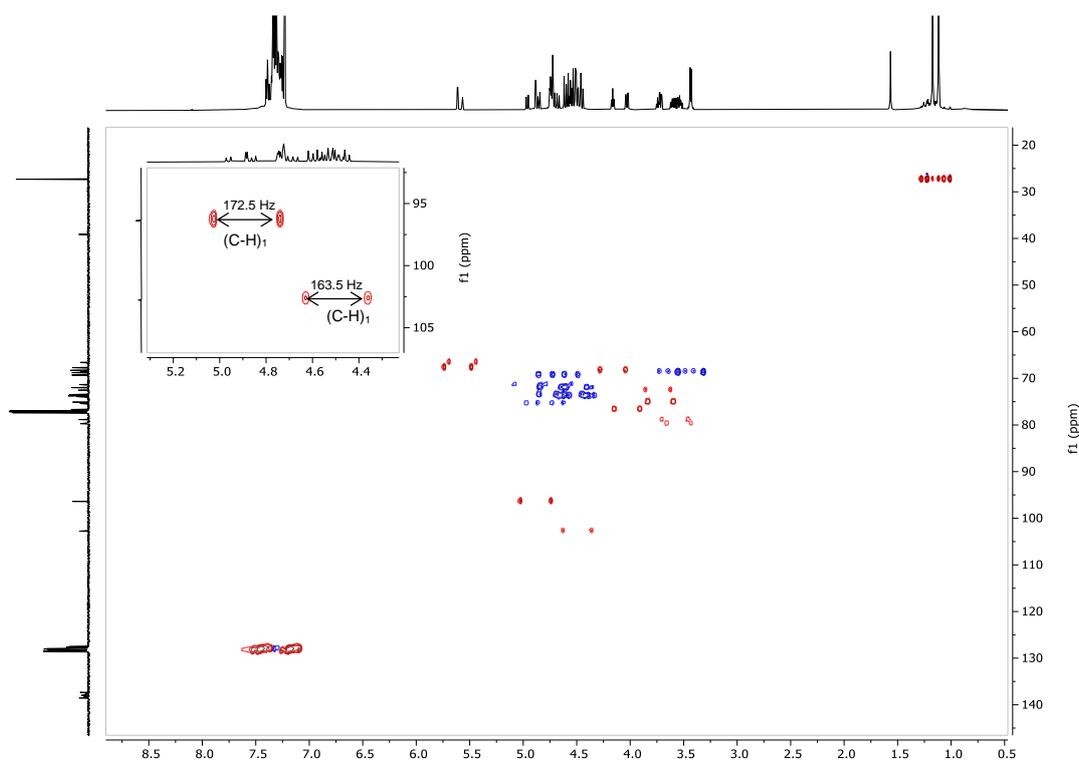
^1H NMR (600 MHz, CDCl_3) of **4Piv-A4**:



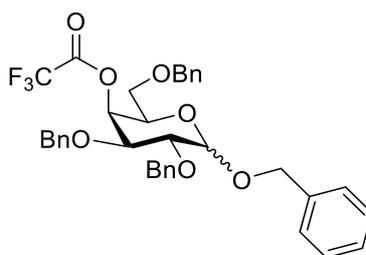
^{13}C NMR (151 MHz, CDCl_3) of **4Piv-A4**:



Coupled $^{13}\text{C}, ^1\text{H}$ HSQC of **4Piv-A4**:



Benzyl 2,3,6-tris-*O*-benzyl-4-(trifluoroacetate)- α/β -D-galactopyranoside (4TFA-A4)



The title compound was prepared according to general procedure for glycosylations. Product **4TFA-A4** (10 mg, 16 μmol , 47%, α/β 54:46) was obtained as a colorless oil after purification using **Method-2b** (t_R (α/β) = 6.1 min).

Data of the anomeric mixture:

^1H NMR (600 MHz, CDCl_3) δ 7.38 – 7.19 (m, 39H), 5.70 (dd, J = 3.4, 1.4 Hz, 1H), 5.66 (q, J = 1.3 Hz, 1H), 4.92 (d, J = 11.9 Hz, 1H), 4.88 – 4.82 (m, 2H), 4.77 – 4.39 (m, 13H), 4.13 (ddd, J = 7.4, 5.7, 1.4 Hz, 1H), 4.06

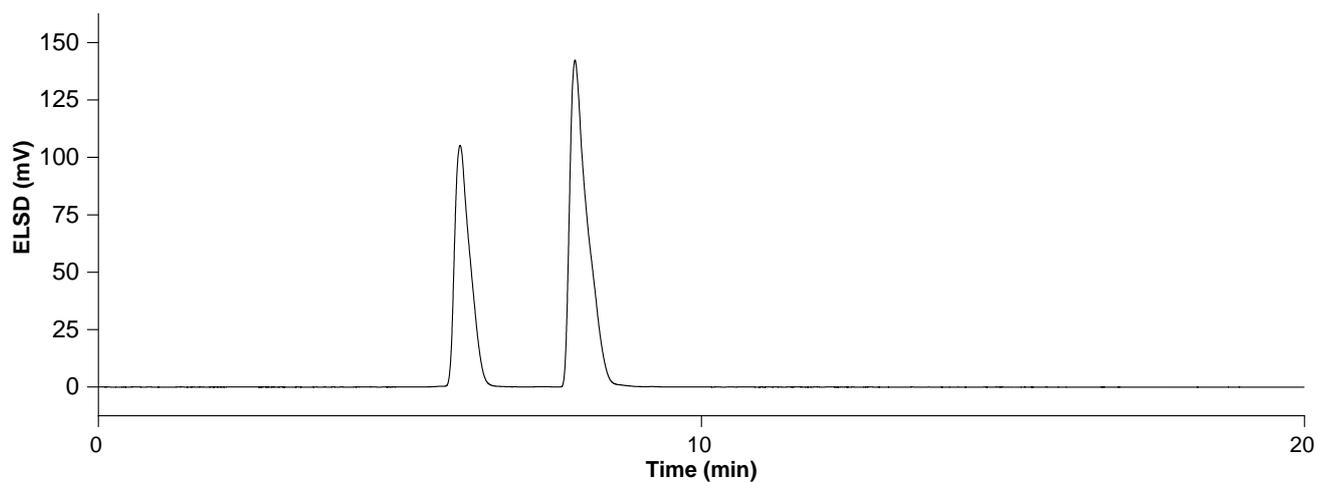
(dd, $J = 10.0, 3.3$ Hz, 1H), 3.76 (ddd, $J = 8.3, 5.5, 1.1$ Hz, 1H), 3.70 (dd, $J = 10.0, 3.7$ Hz, 1H), 3.64 (dd, $J = 9.2, 5.6$ Hz, 1H), 3.60 (dd, $J = 4.5, 1.6$ Hz, 2H), 3.50 (dd, $J = 9.3, 8.2$ Hz, 1H), 3.42 (dd, $J = 9.2, 5.8$ Hz, 1H), 3.37 (dd, $J = 9.3, 7.9$ Hz, 1H) ppm.

^{13}C NMR (151 MHz, CDCl_3) δ 170.2, 170.1, 157.3, 157.2, 157.0, 156.9, 138.4, 138.3, 137.9, 137.5, 137.4, 137.3, 137.2, 137.1, 132.6, 130.1, 128.7, 128.6, 128.6, 128.5, 128.4, 128.3, 128.2, 128.1, 128.0, 127.9, 127.8, 127.8, 102.6, 96.5, 78.7, 75.8, 75.6, 75.2, 74.0, 73.7, 72.9, 72.7, 71.5, 71.2, 69.8, 67.5, 67.2, 67.0 ppm.

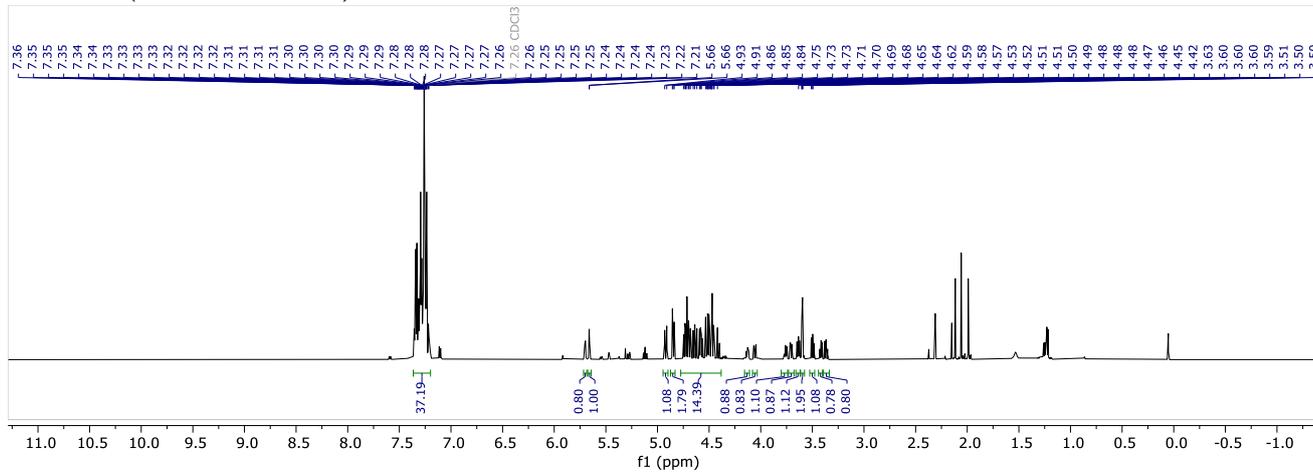
^{19}F NMR (564 MHz, CDCl_3) δ -74.58, -74.80 ppm.

HRMS (QToF): Calcd for $\text{C}_{36}\text{H}_{35}\text{F}_3\text{O}_7\text{Na}$ $[\text{M} + \text{Na}]^+$ 659.2227; found 659.2225.

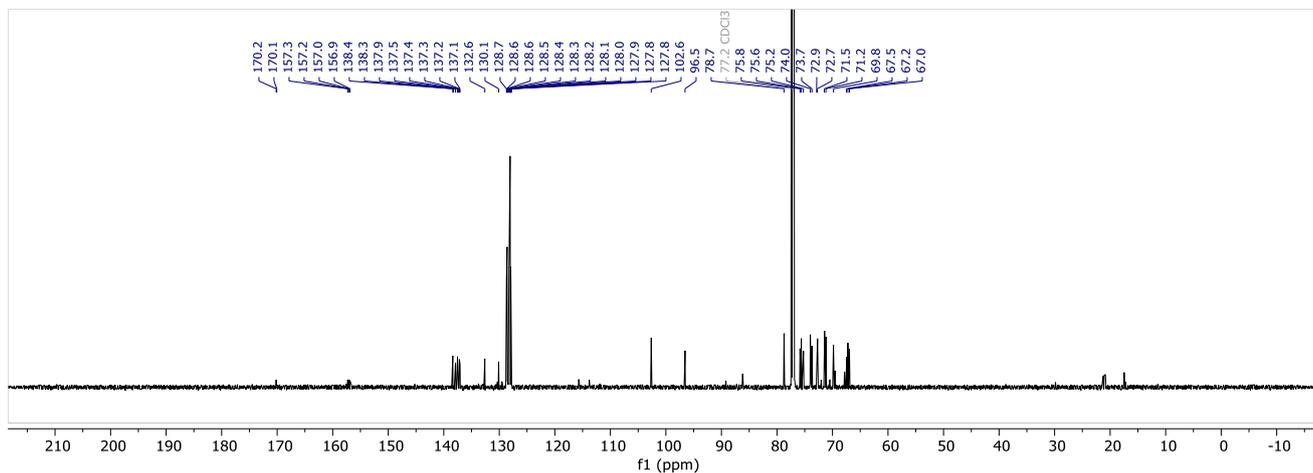
NP-HPLC of **4TFA-A4** (ELSD trace, $t_R(\alpha) = 7.9$ min, $t_R(\beta) = 6.0$ min):



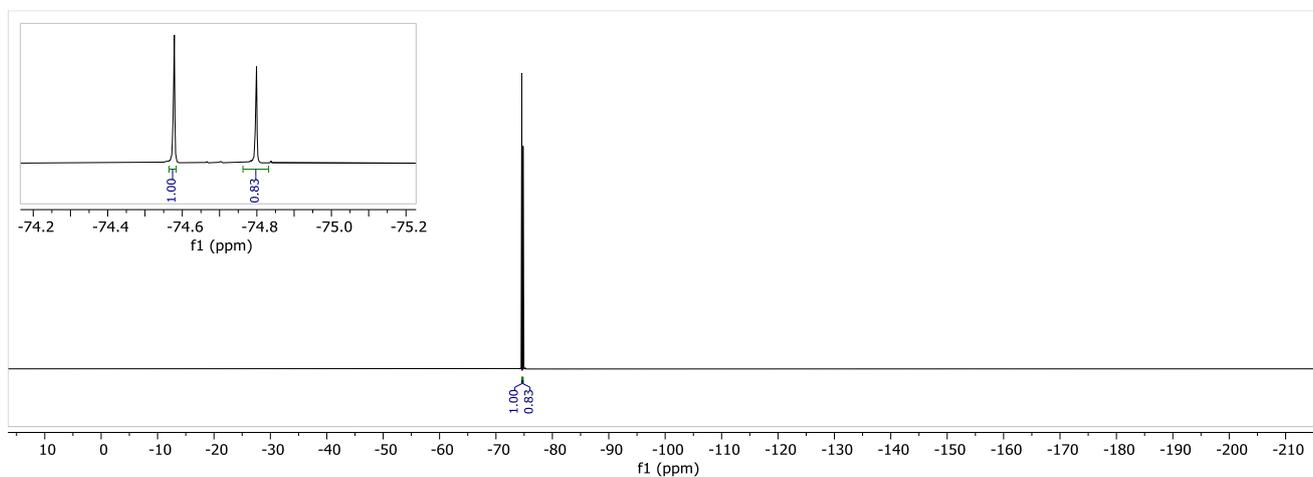
¹H NMR (600 MHz, CDCl₃) of 4TFA-A4:



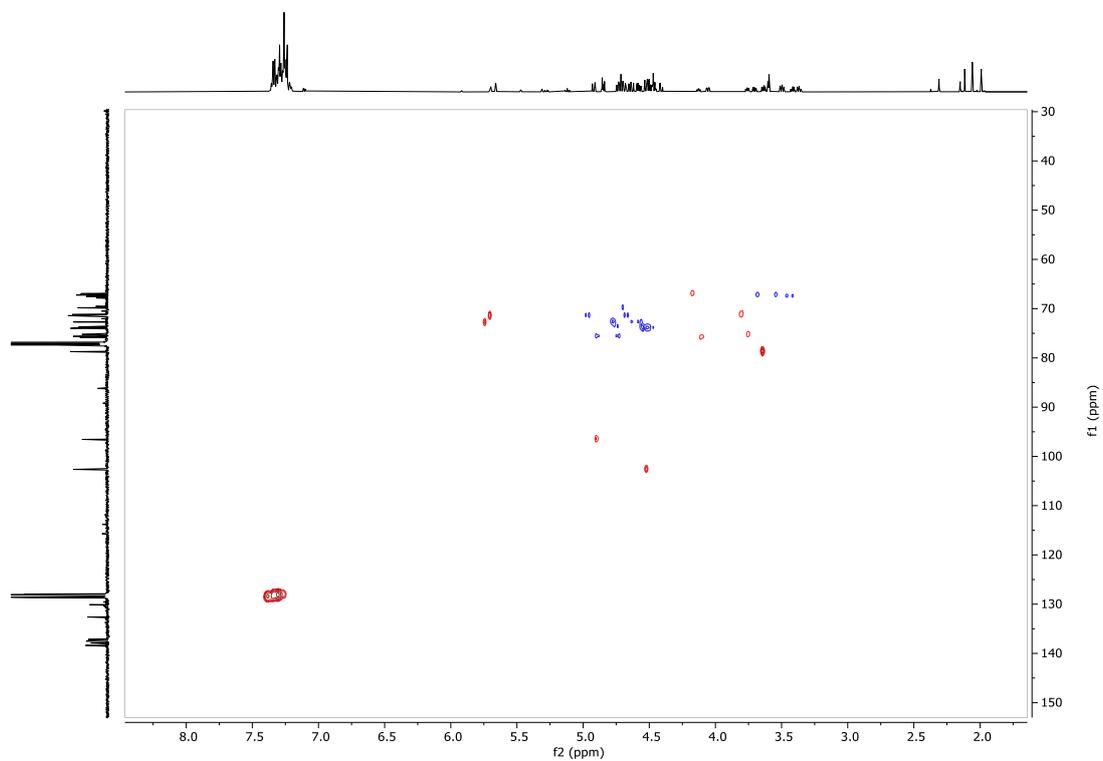
¹³C NMR (151 MHz, CDCl₃) of 4TFA-A4:



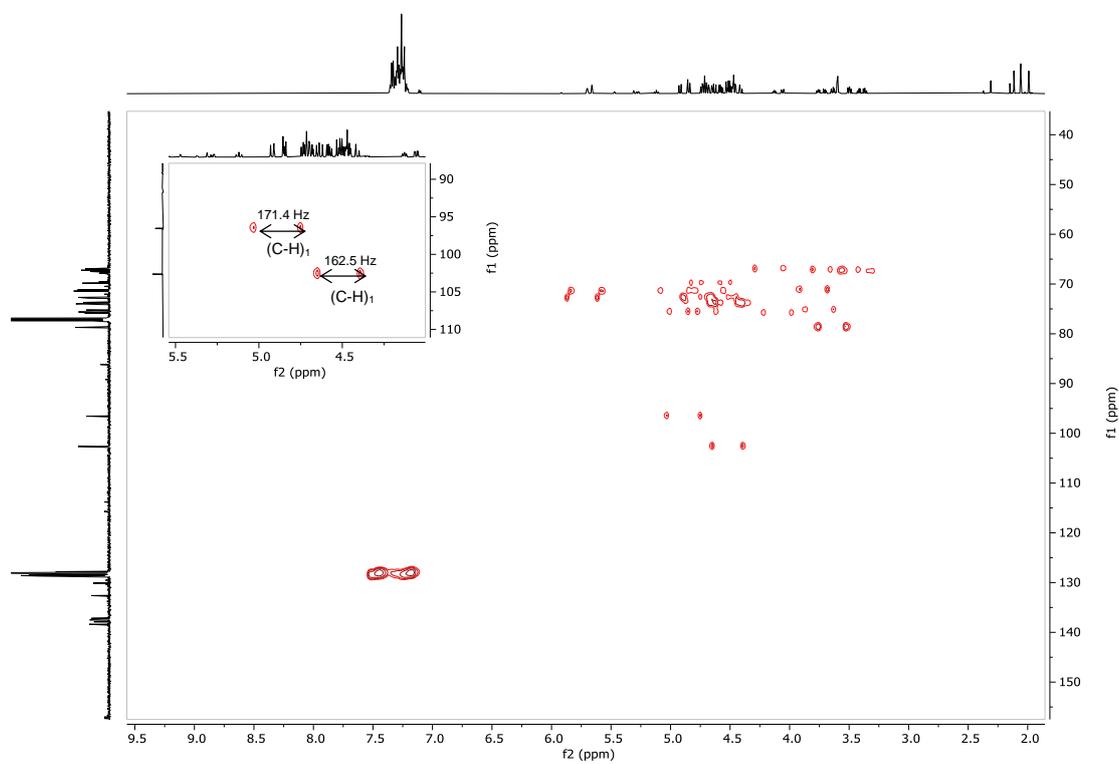
¹⁹F NMR (564 MHz, CDCl₃) of 4TFA-A4:



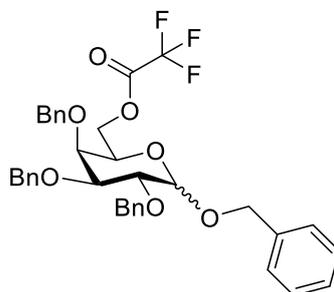
$^{13}\text{C}, ^1\text{H}$ HSQC of **4TFA-A4**:



Coupled $^{13}\text{C}, ^1\text{H}$ HSQC of **4TFA-A4**:



Benzyl 2,3,4-tris-*O*-benzyl-6-(trifluoroacetate)- α/β -D-galactopyranoside (6TFA-A4)



The title compound was prepared according to general procedure for glycosylations. Product **6TFA-A4** (11 mg, 17 μ mol, 50%, α/β 50:50) was obtained as a colorless oil after purification using **Method-2b** ($t_R(\alpha)$ = 5.7 min, $t_R(\beta)$ = 6.0 min).

Data of the anomeric mixture:

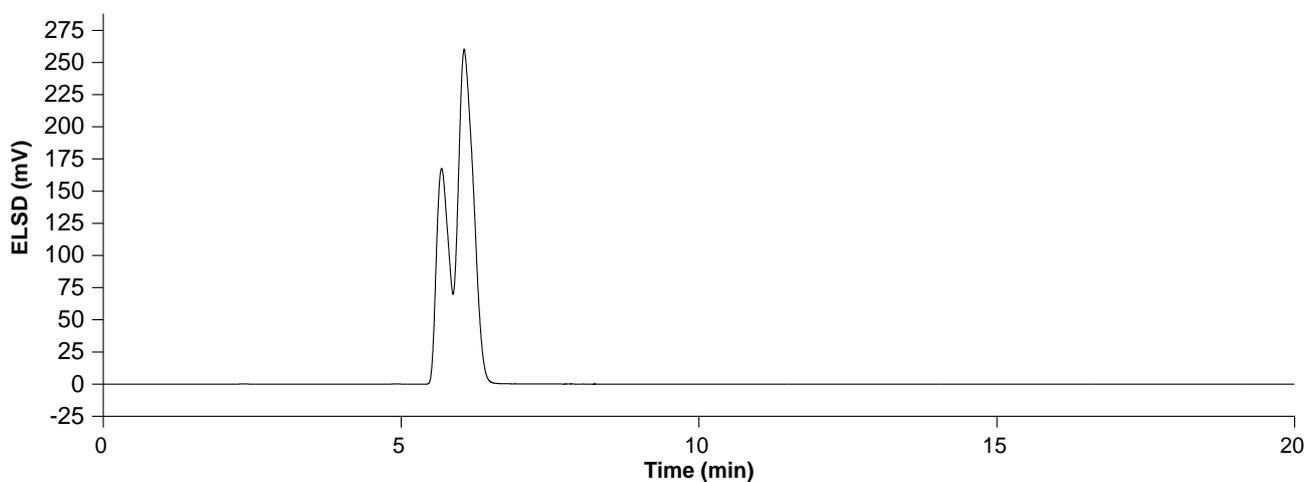
$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.44 – 7.26 (m, 40H), 5.00 (dd, J = 11.7, 6.3 Hz, 2H), 4.98 – 4.89 (m, 4H), 4.84 (d, J = 11.7 Hz, 1H), 4.76 (ddd, J = 20.7, 11.3, 2.1 Hz, 4H), 4.70 – 4.61 (m, 4H), 4.61 – 4.50 (m, 4H), 4.45 (d, J = 7.6 Hz, 1H), 4.41 (dd, J = 11.1, 7.5 Hz, 1H), 4.14 (dd, J = 11.2, 5.1 Hz, 1H), 4.09 – 4.05 (m, 1H), 4.04 – 3.99 (m, 2H), 3.92 (dd, J = 9.7, 7.6 Hz, 1H), 3.86 (dd, J = 2.6, 1.4 Hz, 1H), 3.75 (dd, J = 3.0, 1.3 Hz, 1H), 3.60 (ddd, J = 7.4, 5.1, 1.3 Hz, 1H), 3.53 (dd, J = 9.7, 2.9 Hz, 1H) ppm.

$^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 157.1 (q, J = 42.5 Hz), 138.7, 138.6, 138.5, 138.4, 138.0, 137.9, 137.2, 137.0, 128.7, 128.6, 128.6, 128.6, 128.6, 128.5, 128.5, 128.4, 128.3, 128.2, 128.2, 128.2, 128.1, 128.0, 128.0, 127.9, 127.8, 127.8, 127.8, 127.7, 115.5, 115.5, 113.6, 113.6, 102.4, 95.9, 82.0, 79.4, 79.0, 76.5, 75.4, 74.7, 74.5, 74.0, 73.9, 73.4, 72.9, 71.6, 70.9, 69.1, 68.1, 67.1, 66.8 ppm.

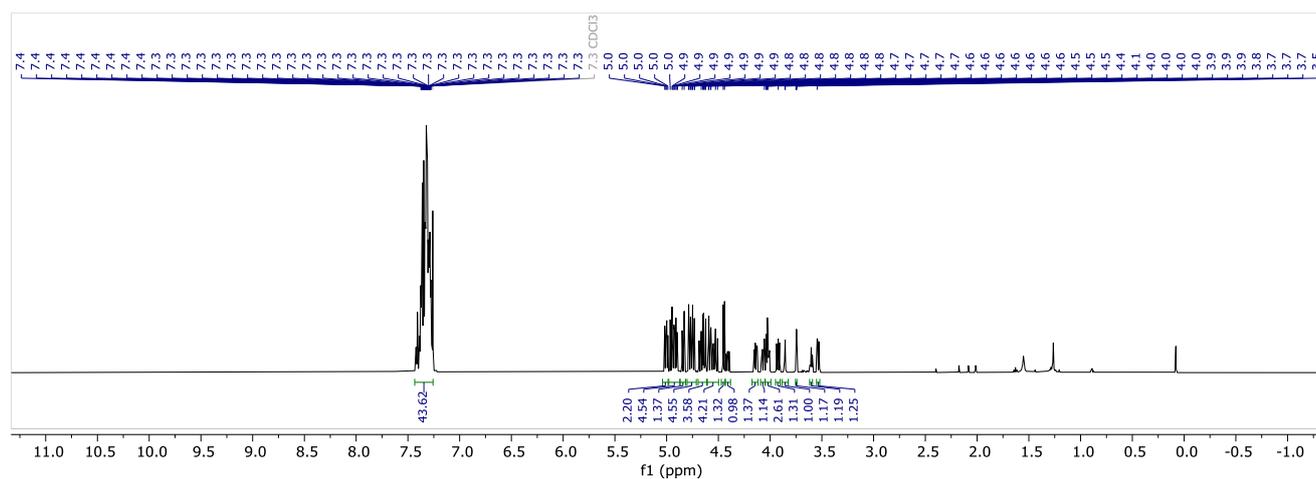
$^{19}\text{F NMR}$ (564 MHz, CDCl_3) δ -74.87, -74.87 ppm.

HRMS (QToF): Calcd for $\text{C}_{36}\text{H}_{35}\text{F}_3\text{O}_7\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 659.2227; found 659.2225.

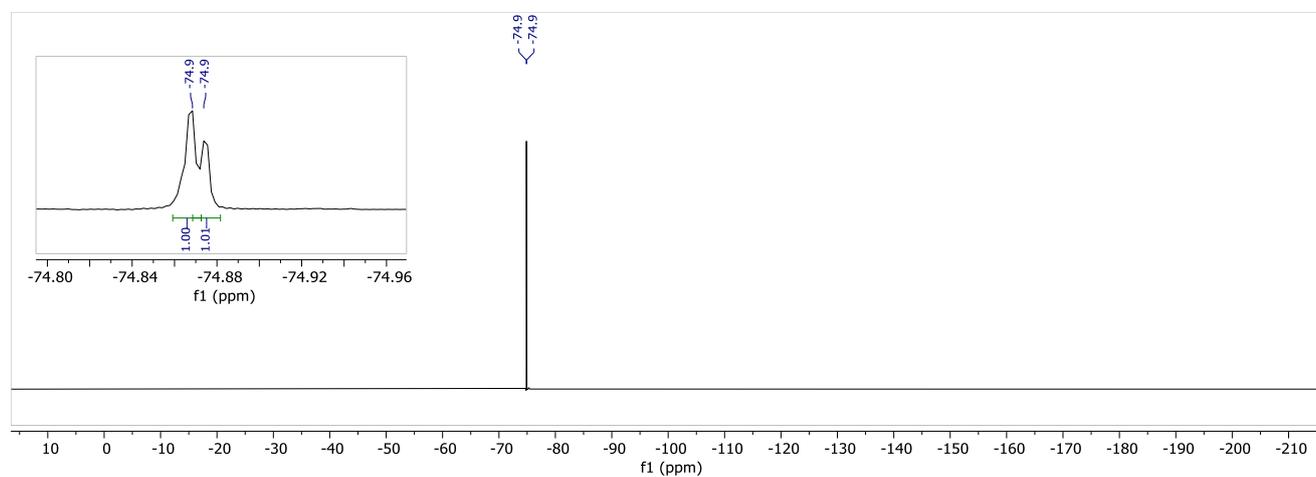
NP-HPLC of **4TFA-A4** (ELSD trace, $t_R(\alpha) = 5.7$ min, $t_R(\beta) = 6.0$ min):



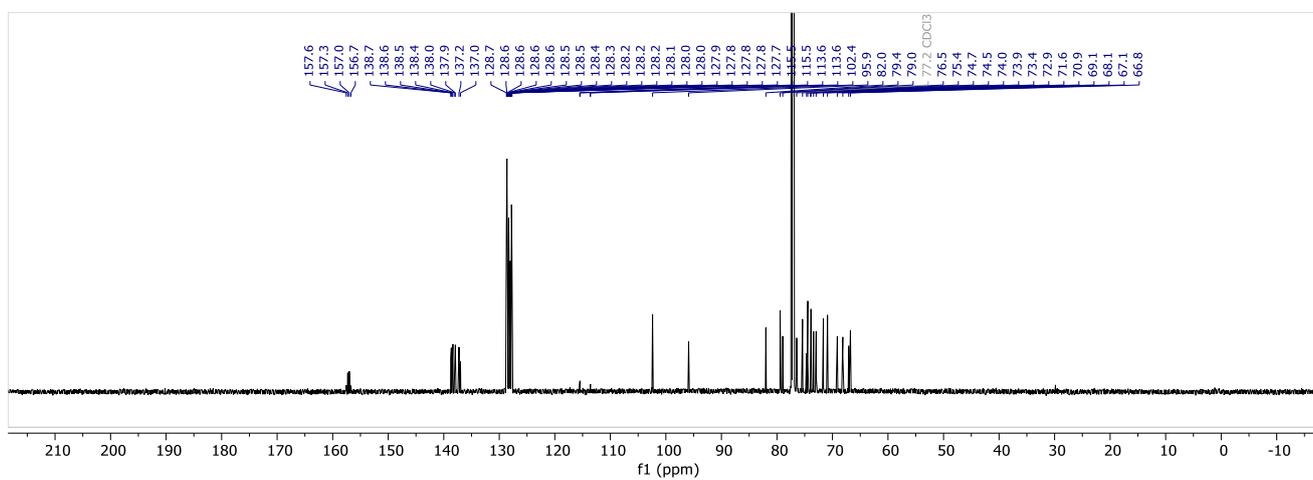
^1H NMR (600 MHz, CDCl_3) of **6TFA-A4**:



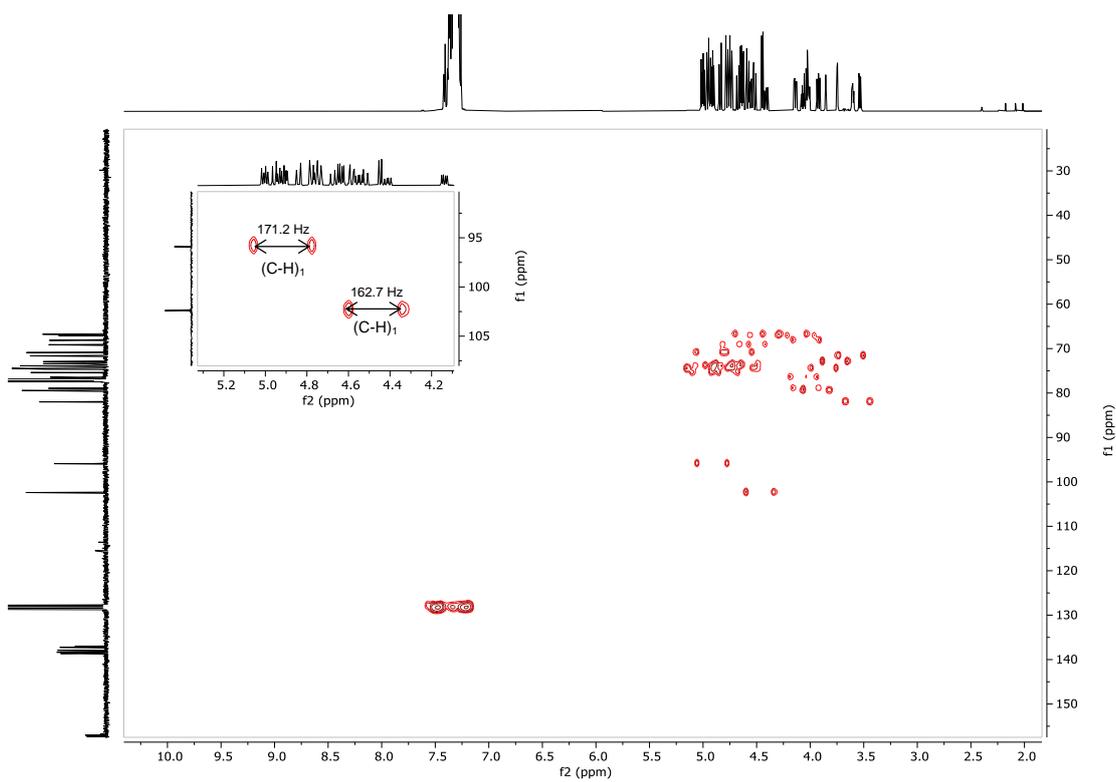
^{19}F NMR (564 MHz, CDCl_3) of **4TFA-A4**:



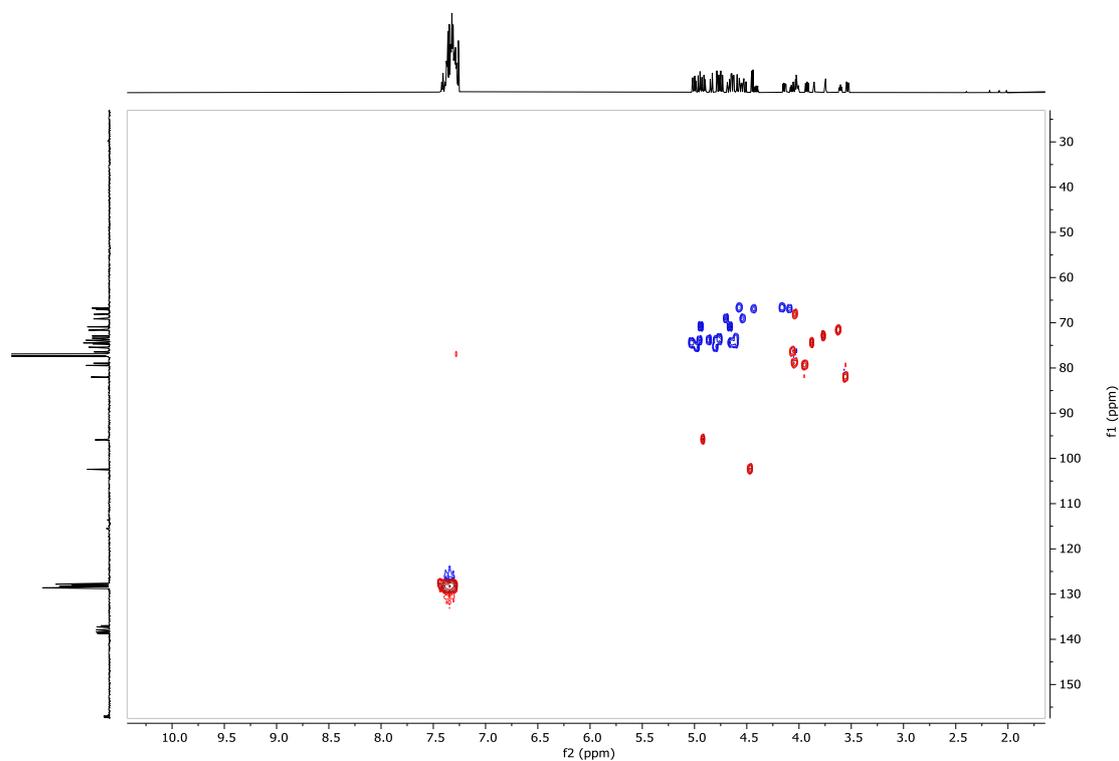
^{13}C NMR (151 MHz, CDCl_3) of **4TFA-A4**:



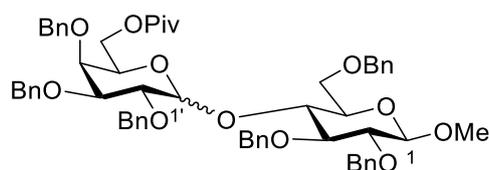
Coupled $^{13}\text{C},^1\text{H}$ HSQC of **6TFA-A4**:



$^{13}\text{C}, ^1\text{H}$ HSQC of **6TFA-A4**:



Methyl *O*-[2,3,4-tris-*O*-benzyl-6-(2,2-dimethylpropanoate)- α/β -D-galactopyranosyl]-(1 \rightarrow 4)-2,3,6-tri-*O*-benzyl- β -D-glucopyranoside (**6Piv-A2**)



The title compound was prepared according to general procedure for glycosylations. Product **6Piv-A2** (28 mg, 28 μmol , 81%, α/β 74:26) was obtained as a colorless oil after purification using **Method-2b** ($t_{\text{R}}(\alpha) = 10.7$ min, $t_{\text{R}}(\beta) = 10.7$ min).

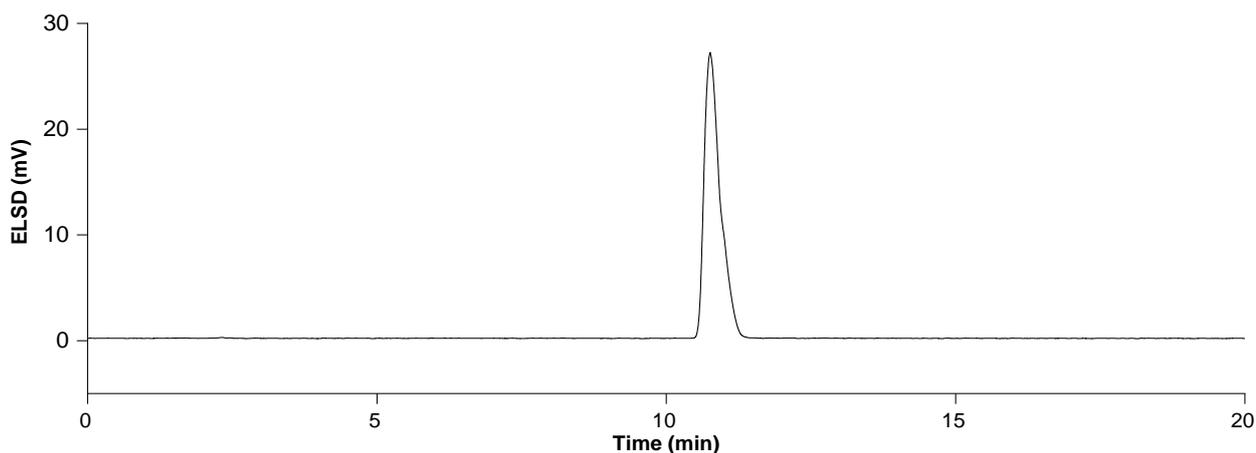
Data of the major isomer (α):

^1H NMR (700 MHz, CDCl_3) δ 7.38 – 7.27 (m, 25H), 7.20 – 7.16 (m, 5H), 5.71 (d, $J = 3.8$ Hz, 1H, H_1'), 4.94 – 4.86 (m, 3H), 4.81 – 4.72 (m, 3H), 4.69 – 4.54 (m, 6H), 4.50 (d, $J = 12.3$ Hz, 1H), 4.36 (d, $J = 7.7$ Hz, 1H, H_1), 4.18 (dd, $J = 11.1, 7.2$ Hz, 1H), 4.00 (d, $J = 3.8$ Hz, 2H), 3.90 – 3.66 (m, 6H), 3.63 (ddd, $J = 9.8, 5.9, 2.4$ Hz, 1H), 3.60 (s, 3H), 3.45 (dd, $J = 9.1, 7.7$ Hz, 1H), 1.21 (s, 9H) ppm.

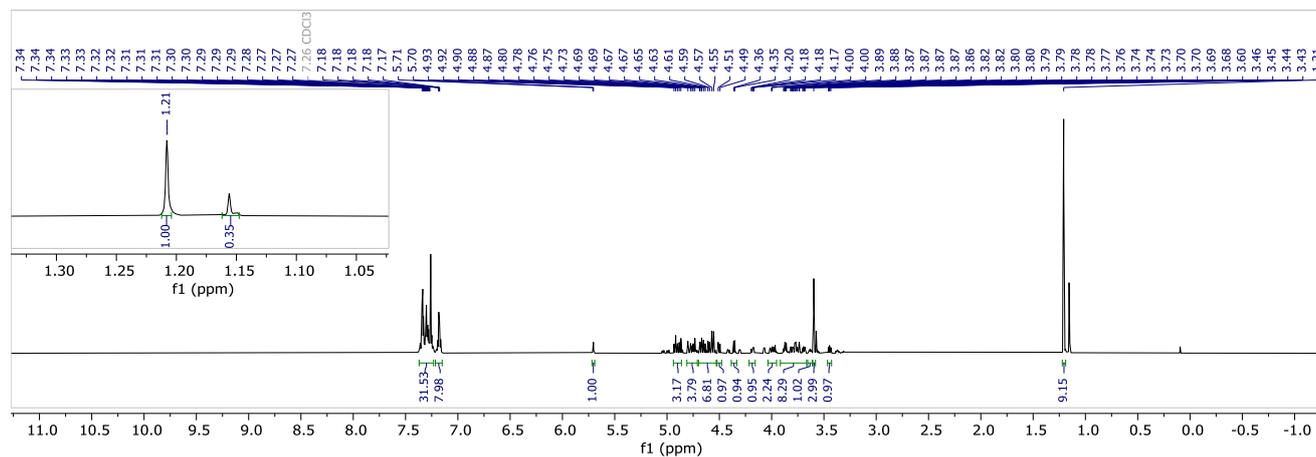
^{13}C NMR (176 MHz, CDCl_3) δ 178.1, 138.8, 138.7, 138.5, 138.4, 138.3, 138.3, 128.5, 128.5, 128.5, 128.4, 128.4, 128.3, 128.2, 128.0, 127.9, 127.7, 127.6, 127.6, 126.8, 104.6 (C_1), 97.2 (C_1'), 84.8, 82.5, 78.9, 75.7, 75.1, 74.8, 74.7, 74.6, 74.1, 73.8, 73.4, 73.3, 70.2, 69.1, 63.5, 57.1, 38.9, 27.5 ppm.

HRMS (QToF): Calcd for $\text{C}_{60}\text{H}_{68}\text{O}_{12}\text{Na}$ $[\text{M} + \text{Na}]^+$ 1003.4603; found 1003.4612.

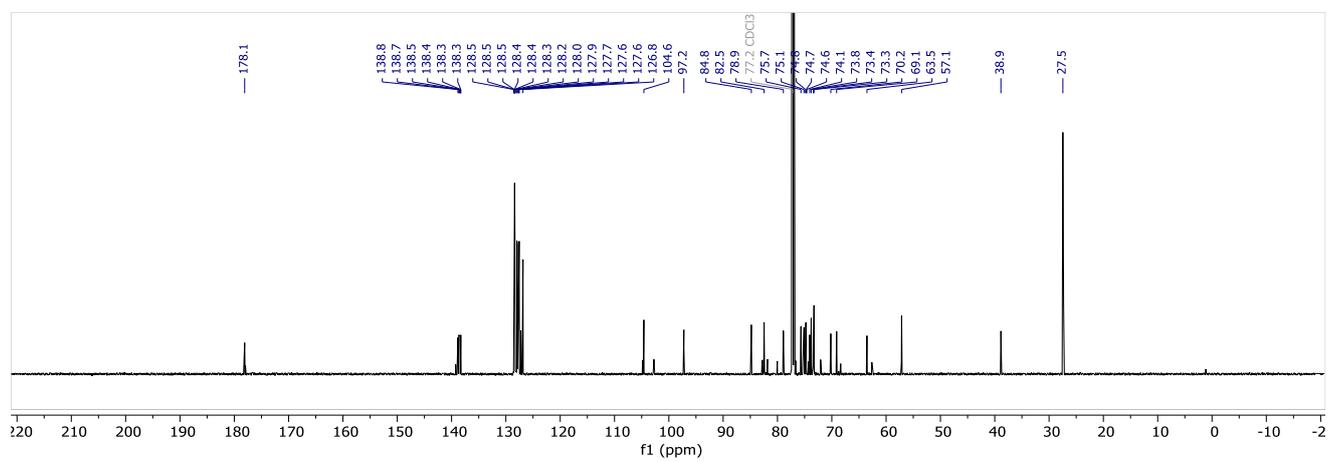
NP-HPLC of **6Piv-A2** (ELSD trace, $t_R(\alpha) = 10.7$ min, $t_R(\beta) = 10.7$ min):



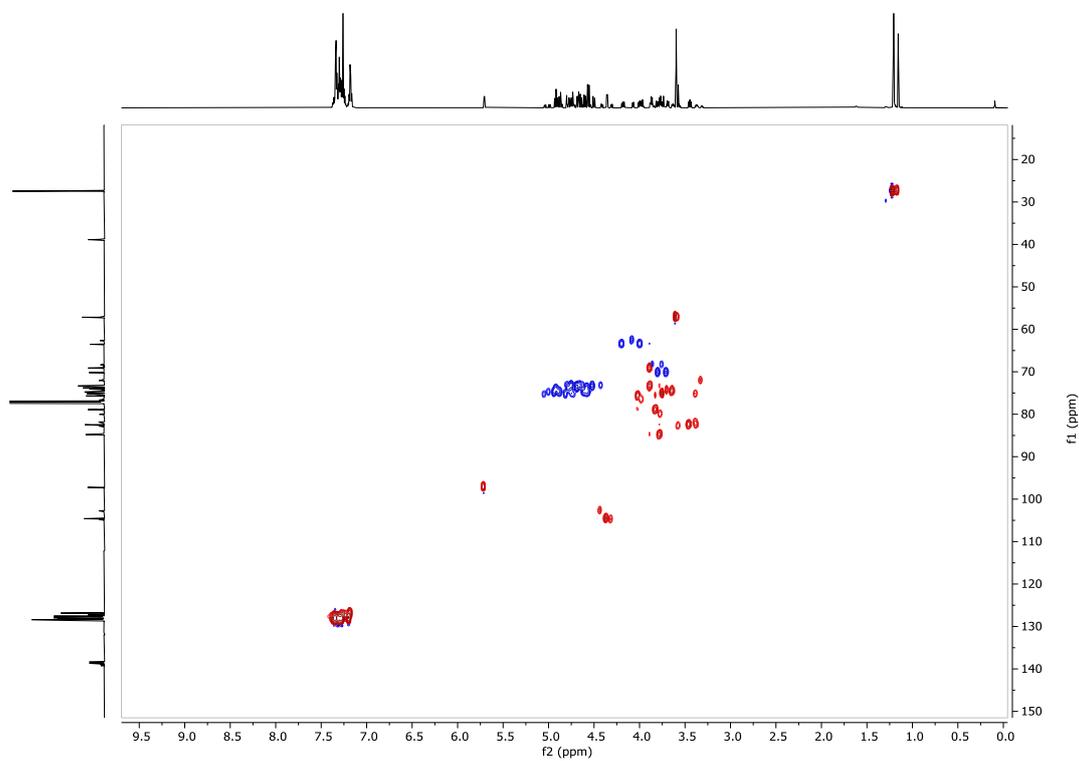
^1H NMR (700 MHz, CDCl_3) of **6Piv-A2**:



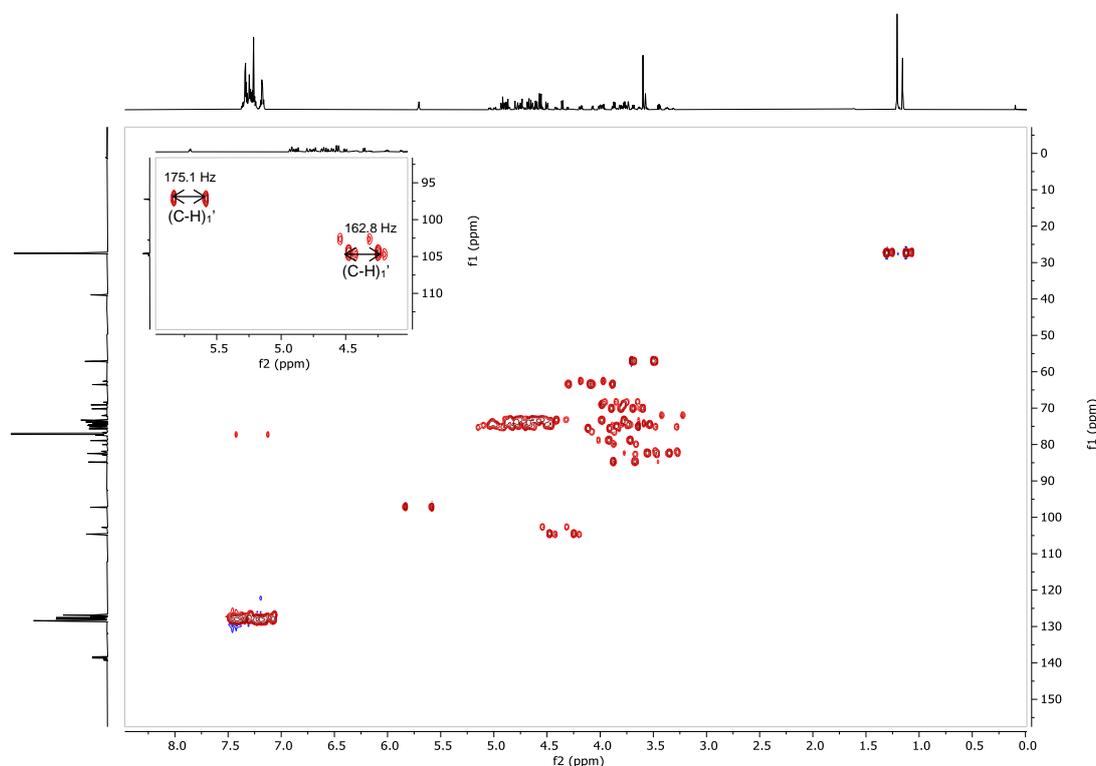
^{13}C NMR (176 MHz, CDCl_3) of **6Piv-A2**:



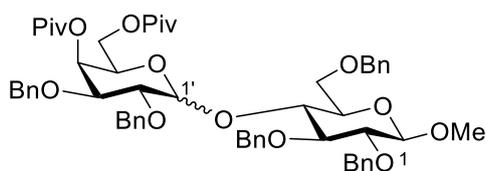
$^{13}\text{C},^1\text{H}$ HSQC of **6Piv-A2**:



Coupled ^{13}C , ^1H HSQC of **6Piv-A2**:



Methyl *O*-[2,3-bis-*O*-benzyl-4,6-bis-(2,2-dimethylpropanoate)- α/β -D-galactopyranosyl]-(1 \rightarrow 4)-2,3,6-tri-*O*-benzyl- β -D-glucopyranoside (4,6Piv-A2)



The title compound was prepared according to general procedure for glycosylations. Product **4,6Piv-A2** (26 mg, 27 μmol , 76%, α/β 89:11) was obtained as a colorless oil after purification using **Method-2b** ($t_{\text{R}}(\alpha) = 9.6$ min, $t_{\text{R}}(\beta) = 9.6$ min).

Data of the major isomer (α):

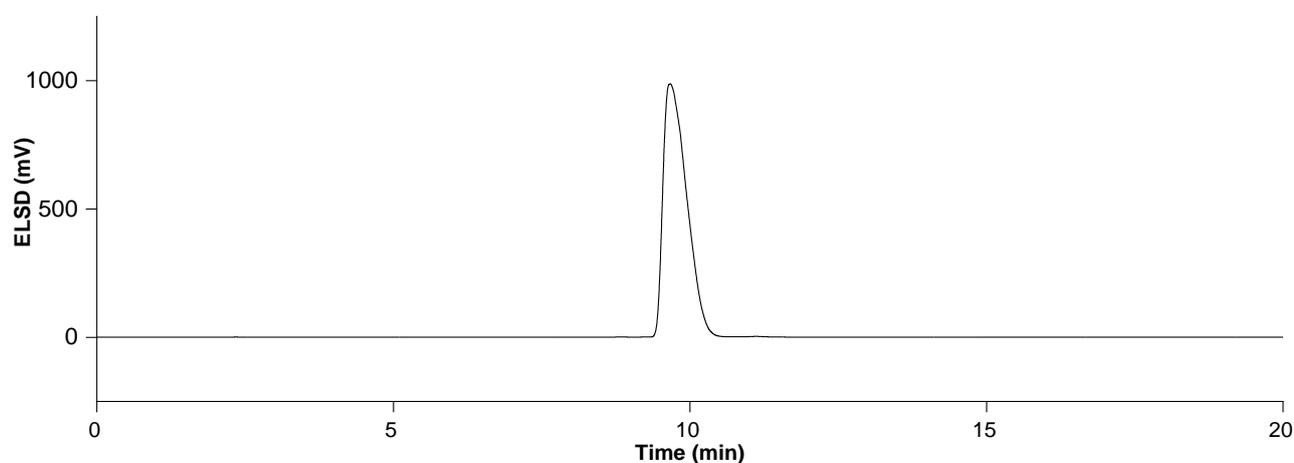
^1H NMR (700 MHz, CDCl_3) δ 7.43 – 7.17 (m, 20H), 7.16 – 7.01 (m, 5H), 5.65 (d, $J = 3.7$ Hz, 1H, $\text{H}_{1'}$), 5.39 (dd, $J = 3.3, 1.7$ Hz, 1H), 4.95 – 4.82 (m, 2H), 4.79 (d, $J = 11.7$ Hz, 1H), 4.67 (d, $J = 12.3$ Hz, 1H), 4.61 (dd, $J = 11.1, 4.9$ Hz, 2H), 4.57 (d, $J = 11.0$ Hz, 1H), 4.51 (t, $J = 11.7$ Hz, 2H), 4.38 – 4.31 (m, 2H, H_4), 4.07 (td, $J =$

6.9, 1.8 Hz, 1H), 3.95 – 3.92 (m, 2H), 3.80 (dd, $J = 10.3, 3.3$ Hz, 1H), 3.77 – 3.72 (m, 2H), 3.69 (dd, $J = 10.7, 5.1$ Hz, 1H), 3.65 – 3.59 (m, 1H), 3.57 (s, 3H), 3.46 – 3.39 (m, 1H), 1.20 (s, 9H), 1.10 (s, 9H) ppm.

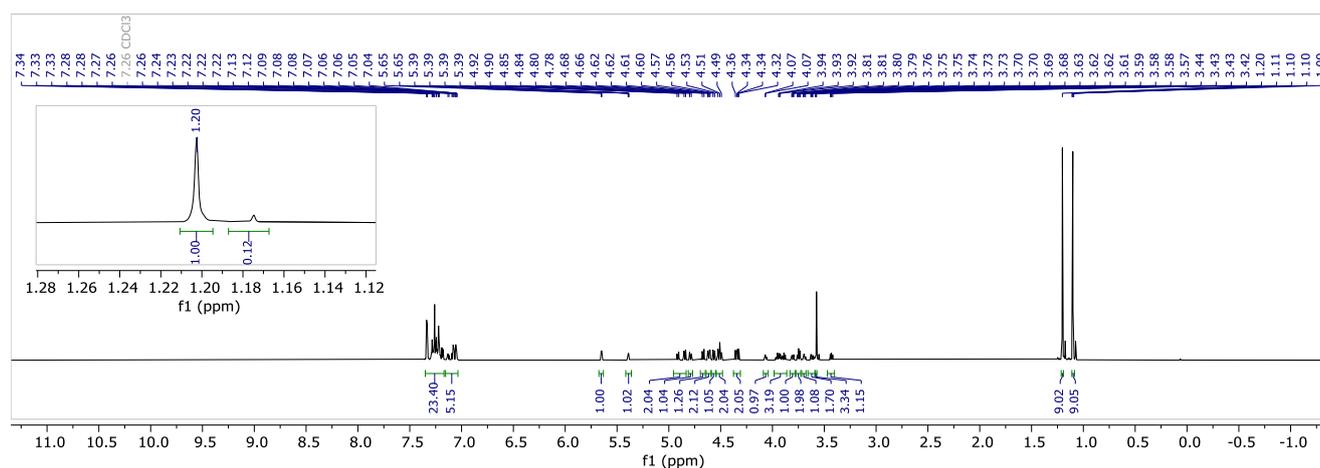
^{13}C NMR (176 MHz, CDCl_3) δ 178.0, 177.5, 139.0, 138.5, 138.2, 138.2, 138.0, 128.6, 128.4, 128.4, 128.3, 128.2, 128.2, 128.1, 128.0, 127.9, 127.7, 127.6, 127.6, 127.2, 126.7, 104.7 (C_1), 97.3 (C_1'), 84.8, 82.5, 76.1, 74.8, 74.4, 74.3, 74.0, 73.8, 73.5, 73.3, 72.0, 69.8, 67.3, 66.9, 62.3, 57.2, 39.1, 38.9, 27.4, 27.3 ppm.

HRMS (QToF): Calcd for $\text{C}_{58}\text{H}_{70}\text{O}_{13}\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 997.4709; found 997.4714.

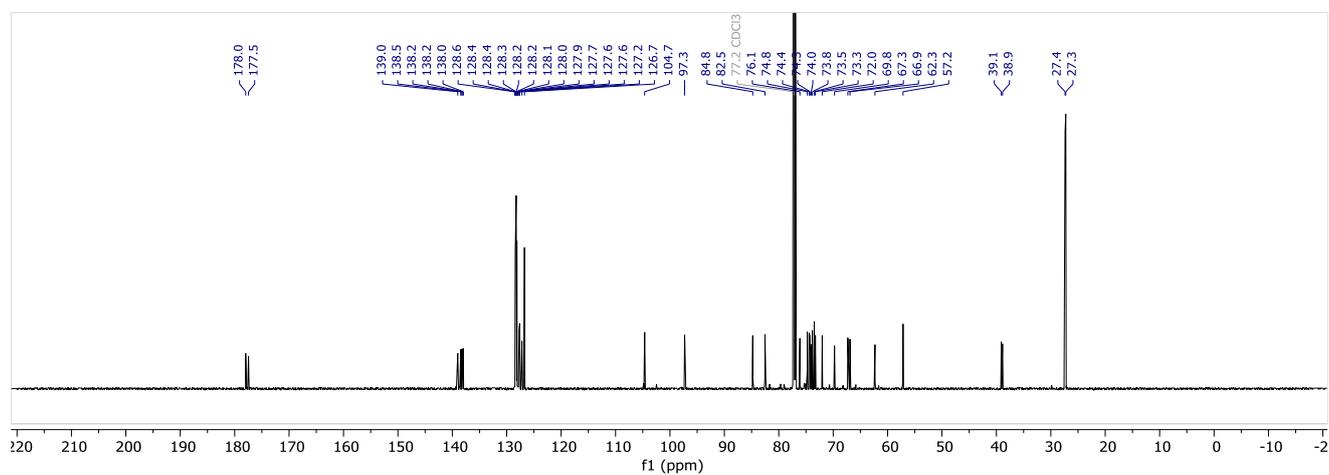
NP-HPLC of **4,6Piv-A2** (ELSD trace, $t_R(\alpha) = 9.6$ min, $t_R(\beta) = 9.6$ min):



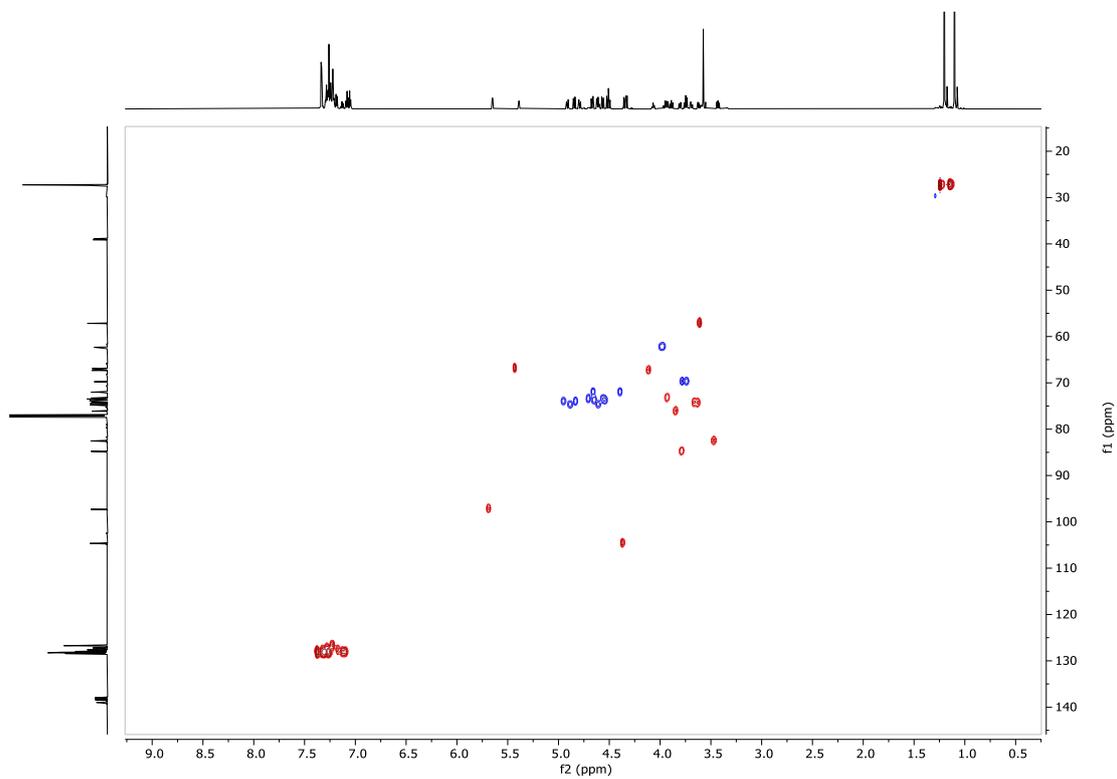
^1H NMR (700 MHz, CDCl_3) of **4,6Piv-A2**:



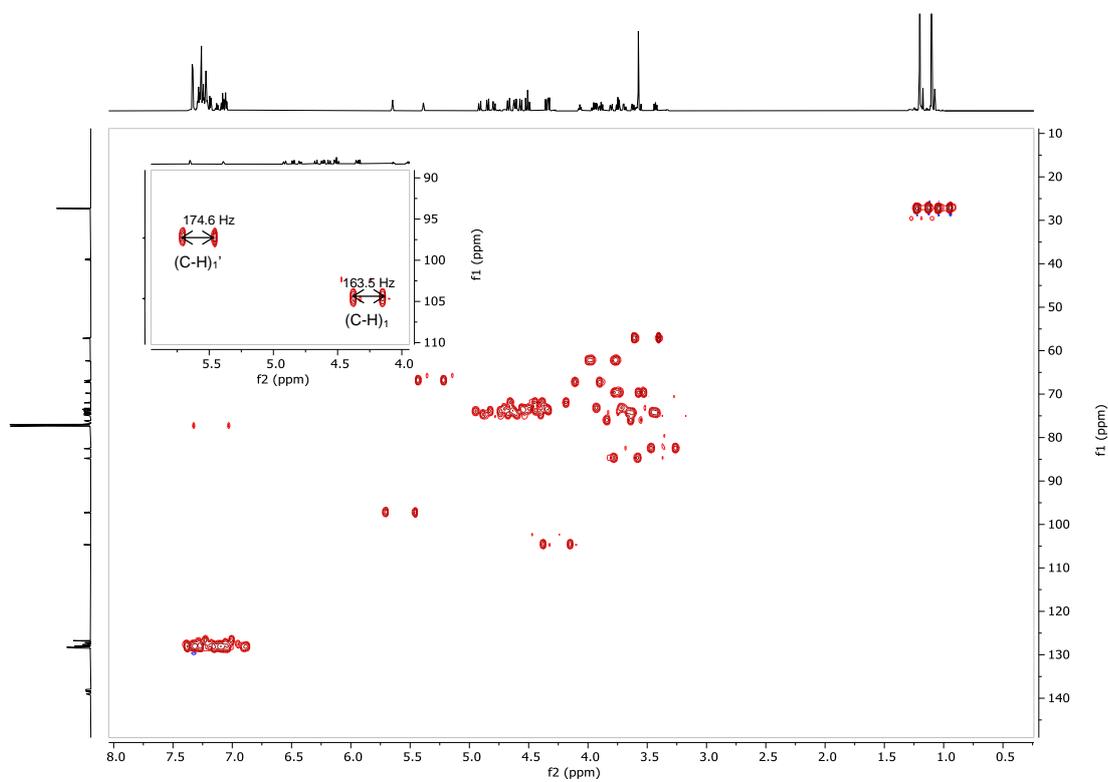
^{13}C NMR (176 MHz, CDCl_3) of **4,6Piv-A2**:



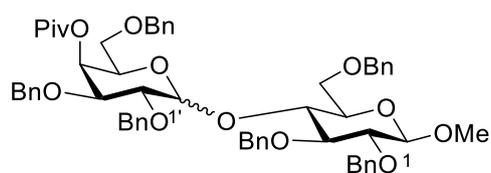
$^{13}\text{C}, ^1\text{H}$ HSQC of **4,6Piv-A2**:



Coupled ^{13}C , ^1H HSQC of **4,6Piv-A2**:



Methyl *O*-[2,3,6-tris-*O*-benzyl-4-(2,2-dimethylpropanoate)- α/β -D-galactopyranosyl]-(1 \rightarrow 4)-2,3,6-tri-*O*-benzyl- β -D-glucopyranoside (**4Piv-A2**)



The title compound was prepared according to general procedure for glycosylations. Product **4Piv-A2** (29 mg, 30 μmol , 87%, α/β 96:4) was obtained as a colorless oil after purification using **Method-2b** ($t_{\text{R}}(\alpha) = 11.9$ min, $t_{\text{R}}(\beta) = 11.9$ min).

Data of the major isomer (α):

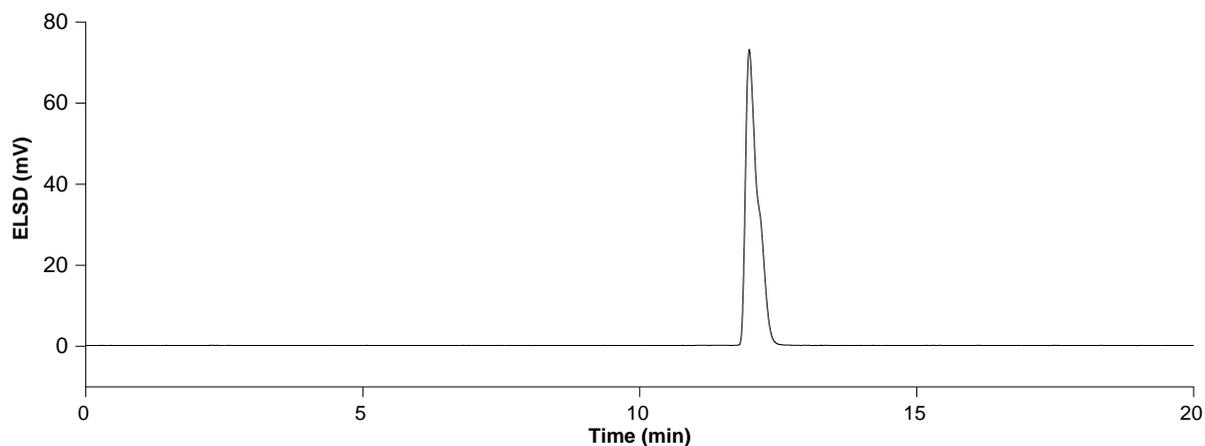
^1H NMR (700 MHz, CDCl_3) δ 7.38 – 7.15 (m, 25H), 7.14 – 7.01 (m, 5H), 5.68 (d, $J = 3.9$ Hz, 1H, $\mathbf{H}_{1'}$), 5.52 (dd, $J = 3.3, 1.5$ Hz, 1H), 4.92 (d, $J = 11.8$ Hz, 1H), 4.83 (dd, $J = 21.5, 11.3$ Hz, 2H), 4.66 – 4.59 (m, 3H), 4.57 (d, $J = 11.0$ Hz, 1H), 4.48 (dd, $J = 12.0, 5.9$ Hz, 2H), 4.41 (d, $J = 11.7$ Hz, 1H), 4.33 (d, $J = 2.7$ Hz, 1H), 4.32 (s, 1H, \mathbf{H}_1), 4.29 (d, $J = 11.7$ Hz, 1H), 4.03 (t, $J = 6.5$ Hz, 1H), 3.96 (t, $J = 9.1$ Hz, 1H), 3.80 (dd, $J = 10.2, 3.2$

Hz, 1H), 3.79 – 3.74 (m, 2H), 3.71 (dd, $J = 10.8, 5.0$ Hz, 1H), 3.63 (dd, $J = 10.2, 3.9$ Hz, 1H), 3.59 (td, $J = 4.8, 2.5$ Hz, 1H), 3.57 (s, 3H), 3.45 (dd, $J = 9.1, 7.6$ Hz, 1H), 3.41 – 3.35 (m, 1H), 3.31 (dd, $J = 9.1, 7.0$ Hz, 1H), 1.07 (s, 9H) ppm.

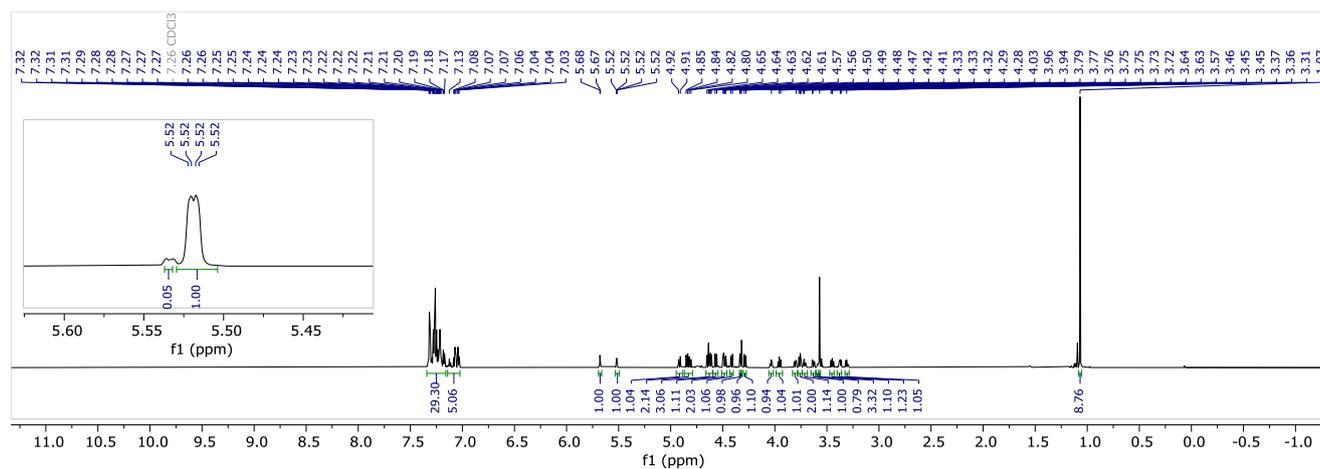
^{13}C NMR (176 MHz, CDCl_3) δ 177.5, 139.2, 138.5, 138.5, 138.3, 138.1, 138.0, 128.5, 128.5, 128.4, 128.3, 128.3, 128.2, 128.1, 128.1, 127.9, 127.9, 127.7, 127.7, 127.6, 127.5, 127.1, 126.7, 104.6 (C_1), 97.7 (C_1'), 84.9, 82.6, 76.6, 74.8, 74.5, 74.4, 74.0, 73.9, 73.8, 73.4, 73.3, 71.8, 69.8, 68.5, 67.3, 57.1, 39.1, 27.3 ppm.

HRMS (QToF): Calcd for $\text{C}_{60}\text{H}_{68}\text{O}_{12}\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 1003.4603; found 1003.4614.

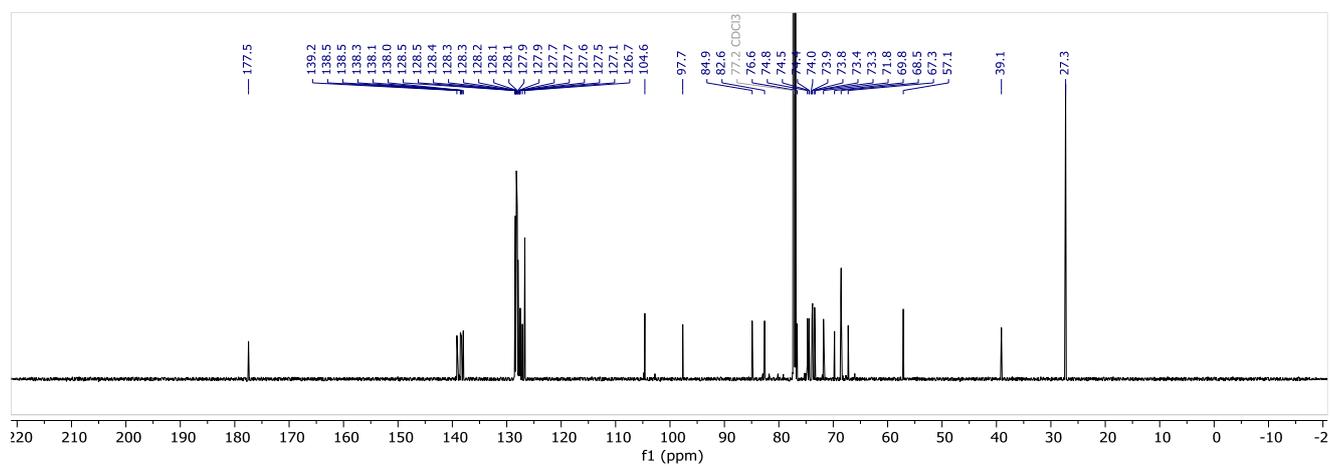
NP-HPLC of **4Piv-A2** (ELSD trace, $t_R(\alpha) = 11.9$ min, $t_R(\beta) = 11.9$ min):



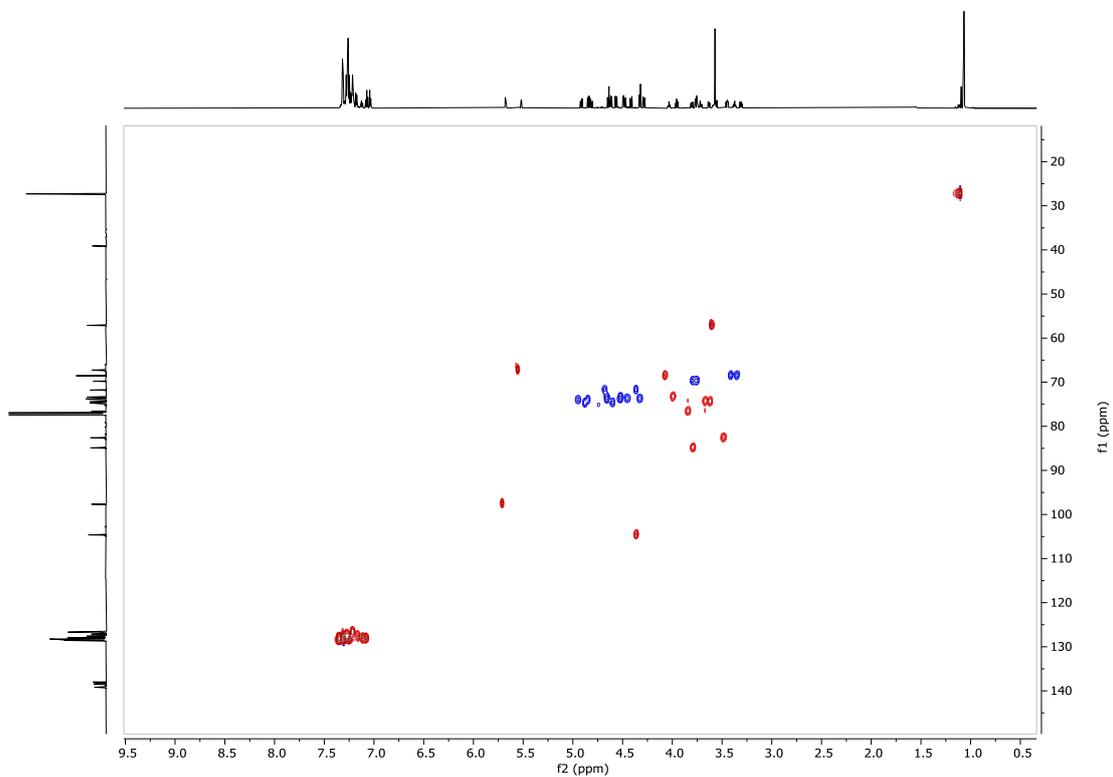
^1H NMR (700 MHz, CDCl_3) of **4Piv-A2**:



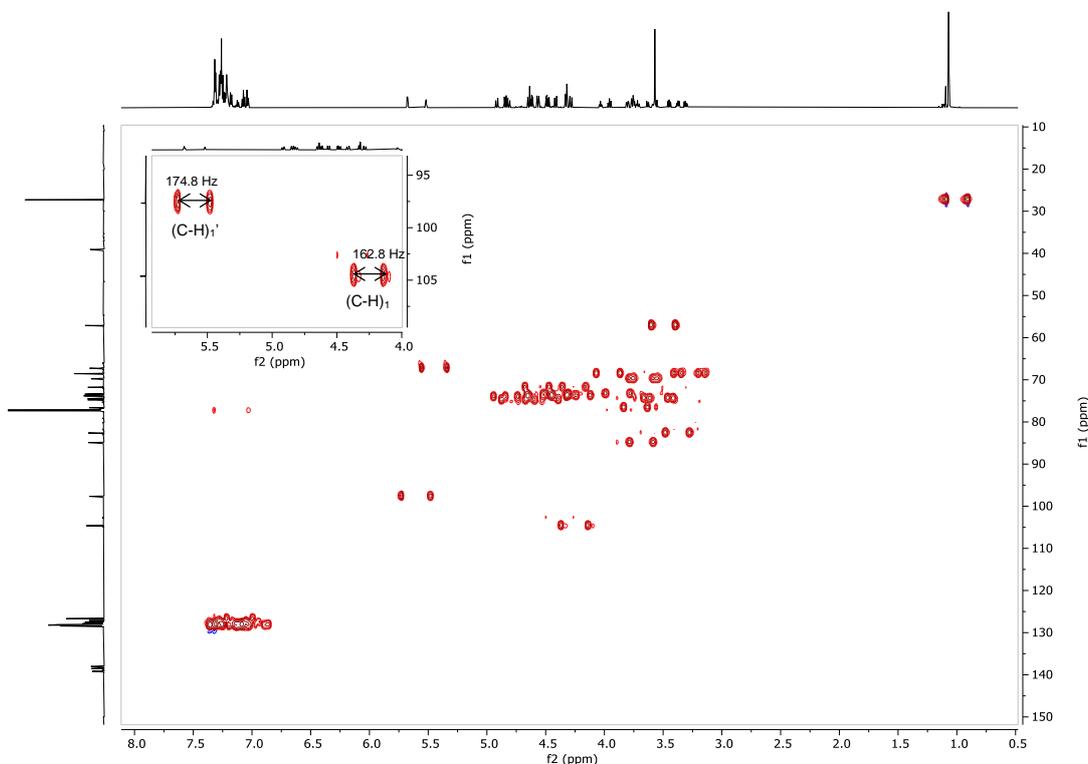
^{13}C NMR (176 MHz, CDCl_3) of **4Piv-A2**:



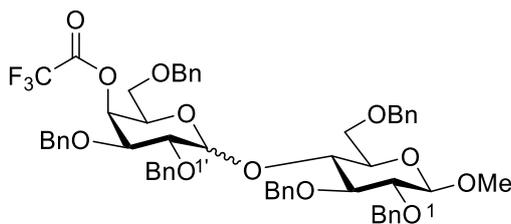
$^{13}\text{C},^1\text{H}$ HSQC of **4Piv-A2**:



Coupled $^{13}\text{C}, ^1\text{H}$ HSQC of **4Piv-A2**:



Methyl *O*-[2,3,6-tris-*O*-benzyl-4-(trifluoroacetate)- α/β -D-galactopyranosyl]-(1 \rightarrow 4)-2,3,6-tri-*O*-benzyl- β -D-glucopyranoside (4TFA-A2**)**



The title compound was prepared according to general procedure for glycosylations. Product **4TFA-A2** (23 mg, 23 μmol , 69%, α/β 97:3) was obtained as a colorless oil after purification using **Method-2b** ($t_R(\alpha) = 11.2$ min, $t_R(\beta) = 11.2$ min).

Data of the major isomer (α):

$^1\text{H NMR}$ (700 MHz, CDCl_3) δ 7.39 – 7.21 (m, 23H), 7.22 – 7.12 (m, 5H), 7.12 – 7.08 (m, 2H), 5.72 (d, $J = 3.8$ Hz, 1H, $\text{H}_{1'}$), 5.68 (dd, $J = 3.3, 1.4$ Hz, 1H), 4.91 (dd, $J = 35.1, 11.3$ Hz, 2H), 4.78 (d, $J = 11.6$ Hz, 1H), 4.70 – 4.57 (m, 4H), 4.56 (d, $J = 12.0$ Hz, 1H), 4.50 (d, $J = 11.6$ Hz, 1H), 4.40 (dd, $J = 25.0, 11.2$ Hz, 2H), 4.35

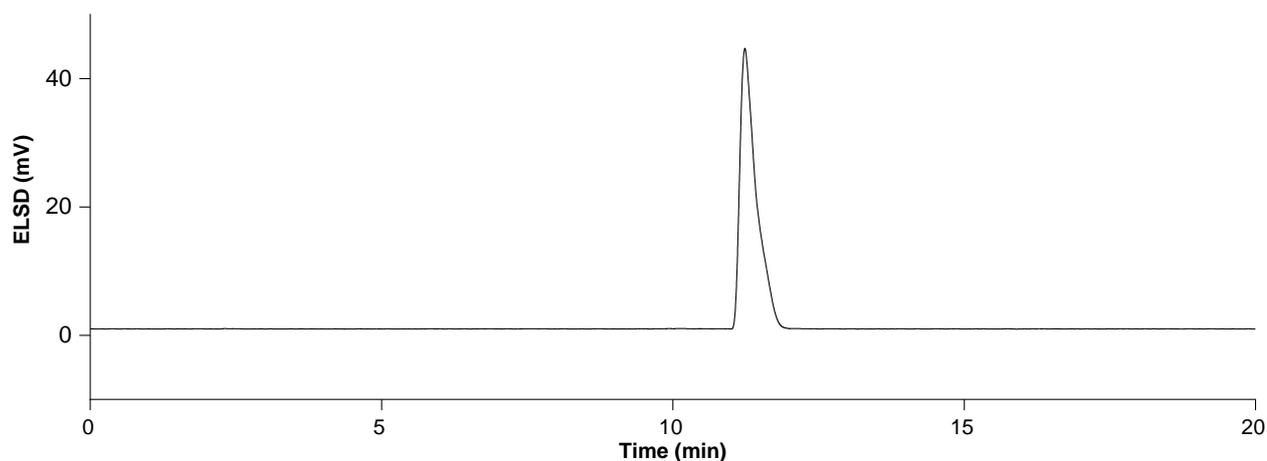
(d, $J = 7.7$ Hz, 1H), 4.26 (d, $J = 11.6$ Hz, 1H, **H_i**), 4.17 (ddd, $J = 8.5, 5.4, 1.5$ Hz, 1H), 4.04 (t, $J = 9.1$ Hz, 1H), 3.92 (dd, $J = 10.2, 3.1$ Hz, 1H), 3.81 – 3.71 (m, 3H), 3.67 (dd, $J = 10.2, 3.8$ Hz, 1H), 3.60 (s, 3H), 3.58 (ddd, $J = 9.6, 4.1, 2.6$ Hz, 1H), 3.48 (dd, $J = 9.1, 7.7$ Hz, 1H), 3.41 (dd, $J = 8.8, 5.5$ Hz, 1H), 3.31 (t, $J = 8.7$ Hz, 1H) ppm.

¹³C NMR (176 MHz, CDCl₃) δ 157.00 (q, $J = 42.6$ Hz), 138.8, 138.4, 138.4, 137.8, 137.7, 137.5, 128.6, 128.5, 128.5, 128.4, 128.3, 128.3, 128.3, 128.2, 128.1, 128.1, 128.0, 128.0, 127.9, 127.8, 127.8, 127.7, 127.3, 126.8, 114.73 (q, $J = 286.3$ Hz), 104.7 (**C_i**), 97.3 (**C_i'**), 84.9, 82.6, 76.0, 74.8, 74.4, 74.3, 74.2, 73.8, 73.6, 72.8, 72.5, 72.4, 69.5, 67.3, 67.1, 57.1 ppm.

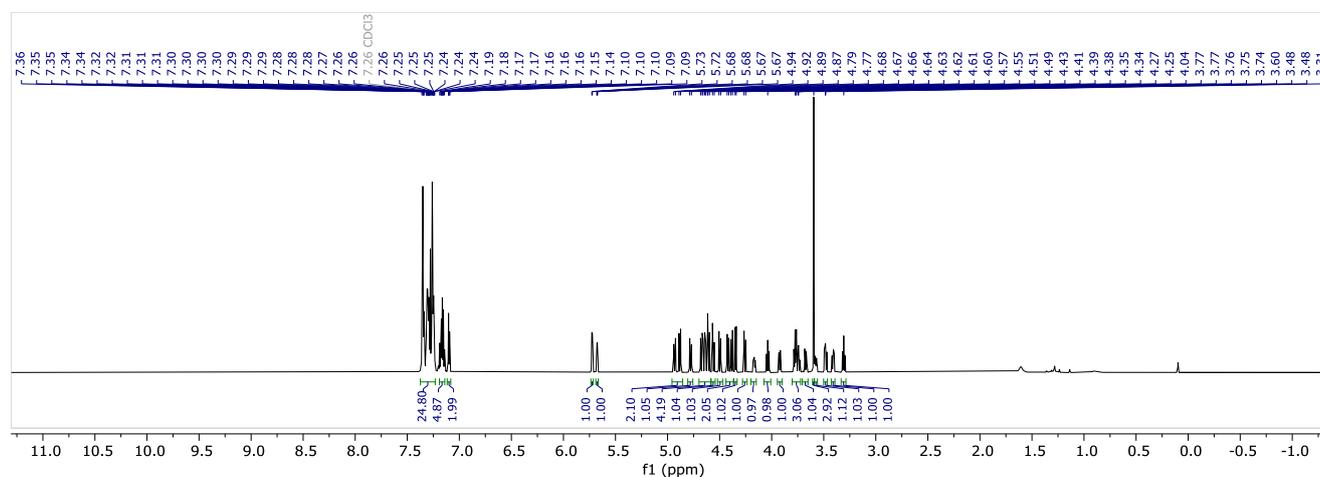
¹⁹F NMR (659 MHz, CDCl₃) δ -74.79 (s) ppm.

HRMS (QToF): Calcd for C₅₇H₅₉F₃O₁₂Na [M + Na]⁺ 1015.3851; found 1015.3853.

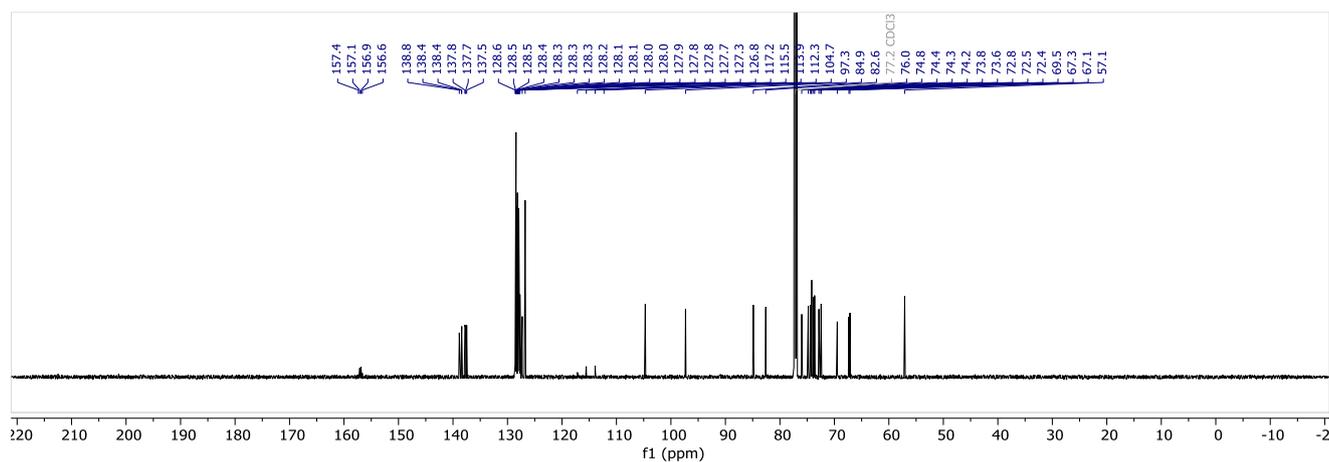
NP-HPLC of **4TFA-A2** (ELSD trace, $t_R(\alpha) = 11.2$ min, $t_R(\beta) = 11.2$ min):



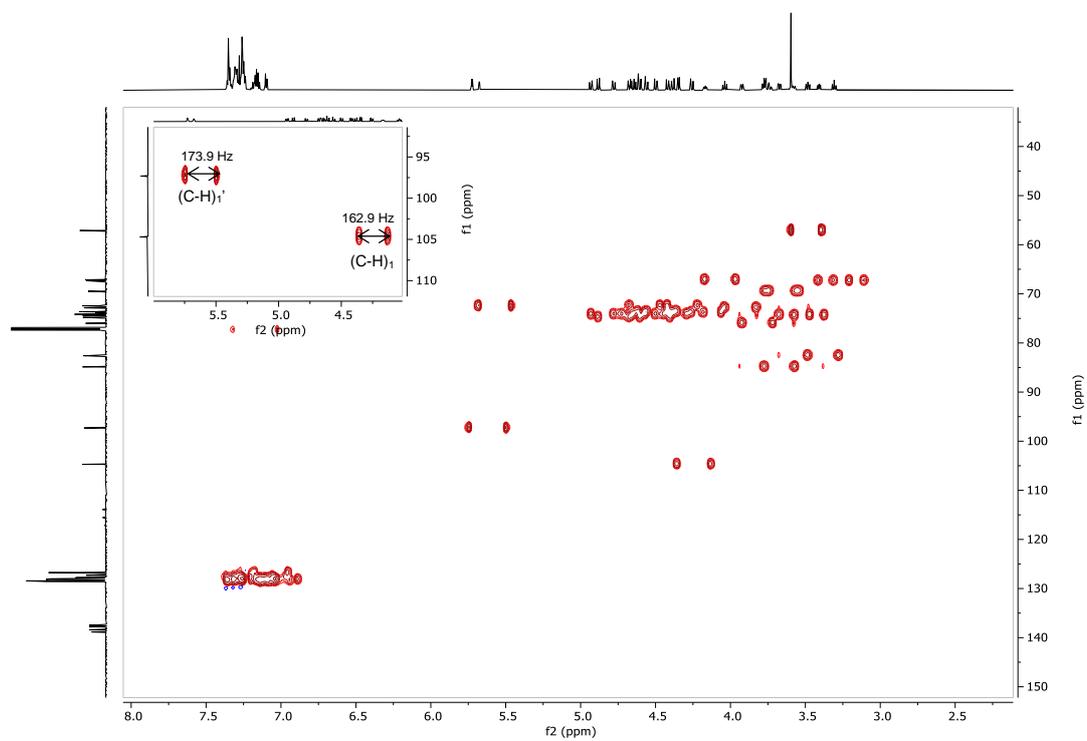
¹H NMR (700 MHz, CDCl₃) of 4TFA-A2:



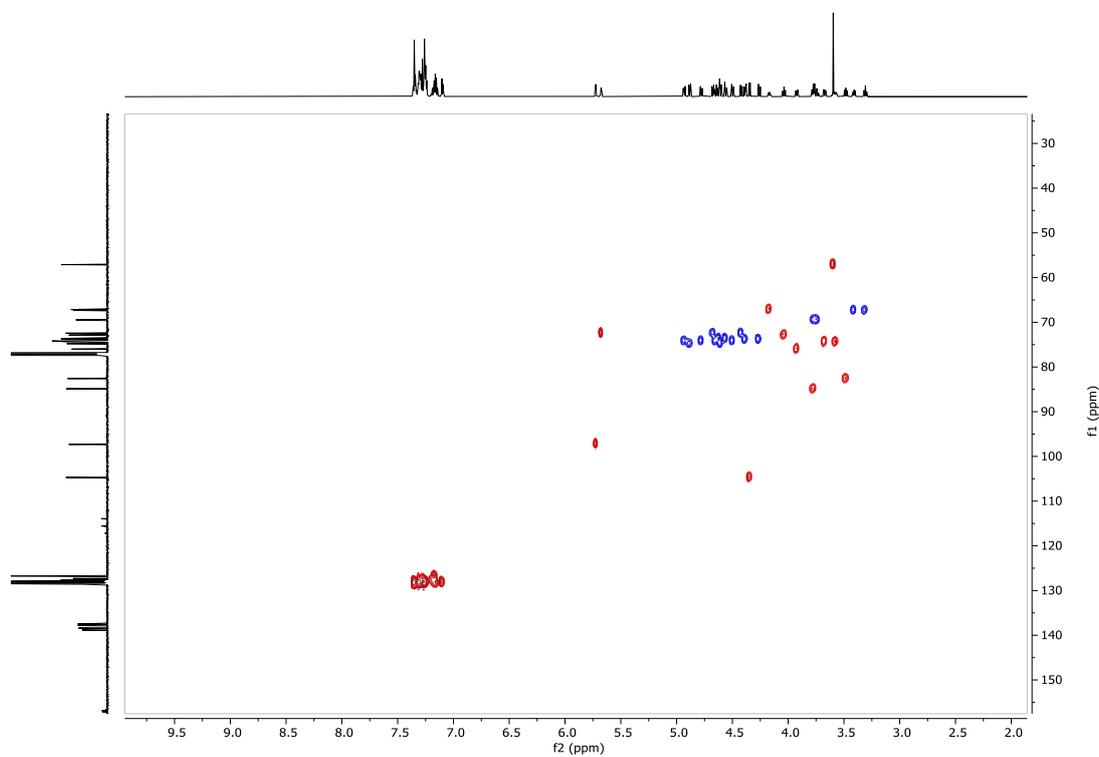
¹³C NMR (176 MHz, CDCl₃) of 4TFA-A2:



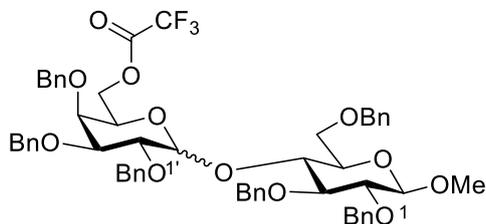
Coupled $^{13}\text{C},^1\text{H}$ HSQC of **4TFA-A2**:



$^{13}\text{C},^1\text{H}$ HSQC of **4TFA-A2**:



Methyl *O*-[2,3,4-tris-*O*-benzyl-6-(trifluoroacetate)- α/β -D-galactopyranosyl]-(1 \rightarrow 4)-2,3,6-tri-*O*-benzyl- β -D-glucopyranoside (6TFA-A2**)**



The title compound was prepared according to general procedure for glycosylations. Product **6TFA-A2** (20 mg, 20 μ mol, 59%, α/β **74:26**) was obtained as a colorless oil after purification using **Method-2b** ($t_R(\alpha)$ = 9.9 min, $t_R(\beta)$ = 9.9 min).

Data of the major isomer (α):

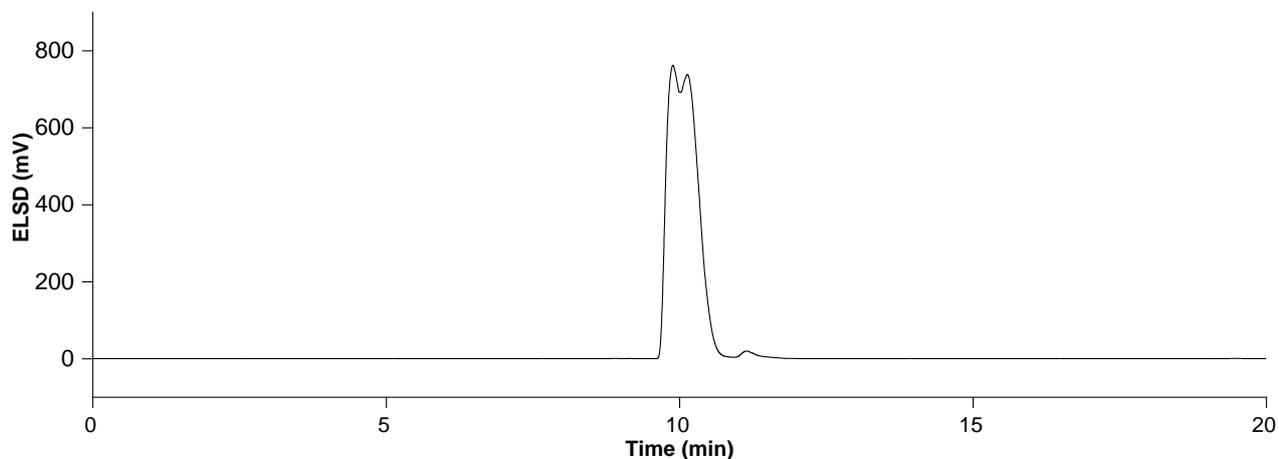
$^1\text{H NMR}$ (700 MHz, CDCl_3) δ 7.38 – 7.26 (m, 20H), 7.23 – 7.13 (m, 10H), 5.73 (d, J = 3.8 Hz, 1H, $\mathbf{H}_{1'}$), 4.94 – 4.87 (m, 3H), 4.77 – 4.72 (m, 3H), 4.67 (dd, J = 15.7, 11.7 Hz, 3H), 4.60 (d, J = 11.1 Hz, 1H), 4.54 (dd, J = 14.0, 11.6 Hz, 2H), 4.48 (d, J = 12.3 Hz, 1H), 4.36 (d, J = 7.7 Hz, 1H, \mathbf{H}_1), 4.32 – 4.25 (m, 1H), 3.98 (dd, J = 10.2, 3.9 Hz, 1H), 3.91 (dd, J = 11.0, 5.2 Hz, 1H), 3.88 – 3.76 (m, 4H), 3.68 – 3.64 (m, 1H), 3.60 (s, 3H), 3.46 (dd, J = 9.1, 7.6 Hz, 1H) ppm.

$^{13}\text{C NMR}$ (176 MHz, CDCl_3) δ 156.95 (q, J = 42.4 Hz), 138.8, 138.5, 138.5, 138.5, 138.2, 137.9, 128.6, 128.6, 128.5, 128.4, 128.4, 128.2, 128.1, 128.0, 127.9, 127.8, 127.7, 127.7, 127.6, 127.6, 126.7, 114.55 (q, J = 285.5 Hz), 104.6 (\mathbf{C}_1), 97.3 ($\mathbf{C}_{1'}$), 84.8, 82.4, 78.7, 75.5, 74.7, 74.5, 74.5, 74.0, 73.9, 73.5, 73.4, 73.4, 69.9, 68.4, 66.9, 57.2 ppm.

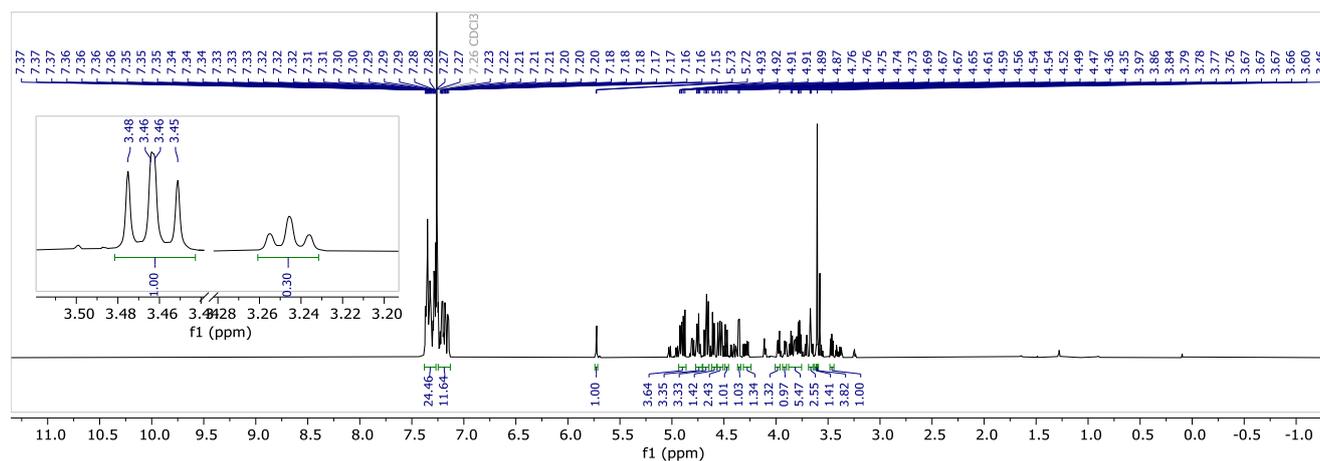
$^{19}\text{F NMR}$ (659 MHz, CDCl_3) δ -74.57 (s).

HRMS (QToF): Calcd for $\text{C}_{57}\text{H}_{59}\text{F}_3\text{O}_{12}\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 1015.3851; found 1015.3860.

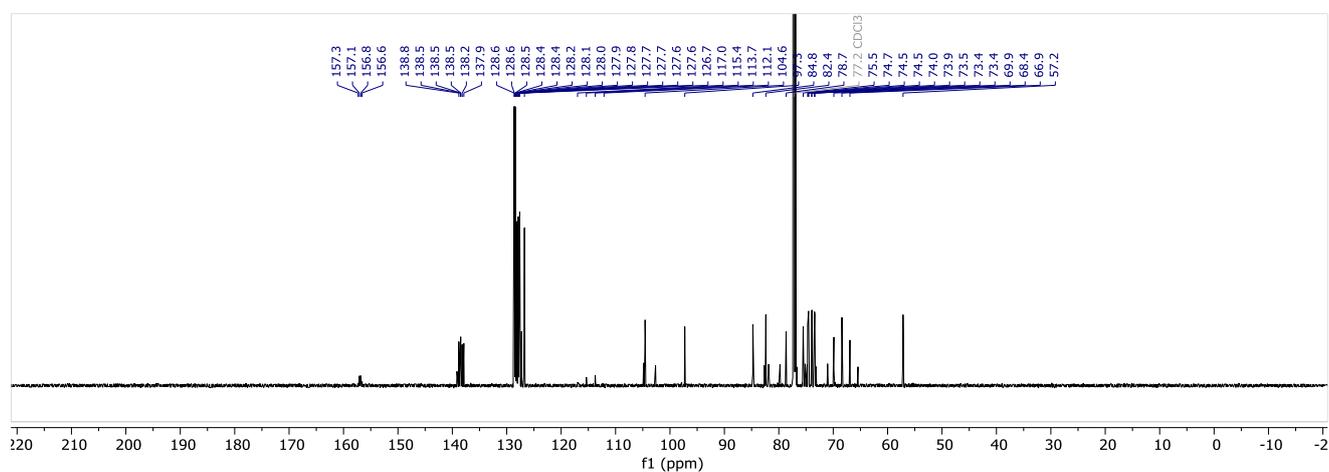
NP-HPLC of **6TFA-A2** (ELSD trace, $t_R(\alpha) = 9.9$ min, $t_R(\beta) = 9.9$ min):



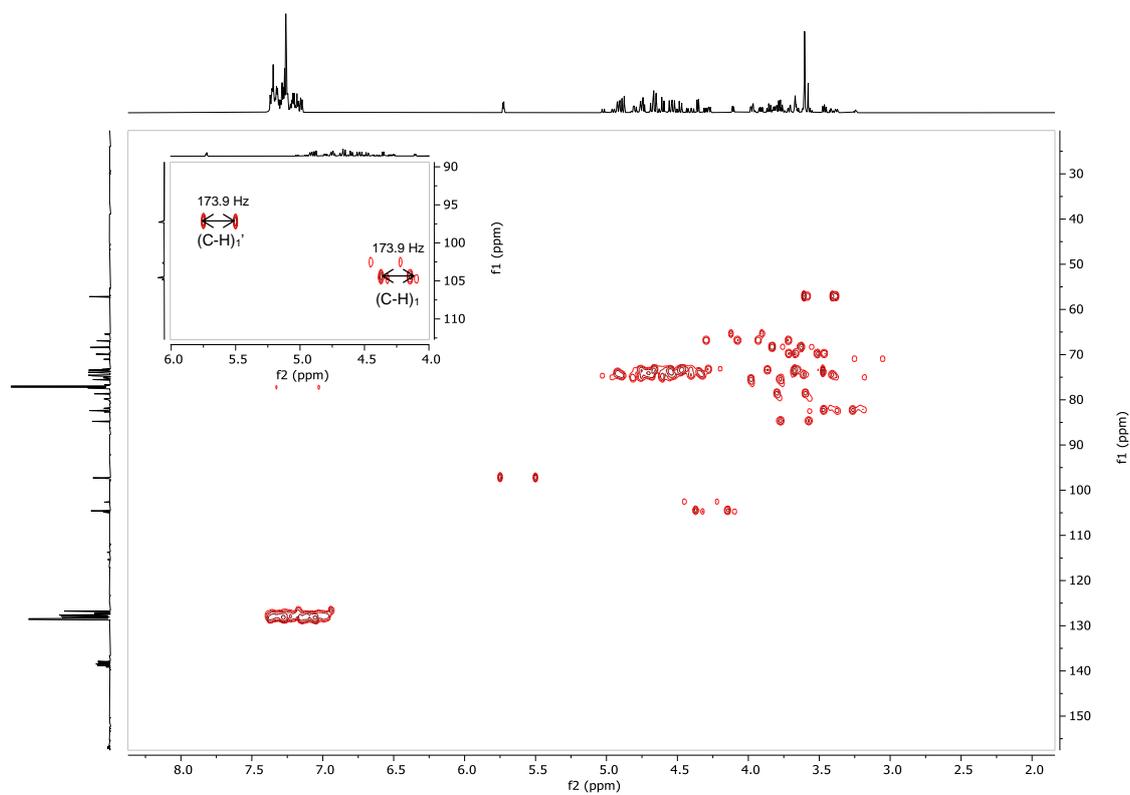
^1H NMR (700 MHz, CDCl_3) of **6TFA-A2**:



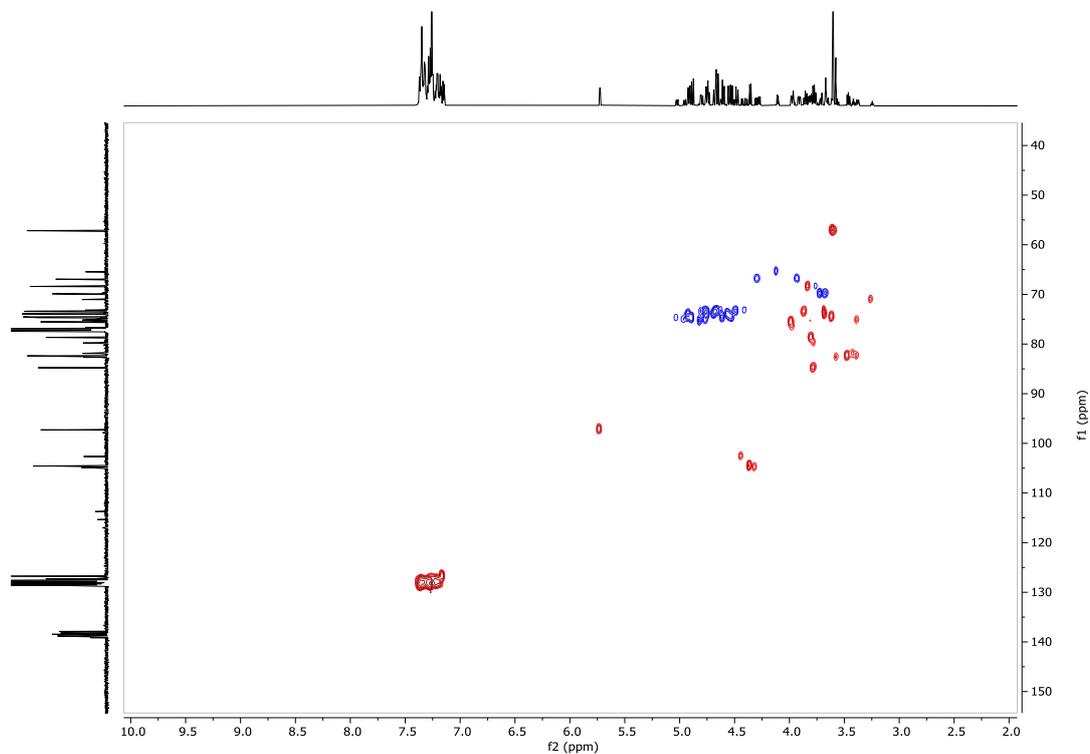
^{13}C NMR (176 MHz, CDCl_3) of **6TFA-A2**:



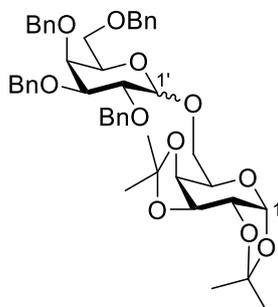
Coupled $^{13}\text{C}, ^1\text{H}$ HSQC of **6TFA-A2**:



$^{13}\text{C}, ^1\text{H}$ HSQC of **6TFA-A2**:



1,2:3,4-Bis-*O*-(1-methylethylidene)-6-*O*-[2,3,4,6-tetra-*O*-benzyl- α/β -D-galactopyranosyl]- α -D-galactopyranoside (4Bn-A1**)**



The title compound was prepared according to general procedure for glycosylations. Product **4Bn-A1** (21 mg, 27 μ mol, 78%, α/β 60:40) was obtained as a colorless oil after purification using **Method-2b** (tR (α) = 10.4 min, tR (β) = 11.4 min).

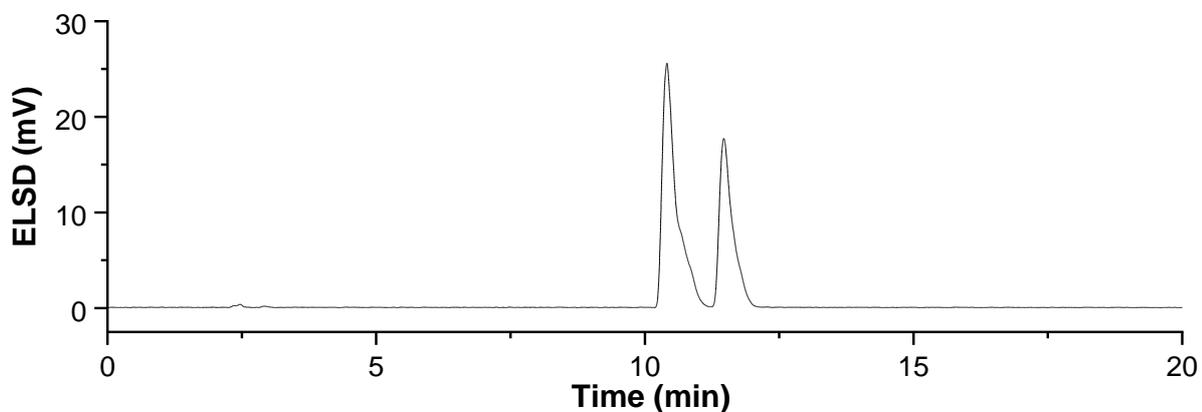
Data of the anomeric mixture:

$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.41 – 7.27 (m, 36H), 7.26 – 7.21 (m, 2H), 5.56 (d, J = 5.0 Hz, 1H, $\text{H}_{1'}(\alpha)$), 5.52 (d, J = 5.0 Hz, 1H, $\text{H}_1(\alpha)$), 5.06 (d, J = 11.0 Hz, 1H), 5.01 (d, J = 3.6 Hz, 1H $\text{H}_{1'}(\alpha)$), 4.94 (dd, J = 11.5, 3.9 Hz, 2H), 4.89 – 4.66 (m, 8H), 4.66 – 4.52 (m, 5H), 4.47 (t, J = 11.9 Hz, 2H), 4.45 – 4.37 (m, 4H $\text{H}_{1'}(\beta)$, CH_2Ph), 4.37 – 4.29 (m, 4H), 4.25 – 4.20 (m, 1H), 4.13 (dd, J = 10.7, 3.5 Hz, 1H), 4.08 – 4.01 (m, 6H), 3.96 (dd, J = 10.0, 2.8 Hz, 1H), 3.89 (d, J = 2.6 Hz, 1H), 3.81 (ddd, J = 23.0, 10.1, 7.1 Hz, 2H), 3.74 (dd, J = 10.5, 7.1 Hz, 1H), 3.69 (dd, J = 10.6, 7.5 Hz, 1H), 3.57 (s, J = 9.0, 3.4 Hz, 3H), 3.54 – 3.48 (m, 3H), 1.52 (s, 3H), 1.49 (s, 3H), 1.44 (s, 3H), 1.43 (s, 3H), 1.33 (s, 4H), 1.31 (s, 4H), 1.30 (s, 4H) ppm.

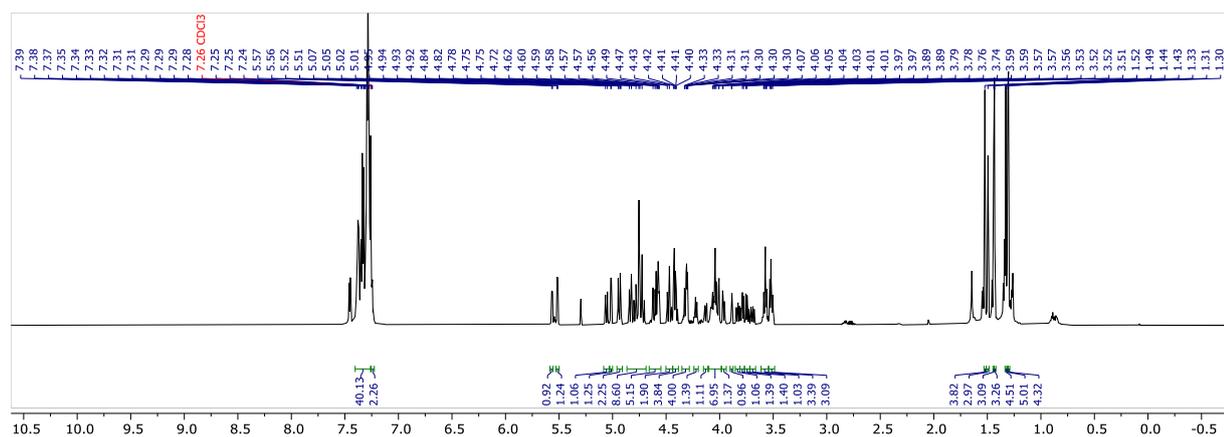
$^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 139.1, 139.0, 138.8, 138.7, 138.7, 138.1, 137.9, 128.6, 128.4, 128.4, 128.4, 128.3, 128.3, 128.3, 128.3, 128.2, 128.1, 128.1, 127.9, 127.8, 127.8, 127.8, 127.7, 127.5, 127.5, 127.5, 127.4, 127.4, 127.3, 109.3, 109.2, 108.6, 108.5, 104.7($\text{C}_{1'}(\beta)$), 97.6($\text{C}_{1'}(\alpha)$), 96.4($\text{C}_1(\alpha)$), 96.3($\text{C}_1(\alpha)$), 82.0, 79.1, 79.0, 77.3, 77.1, 76.8, 76.4, 74.9, 74.8, 74.8, 74.5, 73.6, 73.5, 73.4, 73.3, 73.1, 73.1, 72.7, 71.5, 70.9, 70.8, 70.7, 70.6, 70.5, 69.6, 69.2, 68.7, 68.7, 67.4, 66.3, 65.8, 26.2, 26.1, 26.0, 26.0, 25.1, 25.0, 24.6, 24.5 ppm.

HRMS (ESI): Calcd for $\text{C}_{46}\text{H}_{54}\text{O}_{11}\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 805.3564; found 805.3688.

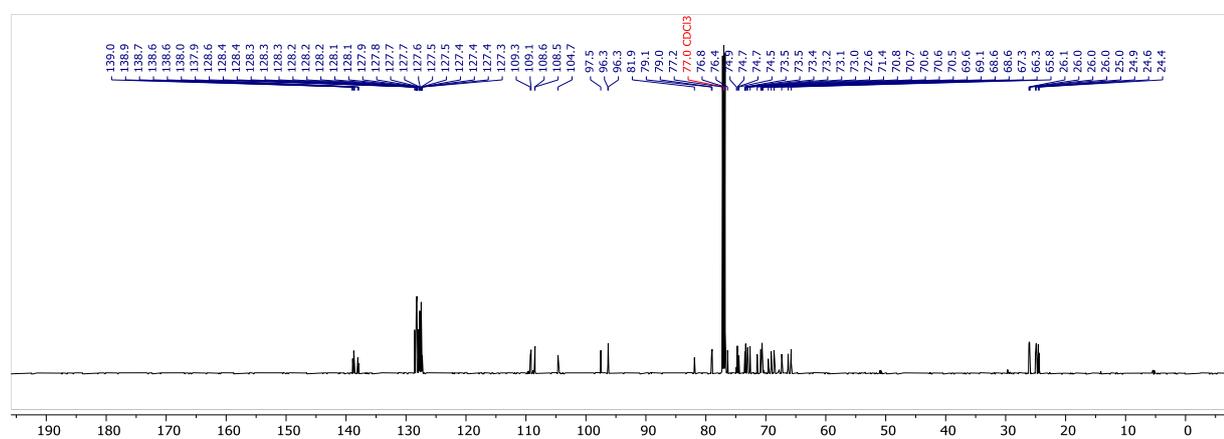
NP-HPLC of **4Bn-A1** (ELSD trace, $t_R(\alpha) = 10.4$ min, $t_R(\beta) = 11.4$ min):



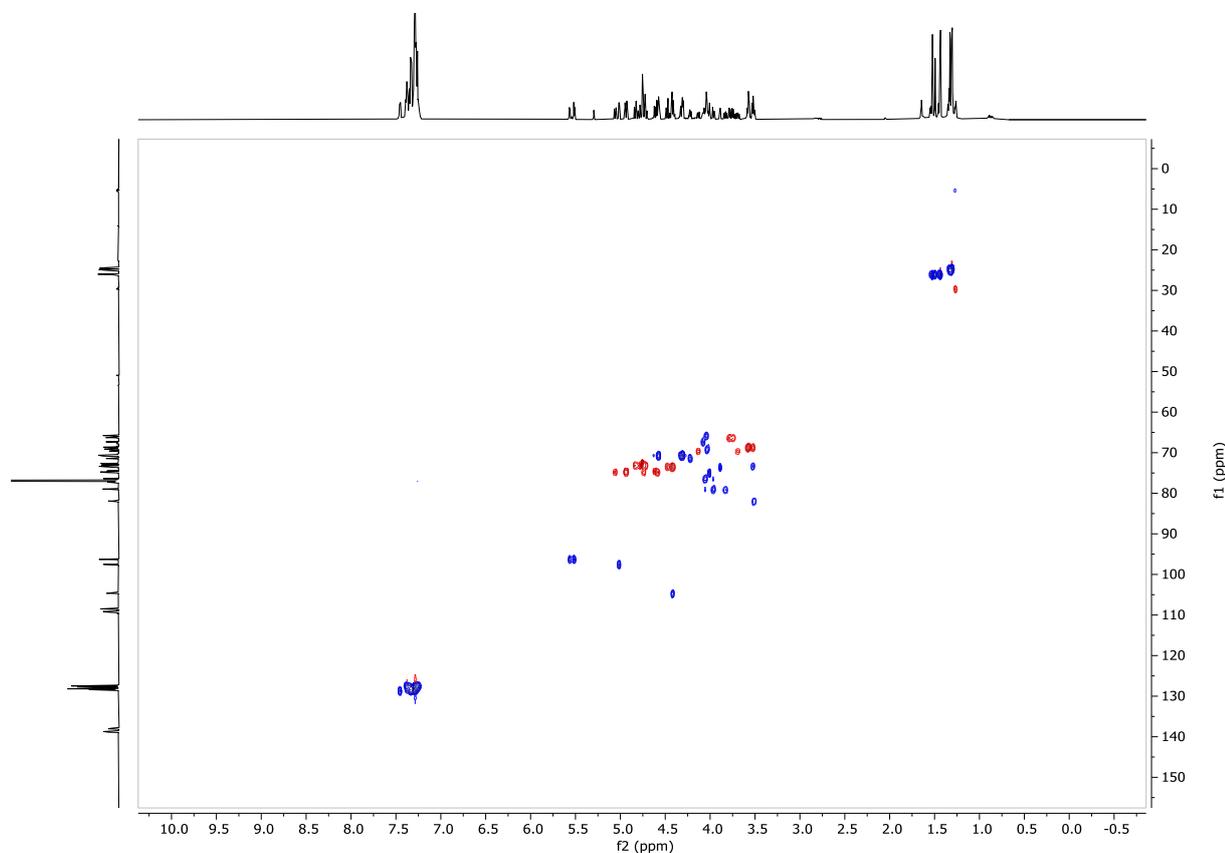
^1H NMR (600 MHz, CDCl_3) of **4Bn-A1**:



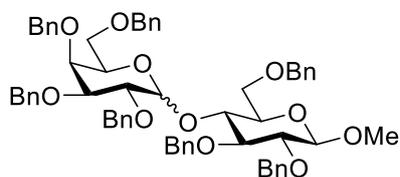
^{13}C NMR (151 MHz, CDCl_3) of **4Bn-A1**:



$^{13}\text{C}, ^1\text{H}$ HSQC of **4Bn-A1**:



Methyl *O*-[2,3,4,6-tetra-*O*-benzyl- α/β -D-galactopyranosyl]-(1 \rightarrow 4)-2,3,6-tri-*O*-benzyl- β -D-glucopyranoside (**4Bn-A2**)



The title compound was prepared according to general procedure for glycosylations. Product **4Bn-A2** (27 mg, 28 μmol , 79%, α/β 84:16) was obtained as a colorless oil after purification using **Method-2b** (t_R (α) = 11.9 min, t_R (β) = 12.2 min).

Data of the α -anomer (major product):

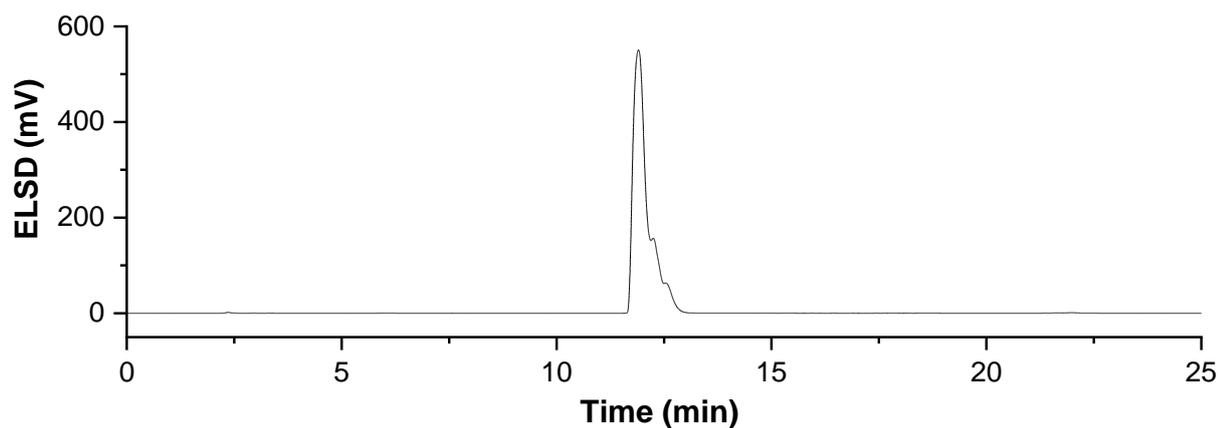
$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.44 – 7.21 (m, 30H), 7.14 (m, 5H), 5.73 (d, $J = 3.9$ Hz, 1H, $\text{H}_1'(\alpha)$), 4.91 (d, $J = 11.5$ Hz, 1H), 4.86 (dd, $J = 11.2, 2.0$ Hz, 2H), 4.76 (d, $J = 11.6$ Hz, 1H), 4.69 – 4.56 (m, 6H), 4.53 (d, $J = 11.6$ Hz, 2H), 4.47 (d, $J = 12.2$ Hz, 1H), 4.38 (d, $J = 11.7$ Hz, 1H), 4.33 (d, $J = 7.7$ Hz, 1H, $\text{H}_1(\beta)$), 4.29 (d, $J = 11.5$ Hz, 1H), 4.02 – 3.94 (m, 2H), 3.93 (s, 1H), 3.88 (t, $J = 6.6$ Hz, 1H), 3.85 – 3.73 (m, 4H), 3.70 (dd, $J = 10.8, 5.2$ Hz, 1H), 3.63 – 3.58 (m, 1H), 3.57 (s, 3H), 3.47 (m, 3H) ppm.

Data of the anomeric mixture:

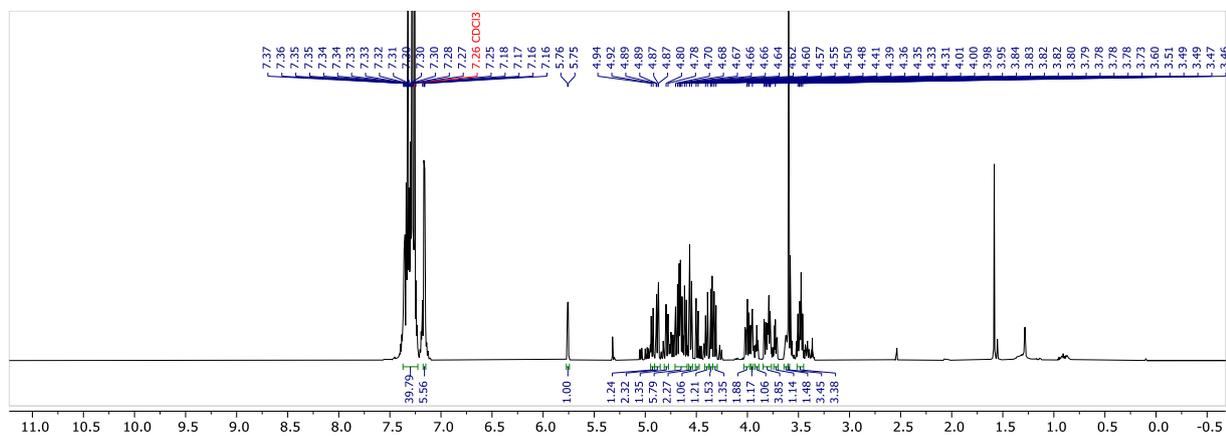
$^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 138.7, 138.6, 138.4, 138.3, 138.2, 138.0, 128.5, 128.4, 128.4, 128.3, 128.3, 128.3, 128.3, 128.2, 128.2, 128.1, 128.1, 128.0, 128.0, 127.9, 127.9, 127.8, 127.8, 127.7, 127.6, 127.5, 127.5, 127.4, 127.4, 127.3, 127.0, 126.6, 104.6 $\text{C}_1'(\beta)$, 104.4 $\text{C}_1(\beta)$, 102.8 $\text{C}_1(\beta)$, 97.4 $\text{C}_1'(\alpha)$, 84.8, 83.9, 82.9, 82.5, 82.4, 81.8, 81.8, 79.9, 79.1, 77.2, 77.0, 76.8, 75.4, 75.4, 75.3, 75.2, 75.1, 74.9, 74.7, 74.7, 74.7, 74.6, 74.5, 74.3, 74.0, 74.0, 73.8, 73.7, 73.5, 73.5, 73.4, 73.2, 73.1, 72.9, 72.7, 72.6, 72.5, 71.6, 70.3, 69.8, 69.6, 68.7, 68.3, 68.0, 57.1, 57.0 ppm.

HRMS (ESI): Calcd for $\text{C}_{62}\text{H}_{66}\text{O}_{11}\text{Na}$ $[\text{M} + \text{Na}]^+$ 1009.4503; found 1009.4631.

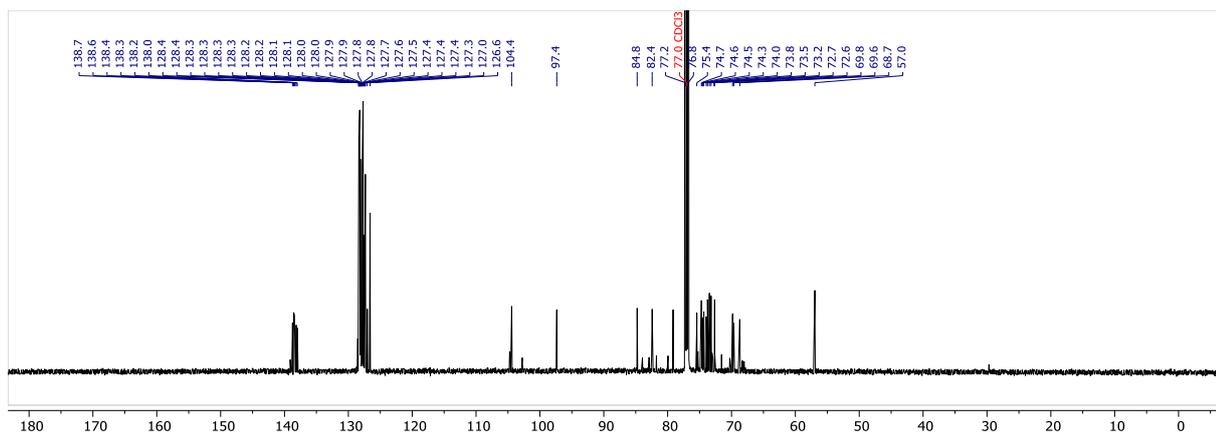
NP-HPLC of **4Bn-A2** (ELSD trace, $t_R(\alpha) = 11.9$ min, $t_R(\beta) = 12.2$ min):



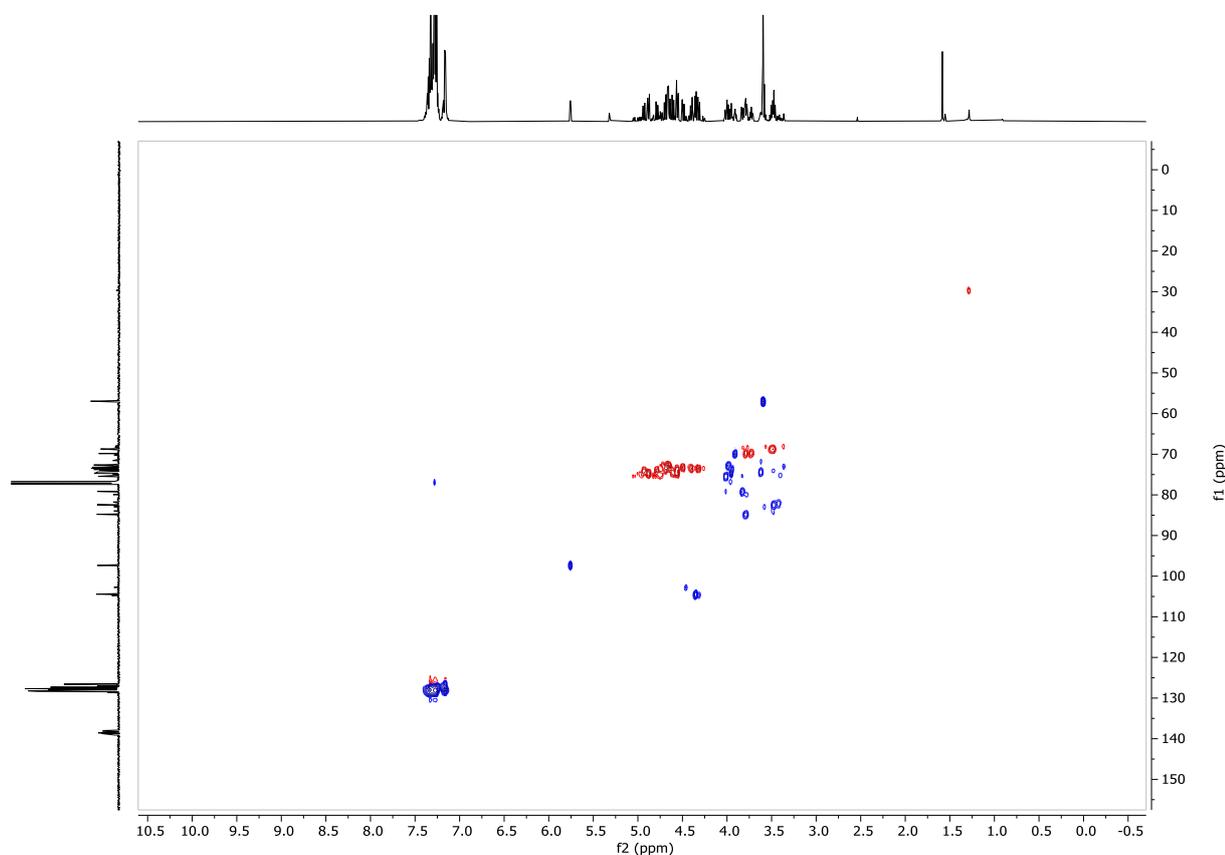
^1H NMR (600 MHz, CDCl_3) of **4Bn-A2**:



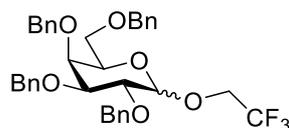
^{13}C NMR (151 MHz, CDCl_3) of **4Bn-A2**:



$^{13}\text{C}, ^1\text{H}$ HSQC of **4Bn-A2**:



2,2-Trifluoroethyl 2,3,4,6-tetra-*O*-benzyl- α/β -D-galactopyranoside (**4Bn-A3**)



The title compound was prepared according to general procedure for glycosylations. Product **4Bn-A3** (17 mg, 28 μmol , 80%, α/β 68:32) was obtained as a colorless oil after purification using **Method-2b** (t_R (α) = 6.1 min, t_R (β) = 7.3 min).

Data of the anomeric mixture:

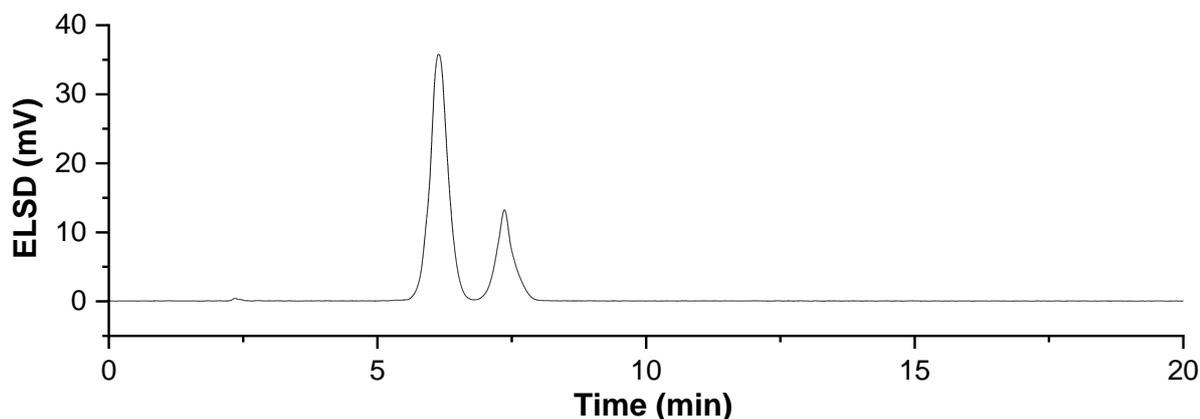
^1H NMR (600 MHz, CDCl_3) δ 7.40 – 7.26 (m, 38H), 7.25 (s, 2H), 4.94 (d, J = 11.5 Hz, 2H), 4.90 (d, J = 10.5 Hz, 1H), 4.88 (d, J = 3.7 Hz, 1H, $\mathbf{H}_1(\alpha)$), 4.85 (d, J = 11.6 Hz, 1H), 4.82 (d, J = 11.9 Hz, 1H), 4.77 (d, J = 11.8

Hz, 1H), 4.75 – 4.69 (m, 3H), 4.66 (d, $J = 11.9$ Hz, 2H), 4.62 (d, $J = 11.6$ Hz, 1H), 4.56 (d, $J = 11.4$ Hz, 1H), 4.47 (d, $J = 7.0$ Hz, 1H, $\mathbf{H}_1(\beta)$), 4.45 (d, $J = 7.5$ Hz, 1H), 4.43 – 4.38 (m, 3H), 4.17 (dq, $J = 12.4, 8.8$ Hz, 1H), 4.08 (d, $J = 3.7$ Hz, 1H), 4.06 (d, $J = 3.7$ Hz, 1H), 3.97 (d, $J = 2.0$ Hz, 2H), 3.96 – 3.83 (m, 8H), 3.60 – 3.52 (m, 3H), 3.51 (s, 2H), 3.50 (s, 1H) ppm.

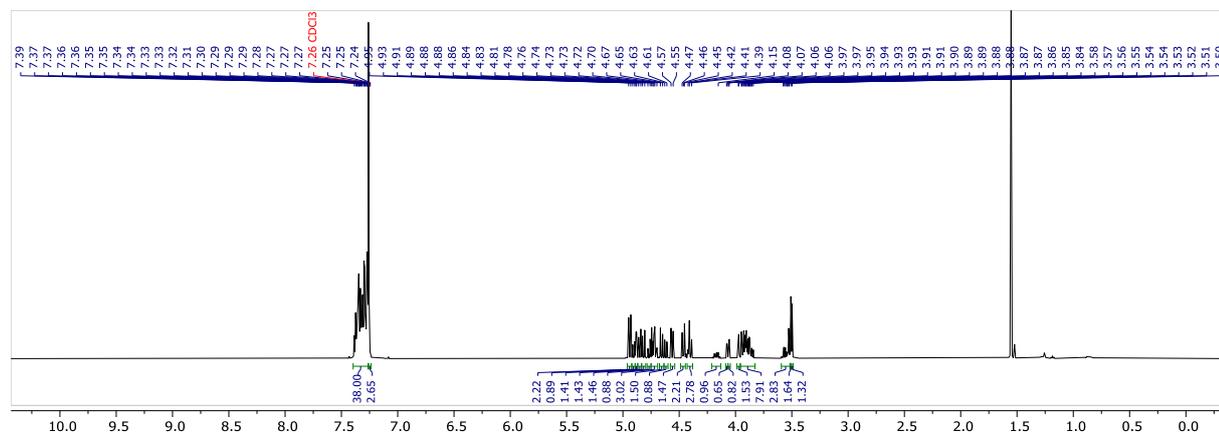
^{13}C NMR (151 MHz, CDCl_3) δ 138.7, 138.4, 138.4, 137.8, 137.7, 128.5, 128.5, 128.4, 128.4, 128.4, 128.4, 128.3, 128.2, 128.2, 127.9, 127.9, 127.8, 127.7, 127.7, 127.6, 127.5, 127.5, 127.5, 103.9 ($\mathbf{C}_1(\beta)$), 98.2 ($\mathbf{C}_1(\alpha)$), 81.8, 79.0, 78.6, 77.2, 77.0, 76.8, 76.1, 75.3, 74.8, 74.8, 74.6, 73.7, 73.6, 73.5, 73.4, 73.3, 73.2, 70.0, 68.7, 64.6 ppm.

HRMS (ESI): Calcd for $\text{C}_{36}\text{H}_{37}\text{F}_3\text{O}_6\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 645.2440; found 645.2419.

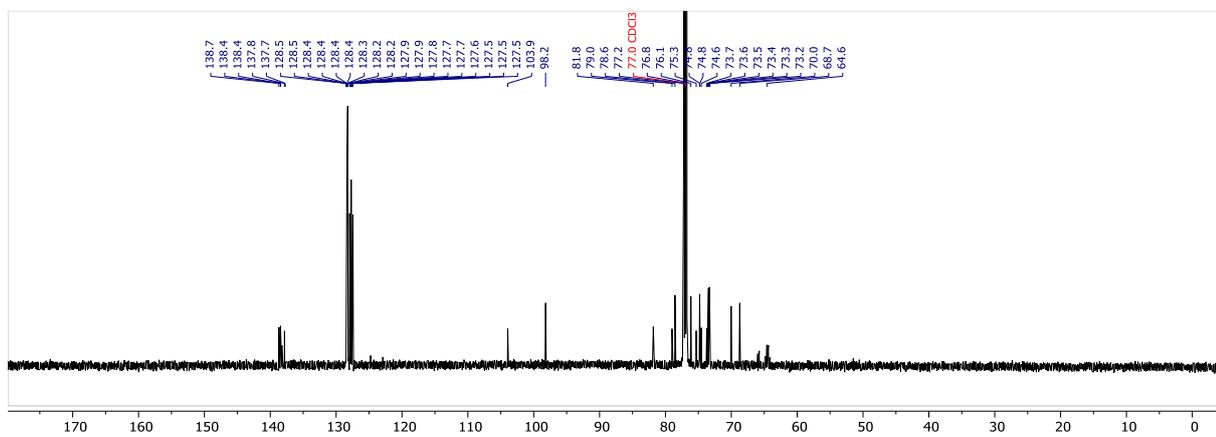
NP-HPLC of **4Bn-A3** (ELSD trace, $t_R(\alpha) = 6.1$ min, $t_R(\beta) = 7.3$ min):



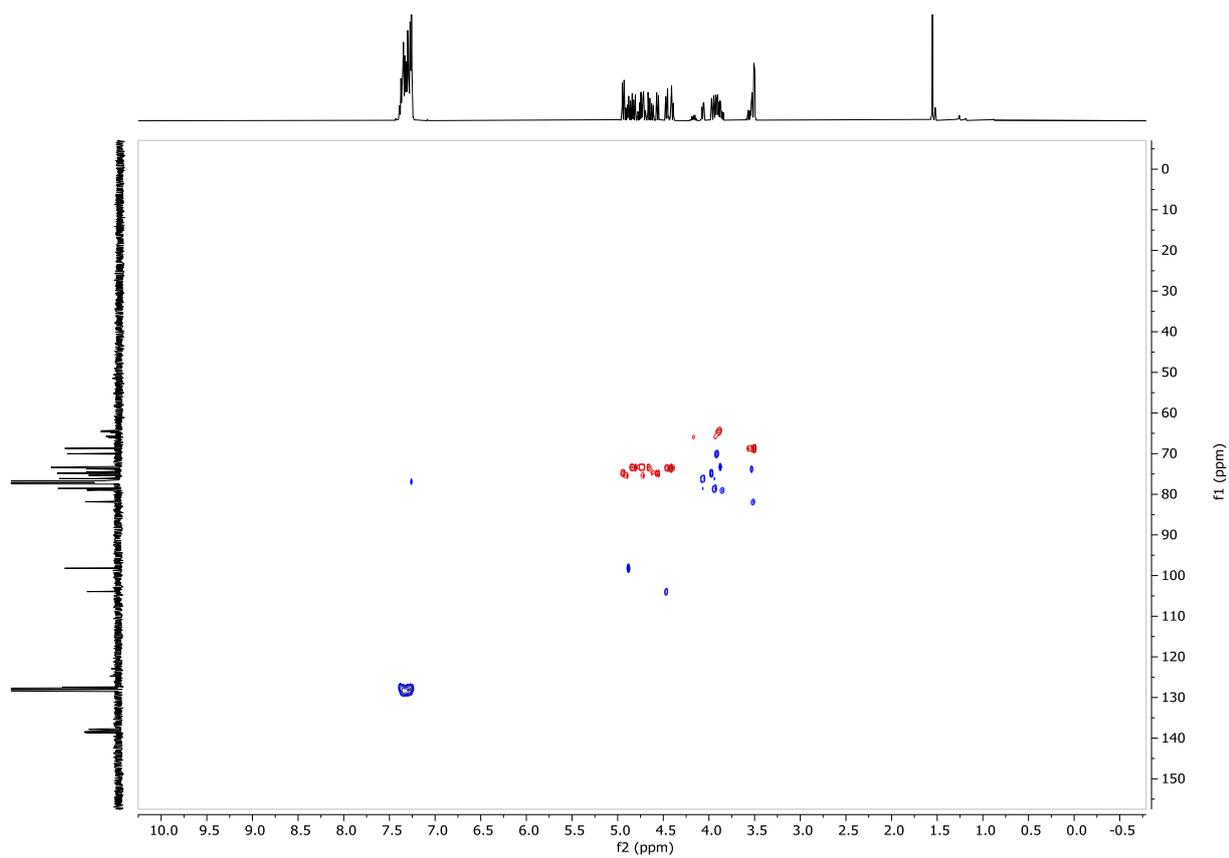
^1H NMR (600 MHz, CDCl_3) of **4Bn-A3**:



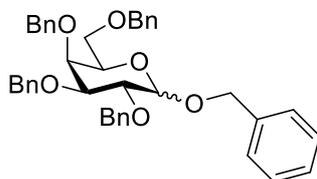
^{13}C NMR (151 MHz, CDCl_3) of **4Bn-A3**:



$^{13}\text{C}, ^1\text{H}$ HSQC of **4Bn-A3**:



2,3,4,6-tetra-*O*-benzyl- α/β -D-galactopyranosyl (**4Bn-A4**)



The title compound was prepared according to general procedure for glycosylations. Product **4Bn-A4** (17 mg, 27 μ mol, 77%, α/β 18:82) was obtained as a colorless oil after purification using **Method-2b** (t_R (α) = 8.1 min, t_R (β) = 8.6 min).

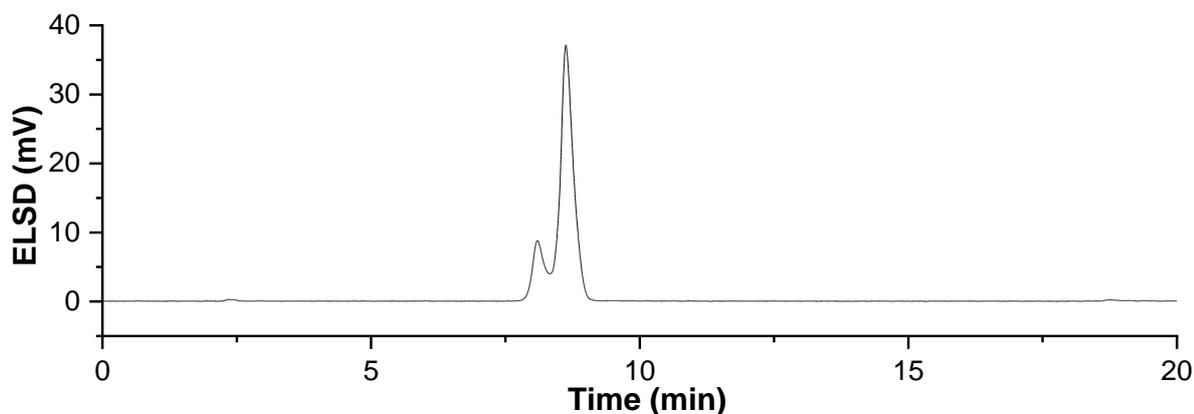
Data of the anomeric mixture:

$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.41 – 7.23 (m, 50H), 4.97 – 4.91 (m, 6H), 4.90 (d, J = 3.5 Hz, 1H, **H1**(α)), 4.85 (d, J = 11.7 Hz, 1H), 4.79 – 4.67 (m, 8H), 4.63 (dd, J = 11.9, 2.7 Hz, 4H), 4.60 – 4.54 (m, 3H), 4.49 – 4.42 (m, 6H, **H1**(β), CH_2Ph), 4.42 – 4.38 (m, 2H), 4.06 – 3.94 (m, 4H), 3.92 – 3.86 (m, 4H), 3.66 – 3.57 (m, 4H), 3.56 – 3.49 (m, 5H), 3.46 (dd, J = 9.3, 6.0 Hz, 1H) ppm.

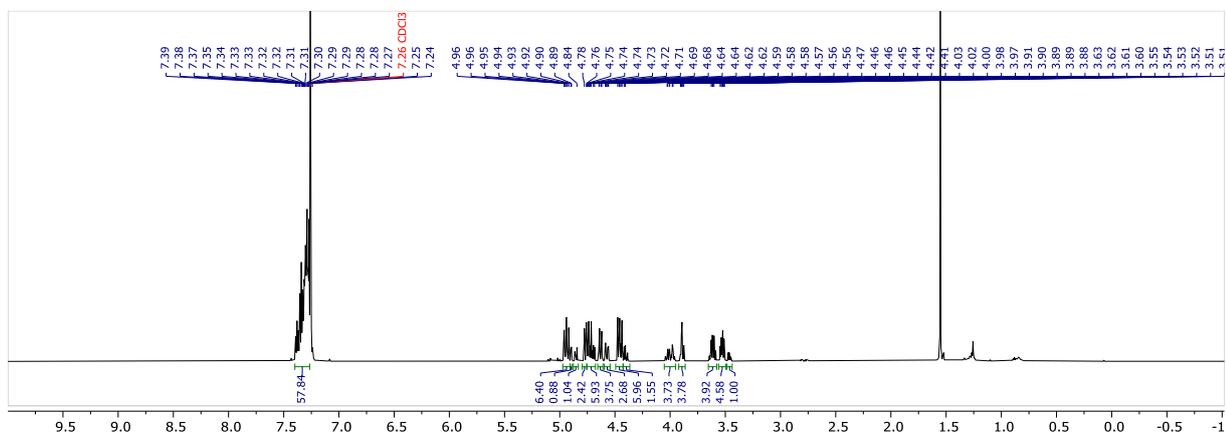
$^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 138.5, 137.6, 128.4, 128.4, 128.4, 128.3, 128.3, 128.2, 128.2, 128.2, 127.9, 127.8, 127.8, 127.8, 127.7, 127.7, 127.6, 127.6, 127.5, 127.4, 102.8 (**C1**(β)), 96.1 (**C1**(α)), 82.3, 79.6, 79.2, 77.2, 77.0, 76.8, 76.5, 75.2, 75.0, 74.7, 74.5, 73.5, 73.5, 73.4, 73.1, 70.9, 69.5, 69.0, 68.9 ppm.

HRMS (ESI): Calcd for $\text{C}_{41}\text{H}_{42}\text{O}_6\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 653.2879; found 653.2873.

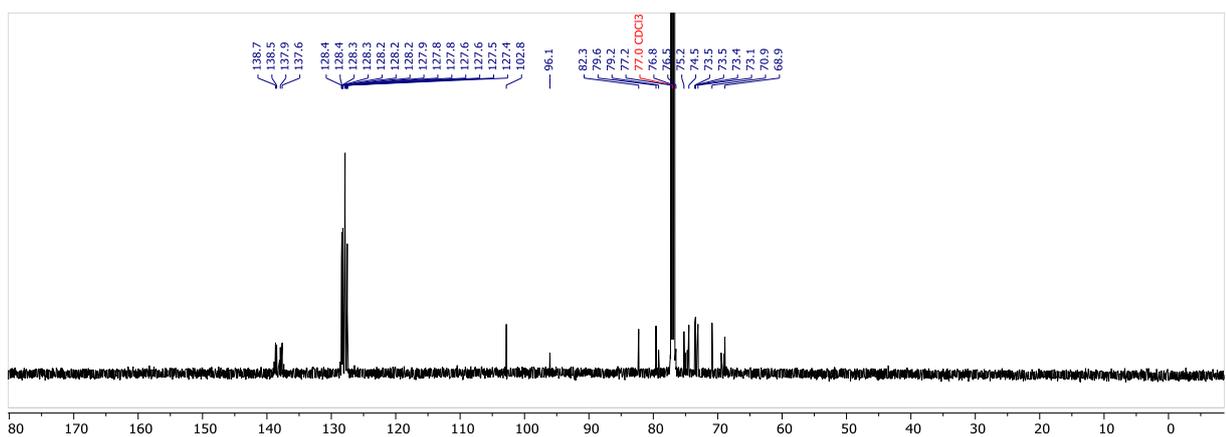
NP-HPLC of **4Bn-A4** (ELSD trace, t_R (α) = 8.1 min, t_R (β) = 8.6 min):



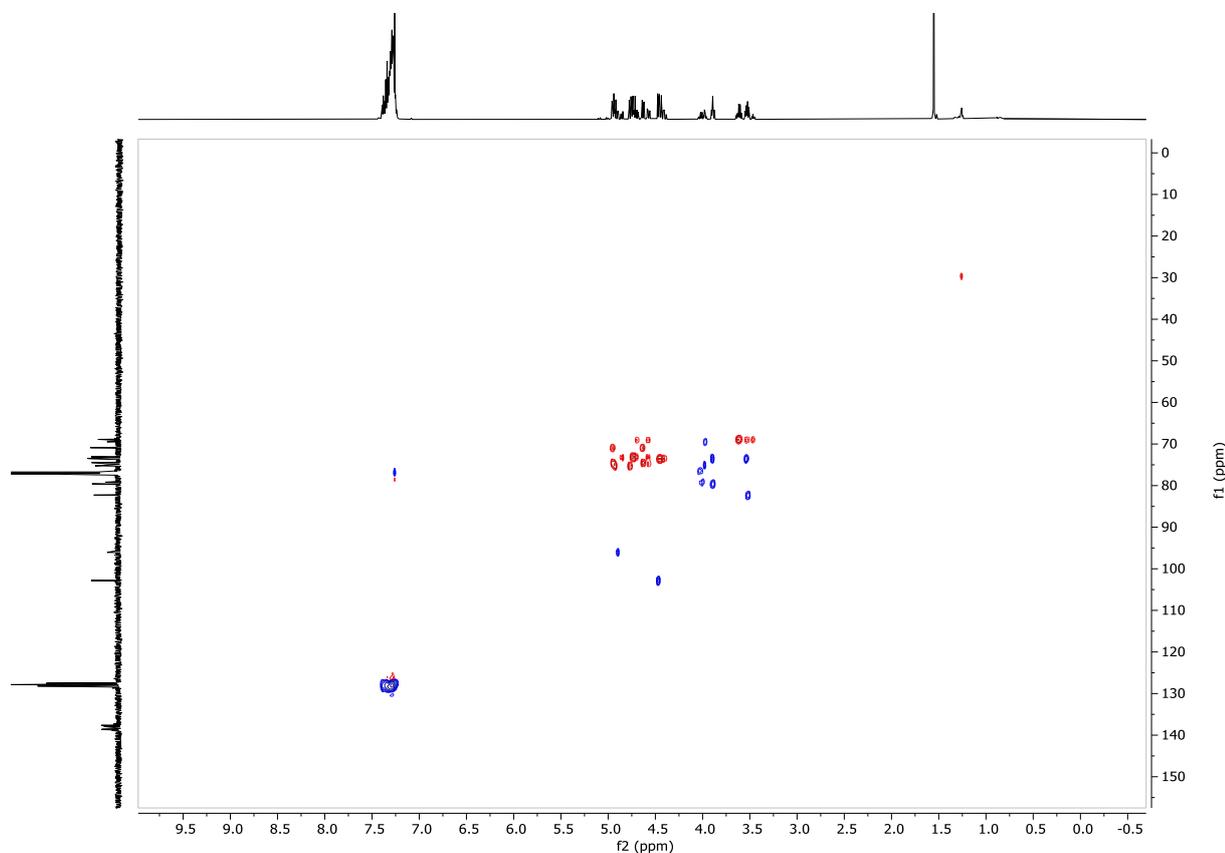
^1H NMR (600 MHz, CDCl_3) of **4Bn-A4**:



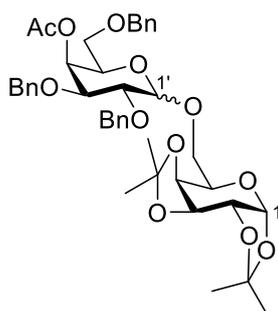
^{13}C NMR (151 MHz, CDCl_3) of **4Bn-A4**:



$^{13}\text{C}, ^1\text{H}$ HSQC of **4Bn-A4**:



1,2:3,4-Bis-*O*-(1-methylethylidene)-6-*O*-[4-*O*-acetyl-2,3,6-tris-*O*-benzyl- α/β -D-galactopyranosyl]- α -D-galactopyranoside (**4Ac-A1**)



The title compound was prepared according to general procedure for glycosylations. Product **4Ac-A1** (21 mg, 28 μmol , 80%, α/β 63:37) was obtained as a colorless oil after purification using **Method-2b** (t_R (α) = 11.5 min, t_R (β) = 12.8 min).

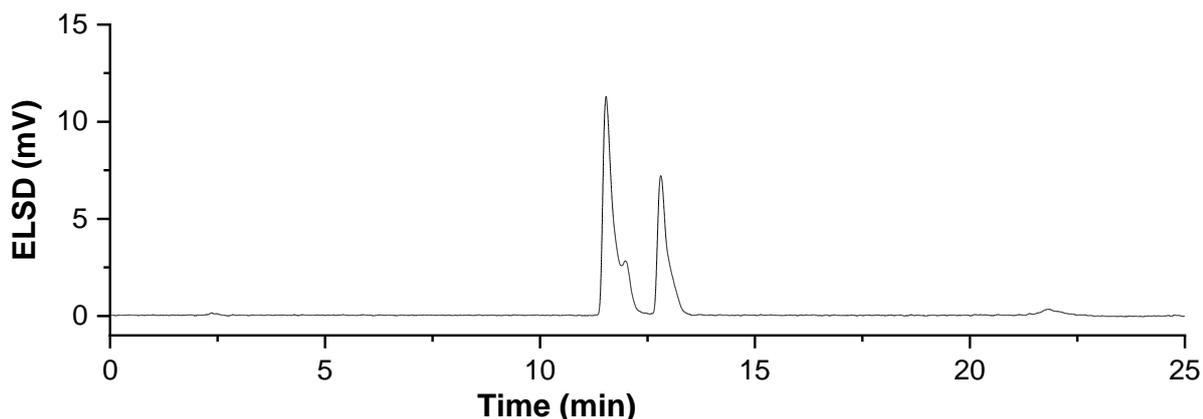
Data of the anomeric mixture:

^1H NMR (600 MHz, CDCl_3) δ 7.43 (d, $J = 6.9$ Hz, 2H), 7.40 – 7.23 (m, 28H), 5.65 (d, $J = 2.8$ Hz, 1H), 5.59 – 5.54 (m, 2H, **H1(α)**), 5.51 (d, $J = 5.0$ Hz, 1H, **H1(α)**), 5.05 – 4.99 (m, 2H, **H1'(α)**), 4.82 – 4.73 (m, 3H), 4.72 – 4.65 (m, 2H), 4.62 (dd, $J = 7.9, 2.4$ Hz, 1H), 4.60 – 4.51 (m, 5H), 4.50 – 4.43 (m, 2H, **H1'(β)**), 4.34 (dd, $J = 5.0, 2.4$ Hz, 1H), 4.33 – 4.29 (m, 2H), 4.28 (dd, $J = 7.9, 1.4$ Hz, 1H), 4.22 (dd, $J = 7.9, 1.9$ Hz, 1H), 4.20 (t, $J = 6.8$ Hz, 1H), 4.14 (dd, $J = 10.8, 3.6$ Hz, 1H), 4.11 – 4.07 (m, 1H), 4.05 – 4.01 (m, 1H), 3.97 (dd, $J = 10.0, 3.4$ Hz, 1H), 3.92 – 3.84 (m, 2H), 3.84 – 3.77 (m, 2H), 3.77 – 3.69 (m, 3H), 3.63 – 3.53 (m, 2H), 3.53 – 3.43 (m, 2H), 2.06 (s, 3H), 2.04 (s, 3H), 1.54 (s, 2H), 1.52 (s, 3H), 1.49 (s, 2H), 1.47 (s, 2H), 1.44 (s, 2H), 1.43 (s, 2H), 1.34 (s, 3H), 1.33 (s, 2H), 1.31 (s, 3H), 1.31 (s, 3H).

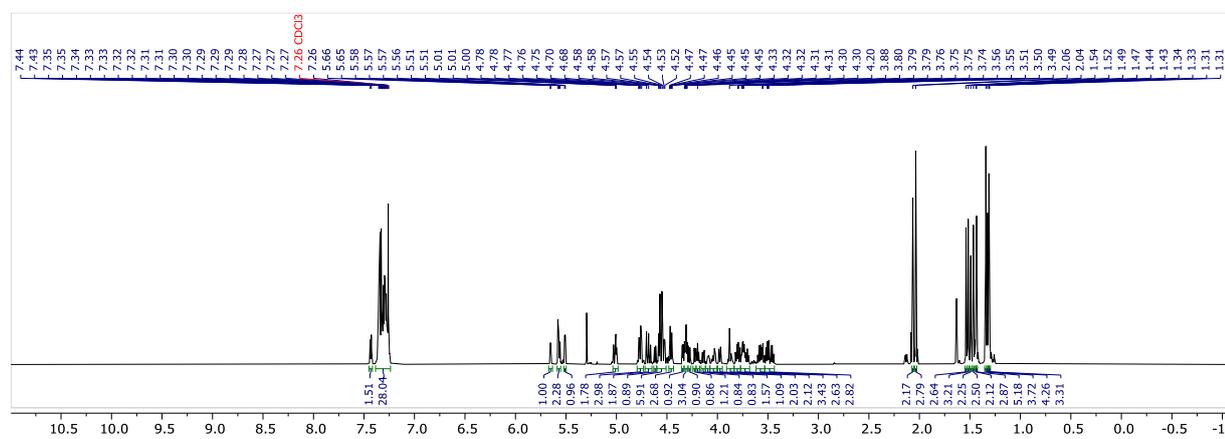
^{13}C NMR (151 MHz, CDCl_3) δ 170.4, 170.3, 138.9, 138.6, 138.2, 137.9, 137.8, 137.6, 128.4, 128.4, 128.4, 128.3, 128.3, 128.3, 128.2, 128.2, 128.1, 128.0, 128.0, 128.0, 127.9, 127.9, 127.8, 127.7, 127.6, 127.5, 127.5, 127.5, 127.3, 109.4, 109.4, 109.2, 108.7, 108.6, 108.5, 104.6 (**C1'(β)**), 97.8 (**C1'(α)**), 96.3 (**C1(α)**), 96.3 (**C1(α)**), 78.9, 78.5, 77.2, 77.0, 76.8, 76.2, 75.5, 74.9, 73.7, 73.5, 72.9, 72.1, 72.0, 71.6, 71.4, 70.8, 70.7, 70.6, 70.6, 70.5, 70.4, 70.0, 68.2, 68.1, 68.0, 68.0, 67.5, 67.3, 66.9, 66.6, 66.0, 62.4, 26.1, 26.0, 26.0, 26.0, 25.9, 25.0, 24.9, 24.9, 24.6, 24.4, 24.3, 20.9, 20.9.

HRMS (ESI): Calcd for $\text{C}_{41}\text{H}_{50}\text{O}_{12}\text{Na}$ [$\text{M} + \text{Na}$] $^+$ 757.3200; found 757.3296.

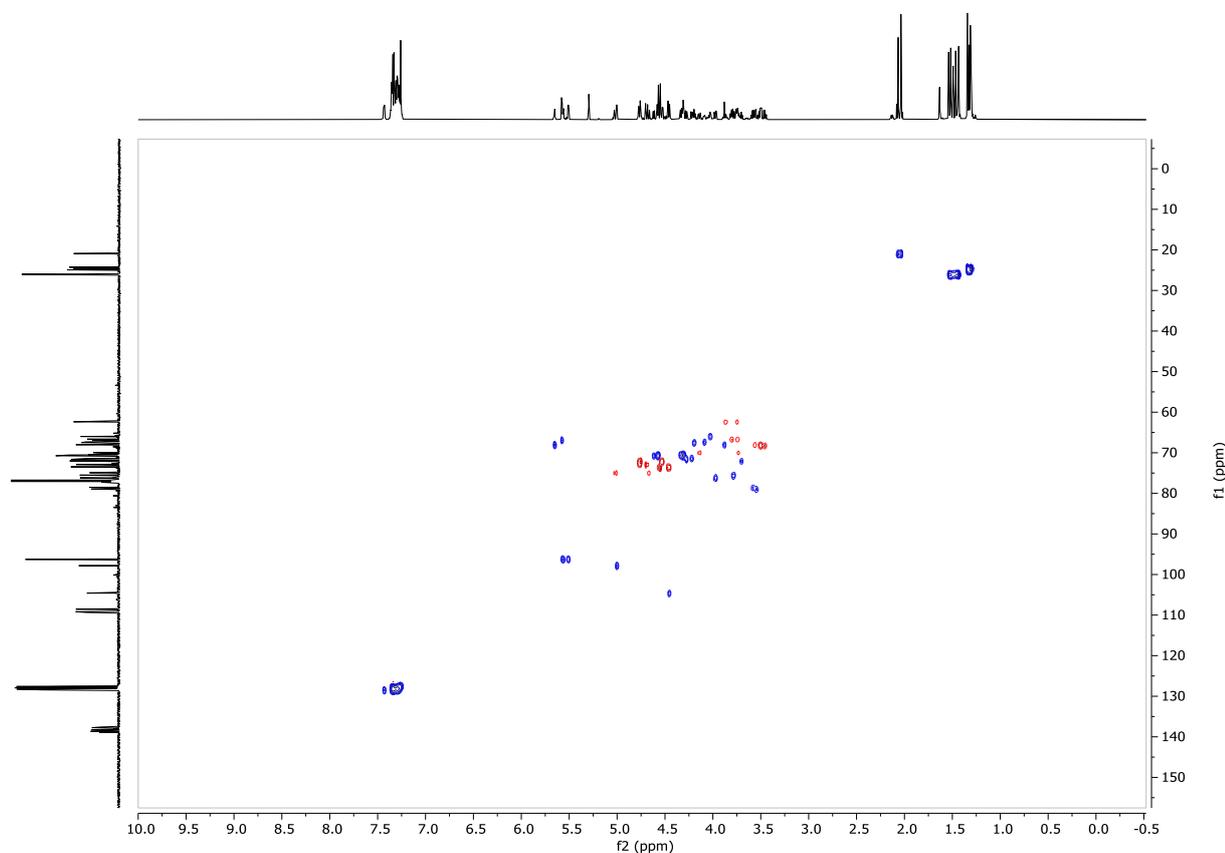
NP-HPLC of **4Ac-A1** (ELSD trace, $t_R(\alpha) = 11.5$ min, $t_R(\beta) = 12.8$ min):



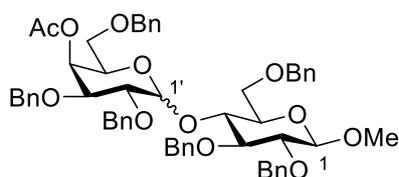
^1H NMR (600 MHz, CDCl_3) of **4Ac-A1**:



$^{13}\text{C}, ^1\text{H}$ HSQC of **4Ac-A1**:



Methyl *O*-[4-*O*-acetyl-2,3,6-tris-*O*-benzyl- α/β -D-galactopyranosyl]-(1 \rightarrow 4)-2,3,6-tri-*O*-benzyl- β -D-glucopyranoside (**4Ac-A2**)



The title compound was prepared according to general procedure for glycosylations. Product **4Ac-A2** (24 mg, 25 μmol , 72%, α/β 60:40) was obtained as a colorless oil after purification using **Method-2b** (t_R (α) = 12.5 min, t_R (β) = 13.0 min).

Data of the anomeric mixture (α -major product):

^1H NMR (600 MHz, CDCl_3) δ 7.50 – 7.24 (m, 53H), 7.22 – 7.10 (m, 10H), 5.77 (d, J = 3.9 Hz, 1H, **H1'**(α)), 5.58 (s, 2H), 4.99 (d, J = 10.5 Hz, 1H), 4.93 (d, J = 11.8 Hz, 1H), 4.89 (d, J = 10.9 Hz, 2H), 4.84 – 4.75 (m,

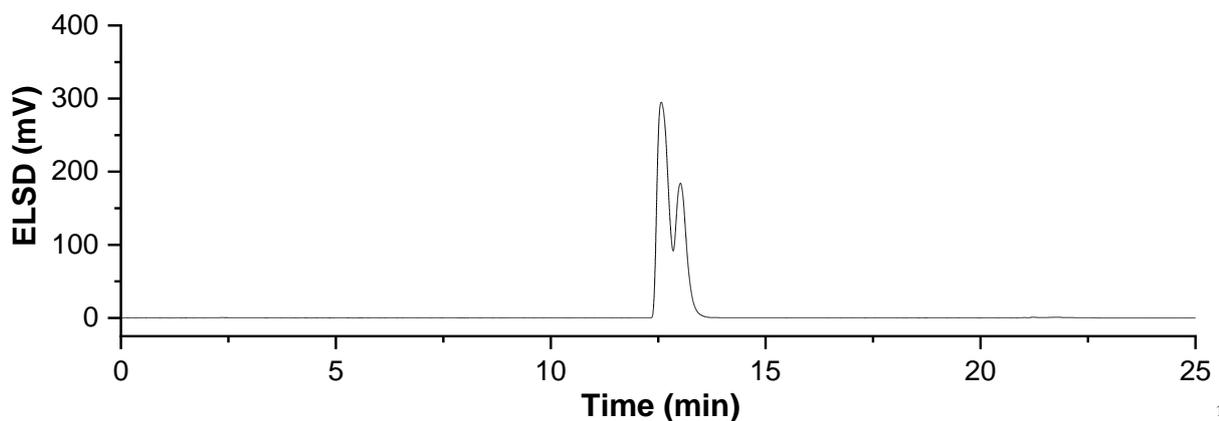
138

3H), 4.75 – 4.67 (m, 4H), 4.66 (d, $J = 12.2$ Hz, 1H), 4.64 – 4.56 (m, 3H), 4.56 – 4.44 (m, 6H, **H1'(β)**), 4.44 – 4.34 (m, 3H, **H1(β)**), 4.33 – 4.25 (m, 3H, **H1(β)**), 4.06 (t, $J = 6.4$ Hz, 1H), 4.04 – 3.96 (m, 2H), 3.85 (dd, $J = 10.3, 3.2$ Hz, 1H), 3.83 – 3.70 (m, 6H), 3.64 – 3.55 (m, 9H), 3.53 – 3.44 (m, 4H), 3.44 – 3.37 (m, 3H), 3.38 – 3.30 (m, 2H), 2.06 (s, 3H), 2.01 (s, 3H) ppm.

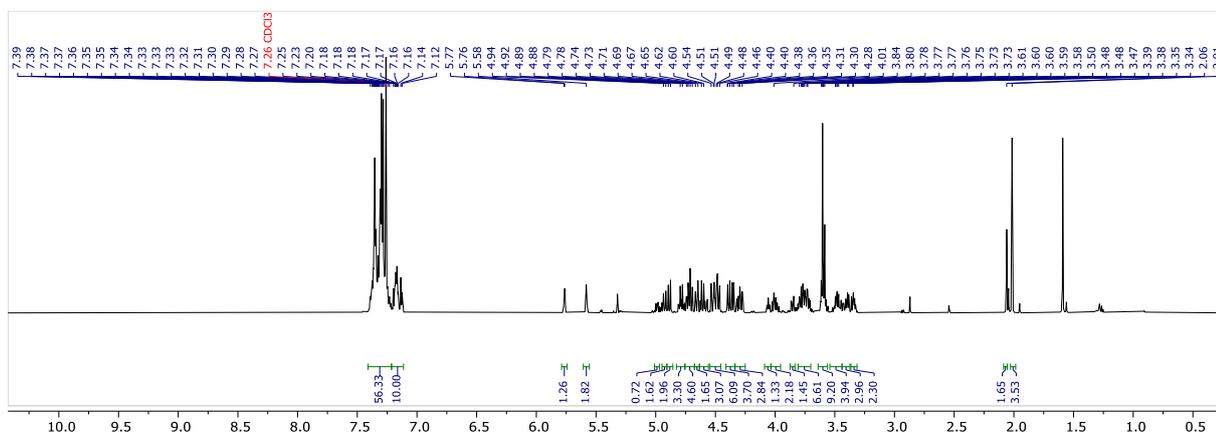
^{13}C NMR (151 MHz, CDCl_3) δ 170.2, 170.2, 139.1, 138.8, 138.7, 138.6, 138.6, 138.4, 138.3, 138.3, 138.2, 138.1, 138.0, 137.9, 137.8, 137.7, 128.4, 128.4, 128.3, 128.3, 128.3, 128.3, 128.2, 128.2, 128.1, 128.1, 128.0, 128.0, 127.8, 127.7, 127.7, 127.7, 127.6, 127.6, 127.6, 127.5, 127.4, 127.0, 126.6, 104.7 (**C1(β)**), 104.5 (**C1(β)**), 102.3 (**C1'(β)**), 97.4 (**C1'(α)**), 84.8, 82.7, 82.4, 81.8, 79.8, 79.4, 77.2, 77.0, 76.8, 76.5, 76.4, 75.3, 75.2, 75.0, 74.9, 74.6, 74.6, 74.2, 73.9, 73.9, 73.6, 73.5, 73.3, 73.1, 72.8, 71.8, 71.8, 71.6, 71.6, 69.5, 68.1, 68.0, 67.7, 57.1, 57.0, 20.9 ppm.

HRMS (ESI): Calcd for $\text{C}_{57}\text{H}_{62}\text{O}_{12}\text{Na}$ $[\text{M} + \text{Na}]^+$ 961.4139; found 961.4299.

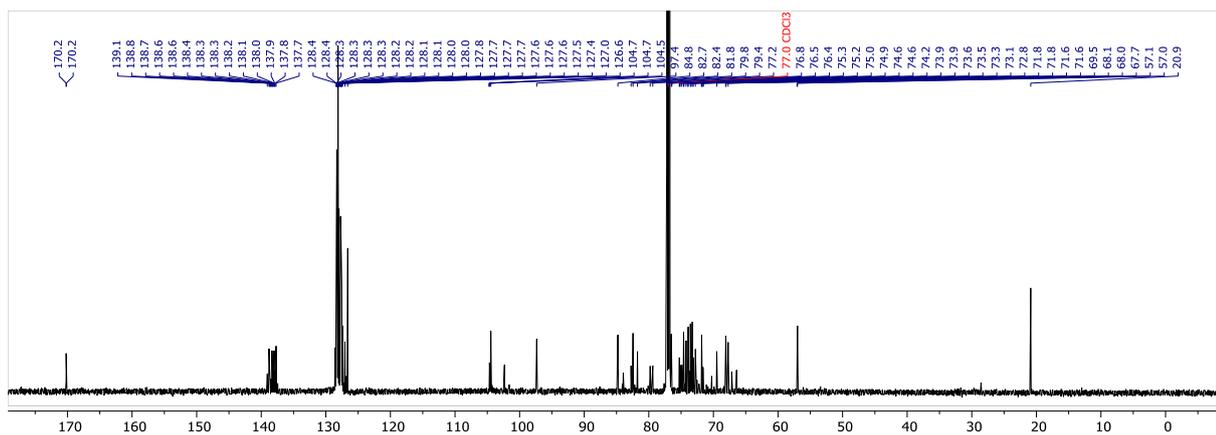
NP-HPLC of **4Ac-A2** (ELSD trace, $t_R(\alpha) = 12.5$ min, $t_R(\beta) = 13.0$ min):



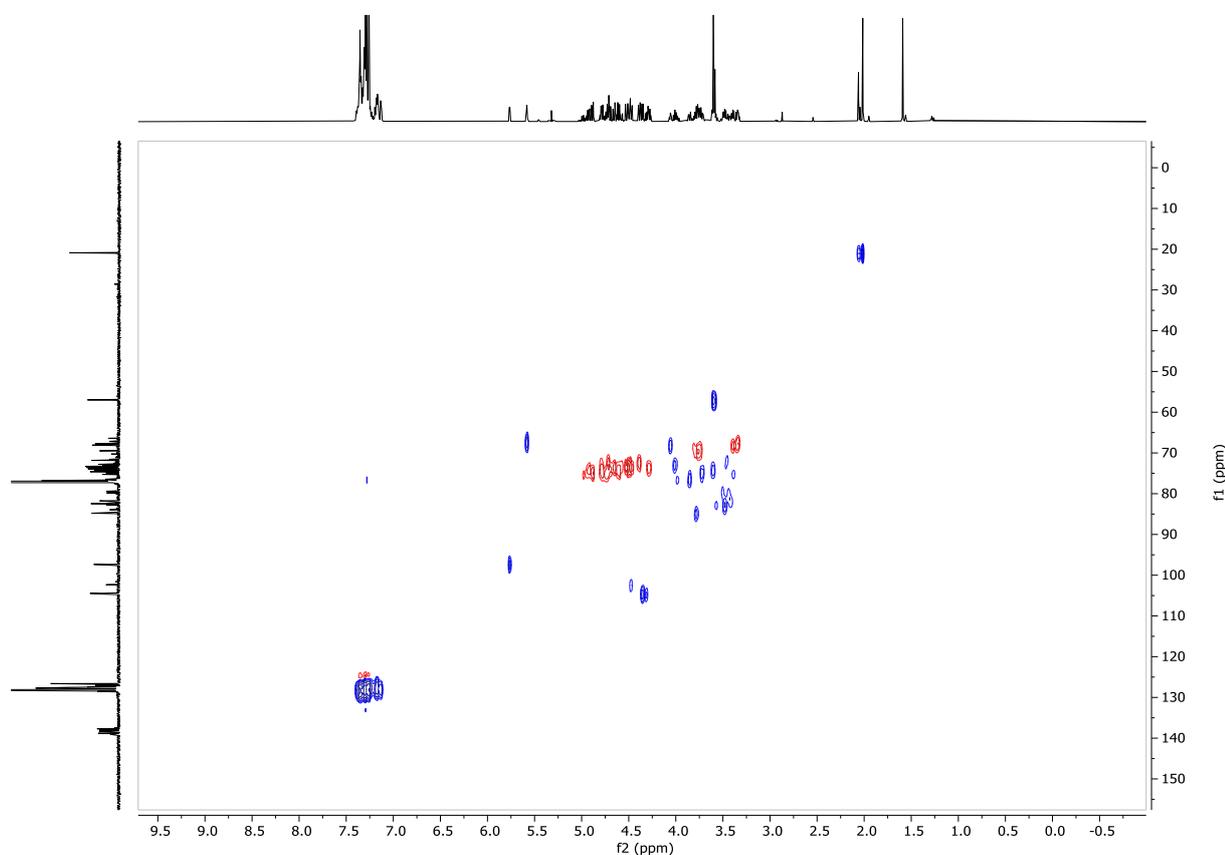
¹H NMR (600 MHz, CDCl₃) of **4Ac-A2**:



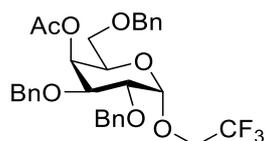
¹³C NMR (151 MHz, CDCl₃) of **4Ac-A2**:



$^{13}\text{C}, ^1\text{H}$ HSQC of **4Ac-A2**:



2,2,2-Trifluoroethyl-4-*O*-Acetyl-2,3,6-tri-*O*-benzyl- α -D-galactopyranoside (4Ac-A3)



The title compound was prepared according to general procedure for glycosylations. Product **4Ac-A3** (13 mg, 23 μmol , 66%, α/β 100:0) was obtained as a colorless oil after purification using **Method-2b** (t_R (α) = 8.2 min).

Data of the α -anomer:

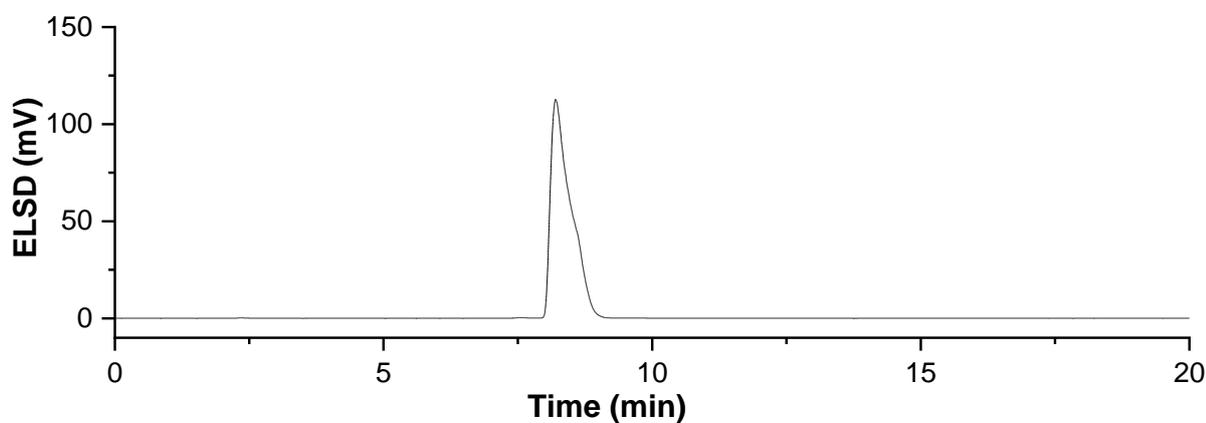
^1H NMR (600 MHz, CDCl_3) δ 7.38 – 7.26 (m, 15H), 5.63 (d, $J = 2.7$ Hz, 1H), 4.90 (d, $J = 3.7$ Hz, 1H, **H1**), 4.83 (d, $J = 11.9$ Hz, 1H), 4.75 (d, $J = 10.9$ Hz, 1H), 4.62 (d, $J = 11.9$ Hz, 1H), 4.55 (dd, $J = 11.4, 3.7$ Hz, 2H),

4.45 (d, $J = 11.9$ Hz, 1H), 4.08 (t, $J = 6.3$ Hz, 1H), 3.96 (dd, $J = 10.1, 3.4$ Hz, 1H), 3.92 (m, 3H), 3.80 (dd, $J = 10.1, 3.7$ Hz, 1H), 3.47 (m, 2H), 2.05 (s, 3H).

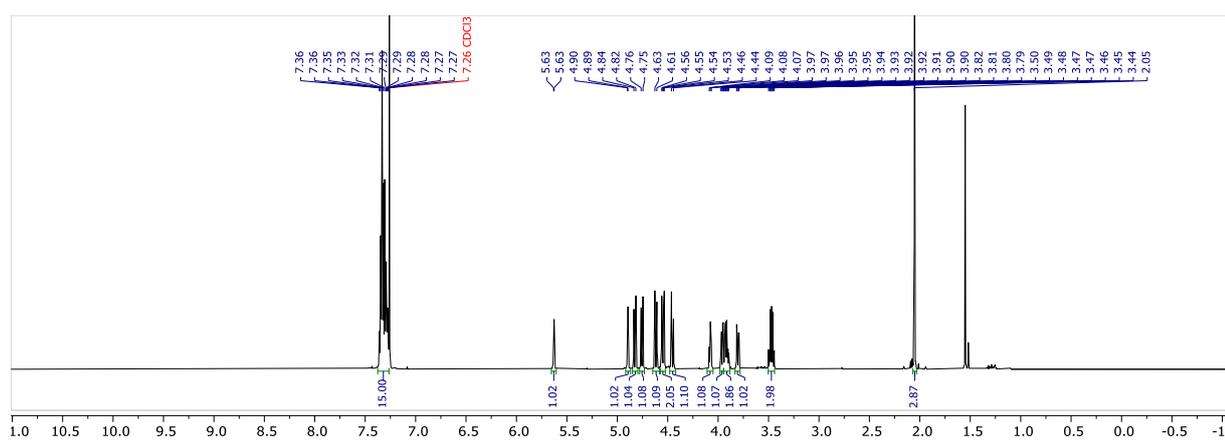
^{13}C NMR (151 MHz, CDCl_3) δ 170.2, 138.4, 138.0, 137.6, 128.4, 128.4, 128.3, 128.0, 127.8, 127.8, 127.6, 98.4 (C1), 77.2, 77.0, 76.8, 75.8, 75.1, 73.6, 73.5, 72.2, 68.5, 68.3, 67.8, 64.8, 64.6, 20.9.

HRMS (ESI): Calcd for $\text{C}_{31}\text{H}_{33}\text{F}_3\text{O}_7\text{Na}$ $[\text{M} + \text{Na}]^+$ 597.2076; found 597.2153.

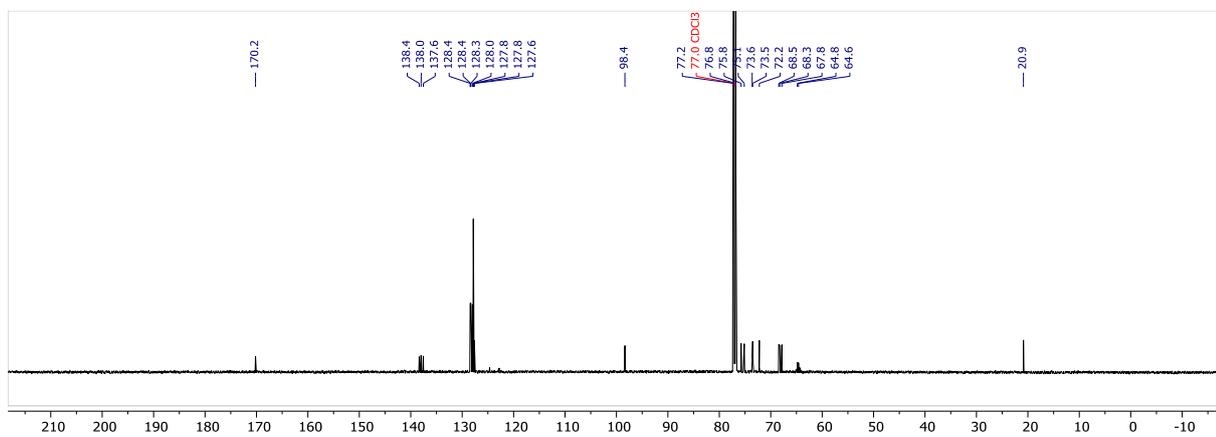
NP-HPLC of **4Ac-A3** (ELSD trace, $t_R(\alpha) = 8.2$ min):



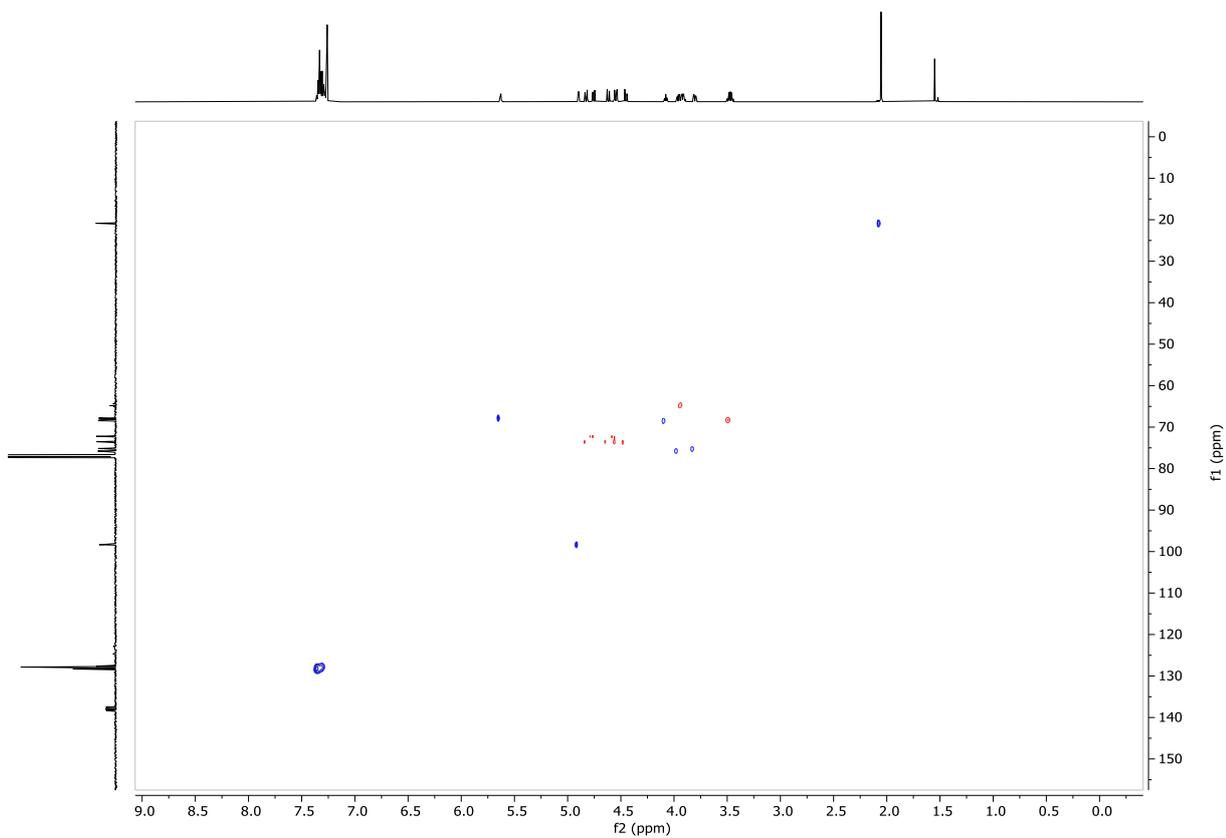
^1H NMR (600 MHz, CDCl_3) of **4Ac-A3**:



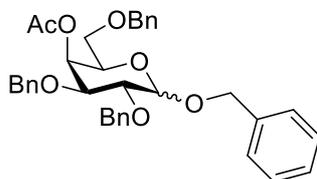
^{13}C NMR (151 MHz, CDCl_3) of **4Ac-A3**:



$^{13}\text{C}, ^1\text{H}$ HSQC of **4Ac-A3**:



4-*O*-Acetyl-2,3,6-tris-*O*-benzyl- α/β -D-galactopyranosyl (**4Ac-A4**)



The title compound was prepared according to general procedure for glycosylations. Product **4Ac-A4** (14 mg, 24 μ mol, 68%, α/β 50:50) was obtained as a colorless oil after purification using **Method-2b** (tR (α) = 7.9 min, tR (β) = 8.8 min).

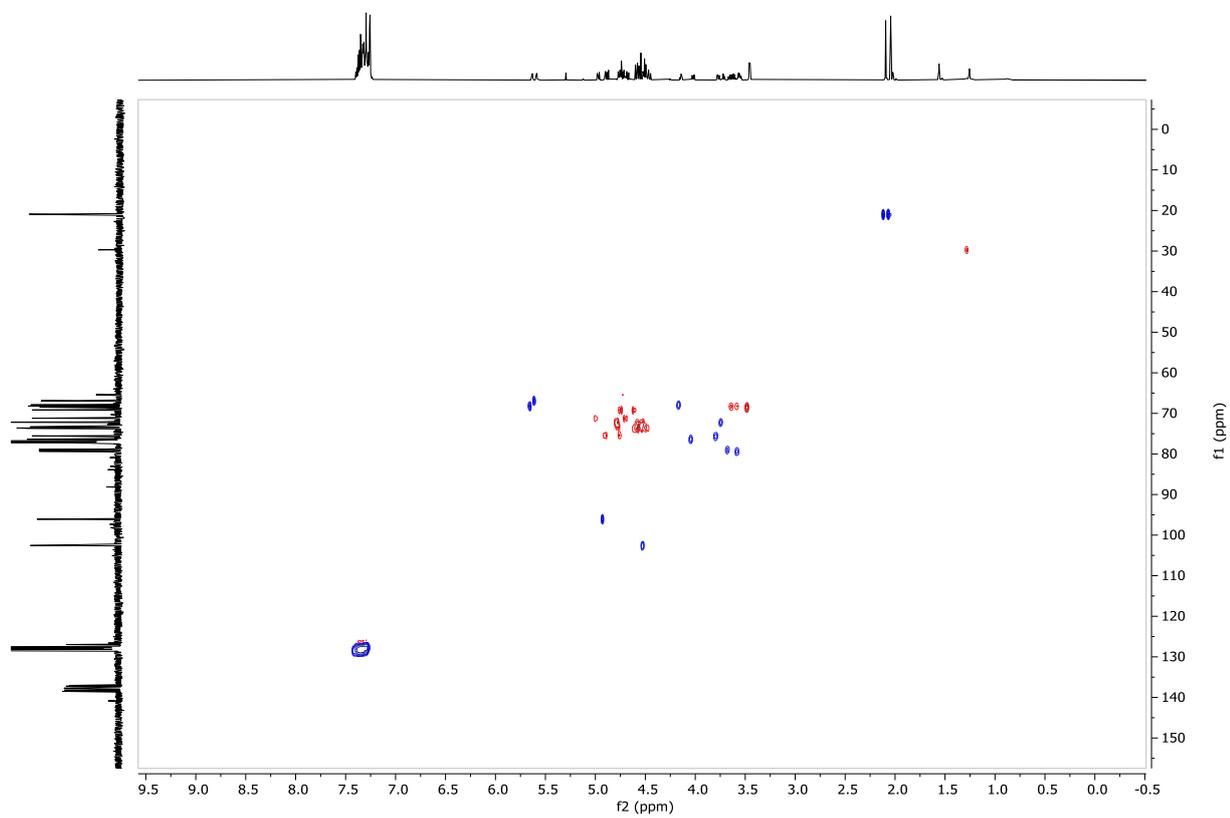
Data of the anomeric mixture:

¹H NMR (600 MHz, CDCl₃) δ 7.45 – 7.21 (m, 46H), 5.63 (d, J = 2.9 Hz, 1H), 5.59 (d, J = 3.1 Hz, 1H), 4.97 (d, J = 12.0 Hz, 1H), 4.90 (d, J = 3.8 Hz, 1H, **H1(α)**), 4.88 (d, J = 10.7 Hz, 1H), 4.82 – 4.65 (m, 5H), 4.62 – 4.42 (m, 7H, **H1(β)**, CH₂Ph), 4.14 (t, J = 6.3 Hz, 1H), 4.02 (dd, J = 10.0, 3.4 Hz, 1H), 3.77 (dd, J = 10.0, 3.8 Hz, 1H), 3.72 (t, J = 6.3 Hz, 1H), 3.64 (ddd, J = 18.3, 9.6, 6.9 Hz, 2H), 3.56 (ddd, J = 9.5, 5.0, 1.4 Hz, 2H), 3.46 (s, 1H), 3.45 (s, 1H), 2.09 (s, 3H), 2.04 (s, 3H) ppm.

¹³C NMR (151 MHz, CDCl₃) δ 170.4, 170.3, 138.5, 138.2, 137.8, 137.7, 137.6, 137.2, 137.0, 128.6, 128.5, 128.4, 128.4, 128.4, 128.3, 128.3, 128.3, 128.2, 128.1, 128.0, 128.0, 127.9, 127.9, 127.8, 127.8, 127.7, 127.7, 127.5, 127.0, 102.6 (**C1(β)**), 96.1 (**C1(α)**), 79.3, 78.9, 77.2, 77.0, 76.8, 76.4, 75.5, 75.4, 73.7, 73.6, 73.3, 72.2, 72.1, 71.2, 69.2, 68.5, 68.2, 68.2, 67.9, 66.9, 65.4, 29.7, 21.0, 20.9 ppm.

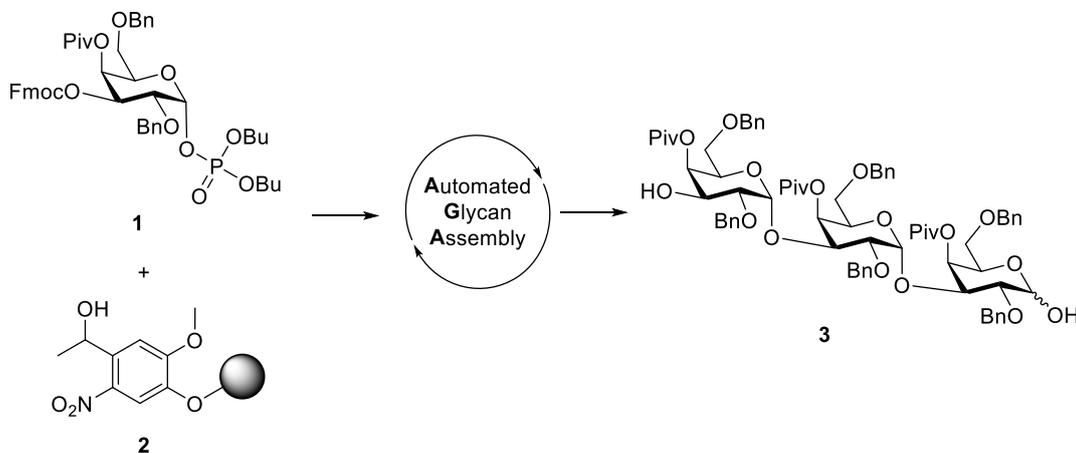
HRMS (ESI): Calcd for C₃₆H₃₈O₇Na [M + Na]⁺ 605.2515; found 605.2466.

$^{13}\text{C}, ^1\text{H}$ HSQC of 4Ac-A4:



7 Automated Glycan Assembly of Building Block 1

α -(1 \rightarrow 3)-D-Trigalactopyranoside (**3**)



Repeat	Building Blocks	Modules	Notes
		I – Acidic Wash	
2x	1 (2 x 4.7 equiv.)	IIb – Glycosylation with glycosyl phosphate	-35 °C (T ₁) 5min (t ₁) -20 °C (T ₂) 50 min (t ₂)
		III – Capping IVc – Fmoc Deprotection	
		I – Acidic Wash	
1x	1 (9.4 equiv.)	IIb – Glycosylation with glycosyl phosphate – 2 cycles	-35 °C (T ₁) 5min (t ₁) -20 °C (T ₂) 50 min (t ₂)
		III – Capping IVc – Fmoc Deprotection	

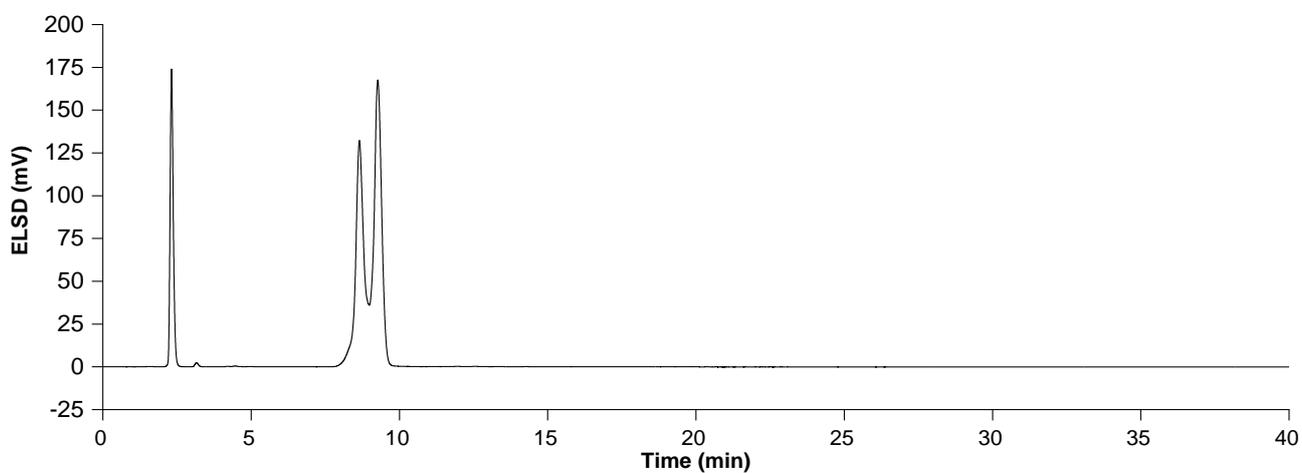
Trisaccharide **3** (12.5 mg, 0.014 mmol, 69%, α -only) was obtained as a colorless oil after photocleavage from solid support following **Method A** and purification by normal-phase HPLC (**Method B-1b**, t_R = 9.1 to 9.8 min).

¹H NMR (700 MHz, CDCl₃) δ 7.40 – 7.20 (m, 30H), 5.54 (d, J = 3.2 Hz, 1H), 5.27 (dd, J = 6.6, 3.5 Hz, 3H), 5.23 (d, J = 3.4 Hz, 1H), 5.12 (d, J = 3.4 Hz, 1H), 4.77 – 4.70 (m, 2H), 4.59 – 4.56 (m, 3H), 4.54 – 4.42 (m, 5H), 4.38 (s, 1H), 4.36 – 4.32 (m, 4H), 4.30 – 4.25 (m, 3H), 3.91 (dd, J = 10.1, 3.5 Hz, 1H), 3.84 (dd, J = 10.3, 3.4 Hz, 1H), 3.78 (dd, J = 10.1, 3.5 Hz, 1H), 3.57 – 3.26 (m, 6H), 1.15 (s, 9H), 1.09 (s, 9H), 1.02 (s, 9H) ppm.

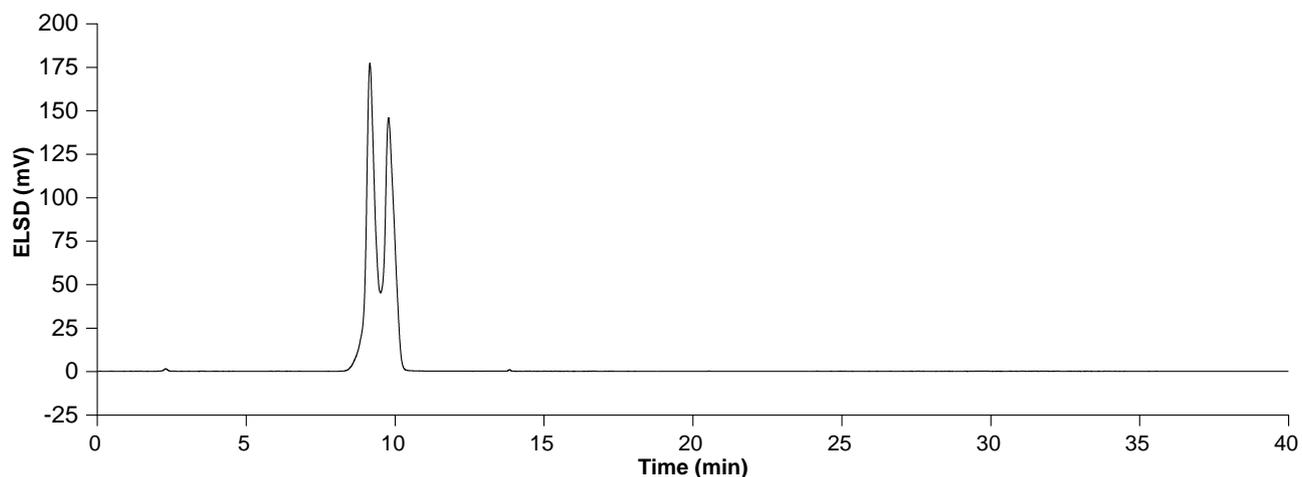
^{13}C NMR (176 MHz, CDCl_3) δ 178.1, 177.4, 177.3, 138.3, 138.2, 138.1, 138.0, 137.9, 137.6, 128.7, 128.6, 128.5, 128.5, 128.5, 128.4, 128.3, 128.0, 128.0, 127.9, 127.7, 127.5, 93.1, 92.9, 91.9, 75.3, 74.7, 73.7, 73.2, 73.0, 72.7, 72.3, 70.5, 70.1, 69.8, 69.0, 68.6, 68.5, 68.3, 68.1, 68.0, 67.5, 66.8, 66.3, 39.2, 39.1, 39.0, 27.3, 27.2, 27.2 ppm.

HRMS (QToF): Calcd for $\text{C}_{75}\text{H}_{92}\text{O}_{19}\text{Na}$ $[\text{M} + \text{Na}]^+$ 1319.6125; found 1319.6166.

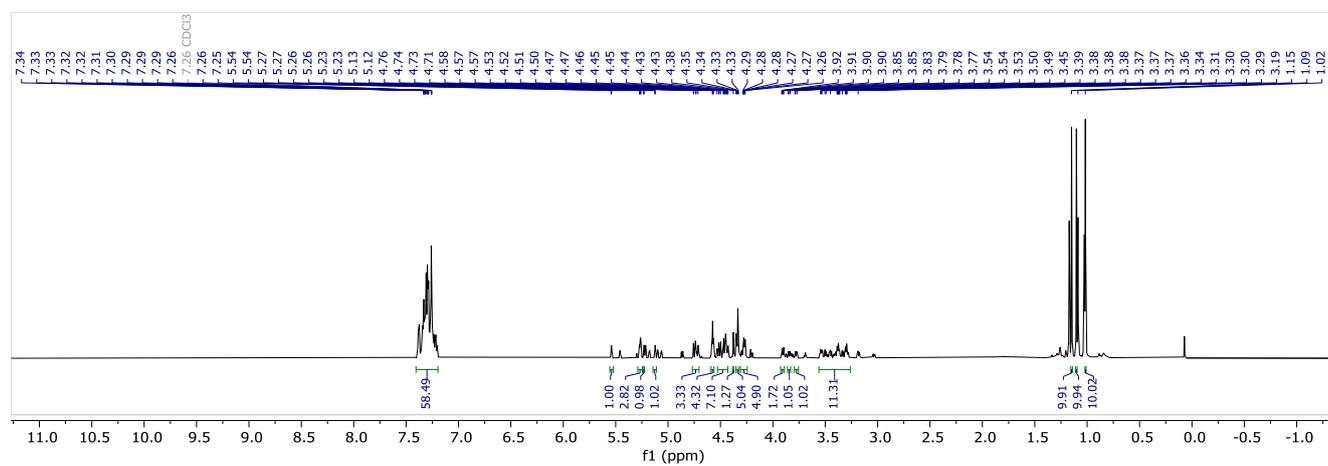
Crude NP-HPLC of **3** (ELSD trace):



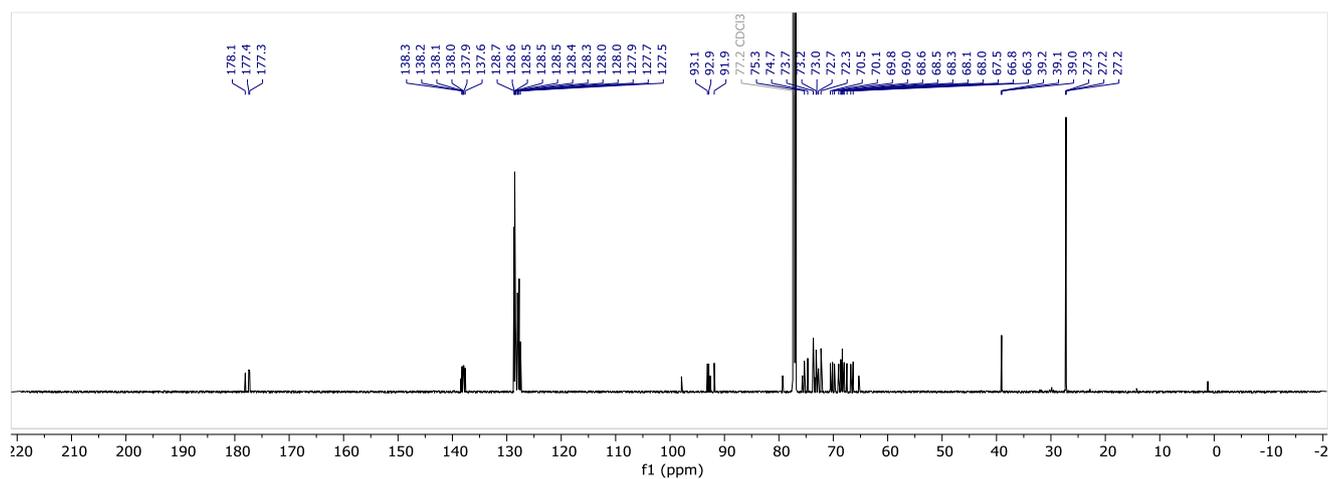
Purified NP-HPLC of **3** (ELSD trace, $t_R(\alpha) = 9.1$ min, $t_R(\beta) = 9.8$ min):



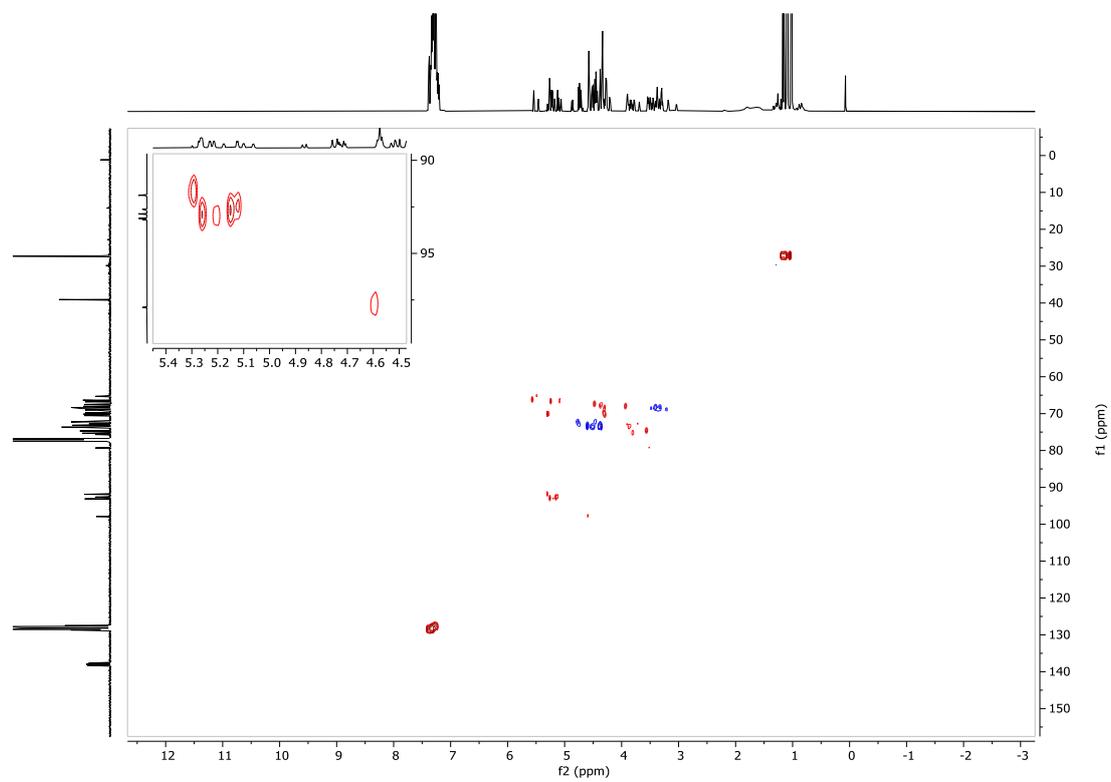
^1H NMR (700 MHz, CDCl_3) of **3** (mutarotates):



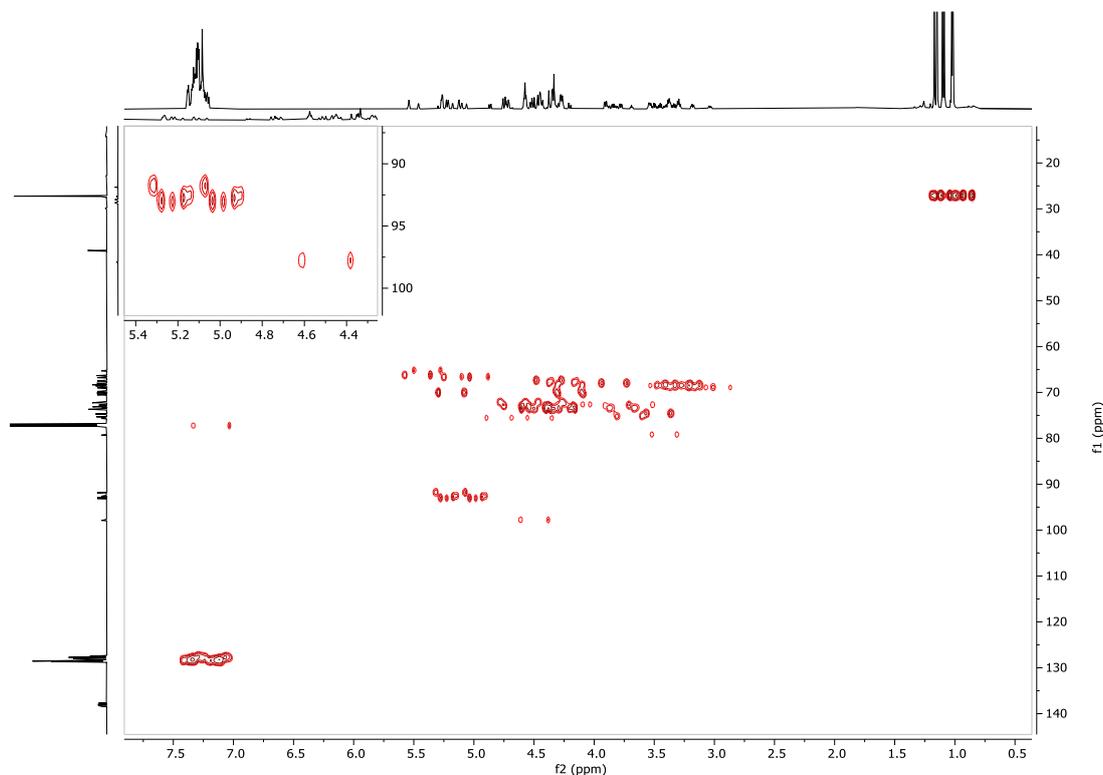
^{13}C NMR (176 MHz, CDCl_3) of **3** (mutarotates):



$^{13}\text{C},^1\text{H}$ HSQC of **3**:



Coupled $^{13}\text{C}, ^1\text{H}$ HSQC of **3**:



8 Literature

1. Schöllkopf, W.; Gewinner, S.; Junkes, H.; Paarmann, A.; von Helden, G.; Bluem, H. P.; Todd, A. M. M., The new IR and THz FEL facility at the Fritz Haber Institute in Berlin. *Proc. SPIE Int. Soc. Opt. Eng.* **2015**, 9512, 95121L.
2. Supady, A.; Blum, V.; Baldauf, C., First-Principles Molecular Structure Search with a Genetic Algorithm. *J. Chem. Inf. Model.* **2015**, 55 (11), 2338-48.
3. Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M., Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **2009**, 180 (11), 2175-2196.
4. Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, 77 (18), 3865-3868.
5. Tkatchenko, A.; Scheffler, M., Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.* **2009**, 102 (7), 073005.
6. Adamo, C.; Barone, V., Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, 110 (13), 6158-6170.
7. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H., A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, 132 (15), 154104.
8. Hehre, W. J.; Ditchfield, R.; Pople, J. A., Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, 56 (5), 2257-2261.
9. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.;

- Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16 Rev. A.03*, Wallingford, CT, 2016.
10. Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F., Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *J. Chem. Phys.* **2013**, *139* (13), 134101.
 11. Weigend, F.; Ahlrichs, R., Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297-305.
 12. Neese, F., Software update: The ORCA program system—Version 5.0. *WIREs Comput. Mol. Sci.* **2022**.
 13. Le Mai Hoang, K.; Pardo-Vargas, A.; Zhu, Y.; Yu, Y.; Loria, M.; Delbianco, M.; Seeberger, P. H., Traceless Photolabile Linker Expedites the Chemical Synthesis of Complex Oligosaccharides by Automated Glycan Assembly. *Journal of the American Chemical Society* **2019**, *141* (22), 9079-9086.
 14. Gude, M.; Ryf, J.; White, P. D., An accurate method for the quantitation of Fmoc-derivatized solid phase supports. *Lett. Pept. Sci.e* **2002**, *9* (4-5), 203-206.
 15. Hahm, H. S.; Hurevich, M.; Seeberger, P. H., Automated assembly of oligosaccharides containing multiple cis-glycosidic linkages. *Nat. Commun.* **2016**, *7*, 12482.
 16. Chatterjee, S.; Moon, S.; Hentschel, F.; Gilmore, K.; Seeberger, P. H., An Empirical Understanding of the Glycosylation Reaction. *J. Am. Chem. Soc.* **2018**, *140* (38), 11942-11953.
 17. Alex, C.; Visansirikul, S.; Demchenko, A. V., A versatile approach to the synthesis of mannosamine glycosides. *Org. Biomol. Chem.* **2020**, *18* (34), 6682-6695.
 18. Scanlan, E. M.; Mackeen, M. M.; Wormald, M. R.; Davis, B. G., Synthesis and Solution-Phase Conformation of the RG-I Fragment of the Plant Polysaccharide Pectin Reveals a Modification-Modulated Assembly Mechanism. *J. Am. Chem. Soc.* **2010**, *132* (21), 7238-7239.
 19. Daragics, K.; Fügedi, P., Regio- and chemoselective reductive cleavage of 4,6-O-benzylidene-type acetals of hexopyranosides using BH₃·THF–TMSOTf. *Tetrahedron Lett.* **2009**, *50* (24), 2914-2916.
 20. Sherman, A. A.; Mironov, Y. V.; Yudina, O. N.; Nifantiev, N. E., The presence of water improves reductive openings of benzylidene acetals with trimethylaminoborane and aluminium chloride. *Carbohydr. Res.* **2003**, *338* (8), 697-703.
 21. Tian, G.; Hu, J.; Qin, C.; Li, L.; Zou, X.; Cai, J.; Seeberger, P. H.; Yin, J., Chemical Synthesis and Immunological Evaluation of Helicobacter pylori Serotype O6 Tridecasaccharide O-Antigen Containing a dd-Heptoglycan. *Angew. Chem. Int. Ed.* **2020**, *59* (32), 13362-13370.
 22. Hofmann, J.; Hahm, H. S.; Seeberger, P. H.; Pagel, K., Identification of carbohydrate anomers using ion mobility–mass spectrometry. *Nature* **2015**, *526* (7572), 241-244.

Appendix B

Neighboring-Group Participation of Benzoyl Protecting Groups in Fluorinated Glucose

This appendix contains the supporting information of the publication related to Chapter 5. It contains mass spectra, energetics, energy hierarchies, and 3D-structures of glucosyl cations. Additionally the synthesis of a building block that has previously not been published is described as well as the characterization of that compound. The publication, the supporting information, and xyz-coordinates of the computed structures can be found online:

<https://doi.org/10.1002/ejoc.202200255>.^[297]

Contents

Experimental Details	2
Mass Spectrometry and Infrared Spectroscopy	2
Mass Spectra	4
Computational Methods	7
Energetics	9
Low-Energy Structures	16
xyz-Coordinates of Reoptimized Structures	20
Synthesis of Building Blocks	21
References	29

Experimental Details

Mass Spectrometry and Infrared Spectroscopy

Infrared (IR) and mass spectra (Figures 1-2 and S2-S6) of glycosyl cations formed by in-source fragmentation of precursor ions were obtained using a custom setup that combines mass spectrometry with IR spectroscopy in helium droplets (Figure S1). The instrument is described in the following paragraphs. The precursor building blocks were dissolved in a 9:1 (V:V) mixture of acetonitrile and deionized water to yield 0.1 mM solutions. The solutions were ionized *via* nanoelectrospray ionization (nESI) using Pd/Pt coated glass capillaries (Sputter Coater HR 208, *Cressington*), pulled to a tip with an inner diameter of 1-2 μm with a micropipette puller (Model P-1000, *Sutter Instrument*). Bare glycosyl cations are generated by applying a voltage of 1 kV to the tip of the capillary using a Z-spray source.

The generated beam of ions traverses two ring-electrode ion guides and a quadrupole mass filter that allows mass-to-charge selection of the ions of interest. Mass-selected ions are then guided into a hexapole ion trap by a quadrupole bender. The trapped ions are thermalized by collisions with helium buffer gas. The ion trap is cooled to ca. 90 K by liquid nitrogen.

Superfluid helium droplets, generated by a pulsed Even-Lavie valve, traverse the ion trap to pick up ions and rapidly cool them to 0.4 K. They guide the embedded ions to a detection region, where vibrational modes of the embedded ions are excited by an IR beam of the Fritz Haber Institute free-electron laser (FHI FEL^[1]). Absorption of multiple resonant photons eventually leads to the release of the ions from the helium droplet and subsequent detection by a time-of-flight detector. Plotting the ion signal as a function of the IR wavenumber yields an IR spectrum. The observed intensities scale non-linearly with the energy of the IR beam because of the multiphoton absorption process. A first-order correction is performed by dividing the ion signal by the energy of the IR macropulse.

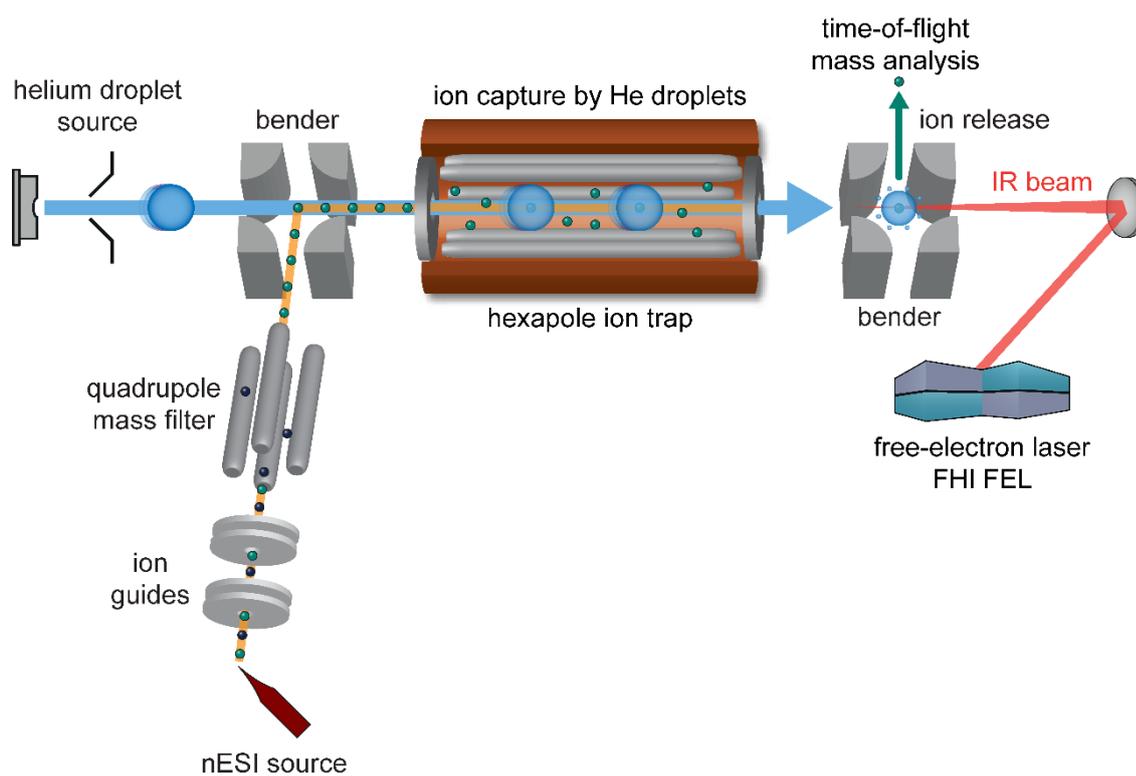


Figure S1. Schematic representation of the helium droplet instrument that combines mass spectrometry with infrared spectroscopy used to generate glycosyl cations and probe their vibrational modes.

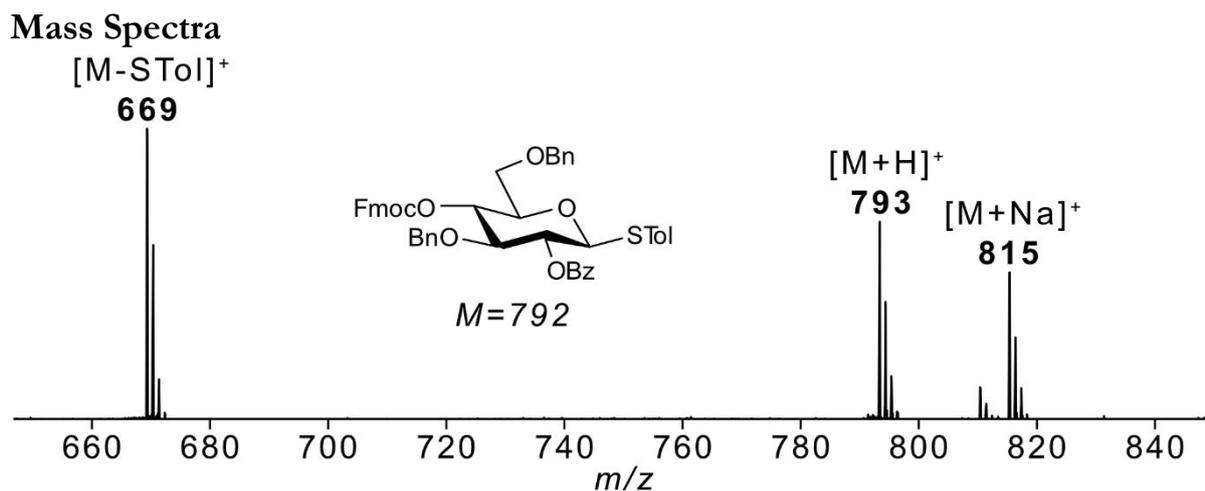


Figure S2. Mass spectrum of 2-*O*-benzoyl-3,6-di-*O*-benzyl-4-*O*-fluorenylmethoxycarbonyl-*D*-glucopyranoside (**Glc1**) generated from β -thiotolyl precursor recorded on the helium droplet instrument. In-source fragmentation of the precursor ions $[M+H]^+$ ($m/z = 793$), $[M+NH_4]^+$ ($m/z = 810$) and $[M+Na]^+$ ($m/z = 815$) leads to glycosyl cations ($m/z = 669$).

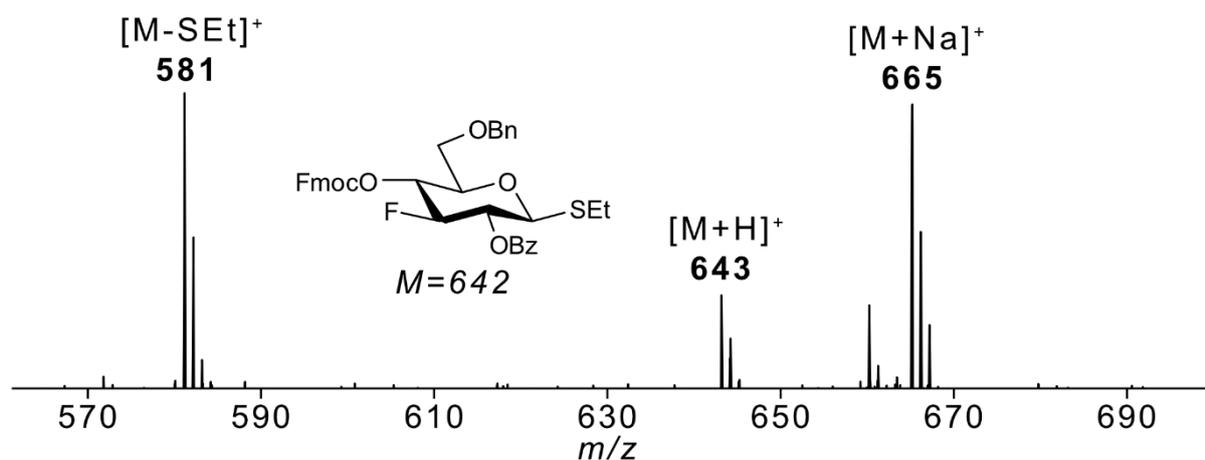


Figure S3. Mass spectrum of 2-*O*-benzoyl-3-*F*-4-*O*-fluorenylmethoxycarbonyl-6-*O*-benzyl-*D*-glucopyranoside (**3F-Glc1**) generated from β -thioethyl precursor recorded on the helium droplet instrument. In-source fragmentation of the precursor ions $[M+H]^+$ ($m/z = 643$), $[M+NH_4]^+$ ($m/z = 660$) and $[M+Na]^+$ ($m/z = 665$) leads to glycosyl cations ($m/z = 581$).

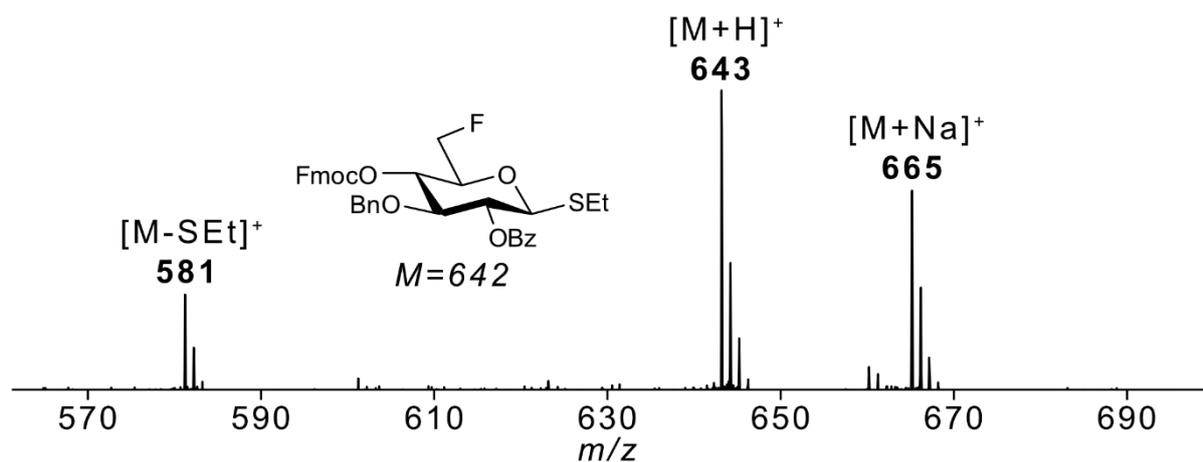


Figure S4. Mass spectrum of 2-*O*-benzoyl-3-*O*-benzyl-4-*O*-fluorenylmethoxycarbonyl-6-*F*-*D*-glucopyranoside (**6F-Glc1**) generated from β -thioethyl precursor recorded on the helium droplet instrument. In-source fragmentation of the precursor ions $[M+H]^+$ ($m/z = 643$), $[M+NH_4]^+$ ($m/z = 660$) and $[M+Na]^+$ ($m/z = 665$) leads to glycosyl cations ($m/z = 581$).

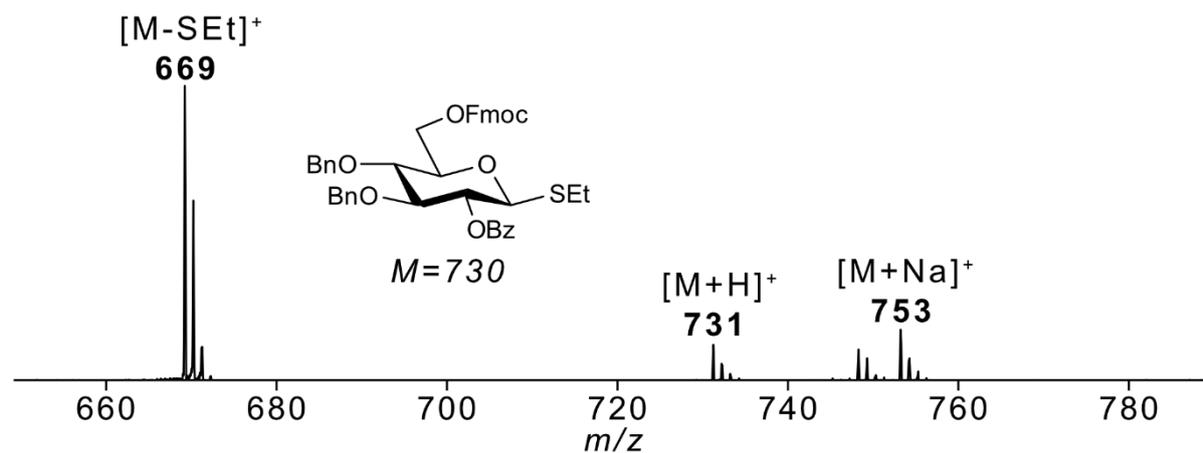


Figure S5. Mass spectrum of 2-*O*-benzoyl-3,4-di-*O*-benzyl-6-*O*-fluorenylmethoxycarbonyl-*D*-glucopyranoside (**Glc2**) generated from β -thioethyl precursor recorded on the helium droplet instrument. In-source fragmentation of the precursor ions $[M+H]^+$ ($m/z = 731$), $[M+NH_4]^+$ ($m/z = 748$) and $[M+Na]^+$ ($m/z = 753$) leads to glycosyl cations ($m/z = 669$).

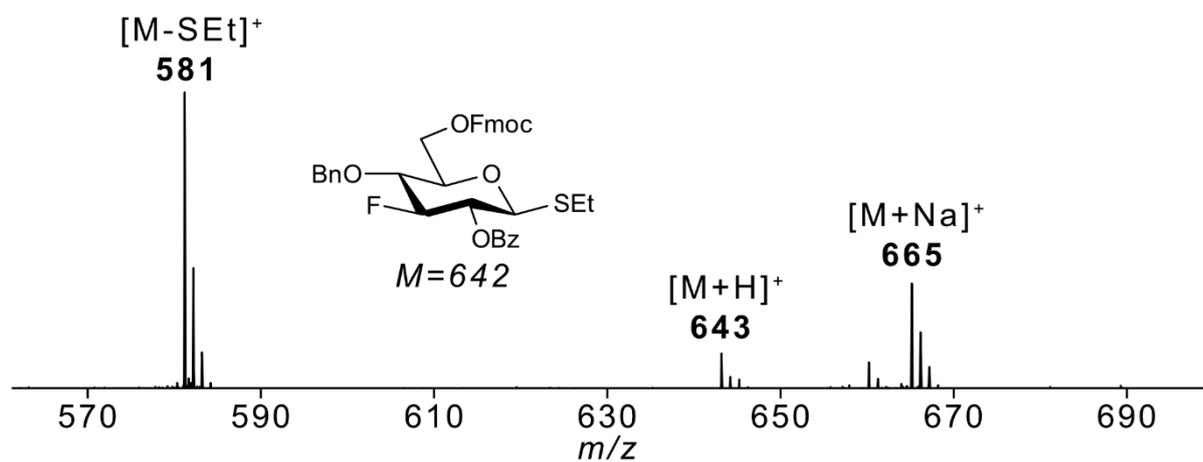


Figure S6. Mass spectrum of 2-*O*-benzoyl-3-*F*-4-*O*-benzyl-6-*O*-fluorenylmethoxycarbonyl- β -D-glucopyranoside (**3F-Glc2**) generated from β -thioethyl precursor recorded on the helium droplet instrument. In-source fragmentation of the precursor ions $[M+H]^+$ ($m/z = 643$), $[M+NH_4]^+$ ($m/z = 660$) and $[M+Na]^+$ ($m/z = 665$) leads to glycosyl cations ($m/z = 581$).

Computational Methods

The conformational space of fluorinated glycosyl cations and that of their non-fluorinated counterparts was sampled using the genetic algorithm (GA) FAFOOM.^[2] It allows sampling of all rotatable bonds and pyranose puckers. In this study, an interface of the GA with ORCA 4.1.1^[3] was used for geometry optimization of each generated structure at the PBE/def2-SVP^[4] level of theory. For each glycosyl cation, ten individual GA runs with the settings specified in Table S1 were carried out. The total number of generated structures can be found in Table S2 and their energy hierarchies in Figure S7. The GA yielded structures with four distinct modes of participation (Scheme 1): **(I)** dioxolenium-type structures exhibiting neighboring group participation of the C2-benzoyl protecting group, **(II)** dioxolenium-type structures exhibiting remote participation of the C4- (or C6-)Fmoc protecting group, **(III)** oxonium-type structures exhibiting *non-classical* remote participation of a C6-benzyl protecting group (only visible for **Glc1** and **3F-Glc1** glycosyl cations), and **(IV)** oxocarbenium-type structures exhibiting no participation (3D structures in Figures S9-S13).

Table S1. GA parameters used in initial search.

	Parameter	Value
Molecule	Distance_cutoff_1	1.2
	Distance_cutoff_2	2.15
	Rmsd_cutoff_uniq	0.25
GA settings	Popsize	10
	Prob_for_crossing	0.95
	Prob_for_mut_pyranosering	0.6
	Prob_for_mut_torsion	0.8
	Fitness_sum_limit	1.2
	Selection	Roulette wheel
	Max_mutations_torsion	3
	Max_mutations_pyranosering	1

For each mode of participation a certain number of low-energy structures, specified in Table S2, were reoptimized and frequencies computed at the PBE0+D3/6-311+G(d,p)^[5] level of theory with default settings using Gaussian 16, Revision A.03.^[6] Energies including zero-point vibrational energies and free energies at 90 K (according to experimental conditions in the hexapole ion trap) of the reoptimized geometries can be found in Tables S3–S7. Free energy hierarchies are shown in Figure S8. It is clearly distinguishable that structures exhibiting neighboring group participation of C2-benzoyl protecting groups are the most favored structural motif. All computed IR spectra were normalized and scaled by a factor of 0.965 (Figures 1 and 2).

Table S2. Number of generated structures for during GA and reoptimized structures. (I) neighboring group, (II) remote participation, (III) *non-classical* remote participation, and (IV) no participation.

Glycosyl Cations	#(GA Structures)	#(Reoptimized Structures)			
		I	II	III	IV
Glc1	214	10	5	5	5
3F-Glc1	314	10	5	5	5
6F-Glc1	316	10	5	/	5
Glc2	245	10	3	/	5
3F-Glc2	310	10	5	/	5

Energetics

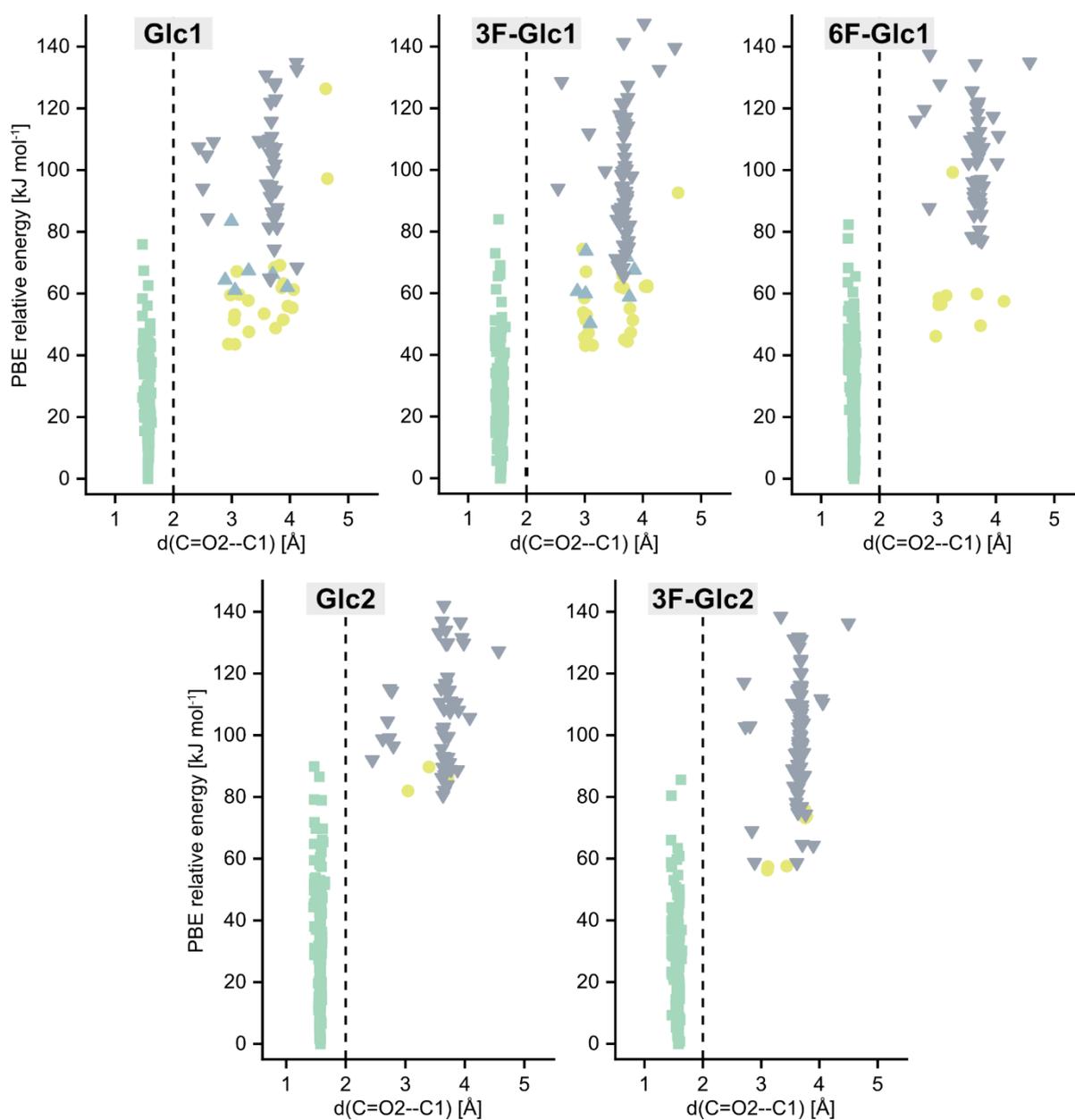


Figure S7. Energy hierarchies of sampled **Glc1**, **3F-Glc1**, **6F-Glc1**, **Glc2**, and **3F-Glc2** glycosyl cations as a function of the distance between the carbonyl oxygen of the C2-benzoyl protecting group and the anomeric carbon (C1). Energies were computed based on optimized geometries at the PBE/def2-SVP level of theory. Green squares indicate structures exhibiting C2-benzoyl neighboring group participation (I, dioxolenium), yellow circles C4-Fmoc remote participation (II, dioxolenium), blue triangles *non-classical* C6-benzyl remote participation (III, oxonium, only possible for **Glc1** and **3F-Glc1**), and gray triangles no participation (IV, oxocarbenium).

Table S3. List of reoptimized geometries of the **Glc1** glycosyl cation at the PBE0+D3/6-311+G(d,p) level of theory. Ring puckers, bond distances between the carbonyl oxygen of the C2-benzoyl group and the anomeric carbon (C1), energies (ΔE , including zero-point-vibrational energy) and free energies (ΔF) at 90 K are assigned to each structure. The infrared spectra of the structures labelled with a roman number are represented in the manuscript.

ID	Ring Pucker	d(C2=O—C1) [Å]	ΔE (PBE0+D3) [kJ mol ⁻¹]	ΔF (PBE0+D3) [kJ mol ⁻¹]
C2_Bz_NGP/conf_00 (I)	O,3B	1.51	0.00	0.00
C2_Bz_NGP/conf_01	3S1	1.51	3.45	3.38
C2_Bz_NGP/conf_02	O,3B	1.52	7.73	8.34
C2_Bz_NGP/conf_03	3S1	1.51	9.19	7.06
C2_Bz_NGP/conf_04	3S1	1.52	23.94	21.91
C2_Bz_NGP/conf_05	3S1	1.52	19.02	17.66
C2_Bz_NGP/conf_06	3S1	1.52	16.20	14.92
C2_Bz_NGP/conf_07	OS2	1.53	23.04	21.71
C2_Bz_NGP/conf_08	O,3B	1.52	34.56	31.66
C2_Bz_NGP/conf_09	OH5	1.51	18.81	19.05
C4_Fmoc_RP/conf_00	5S1	3.04	61.49	61.49
C4_Fmoc_RP/conf_01	B1,4	3.24	61.90	61.84
C4_Fmoc_RP/conf_02	5S1	3.61	76.87	74.13
C4_Fmoc_RP/conf_03 (II)	5S1	2.99	59.86	59.44
C4_Fmoc_RP/conf_04	5S1	3.00	87.16	82.03
C6_OBn_RP/conf_00	EO	2.95	58.52	57.63
C6_OBn_RP/conf_01 (III)	1C4	4.00	58.72	57.28
C6_OBn_RP/conf_02	BO,3	2.88	81.01	80.14
C6_OBn_RP/conf_03	BO,3	3.76	75.97	72.94
C6_OBn_RP/conf_04	1C4	2.99	65.68	64.89
oxocarbenium/conf_00 (IV)	2SO	3.72	83.38	80.03
oxocarbenium/conf_01	2H3	3.76	94.40	91.63
oxocarbenium/conf_02 ^(a)	5S1	3.77	83.99	81.17
oxocarbenium/conf_03	5H4	3.54	116.06	108.67
oxocarbenium/conf_04	2H3	3.74	112.00	108.98

(a) The structure "oxocarbenium/conf_02" converged into a dioxolenium-type structure exhibiting remote participation of the C4-Fmoc protecting group after reoptimization of the geometry at the PBE0+D3/6-311+G(d,p) level of theory and was therefore not further considered.

Table S4. List of reoptimized geometries of the **3F-Glc1** glycosyl cation at the PBE0+D3/6-311+G(d,p) level of theory. Ring puckers, bond distances between the carbonyl oxygen of the C2-benzoyl group and the anomeric carbon (C1), energies (ΔE , including zero-point-vibrational energy) and free energies (ΔF) at 90 K are assigned to each structure. The infrared spectra of the structures labelled with a roman number are represented in the manuscript.

ID	Ring Pucker	d(C2=O—C1) [Å]	ΔE (PBE0+D3) [kJ mol ⁻¹]	ΔF (PBE0+D3) [kJ mol ⁻¹]
C2_Bz_NGP/conf_00 (I)	3S1	1.50	0.00	0.00
C2_Bz_NGP/conf_01	3S1	1.50	0.00	0.00
C2_Bz_NGP/conf_02	3S1	1.51	8.31	6.09
C2_Bz_NGP/conf_03	3S1	1.51	4.86	4.23
C2_Bz_NGP/conf_04	O,3B	1.50	11.31	9.09
C2_Bz_NGP/conf_05	5E	1.46	20.59	19.95
C2_Bz_NGP/conf_06	3S1	1.51	7.12	6.75
C2_Bz_NGP/conf_07	5H4	1.49	29.67	26.37
C2_Bz_NGP/conf_08	3S1	1.51	7.13	6.76
C2_Bz_NGP/conf_09	OH5	1.50	13.49	12.72
C4_Fmoc_RP/conf_00	5S1	2.98	72.22	68.65
C4_Fmoc_RP/conf_01	B1,4	4.08	57.06	55.23
C4_Fmoc_RP/conf_02	3S1	2.92	56.48	55.94
C4_Fmoc_RP/conf_03 (II)	5S1	4.11	54.64	54.54
C4_Fmoc_RP/conf_04	3S1	2.92	81.18	76.80
C6_OBn_RP/conf_00 (III)	1C4	2.92	54.41	51.67
C6_OBn_RP/conf_01	1C4	3.71	69.85	67.04
C6_OBn_RP/conf_02	1C4	3.03	72.97	69.27
C6_OBn_RP/conf_03	BO,3	2.84	83.01	77.84
C6_OBn_RP/conf_04	1C4	4.02	67.80	65.57
oxocarbenium/conf_00 (IV)	2SO	3.68	75.16	71.78
oxocarbenium/conf_01	5H4	3.56	80.98	76.03
oxocarbenium/conf_02	5H4	3.54	90.06	85.90
oxocarbenium/conf_03 ^(a)	1C4	3.99	66.71	65.13
oxocarbenium/conf_04	3H4	3.33	93.88	90.99

(a) The structure “oxocarbenium/conf_03” converged into an oxonium-type structure exhibiting *non-classical* remote participation of a C6-benzyl group after reoptimization of the geometry at the PBE0+D3/6-311+G(d,p) level of theory and was therefore not further considered.

Table S5. List of reoptimized geometries of the **6F-Glc1** glycosyl cation at the PBE0+D3/6-311+G(d,p) level of theory. Ring puckers, bond distances between the carbonyl oxygen of the C2-benzoyl group and the anomeric carbon (C1), energies (ΔE , including zero-point-vibrational energy) and free energies (ΔF) at 90 K are assigned to each structure. The infrared spectra of the structures labelled with a roman number are represented in the manuscript.

ID	Ring Pucker	d(C2=O—C1) [Å]	ΔE (PBE0+D3) [kJ mol ⁻¹]	ΔF (PBE0+D3) [kJ mol ⁻¹]
C2_Bz_NGP/conf_00	3S1	1.52	2.14	2.00
C2_Bz_NGP/conf_01 (IA)	3S1	1.51	0.00	0.00
C2_Bz_NGP/conf_02	3S1	1.50	0.76	0.59
C2_Bz_NGP/conf_03	3S1	1.50	2.22	2.41
C2_Bz_NGP/conf_04	3S1	1.52	5.14	4.24
C2_Bz_NGP/conf_05	3S1	1.53	0.97	2.88
C2_Bz_NGP/conf_06	3S1	1.52	5.14	4.28
C2_Bz_NGP/conf_07 (IB)	O,3B	1.50	0.62	0.12
C2_Bz_NGP/conf_08	3S1	1.51	7.85	7.39
C2_Bz_NGP/conf_09	3S1	1.52	8.18	7.24
C4_Fmoc_RP/conf_00 (II)	5S1	3.09	56.36	56.28
C4_Fmoc_RP/conf_01	5S1	3.73	66.10	67.72
C4_Fmoc_RP/conf_02	3S1	2.88	87.33	83.42
C4_Fmoc_RP/conf_03	3S1	4.10	71.92	70.57
C4_Fmoc_RP/conf_04	3S1	2.96	66.65	65.53
oxocarbenium/conf_00 (IV)	E4	3.61	87.66	83.94
oxocarbenium/conf_01	2H3	3.74	91.25	88.21
oxocarbenium/conf_02	5H4	3.56	86.24	84.68
oxocarbenium/conf_03	E4	3.61	93.06	88.60
oxocarbenium/conf_04	5H4	3.58	131.71	124.56

Table S6. List of reoptimized geometries of the **Glc2** glycosyl cation at the PBE0+D3/6-311+G(d,p) level of theory. Ring puckers, bond distances between the carbonyl oxygen of the C2-benzoyl group and the anomeric carbon (C1), energies (ΔE , including zero-point-vibrational energy) and free energies (ΔF) at 90 K are assigned to each structure. The infrared spectra of the structures labelled with a roman number are represented in the manuscript.

ID	Ring Pucker	d(C2=O—C1) [Å]	ΔE (PBE0+D3) [kJ mol ⁻¹]	ΔF (PBE0+D3) [kJ mol ⁻¹]
C2_Bz_NGP/conf_00 (IA)	3S1	1.51	0.00	0.00
C2_Bz_NGP/conf_01	3S1	1.52	24.18	21.56
C2_Bz_NGP/conf_02	3S1	1.52	14.29	12.90
C2_Bz_NGP/conf_03 (IB)	3S1	1.51	4.53	4.89
C2_Bz_NGP/conf_04	3S1	1.52	24.30	22.00
C2_Bz_NGP/conf_05	3S1	1.52	25.66	23.12
C2_Bz_NGP/conf_06	OS2	1.52	8.12	8.44
C2_Bz_NGP/conf_07	3S1	1.51	12.66	11.52
C2_Bz_NGP/conf_08	3S1	1.51	19.76	19.10
C2_Bz_NGP/conf_09	O,3B	1.52	23.49	22.32
C6_Fmoc_RP/conf_00	5HO	3.02	96.30	95.54
C6_Fmoc_RP/conf_01	2SO	3.72	105.14	103.81
C6_Fmoc_RP/conf_02 (II)	2SO	3.15	95.80	94.88
oxocarbenium/conf_00	E4	2.93	104.93	98.63
oxocarbenium/conf_01	5H4	3.65	86.22	83.25
oxocarbenium/conf_02 (IV)	3H4	3.96	84.98	82.85
oxocarbenium/conf_03	E4	3.71	99.05	95.39
oxocarbenium/conf_04	E3	3.70	130.36	122.22

Table S7. List of reoptimized geometries of the **3F-Glc2** glycosyl cation at the PBE0+D3/6-311+G(d,p) level of theory. Ring puckers, bond distances between the carbonyl oxygen of the C2-benzoyl group and the anomeric carbon (C1), energies (ΔE , including zero-point-vibrational energy) and free energies (ΔF) at 90 K are assigned to each structure. The infrared spectra of the structures labelled with a roman number are represented in the manuscript.

ID	Ring Pucker	d(C2=O—C1) [Å]	ΔE (PBE0+D3) [kJ mol ⁻¹]	ΔF (PBE0+D3) [kJ mol ⁻¹]
C2_Bz_NGP/conf_00	3S1	1.51	10.93	10.27
C2_Bz_NGP/conf_01	4H5	1.54	10.90	10.75
C2_Bz_NGP/conf_02 (IA)	5H4	1.53	0.00	0.00
C2_Bz_NGP/conf_03	3S1	1.52	11.98	11.68
C2_Bz_NGP/conf_04	3S1	1.52	16.49	17.07
C2_Bz_NGP/conf_05	3S1	1.51	7.92	8.76
C2_Bz_NGP/conf_06	5H4	1.52	6.39	4.72
C2_Bz_NGP/conf_07	OS2	1.49	7.14	6.75
C2_Bz_NGP/conf_08	5H4	1.53	0.01	0.00
C2_Bz_NGP/conf_09 (IB)	OS2	1.50	3.64	3.22
C6_Fmoc_RP/conf_00 (II)	1C4	3.05	66.09	65.66
C6_Fmoc_RP/conf_01	1C4	3.02	77.83	77.11
C6_Fmoc_RP/conf_02	2SO	3.35	75.75	74.67
C6_Fmoc_RP/conf_03	2SO	3.52	93.28	91.74
C6_Fmoc_RP/conf_04	BO,3	3.37	98.23	95.88
oxocarbenium/conf_00 (IV)	5H4	3.73	65.81	64.03
oxocarbenium/conf_01	5H4	3.81	67.48	65.86
oxocarbenium/conf_02	4H3	3.64	98.81	95.56
oxocarbenium/conf_03 ^(a)	5HO	4.01	67.64	67.64
oxocarbenium/conf_04	E4	3.59	97.75	93.99

(a) The structure "oxocarbenium/conf_03" converged into a dioxolenium-type structure exhibiting remote participation of the C6-Fmoc protecting group after reoptimization of the geometry at the PBE0+D3/6-311+G(d,p) level of theory and was therefore not further considered.

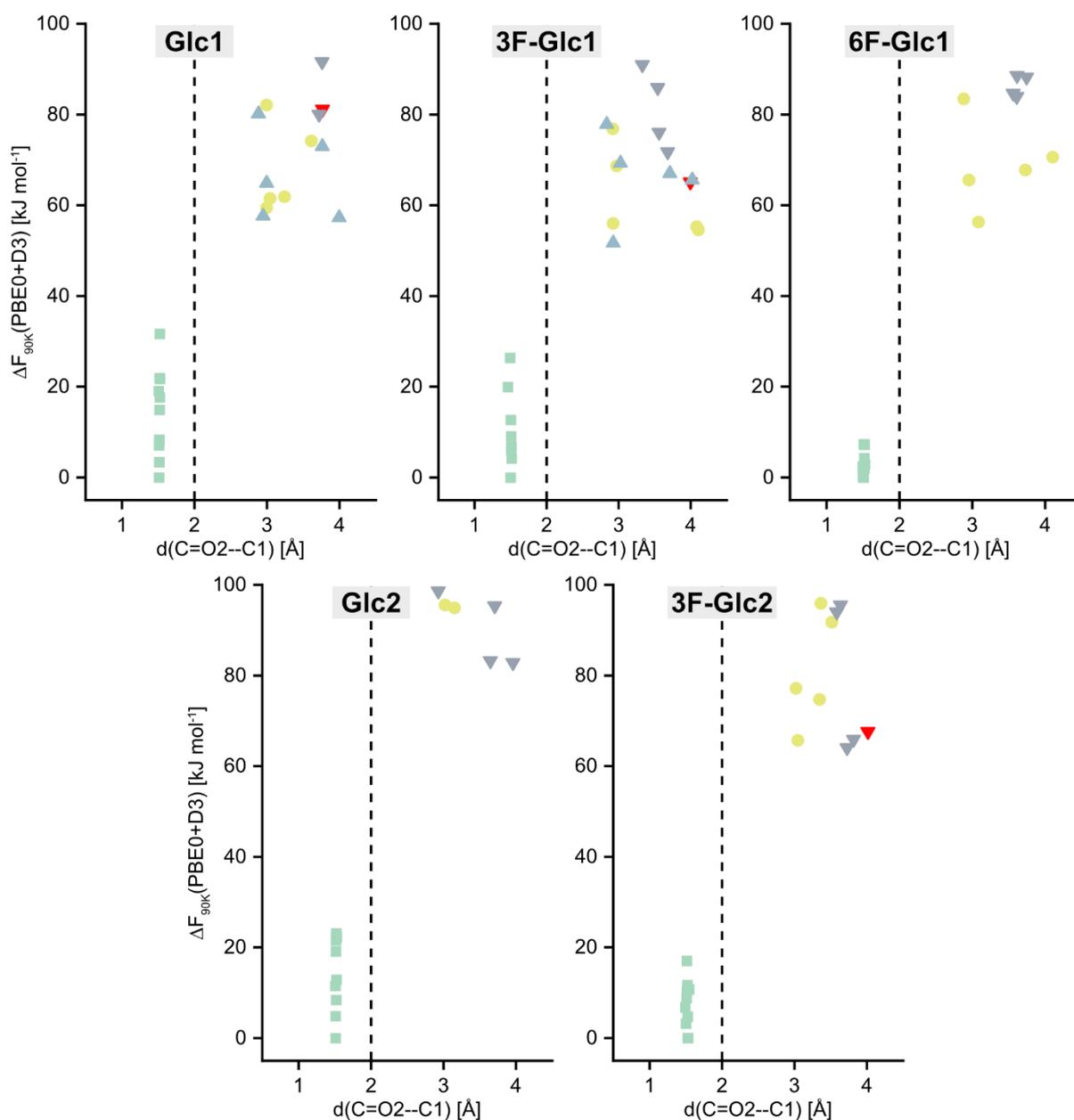


Figure S8. Energy hierarchies of reoptimized **Glc1**, **3F-Glc1**, **6F-Glc1**, **Glc2**, and **3F-Glc2** glycosyl cations as a function of the distance between the carbonyl oxygen of the C2-benzoyl protecting group and the anomeric carbon (C1). Free energies at 90 K were computed based on optimized geometries at the PBE0+D3/6-311+G(d,p) level of theory. Green squares indicate structures exhibiting C2-benzoyl neighboring group participation (**I**, dioxolenium), yellow circles C4-Fmoc remote participation (**II**, dioxolenium), blue triangles *non-classical* C6-benzyl remote participation (**III**, oxonium), and gray triangles no participation (**IV**, oxocarbenium). Red triangles represent former oxocarbenium-type structures that converged into another structural motif during reoptimization of the geometry. As they were not the lowest-energy structure of that motif, they were not further considered.

Low-Energy Structures

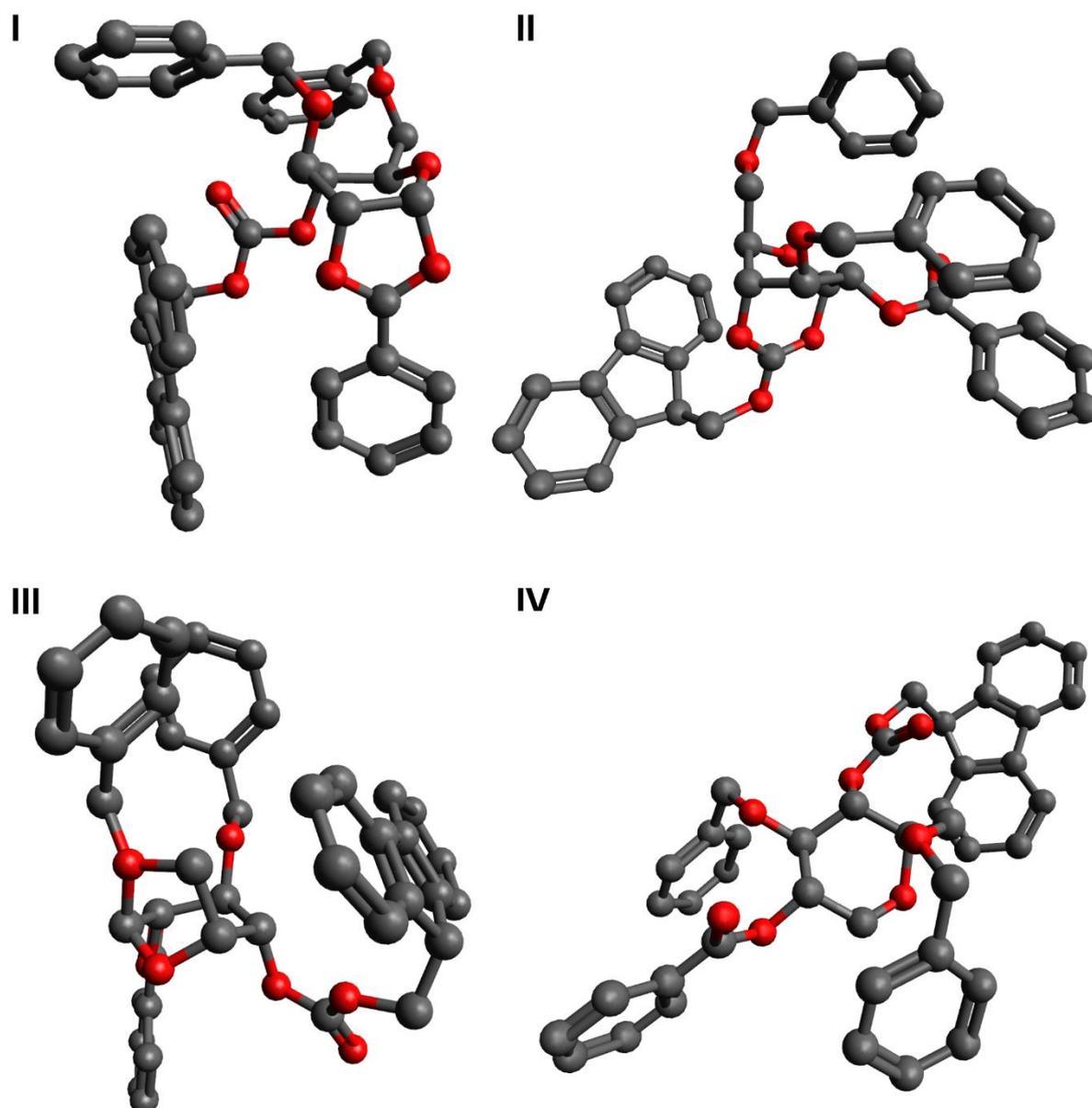


Figure S9. Reoptimized geometries of **Glc1** glycosyl cations exhibiting (I) C2-benzoyl neighboring group participation (C2_Bz_NGP/conf_00), (II) C4-Fmoc remote participation (C4_Fmoc_RP/conf_03), (III) C6-OBn *non-classical* remote participation (C6_OBn_RP/conf_01), and (IV) no participation (oxocarbenium/conf_00). Hydrogens are omitted for clarity.

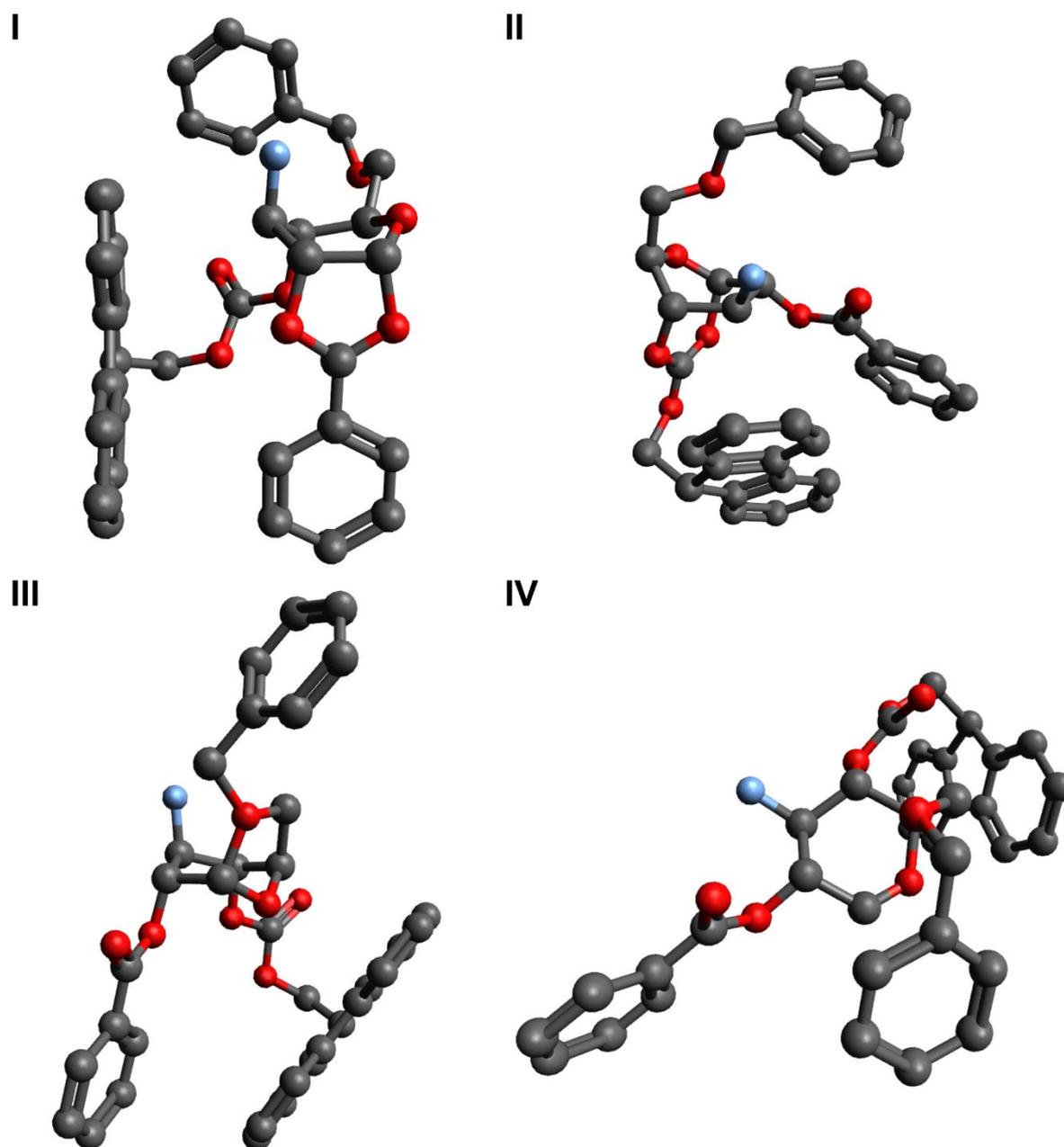


Figure S10. Reoptimized geometries of **3F-Glc1** glycosyl cations exhibiting (I) C2-benzoyl neighboring group participation (C2_Bz_NGP/conf_00), (II) C4-Fmoc remote participation (C4_Fmoc_RP/conf_03), (III) C6-OBn *non-classical* remote participation (C6_OBn_RP/conf_00), and (IV) no participation (oxocarbenium/conf_00). Hydrogens are omitted for clarity.

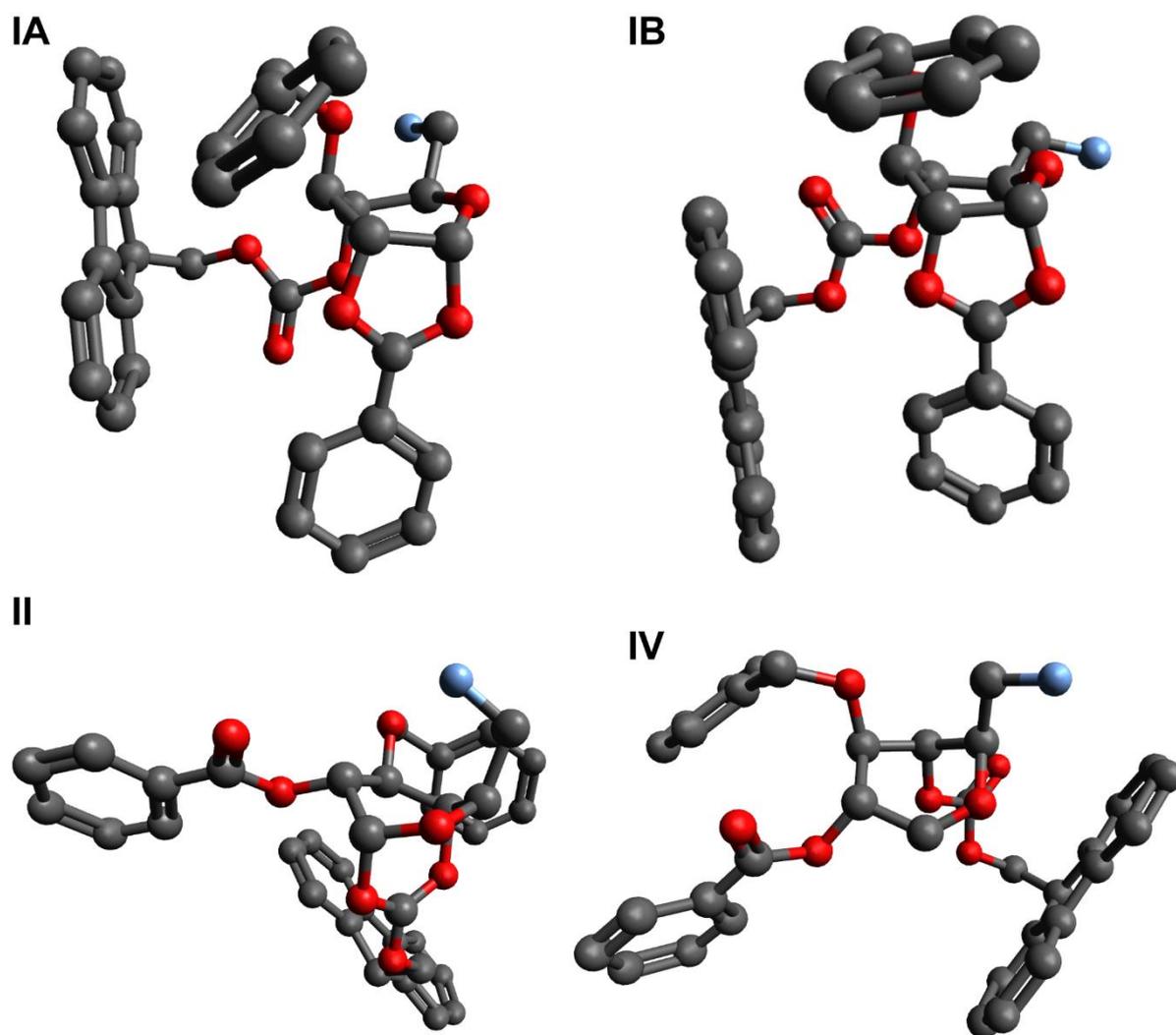


Figure S11. Reoptimized geometries of **6F-Glc1** glycosyl cations exhibiting (I) C2-benzoyl neighboring group participation (**A** C2_Bz_NGP/conf_01 and **B** C2_Bz_NGP/conf_07), (II) C4-Fmoc remote participation (C4_Fmoc_RP/conf_00), and (IV) no participation (oxocarbenium/conf_00). Hydrogens are omitted for clarity.

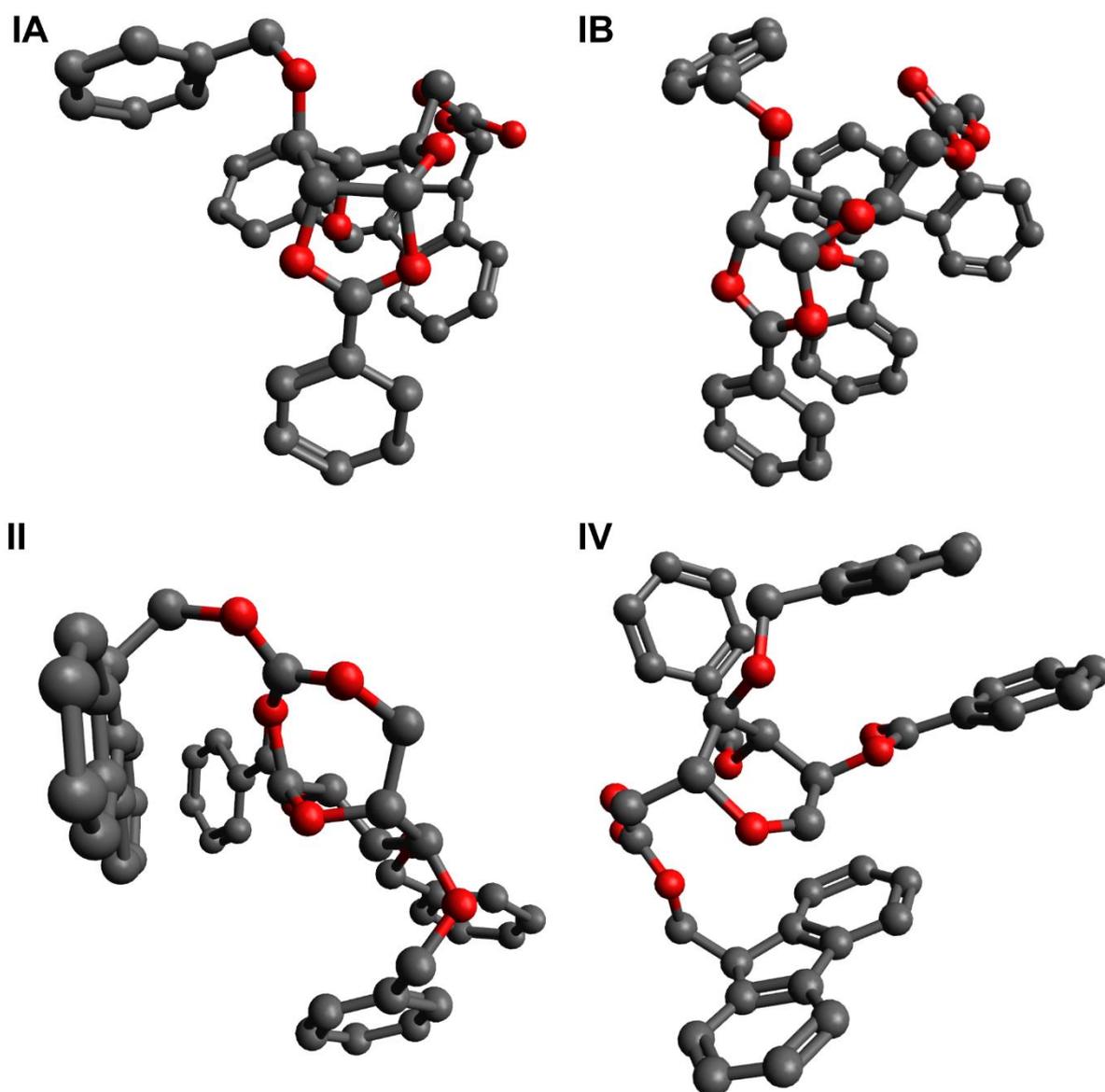


Figure S12. Reoptimized geometries of **Glc2** glycosyl cations exhibiting (I) C2-benzoyl neighboring group participation (A C2_Bz_NGP/conf_00 and B C2_Bz_NGP/conf_03), (II) C6-Fmoc remote participation (C4_Fmoc_RP/conf_02), and (IV) no participation (oxocarbenium/conf_02). Hydrogens are omitted for clarity.

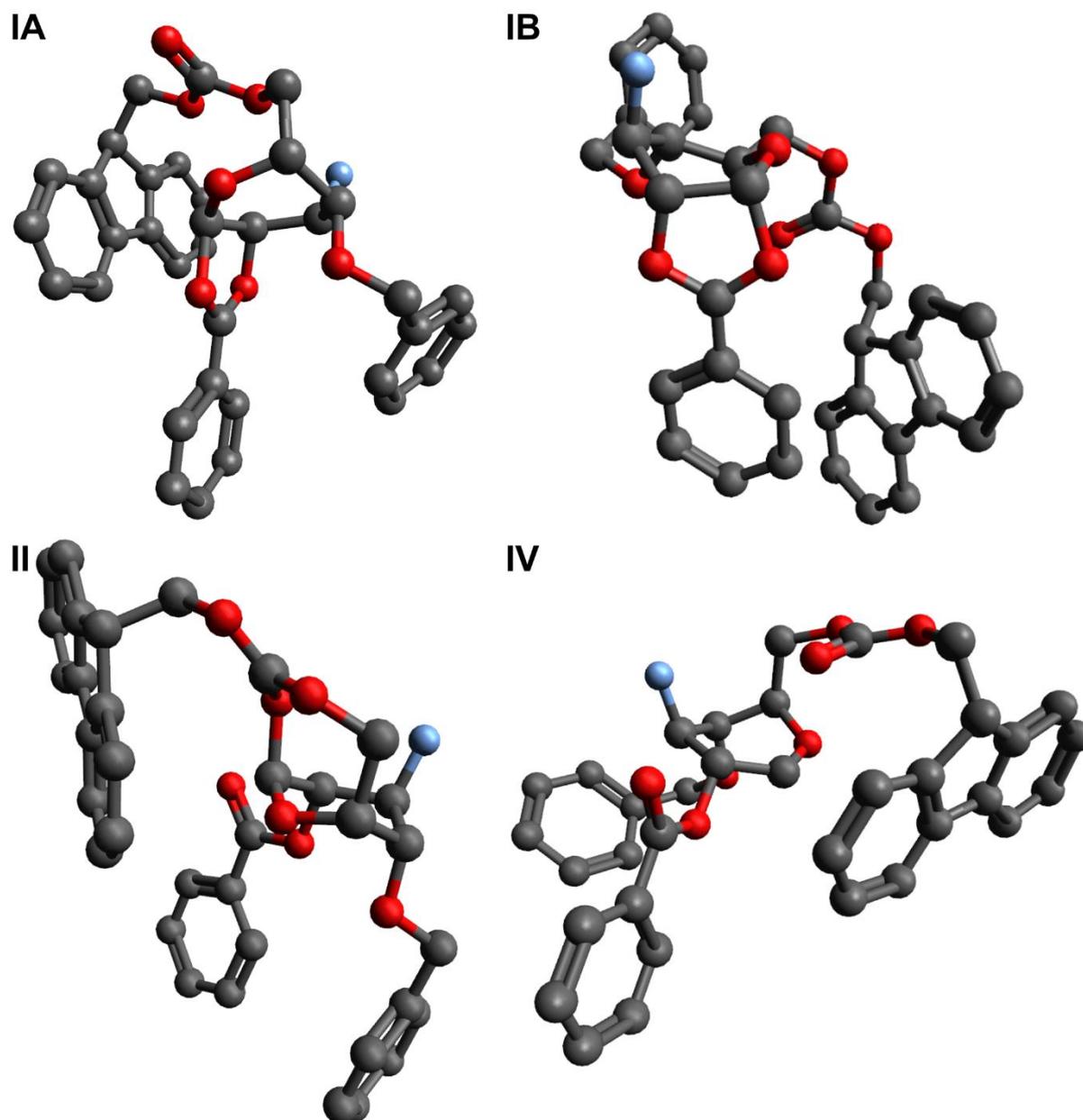
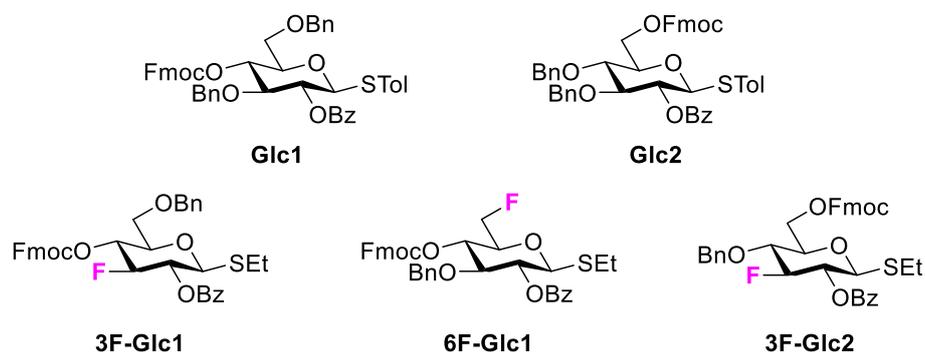


Figure S13. Reoptimized geometries of **3F-Glc2** glycosyl cations exhibiting (I) C2-benzoyl neighboring group participation (**A** C2_Bz_NGP/conf_02 and **B** C2_Bz_NGP/conf_09), (II) C6-Fmoc remote participation (C4_Fmoc_RP/conf_00), and (IV) no participation (oxocarbenium/conf_00). Hydrogens are omitted for clarity.

xyz-Coordinates of Reoptimized Structures

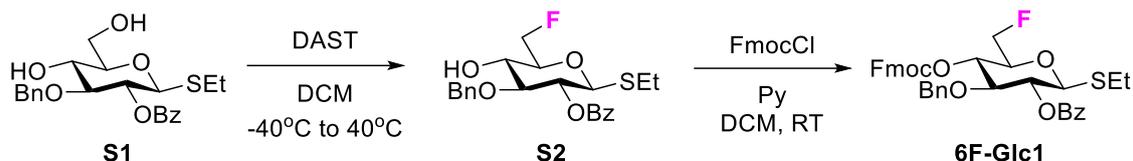
xyz-Coordinates of all reoptimized geometries can be found in a separate document “coordinates.xyz”.

Synthesis of Building Blocks



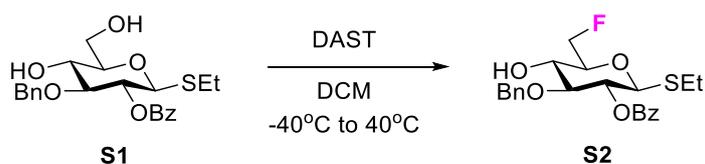
Scheme S1. Building blocks (BBs) used in this study.

Building blocks **Glc1** and **Glc2** were purchased from GlycoUniverse (Germany). **3F-Glc1** and **3F-Glc2** were synthesized as previously reported.^[7] The synthesis of **6F-Glc1** is reported hereafter.



Scheme S2. Synthesis of fluorinated BBs **6F-Glc1**.

Synthesis of ethyl 2-*O*-benzoyl-3-*O*-benzyl-6-deoxy-6-fluoro-1-thio- β -D-glucopyranoside, **S2**



S1 was prepared according to previously established procedures.^[8]

Ethyl 2-*O*-benzoyl-3-*O*-benzyl-6-deoxy-1-thio- β -D-glucopyranoside **S1** (490 mg, 1.17 mmol) was dissolved in anhydrous DCM (10 mL) and cooled to -40°C (dry ice/ACN bath) under Ar atmosphere. DAST (171 μL , 1.29 mmol) was dissolved in anhydrous DCM (200 μL) and added dropwise to the reaction mixture. After 30 min the cooling bath was removed and the reaction heated up to 40°C . The solution was stirred for additional 5 h and quenched with MeOH at 0°C . The crude reaction mixture was diluted with DCM and washed once with brine. The crude

compound was purified by silica gel flash column chromatography (Hexane : EtOAc = 3:1 → 1:1) to give **S2** as a colorless oil (167 mg, 34%).

^1H NMR (400 MHz, Chloroform-*d*) δ 8.11 – 8.02 (m, 2H), 7.64 – 7.56 (m, 1H), 7.47 (tt, $J = 6.7, 1.2$ Hz, 2H), 7.32 – 7.15 (m, 5H), 5.36 – 5.24 (m, 1H), 4.81 – 4.51 (m, 5H), 3.78 – 3.67 (m, 2H), 3.58 (dddd, $J = 21.5, 8.3, 5.0, 2.7$ Hz, 1H), 2.82 – 2.63 (m, 2H), 2.34 (s, 1H), 1.24 (t, $J = 7.5$ Hz, 5H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 165.34, 137.77, 133.51, 129.99, 129.79, 128.78, 128.66, 128.29, 128.17, 84.14, 83.80, 82.19 (d, $J = 173.4$ Hz), 78.55 (d, $J = 18.6$ Hz), 74.94, 72.25, 69.07 (d, $J = 7.2$ Hz), 24.12, 14.92. ^{19}F NMR (376 MHz, Chloroform-*d*) δ -233.66 (td, $J = 47.3, 22.8$ Hz). $[\alpha]_{\text{D}}^{20}$ -18.12 (c 1.2 g/100 mL, CHCl_3). IR $\nu = 3482, 2927, 1724, 1268, 1086, 1070, 1027, 710, 700$ cm^{-1} . (ESI-HRMS) m/z 443.1291 $[\text{M}+\text{Na}]^+$ ($\text{C}_{22}\text{H}_{25}\text{FO}_5\text{SNa}$ requires 443.1302).

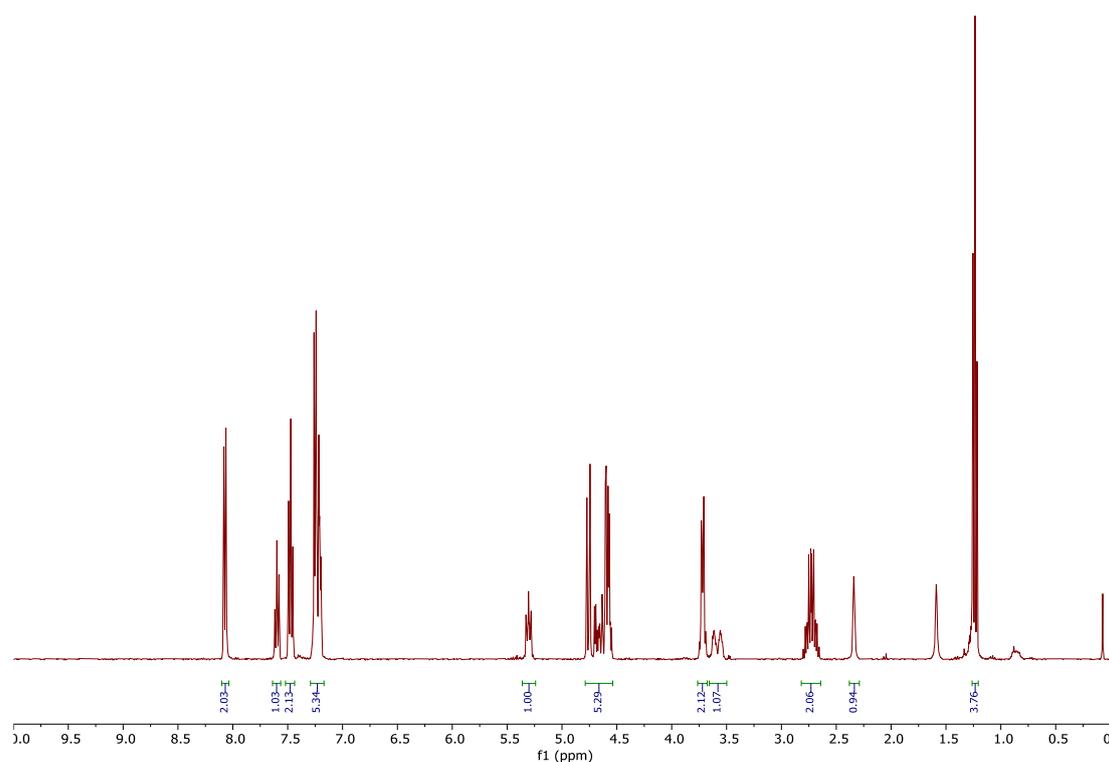


Figure S14. ^1H -NMR spectrum of **S2** (400 MHz, Chloroform-*d*).

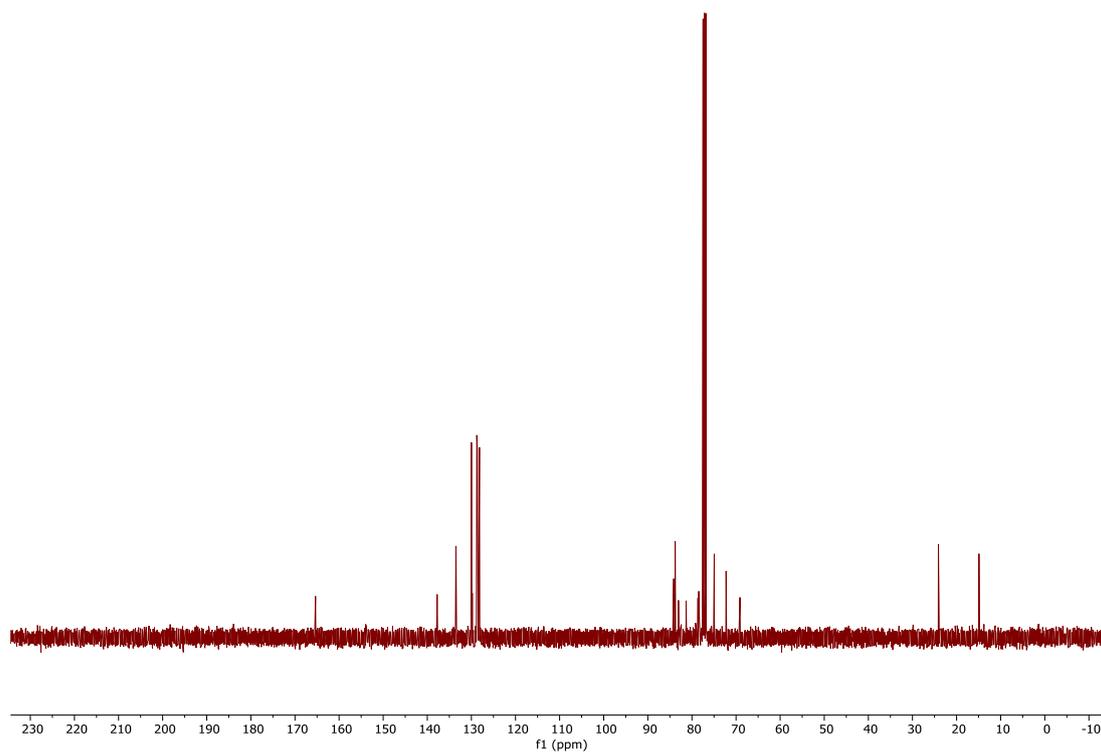


Figure S15. ^{13}C -NMR spectrum of **S2** (101 MHz, Chloroform-*d*).

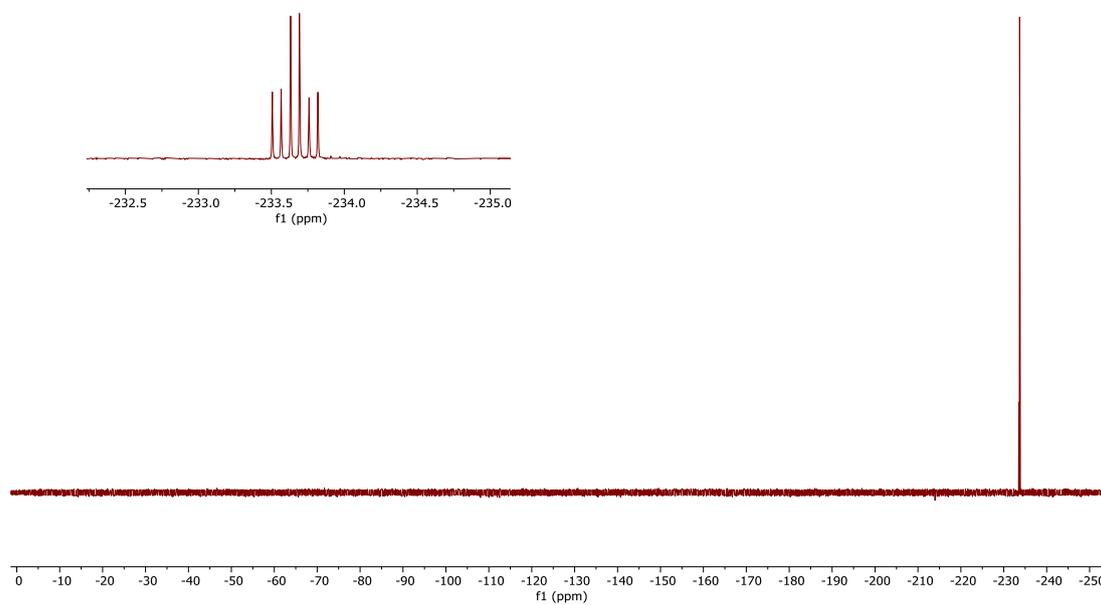


Figure S16. ^{19}F -NMR spectrum of **S2** (376 MHz, Chloroform-*d*).

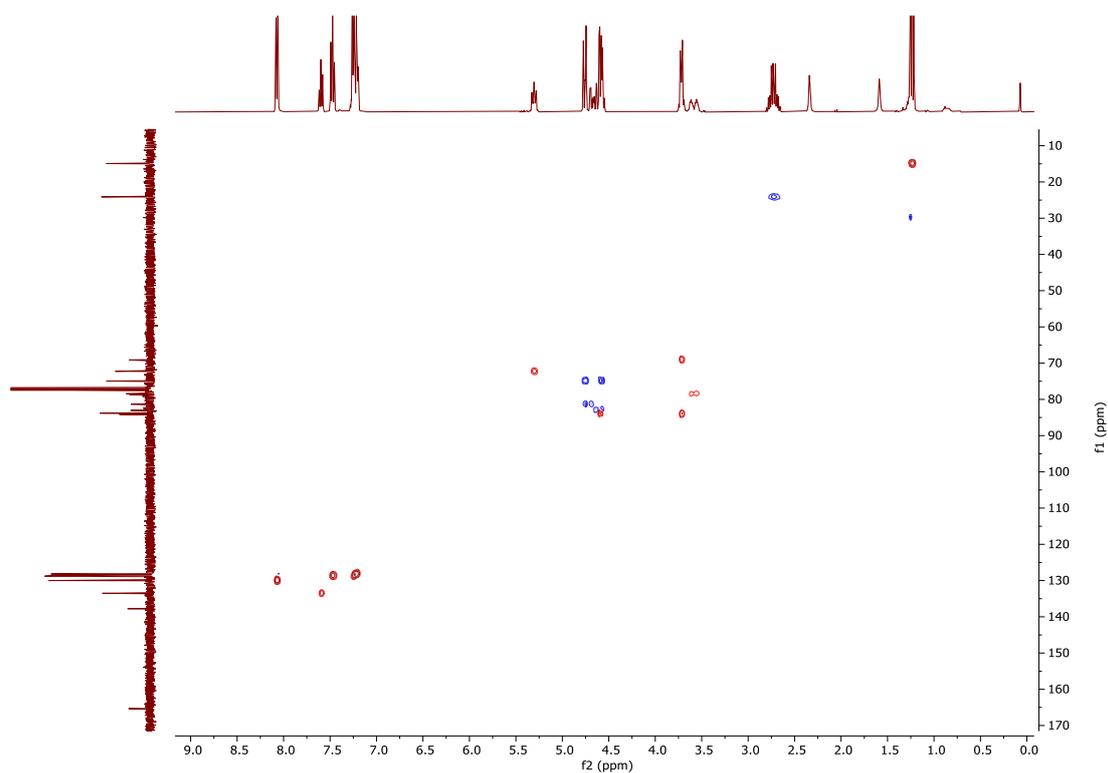


Figure S17. HSQC-NMR spectrum of **S2** (Chloroform-*d*).

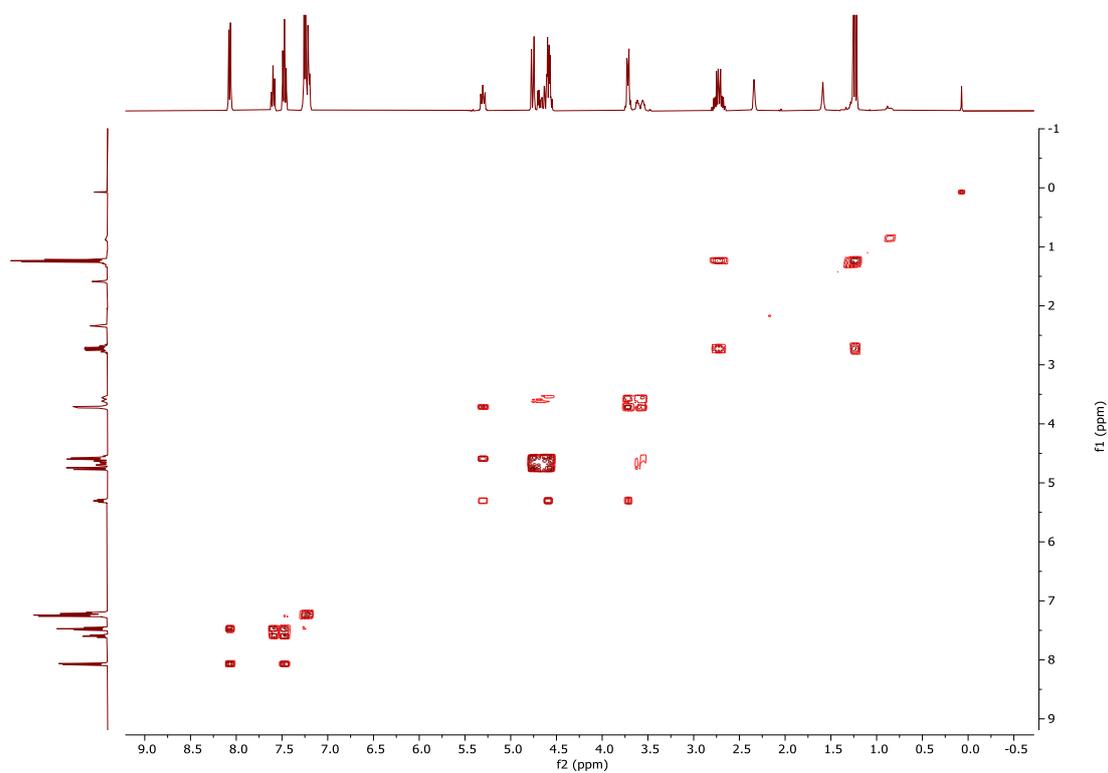
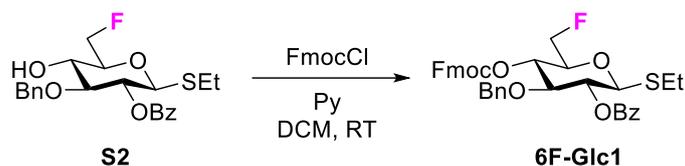


Figure S18. COSY-NMR spectrum of **S2** (Chloroform-*d*).

Synthesis of ethyl 2-*O*-benzoyl-3-*O*-benzyl-4-*O*-(9-fluorenylmethoxycarbonyl)-6-deoxy-6-fluoro-1-thio- β -D-glucopyranoside, **6F-Glc1**



Ethyl 2-*O*-benzoyl-3-*O*-benzyl-6-deoxy-6-fluoro-1-thio- β -D-glucopyranoside **S2** (167 mg, 0.40 mmol) was dissolved in DCM (5 mL) and pyridine was added (100 μ L, 1.2 mmol). FmocCl (200 mg, 0.77 mmol) was dissolved in DCM (1.5 mL) and added to the reaction mixture at RT under Ar atmosphere. The solution was stirred for 3 h and then quenched with a 1 M solution of HCl. The crude reaction mixture was diluted with DCM, washed once with 1 M HCl, and once with brine. The crude compound was purified by silica gel flash column chromatography (Toluene : DCM = 4:1 \rightarrow 3:1 then Toluene : EtOAc = 4:1) and precipitated from DCM : Hexane to give the **6F-Glc1** as a white solid (186 mg, 72%).

^1H NMR (400 MHz, Chloroform-*d*) δ 8.05 – 7.96 (m, 2H), 7.81 – 7.72 (m, 2H), 7.66 – 7.53 (m, 3H), 7.50 – 7.36 (m, 4H), 7.30 (tt, J = 7.5, 1.0 Hz, 2H), 7.20 – 7.00 (m, 5H), 5.33 (dd, J = 10.0, 9.1 Hz, 1H), 4.95 (dd, J = 10.2, 9.2 Hz, 1H), 4.63 – 4.49 (m, 5H), 4.48 – 4.40 (m, 2H), 4.20 (t, J = 6.8 Hz, 1H), 3.91 (t, J = 9.1 Hz, 1H), 3.82 – 3.69 (m, 1H), 2.82 – 2.65 (m, 2H), 1.23 (t, J = 7.5 Hz, 3H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 164.99, 154.17, 143.08, 143.04, 141.36, 137.18, 133.38, 129.91, 129.54, 128.48, 128.21, 128.00, 127.88, 127.73, 127.24, 125.04, 124.91, 120.17, 120.15, 83.69, 81.56 (d, J = 175.3 Hz), 80.81, 76.73, 74.45, 73.92 (d, J = 6.2 Hz), 71.64, 70.14, 46.79, 24.00, 14.80. ^{19}F NMR (376 MHz, Chloroform-*d*) δ -230.76 (td, J = 47.0, 20.0 Hz). $[\alpha]_{\text{D}}^{20}$ 23.35 (c 0.6 g/100 mL, CHCl_3). IR ν = 2928, 1754, 1729, 1248, 1028, 742, 710 cm^{-1} . (ESI-HRMS) m/z 665.1992 $[\text{M}+\text{Na}]^+$ ($\text{C}_{37}\text{H}_{35}\text{FO}_7\text{SNa}$ requires 665.1980).

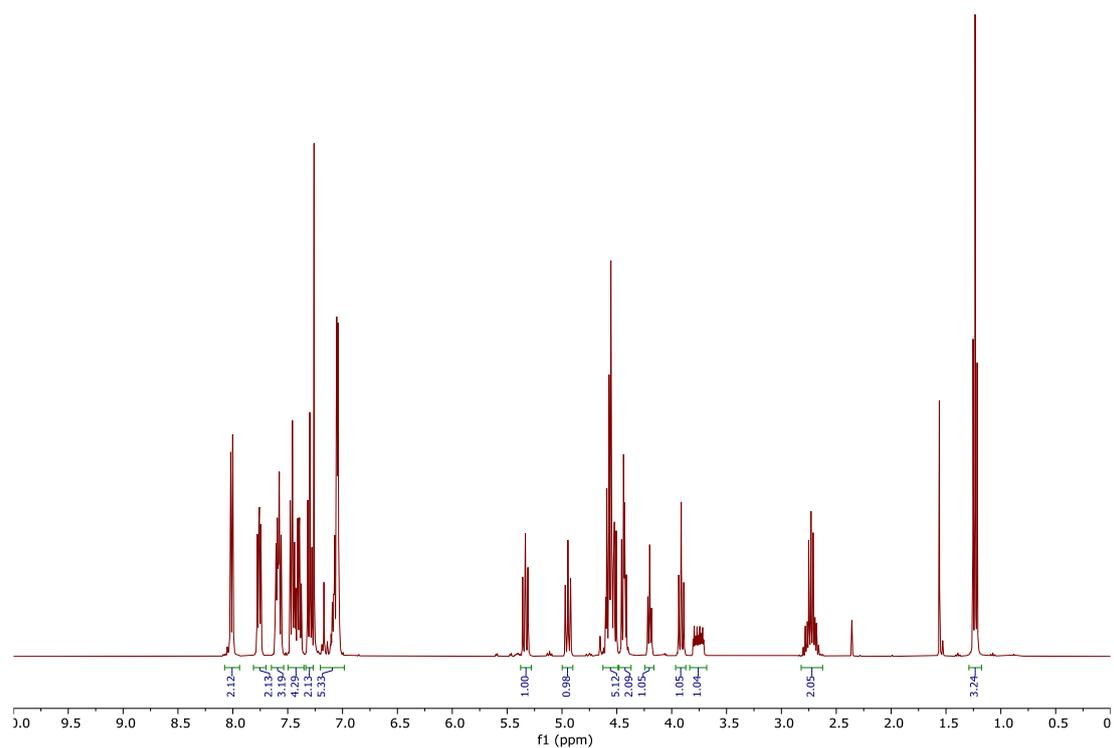


Figure S19. ¹H-NMR spectrum of **6F-Glc1** (400 MHz, Chloroform-*d*).

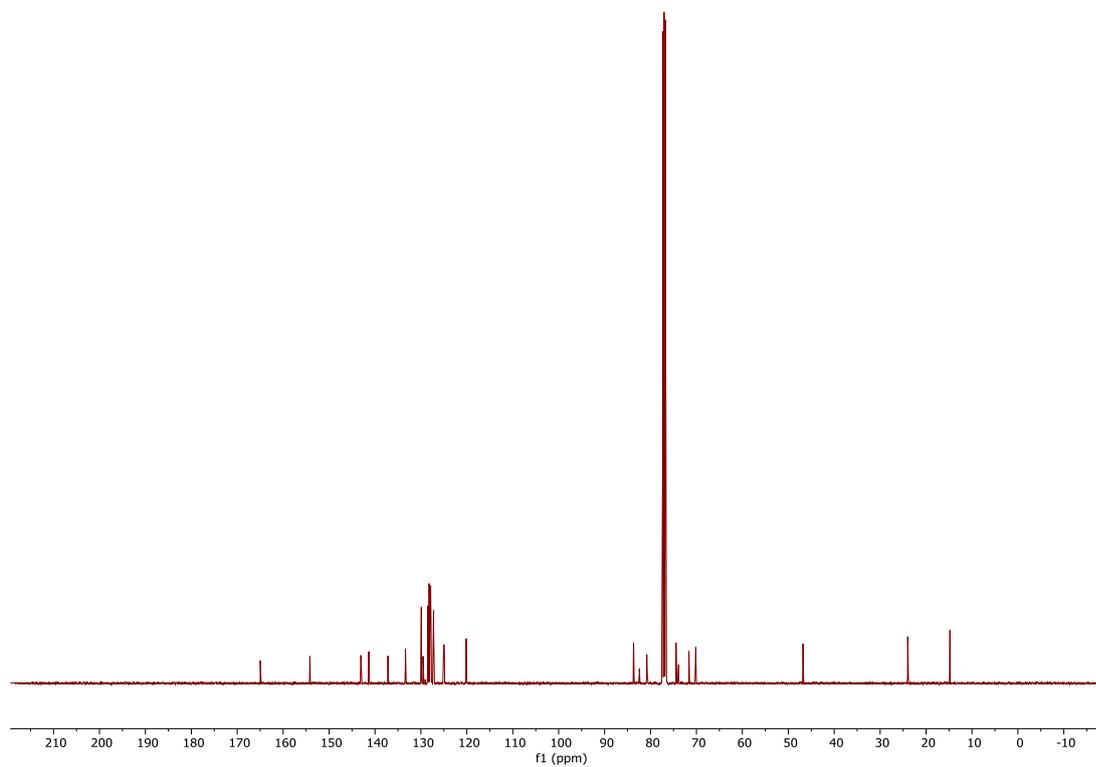


Figure S20. ¹³C-NMR spectrum of **6F-Glc1** (101 MHz, Chloroform-*d*).

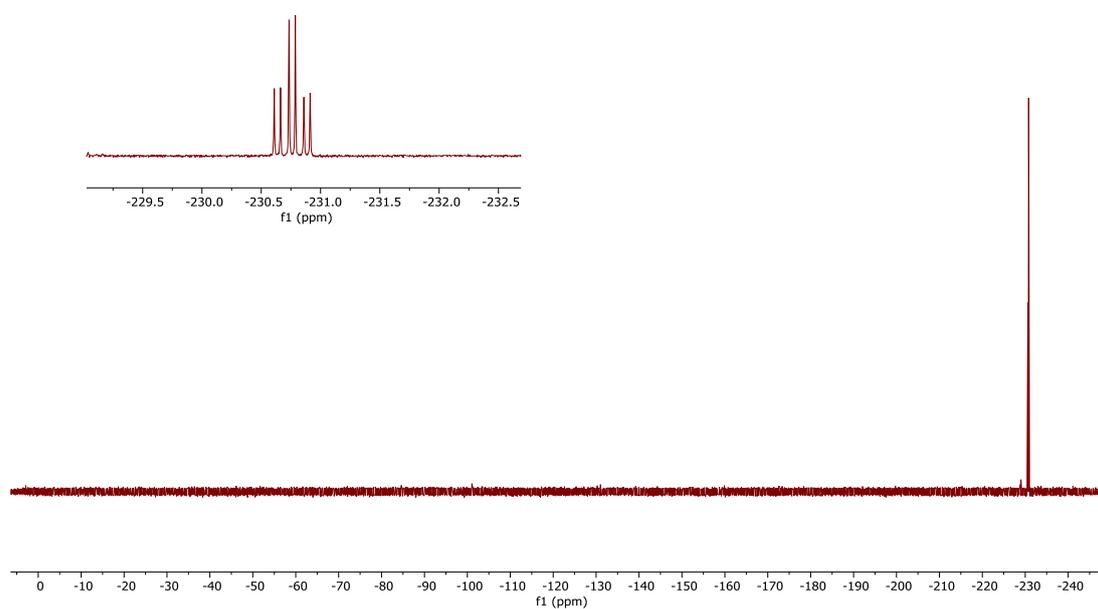


Figure S21. ^{19}F -NMR spectrum of **6F-Glc1** (376 MHz, Chloroform-*d*).

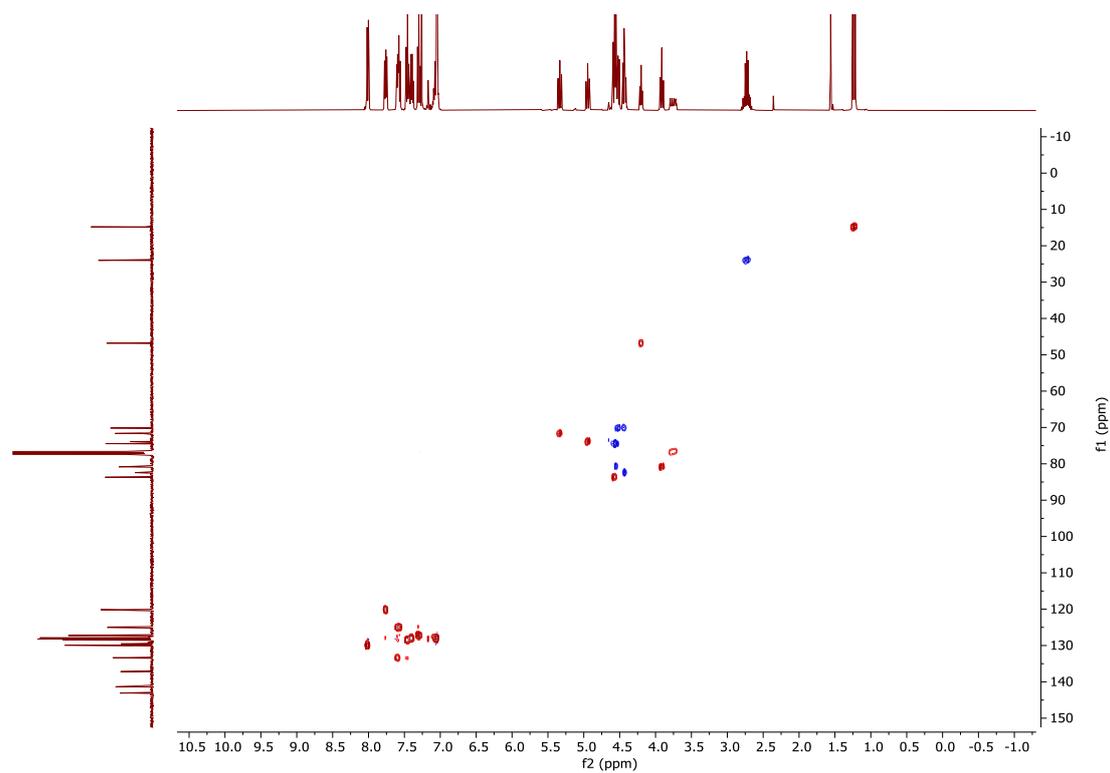


Figure S22. HSQC-NMR spectrum of **6F-Glc1** (Chloroform-*d*).

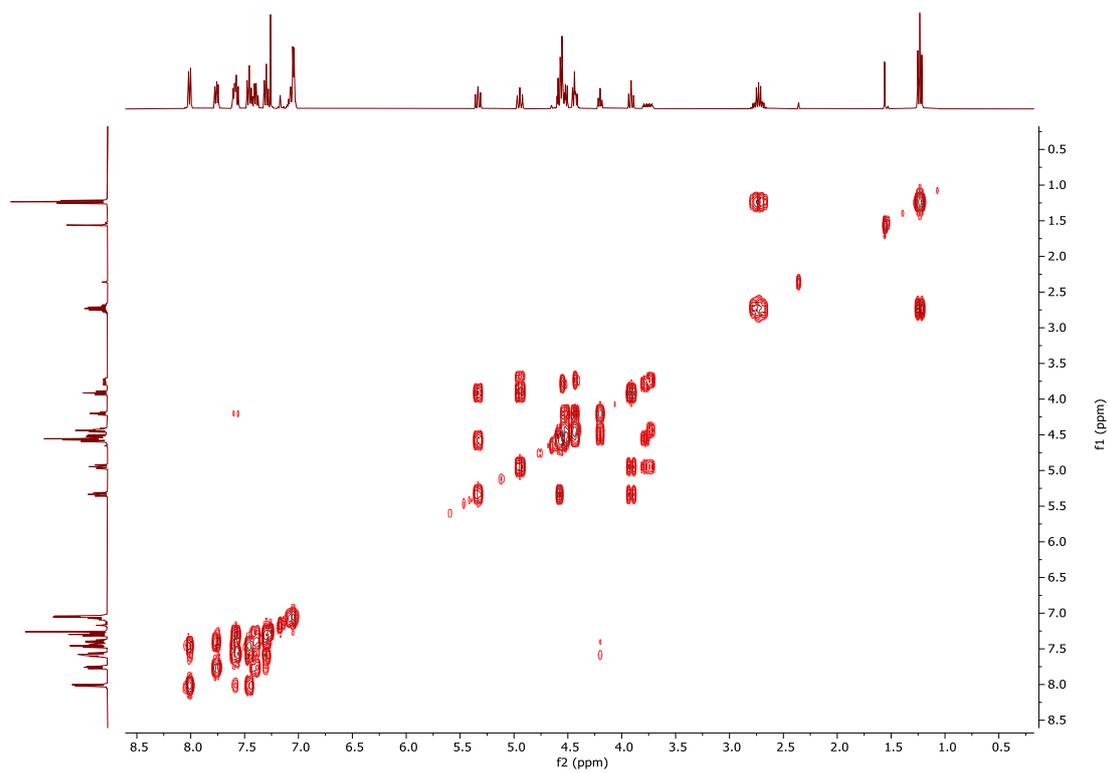


Figure 23. COSY-NMR spectrum of **6F-Glc1** (Chloroform-*d*).

References

- [1] W. Schöllkopf, S. Gewinner, H. Junkes, A. Paarmann, G. von Helden, H. P. Bluem, A. M. M. Todd, *Proc. SPIE Int. Soc. Opt. Eng.* **2015**, 9512, 95121L.
- [2] A. Supady, V. Blum, C. Baldauf, *J. Chem. Inf. Model.* **2015**, 55, 2338.
- [3] F. Neese, *WIREs Comput. Mol. Sci.* **2012**, 2, 73.
- [4] a) J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, 77, 3865; b) F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, 7, 3297.
- [5] a) C. Adamo, V. Barone, *J. Chem. Phys.* **1999**, 110, 6158; b) S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, 132, 154104.
- [6] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Wallingford, CT, **2016**.
- [7] a) Y. Yu, T. Tyrikos-Ergas, Y. Zhu, G. Fittolani, V. Bordoni, A. Singhal, R. J. Fair, A. Grafmuller, P. H. Seeberger, M. Delbianco, *Angew. Chem. Int. Ed.* **2019**, 58, 13127; b) S. Gim, G. Fittolani, Y. Yu, Y. Zhu, P. H. Seeberger, Y. Ogawa, M. Delbianco, *Chem. Eur. J.* **2021**, 27, 13139.
- [8] K. Zegelaar-Jaarsveld, S. A. W. Smits, G. A. van der Marel, J. H. van Boom, *Bioorg. Med. Chem.* **1996**, 4, 1819.

Appendix C

The Ferrier Glycosyl Cation in the Gas Phase

This appendix contains the supporting information of the publication related to Chapter 6. It contains mass spectra, infrared spectra, energetics, energy hierarchies, 3D-structures, and xyz-coordinates of Ferrier cations. The publication and the supporting information can be found online: <https://doi.org/10.1021/acs.orglett.0c03301>.^[305]

Supplementary Information

Direct Experimental Characterization of the Ferrier Glycosyl Cation in the Gas Phase

Kim Greis,^{a,b} Carla Kirschbaum,^{a,b} Sabrina Lechnitz,^{a,c} Sandy Gewinner,^b Wieland Schöllkopf,^b Gert von Helden,^b Gerard Meijer,^b Peter H. Seeberger,^{a,c} Kevin Pagel^{*a,b}

- (a) Institut für Chemie und Biochemie, Freie Universität Berlin, Arnimallee 22, 14195 Berlin, Germany
- (b) Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany
- (c) Max-Planck-Institut für Kolloid- und Grenzflächenforschung, Am Mühlenberg 1, 14476 Potsdam, Germany

Correspondence to: kevin.pagel@fu-berlin.de

Experimental Details

Materials

D-Glucal (Sigma-Aldrich), 3,4,6-tri-*O*-acetyl-D-glucal (Acros Organics) and 3,4,6-tri-*O*-acetyl-D-galactal (Senn) were commercially available. D-Galactal was synthesized from 3,4,6-tri-*O*-acetyl-D-galactal according to a previously reported procedure.¹

Mass Spectrometry and Infrared (IR) Spectroscopy

The precursors were dissolved in acetonitrile and water (9:1, V:V) to yield 0.1 mM solutions. The samples were ionized *via* nano electrospray ionization (nESI) with Pd/Pt coated glass capillaries (Sputter Coater HR 208, *Cressington*), which are pulled to a tip with an inner diameter of 1–2 μm using a micropipette puller (Model P-1000, *Sutter Instrument*).

Mass and IR spectra (Figures S1–S4 and 1–3) were recorded using a custom helium droplet instrument, which is briefly described in the following paragraphs. A Z-spray source was used to generate bare Ferrier glycosyl cations in the gas phase, by employing a voltage of 1 kV to the tip of the capillary.

A quadrupole mass filter allows mass-to-charge selection of the Ferrier glycosyl cations. Then, the ions are guided to a quadrupole bender, which can either lead them to a time-of-flight detector, allowing to monitor the ion signal and record mass spectra or to a hexapole ion trap. Subsequently, the hexapole ion trap is filled with the selected ions, which are then thermalized by collisions with helium buffer gas. The ion trap is cooled with liquid nitrogen to 90 K.

A pulsed Even-Lavie valve is used to generate a beam of superfluid helium droplets by expansion of ^4He from high pressure and low temperature into vacuum. The droplets traverse the ion trap, where they picked up ions and rapidly cool them to 0.4 K. Then the ions are guided to a detection region, where the beam of helium droplets overlaps with an IR beam of the Fritz Haber Institute free-electron laser (FHI FEL²), exciting vibrational modes of the analyte ions. The vibrational energy of the embedded ions is dissipated through the helium matrix, which subsequently evaporates until bare analyte ions are released and guided to a time-of-flight detector. The ion count as a function of the wavenumber of the IR pulse leads to an IR spectrum. Due to the multiphoton absorption process, the intensities observed in the IR spectrum scale non-linearly with the energy of the FEL IR beam. As a first-order correction, the ion count is therefore divided by the energy of the IR beam.

Mass Spectra

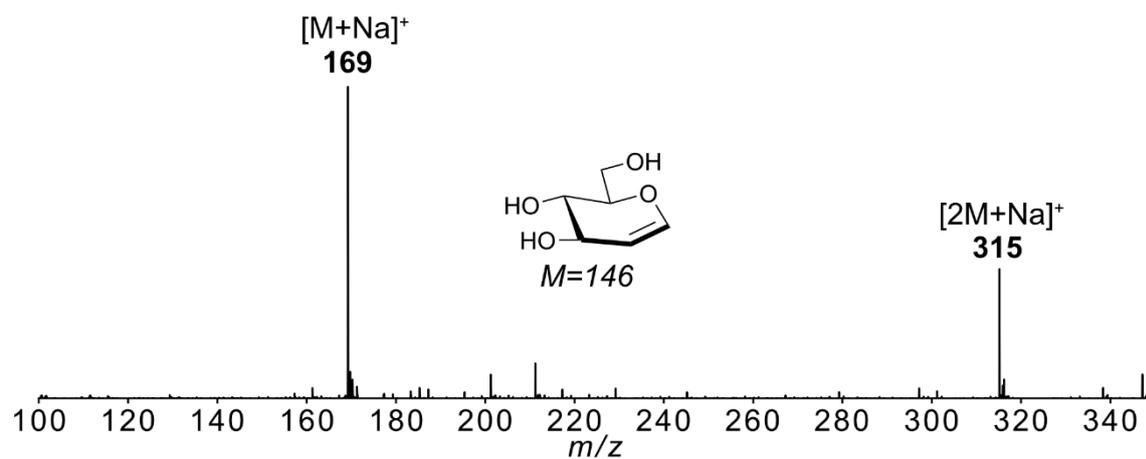


Figure S1. Mass spectrum of D-glucal recorded on the helium droplet instrument. Sodium adducts $[M+Na]^+$ ($m/z = 169$) and sodium bound dimers $[2M+Na]^+$ ($m/z = 315$) are formed.

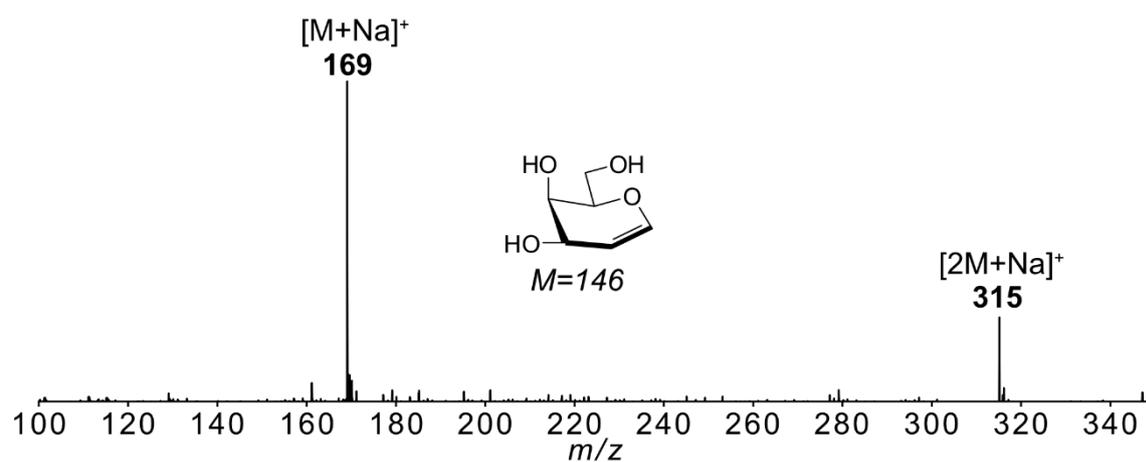


Figure S2. Mass spectrum of D-galactal recorded on the helium droplet instrument. Sodium adducts $[M+Na]^+$ ($m/z = 169$) and sodium bound dimers $[2M+Na]^+$ ($m/z = 315$) are formed.

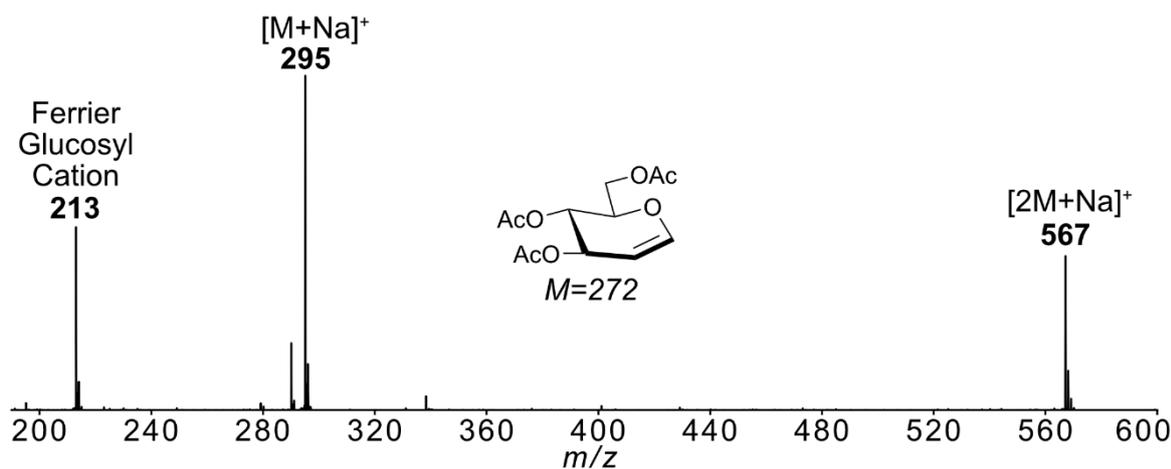


Figure S3. Mass spectrum of 3,4,6-tri-*O*-acetyl-D-glucal recorded on the helium droplet instrument. Ferrier cations ($m/z = 213$) are generated by in-source fragmentation of the precursor ion $[M+Na]^+$ ($m/z = 295$). Besides, ammonium adducts $[M+NH_4]^+$ ($m/z = 290$) and sodium bound dimers $[2M+Na]^+$ ($m/z = 567$) can be observed.

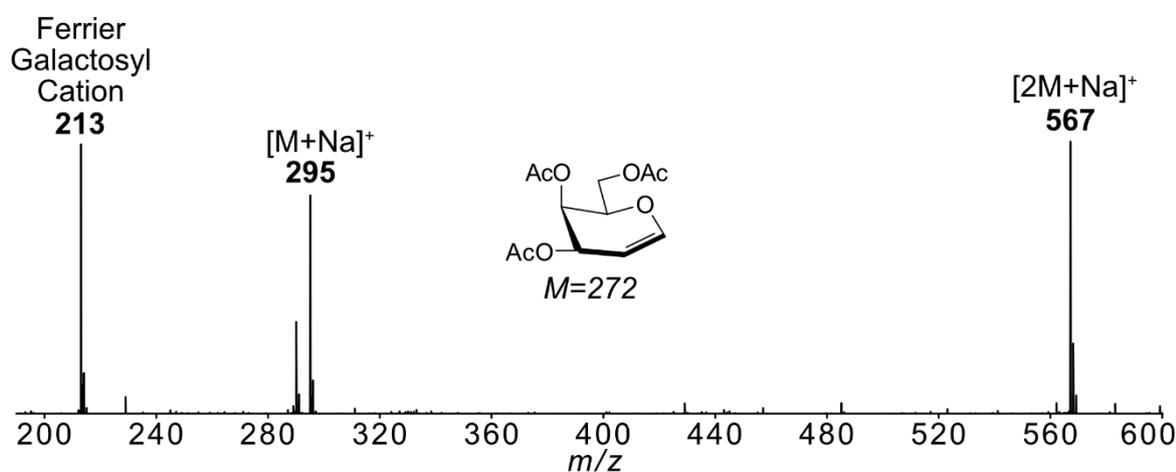


Figure S4. Mass spectrum of 3,4,6-tri-*O*-acetyl-D-galactal recorded on the helium droplet instrument. Ferrier cations ($m/z = 213$) are generated by in-source fragmentation of the precursor ion $[M+Na]^+$ ($m/z = 295$). Besides, ammonium adducts $[M+NH_4]^+$ ($m/z = 290$) and sodium bound dimers $[2M+Na]^+$ ($m/z = 567$) can be observed.

Computational Methods

The conformational space of Ferrier glucosyl and galactosyl cations was sampled using FAFOOM,³ which is a genetic algorithm-based (GA) tool that allows sampling of all rotatable bonds and ring puckers. The tool is interfaced with FHI-aims⁴ (version 171221), which is used for optimizing the geometry of each sampled structure with the dispersion-corrected PBE+vdW^{TS5-6} density functional and *light* basis set settings as implemented in FHI-aims. The settings used for the GA runs are shown in Table S1. For each Ferrier cation, ten individual GA runs were performed leading to 203 structures for the Ferrier glucosyl and 213 structures for the Ferrier galactosyl cation (Figure S5). The conformational search mainly yielded dioxolenium-type **I** structures, featuring a covalent bond between the carbonyl oxygen of the C4-acetyl group and the C3-atom (e.g. Figures S9 and S12) and oxocarbenium-type **II** structures, where no interactions between a protecting group and the pyranose ring can be observed (e.g. Figures S10 and S13). For both the Ferrier glycosyl and galactosyl cation, a further structural motif, where the carbonyl oxygen of the C4-acetyl group bonds covalently to the anomeric carbon, was sampled (e.g. Figures S11 and S14). However, it is rather an artifact, since only two structures exhibiting this structural motif were generated.

Table S1. GA parameters used in initial search.

	Parameter	Value
Molecule	Distance_cutoff_1	1.2
	Distance_cutoff_2	2.15
	Rmsd_cutoff_uniq	0.25
GA settings	Popsize	10
	Prob_for_crossing	0.95
	Prob_for_mut_pyranosering	0.6
	Prob_for_mut_torsion	0.8
	Fitness_sum_limit	1.2
	Selection	Roulette wheel
	Max_mutations_torsion	3

The lowest-energy structures for each type within an energy window of approximately 15 kJ mol⁻¹ above the lowest-energy structure were reoptimized and frequencies computed at PBE0+D3/6-311+G(d,p)⁷⁻⁹ level of theory with default convergence criteria and grid settings using Gaussian 16, Revision A.03.¹⁰ Final energies including zero-point-vibrational energies and free energies at 90 K of the reoptimized structures are listed in Tables S2-S3 and Figure S6. The obtained IR spectra were normalized and scaled by a factor of 0.965 (Figures 2-3 and S7-S8).

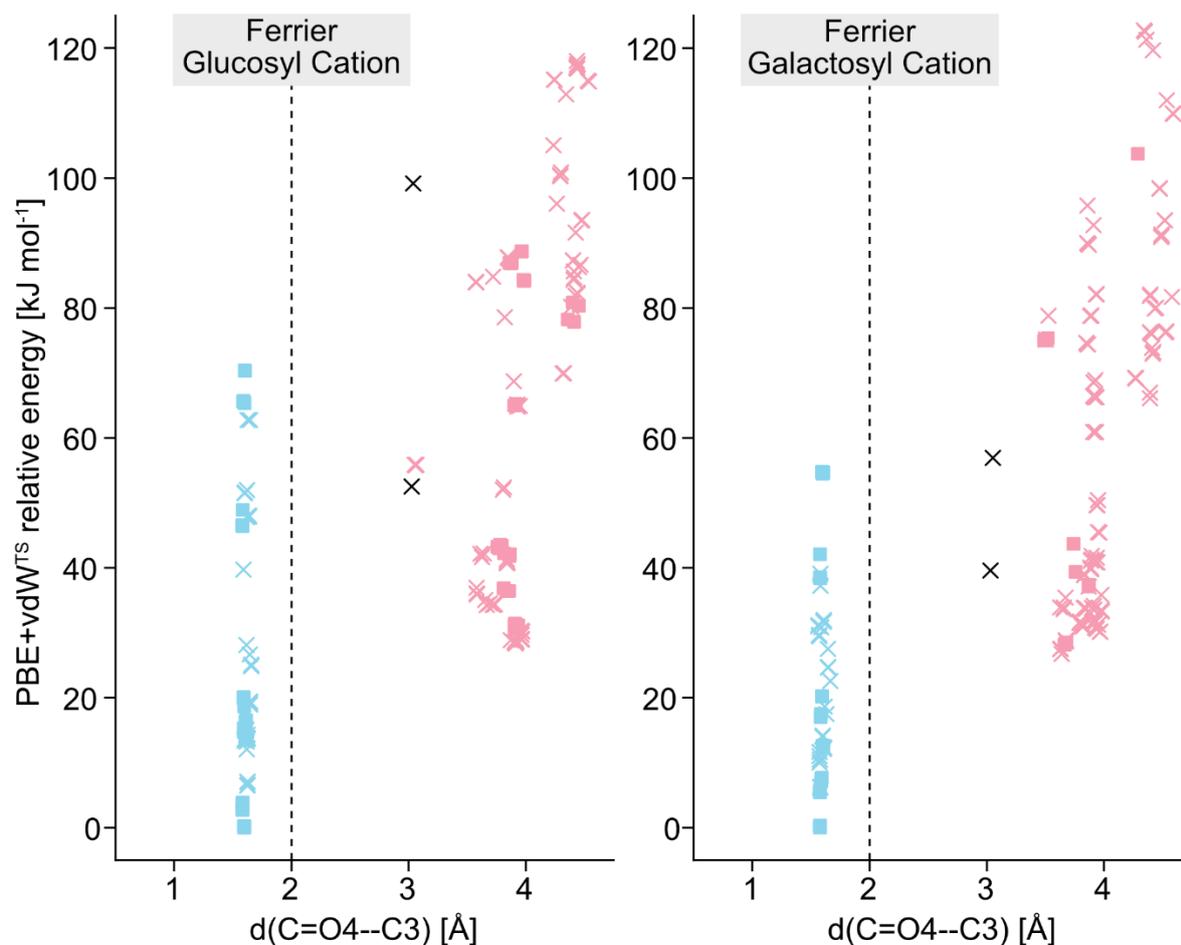


Figure S5. Energy hierarchies of sampled Ferrier glucosyl (left) and galactosyl (right) cations as a function of the distance between the carbonyl oxygen of the C4-acetyl protecting group and the C3-atom. Bold markers indicate structures bearing the same ring pucker as the lowest-energy structure. Blue markers indicate neighboring group participation, while pink markers indicate no participation. Black labels indicate the structures where remote participation between the C4-acetyl group and the anomeric carbon can be observed.

Table S2. List of conformations of the Ferrier glucosyl cation reoptimized at PBE0+D3/6-311+G(d,p) level of theory. Ring puckers, bond distances between the carbonyl oxygen of the C4-acetyl group and the C3-atom, energies (ΔE , including zero-point-vibrational energy) and free energies (ΔF) at 90 K are assigned to each structure. The spectra of the structures labelled with an asterisk (*) are represented in the main manuscript.

ID	Ring Pucker	d(C4=O—C3) [Å]	$\Delta E(\text{PBE0+D3})$ [kJ mol ⁻¹]	$\Delta F(\text{PBE0+D3})$ [kJ mol ⁻¹]
C4_NGP/conf_00	5E	1.53	-0.09977	0.41142
C4_NGP/conf_01*	5E	1.53	0	0
C4_NGP/conf_02	5E	1.52	1.87461	2.52449
C4_NGP/conf_03	5E	1.52	1.87461	2.53162
C4_NGP/conf_04	E5	1.53	4.79941	6.21489
C4_NGP/conf_05	E5	1.53	13.12225	13.08896
C4_NGP/conf_06	E5	1.53	13.11962	13.08396
C4_NGP/conf_07	5H4	1.53	14.58465	14.25733
C4_NGP/conf_08	5H4	1.53	14.58465	14.28111
C4_NGP/conf_09	E5	1.53	11.46293	12.78806
oxocarbenium/conf_00	5H4	2.97	30.78924	29.14393
oxocarbenium /conf_01*	5E	2.99	29.92282	28.19905
oxocarbenium /conf_02	5E	3.99	31.23557	29.66793
oxocarbenium /conf_03	5H4	3.99	37.91747	35.8434
oxocarbenium /conf_04	5H4	3.99	37.91747	35.83865
oxocarbenium /conf_05	4H5	2.96	39.05694	37.9569
oxocarbenium /conf_06	5E	2.96	42.50422	40.24311
oxocarbenium /conf_07	4H5	3.51	39.05694	37.96165
oxocarbenium /conf_08	4H5	3.90	42.84816	40.90961
oxocarbenium /conf_09	5E	3.51	42.8114	38.42154
C4_RP/conf_00	5S1	3.55	48.31708	48.96854
C4_RP/conf_01	B1,4	3.97	98.07555	98.07476

Table S3. List of conformations of the Ferrier galactosyl cation reoptimized at PBE0+D3/6-311+G(d,p) level of theory. Ring puckers, bond distances between the carbonyl oxygen of the C4-acetyl group and the C3-atom, energies (ΔE , including zero-point-vibrational energy) and free energies (ΔF) at 90 K are assigned to each structure. The spectra of the structures labelled with an asterisk (*) are represented in the main manuscript.

ID	Ring Pucker	d(C4=O—C3) [Å]	$\Delta E(\text{PBE0+D3})$ [kJ mol ⁻¹]	$\Delta F(\text{PBE0+D3})$ [kJ mol ⁻¹]
C4_NGP/conf_00*	E5	1.52	0	0
C4_NGP/conf_01	E5	1.52	4.80992	5.42572
C4_NGP/conf_02	5E	1.50	7.08622	8.14981
C4_NGP/conf_03	E5	1.52	7.17287	6.52774
C4_NGP/conf_04	E5	1.52	7.17024	6.52036
C4_NGP/conf_05	5E	1.51	11.59683	13.08918
C4_NGP/conf_06	5E	1.51	11.59421	13.08339
C4_NGP/conf_07	5E	1.52	13.797	14.99453
C4_NGP/conf_08	E5	1.52	14.06743	12.86198
C4_NGP/conf_09	OH5	1.52	13.69986	14.61682
oxocarbenium/conf_00	4H5	2.98	33.75343	32.61931
oxocarbenium /conf_01	4H5	3.01	32.97891	31.67994
oxocarbenium /conf_02*	5E	3.61	31.3931	29.87381
oxocarbenium /conf_03	4H5	3.63	36.42356	33.28749
oxocarbenium /conf_04	5E	4.02	31.96021	30.41318
oxocarbenium /conf_05	5H4	3.84	36.49182	35.01771
oxocarbenium /conf_06	5H4	4.01	37.57091	36.26401
oxocarbenium /conf_07	4H5	3.99	41.48815	39.49175
oxocarbenium /conf_08	4H5	3.99	40.30405	37.29558
oxocarbenium /conf_09	5H4	3.11	36.4892	35.01032
C4_RP/conf_00	1,4B	3.88	31.36947	33.5434
C4_RP/conf_01	1,4B	3.99	49.51693	51.33026

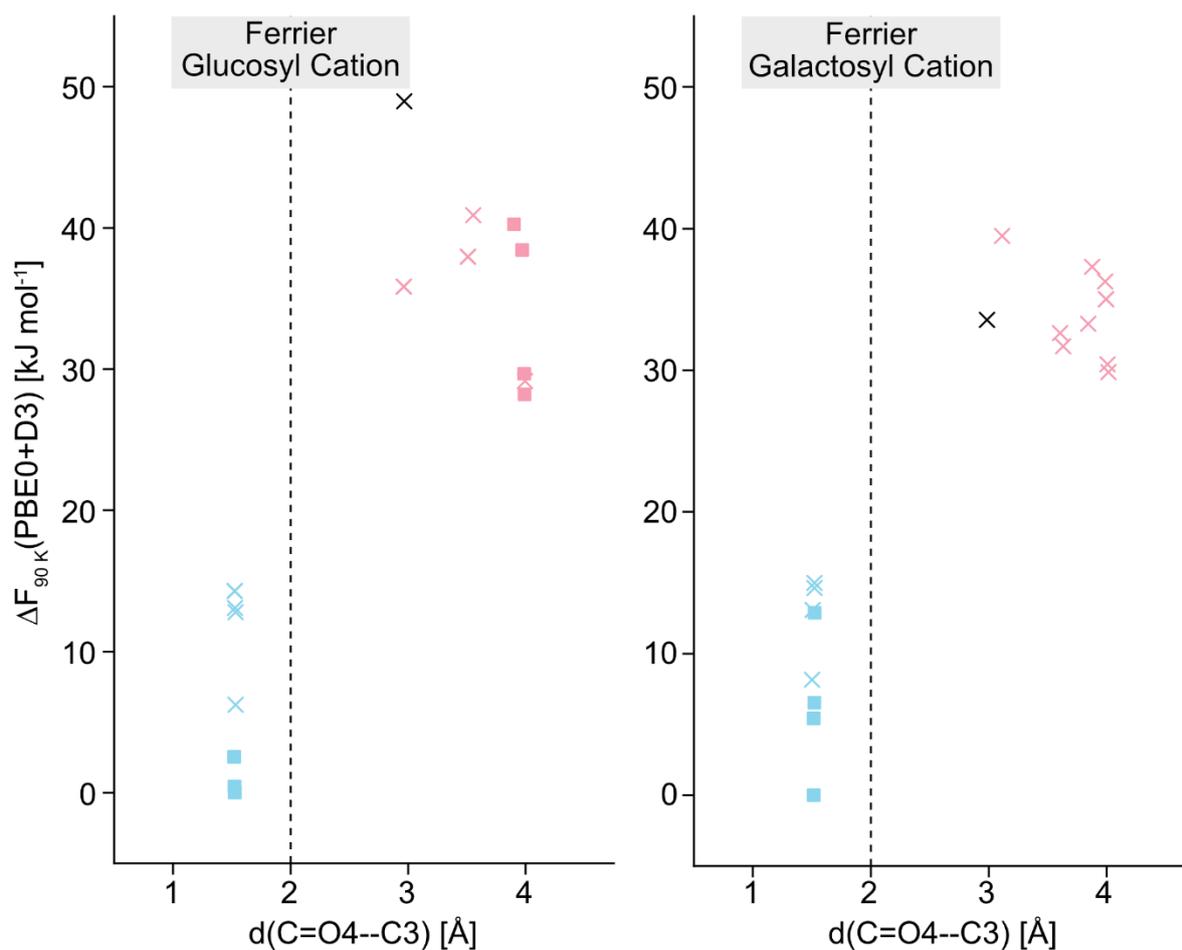


Figure S6. Energy hierarchies of reoptimized Ferrier glucosyl (left) and galactosyl (right) cations as a function of the distance between the carbonyl oxygen of the C4-acetyl protecting group and the C3-atom. Bold markers indicate structures bearing the same ring pucker as the lowest-energy structure. Blue markers indicate neighboring group participation, while pink markers indicate no participation. Black labels indicate the structures where remote participation between the C4-acetyl group and the anomeric carbon can be observed.

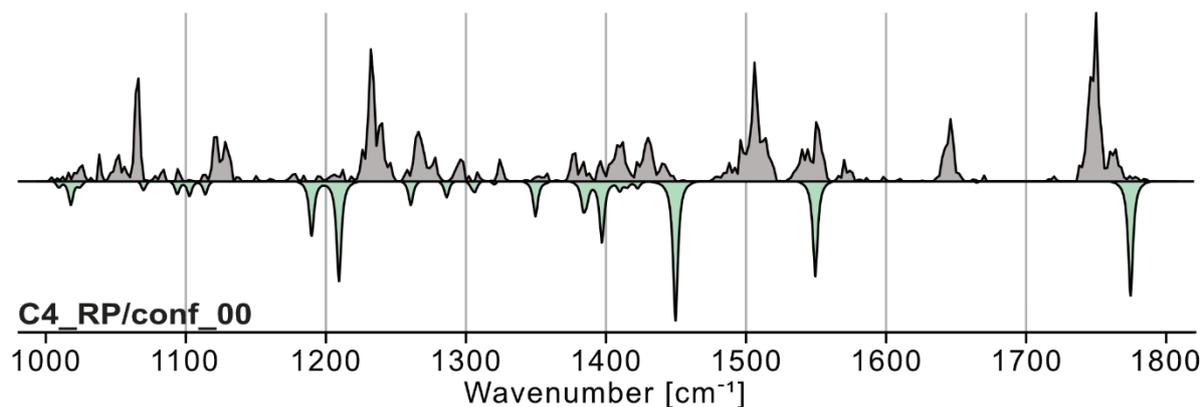


Figure S7. Infrared spectra of Ferrier glucosyl cations generated from 3,4,6-tri-*O*-acetyl-D-glucal precursor. The experimental IR spectrum is shown as grey trace, while the computed spectrum of a dioxolenium structure exhibiting remote participation ($+49 \text{ kJ mol}^{-1}$) is represented in the green inverted trace below. The computed spectra of species exhibiting remote participation of the C4-acetyl group and the anomeric carbon do not agree with the experimental spectrum.

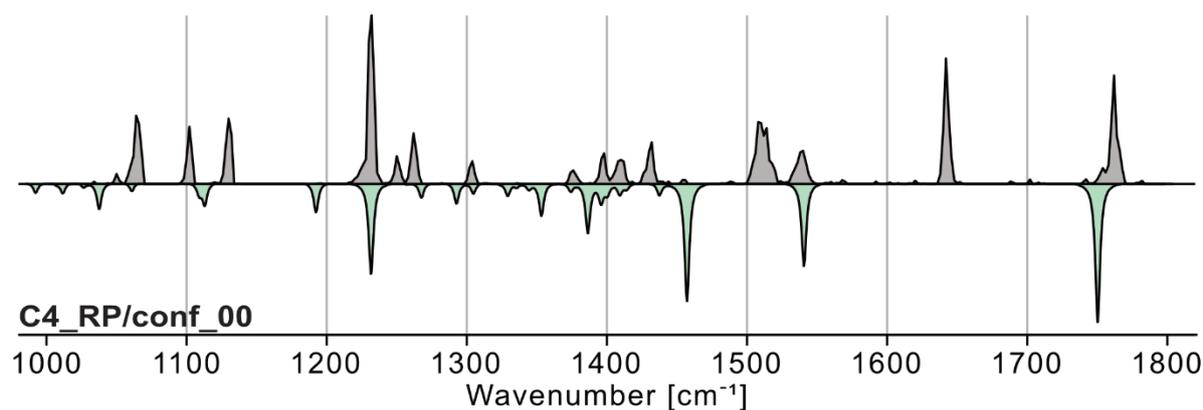


Figure S8. Infrared spectra of Ferrier galactosyl cations generated from 3,4,6-tri-*O*-acetyl-D-galactal precursor. The experimental IR spectrum is shown as grey trace, while the computed spectrum of a dioxolenium structure exhibiting remote participation ($+34 \text{ kJ mol}^{-1}$) is represented in the green inverted trace below. The computed spectra of species exhibiting remote participation of the C4-acetyl group and the anomeric carbon do not agree with the experimental spectrum.

Low-energy Structures

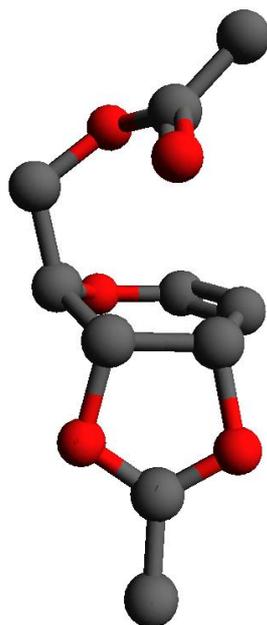


Figure S9. Calculated structure for Ferrier glucosyl cation (C4_NGP/conf_01) with C4-acetyl neighboring group participation at the C3-atom. Hydrogens are omitted for clarity.

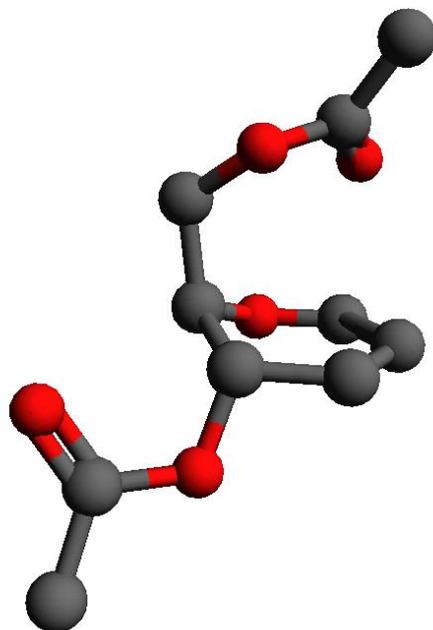


Figure S10. Calculated structure for Ferrier glucosyl cation (oxocarbenium/conf_01) with no participation of acetyl protecting groups. Hydrogens are omitted for clarity.

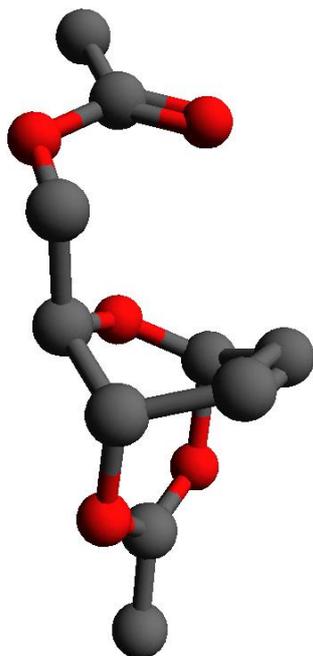


Figure S11. Calculated structure for Ferrier glucosyl cation (C4_RP/conf_00) with C4-acetyl remote participation at the anomeric carbon. Hydrogens are omitted for clarity.

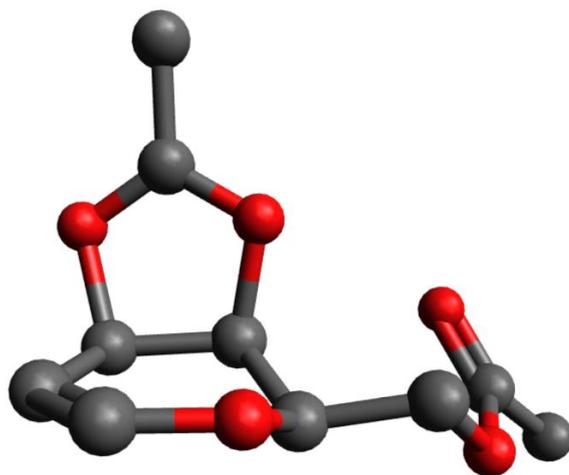


Figure S12. Calculated structure for Ferrier galactosyl cation (C4_NGP/conf_00) with C4-acetyl neighboring group participation at the C3-atom. Hydrogens are omitted for clarity.

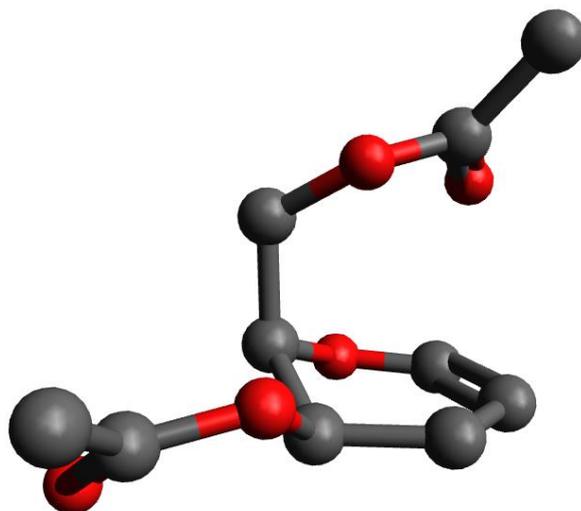


Figure S13. Calculated structure for Ferrier galactosyl cation (oxocarbenium/conf_02) with no participation of acetyl protecting groups. Hydrogens are omitted for clarity.

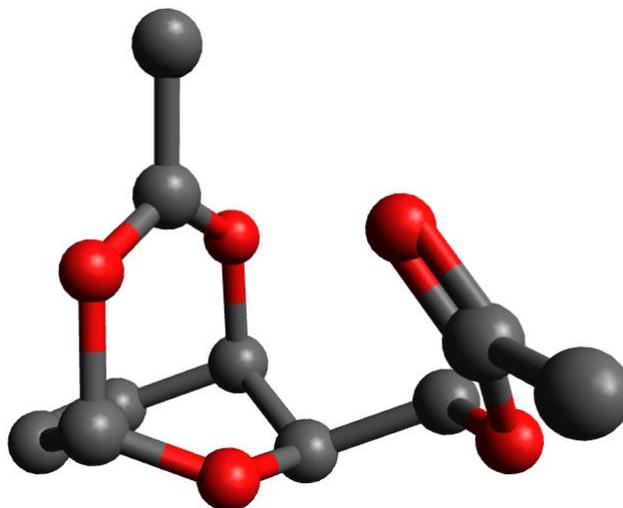


Figure S14. Calculated structure for Ferrier galactosyl cation (C4_RP/conf_00) with C4-acetyl remote participation at the anomeric carbon. Hydrogens are omitted for clarity.

xyz-Coordinates of Reoptimized Structures

Ferrier_Glucosyl_Cation/C4_NGP/conf_00

Charge=+1, Multiplicity=+1

C	-1.58065	1.55013	-0.73163
C	-0.13036	1.43323	-0.26269
C	0.64202	0.23038	-0.77502
C	0.64862	-1.01440	0.11688
C	0.17130	-0.76505	1.48249
C	-0.11441	0.47547	1.89786
O	-0.07453	1.58636	1.14991
O	-2.50692	0.71574	-0.05216
C	-2.57308	-0.56761	-0.44893
O	-1.78645	-1.02750	-1.24335
C	-3.69846	-1.29286	0.20035
H	-1.91460	2.56228	-0.50647
H	-1.63138	1.38172	-1.81120
H	0.37824	2.31336	-0.66830
H	0.23307	-1.89127	-0.37083
H	0.08614	-1.59244	2.17412
H	-0.40562	0.69661	2.91893
H	-3.62373	-2.35724	-0.01125
H	-4.64386	-0.90853	-0.19305
H	-3.69756	-1.10858	1.27637
O	2.07396	0.60597	-0.81810
C	2.77639	-0.32726	-0.29226
O	2.14256	-1.31587	0.18364
C	4.23818	-0.22906	-0.22360
H	4.67812	-1.22500	-0.18503
H	4.49300	0.29792	0.70400
H	4.61988	0.34845	-1.06492
H	0.36206	0.00736	-1.80261

Ferrier_Glucosyl_Cation/C4_NGP/conf_01

Charge=+1, Multiplicity=+1

C	-1.58060	1.54116	-0.75051
C	-0.13409	1.43385	-0.26780
C	0.64690	0.22940	-0.76277
C	0.65143	-1.00671	0.14131
C	0.15935	-0.74791	1.49963
C	-0.13482	0.49516	1.90120
O	-0.09137	1.59948	1.14381
O	-2.50966	0.70880	-0.07229
C	-2.56701	-0.57822	-0.45819
O	-1.77156	-1.04193	-1.24158
C	-3.69487	-1.30250	0.18787
H	-1.92052	2.55381	-0.53683
H	-1.62116	1.36353	-1.82904
H	0.37491	2.31226	-0.67663
H	0.24543	-1.89028	-0.34245
H	0.07025	-1.56960	2.19749

H	-0.43686	0.72428	2.91734
H	-3.61505	-2.36812	-0.01544
H	-4.63864	-0.92407	-0.21500
H	-3.70231	-1.11057	1.26254
O	2.07735	0.61013	-0.79914
C	2.78044	-0.31992	-0.26756
O	2.14684	-1.29984	0.22522
C	4.24544	-0.24561	-0.25491
H	4.64893	-0.89140	0.52302
H	4.56327	0.78950	-0.12341
H	4.61133	-0.58926	-1.22954
H	0.37446	-0.00538	-1.78978

Ferrier_Glucosyl_Cation/C4_NGP/conf_02

Charge=+1, Multiplicity=+1

C	-1.66419	1.27452	-0.41882
C	-0.26362	0.93686	-0.94378
C	0.33043	-0.30473	-0.32203
C	0.85783	-0.12326	1.09731
C	1.20616	1.27277	1.41030
C	1.07566	2.23769	0.49228
O	0.54645	2.08752	-0.73411
O	-2.62425	0.27315	-0.71909
C	-2.81839	-0.68178	0.21955
O	-2.07874	-0.80131	1.16603
C	-4.01797	-1.51082	-0.07744
H	-1.63495	1.46310	0.65646
H	-1.99344	2.17344	-0.93876
H	-0.30539	0.79605	-2.02597
H	0.27071	-0.64824	1.84641
H	1.61261	1.51110	2.38427
H	1.39136	3.26042	0.66666
H	-3.99544	-2.42078	0.51858
H	-4.07869	-1.73903	-1.14258
H	-4.91102	-0.93462	0.18352
O	1.54872	-0.68630	-1.06629
C	2.44523	-1.04663	-0.22435
O	2.15641	-0.90949	1.00470
C	3.74396	-1.56034	-0.66847
H	4.16619	-2.21559	0.09296
H	4.41431	-0.70071	-0.79246
H	3.64677	-2.06530	-1.62883
H	-0.34233	-1.15418	-0.41742

Ferrier_Glucosyl_Cation/C4_NGP/conf_03

Charge=+1, Multiplicity=+1

C	-1.66433	1.27442	-0.41891
C	-0.26364	0.93702	-0.94376
C	0.33059	-0.30455	-0.32223
C	0.85780	-0.12334	1.09721
C	1.20629	1.27260	1.41036

C	1.07557	2.23771	0.49255
O	0.54600	2.08791	-0.73372
O	-2.62438	0.27302	-0.71908
C	-2.81835	-0.68200	0.21958
O	-2.07850	-0.80154	1.16591
C	-4.01803	-1.51087	-0.07729
H	-1.63507	1.46314	0.65635
H	-1.99365	2.17326	-0.93895
H	-0.30525	0.79641	-2.02599
H	0.27049	-0.64843	1.84608
H	1.61314	1.51074	2.38422
H	1.39132	3.26040	0.66708
H	-3.99642	-2.42022	0.51969
H	-4.07806	-1.74013	-1.14225
H	-4.91103	-0.93390	0.18221
O	1.54896	-0.68595	-1.06643
C	2.44536	-1.04648	-0.22443
O	2.15635	-0.90970	1.00463
C	3.74403	-1.56037	-0.66853
H	4.16579	-2.21631	0.09259
H	4.41478	-0.70095	-0.79172
H	3.64691	-2.06470	-1.62923
H	-0.34217	-1.15402	-0.41779

Ferrier_Glucosyl_Cation/C4_NGP/conf_04

Charge=+1, Multiplicity=+1

C	1.16251	-1.04364	1.16857
C	-0.09076	-0.96471	0.31792
C	-1.23533	-0.19713	0.97352
C	-2.38682	-0.09107	-0.00548
C	-2.51189	-1.25614	-0.88610
C	-1.65739	-2.28419	-0.76055
O	-0.54255	-2.26785	-0.01637
O	2.34372	-1.04200	0.37762
C	2.64746	0.12487	-0.22419
O	1.90883	1.08255	-0.17321
C	3.96278	0.06124	-0.91826
H	1.19383	-0.20311	1.86652
H	1.17840	-1.98298	1.72137
H	0.14175	-0.45096	-0.62232
H	-3.33374	0.18911	0.46112
H	-3.38957	-1.36347	-1.50868
H	-1.81120	-3.24261	-1.24360
H	4.08648	0.93315	-1.55693
H	4.04562	-0.86007	-1.49737
H	4.75829	0.04548	-0.16748
O	-0.80365	1.18609	1.11444
C	-1.16426	1.81643	0.04937
O	-1.97467	1.20310	-0.71537
C	-0.62804	3.14196	-0.24648
H	-1.17526	3.61089	-1.06092

H	0.42609	2.99314	-0.51649
H	-0.64785	3.75457	0.65782
H	-1.51579	-0.56399	1.96015

Ferrier_Glucosyl_Cation/C4_NGP/conf_05

Charge=+1, Multiplicity=+1

C	1.35698	0.10906	1.42575
C	0.07597	0.67495	0.84979
C	-0.43209	-0.02378	-0.41208
C	-1.63824	0.68090	-1.01484
C	-1.71406	2.10846	-0.70576
C	-0.73304	2.69443	-0.00343
O	0.26780	2.05782	0.61369
O	2.47858	0.28526	0.57518
C	2.78309	-0.73436	-0.26399
O	2.03768	-1.66552	-0.43507
C	4.11429	-0.53209	-0.90096
H	1.59114	0.65745	2.33895
H	1.21291	-0.94979	1.65571
H	-0.70893	0.57548	1.61689
H	-1.75618	0.47623	-2.08134
H	-2.46248	2.72107	-1.18986
H	-0.64838	3.77054	0.09966
H	4.22635	-1.21067	-1.74396
H	4.89161	-0.74657	-0.16145
H	4.24049	0.50586	-1.21280
O	-0.98539	-1.30999	0.01567
C	-2.26644	-1.22254	0.03162
O	-2.77129	-0.14383	-0.40774
C	-3.09069	-2.33541	0.51193
H	-4.03655	-1.95960	0.90139
H	-2.54697	-2.91851	1.25434
H	-3.30667	-2.98384	-0.34627
H	0.34996	-0.24533	-1.13309

Ferrier_Glucosyl_Cation/C4_NGP/conf_06

Charge=+1, Multiplicity=+1

C	1.35696	0.10942	1.42573
C	0.07593	0.67518	0.84979
C	-0.43213	-0.02375	-0.41200
C	-1.63836	0.68082	-1.01471
C	-1.71438	2.10838	-0.70562
C	-0.73333	2.69453	-0.00357
O	0.26774	2.05801	0.61345
O	2.47850	0.28553	0.57498
C	2.78304	-0.73441	-0.26400
O	2.03769	-1.66563	-0.43473
C	4.11417	-0.53213	-0.90103
H	1.59122	0.65794	2.33883
H	1.21295	-0.94939	1.65588
H	-0.70898	0.57579	1.61690

H	-1.75627	0.47631	-2.08125
H	-2.46294	2.72085	-1.18969
H	-0.64861	3.77063	0.09944
H	4.22621	-1.21080	-1.74398
H	4.89157	-0.74655	-0.16160
H	4.24023	0.50578	-1.21303
O	-0.98515	-1.30991	0.01569
C	-2.26627	-1.22279	0.03163
O	-2.77126	-0.14409	-0.40768
C	-3.09018	-2.33582	0.51189
H	-4.03640	-1.96040	0.90083
H	-2.54650	-2.91849	1.25466
H	-3.30547	-2.98461	-0.34621
H	0.34996	-0.24497	-1.13306

Ferrier_Glucosyl_Cation/C4_NGP/conf_07

Charge=+1, Multiplicity=+1

C	1.44474	-1.45076	-0.82669
C	0.07952	-1.47643	-0.15790
C	-0.84751	-0.46673	-0.79868
C	-0.72880	0.96018	-0.26524
C	-0.18309	1.03431	1.09567
C	0.19813	-0.07392	1.74500
O	0.18782	-1.31646	1.25183
O	1.91142	-0.11900	-1.01253
C	2.88628	0.31725	-0.14981
O	3.27768	-0.35003	0.76139
C	3.32904	1.69459	-0.51030
H	2.14844	-2.02028	-0.21699
H	1.36894	-1.89779	-1.82108
H	-0.34537	-2.47398	-0.29825
H	-0.31224	1.66333	-0.98333
H	-0.10432	1.99640	1.58406
H	0.56573	-0.05566	2.76485
H	4.20352	1.95904	0.08038
H	2.52263	2.40176	-0.29332
H	3.55243	1.76059	-1.57650
O	-2.24678	-0.80528	-0.48970
C	-2.88794	0.27519	-0.22914
O	-2.20405	1.34357	-0.20009
C	-4.32597	0.25601	0.04973
H	-4.76762	1.22098	-0.19751
H	-4.45020	0.08666	1.12682
H	-4.80844	-0.56035	-0.48662
H	-0.76876	-0.52125	-1.88566

Ferrier_Glucosyl_Cation/C4_NGP/conf_08

Charge=+1, Multiplicity=+1

C	1.44477	-1.45110	-0.82613
C	0.07940	-1.47643	-0.15754
C	-0.84739	-0.46676	-0.79878

C	-0.72889	0.96030	-0.26554
C	-0.18266	1.03479	1.09516
C	0.19825	-0.07337	1.74487
O	0.18731	-1.31605	1.25213
O	1.91163	-0.11952	-1.01248
C	2.88646	0.31706	-0.14989
O	3.27792	-0.34989	0.76154
C	3.32883	1.69448	-0.51070
H	2.14832	-2.02048	-0.21611
H	1.36903	-1.89863	-1.82032
H	-0.34560	-2.47395	-0.29782
H	-0.31272	1.66346	-0.98387
H	-0.10326	1.99706	1.58313
H	0.56600	-0.05481	2.76466
H	4.20435	1.95865	0.07856
H	2.52284	2.40156	-0.29169
H	3.55012	1.76111	-1.57729
O	-2.24686	-0.80539	-0.48993
C	-2.88794	0.27496	-0.22915
O	-2.20413	1.34344	-0.20001
C	-4.32601	0.25595	0.04991
H	-4.76774	1.22053	-0.19871
H	-4.45011	0.08826	1.12728
H	-4.80848	-0.56121	-0.48522
H	-0.76859	-0.52155	-1.88575

Ferrier_Glucosyl_Cation/C4_NGP/conf_09

Charge=+1, Multiplicity=+1

C	0.66312	-2.02042	0.03709
C	-0.52249	-1.09162	-0.19533
C	-0.79161	-0.06684	0.90767
C	-2.07213	0.69728	0.61203
C	-3.08245	-0.07439	-0.11165
C	-2.83110	-1.34608	-0.45274
O	-1.64445	-1.95257	-0.32982
O	1.82400	-1.38724	0.55887
C	2.45873	-0.49930	-0.23491
O	1.97124	-0.08633	-1.25842
C	3.78608	-0.12205	0.32757
H	0.38606	-2.76748	0.78181
H	0.89895	-2.52343	-0.90359
H	-0.40039	-0.54597	-1.13812
H	-2.47711	1.20909	1.48794
H	-4.07890	0.32797	-0.23178
H	-3.59137	-2.01742	-0.83597
H	4.17843	0.74711	-0.19647
H	4.47569	-0.96055	0.19232
H	3.71291	0.06533	1.40024
O	0.22399	0.97763	0.80004
C	-0.24878	1.91733	0.05519
O	-1.48954	1.85748	-0.20200

C	0.63142	2.95661	-0.47649
H	0.05452	3.76073	-0.92751
H	1.27282	2.47303	-1.22520
H	1.28110	3.32870	0.31925
H	-0.73943	-0.47195	1.91726

Ferrier_Glucosyl_Cation/oxocarbenium/conf_00

Charge=+1, Multiplicity=+1

C	1.00141	-0.78637	1.46235
C	-0.20559	0.05628	1.05772
C	-0.71959	-0.12248	-0.36032
C	0.05686	0.68403	-1.34669
C	0.71412	1.79699	-0.94073
C	0.62566	2.16842	0.41735
O	0.13214	1.45011	1.34379
O	2.24326	-0.36164	0.91835
C	2.56963	-0.80467	-0.31986
O	1.75468	-1.32939	-1.03703
C	3.99935	-0.54934	-0.63821
H	1.13218	-0.71422	2.54156
H	0.78764	-1.82291	1.18651
H	-1.00795	-0.17159	1.75946
H	-0.01077	0.41370	-2.39514
H	1.22253	2.46493	-1.62459
H	0.94890	3.14805	0.76647
H	4.16252	-0.64751	-1.70948
H	4.61049	-1.29181	-0.11546
H	4.30988	0.43354	-0.27954
O	-2.05438	0.34251	-0.50928
C	-3.02228	-0.55841	-0.10747
O	-2.71805	-1.59888	0.39734
C	-4.38615	-0.03871	-0.39203
H	-5.12475	-0.69332	0.06549
H	-4.54005	-0.01055	-1.47440
H	-4.49193	0.98176	-0.01893
H	-0.66035	-1.18622	-0.59984

Ferrier_Glucosyl_Cation/oxocarbenium/conf_01

Charge=+1, Multiplicity=+1

C	0.77273	-1.22882	0.99098
C	-0.27021	-0.12971	0.90982
C	-0.87301	0.06469	-0.47770
C	-0.13610	1.06648	-1.29676
C	0.80094	1.86188	-0.75486
C	0.96011	1.85704	0.66418
O	0.31879	1.09708	1.45065
O	1.67067	-1.32404	-0.11974
C	2.77303	-0.55261	-0.12565
O	2.87660	0.43838	0.56498
C	3.80198	-1.05949	-1.07088
H	1.33986	-1.12789	1.91980

H	0.23612	-2.17897	0.99555
H	-1.05181	-0.36775	1.62971
H	-0.40959	1.14707	-2.34429
H	1.36826	2.58626	-1.32512
H	1.52601	2.63464	1.17245
H	4.56480	-0.30112	-1.23248
H	3.33898	-1.35921	-2.01326
H	4.26392	-1.95375	-0.64131
O	-2.22036	0.50403	-0.44314
C	-3.12586	-0.48986	-0.14596
O	-2.75188	-1.59741	0.11691
C	-4.52222	0.01678	-0.20277
H	-5.20253	-0.75757	0.14473
H	-4.76610	0.28870	-1.23337
H	-4.62323	0.91797	0.40557
H	-0.82357	-0.91108	-0.97762

Ferrier_Glucosyl_Cation/oxocarbenium/conf_02

Charge=+1, Multiplicity=+1

C	0.92324	-0.90633	-0.84869
C	-0.13473	-0.42517	0.13966
C	-1.21992	0.34652	-0.59882
C	-0.77157	1.73273	-0.92256
C	0.32558	2.26109	-0.34974
C	0.88968	1.57103	0.76366
O	0.51788	0.41139	1.13974
O	2.10547	-1.36271	-0.20568
C	3.07454	-0.45703	0.01039
O	2.89446	0.73768	-0.12483
C	4.35458	-1.09404	0.41377
H	0.51324	-1.76388	-1.38446
H	1.18309	-0.12425	-1.56674
H	-0.54485	-1.25022	0.71716
H	-1.38128	2.30948	-1.61134
H	0.69635	3.25238	-0.57644
H	1.52269	2.08535	1.48204
H	5.05443	-0.33590	0.75751
H	4.77573	-1.62005	-0.44804
H	4.17369	-1.83952	1.19099
O	-2.43081	0.46703	0.11965
C	-3.23508	-0.65265	0.06182
O	-2.87651	-1.63143	-0.52648
C	-4.50761	-0.43790	0.79939
H	-5.08958	-1.35676	0.79133
H	-5.07342	0.36698	0.32302
H	-4.29901	-0.12748	1.82564
H	-1.40738	-0.20133	-1.53317

Ferrier_Glucosyl_Cation/oxocarbenium/conf_03

Charge=+1, Multiplicity=+1

C	-1.73533	1.03500	-0.91381
---	----------	---------	----------

C	-0.20369	0.96453	-1.02900
C	0.46760	-0.03728	-0.12636
C	0.38005	0.40706	1.30359
C	0.42802	1.73901	1.57197
C	0.48963	2.63131	0.48801
O	0.26133	2.31906	-0.73017
O	-2.32860	-0.24341	-1.01050
C	-2.49463	-0.91355	0.15794
O	-2.02500	-0.50258	1.19263
C	-3.29822	-2.15010	-0.02123
H	-2.01241	1.50293	0.03462
H	-2.12714	1.62095	-1.74425
H	0.11291	0.79926	-2.05885
H	0.42275	-0.34257	2.08510
H	0.50109	2.13641	2.57644
H	0.74848	3.68143	0.61283
H	-3.23410	-2.76339	0.87488
H	-2.95890	-2.70416	-0.89849
H	-4.34103	-1.86946	-0.19725
O	1.83946	-0.09653	-0.50784
C	2.50225	-1.22245	-0.07125
O	1.95797	-2.01328	0.64311
C	3.88887	-1.27190	-0.60718
H	4.41188	-2.12175	-0.17434
H	4.41560	-0.34296	-0.37826
H	3.85674	-1.37091	-1.69542
H	0.01539	-1.02169	-0.25093

Ferrier_Glucosyl_Cation/oxocarbenium/conf_04
Charge=+1, Multiplicity=+1

C	-1.73560	1.03515	-0.91345
C	-0.20401	0.96440	-1.02892
C	0.46739	-0.03745	-0.12638
C	0.38004	0.40689	1.30361
C	0.42837	1.73885	1.57190
C	0.48997	2.63111	0.48789
O	0.26131	2.31886	-0.73022
O	-2.32917	-0.24311	-1.01026
C	-2.49501	-0.91346	0.15810
O	-2.02517	-0.50270	1.19276
C	-3.29775	-2.15050	-0.02159
H	-2.01244	1.50296	0.03511
H	-2.12744	1.62133	-1.74371
H	0.11238	0.79901	-2.05882
H	0.42244	-0.34272	2.08513
H	0.50160	2.13627	2.57635
H	0.74903	3.68118	0.61263
H	-3.24689	-2.75611	0.88052
H	-2.94502	-2.71267	-0.88850
H	-4.33739	-1.87124	-0.21691
O	1.83918	-0.09658	-0.50832

C	2.50256	-1.22163	-0.07065
O	1.95900	-2.01165	0.64520
C	3.88839	-1.27236	-0.60853
H	4.41697	-2.11201	-0.16271
H	4.41080	-0.33667	-0.39859
H	3.85332	-1.39189	-1.69473
H	0.01521	-1.02190	-0.25087

Ferrier_Glucosyl_Cation/oxocarbenium/conf_05

Charge=+1, Multiplicity=+1

C	-0.07726	-1.40594	-1.18758
C	-1.22446	-0.47876	-0.88633
C	-0.89325	0.64420	0.07881
C	-2.12504	1.33938	0.52504
C	-3.28764	0.65295	0.64692
C	-3.26990	-0.71976	0.31312
O	-2.32917	-1.29181	-0.31864
O	0.58594	-1.78158	0.00157
C	1.91324	-1.43379	0.09390
O	2.46384	-0.80624	-0.76592
C	2.50037	-1.89378	1.38023
H	-0.45561	-2.30474	-1.68074
H	0.61488	-0.89094	-1.85940
H	-1.63472	-0.08496	-1.82358
H	-2.04557	2.38104	0.82627
H	-4.18976	1.08055	1.06566
H	-4.07951	-1.39607	0.58642
H	3.57599	-2.01374	1.26165
H	2.03781	-2.81805	1.72470
H	2.32263	-1.11239	2.12654
O	0.00579	1.50925	-0.56715
C	1.16338	1.79659	0.14386
O	1.30297	1.41235	1.26784
C	2.12052	2.56529	-0.68692
H	2.78841	3.13202	-0.04065
H	1.60712	3.21518	-1.39492
H	2.71278	1.83511	-1.24914
H	-0.40921	0.24292	0.98943

Ferrier_Glucosyl_Cation/oxocarbenium/conf_06

Charge=+1, Multiplicity=+1

C	-0.24295	-1.92916	-0.04805
C	0.54214	-0.74084	0.52433
C	0.87912	0.30994	-0.52079
C	2.15990	0.05124	-1.22508
C	3.06014	-0.83595	-0.74406
C	2.81828	-1.39820	0.53145
O	1.72547	-1.30931	1.16787
O	-1.38830	-1.45828	-0.71864
C	-2.41473	-1.04042	0.09802
O	-2.29185	-1.04295	1.29005

C	-3.58633	-0.59058	-0.69691
H	0.34617	-2.46810	-0.79406
H	-0.51044	-2.59901	0.77206
H	-0.02178	-0.30617	1.34888
H	2.37687	0.64833	-2.10616
H	4.01639	-1.02254	-1.21650
H	3.59883	-1.91488	1.08870
H	-4.45947	-0.53310	-0.05011
H	-3.76782	-1.25802	-1.54001
H	-3.36620	0.40633	-1.09069
O	1.00025	1.61212	0.02506
C	-0.17016	2.35021	0.02763
O	-1.18136	1.90095	-0.42538
C	0.04188	3.67850	0.65964
H	-0.85321	4.28418	0.53657
H	0.90778	4.17469	0.21706
H	0.24977	3.54164	1.72453
H	0.04564	0.29793	-1.23896

Ferrier_Glucosyl_Cation/oxocarbenium/conf_07

Charge=+1, Multiplicity=+1

C	-0.07681	-1.40515	-1.18765
C	-1.22457	-0.47879	-0.88623
C	-0.89384	0.64434	0.07894
C	-2.12598	1.33915	0.52466
C	-3.28831	0.65220	0.64643
C	-3.26978	-0.72059	0.31319
O	-2.32864	-1.29237	-0.31833
O	0.58628	-1.78079	0.00158
C	1.91371	-1.43335	0.09365
O	2.46427	-0.80608	-0.76641
C	2.50098	-1.89330	1.37991
H	-0.45435	-2.30394	-1.68145
H	0.61524	-0.88932	-1.85895
H	-1.63515	-0.08524	-1.82346
H	-2.04713	2.38093	0.82561
H	-4.19076	1.07946	1.06483
H	-4.07885	-1.39736	0.58691
H	3.57652	-2.01376	1.26103
H	2.03811	-2.81727	1.72476
H	2.32385	-1.11161	2.12606
O	0.00525	1.50971	-0.56661
C	1.16313	1.79631	0.14410
O	1.30323	1.41132	1.26780
C	2.12017	2.56545	-0.68639
H	2.78816	3.13177	-0.03988
H	1.60671	3.21575	-1.39398
H	2.71234	1.83555	-1.24907
H	-0.40994	0.24308	0.98966

Ferrier_Glucosyl_Cation/oxocarbenium/conf_08

Charge=+1, Multiplicity=+1

C	-0.43354	0.84656	1.66757
C	-0.87915	-0.47424	1.04892
C	-0.05340	-0.94188	-0.13953
C	-0.75274	-1.99638	-0.91491
C	-2.10262	-2.03246	-0.96669
C	-2.82119	-1.09066	-0.19014
O	-2.30810	-0.35773	0.70169
O	0.33751	1.67858	0.82221
C	-0.29717	2.22227	-0.24898
O	-1.41685	1.88987	-0.54325
C	0.57788	3.18588	-0.96718
H	-1.31999	1.37225	2.03386
H	0.22296	0.63022	2.50957
H	-0.87523	-1.24882	1.82400
H	-0.14386	-2.67515	-1.50672
H	-2.65381	-2.70829	-1.60802
H	-3.89473	-0.94379	-0.30154
H	0.03149	3.62812	-1.79713
H	1.46242	2.65796	-1.33513
H	0.92190	3.96227	-0.28032
O	1.18004	-1.39211	0.37229
C	2.30605	-0.82440	-0.20256
O	2.22136	-0.06387	-1.12023
C	3.53680	-1.29274	0.48636
H	4.40824	-1.01979	-0.10486
H	3.50134	-2.37073	0.65132
H	3.59900	-0.80985	1.46618
H	0.14295	-0.12422	-0.85001

Ferrier_Glucosyl_Cation/oxocarbenium/conf_09

Charge=+1, Multiplicity=+1

C	-0.97435	-0.82013	-0.87882
C	0.26632	0.03520	-0.91796
C	0.93230	0.22135	0.44439
C	0.51152	1.45061	1.16742
C	-0.17650	2.43419	0.55382
C	-0.38702	2.32802	-0.84141
O	-0.11788	1.31185	-1.54610
O	-1.86288	-0.23096	0.05751
C	-3.06218	-0.89588	0.18920
O	-3.29400	-1.87757	-0.44834
C	-3.93933	-0.23171	1.19508
H	-1.43654	-0.88610	-1.86710
H	-0.68245	-1.83015	-0.57182
H	0.97107	-0.39077	-1.63108
H	0.83262	1.54650	2.20042
H	-0.47219	3.34908	1.05115
H	-0.75344	3.16924	-1.42922
H	-4.91116	-0.72050	1.20091
H	-4.05404	0.82939	0.96386

H	-3.48650	-0.31124	2.18702
O	2.34483	0.31218	0.36231
C	2.97701	-0.90743	0.21233
O	2.33748	-1.91315	0.10739
C	4.45471	-0.74980	0.20257
H	4.91844	-1.70692	-0.02557
H	4.78718	-0.40135	1.18422
H	4.75068	0.00434	-0.52963
H	0.66099	-0.65681	1.04495

Ferrier_Glucosyl_Cation/C4_RP/conf_00

Charge=+1, Multiplicity=+1

C	1.43979	1.46989	-0.14089
C	0.10850	0.88288	-0.61627
C	-1.04201	1.37206	0.23891
C	-0.92430	0.83943	1.63192
C	-0.62795	-0.45293	1.70770
C	-0.52698	-1.16432	0.39737
O	0.15636	-0.55009	-0.58643
O	2.51416	0.77625	-0.73806
C	3.00580	-0.27699	-0.02750
O	2.53486	-0.61972	1.02438
C	4.16545	-0.89379	-0.73021
H	1.51415	2.50862	-0.46903
H	1.52665	1.41389	0.94599
H	-0.05940	1.13668	-1.66615
H	-1.05174	1.48769	2.49036
H	-0.47833	-1.00351	2.62790
H	-0.23170	-2.20744	0.44707
H	4.53436	-1.73934	-0.15410
H	4.95719	-0.15137	-0.85586
H	3.86330	-1.21803	-1.72889
O	-2.32357	0.91576	-0.36866
C	-2.70295	-0.31272	-0.41365
O	-2.00158	-1.31848	-0.09499
C	-4.08861	-0.54774	-0.86133
H	-4.17944	-1.53830	-1.30456
H	-4.40223	0.23394	-1.55263
H	-4.73512	-0.50171	0.02357
H	-1.13984	2.45458	0.19004

Ferrier_Glucosyl_Cation/C4_RP/conf_01

Charge=+1, Multiplicity=+1

C	1.51528	0.32869	-0.59903
C	0.11459	-0.23800	-0.44834
C	-0.94626	0.63651	-1.08427
C	-1.19648	1.87262	-0.28216
C	-1.33271	1.66329	1.02333
C	-1.20132	0.23960	1.46287
O	-0.16845	-0.44848	0.94467
O	2.36923	-0.70141	-0.17194

C	3.71551	-0.49028	0.02812
O	4.36870	-1.42061	0.37648
C	4.23507	0.89815	-0.20439
H	1.69319	0.58335	-1.65228
H	1.61920	1.23360	0.00899
H	0.09847	-1.23049	-0.90960
H	-1.26537	2.84414	-0.75680
H	-1.53077	2.43617	1.75645
H	-1.23662	0.07326	2.53535
H	5.31723	0.87602	-0.09646
H	3.82338	1.59503	0.53206
H	3.97983	1.26716	-1.20147
O	-2.20971	-0.14524	-1.18922
C	-2.87982	-0.56333	-0.17518
O	-2.52406	-0.46303	1.03776
C	-4.17914	-1.19633	-0.46573
H	-4.42024	-1.92867	0.30399
H	-4.17270	-1.64252	-1.45946
H	-4.94376	-0.40997	-0.44434
H	-0.71577	0.82746	-2.12975

Ferrier_Galactosyl_Cation/C4_NGP/conf_00

Charge=+1, Multiplicity=+1

C	1.45742	0.92208	-1.00807
C	0.28409	1.21139	-0.07687
C	-0.35380	-0.00182	0.55613
C	-1.67379	0.29732	1.26791
C	-2.34771	1.50452	0.76150
C	-1.83609	2.19861	-0.26309
O	-0.64692	1.96951	-0.83893
O	-0.76746	-0.96604	-0.47559
C	-1.91756	-1.43025	-0.16703
O	-2.48633	-0.91129	0.84692
C	-2.55046	-2.48672	-0.96053
O	2.67124	0.73757	-0.29565
C	2.95271	-0.51612	0.12991
O	2.14433	-1.40827	0.06585
C	4.34438	-0.61523	0.65085
H	1.25270	0.06115	-1.64720
H	1.60637	1.80924	-1.62271
H	0.63724	1.84059	0.75351
H	0.36838	-0.53648	1.17101
H	-1.62667	0.20836	2.35181
H	-3.29900	1.79765	1.18571
H	-2.35029	3.03651	-0.72125
H	-1.79212	-3.14769	-1.37968
H	-3.27105	-3.03154	-0.35163
H	-3.08814	-2.00549	-1.78685
H	4.46495	-1.54394	1.20453
H	4.58740	0.24649	1.27441
H	5.03627	-0.60925	-0.19670

Ferrier_Galactosyl_Cation/C4_NGP/conf_01

Charge=+1, Multiplicity=+1

C	-1.37759	-1.82564	0.21848
C	0.11301	-1.56986	0.31707
C	0.54455	-0.23548	0.87628
C	2.06593	-0.09395	0.92243
C	2.74663	-0.90841	-0.10108
C	2.04705	-1.64037	-0.97979
O	0.72216	-1.83800	-0.93778
O	0.16880	0.87857	-0.01333
C	1.13909	1.71025	-0.06012
O	2.21232	1.36262	0.52833
C	1.02016	2.98021	-0.78053
O	-2.10134	-0.86993	-0.53916
C	-2.61160	0.17724	0.15051
O	-2.32912	0.38697	1.30268
C	-3.53733	0.98275	-0.69483
H	-1.51564	-2.77894	-0.29262
H	-1.78689	-1.88828	1.22969
H	0.50546	-2.30728	1.03735
H	0.03761	-0.03486	1.82091
H	2.49358	-0.13463	1.92263
H	3.82380	-0.85171	-0.18800
H	2.50955	-2.16375	-1.80962
H	-0.00211	3.35323	-0.72285
H	1.73600	3.70158	-0.38806
H	1.26217	2.78814	-1.83326
H	-3.72314	1.94635	-0.22450
H	-4.48543	0.44422	-0.78563
H	-3.13687	1.10359	-1.70261

Ferrier_Galactosyl_Cation/C4_NGP/conf_02

Charge=+1, Multiplicity=+1

C	0.97773	0.37657	-1.62938
C	1.01757	-1.08093	-1.16393
C	-0.22034	-1.68655	-0.46778
C	-0.29501	-1.51173	1.04874
C	0.95742	-1.06157	1.66265
C	2.08307	-1.04519	0.93909
O	2.17922	-1.28977	-0.37793
O	-1.43486	-1.03344	-0.93390
C	-1.97761	-0.39217	0.04544
O	-1.42841	-0.53634	1.18274
C	-3.23480	0.33433	-0.13442
O	1.32179	1.31060	-0.61385
C	0.33873	2.01698	-0.03731
O	-0.83143	1.79904	-0.25723
C	0.87432	3.06133	0.87707
H	1.76002	0.48533	-2.38069
H	0.01228	0.62312	-2.07477

H	1.18135	-1.66665	-2.07142
H	-0.31634	-2.73527	-0.74630
H	-0.69440	-2.40195	1.54101
H	1.00814	-0.92217	2.73427
H	3.05927	-0.88424	1.38245
H	-4.05472	-0.37590	0.02795
H	-3.30124	1.13422	0.60030
H	-3.29620	0.72757	-1.14729
H	0.05408	3.58807	1.35917
H	1.49001	3.76400	0.31011
H	1.52084	2.59678	1.62568

Ferrier_Galactosyl_Cation/C4_NGP/conf_03

Charge=+1, Multiplicity=+1

C	1.39381	0.21071	-1.21732
C	0.39765	0.84578	-0.25772
C	-0.42688	-0.15422	0.51044
C	-1.59579	0.48042	1.26657
C	-2.01364	1.77862	0.71404
C	-1.44810	2.27099	-0.39675
O	-0.39668	1.73398	-1.02906
O	-1.10517	-1.09515	-0.39889
C	-2.31734	-1.24492	-0.01265
O	-2.69154	-0.53277	0.97067
C	-3.22468	-2.17068	-0.69571
O	2.19814	-0.75600	-0.56009
C	3.22058	-0.25594	0.19573
O	3.32473	0.91862	0.41607
C	4.13445	-1.33902	0.65702
H	0.86827	-0.31952	-2.01185
H	2.00900	1.00349	-1.64665
H	0.94127	1.43416	0.49417
H	0.21902	-0.77126	1.13591
H	-1.50279	0.44216	2.35044
H	-2.83265	2.31035	1.18004
H	-1.79717	3.17511	-0.88377
H	-2.66744	-3.00128	-1.12751
H	-3.99610	-2.51442	-0.00704
H	-3.71183	-1.61674	-1.50800
H	4.73505	-0.97978	1.49036
H	4.80009	-1.60541	-0.16969
H	3.57897	-2.23659	0.93190

Ferrier_Galactosyl_Cation/C4_NGP/conf_04

Charge=+1, Multiplicity=+1

C	1.39372	0.21132	-1.21757
C	0.39743	0.84623	-0.25797
C	-0.42644	-0.15385	0.51073
C	-1.59578	0.48045	1.26655
C	-2.01378	1.77868	0.71417
C	-1.44871	2.27088	-0.39694

O	-0.39763	1.73357	-1.02958
O	-1.10417	-1.09562	-0.39835
C	-2.31639	-1.24541	-0.01258
O	-2.69120	-0.53297	0.97032
C	-3.22347	-2.17152	-0.69558
O	2.19806	-0.75555	-0.56063
C	3.22004	-0.25573	0.19594
O	3.32377	0.91870	0.41716
C	4.13410	-1.33886	0.65671
H	0.86815	-0.31872	-2.01220
H	2.00892	1.00417	-1.64676
H	0.94079	1.43528	0.49356
H	0.21971	-0.77036	1.13644
H	-1.50309	0.44212	2.35043
H	-2.83250	2.31051	1.18056
H	-1.79796	3.17490	-0.88402
H	-2.66555	-2.99918	-1.13201
H	-3.99150	-2.51959	-0.00523
H	-3.71540	-1.61616	-1.50401
H	4.73364	-0.98044	1.49117
H	4.80083	-1.60354	-0.16968
H	3.57898	-2.23724	0.92964

Ferrier_Galactosyl_Cation/C4_NGP/conf_05

Charge=+1, Multiplicity=+1

C	-0.10383	-1.50080	-0.44045
C	0.87574	-1.22140	0.70341
C	1.11187	0.24322	1.09412
C	1.90565	1.02576	0.05460
C	2.71633	0.19257	-0.83839
C	2.84465	-1.11167	-0.56838
O	2.13619	-1.78343	0.35852
O	-0.13879	0.99280	1.16477
C	-0.20649	1.79641	0.15688
O	0.80376	1.80448	-0.61849
C	-1.30699	2.74789	0.00115
O	-1.42400	-1.64333	0.06515
C	-2.34691	-0.74055	-0.31626
O	-2.05830	0.25886	-0.93155
C	-3.71494	-1.14353	0.11002
H	-0.07997	-0.75593	-1.23661
H	0.17465	-2.47274	-0.85120
H	0.53776	-1.78022	1.57741
H	1.56324	0.30319	2.08366
H	2.50211	1.82124	0.50974
H	3.35515	0.65828	-1.57657
H	3.57768	-1.74677	-1.05304
H	-0.93817	3.72351	0.34183
H	-1.57206	2.82604	-1.05303
H	-2.16594	2.44529	0.59337
H	-4.40596	-0.31436	-0.02581

H	-3.70959	-1.48212	1.14761
H	-4.04028	-1.98792	-0.50479

Ferrier_Galactosyl_Cation/C4_NGP/conf_06

Charge=+1, Multiplicity=+1

C	-0.10393	-1.50071	-0.44052
C	0.87582	-1.22163	0.70325
C	1.11177	0.24294	1.09425
C	1.90581	1.02579	0.05512
C	2.71653	0.19287	-0.83810
C	2.84471	-1.11149	-0.56861
O	2.13622	-1.78354	0.35805
O	-0.13905	0.99235	1.16460
C	-0.20646	1.79626	0.15696
O	0.80414	1.80484	-0.61794
C	-1.30710	2.74755	0.00096
O	-1.42403	-1.64308	0.06527
C	-2.34698	-0.74042	-0.31638
O	-2.05841	0.25872	-0.93210
C	-3.71496	-1.14318	0.11032
H	-0.08004	-0.75567	-1.23653
H	0.17435	-2.47261	-0.85151
H	0.53799	-1.78066	1.57716
H	1.56279	0.30287	2.08395
H	2.50227	1.82103	0.51068
H	3.35537	0.65883	-1.57609
H	3.57766	-1.74646	-1.05355
H	-2.16542	2.44586	0.59455
H	-0.93793	3.72379	0.33941
H	-1.57334	2.82382	-1.05307
H	-4.40603	-0.31413	-0.02601
H	-3.70947	-1.48090	1.14820
H	-4.04034	-1.98806	-0.50377

Ferrier_Galactosyl_Cation/C4_NGP/conf_07

Charge=+1, Multiplicity=+1

C	0.64984	0.02162	-1.68222
C	0.38057	-1.37664	-1.14695
C	-0.94028	-1.53061	-0.36478
C	-0.86279	-1.22998	1.13233
C	0.48583	-1.24153	1.69159
C	1.52641	-1.58830	0.92208
O	1.48508	-1.84897	-0.38739
O	-1.89076	-0.52943	-0.84734
C	-2.04486	0.37261	0.05985
O	-1.52246	0.14001	1.18830
C	-2.79650	1.59561	-0.22736
O	0.42559	1.04025	-0.70971
C	1.53546	1.54077	-0.06540
O	2.63234	1.11760	-0.26933
C	1.14904	2.63241	0.87388

H	1.68317	0.05688	-2.03210
H	-0.02493	0.23175	-2.51385
H	0.34758	-2.03849	-2.01556
H	-1.39729	-2.50067	-0.55200
H	-1.56398	-1.83820	1.70804
H	0.62512	-1.12356	2.75760
H	2.52174	-1.74385	1.32252
H	-3.59809	1.38709	-0.93652
H	-3.17753	2.03477	0.69304
H	-2.09329	2.29368	-0.69654
H	2.03365	2.98789	1.39762
H	0.41393	2.26333	1.59380
H	0.69526	3.45798	0.31959

Ferrier_Galactosyl_Cation/C4_NGP/conf_08

Charge=+1, Multiplicity=+1

C	1.29684	-0.51586	1.08476
C	0.30373	0.58883	0.77607
C	-1.11788	0.20023	1.10488
C	-2.16562	1.16498	0.53725
C	-1.66488	2.00647	-0.55794
C	-0.44333	1.80822	-1.07396
O	0.46948	0.96466	-0.58274
O	-1.46624	-1.07807	0.46731
C	-2.60966	-0.96896	-0.10296
O	-3.16347	0.17144	-0.04673
C	-3.21518	-2.10239	-0.80592
O	2.60481	-0.04124	0.86559
C	3.24024	-0.50937	-0.25542
O	2.72897	-1.28907	-1.00632
C	4.60449	0.08108	-0.36291
H	1.21541	-0.79180	2.13978
H	1.10408	-1.38682	0.45418
H	0.52625	1.47351	1.39128
H	-1.22599	0.03210	2.17786
H	-2.74870	1.68505	1.29515
H	-2.32301	2.74310	-0.99904
H	-0.08279	2.33387	-1.95174
H	-2.94078	-3.03972	-0.32224
H	-4.29554	-1.97597	-0.86350
H	-2.80933	-2.11086	-1.82558
H	5.09291	-0.29394	-1.25947
H	5.19061	-0.18047	0.52129
H	4.53605	1.17087	-0.39783

Ferrier_Galactosyl_Cation/C4_NGP/conf_09

Charge=+1, Multiplicity=+1

C	-1.29859	-0.53664	1.58982
C	-0.08511	-1.24958	0.93663
C	1.23845	-0.49539	1.10208
C	2.23712	-0.71743	-0.04199

C	1.85007	-1.73853	-1.01983
C	0.55624	-2.03630	-1.18151
O	-0.42773	-1.57234	-0.40222
O	1.06761	0.95434	1.11623
C	1.53803	1.44815	0.02863
O	2.19820	0.65130	-0.70698
C	1.31103	2.84669	-0.32813
O	-2.39610	-0.40771	0.70654
C	-2.28033	0.51133	-0.26795
O	-1.28826	1.18957	-0.39977
C	-3.50012	0.55879	-1.12013
H	-1.67389	-1.13837	2.41521
H	-1.01209	0.44755	1.96248
H	0.08690	-2.20992	1.43869
H	1.66647	-0.74025	2.07342
H	3.26903	-0.79260	0.29872
H	2.59473	-2.15858	-1.68122
H	0.18423	-2.67370	-1.97666
H	2.03216	3.17593	-1.07332
H	0.29371	2.89118	-0.73707
H	1.34069	3.46791	0.56807
H	-4.38968	0.67753	-0.49828
H	-3.60150	-0.39071	-1.65269
H	-3.42115	1.37612	-1.83339

Ferrier_Galactosyl_Cation/oxocarbenium/conf_00

Charge=+1, Multiplicity=+1

C	0.10852	-1.67103	0.58159
C	-1.08710	-1.12890	-0.17644
C	-0.96542	0.30715	-0.64425
C	-2.29803	0.87744	-0.99470
C	-3.43454	0.35398	-0.47729
C	-3.30399	-0.66898	0.49043
O	-2.23811	-1.31974	0.71231
O	-0.44006	1.16403	0.35132
C	0.73195	1.83527	0.01235
O	1.20632	1.72221	-1.07828
C	1.24364	2.61795	1.16323
O	1.19482	-1.66592	-0.32284
C	2.32889	-1.00507	0.07490
O	2.39353	-0.43047	1.12519
C	3.38240	-1.07763	-0.97240
H	0.32877	-1.05426	1.45538
H	-0.09157	-2.69825	0.89433
H	-1.31132	-1.75253	-1.04948
H	-0.29720	0.31790	-1.51547
H	-2.31927	1.78887	-1.58607
H	-4.41325	0.78720	-0.64223
H	-4.12096	-0.93737	1.15928
H	1.80870	1.92890	1.79926
H	1.91424	3.39496	0.80099

H	0.42787	3.04124	1.74921
H	4.35556	-0.90223	-0.51745
H	3.18854	-0.28237	-1.69967
H	3.35993	-2.03407	-1.49423

Ferrier_Galactosyl_Cation/oxocarbenium/conf_01

Charge=+1, Multiplicity=+1

C	0.19709	-1.24280	1.25357
C	-0.86695	-1.03119	0.17948
C	-0.77710	0.27510	-0.57951
C	-2.06819	0.61029	-1.24481
C	-3.23377	0.07008	-0.81977
C	-3.20295	-0.71236	0.35872
O	-2.14677	-1.16336	0.89408
O	-0.47033	1.37538	0.25931
C	0.71381	2.03587	-0.02713
O	1.36700	1.75277	-0.98627
C	1.00540	3.06176	1.00733
O	1.47341	-0.77120	0.86352
C	2.03711	-1.39768	-0.20730
O	1.41284	-2.19642	-0.85527
C	3.42886	-0.93322	-0.44603
H	-0.06739	-0.66381	2.13858
H	0.24154	-2.30416	1.50910
H	-0.87062	-1.84969	-0.54458
H	0.01391	0.16451	-1.33288
H	-2.05373	1.37455	-2.01712
H	-4.19619	0.33909	-1.23688
H	-4.10651	-0.93751	0.92424
H	1.40172	2.55389	1.89224
H	1.75462	3.75467	0.62969
H	0.09795	3.58981	1.30234
H	3.88951	-1.55036	-1.21426
H	4.00782	-0.97180	0.47845
H	3.39839	0.10839	-0.77918

Ferrier_Galactosyl_Cation/oxocarbenium/conf_02

Charge=+1, Multiplicity=+1

C	0.30340	0.48982	-1.38905
C	-0.25338	-0.82741	-0.89028
C	-0.92743	-0.81979	0.48093
C	0.06584	-0.87564	1.58548
C	1.28106	-1.42019	1.39530
C	1.60151	-1.91324	0.09648
O	0.83099	-1.82347	-0.90995
O	-1.72310	0.33368	0.66228
C	-2.93343	0.30011	0.01066
O	-3.21053	-0.60704	-0.72063
C	-3.76322	1.48516	0.35456
O	0.96131	1.31960	-0.42873
C	2.26144	1.10490	-0.16943

O	2.81773	0.06401	-0.44902
C	2.89513	2.27080	0.50087
H	0.97728	0.28540	-2.22634
H	-0.53102	1.09170	-1.74991
H	-0.94983	-1.19983	-1.64057
H	-1.56892	-1.71316	0.55280
H	-0.25412	-0.51289	2.55817
H	2.01584	-1.52931	2.18219
H	2.48688	-2.51815	-0.08578
H	-3.16723	2.39859	0.31638
H	-4.60947	1.54464	-0.32647
H	-4.13058	1.37503	1.37916
H	3.86450	1.98628	0.90404
H	2.24185	2.65922	1.28495
H	3.02816	3.07124	-0.23328

Ferrier_Galactosyl_Cation/oxocarbenium/conf_03

Charge=+1, Multiplicity=+1

C	0.81150	-0.71059	-0.51529
C	0.26652	0.54167	0.13421
C	-1.14588	0.41545	0.66382
C	-1.73630	1.75534	0.94975
C	-1.26358	2.86384	0.33633
C	-0.28202	2.69166	-0.67117
O	0.38497	1.63089	-0.84870
O	-2.04717	-0.19459	-0.24572
C	-2.39446	-1.50987	0.03809
O	-1.92082	-2.08426	0.97026
C	-3.37966	-2.01569	-0.95356
O	2.17964	-0.59164	-0.84480
C	3.03915	-0.63740	0.21301
O	2.63641	-0.65122	1.34577
C	4.45693	-0.67616	-0.23781
H	0.64697	-1.53776	0.18189
H	0.29240	-0.91170	-1.45290
H	0.91714	0.84704	0.95993
H	-1.08485	-0.19293	1.57589
H	-2.61597	1.79871	1.58570
H	-1.70892	3.84289	0.46223
H	-0.06964	3.46778	-1.40579
H	-3.01669	-1.84872	-1.96981
H	-3.55532	-3.07493	-0.77950
H	-4.31653	-1.46230	-0.84576
H	5.11353	-0.46833	0.60443
H	4.62566	0.03306	-1.04963
H	4.67476	-1.67578	-0.62555

Ferrier_Galactosyl_Cation/oxocarbenium/conf_04

Charge=+1, Multiplicity=+1

C	0.47879	-0.91250	-0.42587
C	-0.15116	-0.07455	0.68201

C	-1.26771	0.82598	0.16966
C	-0.69958	1.90803	-0.68577
C	0.57762	2.30524	-0.53683
C	1.32655	1.73435	0.53582
O	0.89350	0.77652	1.26361
O	-2.19237	0.09264	-0.61028
C	-3.14846	-0.58135	0.12246
O	-3.11260	-0.58443	1.31773
C	-4.14613	-1.23138	-0.76824
O	1.71607	-1.48425	-0.01235
C	2.83008	-0.77407	-0.23860
O	2.80601	0.39101	-0.59326
C	4.06186	-1.57484	-0.01932
H	-0.18221	-1.74825	-0.65191
H	0.63616	-0.33011	-1.33601
H	-0.48021	-0.68659	1.51879
H	-1.78990	1.27774	1.02810
H	-1.36533	2.38313	-1.40077
H	1.02762	3.10082	-1.11603
H	2.21943	2.21465	0.92689
H	-3.65217	-1.75041	-1.59138
H	-4.75523	-1.91978	-0.18637
H	-4.78775	-0.45988	-1.20407
H	4.93383	-0.92494	-0.02759
H	4.14885	-2.31662	-0.81873
H	3.99090	-2.12119	0.92353

Ferrier_Galactosyl_Cation/oxocarbenium/conf_05

Charge=+1, Multiplicity=+1

C	0.22501	0.43083	-1.47594
C	0.80281	-0.85951	-0.95898
C	0.37865	-1.26601	0.46476
C	1.29314	-0.78097	1.52600
C	2.58022	-0.46535	1.27342
C	3.02092	-0.50802	-0.06850
O	2.27901	-0.72340	-1.07028
O	-0.90834	-0.78527	0.79141
C	-1.93341	-1.43071	0.13572
O	-1.68610	-2.28261	-0.67032
C	-3.25658	-0.90113	0.54804
O	0.49151	1.48062	-0.56272
C	-0.59257	2.21258	-0.14559
O	-1.70966	1.94799	-0.48333
C	-0.16480	3.33893	0.73037
H	0.68843	0.67097	-2.43690
H	-0.85037	0.31783	-1.62975
H	0.57671	-1.66278	-1.65866
H	0.37141	-2.36732	0.51977
H	0.88615	-0.75106	2.53337
H	3.28728	-0.19244	2.04599
H	4.06435	-0.34639	-0.33823

H	-4.04313	-1.48474	0.07515
H	-3.35427	-0.93427	1.63521
H	-3.32332	0.14691	0.24093
H	-1.03375	3.75973	1.23186
H	0.29919	4.11159	0.11030
H	0.57973	3.00959	1.45705

Ferrier_Galactosyl_Cation/oxocarbenium/conf_06

Charge=+1, Multiplicity=+1

C	0.52330	-0.22235	1.57516
C	-0.20716	0.91906	0.88048
C	-0.82070	0.64459	-0.49115
C	0.22716	0.57664	-1.53405
C	1.28912	1.43025	-1.43778
C	1.42033	2.19244	-0.27158
O	0.70716	2.05428	0.78457
O	-1.61039	-0.51686	-0.48860
C	-2.88419	-0.33612	-0.00103
O	-3.22390	0.71734	0.45760
C	-3.69399	-1.57452	-0.14500
O	1.84798	-0.48224	1.11919
C	2.00769	-1.29232	0.05255
O	1.09110	-1.55137	-0.69130
C	3.40792	-1.77519	-0.07412
H	0.64217	0.02764	2.62884
H	-0.09450	-1.11933	1.47704
H	-1.00256	1.26222	1.54283
H	-1.46384	1.50791	-0.73935
H	0.04754	-0.02743	-2.41735
H	2.01640	1.55371	-2.22957
H	2.15698	2.98812	-0.17731
H	-3.11460	-2.45285	0.14374
H	-4.59781	-1.48853	0.45439
H	-3.96927	-1.69476	-1.19712
H	3.55204	-2.24376	-1.04518
H	3.60493	-2.50598	0.71590
H	4.11007	-0.95107	0.07067

Ferrier_Galactosyl_Cation/oxocarbenium/conf_07

Charge=+1, Multiplicity=+1

C	-0.14699	-1.74965	1.11171
C	1.12508	-0.98680	0.81824
C	1.05050	0.51017	0.61355
C	2.33353	1.03821	0.05328
C	3.15778	0.23677	-0.65991
C	2.72970	-1.08789	-0.91623
O	1.77181	-1.66267	-0.31891
O	0.03235	0.89549	-0.29367
C	-0.54320	2.12534	-0.00937
O	-0.09049	2.82310	0.84834
C	-1.72375	2.38483	-0.87209

O	-1.00599	-1.81401	-0.00465
C	-2.13955	-1.05810	0.04480
O	-2.37148	-0.30550	0.94989
C	-2.99346	-1.31465	-1.14875
H	0.11614	-2.77714	1.37070
H	-0.64653	-1.27595	1.96225
H	1.82087	-1.15192	1.65089
H	0.86000	0.97138	1.59020
H	2.51260	2.10628	0.14359
H	4.04692	0.59890	-1.16118
H	3.18554	-1.70195	-1.69230
H	-2.59504	1.94914	-0.37300
H	-1.87331	3.45927	-0.96414
H	-1.61460	1.91901	-1.85143
H	-3.32778	-2.35524	-1.13905
H	-2.41503	-1.16895	-2.06380
H	-3.85724	-0.65358	-1.13219

Ferrier_Galactosyl_Cation/oxocarbenium/conf_08

Charge=+1, Multiplicity=+1

C	-0.74415	-0.75311	-0.63789
C	-0.09159	0.60559	-0.76360
C	1.42168	0.59872	-0.65110
C	1.95705	1.96043	-0.35755
C	1.17924	2.89894	0.22494
C	-0.12616	2.51971	0.62908
O	-0.71557	1.46735	0.25204
O	1.91303	-0.20960	0.40359
C	2.37850	-1.46440	0.03345
O	2.32981	-1.82658	-1.10303
C	2.88860	-2.20241	1.21798
O	-2.13137	-0.64816	-0.86444
C	-2.93161	-0.68001	0.24981
O	-2.47879	-0.71619	1.35812
C	-4.36975	-0.67900	-0.13750
H	-0.32445	-1.40419	-1.40787
H	-0.55051	-1.17161	0.35234
H	-0.37356	1.08036	-1.71010
H	1.80830	0.21958	-1.60581
H	3.01415	2.13778	-0.53323
H	1.55170	3.87017	0.52607
H	-0.70449	3.10941	1.33972
H	2.10175	-2.28457	1.97181
H	3.22273	-3.19095	0.91119
H	3.71545	-1.64776	1.66851
H	-4.98484	-0.52313	0.74611
H	-4.61699	-1.64503	-0.58678
H	-4.56703	0.08850	-0.88803

Ferrier_Galactosyl_Cation/oxocarbenium/conf_09

Charge=+1, Multiplicity=+1

C	0.22515	0.43087	-1.47592
C	0.80267	-0.85960	-0.95898
C	0.37845	-1.26606	0.46473
C	1.29296	-0.78111	1.52600
C	2.58005	-0.46555	1.27345
C	3.02083	-0.50842	-0.06845
O	2.27896	-0.72378	-1.07025
O	-0.90847	-0.78512	0.79132
C	-1.93363	-1.43053	0.13572
O	-1.68642	-2.28257	-0.67019
C	-3.25671	-0.90070	0.54799
O	0.49177	1.48057	-0.56271
C	-0.59225	2.21268	-0.14561
O	-1.70934	1.94822	-0.48339
C	-0.16433	3.33896	0.73037
H	0.68863	0.67090	-2.43688
H	-0.85023	0.31801	-1.62979
H	0.57648	-1.66279	-1.65871
H	0.37107	-2.36736	0.51974
H	0.88596	-0.75119	2.53336
H	3.28710	-0.19263	2.04602
H	4.06428	-0.34686	-0.33811
H	-3.32325	0.14736	0.24093
H	-4.04336	-1.48415	0.07506
H	-3.35444	-0.93388	1.63515
H	-1.03322	3.75983	1.23191
H	0.29969	4.11158	0.11029
H	0.58021	3.00952	1.45699

Ferrier_Galactosyl_Cation/C4_RP/conf_00

Charge=+1, Multiplicity=+1

C	1.03259	-1.18884	-1.26971
C	-0.23236	-1.47902	-0.43508
C	-1.46705	-0.82648	-1.02300
C	-2.65398	-1.06208	-0.14273
C	-2.43826	-0.82519	1.14599
C	-1.05803	-0.35446	1.48625
O	-0.04293	-1.04974	0.91666
O	-1.23279	0.63306	-1.16995
C	-1.01709	1.43282	-0.18598
O	-0.97231	1.10119	1.04052
C	-0.82454	2.85198	-0.51805
O	2.18930	-1.08327	-0.46695
C	2.43047	0.14160	0.04332
O	1.69375	1.07814	-0.15910
C	3.68270	0.16211	0.84670
H	1.20872	-2.02660	-1.94425
H	0.91007	-0.27463	-1.85292
H	-0.42011	-2.55753	-0.40454
H	-1.61900	-1.13934	-2.05438
H	-3.59821	-1.39573	-0.55582

H	-3.18024	-0.93149	1.92799
H	-0.85286	-0.25391	2.54787
H	-1.24152	3.48060	0.26876
H	0.25982	3.01139	-0.54870
H	-1.25321	3.08234	-1.49160
H	3.86399	1.16641	1.22260
H	4.52249	-0.16927	0.23157
H	3.59328	-0.54199	1.67783

Ferrier_Galactosyl_Cation/C4_RP/conf_01

Charge=+1, Multiplicity=+1

C	0.93653	-1.54018	-0.48658
C	-0.56718	-1.53398	-0.27278
C	-1.37955	-0.62373	-1.18564
C	-2.80599	-0.56814	-0.73133
C	-2.95655	-0.33267	0.56681
C	-1.68475	-0.23257	1.35173
O	-0.79222	-1.22505	1.10565
O	-0.84792	0.75421	-1.20659
C	-0.67885	1.45181	-0.14258
O	-1.03193	1.10366	1.03291
C	-0.00470	2.74436	-0.29659
O	1.39988	-0.22808	-0.23532
C	2.69439	0.07828	0.08780
O	2.93590	1.23573	0.26619
C	3.67227	-1.04596	0.19379
H	1.37127	-2.26256	0.20877
H	1.16784	-1.84694	-1.51294
H	-0.94142	-2.54872	-0.45826
H	-1.26279	-0.93388	-2.22195
H	-3.62184	-0.68241	-1.43471
H	-3.90991	-0.22584	1.06998
H	-1.81419	-0.14663	2.42637
H	-0.13097	3.12747	-1.30803
H	-0.34929	3.45004	0.45810
H	1.06499	2.53925	-0.12429
H	4.66368	-0.62431	0.34368
H	3.67111	-1.66760	-0.70537
H	3.42825	-1.68500	1.04811

References

1. Keith, D. J.; Townsend, S. D., Total Synthesis of the Congested, Bisphosphorylated *Morganella morganii* Zwitterionic Trisaccharide Repeating Unit. *J. Am. Chem. Soc.* **2019**, *141* (32), 12939-12945.
2. Schöllkopf, W.; Gewinner, S.; Junkes, H.; Paarmann, A.; von Helden, G.; Bluem, H. P.; Todd, A. M. M., The new IR and THz FEL facility at the Fritz Haber Institute in Berlin. *Proc. SPIE Int. Soc. Opt. Eng.* **2015**, 95121L.
3. Supady, A.; Blum, V.; Baldauf, C., First-Principles Molecular Structure Search with a Genetic Algorithm. *J. Chem. Inf. Model.* **2015**, *55* (11), 2338-48.

4. Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M., Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **2009**, *180* (11), 2175-2196.
5. Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865-3868.
6. Tkatchenko, A.; Scheffler, M., Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.* **2009**, *102* (7), 073005.
7. Adamo, C.; Barone, V., Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110* (13), 6158-6170.
8. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H., A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104.
9. Grimme, S.; Ehrlich, S.; Goerigk, L., Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32* (7), 1456-65.
10. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16 Rev. A.03*, Wallingford, CT, 2016.

Appendix D

The Fate of a Septanosyl Ferrier Cation in the Gas Phase

This appendix contains the supporting information of the publication related to Chapter 7. It contains mass spectra, energetics, and 3D-structures of septanosyl Ferrier cations. Furthermore, it contains NMR data related to the glycosylation reactions performed in the scope of this publication. The publication, the supporting information, and xyz-coordinates of the computed structures can be found online: <https://doi.org/10.1021/acs.joc.3c00079>.^[312]

Supporting Information

Characterization and fate of a septanosyl Ferrier cation in the gas and solution phases

Kim Greis^{1,2†}, Caleb Griesbach^{3†}, Carla Kirschbaum^{1,2}, Gerard Meijer¹, Gert von Helden¹, Kevin Pagel^{1,2*}, Mark W. Pecuh^{3*}

¹ Fritz Haber Institute of the Max Planck Society, 14195 Berlin, Germany

² Freie Universität Berlin, Institute of Chemistry and Biochemistry, 14195 Berlin, Germany

³ Department of Chemistry, University of Connecticut, 55 N. Eagleville Road, U3060, Storrs, CT 06269 USA

† Equal contribution

Correspondence to: Prof. Dr. Mark W. Pecuh
mark.pecuh@uconn.edu

Prof. Dr. Kevin Pagel
kevin.pagel@fu-berlin.de

Contents

Figure S1. Electrospray ionization mass spectra of glucose- (top) and mannose-like (bottom) acetylated oxepines in positive ion mode. The main signal ($m/z = 367$) corresponds to the sodiated oxepines $[M + Na^+]$. Cleavage of an acetoxy group leads to $[M - OAc]^+$ ions ($m/z = 285$). Furthermore, sodium bound dimers of the precursor $[2M + Na]^+$ are formed ($m/z = 711$).....	S3
Table S1. List of structures of the septanosyl Ferrier cation generated after fragmentation of 8 or 10. Each structure was optimized at the PBE0+D3/6-311+G(d,p) level of theory. Only structures within 20 kJ mol ⁻¹ of the lowest-energy structure of the respective structural motif are represented. Relative energies with and without zero-point vibrational energy (ΔE and $\Delta E+ZPVE$) as well as relative free energies (ΔF) at 90 K (the temperature of the ion trap) are indicated in kJ mol ⁻¹ . The vibrational spectra of the structures labelled with a roman number are shown in the manuscript.	S4
Table S2. DLPNO-CCSD(T)/Def2-TZVPP single-point energies of selected structures of septanosyl Ferrier cation optimized at the PBE0+D3/6-311+G(d,p) level of theory. The structures whose spectra are shown in the manuscript were selected for calculation of high-level single-point energies. Zero-point vibrational energy and free energy corrections are derived from the PBE0 calculations. All energies are indicated in kJ mol ⁻¹	S8
Figure S2. The experimental infrared spectrum (gray) of the septanosyl Ferrier cation $[M - OAc]^+$ compared to computed spectra (red, inverted traces) of structures exhibiting rearrangement by attack of the (VII') C5-, (VIII') C7-, (IX') C4-acetyl group at the C6 position leading to ring opening. The structures are diastereomeric to the ones shown in Figure 5 and generally less stable (except for the C4_rearranged diastereomer) and their harmonic frequencies generally match less well to the experiment. R/S/R is referring to the stereoconfiguration at C4/C5/C6. The relative free energy at 90 K as well as schematic depictions of each structure are indicated.	S9
Figure S3. Energy diagrams for (a) rearrangement of oxocarbenium- or C4_C3_NGP-structures to the C5- and C7-rearranged structures (RSS diastereomer), respectively, and (b) reaction of C5/C7-rearranged structure (RSS diastereomer) to C7/C5-rearranged structure (RSR diastereomer). All energies are computed at the DLPNO-CCSD(T)/Def2-TZVPP level of theory with ZPVE and free-energy correction at the PBE0+D3/6-311+G(d,p) level of theory. All energies are relative to those of C4_C3_NGP/conf_7 (I).	S10

Figure S4. Selected reoptimized 3D-geometries of low-energy structures of (I) C4_C3_NGP, (II) C5_C3_LRP, (III) C7_C1_LRP, (IV) C5_C1_LRP, (V) C4_C1_LRP, and (VI) C7_C3_LRP septanosyl Ferrier cations. Hydrogen atoms are omitted for clarity.	S11
Figure S5. Selected reoptimized 3D-geometries of low-energy structures of (VII) C5_rearranged, (VIII) C7_rearranged, (IX) C4_rearranged, (X) oxocarbenium, and (XI) oxocarbenium “sandwich” septanosyl Ferrier cations. All rearranged structures are shown as R/S/S (C4/C5/C6) diastereomers. Hydrogen atoms are omitted for clarity.	S12
Figure S6. 3D structures of C4_C3_NGP (I) and the oxocarbenium “sandwich” (XI) with nucleophilic path of attack.	S13
Table S3. Tabulated NMR data for compounds 15, 16, and 18	S14
Figure S7. NOESY NMR of the 15/16 mixture showing H1-H6 NOE interaction.....	S15
Figure S8. Zoomed version of the HSQC with assignments.....	S16
Figure S9. COSY view of nonaromatic protons with correlations of the major isomer.	S17
Figure S10. TOCSY experiment showing H4 and H5 cross peak.	S18
Table S4. Tabulated chemical shifts of the 16.	S19
Figure S11. HMBC spectrum showing acetate cross peaks (red circles), benzyl-acetal cross peak (green circle), and C6-H3 cross peak.	S19
Figure S12. Zoomed version of the NOESY spectrum.....	S20
Table S5. Common abbreviations used in the main text.....	S21
¹ H NMR of compound 8	S22
¹ H NMR of compound 10	S23
NMR spectra used to characterize compound 15	S24
NMR characterization data of 16 – wet Ferrier product	S30
NMR characterization data of 17	S36
NMR spectra used to characterize compound 18	S40

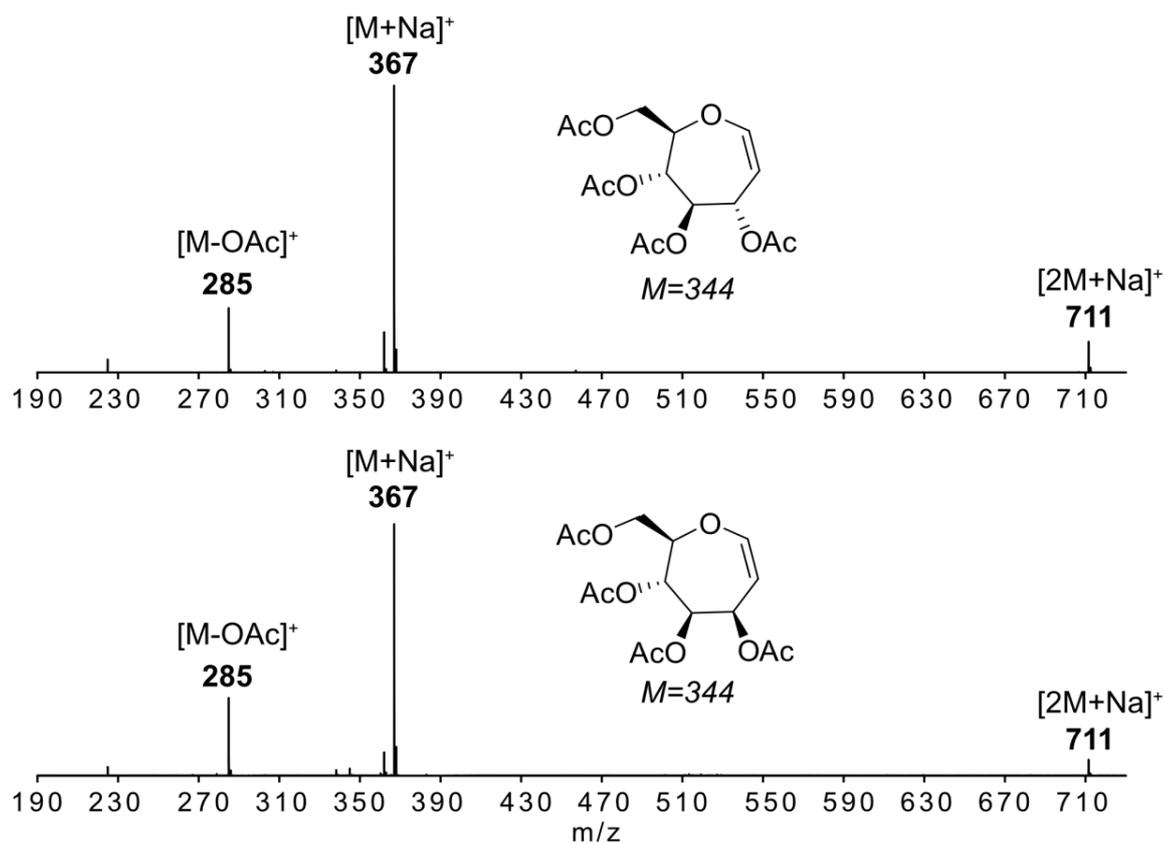


Figure S1. Electrospray ionization mass spectra of glucose- (top) and mannose-like (bottom) acetylated oxepines in positive ion mode. The main signal ($m/z = 367$) corresponds to the sodiated oxepines $[M + Na]^+$. Cleavage of an acetoxy group leads to $[M - OAc]^+$ ions ($m/z = 285$). Furthermore, sodium bound dimers of the precursor $[2M + Na]^+$ are formed ($m/z = 711$).

Table S1. List of structures of the septanosyl Ferrier cation generated after fragmentation of **8** or **10**. Each structure was optimized at the PBE0+D3/6-311+G(d,p) level of theory. Only structures within 20 kJ mol⁻¹ of the lowest-energy structure of the respective structural motif are represented. Relative energies with and without zero-point vibrational energy (ΔE and $\Delta E+ZPVE$) as well as relative free energies (ΔF) at 90 K (the temperature of the ion trap) are indicated in kJ mol⁻¹. The vibrational spectra of the structures labelled with a roman number are shown in the manuscript.

ID	$\Delta E(\text{PBE0})$ [kJ mol ⁻¹]	$\Delta E+ZPVE(\text{PBE0})$ [kJ mol ⁻¹]	$\Delta F(\text{PBE0}, 90\text{K})$ [kJ mol ⁻¹]
C4_C3_NGP/conf_0 ^(a)	8.71	9.91	10.19
C4_C3_NGP/conf_1 ^(a)	8.35	9.91	10.77
C4_C3_NGP/conf_2 ^(a)	5.49	7.76	8.25
C4_C3_NGP/conf_3 ^(a)	2.16	2.17	3.47
C4_C3_NGP/conf_4 ^(a)	9.41	10.95	11.35
C4_C3_NGP/conf_5 ^(a)	5.08	5.10	5.15
C4_C3_NGP/conf_6 ^(a)	13.63	15.82	15.73
C4_C3_NGP/conf_7 (I) ^(a)	0.00	0.00	0.00
C4_C3_NGP/conf_8 ^(a)	4.78	5.65	6.34
C4_C3_NGP/conf_9 ^(a)	4.74	6.76	7.09
C4_C3_NGP/conf_10 ^(a)	4.42	5.15	4.48
C4_C3_NGP/conf_11 ^(a)	14.62	16.70	16.12
C4_C3_NGP/conf_12 ^(a)	3.93	4.05	3.87
C4_C3_NGP/conf_13 ^(a)	15.76	15.79	17.02
C4_C3_NGP/conf_14 ^(a)	10.49	11.23	10.16
C4_C3_NGP/conf_15 ^(a)	6.24	6.14	5.70
C4_C3_NGP/conf_16 ^(a)	19.45	19.20	19.62
C4_C3_NGP/conf_17 ^(a)	10.94	10.14	10.12
C4_C3_NGP/conf_18 ^(a)	12.94	12.65	11.72
C4_C1_LRP/conf_0 ^(a)	42.39	43.29	43.75
C4_C1_LRP/conf_1 ^(a)	44.74	45.04	44.77
C4_C1_LRP/conf_2 ^(a)	42.25	42.75	42.82
C4_C1_LRP/conf_3 ^(a)	34.45	34.71	35.60
C4_C1_LRP/conf_4 ^(a)	43.03	42.92	42.90
C4_C1_LRP/conf_5 ^(a)	43.17	42.88	42.11
C4_C1_LRP/conf_6 ^(a)	47.31	47.55	48.05
C4_C1_LRP/conf_7 (V) ^(a)	36.44	35.41	34.70
C4_C1_LRP/conf_8 ^(a)	42.42	42.48	43.18
C4_C1_LRP/conf_9 ^(a)	39.40	38.89	39.28
C4_C1_LRP/conf_10 ^(a)	51.24	51.32	52.15
C4_C1_LRP/conf_11 ^(a)	51.02	50.30	49.70
C4_C1_LRP/conf_12 ^(a)	46.64	46.18	46.14
C4_C1_LRP/conf_13 ^(a)	49.22	47.88	46.71
C4_C1_LRP/conf_14 ^(a)	45.09	44.28	44.04
C4_C1_LRP/conf_15 ^(a)	45.28	44.36	43.95
C5_C3_LRP/conf_0 ^(a)	23.92	24.92	25.75
C5_C3_LRP/conf_1 ^(a)	21.76	23.24	23.95
C5_C3_LRP/conf_2 ^(a)	27.07	28.05	28.88
C5_C3_LRP/conf_3 (II) ^(a)	19.04	20.65	21.72
C5_C3_LRP/conf_4 ^(a)	24.97	26.47	27.25
C5_C3_LRP/conf_5 ^(a)	21.53	22.05	22.05
C5_C3_LRP/conf_6 ^(a)	22.45	22.47	23.87

Table S1 continued. List of structures of the septanosyl Ferrier cation generated after fragmentation of **8** or **10**. Each structure was optimized at the PBE0+D3/6-311+G(d,p) level of theory. Only structures within 20 kJ mol⁻¹ of the lowest-energy structure of the respective structural motif are represented. Relative energies with and without zero-point vibrational energy (ΔE and ΔE +ZPVE) as well as relative free energies (ΔF) at 90 K (the temperature of the ion trap) are indicated in kJ mol⁻¹. The vibrational spectra of the structures labelled with a roman number are shown in the manuscript.

ID	ΔE (PBE0) [kJ mol ⁻¹]	ΔE +ZPVE(PBE0) [kJ mol ⁻¹]	ΔF (PBE0, 90K) [kJ mol ⁻¹]
C5_C3_LRP/conf_7 ^(a)	22.66	23.48	23.95
C5_C1_LRP/conf_0 ^(a)	31.45	31.64	32.25
C5_C1_LRP/conf_1 (IV) ^(a)	31.57	31.67	31.97
C5_C1_LRP/conf_2 ^(a)	37.01	36.36	35.70
C5_C1_LRP/conf_3 ^(a)	37.21	37.62	37.55
C5_C1_LRP/conf_4 ^(a)	42.52	41.69	41.81
C5_C1_LRP/conf_5 ^(a)	42.43	41.66	41.86
C5_C1_LRP/conf_6 ^(a)	40.61	40.14	39.65
C5_C1_LRP/conf_7 ^(a)	40.55	40.16	39.67
C7_C3_LRP/conf_0 (VI) ^(a)	52.95	55.68	58.44
C7_C1_LRP/conf_0 ^(a)	42.19	40.81	40.35
C7_C1_LRP/conf_1 ^(a)	46.80	44.56	44.32
C7_C1_LRP/conf_2 ^(a)	27.62	26.83	26.38
C7_C1_LRP/conf_3 ^(a)	42.29	39.72	38.18
C7_C1_LRP/conf_4 ^(a)	30.42	29.28	28.98
C7_C1_LRP/conf_5 ^(a)	30.03	29.08	28.40
C7_C1_LRP/conf_6 ^(a)	28.83	26.86	26.04
C7_C1_LRP/conf_7 (III) ^(a)	29.27	27.03	25.60
C7_C1_LRP/conf_8 ^(a)	40.40	37.56	37.42
oxocarbenium/conf_0 ^(c)	26.77	22.27	20.72
oxocarbenium/conf_1 (X) ^(c)	26.41	19.91	16.28
oxocarbenium/conf_2 ^(c)	23.59	18.83	16.97
oxocarbenium/conf_3 (XI) ^(c)	25.07	20.95	20.19
oxocarbenium/conf_4 ^(c)	30.13	24.66	22.74
oxocarbenium/conf_5 ^(b)	27.28	21.07	17.65
oxocarbenium/conf_6 ^(b)	41.77	35.16	32.46
C4_rearranged_RSS/conf_0 (IX) ^(b)	15.20	6.15	3.76
C4_rearranged_RSS/conf_1 ^(b)	19.32	9.58	5.25
C4_rearranged_RSS/conf_2 ^(b)	18.27	9.23	6.25
C4_rearranged_RSS/conf_3 ^(b)	22.05	11.70	6.63
C4_rearranged_RSS/conf_4 ^(b)	18.45	9.33	6.17
C4_rearranged_RSS/conf_5 ^(b)	26.06	16.63	13.47
C4_rearranged_RSS/conf_6 ^(b)	27.84	18.71	15.61
C5_rearranged_RSS/conf_0 ^(c)	-0.97	-8.99	-11.76
C5_rearranged_RSS/conf_1 ^(c)	1.89	-5.48	-8.10
C5_rearranged_RSS/conf_2 ^(c)	-2.77	-11.11	-14.16
C5_rearranged_RSS/conf_3 ^(c)	7.34	-0.49	-3.35
C5_rearranged_RSS/conf_4 ^(c)	7.18	0.66	-0.84
C5_rearranged_RSS/conf_5 ^(c)	10.01	1.60	-1.94
C5_rearranged_RSS/conf_6 ^(c)	9.85	2.32	-1.05
C5_rearranged_RSS/conf_7 ^(c)	11.13	3.82	1.48
C5_rearranged_RSS/conf_8 ^(c)	8.82	2.38	0.48
C5_rearranged_RSS/conf_9 ^(c)	12.71	3.68	-0.75

Table S1 continued. List of structures of the septanosyl Ferrier cation generated after fragmentation of **8** or **10**. Each structure was optimized at the PBE0+D3/6-311+G(d,p) level of theory. Only structures within 20 kJ mol⁻¹ of the lowest-energy structure of the respective structural motif are represented. Relative energies with and without zero-point vibrational energy (ΔE and ΔE +ZPVE) as well as relative free energies (ΔF) at 90 K (the temperature of the ion trap) are indicated in kJ mol⁻¹. The vibrational spectra of the structures labelled with a roman number are shown in the manuscript.

ID	ΔE (PBE0) [kJ mol ⁻¹]	ΔE +ZPVE(PBE0) [kJ mol ⁻¹]	ΔF (PBE0, 90K) [kJ mol ⁻¹]
C5_rearranged_RSS/conf_10 ^(c)	16.18	7.36	4.50
C5_rearranged_RSS/conf_11 ^(c)	12.04	4.20	1.03
C5_rearranged_RSS/conf_12 ^(c)	-4.07	-11.10	-13.22
C5_rearranged_RSS/conf_13 ^(c)	9.15	0.28	-3.94
C5_rearranged_RSS/conf_14 ^(c)	-0.24	-6.67	-8.38
C5_rearranged_RSS/conf_15 ^(c)	9.94	1.19	-2.24
C5_rearranged_RSS/conf_16 ^(c)	0.75	-7.24	-10.12
C5_rearranged_RSS/conf_17 ^(c)	14.25	6.78	3.52
C5_rearranged_RSS/conf_18 ^(c)	10.51	2.02	-2.21
C5_rearranged_RSS/conf_19 ^(c)	4.62	-1.72	-3.08
C5_rearranged_RSS/conf_20 ^(c)	14.33	5.81	1.50
C5_rearranged_RSS/conf_21 ^(c)	-4.65	-11.33	-13.28
C5_rearranged_RSS/conf_22 ^(c)	6.97	-0.96	-4.19
C5_rearranged_RSS/conf_23 ^(c)	11.76	1.83	-3.13
C5_rearranged_RSS/conf_24 ^(c)	17.65	7.61	1.24
C5_rearranged_RSS/conf_25 ^(c)	13.46	5.42	1.24
C5_rearranged_RSS/conf_26 ^(c)	9.28	0.61	-3.40
C5_rearranged_RSS/conf_27 ^(a)	-7.00	-13.57	-15.26
C5_rearranged_RSS/conf_28 ^(a)	2.18	-5.26	-8.03
C5_rearranged_RSS/conf_29 ^(a)	-0.97	-8.99	-11.75
C5_rearranged_RSS/conf_30 (VII) ^(a)	-8.06	-14.07	-15.41
C5_rearranged_RSS/conf_31 ^(b)	9.85	1.11	-2.22
C5_rearranged_RSS/conf_32 ^(b)	7.18	0.65	-0.86
C5_rearranged_RSS/conf_33 ^(b)	8.82	2.38	0.50
C7_rearranged_RSS/conf_0 ^(c)	8.47	1.38	-1.19
C7_rearranged_RSS/conf_1 (VIII) ^(c)	-5.22	-12.42	-15.11
C7_rearranged_RSS/conf_2 ^(c)	11.73	4.27	1.36
C7_rearranged_RSS/conf_3 ^(c)	10.60	4.93	3.26
C7_rearranged_RSS/conf_4 ^(c)	3.14	-2.81	-3.92
C7_rearranged_RSS/conf_5 ^(c)	4.76	-0.95	-2.69
C7_rearranged_RSS/conf_6 ^(c)	10.95	3.55	1.01
C4_rearranged/conf_0 ^(c)	6.13	-0.55	-2.40
C4_rearranged_RSR/conf_1 (IX') ^(c)	4.69	-2.94	-5.06
C4_rearranged_RSR/conf_2 ^(c)	21.28	12.66	8.74
C4_rearranged_RSR/conf_3 ^(c)	16.83	6.73	1.97
C4_rearranged_RSR/conf_4 ^(c)	5.64	-1.50	-4.49
C4_rearranged_RSR/conf_5 ^(c)	9.08	1.93	-0.83
C4_rearranged_RSR/conf_6 ^(c)	21.20	13.05	9.59
C4_rearranged_RSR/conf_7 ^(c)	20.85	13.74	11.95
C4_rearranged_RSR/conf_8 ^(c)	15.74	8.78	7.16
C4_rearranged_RSR/conf_9 ^(c)	24.27	14.33	9.85
C4_rearranged_RSR/conf_10 ^(c)	11.95	3.88	1.02
C4_rearranged_RSR/conf_11 ^(c)	21.27	13.07	9.72

Table S1 continued. List of structures of the septanosyl Ferrier cation generated after fragmentation of **8** or **10**. Each structure was optimized at the PBE0+D3/6-311+G(d,p) level of theory. Only structures within 20 kJ mol⁻¹ of the lowest-energy structure of the respective structural motif are represented. Relative energies with and without zero-point vibrational energy (ΔE and $\Delta E+ZPVE$) as well as relative free energies (ΔF) at 90 K (the temperature of the ion trap) are indicated in kJ mol⁻¹. The vibrational spectra of the structures labelled with a roman number are shown in the manuscript.

ID	$\Delta E(\text{PBE0})$ [kJ mol ⁻¹]	$\Delta E+ZPVE(\text{PBE0})$ [kJ mol ⁻¹]	$\Delta F(\text{PBE0}, 90\text{K})$ [kJ mol ⁻¹]
C4_rearranged_RSR/conf_12 ^(c)	23.45	14.16	9.96
C4_rearranged_RSR/conf_13 ^(c)	24.63	15.72	10.98
C4_rearranged_RSR/conf_14 ^(c)	18.51	11.36	8.95
C4_rearranged_RSR/conf_15 ^(b)	20.98	12.15	8.81
C4_rearranged_RSR/conf_16 ^(b)	14.77	5.32	1.24
C4_rearranged_RSR/conf_17 ^(b)	14.96	5.38	1.07
C5_rearranged_RSR/conf_0 ^(b)	26.20	17.65	14.21
C5_rearranged_RSR/conf_1 ^(b)	26.10	17.64	14.35
C5_rearranged_RSR/conf_2 ^(b)	20.94	12.55	9.10
C5_rearranged_RSR/conf_3 ^(b)	25.56	18.50	16.32
C5_rearranged_RSR/conf_4 ^(b)	25.38	15.85	11.15
C5_rearranged_RSR/conf_5 (VII') ^(b)	17.01	8.98	6.22
C5_rearranged_RSR/conf_6 ^(b)	22.23	14.98	13.05
C5_rearranged_RSR/conf_7 ^(b)	21.71	14.64	12.49
C5_rearranged_RSR/conf_8 ^(b)	26.21	19.31	17.49
C5_rearranged_RSR/conf_9 ^(b)	23.10	16.05	14.03
C5_rearranged_RSR/conf_10 ^(b)	23.41	15.20	11.27
C7_rearranged_RSR/conf_0 ^(b)	7.46	-0.13	-4.01
C7_rearranged_RSR/conf_1 ^(b)	12.45	4.37	0.54
C7_rearranged_RSR/conf_2 ^(b)	2.99	-4.21	-7.60
C7_rearranged_RSR/conf_3 ^(b)	5.21	-2.26	-5.16
C7_rearranged_RSR/conf_4 (VIII') ^(b)	0.52	-6.93	-10.38
C7_rearranged_RSR/conf_5 ^(b)	20.36	12.77	10.20
C7_rearranged_RSR/conf_6 ^(b)	11.82	4.71	2.39
C5_rearrangement/IRCb ^(d)	27.81	26.83	25.98
C5_rearrangement/TS ^(d)	106.87	93.47	92.11
C5_rearrangement/IRCF ^(d)	20.31	11.67	7.46
C7_rearrangement/IRCb ^(d)	23.39	22.58	21.86
C7_rearrangement/TS ^(d)	109.92	98.07	96.54
C7_rearrangement/IRCF ^(d)	20.56	12.77	10.52
C5_RSS_C7_RSR_swap/IRCb ^(d)	12.71	3.66	-0.80
C5_RSS_C7_RSR_swap/TS ^(d)	69.49	56.90	54.56
C5_RSS_C7_RSR_swap/IRCF ^(d)	11.10	3.51	1.36
C7_RSS_C5_RSR_swap/IRCb ^(d)	3.14	-2.82	-3.93
C7_RSS_C5_RSR_swap/TS ^(d)	70.55	58.28	55.37
C7_RSS_C5_RSR_swap/IRCF ^(d)	20.94	12.55	9.10

The structures were obtained after sampling with (a) CREST-GFN2, (b) CREST-GFNFF, or (c) Maestro. (d) The labelled structures were used to construct the energy diagrams in Figure S3. IRC: intrinsic reaction coordinate, TS: transition state.

Table S2. DLPNO-CCSD(T)/Def2-TZVPP single-point energies of selected structures of septanosyl Ferrier cation optimized at the PBE0+D3/6-311+G(d,p) level of theory. The structures whose spectra are shown in the manuscript were selected for calculation of high-level single-point energies. Zero-point vibrational energy and free energy corrections are derived from the PBE0 calculations. All energies are indicated in kJ mol⁻¹.

ID	$\Delta E(\text{CCSD(T)})$ [kJ mol ⁻¹]	$\Delta E+\text{ZPVE}(\text{CCSD(T)})$ [kJ mol ⁻¹]	$\Delta F(\text{CCSD(T)}, 90\text{K})$ [kJ mol ⁻¹]
C4_C3_NGP/conf_7 (I)	0.00	0.00	0.00
C5_C3_LRP/conf_3 (II)	12.69	14.30	15.36
C7_C1_LRP/conf_7 (III)	24.61	22.37	20.93
C5_C1_LRP/conf_1 (IV)	27.42	27.52	27.83
C4_C1_LRP/conf_7 (V)	30.99	29.96	29.25
C7_C3_LRP/conf_0 (VI)	56.03	58.76	61.52
oxocarbenium/conf_1 (X)	25.38	18.88	15.24
oxocarbenium/conf_3 (XI)	31.00	26.88	26.12
C4_rearranged_RSS/conf_2 (IX)	2.07	-6.99	-9.38
C5_rearranged_RSS/conf_30 (VII)	-25.34	-31.36	-32.69
C7_rearranged_RSS/conf_1 (VIII)	-22.33	-29.52	-32.22
C4_rearranged_RSR/conf_1 (IX')	-9.01	-16.64	-18.77
C5_rearranged_RSR/conf_5 (VII')	1.10	-6.93	-9.69
C7_rearranged_RSR/conf_4 (VIII')	-15.05	-22.50	-25.95
C5_rearrangement/IRCb ^(a)	27.34	26.36	25.51
C5_rearrangement/TS ^(a)	108.38	94.98	93.62
C5_rearrangement/IRCF ^(a)	5.02	-3.62	-7.82
C7_rearrangement/IRCb ^(a)	11.86	11.05	10.33
C7_rearrangement/TS ^(a)	107.19	95.34	93.81
C7_rearrangement/IRCF ^(a)	3.25	-4.54	-6.78
C5_RSS_C7_RSR_swap/IRCb ^(a)	-3.94	-12.99	-17.45
C5_RSS_C7_RSR_swap/TS ^(a)	58.43	45.84	43.50
C5_RSS_C7_RSR_swap/IRCF ^(a)	-4.97	-12.55	-14.71
C7_RSS_C5_RSR_swap/IRCb ^(a)	-10.18	-16.14	-17.25
C7_RSS_C5_RSR_swap/TS ^(a)	60.37	48.11	45.19
C7_RSS_C5_RSR_swap/IRCF ^(a)	3.82	-4.57	-8.02

(a) The labelled structures were used to construct the energy diagrams in Figure S3. IRC: intrinsic reaction coordinate, TS: transition state.

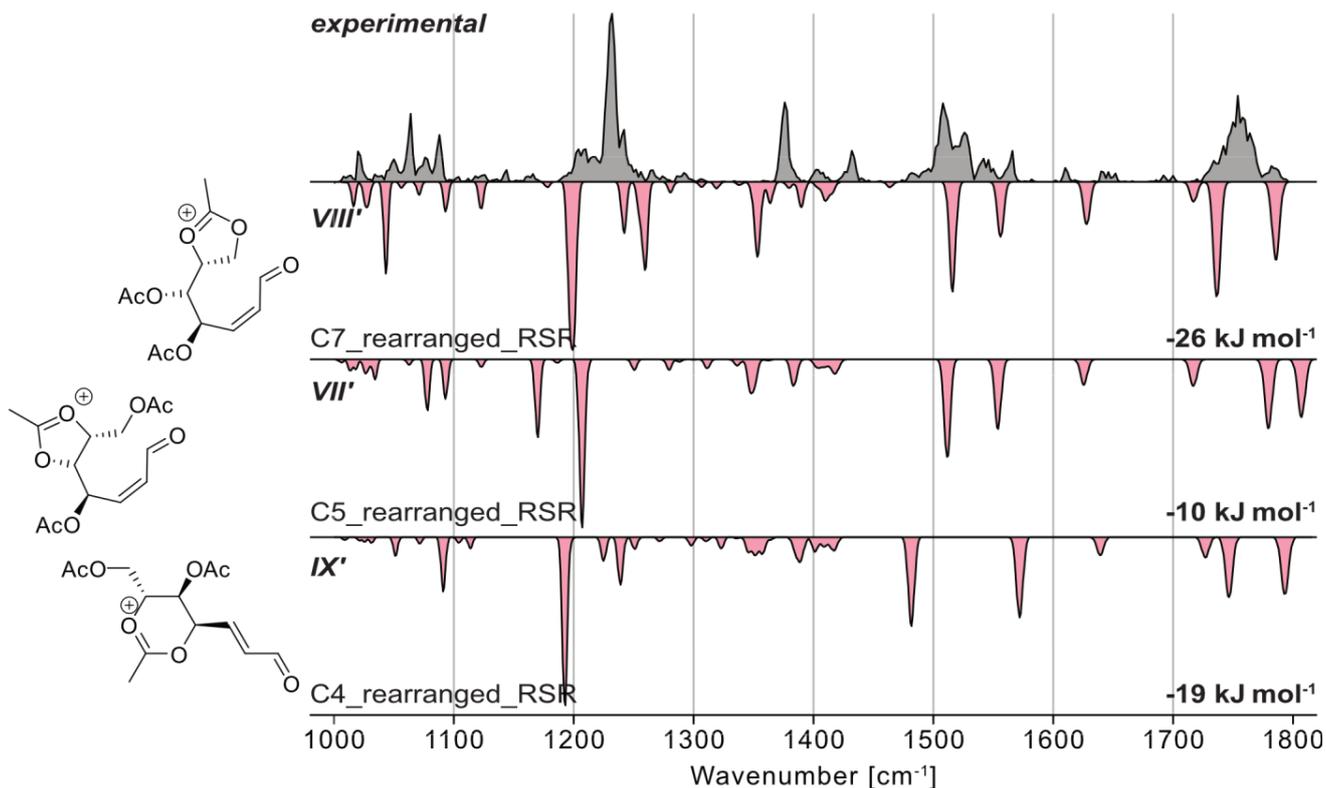


Figure S2. The experimental infrared spectrum (gray) of the septanosyl Ferrier cation $[M - OAc]^+$ compared to computed spectra (red, inverted traces) of structures exhibiting rearrangement by attack of the (VII') C5-, (VIII') C7-, (IX') C4-acetyl group at the C6 position leading to ring opening. The structures are diastereomeric to the ones shown in Figure 5 and generally less stable (except for the C4_rearranged diastereomer) and their harmonic frequencies generally match less well to the experiment. R/S/R is referring to the stereoconfiguration at C4/C5/C6. The relative free energy at 90 K as well as schematic depictions of each structure are indicated.

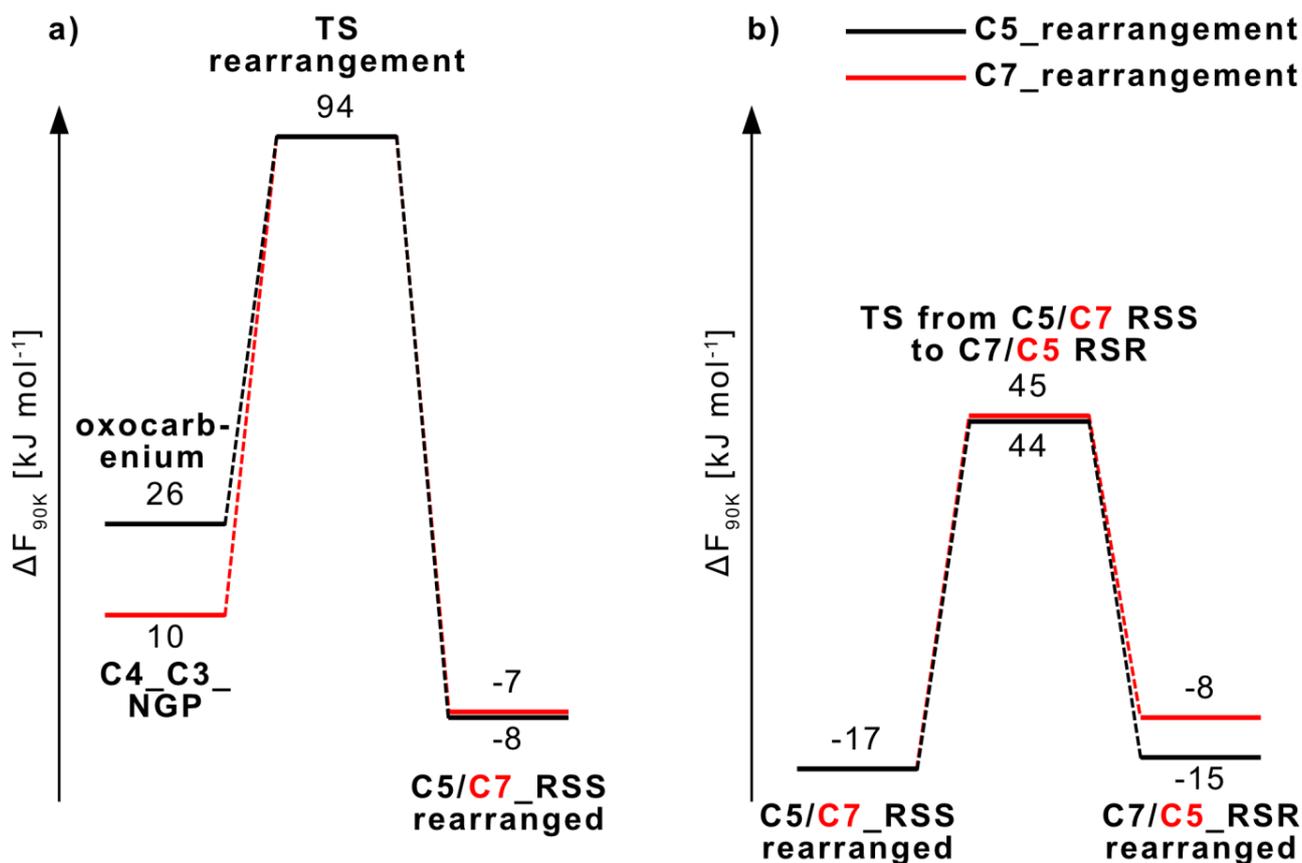


Figure S3. Energy diagrams for (a) rearrangement of oxocarbenium- or C4_C3_NGP-structures to the C5- and C7-rearranged structures (RSS diastereomer), respectively, and (b) reaction of C5/C7-rearranged structure (RSS diastereomer) to C7/C5-rearranged structure (RSR diastereomer). All energies are computed at the DLPNO-CCSD(T)/Def2-TZVPP level of theory with ZPVE and free-energy correction at the PBE0+D3/6-311+G(d,p) level of theory. All energies are relative to those of C4_C3_NGP/conf_7 (I).

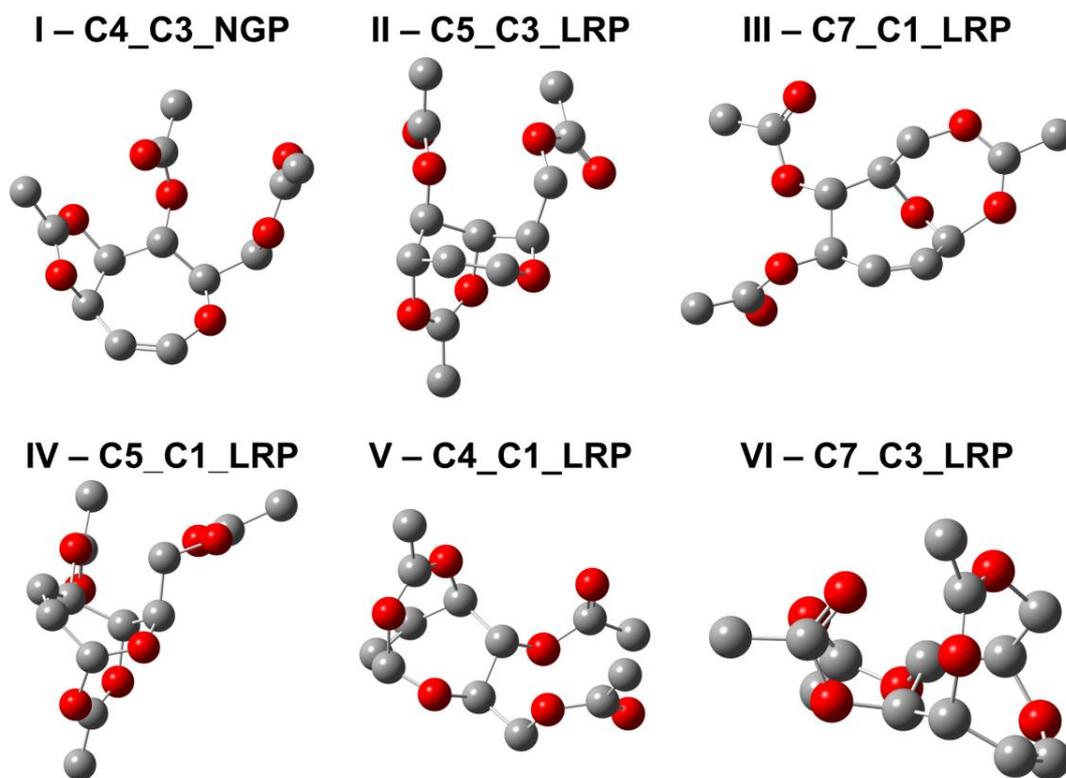


Figure S4. Selected reoptimized 3D-geometries of low-energy structures of (I) C4_C3_NGP, (II) C5_C3_LRP, (III) C7_C1_LRP, (IV) C5_C1_LRP, (V) C4_C1_LRP, and (VI) C7_C3_LRP septanosyl Ferrier cations. Hydrogen atoms are omitted for clarity.

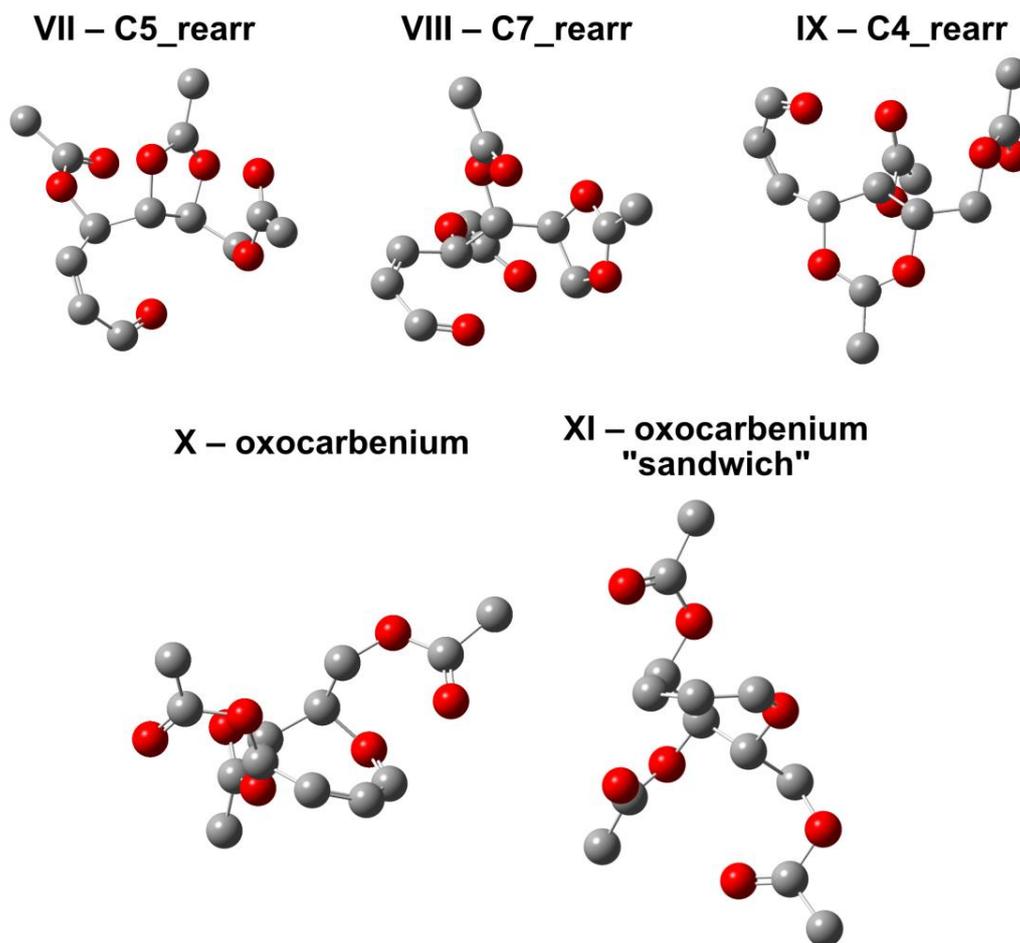


Figure S5. Selected reoptimized 3D-geometries of low-energy structures of (VII) C5_rearranged, (VIII) C7_rearranged, (IX) C4_rearranged, (X) oxocarbenium, and (XI) oxocarbenium "sandwich" septanosyl Ferrier cations. All rearranged structures are shown as R/S/S (C4/C5/C6) diastereomers. Hydrogen atoms are omitted for clarity.

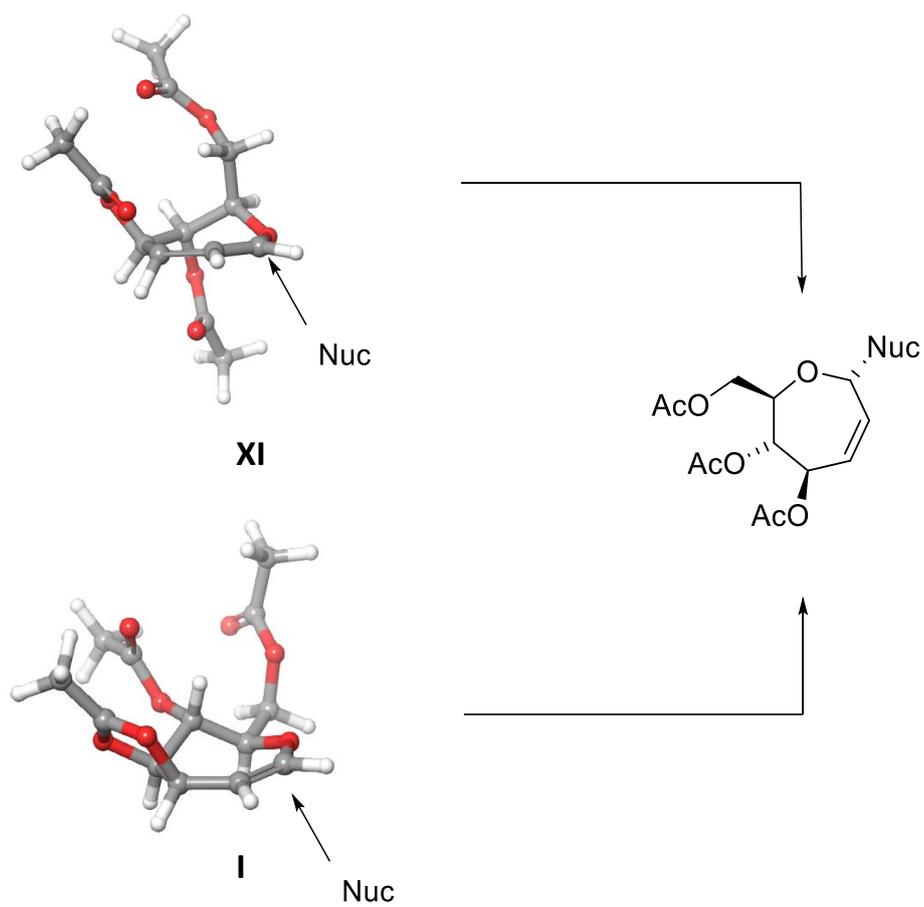
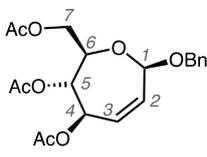
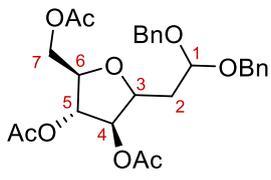
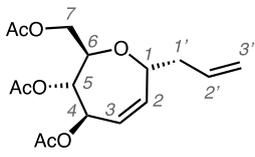


Figure S6. 3D structures of C4_C3_NGP (**I**) and the oxocarbenium “sandwich” (**XI**) with nucleophilic path of attack.

Table S3. Tabulated NMR data for compounds 15, 16, and 18

Compound	Key chemical shifts (δ ppm)				Key Observations
	^1H		$^{13}\text{C}\{^1\text{H}\}$		
 <p>15 CDCl₃</p>	H1 5.40 H2 5.76 H3 5.91 H4 5.27 H5 5.20 H6 4.37 H7/7' 4.14-4.26		C1 98.3 C2 132.3 C3 126.6 C4 70.8 C5 73.1 C6 68.9 C7 63.7		H1-H6 NOE
 <p>16 CDCl₃</p>	major H1 4.93 H2 2.04 H3 4.17 H4 5.15 H5 4.91 H6 3.92 H7/7' 4.33/4.16	minor H1 4.97 H2 2.11 H3 4.24 H4 5.10 H5 5.08 H6 4.16 H7/7' 4.24	major C1 99.92 C2 32.8 C3 77.14 C4 77.41 C5 78.97 C6 81.23 C7 63.84	minor C1 99.74 C2 36.70 C3 80.00 C4 80.53 C5 78.76 C6 80.61 C7 63.52	H3-C6 HMBC Major: H3-H6 NOE
 <p>18 acetone- d₆</p>	H1 4.51 H2 5.74 H3 5.53 H4 5.79 H4 5.79 H5 5.02 H6 4.09 H7 4.08 H7 4.42 H1' 2.33 H2' 5.9 H3' 5.04-5.17				H1-H5 NOE

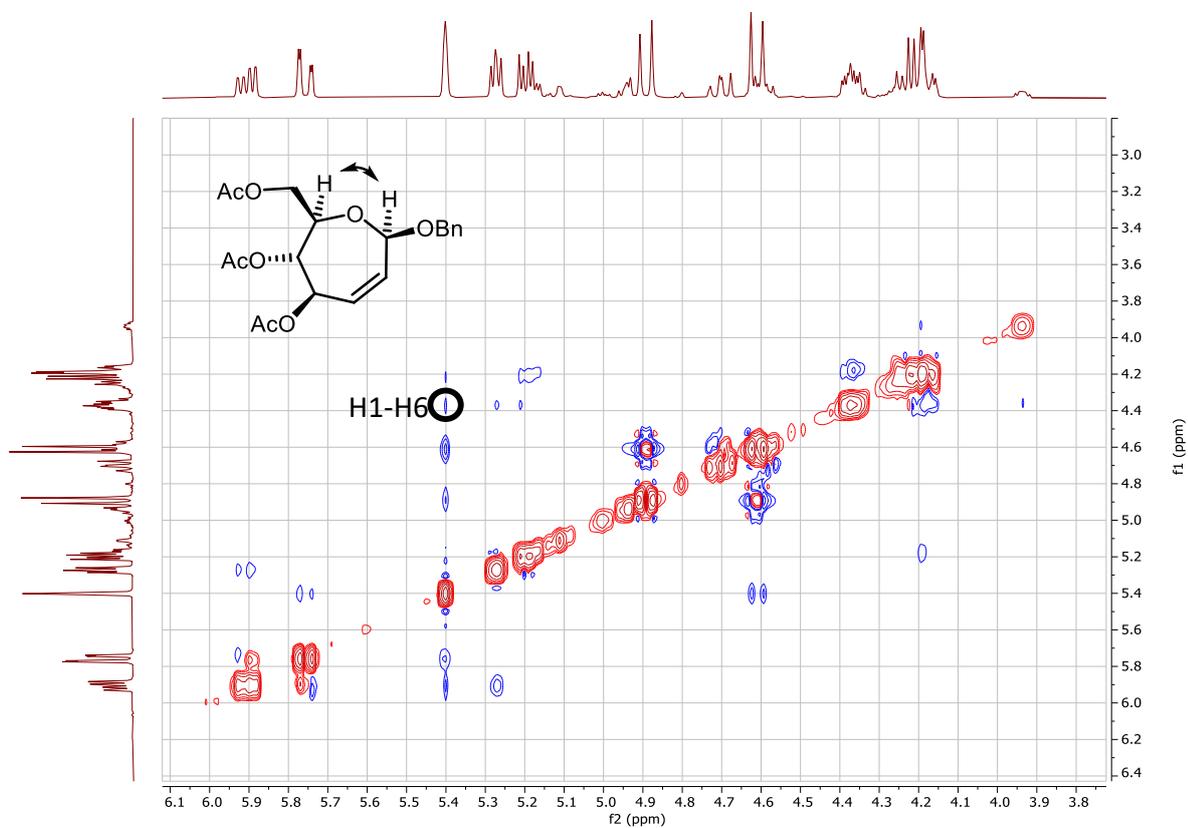


Figure S7. NOESY NMR of the 15/16 mixture showing H1-H6 NOE interaction.

Analysis of NMR spectra to support structure of C-methylene-acetal arabinofuranoside 16

HSQC data

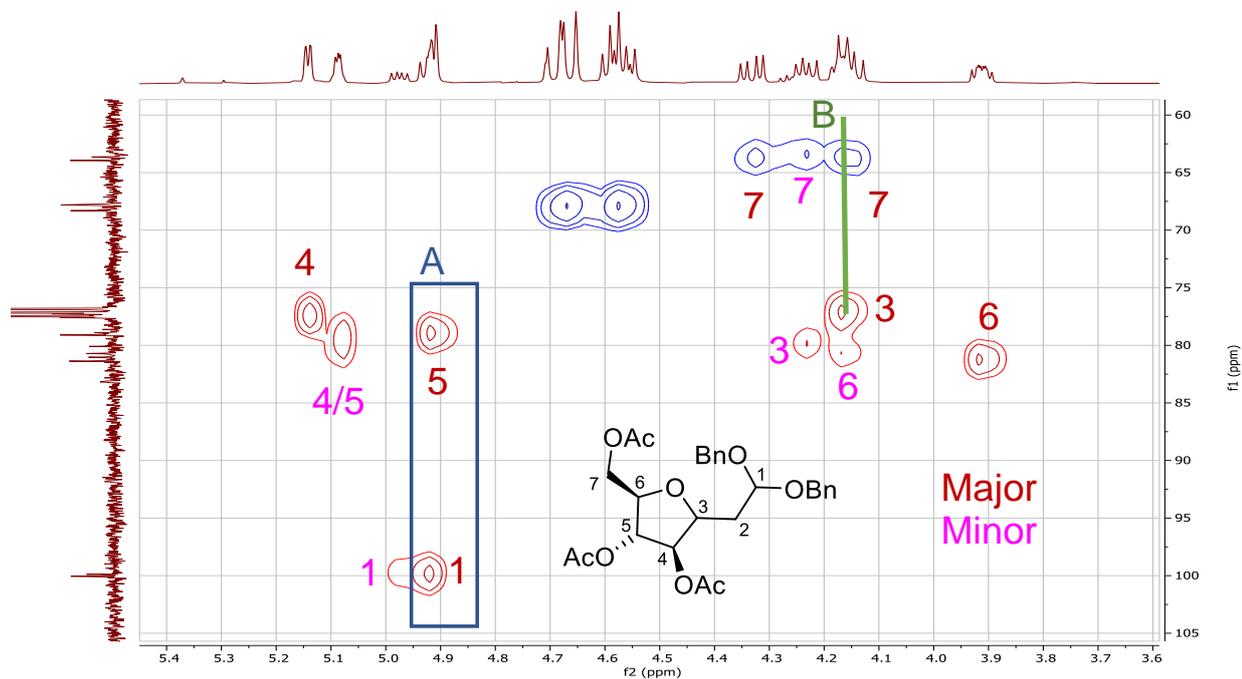


Figure S8. Zoomed version of the HSQC with assignments.

- A) Acetal proton (H1; 4.9 ppm) overlaps with another CH group, but it is slightly more downfield. This multiplet integrated to 2H suggesting it is from the same molecule.
- B) H3 of the major product is slightly more downfield than one of the H7 diastereotopic protons.

COSY data

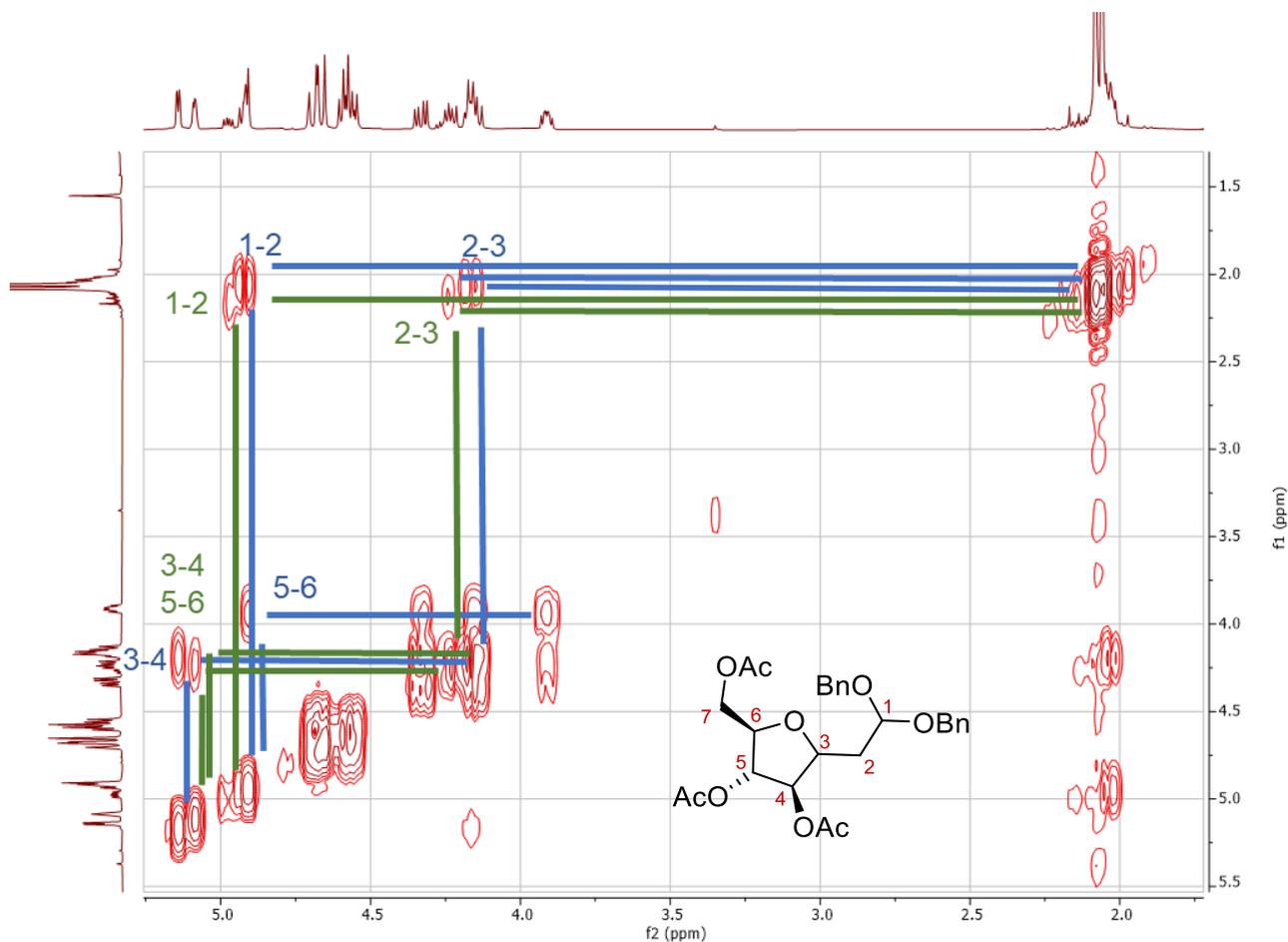


Figure S9. COSY view of nonaromatic protons with correlations of the major isomer.

Blue line: major isomer correlations

Starting at the major acetal proton (H1) at 4.95, there is a correlation to the CH₂ (H2) hidden under the acetate methyl groups. The hidden methylene also correlates to a CH group hidden in a multiplet at 4.1 ppm (H3). This, in turn correlates to another CH (H4). COSY correlation ends here after 4 groups. However, H4 and H5 are shown to correlate in TOCSY experiments (Figure S6). Based on that experiment, H5 is the proton that overlaps with the acetal proton and correlates to H6. H6 is connected to the acetoxy methyl rotamer since it correlates to H7 and H7'.

Green line: minor isomer correlations

In the minor product, the acetal proton is more downfield than the major. The dd at ~5 ppm is H1. This correlates to the hidden CH₂ group (H2). H2 correlates to the multiplet at 4.26 ppm (H3) which, in turn, correlates to the H4 at 5.1 ppm. Again, H4 and H5 do not correlate by COSY. In this case H4 and H5 overlap at 5.1 ppm. This peak correlates to the multiplet at 4.17 ppm (H6). In this case, the H6 proton is overlapped by the major H7 and H3 protons but can be observed to be unique in the HSQC.

TOCSY data

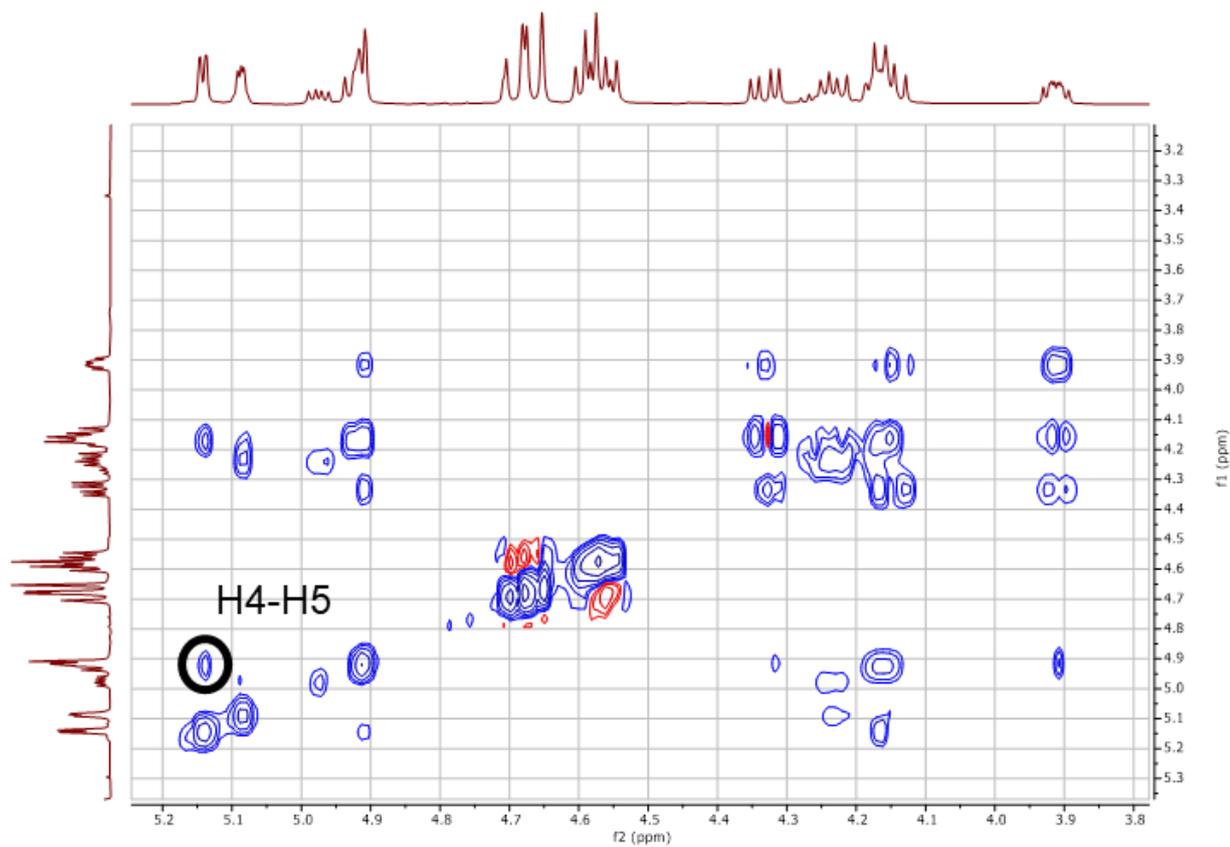


Figure S10. TOCSY experiment showing H4 and H5 cross peak.

Table S4. Tabulated chemical shifts of the 16.

Position	Degree of substitution	Major		Minor	
		^1H	$^{13}\text{C}\{^1\text{H}\}$	^1H	$^{13}\text{C}\{^1\text{H}\}$
1	CH (acetal)	4.93	99.92	4.97	99.74
2	CH ₂ (diastereotopic)	2.04	32.8	2.11	36.70
3	CH	4.17	77.14	4.24	80.00
4	CH	5.15	77.41	5.10	80.53
5	CH	4.91	78.97	5.08	78.76
6	CH	3.92	81.23	4.16	80.61
7	CH ₂ (diastereotopic)	4.33	63.84	4.24	63.52
7'	CH ₂ (diastereotopic)	4.16	63.84	4.24	63.52
Benzyl	CH ₂ (diastereotopic)	4.54-4.72	67.82	4.54-4.72	67.82

HMBC data

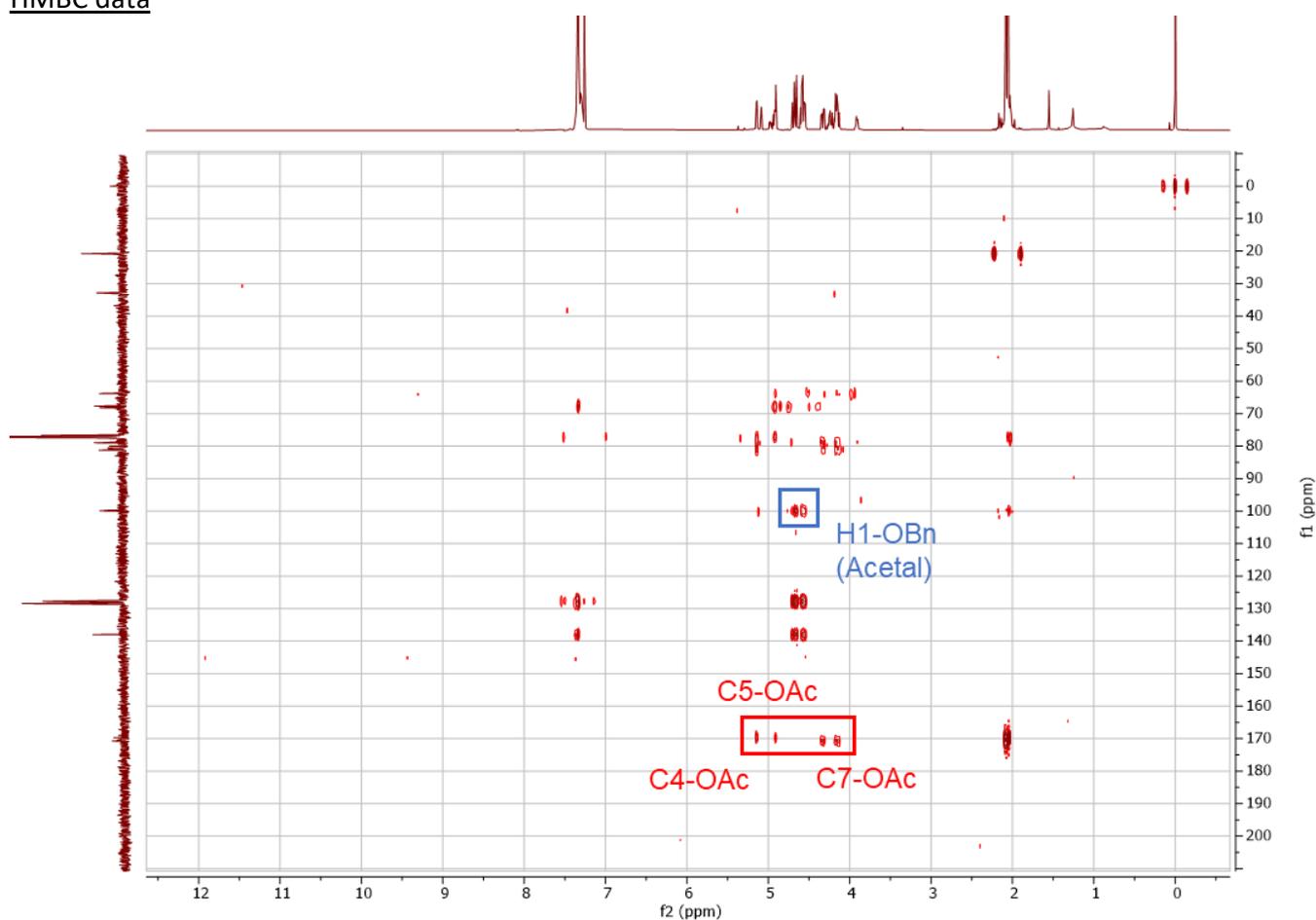
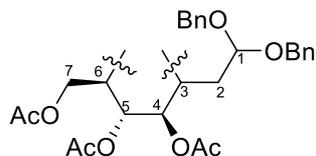


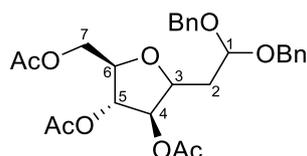
Figure S11. HMBC spectrum showing acetate cross peaks (red circles), benzyl-acetal cross peak (green circle), and C6-H3 cross peak.

Carbonyls (acetates) correlate with the H4, H5, and H7 (red square). Also, the HMBC shows that the acetal is a dibenzyl acetal (blue square).

Using the NMR information, the molecule is constructed as below. C6 and C3 still have one site to fill.



Its molecular weight is 484.54 and m/z 523.1927 [M+Na] was found. A mass of 16.00 is missing. Linking the C6 and C3 with an oxygen will account for the mass discrepancy to create the below molecule.



Stereochemistry at C3

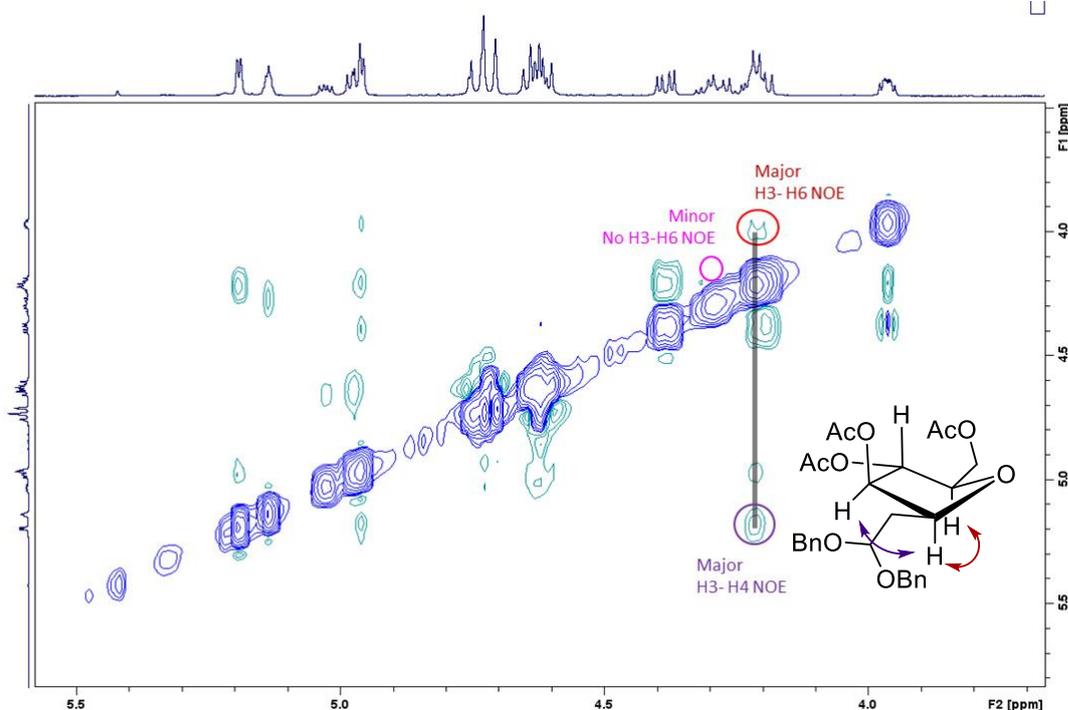
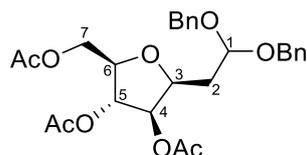
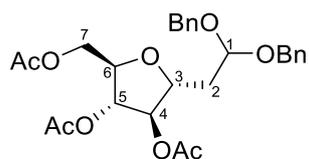


Figure S12. Zoomed version of the NOESY spectrum.

In the NOESY above, two different cross peaks suggest the major product is the *S* stereoisomer. The gray line corresponds to the major H3 chemical shift which is demonstrated in Figure S7. Along it there is a H3-H4 NOE (purple), a H3-H6 NOE (red). Together, this suggests that the H3 is on the same face as H4 and H6.



For the minor product, there is no NOE cross peak between H3 and H6 (pink circle), suggesting that they are on different faces of the molecule.

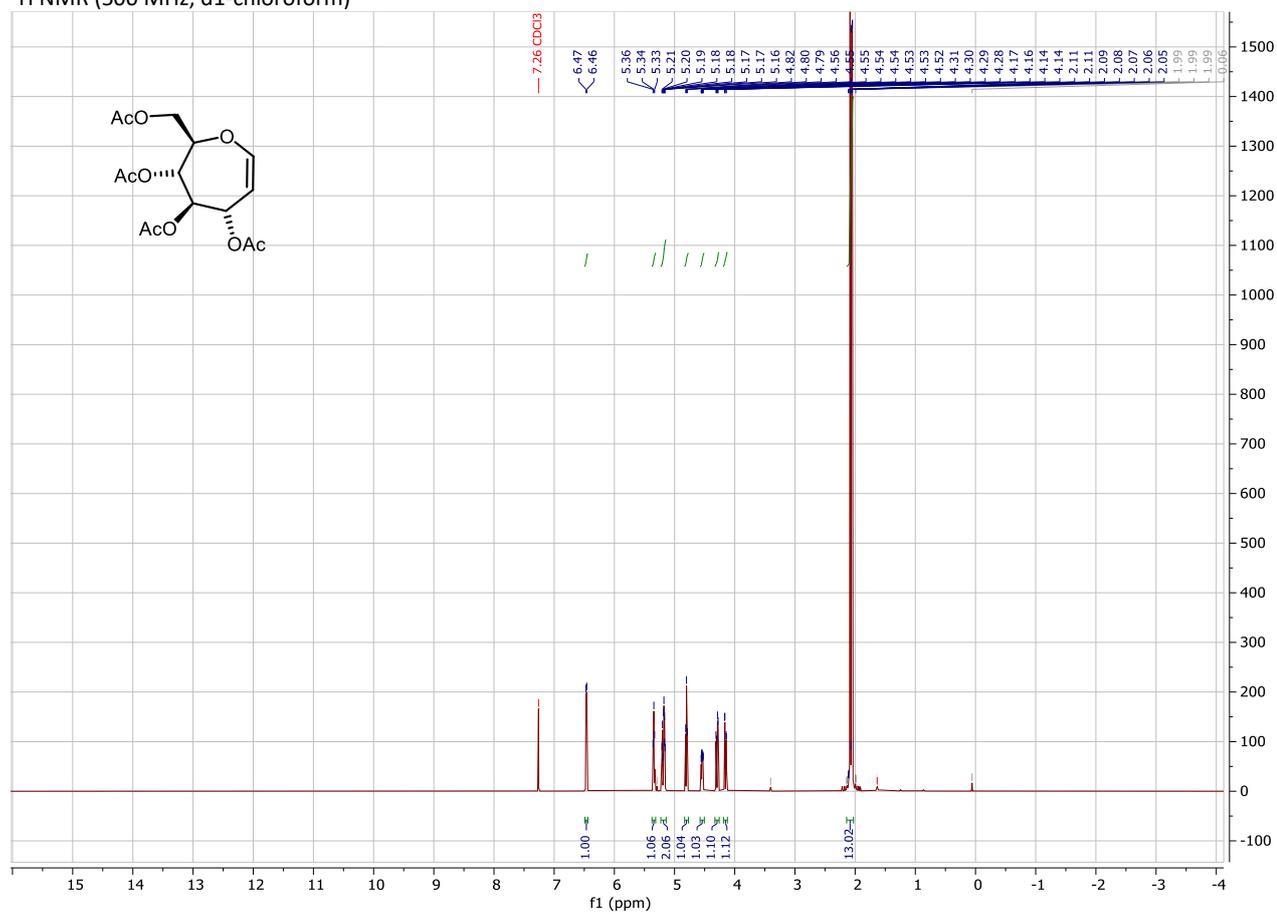


There is a 2:1 ratio between isomers. This means that 15 benzyl protons and 16.5 acetate/CH₂ protons are expected. Which matches the ¹H NMR.

Table S5. Common abbreviations used in the main text.

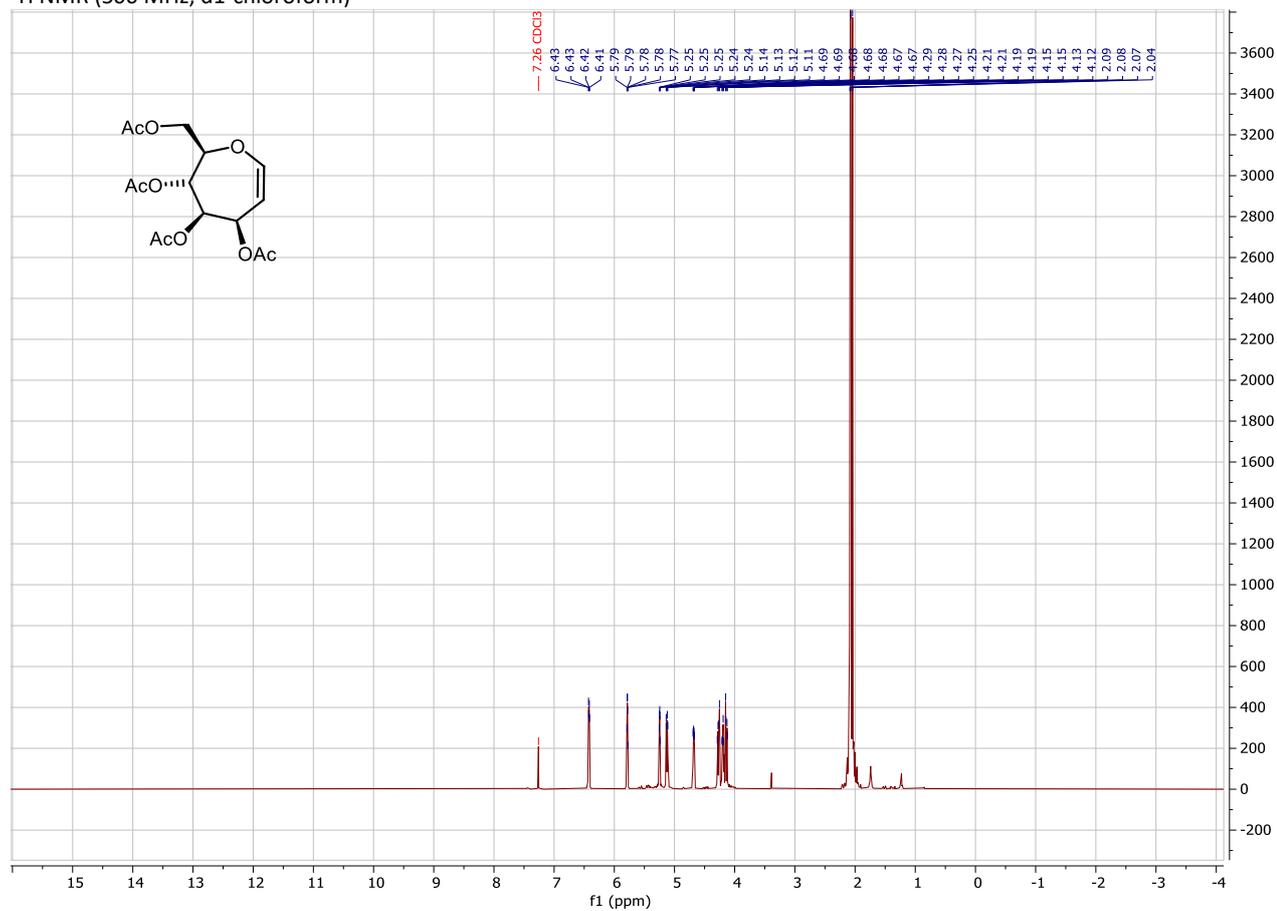
Abbreviations	Definitions
DMDO	Dimethyldioxirane
HFIP	Hexafluoroisopropanol
NGP	Neighboring-group participation
LRP	Long-range participation
TOCSY	Total correlation spectroscopy
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
HSQC	Heteronuclear single quantum coherence
HMBC	Heteronuclear multiple bond correlation
BRSM	Based on recovered starting material

¹H NMR of compound 8
¹H NMR (500 MHz, d1-chloroform)



S22

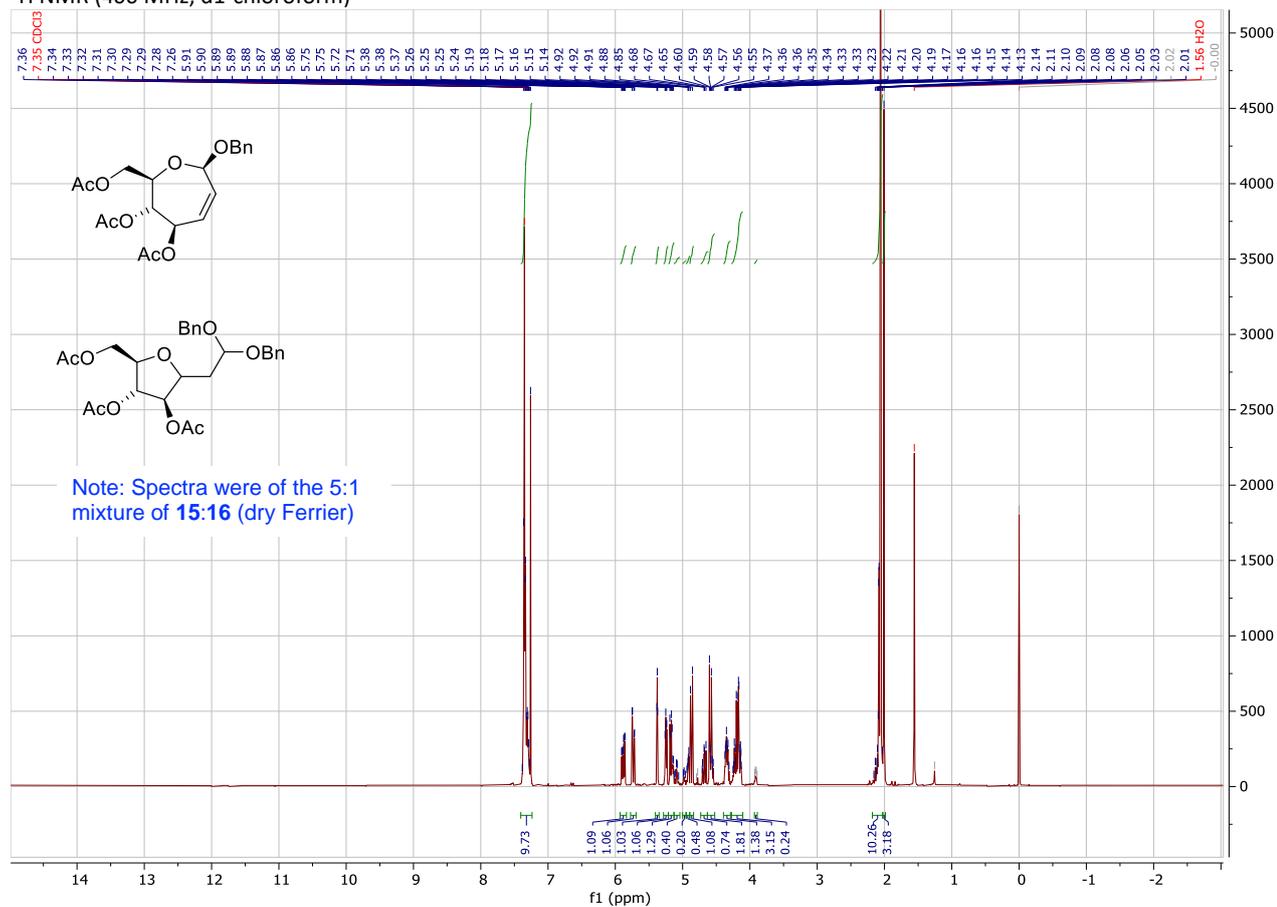
¹H NMR of compound 10
¹H NMR (500 MHz, d1-chloroform)



S23

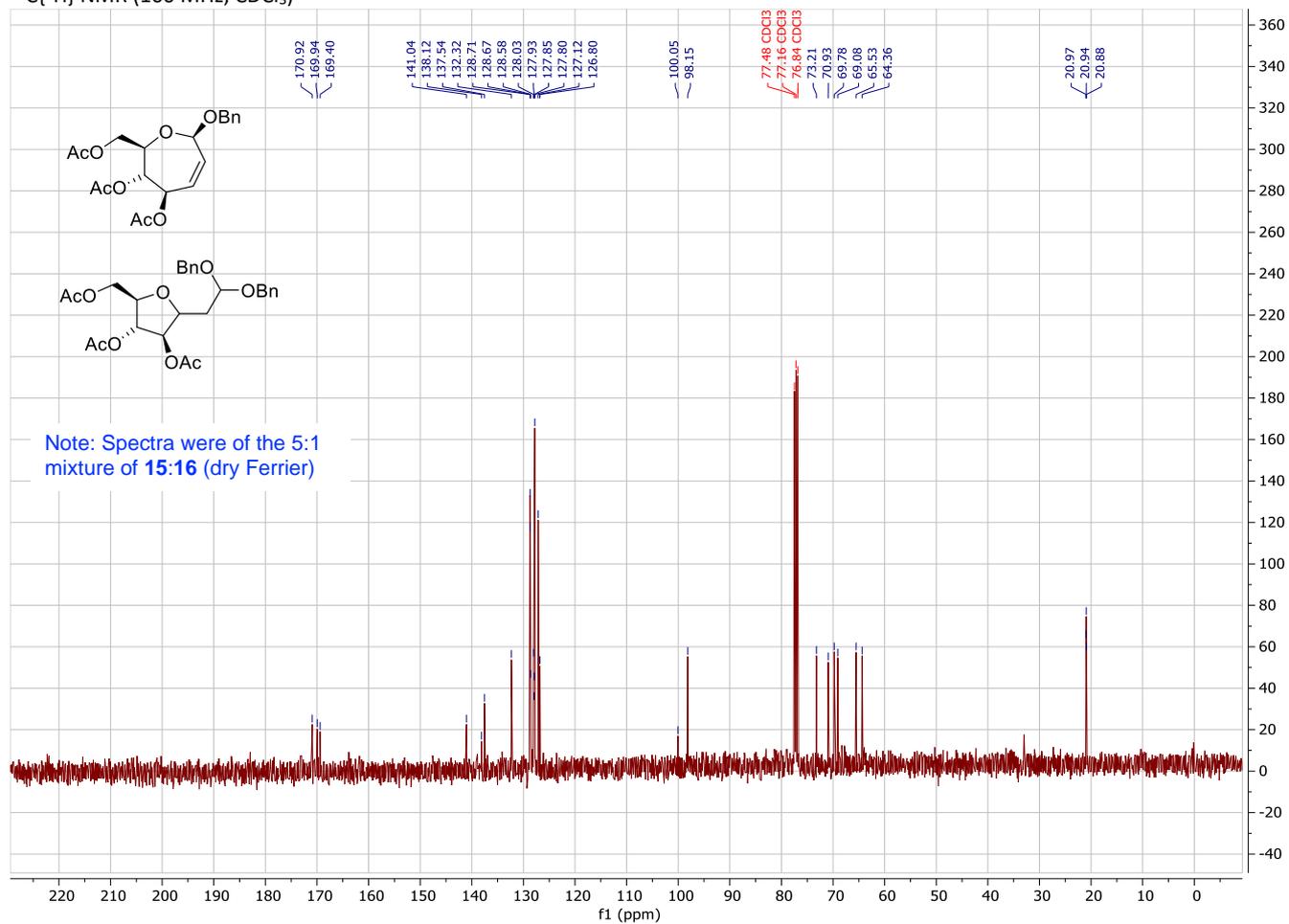
NMR spectra used to characterize compound 15

¹H NMR (400 MHz, d1-chloroform)

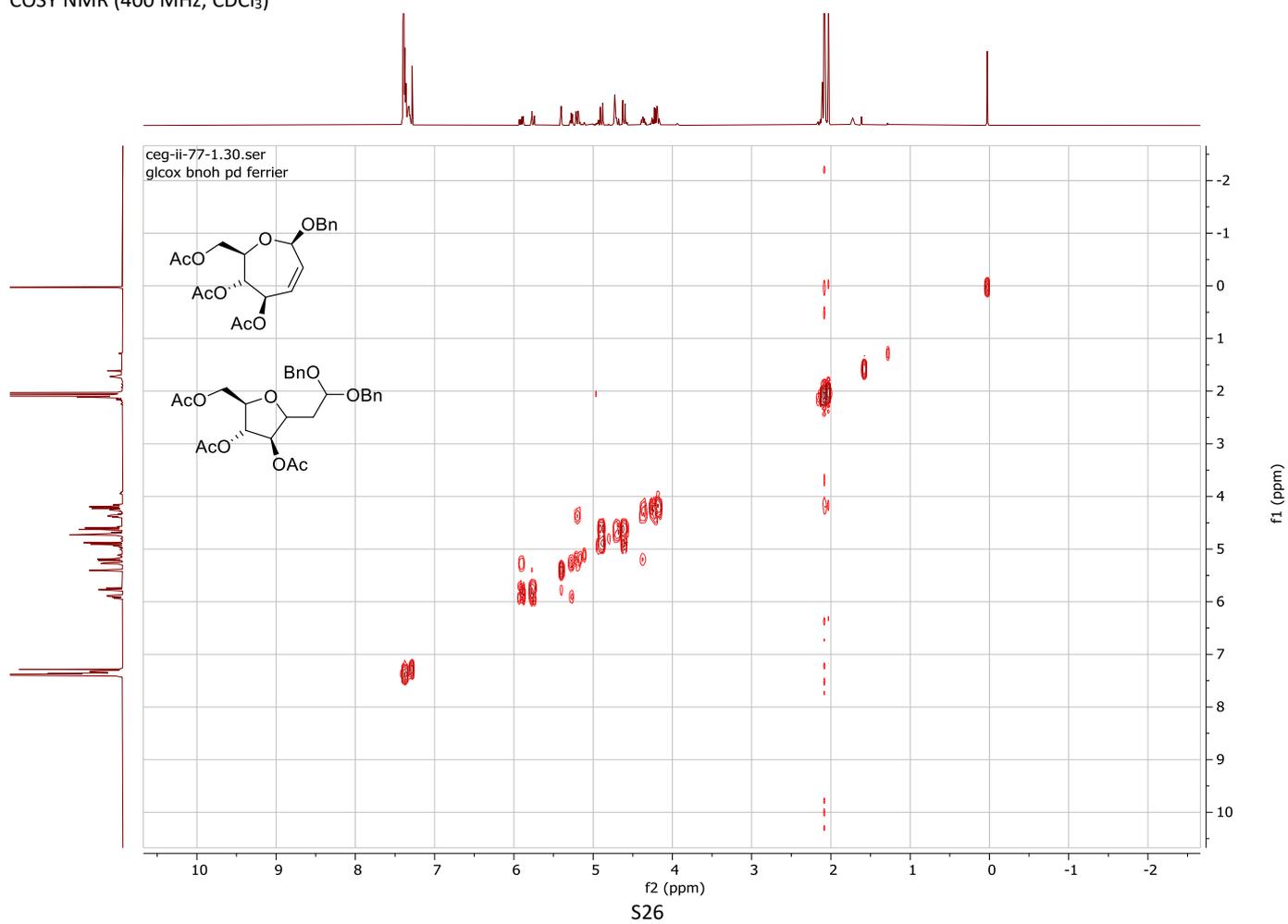


S24

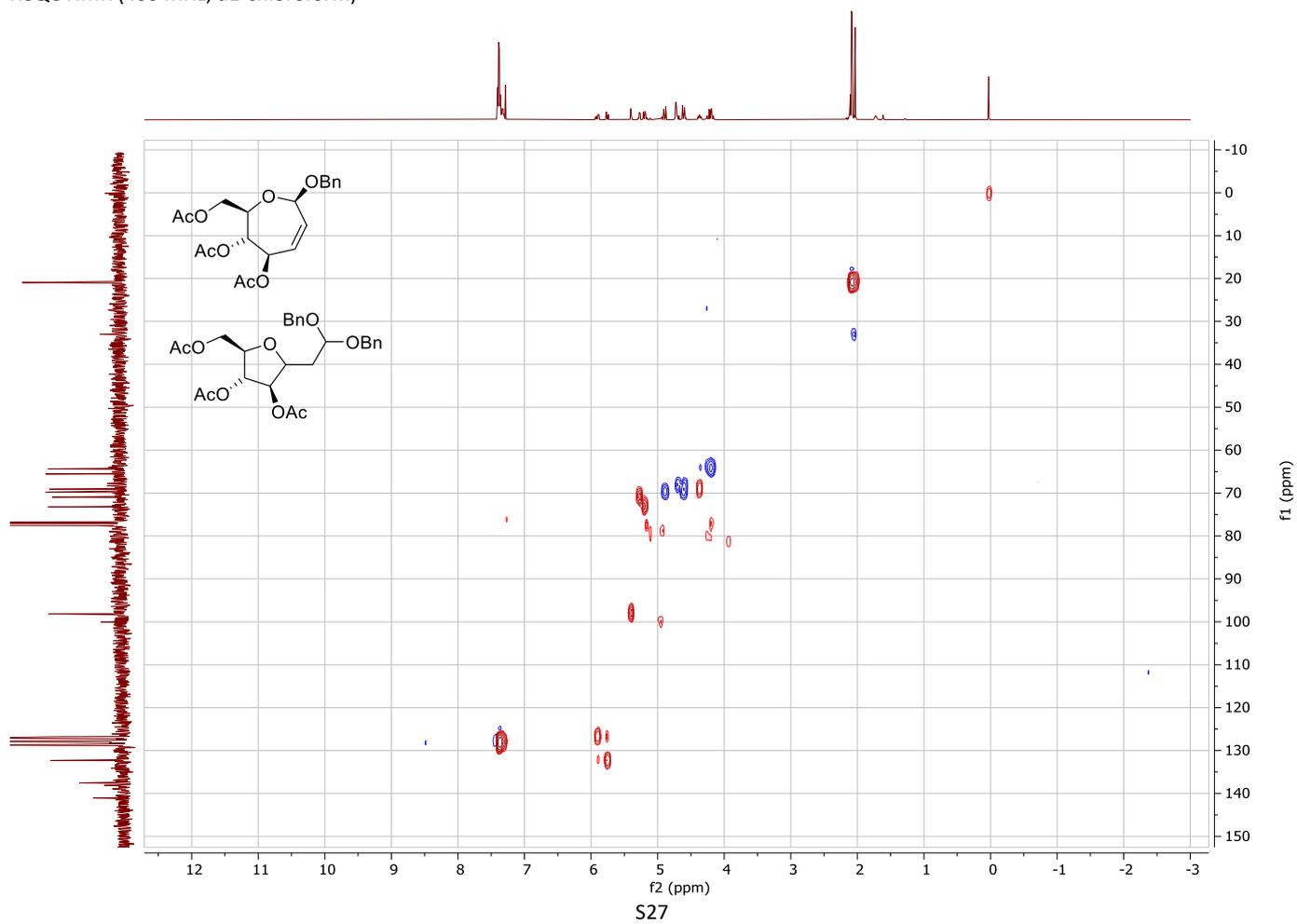
$^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, CDCl_3)



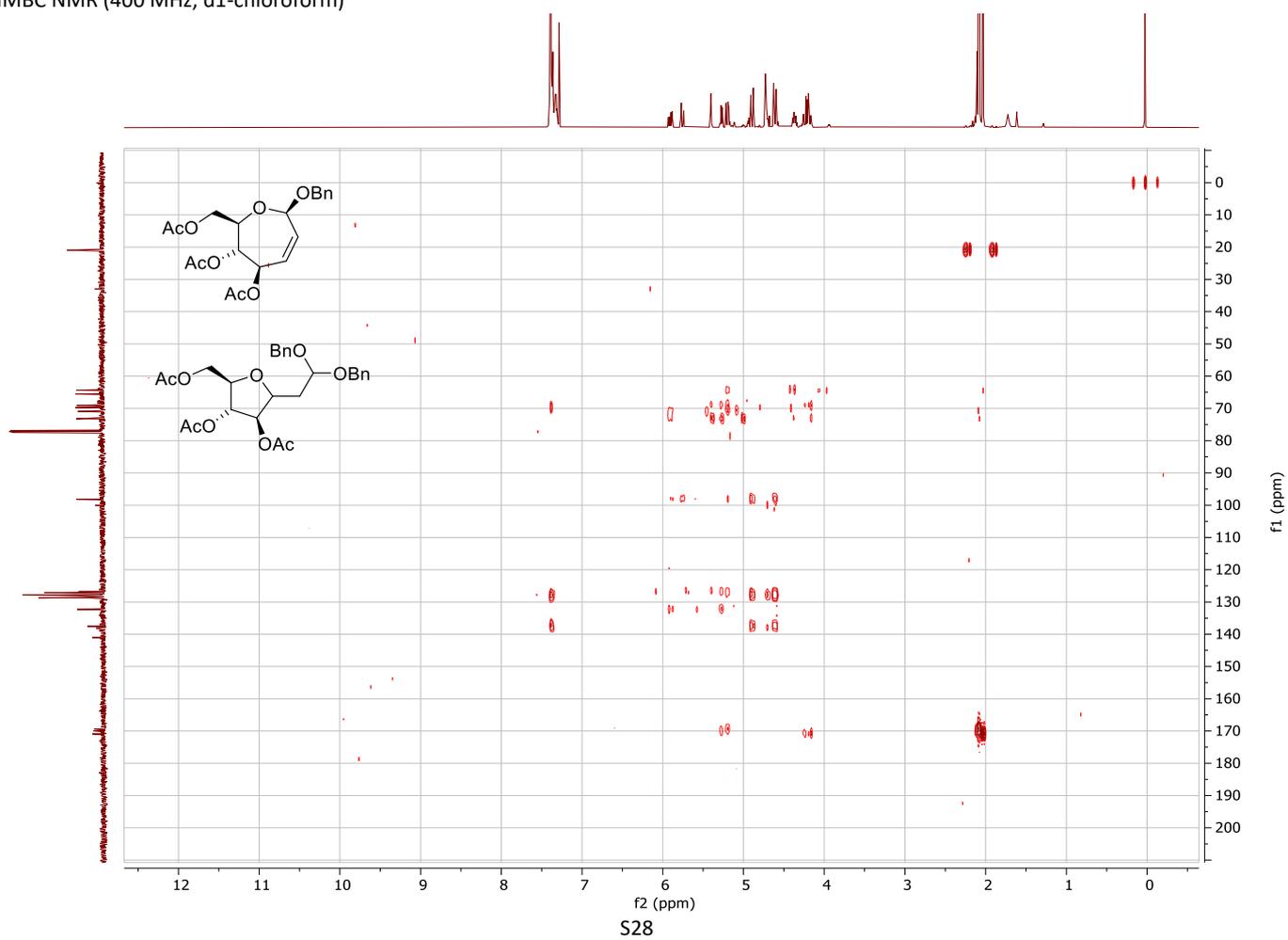
COSY NMR (400 MHz, CDCl₃)



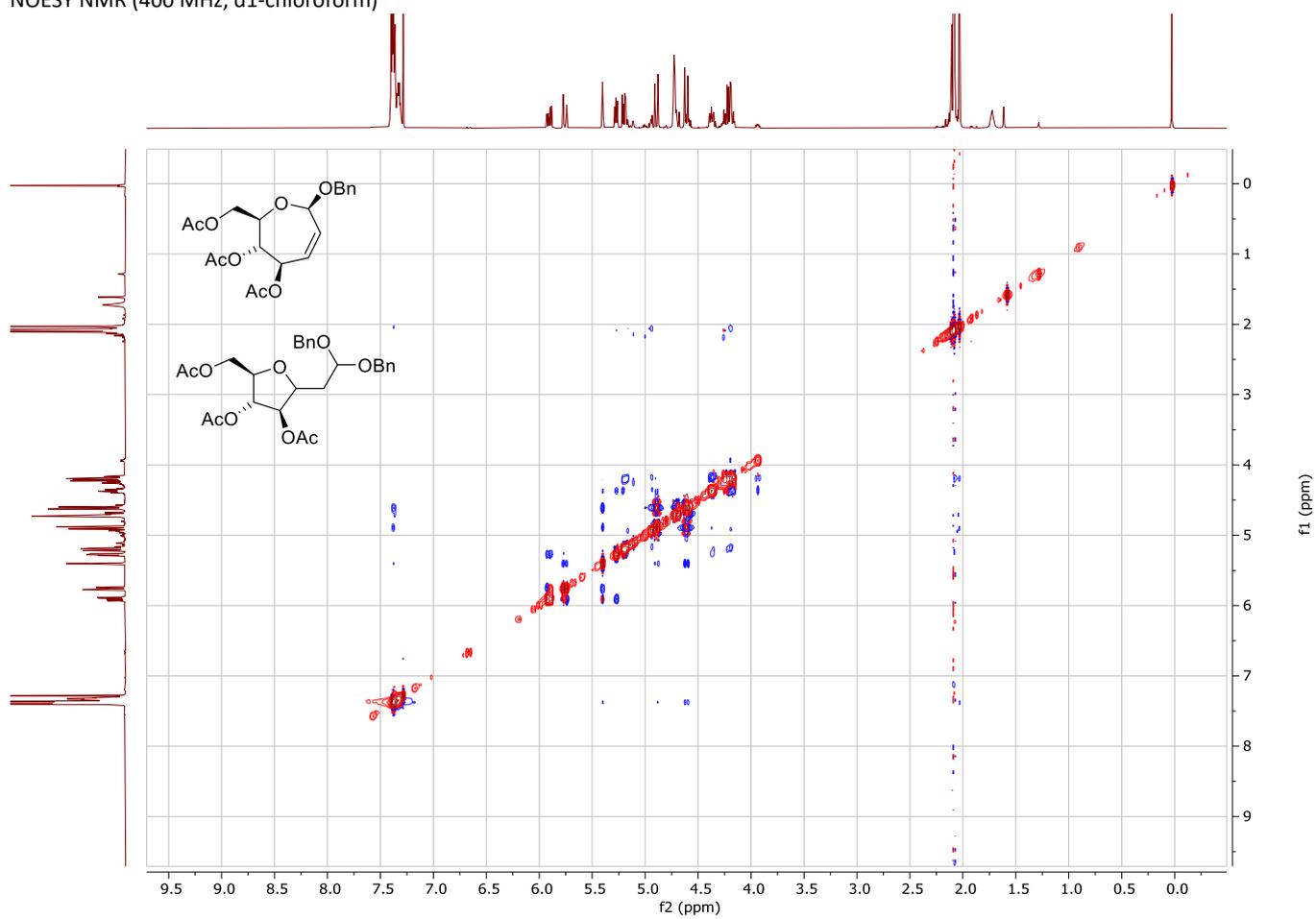
HSQC NMR (400 MHz, d1-chloroform)



HMBC NMR (400 MHz, d1-chloroform)



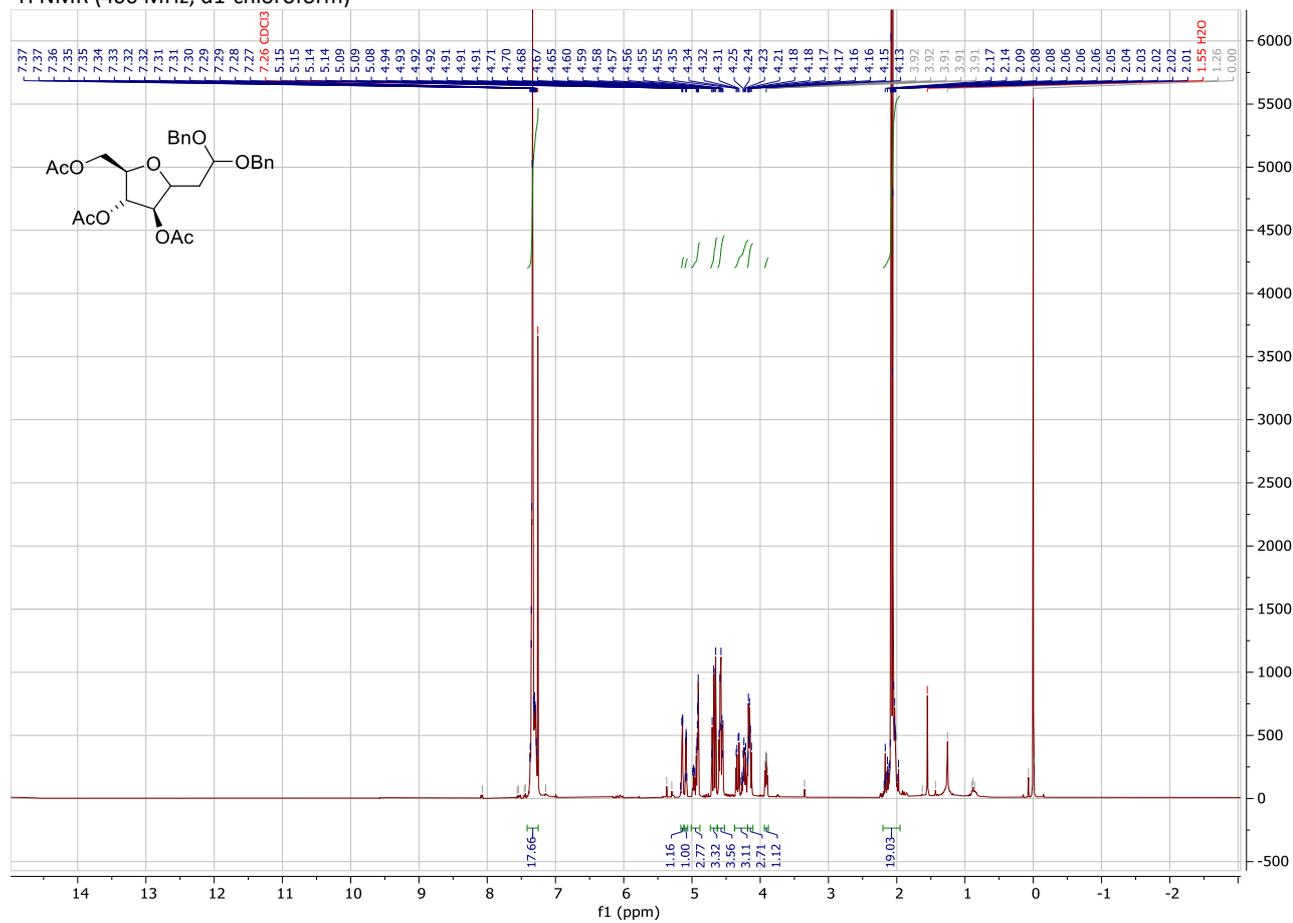
NOESY NMR (400 MHz, d1-chloroform)



S29

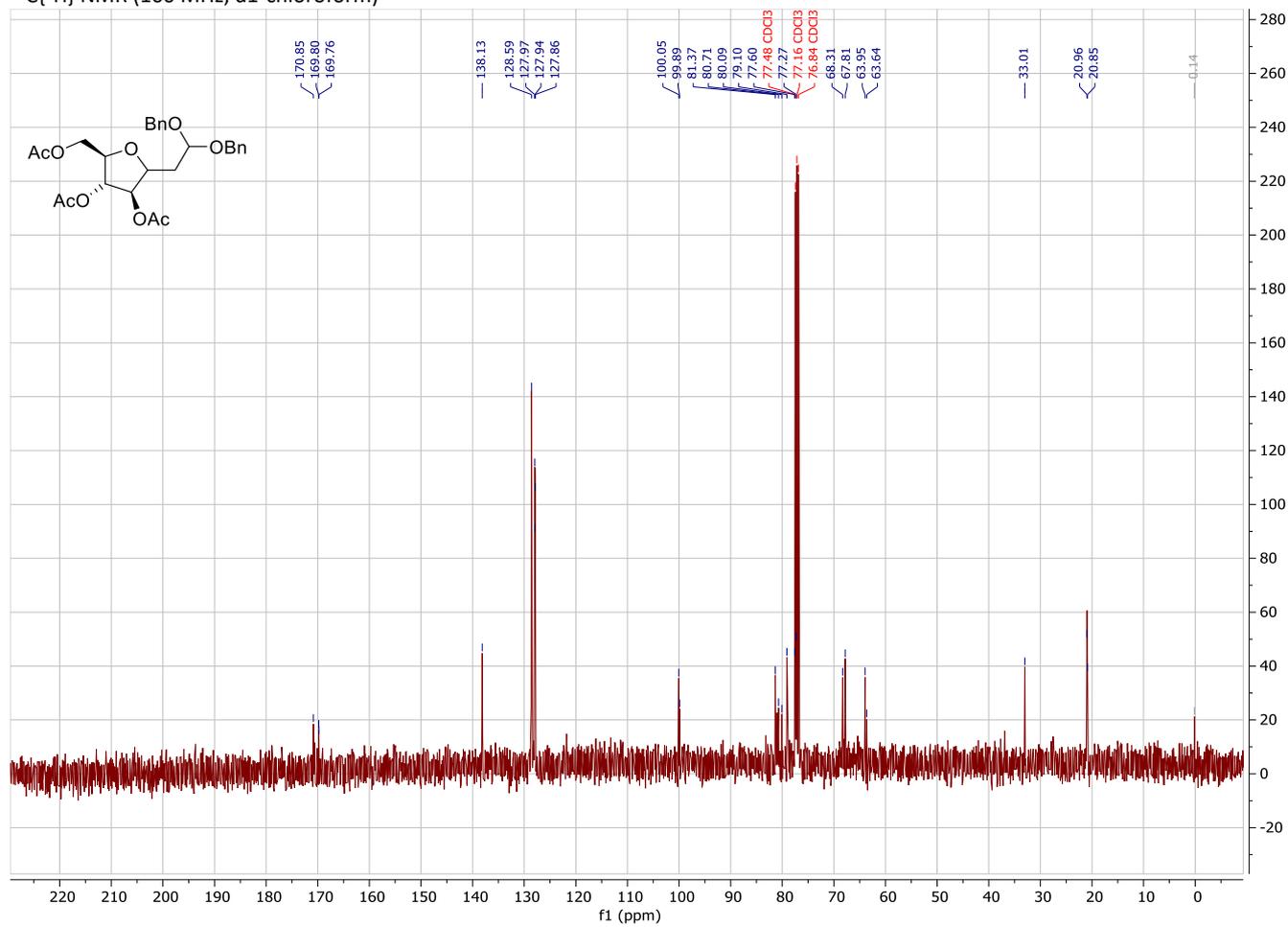
NMR characterization data of 16 – wet Ferrier product

¹H NMR (400 MHz, d1-chloroform)



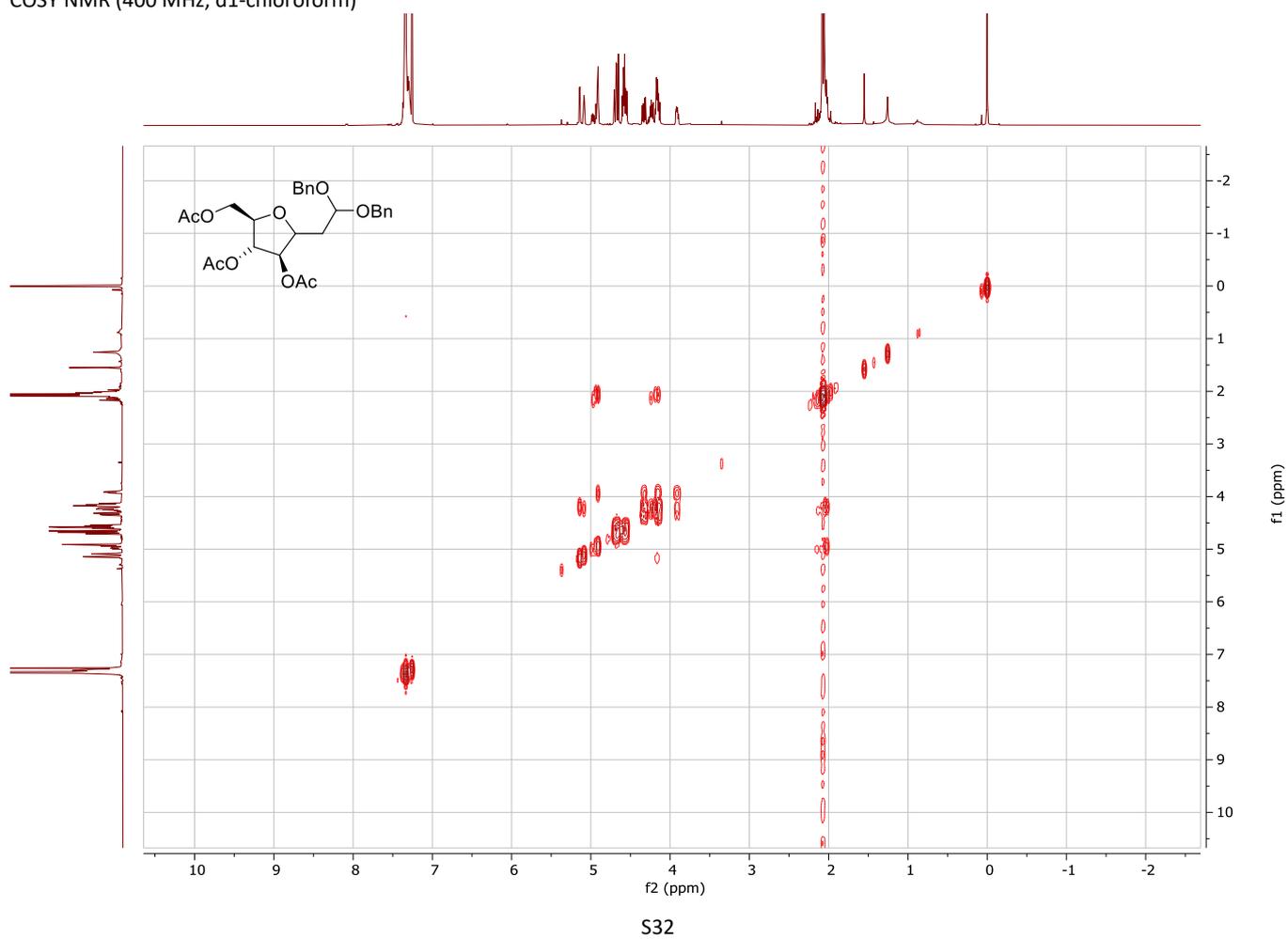
S30

$^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, $\text{d}_1\text{-chloroform}$)

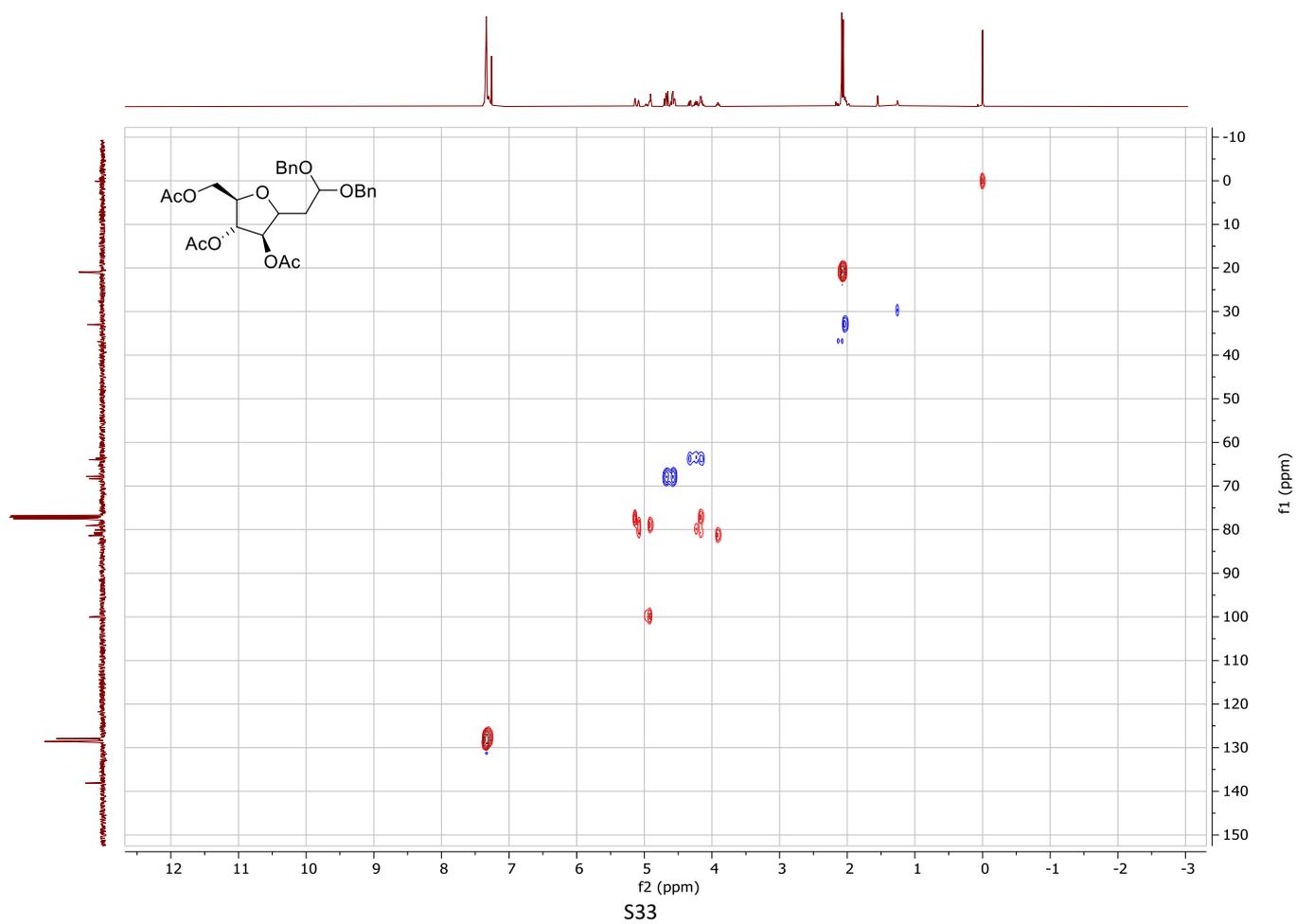


S31

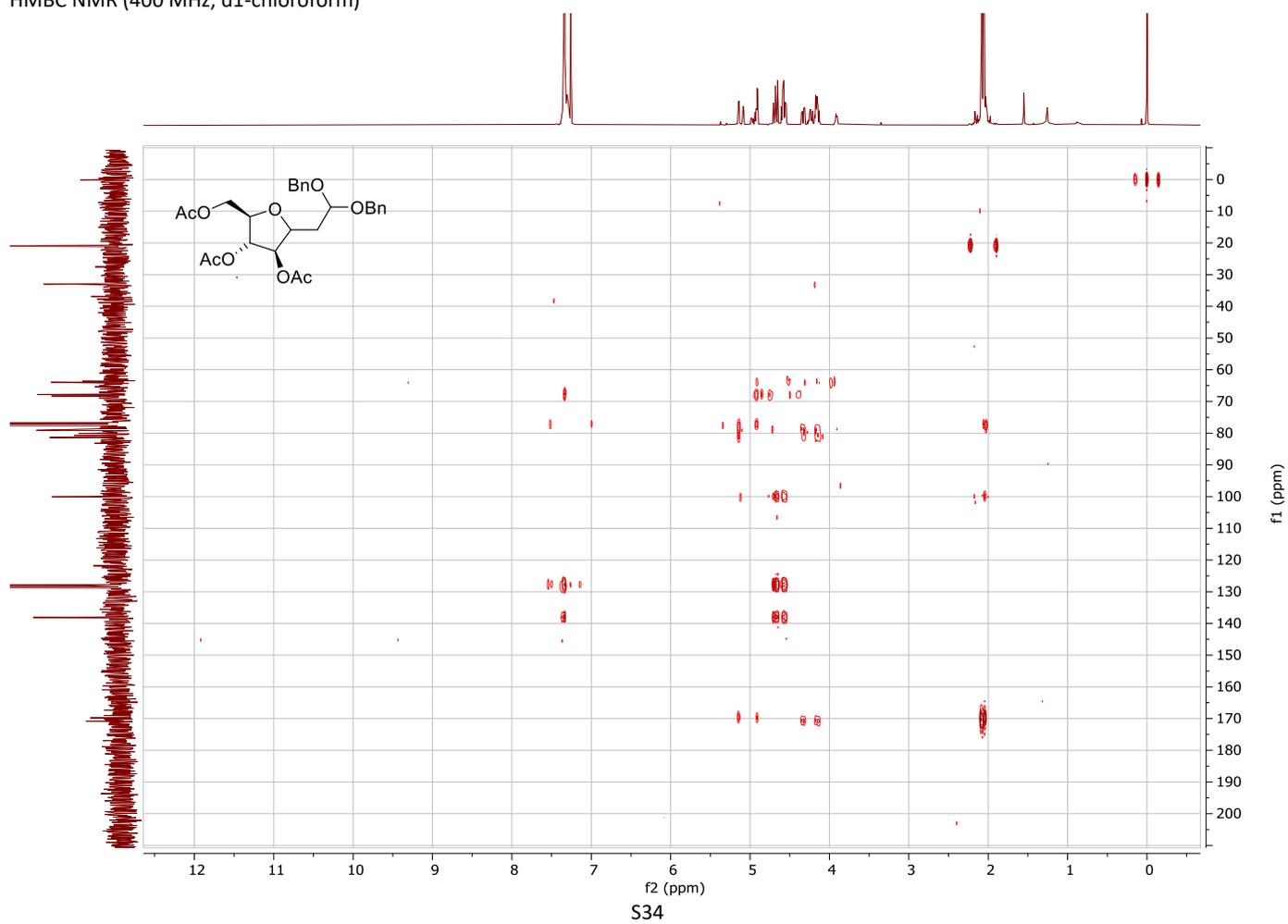
COSY NMR (400 MHz, d1-chloroform)



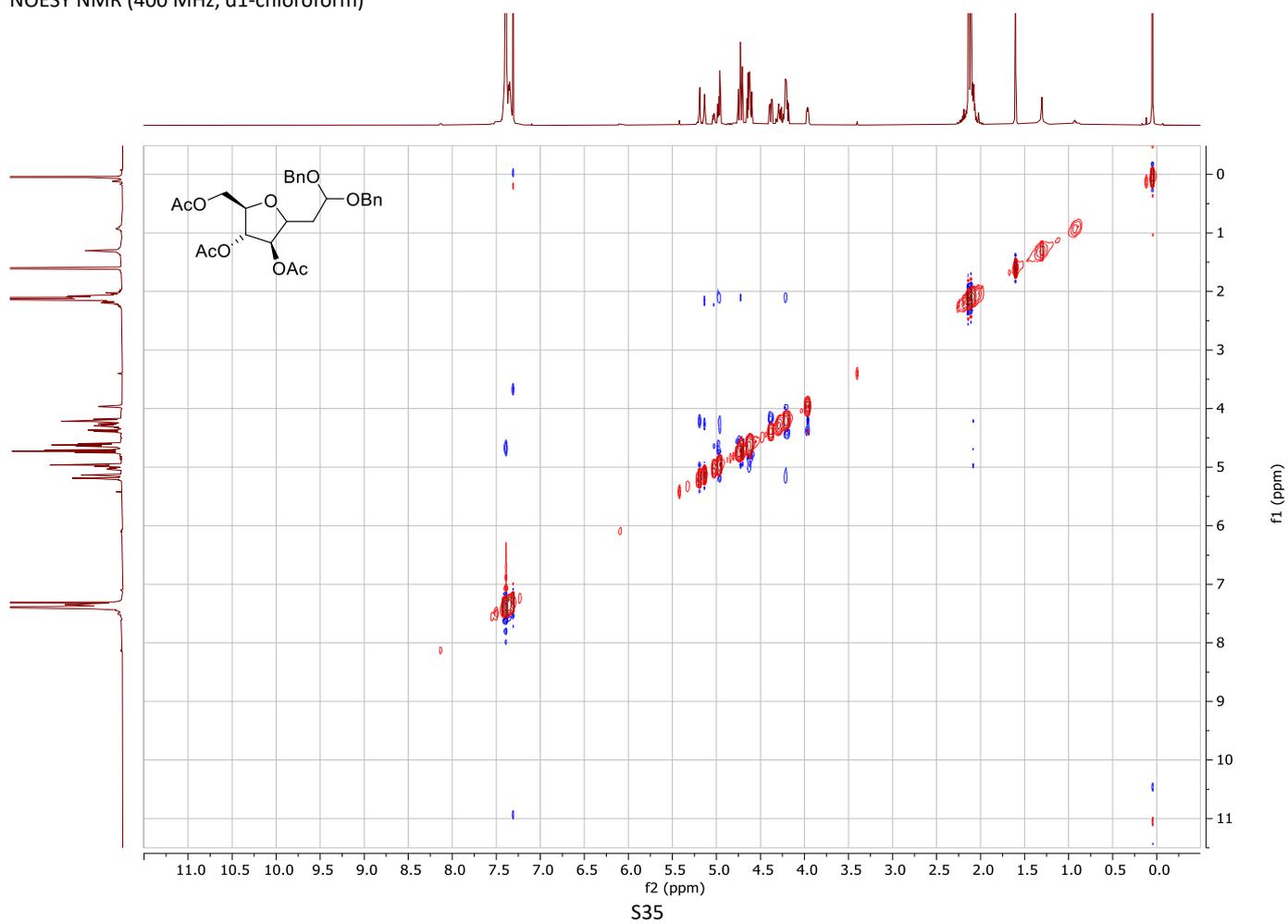
HSQC NMR (400 MHz, d1-chloroform)



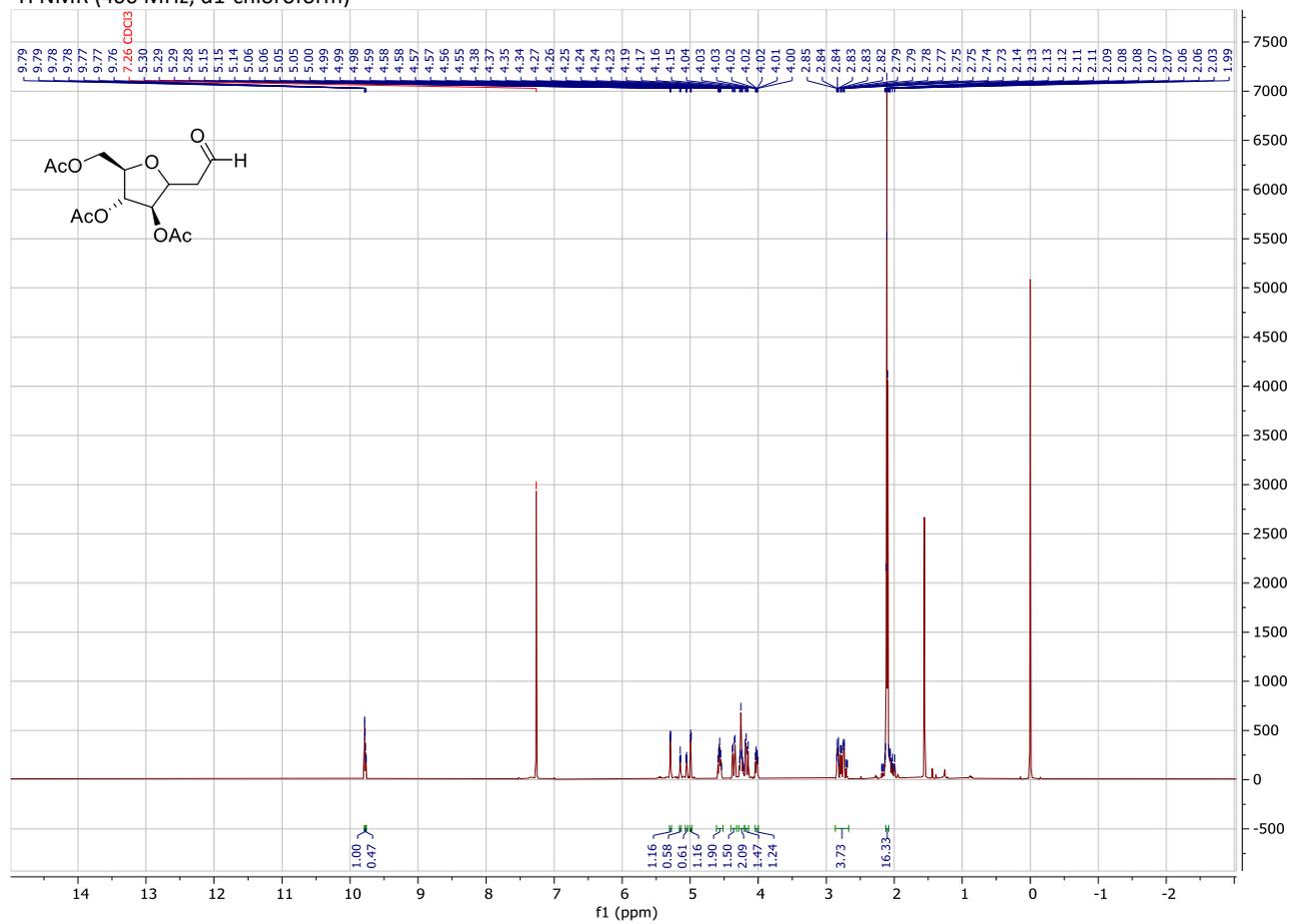
HMBC NMR (400 MHz, d1-chloroform)



NOESY NMR (400 MHz, d1-chloroform)

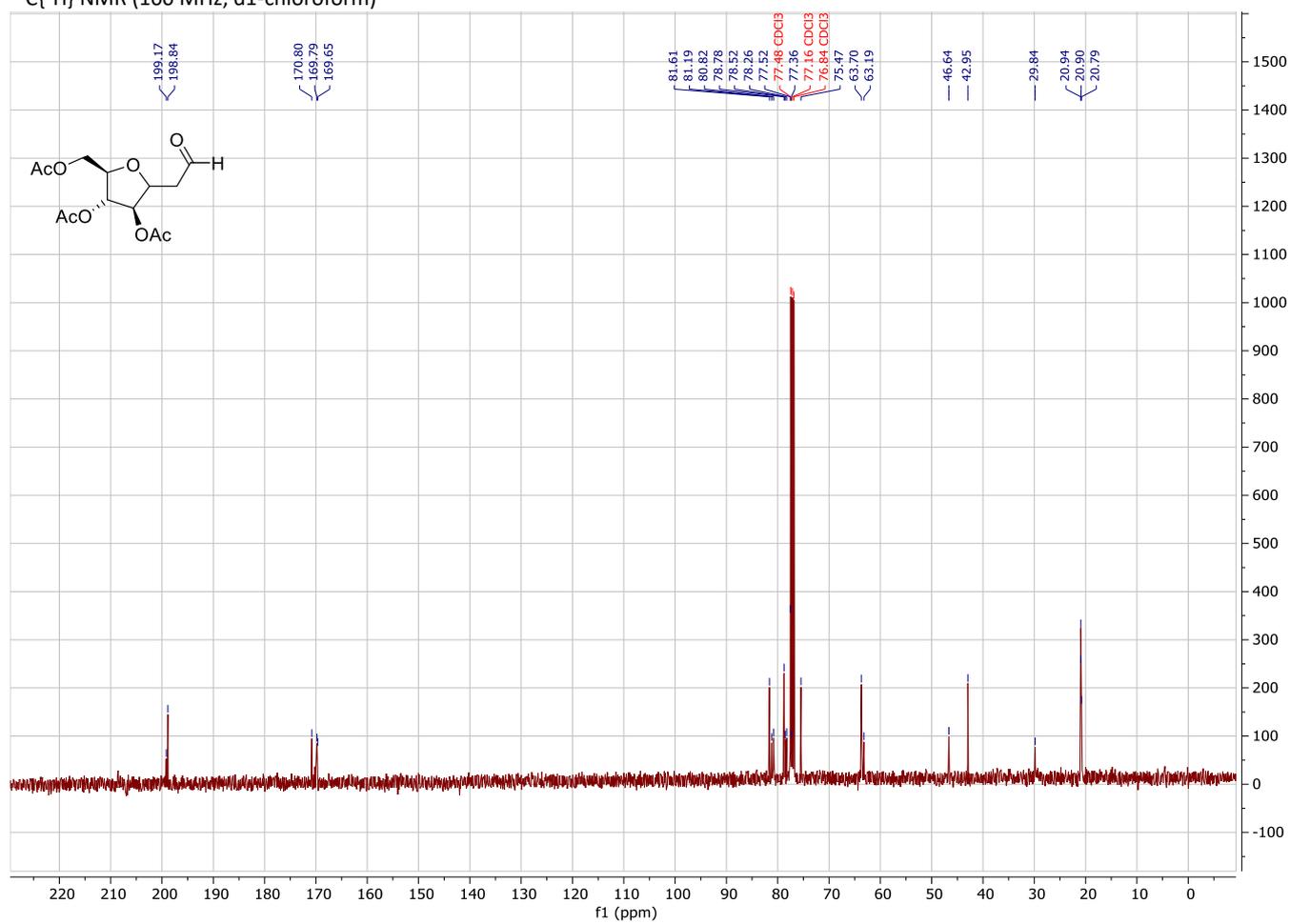


NMR characterization data of 17
¹H NMR (400 MHz, d1-chloroform)



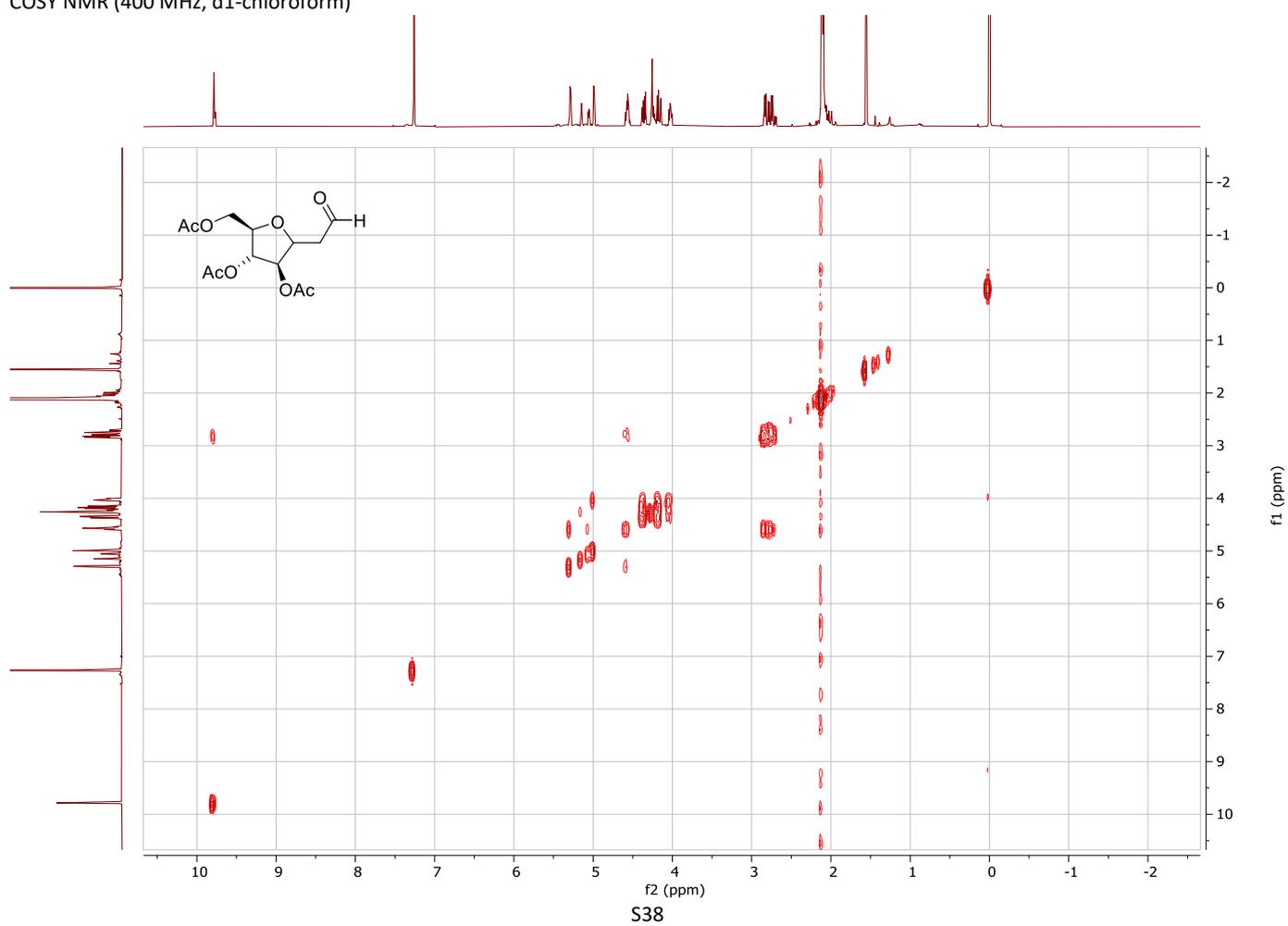
S36

$^{13}\text{C}\{^1\text{H}\}$ NMR (100 MHz, $\text{d}_1\text{-chloroform}$)

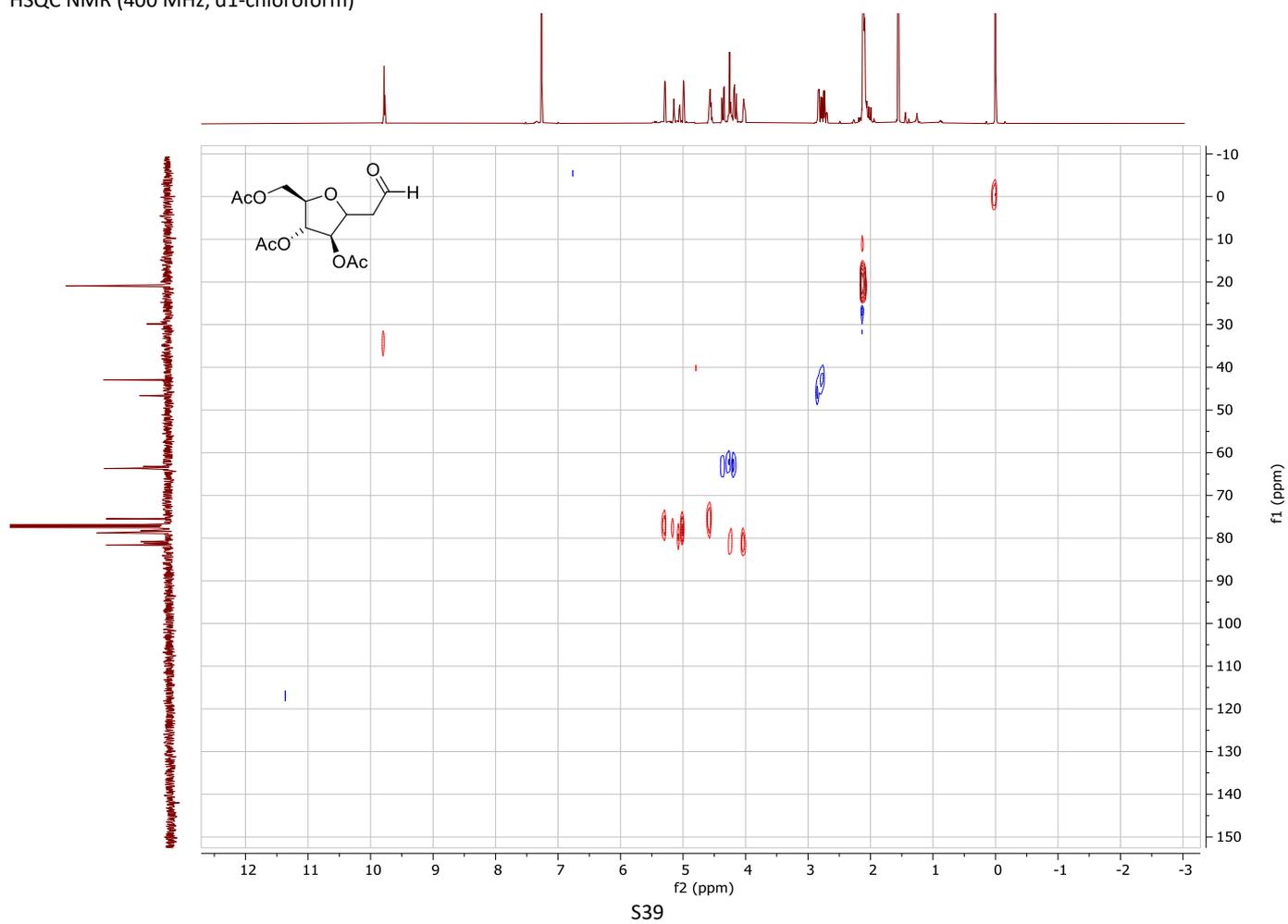


S37

COSY NMR (400 MHz, d1-chloroform)

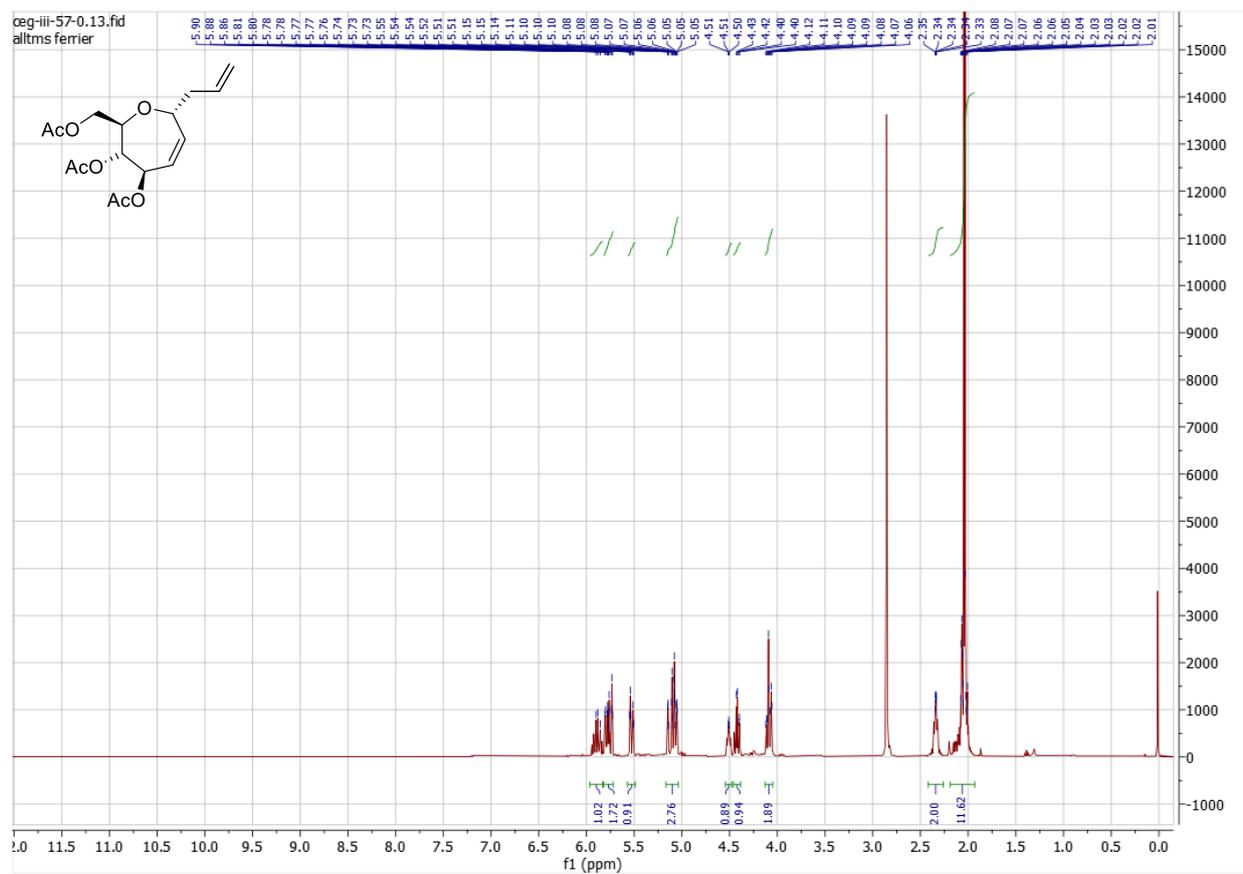


HSQC NMR (400 MHz, d1-chloroform)



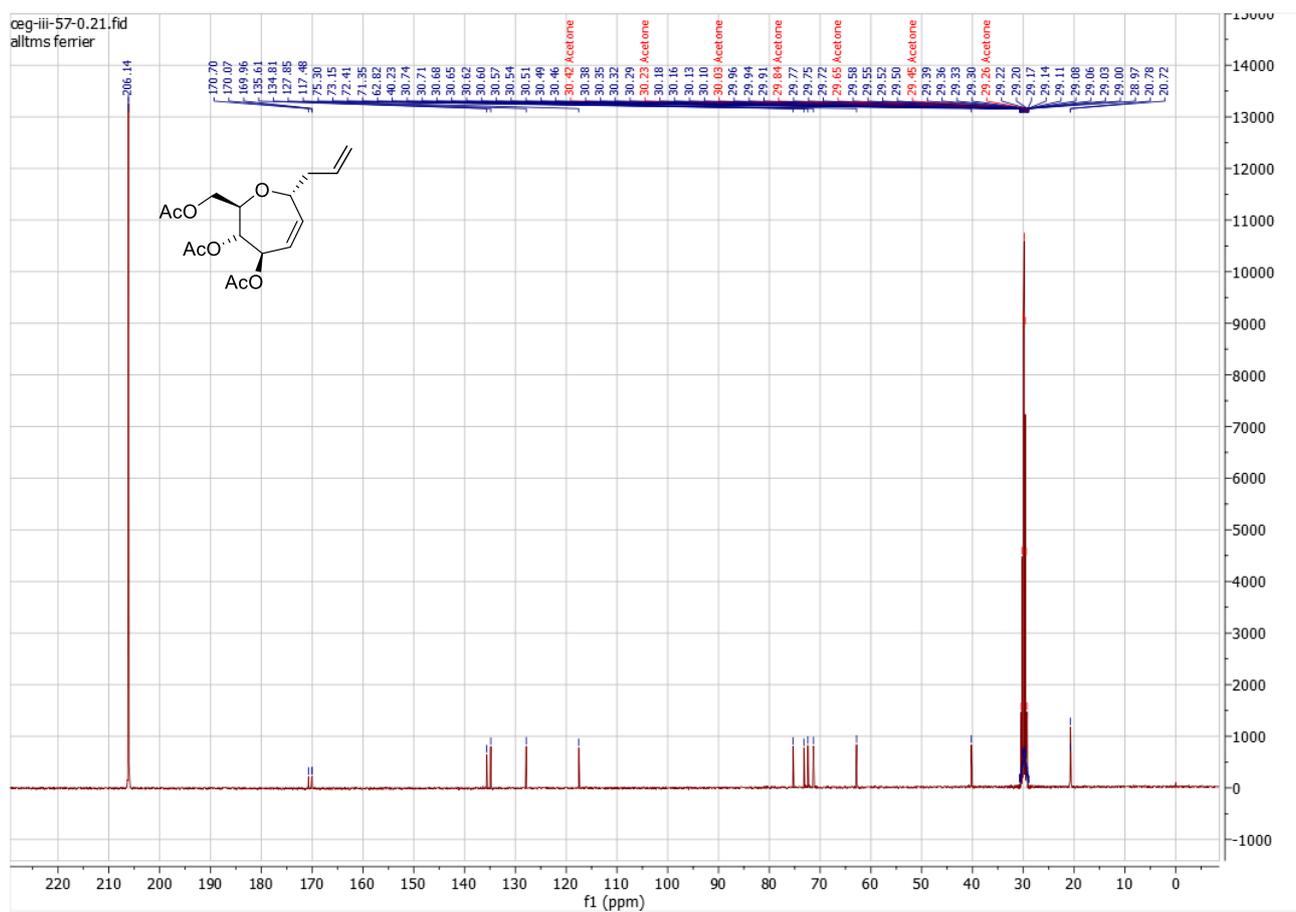
NMR spectra used to characterize compound 18

¹H NMR (400 MHz, d6-acetone)



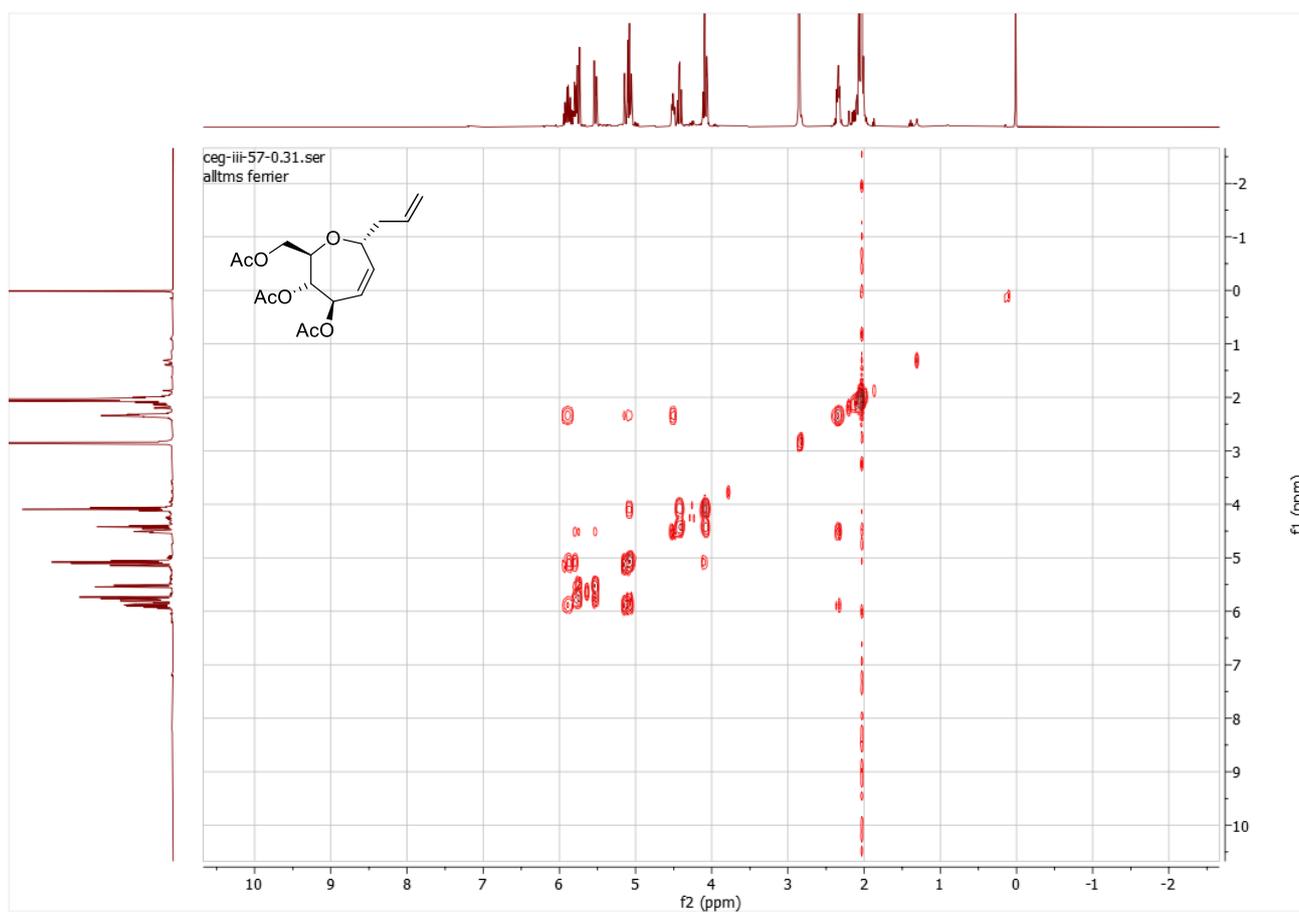
S40

$^{13}\text{C}\{^1\text{H}\}$ NMR (400 MHz, d_6 -acetone)



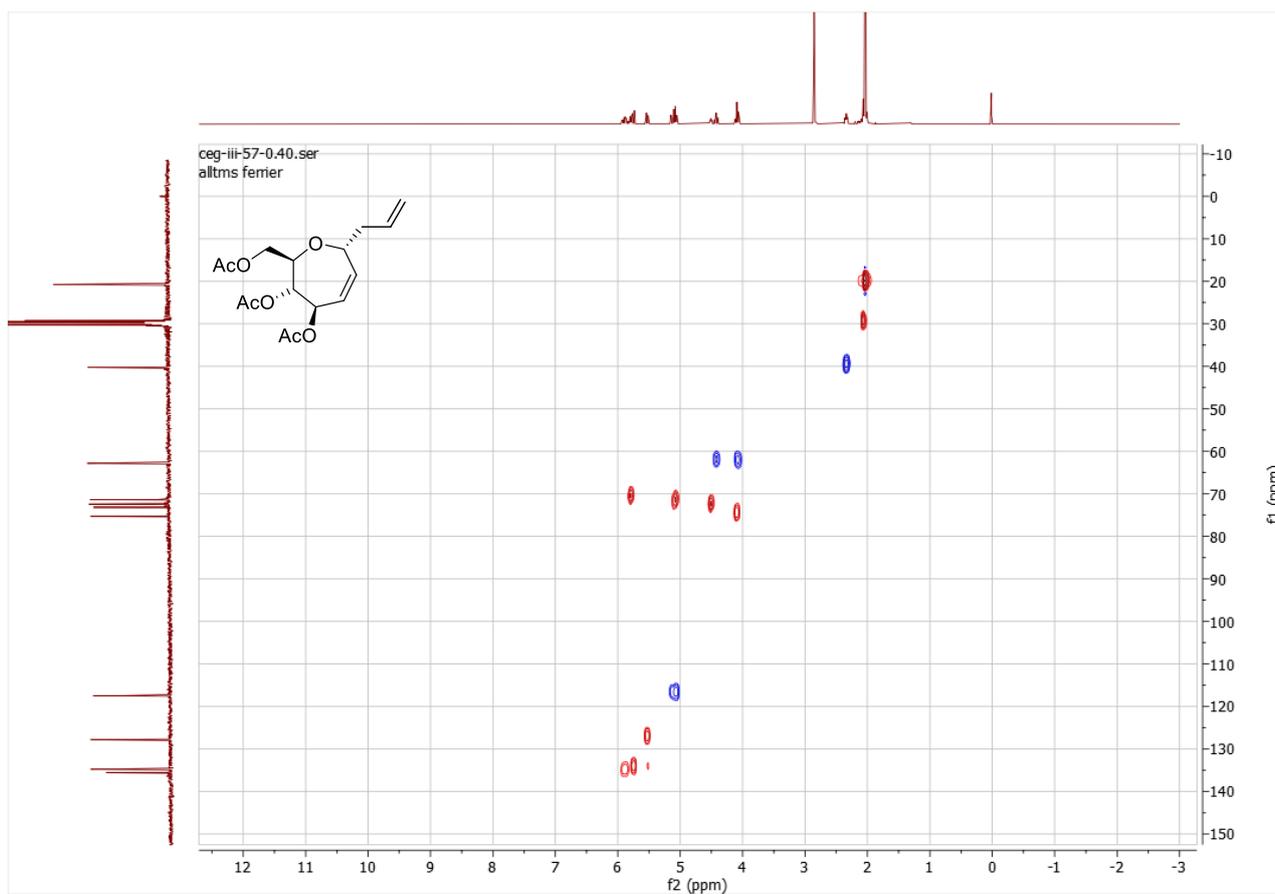
S41

COSY NMR (400 MHz, d6-acetone)



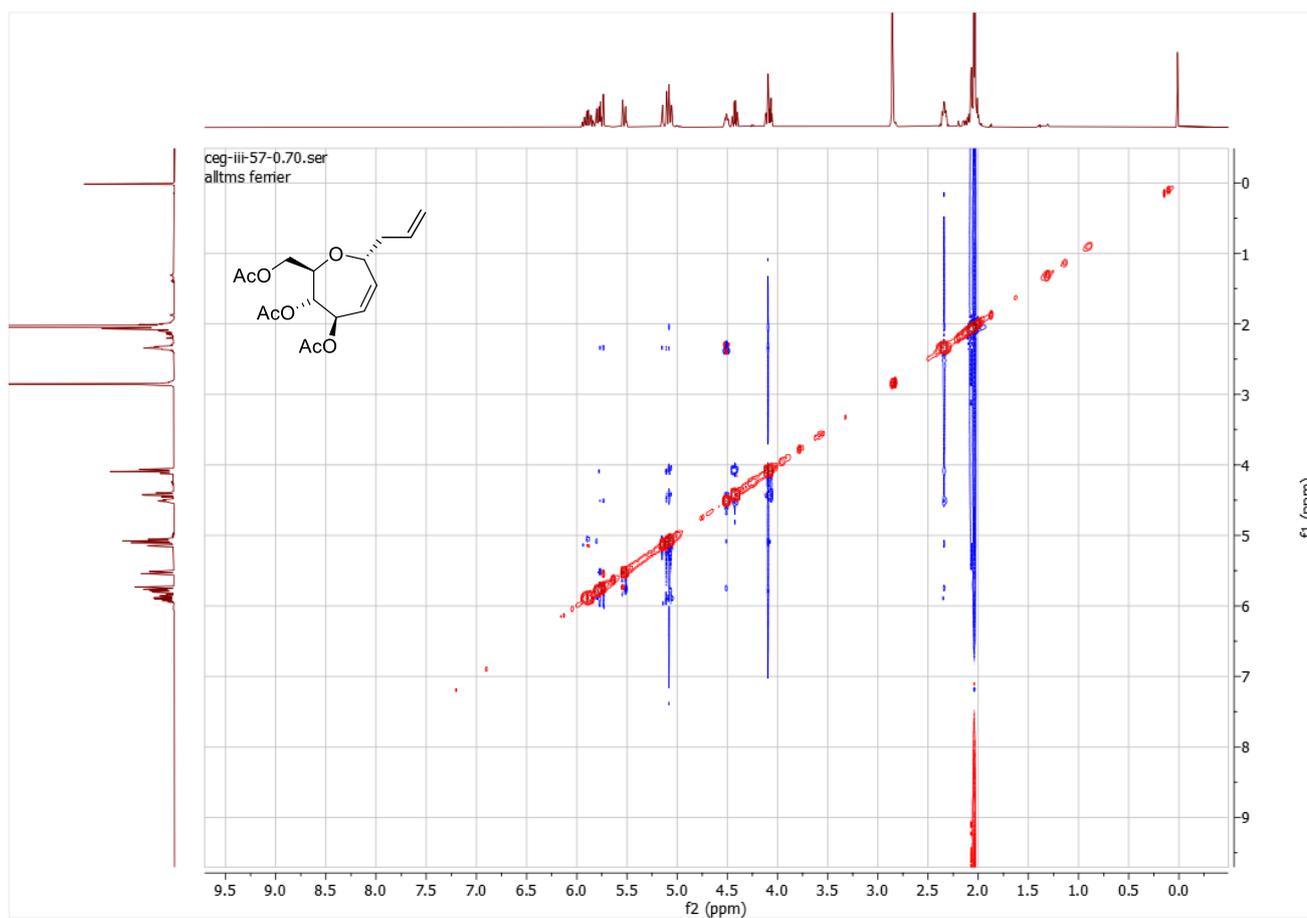
S42

HSQC NMR (400 MHz, d6-acetone)



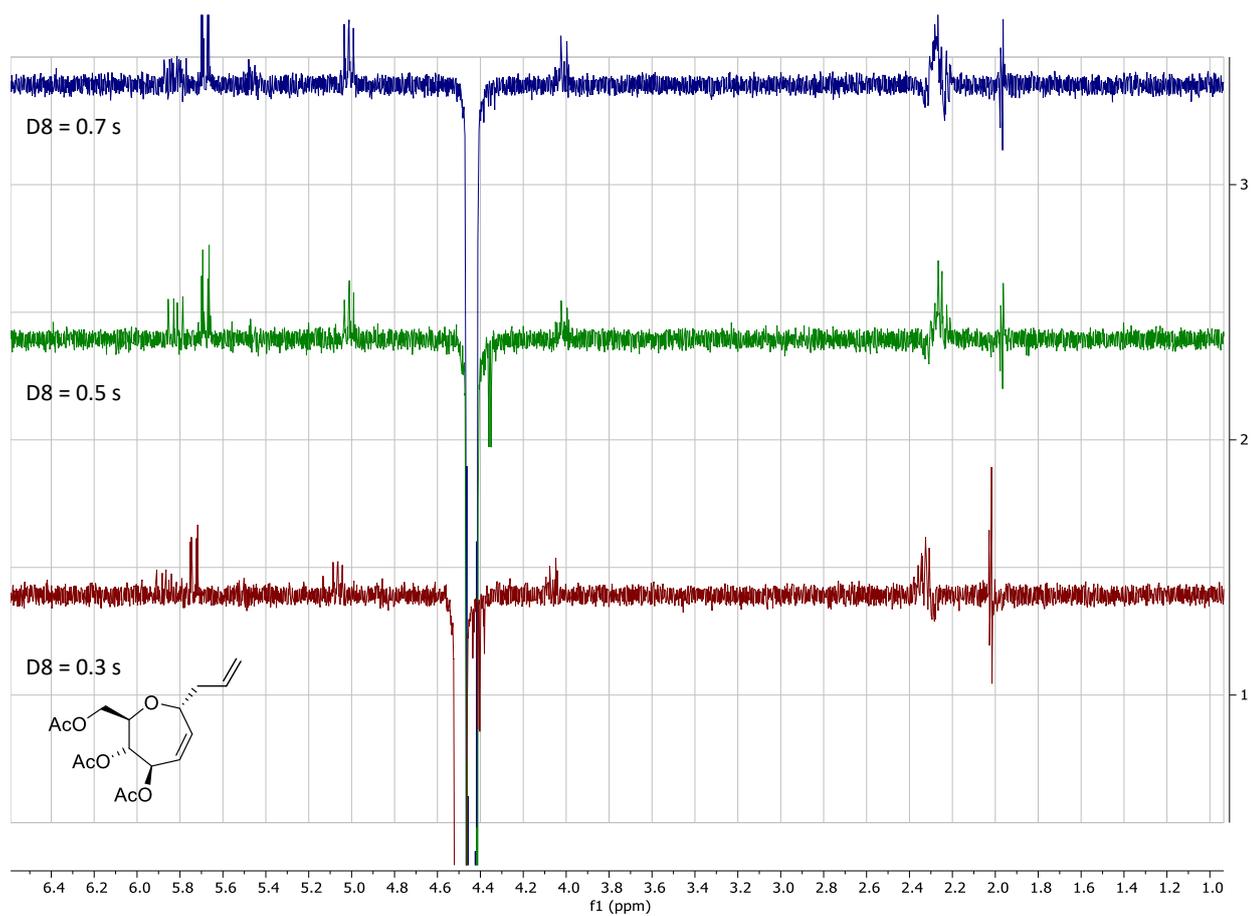
S43

NOESY NMR (400 MHz, d6-acetone)



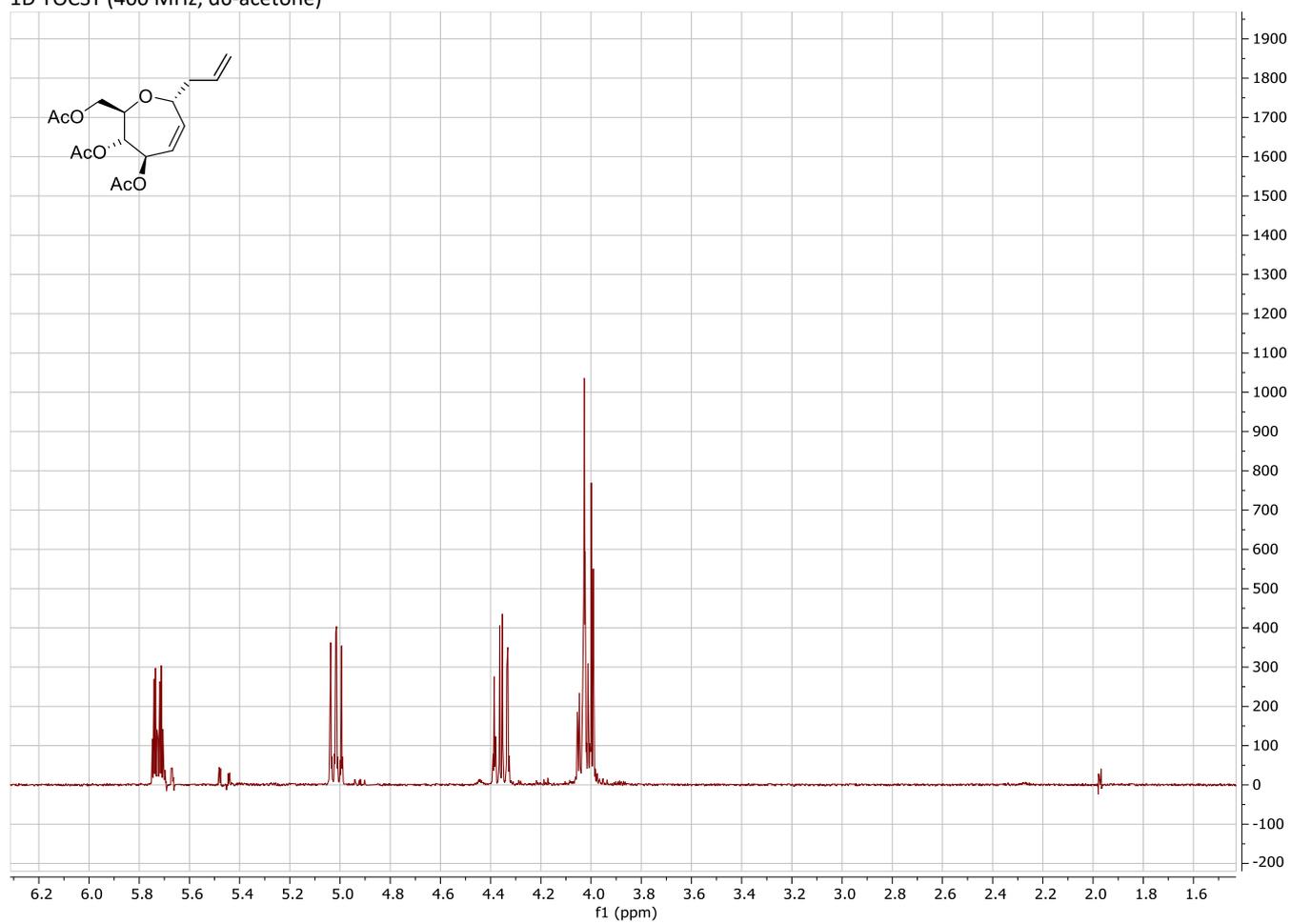
S44

1D NOE NMR (400 MHz, d6-acetone)



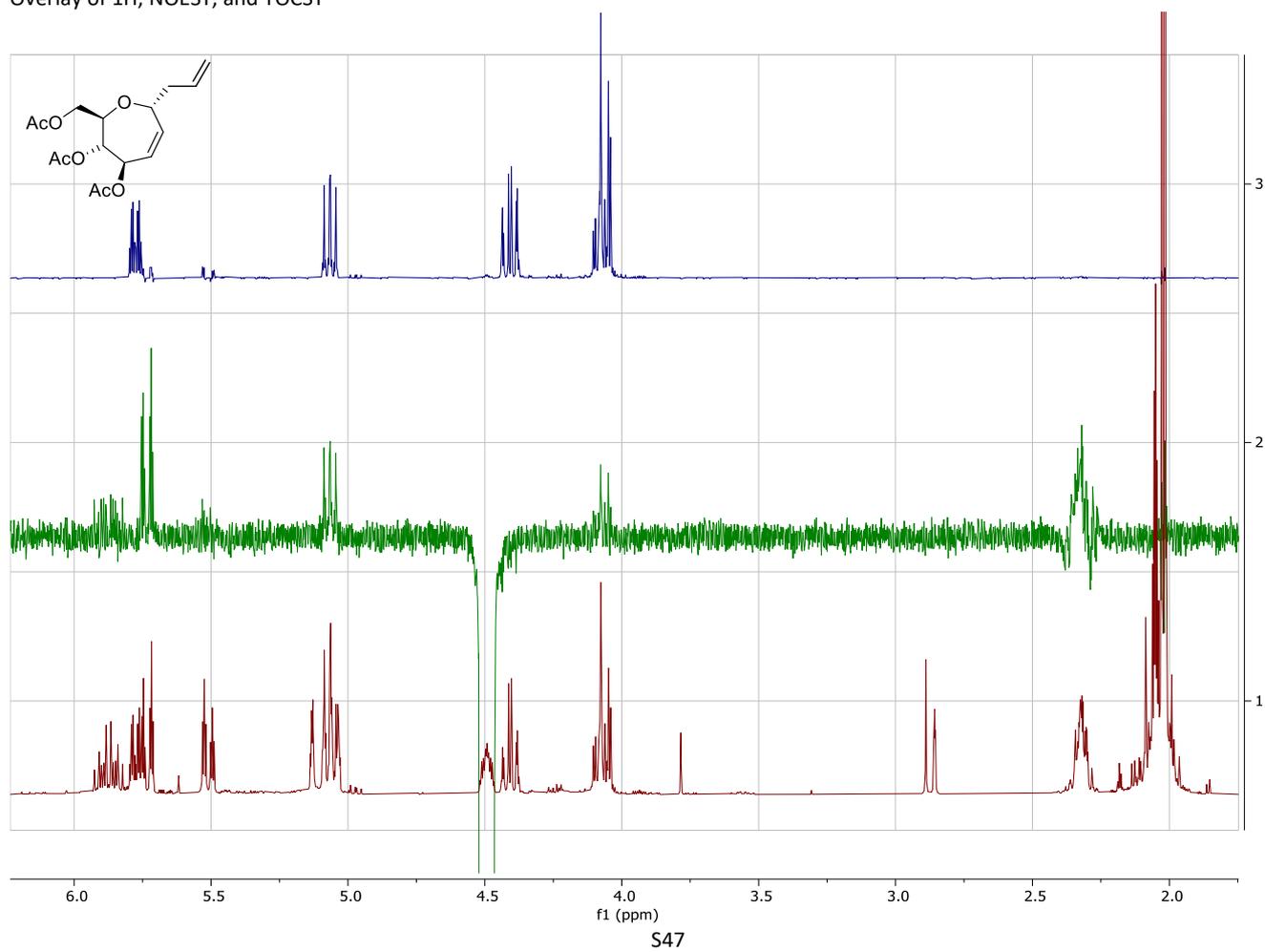
S45

1D TOCSY (400 MHz, d6-acetone)



S46

Overlay of 1H, NOESY, and TOCSY



Appendix E

The Key Intermediate of RNA Autohydrolysis

This appendix contains the supporting information of the publication related to Chapter 8. It contains voltage diagrams, mass spectra, infrared spectra, energetics, energy diagrams, and 3D-structures of *c*-fragments of RNA dinucleotides. The publication, the supporting information, and xyz-coordinates of the computed structures can be found online:

<https://doi.org/10.1002/anie.202115481>.^[268]

Supporting Information

Studying the Key Intermediate of RNA Autohydrolysis by Cryogenic Gas-Phase Infrared Spectroscopy

*K. Greis, C. Kirschbaum, M. I. Taccone, M. Götze, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, K. Pagel**

Table of Contents

Experimental Details	3
Materials	3
Mass Spectrometry and Infrared Spectroscopy.....	3
Mass Spectra	6
Computational Methods	8
Energetics.....	9
Computed Infrared Spectra.....	11
Low-Energy Structures	18
xyz-Coordinates of Reoptimized Structures	20
References	20

Experimental Details

Materials

Sodium salts of ApA, UpG and GpG RNA dinucleotides and cytidine-2',3'-cyclic monophosphate (2',3'-cCMP) were purchased from Jena Bioscience ($\geq 95\%$ purity). The sodium salt of the CpG RNA dinucleotide ($\geq 95\%$ purity) was purchased from TriLink Biotechnologies. Adenosine-2',3'-cyclic monophosphate (2',3'-cAMP, $\geq 93\%$ purity), adenosine 3',5'-cyclic monophosphate (3',5'-cAMP, $\geq 98.5\%$ purity) and cytidine-3',5'-cyclic monophosphate (3',5'-cCMP, $\geq 95\%$ purity) were obtained from Sigma-Aldrich.

Mass Spectrometry and Infrared Spectroscopy

ApA, GpG, UpG, CpG, 2',3'-cAMP, 3',5'-cAMP, 2',3'-cCMP and 3',5'-cCMP were dissolved in a mixture of acetonitrile and water (V:V. 1:1) to yield 250 μM solutions. The samples were ionized by nano-electrospray ionization (nESI) using Pd/Pt coated glass capillaries (Sputter Coater HR 208, *Cressington*). The capillaries were pulled to a tip with an inner diameter of 1–2 μm using a micropipette puller (Model P-1000, *Sutter Instrument*).

Mass and infrared (IR) spectra of the generated ions were recorded using a custom-built helium droplet instrument that is described in the following paragraphs (Figure S1). The RNA dinucleotide precursors were ionized via a Z-spray nESI source (tip voltage of 1 kV) and yielded autohydrolysis-like RNA fragments after in-source fragmentation in negative ion mode. Here, the generated ions collide with the residual gas in the source region, allowing for MS/MS. An exemplary potential diagram for activating in-source fragmentation conditions is shown in Figure S2a. The sodiated cAMP and cCMP precursors were measured at non-activating ionization conditions (Figure S2b).

After ionization, the ion beam is focused by two ring-electrode ion guides and the ions of interest are mass-to-charge selected by a quadrupole mass filter. Subsequently, the ions enter a quadrupole bender that either directs the ions to a time-of-flight detector to monitor the ion signal or to a hexapole ion trap, where the ions are thermalized by collisions with the helium buffer gas. In this experiment, the hexapole ion trap was cooled by liquid nitrogen to ca. 90 K.

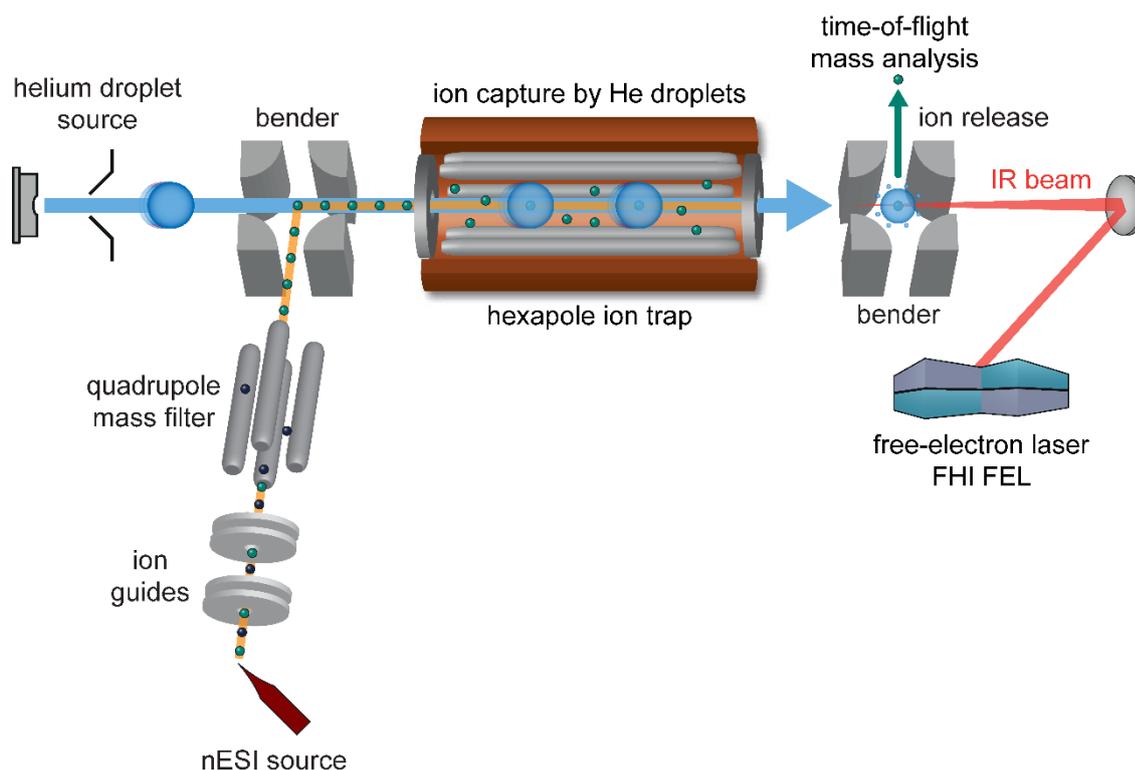


Figure S1. Schematic representation of the custom helium droplet instrument to record mass spectra and cryogenic gas-phase ion infrared spectra.

A beam of superfluid helium nanodroplets, formed by the expansion of pressurized helium through a pulsed Even-Lavie valve (nozzle temperature: 21 K), traverses the hexapole ion trap, where they pick up ions, rapidly cool them to 0.37 K and guide them to a detection region. Here, the beam of doped helium nanodroplets overlaps with an infrared beam produced by the Fritz Haber Institute free-electron laser (FHI-FEL^[1]), leading to the excitation of vibrational modes of the probed ions. The helium matrix acts as a cryostat with a constant temperature of 0.37 K that takes up the energy of the vibrationally excited ions, leading in turn to evaporation of the helium matrix. Subsequently, the ions are released from the matrix and detected by a time-of-flight detector. The ion signal as a function of the wavenumber of the IR laser leads to an IR spectrum. The IR signal scales non-linearly with the energy of the IR pulse, due to the absorption of multiple photons. As a first-order correction, the ion count (IR signal) is divided by the energy of the IR beam. Each spectrum is averaged from two measurements in the 1000–1800 cm^{-1} region.

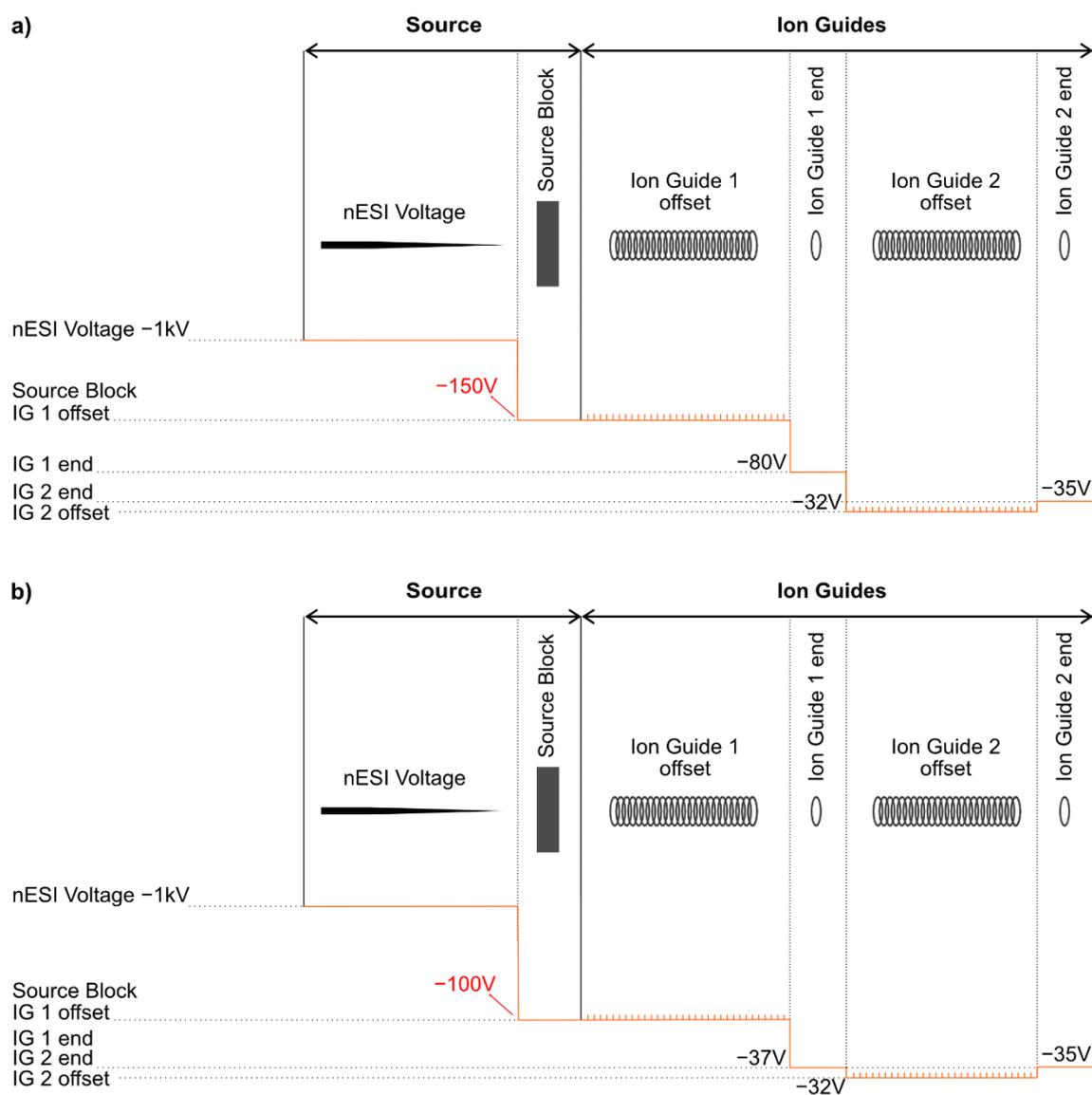


Figure S2. Schematic voltage diagram for the source region of the custom helium droplet instrument. Exemplary voltages for (a) activating in-source fragmentation (MS/MS) and for (b) non-activating conditions are shown. The diagram is not to scale.

Mass Spectra

Mass spectra of ApA, GpG, UpG and CpG RNA dinucleotides were recorded using non-activating (MS) and activating (MS/MS) in-source fragmentation settings in negative ion mode (Figure S3). With non-activating ionization conditions mainly $[M-H]^-$ ions are formed. Under activating in-source fragmentation conditions, $[M-H]^-$ ions yield a variety of less abundant fragment ions.

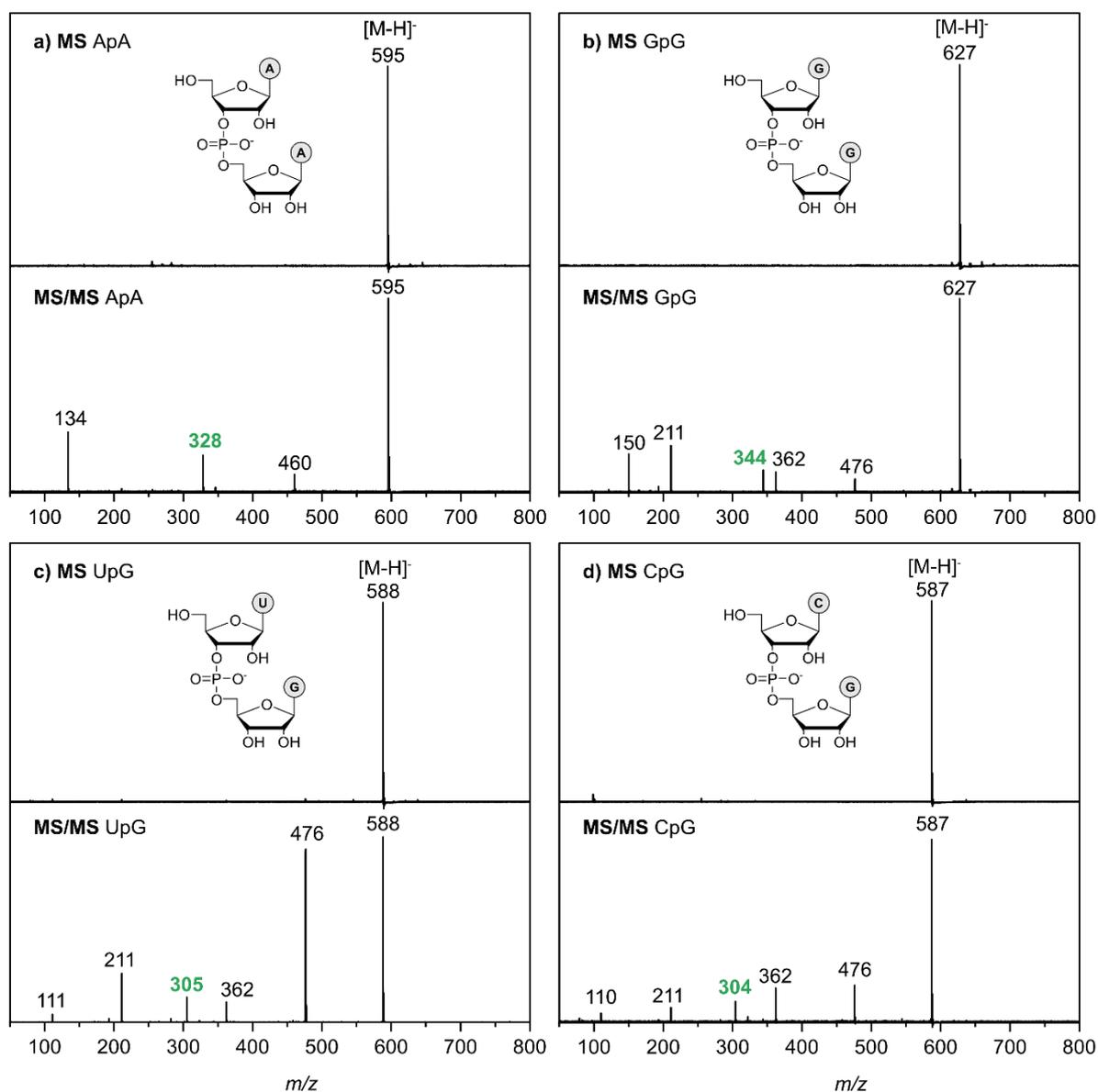


Figure S3. (a) Mass spectrum of the ApA RNA dinucleotide ($m/z = 595$) with (top) non-activating and (bottom) activating in-source fragmentation settings in negative ion mode, leading to autohydrolysis-like RNA ϵ -fragments ($m/z = 328$). The mass spectra of the (b) GpG ($m/z = 627$), the (c) UpG ($m/z = 588$) and (d) CpG RNA dinucleotides ($m/z = 587$) exhibit deprotonated intact ions and MS/MS leads to autohydrolysis-like RNA fragments as well, highlighted in green.

The ions of interest in this study, autohydrolysis-like RNA fragments, are formed for each dinucleotide ($m/z = 328, 344, 305, 303$), and correspond to ϵ -fragments in RNA (cf. Figure S4 for nomenclature of oligonucleotide fragmentation in tandem MS). The fragment ions at $m/z = 460$ and 476 correspond to the elimination product ions $[M-H-N_1]^-$, where the 5'-nucleobase is eliminated. The fragment ions at $m/z = 134, 150, 111, 110$ correspond to the negatively charged nucleobases $[N_1-H]^-$. For the fragment ions $m/z = 362$ and 211 , no corresponding fragments were observed for the ApA RNA dinucleotide. They correspond to w -fragments and anionic ribose phosphate.

For the homodinucleotides ApA and GpG RNA dinucleotide the formation of autohydrolysis-like RNA fragments proceeding via both ϵ - and α -fragments, leading to either 2',3'- or 3',5'-bridged cyclic nucleoside monophosphates with the same m/z , is conceivable. However, as α -fragments are not observed for the heterodinucleotides CpG and UpG, their formation can be ruled out for ApA and GpG.

The mass spectra of the cyclic nucleoside monophosphates 2',3'-cAMP, 3',5'-cAMP, 2',3'-cCMP and 3',5'-cCMP using non-activating ionization settings in negative ion mode are shown in Figure S5. Mainly $[M-H]^-$ and $[N-H]^-$ ions can be observed.

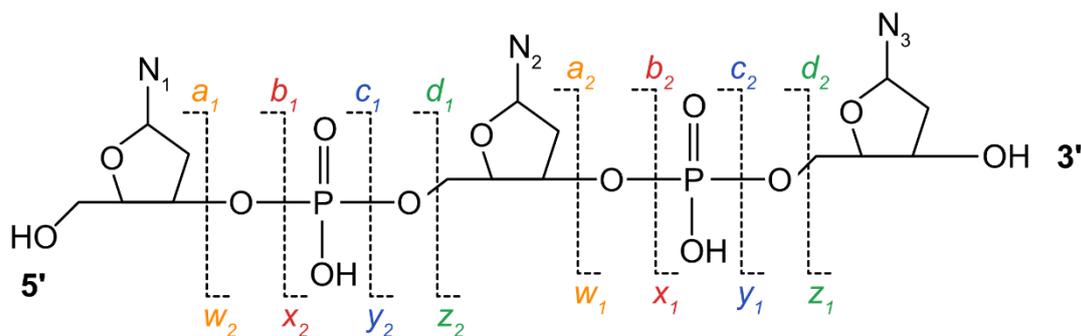


Figure S4. Nomenclature for fragmentation of DNA and RNA oligonucleotides in tandem mass spectrometry as introduced by McLuckey et al.^[2]

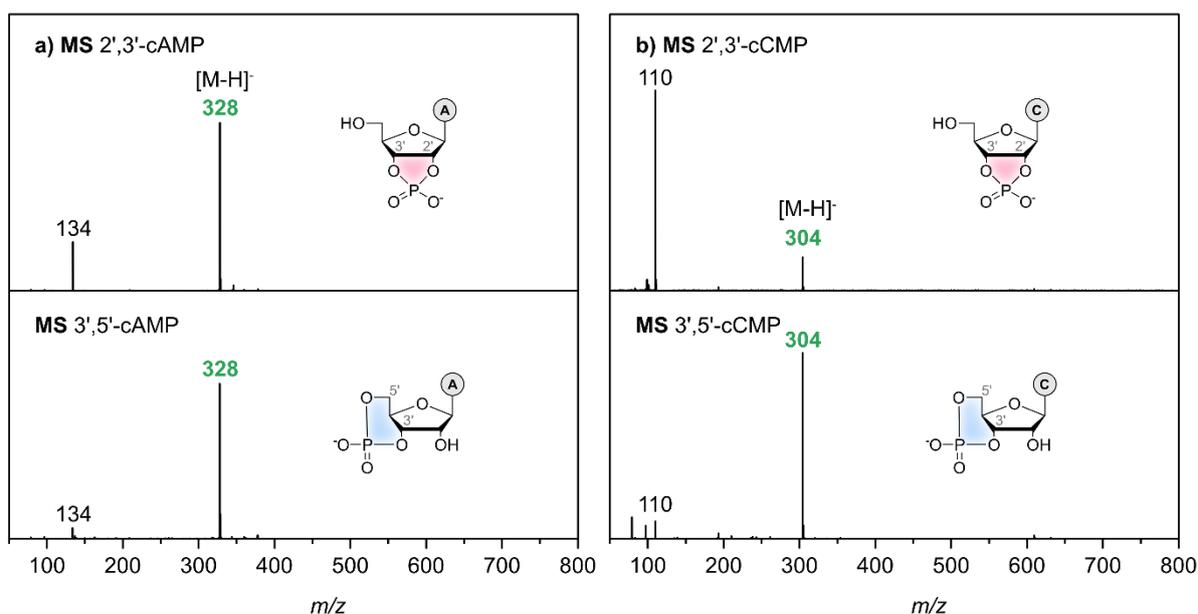


Figure S5. Mass spectra of (a) 2',3'-cAMP (top), 3',5'-cAMP (bottom), (b) 2',3'-cCMP (top), 3',5'-cCMP (bottom) using non-activating settings in negative ion mode. The most abundant ions are $[M-H]^-$ ($m/z = 328$ and 304) and $[N-H]^-$ ions ($m/z = 134$ and 110).

Computational Methods

Initial geometries of 2',3'-cNMP and 3',5'-cNMP anions (with N = A, G, U, C) were constructed using GaussView 6. Their conformational space was sampled using CREST^[3] with the semi-empirical method GFN2-xTB^[4] using default settings, except for the charge (-1). Subsequently, unique conformers below a threshold of 15 kJ mol⁻¹ (relative to the lowest-energy conformer) were selected using principal component analysis (PCA) of all bond lengths except hydrogen atoms using the module sklearn.decomposition.PCA and clustered by k-means clustering.^[5]

The unique conformers were reoptimized and their harmonic frequencies computed using Gaussian 16^[6] at the PBE0+D3/def2-TZVPP^[7] level of theory. Relative free energies at 90 K (approximate temperature of the ion trap) were extracted from the frequency calculation and are represented with the energy ΔE (including zero-point vibrational energy) in Tables S1-S4. Generally, 2',3'-cNMP anions with a five-membered cyclic phosphate moiety are energetically favored over their six-membered counterparts, 3',5'-cNMP anions, by 5-27 kJ mol⁻¹ depending on the nucleobase.

Energetics

Table S1. List of conformations of five- (A5) and six-membered (A6) 2',3'-cAMP and 3',5'-cAMP anions optimized at PBE0+D3/def2-TZVPP level of theory. Relative energies (ΔE , including zero-point vibrational energy) and free energies (ΔF) at 90 K are assigned to each conformer. The spectra of the structures highlighted with a dagger (†) are shown in the manuscript or SI.

ID	$\Delta E(\text{PBE0+D3})$ [kJ mol ⁻¹]	$\Delta F(\text{PBE0+D3})$ [kJ mol ⁻¹]
A5/conf_0 (†)	0.00	0.00
A5/conf_1	19.82	18.37
A5/conf_3	15.43	14.07
A5/conf_4	14.92	14.50
A5/conf_5	26.00	24.12
A6/conf_0 (†)	27.77	26.89
A6/conf_1	41.30	40.37
A6/conf_2	37.23	35.84
A6/conf_3	47.30	47.61
A6/conf_4	44.82	44.23

Table S2. List of conformations of five- (G5) and six-membered (G6) 2',3'-cGMP and 3',5'-cGMP anions optimized at PBE0+D3/def2-TZVPP level of theory. Relative energies (ΔE , including zero-point vibrational energy) and free energies (ΔF) at 90 K are assigned to each conformer. The spectra of the structures highlighted with a dagger (†) are shown in the manuscript or SI.

ID	$\Delta E(\text{PBE0+D3})$ [kJ mol ⁻¹]	$\Delta F(\text{PBE0+D3})$ [kJ mol ⁻¹]
G5/conf_0 (†)	0.00	0.00
G5/conf_1	17.43	16.91
G5/conf_2	25.55	23.05
G5/conf_3	1.15	0.98
G6/conf_0 (†)	3.43	4.90

Table S3. List of conformations of five- (U5) and six-membered (U6) 2',3'-cUMP and 3',5'-cUMP anions optimized at PBE0+D3/def2-TZVPP level of theory. Relative energies (ΔE , including zero-point vibrational energy) and free energies (ΔF) at 90 K are assigned to each conformer. The spectra of the structures highlighted with a dagger (\dagger) are shown in the manuscript or SI.

ID	$\Delta E(\text{PBE0+D3})$ [kJ mol ⁻¹]	$\Delta F(\text{PBE0+D3})$ [kJ mol ⁻¹]
U5/conf_0 (\dagger)	0.00	0.00
U5/conf_1	16.49	15.05
U5/conf_3	13.07	11.93
U5/conf_4	18.77	17.11
U6/conf_0 (\dagger)	19.66	18.51
U6/conf_1	29.81	28.75
U6/conf_2	35.27	34.93
U6/conf_3	47.76	48.39

Table S4. List of conformations of five- (C5) and six-membered (C6) 2',3'-cCMP and 3',5'-cCMP anions optimized at PBE0+D3/def2-TZVPP level of theory. Relative energies (ΔE , including zero-point vibrational energy) and free energies (ΔF) at 90 K are assigned to each conformer. The spectra of the structures highlighted with a dagger (\dagger) are shown in the manuscript or SI.

ID	$\Delta E(\text{PBE0+D3})$ [kJ mol ⁻¹]	$\Delta F(\text{PBE0+D3})$ [kJ mol ⁻¹]
C5/conf_0 (\dagger)	0.00	0.00
C5/conf_1	12.51	12.32
C5/conf_3	20.73	19.33
C5/conf_5	25.49	23.33
C6/conf_0 (\dagger)	17.01	16.97
C6/conf_1	28.33	27.93
C6/conf_2	35.43	35.60

Computed Infrared Spectra

The computed harmonic infrared spectra of the lowest-energy structures of 2',3'-cNMP and 3',5'-cNMP anions are shown along with the experimental spectra of the respective nucleotide ϵ -fragments in Figure S6. All shown harmonic frequencies are scaled by an empirical factor of 0.965. Anharmonic frequencies were calculated using the GVPT2 method^[8] at PBE0+D3/def2-TZVP level of theory. The size of the basis set was slightly reduced, as anharmonic frequency calculations are computationally expensive. Switching between the def2-TZVP and the def2-TZVPP basis sets did not alter the computed harmonic infrared signature (Figure S7).

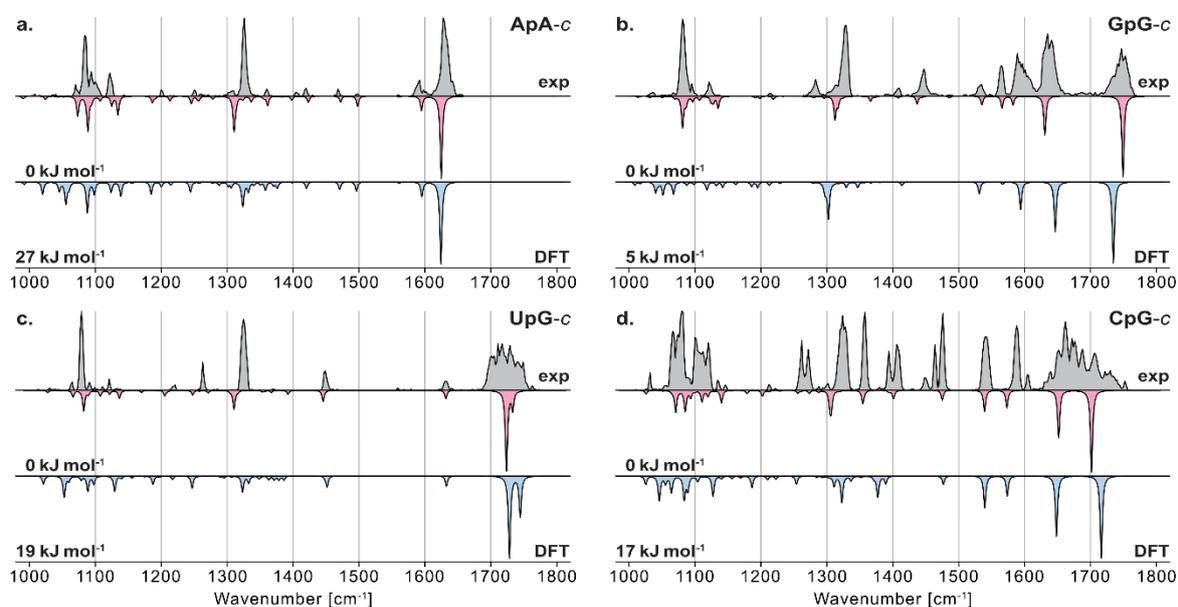


Figure S6. Harmonic infrared spectra of 2',3'-cNMP (red trace) and 3',5'-cNMP anions (blue trace) for N = (a) A, (b) G, (c) U, and (d) C, along with experimental spectra (gray trace) of anionic nucleobase monophosphate ϵ -fragments generated from ApA, GpG, UpG, and CpG RNA dinucleotides respectively. Harmonic spectra are computed at the PBE0+D3/def2-TZVPP level of theory and scaled by an empirical factor of 0.965.

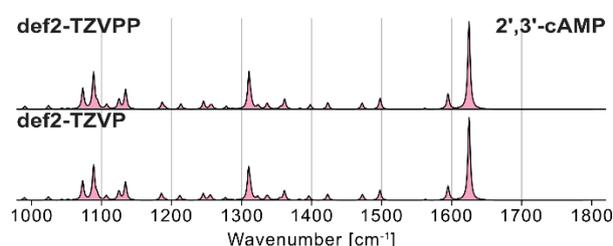


Figure S7. Harmonic infrared spectra of 2',3'-cAMP anion at (top) PBE0+D3/def2-TZVPP and (bottom) PBE0+D3/def2-TZVP levels of theory. The spectra are scaled by an empirical factor of 0.965. The comparison shows that the spectra look almost identical at both levels of theory.

The computed anharmonic infrared spectra of the lowest-energy structures of 2',3'-cNMP and 3',5'-cNMP anions are shown along with the experimental spectra of the respective nucleotide fragments in Figure S8. The unscaled anharmonic spectra are blueshifted by ca. 20 cm⁻¹ compared to the experiment. Interestingly, the absorption band originating from the antisymmetric phosphate stretch (ca. 1326 cm⁻¹) that is comparably intense in the experimental spectrum and the computed harmonic frequencies decreases substantially in intensity in the anharmonic spectra. For 2',3'-cUMP it appears to vanish completely. The depletion of the intense absorption band originating from the antisymmetric phosphate stretch is due to Darling-Dennison resonances. If these are deactivated in Gaussian 16, the antisymmetric phosphate stretch becomes the most intense absorption band in the spectral region (Figure S9). Furthermore, in the anharmonic spectra it becomes even more apparent that for 3',5'-cNMP anions intense absorption bands should be visible around 1020-1060 cm⁻¹ that are missing in the experimental spectra of the *c*-fragments. In Figure S10, the experimental spectra of 2',3'- and 3',5'-cAMP anions are compared to harmonic and anharmonic frequencies of their computed counterparts.

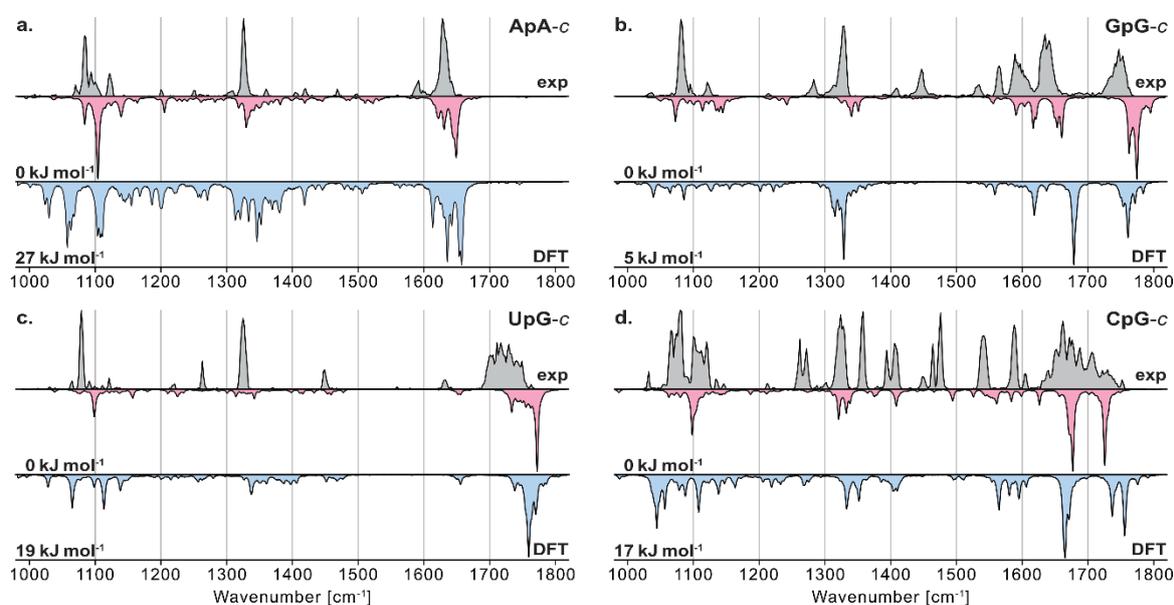


Figure S8. Anharmonic infrared spectra of 2',3'-cNMP (red trace) and 3',5'-cNMP anions (blue trace) for N = (a) A, (b) G, (c) U, and (d) C, along with experimental spectra (gray trace) of anionic nucleobase monophosphate fragments generated from ApA, GpG, UpG, and CpG RNA dinucleotides respectively. The anharmonic spectra are computed using the GVPT2 method at the PBE0+D3/def2-TZVP level of theory.

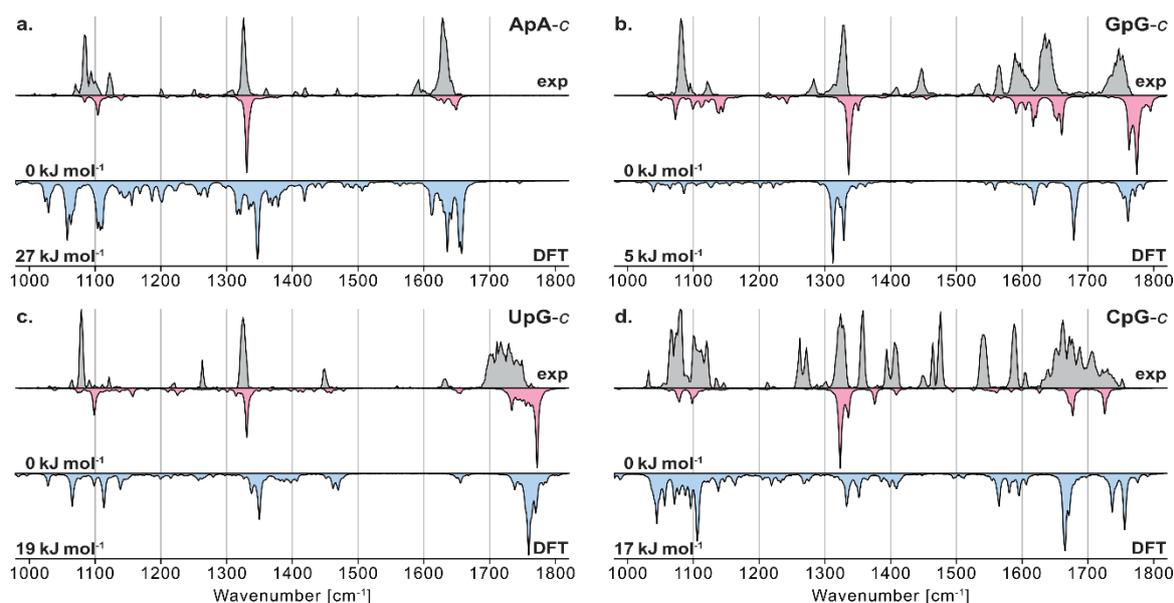


Figure S9. Anharmonic infrared spectra of 2',3'-cNMP (red trace) and 3',5'-cNMP anions (blue trace) for N = (a) A, (b) G, (c) U, and (d) C, along with experimental spectra (gray trace) of anionic nucleobase monophosphate fragments generated from ApA, GpG, UpG, and CpG RNA dinucleotides respectively. The anharmonic spectra are computed using the GVPT2 method at the PBE0+D3/def2-TZVP level of theory, with Darling-Dennison resonances deactivated.

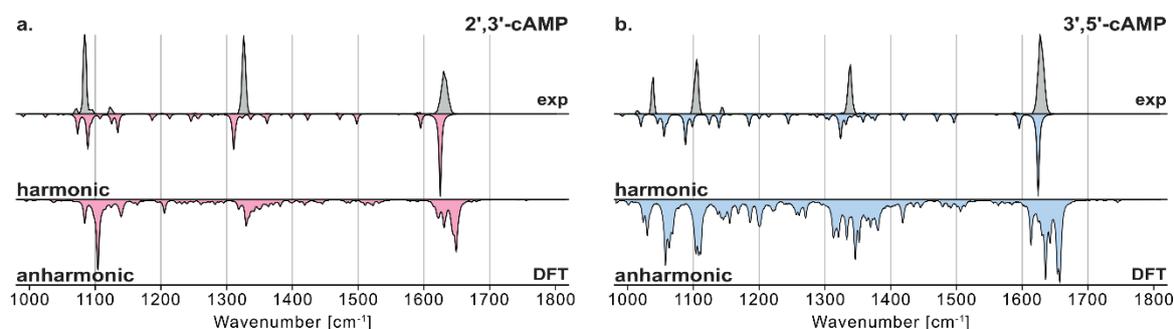


Figure S10. Experimental and computed infrared spectra of (a) 2',3'-cAMP and (b) 3',5'-cAMP anions generated using non-activating ionization settings. The harmonic spectra are computed at the PBE0+D3/def2-TZVPP level of theory and scaled by an empirical factor of 0.965 and anharmonic spectra are computed using the GVPT2 method at the PBE0+D3/def2-TZVP level of theory.

In Figure S11, the experimental spectra of 2',3'- and 3',5'-cCMP anions are compared to harmonic and anharmonic frequencies of their computed counterparts. The conformational space of the imino-oxo tautomers of 2',3'- and 3',5'-cCMPs was sampled, geometries reoptimized and frequencies computed according to the previously described methodology (Table S5). This table also includes the energetics of an intramolecular transition state that is connecting the tautomerized and non-tautomerized 2',3'-cCMP. In a second transition state, the tautomerization is catalyzed by a water molecule, which is lowering the activation barrier for tautomerization, as shown in the energy diagram in Figure S12. The experimentally resolved IR spectra of 2',3'- and 3',5'-cCMPs anions generated with activating ionization settings are compared to harmonic and anharmonic frequencies of their tautomerized and non-tautomerized computed counterparts (Figures S13 and S14).

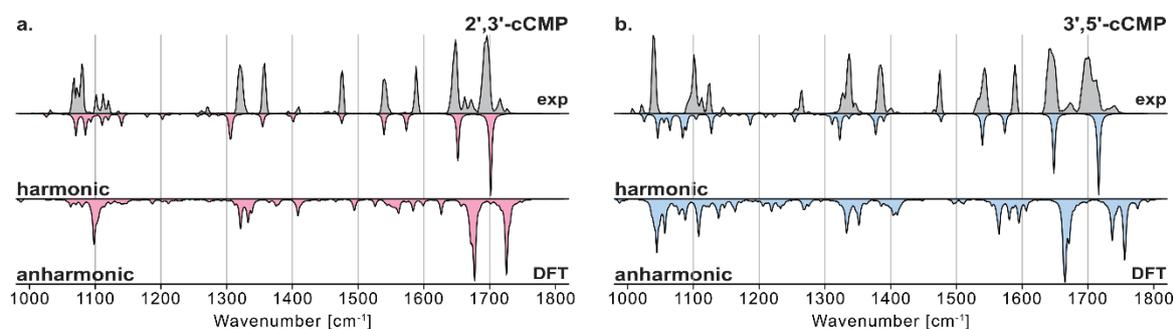


Figure S11. Experimental and computed infrared spectra of (a) 2',3'-cCMP and (b) 3',5'-cCMP anions generated using non-activating ionization settings. The harmonic spectra are computed at the PBE0+D3/def2-TZVPP level of theory and scaled by an empirical factor of 0.965 and anharmonic spectra are computed using the GVPT2 method at the PBE0+D3/def2-TZVP level of theory.

Table S5. List of conformations of tautomerized five- (C5*) and six-membered (C6*) 2',3'-cCMP and 3',5'-cCMP anions optimized at PBE0+D3/def2-TZVPP level of theory. Energies (ΔE , including zero-point vibrational energy) and free energies (ΔF) at 90 K are assigned to each conformer, relative to the non-tautomerized structure C5/conf_00. The spectra of the structures highlighted with a dagger (\dagger) are shown in the manuscript or SI. Furthermore, the energetics of a transition state (TS) involved in intramolecular tautomerization are shown. The energetics of another TS, where the tautomerization is catalyzed by a water molecule, along the reactant and product are included as well.

ID	$\Delta E(\text{PBE0+D3})$ [kJ mol ⁻¹]	$\Delta F(\text{PBE0+D3})$ [kJ mol ⁻¹]
C5/conf_0 (\dagger)	0.00	0.00
C6/conf_0 (\dagger)	17.01	16.97
C5*/conf_0 (\dagger)	2.89	3.15
C5*/conf_1	19.69	18.53
C5*/conf_3	15.90	15.10
C5*/conf_4	21.94	20.56
C5*/conf_5	17.61	16.75
C6*/conf_0 (\dagger)	20.97	20.42
C6*/conf_1	31.77	30.85
C6*/conf_2	37.53	37.72
C5_TS/conf_0	156.29	156.53
C5_H ₂ O/conf_0 ^a	0.00	0.00
C5*_H ₂ O/conf_0 ^a	10.29	10.61
C5_TS_H ₂ O/conf_0 ^a	47.13	49.84

^a The energetics of these structures are relative to C5_H₂O/conf_0

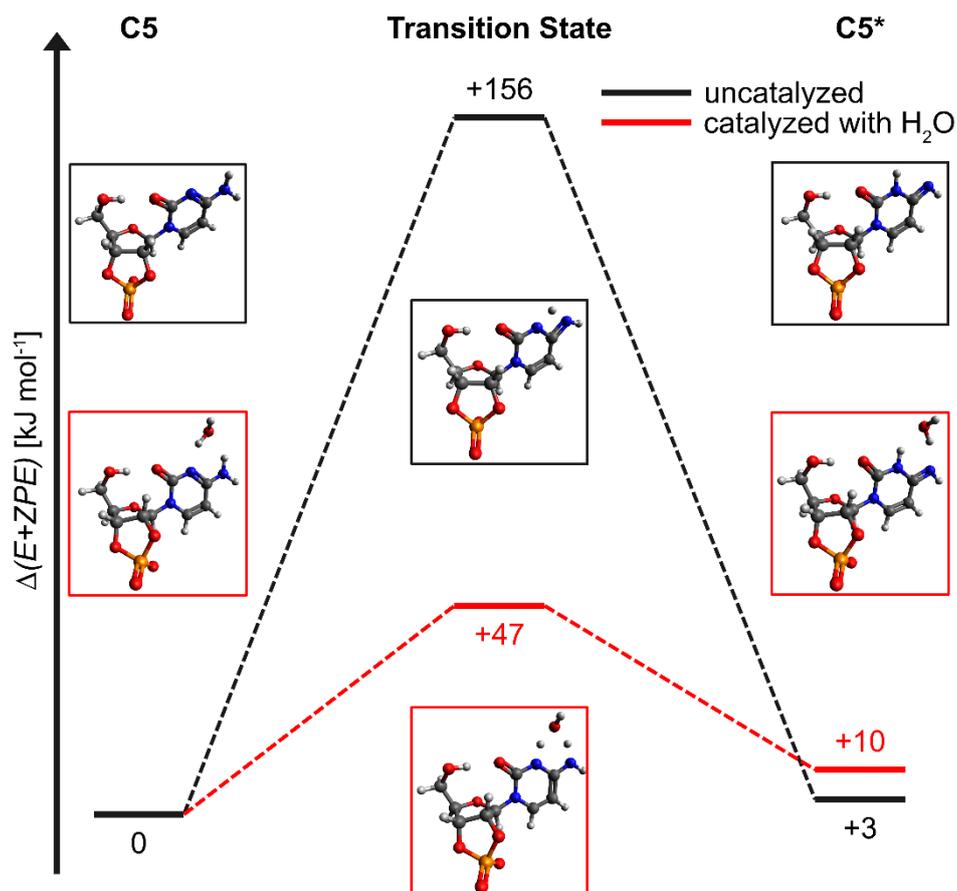


Figure S12. Energy diagram representing the reaction coordinate for tautomerization of 2',3'-cCMP anions from the amino-oxo (C5) to the imino-oxo (C5*) form. The energy of the transition state is lowered from +156 kJ mol^{-1} (black, uncatalyzed) to +47 kJ mol^{-1} (red) when catalyzed by a water molecule.

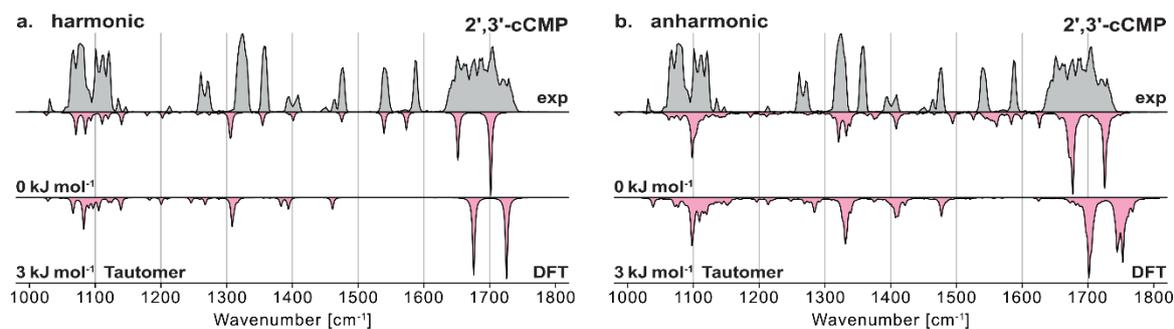


Figure S13. The experimental spectrum of the 2',3'-cCMP anion generated using activating ionization settings is shown in the gray traces. Computed (a) harmonic and (b) anharmonic infrared signatures of the 2',3'-cCMP anion and its tautomer are shown in the inverted red traces.

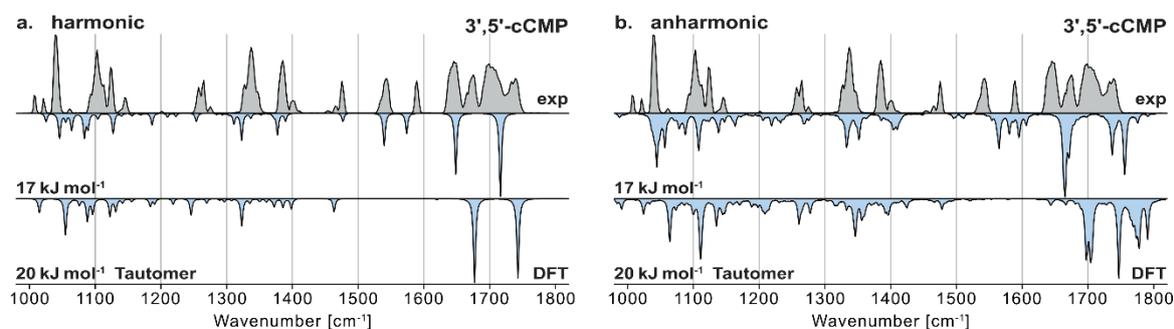


Figure S14. The experimental spectrum of the 3',5'-cCMP anion generated using activating ionization settings is shown in the gray traces. Computed (a) harmonic and (b) anharmonic infrared signatures of the 3',5'-cCMP anion and its tautomer are shown in the inverted blue traces.

Low-Energy Structures

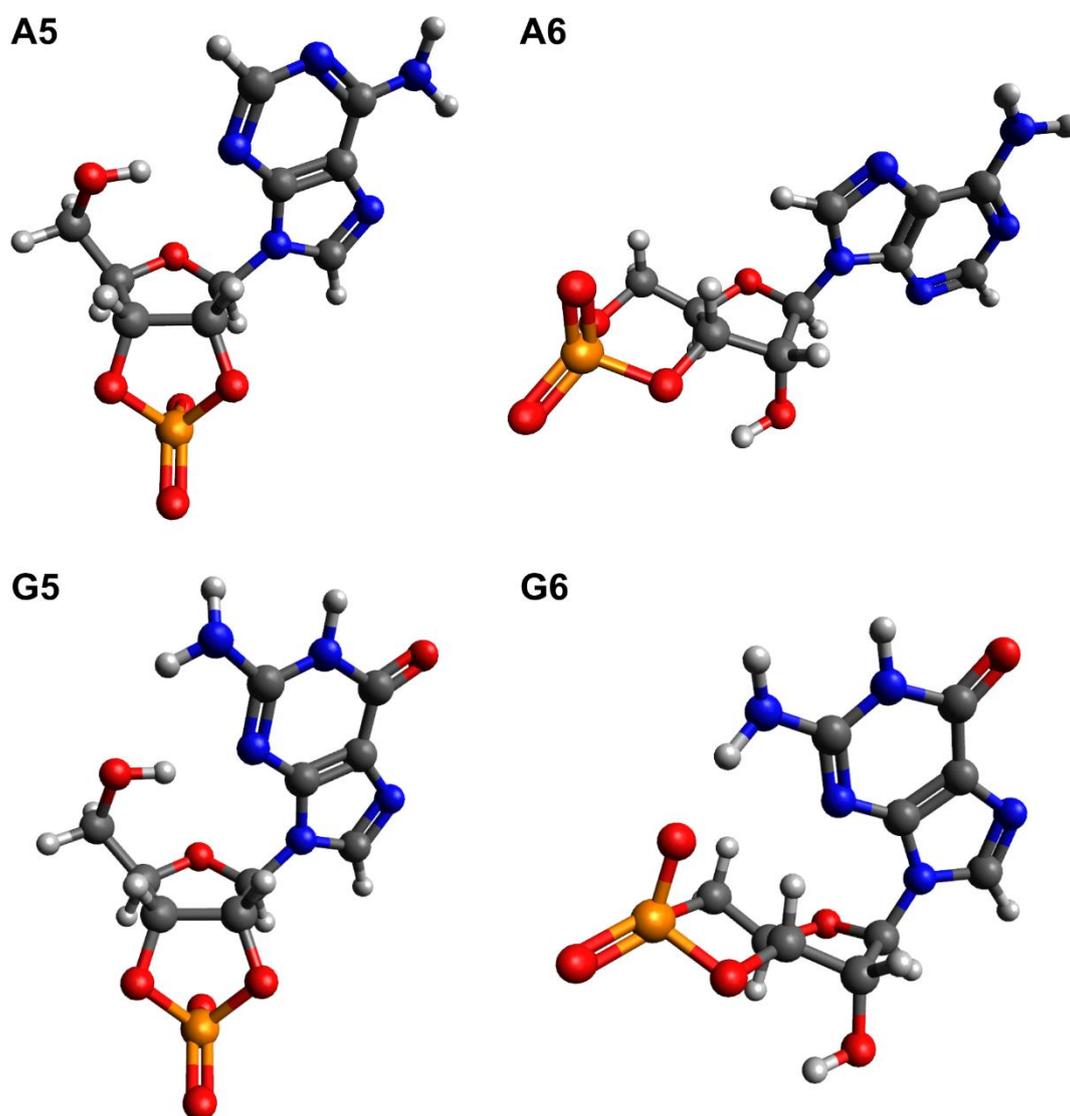


Figure S15. Reoptimized geometries of 2',3'-cAMP (A5), 3',5'-cAMP (A6), 2',3'-cGMP (G5) and 3',5'-cGMP (G6) anions. The structures correspond to those that are highlighted with a dagger (†) in Tables S1-S2.

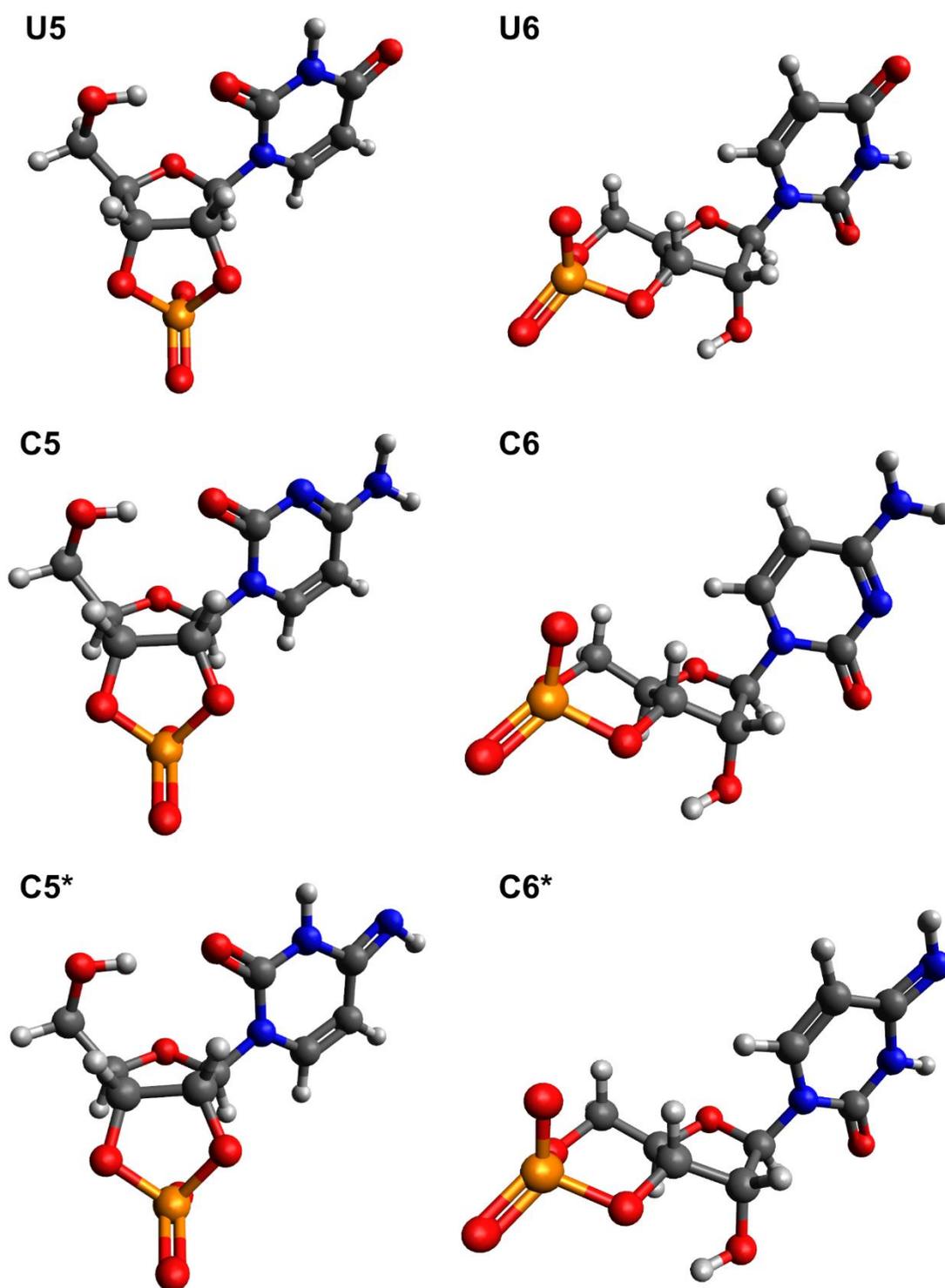


Figure S16. Reoptimized geometries of 2',3'-cUMP (U5), 3',5'-cUMP (U6), 2',3'-cCMP (C5) and 3',5'-cCMP (C6) anions. C5* and C6* correspond to C5 and C6, respectively, with a tautomerized nucleobase. The structures correspond to those that are highlighted with a dagger (†) in Tables S3-S5.

xyz-Coordinates of Reoptimized Structures

xyz-Coordinates of all reoptimized geometries at the PBE0+D3/def2-TZVPP level of theory can be found in a separate document “coordinates.xyz”.

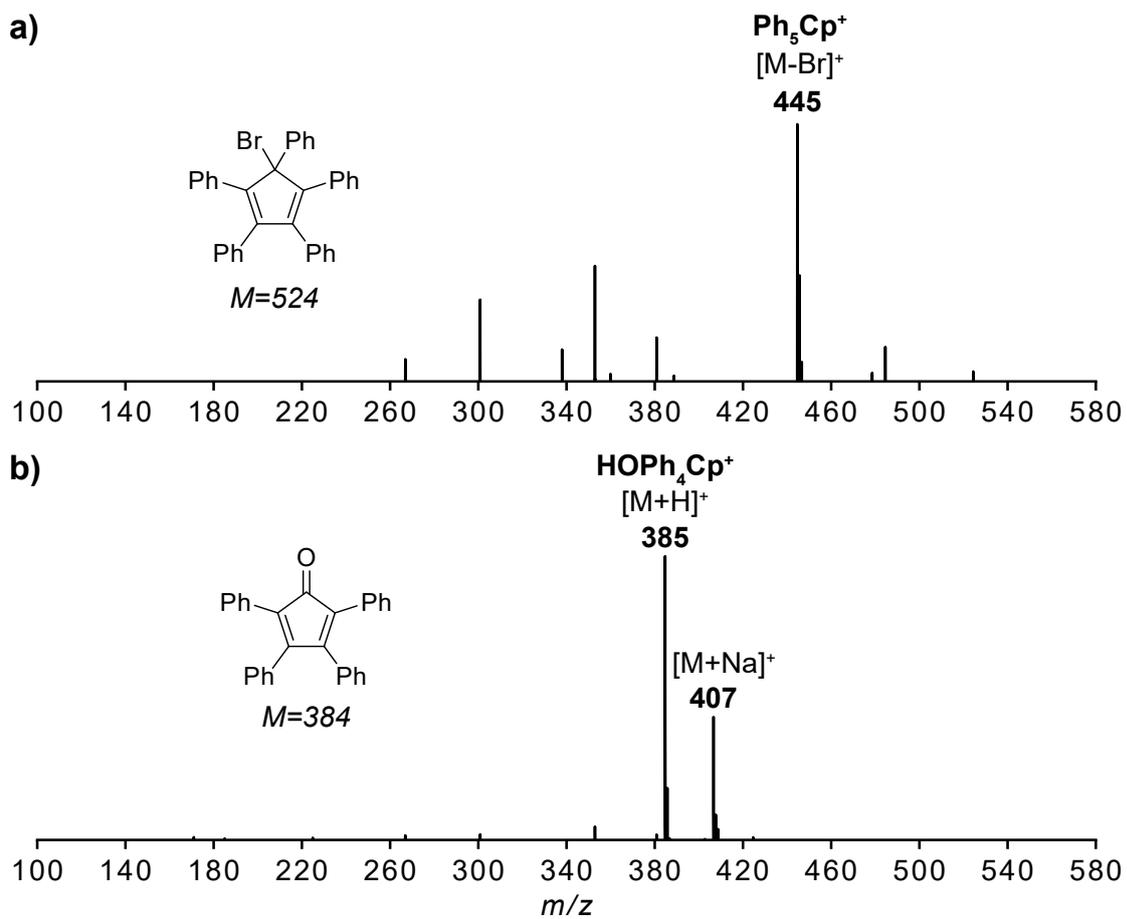
References

- [1] W. Schöllkopf, S. Gewinner, H. Junkes, A. Paarmann, G. von Helden, H. P. Bluem, A. M. M. Todd, *Proc. SPIE Int. Soc. Opt. Eng.* **2015**, 95121L.
- [2] S. A. McLuckey, G. J. Berkel, G. L. Glish, *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60-70.
- [3] P. Pracht, F. Bohle, S. Grimme, *Phys. Chem. Chem. Phys.* **2020**, *22*, 7169-7192.
- [4] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652-1671.
- [5] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, *J. Mach. Learn. Res.* **2011**, *12*, 2825-2830.
- [6] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Wallingford, CT, **2016**.
- [7] a) C. Adamo, V. Barone, *J. Chem. Phys.* **1999**, *110*, 6158-6170; b) S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104; c) F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297-3305.
- [8] a) V. Barone, *J. Chem. Phys.* **2005**, *122*, 14108; b) V. Barone, J. Bloino, C. A. Guido, F. Lipparini, *Chem. Phys. Lett.* **2010**, *496*, 157-161; c) J. Bloino, V. Barone, *J. Chem. Phys.* **2012**, *136*, 124108.

Appendix F

Cyclopentadienyl and Fluorenyl Cations

This appendix contains mass spectra, infrared spectra, energetics, tables with exact positions of experimental and computed infrared absorption bands, and 3D-structures of cyclopentadienyl and fluorenyl cations described in Chapter 9.



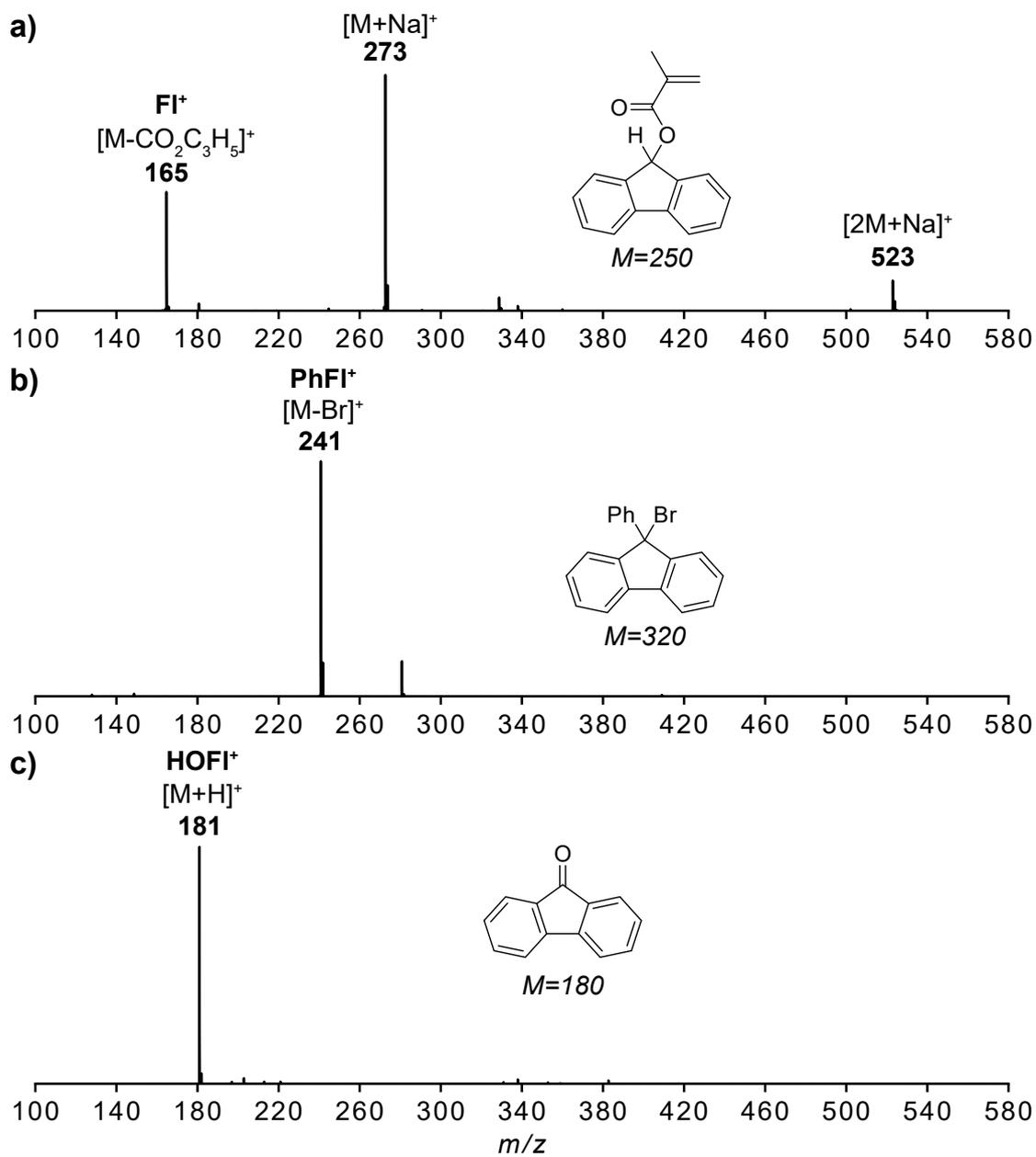


Table F1: Relative energies of the singlet and triplet states of the PhCp_5^+ , HOPh_4Cp^+ , FI^+ , PhFI^+ , and HOFI^+ cations in kJ mol^{-1} . The electronic energies are indicated as single-point energies of optimized structures at the respective level of theory (except for MP2). All computations were carried out using the D3 dispersion correction and the def2-TZVPP basis set. MP2 single-point energies were computed using the structure optimized at CAM-B3LYP. For the triplet state ROMP2 was used. The absolute ZPVE computed at CAM-B3LYP is shown as well.

Ion	Method	Singlet (dienyl)	Singlet (allyl)	Triplet
PhCp_5^+	PBE0	0.0	+0.8	-17.6
	B3LYP	0.0	+0.9	-12.0
	CAM-B3LYP	0.0	+0.1	-13.8
	MP2	0.0	-0.4	-15.2
	ZPVE	1293.0	1293.0	1293.1
HOPh_4Cp^+	PBE0	0.0	-	+6.6
	B3LYP	0.0	-	+8.9
	CAM-B3LYP	0.0	-	+12.7
	ZPVE	1092.9	-	1091.7
FI^+	PBE0	0.0	-	+72.3
	B3LYP	0.0	-	+75.3
	CAM-B3LYP	0.0	-	+77.4
	ZPVE	469.5	-	465.3
PhFI^+	PBE0	0.0	-	+99.1
	B3LYP	0.0	-	+101.1
	CAM-B3LYP	0.0	-	+103.3
	ZPVE	686.8	-	682.2
HOFI^+	PBE0	0.0	-	+126.6
	B3LYP	0.0	-	+125.8
	CAM-B3LYP	0.0	-	+133.7
	ZPVE	484.4	-	478.4

Table F2: Position of the experimental infrared absorption bands for the pentaphenylcyclopentadienyl cation Ph_5Cp^+ (C_2 symmetry) and comparison to the position of scaled harmonic frequencies of the singlet state computed at CAM-B3LYP+D3/def2-TZVPP (scaling factor: 0.965). Only the bands in the experimentally probed region (600–1700 cm^{-1}) are indicated. All frequencies are given in cm^{-1} and computed intensities (I) in km mol^{-1} . Computed absorptions are indicated for the singlet state if its intensity is higher than 10 km mol^{-1} .

$\tilde{\nu}_{\text{exp}}$	$\tilde{\nu}_{\text{calc}}$ (I)	Symmetry	$\tilde{\nu}_{\text{exp}}$	$\tilde{\nu}_{\text{calc}}$ (I)	Symmetry
618	607 (28)	B	1189	1181 (193)	A
625	629 (98)	B	–	1188 (92)	A
674	680 (29)	B	–	1246 (233)	A
–	692 (17)	B	–	1268 (75)	A
700	702 (46)	B	–	1281 (17)	A
709	710 (37)	A	1304	1305 (1012)	A
725	714 (54)	B	1310	–	–
732	746 (112)	B	1321	–	–
755	766 (579)	B	1330	1334 (215)	A
789	788 (24)	B	–	1334 (14)	B
–	796 (46)	A	1343	1348 (132)	B
–	807 (16)	A	1382	–	–
823	818 (19)	B	1408	1407 (625)	A
834	845 (78)	A	1438	1443 (88)	B
844	851 (22)	A	–	1447 (13)	B
861	–	–	–	1451 (12)	B
875	–	–	1469	–	–
885	–	–	1481	1486 (242)	A
983	–	–	–	1492 (15)	B
–	993 (35)	A	1495	1496 (134)	A
993	993 (67)	A	–	1501 (48)	B
–	997 (27)	A	–	1508 (112)	A
–	1003 (13)	A	–	1567 (34)	B
–	1027 (20)	B	–	1580 (183)	A
1082	1066 (179)	B	1587	1589 (495)	A
–	1082 (17)	B	–	1593 (31)	B
–	1091 (12)	B	–	1599 (170)	A
1118	1106 (55)	B	–	1601 (26)	B
–	1107 (28)	A	–	1611 (187)	A
–	1154 (11)	A	–	1615 (34)	B
–	1160 (10)	B	–	1629 (26)	A

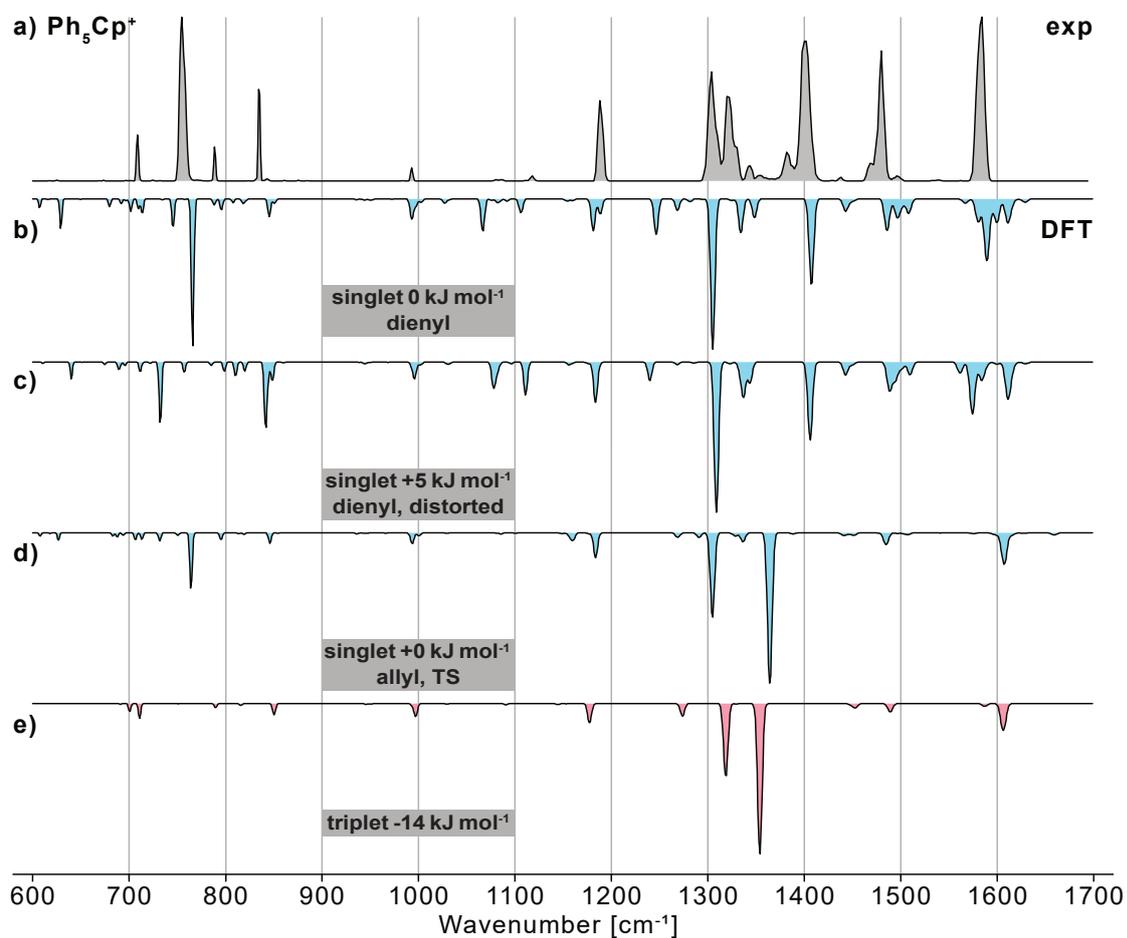


Figure F3: (a) The experimental cryogenic IR spectrum of the pentaphenylcyclopentadienyl cation Ph_5Cp^+ is compared to computed scaled harmonic frequencies of the (b) ground-state and (c) a distorted singlet state geometry, (d) an allyl singlet state geometry, and (e) the triplet at the CAM-B3LYP+D3/def2-TZVPP level of theory. The energies including ZPVE are indicated. The distorted structure displayed in (c) matches the experiment less well and is higher in energy and was therefore not further considered.

Table F3: Position of the experimental infrared absorption bands for the hydroxytetraphenylcyclopentadienyl cation HOPh_4Cp^+ (C_1 symmetry) and comparison to the position of scaled harmonic frequencies of the singlet state computed at CAM-B3LYP+D3/def2-TZVPP (scaling factor: 0.965). Only the bands in the experimentally probed region ($600\text{--}1700\text{ cm}^{-1}$) are indicated. All frequencies are given in cm^{-1} and computed intensities (I) in km mol^{-1} . Computed absorptions are indicated for the singlet state if its intensity is higher than 10 km mol^{-1} .

$\tilde{\nu}_{\text{exp}}$	$\tilde{\nu}_{\text{calc}}$ (I)	$\tilde{\nu}_{\text{exp}}$	$\tilde{\nu}_{\text{calc}}$ (I)
632	638 (34)	1110	1103 (70)
–	639 (21)	–	1182 (27)
685	–	–	1184 (13)
692	693 (21)	–	1272 (79)
697	701 (43)	1304	1294 (242)
703	707 (53)	1324	1306 (500)
712	711 (81)	–	1327 (10)
751	761 (14)	–	1336 (108)
755	762 (28)	1384	1372 (409)
797	805 (83)	1436	1447 (29)
–	809 (10)	–	1448 (17)
829	840 (14)	–	1452 (13)
838	851 (35)	1478	1485 (156)
–	852 (12)	1485	1497 (59)
860	856 (147)	–	1503 (31)
–	979 (22)	1519, 1527	1538 (698)
–	999 (18)	1589	1609 (110)
–	1065 (17)	–	1614 (49)
–	1082 (13)	–	1618 (23)
–	1088 (39)	–	1628 (20)
–	1099 (39)	–	–

Table F4: Position of the experimental infrared absorption bands for the 9-fluorenyl cation FI^+ (C_{2v} symmetry) and comparison to data from matrix isolation spectroscopy (measured in low-density amorphous (LDA) ice),^[343] computed scaled harmonic (using the hybrid DFT functionals PBE0, B3LYP, CAM-B3LYP with the def2-TZVPP basis set and a scaling factor of 0.965), and unscaled anharmonic frequencies (using the GVPT2 method at the CAM-B3LYP/def2-TZVPP level of theory) of the singlet state. Only the bands in the experimentally probed region (600–1700 cm^{-1}) are indicated. All frequencies are given in cm^{-1} and computed intensities (I) in km mol^{-1} . Computed absorptions are indicated for the singlet state if its intensity is higher than 10 km mol^{-1} .

$\tilde{\nu}_{\text{exp}}$	$\tilde{\nu}_{\text{LDA}}$	$\tilde{\nu}_{\text{PBE0}} (I)$	$\tilde{\nu}_{\text{B3LYP}} (I)$	$\tilde{\nu}_{\text{CAM-B3LYP}} (I)$	$\tilde{\nu}_{\text{GVPT2}} (I)$	Symmetry
700	–	703 (55)	702 (54)	714 (55)	688 (32)	B1
770	–	770 (59)	769 (58)	781 (63)	786 (77)	B1
985	986	984 (63)	979 (70)	991 (99)	1013 (55)	B2
–	–	988 (10)	–	–	–	B1
1009	–	–	1003 (16)	1009 (70)	1031 (57)	B2
1072	1077	1072 (230)	1067 (210)	1065 (404)	1088 (199)	B2
1115	1117	1117 (48)	1113 (55)	1112 (14)	–	B2
1166	1164	1153 (68)	1155 (77)	1160 (115)	1187 (58)	B2
1239	1235	1234 (84)	1223 (67)	1233 (122)	1254 (67)	B2
1267	–	1284 (17)	1270 (24)	1262 (60)	1285 (22)	B2
–	–	1286 (16)	1265 (13)	1279 (11)	–	A1
–	–	–	–	1294 (10)	1321 (11)	A1
1343	1344	1330 (141)	1327 (98)	1336 (203)	1363 (161)	B2
–	–	1399 (11)	1386 (13)	1400 (18)	1424 (22)	A1
–	–	1428 (10)	1424 (18)	–	–	B2
–	–	1446 (10)	–	–	–	A1
1473	1469	1468 (76)	1460 (55)	1480 (105)	1501 (55)	B2
1500	1491	1503 (92)	1478 (90)	1504 (108)	1521 (53)	A1
1572	1576	1585 (468)	1562 (393)	1591 (618)	1613 (215)	B2
1583	–	1607 (127)	1588 (161)	1621 (99)	1641 (80)	B2
1599	1601	1618 (37)	1593 (35)	1632 (42)	1652 (32)	A1

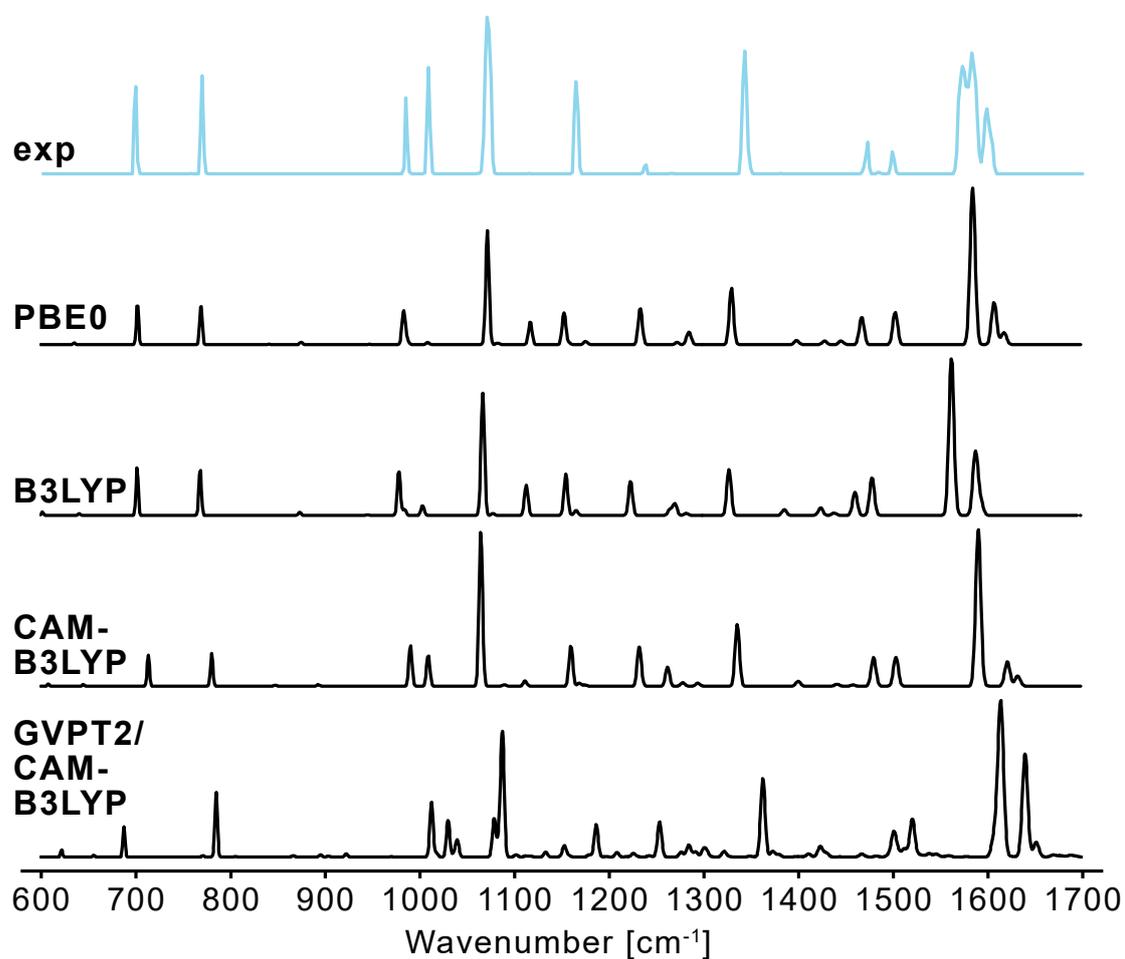


Figure F4: The experimental cryogenic IR spectrum of the 9-fluorenyl cation FI^+ is compared to computed scaled harmonic frequencies of the singlet state at the PBE0+D3/def2-TZVPP, B3LYP+D3/def2-TZVPP, and CAM-B3LYP+D3/def2-TZVPP levels of theory and unscaled anharmonic frequencies at CAM-B3LYP+D3/def2-TZVPP using the GVPT2 method. Except for the vibrations below 800 cm^{-1} , anharmonic frequencies are consistently blueshifted by ca. 20 cm^{-1} , compared to the experimental and scaled harmonic traces.

Table F5: Position of the experimental infrared absorption bands for the 9-phenyl-9-fluorenyl cation PhFl^+ (C_2 symmetry) and comparison to the position of scaled harmonic frequencies of the singlet state computed at CAM-B3LYP+D3/def2-TZVPP (scaling factor: 0.965). Only the bands in the experimentally probed region (600–1700 cm^{-1}) are indicated. All frequencies are given in cm^{-1} and computed intensities (I) in km mol^{-1} . Computed absorptions are indicated for the singlet state if its intensity is higher than 10 km mol^{-1} .

$\tilde{\nu}_{\text{exp}}$	$\tilde{\nu}_{\text{calc}} (I)$	Symmetry	$\tilde{\nu}_{\text{exp}}$	$\tilde{\nu}_{\text{calc}} (I)$	Symmetry
623	625 (29)	A	1290	1290 (19)	A
690	698 (32)	A	1322	1310 (62)	A
727	739 (121)	A	1333	1332 (26)	A
–	776 (14)	A	1387	1386 (509)	A
814	831 (19)	A	1399	1391 (70)	A
845	857 (24)	A	–	1454 (24)	A
993	995 (24)	A	1450	1456 (37)	A
1004	1003 (35)	A	1467	1472 (70)	A
1090	1084 (226)	A	1507	1510 (534)	A
1177	1172 (49)	A	–	1576 (14)	A
1190	1186 (20)	A	1577	1594 (215)	A
1216	1207 (183)	A	1586	1598 (120)	A
1260	1251 (59)	A	1597	1608 (184)	A
–	1285 (11)	A	–	1620 (118)	A

Table F6: Position of the experimental infrared absorption bands for the 9-hydroxy-9-fluorenyl cation HOFl^+ (C_S symmetry) and comparison to computed scaled harmonic (scaling factor: 0.965) and unscaled anharmonic frequencies of the singlet state (using the GVPT2 method at the CAM-B3LYP/def2-TZVPP level of theory). Only the bands in the experimentally probed region (600–1700 cm^{-1}) are indicated. All frequencies are given in cm^{-1} and computed intensities (I) in km mol^{-1} . Computed absorptions are indicated for the singlet state if its intensity is higher than 10 km mol^{-1} .

$\tilde{\nu}_{\text{exp}}$	$\tilde{\nu}_{\text{CAM-B3LYP}} (I)$	$\tilde{\nu}_{\text{GVPT2}} (I)$	Symmetry	$\tilde{\nu}_{\text{exp}}$	$\tilde{\nu}_{\text{CAM-B3LYP}} (I)$	$\tilde{\nu}_{\text{GVPT2}} (I)$	Symmetry
615	619 (78)	620 (94)	A ^o	1255	1247 (35)	1267 (27)	A ^o
–	625 (17)	642 (16)	A ^o	–	1283 (35)	1300 (27)	A ^o
725	738 (130)	751 (119)	A ^o	1365	1350 (303)	1366 (85)	A ^o
–	814 (16)	–	A ^o	–	1448 (29)	1468 (12)	A ^o
900	897 (52)	918 (49)	A ^o	–	1455 (35)	1477 (29)	A ^o
1000	1001 (17)	1026 (14)	A ^o	–	1476 (35)	1500 (12)	A ^o
–	1082 (41)	1107 (35)	A ^o	1479	1483 (99)	1507 (81)	A ^o
–	1139 (18)	1162 (33)	A ^o	1538	1541 (555)	1562 (414)	A ^o
–	1168 (55)	–	A ^o	–	1609 (43)	1632 (50)	A ^o
1171	1171 (88)	1187 (102)	A ^o	1602	1613 (316)	1634 (300)	A
1218	–	1190 (360)	–	–	1622 (70)	1644 (85)	A ^o

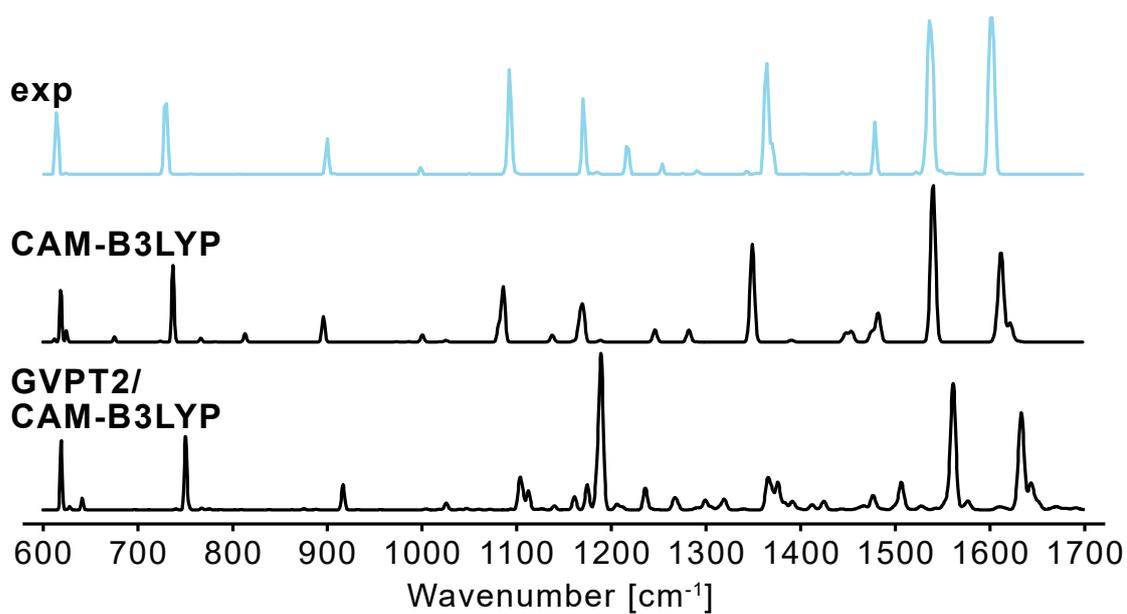


Figure F5: The experimental cryogenic infrared spectrum of the 9-hydroxy-9-fluorenyl cation **HOFI⁺** is compared to computed scaled harmonic frequencies of the singlet state at the CAM-B3LYP+D3/def2-TZVPP level of theory and unscaled anharmonic frequencies at CAM-B3LYP+D3/def2-TZVPP using the GVPT2 method. Except for the vibrations below 800 cm^{-1} , the anharmonic frequencies are consistently blueshifted by ca. 20 cm^{-1} , compared to the experimental and scaled harmonic traces.

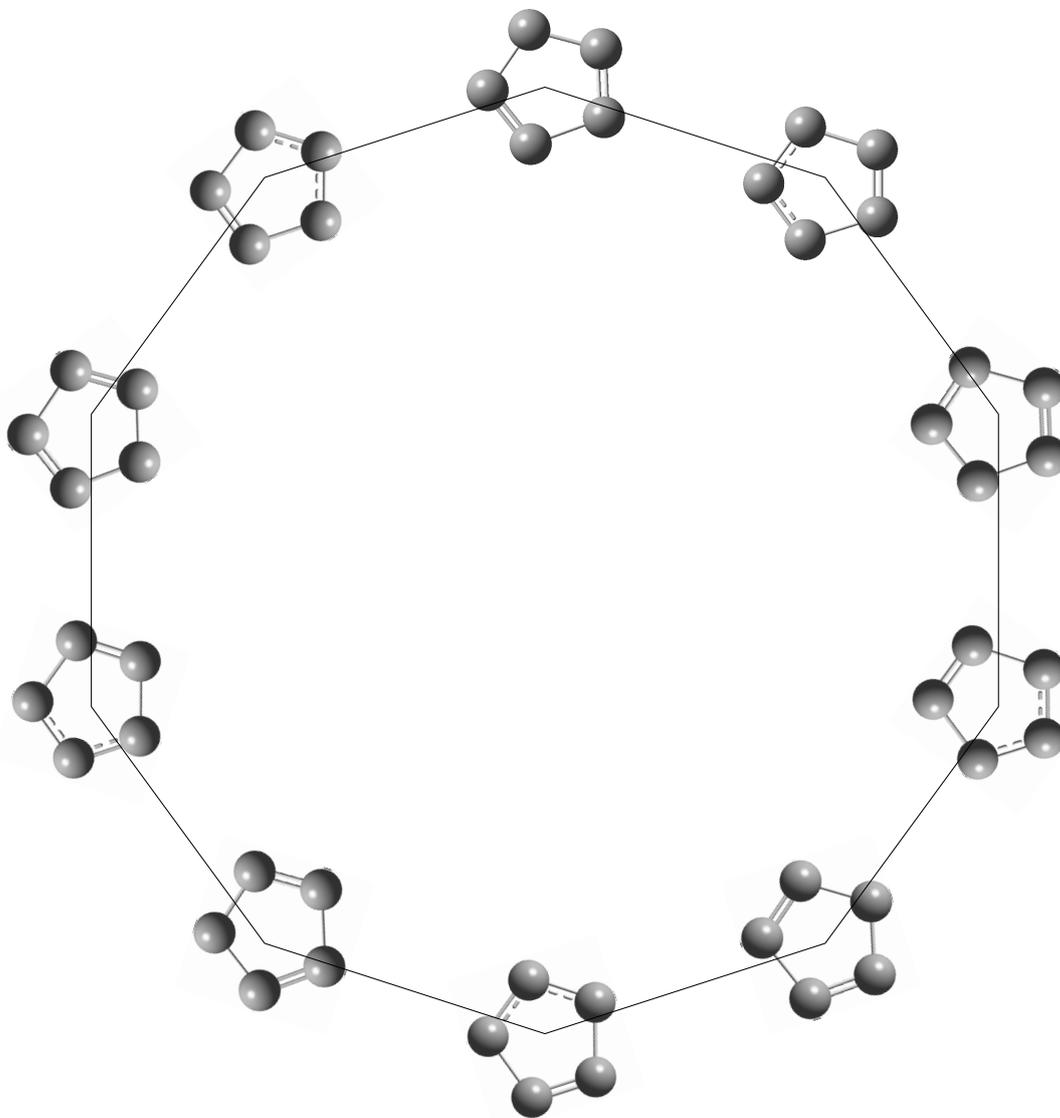


Figure F6: Full cycle showing the structures and isomerization between the diene and the allyl forms of the singlet state cyclopentadienyl cation. The symmetry equivalent diene and allyl structures can be interconverted by slight changes in C-C bond lengths.

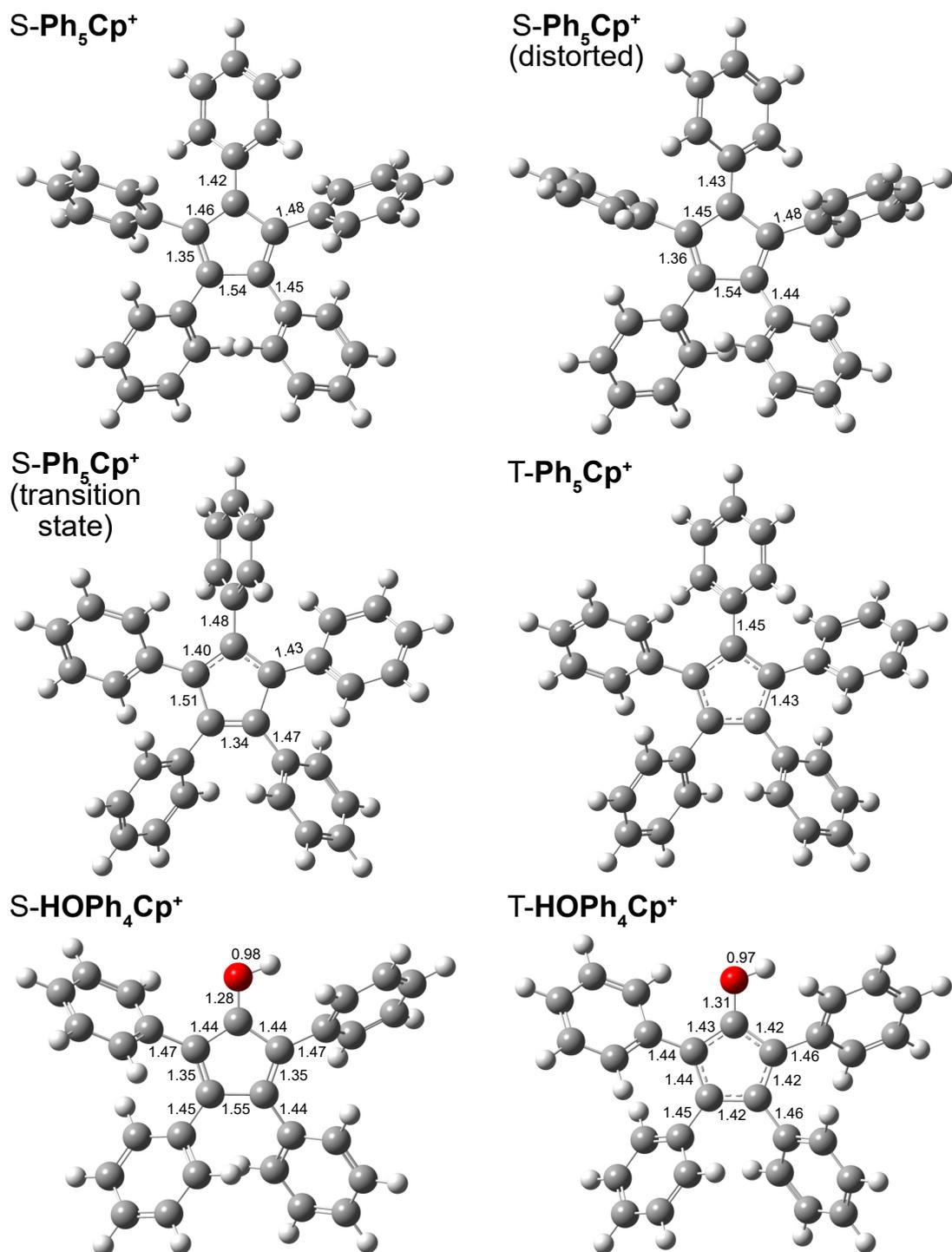


Figure F7: Optimized geometries of the singlet (S) and triplet (T) states of the pentaphenylcyclopentadienyl Ph_5Cp^+ (C_2 symmetry for singlet and D_5 symmetry for triplet) and the hydroxytetraphenylcyclopentadienyl cation HOPh_4Cp^+ (C_1 symmetry) at the CAM-B3LYP+D3/def2-TZVPP level of theory. For Ph_5Cp^+ also a distorted and a transition state structure (allyl) are shown. Bond lengths are indicated in Å.

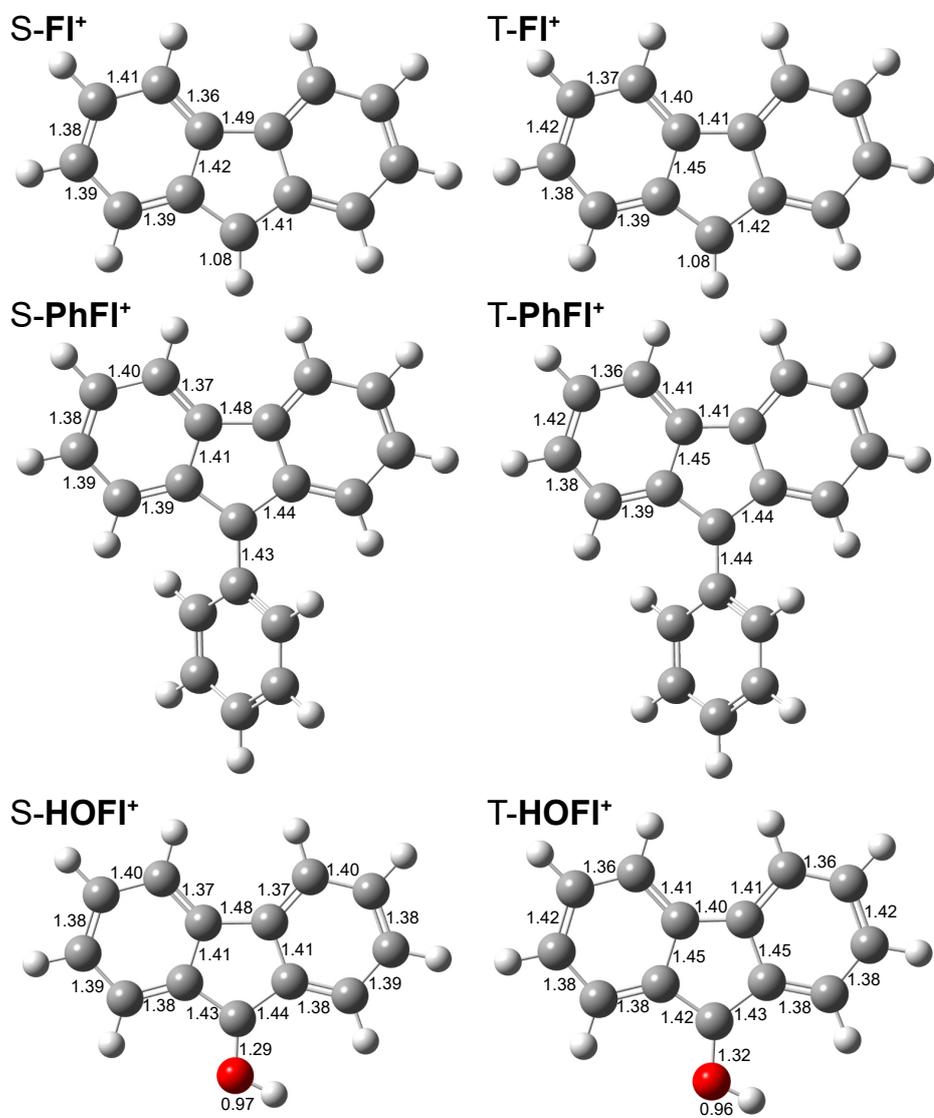


Figure F8: Optimized geometries of the singlet (S) and triplet (T) states of the 9-fluorenyl FI^+ (C_{2V} symmetry), the 9-phenyl-9-fluorenyl PhFI^+ (C_2 symmetry), and the 9-hydroxy-9-fluorenyl cation HOFI^+ (C_S symmetry) at the CAM-B3LYP+D3/def2-TZVPP level of theory. Bond lengths are indicated in Å.

Acknowledgments

I am deeply grateful to my supervisors Kevin and Gert, as well as my colleagues, collaborators, friends, and family, whose unwavering support made this work possible. The past few years have been truly enjoyable, and I look back on them with cheer.

Kevin, I cannot thank you enough for introducing me to the beautiful world of mass spectrometry. Even before joining your group for my master thesis project, you gave me useful advice that would push my career in the right direction. Upon my return from Australia, you entrusted me a very promising research topic that could not only fill a master thesis, but also a whole doctoral thesis. I am honored that you selected me for this topic.

Gert, I thank you for your trust and letting me use the helium droplet instrument. During my time in your research group I learned more about instrumentation than I had imagined before. Although mass spectrometers are already very useful instruments on their own, the techniques from your lab help getting a whole new degree of structural information on ions.

Without the helium droplet crew, the experiments for this thesis would not have been possible. Particularly, I want to thank Carla with whom I spent most of the time during my thesis. We did not only share an office and many FEL beamtimes, but also many cookies, pizza, and oven cheese. I benefited tremendously from your knowledge and enthusiasm. I will certainly miss it to discuss scientific questions with you! I want to thank all past (Eike, Daniel, Kathi, and Maike) and present (América, Katja, and Martín) members of the helium droplet crew for maintaining this almost magic instrument and helping me with the experiments.

Furthermore, I would like to express my gratitude to Gerard Meijer, who allowed me to use the facilities of the Fritz Haber Institute of the Max Planck Society, where all the experiments for this thesis were performed. I thank everyone from the molecular physics department for their generous help in the past years. Thank you Evelyn, Karin, Manuela, Sebastian, and Stefan to keep the department running! I also want to thank the FEL crew, consisting out of Wieland, Sandy, and Marco. Without your valuable FEL photons the experiments would not have been possible. It is truly impressive how efficiently you run the FHI-FEL facility. I want to thank you for your trust and teaching me how to operate the FEL.

I would like to thank everybody from Kevin's lab at the Freie Universität. Anna, I want to thank you for dealing with all the administrative issues. Chun-Wei, thank you for sharing all

your knowledge regarding sugar chemistry with me and doing some work on sugar synthesis side of my projects. Gergö, I want to thank you for your patience and your advice. Micha, thank you for contributing with your expertise to my projects. I also had the opportunity to supervise three students for research and master thesis projects. Jerome, Marc, and Yiran, it was a pleasure working together with you. Moreover, I want to thank all other former and current members of Kevin's and Gert's labs. Thank you Andreas, Christian, Dominika, Jan, Leïla, Łukasz, Márkó, Rayoon, and Tom!

Most of who will read these lines, know that I am a huge fan of computational chemistry. For performing quantum chemical calculations at an appropriate level of theory, high performance computing is necessary. I want to thank Heinz and William from PP&B for maintaining the Q cluster on which all the calculations for this thesis were performed.

Scientific research questions often require knowledge from multiple fields. Therefore, I am very grateful to Peter Seeberger, Mark Peczuh, Martina Delbianco, Sabrina Leichnitz, Giulio Fittolani, and Caleb Griesbach for providing the monosaccharide building blocks that we used in many of our cryogenic infrared studies.

I am very thankful to Mark Johnson, who was willing to host me for a six-month research stay at Yale University in 2022. It was a great experience in many ways and I learned a lot about the fundamentals of messenger-tagging infrared spectroscopy and science in general. I would also like to thank your past and current group members Abhijit, Ahmed, Anna, Anton, Evan, Franziska, James, Joey, Liv, Payten, Santino, Sayoni, Sean, Thien, and Tim. Furthermore, I would like to thank the Yale Tango Club for enhancing my time in the US so much! I am grateful to the Fulbright Program for funding my research stay in the US.

I would also like to thank my friends and family. Thank you for game nights, pool parties, trips inside and outside of Berlin, listening to me, and cheering me up Dustin, Elina, Emil, Jason, Julius, Kevin Parker, Liza, Oliver, and Sillu. Also big thanks to my flatmate Bernd. Living together is not always easy, but I learned a lot and it was an enjoyable time (and thanks for the food). Thank you Flora, Georges, and Ralf for providing a safe haven back in Luxembourg! Koko, I am very grateful for the time that we spent together, you enrich my life both scientifically and emotionally. Your intelligence, warmth, and unique perspective on the world have inspired me in so many ways, and I feel honored to have you in my life. We do not know yet where our journey will take us, but together we can manage it. I am looking forward to our future!

I want to thank all those who proofread my work and provided valuable feedback. Your meticulous attention to detail and insightful suggestions have greatly improved the quality of this thesis. Thank you América, Carla, Franziska, Gergö, Jan, Jerome, Julius, Thomas, Tiffany, and Wieland!

Last but not least, I am grateful to the Luxembourg National Research Fund (Fonds National de la Recherche, FNR) for having faith in my doctoral project GlycoCat (13549747). Thanks for the generous funding and supporting my endeavour to go to the US for a research stay.

List of Publications

Peer-Reviewed Journal Articles

28. K. Greis, C. Kirschbaum, M. I. Taccone, A. Y. Torres-Boy, K. Ober, G. Meijer, K. Pagel, G. von Helden; Infrared Spectroscopy of Cyclopentadienyl and Fluorenyl Cations at Cryogenic Temperatures, *in preparation*.
27. C. A. Lutomski, T. J. El-Baba, J. D. Hinkle, I. Liko, J. L. Bennett, N. V. Kalamankar, A. Dolan, C. Kirschbaum, K. Greis, L. Urner, P. Kapoor, H.-Y. Yen, K. Pagel, C. Mullen, J. E. P. Syka, C. V. Robinson; Infrared Multiphoton Dissociation Enables Top-Down Characterization of Membrane Protein Complexes and G-Protein Coupled Receptors. *Angew. Chem. Int. Ed.* **2023**, *62*, e202305694.
26. S. J. Stropoli, K. Greis, T. Schleif, M. A. Johnson; Characterization of Oxidation Products from HOCl Uptake by Microhydrated Methionine Anions Using Cryogenic Ion Vibrational Spectroscopy. *J. Phys. Chem. A* **2023**, *127*, 4269–4276.
25. K. Greis, C. E. Griesbach, C. Kirschbaum, G. Meijer, G. von Helden, K. Pagel, M. W. Peczu; Characterization and Fate of a Septanosyl Ferrier Cation in the Gas and Solution Phases. *J. Org. Chem.* **2023**, *88*, 5543–5553.
24. M. Lettow, K. Greis, E. Mucha, T. R. Lambeth, M. Yaman, V. Kontodimas, C. Manz, W. Hoffmann, G. Meijer, R. R. Julian, G. von Helden, M. Marianski, K. Pagel; Decoding the Fucose Migration Product during Mass-Spectrometric Analysis of Blood Group Epitopes. *Angew. Chem. Int. Ed.* **2023**, *62*, e202302883.
23. C. Kirschbaum, R. S. E. Young, K. Greis, J. P. Menzel, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, T. Causon, V. R. Narreddula, B. L. J. Poad, S. J. Blanksby, K. Pagel; Establishing Carbon-Carbon Double Bond Position and Configuration in Unsaturated Fatty Acids by Gas-Phase Infrared Spectroscopy. *Chem. Sci.* **2023**, *14*, 2518–2527.

22. K. Greis, S. Lechnitz, C. Kirschbaum, C.-W. Chang, M.-H. Lin, G. Meijer, G. von Helden, P. H. Seeberger, K. Pagel; The Influence of the Electron Density in Acyl Protecting Groups on the Selectivity of Galactose Formation. *J. Am. Chem. Soc.* **2022**, *144*, 20258–20566.
21. T. Khuu, S. J. Stropoli, K. Greis, N. Yang, M. A. Johnson; Microhydration of the Metastable *N*-Protomer of 4-Aminobenzoic Acid by Condensation at 80 K: H/D Exchange without Conversion to the More Stable *O*-Protomer. *J. Chem. Phys.* **2022**, *157*, 131102.
20. E. H. Perez, T. Schleif, J. P. Messinger, A. G. Rullán Buxó, O. C. Moss, K. Greis, M. A. Johnson, Structures and Chemical Rearrangements of Benzoate Derivatives Following Gas Phase Decarboxylation. *J. Am. Soc. Mass Spectrom.* **2022**, *33*, 1914–1920.
19. K. Greis, C. Kirschbaum, G. Fittolani, E. Mucha, R. Chang, G. Meijer, G. von Helden, M. Delbianco, P. H. Seeberger, K. Pagel; Neighboring Group Participation of Benzoyl Protecting Groups in C3- and C6-Fluorinated Glucose. *Eur. J. Org. Chem.* **2022**, e202200255.
18. K. Greis, C. Kirschbaum, M. I. Taccone, M. Götze, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, K. Pagel; Studying the Key Intermediate of RNA Autohydrolysis by Cryogenic Gas-Phase Infrared Spectroscopy. *Angew. Chem. Int. Ed.* **2022**, *61*, e202115481.
17. C. Kirschbaum, K. Greis, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, K. Pagel; Cryogenic Infrared Spectroscopy Provides Mechanistic Insight into the Fragmentation of Ag(I)-Adducted Phospholipid Isomers. *Anal. Bioanal. Chem.* **2022**, *414*, 5275–5285.
16. K. Greis, C. Kirschbaum, G. von Helden, K. Pagel; Gas-Phase Infrared Spectroscopy of Glycans and Glycoconjugates. *Curr. Opin. Chem. Biol.* **2022**, *72*, 194–202.
15. M. Grabarics, M. Lettow, C. Kirschbaum K. Greis, C. Manz, K. Pagel; Mass Spectrometry-Based Techniques to Elucidate the Sugar Code. *Chem. Rev.* **2022**, *122*, 7840–7908.
14. C. Kirschbaum, K. Greis, L. Polewski, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, K. Pagel; Unveiling Glycerolipid Fragmentation by Cryogenic Infrared Spectroscopy. *J. Am. Chem. Soc.* **2021**, *143*, 14827–14834.
13. M. Lettow, K. Greis, M. Grabarics, J. Horlebein, R. L. Miller, G. Meijer, G. von Helden, K. Pagel; Chondroitin Sulfate Disaccharides in the Gas Phase: Differentiation and Conformational Constraints. *J. Phys. Chem. A* **2021**, *125*, 4373–4379.
12. C. Kirschbaum, K. Greis, M. Lettow, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, K. Pagel; Non-Covalent Double Bond Sensors for Gas-Phase Infrared Spectroscopy of Unsaturated Fatty Acids. *Anal. Bioanal. Chem.* **2021**, *413*, 3643–3653.

11. C. Kirschbaum, K. Greis, E. Mucha, L. Kain, S. Deng, A. Zappe, S. Gewinner, W. Schöllkopf, G. von Helden, G. Meijer, P. B. Savage, M. Marianski, L. Teyton, K. Pagel; Unravelling the Structural Complexity of Glycolipids with Cryogenic Infrared Spectroscopy. *Nat. Commun.* **2021**, *12*, 1201.
10. K. Greis, C. Kirschbaum, S. Lechnitz, S. Gewinner, W. Schöllkopf, G. von Helden, G. Meijer, P. H. Seeberger, K. Pagel; Direct Experimental Characterization of the Ferrier Glycosyl Cation in the Gas Phase. *Org. Lett.* **2020**, *22*, 8916-8919.
9. D. A. Thomas, R. Chang, E. Mucha, M. Lettow, K. Greis, S. Gewinner, W. Schöllkopf, G. von Helden; Probing the Conformational Landscape and Thermochemistry of DNA Dinucleotide Anions via Helium Nanodroplet Infrared Action Spectroscopy. *Phys. Chem. Chem. Phys.* **2020**, *22*, 18400–18413.
8. K. Greis, E. Mucha, M. Lettow, D. A. Thomas, C. Kirschbaum, S. Moon, A. Pardo-Vargas, G. von Helden, G. Meijer, K. Gilmore, P. H. Seeberger, K. Pagel; The Impact of Leaving Group Anomerism on the Structure of Glycosyl Cations of Protected Galactosides. *ChemPhysChem* **2020**, *21*, 1905–1907.
7. M. Lettow, M. Grabarics, K. Greis, E. Mucha, D. A. Thomas, P. Chopra, G.-J. Boons, R. Karlsson, J. E. Turnbull, G. Meijer, R. L. Miller, G. von Helden, K. Pagel; Cryogenic Infrared Spectroscopy Reveals Structural Modularity in the Vibrational Fingerprints of Heparan Sulfate Diastereomers. *Anal. Chem.* **2020**, *92*, 10228–10232.
6. C. Kirschbaum, E. M. Saied, K. Greis, E. Mucha, S. Gewinner, W. Schöllkopf, G. Meijer, G. von Helden, B. L. J. Poad, S. J. Blanksby, C. Arenz, K. Pagel; Resolving Sphingolipid Isomers using Cryogenic Infrared Spectroscopy. *Angew. Chem. Int. Ed.* **2020**, *59*, 13638–13642.
5. M. Marianski, E. Mucha, K. Greis, S. Moon, A. Pardo, C. Kirschbaum, D. A. Thomas, G. Meijer, G. von Helden, K. Gilmore, P. H. Seeberger, K. Pagel; Remote Participation during Glycosylation Reactions of Galactose Building Blocks: Direct Evidence from Cryogenic Vibrational Spectroscopy. *Angew. Chem. Int. Ed.* **2020**, *59*, 6166–6171.
4. K. Greis, Y. Yang, A. J. Canty, R. A. J. O’Hair; Gas-Phase Synthesis and Reactivity of Ligated Group 10 Ions in the Formal +1 Oxidation State. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 1867–1880.
3. K. Greis, K. Bethke, J. B. Stückrath, T. T. K. Ingber, S. Valiyaveetil, K. Rademann; One-Pot Synthesis of Xanthate Functionalized Cellulose for the Detection of Micromolar Copper(II) and Nickel(II) Ions. *Clean – Soil, Air, Water* **2019**, *47*, 1900179.

2. [K. Greis](#), A. J. Canty, R. A. J. O'Hair; Gas-Phase Reactions of the Group 10 Organometallic Cations, [(phen)M(CH₃)]⁺ with Acetone: Only Platinum Promotes a Catalytic Cycle via the Enolate [(phen)Pt(OC(CH₂)CH₃)]⁺. *Z. Phys. Chem.* **2019**, 233, 845–864.
1. K. Bethke, S. Palantöken, V. Andrei, M. Roß, V. S. Raghuwanshi, F. Kettemann, [K. Greis](#), T. T. K. Ingber, J. B. Stückrath, S. Valiyaveetil, K. Rademann; Functionalized Cellulose for Water Purification, Antimicrobial Applications, and Sensors. *Adv. Funct. Mater.* **2018**, 28, 1800409.

Preprints

1. S. Ponath, C. Joshi, A. T. Merrill, V. Schmidts, [K. Greis](#), M. Lettow, M. Weber, S. Steinhauer, K. Pagel, C. Thiele, D. J. Tantillo, M. J. Veticatt, M. Christmann: On Stereocontrol in Organocatalytic α -Chlorinations of Aldehydes. *ChemRxiv* **2021**.

Other Publications

1. [K. Greis](#); Emerging Scientists in Analytical Sciences: Kim Greis. *Anal. Sci. Adv.* **2022**, 3, 258–261.