Non-natural oligosaccharides: from structural studies to the design of synthetic carbohydrate materials

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

Giulio Fittolani

from Maragnole, Italy

2022

This work was performed between January 2019 and May 2022 under the supervision of Dr. Martina Delbianco and Prof. Dr. Peter H. Seeberger in the Department of Biomolecular Systems, Max Planck Institute of Colloids and Interfaces Potsdam, and the Institute of Chemistry and Biochemistry, Freie Universität Berlin.

1st reviewer: Prof. Peter H. Seeberger

2nd reviewer: Prof. Beate Koksch

Date of oral defense: 9.9.2022

Acknowledgement

First of all, I would like to thank my supervisor Dr. Martina Delbianco for picking me up when I was still in my Master and for guiding and supporting me throughout the whole PhD. Your help and suggestions have been precious for me in these years.

I would like to extend my gratitude to Prof. Peter H. Seeberger for giving me the opportunity to work in his Department, supported by excellent equipment and a great environment. Thank you for your constant support and for reviewing this thesis.

I thank Prof. Beate Koksch for kindly agreeing to review my thesis.

My time at the Max Planck Institute of Colloids and Interfaces has certainly been great and I must acknowledge all the people that contributed to making this journey unique. Previous and current members of the *Carbohydrate Materials* group such as Dr. Soeun Gim, Dr. Yang Yu, Dr. Vittorio Bordoni, Dr. Yuntao Zhu, Dr. Theodore Tyrikos-Ergas, Dr. Denisa Vargová, Dr. Manishkumar Chaube, Zhouxiang Zhao, Marlene Dal Colle, Jhih-Yi Huang, and Surusch Djalali have contributed immensely to my growth as a scientist, but mostly they have created an amazing atmosphere in the lab. I am thankful to all of you! I would like to thank all the people in the Department of Biomolecular Systems that contributed to this thesis. Everyone had his/her little contribution to this work, someone with discussions on my projects, someone with essential technical support, and someone contributed with beers and company during free time. Thanks also to the staff, Olaf, Eva, Felix, and Dorothee, for their constant support. I would like to extend my acknowledgment also to other members of the MPIKG for being always nice and kind to me, and especially Majd, the first human being I met in Golm.

I am grateful also to the collaborators which gave contributions to some projects of my thesis. Thank you for your hard work and invaluable help Prof. Christoph Rademacher, Dr. Elena Shanina, Prof. Jesús Jiménez-Barbero, Dr. Ana Poveda, Dr. Kelvin Anggara, and Dr. Yu Ogawa.

I want also to thank my friends back in Italy for their key contributions in making my free time fun and for supporting me in their special way.

Finally, I want to thank my family in Italy that shared with me the happy and hard moments of my PhD. I am lucky that I went through this journey with you by my side. Thank you, Mario, Rosa, Francesca, Davide, and Maddalena! A special thought goes also to all my grandparents, but especially to *Nonno Jìjo*, I dedicate this thesis to you! Special gratitude goes also to Juanita, which has been an invaluable guide for me, both from long and short distance.

Herewith I certify that I have prepared and written my thesis independently and that I have not used any sources and aids other than those indicated by me.

Dedicated to Nonno Jijo,

Outline

Summary

Biopolymers self-assemble generating dynamic and complex functional materials that carry out sophisticated tasks. Nature has been a source of inspiration for chemists to create artificial analogs that mimic biopolymer self-assembly. In contrast to peptides and DNAs, carbohydrates are less understood at the molecular level and therefore have rarely been employed as scaffolds to construct self-assembling materials using *bottom-up* approaches. To date, carbohydrates' potential in supramolecular chemistry remains mostly untapped. In this thesis, I established structure-property correlations for oligosaccharides and then translated this knowledge into design principles to access self-assembling carbohydrate materials. The use of non-natural monosaccharides and *bottom-up* approaches were central to deeply understand and design synthetic carbohydrate materials.

In Chapter 2, I developed synthetic routes to fluorinated monosaccharide building blocks (BBs) and employed these BBs to assemble a broad collection of complex fluorinated glycans using Automated Glycan Assembly (AGA). Fluorination was exploited as a method to manipulate the hydrogen bonds of cellulose, a polysaccharide with high propensity to crystallize. The subtle role of the substitution pattern and position of the fluorine atom was investigated using X-ray diffraction analysis (XRD). In the second part of Chapter 2, I explored the use of fluorinated glycans as NMR probes. First, ^{19}F-labelled glycans served for structural analysis *via* NMR revealing the tendency of some oligosaccharides to adopt helical conformations. Last, ^{19}F-labelled Lewis antigens, a class of biologically relevant glycans, were employed to study their binding to proteins and to monitor in real-time enzymatic reactions.

In Chapter 3, I developed a *bottom-up* approach to study the self-assembly and the chirality of supramolecular carbohydrate materials. In the first part, I employed a simple disaccharide that self-assembles into helical fibers to study the transfer of chirality across scales. Site-specific modifications identified key non-covalent interactions stabilizing the assembly. In the second part, I translated the *bottom-up* approach to a more complex natural polysaccharide and used synthetic oligomers to understand chirality transfer across scales in cellulose. Synthetic D- and L-cellulose oligomers were found to assemble into platelets with controlled dimensions that further aggregate into bundles, displaying chiral features directly connected to their monosaccharide composition. The insertion of L-Glc units in the sequence of D-Glc cellulose oligomers drastically impacted macroscopic properties such as solubility, crystallinity, and chirality of bundles.

In Chapter 4, I presented the rational design and synthesis of a glycan adopting a stable secondary structure, challenging the common belief that glycans are not capable of folding due to their flexibility. By combining natural glycan motifs, stabilized by a non-conventional hydrogen bond and hydrophobic interactions, I designed a glycan hairpin, a secondary structure not present in nature. AGA enabled rapid access to synthetic analogs, including site-specific ^{13}C-labelled ones, for NMR conformational analysis. Structural analysis *via* NMR unequivocally confirmed the folded conformation of the synthetic glycan hairpin. This work demonstrates that it is possible to program glycans to adopt defined conformations in aqueous solution.

Overall, the *bottom-up* approaches used in this doctoral thesis allowed for a deeper understanding of the principles dictating carbohydrate self-assembly. The work presented here opens the way to future explorations of glycans as scaffolds for self-assembly or to perform complex functions as catalysis.

Zusammenfassung

Biopolymere sind selbstanordnende, dynamisch-komplexe Funktionsmaterialien, die bedeutende Aufgaben erfüllen. Für die Synthese künstlicher Analoga dienen der Wissenschaft die von der Natur genutzten Mechanismen der Selbstorganisation noch immer als Vorbild und Inspirationsquelle. Im Gegensatz zu Peptiden und DNA sind Kohlenhydrate auf molekularer Ebene bis jetzt nur wenig erforscht. Bisher wurden sie nur selten als Basis für selbstorganisierende Materialien mit *Bottom-up*-Ansätzen verwendet. Bis heute bleibt das Potenzial von Kohlenhydraten in der supramolekularen Chemie weitgehend ungenutzt. In dieser Arbeit habe ich Struktur-Eigenschafts-Korrelationen für Oligosaccharide erforscht und mit Hilfe dieser Designprinzipien gestaltet, die den Zugang zu selbstorganisierenden Kohlenhydratmaterialien ebnen. Dabei war die Verwendung nicht-natürlicher Monosaccharide und *Bottom-up*-Ansätze von zentraler Bedeutung, um ein tieferes Verständnis für die Entwicklung synthetischer Kohlenhydratmaterialien zu gewinnen.

In Kapitel 2 entwickelte ich synthetische Wege zu fluorierten Monosaccharid-Bausteinen (BBs) und erschloss, unter zur Hilfenahme von *Automated Glycan Assembly* (AGA), eine breite Sammlung komplexer fluorierter Glykane. Die Fluorierung wurde als Methode zur Manipulation der Wasserstoffbrückenbindungen von Cellulose, einem Polysaccharid mit hoher Kristallisationsneigung, genutzt. Die Rolle des Substitutionsmusters und der Position des Fluoratoms wurde mit Hilfe der Röntgenbeugungsanalyse (XRD) untersucht. Im zweiten Teil von Kapitel 2 testete ich das Potential der fluorierten Glykane als NMR-Sonden. Zunächst dienten ^{19}F-markierte Glykane zur Strukturanalyse mittels NMR, wobei die Tendenz einiger Oligosaccharide zur Annahme helikaler Konformationen deutlich wurde. Schließlich wurden ^{19}F-markierte Lewis-Antigene, eine Klasse biologisch relevanter Glykane, zur Untersuchung ihrer Bindung an Proteine und zur Überwachung enzymatischer Reaktionen in Echtzeit eingesetzt.

In Kapitel 3 entwickelte ich einen *Bottom-up*-Ansatz zur Untersuchung der Selbstorganisation und der Chiralität von supramolekularen Kohlenhydratmaterialien. Im ersten Teil konnte ich mit Hilfe eines einfachen Disaccharids, dass sich selbst zu spiralförmigen Fasern zusammensetzt, die Übertragung von Chiralität über mehrere Maßstäbe hinweg verfolgen. Durch ortsspezifische Modifikationen wurden wichtige nicht-kovalente Wechselwirkungen identifiziert, die den spezifischen Aufbau stabilisieren. Im zweiten Teil übertrug ich den *Bottom-up*-Ansatz auf das natürlich vorkommende Polysaccharid Cellulose. Die Studien der synthetischen Oligomere erlaubten es den Chiralitätstransfer über mehrere Maßstäbe hinweg zu verstehen. Ich konnte feststellen, dass sich synthetische D- und L-Cellulose-Oligomere zu Plättchen mit kontrollierten Abmessungen zusammensetzen, die sich wiederum zu chiralen Bündeln anordnen. Die Chiralität dieser Aggregate ist dabei direkt von der Monosaccharidzusammensetzung abhängig. Die Einfügung von L-Glc-Einheiten in die Sequenz von D-Glc-Cellulose-Oligomeren wirkte sich drastisch auf die makroskopischen Eigenschaften wie Löslichkeit, Kristallinität und Chiralität der Bündel aus.

In Kapitel 4 habe ich das rationale Design und die Synthese eines Glykans mit einer stabilen Sekundärstruktur vorgestellt und damit die gängige Meinung widerlegt, dass Glykane aufgrund ihrer Flexibilität nicht faltbar sind. Durch die Kombination natürlicher Glykanmotive, die durch eine unkonventionelle Wasserstoffbrückenbindung und hydrophobe Wechselwirkungen stabilisiert werden, habe ich eine Glykan-Haarnadel entworfen, eine Sekundärstruktur, die in der Natur nicht vorkommt. AGA ermöglichte einen schnellen Zugang zu synthetischen Analoga, einschließlich ortsspezifischer 13C-markierter Spezies, die für NMR-Konformationsanalysen genutzt werden können. Die Strukturanalyse mittels NMR bestätigte eindeutig die gefaltete Konformation der synthetischen Glykan-Haarnadel. Diese Arbeit zeigt, dass es möglich ist, Glykane so zu programmieren, dass sie in wässriger Lösung definierte Konformationen annehmen.

Insgesamt ermöglichten die in dieser Dissertation verwendeten Bottom-up-Ansätze ein tieferes Verständnis der Prinzipien, die die Selbstorganisation von Kohlenhydraten bestimmen. Die hier vorgestellte Arbeit öffnet den Weg für künftige Untersuchungen von Glykanen als Gerüste für die Selbstorganisation oder zur Ausführung komplexer Funktionen wie der Katalyse.

List of abbreviations

ACN acetonitrile

AFM atomic force microscopy

AGA automated glycan assembly

BBs building blocks

Bn benzyl

Bz benzoyl

CD circular dichroism

CNC cellulose nanocrystal

CNF cellulose nanofiber

Cryo-EM cryo-electron microscopy

CSP chemical shift perturbation

DAST diethylaminosulfur trifluoride

DCM dichloromethane

DMF dimethylformamide

DP degree of polymerization

ED electron diffraction

EDG electron donating group

ESI electrospray ionization

ES-IBD electrospray ion-beam deposition

EtOAc ethylacetate

EWG electron withdrawing group

Fmoc 9-fluorenylmethoxycarbonyl

FTIR Fourier transfer infrared

Fuc fucose

Gal galactose

Glc glucose

GlcNAc *N*-acetylglucosamine

HFIP 1,1,1,3,3,3-hexafluoroisopropanol

HPLC high-performance liquid chromatography

HSQC heteronuclear single-quantum coherence

Lev levulinoyl

MD molecular dynamics

MicroED microcrystal electron diffraction

MS mass spectrometry

MW molecular weight

ncAA non-canonical amino acid

NIS *N*-iodosuccinimide

NMR nuclear magnetic resonance

NOE nuclear Overhauser effect

NOESY nuclear Overhauser effect spectroscopy

NP normal phase

PG protecting group

Ph phenyl

Py pyridine

RDC residual dipolar coupling

Rha rhamnose

RMSD root-mean-square deviation

ROESY rotating frame nuclear Overhauser effect

R_g radius of gyration

RP reverse phase

RT room temperature

SEM scanning electron microscopy

SNFG symbol nomenclature for glycans

ssNMR solid state nuclear magnetic resonance

STM scanning tunneling microscopy

TBS (also TBDMS) *t*-butyl dimethyl silyl

TCA trichloroacetamide

TEM transmission electron microscopy

TfOH triflic acid

THF tetrahydrofurane

TMSOTf trifluoromethansulfonic triflate

TOCSY total correlation spectroscopy

Tol toluene

UV-Vis ultraviolet-visible

XRD X-ray diffraction

1 Introduction

This chapter has been modified in part from the following articles:

G. Fittolani, P.H. Seeberger, M. Delbianco Helical polysaccharides *Pept. Sci.* **2019**, 112, 1, e24124 https://doi.org/10.1002/pep2.24124

T. Tyrikos-Ergas,[+] **G. Fittolani**,[+] P.H. Seeberger, M. Delbianco Structural Studies Using Unnatural Oligosaccharides: Toward Sugar Foldamers *Biomacromolecules* **2020**, 21, 1, 18-29 https://doi.org/10.1021/acs.biomac.9b01090

G. Fittolani, T. Tyrikos-Ergas, D. Vargová, M. A. Chaube, M. Delbianco Progress and challenges in the synthesis of sequence controlled polysaccharides *Beilstein J. Org. Chem.* **2021**, 17, 1981-2025 https://doi.org/10.3762/bjoc.17.129

[+]equal contribution

1.1 Synthetic materials based on peptide, nucleic acids, and glycans

In nature, biopolymers self-assemble to generate dynamic and complex functional materials.[1] Upon self-assembly peptides, nucleic acids, and carbohydrates are capable of achieving unique properties. For instance, self-assembly is responsible for the excellent thermal stability and mechanical strength of collagen, a structural protein component of human skin.[2] Similarly, polysaccharides such as cellulose and chitin achieve excellent mechanical properties through the hierarchical self-assembly of single chains into fibrils.[3,4] Nucleic acids self-assembly into compact double helical structures enables safe and efficient storage of genetic information. Molecular self- assembly is driven by non-covalent interactions such as van der Waals, electrostatic, and hydrophobic interactions as well as hydrogen and coordination bonds. Non-covalent interactions are typically longer range compared to covalent bonds, reversible, and often benefit from cooperativity.[5]

Nature has been a source of inspiration for chemists to create artificial analogs that mimic biopolymer self-assembly.[6–11] Self-assembling synthetic materials offer all the advantages of a *bottom-up* approach; these materials are programmable, can be responsive to external stimuli, and can be manipulated at the nanoscale to control their 3-D shape, physicochemical, and mechanical properties.[12,13] Sequence engineering and/or chemical modifications enable fine-tuning of the material and the reversibility of the self-assembly process, permitting responsiveness to external stimuli (e.g. pH, temperature, ionic strength), adaptation to the environment, and the ability to exert bioactive functions. Last, biopolymer-based materials are expected to be biocompatible. Using self-assembly, a variety of synthetic materials with diverse macroscopic properties and nanoscale morphologies has been created ranging from fibers, membranes,[14] platelets,[15] hydrogels,[16] as well as more complex architectures such as DNA origami (**Figure 1.1**).[17,18] Synthetic self-assembling materials are attractive for a wide range of applications in the biomedical field.[13] A prominent example is in tissue regeneration in which the reversibility of the self-assembly and the precise geometrical display of bioactive moieties triggering a biological event are great advantages over conventional covalent polymers.[19]

Peptides have been designed to self-assemble and generate a plethora of materials. Amphiphatic peptides were among the first self-assembling peptides discovered in the 1990s; their sequences featured alternating hydrophobic and hydrophilic residues and were capable of forming nanofibers and macroscopic membranes.[8,14] Attaching a hydrophobic alkyl tail to a β-sheet promoting peptide sequence allowed to obtain peptide amphiphiles, capable of assembling into fibers of various shape and properties (**Figure 1.1**A).[20] Peptide sequences responsible for protein amyloid formation were exploited to generate hydrogels, nanotubes, and artificial biofilms.[10,21–23] In addition, short peptides have been programmed to fold into hairpin secondary structures and self-assemble into fibers forming a hydrogel.[16,24] Other self-assembling peptides rely on α-helices scaffolds forming coiled-coil superstructures that can assemble in a predictable manner to achieve precise control over the 3-D shape.[25,26] Collagen has also been a source of inspiration for designing peptides that self-assemble into triple helical fibrils based hydrogels (**Figure 1.1**B, *left*).[27,28] Sequence manipulation of collagen mimetic peptides gave access to an entirely new morphology resembling thin square platelets (**Figure 1.1**B, *right*).[15,29] More recently, it was shown that protein could be designed to self-assemble into well-defined nanoscale architectures (**Figure 1.1**C).[30–33] These advancements were supported by novel computational methods, enabling accurate protein folding prediction.[34]

Other spectacular examples of synthetic self-assembling material are DNA-based assemblies.[17,18,35] The high predictability and specificity of base pairing and the well-studied molecular conformation of DNA fostered the development DNA nanotechnology.[36] This field has advanced so rapidly in the last two decades that it is now possible to generate nanoscale objects with virtually any 2-D or 3-D shape *via* rational design (i.e. DNA origami, **Figure 1.1**D).[18]

In contrast to peptides, proteins, and DNA, synthetic self-assembling carbohydrate materials remain, to date, highly unexplored. Carbohydrates are often exploited merely as a hydrophilic moiety and rarely

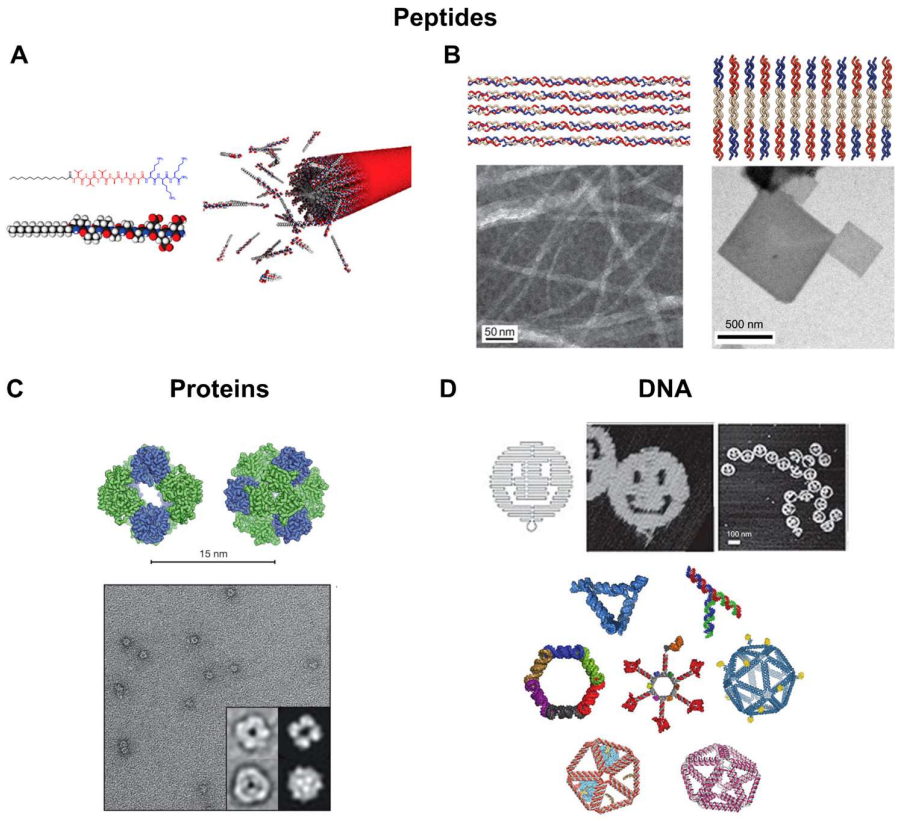


Figure 1.1 A) Peptide amphiphiles and model of the self-assembled fibers. B) Collagen mimetic peptides can be designed to self-assemble into fibers (*left*) or platelets (*right*) as observed in TEM images. C) Protein engineered to self-assemble into defined nano-sized architectures. TEM image (*bottom*) of the designed two-component protein cage-like architecture (*top*). D) The DNA origami technique allows folding a long single-stranded DNA sequence into various nano-sized architectures (*bottom*). AFM images of a disk with three holes (*top*). Reprinted (or adapted) with permission from *Acc. Chem. Res.* **2017**, 50, 2440-2448 Copyright 2017 American Chemical Society, *Nat. Chem.* **2011**, 3, 821-828 Copyright 2011 Nature Publishing Group, *J. Am. Chem. Soc.* **2014**, 136, 4300-4308 Copyright 2014 American Chemical Society, *Nature* **2006**, 440, 297-302 Copyright 2006 Nature Publishing Group, and *Nature* **2014**, 510, 103-108 Copyright 2014 Nature Publishing Group.

employed as scaffolds to construct self-assembling materials using *bottom-up* approaches.[37] Only few examples of self-assembled carbohydrate materials have been reported, mostly based on simple mono- and disaccharides.[38] Still, carbohydrates offer many attractive features. They are highly diverse in nature since they are constructed from a vast pool of more than 100 natural monosaccharides that can be combined into linear or branched polymers, suggesting an enormous potential to access new conformations compared to other biopolymers. Plenty of exceptional polysaccharide-based materials exist in nature that could serve as an inspiration for the generation of synthetic materials *via* a *bottom-up* approach. For instance, cellulose and chitin are mechanically stable structural materials found in plant and fungal cell walls, respectively, with a strong tendency to self-assemble in a hierarchical fashion (**Figure 1.2**A and B). Carrageenans, sulfated polysaccharide found in algae, undergo ion-mediated folding into helical secondary structures and further chain association (i.e. supercoiling) leads to hydrogel formation (**Figure 1.2**A and C).[39,40]

In comparison with peptides and DNA, the chemical structure of carbohydrates is intrinsically more complex and diverse, hampering the development of rationally designed carbohydrate assemblies. Thorough structure-function correlations, enabled by easy access to synthetic compounds and rigorous structural characterizations, have fostered the creation of synthetic materials based on peptides and DNA. In contrast, the synthetic and analytical challenges of carbohydrates have severely limited the development of *bottom-up* approaches in the carbohydrate materials field. *Top-down* approaches, in which carbohydrates are extracted from natural sources and chemically modified, have dominated the scene to date.[41,42] However,

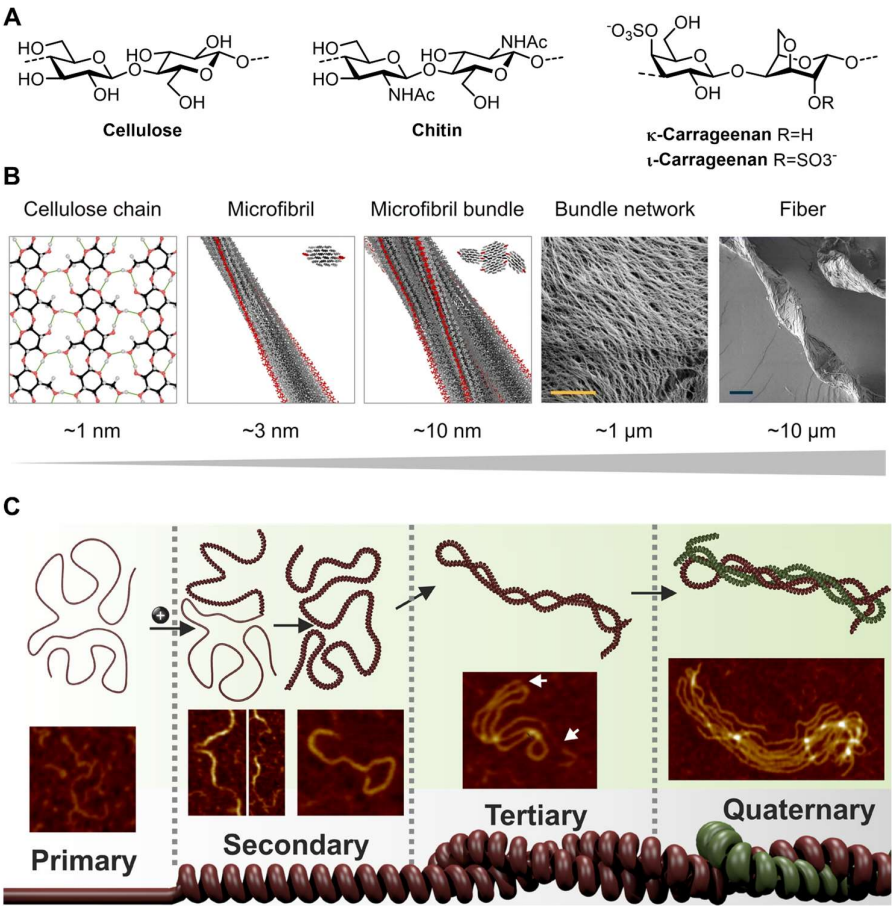


Figure 1.2 A) Chemical structure of cellulose, chitin, κ-, and ι-carrageenans. B) Hierarchical levels of the cellulose fiber structure. Adapted with permission from *Cellulose* **2019**, 27, 5877-5892 Copyright 2019 Nature Publishing Group. C) Primary, secondary, tertiary, and quaternary structures arising in carrageenans. Single carrageenan chains undergo ion-mediated folding into helical conformation structures that can further form intra- or intermolecular supercoiled structures (inserts are AFM images). Adapted with permission from *Biomacromolecules* **2019**, 20, 1731-1739 Copyright 2019 American Chemical Society.

extraction of oligosaccharides from natural sources often results in heterogeneous mixtures hampering structure-function correlations. Chemical synthesis can provide homogeneous oligosaccharide samples but can be challenging and lengthy (see Section 1.3). Constructing glycosidic linkages requires both regio- and stereocontrol, imposing the use of multiple protecting groups and time-consuming protocols. Beside synthesis, also structural analysis is a major challenge (see Section 1.2).[43] The lack of chromophores limits the use of circular dichroism (CD) and difficulties in obtaining single crystals hinder the application of X-ray crystallography. NMR of carbohydrates may provide some structural information (but suffers from other problems discussed more in detail in Section 1.2). Addressing the synthetic and analytical challenge of carbohydrates is highly important to generate design principle to access synthetic carbohydrate materials. Such principles are well-established for peptides[6] and DNA[18] enabling the creation of synthetic materials, but they are still lacking for carbohydrates.

1.2 Conformational analysis of oligo- and polysaccharides

The study of polysaccharide secondary structures, and higher hierarchical assemblies, is hampered by the intrinsic complexity of polysaccharides. The main variable implicated in polysaccharide conformation is the geometry of the glycosidic linkage, with the monosaccharide units generally considered rigid (in the case of pyranose rings).[44] The torsion angles Φ (H_1-C_1-O_x-C_x) and Ψ (C_1-O_x-C_x-H_x) define the relative orientation of two monosaccharides involved in a glycosidic bond. For 1,6-linkages, the ω torsion angle (O_6-C_6-C_5-O_5) provides additional flexibility. The most populated conformation is generally the exo-syn-Φ,[45] due to hyperconjugation between the exocyclic oxygen lone electron pair (n_p) and the antibonding orbital (σ*) of the endocyclic C−O bond (i.e. exo-anomeric effect) (**Figure 1.3**).[46] The Ψ dihedral is more sensitive to sterics, favoring the anti-Ψ conformer.[47] Water interacts with the hydroxyl groups and can disrupt intramolecular hydrogen bonds. As a result, glycosidic linkages possess a significant degree of freedom, often resulting in quite flexible oligosaccharides.[48]

Such complexity has hampered the development of suitable analytical techniques. Even though some crystal structures were obtained for polysaccharide samples using X-ray fiber diffraction, the geometry of many complex oligo- and polysaccharides remains less well understood. The complementary, rather than alternative methods used to study polysaccharide conformations and folding can be divided into three broad categories: i) diffraction techniques, ii) imaging techniques, and iii) spectroscopic techniques.

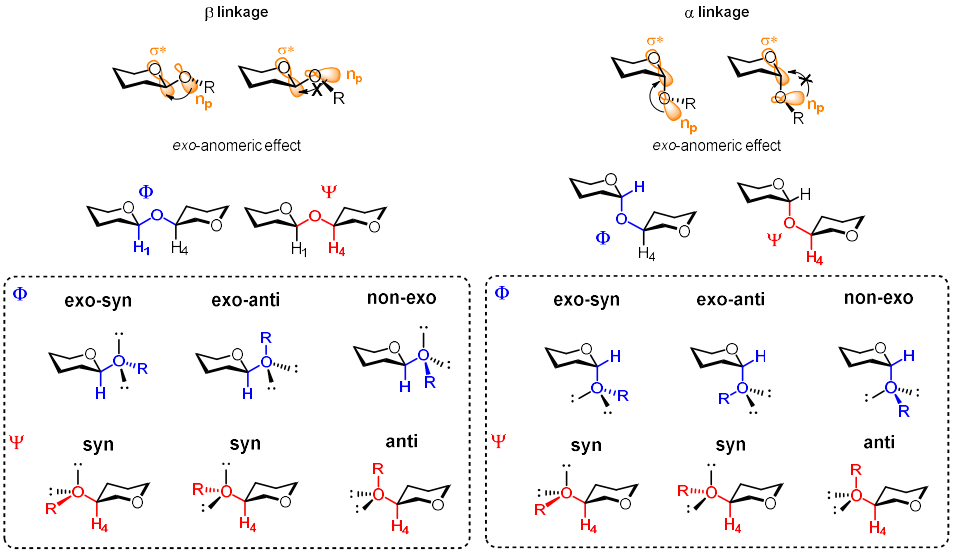


Figure 1.3 Exo-anomeric effect on α and β glycosidic linkages (*top*). Standard definition of dihedral angles used for the description of a glycosidic bond exemplified for a β-1,4 glycosidic linkage (*left*) and an α-1,4 glycosidic linkage (*right*).

One of the most powerful diffraction techniques used to study polysaccharide crystal structures is fiber (or film) X-ray diffraction (XRD) analysis. Due to their nature, sufficiently large single crystals of polysaccharides are hard to obtain, whereas it is possible to prepare oriented fibers (i.e. all the helical axes are parallel to each other). The fiber diffraction pattern, aided by modelling, can give information on the helix type (i.e. pitch and symmetry), on the possible formation of intertwining helices, as well as on the relative orientation (i.e. parallel/antiparallel).[49,50] A complementary technique is selected area electron diffraction (ED), which allows for the measurement of a diffraction pattern of micro-sized crystals (**Figure 1.4**A).[51] A related method is microcrystal electron diffraction (MicroED), a recently developed technique that involves the continuous recording of the electron diffraction pattern while rotating the specimen at cryogenic temperatures.[51–53] This method allows for the acquisition of ED patterns with different crystal orientations to determine the crystal structure. MicroED represents a promising technique for studying self-assembled crystalline materials in their native state, avoiding tedious crystallization trials. Nevertheless,

these methods have an intrinsic limitation since they can only reveal the conformation in solid phase and they require samples in crystal or fiber form.

Among the imaging techniques, Atomic Force Microscopy (AFM) has been applied to directly visualize single polysaccharide molecules, with the significant advantage to operate in "near-native" conditions.[54–57] With this technique, it has been possible to follow secondary and tertiary structure formation of carrageenans, as a function of time or salt concentration.[39,40,58–60] However, a general AFM method to obtain high-resolution images applicable to other classes of polysaccharides is not available to date. Similarly, cryoEM methods are not yet capable of capturing the fine details of glycan conformations, despite the recent advancements.[61] Recently, a single molecule imaging technique has been developed, combining controlled deposition and scanning tunneling microscopy (STM) (**Figure 1.4**B). This technique involves a first step of deposition of mass-selected oligosaccharide ions ($[M-H]^-$) on a Cu(100) surface held at 120 K using the electrospray ion-beam deposition (ES-IBD) method.[62] In a second step, observation of the single oligosaccharide chains is performed by scanning tunneling microscopy (STM) at 11 K.[63] The STM imaging permits subnanometer resolution and direct visualization of the single residues. Although this method is not capable of studying glycans in solution, resolution improvements and image analysis have great potential to provide information on the glycosidic linkage conformation.[64] A major limitation of this method is that it is limited to the study of the conformation in gas phase and interaction with solvent molecules is thus neglected.

NMR is often the technique of choice to gain structural information of glycans in solution.[65–67] With NMR, it is possible to estimate inter-proton distances by measuring peak intensities observed in nuclear Overhauser effect (NOE) experiments (**Figure 1.4**C). The NOE originates from a direct through-space dipolar interaction occurring between two nuclei. However, NOE intensity decreases quite rapidly with the inter-proton distance ($NOE \propto r^{-6}$) and distance measurements are possible only within about 5 Å. When NOE enhancements are close to zero, due to the unfavorable tumbling regime of the glycan ($\omega_0\tau_c \approx 1$ for MW of ca. 1000-2000 Da), rotating frame nuclear Overhauser effect (ROE) experiments can be used.[68] In such experiments, a spin-lock sequence is applied allowing NOE enhancements to be always positive across the entire tumbling regime.[68] However, this experimental advantage comes with the drawback of the large number of artifacts. Other methods, such as the residual dipolar coupling (RDC) analysis, measurements of *J*-couplings across *O*-glycosidic linkages, and hydrogen bond studies are useful complementary tools.[65,69–74] The most important drawbacks of glycan NMR analysis are i) conformational averaging and ii) spectral overlap. Compared to proteins, glycans are flexible and can rapidly interconvert between multiple conformational states complicating the NMR analysis, typically performed over a larger time scale. As a consequence, the different conformational states of a glycan in solution are averaged during NMR experiments. Another limitation is the severe chemical shift degeneracy often observed in ^{1}H NMR spectra. To overcome this issue, the introduction of isotopic labels[69–71,75] or paramagnetic lanthanide complexes[76,77] can be exploited. However, generating a 3-D structure solely from NMR data is rather difficult and molecular dynamics (MD) simulations are often employed as a complementary tool to guide structural analysis.[44,70,78–80] Circular dichroism (CD) is another spectroscopic method which can be used for conformational analysis. CD is well established in peptide science, providing a characteristic fingerprint for each secondary structure. However, the absence of chromophores in polysaccharides limited the use of this technique. Chromophores in polysaccharides can be located in the sidechain and might not be informative of the backbone conformation.[81] An alternative for polysaccharides, lacking chromophores with absorption above 190 nm, is vacuum ultraviolet circular dichroism (VUCD). VUCD enables conformational studies, exploiting higher energy electronic transitions, typical of the ether chromophore. Nevertheless, the need for intense vacuum UV sources, complex instrumentation (i.e. synchrotron radiation), and the rather challenging results interpretation limited the use of this technique.[82]

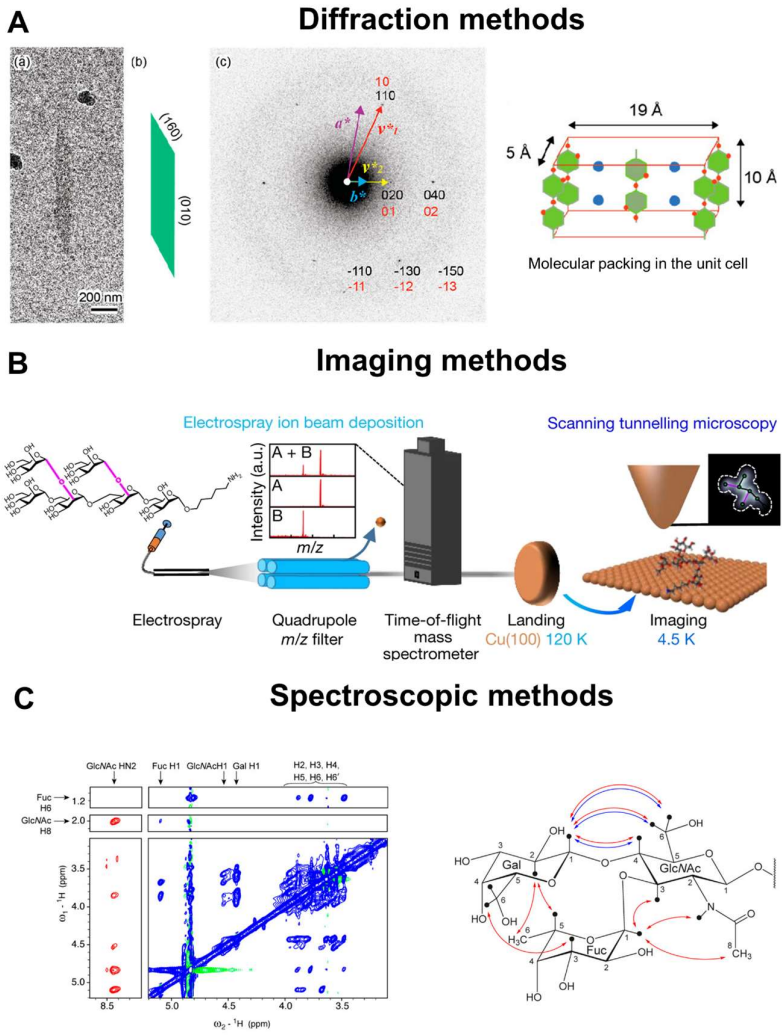


Figure 1.4 A) TEM image and selected area electron diffraction pattern of a xylan nanocrystal (*left*). Molecular packing of the xylan chains in the unit cell (*right*). Adapted with permission from *Biomacromolecules* **2021**, 22, 898-906 Copyright 2021 American Chemical Society. B) Schematic workflow of glycan transfer and deposition via ES-IBD combined with single-glycan STM imaging. Adapted with permission from *Nature* **2020**, 582, 375-378 Copyright 2020 Nature Publishing Group. C) 2D NOESY spectrum allows the identification of cross peaks (*left*). Schematic overview of inter-residue NOEs (*right, arrows*). Adapted with permission from *J. Am. Chem. Soc.* **2013**, 135, 13464-13472 Copyright 2013 American Chemical Society.

1.3 Synthesis of oligo- and polysaccharides

Synthetic oligosaccharides are essential tools to establish structure-property correlations by facilitating the structural analysis.[83] Synthesis offers the opportunity to access poly- and oligosaccharides with well-defined composition, length, and substitution. In particular, the generation of labelled oligosaccharides (e.g. ^{19}F, ^{13}C) could enormously help NMR structural analysis.[66] Nevertheless, the complexity and diversity of poly- and oligosaccharides makes the synthetic process laborious and time consuming. Several aspects are crucial to plan a successful poly- or oligosaccharide synthesis. Properly designed starting materials and/or catalysts are required to ensure regio- and stereocontrol during glycosidic bond formation. Control over the polysaccharide length (degree of polymerization, DP) and regioselective insertion of modifications or branches are additional challenges. Two main approaches are available for obtaining well-defined structures (**Figure 1.5**): i) enzymatic[84–91] and ii) chemical synthesis. Another possible approach is chemical polymerization,[92] however this method lacks control over the length and substitution pattern. Combination of methods can also be exploited to overcome limitations of the individual methods. An example is chemoenzymatic methods in which synthetic BBs are used to direct enzymatic transformations on the oligosaccharide.[93]

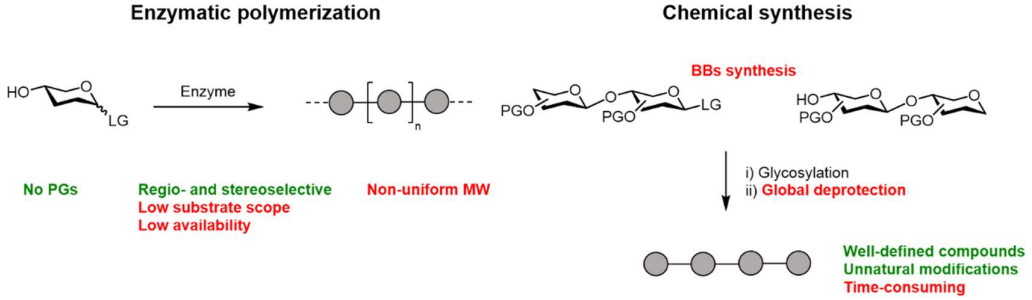


Figure 1.5 Enzymatic (*left*) and chemical (*right*) approaches for the synthesis of polysaccharides. For each method, advantages (*green*) and disadvantages (*red*) are highlighted.

The use of enzymes has undeniable advantages because it offers the possibility to use unprotected sugars as substrates and guarantees remarkable control of the regio- and stereoselectivity during glycosylation. Mono- or oligosaccharides bearing a reactive LG (e.g. phosphate, fluoride, nucleotide) are polymerized by the enzyme to form the desired oligo- or polysaccharide (**Figure 1.5**, *left*). Several classes of enzymes are available, including hydrolases, phosphorylases, sucrases, glycosyltransferases, and glycosynthases.[85,94–97] Despite the numerous advantages of this approach, limited enzyme availability as well as their high specificity narrowed the substrate scope. Generally, enzyme reactive sites are highly specific and tolerate only small modifications, hampering the formation of non-natural polymers.[98] Low glycosylation yields and product hydrolysis represent in some cases additional hurdles associated with enzymatic synthesis of polysaccharides.[99] With this approach, homopolymers are often obtained as non-uniform samples, because the enzymes cannot distinguish between acceptors with different lengths in the reaction mixture.

Chemical synthesis provides compounds with well-defined length and substitution pattern (**Figure 1.5**, *right*), but requires substantial synthetic effort. For this reason, only few examples of long polysaccharides prepared by chemical synthesis are available. Properly designed building blocks (BBs) are needed, often prepared following numerous synthetic steps. In general, BBs are equipped with a reactive anomeric leaving group (LG) to allow for glycosylation and suitable PGs to ensure regio- and stereocontrol.[100] Even though, in most cases, BB preparation follows straightforward protection/deprotection strategies, the low selectivity and yield of certain transformations[101,102] can limit the scope of this approach. The desired polysaccharide is assembled following a linear or a convergent approach. In the former, the desired BBs are added sequentially to the growing polysaccharide chain *via* a series of glycosylation reactions. In contrast, a convergent approach (also known as fragment coupling) allows connecting pre-assembled oligosaccharide

blocks. To decrease the synthetic time required for the chemical synthesis of polysaccharides, automated solid-phase supported or solution-based techniques have been developed.[103–106]

Automated Glycan Assembly (AGA)[107] connects monosaccharide BBs on a solid support following a linear approach (**Figure 1.6**). The solid support is a polystyrene-based Merrifield resin equipped with a cleavable linker.[108] Cycles of glycosylation and selective deprotection are iteratively performed to access the desired oligo- or polysaccharide with full control over the length and the monosaccharide sequence.[107,108] Excess reagents and byproducts can be removed by filtration while the desired product remains anchored to the solid support. The AGA cycle involves i) acidic wash to prepare the resin, ii) coupling step with a glycosyl donor and an activator, iii) capping step (i.e. acetylation) to block unreacted acceptor, and iv) temporary protecting group cleavage to liberate the acceptor for the subsequent coupling step (**Figure 1.6**). When constructing branched structures, orthogonal temporary protecting groups are used (**Figure 1.6**, *bottom*). Washing steps with different solvents are included in the AGA cycle to ensure removal of excess reagents and preparation of the solid support for the next reaction. The whole process is automated and controlled *via* a computer. Upon completion of the assembly, the desired product is released from the solid support and subjected to global deprotection to remove all remaining permanent protecting groups. The reaction used to cleave the desired product from the solid support depends on the linker used. Currently, the most widely used linker consists of a photosensitive methyl-6-nitroveratryl (MeNV) or *o*-nitrobenzyl (oNB) structure that can be cleaved upon irradiation with UV light (366 nm, Hg lamp) in a flow photoreactor (**Figure 1.6**, *bottom left*).[109] AGA offers the possibility to assemble collections of related, well-defined compounds in a short amount of time. Recently, large polysaccharides (i.e. 100mer and 151mer) could be accessed using AGA.[110] Another significant advantage of AGA (and chemical synthesis in general) is the possibility to incorporate Non-natural monosaccharides in a site selective manner. Non-natural residues such as isotopically labelled residues, can simplify structural NMR analysis, opening the way to structure-function correlations (see Section 1.2).

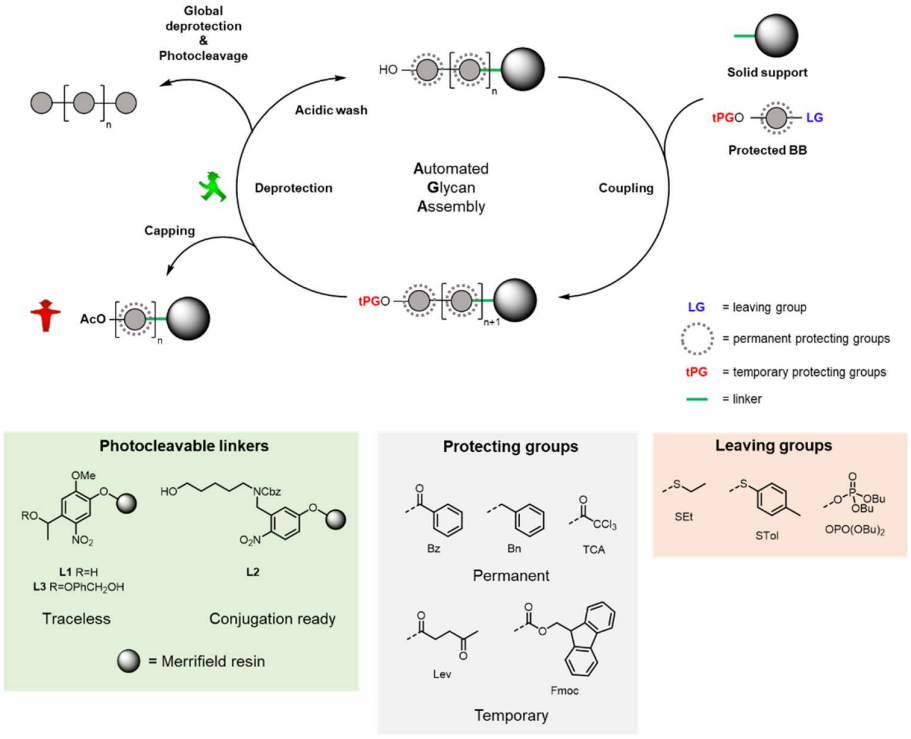


Figure 1.6 Synthetic cycle of AGA exemplified for a linear structure (*top*). Photocleavable linkers, protecting, and leaving groups used in this thesis (*bottom*).

1.4 Non-natural synthetic glycans

1.4.1 The inspiration: non-natural amino acids

In peptide and protein science, the incorporation of non-canonical amino acids (ncAAs) allows for the alteration of the properties of the peptide. Perturbing the primary sequence with ncAAs can serve to understand the role of each residue for the function of a protein.[111] Protein engineering with ncAAs expands the functions normally accessible with the canonical twenty amino acids.[112] For instance, including ncAAs into proteins allowed manipulation of photophysical properties, incorporation of biorthogonal handles (e.g. azide, alkyne)[113] for further functionalization, stabilization of folding, tuning of protein-protein interactions, and offered tools to study protein structure through NMR (e.g. isotopic labelling).[114]

Incorporation of fluorinated ncAAs has drawn considerable attention in the past decades.[115] Fluorine is an element with unique properties rarely occurring in natural products. Fluorine is a small and highly electronegative atom and the C-F bond is chemical inert due to its extreme stability (105.4 kcal/mol).[116] The insertion of fluorine in a molecule is often accompanied by dramatic physicochemical changes (e.g. pK_a, solubility, *in vivo* stability, intermolecular interactions modulation). When multiple fluorine atoms are installed in a molecule its hydrophobicity is enhanced. Therefore, the incorporation of fluorinated ncAAs became a standard strategy to modify peptides and proteins with two major purposes: i) the alteration of physicochemical properties and ii) the use of ^{19}F atoms as reporter groups for NMR studies.[112,115,117,118] Fluorinated ncAAs can be used to generate hyperstable protein folds through subtle mechanisms, as demonstrated for fluorinated collagen analogs. Collagen single-chains self-assemble into tightly packed triple helices that arrange into fibrils of exceptional stability.[2] The replacement of proline with 4*R*-fluoroproline in a collagen model peptide significantly raised the melting temperature of the triple helix. Fluorination of proline altered the puckering of the five-membered ring favoring a pre-organization of the single chains in a conformation suitable to form the triple helices (**Figure 1.7**A).[119,120] Several examples exist that used fluorinated ncAAs to impact peptide conformation and hydrophobicity or to direct highly specific protein-protein interactions through the "fluorous effect".[112,117,118] The fluorous effect has been exploited also to direct the self-assembly of peptide materials.[121] An additional advantage of the introduction of fluorinated ncAAs is the possibility to carry out ^{19}F NMR experiments to monitor binding events, protein (mis)folding, protein-membrane interactions, and protein aggregation (**Figure 1.7**B).[122–124] The use of the ^{19}F NMR channel enables real-time monitoring during *in vitro* experiments as opposed to traditional structural studies involving X-ray crystallography. The absence of background signal of biological samples in the ^{19}F NMR channel allows also to carry out *in cell* NMR experiments on labelled peptides and proteins.[125]

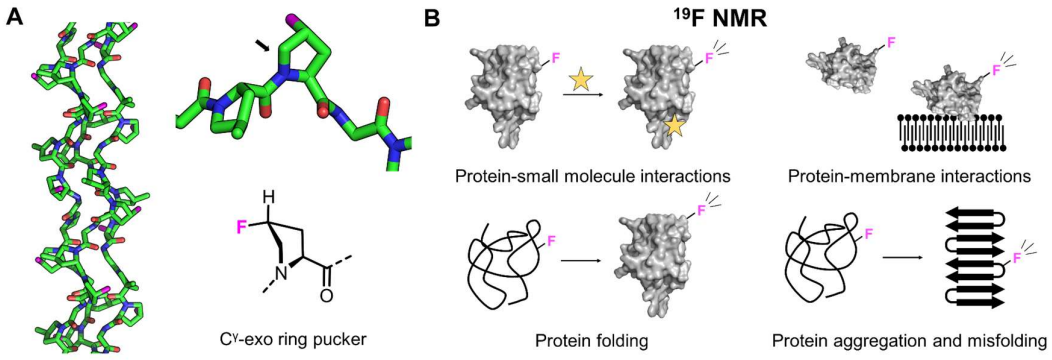


Figure 1.7 A) Structure of a triple helical peptide containing 4*R*-methylproline and 4*R*-fluoroproline (PDB 3IPN). The 4*R*-fluoroproline preferentially adopts an *exo* ring pucker. B) General overview of the applications of ^{19}F-labelled ncAAs for ^{19}F NMR studies of proteins.

Besides fluorinated ncAAs, mirror-image D-amino acids are also a widely explored class of ncAAs employed in the modification for peptide and proteins. D-amino acids have been incorporated into peptides and proteins for many purposes such as enhancing proteolytic[126] and thermal stability,[127] to explore the conformational space of heterochiral proteins,[128] or to facilitate crystallization (i.e. racemic crystallization).[129] Furthermore, mirror image amino acids found many applications in the study of self-assembling peptides and contributed to the elucidation of the origin of supramolecular chirality of amyloid fibrils.[130–132] Synthetic peptides made of L- and D- amino acid sequences suggested that amino acids located at the termini were responsible for the twist adopted by peptide supramolecular fibers.[8,22,133,134] Chirality can also be used as a design tool to control morphology[135,136] and mechanical properties[137–139] of self-assembled peptide materials. For instance, it was found that co-assembling L- and D-peptides into fibrils (**Figure 1.8**A and B) leads to enhanced mechanical rigidity of the resulting hydrogel due to unique favorable interactions taking place between the two enantiomers (**Figure 1.8**C).[137,138] Inter-residue hydrophobic interactions are maximized between the different layers of the peptide in the interior of a fibril leading to higher stiffness. This example demonstrates that mirror image D-amino acids have great potential to tune bulk mechanical properties of self-assembling peptides using a *bottom-up* approach.

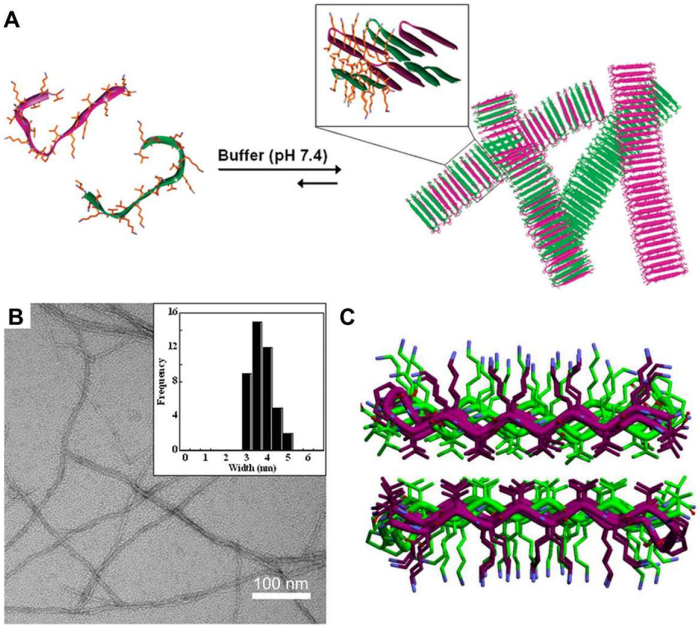


Figure 1.8 A) Assembly mechanism for enantiomeric peptides leading to the formation of a network of fibrils. B) TEM image of fibrils obtained from the racemic gel. C) Model of the two enantiomers in their co-assembled, racemic fibrillary state. Adapted with permission from *J. Am. Chem. Soc.* **2011**, 133, 14975-14977 Copyright 2011 American Chemical Society and *ACS Cent. Sci.* **2017**, 3, 586-597 Copyright 2017 American Chemical Society.

1.4.2 Non-natural monosaccharides

The use of non-natural monosaccharides has been explored mainly in the context of glycomimetics, to improve the performance of natural glycans as drugs (e.g. bioavailability, affinity for the target, enzymatic stability).[140] Modification of monosaccharides typically involve i) functional group modification, ii) replacement of the endocyclic oxygen with other atoms (i.e. C, S, or N), and iii) modification of the glycosidic linkage. Still, to date, the incorporation of non-natural monosaccharides into complex and large oligosaccharides is much less explored as opposed to the use of ncAAs in peptides and proteins.

Replacing C-OH groups with C-F can be used as a strategy to tune hydrogen bonding and interactions with water molecules with potential impact on the conformation, the aggregation propensity, and the interactions with proteins. However, limited studies have been carried out with fluorinated glycans, since the installation of fluorine into an oligosaccharide backbone can be challenging.[141] To date, the site-specific installation of

fluorine proved to be a successful strategy to increase the *in vivo* stability of glycoproteins,[142] to generate enzyme inhibitors,[143] to enhance immunogenicity,[144,145] to selectively disrupt protein-glycan hydrogen bonds,[146] or to enhance CH-π interactions.[147] Furthermore, simple fluorinated glycans have been used as NMR probes to detect binding with proteins or for structural studies.[148] However, fluorinated monosaccharides have not been incorporated in complex glycans nor have they been exploited to rationally alter conformational properties or inter-molecular interactions. The synthesis remains a major hurdle, suggesting that the use of automated synthetic techniques, such as AGA, could be explored to quickly access to collection of oligosaccharides that incorporate fluorinated monosaccharide building blocks.

Other modifications of monosaccharide include the replacement of the endocyclic oxygen with non-natural functionalities, typically used to increase oligosaccharide stability towards enzymatic degradation. The endocyclic oxygen has been replaced with a methylene moiety (i.e. carbasugars), with sulfur (i.e. thiasugars), or with a CF_2 moiety. Although with some differences, these synthetic analogs are generally increasing the flexibility of the glycosidic linkage.[149–151] In addition, modification of the glycosidic linkages by replacing the oxygen with sulfur (i.e. *S*-glycosides) or carbon (*C*-glycosides) are other strategies to improve enzymatic stability, resulting often in increased flexibility of the oligosaccharides.[140,152]

The exploitation of non-natural monosaccharides in longer and complex glycans remains, to date, largely underexplored. The recent advancements of automated platforms and synthetic protocols allowed to access increasingly complex and large polysaccharides,[104,110,153–155] paving the way for the application of non-natural monosaccharides.

1.5 Rational design of synthetic materials

The design of synthetic peptide materials was fostered by the wide availability of automated solid-phase peptide synthesis (SPPS) devices, which gave rapid access to broad collections of well-defined peptides for systematic conformational analyses. Determination of folding propensity of single amino acids[156] and elucidation of the factors that influence α-helix and β-sheet stability[157] were key elements in establishing design principles to create peptides adopting a specific conformation or that self-assemble.[158] In the 1990s, the design and the synthesis of a peptide autonomously folding into a triple-stranded β-sheet was an enormous breakthrough.[157,159,160] Nowadays, entire proteins can be designed to obtain diverse folded conformations and novel functions (i.e. protein *de novo* design).[34] Similarly, the development of DNA nanotechnology was fostered by a thorough understanding of nucleotide base pairing and DNA conformation combined with automated synthetic[161] and amplification (i.e. PCR) methods.[18] Nowadays, DNA sequences can be designed to self-assemble in a wide variety of nanostructures (i.e. DNA origami, see Section 1.1).

ncAAs enabled the exploration of new primary sequences folding into completely new conformations, beyond natural ones. Artificial folded molecules inspired by the structure and function of biopolymers were developed and termed "foldamers".[162–164] Foldamers adopt highly predictable and stable secondary structures with distinct size and shape (**Figure 1.9**A).[163,165] Given the diversity of accessible 3-D shape, foldamers could be used to mimic or expand the properties and functions of natural biopolymers. Foldamers can be designed with conformations that promote specific binding to proteins or small molecules with applications in pharmacology; to perform catalytic transformations mimicking enzymes;[166] or to self-assemble into supramolecular aggregates.[167] Recently, a helical foldamer was employed to catalyze a macrocyclization reaction by providing ideal orientation of two catalytic groups (**Figure 1.9**B).[168] In this seminal study, the modularity of the foldamer scaffold allowed an exploration of different reactive groups and spatial orientation to find an optimal arrangement.[169]

The generation of foldamers, heavily relying on ncAAs, and the demonstration that these scaffolds can accomplish functions beyond the ones found in natural biopolymers, is a solid demonstration of their potential in catalysis, medicinal chemistry, or material science. I imagined that a similar approach to carbohydrates can generate even more structural diversity and further expand foldamer applications.

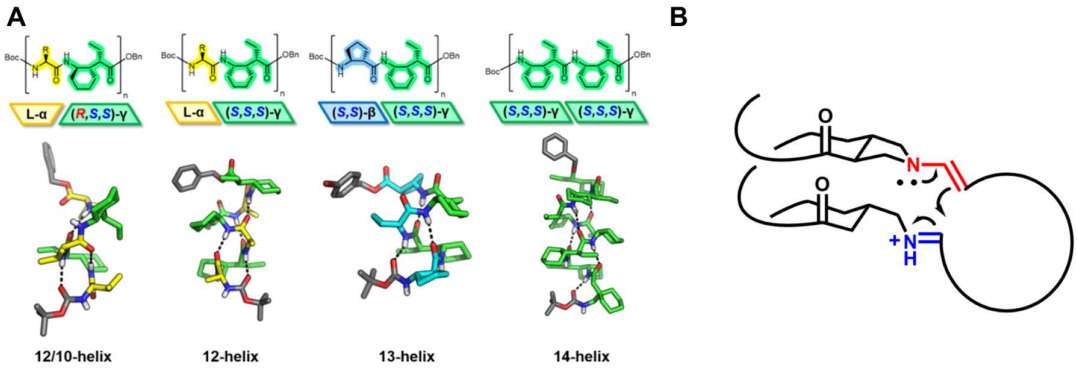


Figure 1.9 A) Foldamer backbones (*top*) consisting of conformationally constrained α-, β-, and γ-amino acids that generate helical folded structures with diverse size and shape (*bottom*). B) Foldamer-catalyzed macrocyclization. Reproduced from www.gellman.chem.wisc.edu.

1.6 Aims of this thesis

The general aim of my doctoral research was to establish structure-property correlations for oligosaccharides and then translate this knowledge into design principles to access self-assembling carbohydrate materials. In contrast to other biopolymers (i.e. peptides, DNA), carbohydrates are less understood at the molecular level and therefore have rarely been employed as scaffolds to construct self-assembling materials using *bottom-up* approaches. Thus, to date, carbohydrates' potential in supramolecular chemistry remains mostly untapped. In this doctoral thesis, I present the use of non-natural monosaccharides to understand structure and function of oligosaccharides. Furthermore, I suggest a rational *bottom-up* approach to understand and design self-assembling carbohydrate materials.

The aim of my first project (Chapter 2) was to understand the impact of single-site fluorination in oligosaccharides and exploit the fluorine atom as an NMR reporter to study oligosaccharides, from a structural to a functional perspective. The project required the implementation of the following steps:

1. Design and optimization of the synthesis of fluorinated monosaccharide BBs
2. Synthesis of a broad collection of complex fluorinated glycans
3. Evaluation of the impact of fluorination on the oligosaccharide conformation *via* MD simulations and single molecule imaging (in collaboration with Dr. Theodore Tyrikos-Ergas and Dr. Kelvin Anggara)
4. Evaluation of the crystallization propensity of the fluorinated glycans *via* XRD
5. Use of ^{19}F-labelled BBs to study the solution conformation of oligosaccharides (in collaboration with Dr. Ana Poveda)
6. Study of ^{19}F-labelled glycans as probes for the screening of protein-glycan interactions (in collaboration with Dr. Elena Shanina)

The aim of my second project (Chapter 3) was to develop a *bottom-up* approach to understand the self-assembly of carbohydrate materials. In particular, the main objective was to understand the origin of the supramolecular chirality found in cellulose fibrils, yet not fully understood at the molecular level. To this end, I had to implement the following steps:

1. Synthesis of non-natural mirror-image L-Glc BBs
2. Synthesis of site-specific modified disaccharides as model systems to understand the rules that govern oligosaccharide aggregations
3. Study of the self-assembly of simple disaccharide model systems (in collaboration with Dr. Soeun Gim and Dr. Yu Ogawa)
4. Preparation of a collection of well-defined cellulose oligomers
5. Study of the self-assembly properties of the synthetic cellulose oligomers through XRD, AFM, TEM, and electron diffraction (in collaboration with Dr. Yu Ogawa)
6. Design and perform experiments to understand the origin of chirality and correlate oligosaccharide sequence with supramolecular chirality

In my last project (Chapter 4), the aim was to create a glycan secondary structure that does not exist in nature. In particular, the objective was to demonstrate the feasibility to program glycans to adopt a defined conformation in aqueous solution. To this end, I had to implement the following steps:

1. Establish principles to rationally design a glycan secondary structure
2. Identify target structures with the aid of MD simulations (in collaboration with Dr. Theodore Tyrikos-Ergas)
3. Design and synthesize suitable BBs for the synthesis of the target structures
4. Optimize the synthesis of the target oligosaccharides
5. Study the conformation of the oligosaccharides *via* NMR (in collaboration with Dr. Ana Poveda)

2 Fluorinated glycans as tools to study carbohydrate structures and functions

This chapter has been modified in part from the following articles:

Y. Yu, T. Tyrikos-Ergas, Y. Zhu, **G. Fittolani**, V. Bordoni, A. Singhal, R.J. Fair, A. Grafmüller, P.H. Seeberger, M. Delbianco Systematic Hydrogen-Bond Manipulations to Establish Polysaccharide Structure-Property Correlations *Angew. Chem., Int. Ed.* **2019**, 58, 37, 13127-13132 https://doi.org/10.1002/ange.201906577

K. Anggara, Y. Zhu, **G. Fittolani**, Y. Yu, T. Tyrikos-Ergaas, M. Delbianco, S. Rauschenbach, S. Abb, P.H. Seeberger, K. Kern Identifying the Origin of Local Flexibility in a Carbohydrate Polymer *Proc. Natl. Acad. Sci.* **2021**, 118, 23, e2102168118 https://doi.org/10.1073/pnas.2102168118

A. Poveda, **G. Fittolani**, P. H. Seeberger, M. Delbianco, J. Jiménez-Barbero The Flexibility of Oligosaccharides Unveiled Through Residual Dipolar Coupling Analysis *Front. Mol. Biosci.* **2021**, 8, 784318 https://doi.org/10.3389/fmolb.2021.784318

G. Fittolani,[+] E. Shanina,[+] M. Guberman, P.H. Seeberger, C. Rademacher, M. Delbianco Automated Glycan Assembly of ^{19}F-labeled Glycan Probes Enables High-Throughput NMR studies of Protein-Glycan Interactions *Angew. Chem. Int. Ed.* **2021**, 60, 24, 13302-13309 https://doi.org/10.1002/anie.202102690

[+]equal contribution

Specific contribution

Section 2.2.1 I performed the optimization of the building blocks synthesis.

Section 2.2.2 I performed synthesis, XRD and NMR analysis. Dr. Theodore Tyrikos-Ergas and Surusch Djalali performed the MD simulations. Dr. Kelvin Anggara performed the STM imaging.

Section 2.2.3 I performed the synthesis of the fluorinated analogs and Dr. Ana Poveda performed the NMR analysis.

Section 2.2.4 I performed the design and the synthesis of all fluorinated probes and Dr. Elena Shanina the NMR analysis.

2.1 Introduction

The intrinsic complexity of polysaccharides poses a bottleneck to the understanding of their structure and function. In order to establish structure-property correlations, tools to manipulate glycans at a molecular level and study their 3-D structure are needed. Well-defined chemical modification and/or introduction of non-natural residues in a glycan give the opportunity to tune the molecular structure. By changing glycans polarity, ability to engage in hydrogen bonds, hydrophobicity, and solubility it is possible to also alter their macroscopic properties. A systematic study of the effect of such manipulations allows to correlate molecular structure with the properties of the natural counterpart.

Fluorine incorporation is widely explored in medicinal and material chemistry.[116,170,171] The features of fluorine offer unique advantages to tune physicochemical properties of molecules. Fluorine is a small, highly electronegative atom and the C-F bond is chemically inert due to its extreme stability (105.4 kcal/mol).[116] The C-F bond has been used to replace functional groups such as C-OH, C-H, and C=O. The installation of fluorine atoms can alter the metabolic stability of a drug or inhibit a certain metabolic pathway.[170] Despite the high polarization of the C-F bond, hydrogen bonds involving the fluorine atom are rare and thus deoxyfluorination can be used to manipulate hydrogen bonds. In addition, when multiple fluorine atoms are installed in a molecule its hydrophobicity is enhanced. Inspired by peptide and protein science, in which fluorinated amino acids have been used to tune intra- and intermolecular interactions (**Figure 2.1**A, see also Section 1.4.1),[115,117,121,172] I explored the properties of fluorinated glycans. Fluorination can tune hydrogen bonding and interactions with water molecules, suggesting that OH/F substitutions in oligosaccharides can have a profound effect on the glycan conformation. Still, to date, the challenging installation of fluorine into an oligosaccharide backbone (i.e. multistep syntheses, stereo-control required, harsh conditions)[141] has prevented broad structural studies using fluorinated glycans.

The presence of fluorine also enables the use of the ^{19}F NMR channel. Among all NMR-active nuclei, the ^{19}F nucleus stands out due to its unique properties such as: i) high sensitivity to local chemical environment, ii) short acquisition times, iii) simple spectra, iv) broad chemical shift range, and v) absence in biological systems (no background signal).[123,173] Thus, fluorinated amino acids have been used as NMR reporters in peptide and protein science.[122,123] For example, a ^{19}F reporter installed on a protein can provide a real-time readout of ligand binding (**Figure 2.1**B, see also Section 1.4.1).[174] Similar studies can be imagined with ^{19}F-labelled glycans, [143,175–177] with the additional advantage that labelled analogs overcome the chemical shift degeneracy of glycans. ^{19}F-labelled glycans have the potential to dissect protein-glycan interactions,[178,179] and simplify conformational studies.[66,180] Still, the labour-intensive multistep synthesis of ^{19}F labelled glycans has limited these studies to small collections of short and relatively simple glycans.[148,181–184] Lengthy synthetic protocols have hampered the synthesis of larger and more complex synthetic fluorinated glycans.

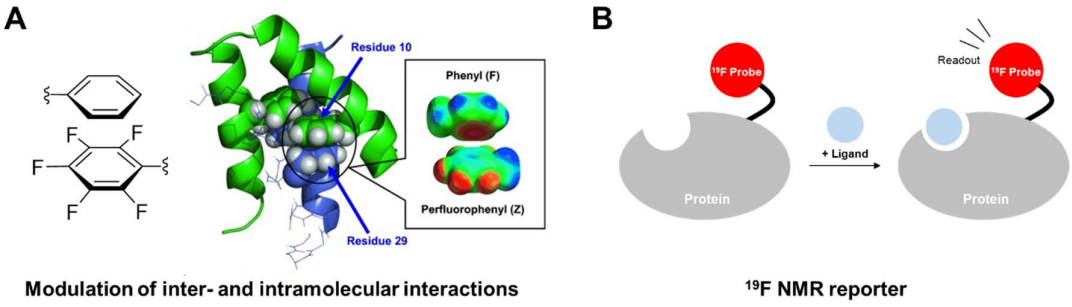


Figure 2.1 A) Fluorinated phenylalanine enables modulation of the strength of inter- and intramolecular stacking interactions. Reprinted with permission from *Angew. Chem. Int. Ed.* **2010**, 49, 8635-8639 Copyright 2010 Wiley-VCH. B) Schematic illustration of the detection of binding using a ^{19}F reporter installed on a protein. Adapted from *Nat. Chem. Biol.* **2009**, 5, 341-343.

In the first part of this chapter, I developed efficient synthetic routes to fluorinated Glc and GlcNAc BBs (Section 2.2.1). Using these non-natural monosaccharides, I assembled a broad collection of complex fluorinated glycans using AGA.[107] These synthetic compounds were designed to manipulate hydrogen bonds in cellulose, a polysaccharide with high propensity to crystallize, and generated more soluble analogs (Section 2.2.2). High-resolution imaging permitted the visualization of single molecules confirming the higher flexibility predicted by MD. The aggregation of fluorinated cellulose was then investigated by XRD, revealing the subtle role of the F substitution pattern. In the second part of this chapter, I explored the use of fluorinated glycans as NMR probes. A collection of ^{19}F-labelled glycans was exploited to confirm the helical conformation of a β-1,6 Glc hexasaccharide *via* residual dipolar coupling (RDC) analysis (Section 2.2.3). Lastly, I prepared ^{19}F-labelled Lewis antigens, a class of biologically relevant glycans, to study their binding to proteins (Section 2.2.4). Protein-glycan interactions were screened in high-throughput and enzymatic reactions were followed in real-time through the ^{19}F NMR channel.

2.2 Results and discussion

2.2.1 Synthesis of fluorinated building blocks

Cellulose is a linear polysaccharide consisting of β-1,4 linked Glc. In cellulose chains, Glc hydroxyl groups are all oriented equatorially and have a strong tendency to engage into a dense network of hydrogen bonds (**Figure 2.2**).[185–187] The propensity of cellulose to crystallize is largely dependent on these hydrogen bonds with additional contributions from other non-covalent interactions such as the hydrophobic interactions occurring between the C-H rich, apolar faces of the Glc units.[188,189] I designed non-natural cellulose analogues to explore the effect of the substitution of specific OH with F atoms on the resulting material. By placing a fluorine at the C-3 of Glc, I aimed at disrupting the key intramolecular hydrogen bond occurring between the OH-3 and the O-5 of the following residue in the cellobiose repeating unit (**Figure 2.2**, *left*). In contrast, placing a fluorine at the C-6 of Glc should disrupt other hydrogen bonds, such as the one between OH-6 and O-2 of the subsequent Glc residue (**Figure 2.2**, *right*). Both substitution are also expected to interfere with intermolecular hydrogen bonds as well as tune the dipolar momentum of the Glc units due to the strong EWG nature of fluorine. I then applied a similar strategy to get insights into chitin, a linear polysaccharide consisting of β-1,4 linked GlcNAc residues, that shares with cellulose the propensity of aggregating into crystalline domains stabilized by inter- and intramolecular hydrogen bonds. Here, I modified the C-6 position of GlcNAc by OH/F substitution.

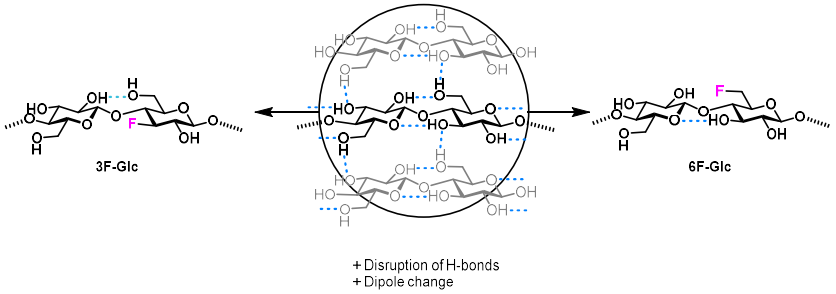


Figure 2.2 The dense network of hydrogen bonds of cellulose can be manipulated by introducing deoxyfluorinated Glc units.

The synthetic cellulose and chitin analogues will be prepared by AGA, requiring the preparation of properly designed fluorinated BBs. To install fluorinated Glc and GlcNAc residues within the β-1,4 cellulose or chitin backbone, I designed and performed the synthesis of **BB2.1**, **BB2.3** and **BB2.4** (**Figure 2.3**). Additionally, I prepared **BB2.2** to construct β-1,6 Glc linkages (see Section 2.2.3). To be compatible with the AGA technology, each BB was equipped with a reactive anomeric leaving group, either a thioether or a phosphate. Permanent protecting groups (Bn, Bz, or TCA) and a temporary protecting group (Fmoc) enabled the regioselective step-wise elongation of β-1,4 glycosidic linkages. β-Stereoselectivity was ensured by participation of the C-2 Bz or *N*-TCA protecting groups.

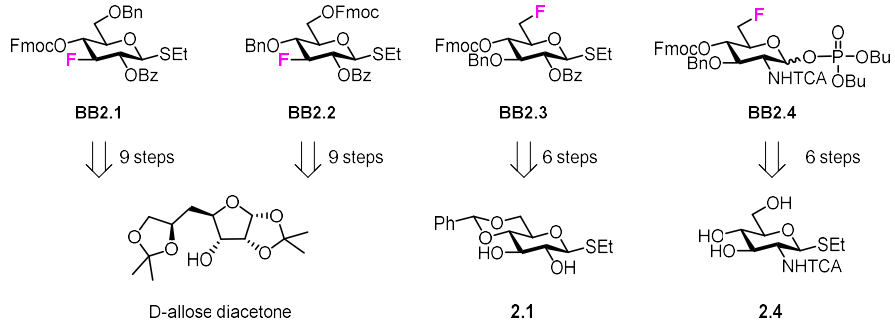
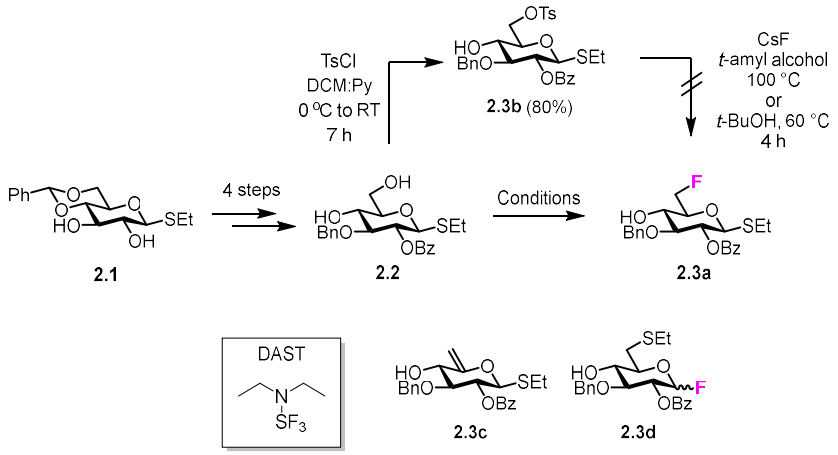


Figure 2.3 Target deoxyfluorinated Glc and GlcNAc BBs.

The syntheses of **BB2.1** and **BB2.2** were accomplished in nine steps starting for D-allose diacetone (performed during my M.Sc. internship preceding the start of my doctorate).[190] In the case of **BB2.3** and **BB2.4**, I devised synthetic routes that start from commercially available starting materials and that minimize the number of synthetic steps required.

First, I performed the synthesis of the 6F-Glc BB (**BB2.3**). To introduce the fluorine moiety, I planned a regioselective C-6 deoxyfluorination of compound **2.2**, readily accessible in four steps from commercially available intermediate **2.1** (**Scheme 2.1**). The first strategy involved the regioselective tosylation of **2.2**,[191] followed by the nucleophilic substitution with CsF on intermediate **2.3b**. However, this route yielded only trace amounts of desired product **2.3a** (**Table 2.1**, entries 1 and 2) together with a significant amount of the elimination side-product **2.3c** (ESI-MS m/z 401.1 $[M+H]^+$ and ^1H NMR ($CDCl_3$) δ 6.02 ppm (d, J = 4.8 Hz), **Scheme 2.1**). Next, I attempted the deoxyfluorination directly on **2.2** using DAST in the presence of pyridine.[192] These conditions yielded trace amount of **2.3a** and decomposition of the starting material (**Table 2.1**, entries 3 and 4). When the reaction was performed in the absence of base, compound **2.3a** was formed as a major product in a modest 34% yield (**Table 2.1**, entry 6 and 7). Performing the reaction at higher temperatures only led to complex mixtures which complicated isolation (**Table 2.1**, entry 5). The low yield of the DAST-mediated deoxyfluorination of thioglycoside **2.2** can be the results of the following side-reactions: i) elimination, to give **2.3c**, is a commonly observed side reaction in nucleophilic subsitutiton[193] using the fluoride ion which can act as a base, ii) a migration reaction involving the anomeric SEt moiety[194] could generate **2.3d** (observed as anomeric fluoride at ca. 130-140 ppm in ^{19}F NMR), and iii) DAST-mediated thioglycoside activation[195–197] could produce reactive species capable of undergoing glycosylation with the nucleophiles present in solution or promote starting material decomposition, as observed when higher temperatures are used. The replacement of the thioglycoside leaving group with a more stable analog (e.g. *O*-based glycoside) could overcome these synthetic hurdles, however this operation would imply many more protecting group manipulations that might result in comparable overall yields.



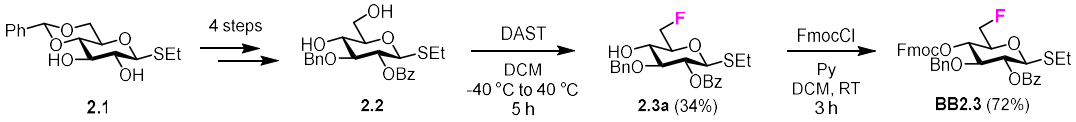
Scheme 2.1 Selective C-6 deoxyfluorination of **2.2** and potentially forming side-products. Refer to **Table 2.1** for the conditions used.

Entry	Reagents	Conditions	Yield for 2.3a	Notes
1	i) TsCl (1.3 equiv.), Py ii) CsF (3 equiv.)	i) DCM, 0 °C to RT, 7 h ii) *t*-amyl alcohol, 100 °C, 1 h	Traces	**2.3c** major product
2	i) TsCl (1.3 equiv.), Py ii) CsF (3 equiv.)	i) DCM, 0 °C to RT, 7 h ii) *t*-BuOH, 60 °C, 1h	Traces	**2.3c** major product
3	DAST (1.2 equiv.)	Py, -20 °C to 70 °C	-	**2.2** decomposition

4	DAST (1.2 equiv), Py (2 equiv.)	DCM, -40 °C to 40°C, 5 h	Traces	*
5	DAST (1.2 equiv.)	DCE, -20 °C to 70 °C	27%	*
6	DAST (1.2 equiv.)	DCM, -40 °C to RT, 5 h	29%	*
7	DAST (1.2 equiv.)	DCM, -40 °C to 40 °C, 5 h	34%	*

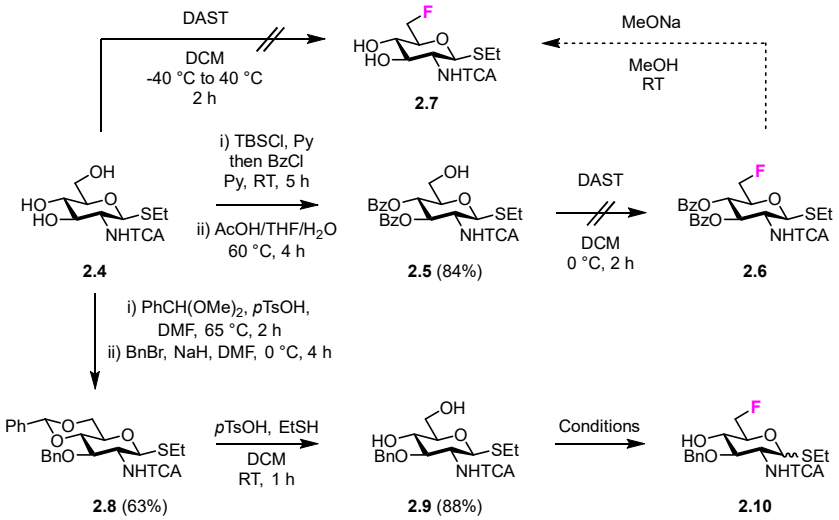
Table 2.1 Screening of conditions for the selective C-6 deoxyfluorination of **2.2**. *Several unidentified side products among which **2.3c** and **2.3d**. Reported are isolated yields.

After the deoxyfluorination step, installation of Fmoc on **2.3a** completed the synthesis of **BB2.3**. In summary, **BB2.3** was synthesized in six steps starting from the commercially available intermediate **2.1** (**Scheme 2.2**).



Scheme 2.2 Synthetic route to **BB2.3**.

Next, I focused on the synthesis of the 6F-GlcNAc **BB2.4**. Initially, I attempted the deoxyfluorination of the primary hydroxyl group using DAST on intermediate **2.4**, bearing three free hydroxyl groups.[198] This strategy gave a complex mixtures of products with only trace amounts of **2.7** (**Scheme 2.3**). In a second attempt, I converted the triol **2.4** into intermediate **2.5** in two steps since an efficient deoxyfluorination procedure was reported for an analogous thioglycoside intermediate.[199] However, the subsequent DAST-mediated deoxyfluorination of **2.5** failed, yielding a complex mixture of side products with no desired product **2.6** (**Scheme 2.3**). Next, inspired by the successful deoxyfluorination of the 3-*O*-Bn intermediate (**2.2**), I converted the triol **2.4** into diol **2.9** in three steps (**Scheme 2.3**). In this case, the DAST-mediated deoxyfluorination provided desired the product **2.10**, albeit in modest yield (**Table 2.2**). The reaction with DAST produced a complex mixture of products and isolation of the desired product proved challenging. Initially, a mixture of DCM and 1,4-dioxane was employed to solubilize the starting material **2.9**, but anomerization[195–197] was observed (**Table 2.2**, entry 1). To simplify purification and characterization, I performed the reaction in ACN. These conditions minimized anomerization,[200] yielding the target compound **2.10** in 23% yield (**Table 2.2**, entry 2). A further improvement of the yield of the deoxyfluorination step could be expected replacing the thioether anomeric leaving group with an *O*-based anomeric LG that should not interact with DAST, however increasing significantly the number of steps necessary.

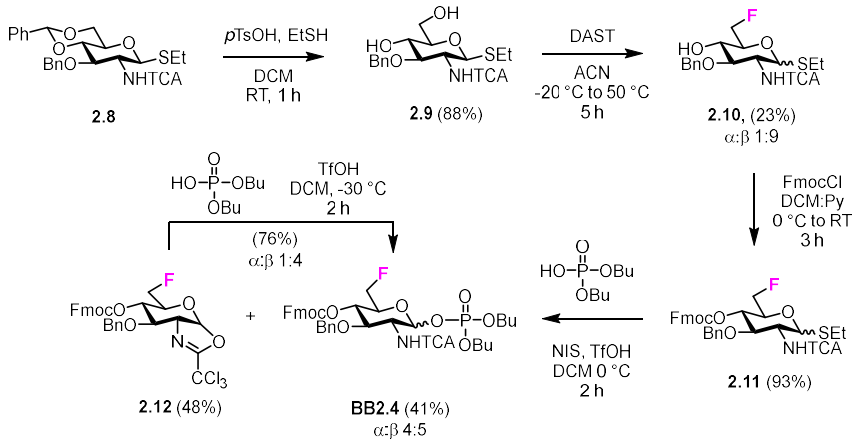


Scheme 2.3 Routes explored to access the C-6 deoxyfluorinated GlcNAc intermediate. For the conversion of **2.9** into **2.10** the conditions are reported in **Table 2.2**.

Entry	Reagents	Conditions	Yield for 2.10
1	DAST (1.1 equiv.)	DCM:1,4-dioxane (1:1), -20 °C to RT, 5 h	26% (α:β 3:2)*
2	DAST (1.1 equiv.)	ACN, -40 °C to 50 °C, 5 h	23% (α:β 4:96)*

Table 2.2 Conditions used for the C-6 deoxyfluorination of intermediate **2.9**. Reported are isolated yields.*Anomeric ratio was measured by integration of the H-1 signals in the ^{1}H NMR.

To complete the synthesis of **BB2.4**, Fmoc was installed and the thioether leaving group exchanged with a dibutyl phosphate one (**Scheme 2.4**). The latter aimed to increase the donor reactivity that could be beneficial when working with unreactive acceptors such as the GlcNAc-OH-4.[200,201] During the installation of the dibutyl phosphate anomeric group, the trichloro oxazoline side product[202] **2.12** was formed alongside to the target compound **BB2.4** (**Scheme 2.4**). An additional step, adapting a reported procedure,[203] permitted the conversion of the trichloro oxazoline into the desired **BB2.4**.



Scheme 2.4 Synthetic route to **BB2.4**.

In summary, **BB2.3** and **BB2.4** were obtained through synthetic routes that minimize overall steps and protecting group manipulations. Further improvements of the deoxyfluorination step could be imagined with new reagents that minimize elimination reaction.[204] It is worth mentioning that the success of the C-6

deoxyfluorination of thioglycoside intermediates seems highly dependent on the protecting group pattern, potentially modulating the rate of the multiple side reactions that can occur.

2.2.2 Deoxyfluorination as tool to manipulate hydrogen bonds

I assembled a collection of cellulose analogs using **BB2.1**, **BB2.3**, and **BB2.5** (**Figure 2.4**) to study the effect of deoxyfluorination and the pattern of substitution on the resulting oligomers. The synthesis of fluorinated cellulose analogs was performed using AGA, following cycles of glycosylation and Fmoc deprotection on solid support **L1** (see Experimental section 6.3.2). Solid-phase removal of Bz esters (Module F1) was followed by cleavage from the solid support (Module G1) and subsequent hydrogenolysis (Module H1 or H2) (for details on AGA modules see Experimental section 6.2). One single final purification step afforded a collection of fluorinated cellulose analogs. The cellulose analogs were assembled with either 3F- or 6F-Glc units. For both analogs, different substitution patterns were explored: alternated (**FAFAFA**), random (**AFAAFA**), block (**FFFAAA**, and **fffAAA**), and fully substituted (**FFFFFF** and **ffffff**) (**Figure 2.4**).

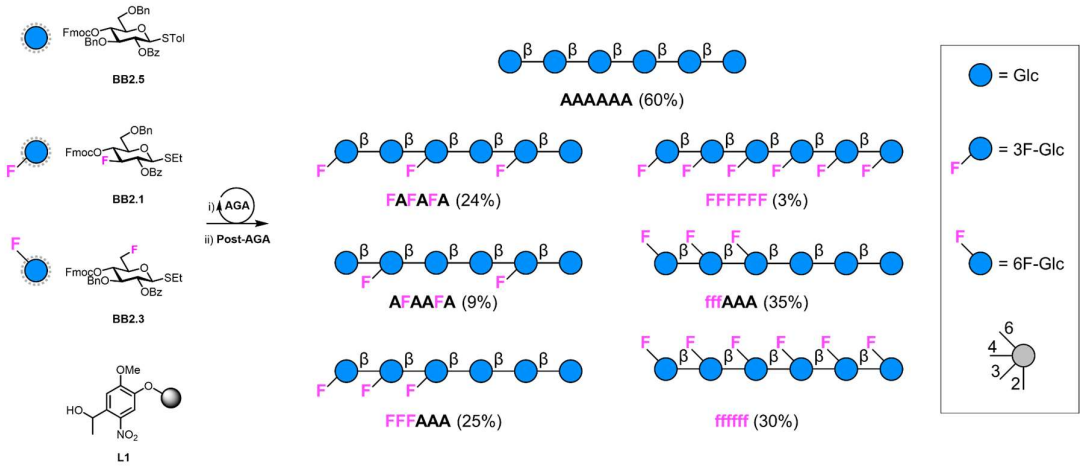


Figure 2.4 AGA of a collection of deoxyfluorinated cellulose analogs. Yields of isolated products after AGA, Post-AGA and purification are reported in parentheses. The Symbol Nomenclature for Glycans (SNFG) was used to represent the oligosaccharides. Capital **F** indicates 3F-Glc while lowercase **f** indicates 6F-Glc.

To understand the impact of deoxyfluorination at a single molecule level, MD simulations were performed employing a modified version of the GLYCAM06 carbohydrate force field (by Dr. Theodore Tyrikos-Ergas).[205,206] The comparison between the unsubstituted **AAAAAA** and the di-substituted **AFAAFA** hexasaccharides showed that both analogs adopted a linear and extended conformation in solution (**Figure 2.5**A). A higher conformational variability was observed for **AFAAFA**, as exemplified in the end-to-end distance plot (**Figure 2.5**B): larger fluctuations and a smaller end-to-end distance average value indicated higher flexibility. To confirm this observation, a recently developed single molecule imaging technique was employed (see Section 1.2).[63,64] The experiments were performed at the Max Planck Institute for Solid State Research by Dr. Kelvin Anggara. The imaging technique involves a first step of deposition of mass-selected oligosaccharide ions ([M-H]$^-$) on a Cu(100) surface held at 120 K using the electrospray ion-beam deposition (ES-IBD) method.[62] In a second step, observation of the single oligosaccharide chains is performed by scanning tunneling microscopy (STM) at 11 K.[63] Subnanometer resolution enabled the direct visualization of the single residues and quantification of the local chain flexibility. We demonstrated that **AAAAAA** chains adopted mostly straight geometries, while **AFAAFA** presented both straight- and bent-chain geometries (**Figure 2.5**C). The increased flexibility of **AFAAFA** single chains translated into a drastically decreased propensity to form crystalline domains. While **AAAAAA** displayed a cellulose II pattern,[186] **AFAAFA** showed an amorphous XRD pattern (**Figure 2.5**D). In addition, the solubility drastically increased to values > 50 mg/mL (**AAAAAA** solubility was < 1 mg/mL). Together, these results suggest that C-3 deoxyfluorination increased the overall flexibility of the single chain by disrupting the intramolecular OH-3/O-5 hydrogen bond. The enhanced flexibility drastically increased solubility and decreased crystallinity of the fluorinated analog.

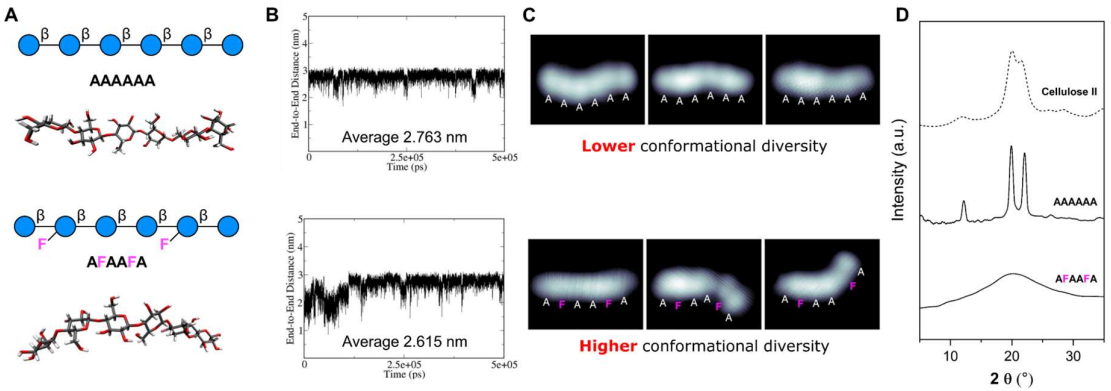


Figure 2.5 A) Representative snapshot from the MD simulations. B) End-to-end distance plot extracted from the MD simulations and average value. C) STM images of the two hexasaccharides obtained after ES-IBD on a Cu(100) surface. D) Powder XRD profile for the two hexasaccharides.

Following these preliminary results (**Figure 2.5**), I studied the impact of the substitution pattern and the effect of the different deoxyfluorination site on cellulose hexasaccharides (i.e. 3F-Glc *versus* 6F-Glc, see **Figure 2.4**). Interestingly, the hexasaccharide with a block pattern of C-3 fluorination (**FFFAAA**) displayed a cellulose II-type XRD pattern, despite the high degree of substitution (**Figure 2.6**A). In contrast, when the pattern was alternated (**FAFAFA**) an amorphous XRD pattern was observed. When 6F-substituted block pattern hexasaccharide **fffAAA** was analyzed, an amorphous XRD was observed indicating that substitution site (6F *versus* 3F) is a key parameter in controlling the aggregation propensity (**Figure 2.6**A). These results are consistent with the similarity observed in the Ramachandran plot extracted from MD simulations (**Figure 2.6**B): while **FFFAAA** resembles **AAAAAA**, **fffAAA** showed a broader distribution, with a wider range of Ψ dihedral angles populated (**Figure 2.6**B, *balck arrow*). The radius of gyration plots confirmed that the 6F substituted analog **fffAAA** has a higher overall flexibility, exploring extended geometries as well as more compact conformations (**Figure 2.6**C). All the other modified hexasaccharides showed amorphous XRDs (**Figure 2.6**A). Overall, the 3F substitution appears to perturb less the typical extended cellulose conformation compared to the 6F substitution in the specific case of a block pattern. However, when the degree of substitution increases, such as in the fully substituted analogs, both 3F and 6F substitution appear to disrupt the typical crystallinity of cellulose.

The introduction of fluorine into oligosaccharides enables the use of the ^{19}F NMR channel to perform structural studies.[143,148] Easily accessible experimental values, such as chemical shifts and coupling

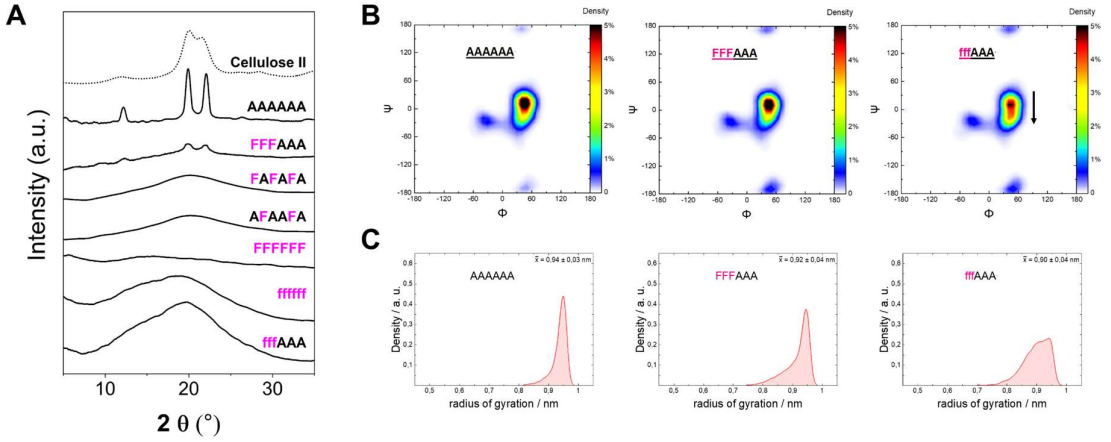


Figure 2.6 A) Powder XRD analysis of synthetic deoxyfluorinated cellulose analogs and the natural counterparts. B) Ramachandran plots for all the glycosidic linkages combined obtained by MD simulations. C) Radius of gyration plots extracted from the MD simulations.

constants, can provide useful structural information. For instance, the ^{19}F NMR chemical shifts of **FFFAAA** indicate that a different chemical environment surrounds the 3F moiety (i.e. internal *versus* external units, ca. 2.5 ppm difference), potentially due to largely different interactions with the solvent molecules (**Figure 2.7**A). In contrast, the ^{19}F NMR chemical shifts of **fffAAA** display a much smaller difference between internal and external residues (ca. 1 ppm, **Figure 2.7**C). In the case of 6F substituted analogs, ^{19}F NMR allows a qualitative analysis[207] of the populations of the ω dihedral angle using as reference the calculated values[208] for the three different ω rotamers (**Figure 2.7**B). I measured different $^{3}J_{H5\text{-}F6}$ values for the external (27.2 Hz) *versus* the internal (25.8 Hz) residues (**Figure 2.7**C). This difference suggests that the *gg* rotamer population is lower in external residues compared to internal ones. A complete estimation of the *gg*, *gt*, and *tg* populations required an additional experimental value (i.e. $^{3}J_{H6\text{-}H5}$) which could not be extracted due to severe spectral overlap in the ^{1}H NMR. MD simulations supported this experimental observation, showing a similar trend of decreased *gg* population for external residues both for **AAAAAA** (**Figure 2.7**D, *top*) and for **fffAAA** (**Figure 2.7**D, *bottom*).

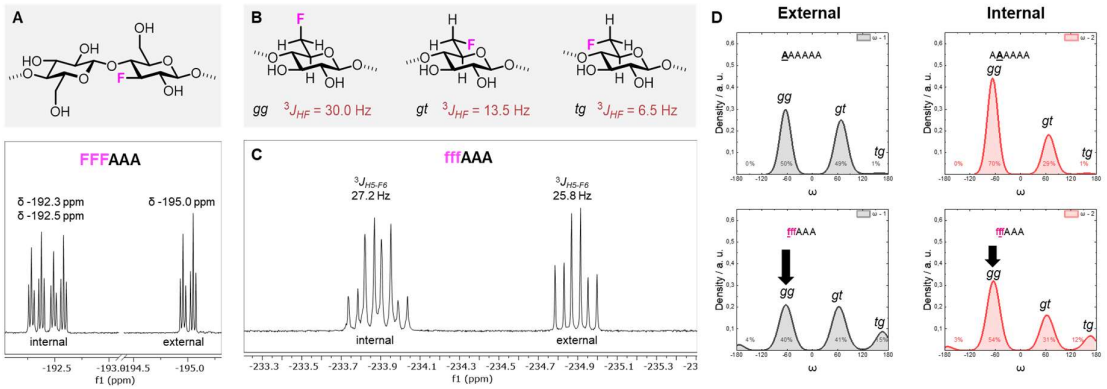


Figure 2.7 A) Fluorinated cellobiose repeating unit (*top*) and excerpt of the ^{19}F NMR of **FFFAAA** (*bottom*). B) Definition of *gg*, *gt*, and *tg* rotamers of the ω dihedral angle. The respective ^{3}J values were previously reported. C) Excerpt of the ^{19}F NMR of **fffAAA** and respective $^{3}J_{H5\text{-}F6}$ values. D) Predicted populations of the rotamers for the ω dihedral angles for external (*left*) and internal (*right*) residues of two analogs.

Overall, the data presented here suggest that the OH/F substitution at the C-6 has a bigger impact on the crystallization tendency than a substitution at the C-3. This indicates that the OH-6 plays a more important role than OH-3 in the crystallization process, potentially due to its higher exposure to solvent molecules, as indicated by NMR on 3F and 6F analogs. Further experiments are ongoing to support this hypothesis.

These results will be compared with analogs of fluorinated chitin prepared with 6F-GlcNAc **BB2.4** (ongoing). With the aid of MD simulations and experimental values of solubility and XRD profiles, we are aiming to understand the impact of C-6 deoxyfluorination in chitin. A comparison between cellulose and chitin will be performed to understand the role of the C-6-OH for the aggregation and crystallinity in these two abundant polysaccharides.

2.2.3 ^{19}F-labelled glycans enable NMR structural studies

In Section 2.2.2, I showed how OH/F substitutions can dramatically alter the macroscopic properties of polysaccharides by altering the conformation through the disruption of selective hydrogen bonds. In this Section, I will discuss how fluorine can be exploited to assist the structural analysis of unlabeled analogues. As extracting conformational information of oligosaccharides using NMR is a challenging task,[66] fluorinated glycans can be conveniently studied using ^{19}F NMR and the structural analysis can be considerably simplified compared to unlabeled analogs.[143,148] Cases where chemical shift degeneracy is extreme, such as in the scenario of highly repetitive oligosaccharides, can benefit from labelling with NMR active nuclei. To achieve this goal, a careful OH/F substitution has to be planned to make sure that solution state conformations are minimally altered.

We aimed to confirm the structure of β-1,6 Glc hexamer **2.13** (**Figure 2.8**A). ^{13}C-labelled analogs, prepared by AGA, were useful probes to provide preliminary data supporting the proposed helical conformation obtained by MD.[83] However, direct experimental evidence of the helical shape of the hexamer remained elusive because the methods commonly employed (i.e. NOEs, J-couplings) are limited to short-range distances (< 5 Å).[65,66] To describe global conformations, residual dipolar coupling (RDC) analysis is often preferred because it can provide information on the relative orientation between monosaccharide rings, regardless of their distance in space.[74,209,210] In particular, RDC delivers information on the relative orientation of specific X-Y bonds between NMR-active nuclei (**Figure 2.8**B). RDCs are generated with samples dissolved in "alignment media" which are anisotropic phases (e.g. liquid crystals, stretched polymer gels) that induce a partial alignment of the oligosaccharide molecules with respect to the magnetic field.[69] This alignment provides RDCs (in the order of Hz), which depend on the orientation between the X-Y bond vector with respect to the magnetic field (**Figure 2.8**B). The RDCs are extracted from HSQC NMR spectra comparing aligned and non-aligned samples.

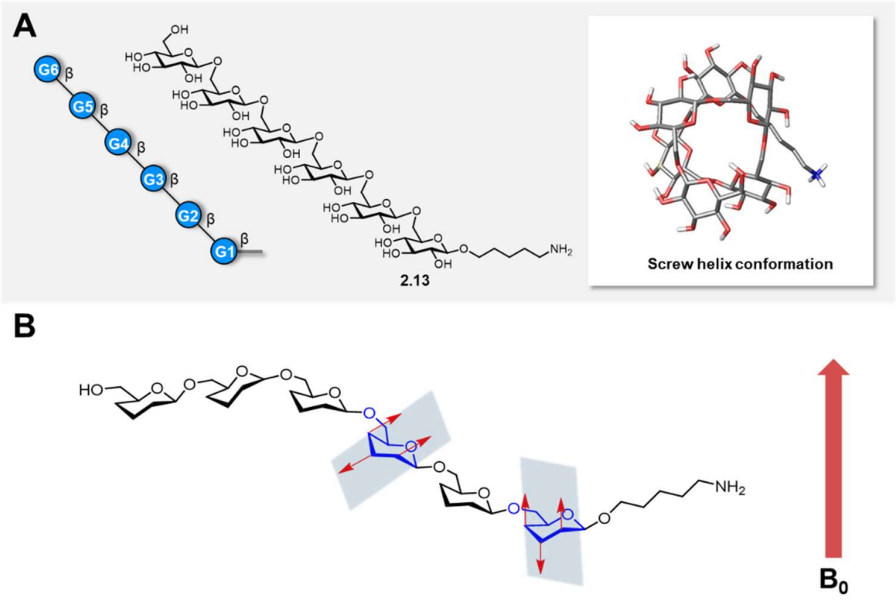


Figure 2.8 A) MD analysis indicated that a β-1,6 Glc hexamer adopts a screw helix conformation. B) Simplified cartoon exemplifying the principle of RDCs analysis. The arrows represent C-H vectors of labelled residues.

In order to obtain experimental evidence of the helical shape of β-1,6 Glc oligosaccharides, site-specific ^{13}C-labelled analogs were initially targeted. In this case, all the intra-ring C-H vectors are parallel to each other and should provide the same RDC value (**Figure 2.9**A). The only C-H vectors that could provide additional information are the ones belonging to the highly flexible methylene moiety of the monosaccharide. I hypothesized that site-specific ^{19}F-labelling could provide an additional vector orientation to be measured through RDC analysis (**Figure 2.9**A). I therefore installed an equatorial F substituent at position C-3, to minimize interference with folding.

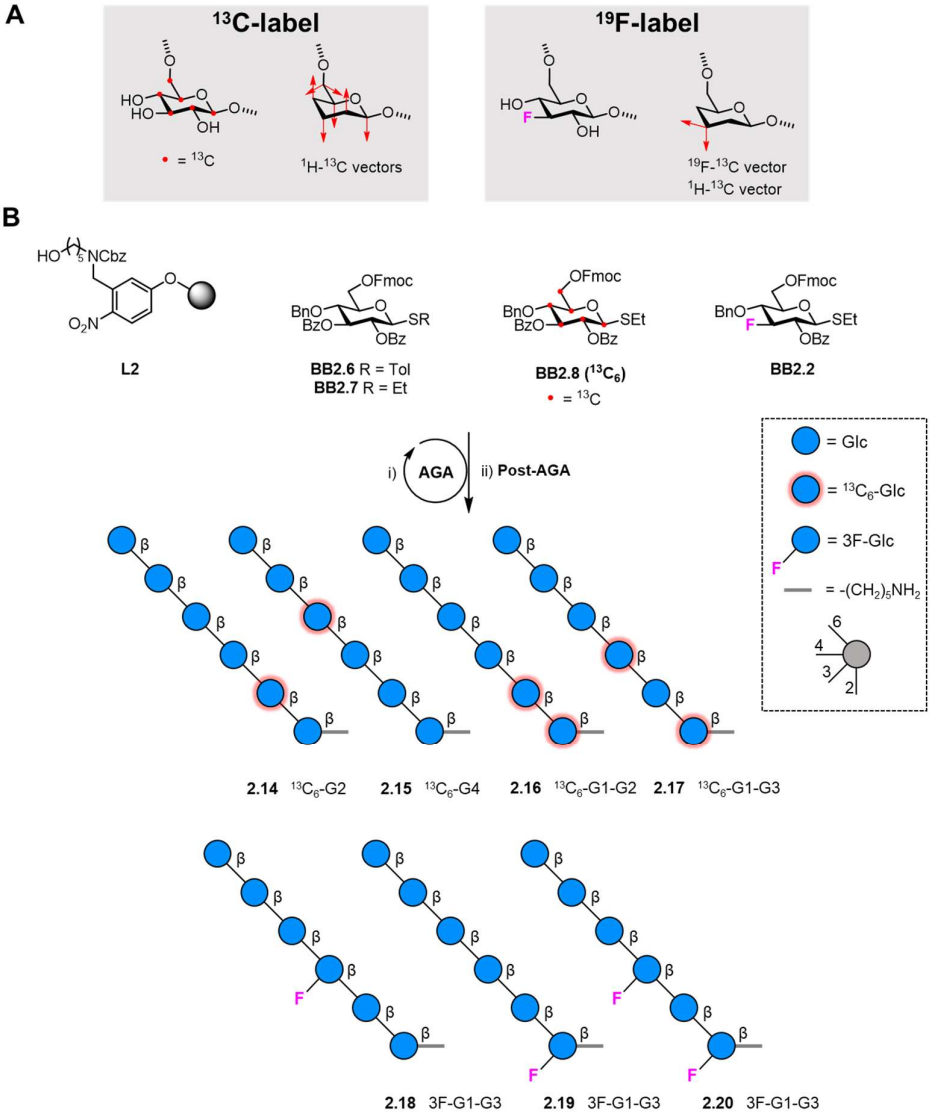


Figure 2.9 A) Comparison between ^{13}C- and ^{19}F-labelled monosaccharides for RDC analysis. ^{13}C-BBs enable RDC measurement on all the C-H vectors (*red arrows*) while ^{19}F-labelled BBs enable RDC measurement on C-F as well as one C-H vector (*red arrows*). B) Synthesis of five hexasaccharide analogues **2.14-20** bearing labelled functionalities in specific position of the glycan chain. The synthesis includes AGA and Post-AGA (detailed procedures are described in Experimental sections 6.2 and 6.3).

Using **BB2.7** and **BB2.8**, a collection of analogs of the β-1,6 Glc hexamer ^{13}C-labelled at different residues was synthesized using AGA (by Dr. Martina Delbianco) (**Figure 2.9**B). To complement this collection, I synthesized three ^{19}F-labelled hexamers by AGA using **BB2.6** and **BB2.2**[190] (**Figure 2.9**B). Two mono-fluorinated compounds (**2.18** and **2.19**) served as chemical shift reference, whereas the di-fluorinated analogue **2.20** was employed in the RDC studies.

To perform the RDC analysis, the labelled analogs were dissolved in an alignment medium, a cromoglycate sodium salt (cromolyn) solution in D_2O.[211,212] Cromolyn is a small rigid aromatic molecule capable of self-aggregating into liquid crystalline phases consisting of columnar aggregates (chromonic phases).[213] These columns are aligned perpendicular to the magnetic field. The RDC NMR analysis was performed by Dr. Ana Poveda (CIC bioGUNE) and further details on the experiments are reported in the original publication (see Experimental section 6.3.3).[214] After confirming similar degree of alignment among the different samples, RDCs were experimentally determined from the HSQC NMR spectra at 800 MHz (**Table S2.1**). These experimental values were compared with calculated ones for different 3-D models. The predicted conformations for the hexasaccharide (**Figure 2.10**, *left*) displayed differences in the conformation around the ω torsional angle (**Figure 2.10**, *right panel*). The helical shape was observed when all ω displayed the *gauche-gauche* conformation (gg_6) while a linear and extended shape was characterized by all *gauche-trans* (gt_6). Intermediate conformers adopted various combinations of *gg* and *gt* rotamers.

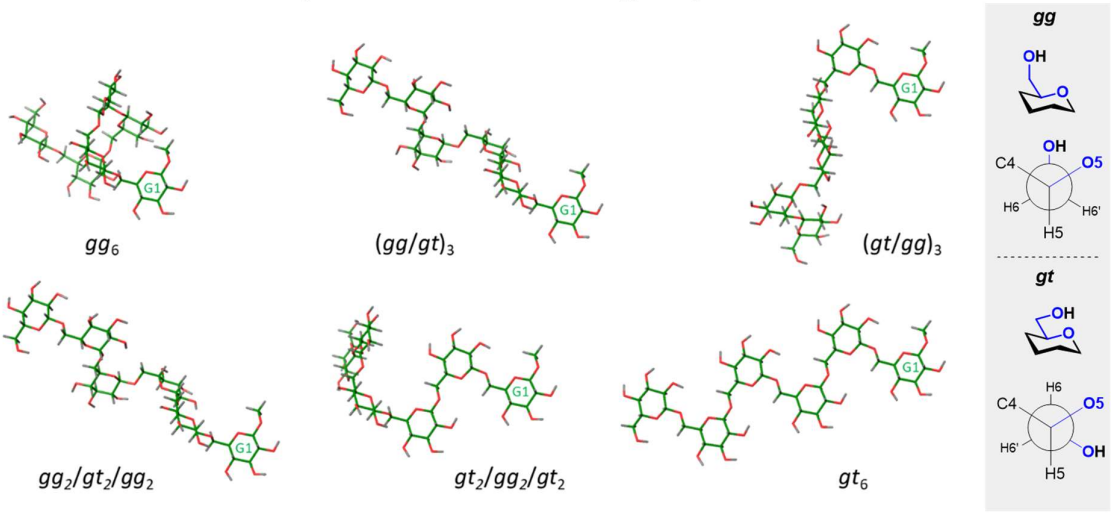


Figure 2.10 3-D models of the hexasaccharide selected for the comparison with the experimental data. G1 is always presented in the same orientation for comparison. The description of the *gg* and *gt* rotamers is shown in the right panel.

A comparison between the experimental RDCs and those predicted for the models (**Figure 2.10**) served to establish the overall conformation of the hexasaccharide. A quality factor (Cornilescu Quality factor, CQf)[215] was calculated to identify which conformer fits best to the experimental data (**Table 2.3**). A perfect fit is observed when the CQf equals zero, while a deviation indicates worse fit. The fitting analysis gathered in **Table 2.3** used different data sets. The fitting is considerably improved when the methylene data are neglected (entries 3 and 4, **Table 2.3**). When the ^{13}C-^{19}F RDC values are taken into consideration, the best fitting is obtained for a mixed *gg*/*gt* form (CQf=0.15, entry 4, **Table 2.3**).

These data suggest that in the cromolyn aligning medium, the hexasaccharide exists as an ensemble of conformers. The extended shape appears to be poorly populated while more compact conformation with larger *gg* populations are more likely to exist, including the helical gg_6. The nature of the β-1,6 glycosidic linkages generates a higher flexibility due to the number of dihedral angles available. No single conformer could be fitted to the experimental data meaning that conformational averaging is taking place, as often observed for oligosaccharides.[65,66] RDC analysis suggested a tendency to adopt compact conformations, including the helical one (gg_6). The ^{19}F-labelled analogs supported the structural elucidation by expanding the number and the orientation of the bonds useful for RDC analysis.

Entry	Analysed compounds	CH_2 included	Alignment medium	Conformers / *CQf*					
				gg$_6$	(*gg*/*gt*)$_3$	(*gt*/*gg*)$_3$	*gg*$_2$/*gt*$_2$/*gg*$_2$	*gt*$_2$ /*gg*$_2$/*gt*$_2$	*gt*$_6$
1	**2.14**, **2.15**, **2.17**	Yes	Cromolyn	**0.46**	0.50	0.80	0.51	0.63	0.53
2	**2.14**, **2.15**, **2.20** (**^{19}F**)	Yes	Cromolyn	0.52	**0.26**	0.79	0.50	0.54	0.71
3	**2.14**, **2.15**, **2.17**	No	Cromolyn	0.22	**0.11**	0.49	0.24	0.52	0.35
4	**2.14**, **2.15**, **2.20** (**^{19}F**)	No	Cromolyn	0.42	**0.15**	0.72	0.48	0.39	0.64

Table 2.3 Cross-fitting of the experimental RDC data obtained in the chromonic medium for the model structures shown in **Figure 2.10**, with ω angles in the *gg* or *gt* conformations. The fitting of the experimental RDCs to other intermediate structures with mixed *gg* and *gt* orientations of the hydroxymethyl groups ((*gg*/*gt*)$_3$, (*gt*/*gg*)$_3$, *gg*$_2$/*gt*$_2$/*gg*$_2$, *gt*$_2$/*gg*$_2$/*gt*$_2$) is also presented. The Cornillescu Quality factor (CQf) derived from the fitting in MSPIN is shown as a quality parameter of the adjustment. The best fitting for each compound is highlighted in green (bold). The following best fitting is light shaded.

2.2.4 ^{19}F-labelled glycan probes to study protein-glycan interactions

In Section 2.2.3, ^{19}F-labelled oligosaccharides have been used for conformational analysis. In this Section, I present how ^{19}F-labelled can serve as useful probes to study functional aspects of glycans, such as binding to proteins and interactions with enzymes.

A challenge in the study of glycan-protein interactions is the need for highly sensitive methods capable of detecting the often low affinities. Several analytical techniques have been developed to quantitatively describe these interactions at the molecular level and in a high-throughput manner.[216–218] Most of these strategies rely on immobilized glycans (e.g. microarray technology)[216–220] or require large amount of samples and analysis time (e.g. ITC[221], SPR[222] or X-ray crystallography[223]). In contrast, NMR allows for the detection of protein-glycan interactions in a fast and reliable manner, providing information on the binding mode in a homogeneous assay format in absence of immobilization protocols.[148,180] ^{19}F-labelled probes have been used to study a plethora of phenomena including (mis)folding, real-time *in vivo* events,[122–124,173,174,224] protein-ligand interactions, and high-throughput ligand screening.[225,226] Inspired by these works, I developed ^{19}F-based glycan probes that enabled the study of protein binding events and enzymatic transformations, while providing guidelines for future design of glycan probes for NMR spectroscopy.

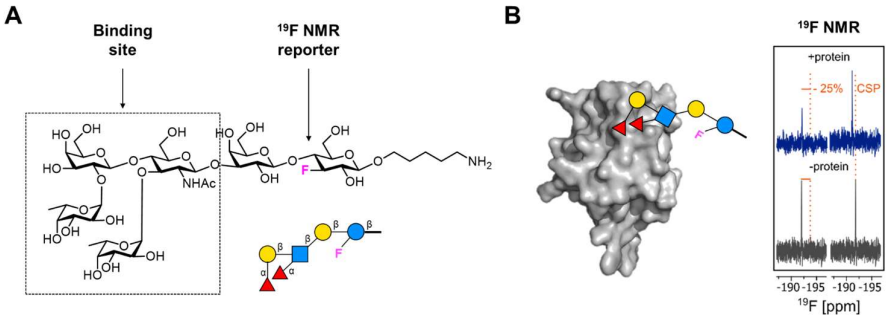


Figure 2.11 A) F-glycans design: the ^{19}F reporter was placed on the Glc unit, distal from what is commonly recognized as the binding site of Lewis type 2 antigens. B) Binding was detected observing changes (decrease in peak intensity or chemical shift perturbation, CSP) in the ^{19}F NMR channel upon addition of protein.

I targeted Lewis type 2 antigens, a class of biologically relevant glycans involved in several physiological and pathological processes, including cancer, where they act as cell adhesion or recognition mediators.[227,228] Subtle differences in the fucosylation pattern strongly impact their interaction with proteins and ultimately can lead to host immune system elusion.[229–232] Recently, an elegant procedure to access a collection of Lewis type 2 antigens by AGA was reported.[233] I envisioned a similar approach to produce a set of ^{19}F labelled analogues to screen protein binding in a simple ^{19}F NMR assay. Since the position of the ^{19}F reporter is thought to be crucial to obtain valuable information, ^{19}F-labelled glycans, hereafter F-glycans, were designed with the ^{19}F reporter in the lactose inner core subunit (**Figure 2.11**A). This position is distal from the binding site (*i.e.* non reducing end) to minimize the effect of the fluorine atom during the binding event.[234,235] I hypothesized that labelling of the inner core glucose unit should maintain sensitivity to the binding event due to overall changes in the correlation time of the glycan in the bound state, reporting changes in the ^{19}F NMR signal (**Figure 2.11**B).[224] In addition a chemical shift perturbation (CSP) could also occur upon binding (**Figure 2.11**B).

^{19}F labelled analogues of Lewis type 2 antigens were assembled on a solid support (functionalized Merrifield resin, **L2**) using **BB2.1** and **BB2.9-12** (**Figure 2.12**A). The BBs are equipped with a thioether or a dibutylphosphate reactive leaving group. Orthogonal cleavage of the 9-fluorenylmethoxycarbonyl (Fmoc) and levulinoyl (Lev) temporary protecting groups permits regioselective chain elongation. Benzyl (Bn), benzoyl (Bz), and *N*-trichloroacetyl (TCA) groups protect the remaining functionalities. β-Stereoselectivity during glycosylation with BBs **BB2.1** and **BB2.9-12** is ensured by anchimeric assistance of the protecting

groups at C-2, while α-stereoselectivity with **BB2.12** was verified in previous studies.[233] **BB2.1** is labelled with the ^{19}F reporter at C-3 position.[190] Each oligosaccharide was assembled in an overnight run following previously reported conditions for unlabelled analogues (see Experimental section 6.3.2).[233] Post-AGA manipulations included solid-phase methanolysis (Module F1),[190] photocleavage[109] from the solid support (Module G1), and hydrogenolysis (Module H1 and H2) (for AGA modules see Experimental section 6.2). A single final purification step afforded the target F-glycans (**F-Lac**, **F-nLac$_4$**, **F-Le^x**, **F-H type 2**, and **F-Le^y**) in overall yields of 5% to 16% over 7 to 15 steps.

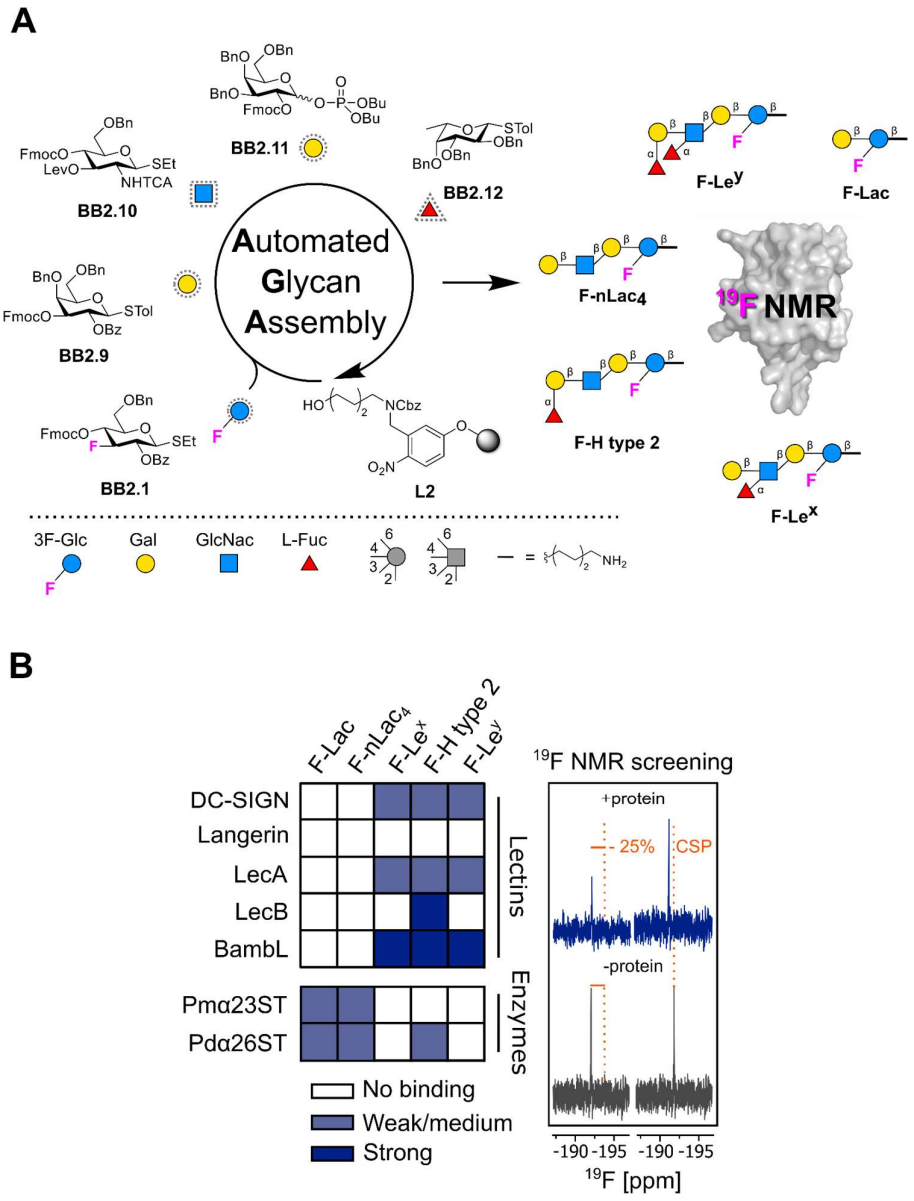


Figure 2.12 A) BBs, including **BB2.1** bearing the ^{19}F reporter, were employed for the AGA of a collection ^{19}F-labelled Lewis type 2 antigen analogues represented following the Symbol Nomenclature for Glycans (SNFG). B) The F-glycans were screened against proteins, including mammalian and bacterial lectins, as well as enzymes. The enzymes were screened in the absence of donor (*i.e.* CMP-Neu5Ac) to probe binding to the substrate. The binding strength was defined depending on the changes observed in the NMR after addition of the protein (*right panel*). Strong binding (*blue*) is defined as a decrease in peak intensity higher than -25% or a chemical shift perturbation (CSP) bigger than 0.01 ppm in the ^{19}F NMR. Weak/medium binding (*light blue*) is defined as a decrease in peak intensity higher than -25% in the CPMG filtered ^{19}F NMR. No binding (*white*) is defined as a decrease in peak intensity lower than -25% in CPMG filtered ^{19}F NMR.

The F-glycans were screened against mammalian and bacterial lectins as well as enzymes (performed by Dr. Elena Shanina, see Experimental section 6.3.4). A ^{19}F and CPMG NMR screening was performed to probe the interactions of five F-glycans (**F-Lac**, **F-n-Lac$_4$**, **F-Le^x**, **F-H type 2**, and **F-Le^y**) with mammalian (Langerin, DC-SIGN) and bacterial (LecA, LecB, BambL) lectins and enzymes (α(2,3)-sialyltransferase from *Pasteurella multocida* (Pmα23ST)[236] and α(2,6)-sialyltransferase from *Photobacterium damsela* (Pdα26ST)[237]) (**Figure 2.12**B). Upon protein binding, the molecular tumbling rate of the glycan is drastically affected resulting in a decrease of the ^{19}F signal intensity.[224] Monitoring ^{19}F chemical shift perturbation (CSP) or change in peak intensity upon addition of protein allowed to qualitatively evaluate the strength of the interaction. A decrease in peak intensity or a CSP in ^{19}F NMR indicates strong binding. Application of a spin echo filter[68] allowed the detection of weak binders. The Carr-Purcell-Meiboom-Gill (CPMG) spin echo filter allows to remove the contribution of resonances of bound ligand, typically having short transverse relaxation time (T_2) and broad signals.[68] As a result, bacterial (LecA, LecB and BambL) and mammalian (DC-SIGN ECD) lectins preferred fucosylated glycans (**Figures S2.2**A, **S2.2**B, **S2.2**C and **S2.2**E). No binding to F-glycans was observed in presence of Langerin ECD (**Figure S2.2**D), in agreement with previous reports.[238] In contrast, the enzymes showed much weaker interactions and a slight preference for shorter non-branched glycans (**Figure S2.3**).

After confirming the success of the F-glycans in revealing binding with proteins, I wanted to answer two questions regarding these NMR glycan probes. First, I wanted to confirm that the fluorine labelling did not affect the binding with the protein counterpart. Second, I wanted to explore the importance of reporter position in revealing binding. To answer these two questions, I focused on the interaction between the lectin DC-SIGN and fucosylated Lewis antigens. DC-SIGN recognizes cellular ligands and pathogens that express Lewis antigens. In particular, Le^x and Le^y present on *Schistosoma mansoni*[239] and *Helicobacter pylori*[240] or endothelial cells[241], respectively, are known binding partners for DC-SIGN.[242] The strong preference of DC-SIGN for fucosylated ligands has also been elucidated with the crystal structure of the carbohydrate-binding site of DC-SIGN bound to Le^x.[243] The qualitative CPMG NMR screening of mammalian lectins confirmed the interaction of DC-SIGN with fucosylated glycans **F-Le^x**, **F-H type 2** and **F-Le^y** (**Figure 2.12**B), as indicated by changes in the NMR peak intensity of the reporter molecule. First, I explored the role of the ^{19}F reporter in F-glycan binding to DC-SIGN. Protein-observed ^{15}N HSQC NMR of DC-SIGN CRD in the presence of **F-Le^x** and **Le^x** were performed (by Dr. Elena Shanina). Both ligands promoted similar changes in the backbone of DC-SIGN CRD, confirming that deoxyfluorination did not affect binding (**Figure 2.13**A). Then, I investigated the effect of the reporter's position on the ability to reveal binding events. I conjugated a CF_3 moiety on the **H type 2** sample using mild conditions to afford **CF$_3$-H type 2** (**Figure 2.13**B). Functionalization of the remote end of the aminopentyl linker enabled ^{19}F-labelling in a position far from the carbohydrate-binding site. The new ligand was tested using ^{19}F and CPMG NMR and, remarkably, its binding was observed with both mammalian (DC-SIGN, **Figure 2.13**C) and bacterial lectins (BambL, **Figure S2.4**). These results indicate that the positioning of the ^{19}F reporter on the Glc unit does not affect the binding of F-glycans with proteins. Furthermore, the ^{19}F reporter can be remote to the glycan binding site to avoid any interference with the binding event, while preserving excellent sensitivity. However, the functionalization of the amino-linker with a CF_3 moiety prevents any further conjugation of the glycan (e.g. to protein, surface, liposome).

Besides monitoring lectin-glycan binding, the ^{19}F NMR assay allowed to monitor the binding of F-glycans (**F-Lac** and **F-nLac$_4$**) to enzymes. Two sialyltransferases (Pmα23ST[236] and Pdα26ST[237]) were screened in the absence of glycosyl donor (i.e. CMP-Neu5Ac) and revealed weak binding to the glycan substrate (**Figure 2.12**B and **S2.3**). This is particularly relevant because acceptor binding sites of transferases usually have a very low affinity and are difficult to detect. Shorter non-branched glycans (**F-Lac** and **F-nLac$_4$**) showed stronger binding than longer branched ones. **F-Le^x** did not show any binding with Pmα23ST or Pdα26ST, matching its known poor reactivity as acceptor (**Figure S2.3**).[244] In contrast, Pdα26ST showed weak binding to **F-H type 2**, in agreement with previously reported enzymatic activity (**Figure S2.3**).[237]

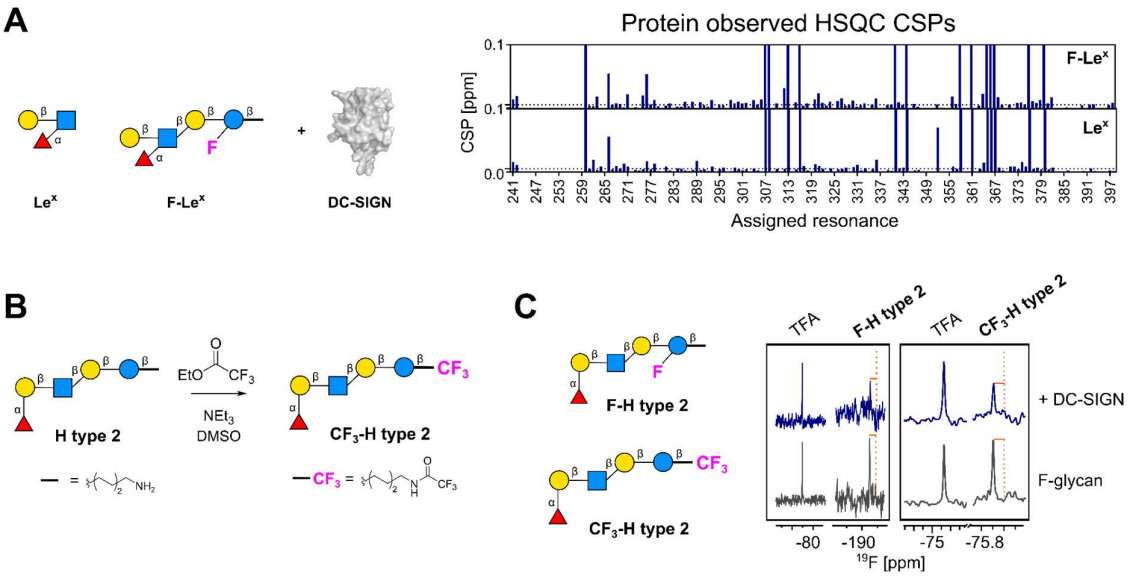


Figure 2.13 A) CSP plot of assigned resonances of DC-SIGN ECD in presence of **F-Le^x** and **Le^x** showing that **F-Le^x** perturbed resonances similarly to unlabelled **Le^x**. B) Conjugation of a CF_3 reporter on the aminopentyl linker of unprotected **H type 2**. Reaction conditions: CF_3COOEt (7.4 equiv.), NEt_3 (3.7 equiv.), DMSO, RT, 16 h, yield 51%. C) CPMG NMR screening of F-glycans alone (*grey*) and in presence of DC-SIGN ECD (*blue*). DC-SIGN ECD binds to **F-H type 2** and **CF_3-H type 2** as shown by a decrease in peak intensity in presence of protein (*orange lines*).

This simple assay could be envisioned as screening platform to identify acceptor substrates for known enzymes and for the discovery of new glycosyltransferases.[245,246]

Next, I wanted to expand the scope of applications of the F-glycans to real-time monitoring of enzymatic reactions. I hypothesized that the high sensitivity of the ^{19}F reporter to subtle modifications in its chemical environment[224] could be exploited to reveal transformations occurring to the glycan. The difference in chemical shift between substrate and product could serve to monitor the reaction in real-time (**Figure 2.14**A). The possibility to place the ^{19}F reporter on a carbohydrate unit in proximity to the functionalization site is crucial to detect a chemical shift perturbation. I selected two enzymes (β-galactosidase[247] and Pmα23ST[236]) and their activity was monitored on a model substrate, **F-Lac**. Glycosidic bond cleavage, mediated by β-galactosidase, was followed by ^{19}F NMR (performed by Dr. Elena Shanina). Cleavage of the terminal β-galactose induced a chemical shift perturbation and allowed real-time tracking of the enzymatic reaction (**Figure 2.14**B). Next, glycosidic bond formation promoted by Pmα23ST[236] was monitored in real-time. Pmα23ST trasnfers *N*-Acetyl-neuraminic acid (Neu5Ac) from an activated cytidine monophosphate donor (CMP-Neu5Ac) to the C-3 OH of the terminal galactose unit of **F-Lac** to yield **F-sLac**. The electron-withdrawing nature of Neu5Ac induced a chemical shift perturbation of 0.2 ppm on the ^{19}F labelled acceptor, allowing to track in real-time the enzymatic sialylation process (**Figure 2.14**C). When the ^{19}F reporter was positioned remotely to the reactive site of the acceptor (> 3 residues away, **F-nLac$_4$**), no chemical shift perturbation was noticed, despite the success of the enzymatic transformation (**Figure S2.6**). Thus, in contrast to what observed for protein binding, the position of the ^{19}F reporter is key to monitor enzymatic reactions.

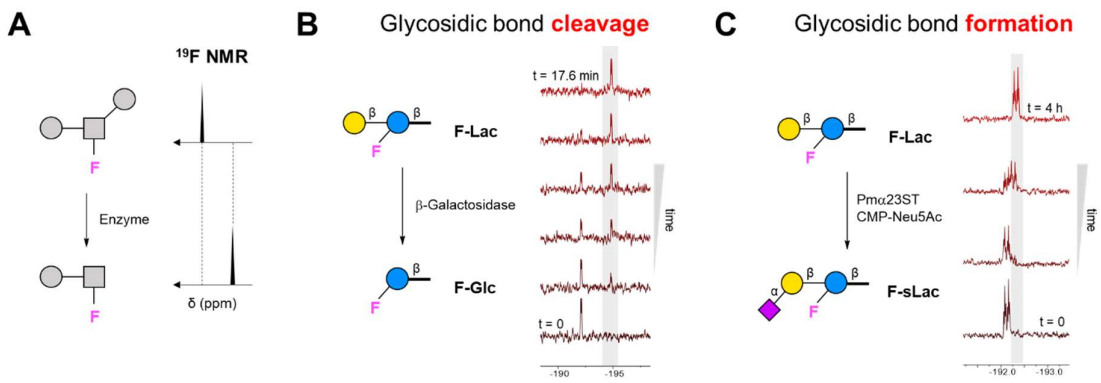


Figure 2.14 A) Principle of ^{19}F NMR monitoring of enzymatic reactions. A ^{19}F NMR readout enables detection of the enzymatic transformation occurring on the glycan. B) **F-Lac** was incubated with β-galactosidase and **^{19}F** NMR was used to perform real-time tracking of **F-Glc** product formation. C) B) **F-Lac** was incubated with Pmα23ST and **^{19}F** NMR was used to perform real-time tracking of **F-sLac** product formation. Product formation was confirmed by HPLC (Figure S2.5).

2.3 Conclusions

In summary, I developed the AGA synthesis of several, diverse, well-defined fluorinated oligosaccharides ranging from cellulose, a natural structural material, to Lewis type 2 antigens, biologically relevant complex glycans. Single-site fluorination, designed to disrupt specific hydrogen bonds, permitted to study how specific modification patterns affect the conformation as well as the macroscopic properties of the material. Single molecule imaging allowed to directly visualize how fluorination impacts cellulose chain conformation. In addition, fluorine enabled the study of a helical hexasaccharide *via* RDC analysis. Furthermore, the unique properties of fluorine enabled the study of glycans from a functional perspective. A collection of complex glycans labelled with a ^{19}F reporter was assembled by AGA and used for the high-throughput NMR screening and characterization of protein-glycan interactions. In addition, these probes allowed for real-time tracking of enzymatic transformations (e.g. sialylation).

Single-site fluorination represents a powerful method to manipulate molecular flexibility through the disruption of hydrogen bonds and is expected to have an impact where fine structural manipulation of glycan conformation is needed (e.g. protein glycan interactions). Such molecular level manipulations enable also fine tuning of macroscopic properties (e.g. crystallization propensity), which could be useful in designing self-assembling carbohydrate materials. In this chapter, I demonstrated that ^{19}F-labelled glycans are versatile probes for NMR studies since both structure (Section 2.2.3) and function (Section 2.2.4) can be studied. Labelled glycans, rapidly assembled through AGA, could be essential tools to study glycans' secondary structures which have long been neglected.[66] Furthermore, the ability of ^{19}F glycan probes to reveal binding or enzymatic transformation in solution and in real-time could open the way to *in cell* NMR applications, often hampered by high background signals.[124,125,148]

3 A bottom-up approach to study the supramolecular assembly and chirality of carbohydrate materials

This chapter has been modified in part from the following articles:

S. Gim, **G. Fittolani**, Y. Nishiyama, P.H. Seeberger, M. Delbianco Supramolecular Assembly and Chirality of Synthetic Carbohydrate Materials *Angew. Chem., Int. Ed.* **2020**, 59, 50, 22577-22583 https://doi.org/10.1002/ange.202008153

S. Gim, **G. Fittolani**, Y. Yu, Y. Zhu, P. H. Seeberger, Y. Ogawa, M. Delbianco Targeted chemical modifications identify key features of carbohydrate assemblies and generate tailored carbohydrate materials *Chem. Eur. J.* **2021**, 27, 52, 13139-13143 https://doi.org/10.1002/chem.202102164

G. Fittolani, D. Vargová, P.H. Seeberger, Y. Ogawa, M. Delbianco Bottom-up approach to understand chirality transfer across scales in cellulose assemblies *J. Am. Chem. Soc.* **2022**, 144, 27, 12469-12475 https://doi.org/10.1021/jacs.2c04522

Specific contribution

Section 3.2.1 I performed the synthesis (building blocks and **LL**, **DD-PFB**, **DD-3F**). Dr. Yang Yu performed the synthesis of **DD-6F**. Dr. Soeun Gim and Dr. Yu Ogawa performed microscopic characterization (TEM, AFM) and powder XRD.

Section 3.2.2 I performed the synthesis (building blocks and $\mathbf{D_6}$, $\mathbf{L_6}$, $\mathbf{L_3D_3}$, $\mathbf{L_2D_4}$, $\mathbf{LD_6}$, $\mathbf{LD_6L}$, $\mathbf{LD_5L}$). Dr. Denisa Vargová performed the synthesis of $\mathbf{D_5}$, $\mathbf{D_7}$, $\mathbf{D_8}$, $\mathbf{D_9}$. I performed the microscopic characterization (TEM, AFM) and XRD analysis. Dr. Yu Ogawa performed the electron diffraction analysis.

3.1 Introduction

Non-natural monosaccharides represent unique tools to study structure and the aggregation of carbohydrates. Site-specific deoxyfluorination enabled the selective disruption of hydrogen bonds in cellulose and gave access to the ^{19}F NMR channel for NMR studies (Chapter 2). In this chapter, non-natural monosaccharides were exploited to study the chirality of self-assembled carbohydrate materials. Using mirror-image monosaccharides, I aimed at studying the transfer of chirality occurring across different hierarchical levels of carbohydrate materials. The underlying theme in this chapter is the *bottom-up* approach that was used to gain molecular understanding of supramolecular chirality, a common feature of natural carbohydrate assemblies, and ultimately to control it.

Self-assembled systems are widespread in nature at all scales, from micro- to macroscopic level.[1] Biopolymers can form hierarchically self-assembled materials based on non-covalent interactions such as van der Waals, electrostatic, π-π stacking, hydrophobic interactions, as well as hydrogen and coordination bonds.[37,248,249] When discrete molecules assemble to form a supramolecular system, new properties emerge such as supramolecular chirality.[250,251] Molecular chirality - based on covalent bonds - is well-understood and many synthetic methods are available to construct asymmetric molecules.[252] In contrast, the driving forces and key interactions imparting chirality at the supramolecular level are harder to elucidate and, therefore, to reproduce. Supramolecular chirality has been observed in inorganic materials, polymers, as well as biological aggregates (e.g. amyloids) and much research has been devoted to describe the origin of these twists.[253] In order to understand such complex systems at a molecular level, synthetic model systems have been developed. For instance, well-defined model peptides have contributed to the long lasting debate on the origin of the twist in amyloid aggregates.[130–132] Synthetic peptides based on defined sequences of L- and D-amino acids suggested that amino acids located at the peptide termini were responsible for the twist adopted by peptide supramolecular fibers,[133,134] with left handed twist associated to the presence of L-amino acids.[8,22] Moreover, fine-tuning of these synthetic analogs generated novel chiral self-assembled materials with tunable shapes and properties.[135–137]

I proposed to study the self-assembly and the chirality of supramolecular carbohydrate materials using a *bottom-up* approach. Simple and well-defined synthetic compounds formed the basis of this approach. First, I studied the transfer of chirality across scales using a simple disaccharide model system that self-assembles into helical fibers. I employed site-specific modifications to identify key non-covalent interactions stabilizing the helical supramolecular assembly. In the second part of the chapter, I expanded this *bottom-up* approach to a more complex natural polysaccharide and used synthetic oligomers to understand chirality transfer across scales in cellulose. Synthetic cellulose oligomers with well-defined sequences of natural D-Glc and its enantiomer L-Glc, constituted the model system. These cellulose oligomers self-assemble into platelets with controlled dimensions that further aggregate into bundles, displaying chiral features directly connected to their monosaccharide composition. The insertion of L-Glc units in the sequence of D-Glc cellulose oligomers drastically impacted the macroscopic properties, such as solubility, crystallinity, and chirality of the bundles. Overall, this chapter introduces a *bottom-up* approach for the study of carbohydrate materials and highlights the importance of the molecular sequence in dictating supramolecular assembly and chirality. The *bottom-up* approach presented here allowed a molecular level understanding of the self-assembly process opening the way to site-specific modifications that could generate functional carbohydrate materials.

3.2 Results and discussion

3.2.1 A model system to study chirality and supramolecular assembly of carbohydrate materials

3.2.1.1 Manipulating supramolecular chirality with mirror-image monosaccharides

A previous study performed in my group identified that simple oligosaccharides self-assemble into defined supramolecular structures.[254] In particular, upon solvent switch, dimer **DD** (**Figure 3.1**) self-assembled into fibers around micrometers long and nanometers wide (**Figure 3.2**A, for details see Experimental section 6.4.4). The formation of fibers occurred upon injection of a **DD** solution (100 mg/mL in hexafluorisopropanol, HFIP) into water (final concentration 2 mg/mL). The crystal structure of these fibers was determined using X-ray diffraction (XRD), solid-state NMR (ssNMR), and microcrystal electron diffraction (MicroED) by Dr. Yu Ogawa. Combined, these characterization techniques provided molecular detail of the packing of **DD** in the fibers. MicroED was especially useful because it enabled a molecular level characterization of the fibers in their native state avoiding time consuming crystallization trials.[53,255] MicroED was never employed to study oligosaccharides due to their electron beam sensitivity. **DD** was an ideal substrate since the benzyl groups are known to render the crystal more resistant to prolonged irradiation.[256,257]

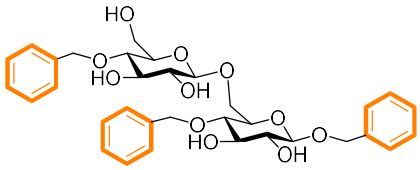


Figure 3.1 Chemical structure of **DD**.

Powder XRD and ssNMR analysis indicated the high crystallinity of the assemblies and the presence of two sets of **DD** in a single unit cell (see Experimental section 6.4.4). MicroED analysis of **DD** fibers at cryogenic temperatures provided a high resolution crystal structure. Electron diffraction patterns obtained from the flat section of the fiber (circled area in **Figure 3.2**A) indicated high crystallinity with a resolution of ca. 1.2 Å. The crystal structure was determined by tilt series MicroED analysis to be an orthorhombic unit cell with a = 5.2 Å, b = 20 Å, c = 37 Å (**Figure 3.2**B, see Experimental section 6.4.4). The tentative molecular packing model in bc and ac projections show a short a-axis indicating that **DD** molecules assume an overall flat conformation and stack along the a-axis (**Figure 3.2**C). The glucose ring planes are oriented roughly in the bc plane. The aromatic rings assemble in a close proximity to each other. The interactions between the aromatic rings are mostly C-H···π type edge-to-face interactions. No face-to-face π-π stacking is present in the packing model, as the molecular spacing in the stacking direction (5.2 Å, a-axis), is larger than the maximum acceptable distance for π-π stacking formation (3.8 Å).[258] The carbohydrate moieties are not in close contact with each other. The relatively low density of the crystal implies that water molecules may be involved in the crystalline lattice. The distances between hydroxyl groups of adjacent molecules allow forming water-bridged hydrogen bonds with a single water molecule between the hydroxyl groups. The **DD** crystalline fibers were subjected to a sequential electron microdiffraction experiment with an electron probe size of about 100 nm. Each ED pattern obtained along the fiber axis corresponds to a different lattice projection (**Figure 3.2**D), revealing left handed twists along their fiber axes. In all ED patterns, the a*-axis is oriented along the fiber axis of the crystal, indicating that the stacking of flat molecular sheets occurs parallel to the fiber axis (**Figure 3.2**E). The crystal twists along the stacking direction, implying that this phenomenon is likely to originate from a slight rotation between the stacked molecules. The apparent half twist pitch (180 degree rotation) is about 5 μm in most crystallites, resulting in a rotation per unit cell of about 0.02 degree. While crystal twists were observed previously for natural carbohydrate crystals such as cellulose and chitin, the mechanism of twisting of carbohydrate crystals is still elusive.[52,259,260] These

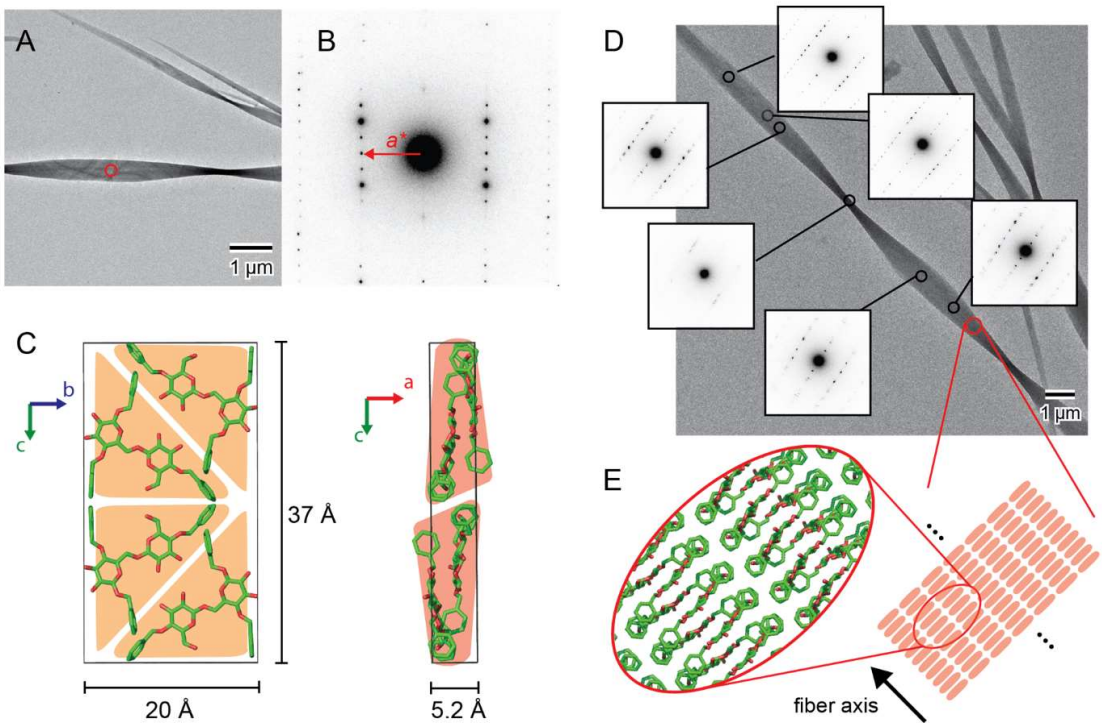
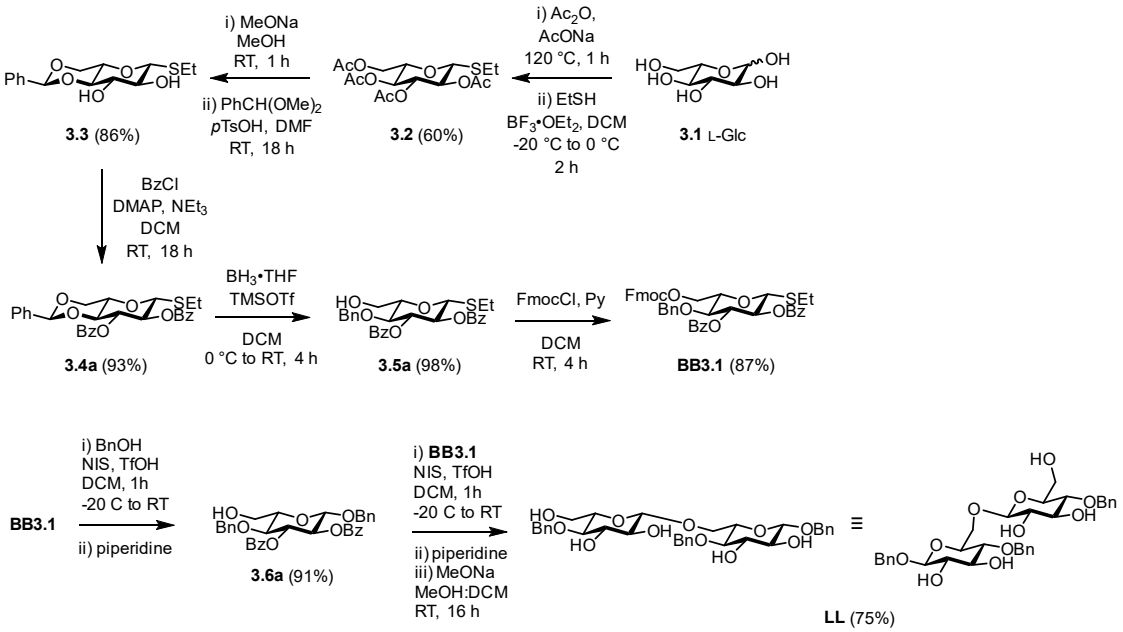


Figure 3.2 MicroED analysis of self-assembled **DD** performed at cryogenic temperature. A) Diffraction contrast image of **DD** crystals. B) Electron diffraction diagram obtained from the circled area in A. C) Tentative molecular packing model of **DD** in the unit cell determined using the MicroED method. Hydrogen atoms are omitted for clarity. D) Twist geometry followed by sequential electron microdiffraction. E) Schematic of molecular packing manner in the fibrillary **DD** crystal.

results suggests that well-defined synthetic systems could shine light on the twisting mechanism of natural systems as well as on the relationship between molecular chirality and supramolecular structures.

After observing that **DD** self-assembles into left handed fibers (**Figure 3.3**, *left*), I hypothesized that the enantiomeric dimer should invert the fiber chirality. This result would demonstrate a direct correlation between molecular and supramolecular chirality. Moreover, being able to control the helicity of the fibers would offer an additional tool to tune the properties of self-assembled materials.[261,262] Indeed, chirality is an important design mode in peptide nanotechnology.[263] In particular, heterochiral peptide-based systems offer many advantages such as increased stiffness of the self-assembled fibers,[137] increased stability towards enzymatic degradation,[264] and access to new morphologies.[135,136,139] Inspired by this work on peptides, I synthesized the enantiomeric disaccharide **LL** starting from L-glucose (**Scheme 3.1**, for details see the Experimental section 6.4.1). The synthesis started from L-Glc **3.1** and, through a seven-step synthesis, led to **BB3.1** (**Scheme 3.1**, further details are reported in Experimental section 6.4.1). The key step is the regio-selective reductive benzylidene ring opening of intermediate **3.3** to give 4-O-Bn protected **3.4a**.[265] **BB3.1** is equipped with a temporary protecting group (Fmoc), a reactive thioglycoside leaving group (SEt), and permanent protecting groups (Bn and Bz) to enable regio- and stereo-selective construction of consecutive β-1,6 linkages between Glc units in **LL** (**Scheme 3.1**).



Scheme 3.1 Synthesis of **BB3.1** and **LL**.

Upon solvent switch, **LL** formed the enantiomeric helical fibers (right handed), confirming the direct correlation between oligosaccharide chirality and fiber helicity (**Figure 3.3**, *right*). Furthermore, the racemic mixture **DD**-**LL**, prepared by solvent switch method, aggregated in a completely new and flat morphology (**Figure 3.3**, *middle* and **Figure S3.3**). AFM analysis of the flat aggregates suggested that the two enantiomers may construct a layer-by-layer supramolecular assembly (see **Figure S3.2**B). The height of the sheets varies from a few hundreds of nanometers to several micrometers, with the single layer measuring 1.5 nm, which is comparable to one dimension of the disaccharide. XRD confirmed that both enantiomers, **DD** and **LL**, have identical crystallinity, whereas the racemic mixture packs in a different manner (see **Figure S3.2**A), as previously observed for heterochiral peptides assemblies.[135,136,139]

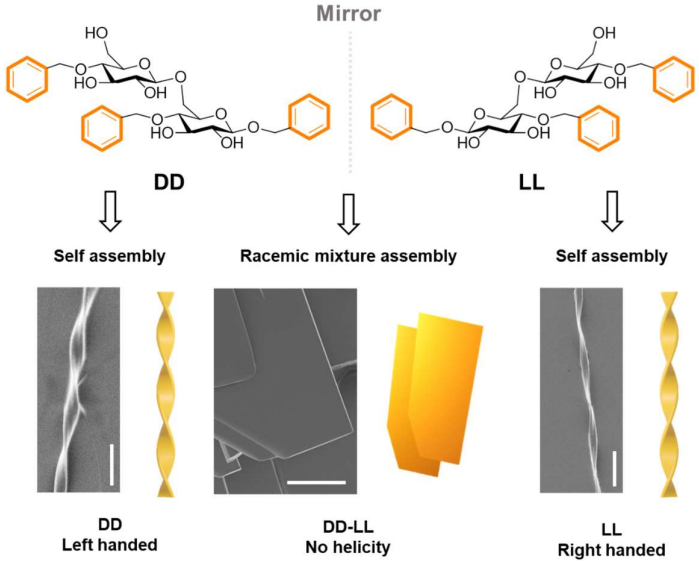


Figure 3.3 Chemical structures of **DD** and **LL** enantiomers and SEM images of their supramolecular assembly. While **DD** (*left*) and **LL** (*right*) show opposite helicity, the racemic mixture **DD**-**LL** (*middle*) generated flat sheet-like assemblies (scale bars 10 μm).

3.2.1.2 Targeted chemical modifications identify key non-covalent interactions that stabilize supramolecular assembly

In a second part of this project, I focused on validating the molecular structure of **DD** fibers (**Figure 3.2**C). According to the MicroED-based molecular packing model (**Figure 3.4**A), the key intermolecular interactions responsible for the formation of the supramolecular fiber (**Figure 3.4**B) are 1) C-H···π edge-to-face type interactions between the aromatic rings and 2) water-bridged hydrogen bonds between specific hydroxyl groups stabilized the assembly. To validate the proposed crystal structure, I designed and synthesized three analogs of **DD** bearing single-site modifications (**Figure 3.4**C-E). Each modification was aimed at disrupting a particular interaction and highlight the importance of the replaced moiety in the assembly process. Compound **DD-PFB** bears a pentafluorobenzyl (PFB) group at the anomeric position, resulting in enhanced hydrophobicity and altered electrostatic interactions. Moreover, the PFB group does not engage in C-H···π edge-to-face stacking because the electron rich para-F does not interact with electron rich phenyl rings.[266] Compounds **DD-3F** and **DD-6F** are the C-3 and C-6 deoxyfluorinated analogues, respectively. Deoxyfluorination was designed to selectively disrupt specific H-bonds.[146,170] Furthermore, this approach will identify functional groups that are non-essential for the self-assembly, that could act as handles to introduce other functionalities for generating functional materials.

The synthesis of **DD-PFB**, **DD-3F**, and **DD-6F** is described in the Experimental section 6.4.2. Upon solvent switch (HFIP to water, see Experimental section 6.4.4), **DD-PFB** and **DD-3F** remained soluble and, upon drying, formed amorphous aggregates (**Figure 3.4**C and D). In contrast, **DD-6F** formed insoluble fibers (**Figure 3.4**E), suggesting that modifications at the non-reducing end C-6 position are tolerated.

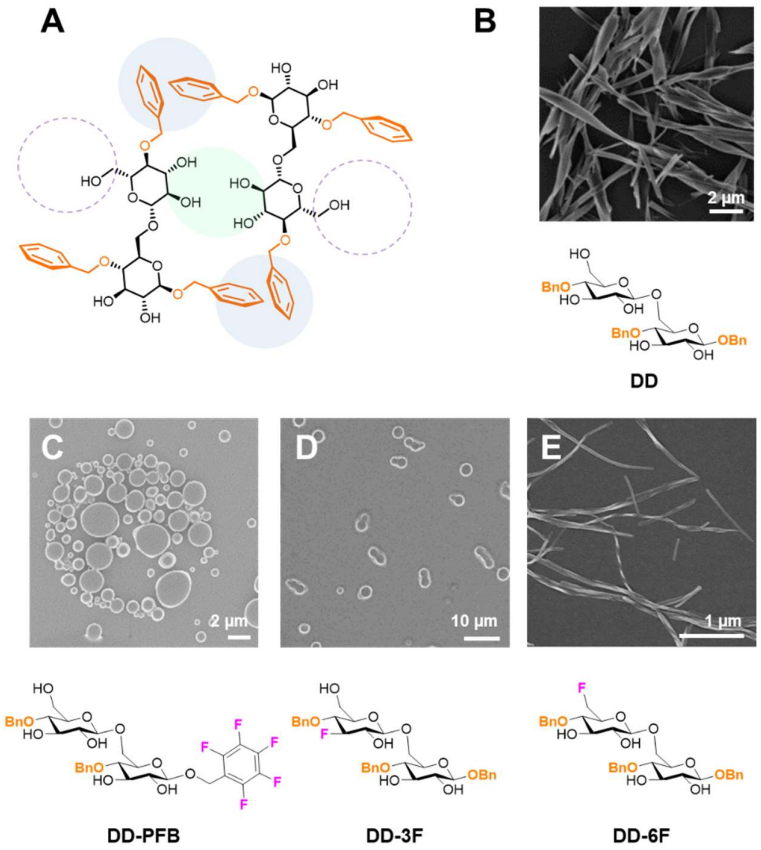


Figure 3.4 A) Schematic representation of the unit cell of Figure 3.2C to highlight the site of the modifications. B-E) Chemical structures and SEM images of **DD**, **DD-PFB**, **DD-3F**, and **DD-6F**, respectively.

The drastically different outcome between **DD-3F** and **DD-6F** demonstrates the importance of the C-3 hydroxyl group for the assembly process. In addition, the lack of self-assembled structures for **DD-PFB** indicates that the C-H···π edge-to-face type stacking interactions are disrupted by the electronegative fluorine atoms of the PFB moiety. **DD-6F** fibers resembled the one formed from **DD** according to powder XRD and electron diffraction patterns (**Figure S3.2C**). However, a higher level of flexibility (i.e. twist and curvature) was observed for **DD-6F** fibers, likely arising from displacement between stacked molecules along the fiber axis.

3.2.2 A bottom-up approach to study cellulose assemblies

In Section 3.2.1, a simplified model system for studying carbohydrate materials, consisting of a disaccharide, was presented. With the characterization methods developed for this model system, I was interested in expanding the *bottom-up* approach to a more complex natural polysaccharide. Cellulose is a perfect candidate to showcase this approach as it is a natural structural material with strong tendency to generate supramolecular hierarchically organized assemblies.

A *bottom-up* approach based on synthetic analogs has never been applied to study cellulose, the most abundant organic material on Earth, despite the growing interest in the chiral properties of cellulose nanofibers (CNFs), cellulose nanocrystals (CNCs) and their assemblies. Native CNFs are intrinsically chiral, possessing nanoscale twists along the fiber direction. Chiral cellulose particles can assemble into larger chiral architectures *in vivo*. Helicoidal arrangements of cellulose fibrils are found in the plant cell walls, and are capable of generating colors in the absence of any pigment through structural coloration.[267,268] Left handed and, in some rare cases, right handed helicoids have been observed.[267] Still, the mechanism for helicoids formation in plants remains debated, invoking the role of other cell-wall constituents (*i.e.* hemicelluloses).[268,269]

Inspired by nature, CNCs have been used to fabricate optical materials such as films[270–272] and structurally colored pigments.[273] Due to their chirality, CNCs have found applications as nanosized chiral inducers for liquid crystals assemblies[274] and heterogeneous enantioselective palladium catalysis[275] or as chiral templates.[276,277] *In vitro*, when a suspension of CNCs transitions from a diluted to a concentrated regime, spontaneous self-organization results in the formation of a chiral nematic liquid crystalline phase. While the intrinsic twisted morphology of single CNCs is considered to govern the chirality of these chiral nematic phases, the mechanism by which nanoscale chirality is transferred to larger scale is debated.[278–280] While single particles (*i.e.* CNCs) are right handed, upon self-assembly, an inversion of chirality occurs, resulting in a left handed chiral nematic phase (**Figure 3.5**).[281] This inversion was quantified using atomic force microscopy (AFM) and electron diffraction methods.[52,259,282] Recently, bundles of naturally sourced cellulose crystallites have been suggested to be responsible for the transfer of chirality across different hierarchical levels.[280,283] However, heterogeneous degrees of polymerization and crystal sizes of cellulose nanomaterials obtained via *top-down* approaches complicated the description of these systems at the molecular level.[284] Computational models have speculated on the origin of the twist in cellulose crystals and are, to date, the only option to study these systems at the molecular level.[285] Thus, the molecular origin of this twist and the mechanism of chirality transfer to higher hierarchical assemblies remain open questions.

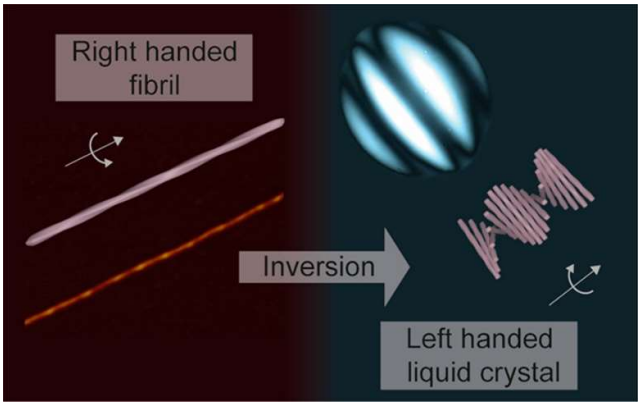
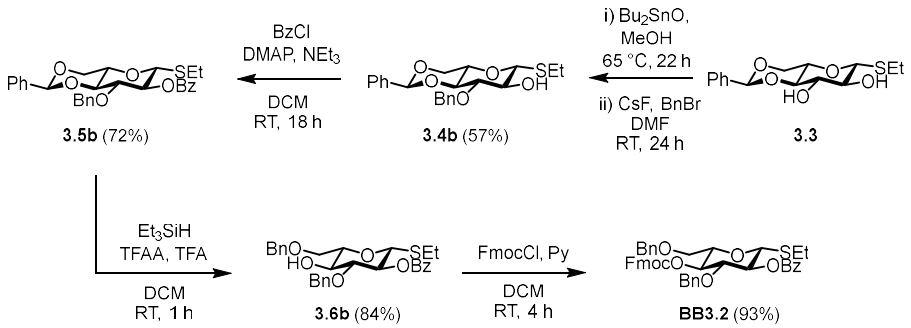


Figure 3.5 Right handed CNCs (*left*) self-assemble into left handed nematic liquid crystals phases. Reprinted with permission from *ACS Nano* **2018**, 12, 5141-5148 Copyright 2018 American Chemical Society.

3.2.2.1 Cellulose oligomers assemble into nanocrystals

I targeted well-defined cellulose oligosaccharides resembling the macroscopic properties of cellulose (*i.e.* crystallinity, mode of assembly) as model system to study cellulose assembly. Cellulose oligomers with

defined length (uniform dispersity, Đ=1) were synthesized by AGA[107,190] on a solid support (functionalized Merrifield resin, see Experimental section 6.2) using protected D-Glc **BB3.3** and **BB3.4** (**Figure 3.6**). The BBs were equipped with an anomeric thioether or dibutylphosphate reactive leaving group, while most hydroxyl groups were masked as benzyl (Bn) ethers and benzoyl (Bz) esters. Upon glycosylation, cleavage of the 9-fluorenylmethoxycarbonyl (Fmoc) temporary protecting group liberated the hydroxyl group to be used in the following regio-selective chain elongation. Iterative cycles of glycosylation and Fmoc deprotection permitted precise control over the length of the oligomer. Neighboring group participation of the ester group at C-2 ensured β-stereoselectivity during glycosylation. Each oligosaccharide was assembled overnight following previously reported conditions.[190] Post-AGA manipulations included solid-phase methanolysis, photocleavage from the solid support, and hydrogenolysis (see Experimental section 6.2). A single final purification step afforded the target cellulose analogs in overall yields of 6 to 60%. Longer cellulose analogs (degree of polymerization, DP > 6) were poorly soluble after the hydrogenolysis step and therefore obtained in drastically decreasing isolated yield. The enantiomeric non-natural analog $\mathbf{L_6}$ was assembled using protected L-Glc **BB3.2** (**Scheme 3.2**, further details are reported in Experimental section 6.4.1). The key step in the synthesis of **BB3.2** involved the selective ring opening of the benzylidene protected intermediate **3.3** to obtain the 6-O-Bn protected compound **3.4b** (**Scheme 3.2**).[265]



Scheme 3.2 Synthesis of L-Glc BB 3.7b.

To evaluate the crystallinity and the aggregation tendency of these oligomers, I employed powder XRD, TEM and AFM. Powder XRD indicated that all analogs $\mathbf{D_5}$, $\mathbf{D_6}$, and $\mathbf{D_7}$ (DP ranging from 5 to 7) assemble with the cellulose II crystal structure (**Figure 3.6**B).[186] $\mathbf{D_5}$ was highly soluble, $\mathbf{D_6}$ showed intermediate solubility, while $\mathbf{D_7}$, $\mathbf{D_8}$, and $\mathbf{D_9}$ were poorly soluble in water (Table S3.1). The mirror-image oligomer $\mathbf{L_6}$ showed a cellulose II-type powder XRD profile and solubility similar to $\mathbf{D_6}$.

Conventional TEM imaging of the negatively stained cellulose oligomers (1 mg/mL in water) indicated the presence of thin platelet-like particles for analogs with DP≥6 (**Figure S3.6**, **S3.9**, **S3.13**, and **S3.15**). No platelets were observed for $\mathbf{D_5}$, suggesting that its high water solubility prevented assembly at the concentration used for this study. The platelets obtained from $\mathbf{D_6}$ are ca. 50-500 nm long and ca. 20-50 nm wide. A similar morphology was observed for the assembly of enzymatically synthesized cellulose oligomers[286] and for mercerized cellulose nanocrystals,[287] both based on the cellulose II-type crystal structure. CryoTEM demonstrated that the platelets existed in aqueous suspension and were not the result of solvent evaporation (**Figure 3.6**C and D). Electron diffraction (ED) analysis of the platelets generated from $\mathbf{D_6}$ and $\mathbf{L_6}$ confirmed the cellulose II type assembly (**Figure 3.6**D) and indicated a molecular packing where the cellulose chains are aligned in an antiparallel manner along the platelet thickness (**Figure 3.6**E). The (001) faces are exposed on the top and bottom sides of the platelets, presenting an alternation of

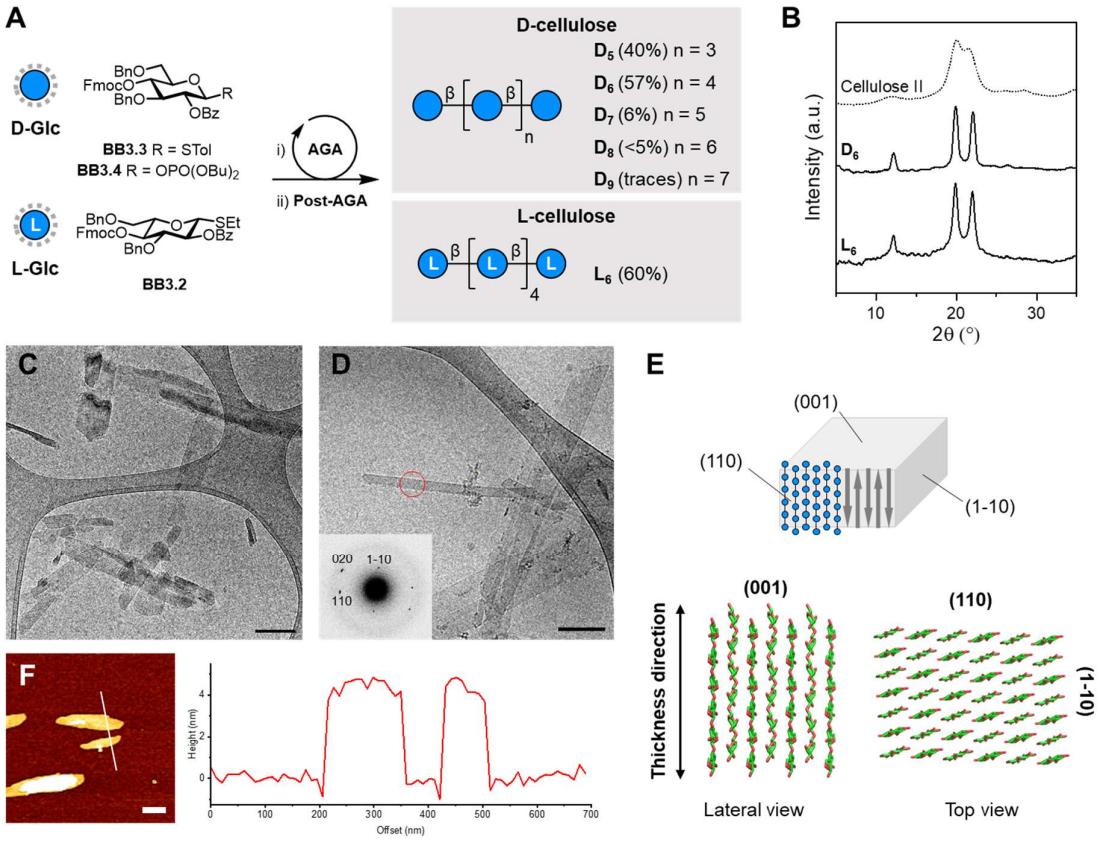


Figure 3.6 A) AGA of cellulose analogs using protected monosaccharide BBs (overall yields reported in parentheses. Solid support and modules used for AGA and Post-AGA are reported in Experimental section 6.2). B) Powder XRD profiles for selected cellulose oligomers indicating a cellulose II-type molecular packing. C) CryoTEM image of $\mathbf{D_6}$ platelets (scale bar 200 nm). D) CryoTEM image of $\mathbf{L_6}$ platelets with the electron diffraction diagram obtained from the red circled area (*red circle*, scale bar 200 nm). E) 3-D molecular model of the platelets composed of cellulose oligomers ($\mathbf{D_6}$) arranged in antiparallel fashion according to the cellulose II crystal structure. F) AFM image of $\mathbf{L_6}$ platelets and height trace measurement (scale bar 200 nm).

reducing and non-reducing ends at the surface. The (110) face is exposed at the platelet tip and presents the Glc hydrophobic face. In aqueous media, crystal growth along the hydrophobic [110] direction is faster than along the hydrophilic [1-10] direction, resulting in elongated platelets.[286] Owing to the controlled length of the cellulose oligomers, the platelets have a well-defined thickness corresponding to the oligomers length. AFM analysis confirmed the tunable thickness of the platelets (**Figure 3.6**F), increasing of about 0.5 nm across the oligomer series, consistent with the addition of a single Glc unit (**Figure S3.26-S3.30**). This finding suggests that chemical synthesis can generate cellulose materials with tunable and controlled sub-nanometer dimensions.

3.2.2.2 Nanocrystals assemble into chiral bundles

No evident chiral features were observed when inspecting the single platelets. However, TEM and AFM imaging revealed the formation of bundles of platelets alongside with isolated ones. These bundles displayed intrinsic chiral features clearly distinguishable by TEM (**Figure 3.7**A). $\mathbf{D_6}$ bundles showed a right handed twist (**Figure 3.7**B, **S3.7**, and **S3.8**), while $\mathbf{L_6}$ bundles were left handed (**Figure 3.7**C, **S3.10**, and **S3.11**). The absolute chirality of the bundles was confirmed by AFM imaging (**Figure 3.7**D and E and **Figure S3.31**). Spherulite-type assemblies have been observed previously for enzymatically synthesized α-chitin[288] and cellulose II[289] aggregates, although no chiral features were recognized. The formation of bundles in this system reflected the strong tendency of the platelets to interact with each other during the drying process, likely due to the increasing concentration.

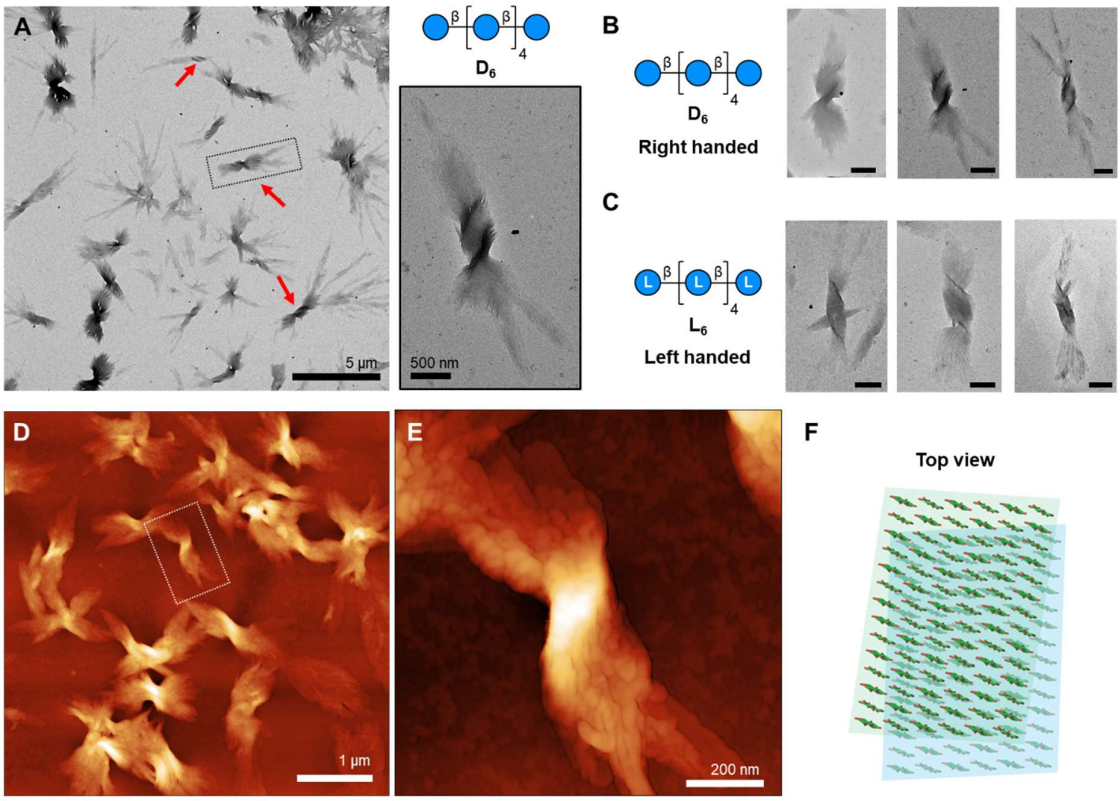


Figure 3.7 A) TEM image of $\mathbf{D_6}$ (1 mg/mL aqueous suspension) shows bundles of platelets with intrinsic chirality (*red arrows*). B) TEM images of $\mathbf{D_6}$ (1 mg/mL aqueous suspension) bundles showing an intrinsic right handed chirality (scale bar 500 nm). C) TEM images of $\mathbf{L_6}$ (1 mg/mL aqueous suspension) bundles showing an intrinsic left handed chirality (scale bar 500 nm). D-E) AFM images of $\mathbf{D_6}$ (1 mg/mL aqueous suspension) bundles showing an intrinsic right handed chirality. F) The fan-like arrangement of the stacking platelets was interpreted as a rotation between the (001) planes.

The analysis of samples at different drying conditions clarified the mechanism of the evaporation induced assembly (**Figure 3.8**). Aqueous suspensions were deposited on TEM grids and the suspension was blotted with a filter paper to a different extent, leading to different amount of water left on the grid surface at the end of the drying process. The sample prepared with extensive blotting showed mainly flat individual platelets (**Figure 3.8**A), while short blotting, hence more remaining water, promoted the formation of large

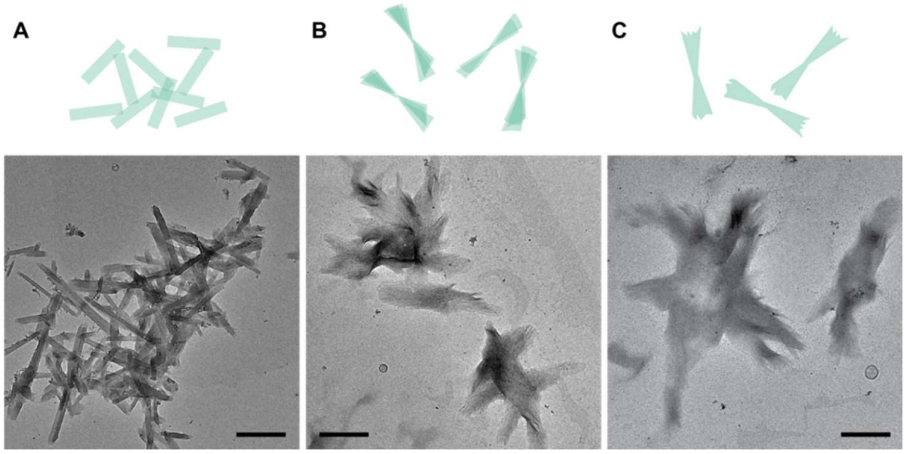


Figure 3.8 TEM images of grids obtained with different blotting strength of $\mathbf{L_6}$. A) Extensive blotting yielded mainly discrete platelets. B) Low blotting strength induces formation of large bundles with platelets stacking. C) At the lowest blotting strength the bundles composed of platelets show no edge features suggesting that the platelets tend to merge (or fuse). Scale bars 200 nm.

bundles (**Figure 3.8**B). With decreasing blotting strength, multiple platelets stack on top of each other and a twist along the main axis appears (**Figure 3.8**C and **Figure 3.7**A). When enough water remained after blotting, no edge features can be seen by TEM suggesting that the platelets tend to merge (or fuse).

3.2.2.3 Chirality transfer from the oligomers to the bundles

During the evaporation induced assembly, the platelets stack on top of each other and rotate in a fan-like manner suggesting a slight rotation between the (001) crystal faces (**Figure 3.7**E and **S3.12**). Thus, I hypothesized that the (001) surface of the platelets (or the interaction between these surfaces) plays a major role in the assembly process, determining the chirality of the bundles (**Figure 3.7**F). This hypothesis is in agreement with amyloid-type assemblies of peptides, in which amino acids at the termini are responsible for the supramolecular twist.[133,134] Chirality transfer processes from different parts of the monomeric units (core and/or termini) also have been reported for supramolecular polymer assemblies.[290]

To gain insight into the origin of the twist, I designed three different experiments. First, I analyzed the assembly of a $\mathbf{D_6}$ + $\mathbf{L_6}$ racemic mixture (**Figure 3.9**A). A racemic mixture of $\mathbf{D_6}$ and $\mathbf{L_6}$ was dissolved in water and lyophilized; the XRD profile showed an amorphous arrangement (**Figure 3.9**B). No platelets nor bundles were observed during TEM analysis of the $\mathbf{D_6}$ + $\mathbf{L_6}$ solution in water (1 mg/mL), suggesting an interaction between the two enantiomers that prevented crystallization. When the $\mathbf{D_6}$ + $\mathbf{L_6}$ racemic mixture was recrystallized using a DMSO to MeOH solvent switch method (see Experimental section 6.4.5), the powder XRD showed a prevalence of the cellulose II pattern and platelets were regenerated, but no chiral features were observed (**Figure 3.9**B and C, and **Figure S3.17**).

We then focused on the assembly of cellulose oligomers with mixed sequence of D- and L-Glc residues. Five new oligomers were synthesized to selectively change the residues placed at the termini of the sequence (**Figure 3.9**D). In the second experiment, we analyzed the assembly of $\mathbf{L_2D_4}$, $\mathbf{L_3D_3}$, and $\mathbf{LD_6}$ in which a D-Glc core was capped at the non-reducing end with L-Glc residues. In this scenario, both D- and L-Glc residues are displayed at the (001) surface due to the antiparallel arrangement of cellulose II (**Figure 3.9**E, *top*). When a large portion of D-Glc residues was replaced by their mirror-image L-Glc, such as in $\mathbf{L_2D_4}$ and $\mathbf{L_3D_3}$, crystallinity was lost and solubility drastically increased (**Table S3.1** and **Figure S3.4** and **S3.5**). In contrast, when only one L-Glc unit was placed at the non-reducing end of the oligomer, such as in $\mathbf{LD_6}$, the solubility was comparable to that of $\mathbf{D_6}$ and the cellulose II crystal packing preserved (**Table S3.1** and **Figure S3.4** and **S3.5**). TEM analysis showed $\mathbf{LD_6}$ platelets and bundles with similar size and shape as for the model system $\mathbf{D_6}$ (**Figure S3.18**). However, the introduction of L-Glc weakened the twisting tendency and no clear handedness could be deduced from TEM images (**Figure 3.9**E, *bottom* and **S3.19**).

In a third experiment, we focused on oligomers bearing L-Glc residues at both termini, such as in $\mathbf{LD_5L}$ and $\mathbf{LD_6L}$. In this scenario, only L-Glc residues are displayed at the (001) surface, while the core of the crystal is based on D-Glc units (**Figure 3.9**F, *top*). The solubility of those oligomers was comparable to the model system $\mathbf{D_6}$ (**Table S3.1**), however, unexpectedly, $\mathbf{LD_5L}$ and $\mathbf{LD_6L}$ displayed a cellulose IV_{II}-type XRD profile[291] (**Figure S3.4**) and assembled into square platelets with no bundle-type aggregates (**Figure S3.20** and **S3.24**). Recrystallization using a solvent switch method from DMSO to MeOH converted $\mathbf{LD_6L}$ to cellulose II, as confirmed by XRD (**Figure S3.5**), and generated the typical platelet morphology (**Figure S3.21** and **S3.22**). Similar to the second scenario, $\mathbf{LD_6L}$ bundles of platelets displayed attenuated chiral features and, in some cases, the bundles appeared to be left handed (**Figure 3.9**F, *bottom* and **S3.23**).

Overall, the data on $\mathbf{LD_6}$ and $\mathbf{LD_6L}$ suggest that perturbing the molecular nature of the surface of cellulose II platelets drastically affects the chirality of their assembly. Bundles with much less pronounced chiral features were observed. No complete inversion of bundle chirality was noticed in the case of $\mathbf{LD_6L}$, suggesting that not only the surface, but also the core of the platelets are responsible for inducing the chirality of the homochiral assemblies (i.e. $\mathbf{D_6}$).

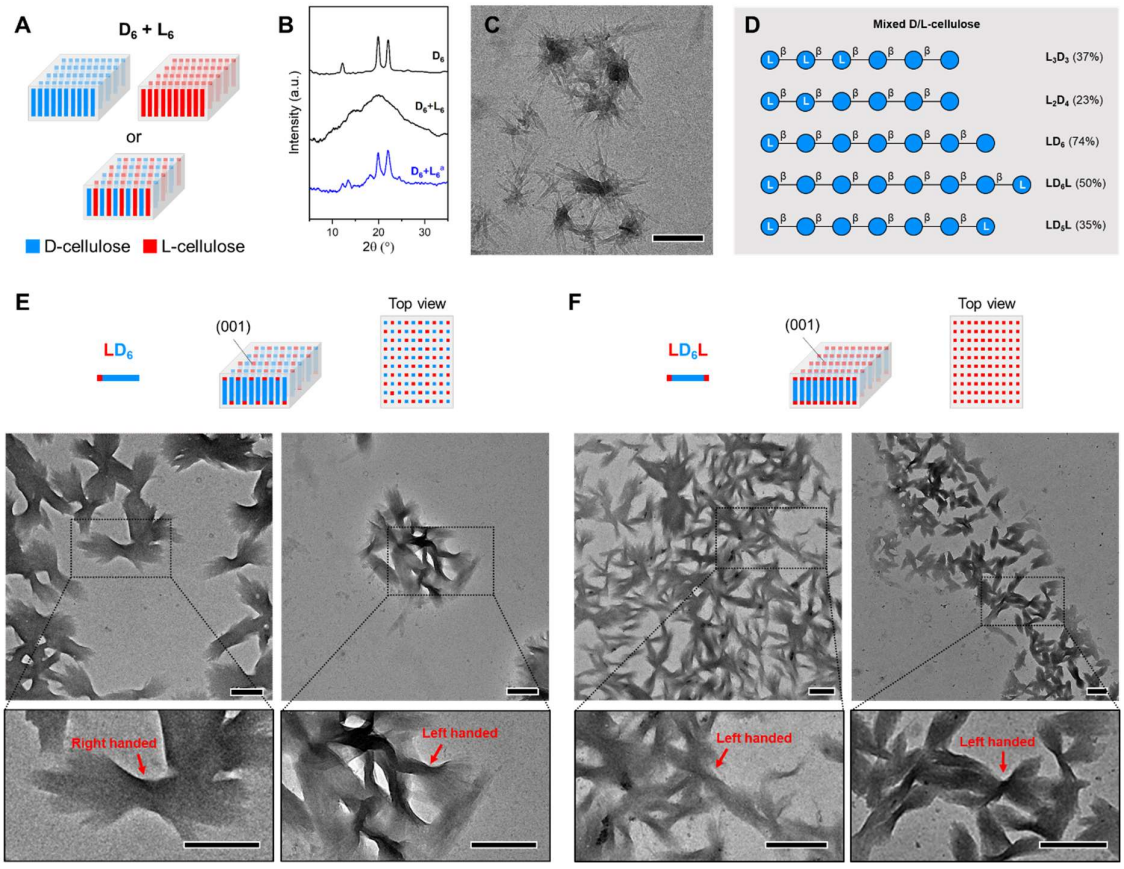


Figure 3.9 A) Simplified model of platelets formed by mixing **D**$_6$ and **L**$_6$. Self-sorting (*top*) or co-assembly (*bottom*) of the two enantiomers are two possible scenarios. B) Powder XRD profiles for **D**$_6$ + **L**$_6$ racemic mixtures. [a]The sample was prepared by recrystallization from DMSO/MeOH (see the Experimental section 6.4.5). C) TEM image of **D**$_6$ + **L**$_6$ (MeOH suspension) after recrystallization (scale bar, 500 nm). D) Hybrid D-/L-cellulose analogs were synthesized using protected monosaccharide BBs (overall yields reported in parentheses). BBs, solid support, and modules used for AGA and Post-AGA are reported in the Experimental section 6.2. E) Simplified model of **LD**$_6$ platelets and top view of the (001) face (*top*). TEM images of **LD**$_6$ bundles (aqueous suspension, *bottom*) showing a twisted morphology (*red arrows*, scale bars 500 nm). F) Simplified model of **LD**$_6$**L** platelets and top view of the (001) face (*top*). TEM images of **LD**$_6$**L** bundles (MeOH suspension, *bottom*) showing a twisted morphology (*red arrows*, scale bars 500 nm).

3.3 Conclusions

In this chapter, I presented how synthetic model systems can shine light on self-assembled carbohydrate materials and their chirality. The use of non-natural mirror-image monosaccharides (i.e. L-Glc) enabled to correlate molecular to supramolecular chirality. Starting from a simple self-assembling disaccharide **DD**, I developed a method to manipulate the chirality of the assemblies by introducing non-natural L-Glc residues. Helical fibers (left or right handed) as well as flat lamellae could be accessed on demand by using the disaccharides **DD**, **LL**, or a racemic mixture **DD-LL**. Furthermore, to validate the proposed crystal structure of **DD** fibers, I designed and synthesized site-specific modified analogs. Using this approach, I identified which modifications are tolerated and which disrupt key non covalent interactions preventing supramolecular assembly.

In the second part of these projects, I translated this *bottom-up* approach to a natural polysaccharide. I established a model system for the study of cellulose aggregation using synthetic oligomers produced by AGA. Cellulose oligomers of defined length and sequence, including D- and L-Glc, self-assembled into thin platelets with thickness matching the length of a single oligomer chain. Within the platelets, the oligomers arranged in a cellulose II fashion (antiparallel chain arrangement). The thickness of the platelet can be manipulated at the nanoscale by tuning the length of the synthetic cellulose oligomer. The synthetic cellulose platelets further assembled into bundles displaying intrinsic chiral features directly connected to the chirality of the cellulose chain (right handed for $\mathbf{D_6}$ and left handed for $\mathbf{L_6}$). Synthetic hybrid D- and L-cellulose oligomers helped to elucidate the origin of the chirality of these bundles. Terminal residues of opposite chirality drastically weakened the twisting tendency of the bundles, suggesting that the surface of the platelets plays an important role in determining the chirality of the bundles. Still, the surface does not solely determine the twist, as complete inversion of chirality was not observed for analogue displaying only L-Glc at the extremities.

Bundles of naturally sourced cellulose crystallites have been recently proposed to be a key element in the transfer of chirality across different hierarchical levels.[280,283] The model system presented herein offers a well-defined approach to validate this hypothesis that could be extended to other polysaccharides, such as chitin. In contrast to cellulose, the chiral nature of chitin nanocrystals (ChNCs) remains elusive[282] even though *in vitro* formation of chiral nematic phases[292] and helicoidal organization of chitin microfibrils have been reported.

4 Design and synthesis of a glycan hairpin

This chapter has been modified in part from the following article:

G. Fittolani, T. Tyrikos-Ergas, A. Poveda, Y. Yu, P.H. Seeberger, J. Jiménez-Barbero, M. Delbianco Synthesis of a glycan hairpin *Manuscript in preparation* **2022**

Specific contribution

Section 4.2.1 I performed the design of the hairpin and Dr. Theodore Tyrikos-Ergas the MD simulations.

Section 4.2.2 I performed the synthesis of the building blocks and the AGA of hairpin analogs. Dr. Yang Yu performed the synthesis of intermediate **S4.7**.

Section 4.2.3 I performed the NMR studies on all unlabelled analogs and Dr. Ana Poveda performed the NMR studies on the labelled analogs.

4.1 Introduction

The strong tendency of cellulose oligomers to interact with each other forming supramolecular assemblies (Chapter 3) made me realize this could be a starting point to develop more complex glycan assemblies. The non-covalent interactions taking place between cellulose chains appear to be strong, often leading to a high propensity of crystallization. I envisioned to exploit this feature of cellulose to design a glycan secondary structure. With similar approaches, synthetic peptides were designed to fold into defined secondary structure and self-assemble into materials with controllable properties (see Section 1.1 and 1.5).

Linear amino acid sequences encode the information needed to generate 3-D motifs such as helices, sheets, and turns, which can combine to give complex functional macromolecules (i.e. proteins). Using the same alphabet of 20 α-amino acids, it is possible to design molecules with engineered 3-D structures.[293] The development of non-natural amino acids expanded the catalogue even further and molecules with size, shape, and arrangement beyond the natural ones could be generated (i.e. foldamers).[162–164,294] Following the principle 'form follows function', foldamers were designed to recognize specific ligands[295] or penetrate cell membranes[164] as well as to perform catalysis[166,168] or to undergo supramolecular assembly[167]. Moreover, these designed systems largely improved our understanding of the factors influencing the stability of protein folding.[293,296]

In contrast, glycan structures capable of folding into defined 3-D shapes have never been designed, despite the potential numerous advantages. Glycans are constructed from a vast pool of more than 100 natural monosaccharides that generates linear or branched polymers, suggesting an enormous potential to access new conformations. The presence of many hydroxyl groups with well-defined orientations provides opportunities for site-specific functionalization. The tendency of naturally occurring polysaccharides to self-assemble into hierarchically organized materials is another attractive feature.[39] However, to date, the complex chemical synthesis of glycans and the lack of design principles have prevented access to glycans with predictable 3-D shapes. Glycans are historically considered flexible molecules, rarely exhibiting stable secondary structures,[44] with few exceptions for polysaccharides capable of adopting helix- (e.g. amylose) or ribbon-like (e.g. cellulose) conformations in solution.[297] Nevertheless, the central role of glycans in a biological setting is often tightly related to the conformation and presentation they adopt.[298] Recent studies report increasing evidence that even small glycans can adopt relatively well-defined conformation in solution, often stabilized by intramolecular hydrogen bonds.[67,71,83,299–306] The increased understanding of glycan behavior suggests that glycans capable of adopting stable 3-D architectures can be created.

Herein, I rationally designed a glycan capable of folding autonomously into a secondary structure motif to challenge the common view that sees glycans exclusively as flexible molecules.[44,65] Inspired by peptide model systems, I constructed a glycan that adopts a hairpin secondary structure motif in aqueous solution. The design was based on the combination of natural glycan structural elements and aided by Molecular Dynamics (MD) simulations. Automated Glycan Assembly (AGA) provided rapid access to a series of well-defined glycan sequences, including ^{13}C-labelled analogues to facilitate the structural and conformational analysis. NMR spectroscopy detected long-range inter-residue nuclear Overhauser effects (NOEs) that unequivocally confirmed the folded conformation of the synthetic glycan hairpin. This work demonstrates that it is possible to program glycans adopting defined conformation in aqueous solution,[307] opening opportunities for endowing glycans with new properties and functions.

4.2 Results and discussion

4.2.1 Design of a glycan hairpin

Peptide model systems have largely improved our understanding of the factors influencing protein folding and stability.[293,296] This is especially true in the case of protein β-sheets, for which folding rules could be deduced from model systems consisting of a hairpin (i.e. strand-loop-strand). In a hairpin, the β-turn acts as a β-sheet nucleator, placing the two strands adjacent to each other and favoring inter-strand interactions in a parallel[308] or antiparallel[309] fashion (**Figure 4.1**A). Intramolecular hydrogen bonds stabilize the turn unit and hold the two strands in a closed hairpin conformation (**Figure 4.1**A). Model systems such as double[160,310] or triple[157,159] stranded β-hairpins served as blueprints to gain molecular insights into protein folding[308,309,311–313] and have proven essential to study structural aspects of amyloid aggregation.[314] Beyond applications in biology, these systems have been employed to generate self-assembling nanomaterials[9,22,23] and hydrogels,[16,24] or as catalysts for asymmetric catalysis[315].

Given the breadth of the research on peptide hairpins, I decided to explore the design of an analogous secondary structure motif in glycans. The molecular nature of glycans, composed of hydroxyl-rich backbones, makes the design of a glycan folded secondary structure highly challenging in comparison to peptides and imposes the exploration of new design rules. Interactions with water are strong and drastically impact the conformation in solution, disrupting intramolecular hydrogen bonds. Long-range inter-residue interactions are scarce and the absence of hydrophobic sidechains limits the stability of glycan secondary structures.[44] Still, within the glycan realm, subtle non-covalent interactions can take place (**Figure 4.1**B), which have a large impact on macroscopic properties[188] or bioactivity.[298] These non-covalent interactions, combined with design rules developed for peptides, served as inspiration for the creation of a glycan hairpin.

A parallel glycan hairpin structure, composed of two strands attached to a turn, is a secondary structure motif absent in nature. The design requires 1) a suitable glycan loop adopting a turn conformation and 2) two stacking strands.[316] The similarity of the Lewis X (Le^x) trisaccharide conformation to that of naturally occurring peptide β-turns prompted us to explore this glycan motif as a starting point for the design of the turn unit. Le^x adopts a fairly rigid closed conformation in aqueous solution stabilized by hydrophobic interactions between the methyl group of L-Fuc and the β-face of Gal. Further stabilization arises due to a non-conventional CH···O hydrogen bond[317] between the H-5 of L-Fuc (L-Fuc-5) and the O-5 of Gal branches (**Figure 4.1**B).[303,304,306] This CH···O hydrogen bond forms a 10-membered atom ring analogous to naturally occurring peptide β-turns. Looking at the 3-D shape of Le^x, I noticed that further monosaccharide extensions at the OH-4 positions of the stacked residues could keep a parallel arrangement between the two new moieties, provided that OH-4 displays an equatorial orientation.[304] Therefore, the canonical Le^x trisaccharide loop was modified by converting β-1,4-Gal into β-1,4-Glc and α-1,3-L-Fuc into α-1,3-L-Rha, both presenting OH-4 in the equatorial orientation (**Figure 4.1**C, **3mer-I**). I hypothesized that the non-conventional hydrogen bond could still act as a key stabilizing force to keep in spatial proximity the two branches.[304] Two additional turn units were also designed: i) **3mer-II** is based on GlcNAc substituted with two Glc residues attached as β-1,4 and β-1,3 branches, lacking the stabilizing non-conventional H-bond, and ii) **3mer-III**, in which the branched GlcNAc residue was substituted by Glc, ideally preserving an optimal spatial arrangement of the two branches to engage in the CH···O hydrogen bond (**Figure 4.1**C).

As model for the stacking strands, I was inspired by the cellulose backbone, a polysaccharide consisting of a β-1,4-Glc repeating sequence (**Figure 4.1**B). Cellulose is a linear polysaccharide adopting a rigid rod conformation in solution stabilized by intramolecular hydrogen bonds between the OH-3 and O-5 of neighboring Glc residues. Strong intermolecular interactions such as hydrogen bonds and hydrophobic interactions between the C-H rich faces of Glc drive amphiphilic cellulose chains to self-assemble into highly ordered (insoluble) crystallites.[188,189] Thus, I selected oligomers of cellulose as strand sequences due to their propensity to assemble in a regular manner.[318,319]

Overall, the design includes a rigid turn-like glycan, stabilized by a non-conventional CH···O hydrogen bond, carrying two cellulose oligomer strands. Natural glycan structural elements are thus combined to generate an non-natural glycan hairpin. As target structures, I set a short pentasaccharide (**5mer-I**) and two nonasaccharides analogs (**9mer-I** and **9mer-II**) (**Figure 4.1**D).

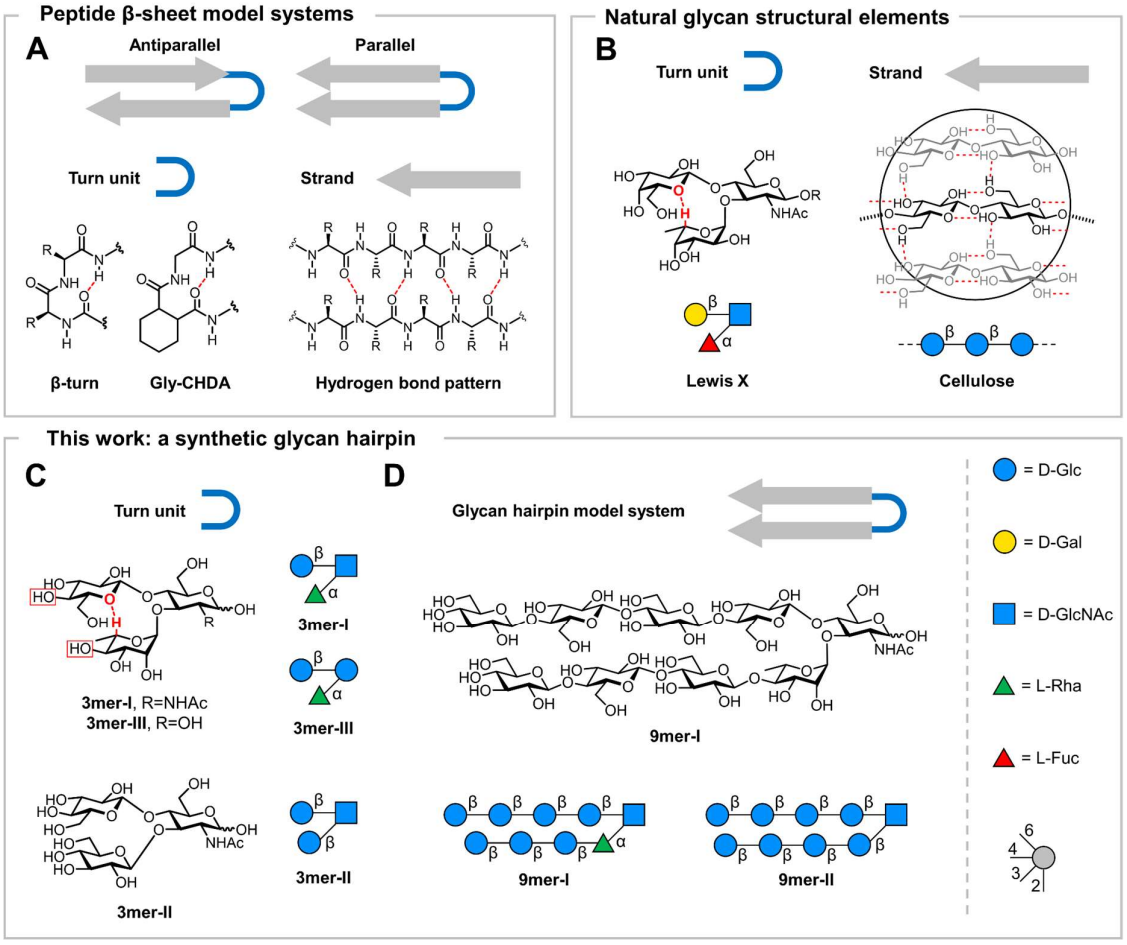


Figure 4.1 A) Autonomously folding anti-parallel or parallel peptide β-sheet model systems are composed of a turn unit (e.g. β-turn or Gly-CHDA) and two strands interacting *via* hydrogen bonds (*red dotted lines*, exemplified for parallel β-sheets). B) Naturally occurring glycans with turn-like conformation (i.e. Le^X trisaccharide) and strand-like conformation (i.e. cellulose) inspired the design of an artificial glycan hairpin. C) The glycan turn unit **3mer-I** was designed to preserve the non-conventional hydrogen bond (*red dotted line*) and optimally align two hydroxyl groups (*red boxes*) to elongate the strands of the hairpin. D) Synthetic glycan hairpin model system composed of two cellulose strands attached as branches to a trisaccharide turn unit. The following abbreviations are used for monosaccharides: Glc = glucose, Gal = galactose, GlcNAc = *N*-acetyl glucosamine, Rha = rhamnose, Fuc = fucose. The monosaccharide residues are represented following the Symbol Nomenclature for Glycans (SNFG) graphical representation.

Atomistic MD simulations were carried out to screen the tendency of the designed glycans to adopt the desired conformation. All the modelled structures were simulated for 500 ns, employing a modified version of the GLYCAM06[205] carbohydrate force field (see Experimental section 6.5.3). The systems were solvated with TIP5P[320] water model to avoid excessive interactions between the monomers.[206] I began by comparing **3mer-I**, **3mer-II**, and **3mer-III** to identify the optimal sequence for the turn unit. The MD trajectory suggested a clear difference in overall conformation between **3mer-I** and **3mer-II**. The latter displayed enhanced flexibility and extended conformation compared to **3mer-I**, as clearly indicated by the large fluctuations in the root mean square deviation (RMSD) plot and the higher value of the radius of gyration (R_g) (**Figure S4.6**). In contrast, **3mer-I** adopted a more rigid and compact turn-like conformation with a single conformer dominating nearly all the simulation time (**Figure 4.2**A, *top*). MD simulations pointed at

typical exo-syn-Φ/syn-Ψ conformations for all glycosidic linkages in **3mer-I** (**Figure 4.2**A, *bottom* and **Figure S4.7**). In contrast, Ramachandran plots for **3mer-II** suggested a two-state situation (i.e. closed *versus* open conformation, **Figure S4.11**) with the β-1,4 glycosidic linkage between Glc and GlcNAc fluctuating between exo-syn-Φ/syn-Ψ and non-exo-Φ/syn-Ψ conformations (**Figure 4.2**A, *bottom* and **Figure S4.8**). Similarly to **3mer-I**, **3mer-III** displayed a rigid closed conformation albeit slightly more flexible, likely due to the lack of the NHAc moiety (**Figure S4.6**, **S4.9**, and **S4.10**). Importantly, **3mer-I** displayed the hydroxyls planned for elongation of the two strands in an ideal orientation (**Figure 4.2**B).

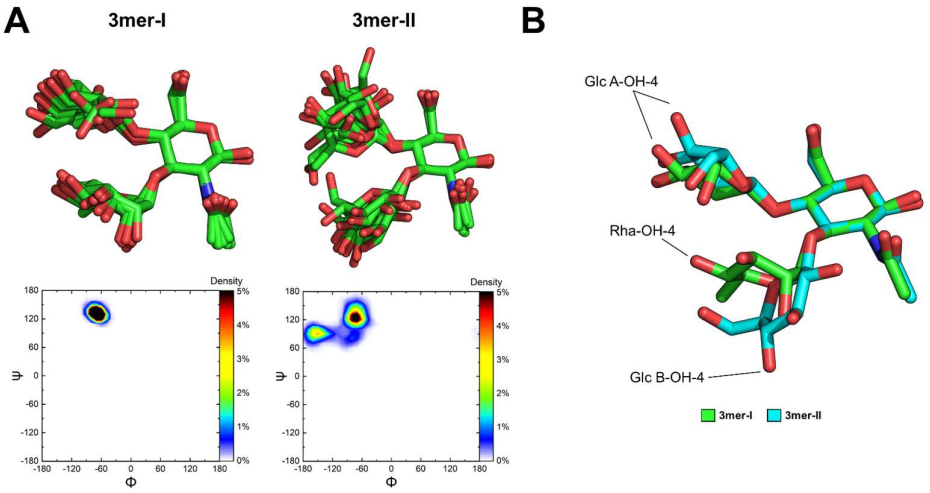


Figure 4.2 A) Over imposition of seven representative snapshots (*top*) and Ramachandran plots for the β-1,4-Glc linkage (*bottom*) extracted from the MD simulation show the larger degree of flexibility of **3mer-II** compared to **3mer-I**. B) Over imposition of the 3-D models (centroid structures) of **3mer-I** and **3mer-II**.

I then analyzed the hairpin analogs, **9mer-I** and **9mer-II**, to monitor how the turn unit impacted the overall conformation of the glycan hairpins. According to MD, only **9mer-I** adopted a closed hairpin conformation with the two cellulose strands in close proximity (**Figure 4.3**A). The turn unit based on **3mer-I** held the two cellulose strands in an optimal orientation resulting in small inter-residue distances across the strands (between 5.0 and 7.2 Å, **Figure 4.3**B, *top* and **Figure S4.12**). In contrast, in **9mer-II** the turn unit based on **3mer-II** did not hold the two strands in proximity and the two cellulose strands tended to stay further apart (up to 18.0 Å, **Figure 4.3**B, *top*). This tendency was also reflected in the R_g and RSMD plots which pointed at a rigid and compact shape for **9mer-I** *versus* a flexible, more extended shape for **9mer-II** (**Figure 4.3**B, *bottom* and **Figure S4.13**). MD simulations suggested also that larger glycan hairpins based on the **3mer-I** turn unit could in principle be obtained, even though larger fluctuations were observed for the residues located at the non-reducing termini (**Figure S4.14** and **S4.15**).

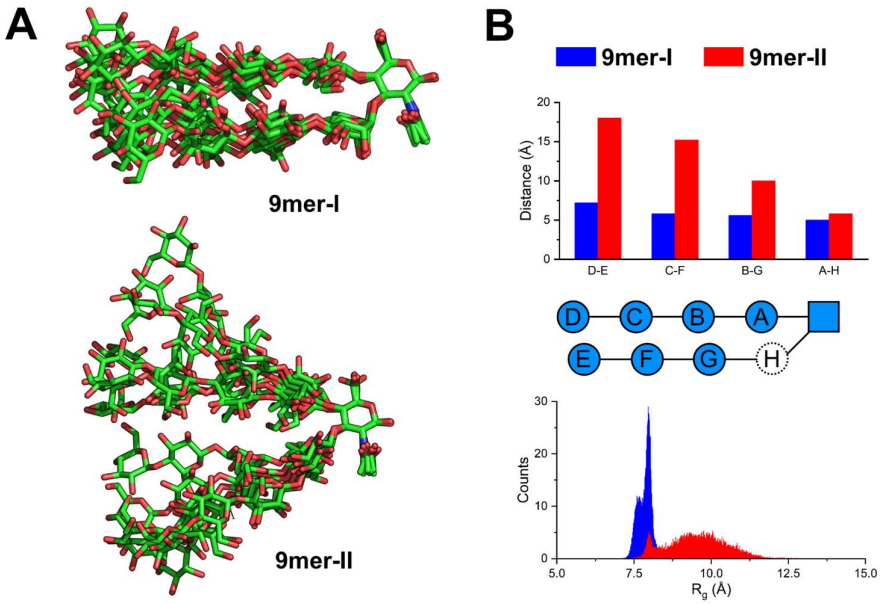


Figure 4.3 A) Over imposition of seven representative snapshot shows a closed hairpin conformation for **9mer-I** (*top*) in contrast to a highly flexible and open conformation for **9mer-II** (*bottom*). B) Inter-residue distances (*top*) and R_g (*bottom*) plots showing opposite trends for the two nonasaccharides regarding overall conformation and flexibility.

4.2.2 Synthesis of a glycan hairpin

To access the target compounds, I planned to use the BBs reported in **Figure 4.4** and AGA. All BBs are equipped with reactive anomeric thioether or phosphate leaving groups and permanent protecting groups (Bn, Bz, and TCA). **BB4.1** is equipped with two orthogonal protecting groups – levulinoyl ester (Lev) and 9-fluorenylmethoxycarbonyl (Fmoc) – allowing for the construction of branched structures. In addition, I designed and synthesized **BB4.4** to install Glc as an alternative branching unit (for further details see Experimental section 6.5.1). **BB4.2a** and **BB4.2b**, bearing a single temporary protecting group, were used to elongate the linear cellulose strands. For the installation of the key α-L-Rha unit of the turn unit, I designed a three-step synthetic route to access **BB4.3a** (**Scheme 4.1**). Solid supports **L1** and **L3** allowed for the synthesis of the target glycans with the free reducing end.[321]

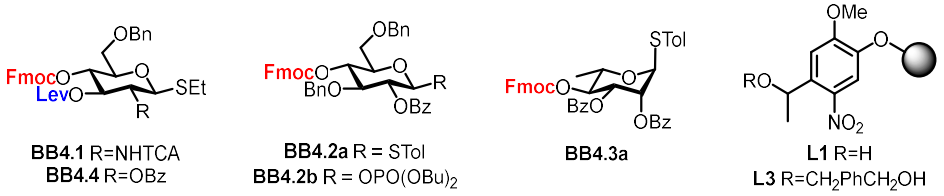
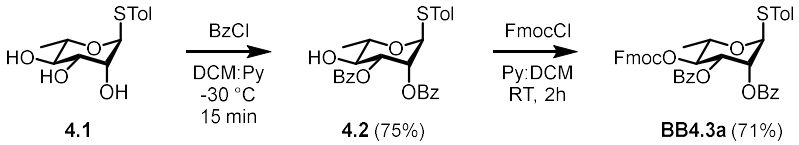


Figure 4.4 Set of BBs and solid supports used for the AGA of glycan hairpin analogs. Temporary protecting groups are highlighted in red and blue.

The synthesis of **BB4.3a** started from triol **4.1**. To achieve regio-selectivity during the benzoylation step, I exploited the higher reactivity of OH-3 and OH-2 of L-Rha when compared to OH-4. Highly regio-selective 2,3-*O*-benzoylation of **4.1** performed at low temperature (-30 °C) afforded **4.2**.[322] Subsequently Fmoc protection of OH-4 yielded **BB4.3a**.



Scheme 4.1 Synthesis of **BB4.3a**.

With these protected monosaccharide BBs in hand I moved on to the AGA of the target hairpin analogs. The synthesis proceeded from the branching monosaccharide towards the non-reducing end, elongating one strand at a time, following cycles of glycosylation and Fmoc deprotections (Module E1, see Experimental section 6.2). Once the first strand of the hairpin was constructed using **BB4.2a** or **BB4.2b**,[190,323] the terminal hydroxyl group was capped (Module D, see the Experimental section 6.2). Elongation of the second strand proceeded after Lev deprotection (Module E2).

Elongation of the α-1,3 branch of the hairpin glycan proved problematic during the synthesis of **5mer-I** and **9mer-I** as observed in the HPLC traces of the crude products after global deprotection (**Figure 4.5**A and **S4.4**). Low reactivity of the L-Rha-OH-4 solid bound acceptor was observed, resulting in the presence of a deletion side product lacking a Glc unit in the second branch, **4mer-I** (**Figure 4.5**A). I hypothesized two possible causes for the low nucleophilicity of L-Rha-OH-4: i) the bulky β-1,4 branch, installed on the GlcNAc branching unit, increased steric hindrance at the hydroxyl group and ii) the EWG nature of the 3-*O*-Bz group installed in **BB4.3a** decreased nucleophilicity through inductive effect. I designed an experiment to establish the cause of the low reactivity and thus devise a strategy to overcome it. In compound **4.3**, the α-1,3 branch of the hairpin glycan is elongated in absence of the bulky β-1,4 branch (**Figure 4.5**B). After subjecting **4.3** to one cycle of glycosylation, incomplete glycosylation was observed indicating that electronic effects play a major role in determining the low reactivity of L-Rha-OH-4 (**Figure 4.5**C).

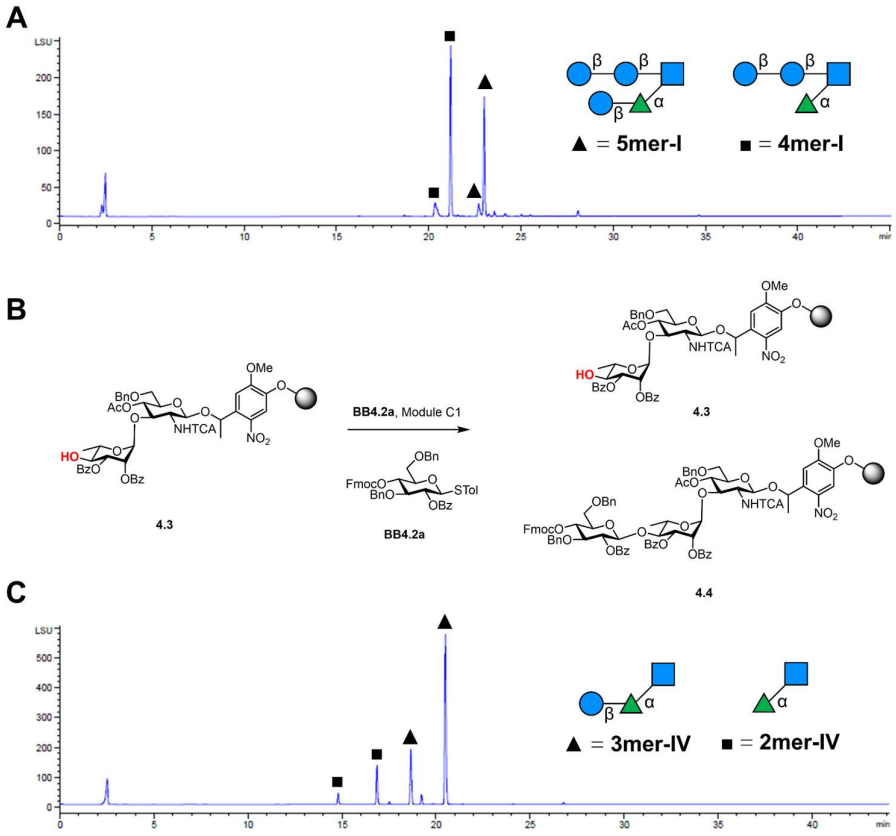
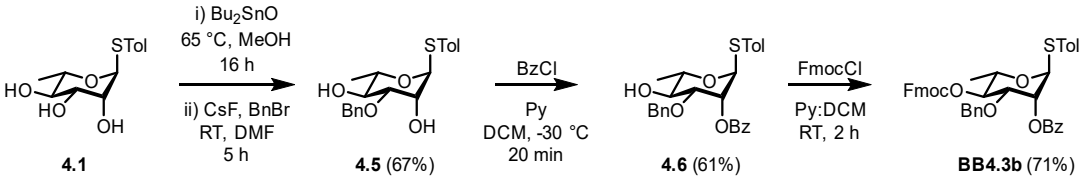


Figure 4.5 A) RP HPLC trace of the crude product after global deprotection for the synthesis of **5mer-I** using **BB4.3a** (Method A1). The two peaks correspond to α and β anomers of the free reducing end. B) Experiment to establish the cause of the low reactivity. C) RP HPLC trace of the crude product after global deprotection for the synthesis of **3mer-IV** using **BB4.3a** (Method A1). The two peaks correspond to α and β anomers of the free reducing end. For further details on Modules for AGA and purification method see Sections 6.2.3 and 6.2.4.

To increase the nucleophilicity of the L-Rha acceptor, I replaced the EWG 3-*O*-Bz group with an EDG 3-*O*-Bn. I designed a four-step synthesis to access **BB4.3b**, minimizing protecting group manipulations (**Scheme 4.2**). The key step is a tin mediated regio-selective benzylation of **4.1** to obtain **4.5**, which I adapted from a previously reported protocol developed for D-Man.[324] Subsequent regio-selective benzoylation of **4.5** to obtain **4.6** could be achieved thanks to the higher reactivity of OH-2 compared to OH-4. Installation of Fmoc afforded **BB4.3b**.



Scheme 4.2 Synthesis of BB4.3b.

The newly synthesized **BB4.3b** was tested in the AGA of **5mer-I** and the absence of deletion side products confirmed its excellent reactivity as a glycosyl acceptor (see Experimental section 6.5.2). With the optimized set of BBs, I synthesized the target compounds using AGA (**Figure 4.7**). A representative example of the AGA of **9mer-I** is reported in **Figure 4.6**.

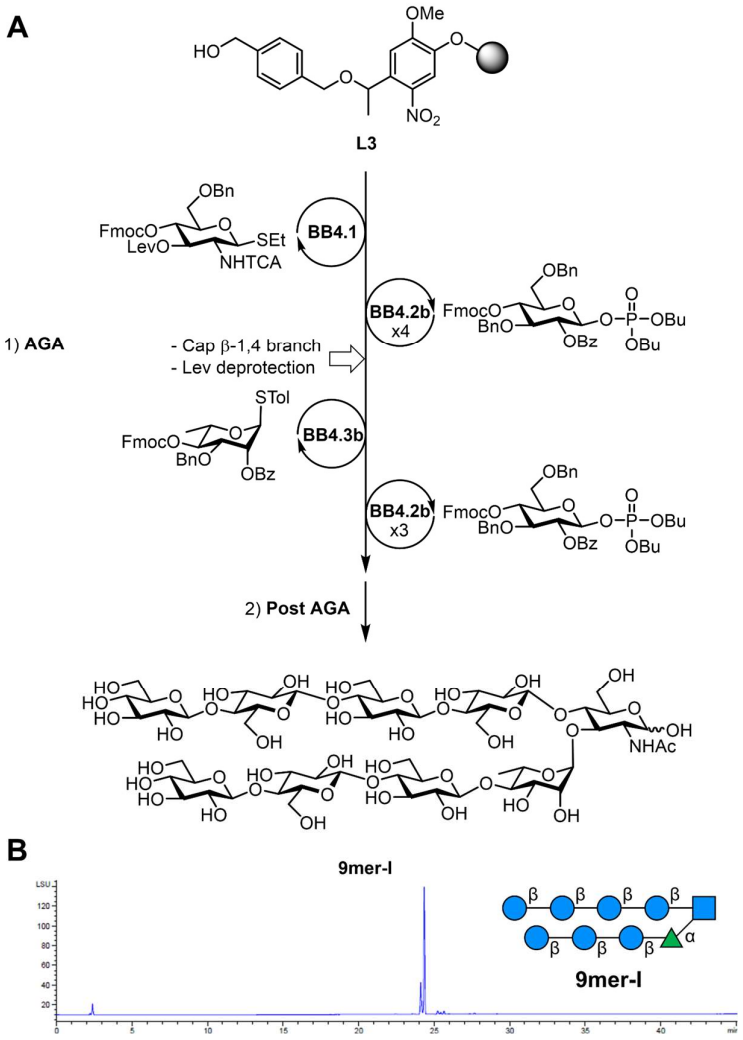


Figure 4.6 A) AGA of **9mer-I**. Post-AGA includes cleavage from the solid support, global deprotection steps, and final purification (see Experimental Section 6.5.2). B) RP HPLC trace of the crude product after global deprotection for the synthesis of **9mer-I** using **BB3b** (Method B1). The two peaks belong to the α and β anomers of **9mer-I**.

Each oligosaccharide was assembled in an overnight run using protected monosaccharide BBs. Post-AGA steps included solid-phase methanolysis, photocleavage from the solid support, and hydrogenolysis (see the Experimental section 6.2.4). The latter proved increasingly challenging for **5mer-I** and **9mer-I**, requiring multiple hydrogenolysis cycles to convert the *N*-trichloroacetyl (TCA) protecting group to the *N*-acetyl moiety. The turn units **3mer-I**, **3mer-II**, and **3mer-III** as well as the longer hairpin analogues **5mer-I**, **9mer-I** and **9mer-II** were obtained after a single final purification step in overall yields of 10 to 27% (**Figure 4.7**).

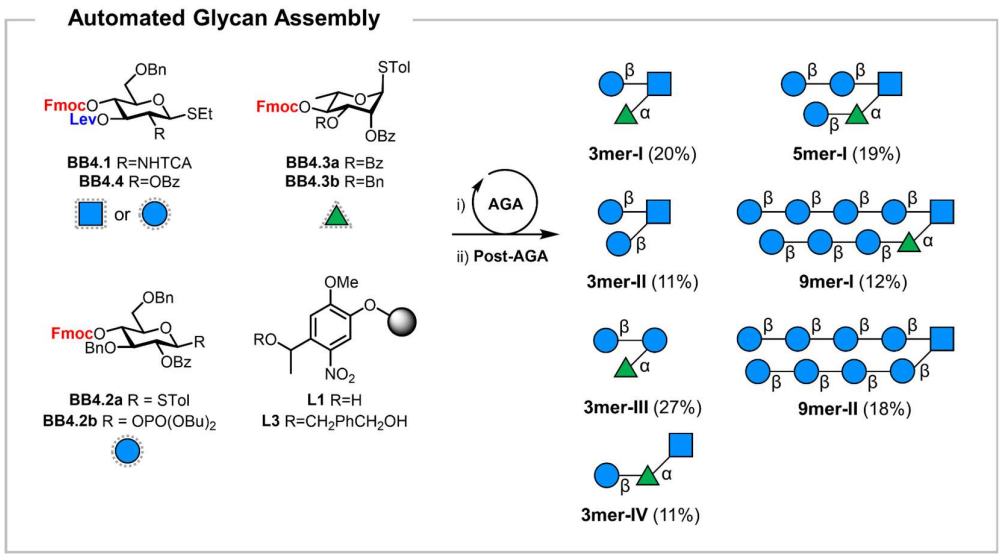


Figure 4.7 AGA of glycan hairpin model structures using protected monosaccharide BBs. Overall yields are reported in parentheses. Reaction conditions for AGA and Post-AGA are reported in the Experimental section 6.2.

4.2.3 NMR structural analysis

The structural analysis has been a major obstacle to the complete description of peptide β–sheet models. Circular dichroism (CD) provides simple and rapid information on the overall conformation, but lacks molecular details. X-ray crystallography affords molecular resolution, but fails to represent the range or ensemble of conformations that may be present in solution. NMR spectroscopy can capture the solution conformation(s) with molecular resolution and is to date the most reliable technique to characterize peptide β-sheet models.[313] Glycan structural studies are even more challenging because the lack of chromophores limits the use of CD, while difficulties in obtaining single crystals hinder the application of X-ray crystallography to glycans beyond a certain size (i.e. tetrasaccharides). NMR may provide some structural information, but suffers from severe overlap between the resonances of different residues and the scarcity of inter-residue and long range NOEs.[65,66,325]

A key feature of our turn is the non-conventional CH···O hydrogen bond forming a 10-membered ring that holds the two branches in an ideal parallel orientation. Such non-conventional hydrogen bonding has been hypothesized for rhamnosylated motifs,[304] but not investigated further.[326,327] I set out to verify this hypothesis using the chemical shift deviation (Δδ) of Rha-5 as an indicator.[302–304] I compared **3mer-I** and **3mer-III** with **3mer-IV**, in which no hydrogen bond can occur, observing a significant downfield shift of Rha-5 for the first two compounds (Δδ = 0.36 ppm, **Figure 4.8**A, see Experimental section 6.5.4 for details) providing solid ground to our hypothesis.

To confirm the spatial proximity between key residues at both sides of the putative hairpin, nuclear Overhauser effect spectroscopy (NOESY, see Section 1.2) experiments were employed. First, complete assignment of the resonances was carried out using a combination of HSQC and selective 1D TOCSY experiments (see Experimental section 6.5.4). Next, the 2D NOESY spectra were inspected for the presence of cross-peaks indicating through-space correlations between different protons (see Experimental section 6.5.4). The analysis of the NOESY spectra provided additional evidence that **3mer-I** adopts a closed turn-like conformation. In particular, besides the typical NOEs between protons across glycosidic linkages (Glc-1/GlcNAc-4 and Rha-1/GlcNAc-3), I observed key NOEs between the Glc and Rha moieties (Glc-2/Rha-5, Glc-2/Rha-6, Glc-4/Rha-5, and Glc-4/Rha-6) (**Figure 4.8**B, *top* and **Figure S4.26** and **S4.27**). Inter-proton distances of 2.9, 3.6, 3.6, and >4 Å, estimated by applying the Isolated Spin Pair Approximation to the NOE intensities, satisfactorily matched those calculated in the MD simulations. Similar NOEs between the key proton pairs were also observed for **3mer-III** suggesting a related conformation (**Figure S4.38** and **S4.39**). The NOE pattern observed for **3mer-I** (and **3mer-III**) is fully consistent with previous reports on analogous glycan motifs (including Le^x)[303,304,326] further corroborating the presence of a closed conformation. In contrast, only the standard NOEs across the glycosidic linkages

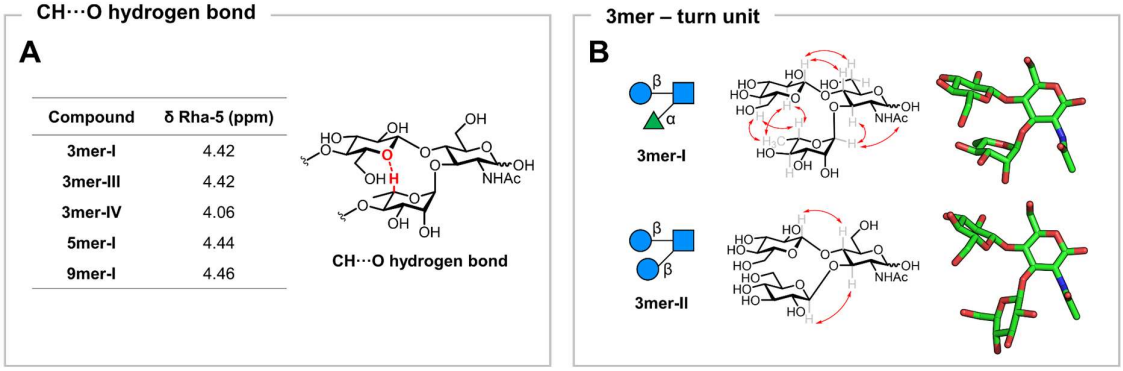


Figure 4.8 A) The chemical shift of Rha-5 reflects the presence of a non-conventional CH···O hydrogen bond. The table reports the chemical shifts for Rha-5. B) Experimental NOEs extracted from NOESY NMR experiments for **3mer-I** and **3mer-II** (*red arrows*).

were observed for **3mer-II**, in agreement with distances calculated in the MD model, indicating that the two branches are further apart compared to **3mer-I** (**Figure 4.8**B, *bottom* and **Figure S4.32** and **S4.33**).

The short hairpin model **5mer-I** was then scrutinized. The observed downfield shift of Rha-5 confirmed the presence of the non-conventional hydrogen bond ($\Delta\delta$ = 0.38 ppm, **Figure 4.8**A). The unfavorable tumbling regime of **5mer-I** prevented the use of NOESY experiments at the spectrometer frequency used in this study (i.e. 600 MHz) and ROESY experiments were necessary to observe positive NOEs. A transverse ROESY (t-ROESY) pulse sequence was used to minimize common artifacts (e.g. TOCSY transfer).[328] Both inter- and intra-residue NOEs are detected in the 2D t-ROESY NMR. To identify intra-residue NOEs, which are not useful for conformational analysis, I recorded a 2D TOCSY spectrum and overimposed it with the 2D ROESY. In **Figure 4.9** (*top*), the blue cross-peaks that are not overlapping with grey cross-peaks can be unambiguously assigned to an inter-residue NOE. The t-ROESY[329] experiment confirmed the presence of the key inter-residue NOEs between Glc A and Rha, identical to those observed for **3mer-I** (**Figure 4.9**A-K and **Figure S4.46**). The observed NOEs are summarized in **Figure 4.9** (*bottom*). These data proved that the closed conformation of the turn unit remains stable upon elongation of the two strands. To confirm that **5mer-I** folds into a hairpin conformation, additional evidence of the proximity between Glc B and Glc C on the two strands was required. The NOE between Glc B-1 and Glc C-2 was evident in the 2D t-ROESY (**Figure 4.9**J and **Figure S4.44**). To corroborate this evidence, I performed a selective 1D t-ROESY experiment and I observed the NOE between Glc B-1 and Glc C-2, corresponding to an inter-proton distance of ca. 3.0 Å, in full agreement with the MD simulations (**Figure 4.10** and **Figure S4.45**). As commonly recognized for peptides, the NOEs between residues not close in sequence are a strong indication that **5mer-I** adopts a hairpin conformation (**Figure 4.9**, *bottom*).

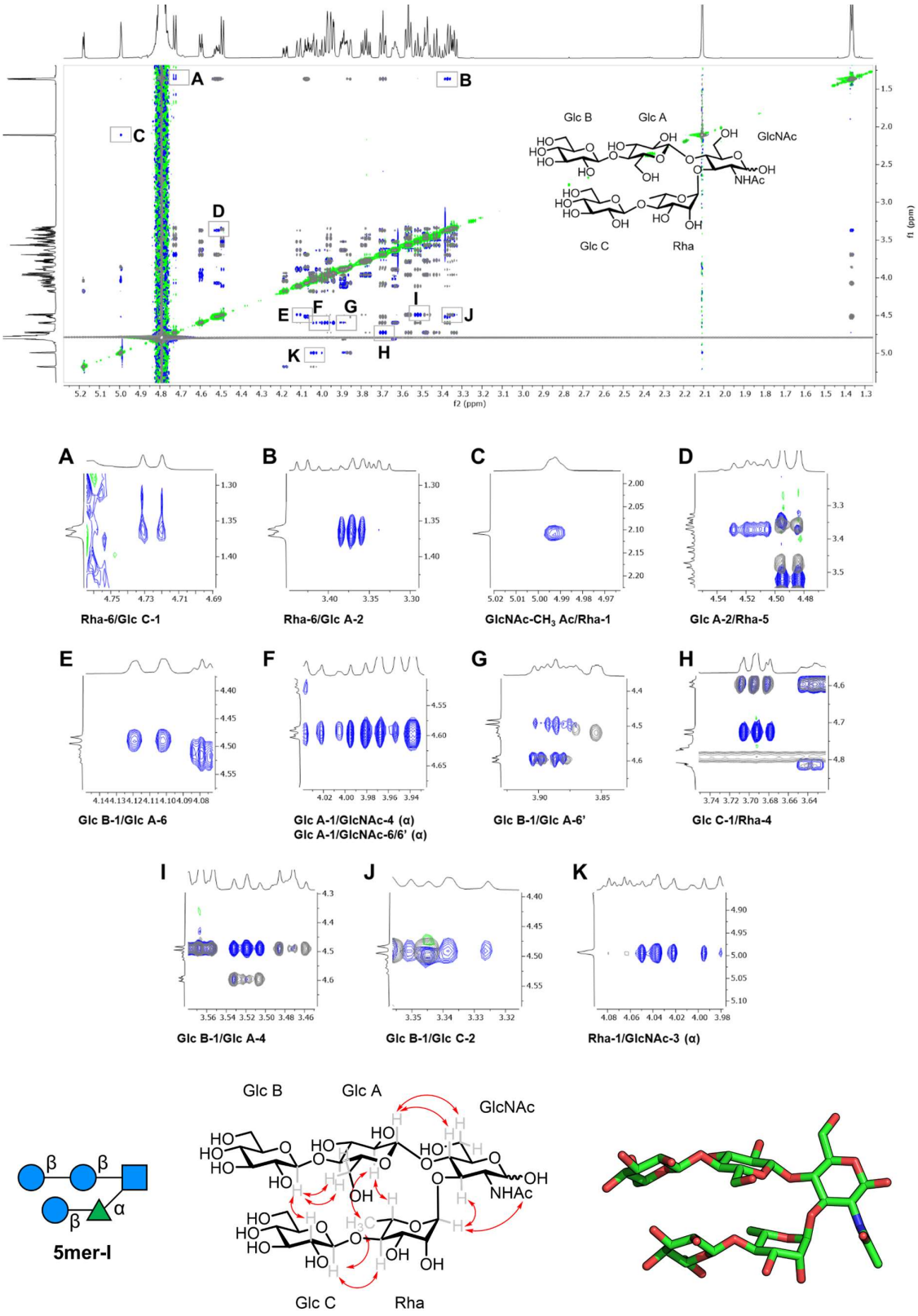


Figure 4.9 Overimposed 2D t-ROESY NMR (green-blue, 600 MHz, mixing time 300 ms, 300 K, D_2O) of **5mer-I** showing inter-residue NOEs and 2D TOCSY NMR (grey, 400 MHz, mixing time 120 ms, 300 K) (*top*). A-K) Excerpts of the 2D t-ROESY NMR. Experimental NOEs extracted from t-ROESY NMR experiments for **5mer-I** (*red arrows*) and snapshot of the MD simulation of **5mer-I** showing a hairpin conformation (*bottom*).

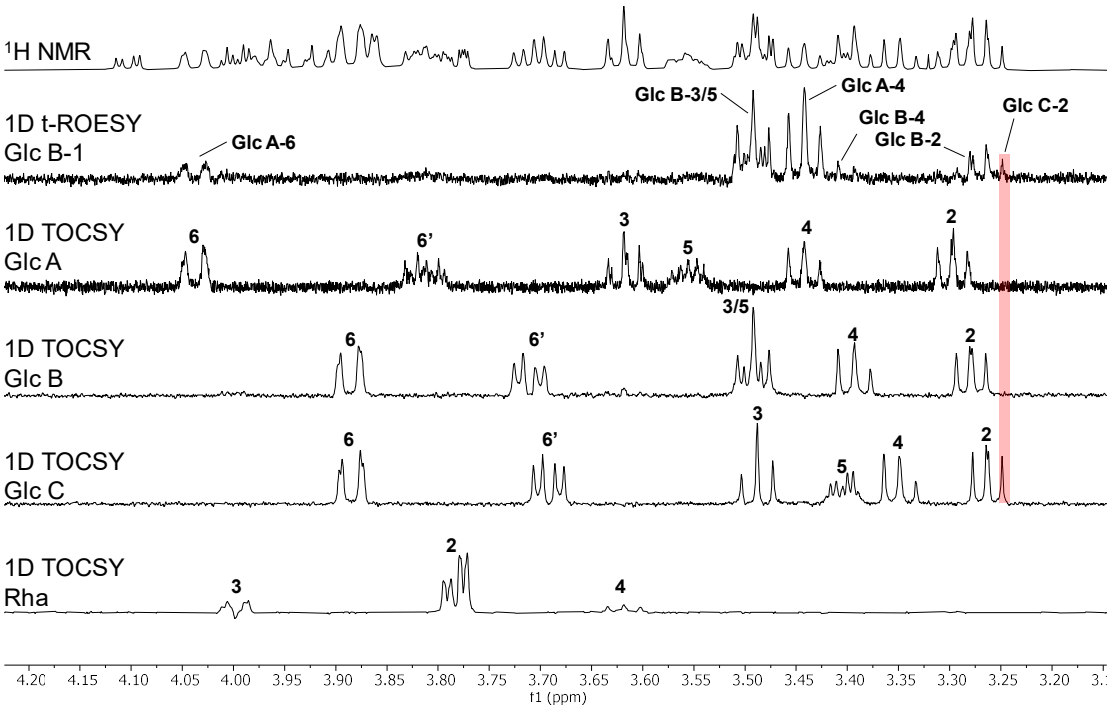


Figure 4.10 Overlay of selective 1D t-ROESY (600 MHz, mixing time 200 ms, 294 K, D_2O) and 1D TOCSY of **5mer-I**. The 1D t-ROESY was obtained by selective excitation of the Glc B-1 resonance (δ 4.41 ppm). The key NOE between Glc B-1/Glc C-2 is highlighted in a red box.

The next challenge was to demonstrate that the hairpin conformation is also kept in longer glycan structures. Thus, long-range inter-residue NOEs were also searched within **9mer-I**, as ultimate prove that this oligomer autonomously folds into a hairpin conformation. However, the severe chemical shift degeneracy hampered any standard structural analysis besides the observation of the hydrogen bond-induced downfield shift of Rha-5 ($\Delta\delta$ = 0.40 ppm, **Figure 4.8**A) and a key Rha-6/Glc A-2 NOE confirming the stability of the closed turn unit conformation (**Figure 4.11**B).

To break the chemical shift degeneracy, and thus perform the required NMR analysis, two selectively ^{13}C-labelled analogs of **9mer-I** were synthesized, employing **^{13}C-BB2b**, using AGA (see Experimental section 6.5.2). **9mer-I-B** and **9mer-I-F** present a $^{13}C_6$-labelled Glc at residues B and F of the hairpin structure, respectively (**Figure 4.11**A). The use of site-specific labelled analogs allowed extracting the key inter-residue NOE information otherwise inaccessible due to spectral overlap. Half-filtered $^{1}H[^{13}C]$, $^{1}H[^{12}C]$ NOESY experiments, only showing inter-residue NOEs, demonstrated the presence of the Glc B-1/Glc G-2 medium size NOE, corresponding to an inter-proton distance of 2.9-3.0 Å on **9mer-I-B** (estimated by applying the Isolated Spin Pair Approximation to the NOE intensities, **Figure 4.11**C), which was further confirmed by HSQC-NOESY (**Figure S4.47**). Next, HSQC-NOESY was recorded for **9mer-I-F** and another key medium size NOE (Glc F-6/Glc C-2, **Figure 4.11**D) was detected, corresponding to an inter-proton distance of ca. 4.0 Å. These data unambiguously show that, even for the nonasaccharide, the hairpin conformation is kept in solution (**Figure 4.11**E and F).

Further demonstration of the different 3-D shape of **9mer-I** *versus* the **9mer-II** analogue was deduced from diffusion-ordered NMR spectroscopy (DOSY) experiments independently carried out for both molecules at the same concentration (**Figure 4.11**G). A significant difference in the diffusion coefficients estimated for both molecules (2.03 10^{-10} m^2/s for **9mer-I** and 1.77 10^{-10} m^2/s for **9mer-II**) was observed, indicating a much more compact structure for **9mer-I**. According to the Stokes-Einstein equation that describes

DOSY experiments, the effective hydrodynamic volume of **9mer-II** is 50% larger than that of **9mer-I** (the ratio of hydrodynamic volumes is 1.52).

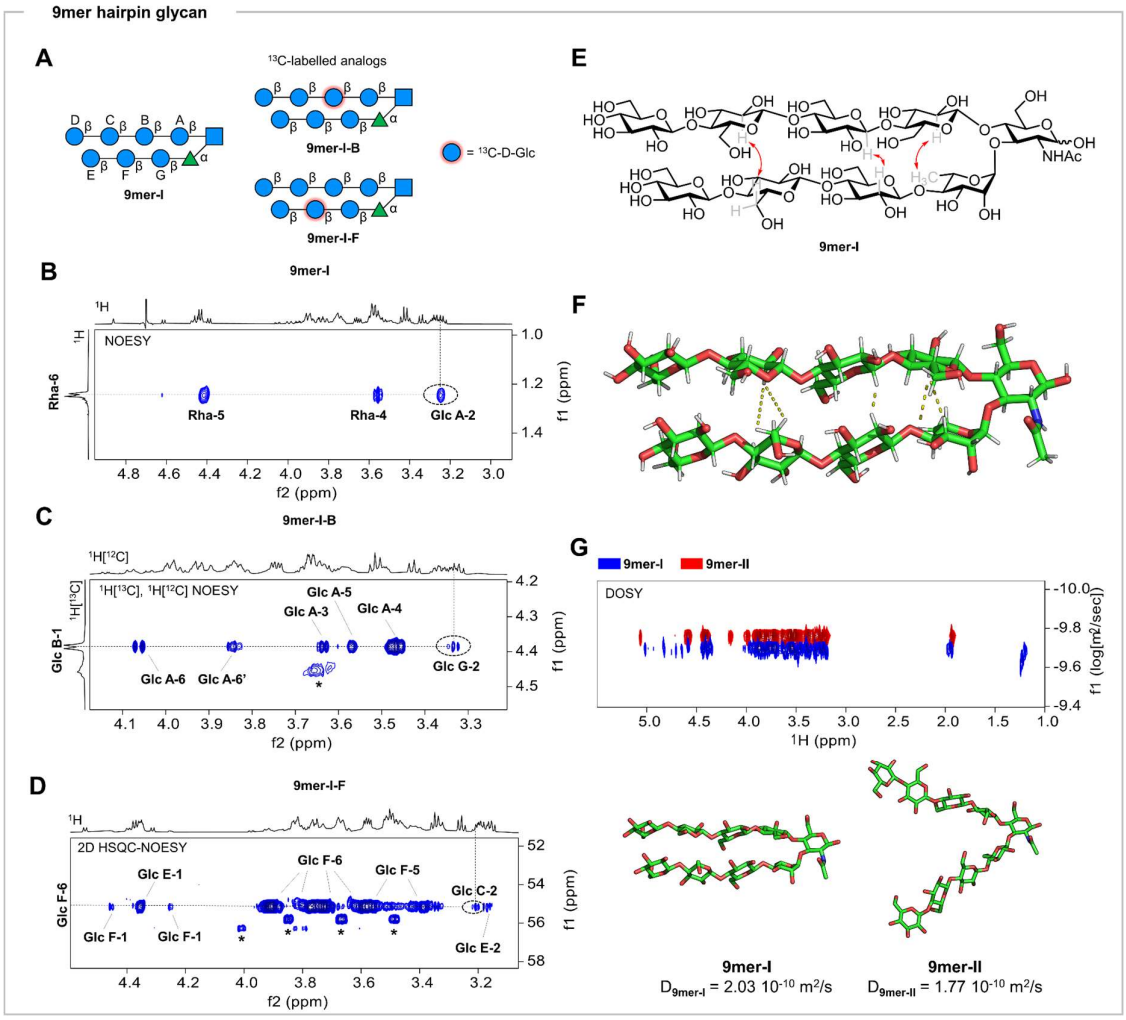


Figure 4.11 A) AGA enabled the rapid synthesis of two labelled analogs of **9mer-I**. For details about BBs, reaction conditions, and AGA see the Experimental section 6.2. B) Excerpt of the 2D NOESY NMR of **9mer-I** showing the key NOE (Rha-6/Glc A-2) at the turn unit (298 K, mixing time 300 ms, D_2O, 800 MHz). C) Excerpt of the half-filtered 2D $^1H[^{13}C]$, $^1H[^{12}C]$ NOESY NMR of **9mer-I-B** showing the typical inter-residue NOEs across the glycosidic linkages and the key NOE (Glc B-1/Glc G-2) between the two strands of the hairpin (298 K, mixing time 300 ms, D_2O, 800 MHz). The symbol ♦ indicates signals belonging to an impurity. D) Excerpt of the 2D HSQC-NOESY NMR of **9mer-I-F** showing the expected intra-residue NOEs and the key NOE (Glc F-6/Glc C-2) between the two strands of the hairpin (288 K, mixing time 300 ms, D_2O, 800 MHz). The symbol ♦ indicates signals belonging to an impurity. E) Selected experimental NOEs extracted from NMR conformational experiments for **9mer-I** (*red arrows*). F) Snapshot of the MD simulation of **9mer-I** showing a hairpin conformation. G) Superimposed DOSY spectra of 9mer-I (*blue*) and 9mer-II (*red*) (295 K, D_2O, 600 MHz). The corresponding diffusion coefficients are indicated.

4.3 Conclusions

In this chapter, I designed a glycan that folds into a secondary structure not found in nature. Employing design principles developed for creating peptide hairpin structures, I designed and synthesized a glycan hairpin. The hairpin design combined natural glycan structural elements to access a turn unit and two stacking strands. The glycan turn unit was inspired by the Le^x trisaccharide motif, adopting a closed conformation stabilized by a non-conventional hydrogen bond. The two strands were based on cellulose chains, adopting a rigid rod conformation stabilized by intramolecular hydrogen bonds and hydrophobic interactions. Atomistic MD simulations aided the design, by providing 3-D models of the glycans as a blueprint for synthesis and structural analysis. A collection of synthetic hairpin analogs, including ^{13}C-labelled oligosaccharides, was rapidly synthesized by AGA. The detection of long range NOEs unambiguously demonstrated the hairpin conformation adopted by the nonasaccharide **9mer-I**, and its shorter analogs. DOSY experiments further confirmed the propensity to adopt a closed and compact conformation.

Our work illustrates that glycan sequences capable of adopting specific secondary structure motifs can be designed. Glycans with predictable folded shapes will expand the catalogue of foldamer scaffolds. The vast alphabet of monosaccharides available might result in many structural motifs, beyond the commonly known natural geometries. New natural polysaccharides are continuously discovered and their unprecedented sequences could provide a source of secondary structure motifs.[330,331] In addition, non-canonical monosaccharides could be used as templates, inducing even shorter glycans to fold, similarly to what happened for peptides.[332] I imagine that the ability to control the conformation of glycans could lead to predictable functions and properties, with applications in catalysis and nanotechnology.

5 Conclusions and outlook

The ultimate goal of my thesis was to establish structure-function correlations for oligosaccharides and then translate this knowledge into design principles to generate self-assembling carbohydrate materials. *Bottom-up* approaches are underlying themes in my doctoral thesis and allowed me to expand our understanding of the principles dictating carbohydrate self-assembly. The knowledge acquired in the study of carbohydrate materials culminated in the rational design of a glycan secondary structure, not found in nature. This work will serve as a starting point for the future design of oligosaccharides with applications in material science and catalysis.

In the first part of this thesis (Chapter 2), I developed efficient synthetic routes to fluorinated Glc and GlcNAc BBs and assembled a broad collection of complex fluorinated glycans using Automated Glycan Assembly (AGA). These non-natural compounds were designed to manipulate hydrogen bonds in cellulose, a polysaccharide with high propensity to crystallize, and generated more flexible and soluble analogs. High-resolution imaging permitted the visualization of single molecules, confirming the higher flexibility predicted by Molecular Dynamics (MD) simulations. Furthermore, by exploiting the ^{19}F NMR channel, fluorinated oligosaccharides were used as probes for conformational studies and to screen for carbohydrate-protein interactions.

Currently, the effect of site-specific fluorination is being studied for chitin, using the 6F-GlcNAc BB I synthesized. Chitin, similarly to cellulose, has a strong tendency to crystallize, and the impact of fluorination pattern has never been explored. Preliminary results obtained in my group showed that chitin crystallinity is much less affected by fluorination compared to cellulose, indicating that these two polysaccharides, while sharing many structural similarities, might differ considerably in the factors influencing crystallinity (**Figure**

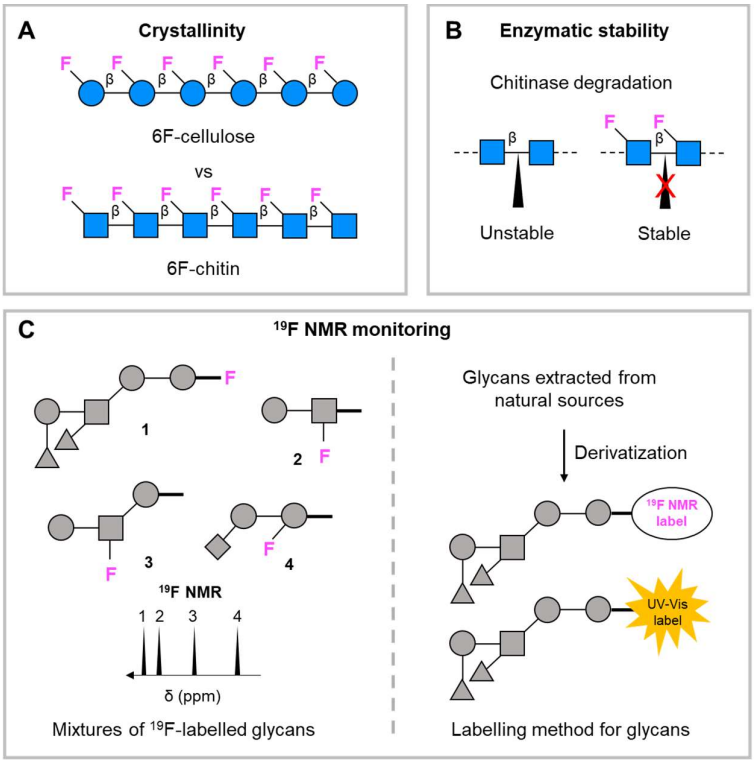


Figure 5.1 A) A comparison between the impact of OH/F substitution in cellulose and chitin could give an insight into the different factors dictating the crystalline behavior. B) Fluorinated chitin could be tested for its stability against chitinases. C) Mixtures of ^{19}F-labelled glycans prepared by AGA could be useful to increase the throughput of ^{19}F NMR binding assays (*left*). Labelling glycans extracted from natural sources with ^{19}F reporters could be exploited as an alternative method to labelling with UV-Vis chromophores (*right*).

5.1A). Site-specific fluorination of cellulose and chitin could also serve as a tool to manipulate enzymatic degradation. Cellulose and chitin are degraded by cellulases[333] and chitinases[334–336] and the specific insertion of fluorine could modulate the rate of degradation or even prevent it (**Figure 5.1**B). Experiments in which fluorination site and pattern are systematically varied might reveal enzyme preferences and identify candidates that are stable to enzymatic degradation. This would be particular interesting in the case of chitin, because chitooligosaccharides (COS) are known to trigger an innate immune response in humans and antifungal defense mechanisms in plants[337–339] in a size-dependent manner.[334] In addition, fluorinated monosaccharides, incorporated in a site-specific manner *via* AGA, will be key tools to provide structural information on other polysaccharides. ^{19}F-labelled oligosaccharides could also serve as probes for high-throughput screenings to identify acceptor substrates for enzymes or for the discovery of new glycosyltransferases.[245,246] Mixtures of site-specific ^{19}F-labelled glycans with different chemical shift could be designed for rapid real-time detection of binding to a protein or enzymatic reaction (**Figure 5.1**C, *left*).[226] The realization that the fluorine labels can be placed remotely from the glycan interacting site (Section 2.2.4) could be exploited for labelling naturally extracted oligosaccharides (**Figure 5.1**C, *right*). Typically, naturally extracted oligosaccharides are labelled (or tagged) with UV-Vis active chromophores to carry out protein binding studies.[340,341] ^{19}F-labelling could serve as an alternative labelling method that allows using NMR spectroscopy for detection and benefits from the minimal impact of ^{19}F reporters.

In Chapter 3, I developed a *bottom-up* approach that uses a simple synthetic model system to correlate molecular and supramolecular chirality of carbohydrate materials. I drew a correlation between molecular and supramolecular chirality by exploiting non-natural mirror image monosaccharides (i.e. L-Glc). I started with a simple disaccharide (Section 3.2.1) and developed a method to manipulate the chirality of the self-assembled twisted fibers. Next, I systematically modified the disaccharide to validate the proposed fiber crystal structure. In the second part of Chapter 3 (Section 3.2.2), I translated this *bottom-up* approach to a natural polysaccharide. I established a model system for the study of cellulose aggregation using synthetic oligomers produced by AGA. I discovered that cellulose oligomers of defined length and sequence, including D- and L-Glc, self-assembled into thin platelets with tunable thickness. The synthetic cellulose platelets further assembled into bundles displaying intrinsic chiral features directly connected to the chirality of the cellulose chain. Synthetic hybrid D- and L-cellulose oligomers helped to elucidate the origin of the chirality of these bundles, associated to both the surface and the core of the platelets.

Future projects can be envisioned from the work described in Chapter 3. In Section 3.2.2.3, I observed that certain hybrid D-/L-cellulose oligomers, in contrast to homochiral D-cellulose oligomers, displayed a square platelet morphology and a cellulose IV_{II}-type XRD pattern (**Figure 5.2**A). The cellulose IV_{II} crystal structure is currently the focus of a collaborative project with Dr. Yu Ogawa (CNRS-Cermav, Grenoble) aimed at using synthetic oligosaccharides to resolve the fine details of molecular association *via* MicroED (see Section 1.2). Revealing the molecular details of the particles obtained from synthetic cellulose oligomers could be interesting to understand the factors (e.g. primary sequence, crystallization conditions) that dictate shape and size of the aggregates. A broader systematic screening of D- and L-cellulose hybrid oligomers combined with electron diffraction and powder XRD analysis could be envisioned to correlate structure and properties. The knowledge gained from such a systematic study could be used in the future to design oligosaccharide sequences that self-assemble into particles with defined 3-D shape.

Another possible follow up project could exploit synthetic cellulose oligomers as a scaffold for designing self-assembling materials. The strong propensity of cellulose to stack and form aggregates could be exploited to direct the self-assembly, while other residues could be used to promote solubility and interactions with water (**Figure 5.2**B). In the field of supramolecular peptide materials, similar concepts have been exploited to design peptide amphiphiles.[6,20] Different length of the cellulose portion and the hydrophilic portion could be used to finely tune the shape and/or the property of the self-assembled object (e.g. fiber, hydrogel, platelet).[342,343]

The cellulose bundles described in Section 3.2.2.2 show a well-defined chirality and can be imagined as chiral dopant to direct the self-assembly of naturally sourced CNCs. The presence of crystallite bundles is emerging as a key factor in controlling CNCs self-assembly.[283] Currently, only left handed chiral nematic phases can be obtained. Synthetic analogs capable of generating enantiomeric bundles could be interesting materials to invert the chirality of CNCs films for photonic applications (**Figure 5.2**C).[270] Last, the *bottom-up* approach presented herein for cellulose could be extended to other polysaccharides, such as chitin, for which supramolecular chirality is not well-studied and the literature is still lacking.

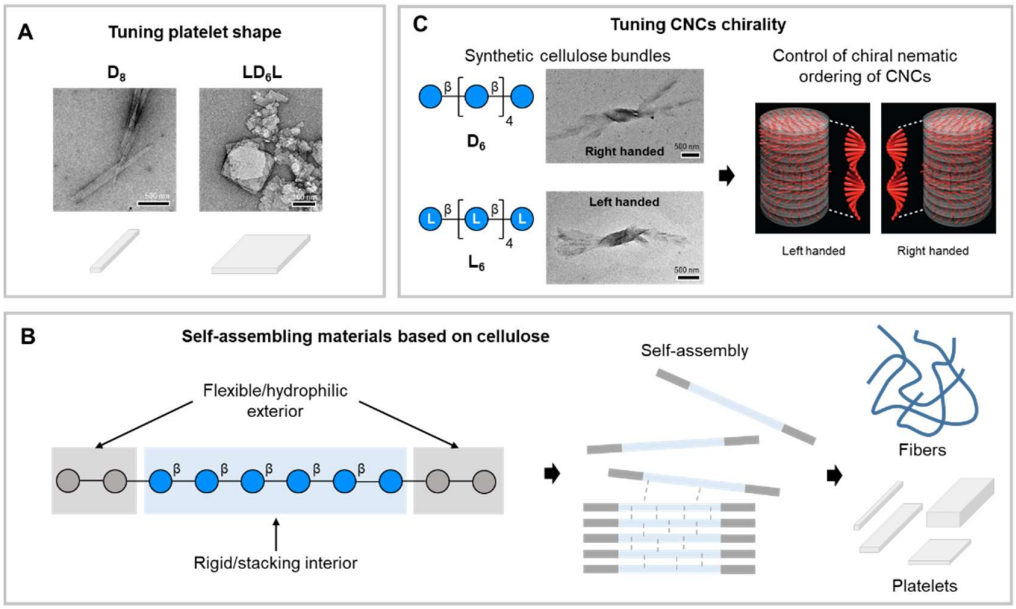


Figure 5.2 A) Platelets with diverse shape and size were observed for hybrid D- and L-cellulose sequences. Understanding the factors governing size and the shape of cellulose oligomer platelets could give access to materials with defined size and shape. B) The cellulose backbone could serve as a scaffold to generate self-assembling materials. The stacking tendency of cellulose could be used to guide the self-assembly. Materials with diverse shapes and properties could be created. C) Synthetic cellulose platelets that assemble into chiral bundles could serve as chiral dopants for naturally extracted CNCs.

In Chapter 4, I applied the knowledge gained in previous projects to rationally design a glycan capable of folding into a secondary structure in a way not found in nature. Employing design principles developed for creating peptide hairpin structures, I designed and synthesized a glycan hairpin that combines natural glycan structural elements. MD simulations assisted the design and guided the synthesis, *via* AGA, of glycan hairpin analogs. Finally, NMR validated the MD simulation and unambiguously demonstrated the hairpin conformation adopted by the nonasaccharide.

Further steps could involve the study the self-assembling propensity of the glycan hairpin (**Figure 5.3**A). The longest analog described in Chapter 4 (a nonasaccharide) is highly soluble in water, however longer analogs could potentially lead to aggregation. Powder XRD and TEM imaging could aid the characterization of the crystal structure and particle morphology obtained.

Another interesting avenue is the stabilization of the glycan hairpin motif (**Figure 5.3**B). A possible strategy is macrocyclization which has been widely explored for peptide hairpins,[309,344] but rarely applied to glycans.[345,346] The introduction of a salt bridge between the two strand of the hairpin could provide further stabilization.[347] Salt bridges are key stabilizing interactions in peptide hairpins, but have not yet been reported to contribute to the stability of glycan secondary structures.

Finally, a stable and simple glycan secondary structure can be imagined as a scaffold to design effective catalysts for organic reactions (**Figure 5.3**C). In the last decade, low molecular weight peptides and foldamers have been exploited as catalysts for a wide range of asymmetric transformations, site-selective functionalizations, and macrocyclizations.[166,315,348] Catalyst-substrate non-covalent interactions and catalyst conformation are crucial elements for efficient catalysis. The development of peptide β-turn catalysts resulted in a wide scope of catalytic transformation (**Figure 5.3**C, *left*).[315,348,349] A similar perspective can be imagined for the glycan turn units described in Chapter 4 (**Figure 5.3**C, *right*). Even though glycans that catalyze reactions are not known in nature, they have a number of attractive features such as the large number of monosaccharides available and the diverse spatial orientation of the hydroxyl groups that can be used to graft reactive functionalities. The high density of stereochemical information of glycans could be useful in asymmetric reactions. This feature could be exploited to design catalysts for asymmetric reactions or site-selective functionalizations.

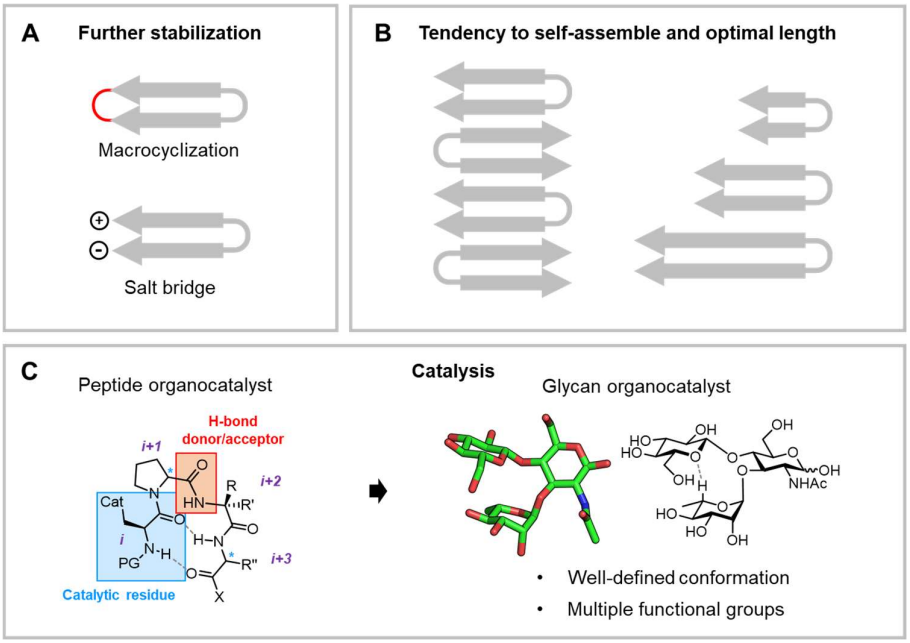


Figure 5.3 A) Further stabilization of the hairpin could be achieved using a covalent link between the two strands (e.g. macrocyclization) or through a non-covalent interaction, such as a salt bridge. B) Further studies are underway to understand the tendency to self-assemble of the glycan hairpin and how the length of the strands impact this behavior. C) Peptide β-turns have been exploited as organocatalysts (*left*). The well-defined conformation of the glycan turn unit could serve as a scaffold for generating organocatalysts (*right*).

6 Experimental section

6.1 General materials and methods

All chemicals used were reagent grade and used as supplied unless otherwise noted. The automated syntheses were performed on a home-built synthesizer developed at the Max Planck Institute of Colloids and Interfaces.[350] Analytical thin-layer chromatography (TLC) was performed on Merck silica gel 60 F254 plates (0.25 mm). Compounds were visualized by UV irradiation or dipping the plate in a staining solution (sugar stain: 10% H_2SO_4 in EtOH; CAM: 48 g/L ammonium molybdate, 60 g/L ceric ammonium molybdate in 6% H_2SO_4 aqueous solution). Flash column chromatography was carried out by using forced flow of the indicated solvent on Fluka Kieselgel 60 M (0.04 – 0.063 mm). Analysis and purification by normal and reverse phase HPLC was performed by using an Agilent 1200 series. Products were lyophilized using a Christ Alpha 2-4 LD plus freeze dryer. ^1H, ^{13}C and HSQC NMR spectra were recorded on a Varian 400-MR (400 MHz), Varian 600-NMR (600 MHz), Bruker Biospin AVANCE700 (700 MHz) Bruker AVANCE III 800 (800 MHz) spectrometer. Spectra were recorded in $CDCl_3$ or MeOD using the solvent residual peak chemical shift as the internal standard in ^1H and ^{13}C NMR ($CDCl_3$: 7.26 ppm ^1H, 77.0 ppm ^{13}C, MeOD: 3.31 ppm ^1H, 49.0 ppm ^{13}C) or in D_2O using the solvent as the internal standard in ^1H NMR (D_2O: 4.79 ppm ^1H). ^1H NMR spectra for all compounds were recorded without ^{13}C decoupling. Weak intensity ^{13}C resonances were derived from the respective HSQC crosspeaks. ^1H NMR integrals of the resonances corresponding to residues at the reducing end are reported as non-integer numbers and the sum of the integrals of α and β anomers is set to 1. High resolution mass spectra were obtained using a 6210 ESI-TOF mass spectrometer (Agilent) and a MALDI-TOF autoflexTM (Bruker). MALDI and ESI mass spectra were run on IonSpec Ultima instruments. IR spectra were recorded on a Perkin-Elmer 1600 FTIR spectrometer. Optical rotations were measured by using a Perkin-Elmer 241 and Unipol L1000 polarimeter.

6.2 Automated Glycan Assembly

6.2.1 General materials and methods

The automated syntheses were performed on a home-built synthesizer developed at the Max Planck Institute of Colloids and Interfaces. All solvents used were HPLC-grade. The solvents used for the building blocks, activator, TMSOTf and capping solutions were taken from an anhydrous solvent system (J.C. Meyer). The building blocks were co-evaporated three times with toluene and dried for 1 h on high vacuum before use. Oven-heated, argon-flushed flasks were used to prepare all moisture-sensitive solutions. Activator, capping, deprotection, acidic wash and building block solutions were freshly prepared and kept under argon during the automation run. All yields of products obtained by AGA were calculated on the basis of resin loading. Resin loading was determined following previously established procedures.[351]

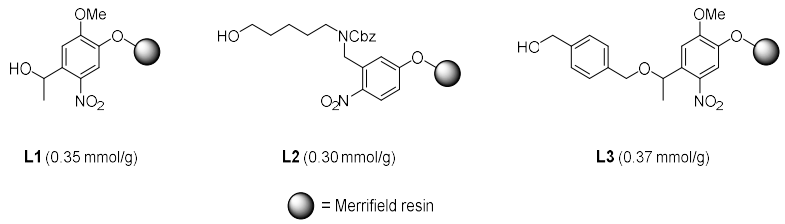
6.2.2 Preparation of stock solutions

- **Building block solution**: between 0.06 and 0.10 mmol of building block (depending on the BB, see Module C1 and C2) was dissolved in DCM (1 mL).
- **NIS/TfOH activator solution**: 1.35 g (6.0 mmol) of recrystallized NIS was dissolved in 40 mL of a 2:1 v/v mixture of anhydrous DCM and anhydrous dioxane. Then triflic acid (55 µL, 0.6 mmol) was added. The solution is kept at 0 °C (ice bath) for the duration of the automation run.
- **Fmoc deprotection solution**: a solution of 20%$_{v/v}$ piperidine in DMF was prepared.
- **Lev deprotection solution**: hydrazine acetate (550 mg, 5.97 mmol) was dissolved in pyridine/AcOH/H_2O (40mL, v/v, 32:8:2) and sonicated for 10 min.
- **TMSOTf solution**: TMSOTf (0.45 mL, 2.49 mmol) was added to DCM (40 mL).
- **Capping solution**: a solution of 10%$_{v/v}$ acetic anhydride and 2%$_{v/v}$ methanesulfunic acid in DCM was prepared.

6.2.3 Modules for automated synthesis

Module A: Resin preparation

All automated syntheses were performed on 0.0125 mmol scale. Resin (**L1**, **L2**, or **L3**) is placed in the reaction vessel and swollen in DCM for 20 min at room temperature prior to the synthesis. During this time, all reagent lines needed for the synthesis are washed and primed. After the swelling, the resin is washed with DMF, THF, and DCM (three times each with 2 mL for 25 s).



Module B: Acidic wash with TMSOTf solution (20 min)

The resin is swollen in 2 mL DCM and the temperature of the reaction vessel adjusted to -20 °C. Upon reaching the low temperature, TMSOTf solution (1 mL) is added dropwise to the reaction vessel. After bubbling for 3 min, the acidic solution is drained and the resin washed with 2 mL DCM for 25 s.

Action	Cycles	Solution	Amount	T (°C)	Incubation time
Cooling	-	-	-	-20	(15 min)*
Deliver	1	DCM	2 mL	-20	-

Deliver	1	TMSOTf solution	1 mL	-20	3 min
Wash	1	DCM	2 mL	-20	25 sec

*Time required to reach the desired temperature.

Module C1: Thioglycoside glycosylation (35 min-55 min)

The building block solution (0.10 mmol of BB in 1 mL of DCM per glycosylation) is delivered to the reaction vessel. After the set temperature is reached, the reaction is started by dropwise addition of the NIS/TfOH activator solution (1.0 mL, excess). The glycosylation conditions (T_1, T_2, t_1, and t_2) are building block dependent and are reported in for each synthesis. After completion of the reaction, the solution is drained and the resin is washed with DCM, DCM:dioxane (1:2, 3 mL for 20 s) and DCM (two times, each with 2 mL for 25 s). The temperature of the reaction vessel is increased to 25 °C for the next module. In case of a double cycle (C1*, *Double cycle), module C1 is repeated twice.

Action	Cycles	Solution	Amount	T (°C)	Incubation time
Cooling	-	-	-	T_1	-
Deliver	1	BB solution	1 mL	T_1	-
Deliver	1	NIS/TfOH activator solution	1 mL	T_1	-
Reaction time (BB dependent)	1	-	-	T_1 to T_2	t_1 t_2
Wash	1	DCM	2 mL	T_2	5 sec
Wash	1	DCM : Dioxane (1:2)	2 mL	T_2	20 sec
Heating	-	-	-	25	-
Wash	2	DCM	2 mL	> 0	25 sec

Module C2: Glycosyl phosphate glycosylation (45 min)

The building block solution (0.06 mmol of BB in 1 mL of DCM per glycosylation) is delivered to the reaction vessel. After the set temperature is reached, the reaction is started by dropwise addition of the TMSOTf solution (1.0 mL, stoichiometric). After completion of the reaction, the solution is drained and the resin washed with DCM (six times, each with 2 mL for 25 s). The temperature of the reaction vessel is increased to 25 °C for the next module. In case of a double cycle (C2*, *Double cycle), module C2 is repeated twice.

Action	Cycles	Solution	Amount	T (°C)	Incubation time
Cooling	-	-	-	-30	-
Deliver	1	BB solution	1 mL	-30	-
Deliver	1	TMSOTf solution	1 mL	-30	-
Reaction time (BB dependent)	1	-	-	-30 to -10	5 min 40 min
Wash	1	DCM	2 mL	-10	5 sec
Heating	-	-	-	25	-

Wash	6	DCM	2 mL	> 0	25 sec

Module D: Capping (30 min)

The resin is washed with DMF (two times with 2 mL for 25 s) and the temperature of the reaction vessel adjusted to 25 °C. A pyridine solution (2 mL, 10%$_{v/v}$ in DMF) is delivered into the reaction vessel. After 1 min, the reaction solution is drained and the resin washed with DCM (three times with 3 mL for 25 s). Capping solution (4 mL) is delivered into the reaction vessel. After 20 min, the reaction solution is drained and the resin washed with DCM (three times with 3 mL for 25 s).

Action	Cycles	Solution	Amount	T (°C)	Incubation time
Heating	-	-	-	25	(5 min)*
Wash	2	DMF	2 mL	25	25 sec
Deliver	1	10% Pyridine in DMF	2 mL	25	1 min
Wash	3	DCM	2 mL	25	25 sec
Deliver	1	Capping Solution	4 mL	25	20 min
Wash	3	DCM	2 mL	25	25 sec

*Time required to reach the desired temperature.

Module E1: Fmoc deprotection (9 min)

The resin is washed with DMF (three times with 2 mL for 25 s) and the temperature of the reaction vessel adjusted to 25 °C. Fmoc deprotection solution (2mL) is delivered to the reaction vessel and kept under Ar bubbling. After 5 min, the reaction solution is drained and the resin washed with DMF (three times with 3 mL for 25 s) and DCM (five times each with 2 mL for 25 s). The temperature of the reaction vessel is decreased to -20 °C for the next module.

Action	Cycles	Solution	Amount	T (°C)	Incubation time
Wash	3	DMF	2 mL	25	25 sec
Deliver	1	Fmoc depr. solution	2 mL	25	5 min
Wash	1	DMF	2 mL		
Cooling	-	-	-	-20	-
Wash	3	DMF	2 mL	< 25	25 sec
Wash	5	DCM	2 mL	< 25	25 sec

Module E2: Lev deprotection (90 min)

The resin is washed with DCM (three times with 2 mL for 25 s). DCM (1.3 mL) is delivered to the reaction vessel and the temperature of the reaction vessel is adjusted to 25 °C. Lev deprotection solution (2mL) is delivered to the reaction vessel, kept under pulsed Ar bubbling for 30 min. This procedure is repeated twice. The reaction solution is drained and the resin washed with DMF (three times with 3 mL for 25 s) and DCM (five times each with 2 mL for 25 s).

Action	Cycles	Solution	Amount	T (°C)	Incubation time
Wash	3	DMF	2 mL	25	25 sec
Deliver	2	Lev depr. solution	2 mL	25	30 min
Wash	1	DMF	2 mL	-	-
Cooling	-	-	-	-20	-
Wash	3	DMF	2 mL	< 25	25 sec
Wash	5	DCM	2 mL	< 25	25 sec

6.2.4 Post-AGA manipulations

Module F1: On-resin methanolysis
The resin is suspended in THF (4 mL). MeONa in MeOH (0.5 M, 0.4 mL) is added and the suspension is gently shaken at room temperature. After micro-cleavage (see *Module G2*) indicates the complete removal of benzoyl groups, the resin is repeatedly washed with MeOH (3 x 2 mL) and DCM (3 x 2 mL).

Module F2: Solution-phase methanolysis
The protected oligosaccharide is dissolved in MeOH : DCM (1.5 mL, 1:1). NaOMe in MeOH (0.5 M, 3 equiv. per benzoyl ester) is added and the solution stirred at room temperature, neutralized with Amberlite IR-120 (H+ form), filtered and concentrated *in vacuo*. The crude compound is used for hydrogenolysis without further purification.

Module G1: Cleavage from solid support
The oligosaccharides are cleaved from the solid support using a continuous-flow photoreactor as described previously.[109]

Module G2: Micro-cleavage from solid support
Trace amount of resin (around 20 beads) is dispersed in DCM (0.1 mL) and irradiated with a UV lamp (6 W, 356 nm) for 10 minutes. ACN (10 µL) is then added to the resin and the resulting solution analyzed by MALDI.

Module H1: Hydrogenolysis
The crude compound obtained from *Module G1* is dissolved in 2 mL of EtOAc:*t*BuOH:H_2O (2:1:1). 100% by weight Pd/C (10%$_w$) or $Pd(OH)_2$/C (10-20%$_w$, moistened with water) is added and the reaction stirred in a pressurized reactor under H_2 pressure (4 bar). The reaction progress is monitored to avoid undesired side products formation (*i.e.* degradation of reducing end).[352] Upon completion, the reaction is filtered (PTFE 0.45 µm 25 mm syringe filter, Fisher scientific) and washed with EtOAc, H_2O, and ACN (4 mL each). The filtrates are concentrated *in vacuo*.

Module H2: Hydrogenolysis at ambient pressure
The crude compound obtained from *Module G1* is dissolved in 2 mL of EtOAc:*t*BuOH:H_2O (2:1:1). 100% by weight Pd/C (10%$_w$) is added to the stirred flask, the reaction purged for 5 min with a N_2 balloon, and equipped with a H_2 balloon. The reaction progress is monitored to avoid undesired side products formation (*i.e.* degradation of reducing end).[352] Upon completion, the reaction is filtered (PTFE 0.45 µm 25 mm syringe filter, Fisher scientific) and washed with EtOAc, H_2O, and ACN (4 mL each). The filtrates are concentrated *in vacuo*.

Module I: Purification

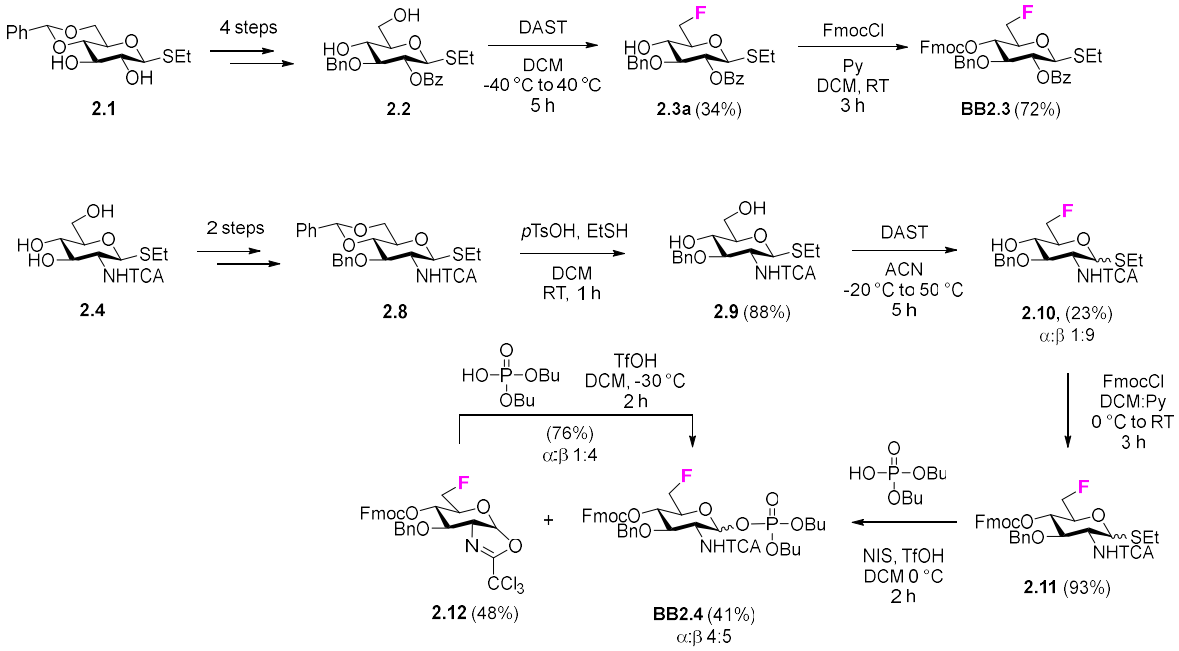
The final compounds are analyzed using analytical normal or reverse phase HPLC (Agilent 1200 Series, Methods A1, B1, and C1). The purification of the crudes is conducted using normal or reverse phase HPLC (Agilent 1200 Series, Method A2, B2, and C2).

- **Method A1**: (Hypercarb column, ThermoFisher scientific, 150 x 4.6 mm, 3 μm) flow rate of 0.7 mL/min with H_2O (0.1% formic acid) and ACN as eluents [isocratic (5 min), linear gradient to 30% ACN (30 min), linear gradient to 100% ACN (5 min), isocratic 100% ACN (5 min)].
- **Method A2 (Prep)**: (Hypercarb column, ThermoFisher scientific, 150 x 10 mm, 5 μm), flow rate of 3 mL /min with H_2O (0.1% formic acid) and ACN as eluents [isocratic (5 min), linear gradient to 30% ACN (30 min), linear gradient to 100% ACN (5 min), isocratic 100% ACN (5 min)].
- **Method B1**: (Hypercarb column, ThermoFisher scientific, 150 x 4.6 mm, 3 μm) flow rate of 0.7 mL/min with H_2O (0.1% formic acid) and ACN as eluents [isocratic (5 min), linear gradient to 60% ACN (30 min), linear gradient to 100% ACN (5 min), isocratic 100% ACN (5 min)].
- **Method B2 (Prep)**: (Hypercarb column, ThermoFisher scientific, 150 x 10 mm, 5 μm), flow rate of 3 mL/min with H_2O (0.1% formic acid) and ACN as eluents [isocratic (5 min), linear gradient to 60% ACN (30 min), linear gradient to 100% ACN (5 min), isocratic 100% ACN (5 min)].
- **Method C1**: (Synergi Hydro RP18 column, Phenomenex, 250 x 4.6 mm), flow rate of 1.0 mL/min with H_2O (0.1% formic acid) and ACN as eluents [isocratic (5 min), linear gradient to 30% ACN (30 min), linear gradient to 100% ACN (5 min), isocratic 100% ACN (5 min)].
- **Method C2 (Prep)**: (Synergi Hydro RP18 column, Phenomenex, 250 x 10 mm) flow rate of 4.0 mL/min with H_2O (0.1% formic acid) and ACN as eluents [isocratic (5 min), linear gradient to 30% ACN (30 min), linear gradient to 100% ACN (5 min), isocratic 100% ACN (5 min)].
- **Method D (Prep)**: (Manual reverse phase C_{18} silica gel column chromatography, 80 x 15 mm): H_2O (0.1% formic acid, 10 mL), 3% MeOH (10 mL), 6% MeOH (10 mL), 9% MeOH (10 mL), 12% MeOH (10 mL), 15% MeOH (10 mL).
- **Method E1**: (YMC-Diol-300 column, 150 x 4.6 mm), flow rate of 1.0 mL/min with Hexane and EtOAc as eluents [isocratic 20% EtOAc (5 min), linear gradient to 55% EtOAc (35 min), linear gradient to 100% EtOAc (5 min)].
- **Method E2 (Prep)**: (YMC-Diol-300 column, 150 x 20 mm), flow rate of 15 mL/min with Hexane and EtOAc as eluents [isocratic 20% EtOAc (5 min), linear gradient to 55% EtOAc (35 min), linear gradient to 100% EtOAc (5 min)].

Following final purification, all deprotected products are lyophilized on a Christ Alpha 2-4 LD plus freeze dryer prior to characterization.

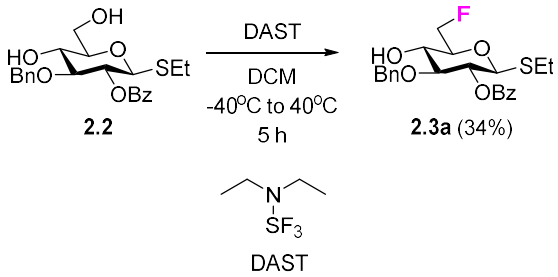
6.3 Experimental section for Chapter 2

6.3.1 Synthesis of fluorinated building blocks



Scheme S2.1 Synthesis of **BB2.3** and **BB2.4**.

Synthesis of **2.3a**

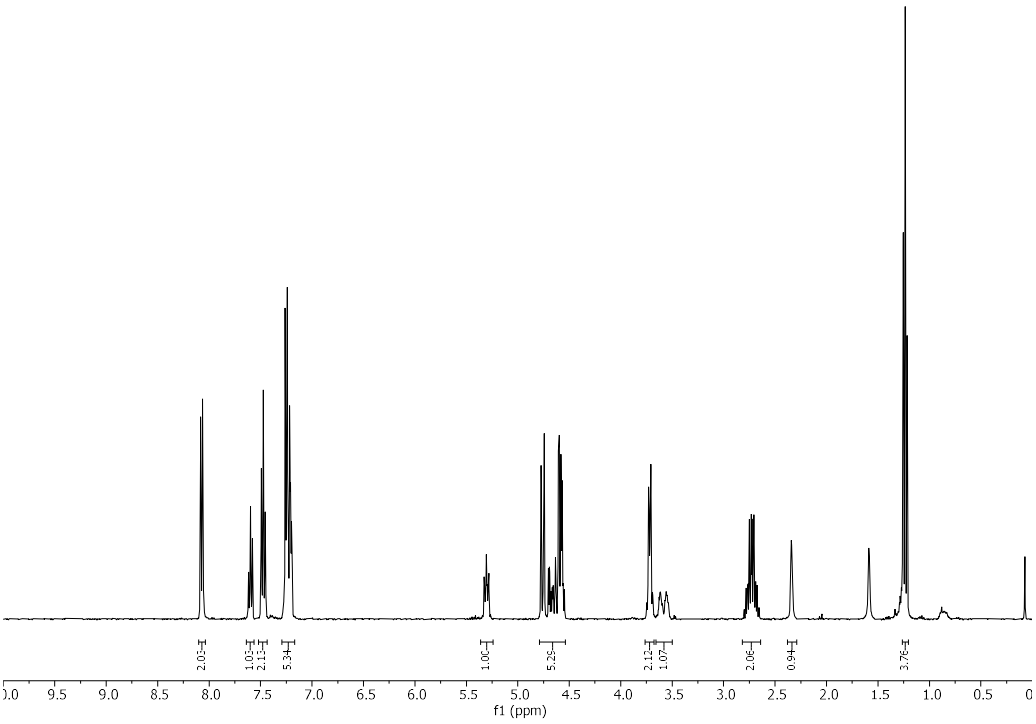


2.2 was prepared according to previously established procedures.[353]

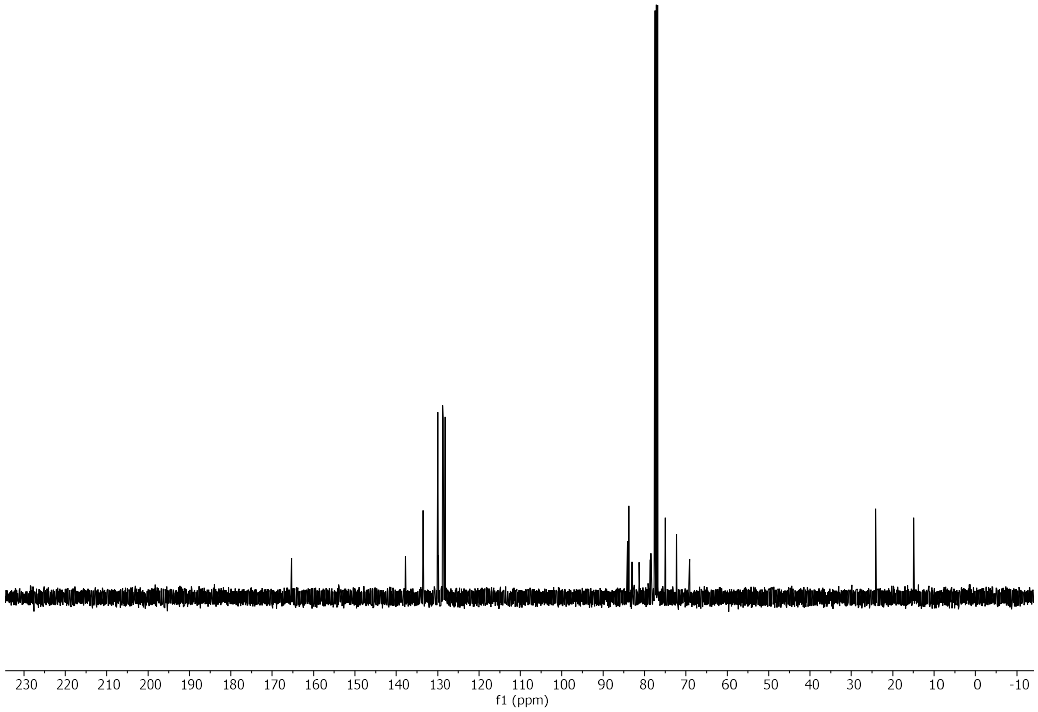
2.2 (490 mg, 1.17 mmol) was dissolved in anhydrous DCM (10 mL) and cooled to -40 °C (dry ice/ACN bath) under Ar atmosphere. DAST (171 µL, 1.29 mmol) was dissolved in anhydrous DCM (200 µL) and added dropwise to the reaction mixture. After 30 min, the cooling bath was removed and the reaction heated up to 40 °C. The solution was stirred for additional 5 h and then 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 **2.3a** as a colorless oil (167 mg, 34%).

^{1}H 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). $[\alpha]_D^{20}$ -18.12 (c 1.2 g/100 mL, $CHCl_3$). IR ν = 3482, 2927, 1724, 1268, 1086, 1070, 1027, 710, 700 cm^{-1}. (ESI-HRMS) m/z 443.1291 $[M+Na]^+$ ($C_{22}H_{25}FO_5SNa$ requires 443.1302).

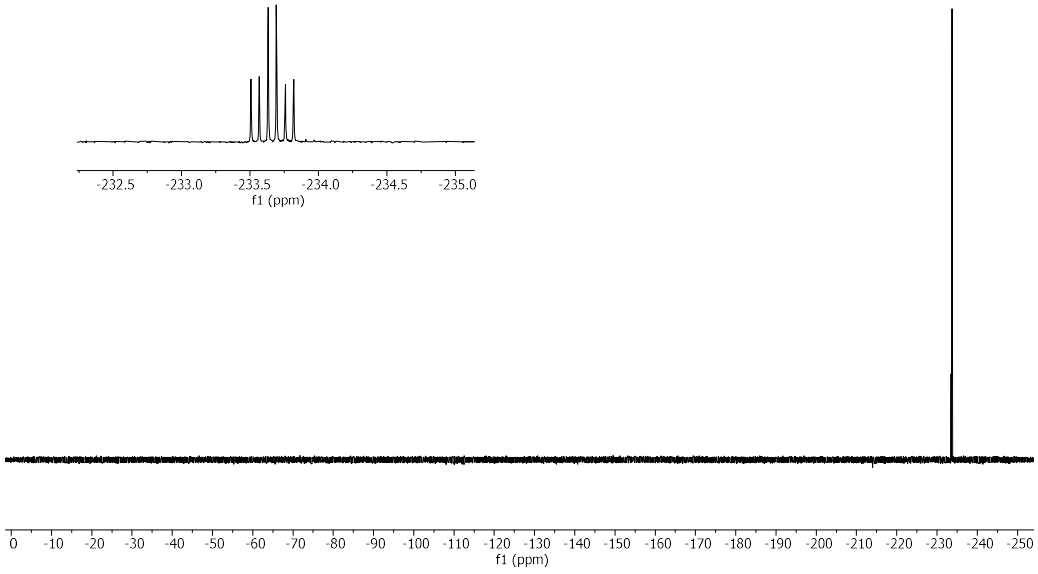
^{1}H NMR of 2.3a (400 MHz, $CDCl_3$)



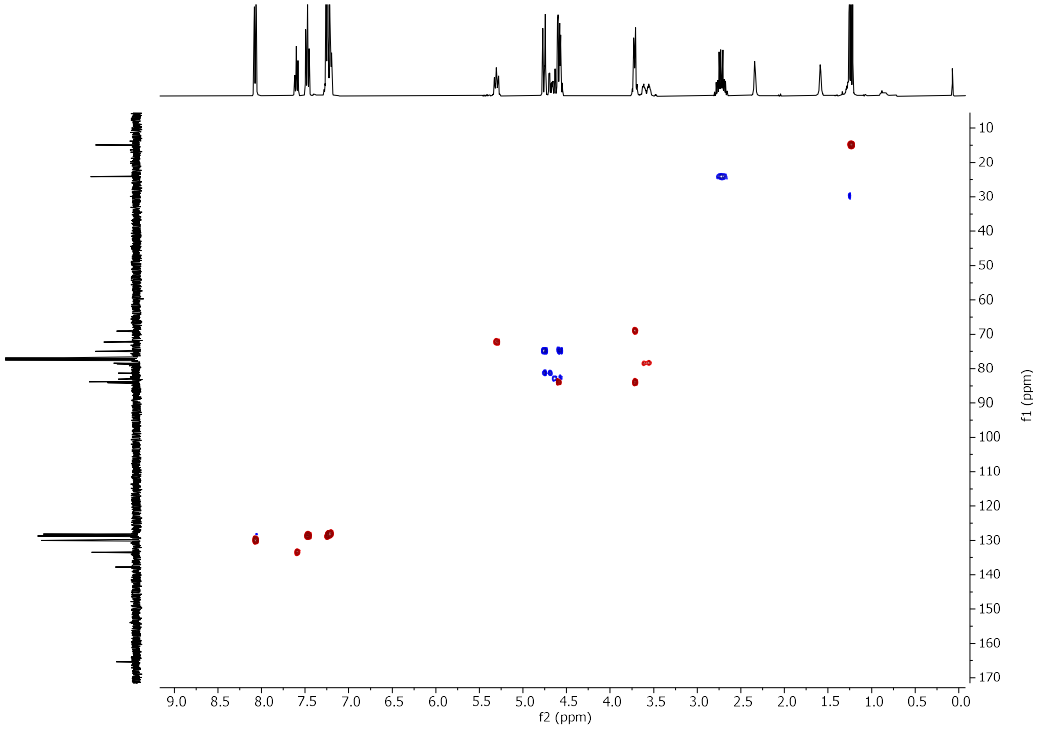
^{13}C NMR of 2.3a (101 MHz, $CDCl_3$)



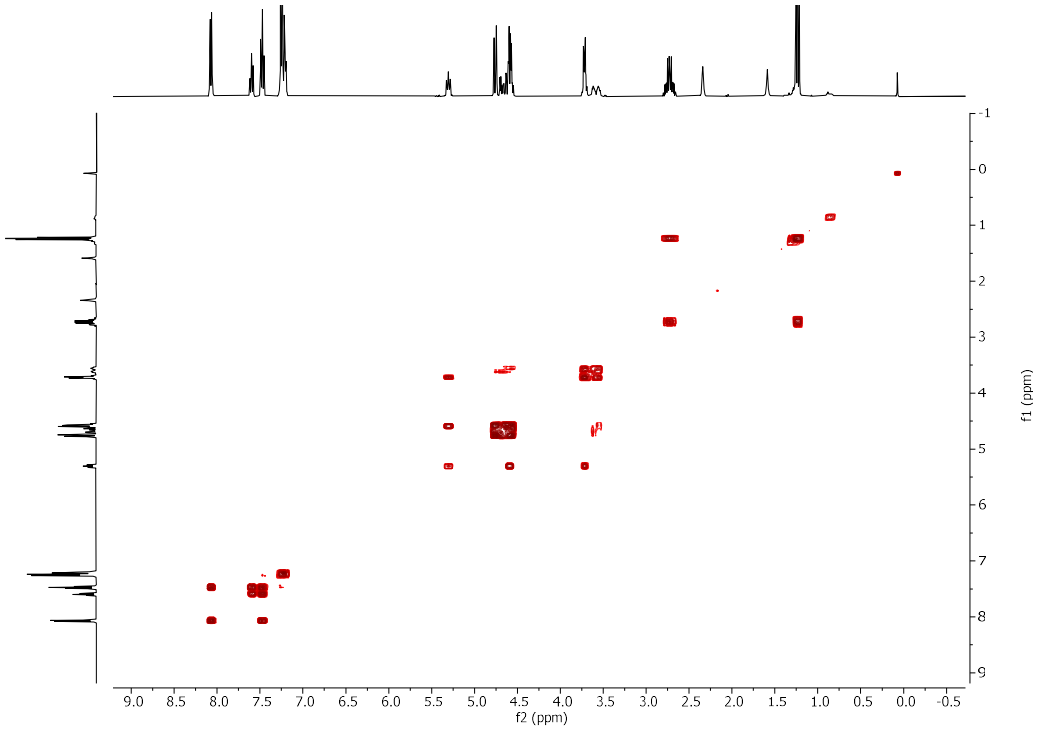
^{19}F NMR of 2.3a (376 MHz, $CDCl_3$)



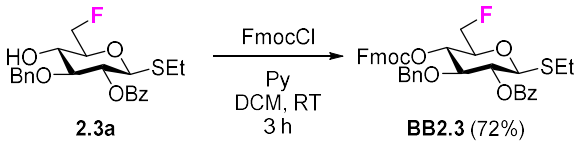
HSQC NMR of 2.3a ($CDCl_3$)



COSY NMR of 2.3a ($CDCl_3$)



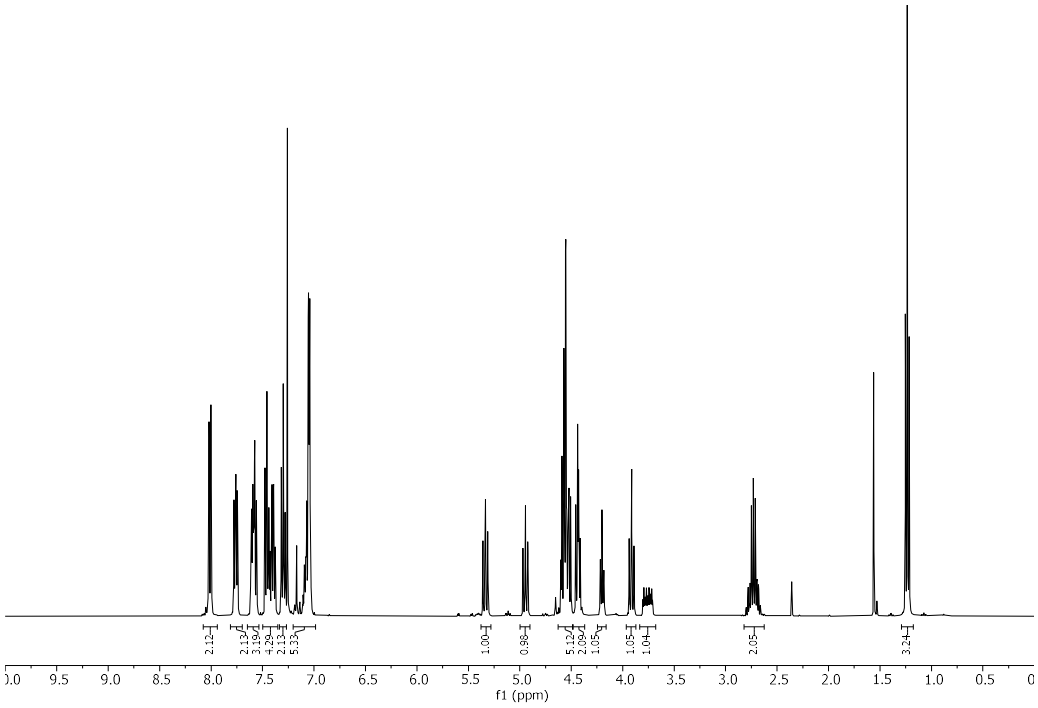
Synthesis of **BB2.3**



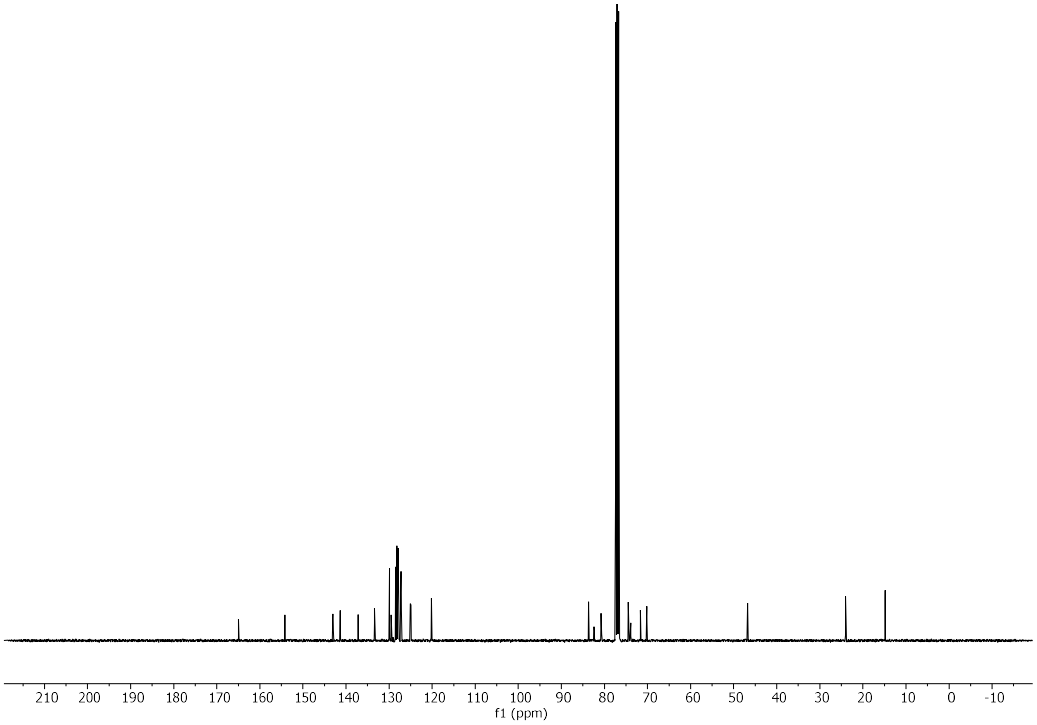
2.3a (167 mg, 0.40 mmol) was dissolved in DCM (5 mL) and pyridine was added (100 µL, 1.2 mmol). FmocCl (200 mg, 0.77 mmol) was dissolved in DCM (1.5 mL) and added to the reaction mixture at RT under Ar atmosphere. The solution was stirred for 3 h and then quenched with a 1 M solution of HCl. The crude 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→3:1 then Toluene : EtOAc = 4:1) and recrystallized from DCM : Hexane to give the **BB2.3** as a white solid (186 mg, 72%).

^{1}H 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). $[\alpha]_D^{20}$ 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 $[M+Na]^+$ ($C_{37}H_{35}FO_7SNa$ requires 665.1980).

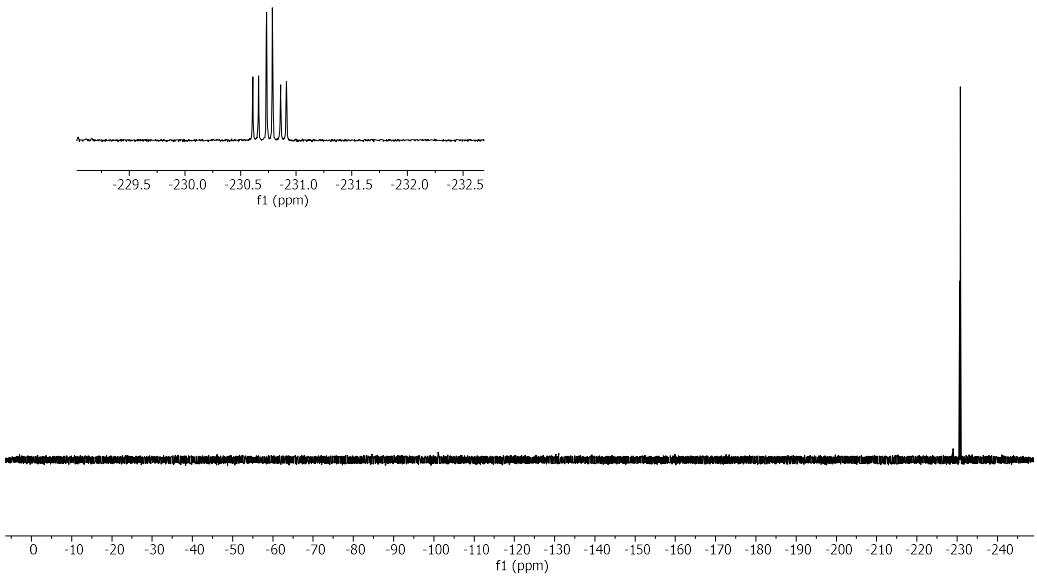
^{1}H NMR of BB2.3 (400 MHz, $CDCl_3$)



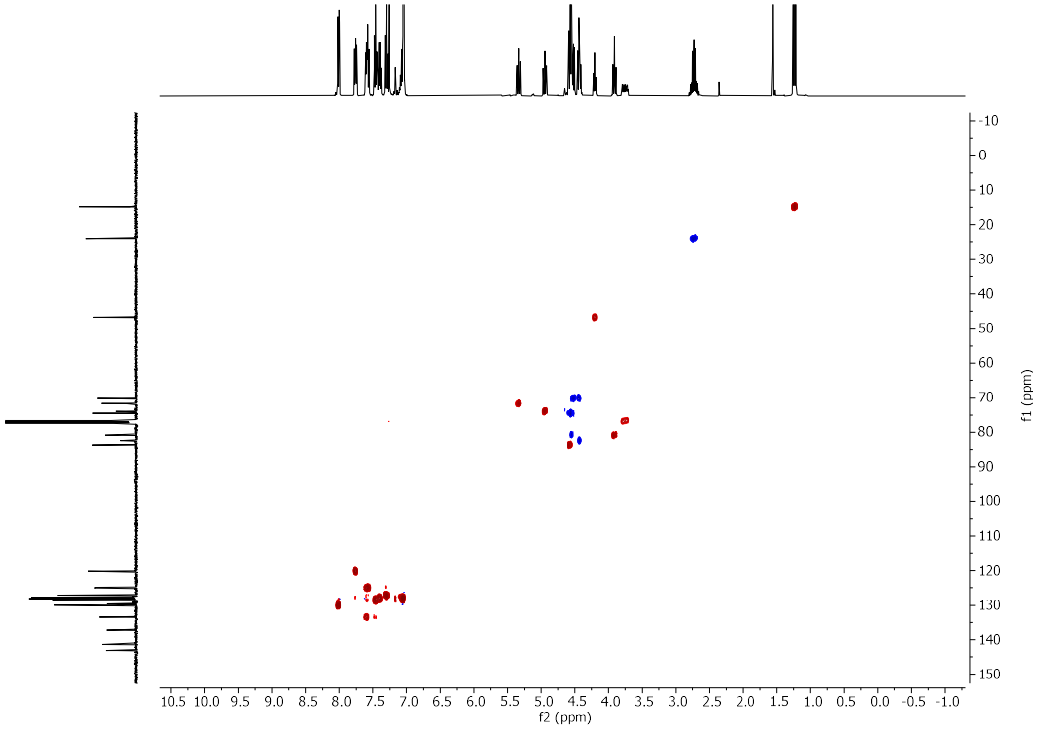
^{13}C NMR of BB2.3 (101 MHz, $CDCl_3$)



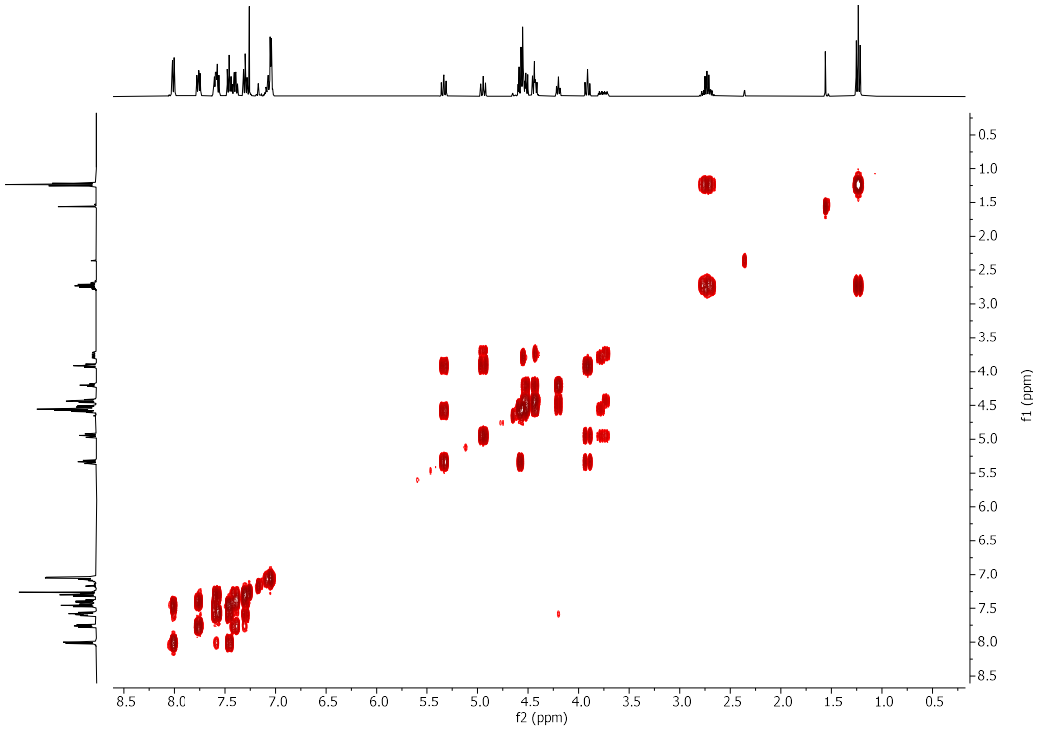
^{19}F NMR of BB2.3 (376 MHz, $CDCl_3$)



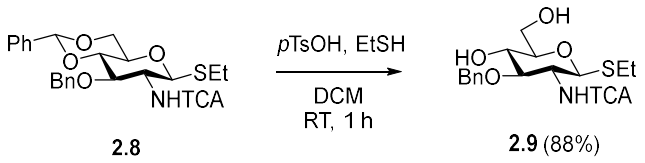
HSQC NMR of BB2.3 ($CDCl_3$)



COSY NMR of BB2.3 ($CDCl_3$)



Synthesis of **2.9**

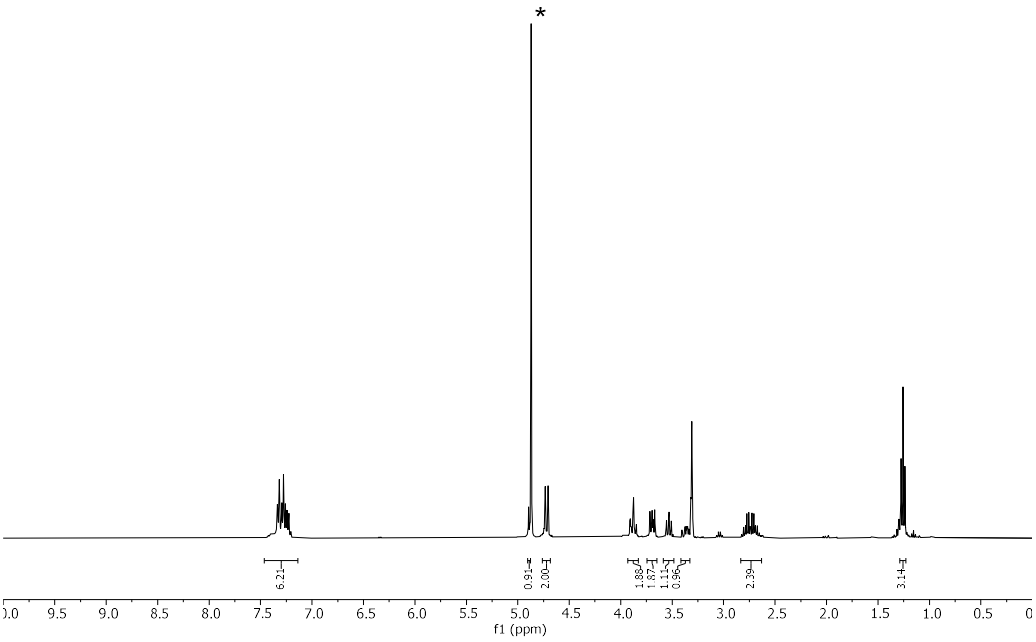


Compound **2.8** was synthesized according to previous reported procedure.[354]

2.8 (8.30 g, 15.2 mmol) was dissolved in anhydrous DCM (200 mL). EtSH (2.20 mL, 29.7 mmol) and p'TsOH (577 mg, 3 mmol) were added to the reaction mixture under Ar atmosphere and stirred at RT. After 10 min, a white precipitate formed. The solution was stirred for 1 h and quenched with NEt_3 (1 mL). The crude reaction mixture was dried under vacuum and purified by flash column chromatography (Hexane : Acetone = 2:1→1:1→1:2) to give **2.9** (6.17 g, 88%).

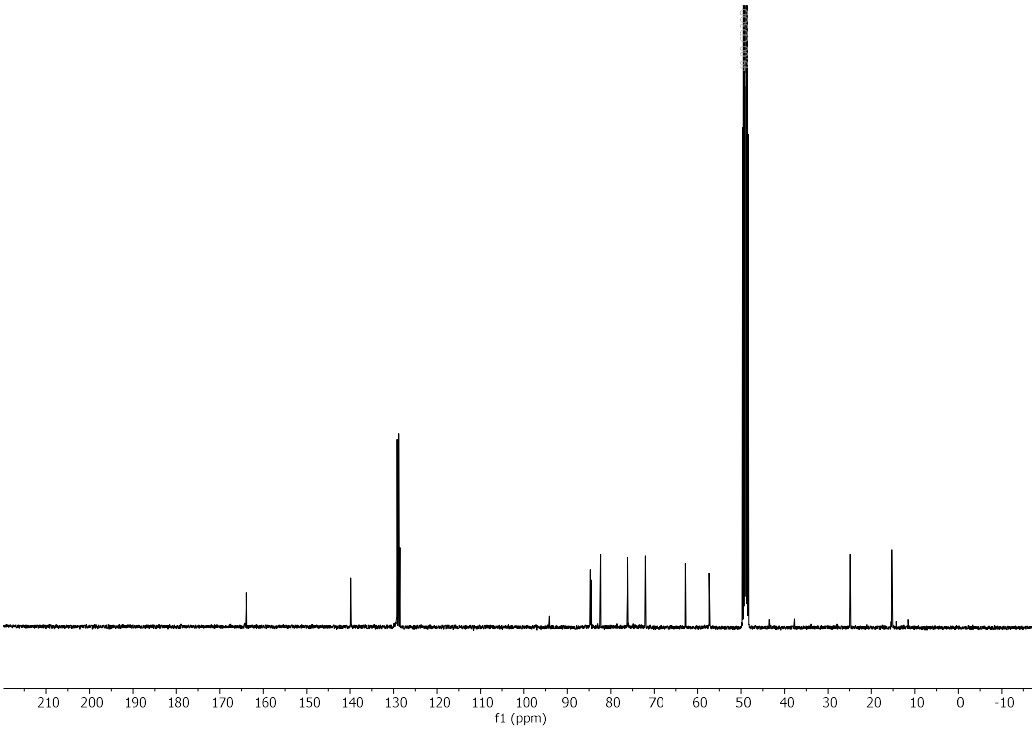
^{1}H NMR (400 MHz, MeOD) δ 7.44 – 7.16 (m, 6H), 4.89 (d, *J* = 8.5 Hz, 1H), 4.77 – 4.66 (m, 2H), 3.95 – 3.81 (m, 2H), 3.75 – 3.64 (m, 2H), 3.57 – 3.49 (m, 1H), 3.42 – 3.33 (m, 1H), 2.84 – 2.65 (m, 2H), 1.26 (t, *J* = 7.5 Hz, 3H). ^{13}C NMR (101 MHz, MeOD) δ 163.88, 139.86, 129.17, 128.82, 128.51, 84.72, 84.59, 82.38, 76.13, 72.07, 62.79, 57.35, 24.89, 15.28. $[\alpha]_D^{20}$ -0.87 (c 1.05 g/100 mL, MeOH). IR ν = 3322, 2931, 1691, 1527, 1071, 1043, 822 cm^{-1}. (ESI-HRMS) m/z 480.0172 $[M+Na]^+$ ($C_{17}H_{22}Cl_3NO_5SNa$ requires 480.0176).

^{1}H NMR of 2.9 (400 MHz, MeOD)

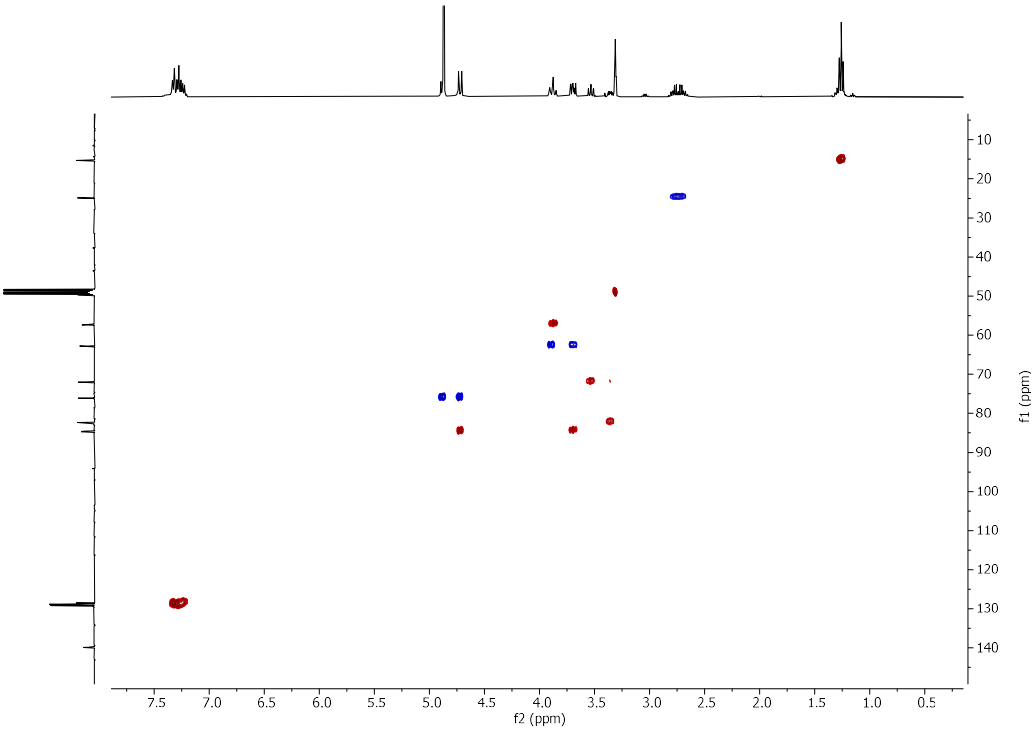


*The peak at 4.87 ppm belongs to residual water.

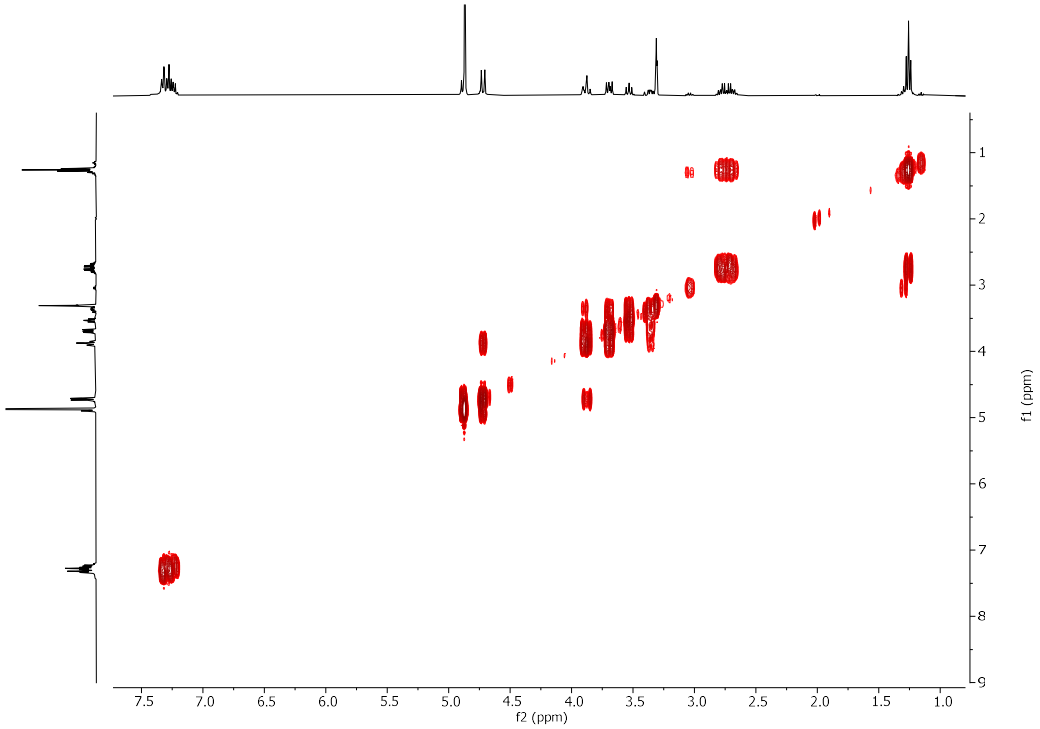
^{13}C NMR of 2.9 (101 MHz, MeOD)



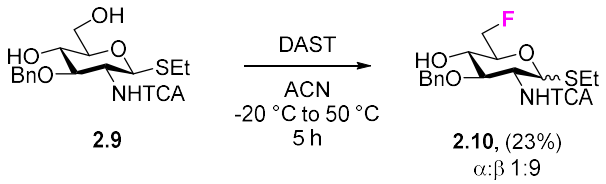
HSQC NMR of 2.9 (MeOD)



COSY NMR of 2.9 (MeOD)



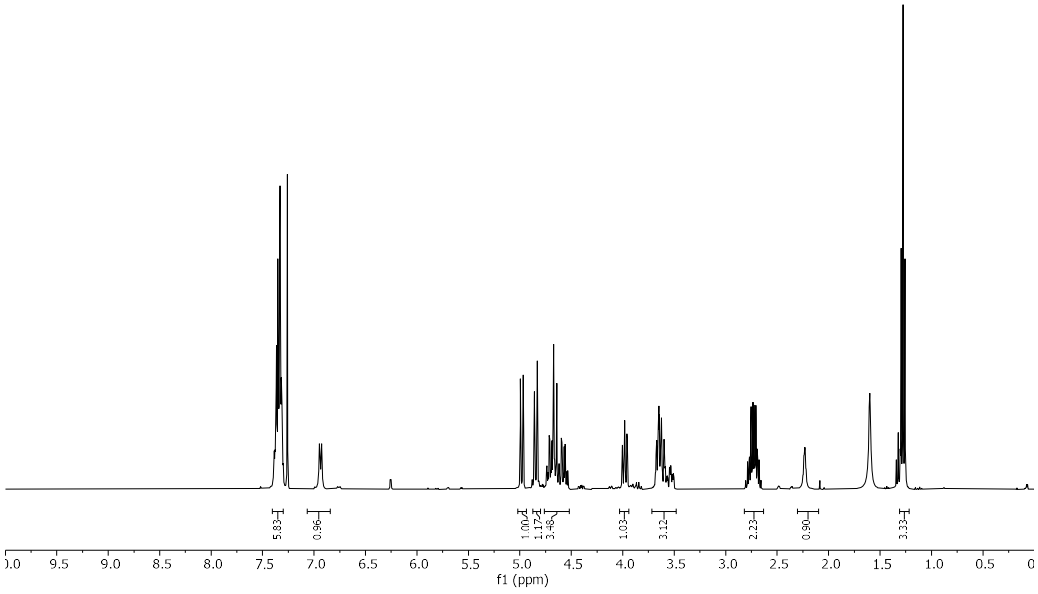
Synthesis of **2.10**



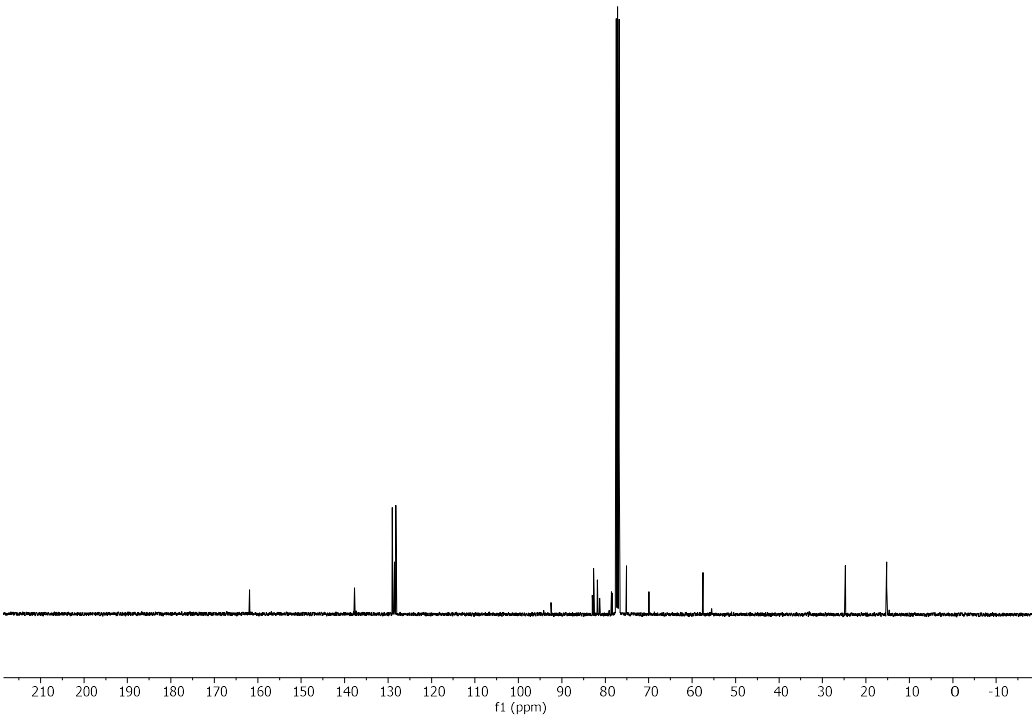
2.9 (6.17 g, 13.4 mmol) was dissolved in anhydrous ACN (130 mL) and cooled down to -20 °C. DAST (1.96 mL, 14.8 mmol) was added dropwise to the reaction mixture under Ar atmosphere. After 20 min the cooling bath was removed and the reaction heated to 50 °C. The solution was stirred for additional 3 h and quenched with MeOH at 0 °C. The crude reaction mixture was dried under vacuum and purified by flash column chromatography (Hexane : EtOAc = 3:1→2:1→1:1) to give **2.10** as an inseparable mixture of anomers (1.42 mg, 23%, α:β ratio 7:93). During the reaction, anomerization was observed.

Analytical data for the β anomer: ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.45 – 7.27 (m, 5H), 6.94 (d, *J* = 8.1 Hz, 1H), 4.98 (d, *J* = 10.3 Hz, 1H), 4.84 (d, *J* = 11.3 Hz, 1H), 4.76 – 4.51 (m, 3H), 3.98 (dd, *J* = 10.0, 8.4 Hz, 1H), 3.71 – 3.48 (m, 3H), 2.82 – 2.63 (m, 2H), 2.23 (s, 1H), 1.28 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 161.87, 137.73, 129.00, 128.53, 128.20, 82.69, 82.13 (d, *J* = 173.4 Hz), 81.78, 78.49 (d, *J* = 18.4 Hz), 75.12, 69.91 (d, *J* = 7.2 Hz), 57.52, 24.74, 15.21. ^{19}F NMR (376 MHz, Chloroform-*d*) δ -234.01 (td, *J* = 47.1, 22.9 Hz, F-6 β), -236.12 (td, *J* = 47.4, 26.7 Hz, F-6 α). $[\alpha]_D^{20}$ -23.91 (c 0.7 g/100 mL, $CHCl_3$). IR ν = 3302, 2928, 1688, 1539, 1084. 834, 825 cm^{-1}. (ESI-HRMS) m/z 482.0163 $[M+Na]^+$ ($C_{17}H_{21}Cl_3FNO_4SNa$ requires 482.0133).

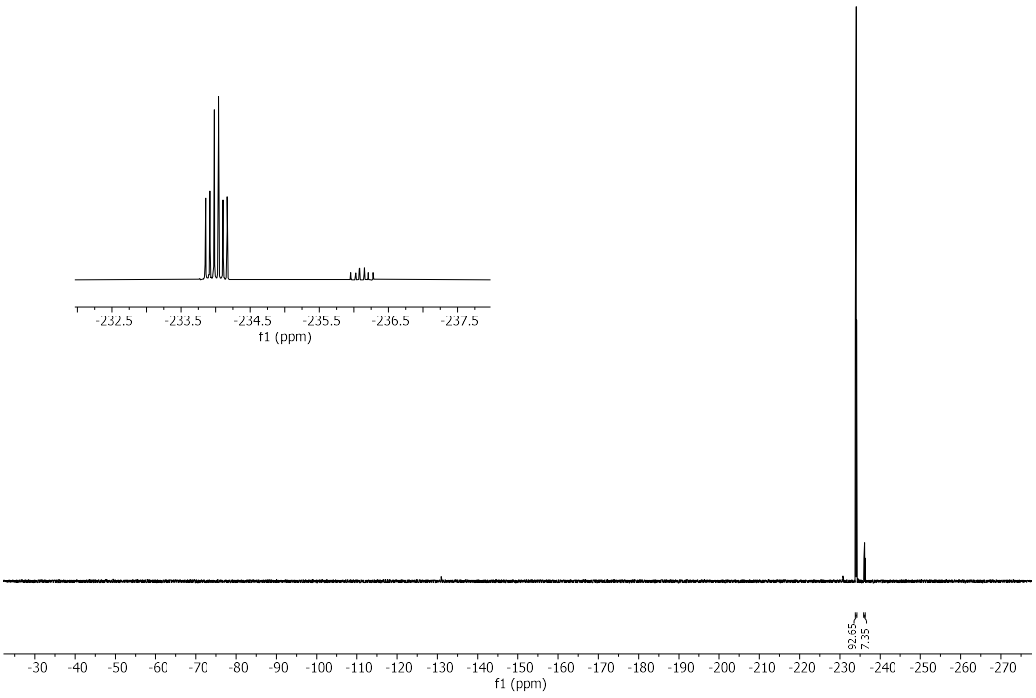
^{1}H NMR of 2.10 (400 MHz, $CDCl_3$)



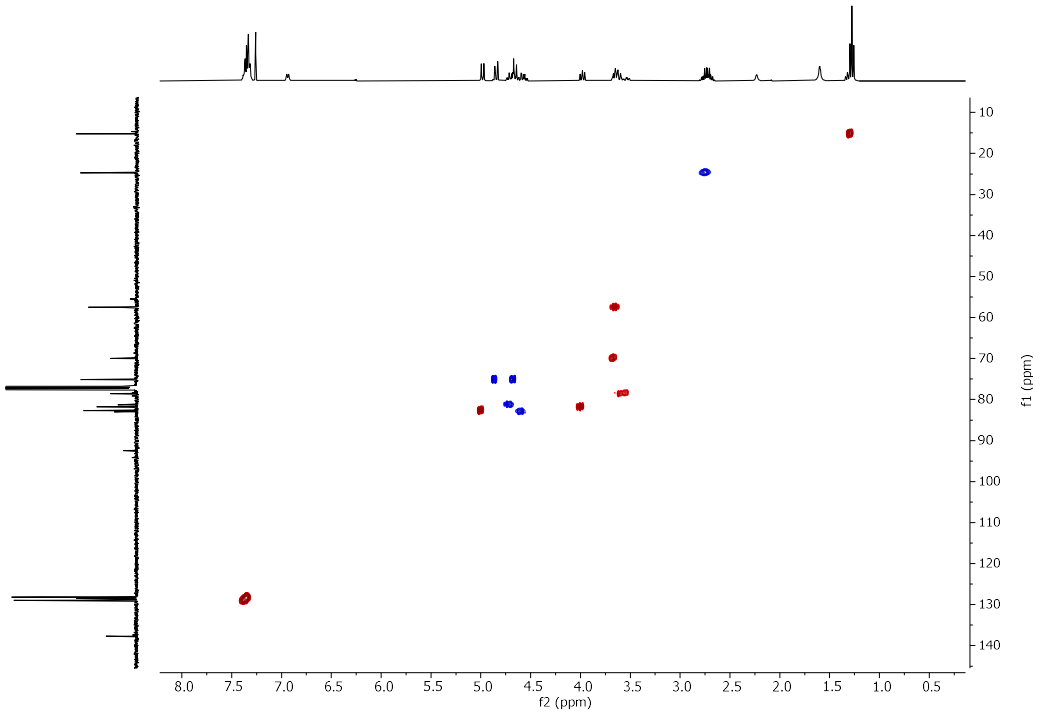
^{13}C NMR of 2.10 (101 MHz, $CDCl_3$)



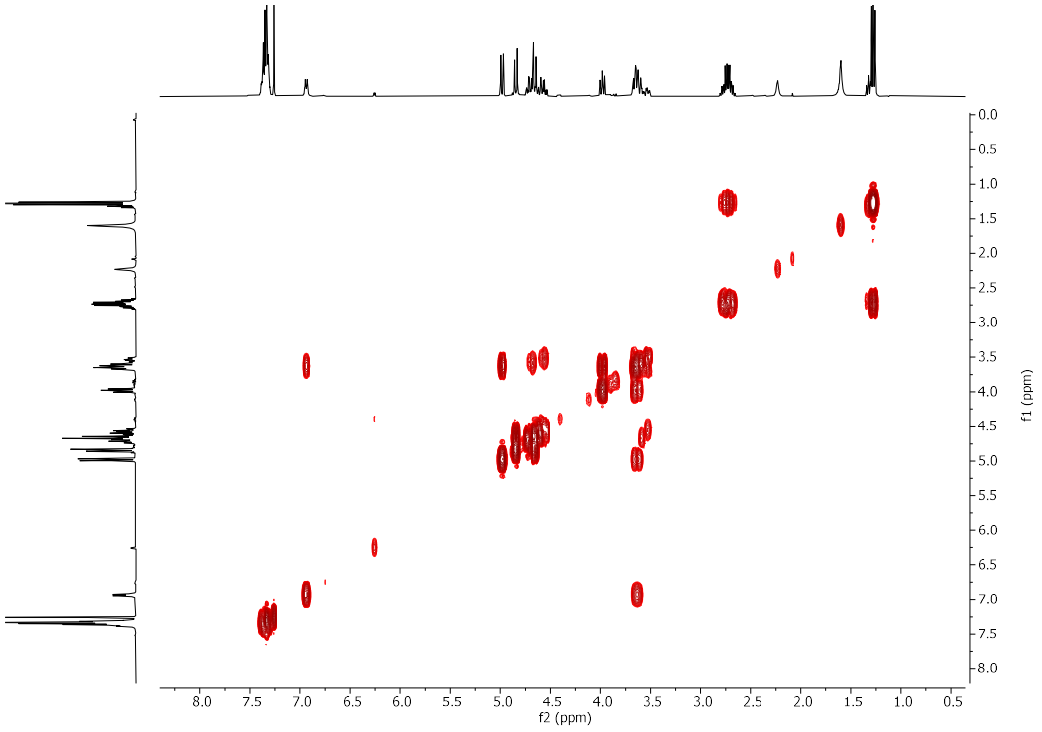
^{19}F NMR of 2.10 (101 MHz, $CDCl_3$)



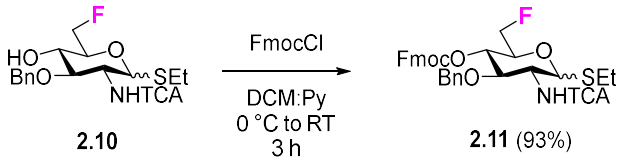
HSQC NMR of 2.10 ($CDCl_3$)



COSY NMR of 2.10 ($CDCl_3$)



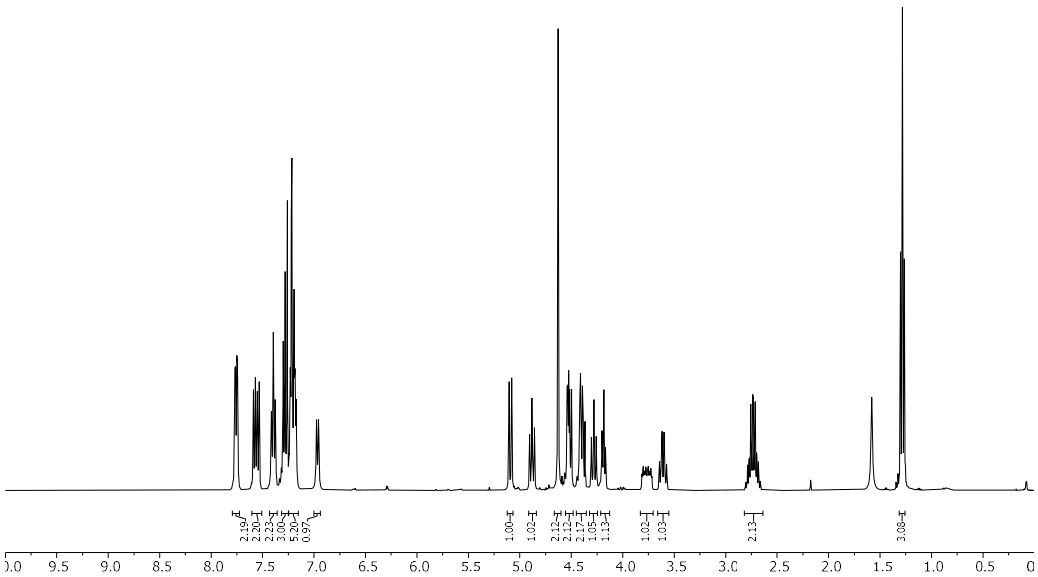
Synthesis of **2.11**



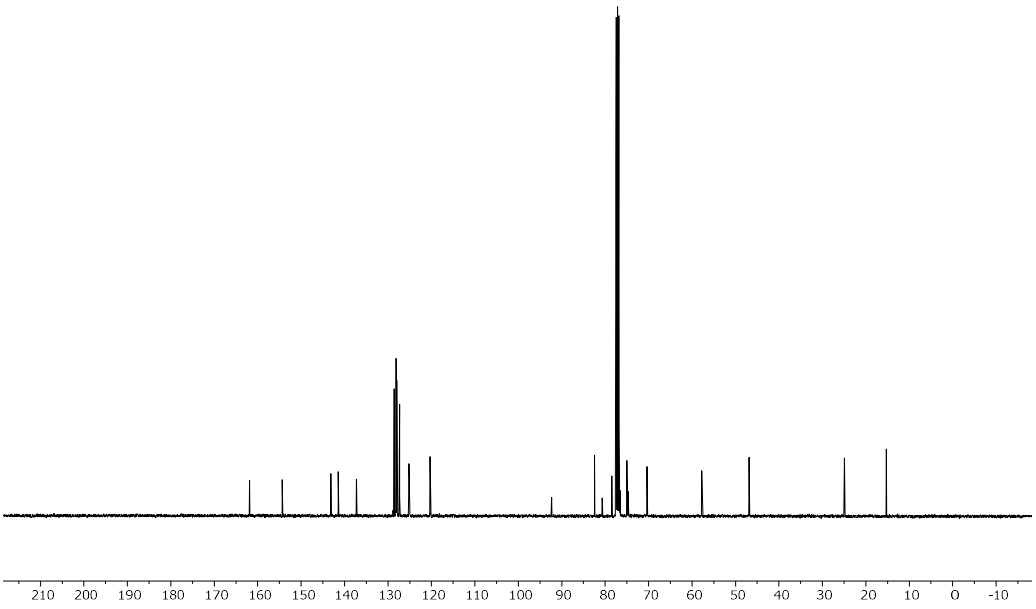
2.10 (1.42 g, 3.1 mmol) was dissolved in DCM:Py (3:1, 15 mL) and cooled to 0 °C. FmocCl (1.60 g, 6.2 mmol) was dissolved in DCM (0.5 mL) and added to the reaction mixture under Ar atmosphere. The solution was stirred for 20 min at 0 °C then warmed up to RT and stirred for additional 3 h, after which time it was quenched with a 0.5 M solution of citric acid. The crude reaction mixture was diluted with DCM, washed once with 0.5 M citric acid and once with brine. The crude compound was purified by silica gel flash column chromatography (Hexane : Acetone = 3:1→2:1→1:1) and recrystallized from DCM : Hexane to give **2.11** as an inseparable mixture of anomers (1.84 g, 87%, α:β ratio 3:97).

Analytical data for the β anomer: ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.76 (dd, J = 7.6, 3.1 Hz, 2H), 7.61 – 7.51 (m, 2H), 7.40 (tt, J = 7.6, 1.7 Hz, 2H), 7.31 – 7.16 (m, 7H), 6.98 (d, J = 8.0 Hz, 1H), 5.09 (d, J = 10.2 Hz, 1H), 4.88 (t, J = 9.5 Hz, 1H), 4.63 (s, 2H), 4.56 – 4.48 (m, 2H), 4.45 – 4.36 (m, 2H), 4.28 (t, J = 9.4 Hz, 1H), 4.19 (t, J = 6.9 Hz, 1H), 3.76 (dddd, J = 20.5, 10.0, 4.6, 2.9 Hz, 1H), 3.68 – 3.53 (m, 1H), 2.83 – 2.65 (m, 2H), 1.28 (t, J = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 161.88, 154.31, 143.13 (d, J = 1.5 Hz), 141.45, 137.25, 128.58, 128.12, 127.96, 127.34, 125.15, 124.99, 120.28, 120.25, 8.44, f81.57 (d, J = 175.4 Hz), 78.50, 76.66 (d, J z= 19.6 Hz), 74.99, 74.70 (d, J = 6.1 Hz), 70.35, 57.73, 46.84, 24.88, 15.24. ^{19}F NMR (376 MHz, Chloroform-*d*) δ -231.05 (td, J = 47.0, 20.5 Hz, F-6 β), -232.92 (td, J = 47.3, 23.9 Hz, F-6 α). $[\alpha]_D^{20}$ +14.06 (c 1.0 g/100 mL, $CHCl_3$). IR ν = 3321, 1754, 1692, 1530, 1257, 758, 742 cm^{-1}. (ESI-HRMS) m/z 704.0836 $[M+Na]^+$ ($C_{32}H_{31}Cl_3FNO_6SNa$ requires 704.0814).

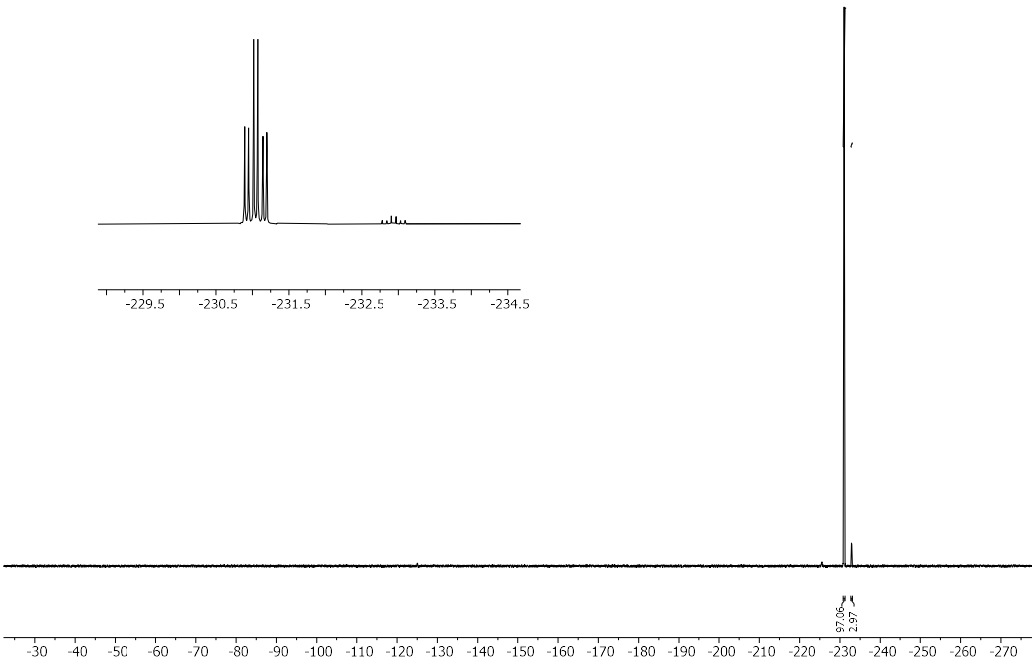
^{1}H NMR of 2.11 (400 MHz, $CDCl_3$)



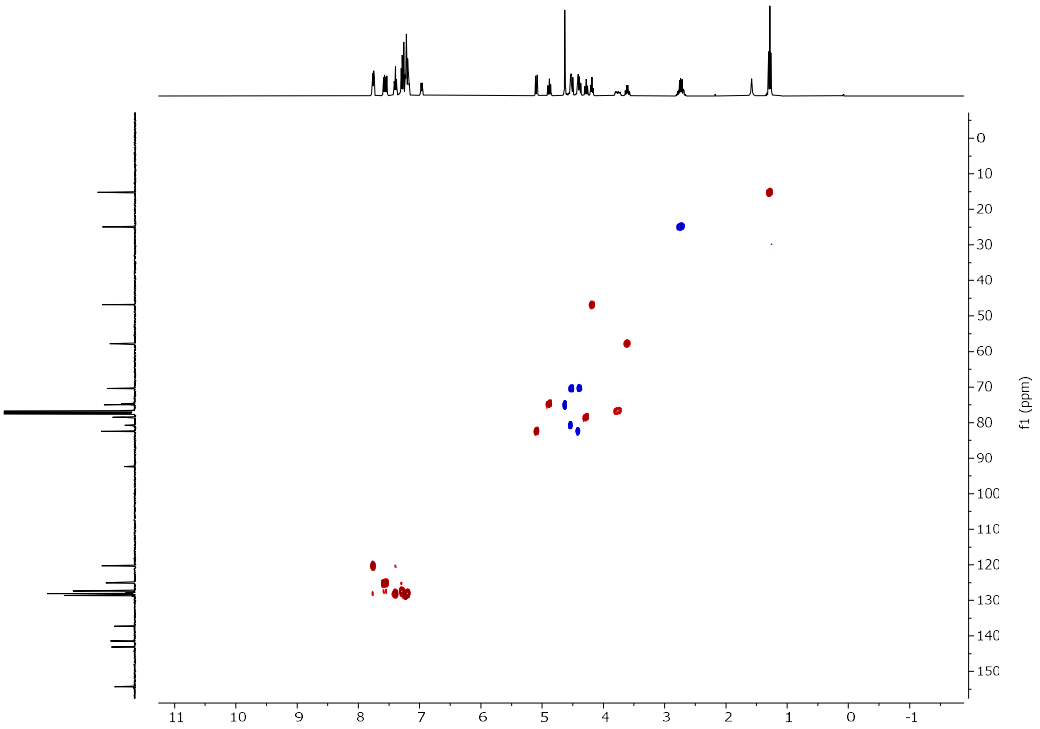
^{13}C NMR of 2.11 (101 MHz, $CDCl_3$)



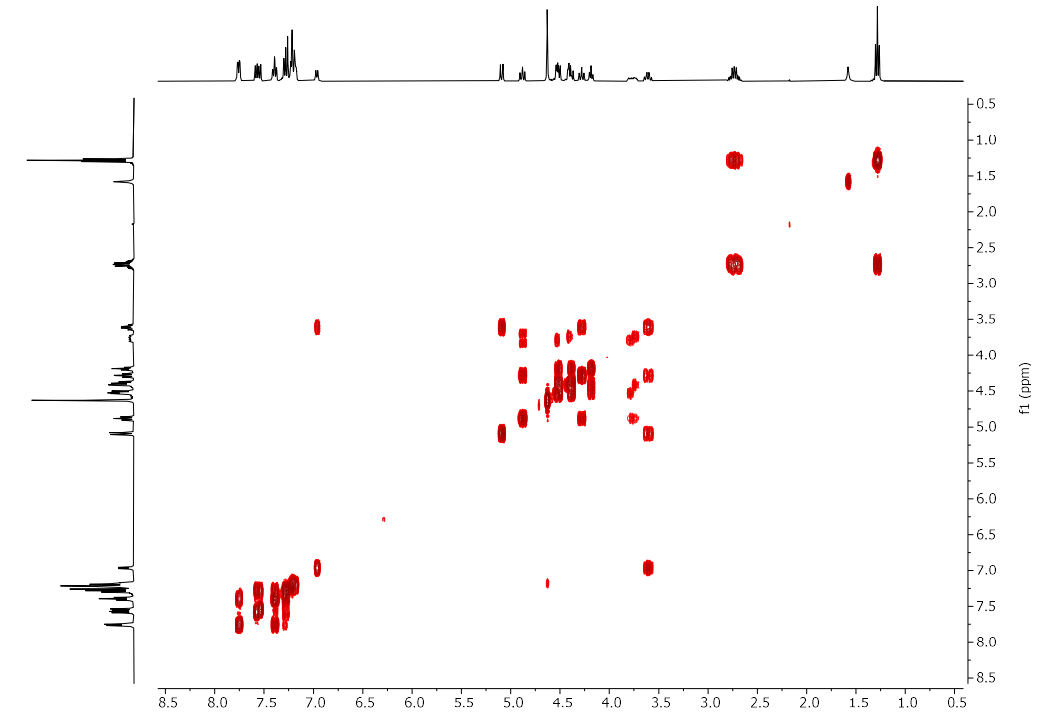
^{19}F NMR of 2.11 (101 MHz, $CDCl_3$)



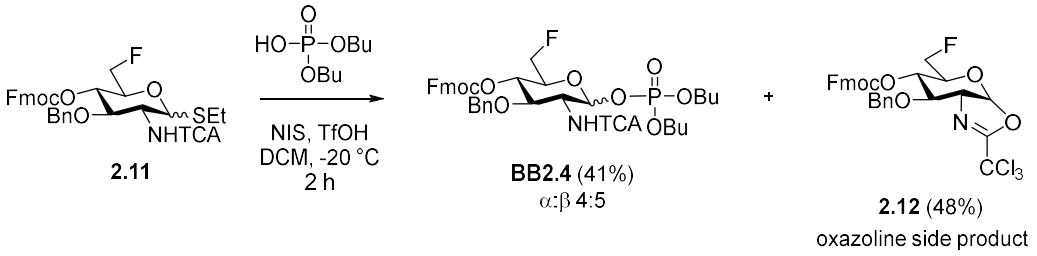
HSQC NMR of 2.11 ($CDCl_3$)



COSY NMR of 2.11 ($CDCl_3$)



Synthesis of **BB2.4**



2.11 (1.90 g, 2.78 mmol) was co-evaporated with toluene three times and dissolved in anhydrous DCM (35 mL). Activated molecular sieves (4 Å) were added to the mixture and the suspension was stirred at RT for 1 h. A solution of dibutyl phosphate (15 mL of a 7%$_v$ solution in DCM stirred in presence of activated 4 Å molecular sieves for 1 h, 5.30 mmol) was added to the solution of **2.11** and cooled down to -20 °C. NIS (751 mg, 3.34 mmol) and TfOH (17.5 μL, 0.20 mmol) were added and the reaction was slowly allowed to RT. After 2 h, the reaction was quenched with $NaHCO_3$ saturated aqueous solution, filtered, diluted with DCM and washed once with 10% aqueous $Na_2S_2O_3$. The organic layer was dried Na_2SO_4, concentrated under vacuum and purified by silica gel column chromatography (Hexane : Acetone 5:1→4:1→3:1→2:1) to obtain **BB2.4** as a sticky colorless solid (952 mg, 41%, α:β ratio 4:5). The oxazoline side product **2.12** formed during the reaction and was isolated as a sticky colorless solid (827 mg, 48%).

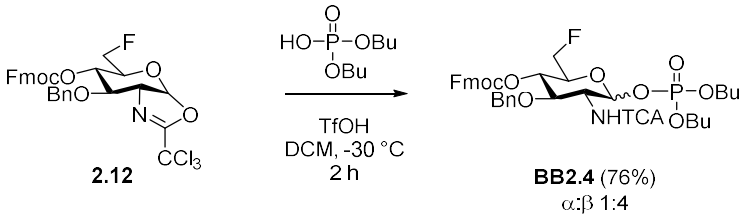
Analytical data for **BB2.4**:

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.00 (d, *J* = 6.3 Hz, 0.5H), 7.81 – 7.72 (m, 2H), 7.62 – 7.49 (m, 2H), 7.40 (td, *J* = 7.4, 4.2 Hz, 2H), 7.32 – 7.27 (m, 2H), 7.25 – 7.17 (m, 5H), 6.88 (d, *J* = 8.6 Hz, 0.5H), 5.74 (dd, *J* = 6.0, 3.3 Hz, 0.5H), 5.60 – 5.54 (m, 0.5H), 4.99 (dd, *J* = 10.4, 9.1 Hz, 0.5H), 4.93 – 4.82 (m, 0.5H), 4.67 (d, *J* = 11.1 Hz, 0.5H), 4.64 – 4.56 (m, 1.5H), 4.56 – 4.44 (m, 2.5H), 4.44 – 4.35 (m, 1.5H), 4.31 (ddt, *J* = 11.7, 8.8, 3.1 Hz, 0.5H), 4.27 – 3.94 (m, 7H), 3.80 (ddt, *J* = 18.6, 8.8, 3.8 Hz, 0.5H), 1.76 – 1.55 (m, 4H), 1.40 (dq, *J* = 14.9, 7.4 Hz, 4H), 1.00 – 0.87 (m, 6H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 162.37, 162.07, 154.33, 154.15, 143.20, 143.12, 143.06, 143.03, 141.49, 141.46, 137.33, 137.05, 128.69, 128.50, 128.18, 128.15, 128.03, 128.01, 127.36, 127.34, 125.11, 125.09, 124.97, 96.32, 95.35 (d, *J* = 6.5 Hz), 82.07 (d, *J* = 38.8 Hz), 80.32 (d, *J* = 39.6 Hz), 78.10, 75.95, 74.41 (d, *J* = 13.7 Hz), 73.57, 73.37, 70.83, 70.63, 70.38, 70.36, 68.88, 68.81, 68.65, 68.58, 68.55, 68.52, 68.50, 56.90 (d, *J* = 8.9 Hz), 54.33 (d, *J* = 7.8 Hz), 46.90, 46.87, 32.37, 32.35, 32.31, 32.29, 32.21, 32.12, 32.04, 18.73, 18.70, 13.73, 13.70, 13.68. ^{31}P NMR (162 MHz, $CDCl_3$) δ -2.75, -3.58. ^{19}F NMR (376 MHz, $CDCl_3$) δ -230.70 (td, *J* = 46.9, 18.9 Hz), -231.98 (td, *J* = 47.0, 22.4 Hz). IR ν = 3252, 2963, 1756, 1719, 1259, 1029, 968 cm^{-1}. (ESI-HRMS) m/z 852.1664 $[M+Na]^+$ ($C_{38}H_{44}Cl_3FNO_{10}PNa$ requires 852.1644).

Analytical data for **2.12**:

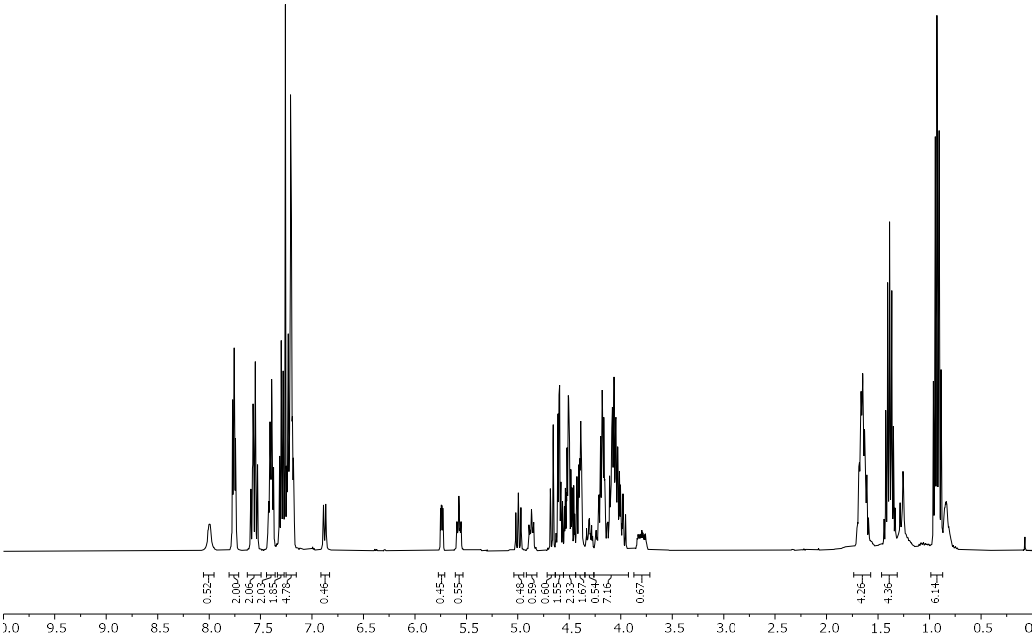
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.82 – 7.76 (m, 2H), 7.65 – 7.59 (m, 2H), 7.44 (t, *J* = 7.5 Hz, 2H), 7.40 – 7.27 (m, 7H), 6.39 (d, *J* = 7.4 Hz, 1H), 4.98 (dt, *J* = 8.0, 1.6 Hz, 1H), 4.86 (d, *J* = 12.2 Hz, 1H), 4.71 (d, *J* = 12.2 Hz, 1H), 4.65 (dd, *J* = 7.7, 3.6 Hz, 1H), 4.53 (dd, *J* = 7.4, 3.6 Hz, 1H), 4.50 – 4.40 (m, 3H), 4.27 (t, *J* = 7.5 Hz, 1H), 4.19 (s, 1H), 3.88 (dddd, *J* = 22.5, 7.4, 4.4, 2.7 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 163.02, 154.16, 143.16, 143.10, 141.43, 141.41, 137.13, 128.69, 128.28, 128.17, 128.12, 127.37, 127.31, 125.34, 125.29, 103.73, 82.24 (d, *J* = 176.8 Hz), 73.48, 71.95, 70.47, 69.89 (d, *J* = 6.7 Hz), 69.40 (d, *J* = 19.1 Hz), 65.52, 46.72. ^{19}F NMR (376 MHz, $CDCl_3$) δ -230.89 (td, *J* = 47.1, 22.5 Hz). IR ν = 2956, 1751, 1259, 981, 742 cm^{-1}. (ESI-HRMS) m/z 620.1012 $[M+Na]^+$ ($C_{30}H_{26}Cl_3FNO_6$ requires 620.0804).

Recovery of the oxazoline side product **2.12**

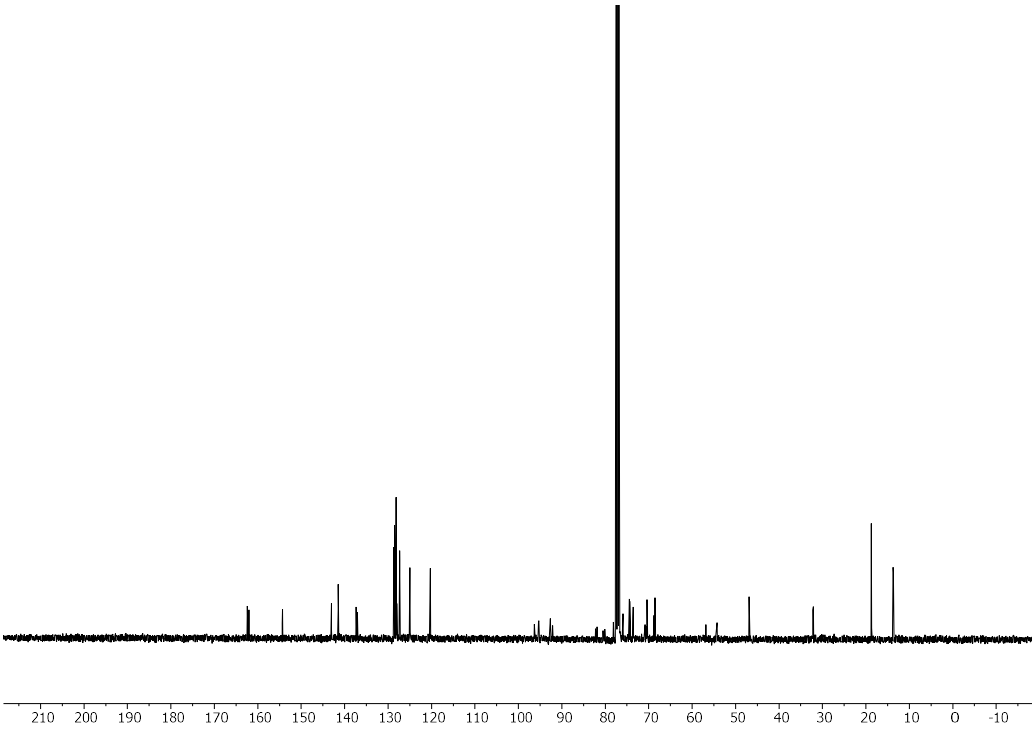


2.12 (781 mg, 1.26 mmol) was co-evaporated with toluene three times and dissolved in anhydrous DCM (15 mL). Activated molecular sieves (4 Å) were added to the mixture and the suspension was stirred at RT for 1h. A solution of dibutyl phosphate (10 mL of a 7.5%$_v$ solution in DCM stirred in presence of activated 4 Å molecular sieves for 1 h, 3.78 mmol) was added to the suspension of **2.12** and cooled down to -30 °C. TfOH (50 µL, 0.57 mmol) was added dropwise and the reaction stirred at -30 °C for 2 h, after which time it was quenched with NEt_3 (90 µL). The reaction was diluted with EtOAc, passed through a short plug of silica, and concentrated under vacuum. The crude was purified by silica gel flash column chromatography (Toluene 100%→10% EtOAc : Toluene) to yield **BB2.4** as a sticky colorless solid (843 mg, 76%, α:β ratio 1:4).

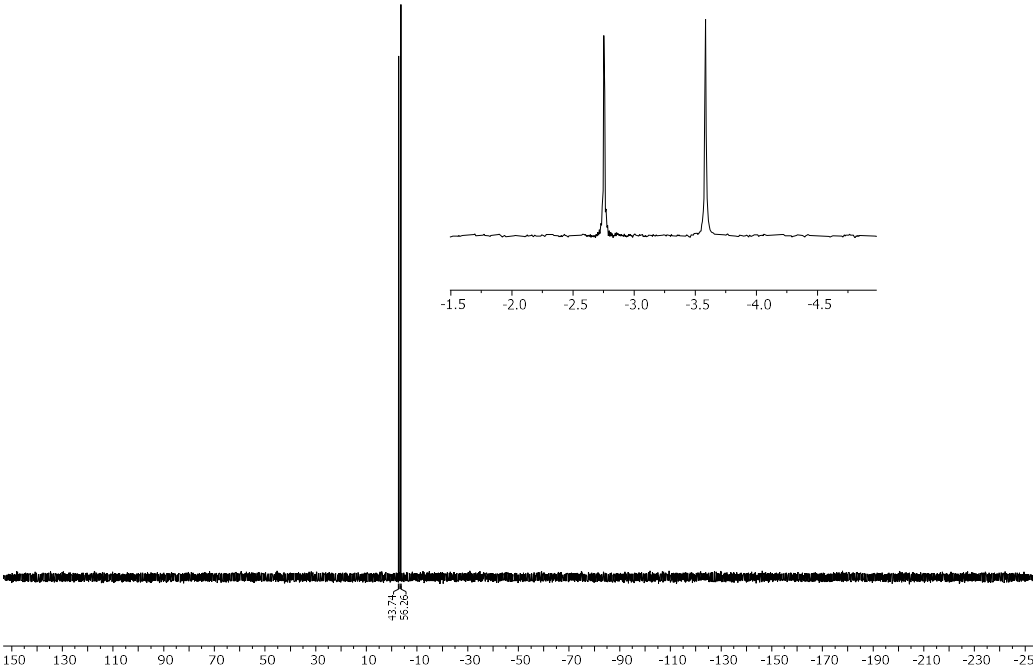
^{1}H NMR of BB2.4 (400 MHz, $CDCl_3$)



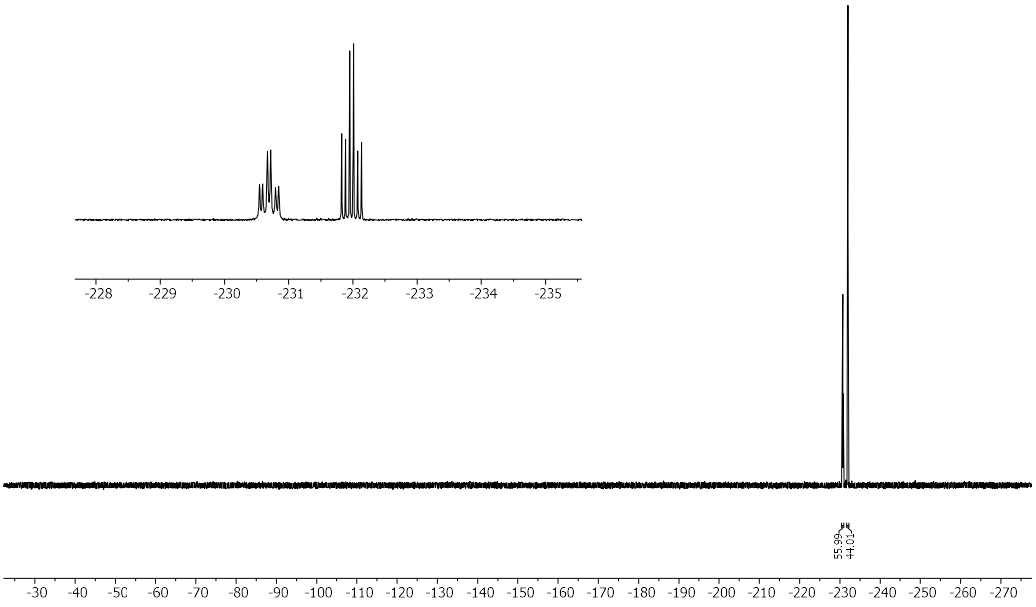
^{13}C NMR of BB2.4 (101 MHz, $CDCl_3$)



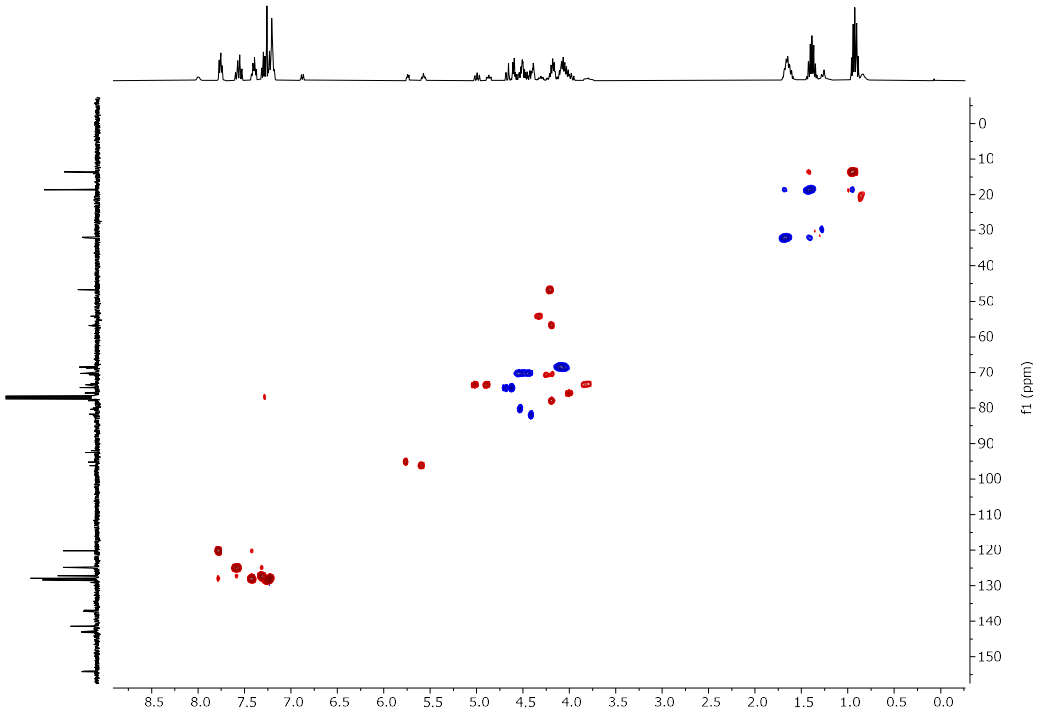
^{31}P NMR of BB2.4 (162 MHz, $CDCl_3$)



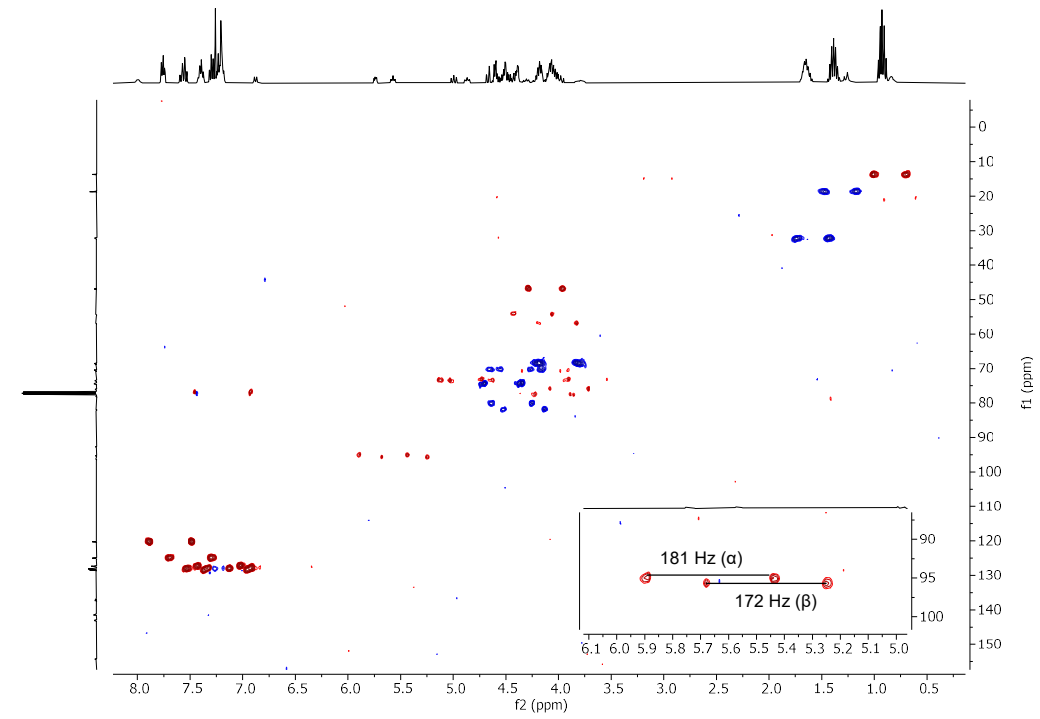
^{19}F NMR of BB2.4 (101 MHz, $CDCl_3$)



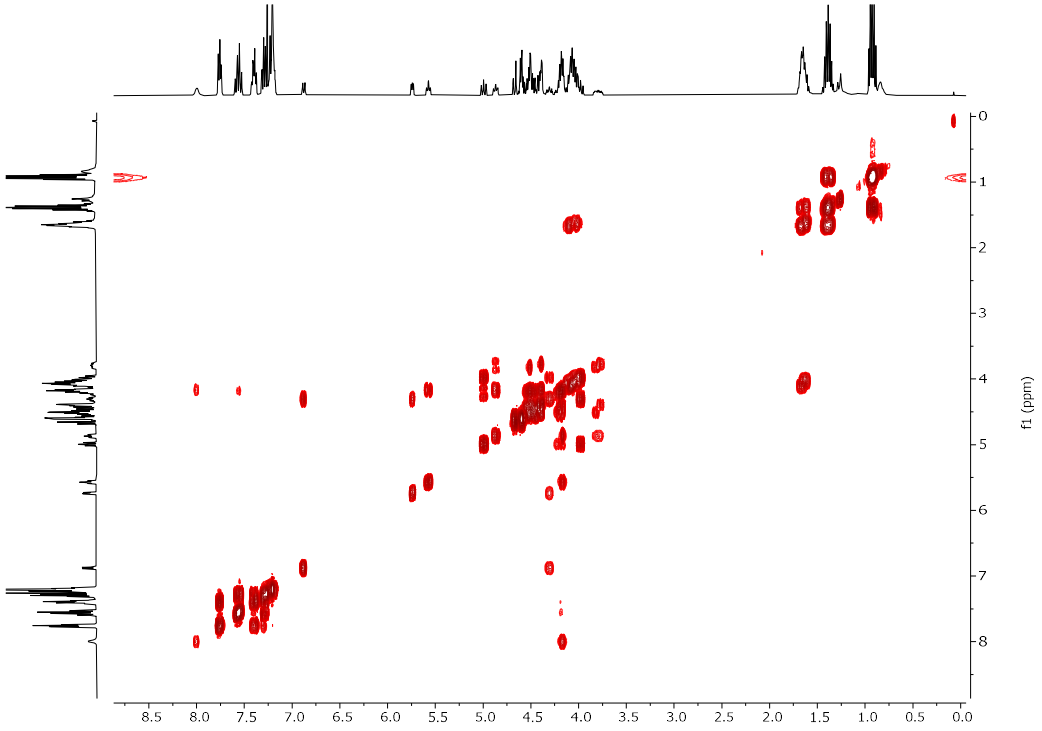
HSQC NMR of BB2.4 ($CDCl_3$)



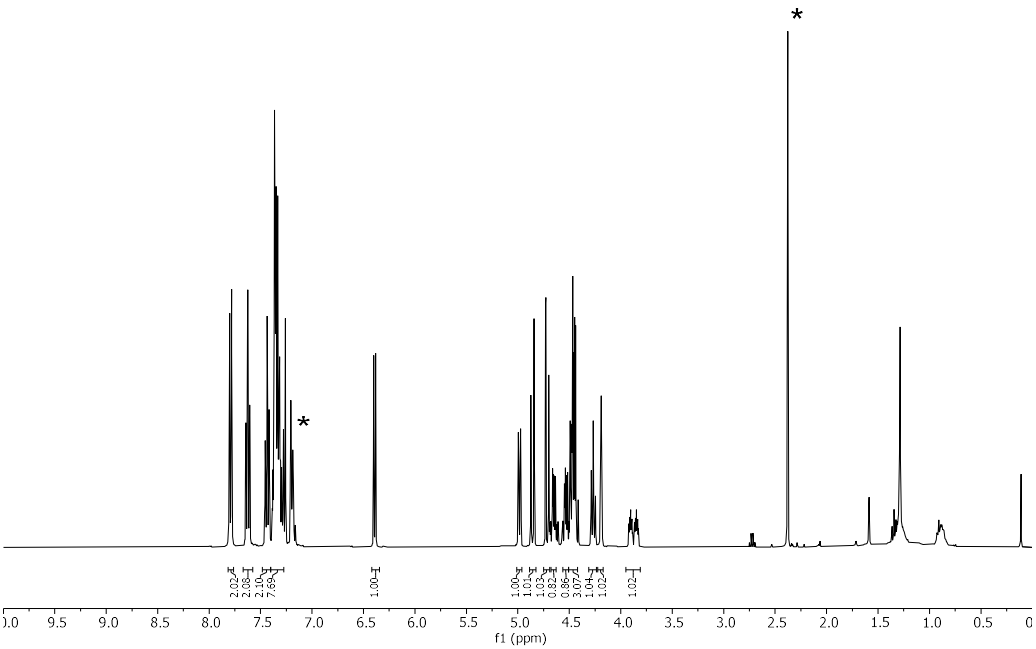
Coupled HSQC NMR of BB2.4 ($CDCl_3$)



COSY NMR of BB2.4 ($CDCl_3$)

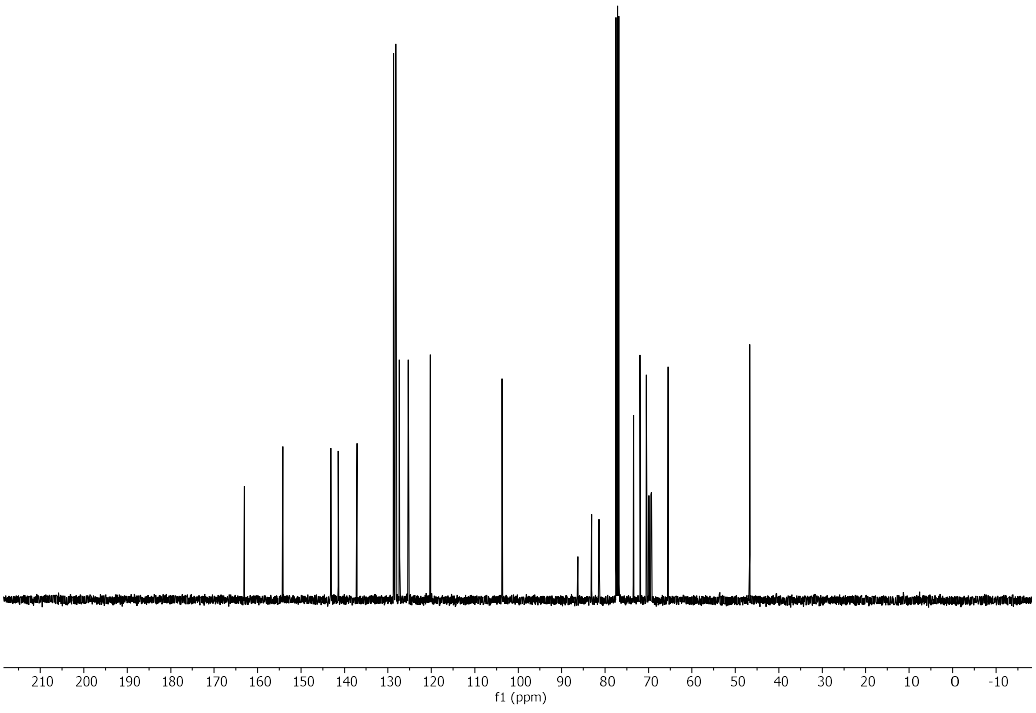


^{1}H NMR of 2.12 (400 MHz, $CDCl_3$)

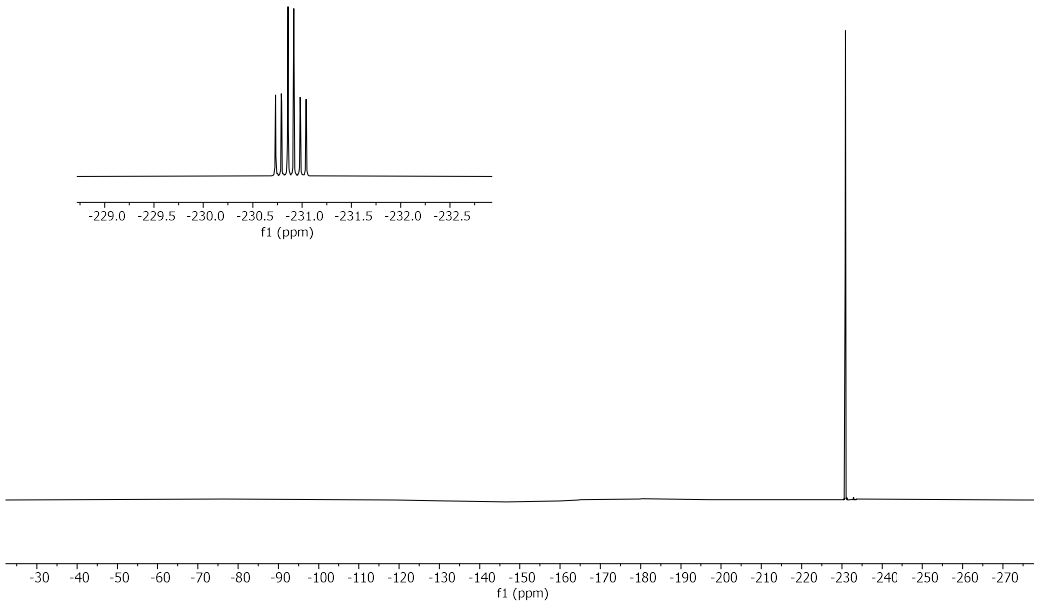


*Peaks belonging to residual Toluene.

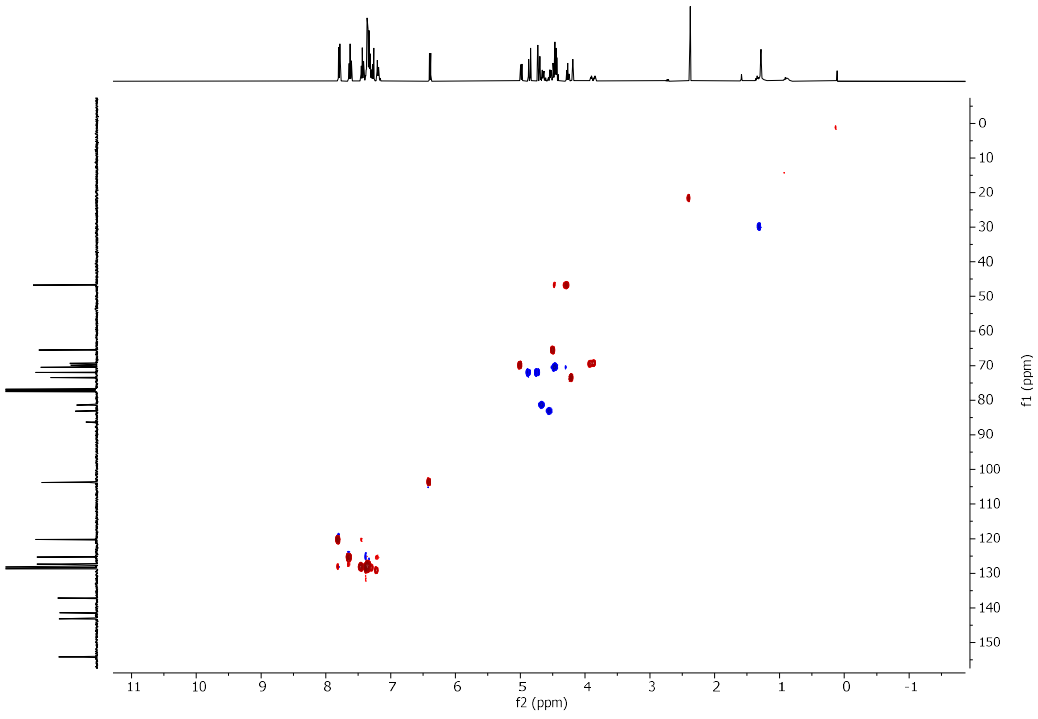
^{13}C NMR of 2.12 (101 MHz, $CDCl_3$)



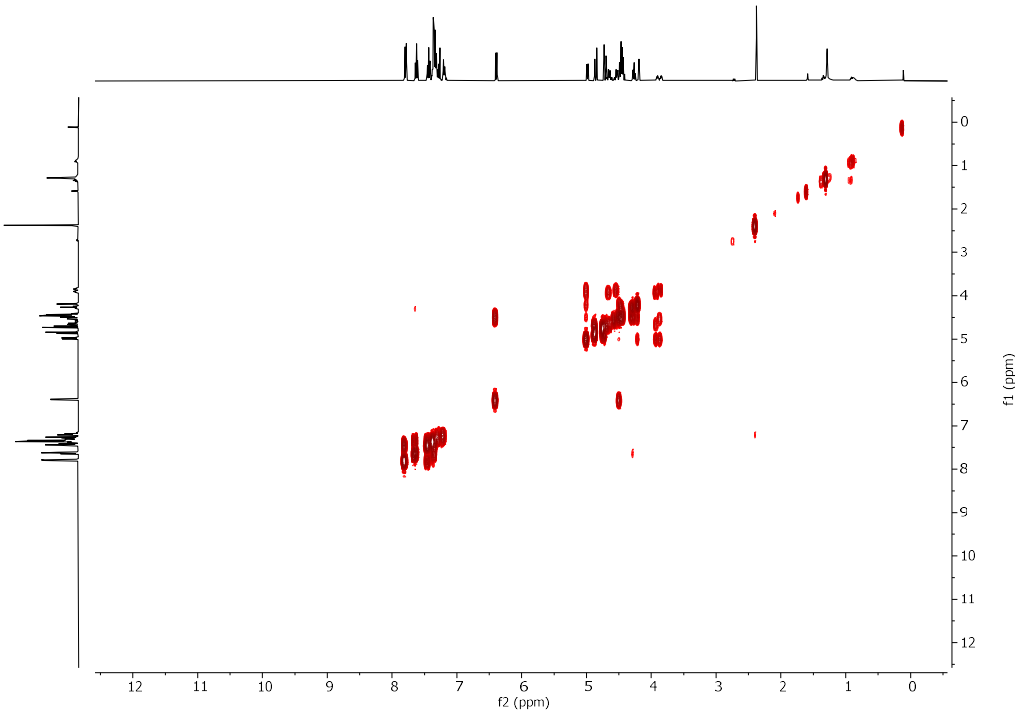
^{19}F NMR of 2.12 (101 MHz, $CDCl_3$)



HSQC NMR of 2.12 ($CDCl_3$)



COSY NMR of 2.12 ($CDCl_3$)



6.3.2 Oligosaccharide Synthesis

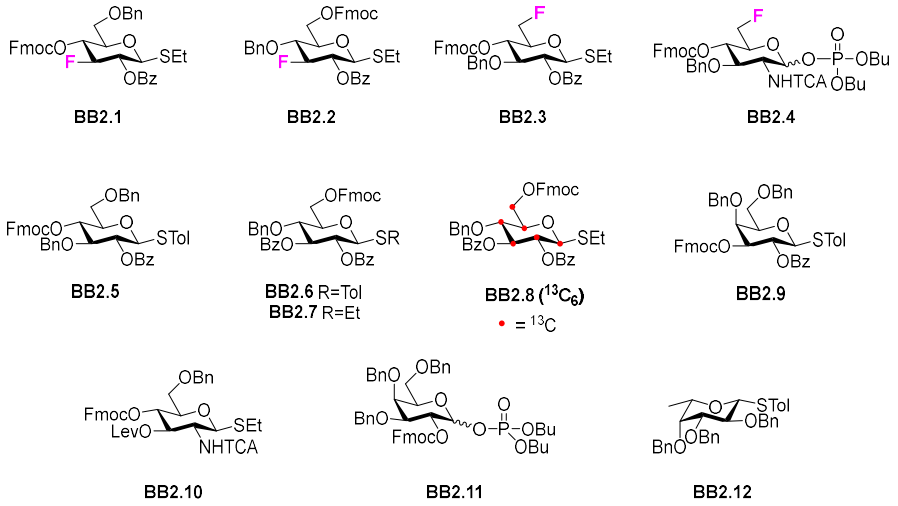
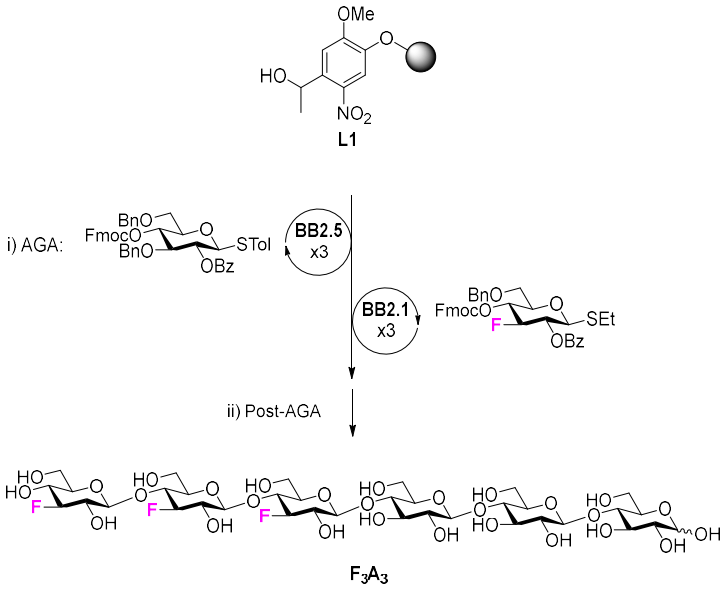


Figure S2.1 BBs used for the AGA of fluorinated oligosaccharides.

Synthesis of **F_3A_3**

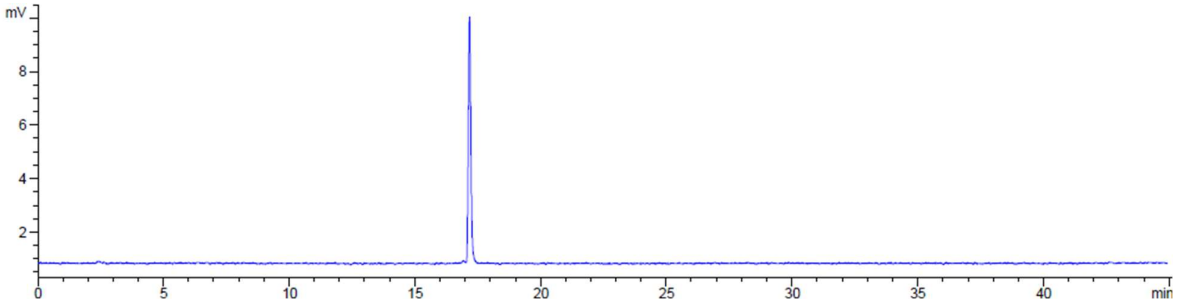


Step	BB	Modules	Notes
AGA	**-**	**A**	**L1** swelling
	(**BB2.5**)x3	(**B**, **C1**, **D**, **E1**)x3	**C1**: (**BB2.5**, -20°C for 5 min, 0°C for 20 min)
	(**BB2.1**)x3	(**B**, **C1**, **D**, **E1**)x3	**C1**: (**BB2.1**, -20°C for 5 min, 0°C for 20 min)
Post-AGA	**-**	**F1**, **G1**, **H2**, **I**	**F1**: (30 h) **H2**: (4 h) **I**: (Method D)

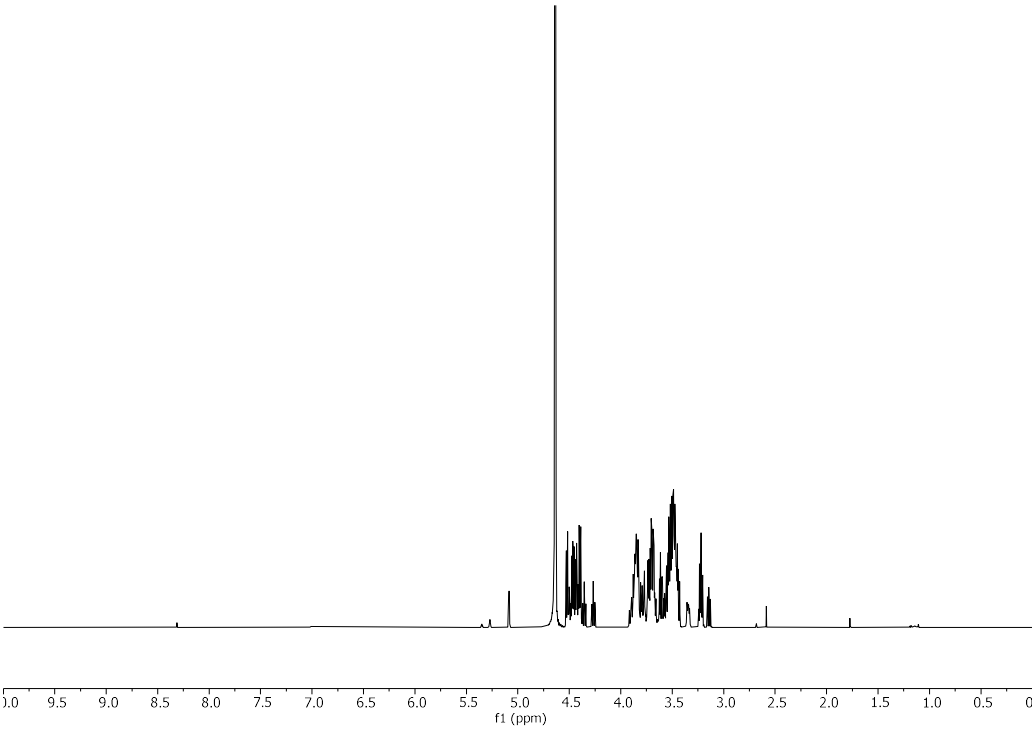
Automated synthesis, global deprotection, and purification afforded **F_3A_3** as a white solid (3.5 mg, 25% overall yield).

^{1}H NMR (600 MHz, D_2O) δ 5.09 (d, *J* = 3.8 Hz, 0.4H), 4.55 – 4.38 (m, 7.6H), 4.31 (dt, *J* = 52.8, 8.9 Hz, 1H), 3.94 – 3.75 (m, 8H), 3.70 (dddt, *J* = 20.1, 11.4, 7.8, 4.3 Hz, 6.4H), 3.64 – 3.41 (m, 15H), 3.34 (dd, *J* = 10.0, 5.3 Hz, 1H), 3.22 (td, *J* = 8.3, 5.1 Hz, 2H), 3.18 – 3.12 (m, 0.6H). ^{13}C NMR (151 MHz, D_2O) δ 102.29, 102.28, 101.54, 101.46, 101.38, 101.32, 96.77, 95.69, 95.57, 94.99, 94.89, 93.77, 93.67, 91.75, 78.56, 78.41, 78.18, 78.08, 75.59, 75.54, 75.48, 75.42, 74.75, 74.73, 74.72, 74.65, 74.59, 74.18, 73.94, 73.86, 73.82, 73.80, 73.75, 73.69, 73.64, 72.88, 71.90, 71.82, 71.81, 71.69, 71.23, 71.16, 70.05, 67.84, 67.72, 60.14, 60.13, 59.94, 59.78, 59.71, 59.50. ^{19}F NMR (564 MHz, D_2O) δ -192.33 (dt, *J* = 52.2, 14.0 Hz), -192.53 (dt, *J* = 52.2, 13.9 Hz), -195.00 (dt, *J* = 52.8, 13.9 Hz). (ESI-HRMS) m/z 1019.3104 $[M+Na]^+$ ($C_{36}H_{59}F_3O_{28}Na$ requires 1019.3037).

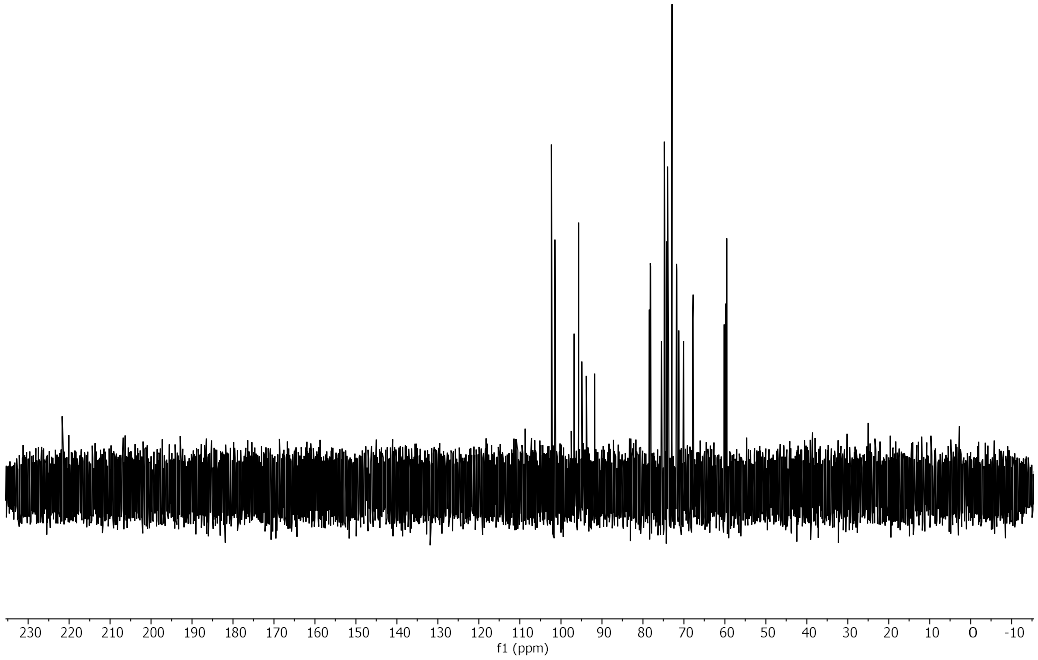
RP HPLC of F_3A_3 (ELSD trace, Method C1, t_R = 17.2 min)



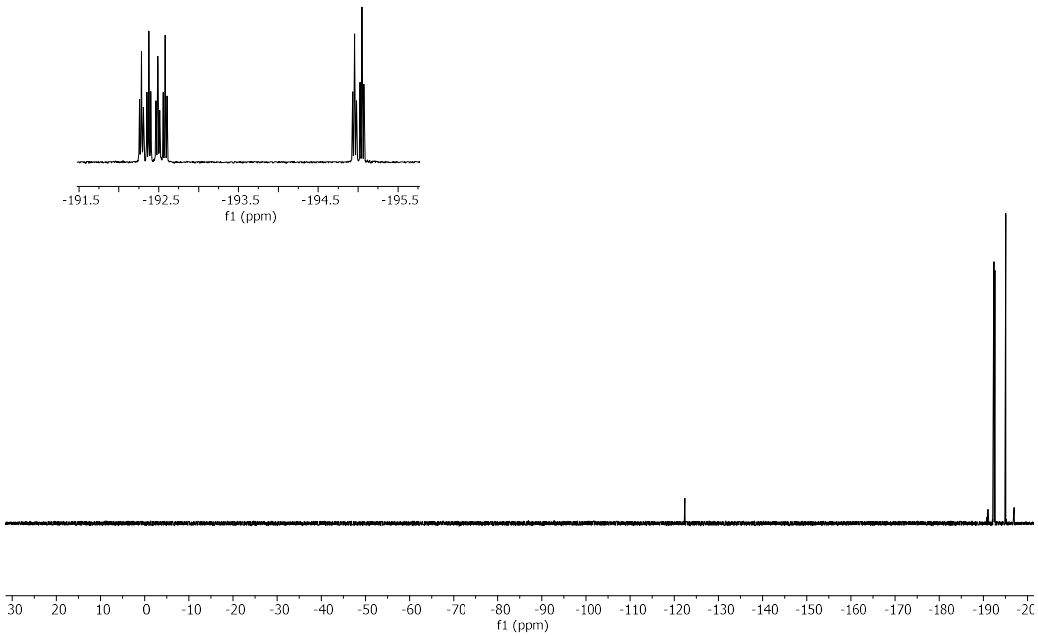
^{1}H NMR of F_3A_3 (600 MHz, Deuterium Oxide)



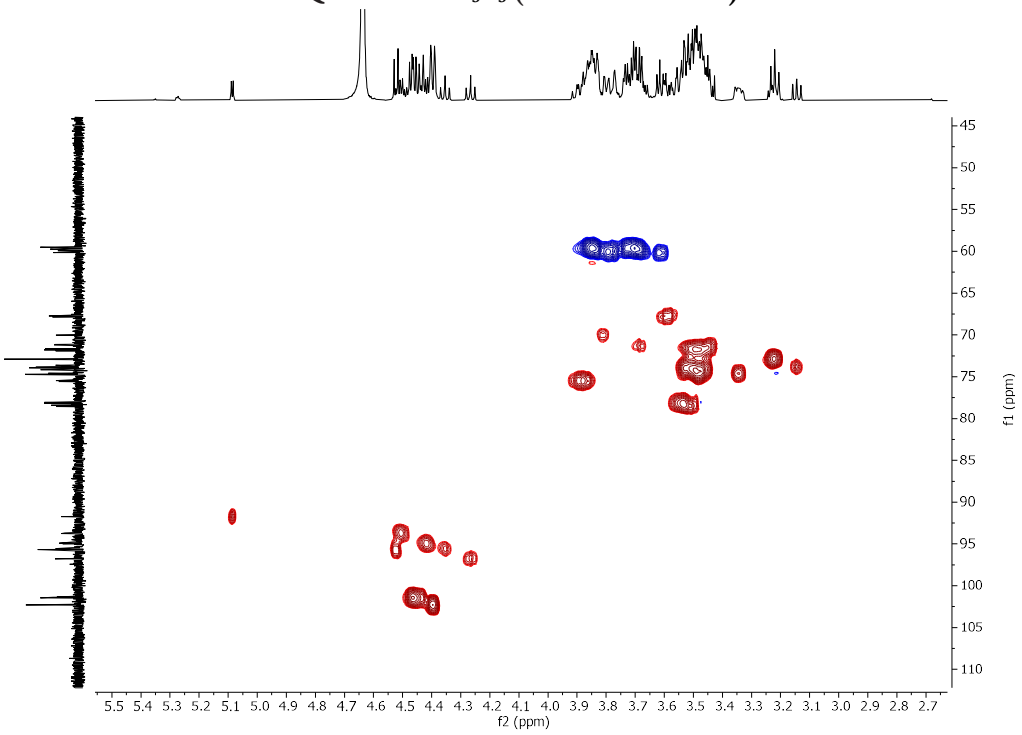
^{13}C NMR of F_3A_3 (151 MHz, Deuterium Oxide)



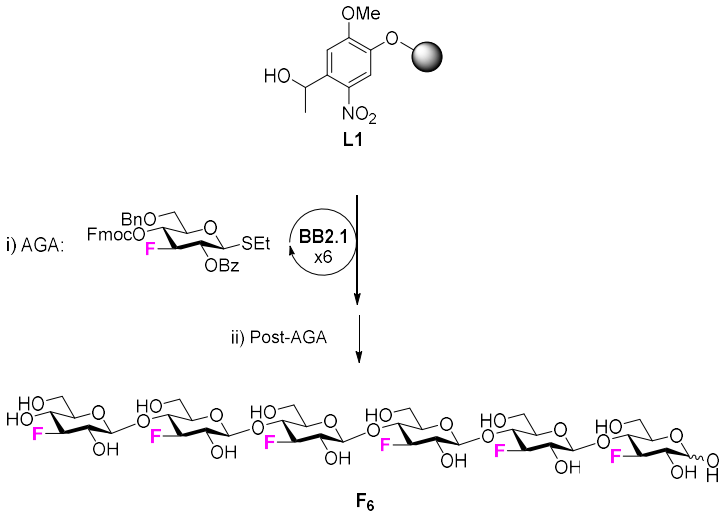
^{19}F NMR of F_3A_3 (564 MHz, Deuterium Oxide)



HSQC NMR of F_3A_3 (Deuterium Oxide)



Synthesis of **F_6**



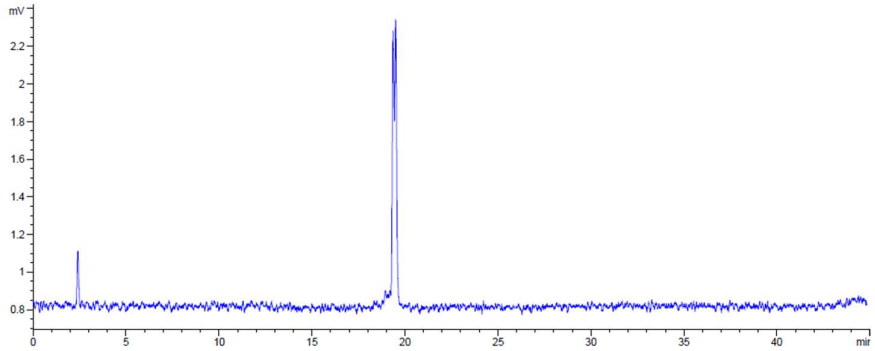
Step	BB	Modules	Notes
AGA	**-** (**BB1**)x6	**A** (**B**, **C1**, **D**, **E1**)x6	**L1** swelling **C1**: (**BB2.1**, -20°C for 5 min, 0°C for 20 min)
Post-AGA	**-**	**F1**, **G1**, **H2**, **I**	**F1**: (16 h) **H2**: (5 h) **I**: (Method D)

Automated synthesis, global deprotection, and purification afforded **F_6** as a white solid (0.4 mg, 3% overall yield).

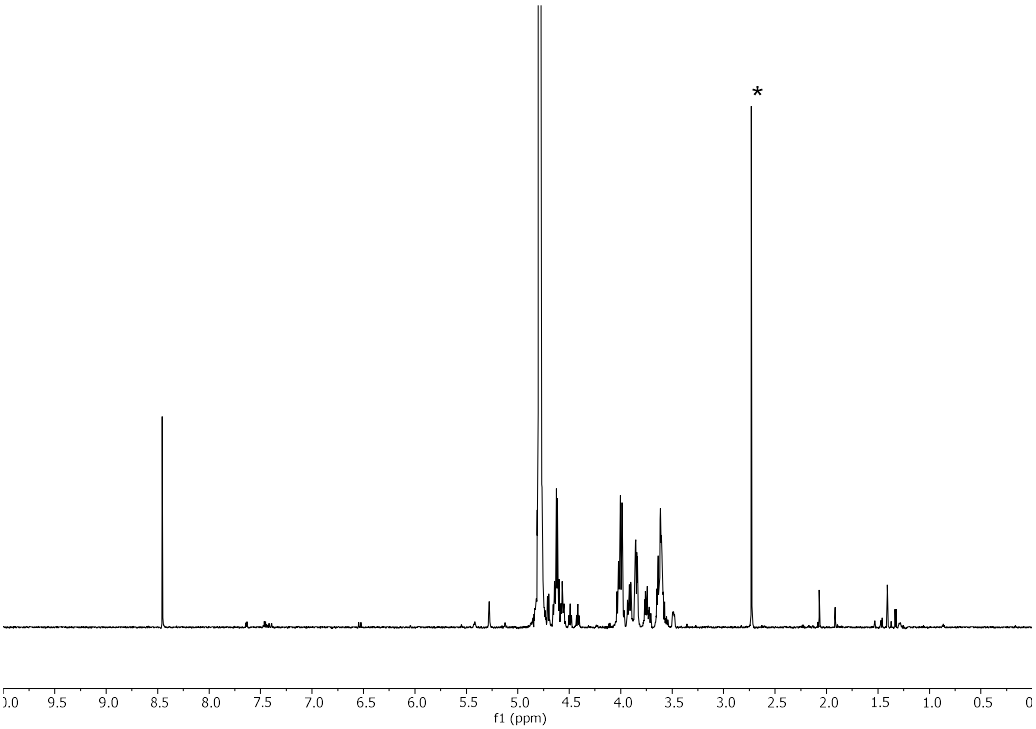
^{1}H NMR (700 MHz, D_2O) δ 5.19 (t, *J* = 3.6 Hz, 0.5H), 4.62 (d, *J* = 8.0 Hz, 1H), 4.58 – 4.45 (m, 9.5H), 4.37 (dt, *J* = 52.8, 8.9 Hz, 1H), 3.92 (dd, *J* = 17.5, 8.2 Hz, 10H), 3.85 – 3.79 (m, 3H), 3.76 (dd, *J* = 12.7, 4.7 Hz, 5H), 3.71 – 3.61 (m, 2H), 3.59 – 3.45 (m, 9H), 3.40 (s, 1H). ^{19}F NMR (564 MHz, D_2O) δ -192.14 – -192.48 (m), -194.93 – -195.09 (m). (ESI-HRMS) m/z 1025.2947 $[M+Na]^+$ ($C_{36}H_{56}F_6O_{25}Na$ requires 1025.2907).

^{13}C NMR was not recorded due to the limited amount of sample.

RP HPLC of F_6 (ELSD trace, Method C1, t_R = 19.5 min)

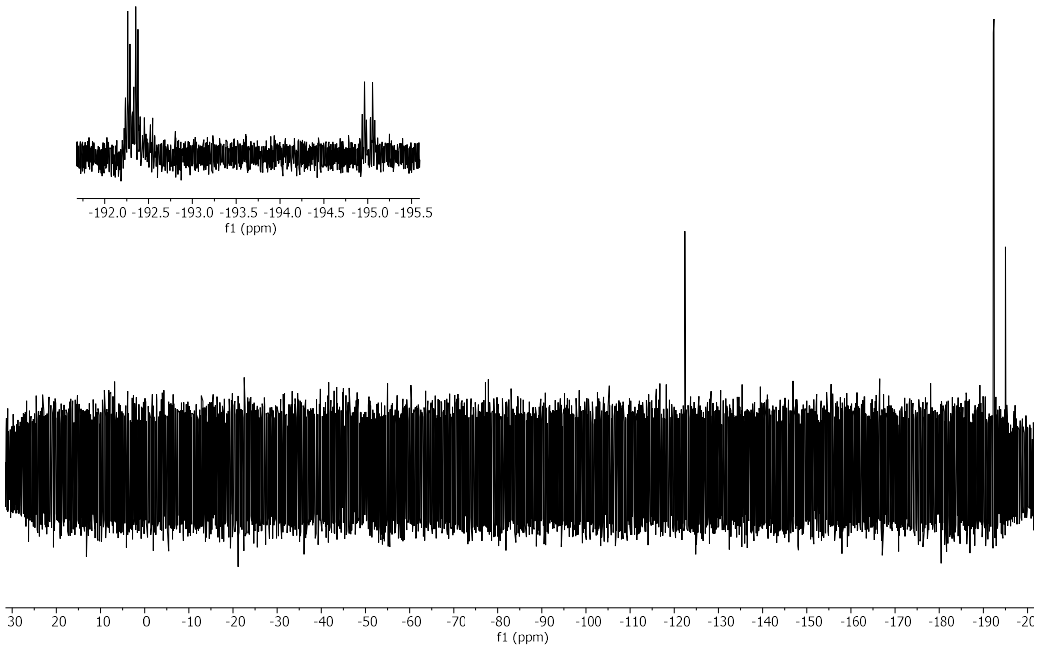


^{1}H NMR of F_6 (600 MHz, Deuterium Oxide)

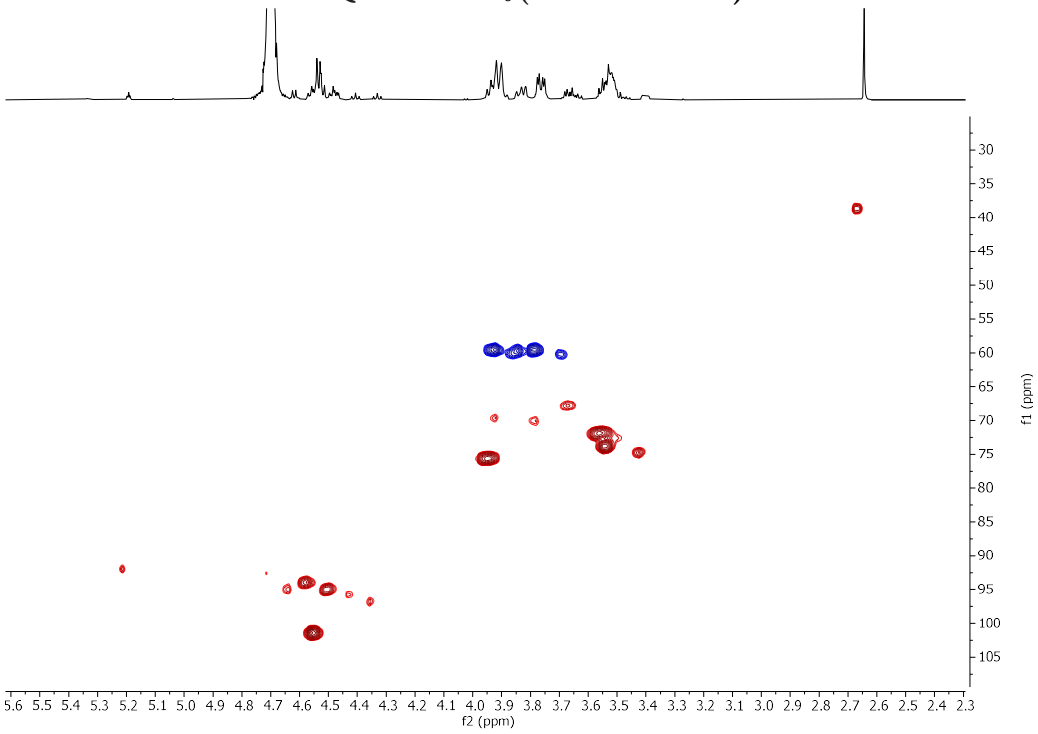


*Unidentified impurity

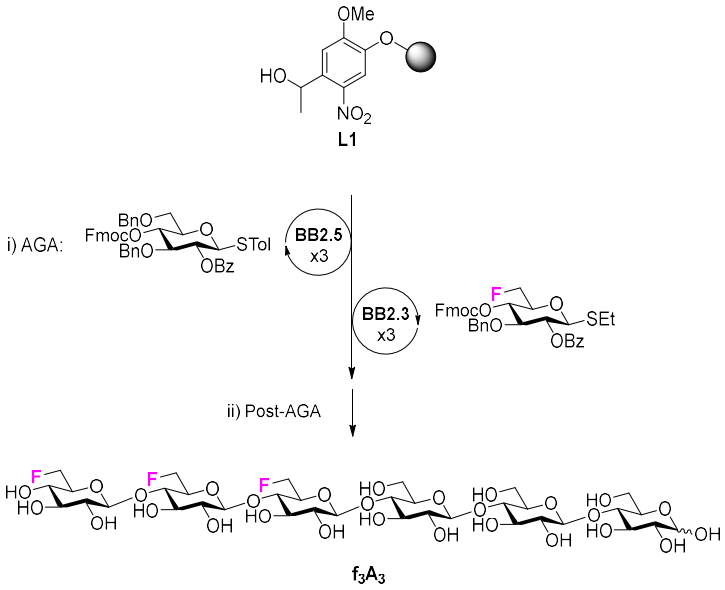
^{19}F NMR of F_6 (376 MHz, Deuterium Oxide)



HSQC NMR of F_6 (Deuterium Oxide)



Synthesis of **f_3A_3**

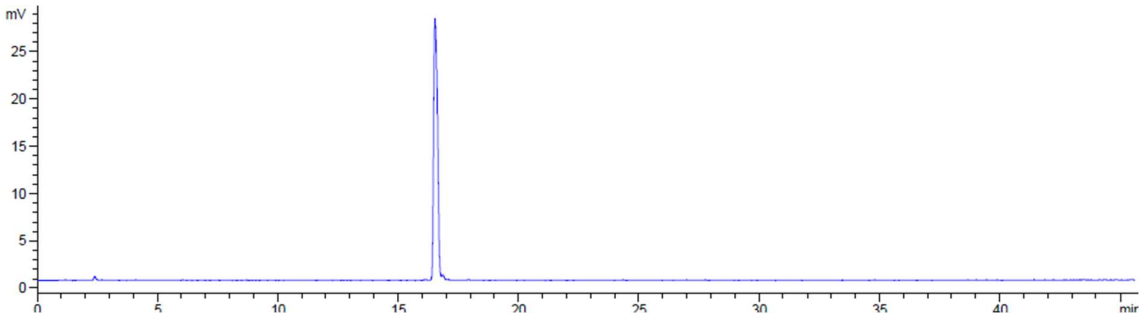


Step	BB	Modules	Notes
AGA	**-**	**A**	**L1** swelling
	(**BB2.5**)x3	(**B**, **C1**, **D**, **E1**)x3	**C1**: (**BB2.5**, -20°C for 5 min, 0°C for 20 min)
	(**BB2.3**)x3	(**B**, **C1**, **D**, **E1**)x3	**C1**: (**BB2.3**, -20°C for 5 min, 0°C for 20 min)
Post-AGA	**-**	**F1**, **G1**, **H2**, **I**	**F1**: (16 h) **H2**: (3 h) **I**: (Method D)

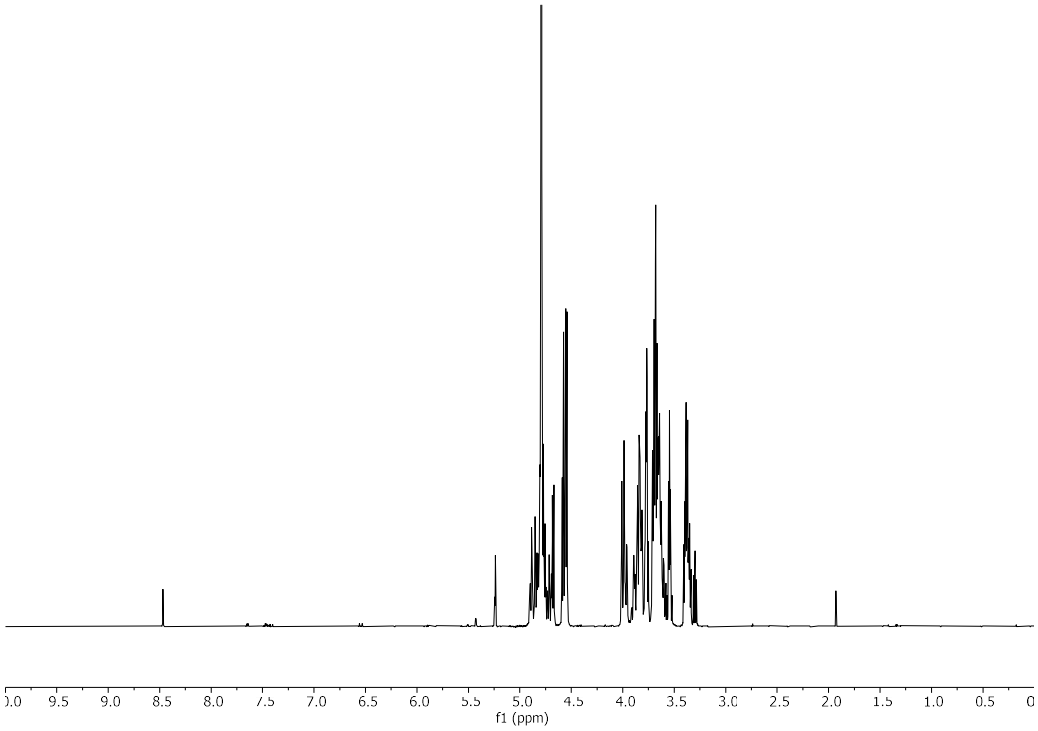
Automated synthesis, global deprotection, and purification afforded **f_3A_3** as a white solid (4.6 mg, 35% overall yield).

^{1}H NMR (600 MHz, D_2O) δ 5.24 (d, *J* = 3.8 Hz, 0.4H), 4.92 – 4.66 (m, 6H), 4.68 (d, *J* = 7.9 Hz, 0.6H), 4.62 – 4.50 (m, 5H), 4.03 – 3.94 (m, 3H), 3.93 – 3.80 (m, 5H), 3.80 – 3.74 (m, 3H), 3.73 – 3.51 (m, 14H), 3.42 – 3.27 (m, 5H). ^{13}C NMR (151 MHz, D_2O) δ 102.72, 102.51, 102.45, 102.27, 95.68, 91.75, 82.07 (d, *J* = 167.6 Hz), 81.49 (d, *J* = 167.6 Hz), 78.56, 78.46, 78.41, 78.20, 77.72, 77.69, 77.59, 77.55, 75.20, 74.75, 74.73, 74.71, 74.35, 74.24, 74.18, 73.94, 73.89, 73.82, 73.77, 73.73, 73.24, 73.20, 73.12, 73.08, 72.95, 72.92, 72.89, 72.88, 72.85, 72.78, 71.23, 71.16, 70.05, 68.29, 68.24, 59.94, 59.76. ^{19}F NMR (564 MHz, D_2O) δ -233.89 (qd, *J* = 47.7, 27.2 Hz), -234.89 (td, *J* = 47.1, 25.8 Hz). (ESI-HRMS) m/z 1019.3206 $[M+Na]^+$ ($C_{36}H_{59}F_3O_{28}Na$ requires 1019.3037).

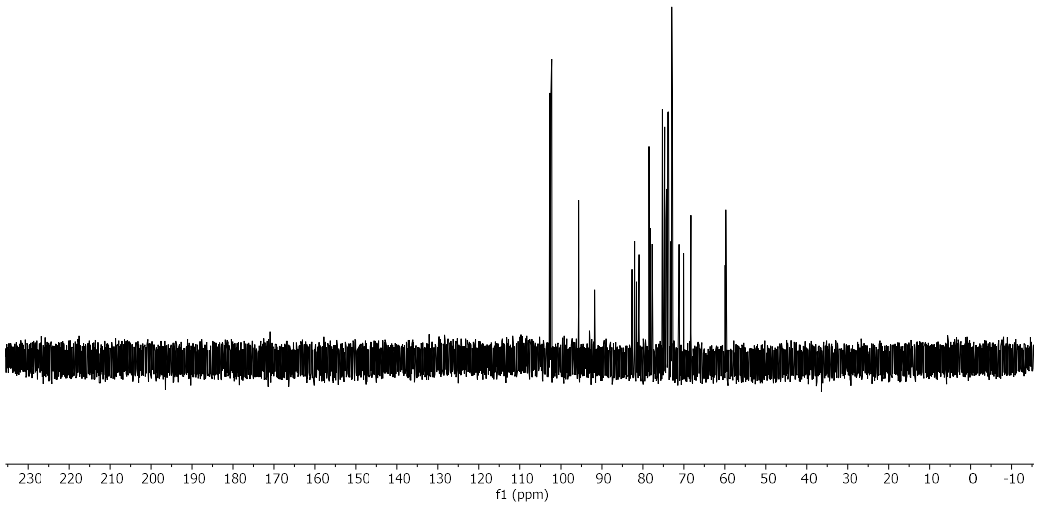
RP HPLC of f_3A_3 (ELSD trace, MethodC1, t_R = 16.5 min)



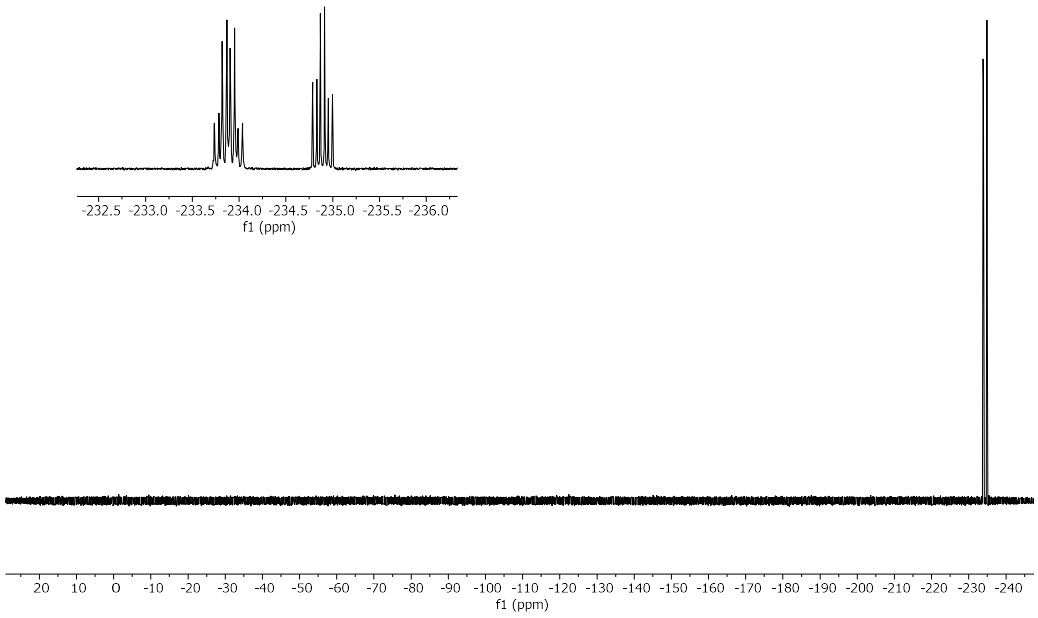
^{1}H NMR of f_3A_3 (600 MHz, Deuterium Oxide)



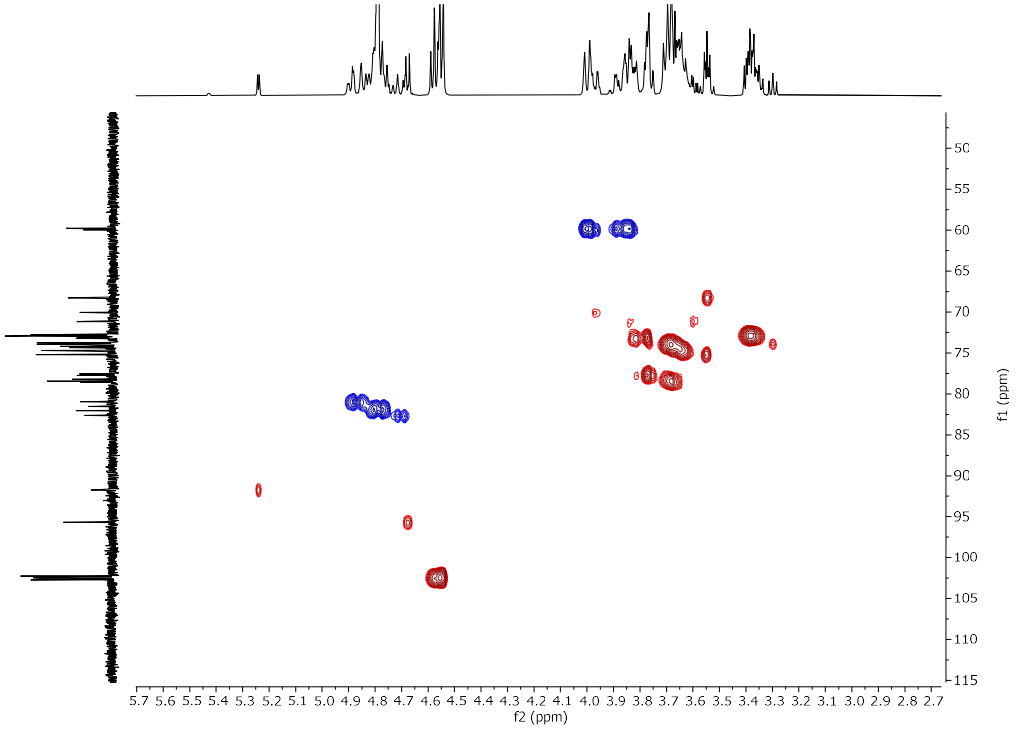
^{13}C NMR of f_3A_3 (151 MHz, Deuterium Oxide)



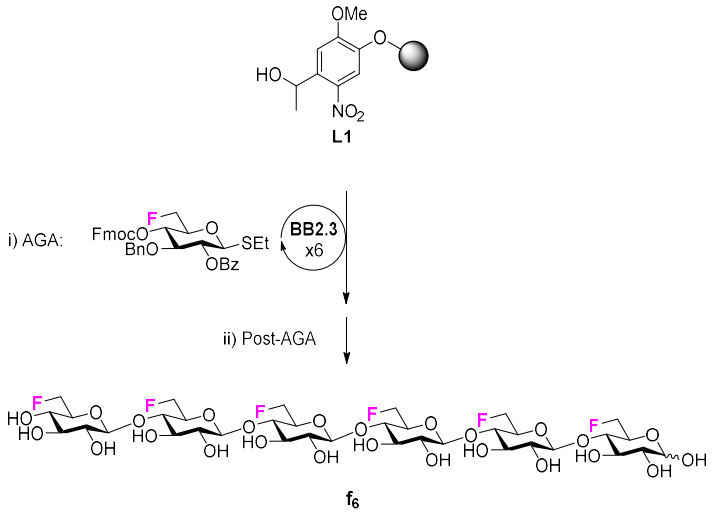
^{19}F NMR of f_3A_3 (564 MHz, Deuterium Oxide)



HSQC NMR of f_3A_3 (Deuterium Oxide)



Synthesis of **f_6**



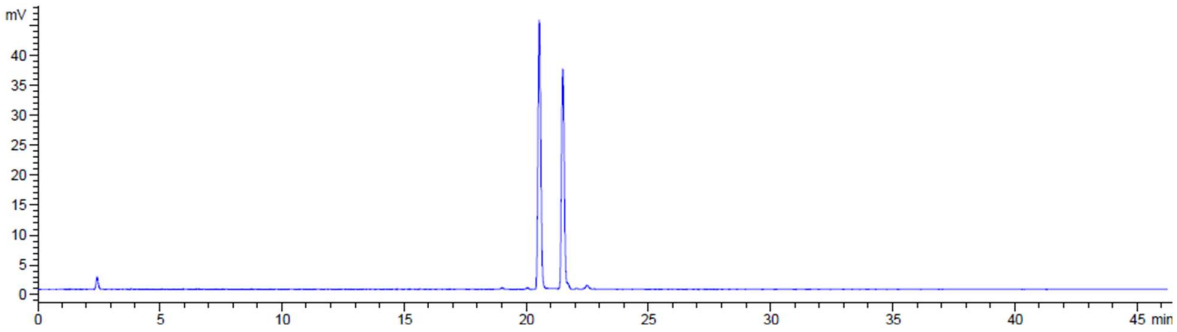
f_6

Step	BB	Modules	Notes
AGA	**-** (**BB2.3**)x6	**A** **B**, **C1**, **D**, **E1**	**L1** swelling **C1**: (**BB2.3**, -20°C for 5 min, 0°C for 20 min)
Post-AGA	**-**	**F1**, **G1**, **H2**, **I**	**F1**: (16 h) **H2**: (3 h) **I**: (Method D)

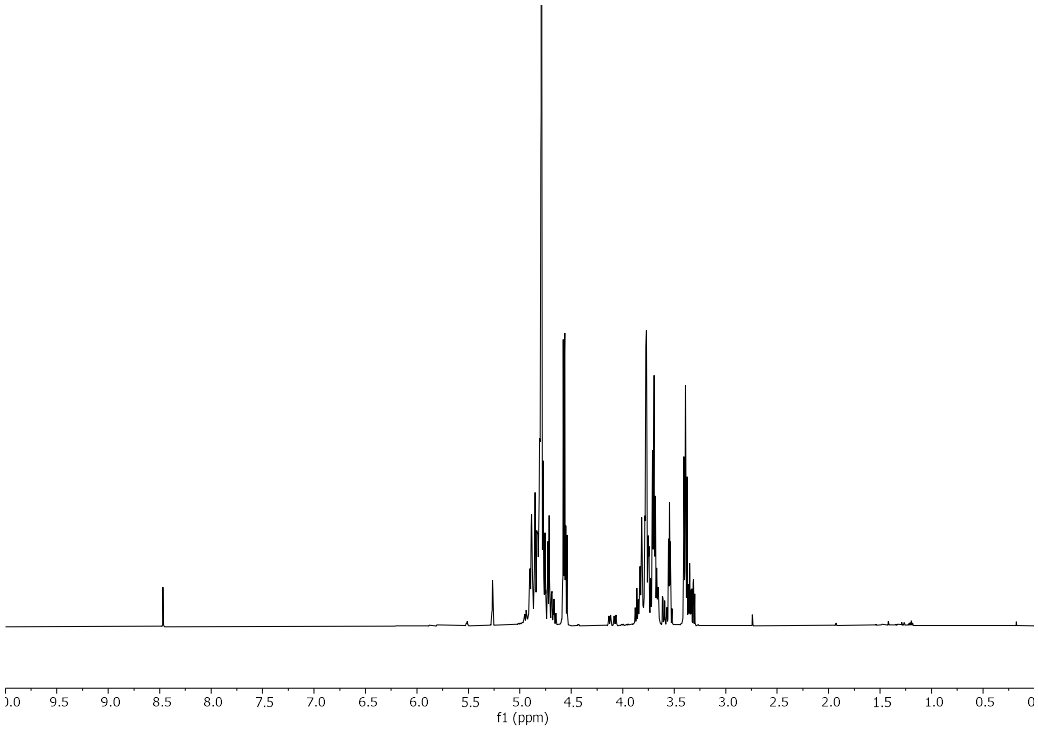
Automated synthesis, global deprotection, and purification afforded **f_6** as a white solid (3.7 mg, 30% overall yield).

^{1}H NMR (600 MHz, D_2O) δ 5.27 (d, *J* = 3.7 Hz, 0.4H), 4.79 (s, 12.6H), 4.56 (dd, *J* = 13.5, 8.0 Hz, 5H), 4.10 (dd, *J* = 31.1, 10.1 Hz, 0.6H), 3.91 – 3.63 (m, 16H), 3.60 (dd, *J* = 9.8, 3.8 Hz, 0.4H), 3.58 – 3.50 (m, 2H), 3.44 – 3.29 (m, 5H). ^{13}C NMR (151 MHz, D_2O) δ 102.71, 102.52, 95.83, 82.07 (d, *J* = 168.3 Hz), 81.47 (d, *J* = 167.8 Hz), 77.68, 77.56, 77.52, 75.19, 74.35, 74.23, 73.96, 73.73, 73.23, 73.11, 72.95, 72.82, 72.78, 71.09, 68.28. ^{19}F NMR (564 MHz, D_2O) δ -233.70 – -234.10 (m), -234.47 (td, *J* = 47.5, 31.2 Hz), -234.89 (td, *J* = 47.2, 25.7 Hz). (ESI-HRMS) m/z 1025.3114 $[M+Na]^+$ ($C_{36}H_{56}F_6O_{25}Na$ requires 1025.2907).

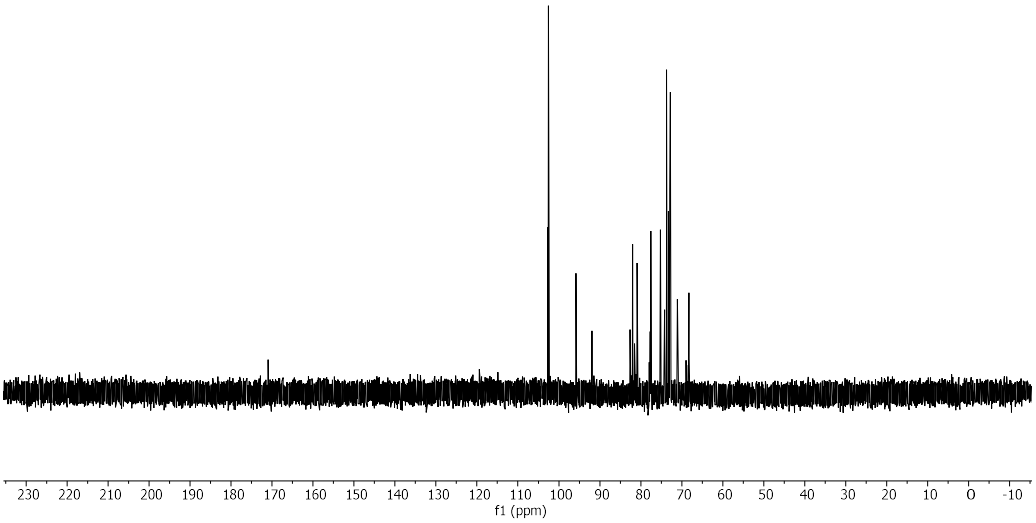
RP HPLC of f_6 (ELSD trace, Method C1, t_R = 20.5, 21.5 min)



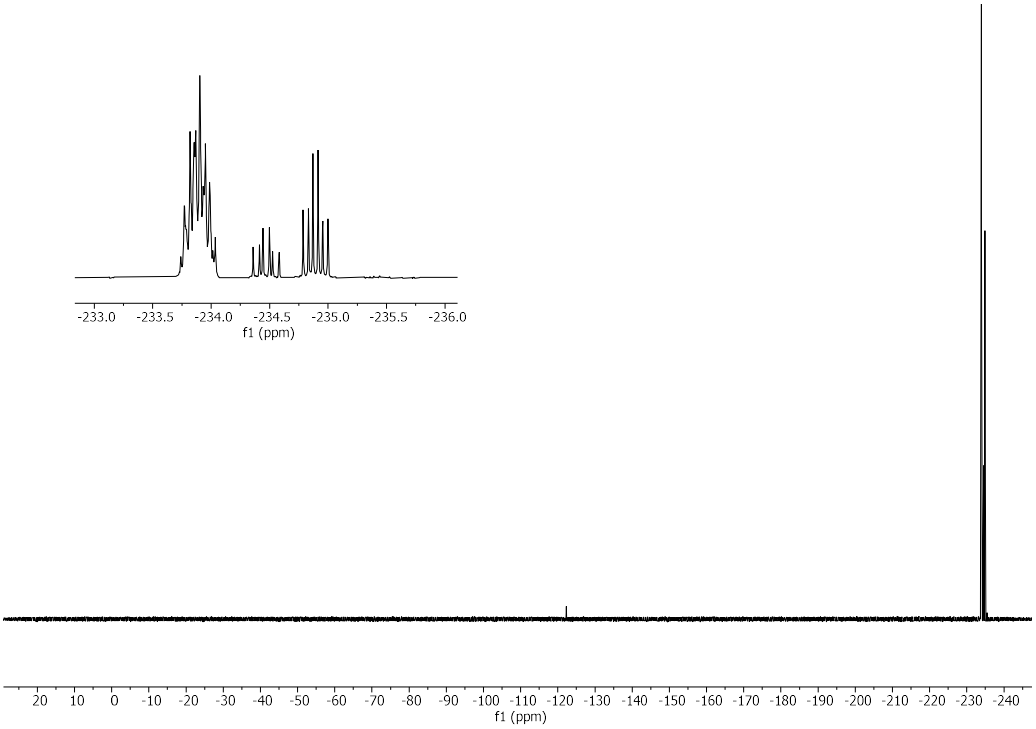
^{1}H NMR of f_6 (600 MHz, Deuterium Oxide)



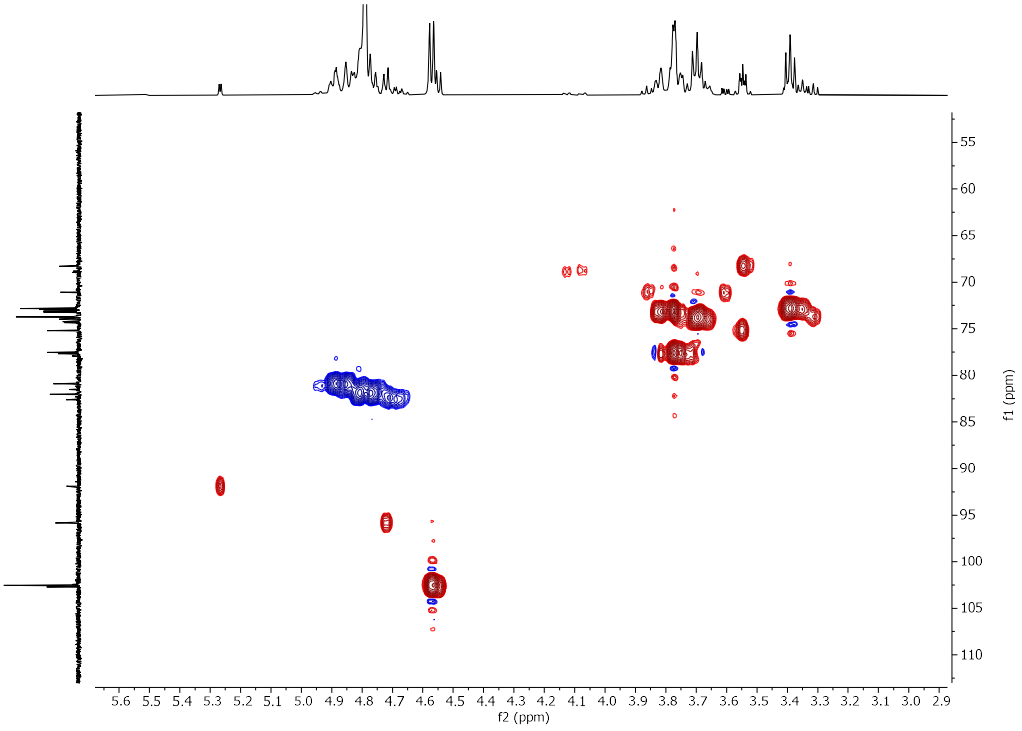
^{13}C NMR of f_6 (151 MHz, Deuterium Oxide)



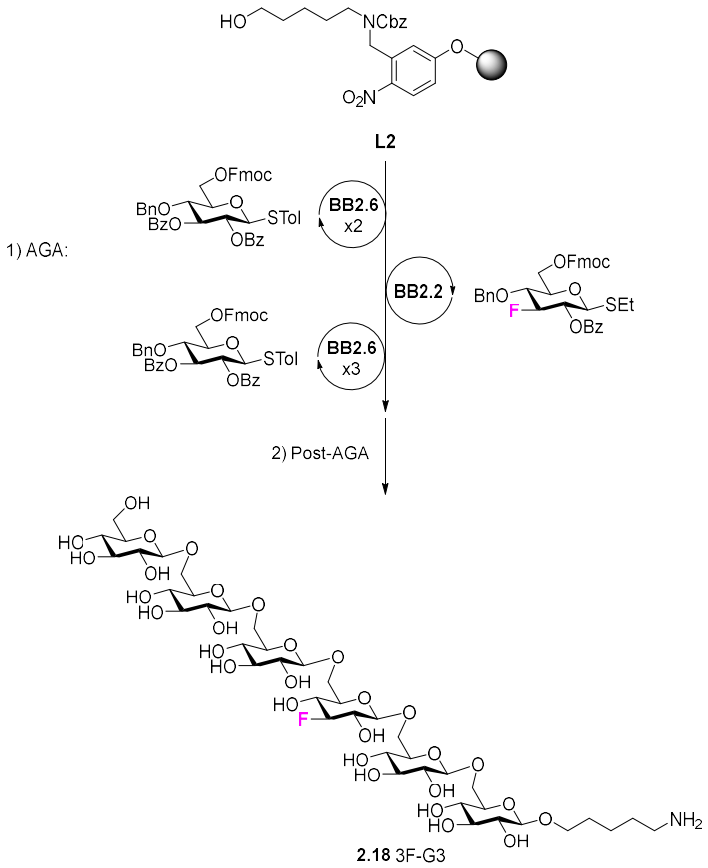
^{19}F NMR of f_6 (564 MHz, Deuterium Oxide)



HSQC NMR of f_3A_3 (Deuterium Oxide)



Synthesis of **2.18**

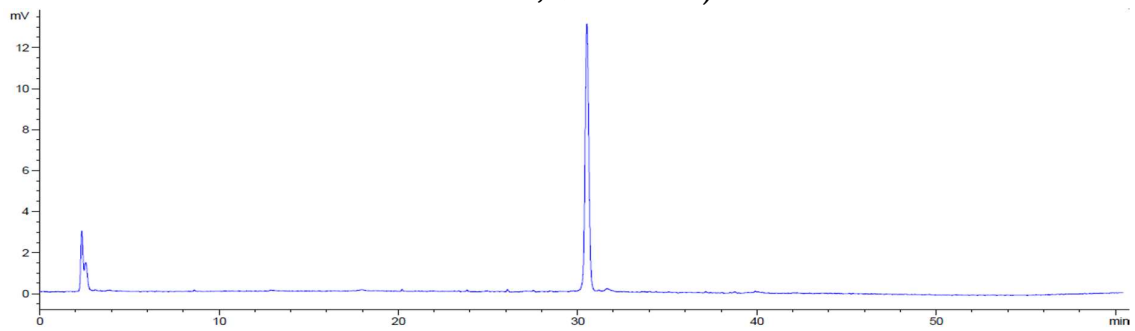


Step	BB	Modules	Notes
AGA	**-**	**A**	**L2** swelling
	(**BB2.6**)x3	(**B**, **C1**, **D**, **E1**)x3	**C1**: (**BB2.6**, -20°C for 5 min, 0°C for 20 min)
	BB2.2	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.2**, -20°C for 5 min, 0°C for 20 min)
	(**BB2.6**)x2	(**B**, **C1**, **D**, **E1**)x2	**C1**: (**BB2.6**, -20°C for 5 min, 0°C for 20 min)
Post-AGA	**-**	**G1**, **I***, **F2**, **H1**, **I****	**I***: (Method E2, t_R = 33.5 min) **F2**: (16 h) **H1**: (20 h) **I****: (Method C2: 15.0 min)

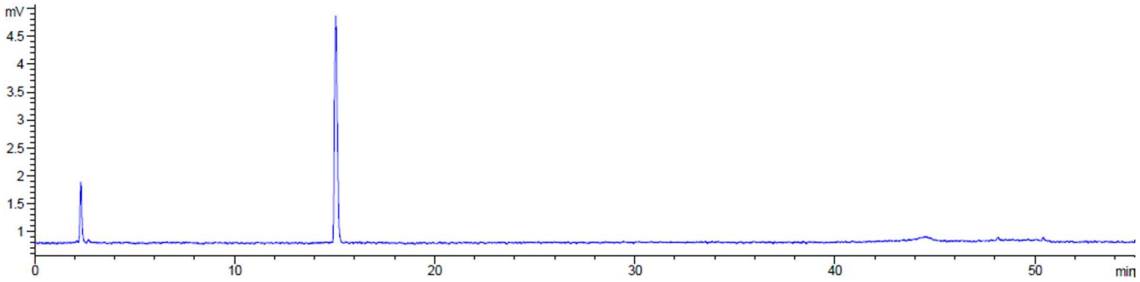
Automated synthesis, global deprotection, and purification afforded **2.18** as a white solid (2.4 mg, 17% overall yield).

^{1}H NMR (600 MHz, Deuterium Oxide) δ 4.60 (d, *J* = 8.0 Hz, 1H), 4.57 – 4.39 (m, 6H), 4.31 – 4.19 (m, 5H), 3.99 – 3.79 (m, 8H), 3.79 – 3.59 (m, 8H), 3.57 – 3.45 (m, 10H), 3.44 – 3.39 (m, 1H), 3.38 – 3.31 (m, 4H), 3.29 (t, *J* = 8.4 Hz, 1H), 3.02 (t, *J* = 7.6 Hz, 2H), 1.76 – 1.64 (m, 4H), 1.52 – 1.43 (m, 2H). ^{13}C NMR (151 MHz, Deuterium Oxide) δ 102.95, 102.83, 102.16, 102.06, 96.15 (d, *J* = 180.6 Hz), 75.85, 75.67, 75.60, 75.51, 74.85, 74.80, 73.56, 73.50, 73.01, 72.97, 71.63, 71.50, 70.13, 69.56, 69.38, 68.86, 68.77, 68.69, 68.54, 68.35, 67.76, 67.64, 60.66, 39.33, 28.13, 26.50, 22.04. ^{19}F NMR (376 MHz, Deuterium Oxide) δ -195.13 (dt, *J* = 52.9, 13.7 Hz). (ESI-HRMS) m/z 1078.420 $[M+H]^+$ ($C_{41}H_{73}FNO_{30}$ requires 1078.420).

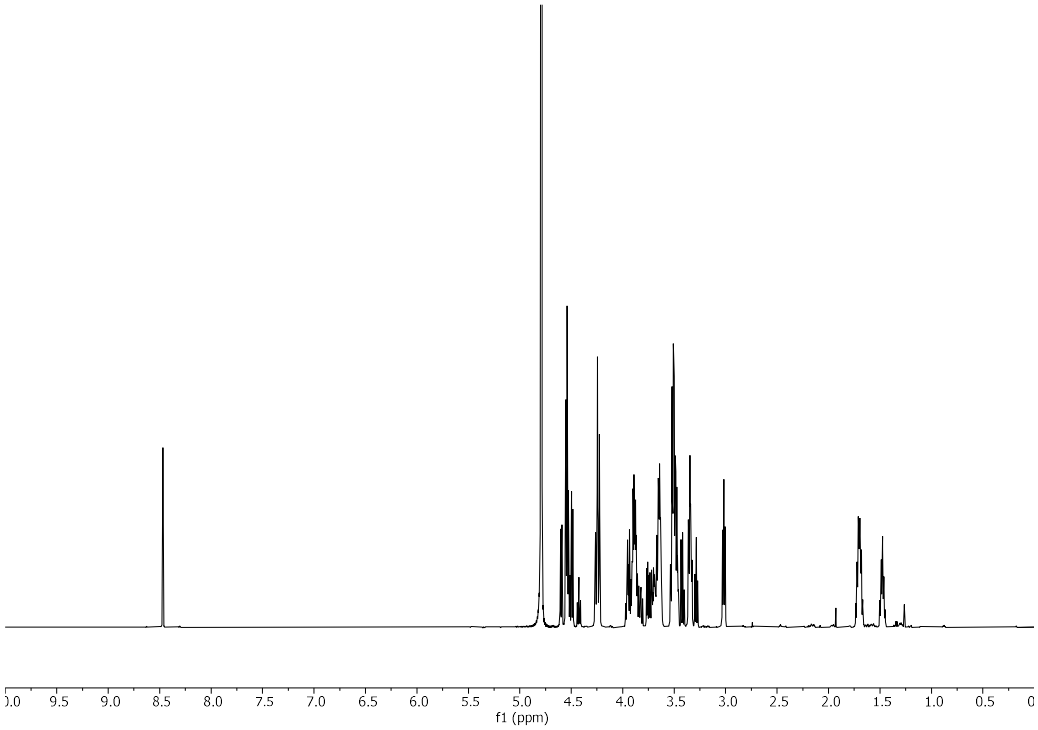
NP HPLC of crude fully protected compound after photocleavage (Module E1) (ELSD trace, Method B, t_R = 30.5 min).



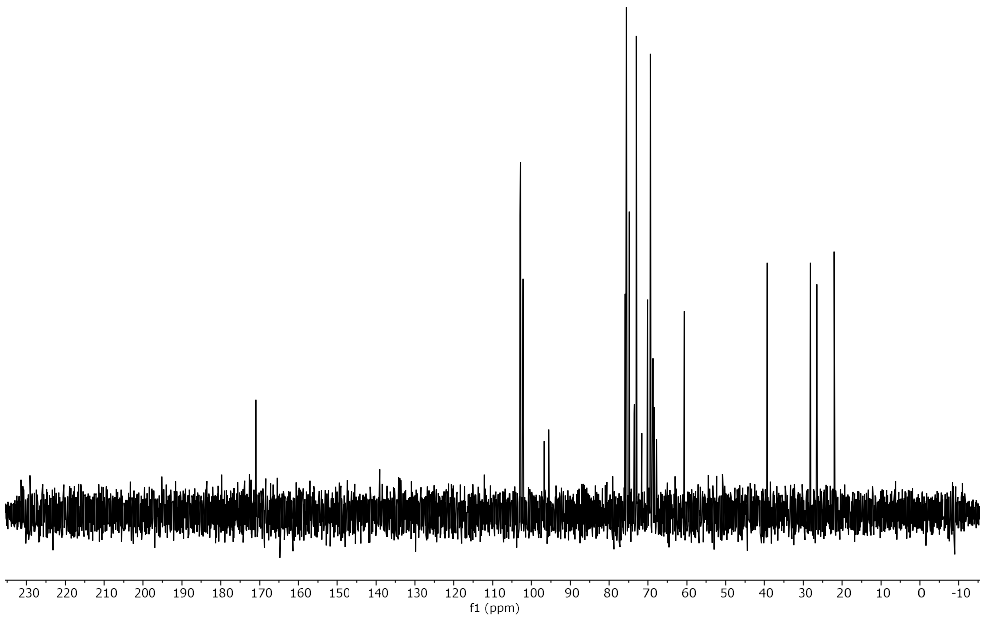
RP HPLC of 2.18 (ELSD trace, Method C1, t_R = 15.0 min)



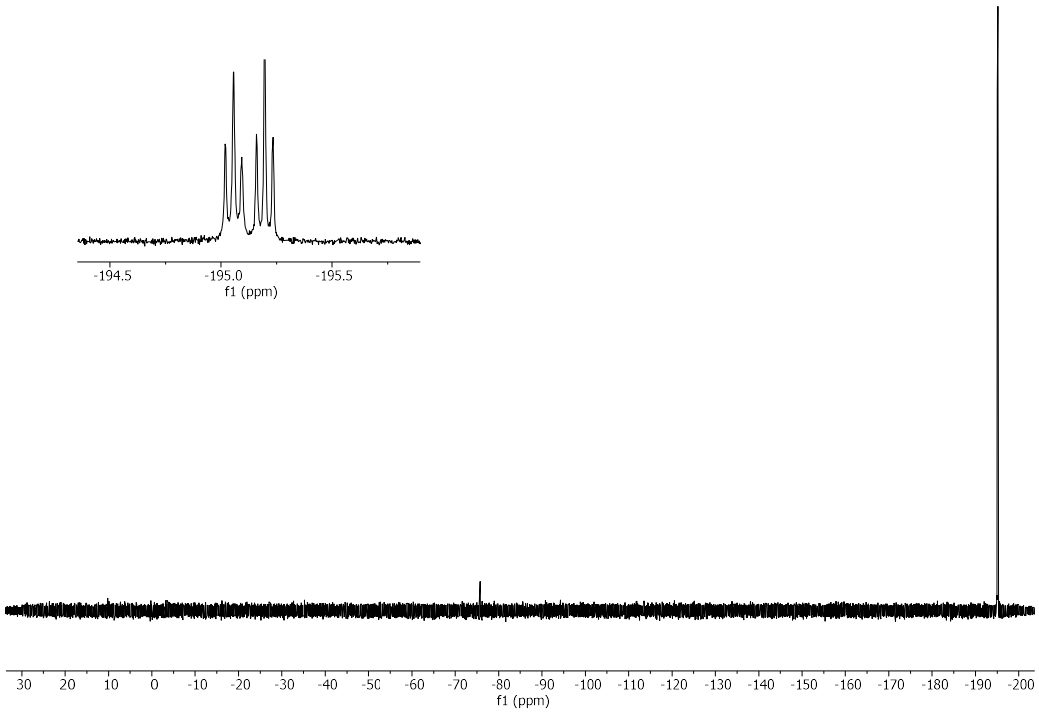
^{1}H NMR of 2.18 (600 MHz, Deuterium Oxide)



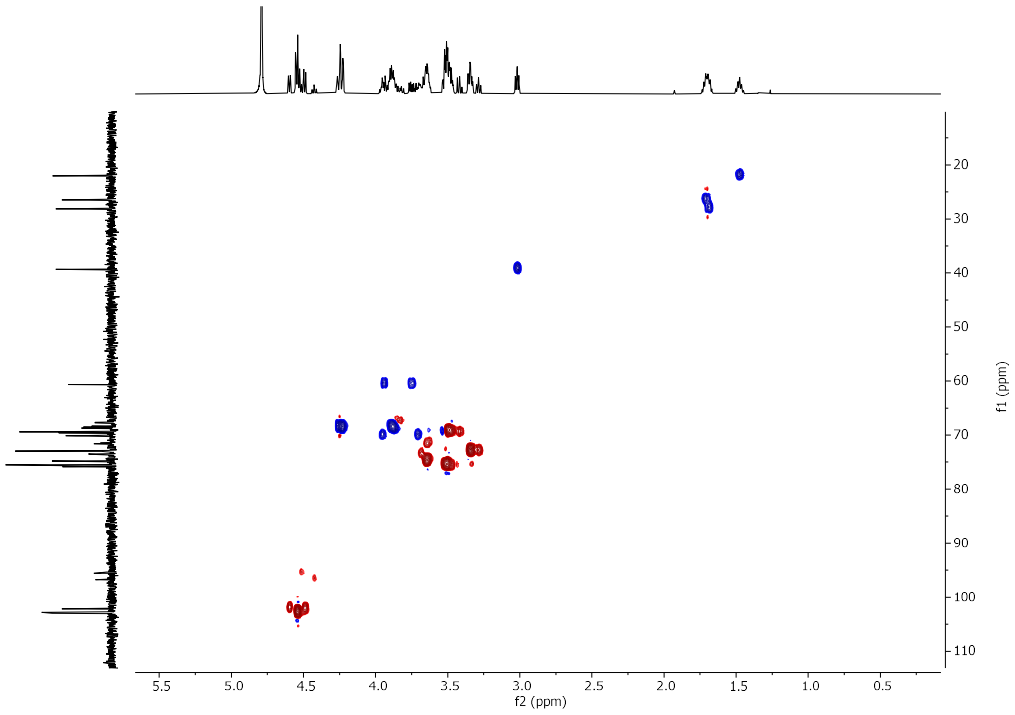
^{13}C NMR of 2.18 (151 MHz, Deuterium Oxide)



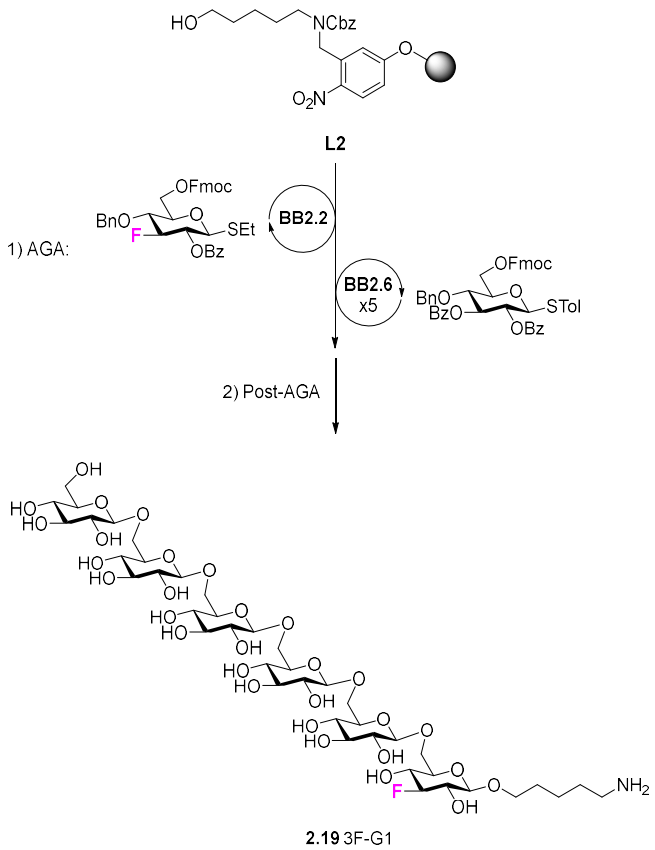
^{19}F NMR of 2.18 (376 MHz, Deuterium Oxide)



HSQC NMR of 2.18 (Deuterium Oxide)



Synthesis of **2.19**

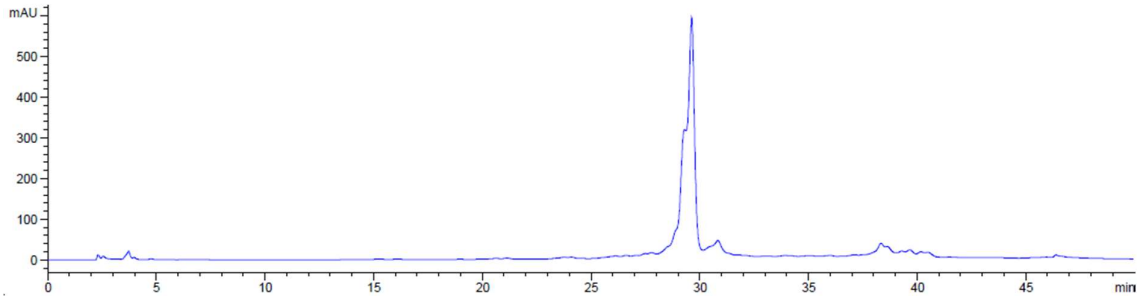


Step	BB	Modules	Notes
AGA	-	**A**	**L2** swelling
	BB2.2	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.2**, -20°C for 5 min, 0°C for 20 min)
	(**BB2.6**)x5	(**B**, **C1**, **D**, **E1**)x5	**C1**: (**BB2.6**, -20°C for 5 min, 0°C for 20 min)
Post-AGA	-	**G1**, **I***, **F2**, **H1**, **I****	**I***: (Method E2, t_R = 29.6 min) **F2**: (16 h) **H1**: (20 h) **I****: (Method C2: 14.4 min)

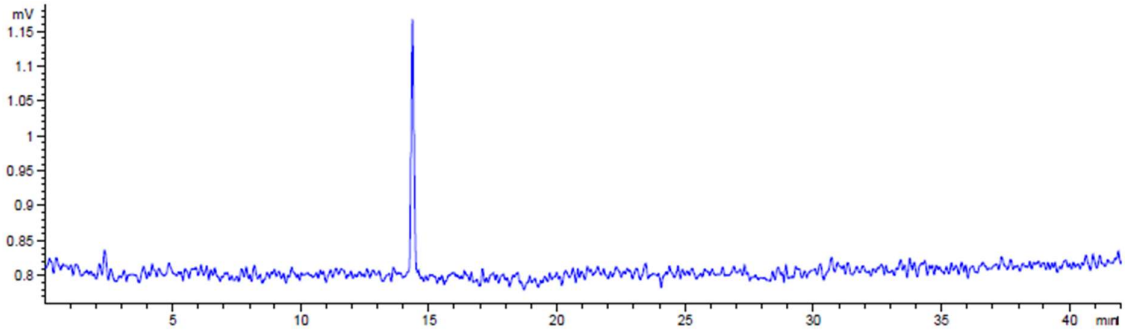
Automated synthesis, global deprotection, and purification afforded **2.19** as a white solid (3.8 mg, 21% overall yield).

^{1}H NMR (600 MHz, Deuterium Oxide) δ 4.57 – 4.52 (m, 6H), 4.46 (dt, *J* = 52.8, 8.9 Hz, 1H, H-3), 4.24 (dq, *J* = 11.3, 3.0, 2.4 Hz, 5H), 3.99 – 3.85 (m, 7H), 3.81 (ddd, *J* = 13.8, 10.1, 8.8 Hz, 1H), 3.77 – 3.62 (m, 7H), 3.62 – 3.55 (m, 1H), 3.55 – 3.45 (m, 10H), 3.45 – 3.39 (m, 1H), 3.34 (ddt, *J* = 9.6, 7.9, 2.1 Hz, 5H), 3.04 – 2.99 (m, 2H), 1.70 (h, *J* = 7.2, 6.7 Hz, 4H), 1.47 (p, *J* = 7.7, 7.2 Hz, 2H). ^{13}C NMR (151 MHz, Deuterium Oxide) δ 102.95, 102.92, 102.88, 102.83, 101.37, 101.29, 96.32 (d, *J* = 181.0 Hz, C-3), 75.86, 75.60, 75.52, 74.87, 74.84, 74.82, 73.56, 73.50, 73.02, 72.98, 71.62, 71.50, 70.31, 69.56, 69.41, 69.39, 69.35, 68.76, 68.72, 68.70, 68.54, 68.35, 67.78, 67.66, 60.66, 39.34, 28.11, 26.53, 22.03. ^{19}F NMR (376 MHz, Deuterium Oxide) δ -195.04 (dt, *J* = 52.7, 13.9 Hz). (ESI-HRMS) m/z 1078.419 $[M+H]^+$ ($C_{41}H_{73}FNO_{30}$ requires 1078.420).

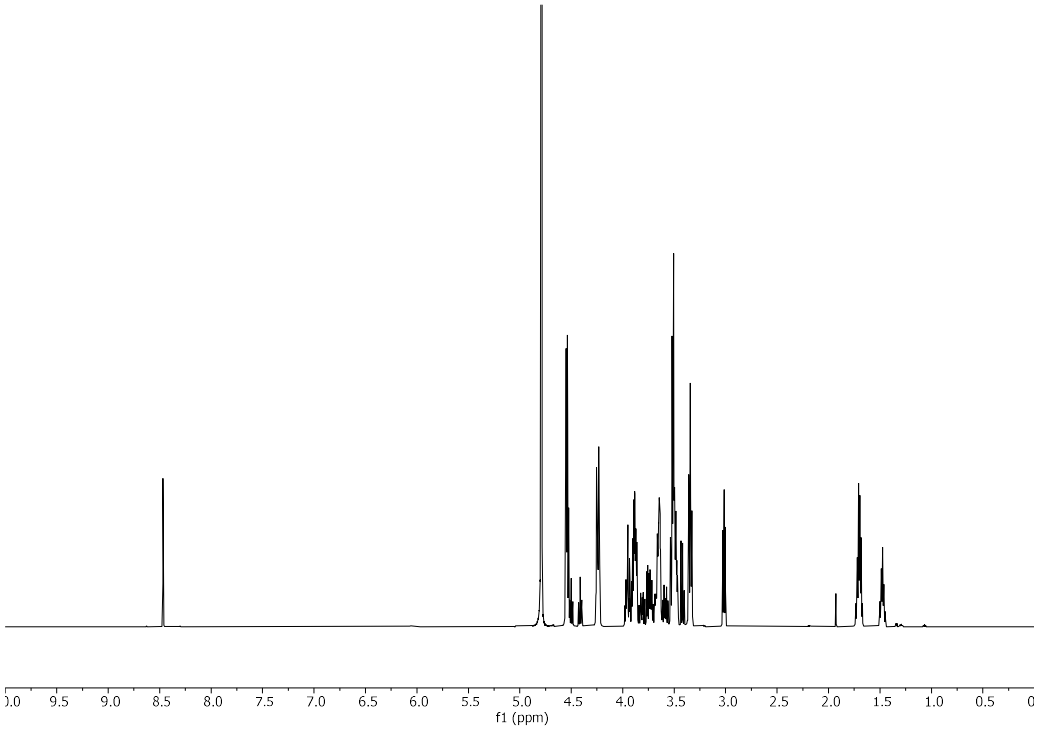
NP HPLC of crude fully protected compound after photocleavage (Module G1) (ELSD trace, Method E1, t_R = 29.6 min)



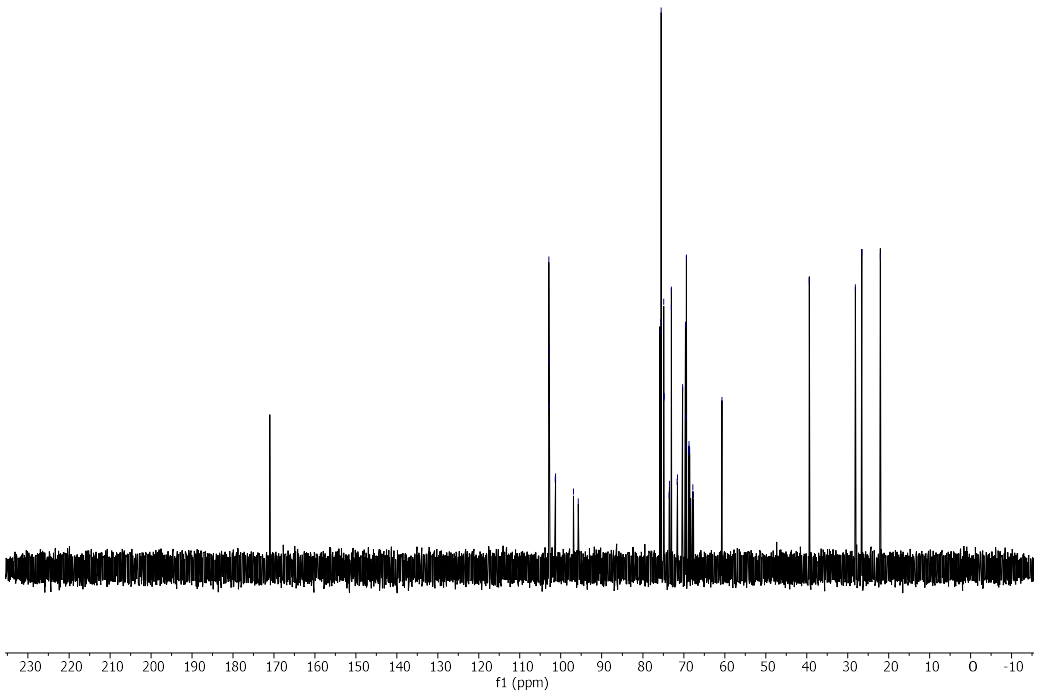
RP HPLC of 2.19 (ELSD trace, Method C1, t_R = 15.3 min)



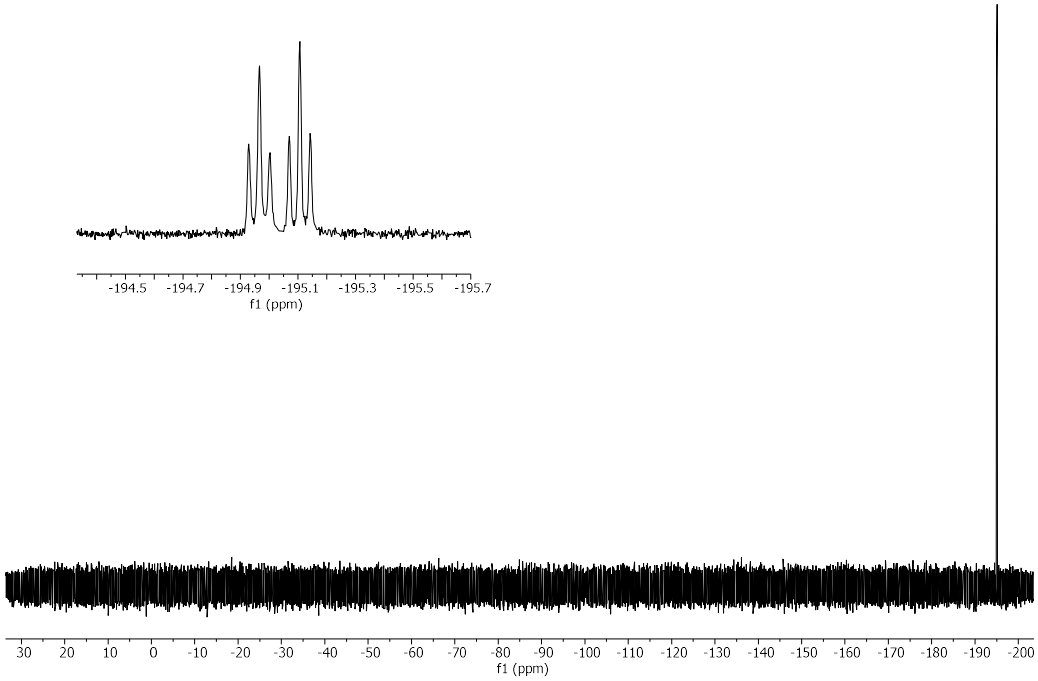
^{1}H NMR of 2.19 (600 MHz, Deuterium Oxide)



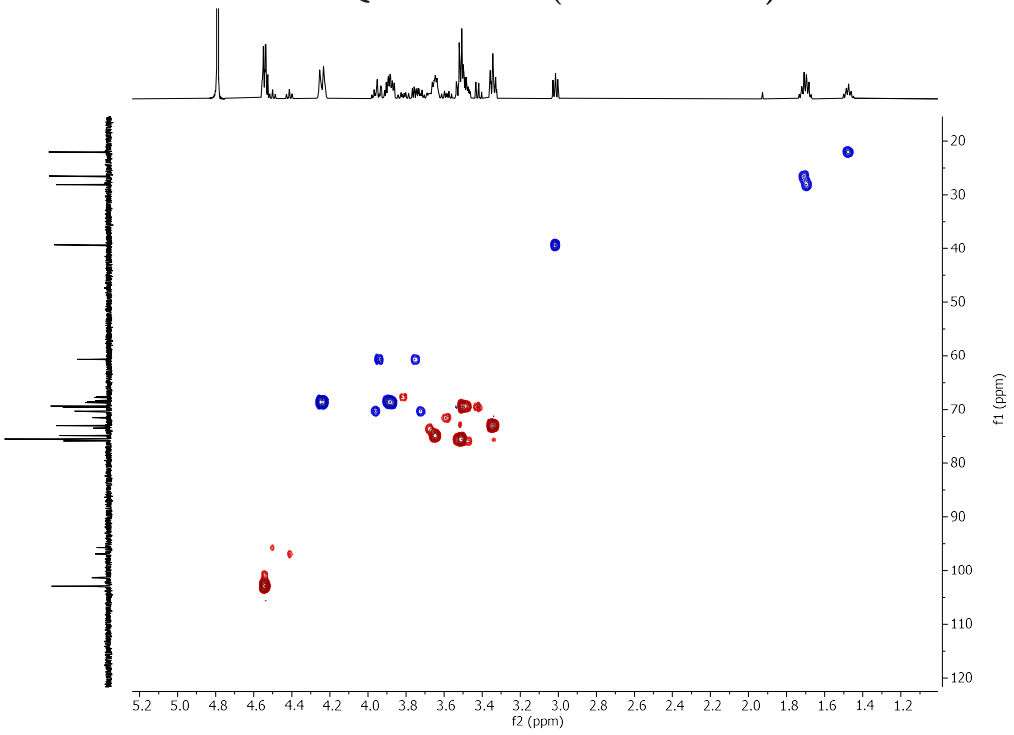
^{13}C NMR of 2.19 (151 MHz, Deuterium Oxide)



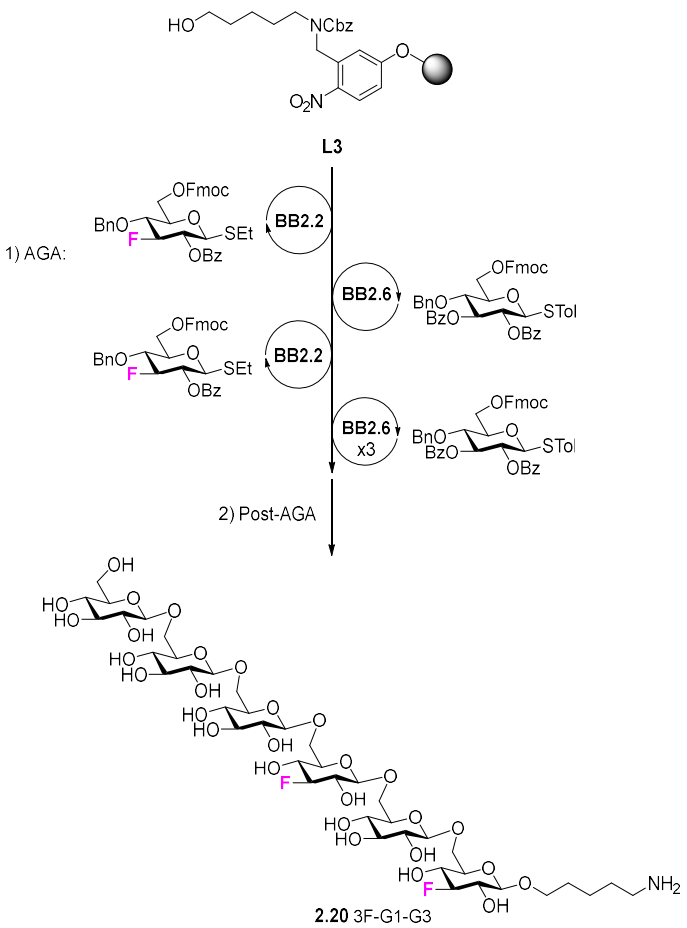
^{19}F NMR of 2.19 (376 MHz, Deuterium Oxide)



HSQC NMR of 2.19 (Deuterium Oxide)



Synthesis of **2.20**



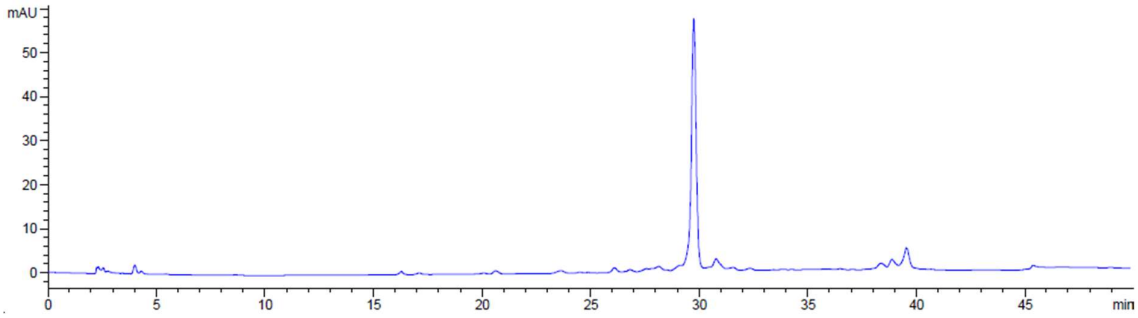
Step	BB	Modules	Notes
AGA	**-**	**A**	**L2** swelling
	BB2.2	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.2**, -20°C for 5 min, 0°C for 20 min)
	BB2.6	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.6**, -20°C for 5 min, 0°C for 20 min)
	BB2.2	**B**, **C1**, **D**, **E1**	**C1**: (**BB1**, -20°C for 5 min, 0°C for 20 min)
	(**BB2.6**)x3	(**B**, **C1**, **D**, **E1**)x3	**C1**: (**BB2.6**, -20°C for 5 min, 0°C for 20 min)
Post-AGA	**-**	**G1**, **I***, **F2**, **H1**, **I****	**I***: (Method E2, t_R = 30.2 min) **F2**: (16 h) **H1**: (20 h) **I****: (Method C2: 16.1 min)

Automated synthesis, global deprotection, and purification afforded **2.20** as a white solid (2.4 mg, 17% overall yield).

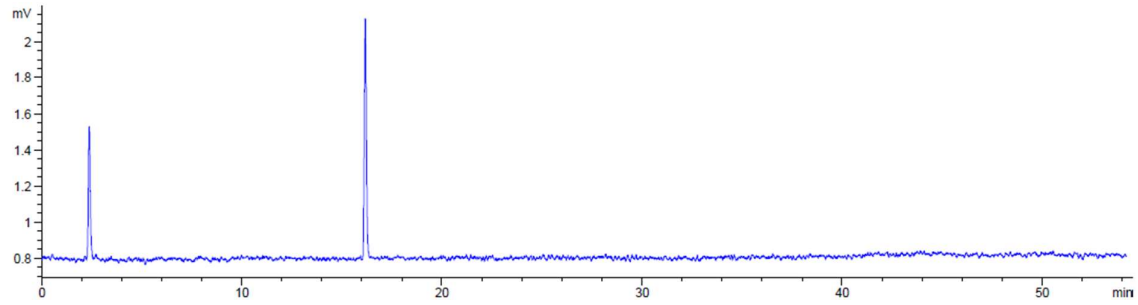
^{1}H NMR (600 MHz, Deuterium Oxide) δ 4.55 (d, *J* = 8.0 Hz, 1H), 4.49 (dd, *J* = 8.1, 5.6 Hz, 6H), 4.40 – 4.31 (m, 1H), 4.20 (dd, *J* = 11.2, 5.4 Hz, 5H), 3.98 – 3.74 (m, 9H), 3.74 – 3.51 (m, 9H), 3.51 – 3.34 (m, 9H), 3.30 (t, *J* = 8.2 Hz, 4H), 2.98 (t, *J* = 7.5 Hz, 2H), 1.66 (h, *J* = 7.2 Hz, 4H), 1.43 (p, *J* = 7.9 Hz, 2H). ^{13}C NMR (151 MHz, Deuterium Oxide) δ 102.96 (d, *J* = 2.4 Hz), 102.84 (d, *J* = 2.4 Hz), 102.15, 102.07, 101.37, 101.29, 96.32 (d, *J* = 180.9 Hz), 95.17 (d, *J* = 180.9 Hz), 75.86, 75.61, 75.53, 74.87, 74.86, 74.80, 73.57, 73.53, 73.51, 73.47, 73.02, 72.98, 72.96, 71.63, 71.51, 70.29, 69.56, 69.40, 69.34, 68.89, 68.79, 68.54, 68.35,

67.77, 67.65, 60.67, 39.31, 28.10, 26.37, 22.02. ^{19}F NMR (376 MHz, Deuterium Oxide) δ -194.87 – -195.28 (m) (ESI-HRMS) m/z 1080.415 $[M+H]^+$ ($C_{41}H_{73}FNO_{30}$ requires 1078.416).

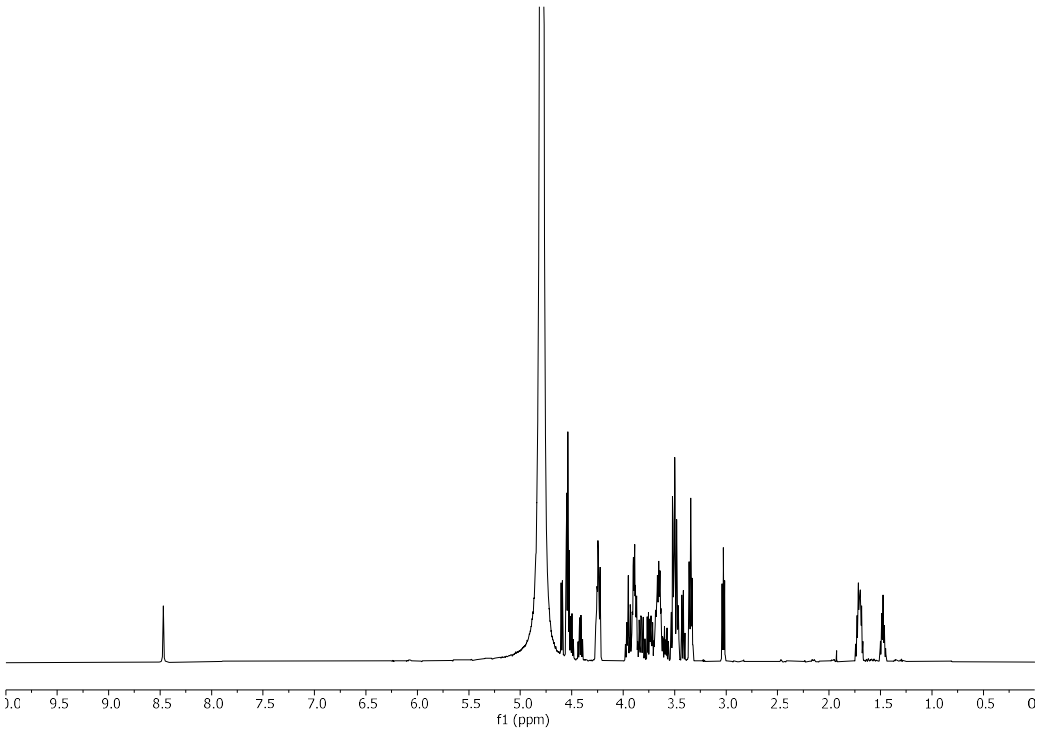
NP HPLC of crude fully protected compound after photocleavage (Module E1) (ELSD trace, Method B, t_R = 29.7 min).



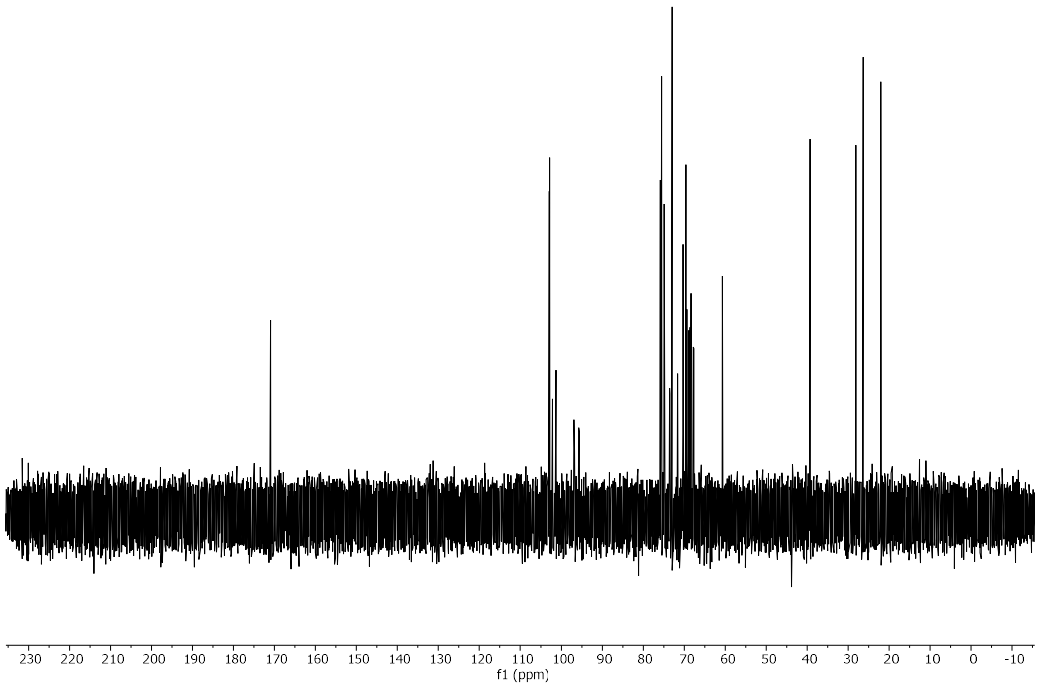
RP HPLC of 2.20 (ELSD trace, Method C1, t_R = 16.2 min).



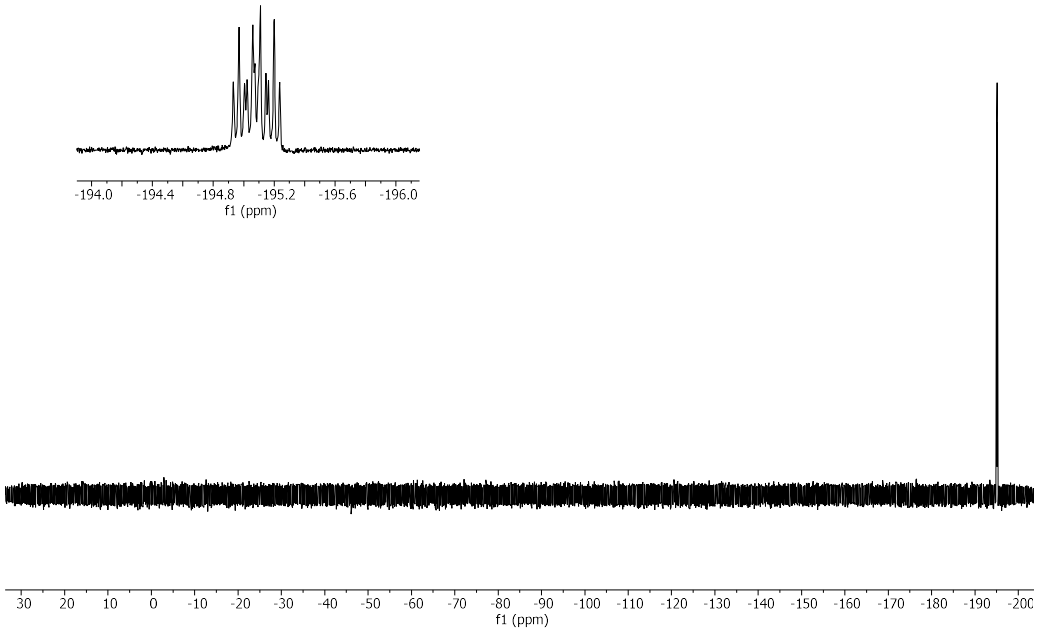
^{1}H NMR of 2.20 (600 MHz, Deuterium Oxide)



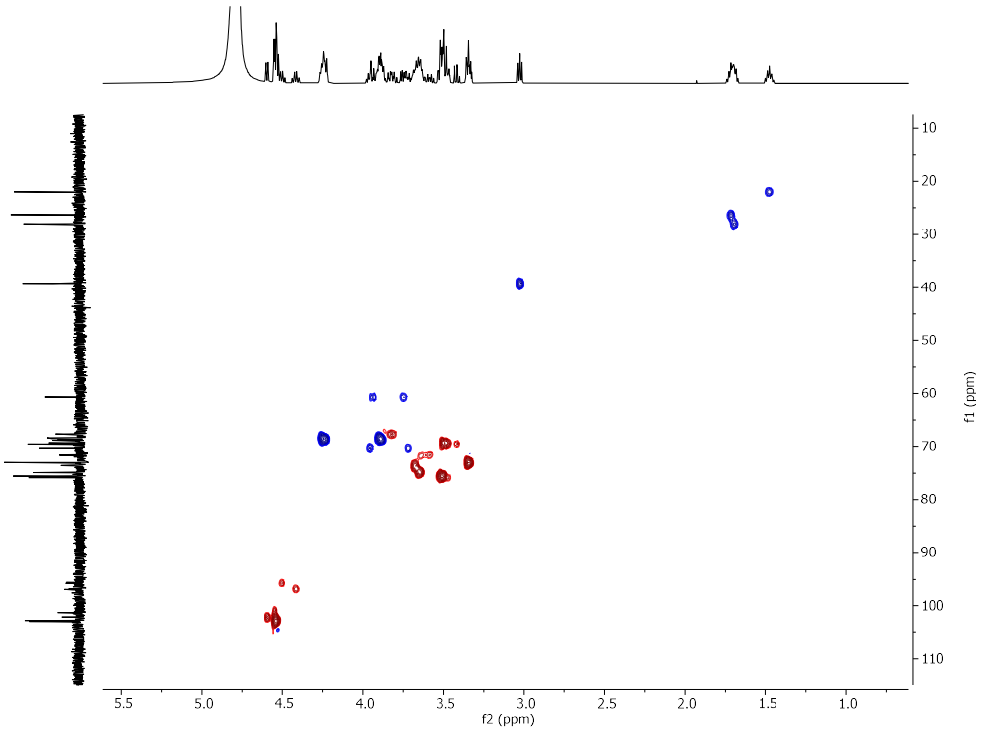
^{13}C NMR of 2.20 (151 MHz, Deuterium Oxide).



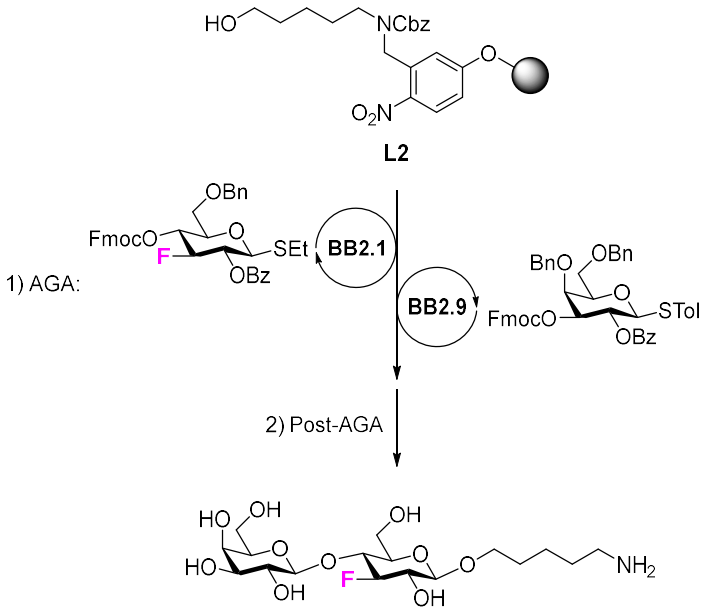
^{19}F NMR of 2.20 (376 MHz, Deuterium Oxide).



HSQC NMR of 2.20 (Deuterium Oxide)



Synthesis of **F-Lac**

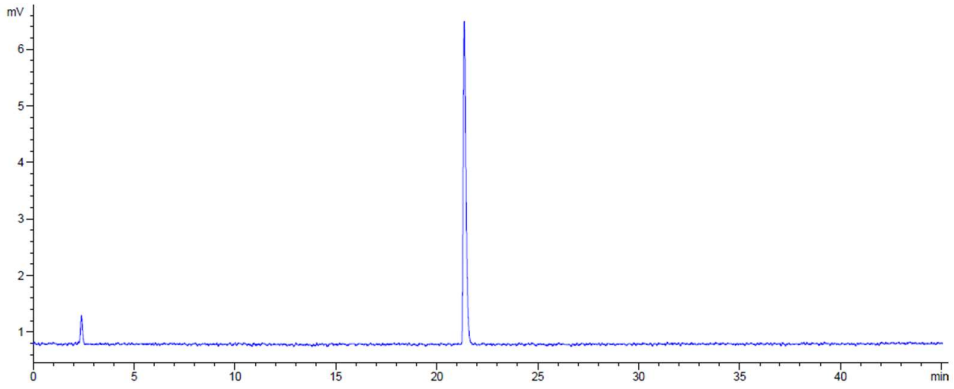


Step	BB	Modules	Notes
AGA	**-**	**A**	**L2** swelling
	BB2.1	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.1**, -20°C for 5 min, 0°C for 20 min)
	BB2.9	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.9**, -20°C for 5 min, 0°C for 20 min)
Post-AGA	**-**	**F1**, **G1**, **H2**, **I**	**F1**: (16 h) **H2**: (3 h) **I**: (Method A2, t_R = 19.6 min)

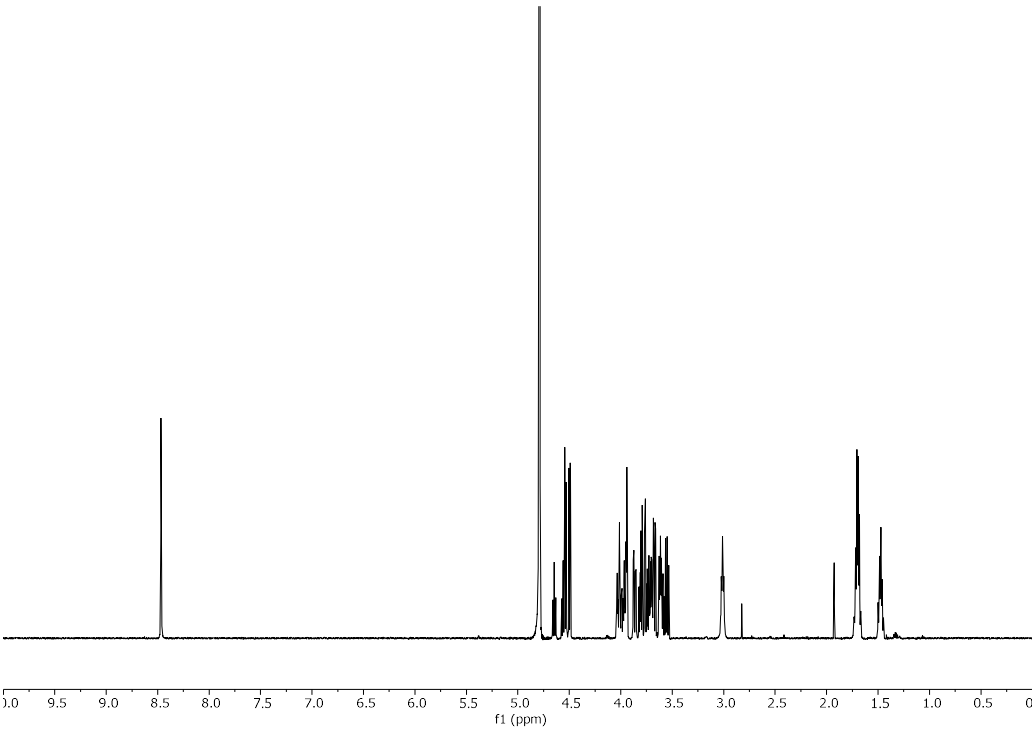
Automated synthesis, global deprotection, and purification afforded **F-Lac** as a white solid (1.17 mg, 20% overall yield).

^{1}H NMR (600 MHz, Deuterium Oxide) δ 4.67 – 4.55 (m, 1H, **H-1 Glc**, H-3 Glc), 4.54 (d, *J* = 8.1 Hz, 1H, **H-1 Gal**), 4.50 (d, *J* = 7.9 Hz, 1H), 4.05 – 3.92 (m, 4H), 3.87 (dd, *J* = 12.3, 5.1 Hz, 1H), 3.81 (dd, *J* = 11.8, 7.7 Hz, 1H), 3.78 – 3.65 (m, 4H), 3.64 – 3.53 (m, 3H), 3.01 (t, *J* = 7.5 Hz, 2H, C*H*$_2$-NH_3^+ linker), 1.70 (h, *J* = 7.4, 6.8 Hz, 4H, 2x C*H*$_2$ linker), 1.47 (p, *J* = 7.7 Hz, 2H, *CH*$_2$ linker). ^{13}C NMR (151 MHz, Deuterium Oxide) δ 102.81 (s, **C-1 Gal**), 101.11 (d, *J* = 12.1 Hz, **C-1 Glc**), 94.77 (d, *J* = 183.6 Hz, C-3 Glc), 75.46 (d, *J* = 16.9 Hz), 75.18, 73.80 (d, *J* = 8.1 Hz), 72.57, 71.72 (d, *J* = 18.1 Hz), 70.92, 70.17 (s, *C*H$_2$-O linker), 68.43, 60.79, 59.77, 39.31 (s, *C*H$_2$-NH_3^+ linker), 28.07 (s, *C*H$_2$ linker), 26.49 (s, *C*H$_2$ linker), 22.01 (s, *C*H$_2$ linker). ^{19}F NMR (564 MHz, Deuterium Oxide) δ -192.23 (dt, *J* = 52.1, 14.0 Hz). (ESI-HRMS) m/z 430.208 $[M+H]^+$ (m/z calcd for $C_{17}H_{33}FNO_{10}$: 430.208).

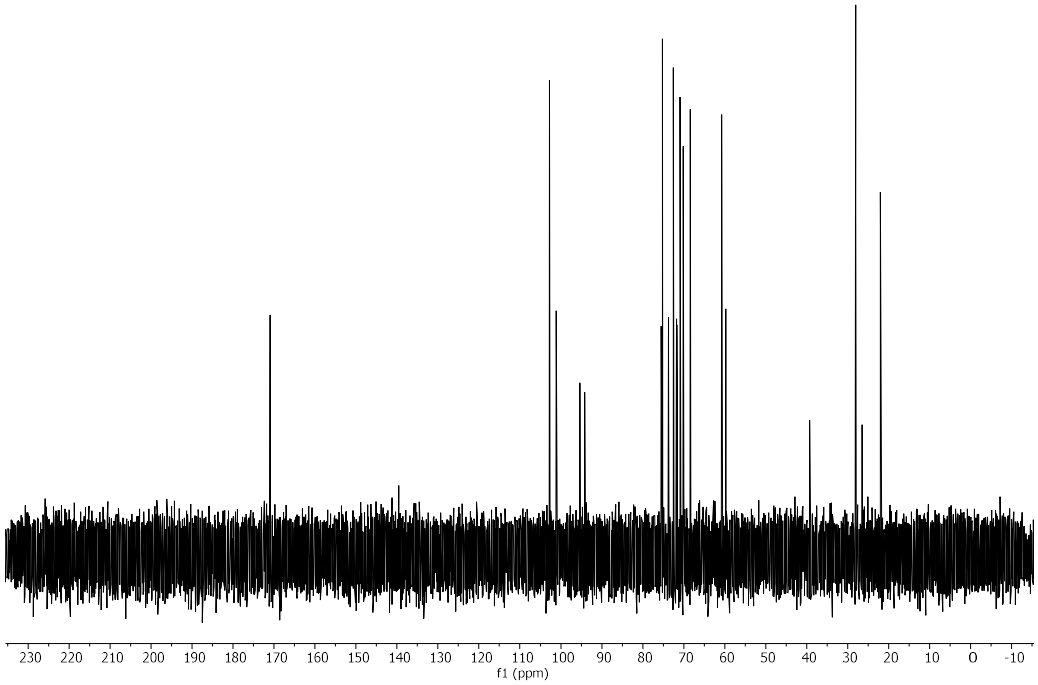
RP-HPLC of F-Lac (ELSD trace, Method A1, t_R = 21.4 min)



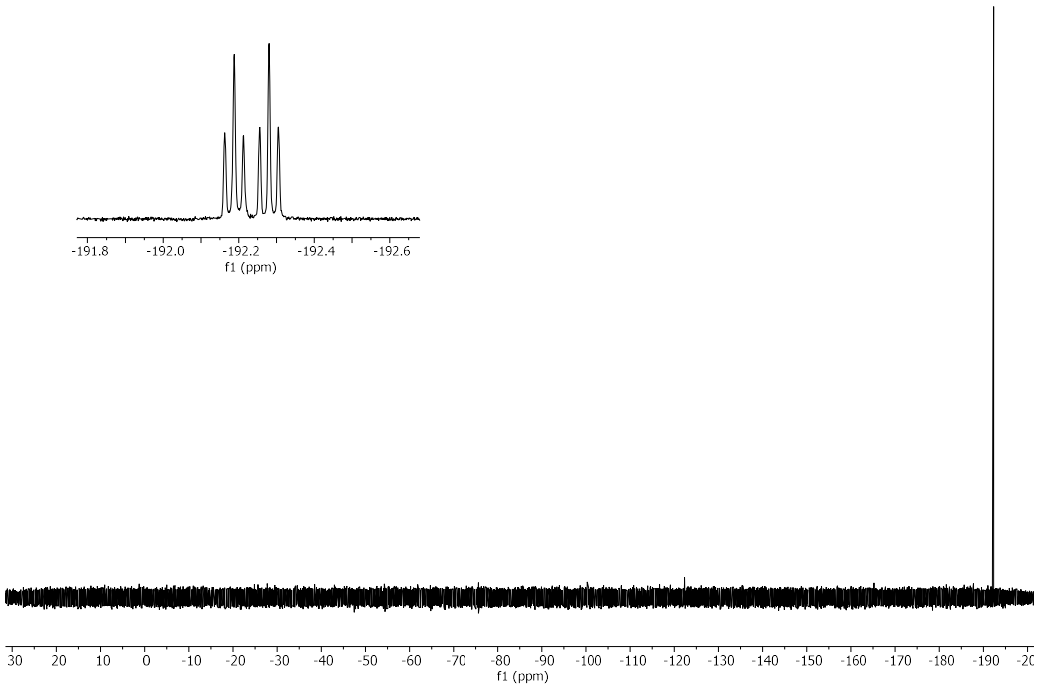
^{1}H NMR of F-Lac (600 MHz, D_2O)



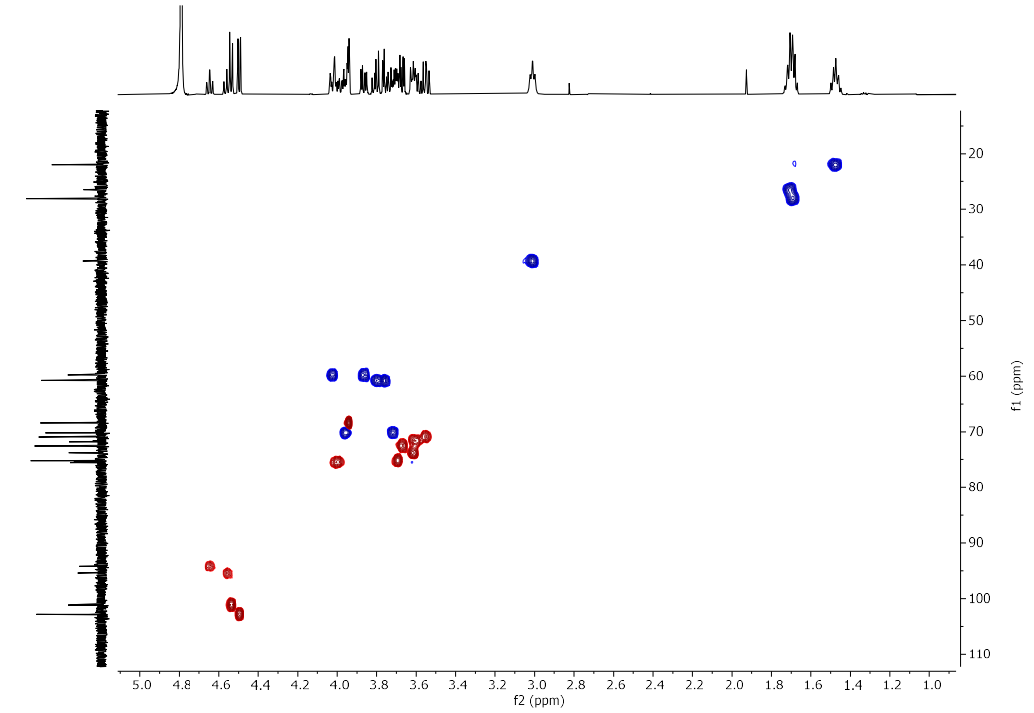
^{13}C NMR of F-Lac (151 MHz, D_2O)



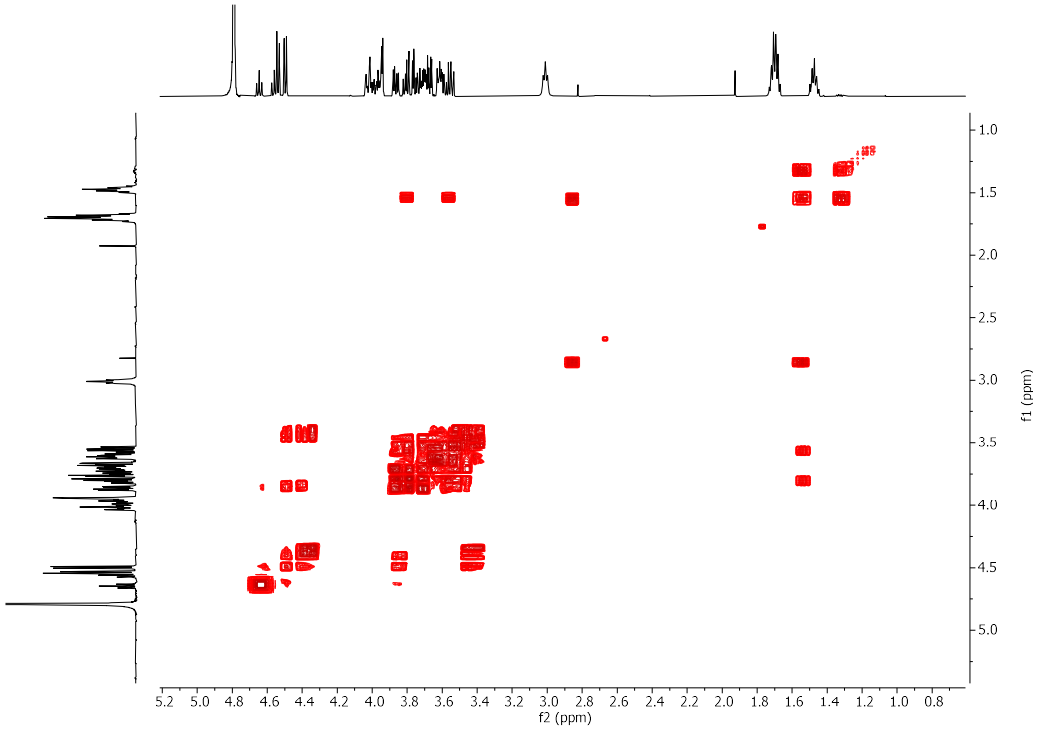
^{19}F NMR of F-Lac (564 MHz, D_2O)



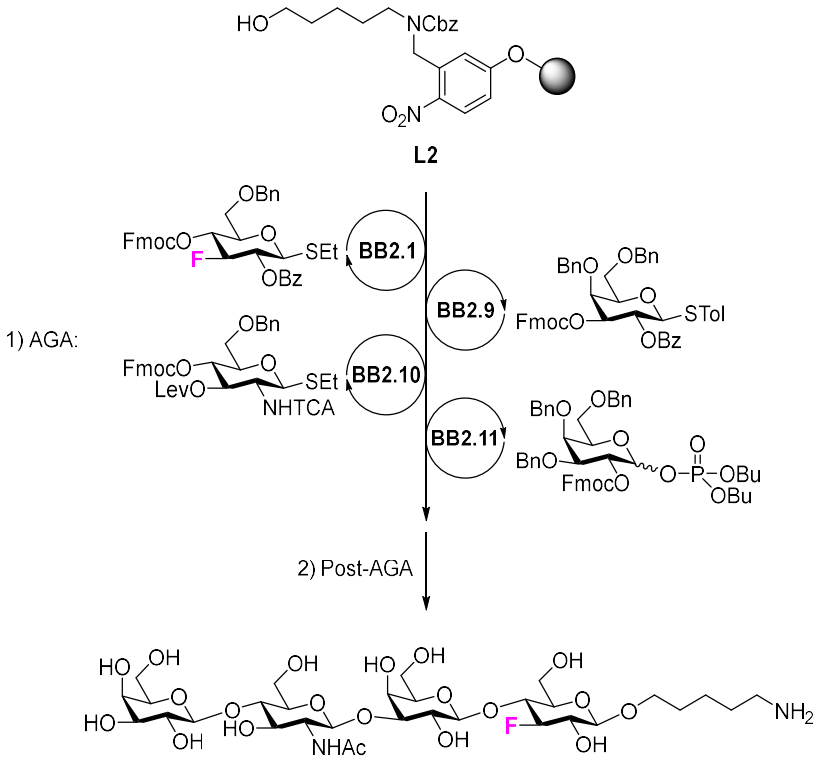
HSQC NMR of F-Lac (D_2O)



COSY NMR of F-Lac (D_2O)



Synthesis of **F-Lac$_4$**



F-nLac$_4$

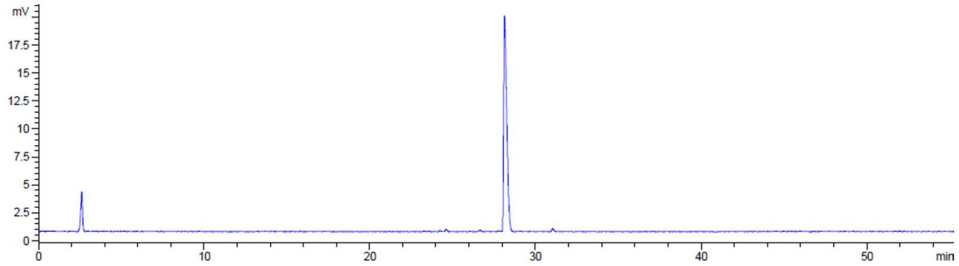
Step	BB	Modules	Notes
AGA	**-**	**A**	**L2** swelling
	BB2.1	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.1**, -20°C for 5 min, 0°C for 20 min)
	BB2.9	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.1**, -20°C for 5 min, 0°C for 20 min)
	BB2.10	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.10**, -20°C for 5 min, 0°C for 40 min)
	BB2.11	**B**, **C2**, **D**, **E1**, **E2**	**C2**: (**BB2.11**, -35°C for 5 min, -15°C for 30 min)
Post-AGA	**-**	**F1**, **G1**, **H1**, **I**	**F1:** (3 d) **H1**: (3 d) **I**: (Method D)

Automated synthesis, global deprotection, and purification afforded **F-nLac$_4$** as white solid (1.64 mg, 16% overall yield).

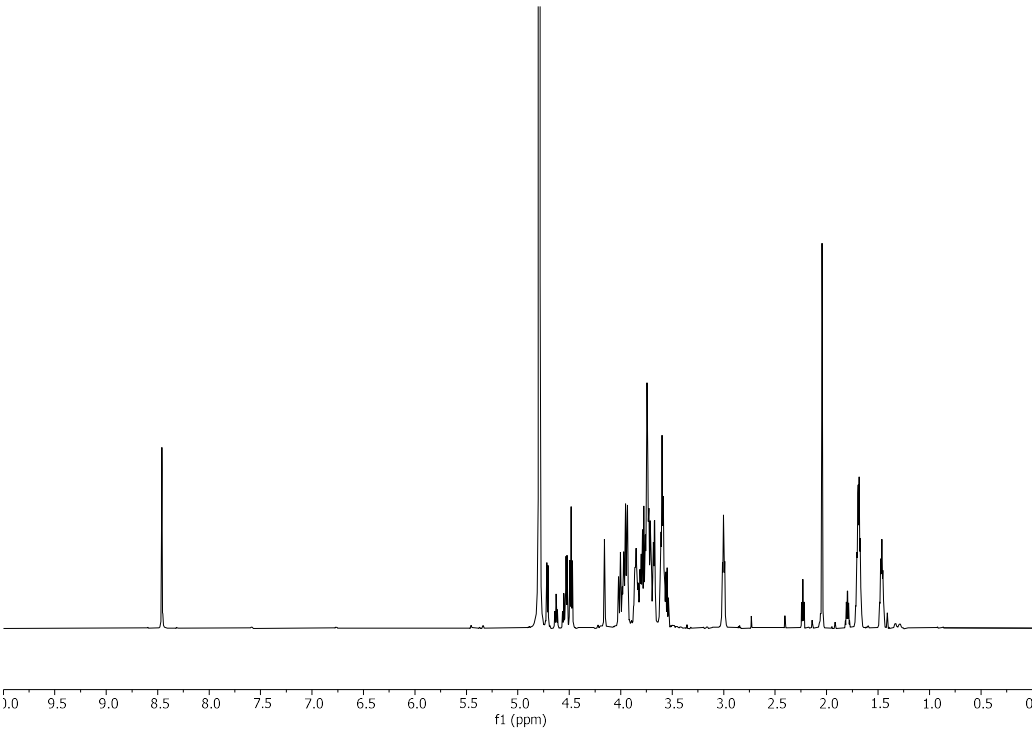
^{1}H NMR (700 MHz, Deuterium oxide) 4.71 (d, *J* = 8.4 Hz, 1H, **H-1**), 4.59 (dt, *J* = 52.0, 7.8 Hz, 1H, H-3 Glc), 4.53 (d, *J* = 8.0 Hz, 1H, **H-1**), 4.48 (t, *J* = 8.8 Hz, 2H, 2x **H-1**), 4.16 (s, 1H), 4.07 – 3.89 (m, 5H), 3.89 – 3.65 (m, 14H), 3.64 – 3.50 (m, 5H), 3.00 (t, *J* = 7.7 Hz, 2H, C*H*$_2$-NH$_3^+$ linker), 2.04 (s, 3H, C*H*$_3$ NHAc), 1.73 – 1.64 (m, 4H, 2x C*H*$_2$ linker), 1.53 – 1.43 (m, 2H, *CH*$_2$ linker). ^{13}C NMR (151 MHz, Deuterium oxide) δ 174.84, 102.83 (s, **C-1**), 102.80 (s, **C-1**), 102.66 (s, **C-1**), 101.13 (s, **C-1**) 94.72 (d, *J* = 184.3 Hz, C-3 Glc), 82.13, 78.14, 75.57, 75.29, 74.72, 74.50, 73.81, 72.44, 72.12, 71.66, 70.90, 70.15 (s, *C*H$_2$-O linker), 69.92, 68.48, 68.18, 60.96, 60.72, 59.78, 55.14, 39.28 (s, *C*H$_2$ linker), 28.06 (s, *C*H$_2$ linker), 26.33 (s, *C*H$_2$ linker),

22.11 (s, *C*H$_3$ NHAc), 21.99 (s, *C*H$_2$ linker). ^{19}F NMR (564 MHz, Deuterium oxide) δ -192.16 (dt, *J* = 52.1, 14.0 Hz). (ESI-HRMS) m/z 795.344 $[M+H]^+$ (m/z calcd for $C_{31}H_{56}FN_2O_{20}$: 795.341).

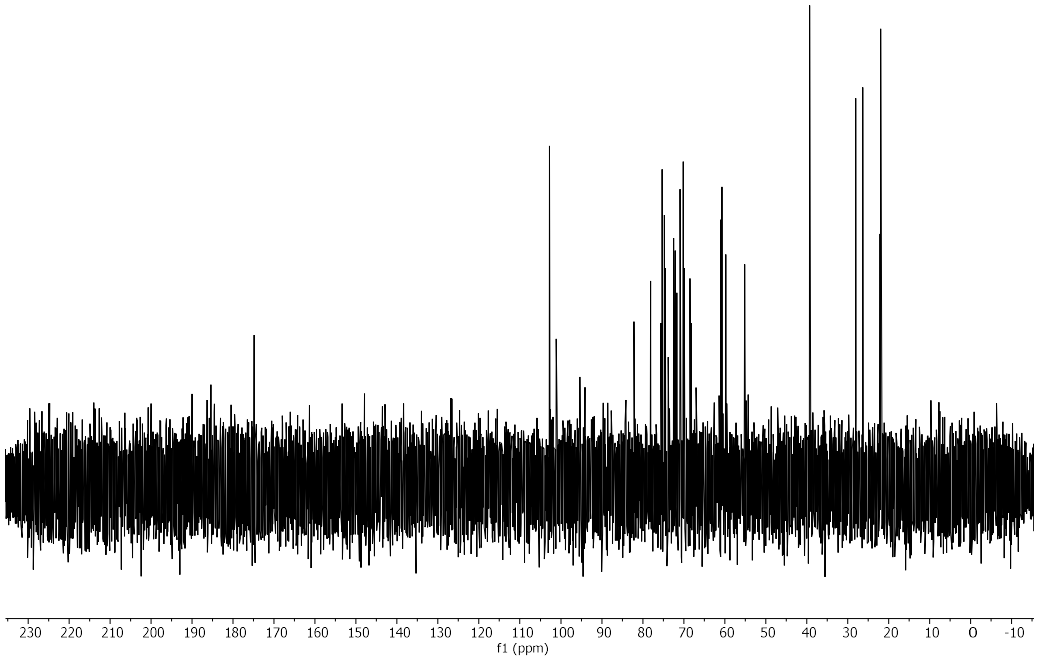
RP-HPLC of F-nLac$_4$ (ELSD trace, Method A1, t_R = 28.1 min)



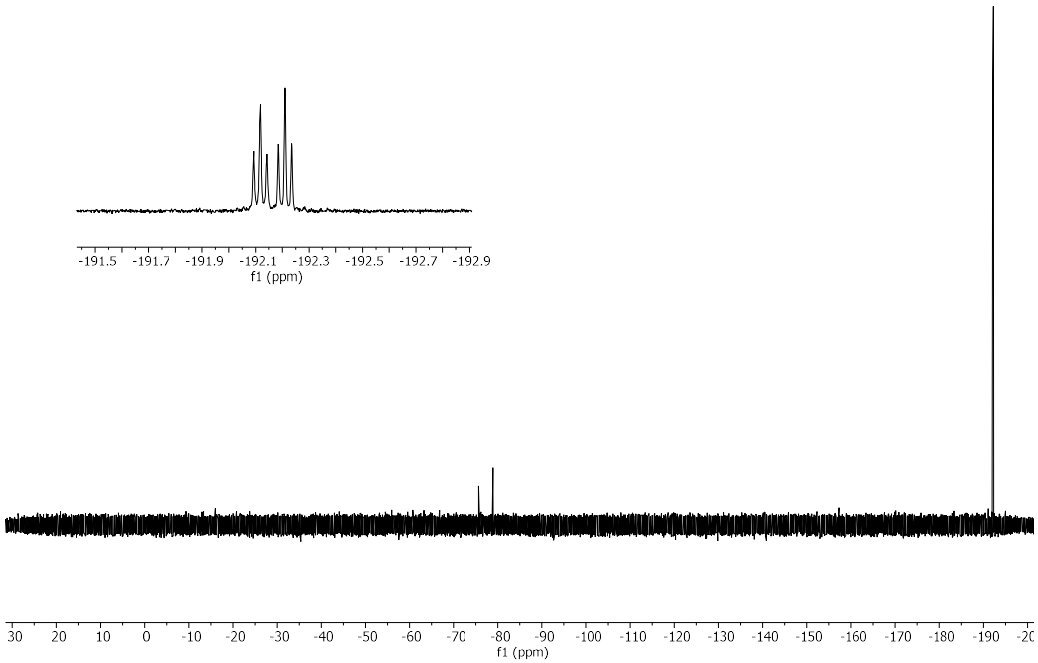
^{1}H NMR of F-nLac$_4$ (700 MHz, D$_2$O)



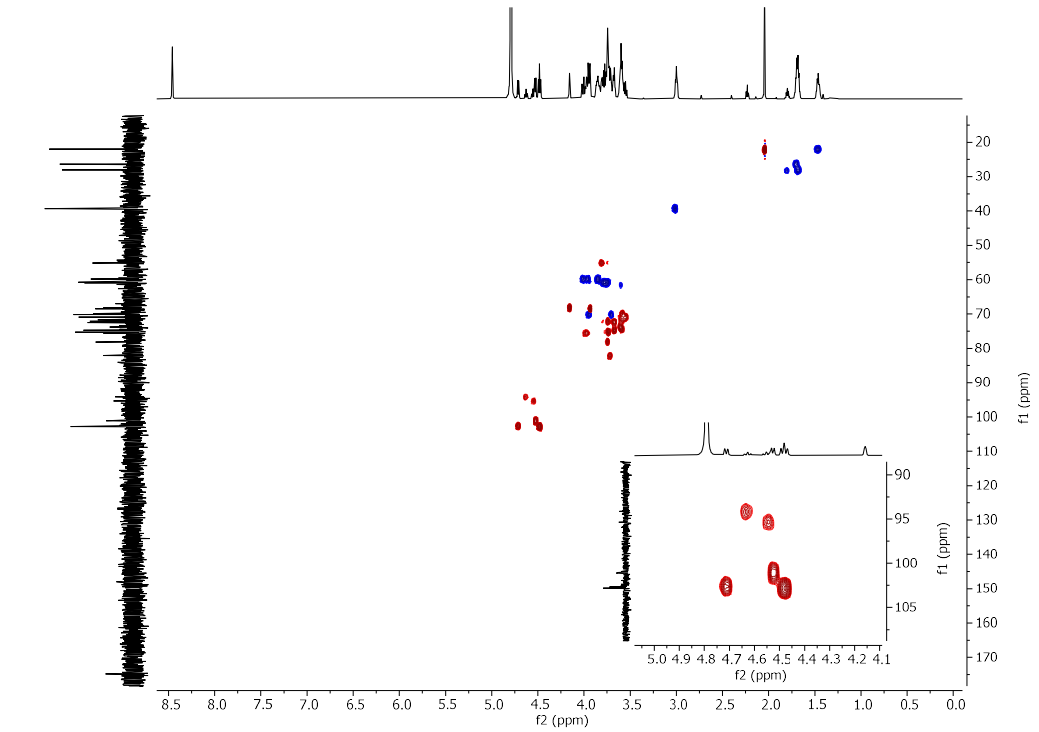
^{13}C NMR of F-nLac$_4$ (151 MHz, D_2O)



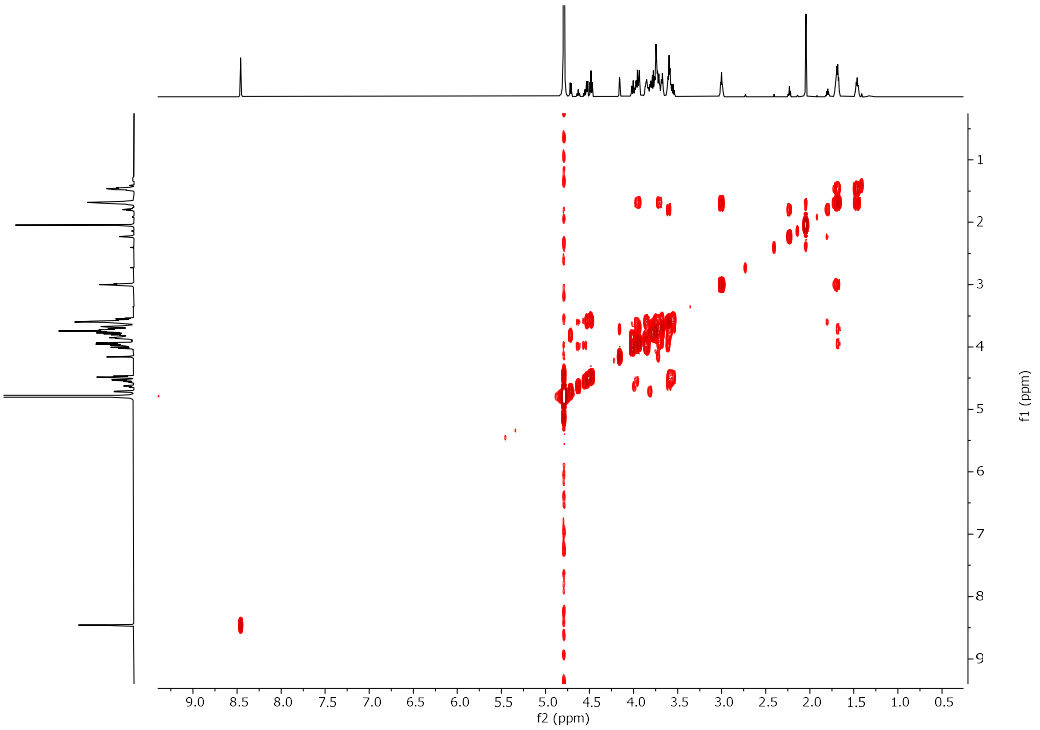
^{19}F NMR of F-nLac$_4$ (376 MHz, D_2O)



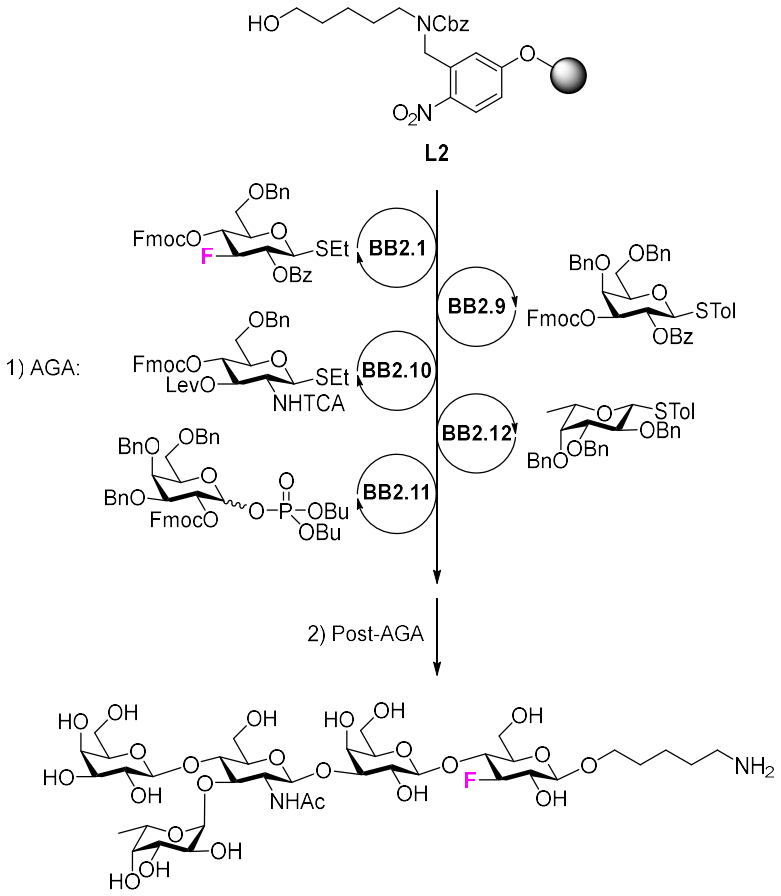
HSQC NMR of F-nLac$_4$ (D_2O)



COSY NMR of F-nLac$_4$ (D_2O)



Synthesis of **F-Le^x**



F-Le^x

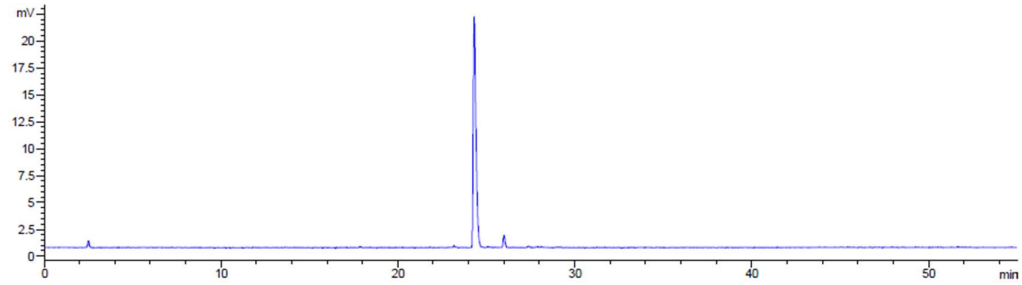
Step	BB	Modules	Notes
AGA		**A**	**L2** swelling
	BB2.1	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.1**, -20°C for 5 min, 0°C for 20 min)
	BB2.9	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.9**, -20°C for 5 min, 0°C for 20 min)
	BB2.10	**B**, **C1**, **D**, **E2**	**C1**: (**BB2.10**, -20°C for 5 min, 0°C for 40 min)
	BB2.12	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.12**, -40°C for 5 min, -20°C for 20 min)
	BB2.11	**B**, **C2**, **D**, **E1**	**C2**: (**BB2.11**, -35°C for 5 min, -15°C for 30 min)
Post-AGA		**F1**, **G1**, **H1**, **I**	**F1:** (3 d)
			H1: (3 d)
			I: (Method D)

Automated synthesis, global deprotection, and purification afforded **F-Le^x** as a white solid (1.1 mg, 8% overall yield).

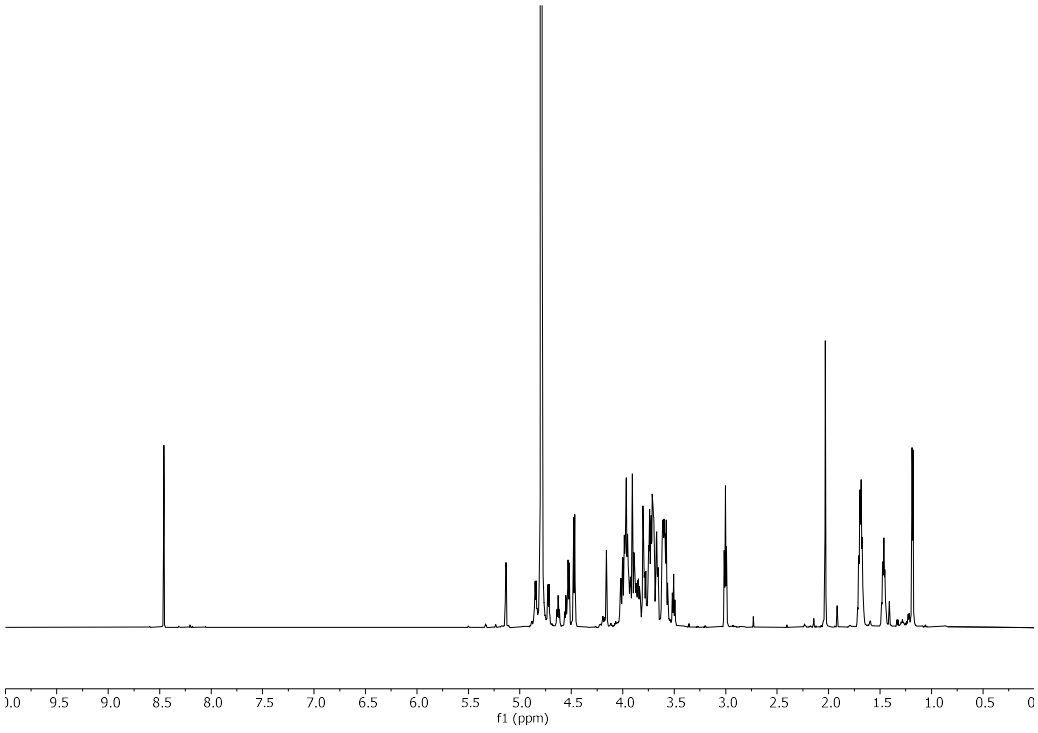
^{1}H NMR (700 MHz, Deuterium Oxide) δ 5.14 (s, 1H, **H-1 Fuc α-1,3**), 4.89 – 4.83 (m, 1H, H-5 Fuc α-1,3), 4.72 (d, *J* = 8.4 Hz, 1H, **H-1**), 4.67 – 4.51 (m, 2H, H-3 Glc, **H-1**), 4.50 – 4.44 (m, 2H, 2x **H-1**), 4.21 – 4.13 (m, 1H), 4.06 – 3.82 (m, 11H), 3.82 – 3.64 (m, 10H), 3.59 (ddt, *J* = 17.3, 10.2, 3.8 Hz, 5H), 3.51 (t, *J* = 8.9

Hz, 1H), 3.00 (t, *J* = 7.6 Hz, 2H, C*H*$_2$-NH_3^+ linker), 2.03 (s, 3H, C*H*$_3$ NHAc), 1.68 (p, *J* = 7.8, 7.1 Hz, 4H, 2x C*H*$_2$ linker), 1.46 (p, *J* = 7.5 Hz, 2H, C*H*$_2$ linker), 1.18 (d, *J* = 6.7 Hz, 3H, C*H*$_3$ Fuc). ^{13}C NMR (151 MHz, Deuterium oxide) δ 170.97 (s, *C*=O NHAc), 102.82 (s, **C-1**), 102.46 (s, **C-1**), 101.71 (s, **C-1**), 101.13 (s, **C-1**), 98.54 (s, **C-1 Fuc α-1,3**), 94.73 (d, *J* = 178.3 Hz, C-3 Glc), 82.16, 75.07, 74.85, 74.71, 73.81, 73.01, 72.42, 71.85, 70.99, 70.15 (s, *C*H_2-O linker), 69.92, 69.14, 68.29, 68.16, 67.65, 66.63 (s, C-5 Fuc α-1,3), 61.44, 60.73, 59.58, 39.29 (s, *C*H_2-NH_3^+ linker), 28.06 (s, *C*H_2 linker), 26.35 (s, *C*H_2 linker), 22.19 (s, *C*H_3 NHAc), 21.99 (s, *C*H_2 linker), 15.24 (s, *C*H_3 Fuc). ^{19}F NMR (564 MHz, Deuterium oxide) δ -192.15 (dt, *J* = 52.2, 14.0 Hz). (ESI-HRMS) m/z 941.403 $[M+H]^+$ (m/z calcd for $C_{37}H_{66}FN_2O_{24}$: 941.398).

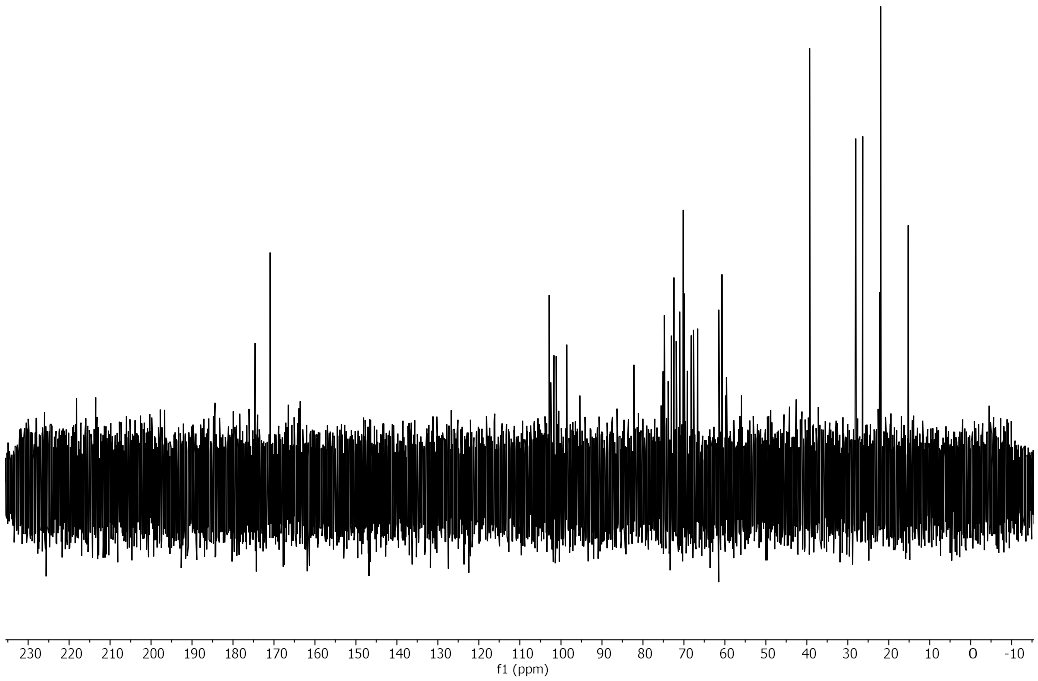
RP-HPLC of F-Le^x (ELSD trace, Method A1, t_R = 24.3 min)



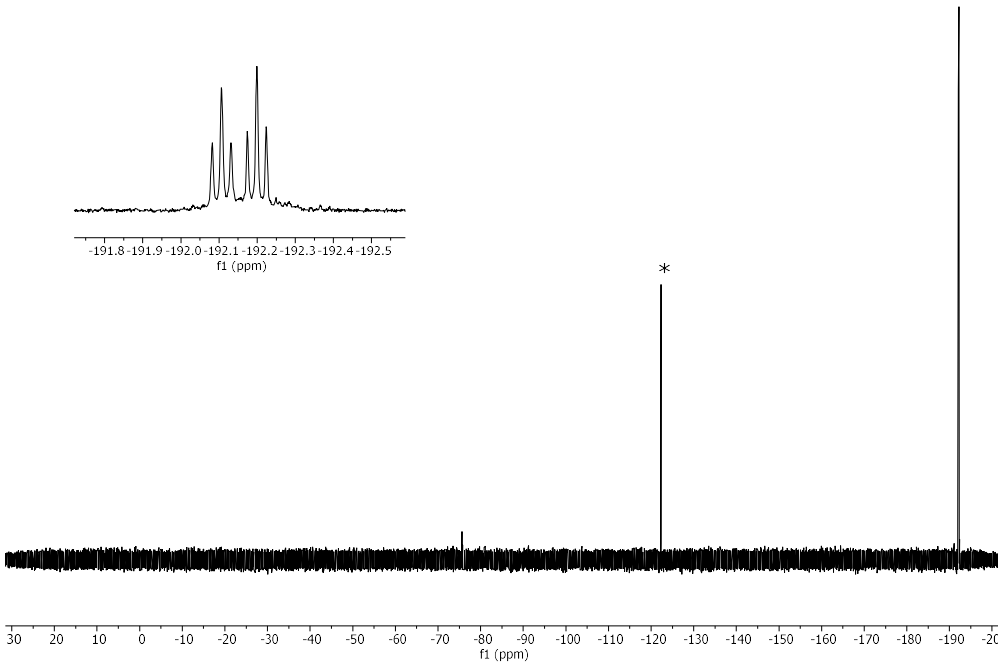
^{1}H NMR of F-Le^x (700 MHz, D_2O)



^{13}C NMR of F-Le^{x} (151 MHz, D_2O)

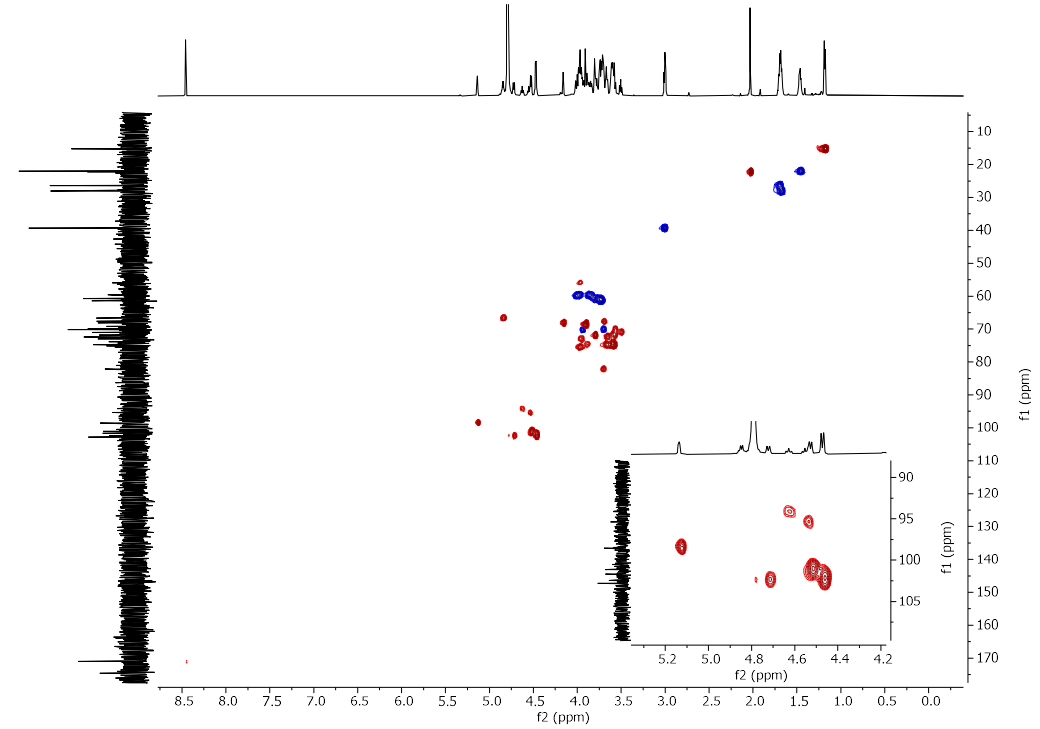


^{19}F NMR of F-Le^{x} (376 MHz, D_2O)

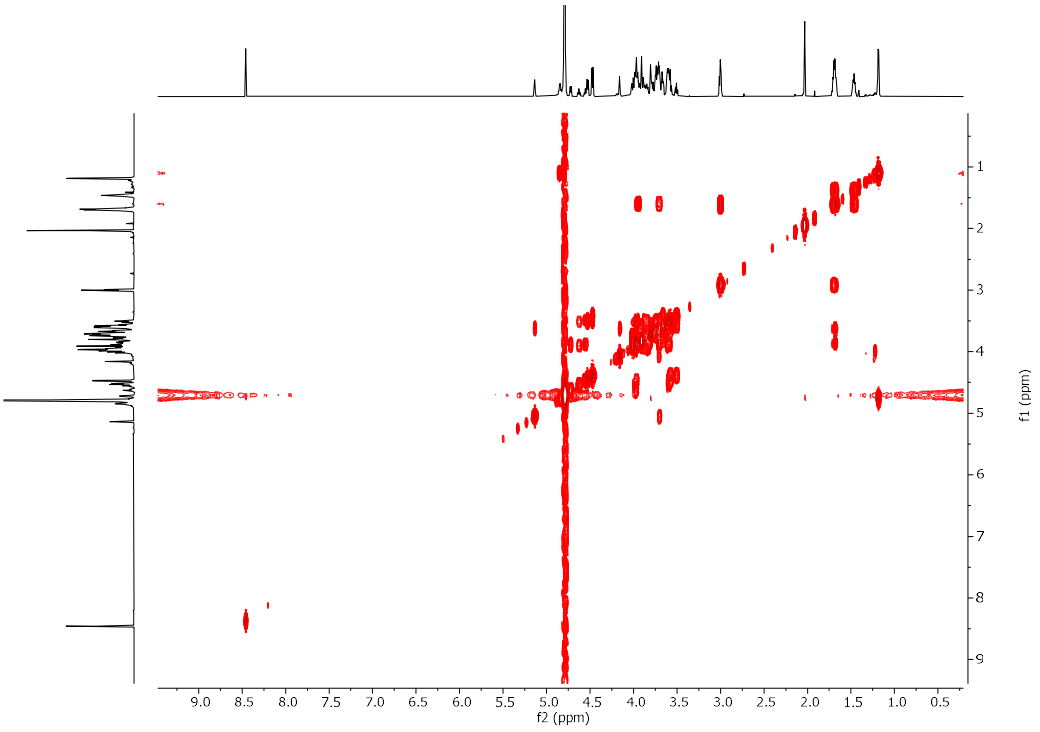


* The peak at -122.29 ppm is a fluoride impurity

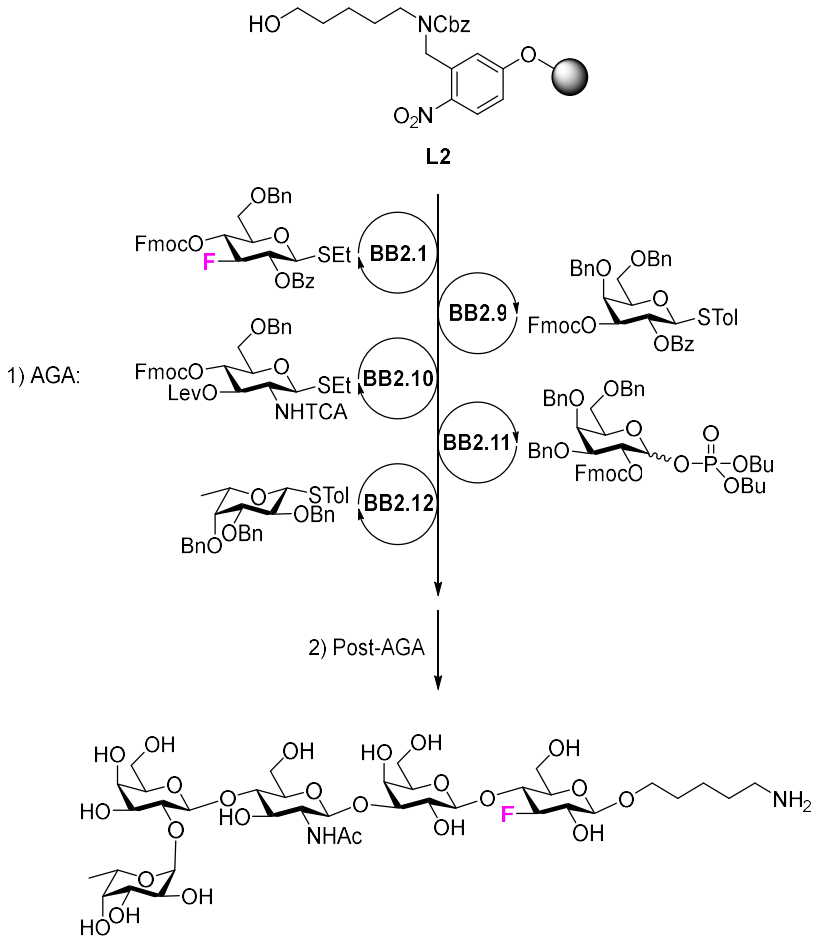
HSQC NMR of F-Le^x (D_2O)



COSY NMR of F-Le^x (D_2O)



Synthesis of **F-H type 2**



F-H type 2

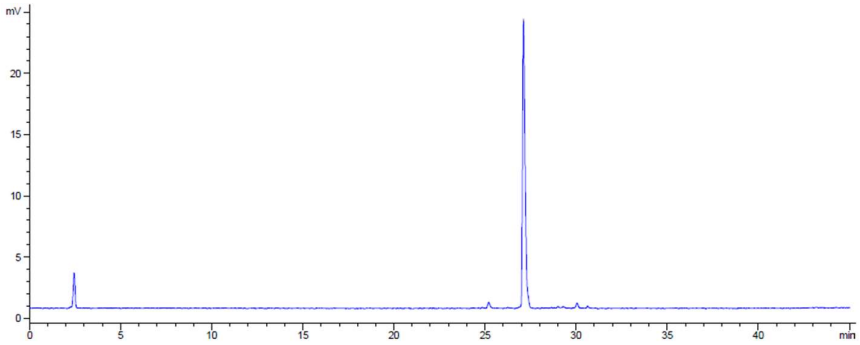
Step	BB	Modules	Notes
AGA	-	**A**	**L2** swelling
	BB2.1	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.1**, -20°C for 5 min, 0°C for 20 min)
	BB2.9	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.9**, -20°C for 5 min, 0°C for 20 min)
	BB2.10	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.10**, -20°C for 5 min, 0°C for 40 min)
	BB2.11	**B**, **C2**, **D**, **E1**	**C2**: (**BB2.11**, -35°C for 5 min, -15°C for 30 min)
	BB2.12	**B**, **C1**, **D**, **E2**	**C1**: (**BB2.12**, -40°C for 5 min, -20°C for 20 min)
Post-AGA	-	**F1**, **G1**, **H1**, **I**	**F1:** (2 d) **H1**: (3 d) **I**: (Method D)

Automated synthesis, global deprotection, and purification afforded **F-H type 2** as a white solid (1.3 mg, 10% overall yield).

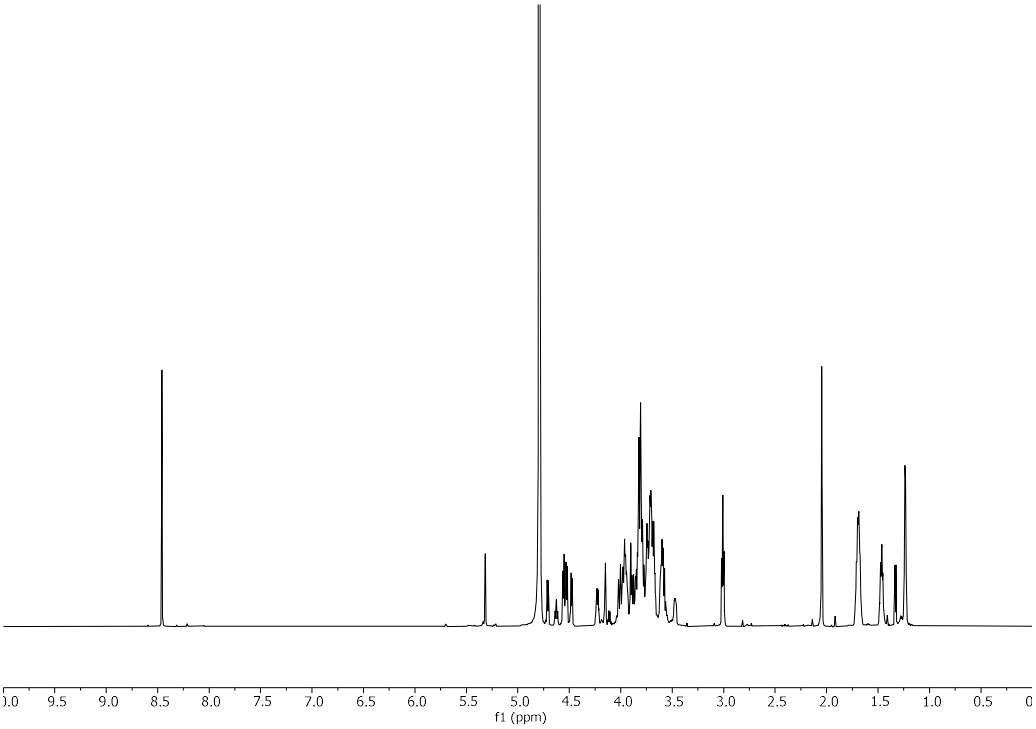
^{1}H NMR (700 MHz, Deuterium Oxide) δ 5.32 (s, 1H, **H-1 Fuc α-1,2**), 4.71 (d, *J* = 8.5 Hz, 1H, **H-1**), 4.68 – 4.44 (m, 4H, 4x **H-1**, H-3 Glc), 4.23 (q, *J* = 6.5 Hz, 1H, H-5 Fuc α-1,2), 4.20 – 3.64 (m, 25H), 3.64 – 3.50

(m, 2H), 3.47 (m, 1H), 3.01 (t, *J* = 7.6 Hz, 2H, C*H*$_2$-NH$_3^+$ linker), 2.05 (s, 3H, C*H*$_3$ NHAc), 1.80 – 1.62 (m, 4H, 2x C*H*$_2$ linker), 1.46 (p, *J* = 8.0 Hz, 2H, C*H*$_2$ linker), 1.27 – 1.20 (d, *J* = 6.1 Hz, 3H, C*H*$_3$ Fuc). ^{13}C NMR (151 MHz, Deuterium oxide) δ 170.79 (s, *C*=O NHAc), 102.82 (s, **C-1**), 102.68 (s, **C-1**), 101.13 (s, **C-1**), 100.19 (s, **C-1**), 99.35 (s, **C-1 Fuc α-1,2**), 94.75 (d, *J* = 196.3 Hz, C-3 Glc), 82.07, 76.39, 75.85, 75.58, 75.19, 75.04, 74.70, 73.76, 73.46, 72.00, 71.61, 70.15 (s, *C*H$_2$-O linker), 69.94, 69.55, 69.05, 68.14, 66.87 (s, C-5 Fuc α-1,2), 61.05, 60.68, 59.78, 55.33, 39.28 (s, *C*H$_2$-NH$_3^+$ linker), 28.06 (s, *C*H$_2$ linker), 26.33 (s, *C*H$_2$ linker), 22.13 (s, *C*H$_3$ NHAc), 21.99 (s, *C*H$_2$ linker), 15.24 (s, *C*H$_3$ Fuc). ^{19}F NMR (564 MHz, Deuterium oxide) δ -192.17 (dt, *J* = 52.1, 14.1 Hz). (ESI-HRMS) m/z 941.404 [M+H]$^+$ (m/z calcd for $C_{37}H_{66}FN_2O_{24}$: 941.398).

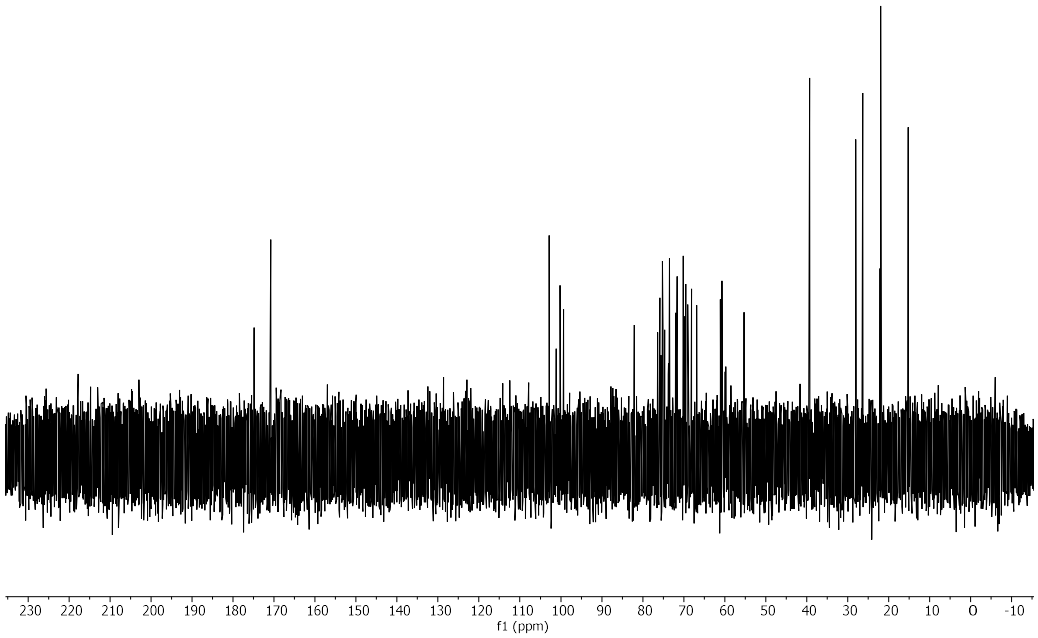
RP-HPLC of F-H type 2 (ELSD trace, Method A1, t_R = 27.4 min)



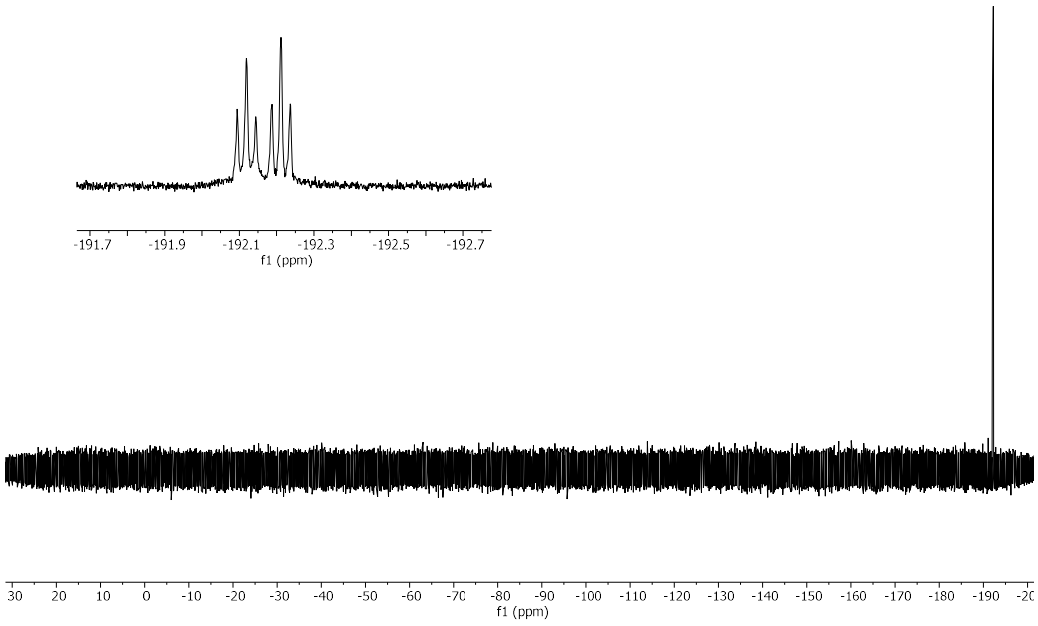
^{1}H NMR of F-H type 2 (700 MHz, D_2O)



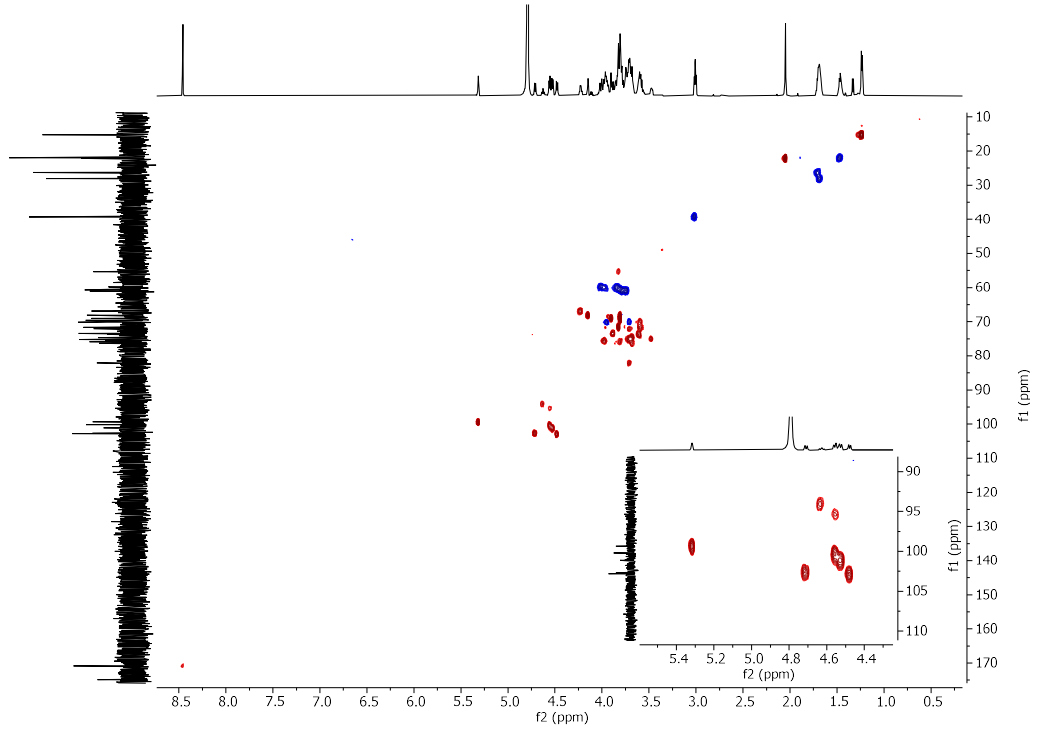
^{13}C NMR of F-H type 2 (151 MHz, D_2O)



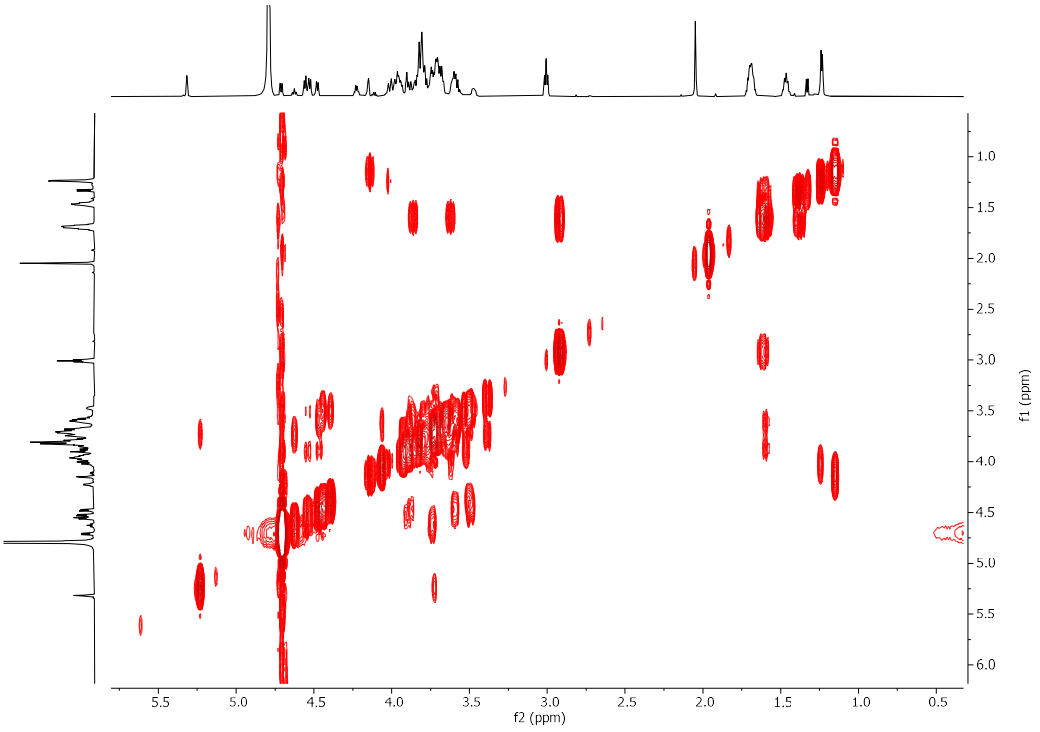
^{19}F NMR of F-H type 2 (564 MHz, D_2O)



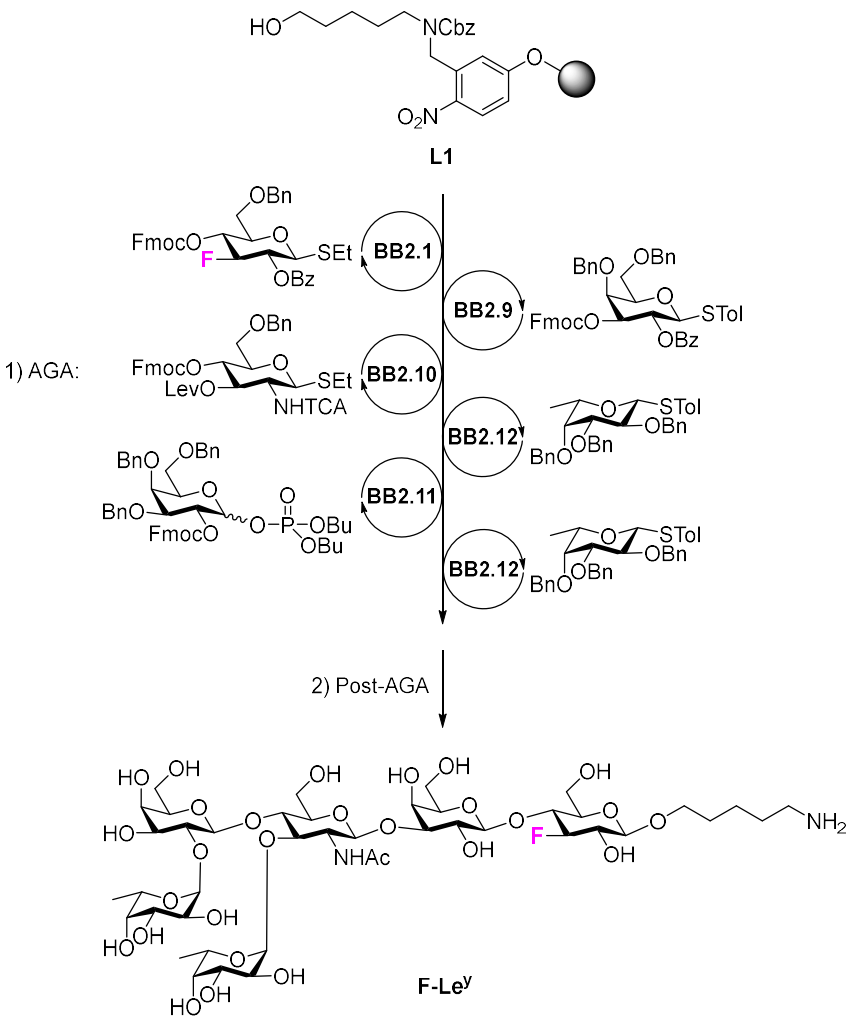
HSQC NMR of F-H type 2 (D_2O)



COSY NMR of F-H type 2 (D_2O)



Synthesis of **F-Le^y**

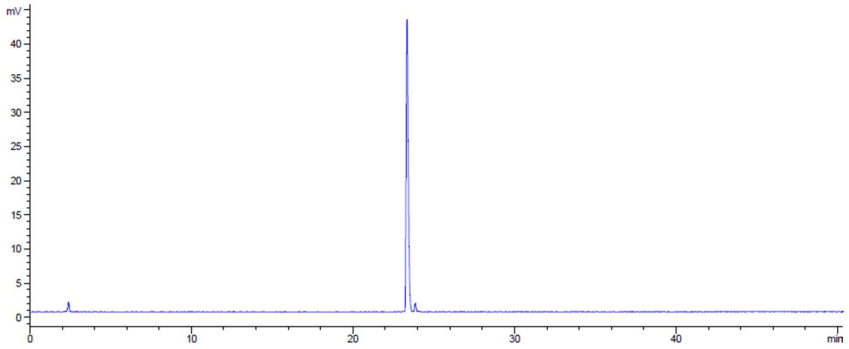


Step	BB	Modules	Notes
AGA	-	**A**	**L2** swelling
	BB2.1	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.1**, -20°C for 5 min, 0°C for 20 min)
	BB2.9	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.9**, -20°C for 5 min, 0°C for 20 min)
	BB2.10	**B**, **C1**, **D**, **E2**	**C1**: (**BB2.10**, -20°C for 5 min, 0°C for 40 min)
	BB2.12	**B**, **C1**, **D**, **E1**	**C1**: (**BB2.12**, -40°C for 5 min, -20°C for 20 min)
	BB2.11	**B**, **C2**, **D**, **E1**	**C2**: (**BB2.11**, -35°C for 5 min, -15°C for 30 min)
	BB2.12	**B**, **C1**, **D**	**C1**: (**BB2.12**, -40°C for 5 min, -20°C for 20 min)
Post-AGA	-	**F1**, **G1**, **H1**, **I**	**F1:** (3 d)
			H1: (3 d)
			I: (Method A2, t_R = 22.0 min)

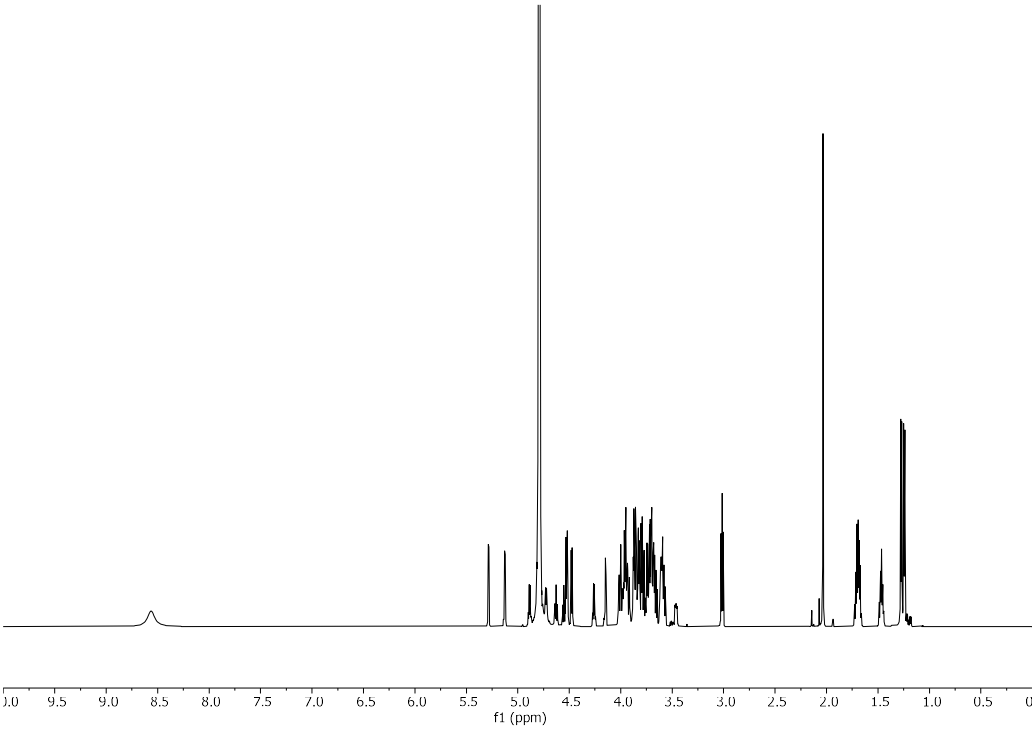
Automated synthesis, global deprotection, and purification afforded **F-Le^y** as a white solid (0.8 mg, 5% overall yield).

^{1}H NMR (700 MHz, Deuterium Oxide) δ 5.28 (d, *J* = 3.6 Hz, 1H, **H-1 Fuc α-1,2**), 5.12 (d, *J* = 4.0 Hz, 1H, **H-1 Fuc α-1,3**), 4.89 (q, *J* = 6.8 Hz, 1H, H-5 Fuc **α-1,3**), 4.73 (d, *J* = 8.4 Hz, 1H, **H-1**), 4.59 (dt, *J* = 52.3, 8.8 Hz, 1H, H-3 Glc), 4.54 – 4.50 (m, 2H, 2x **H-1**), 4.48 (d, *J* = 7.9 Hz, 1H, **H-1**), 4.26 (q, *J* = 6.6 Hz, 1H, H-5 Fuc α-1,2), 4.16 – 4.13 (m, 1H), 4.05 – 3.90 (m, 7H), 3.90 – 3.64 (m, 17H), 3.63 – 3.55 (m, 5H), 3.46 (ddd, *J* = 10.1, 5.2, 2.4 Hz, 1H), 3.05 – 2.99 (m, 2H, C*H*$_2$-NH_3^+ linker), 2.03 (s, 3H, C*H*$_3$ NHAc), 1.69 (dp, *J* = 13.8, 7.1 Hz, 4H, 2x C*H*$_2$ linker), 1.47 (qd, *J* = 9.6, 8.9, 6.5 Hz, 2H, C*H*$_2$ linker), 1.27 (d, *J* = 6.6 Hz, 3H, C*H*$_3$ Fuc), 1.24 (d, *J* = 6.6 Hz, 3H, C*H*$_3$ Fuc). ^{13}C NMR (176 MHz, Deuterium oxide) δ 170.92 (s, *C*=O NHAc), 102.88 (s, **C-1**), 101.20 (s, **C-1**), 101.13 (s, **C-1**), 100.22 (s, **C-1**), 99.44 (s, **C-1 Fuc α-1,2**), 98.60 (s, **C-1 Fuc α-1,3**), 94.83 (d, *J* = 182.8 Hz, C-3 Glc), 82.16, 76.40, 75.63, 75.39, 74.87, 74.76, 73.88, 73.83, 73.56, 73.09, 71.95, 71.85, 71.72, 70.22 (s, *C*H$_2$-O linker), 70.02, 69.74, 69.19, 68.75, 68.29, 68.22, 67.72, 66.92 (s, C-5 Fuc α-1,2), 66.80 (s, C-5 Fuc α-1,3), 61.48, 60.75, 59.86, 59.80, 56.15, 39.35 (s, *C*H$_2$-NH_3^+ linker), 28.13 (s, *C*H$_2$ linker), 26.40 (s, *C*H$_2$ linker), 22.28 (s, *C*H$_3$ NHAc), 22.06 (s, *C*H$_2$ linker), 15.46 (s, *C*H$_3$ Fuc), 15,44 (s, *C*H$_3$ Fuc). ^{19}F NMR (564 MHz, Deuterium oxide) δ -192.16 (dt, *J* = 52.4, 14.1 Hz). (ESI-HRMS) m/z 1087.462 $[M+H]^+$ (m/z calcd for $C_{43}H_{76}FN_2O_{28}$: 1087.456).

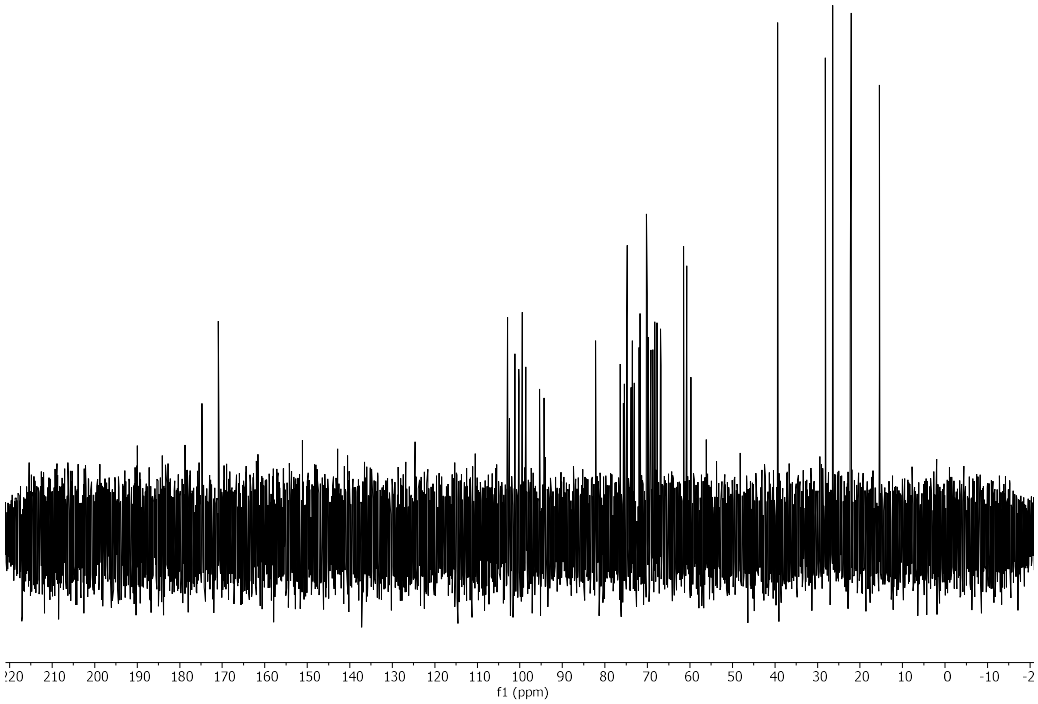
RP-HPLC of F-Le^y (ELSD trace, Method A1, t_R = 23.0 min)



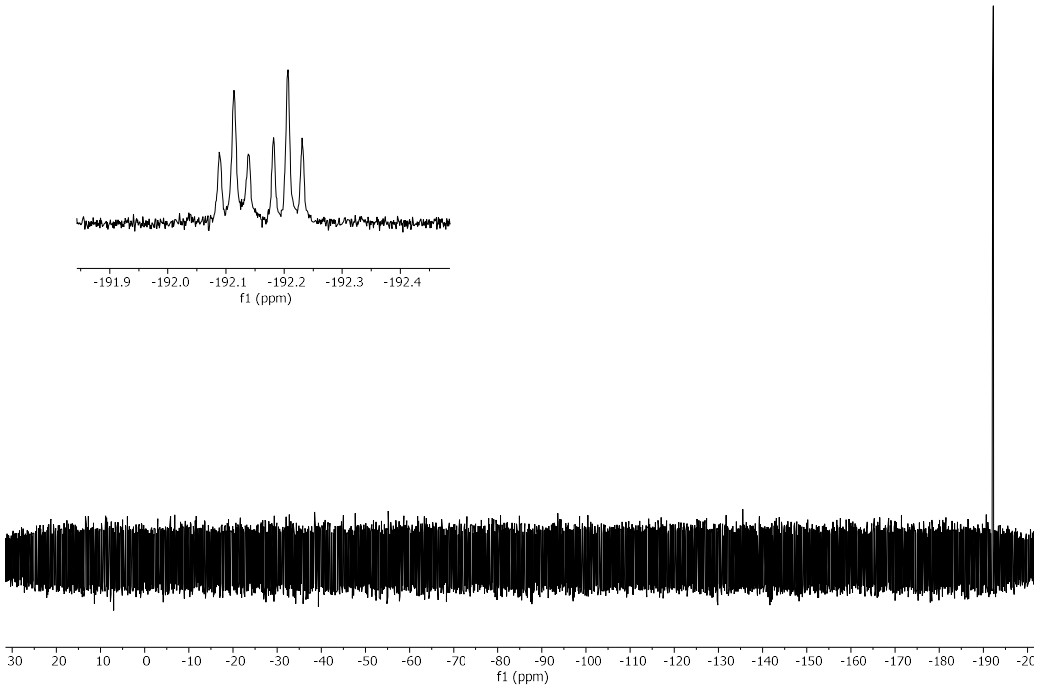
^{1}H NMR of F-Le^y (700 MHz, D_2O)



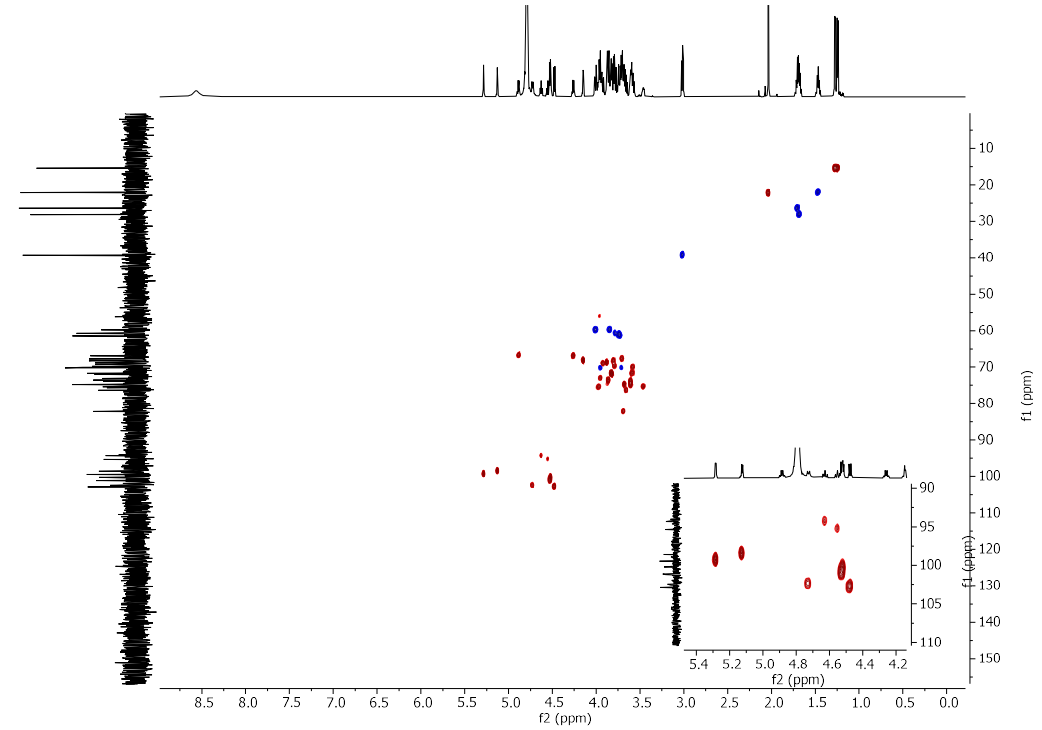
^{13}C NMR of F-Le^{y} (151 MHz, D_2O)



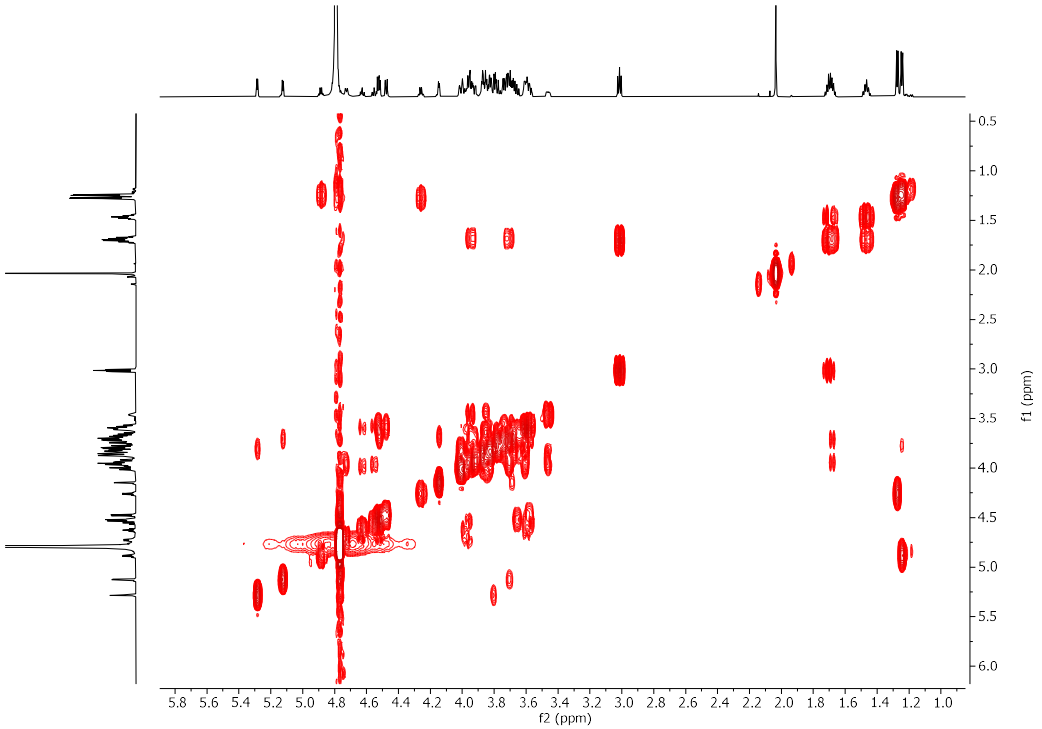
^{19}F NMR of F-Le^{y} (564 MHz, D_2O)



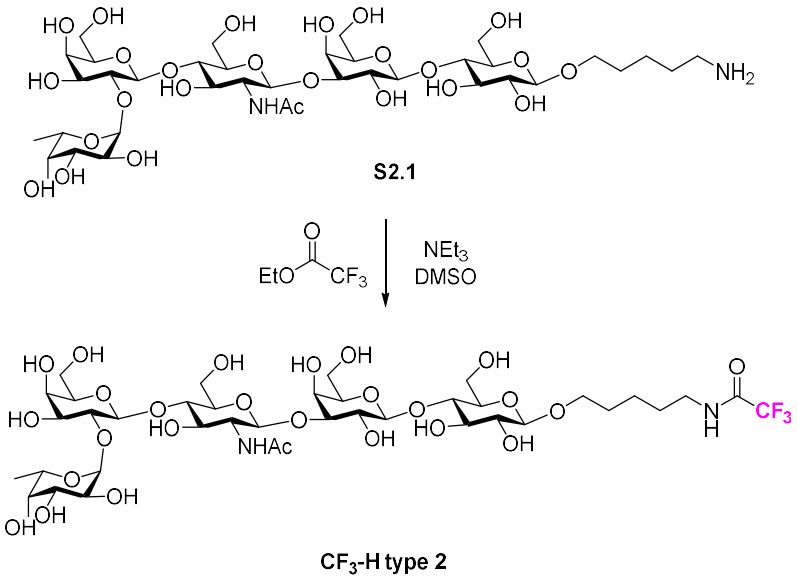
HSQC NMR of F-Le^y (D_2O)



COSY NMR of F-Le^y (D_2O)



Synthesis of **CF_3-H type 2**

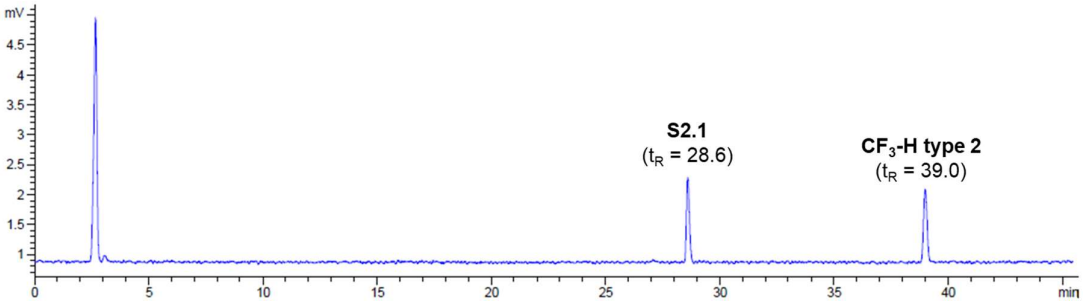


S2.1 was prepared by automated synthesis according to previously established protocols.[233]

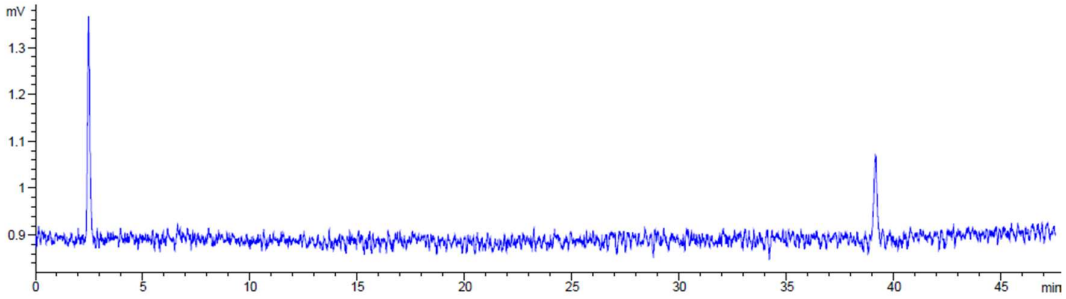
S2.1 (250 µg, 0.27 µmol) was dissolved in DMSO (150 µL) and a triethylamine solution (10 µL, 0.1 M in DMSO, 1 µmol) was added. A solution of ethyltrifluoroacetate (20 µL, 0.1 M in DMSO, 2 µmol) was added and the mixture stirred at RT overnight. The reaction was then diluted with water, lyophilized, and purified by RP HPLC (*Module I*, Method A2, t_R = 38.6 min). **CF_3-H type 2** was obtained as a white solid (140 µg, 51% yield).

^{1}H NMR (600 MHz, Deuterium Oxide) δ 5.32 (d, *J* = 3.1 Hz, 1H, **H-1 Fuc α-1,2**), 4.72 (d, *J* = 8.4 Hz, 1H, **H-1**), 4.56 (d, *J* = 7.7 Hz, 1H, **H-1**), 4.49 (d, *J* = 8.0 Hz, 1H, **H-1**), 4.46 (d, *J* = 7.9 Hz, 1H, **H-1**), 4.23 (q, *J* = 6.6 Hz, 1H, H-5 Fuc α-1,2), 4.16 (d, *J* = 3.4 Hz, 1H), 4.02 – 3.87 (m, 5H), 3.87 – 3.57 (m, 22H), 3.51 – 3.45 (m, 1H), 3.35 (t, *J* = 7.0 Hz, 2H, C*H*$_2$-NH_3^+ linker), 3.33 – 3.29 (m, 1H), 2.06 (s, 3H, C*H*$_3$ NHAc), 1.70 – 1.59 (m, 4H, 2x C*H*$_2$ linker), 1.42 (p, *J* = 7.7 Hz, 2H, C*H*$_2$ linker), 1.24 (d, *J* = 6.6 Hz, 3H, C*H*$_3$ Fuc). ^{19}F NMR (564 MHz, Deuterium oxide) δ -75.92. (ESI-HRMS) m/z 1057.372 $[M+Na]^+$ (m/z calcd for $C_{39}H_{65}F_3N_2O_{26}Na$: 1057.367).

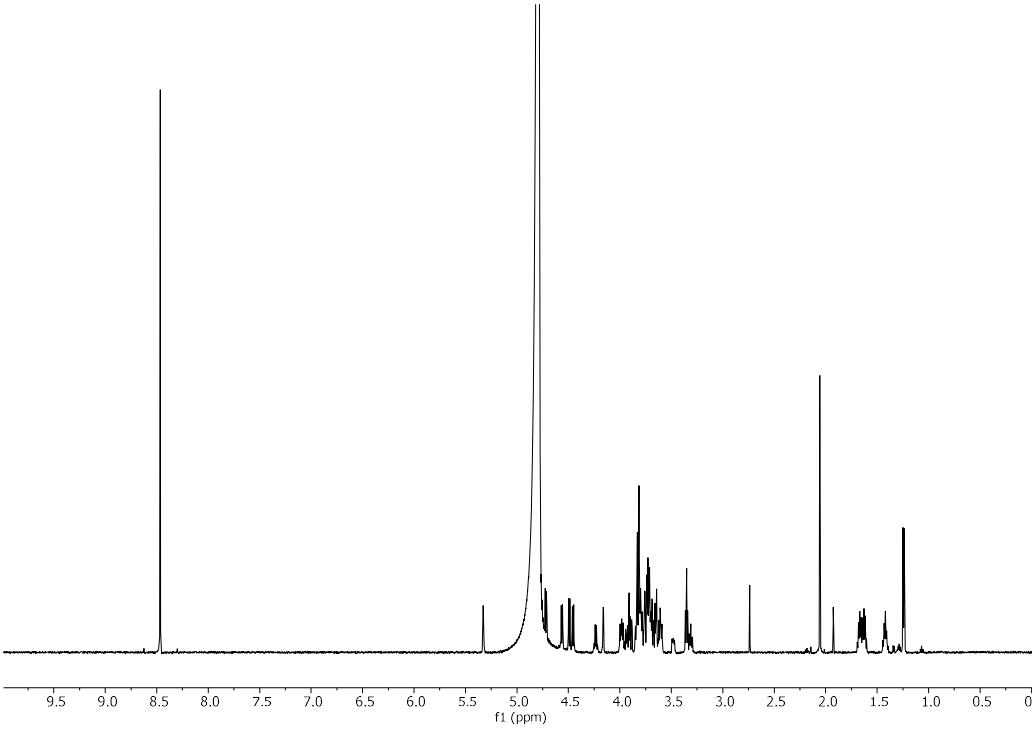
RP-HPLC of the crude reaction mixture (ELSD trace, Method A1)



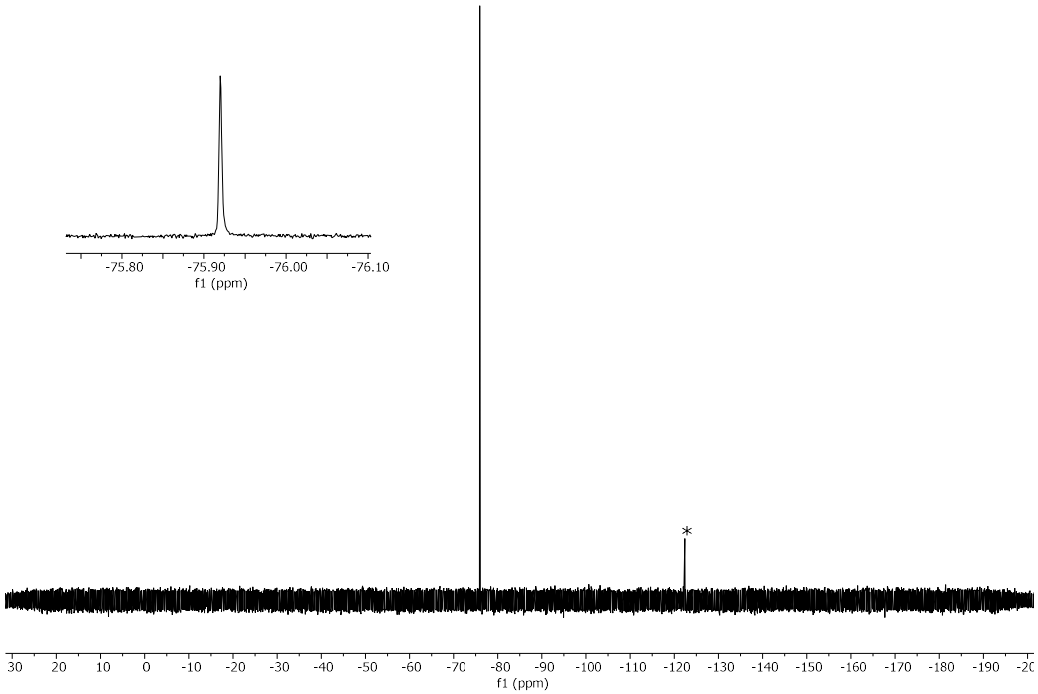
RP-HPLC of pure CF_3-H type 2 (ELSD trace, Method A1, t_R = 39.1 min)



^{1}H NMR of CF_3-H type 2 (600 MHz, D_2O)



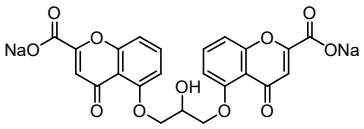
^{19}F NMR of CF_3-H type 2 (564 MHz, D_2O)



* The peak at -122.29 ppm is a fluoride impurity

6.3.3 RDC analysis

Compound	**2.17** [$^{13}C_6$]-G1-G3	**2.14** [$^{13}C_6$]-G2	**2.17** [$^{13}C_6$]-G1-G3	**2.15** [$^{13}C_6$]-G4	**2.20** 3F-G1-G3	**2.20** 3F-G1-G3
Residue	**G1**	**G2**	**G3**	**G4**	**3F-G1**	**3F-G3**
C1-H1	23	-19	-25	-4.2	18	-30
C2-H2	25.2	-23	-19.7	-1.2		
C3-H3		-21.9		-5	23	-24
C3-F	-	-	-	-	-13	13
C4-H4		-23.9		-5.3		
C5-H5		-19.5		-4.5		
C6-H6a	10.3	21.3	-0.1	-1.3		
C6-H6b	3.6	-10.2	2.9	4.3		
^{2}H Q splitting	89	85	89	84	81	81



cromolyn sodium salt

Table S2.1 RDCs (Hz) values measured in the cromolyn medium for the different ^{13}C-labelled or ^{19}F-containing hexasaccharides. The ^{19}F-substituted and ^{13}C-labelled residues are indicated (refer to **Figure 2.8**A for residue labelling). The specific RDC values in the Table correspond to the specified residue. The deuterium residual Quadrupolar splitting (Hz) for every measurement is also shown. The estimated error in the RDC values is ca. 1Hz. The splitting in the deuterium lock signal (^{2}H Q splitting) confirmed a similar degree of alignment for each sample.

Two set of spectra were recorded: one in isotropic conditions, and the second in anisotropic media. Both isotropic and anisotropic data were obtained in the same sample, just varying the temperature: 293 K for anisotropic conditions and 308 K for isotropic conditions. The residual dipolar couplings (D) were obtained by subtracting the ^{1}J C-H splitting measured in a F2-coupled HSQC spectrum acquired in isotropic conditions ($^{1}J_{CH}$) to the same coupling obtained in anisotropic conditions ($^{1}T_{CH}$):

$$^{1}D_{CH} = {}^{1}T_{CH} - {}^{1}J_{CH}$$

6.3.4 NMR studies of glycan-protein interactions

Recombinant proteins Detailed experimental information on expression of recombinant LecA, LecB, BambL, Langerin, and DC-SIGN can be found in the original publication.[355]

^{19}F NMR screening All experiments were performed on a Bruker Ascend700 (AvanceIII HD) spectrometer equipped with a 5 mm TCI700 CryoProbe™ in 3 mm tubes (Norell S-3-800-7) at 298 K. The proteins were screened at 10 μM (or 67 mU/mL for enzymes) against F-glycans: **F-Lac**, **F-nLac$_4$**, **F-Le^x**, **F-H type 2** and **F-Le^y** in 25 mM Tris-HCl pH 7.8, 150 mM NaCl, 10% D_2O, 25 μM TFA and 5 mM $CaCl_2$ at 298 K. For this, proteins were mixed at 1:1 in volume ratio with the F-glycans resulting in the concentrations of 10 μM (or 67 mU/mL for enzymes) and 50 μM, respectively. The enzymes were screened in the absence of donor (*i.e.* CMP-Neu5Ac). ^{19}F spectra were recorded with 512 scans, a spectral width of 10 ppm, a transmitter offset at -190 ppm, acquisition time of 2 s and 1 s relaxation time. T2-filtered spectra were recorded using a CPMG pulse sequence with a 180° pulse repetition rate of 0.38 s using same acquisition and relaxation times.[356,357] ^{19}F R_2-filtered experiments for 0.6 mM **F-H type 2** and 0.15 mM LecA were performed with a CPMG pulse sequence with a 180° pulse repetition rate of 0, 0.24, 0.48, 0.72, 0.96 and 1.2 s using the same buffer and acquisition conditions as described above. ^{19}F and CPMG NMR spectra of 50 μM **CF$_3$-H type 2** were recorded in presence of 10 μM DC-SIGN ECD and 5 μM BambL with 16 scans, a spectral width of 10 ppm, a transmitter offset at -75 ppm, acquisition time of 2 s and 1 s relaxation time. Data were recorded without proton decoupling. All spectra were analyzed in MestReNova 11.0.0 (Mestrelab Research SL). The binding strength was defined depending on the changes observed in the NMR after addition of the protein. A decrease in peak intensity higher than -25% or a CSP higher than 0.01 ppm in the normal ^{19}F NMR was interpreted as strong binding, a decrease in peak intensity higher than -25% in the CPMG filtered ^{19}F NMR as weak/medium binding, and a decrease in peak lower than -25% in normal or CPMG filtered ^{19}F NMR as no binding.

Real-time ^{19}F NMR kinetic measurement The enzymes β-galactosidase (*E. coli*, CAS: 9031-11-2), and α(2,3)-sialyltransferase from *Pasteurella multocida* (Pmα23ST, *E. coli* EC no. 2.4.99.4) were purchased from Sigma Aldrich. The enzymatic reactions with β-galactosidase (0.15 μM), and Pmα23ST (100 mU/mL) were performed at 310 K in 25 mM Tris-HCl pH 7.8, 150 mM NaCl with 2 mM $MgCl_2$, 100 μM TFA and 10% D_2O. The concentration of **F-Lac** was 250 μM. The reactions with β-galactosidase were monitored every 149 s. All data were recorded without proton decoupling with 32 scans, acquisition time of 2 s and relaxation delay of 1 s. All spectra were referenced to TFA at -75.6 ppm. To monitor the sialylation of 250 μM **F-Lac** with Pmα23ST, we used the buffer conditions reported above with 275 μM CMP-Neu5Ac and recorded data at time points: 0 min, 13 min, 32 min, and 246 min. All data were recorded without proton decoupling with 128 scans, acquisition time of 2 s and relaxation delay of 1 s. All spectra were referenced to TFA at -75.6 ppm. To determine the yield of this reaction, we used the ratio of peak integrals of substrate and product at -192.1 ppm and -192.3 ppm, respectively.

2D NMR experiments Detailed experimental information on protein observed 2D NMR experiments is reported in the original publication.[355]

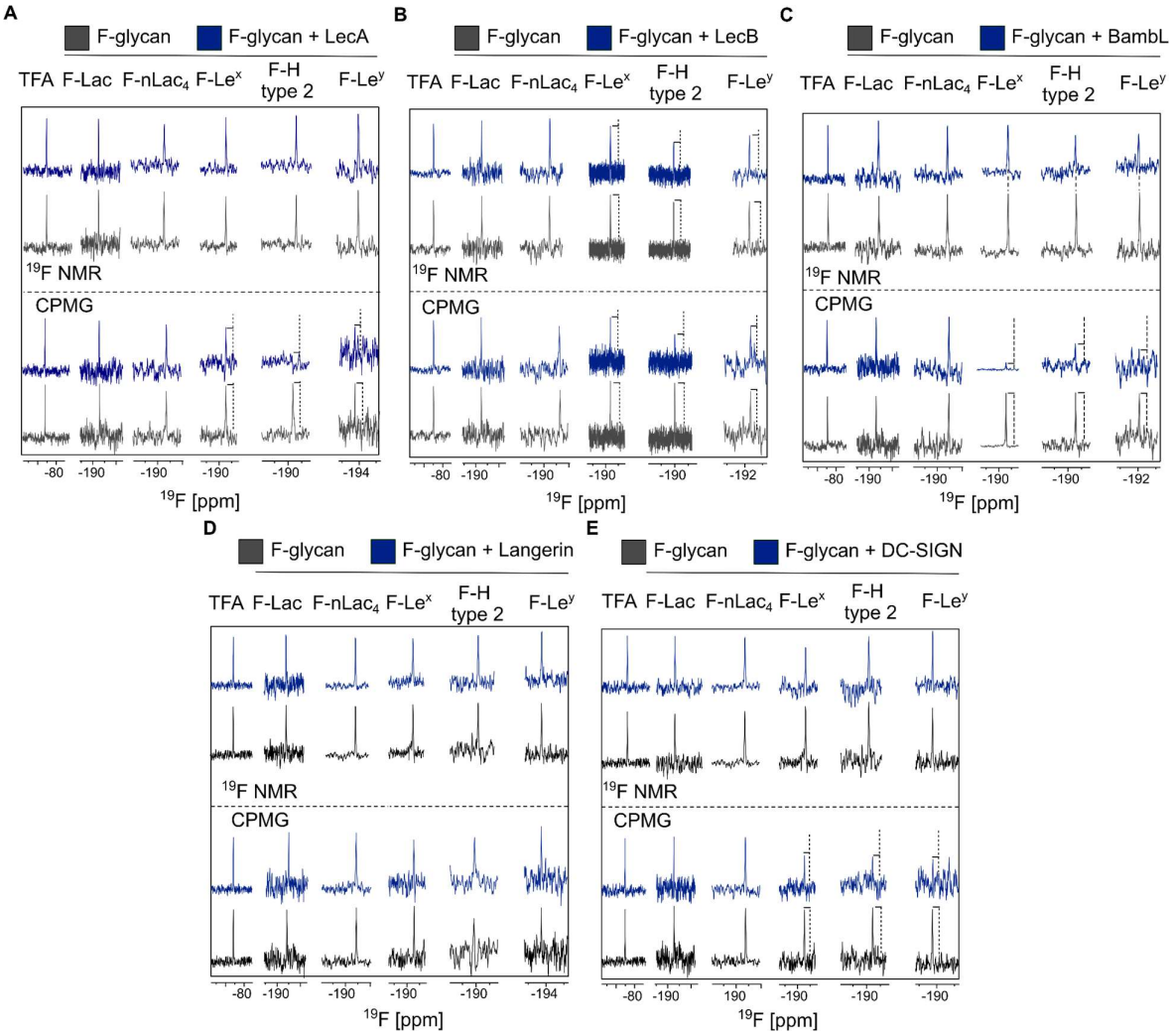


Figure S2.2 ^{19}F and CPMG NMR spectra of F-glycans alone (*gray*) and in presence of DC-SIGN ECD (*blue*). A chemical shift perturbation of fucosylated Lewis antigens in presence of BambL (**C**, *upper panel*) in ^{19}F NMR is shown (*dashed line*). Decrease in peak intensity (*dashed line*) in CPMG NMR indicates binding of fucosylated Lewis antigens to bacterial: LecA (**A**), LecB (**B**) and BambL (**C**, *lower panel*) and mammalian: Langerin ECD (**D**) DC-SIGN ECD (**E**) lectins. No binding to F-glycans has been observed in presence of Langerin ECD. Differences in signal-to-noise ratio were observed for the spectra acquired with different quality of shimming.

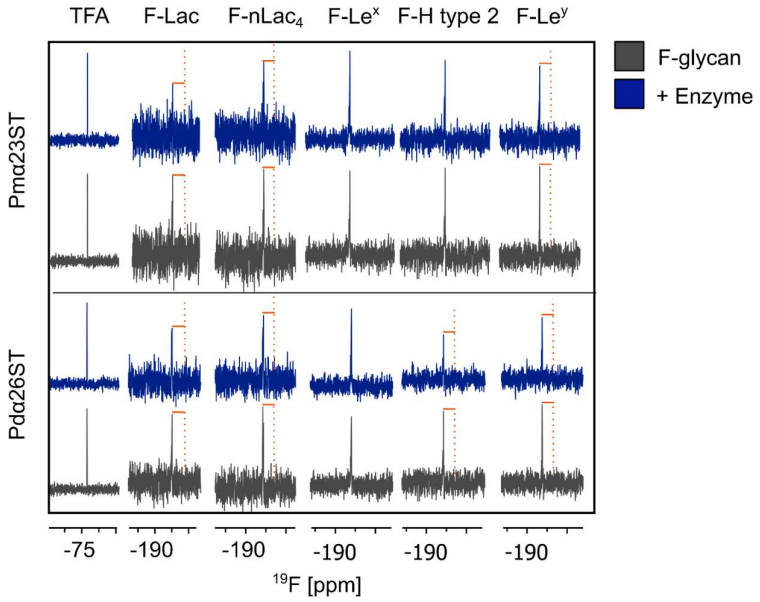


Figure S2.3 CPMG NMR of F-glycans alone and in presence of Pmα23ST or Pdα26ST in the absence of CMP-Neu5Ac donor. F-glycan binding to both enzymes is shown as a reduction in peak intensity in presence of enzymes (*orange*).

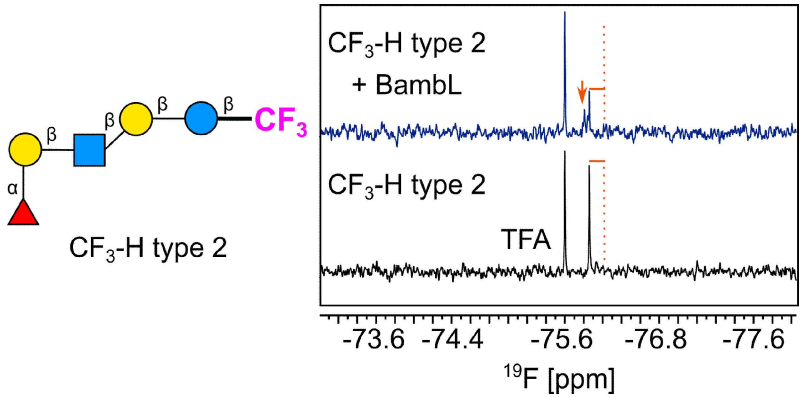


Figure S2.4 ^{19}F NMR spectrum of **CF_3-H type 2** alone and in presence of BambL (6 μM). Given that **CF_3-H type 2** undergoes a slow exchange in presence of BambL a new peak for the bound ligand arises (*arrow*). This proves that the CF_3 reporter does not have to be near the carbohydrate-binding site of a lectin and can be placed even in remote position (*i.e.* aminopentyl linker attached to the reducing end) of the glycan in order to be detected in ^{19}F NMR.

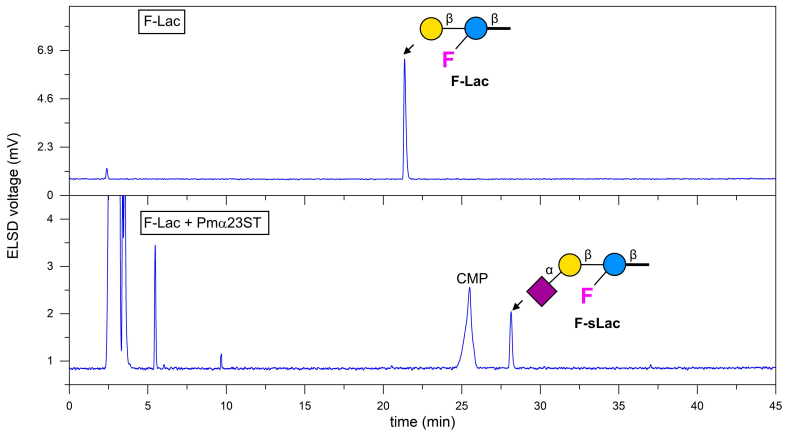


Figure S2.5 RP-HPLC trace (ELSD trace, Method A) shows consumption of the starting material (**F-Lac**, t_R = 21.4 min, *top*) and formation of a new peak in the crude reaction mixture after 246 min (**F-sLac**, t_R = 28.1 min, *bottom*).

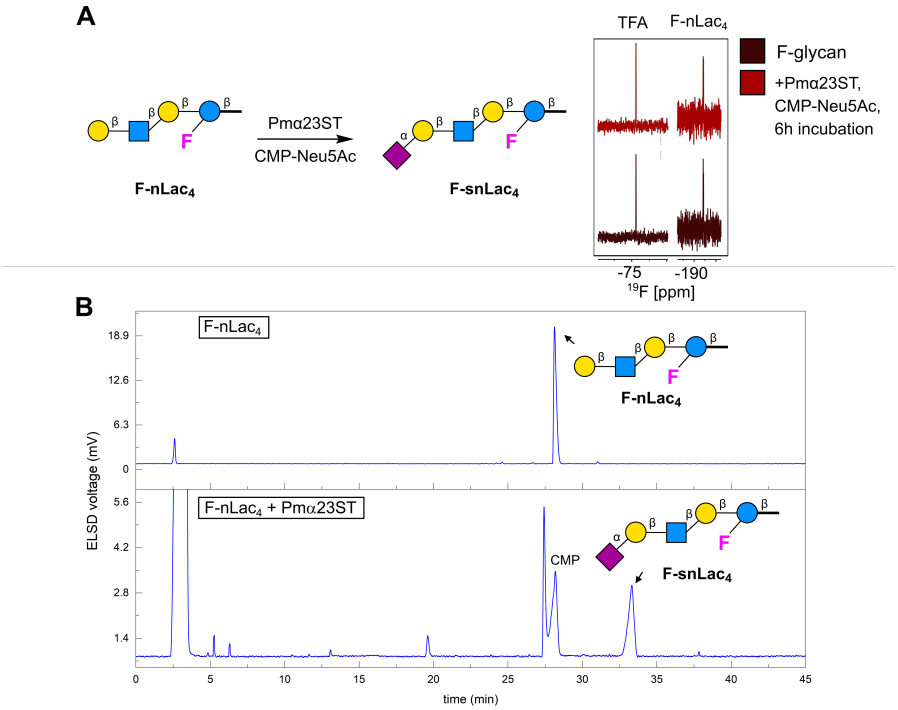
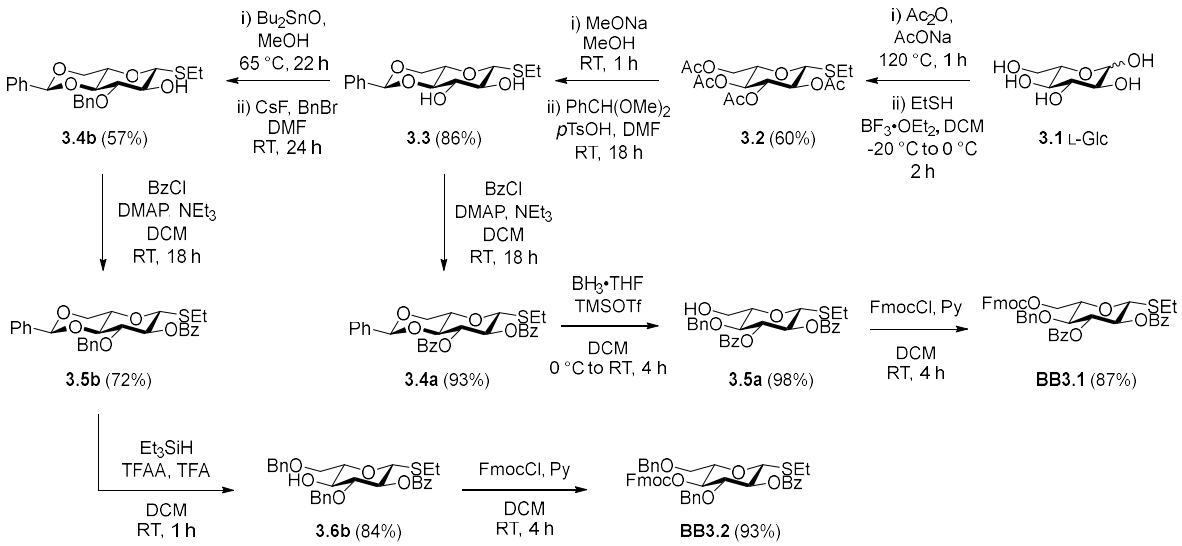


Figure S2.6 A) Real-time ^{19}F NMR monitoring of **F-nLac$_4$** incubated with Pmα23ST in the presence of CMP-Neu5Ac. No chemical shift perturbation was observed after 6 h incubation time. B) RP-HPLC trace (ELSD trace, Method A) shows incomplete consumption of the starting material (**F-nLac$_4$**, t_R = 28.1 min, *top*) and formation of the sialylated product (**F-snLac$_4$**, t_R = 33.3 min, *bottom*) in the crude reaction mixture.

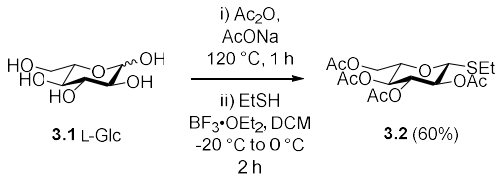
6.4 Experimental section for Chapter 3

6.4.1 Synthesis of L-Glc building blocks



Scheme S3.1 Synthesis of L-Glc **BB3.1** and **BB3.2**.

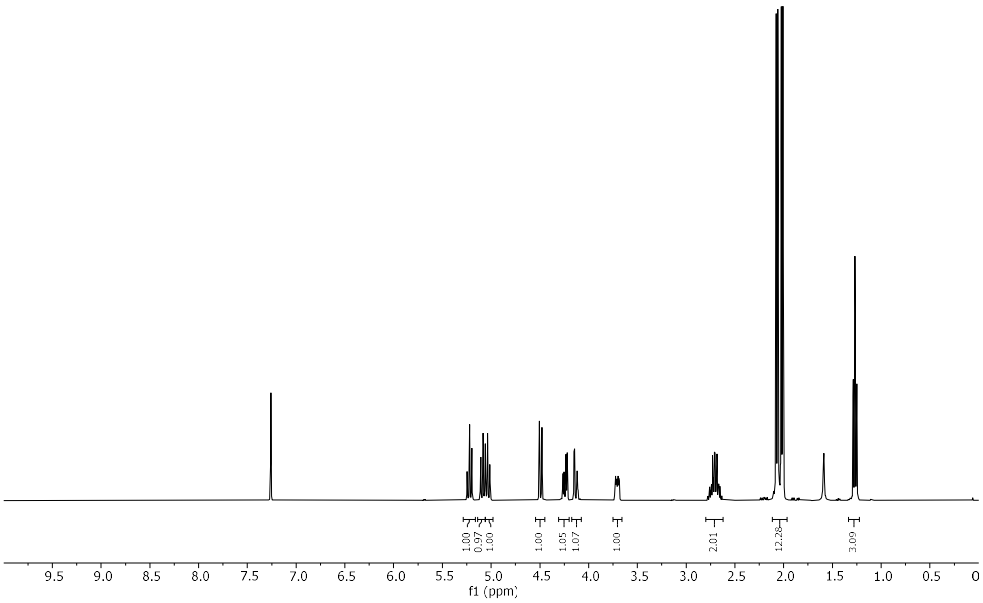
Synthesis of **3.2**



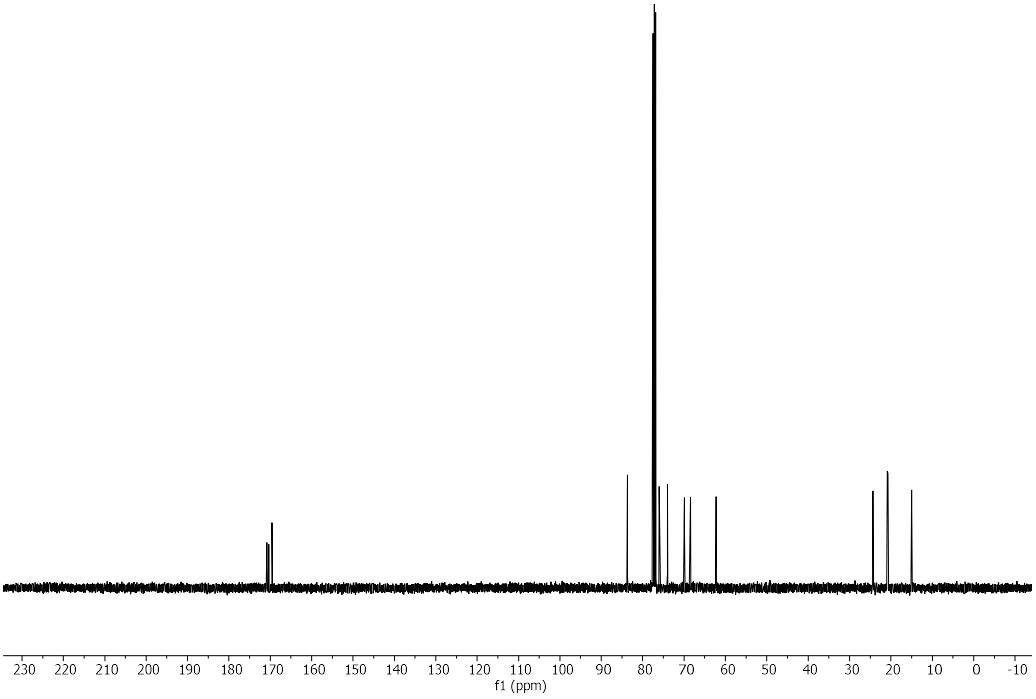
L-glucose **3.1** (4.95 g, 27.5 mmol) was slowly added (over 10 min) to a stirred solution of sodium acetate (AcONa) (1.9 g, 13.8 mmol) in acetic anhydride (Ac_2O) (45 mL) at 120°C. The mixture was stirred for 1 h and then cooled to RT. The reaction was quenched with ice (100 g) and diluted with EtOAc. The reaction mixture was washed five times with saturated aqueous solution of $NaHCO_3$ and one time with brine. The crude was passed through a short plug of silica with EtOAc, dried over Na_2SO_4, and concentrated *in vacuo*. The peracetylated intermediate was obtained as a white solid and used in the following step without further purification (11.3 g, quantitative yield, α:β ratio 1:4). $BF_3{\cdot}EtO_2$ complex (5.4 mL, 42.6 mmol) was slowly added to a stirred solution of the peracetylated intermediate (11.3 g, 29.1 mmol) with EtSH (5.4 mL, 73.1 mmol) in DCM (100 mL) at 0 °C under Ar atmosphere. The reaction was slowly brought to RT and, after 2 h, quenched with saturated aqueous solution of $NaHCO_3$ at 0 °C. The mixture was diluted with DCM and washed three times with saturated aqueous solution of $NaHCO_3$ and once with brine. The organic phase was dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by silica gel flash column chromatography (Hexane : EtOAc = 4:1→2:1) to yield **3.2** as a white solid (6.78 g, 60%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 5.22 (t, *J* = 9.4 Hz, 1H), 5.12 – 5.00 (m, 2H), 4.49 (d, *J* = 10.0 Hz, 1H), 4.24 (dd, *J* = 12.4, 4.9 Hz, 1H), 4.13 (dd, *J* = 12.4, 2.4 Hz, 1H), 3.71 (ddd, *J* = 10.0, 4.9, 2.3 Hz, 1H), 2.79 – 2.62 (m, 2H), 2.10 – 1.98 (m, 12H), 1.27 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 170.84, 170.38, 169.58, 169.58, 83.66, 82.49, 75.98, 74.02, 69.92, 68.42, 24.33, 20.90, 20.89, 20.78, 20.75, 14.96. $[\alpha]_D^{20}$ 20.24 (c 0.42 g/100 mL, $CHCl_3$). IR ν = 1752, 1370, 1223, 1038 cm^{-1}. (ESI-HRMS) m/z 415.098 $[M+Na]^+$ ($C_{16}H_{24}O_9SNa$ requires 415.103).

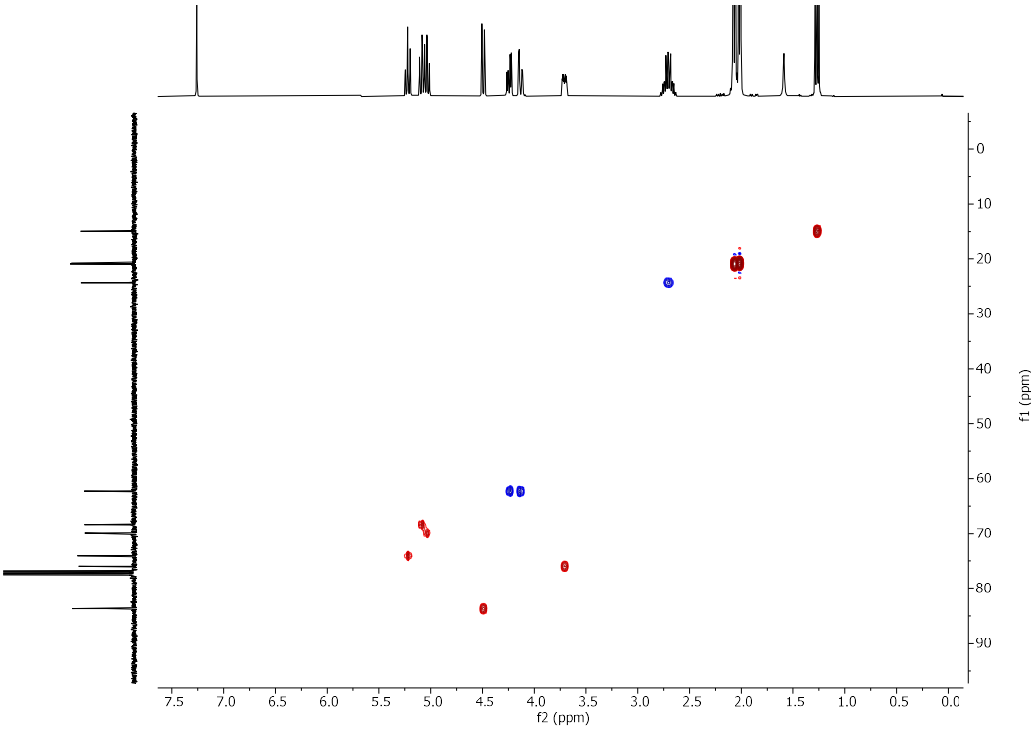
^{1}H NMR of 3.2 (400 MHz, $CDCl_3$)



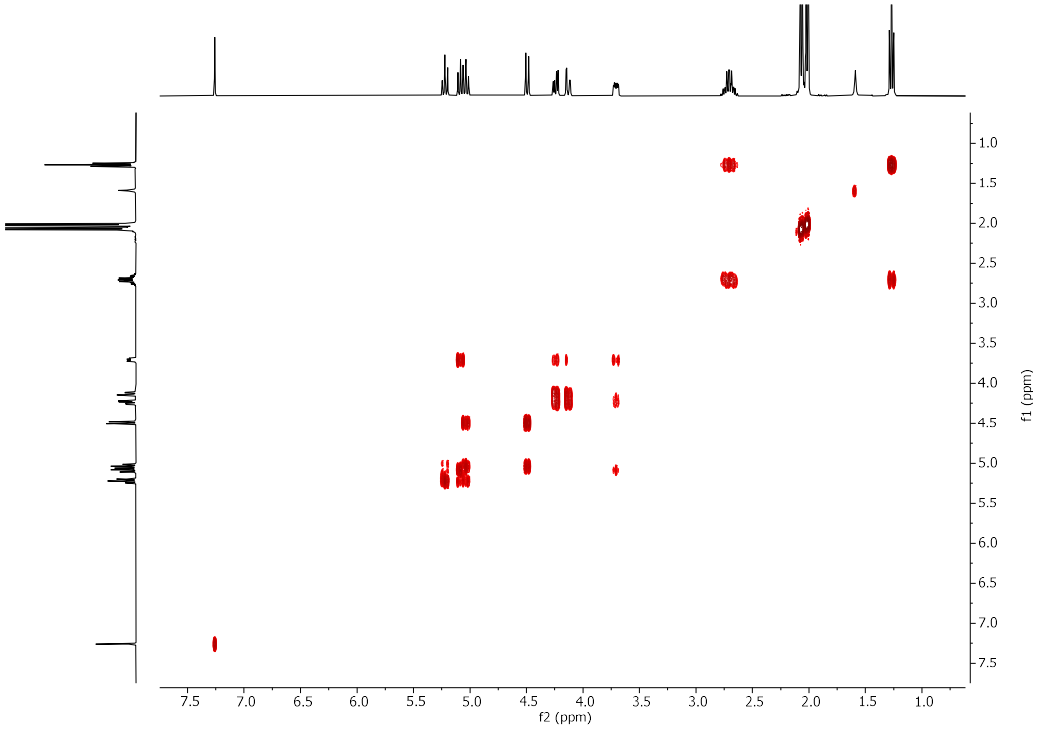
^{13}C NMR of 3.2 (101 MHz, $CDCl_3$)



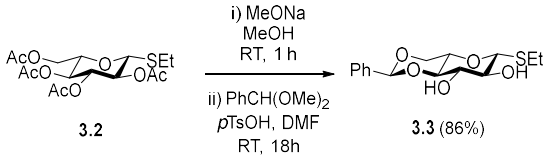
HSQC NMR of 3.2 ($CDCl_3$)



COSY NMR of 3.2 ($CDCl_3$)



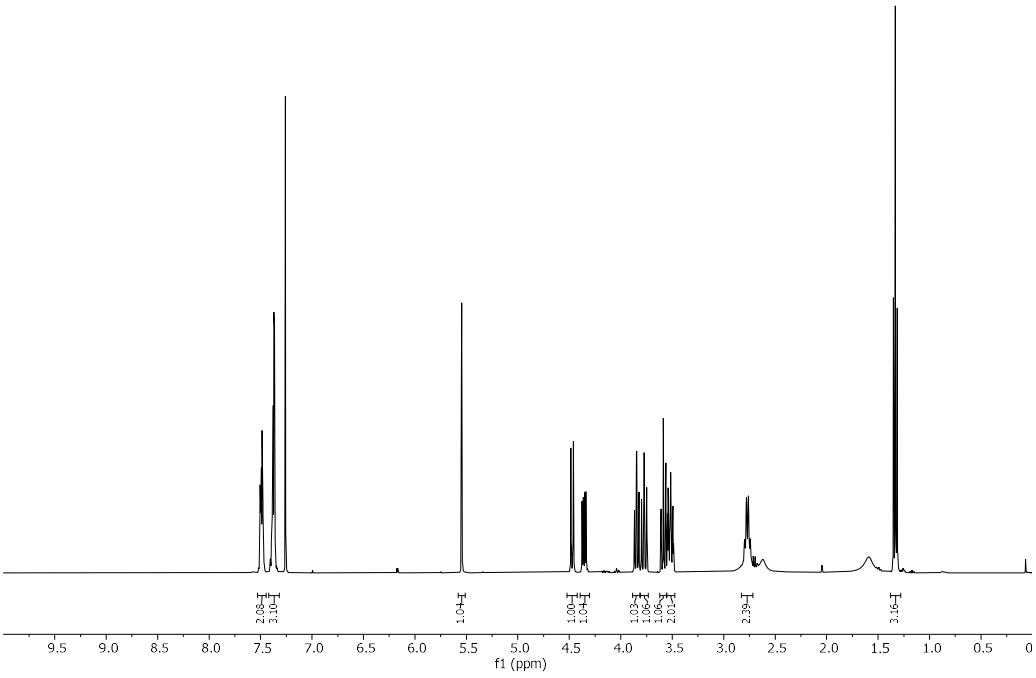
Synthesis of **3.3**



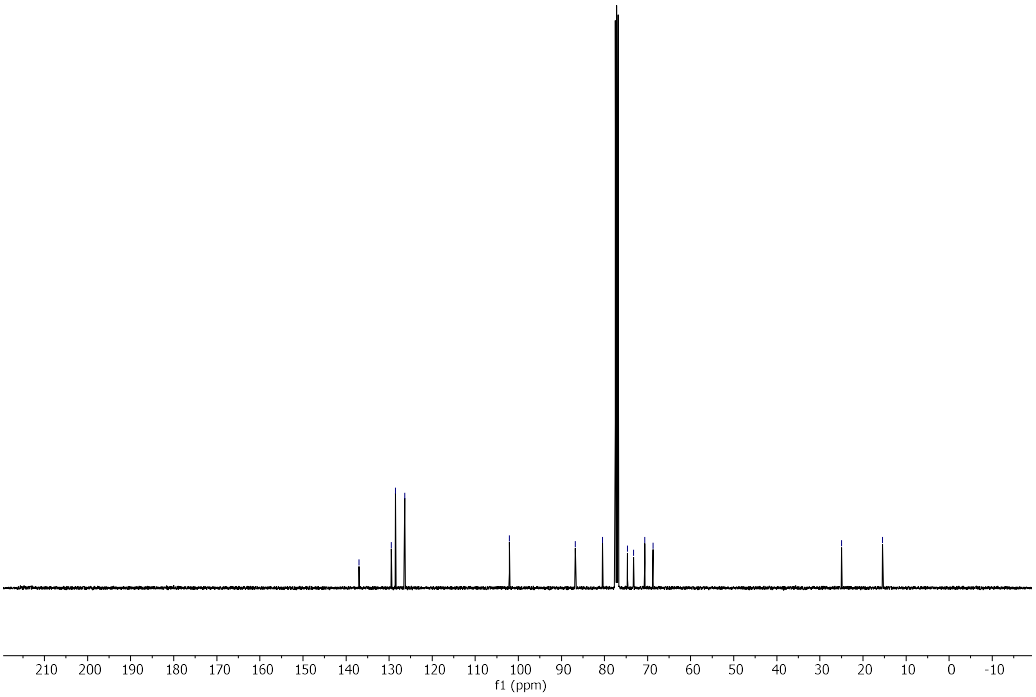
3.2 (6.78 g, 17.3 mmol) was dissolved in MeOH (60 mL) and a 0.5 M solution of MeONa in MeOH (7.0 mL, 3.5 mmol) was slowly added. The mixture was stirred at RT for 1 h. The reaction was neutralized with Amberlite IR-120 (H^+ form), filtered, and concentrated under reduced pressure. The product was used in the next step without any further purification assuming quantitative conversion. The crude product was dissolved in DMF and a catalytic amount of *p*-toluensulfonic acid monohydrate (*p*TsOH) (356 mg, 1.7 mmol) was added to the mixture. Benzaldehyde dimethyl acetal ($PhCH(OMe)_2$) (5.2 mL, 34.6 mmol) was then added dropwise at RT to the stirred solution. The reaction was heated to 45°C overnight, after which time it was quenched with triethylamine (2 mL) at 0 °C. The reaction mixture was diluted with EtOAc and washed three times with saturated aqueous solution of $NaHCO_3$ and once with brine. The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The sticky oil was recrystallized from Hexane : EtOAc to yield **3.3** as a white solid (4.65 g, 86% over 2 steps).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.53 – 7.45 (m, 2H), 7.42 – 7.33 (m, 3H), 5.55 (s, 1H), 4.47 (d, *J* = 9.8 Hz, 1H), 4.36 (dd, *J* = 10.5, 4.8 Hz, 1H), 3.85 (t, *J* = 8.8 Hz, 1H), 3.81 – 3.74 (m, 1H), 3.59 (t, *J* = 9.2 Hz, 1H), 3.55 – 3.48 (m, 2H), 2.91 – 2.66 (m, 2H), 1.33 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ136.96, 129.49, 128.52, 126.40, 102.08, 86.78, 80.47, 74.65, 73.28, 70.68, 68.72, 24.97, 15.47. $[\alpha]_D^{20}$ 48.95 (c 0.39 g/100 mL, $CHCl_3$). IR ν = 3406, 2926, 1074, 1029, 1004, 700 cm^{-1}. (ESI-HRMS) m/z 335.083 $[M+Na]^+$ ($C_{15}H_{20}O_5SNa$ requires 335.092).

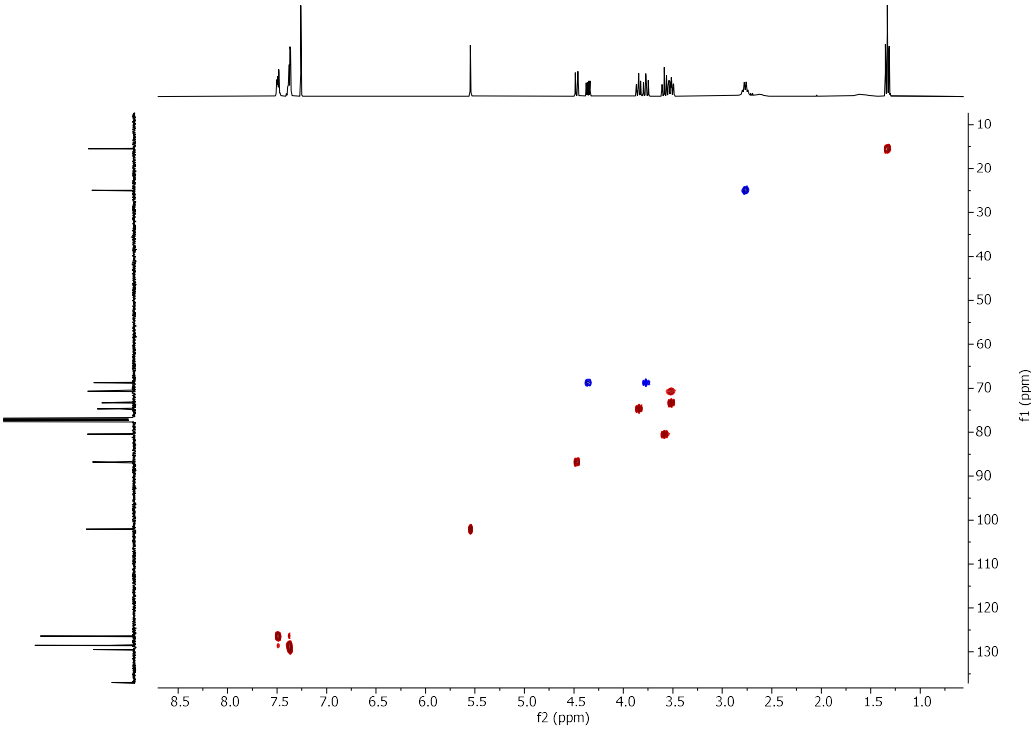
^{1}H NMR of 3.3 (400 MHz, $CDCl_3$)



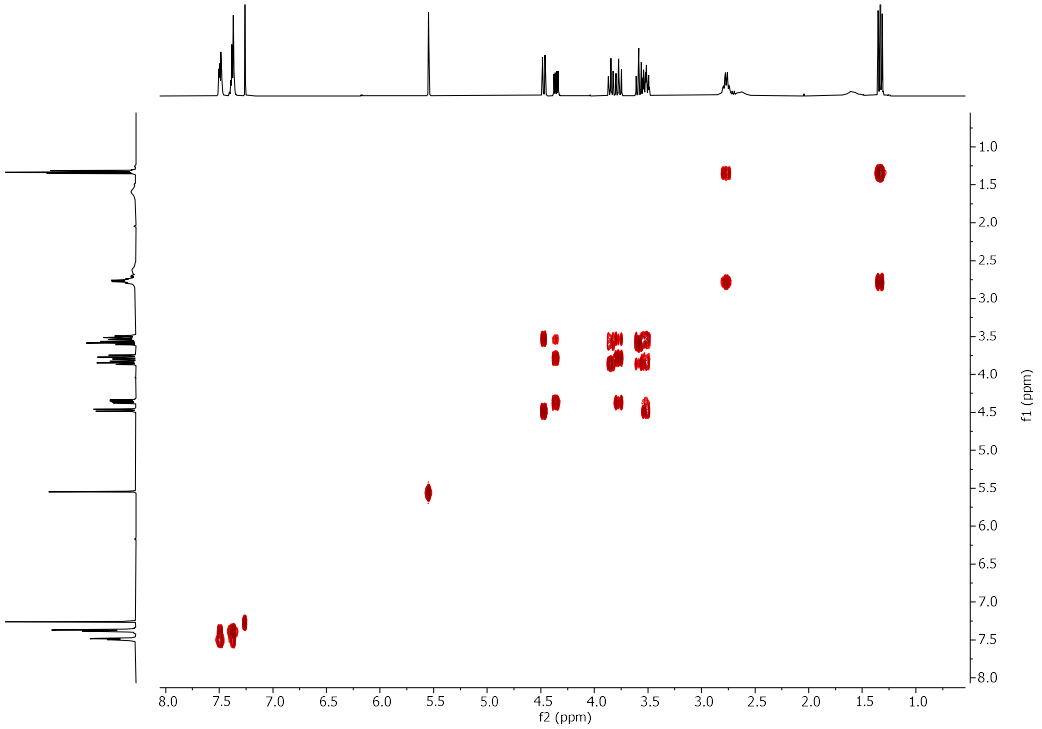
^{13}C NMR of 3.3 (101 MHz, $CDCl_3$)



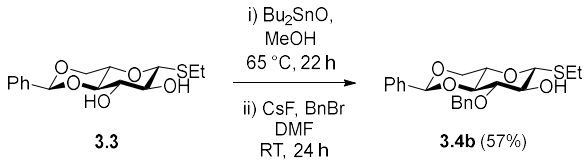
HSQC NMR of 3.3 ($CDCl_3$)



COSY NMR of 3.3 ($CDCl_3$)



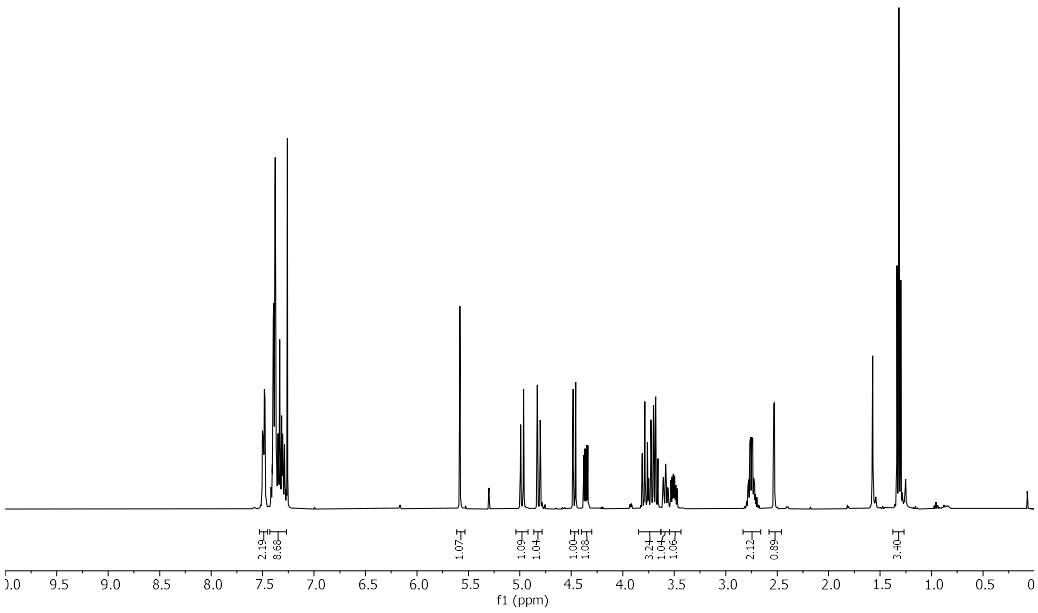
Synthesis of **3.4b**



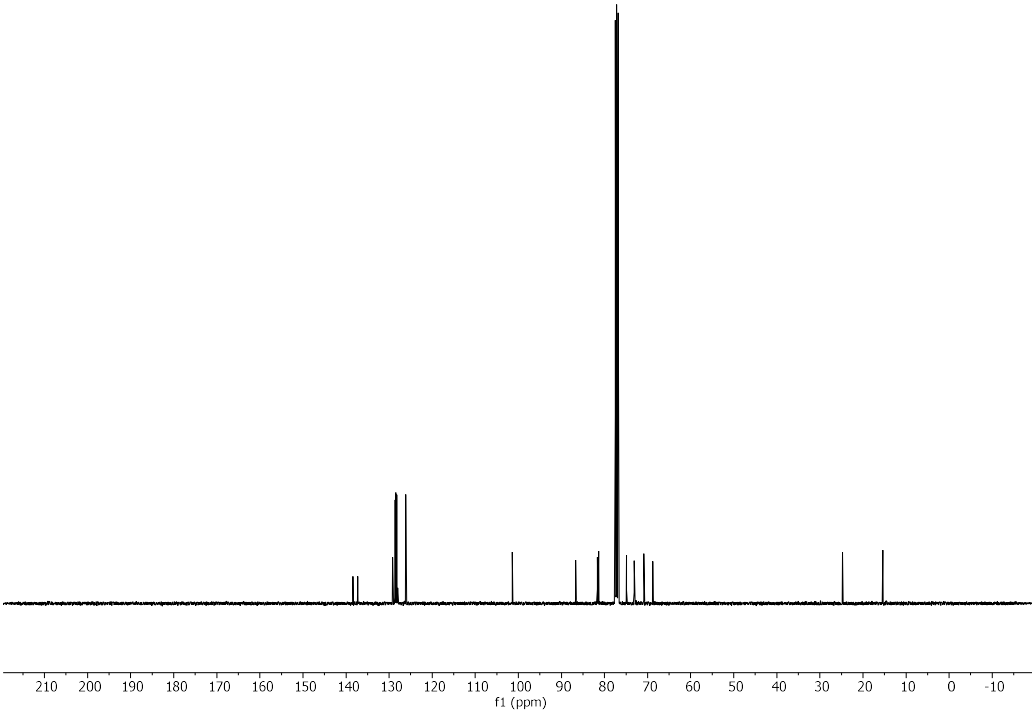
3.3 (1.06 g, 3.40 mmol) was dissolved in MeOH (40 mL), di-*n*-butyltin oxide (Bu_2SnO) (1.02 g, 4.10 mmol) was added and the reaction mixture (white suspension) heated at reflux (65 °C) under vigorous stirring for 22 h. The reaction mixture (clear solution) was then cooled, concentrated under reduced pressure and the crude product was used in the next step without further purification. The crude was dissolved in DMF (20 mL). Benzyl bromide (0.49 mL, 4.13 mmol) and cesium (I) fluoride (CsF) (670 g, 4.41 mmol) were added and the clear solution stirred at RT for 24 h under Ar atmosphere. The cloudy reaction mixture was diluted with EtOAc and the organic layer was passed through a short plug of silica gel and concentrated under reduced pressure. The crude product was diluted with EtOAc and the organic layer washed once with an aqueous solution of KF (1 M), once with water, once with brine, dried over Na_2SO_4, and concentrated under reduced pressure. The crude product was purified by silica gel flash column chromatography (Hexane : Acetone = 3:1→2:1→1:1) to yield **3.4b** as a white solid (0.78 g, 57%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.52 – 7.46 (m, 2H), 7.43 – 7.27 (m, 8H), 5.58 (s, 1H), 4.98 (d, *J* = 11.6 Hz, 1H), 4.82 (d, *J* = 11.6 Hz, 1H), 4.47 (d, *J* = 9.6 Hz, 1H), 4.36 (dd, *J* = 10.5, 5.0 Hz, 1H), 3.79 (t, *J* = 10.3 Hz, 1H), 3.75 – 3.64 (m, 2H), 3.58 (ddd, *J* = 9.9, 8.1, 1.9 Hz, 1H), 3.50 (ddd, *J* = 10.0, 8.7, 4.9 Hz, 1H), 2.75 (qd, *J* = 7.4, 2.8 Hz, 2H), 2.53 (d, *J* = 2.0 Hz, 1H), 1.32 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 138.38, 137.30, 129.16, 128.62, 128.42, 128.21, 128.01, 126.13, 101.38, 86.69, 81.65, 81.35, 74.87, 73.09, 70.88, 68.77, 24.73, 15.38. $[\alpha]_D^{20}$ +39.51 (c 0.63 g/100 mL, $CHCl_3$). IR ν = 3362, 1367, 1088, 1063, 1009, 746, 697 cm^{-1}. R_f = 0.31 (Hexane : EtOAc 3:1). (ESI-HRMS) m/z 425.139 $[M+Na]^+$ ($C_{22}H_{26}O_5SNa$ requires 425.139).

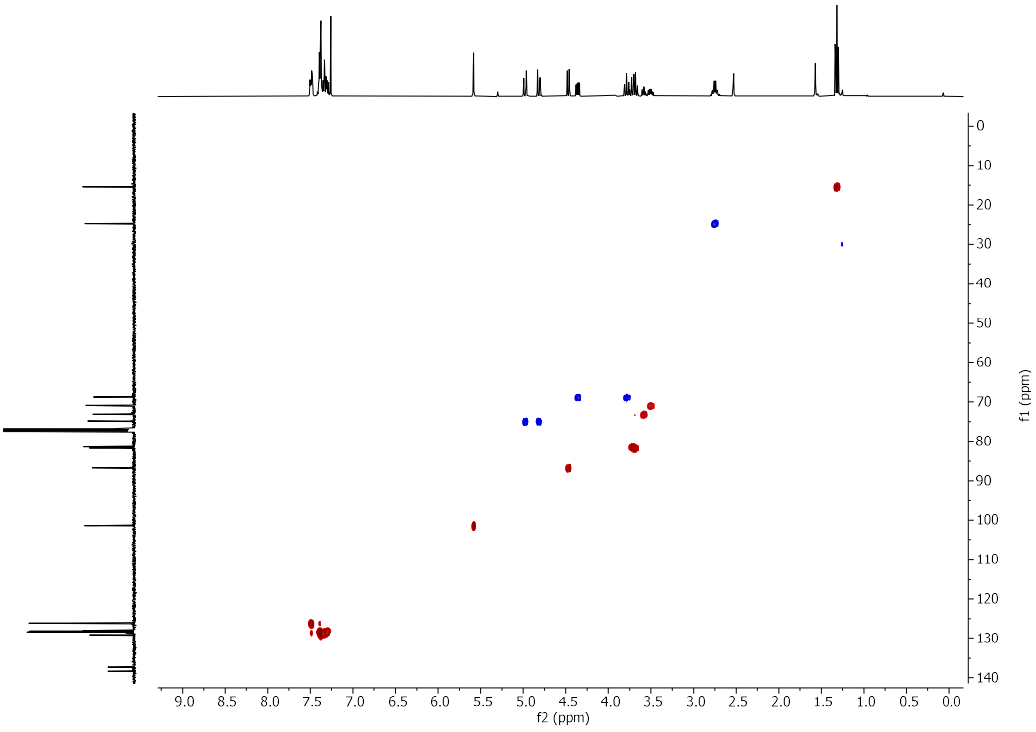
^{1}H NMR of 3.4b (400 MHz, $CDCl_3$)



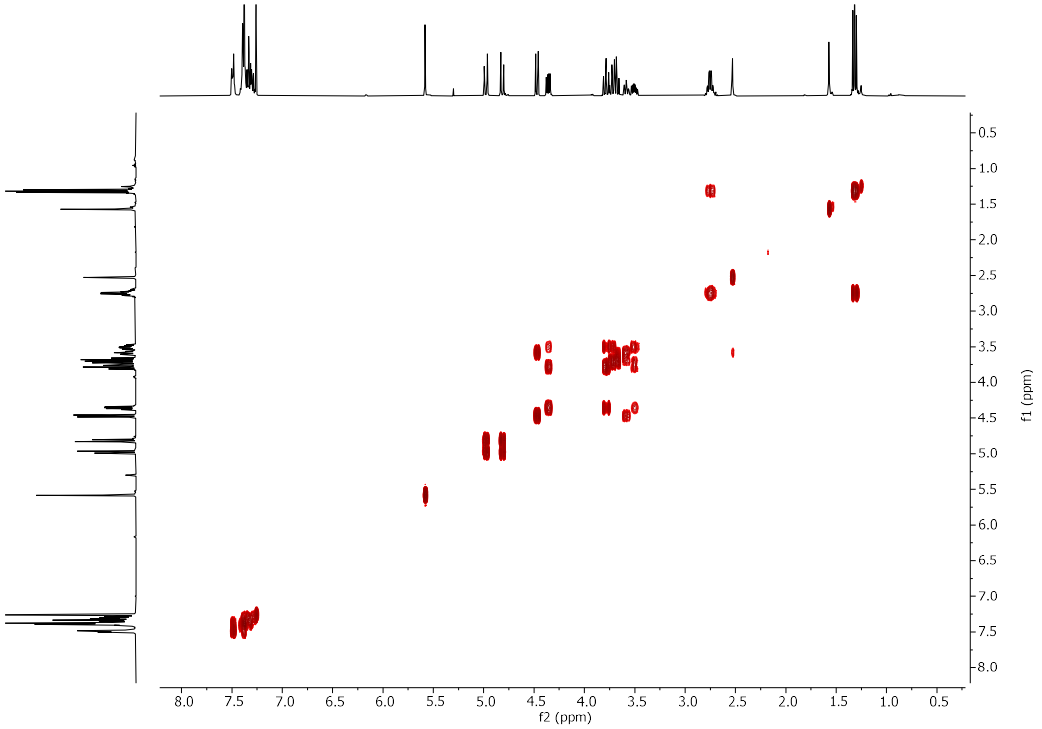
^{13}C NMR of 3.4b (101 MHz, $CDCl_3$)



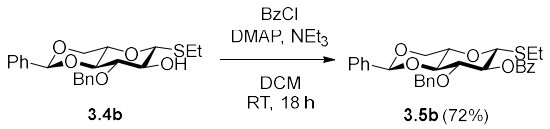
HSQC NMR of 3.4b ($CDCl_3$)



COSY NMR of 3.4b ($CDCl_3$)



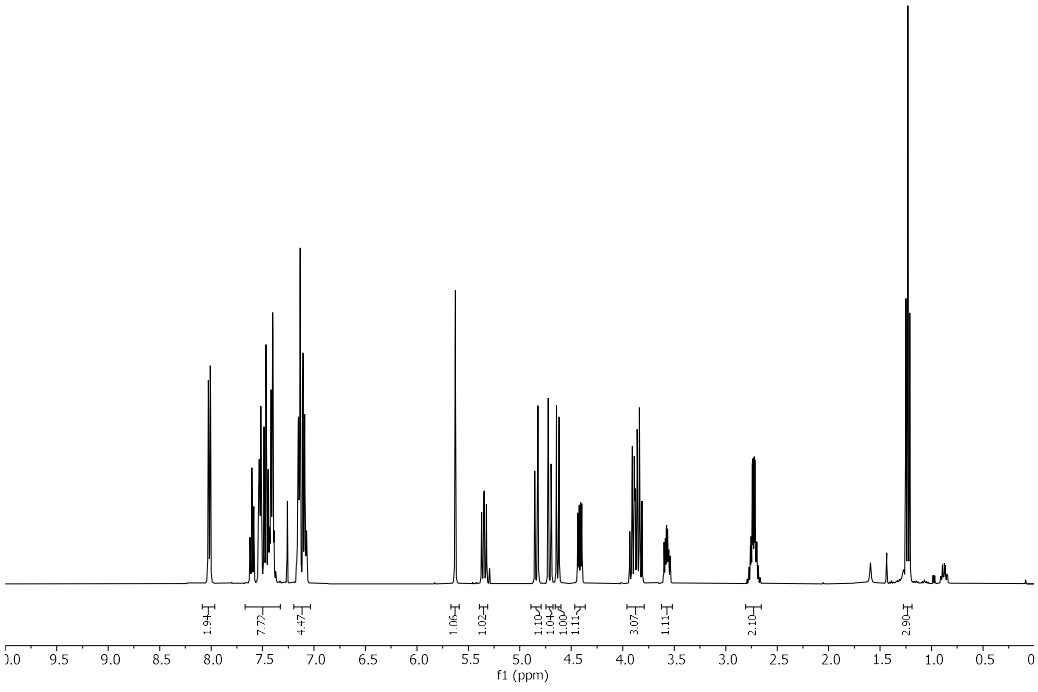
Synthesis of **3.5b**



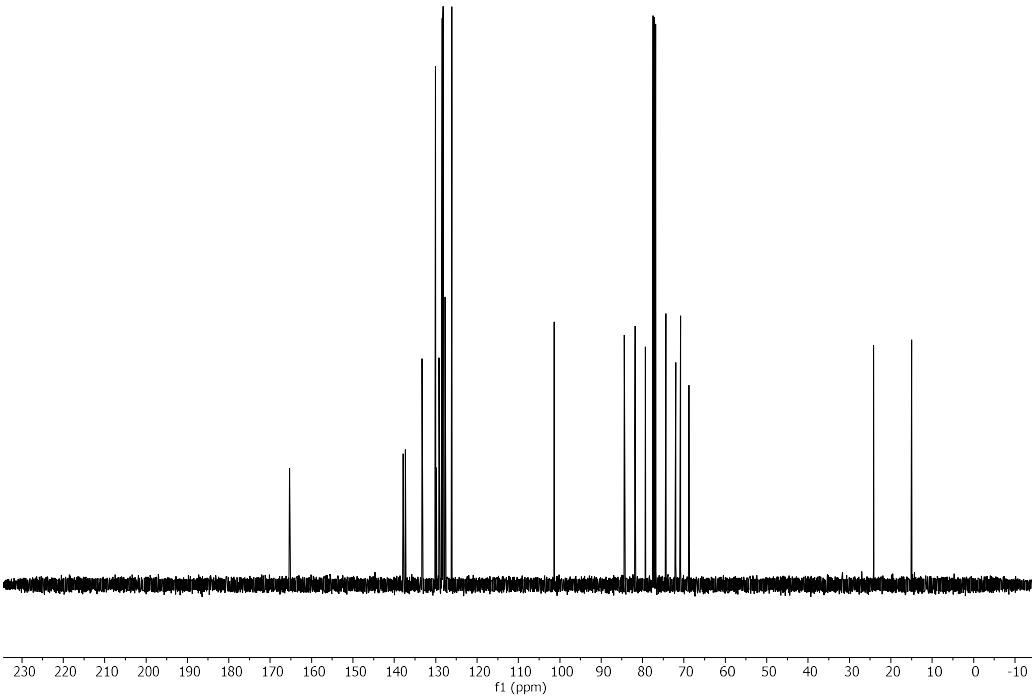
3.4b (775 mg, 1.93 mmol) was dissolved in anhydrous DCM (40 mL) under Ar atmosphere. Triethylamine (NEt_3) (2.86 mL, 0.58 mmol) and 4-dimethylaminopyridine (DMAP) (70 mg, 0.58 mmol) were added to the solution, while stirring. Benzoyl chloride (BzCl) (350 µL, 3.03 mmol) was slowly added at 0°C and the reaction allowed to RT. After 18 h the reaction was diluted with DCM and quenched with a saturated aqueous solution of $NaHCO_3$. The organic layer was washed three times with a saturated aqueous solution of $NaHCO_3$ and once with brine. The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified through a short plug of silica (EtOAc isocratic) and recrystallized from Hexane : EtOAc to yield **3.5b** as a white solid (708 mg, 72%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.06 – 7.96 (m, 2H), 7.66 – 7.33 (m, 8H), 7.19 – 7.04 (m, 5H), 5.63 (s, 1H), 5.35 (dd, *J* = 10.0, 8.3 Hz, 1H), 4.84 (d, *J* = 11.9 Hz, 1H), 4.71 (d, *J* = 11.9 Hz, 1H), 4.63 (d, *J* = 10.1 Hz, 1H), 4.42 (dd, *J* = 10.5, 5.0 Hz, 1H), 3.96 – 3.79 (m, 3H), 3.57 (ddd, *J* = 10.0, 8.8, 4.9 Hz, 1H), 2.73 (qd, *J* = 7.5, 3.6 Hz, 2H), 1.23 (t, *J* = 7.5 Hz, 3H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 165.29, 137.88, 137.30, 133.33, 130.05, 129.85, 129.18, 128.50, 128.42, 128.29, 128.18, 127.70, 126.13, 101.39, 84.42, 81.79, 79.31, 74.34, 71.97, 70.84, 68.77, 24.15, 14.94. $[\alpha]_D^{20}$ -16.49 (c 1.77 g/100 mL, $CHCl_3$). IR ν = 1728, 1269, 1094, 1070, 699 cm^{-1}. R_f = 0.47 (Hexane : EtOAc = 3:1). (ESI-HRMS) m/z 529.166 $[M+Na]^+$ ($C_{29}H_{30}O_6SNa$ requires 529.166).

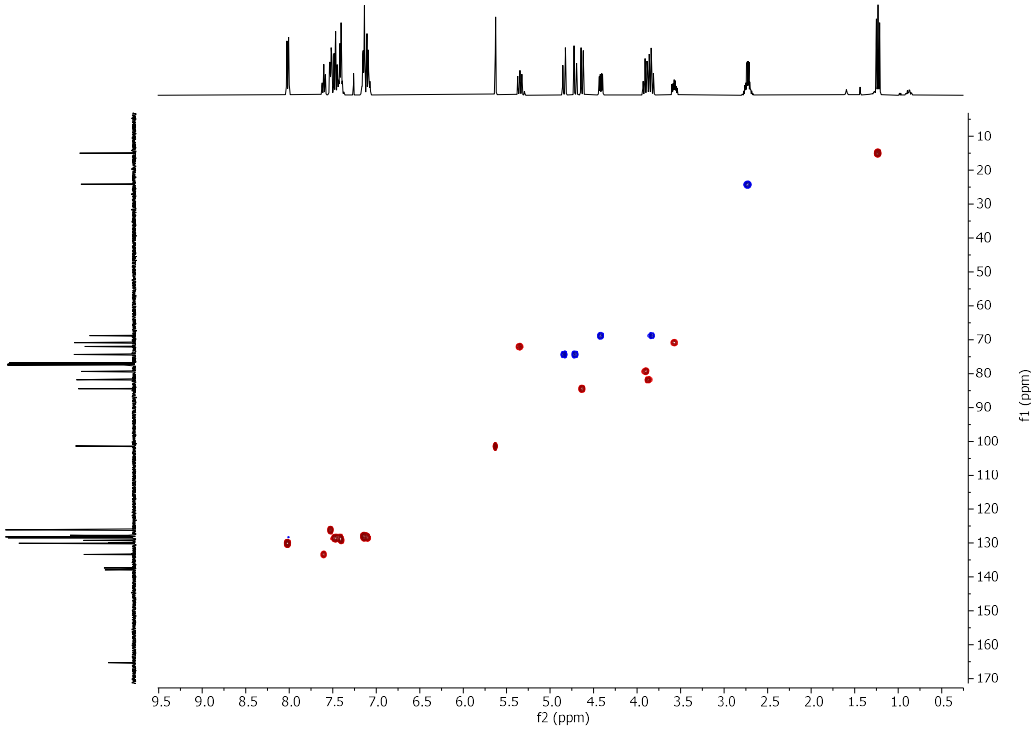
^{1}H NMR of 3.5b (400 MHz, $CDCl_3$)



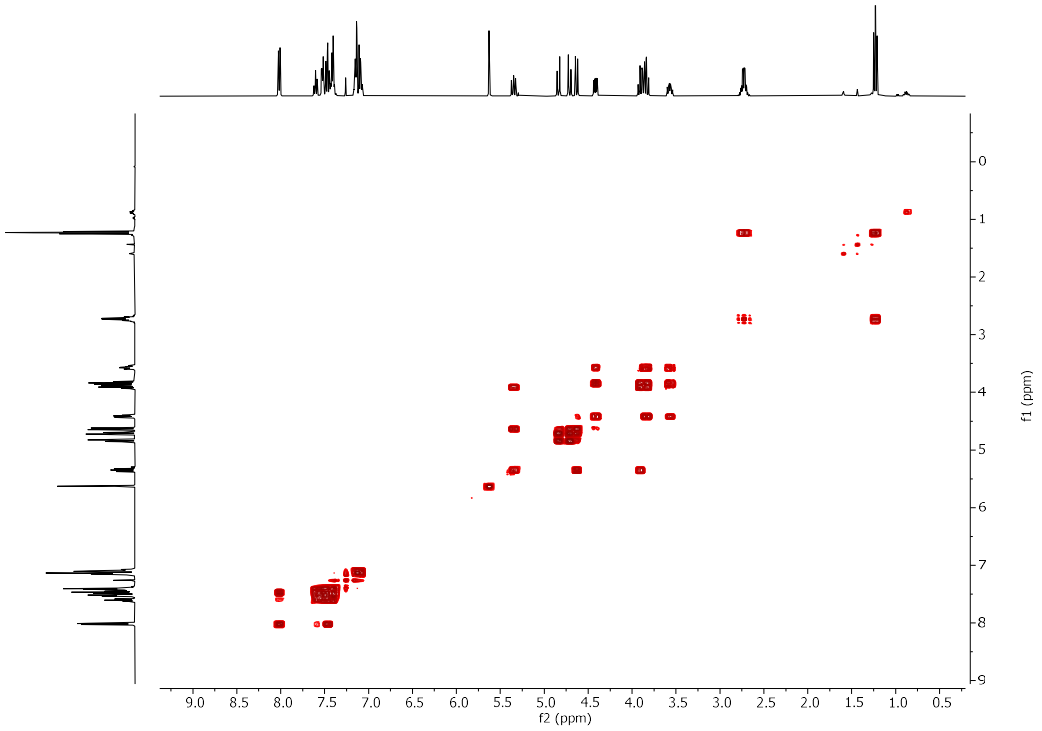
^{13}C NMR of 3.5b (101 MHz, $CDCl_3$)



HSQC NMR of 3.5b ($CDCl_3$)



COSY NMR of 3.5b ($CDCl_3$)

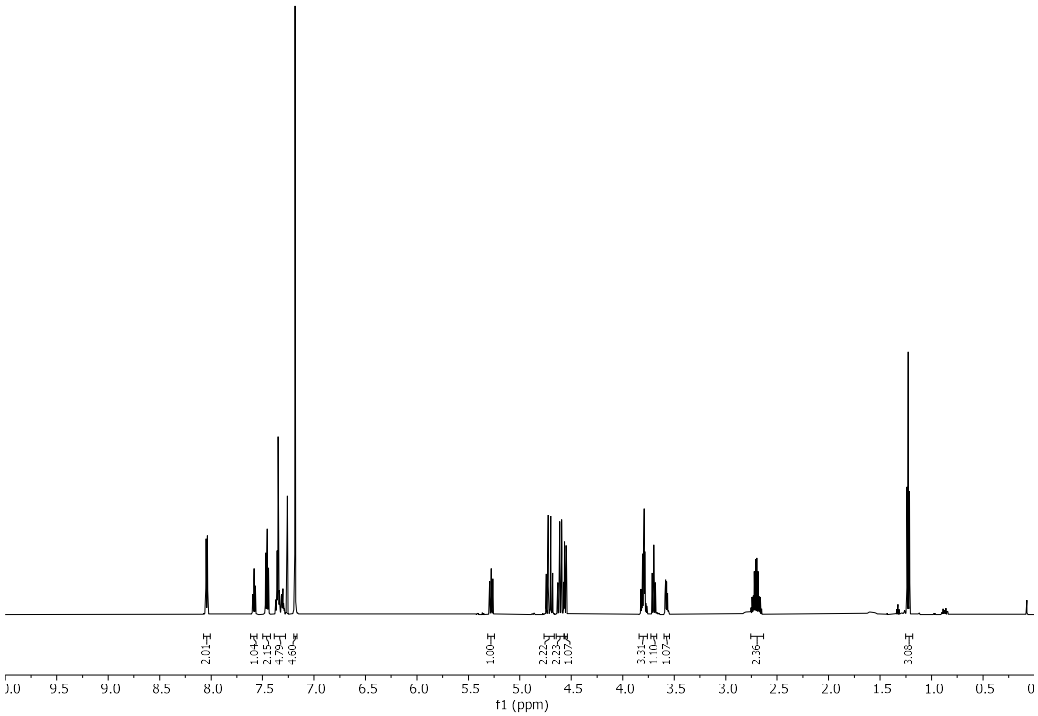


Synthesis of **3.6b**

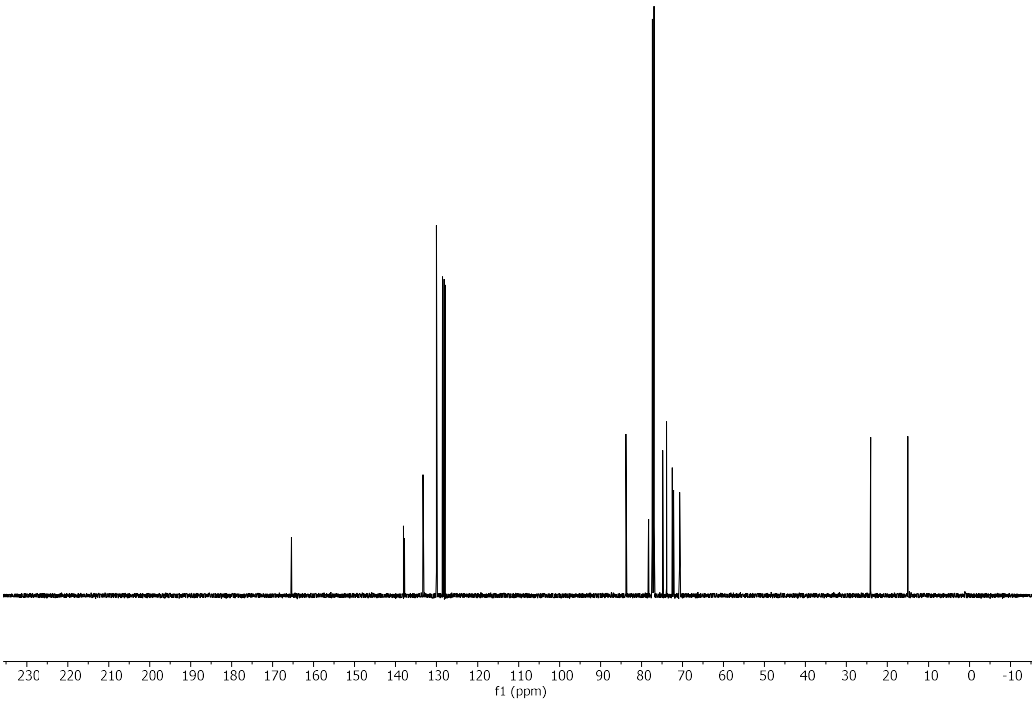
3.5b (914 mg, 1.81 mmol) was dissolved in anhydrous DCM (30 mL) under Ar atmosphere. Triethylsilane (Et_3SiH) (1.72 mL, 10.8 mmol) and trifluoroacetic anhydride (TFAA) (250 µL, 1.80 mmol) were sequentially added to the stirred solution at 0°C. After 20 min, trifluoroacetic acid (TFA) (830 µL, 11.2 mmol) was added dropwise at 0°C. The reaction was allowed to RT and after 1 h quenched with a saturated aqueous solution of $NaHCO_3$. The organic layer was washed twice with a saturated aqueous solution of $NaHCO_3$, and once with brine. The crude product was purified by silica gel flash column chromatography (Hexane : EtOAc = 5:1→2:1) to give **3.6b** as a white solid (0.77 g, 84%).

^{1}H NMR (600 MHz, Chloroform-*d*) δ 8.08 – 8.02 (m, 2H), 7.58 (ddt, *J* = 8.6, 7.2, 1.3 Hz, 1H), 7.50 – 7.43 (m, 2H), 7.39 – 7.28 (m, 5H), 7.18 (s, 5H), 5.28 (dd, *J* = 10.0, 9.1 Hz, 1H), 4.73 (d, *J* = 11.5 Hz, 1H), 4.69 (d, *J* = 11.5 Hz, 1H), 4.62 (d, *J* = 12.0 Hz, 1H), 4.58 (d, *J* = 12.0 Hz, 1H), 4.56 (d, *J* = 10.0 Hz, 1H), 3.85 – 3.75 (m, 4H), 3.70 (t, *J* = 9.0 Hz, 1H), 3.57 (dt, *J* = 9.6, 4.8 Hz, 1H), 2.76 – 2.63 (m, 2H), 1.23 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (151 MHz, Chloroform-*d*) δ 165.41, 138.08, 137.84, 133.33, 130.00, 128.63, 128.56, 128.54, 128.13, 128.00, 127.93, 127.91, 83.78, 83.68, 78.31, 74.80, 73.92, 72.48, 72.15, 70.69, 24.10, 15.02. $[\alpha]_D^{20}$ -3.03 (c 1.45 g/100 mL, $CHCl_3$). IR ν = 3479, 1725, 1269, 1068, 710, 698 cm^{-1}. R_f = 0.39 (Hexane : EtOAc = 3:1). (ESI-HRMS) m/z 531.181 $[M+Na]^+$ ($C_{29}H_{32}O_6SNa$ requires 531.181).

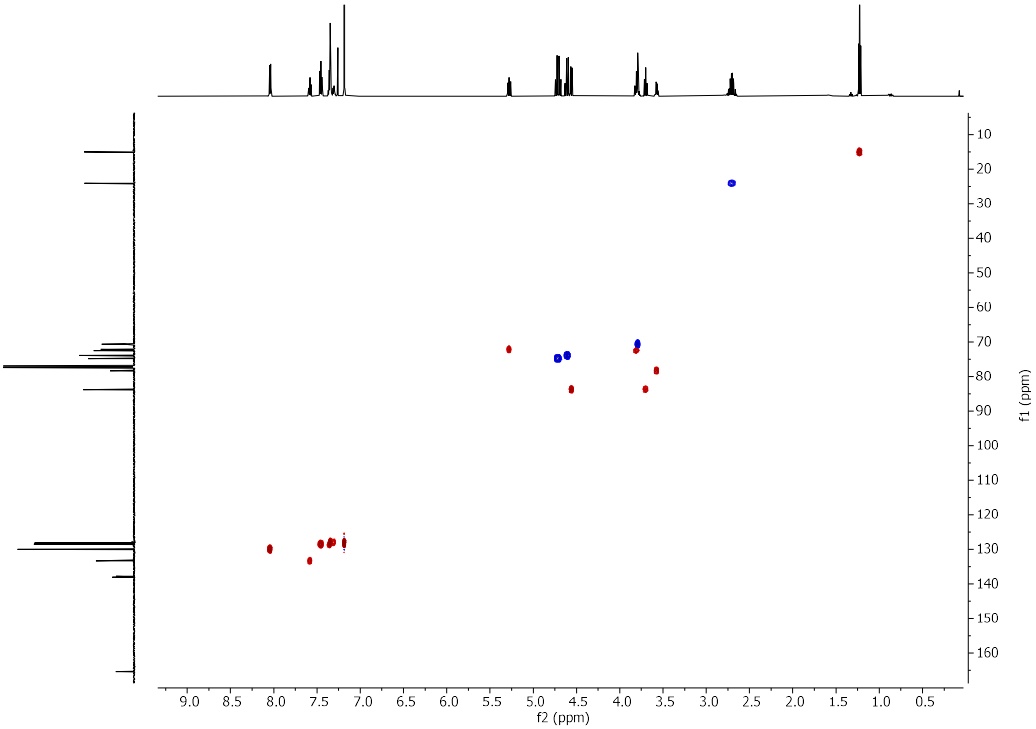
^{1}H NMR of 3.6b (600 MHz, $CDCl_3$)



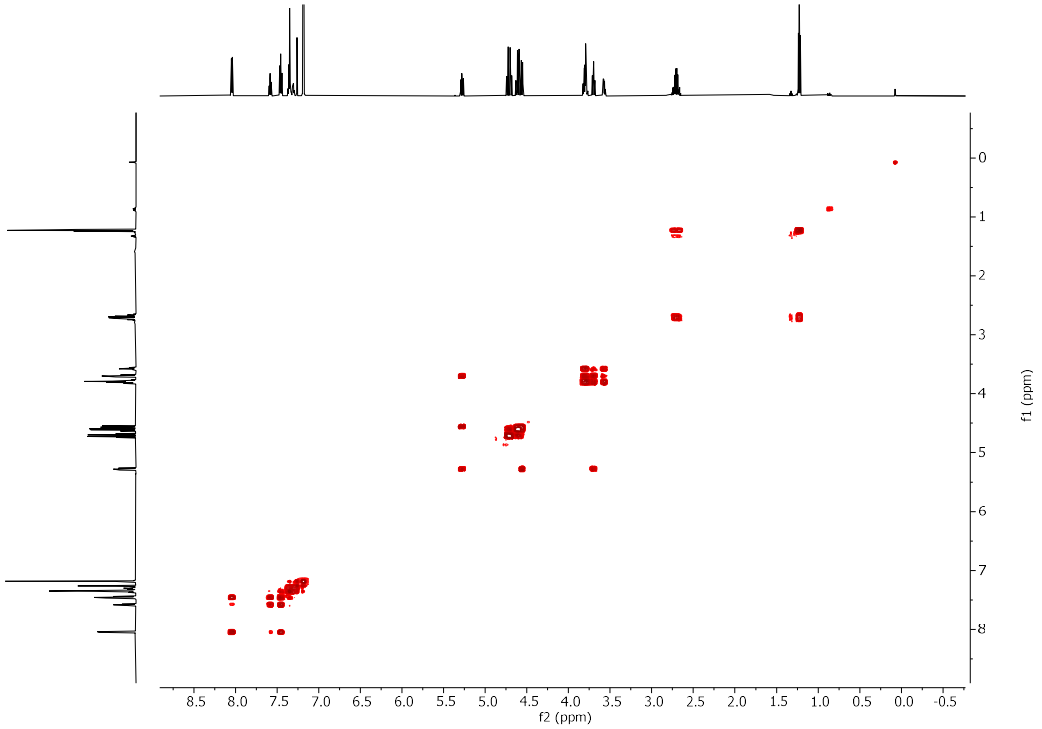
^{13}C NMR of 3.6b (151 MHz, $CDCl_3$)



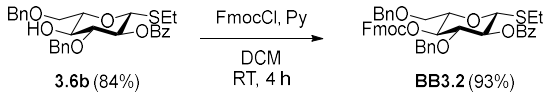
HSQC NMR of 3.6b ($CDCl_3$)



COSY NMR of 3.6b ($CDCl_3$)



Synthesis of **BB3.2**

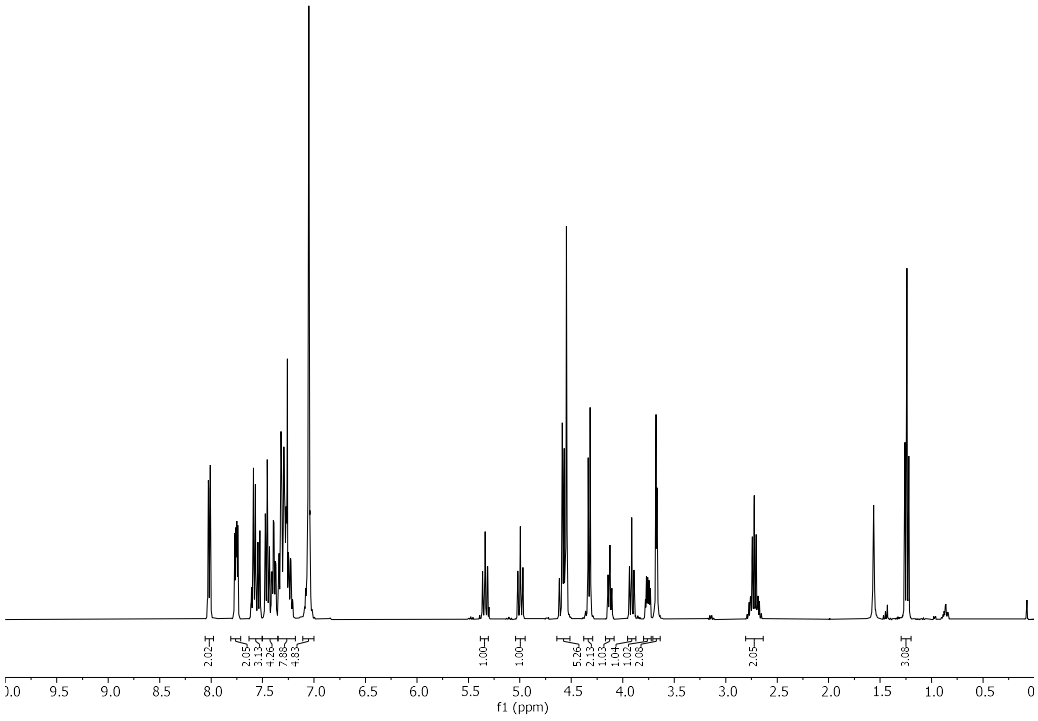


3.6b (84%) **BB3.2** (93%)

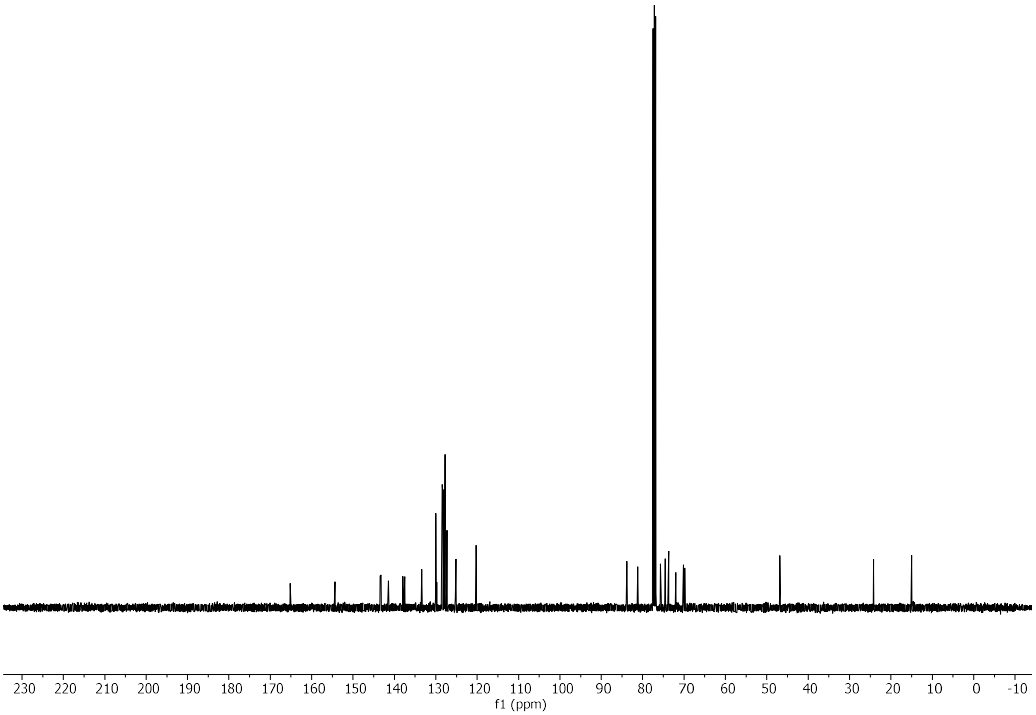
3.6b (769 mg, 1.52 mmol) was dissolved in anhydrous DCM (15 mL) and pyridine (Py) was added (370 µL, 4.6 mmol) to the stirred solution under Ar atmosphere. 9-Fluorenylmethyl chloroformate (FmocCl) (798 mg, 3.1 mmol) was dissolved in anhydrous DCM (5 mL) and added to the reaction mixture. The yellow solution was stirred for 4 h then quenched with an aqueous solution of HCl (1 M). The organic layer was washed three times with an aqueous solution of HCl (1 M), and once with brine. The crude product was purified by silica gel flash column chromatography (Toluene : DCM = 4:1→3:1 then Toluene : EtOAc = 4:1) to give **BB3.2** as a white solid (1.04 g, 93%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.06 – 7.97 (m, 2H), 7.80 – 7.72 (m, 2H), 7.63 – 7.51 (m, 3H), 7.49 – 7.20 (m, 11H), 7.05 (d, *J* = 3.1 Hz, 5H), 5.34 (t, *J* = 9.6 Hz, 1H), 5.00 (t, *J* = 9.6 Hz, 1H), 4.64 – 4.52 (m, 5H), 4.33 (d, *J* = 7.2 Hz, 2H), 4.13 (t, *J* = 7.1 Hz, 1H), 3.91 (t, *J* = 9.2 Hz, 1H), 3.76 (ddd, *J* = 9.6, 5.0, 3.9 Hz, 1H), 3.67 (d, *J* = 4.2 Hz, 2H), 2.81 – 2.63 (m, 2H), 1.24 (t, *J* = 7.5 Hz, 2H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 165.17, 154.37, 143.42, 143.24, 141.44, 141.41, 138.00, 137.50, 133.40, 130.02, 129.79, 128.56, 128.47, 128.29, 128.05, 127.97, 127.78, 127.76, 127.31, 125.24, 125.15, 120.22, 83.82, 81.22, 77.53, 75.69, 74.52, 73.73, 72.02, 70.17, 69.83, 46.82, 24.20, 15.02. $[\alpha]_D^{20}$ -26.84 (c 1.01 g/100 mL, $CHCl_3$). IR ν = 1754, 1729, 1248, 1070, 739, 710 cm^{-1}. R_f = 0.53 (Hexane : EtOAc = 3:1). (ESI-HRMS) m/z 753.254 $[M+Na]^+$ ($C_{44}H_{42}O_8SNa$ requires 753.249).

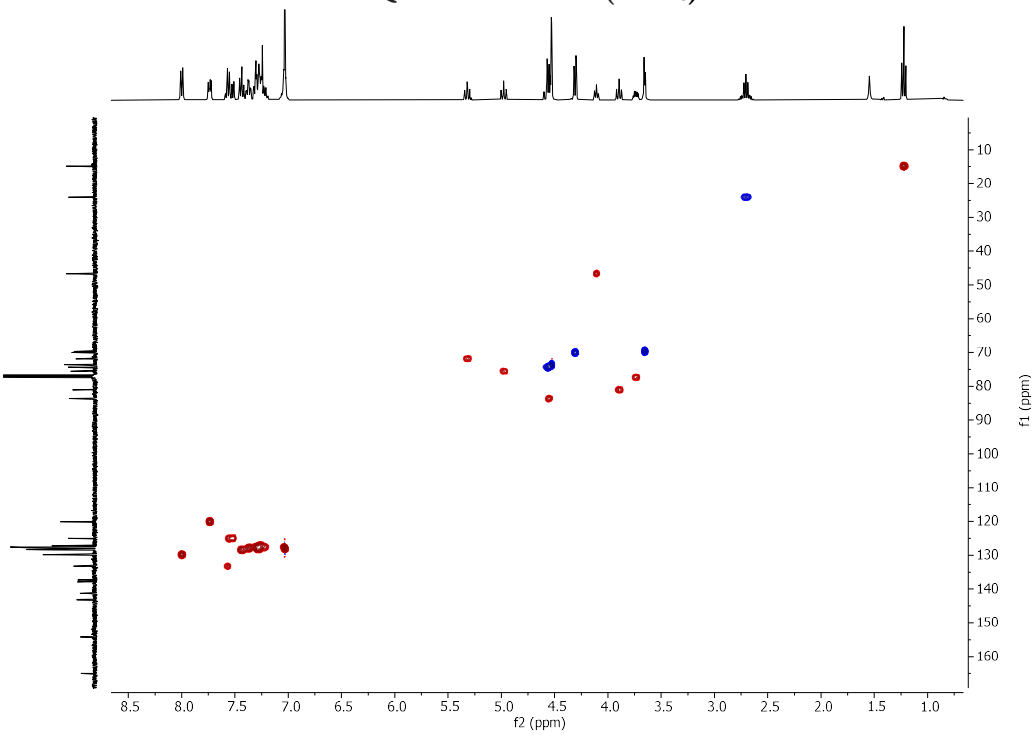
^{1}H NMR of BB3.2 (400 MHz, $CDCl_3$)



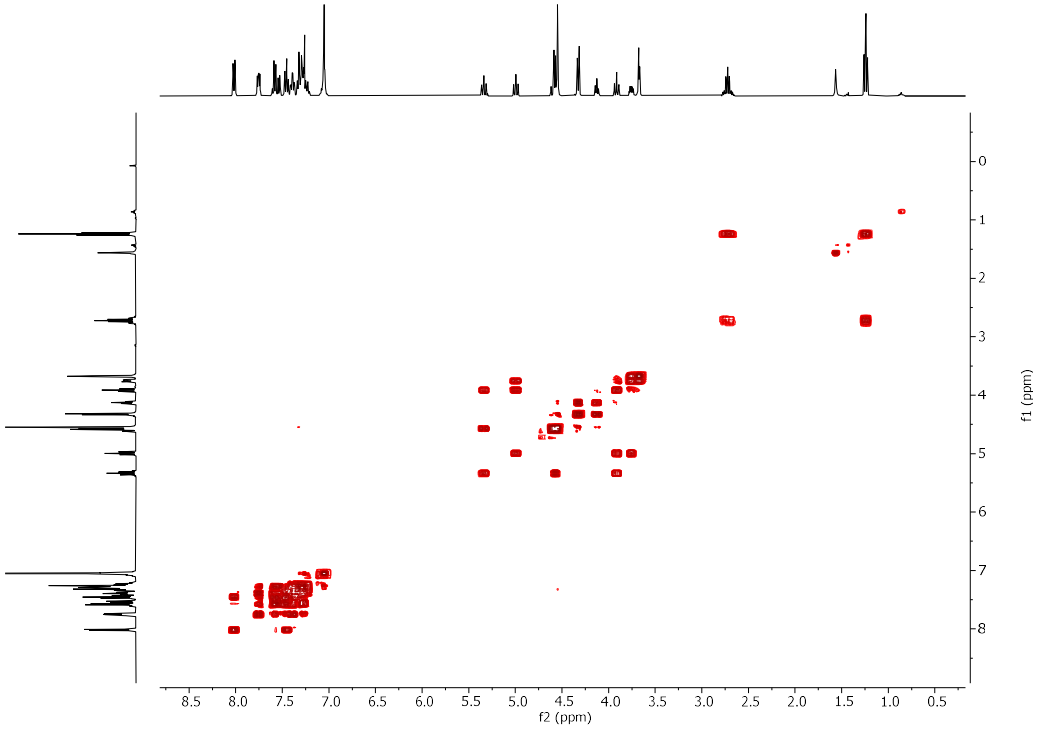
^{13}C NMR of BB3.2 (101 MHz, $CDCl_3$)



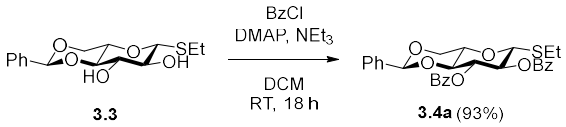
HSQC NMR of BB3.2 ($CDCl_3$)



COSY NMR of BB3.2 ($CDCl_3$)



Synthesis of **3.4a**

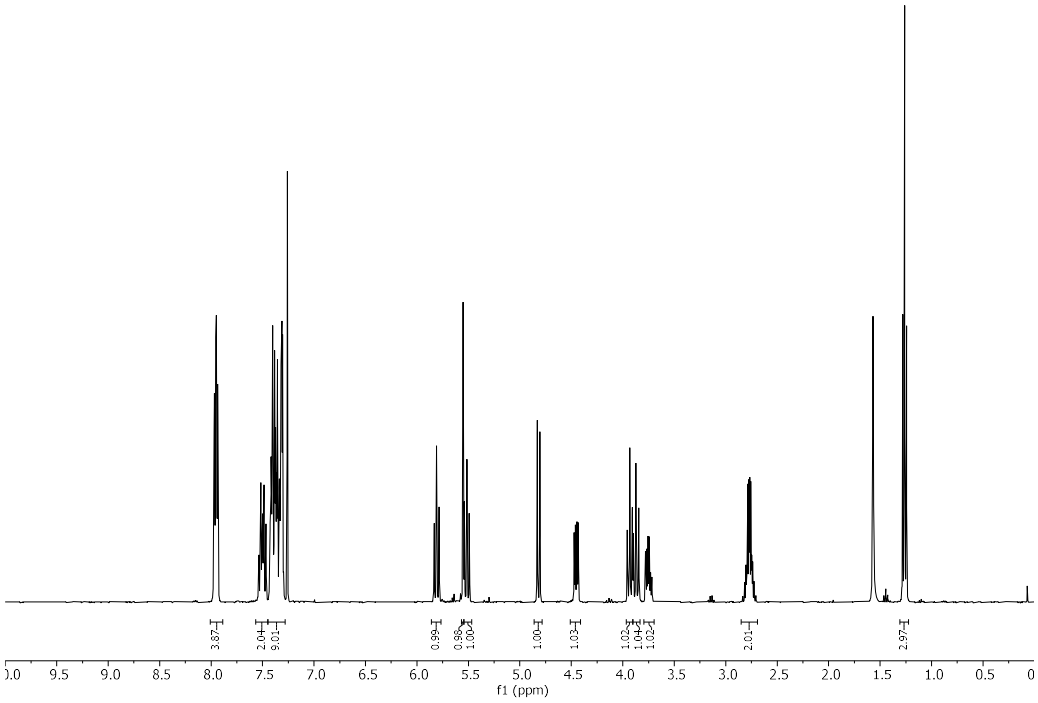


3.3 **3.4a** (93%)

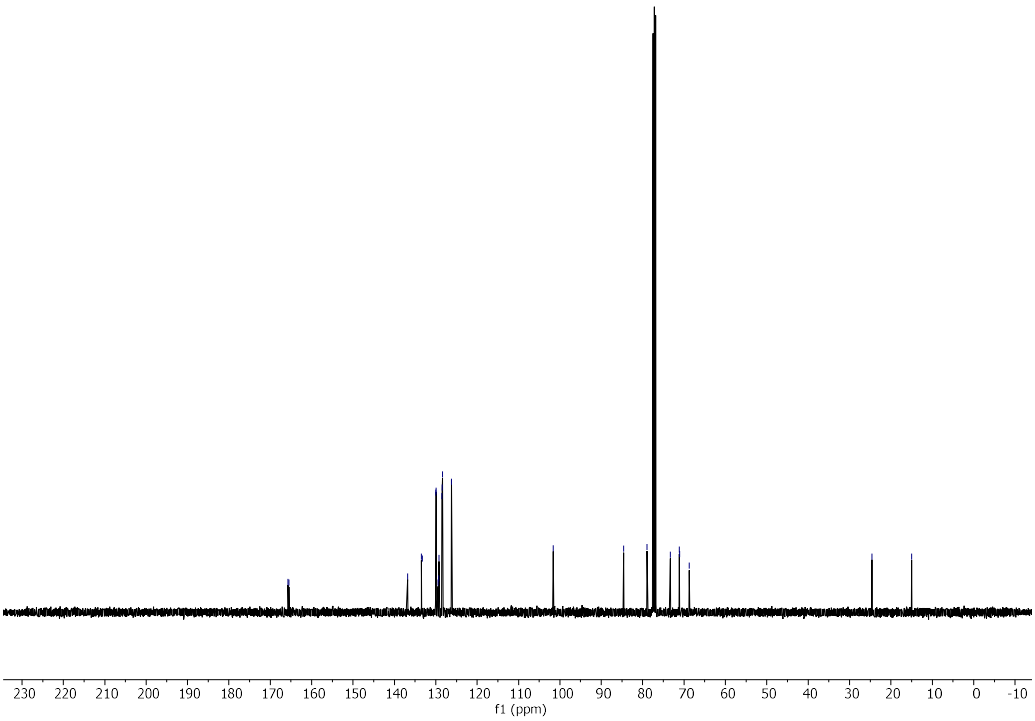
3.3 (0.85 mg, 2.72 mmol) was dissolved in anhydrous DCM (40 mL). Triethylamine (NEt_3) (1.1 mL, 7.9 mmol) and 4-dimethylaminopyridine (DMAP) (200 mg, 1.6 mmol) were added to the solution, while stirring. Benzoyl chloride (BzCl) (950 µL, 8.2 mmol) was slowly added at 0 °C and the reaction allowed to RT. Upon completion (18 h), the reaction was quenched with saturated aqueous solution of $NaHCO_3$. The mixture was washed three times with saturated aqueous solution of $NaHCO_3$ and once with brine. The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified through a short plug of silica with EtOAc followed by recrystallization from Hexane : EtOAc to yield **3.4a** as a white solid (1.31 g, 93%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.01 – 7.89 (m, 4H), 7.56 – 7.45 (m, 2H), 7.45 – 7.28 (m, 9H), 5.81 (t, *J* = 9.4 Hz, 1H), 5.55 (s, 1H), 5.52 (dd, *J* = 10.0, 9.2 Hz, 1H), 4.82 (d, *J* = 10.0 Hz, 1H), 4.45 (dd, *J* = 10.4, 4.8 Hz, 1H), 3.93 (t, *J* = 9.5 Hz, 1H), 3.87 (t, *J* = 10.2 Hz, 1H), 3.75 (td, *J* = 9.7, 4.8 Hz, 1H), 2.77 (qd, *J* = 7.5, 4.5 Hz, 2H), 1.26 (t, *J* = 7.5 Hz, 3H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 165.73, 165.46, 136.84, 133.46, 133.27, 130.02, 129.93, 129.48, 129.26, 129.20, 128.53, 128.45, 128.35, 126.24, 101.60, 84.63, 78.91, 73.30, 71.15, 71.09, 68.72, 24.58, 14.98. $[\alpha]_D^{20}$ -16.05 (c 0.54 g/100 mL, $CHCl_3$). IR ν = 2926, 1719, 1258, 1094, 1069, 709 cm^{-1}. (ESI-HRMS) m/z 543.145 $[M+Na]^+$ ($C_{29}H_{28}O_7SNa$ requires 543.145).

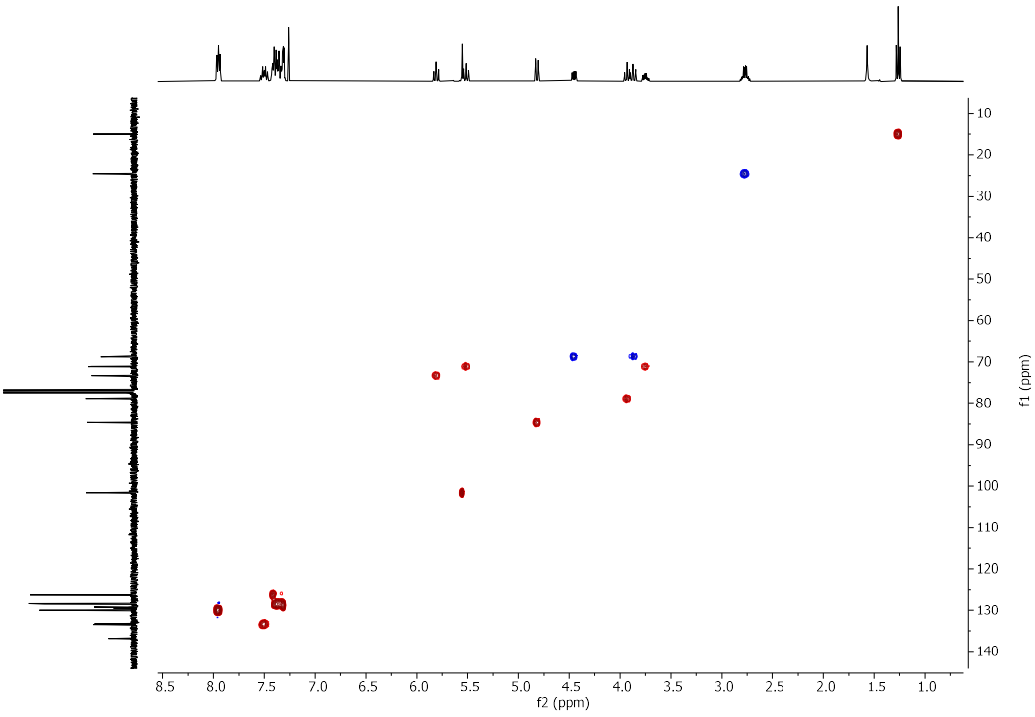
^{1}H NMR of 3.4a (400 MHz, $CDCl_3$)



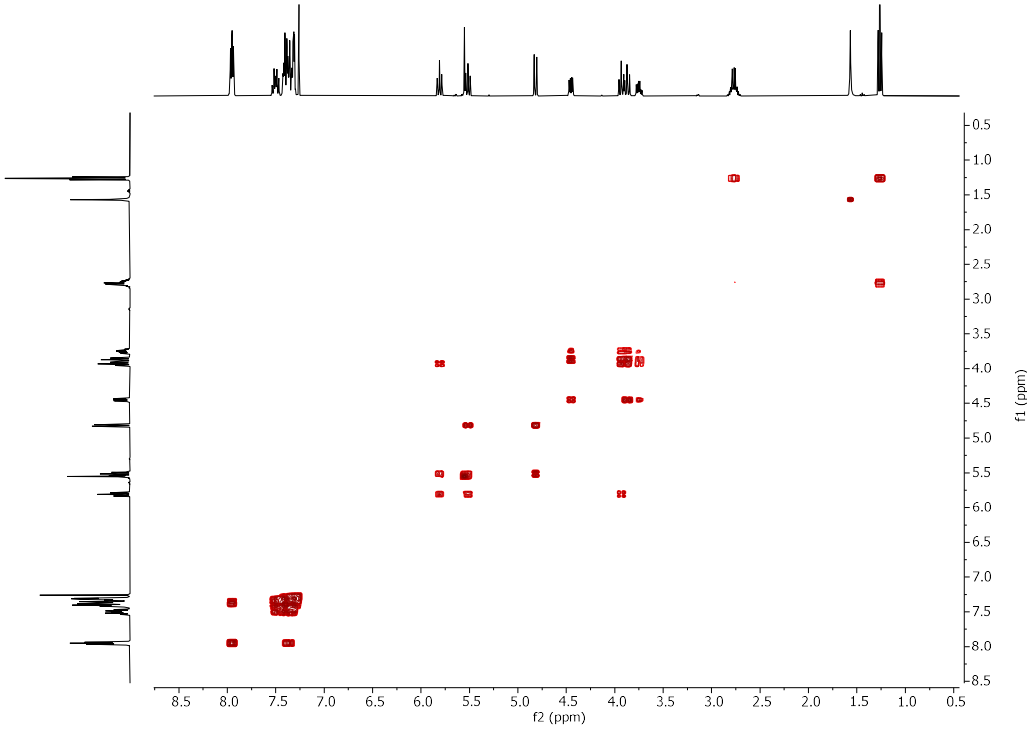
^{13}C NMR of 3.4a (101 MHz, $CDCl_3$)



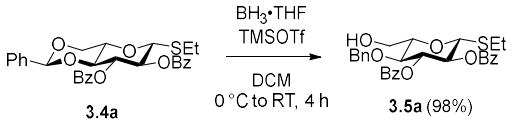
HSQC NMR of 3.4a ($CDCl_3$)



COSY NMR of 3.4a ($CDCl_3$)



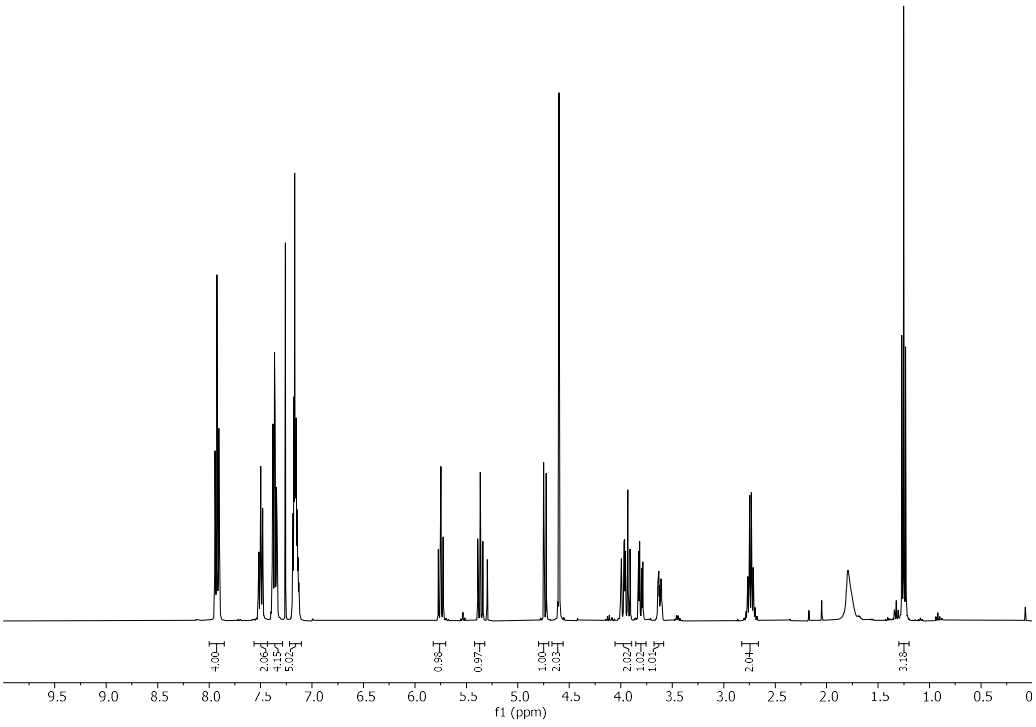
Synthesis of **3.5a**



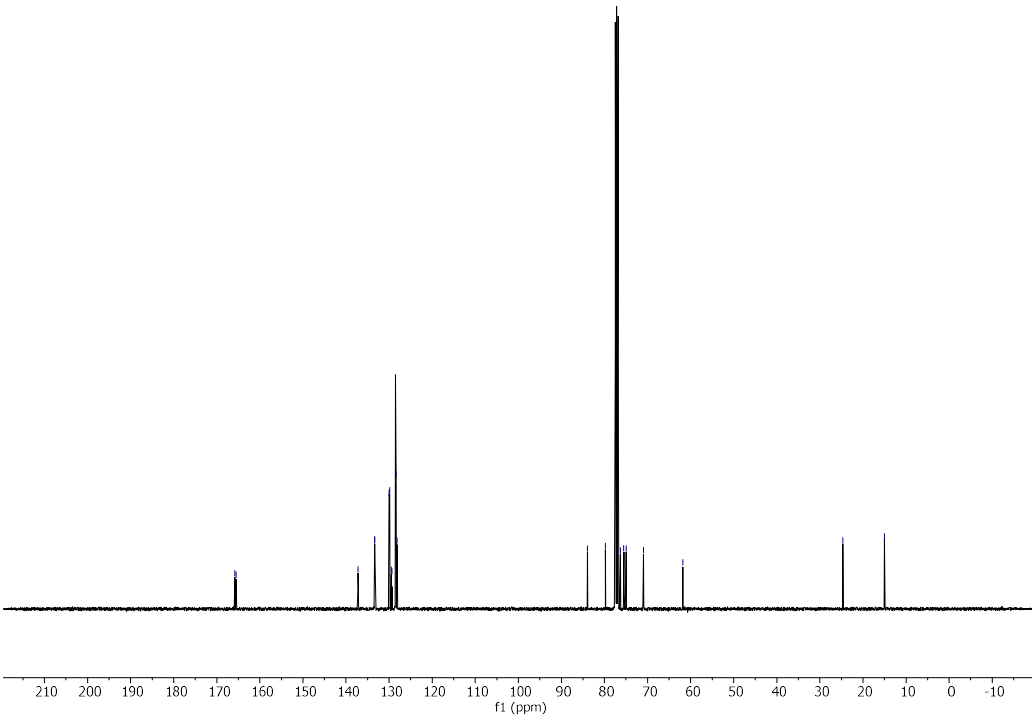
3.4a (1.31 g, 2.52 mmol) was dissolved in DCM (30 mL). BH_3·THF (10 mL, 10 mmol) and TMSOTf (230 µL, 0.5 mmol) were sequentially added to the stirred solution at 0 °C. The reaction was allowed to RT and quenched with saturated aqueous solution of $NaHCO_3$ after 4 h. The organic layer was washed three times with saturated aqueous solution of $NaHCO_3$ and once with brine, dried with Na_2SO_4 and concentrated in vacuum. The crude product was purified by silica gel flash column chromatography (Hexane : EtOAc = 3:1→3:2) to give **3.5a** as a colorless oil (1.29 g, 98%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.99 – 7.86 (m, 4H), 7.56 – 7.45 (m, 2H), 7.42 – 7.29 (m, 4H), 7.22 – 7.10 (m, 5H), 5.75 (t, *J* = 9.4 Hz, 1H), 5.36 (t, *J* = 9.8 Hz, 1H), 4.74 (d, *J* = 10.0 Hz, 1H), 4.60 (s, 2H), 3.98 (dd, *J* = 12.3, 2.5 Hz, 1H), 3.93 (t, *J* = 9.5 Hz, 1H), 3.81 (dd, *J* = 12.3, 4.0 Hz, 1H), 3.62 (ddd, *J* = 9.7, 4.1, 2.5 Hz, 1H), 2.74 (qd, *J* = 7.5, 1.9 Hz, 2H), 1.25 (t, *J* = 7.5 Hz, 3H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 165.81, 165.55, 137.22, 133.37, 133.36, 129.99, 129.85, 129.46, 129.31, 128.48, 128.37, 128.15, 83.95, 79.77, 76.32, 75.52, 74.98, 70.94, 61.84, 24.67, 15.02. $[\alpha]_D^{20}$ -54.95 (c 0.89 g/100 mL, $CHCl_3$). IR ν = 1726, 1273, 1090, 1070, 709 cm^{-1}. (ESI-HRMS) m/z 545.161 $[M+Na]^+$ ($C_{29}H_{30}O_7SNa$ requires 545.160).

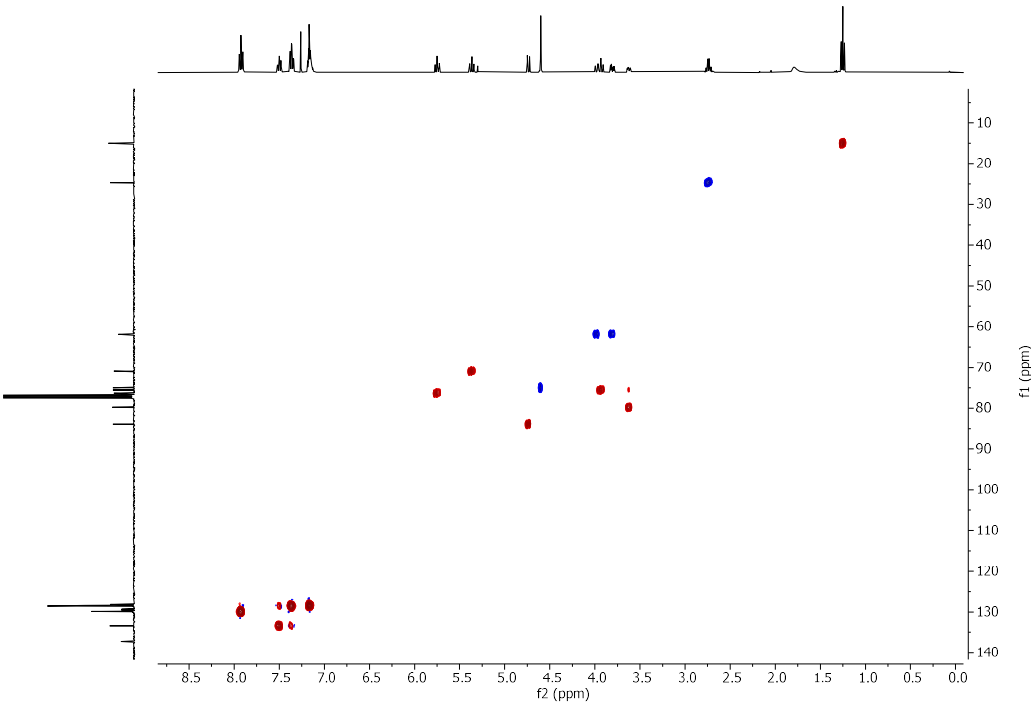
^{1}H NMR of 3.5a (400 MHz, $CDCl_3$)



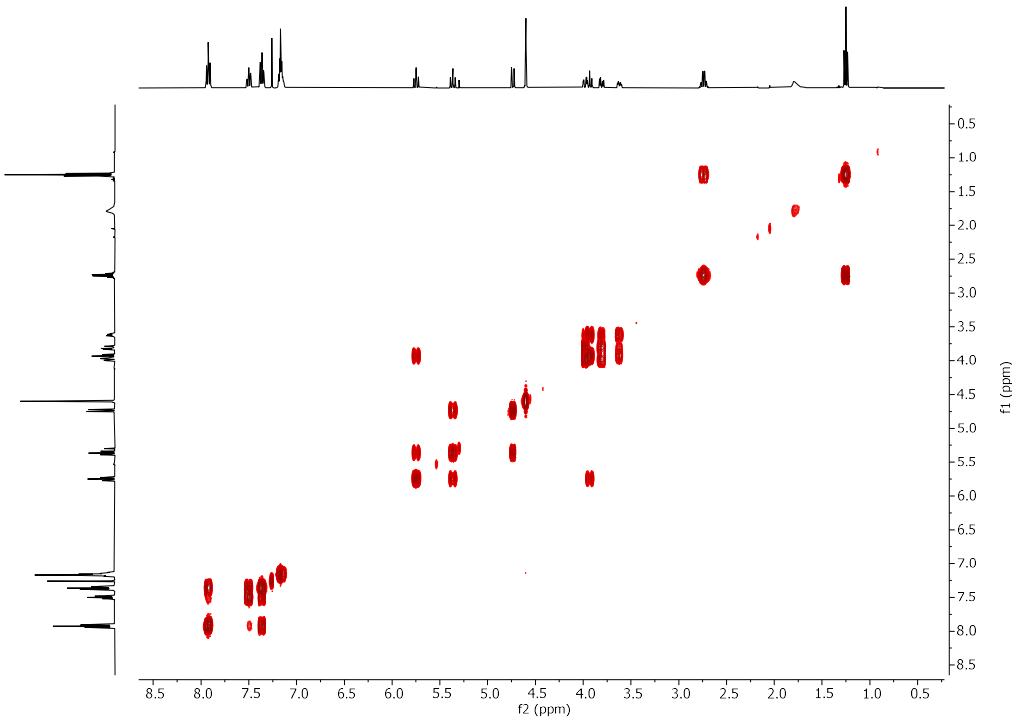
^{13}C NMR of 3.5a (101 MHz, $CDCl_3$)



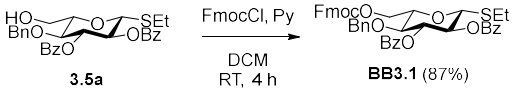
HSQC NMR of 3.5a ($CDCl_3$)



COSY NMR of 3.5a ($CDCl_3$)



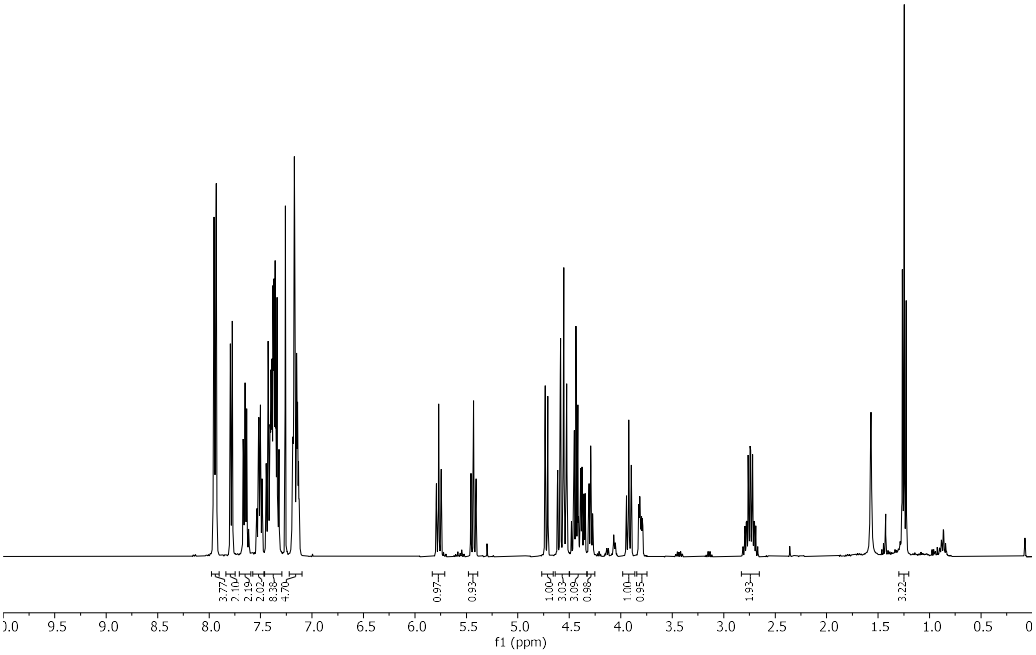
Synthesis of **BB3.2**



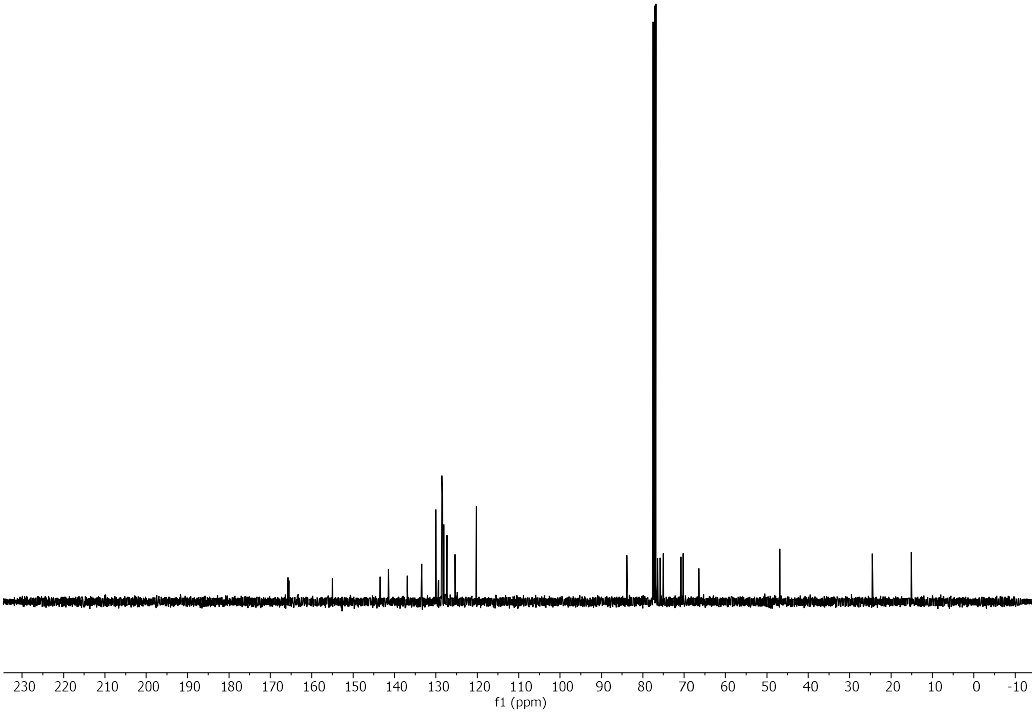
3.6a (1.29 g, 2.46 mmol) was dissolved in DCM (30 mL) and pyridine was added (600 µL, 7.43 mmol). 9-Fluorenylmethyl chloroformate (FmocCl) (1.28 g, 4.95 mmol) was dissolved in DCM (5 mL) and added to the reaction mixture under Ar atmosphere. The yellow solution was stirred for 4 h then quenched with a 1 M solution of HCl. The organic layer was washed three times with a 1 M solution of HCl and once with brine. The crude compound was purified by silica gel flash column chromatography (Toluene : DCM = 4:1→3:1 then Toluene : EtOAc = 4:1) to give **BB3.1** as a white solid (1.60 g, 87%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.00 – 7.89 (m, 4H), 7.78 (d, *J* = 7.5 Hz, 2H), 7.70 – 7.59 (m, 2H), 7.51 (td, *J* = 7.3, 5.5 Hz, 2H), 7.46 – 7.31 (m, 8H), 7.21 – 7.10 (m, 5H), 5.77 (t, *J* = 9.3 Hz, 1H), 5.43 (t, *J* = 9.8 Hz, 1H), 4.72 (d, *J* = 10.0 Hz, 1H), 4.64 – 4.51 (m, 3H), 4.50 – 4.33 (m, 3H), 4.29 (t, *J* = 7.4 Hz, 1H), 3.92 (t, *J* = 9.4 Hz, 1H), 3.81 (ddd, *J* = 9.9, 4.9, 2.1 Hz, 1H), 2.83 – 2.65 (m, 2H), 1.25 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 165.77, 165.53, 155.04, 143.50, 143.39, 141.45, 136.91, 133.44, 133.37, 130.01, 129.88, 129.40, 129.34, 128.59, 128.58, 128.48, 128.43, 128.25, 128.08, 127.36, 127.34, 125.36, 125.30, 120.24, 83.86, 77.48, 77.26, 76.43, 75.77, 75.00, 70.78, 70.22, 66.45, 46.85, 24.52, 15.06. $[\alpha]_D^{20}$ -44.69 (c 1.03 g/100 mL, $CHCl_3$). IR ν = 1727, 1251, 1088, 1069, 741, 708 cm^{-1}. (ESI-HRMS) m/z 767.232 $[M+Na]^+$ ($C_{44}H_{40}O_9SNa$ requires 767.229).

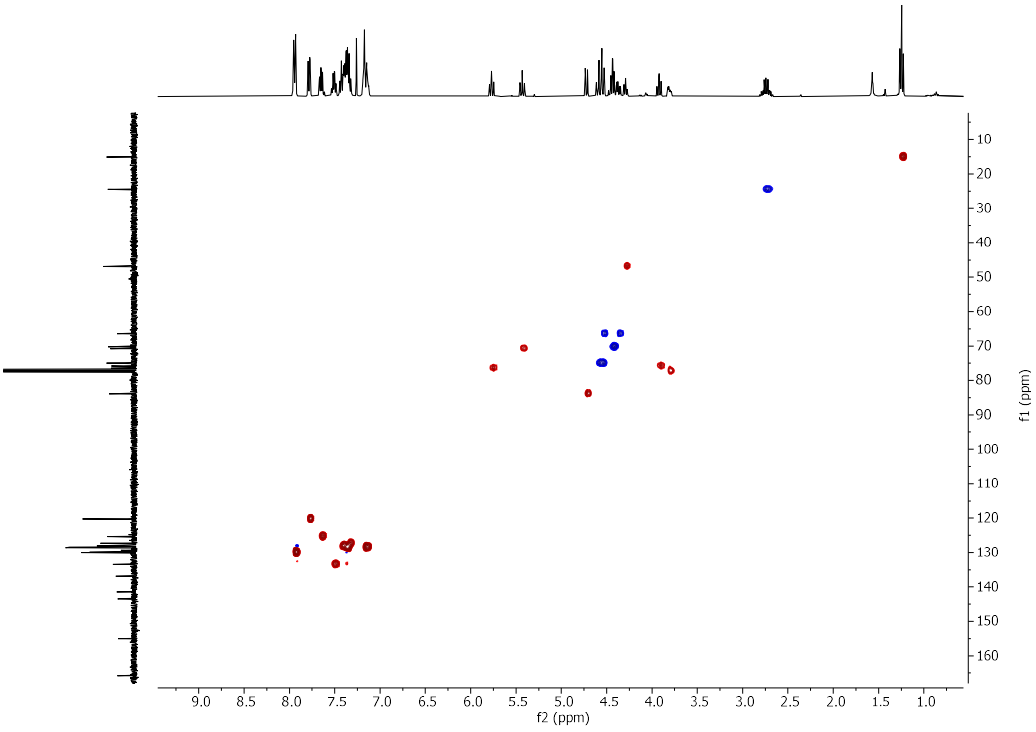
^{1}H NMR of BB3.1 (400 MHz, $CDCl_3$)



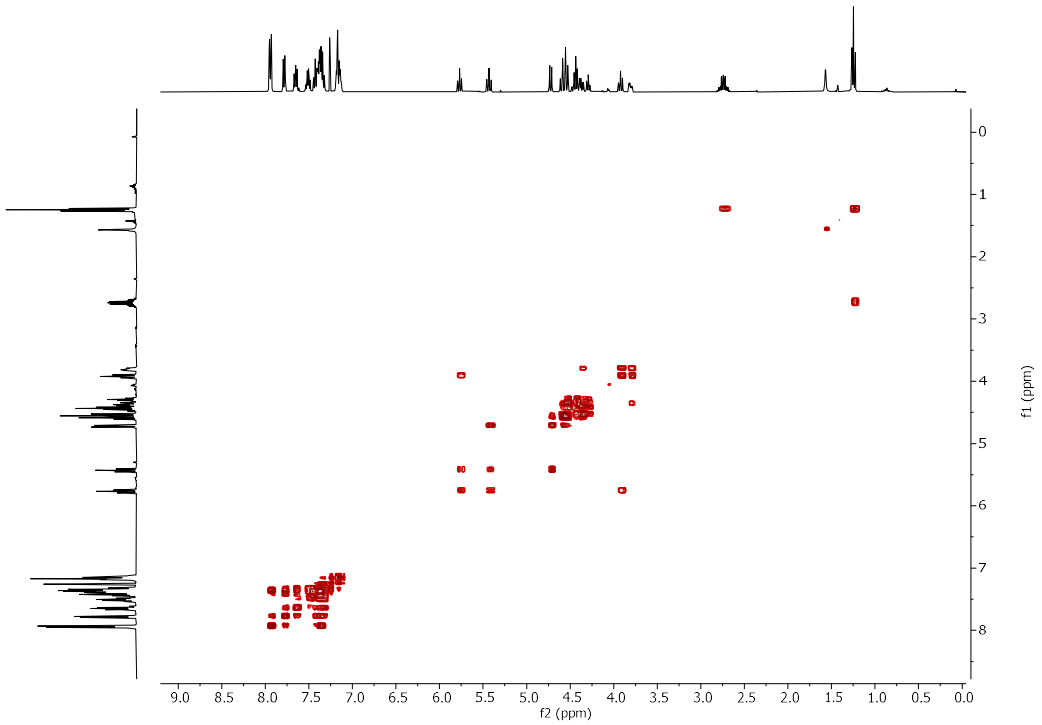
^{13}C NMR of BB3.1 (101 MHz, $CDCl_3$)



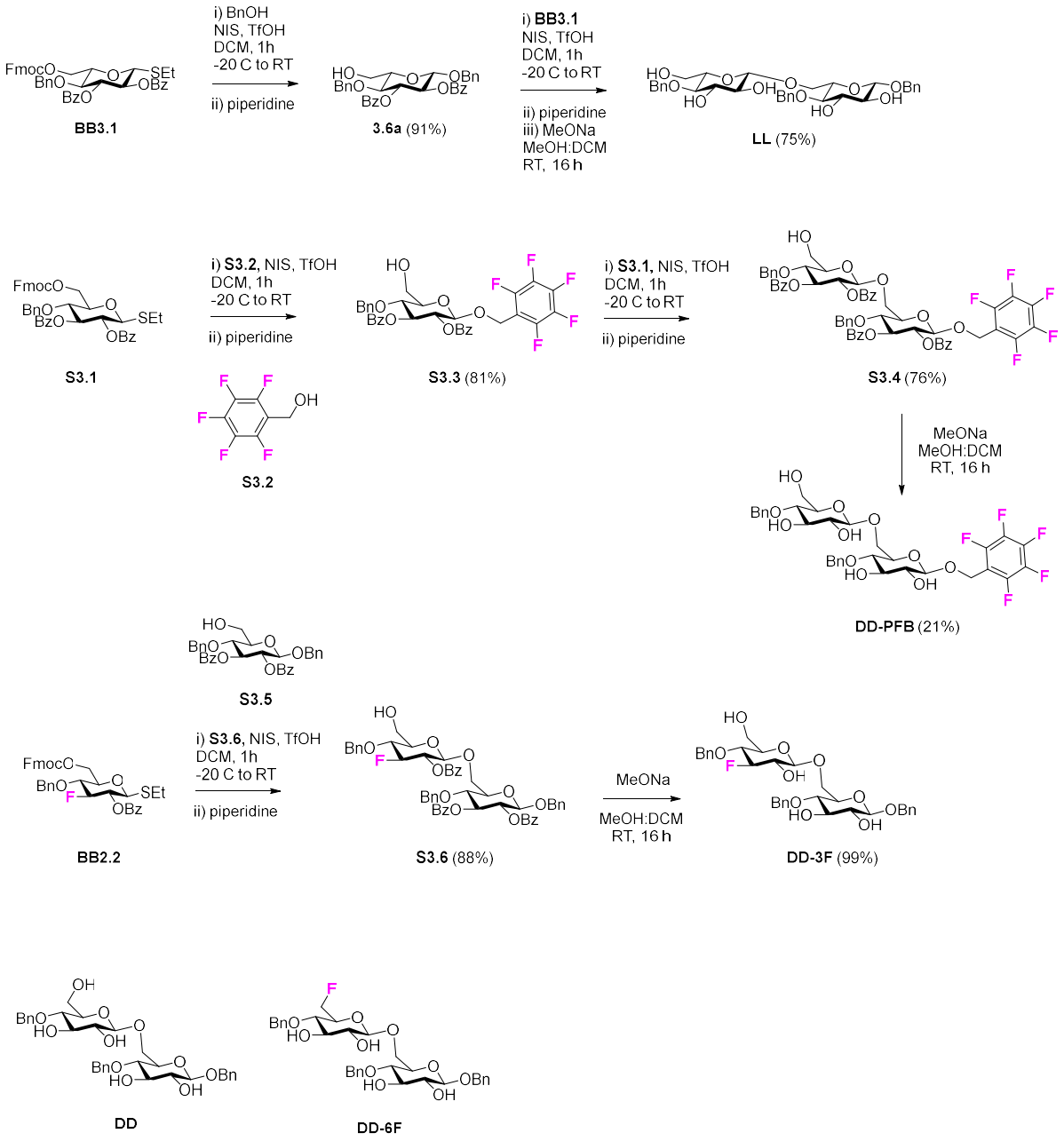
HSQC NMR of BB3.1 ($CDCl_3$)



COSY NMR of BB3.1 ($CDCl_3$)

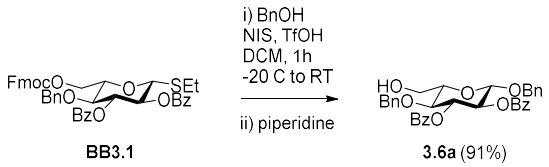


6.4.2 Synthesis of dimers



Scheme S3.2 Synthesis of dimers **LL**, **DD-PFB**, and **DD-3F**. Compounds **DD** and **DD-6F** were prepared by Dr. Yang Yu.[358] The synthesis of **BB2.2** is described in Chapter 2.

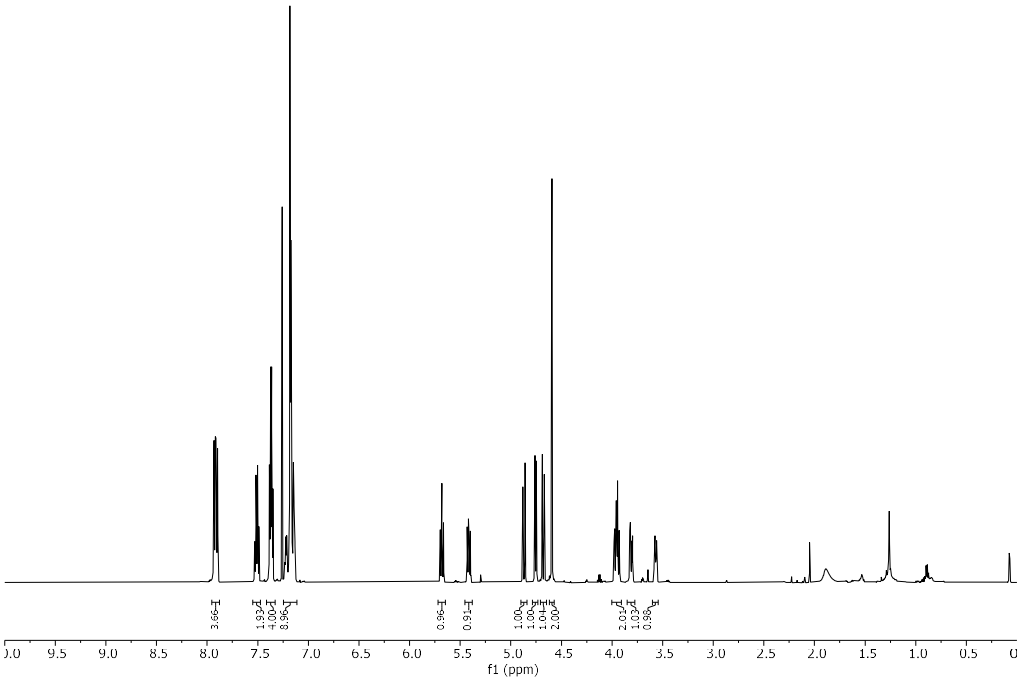
Synthesis of **3.6a**



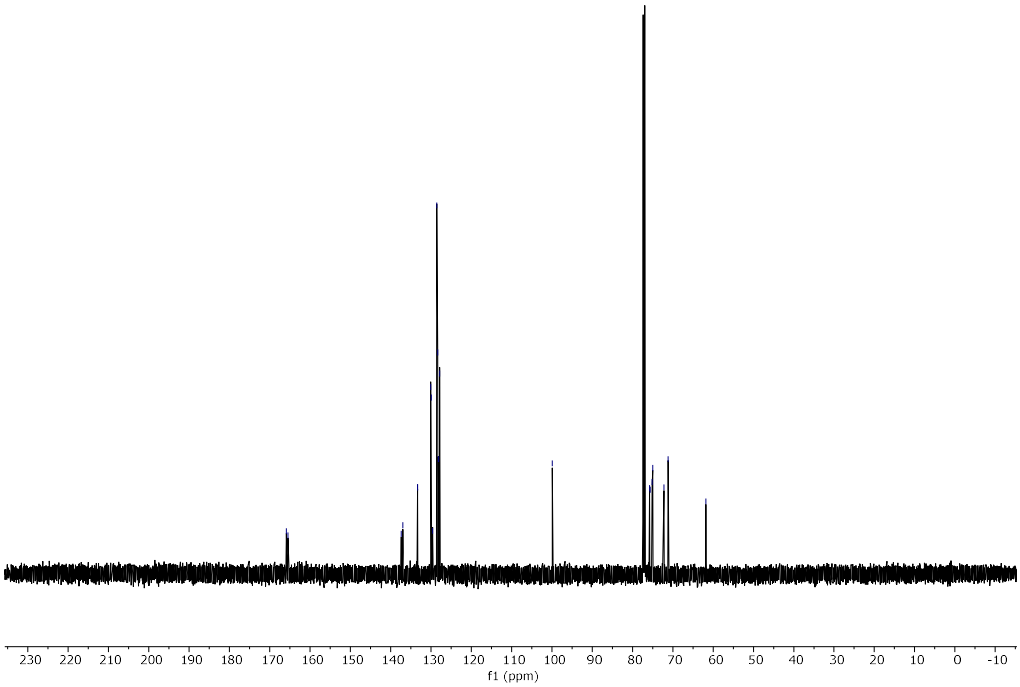
BB3.1 (60 mg, 0.08 mmol) and benzyl alcohol (BnOH) (17 µL, 0.16 mmol) were dissolved in anhydrous DCM (800 µL). The solution was then stirred with molecular sieves (4 Å) for 1 h at RT under Ar atmosphere and then cooled to -20 °C. The activator solution (300 µL of a 200 mg/mL solution of NIS in DCM : Dioxane 2:1 with 1% TfOH) was added dropwise and the reaction was stirred for 5 min at -20 °C, after which time the cooling bath was removed and the reaction allowed to RT. After 1 h, ESI-MS indicated the disappearance of **BB3.1**. Piperidine (0.5 mL) was added and the reaction was stirred at RT for an additional 1h. The reaction was diluted with DCM and washed once with a 1 M aqueous solution of HCl, once with saturated aqueous solution of $NaHCO_3$, and once with brine. The crude compound was purified by column chromatography (Hexane : EtOAc = 3:1→2:1→1:1) to give **3.6a** as a white solid (42 mg, 91%).

^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.92 (ddt, J = 12.7, 7.0, 1.5 Hz, 3H), 7.55 – 7.47 (m, 2H), 7.41 – 7.33 (m, 4H), 7.25 – 7.11 (m, 9H), 5.68 (t, J = 9.6 Hz, 1H), 5.45 – 5.38 (m, 1H), 4.87 (d, J = 12.6 Hz, 1H), 4.76 (d, J = 7.9 Hz, 1H), 4.68 (dd, J = 12.7, 1.7 Hz, 1H), 4.59 (s, 2H), 4.00 – 3.91 (m, 2H), 3.81 (dd, J = 12.1, 4.0 Hz, 1H), 3.57 (ddd, J = 9.7, 4.0, 2.5 Hz, 1H). ^{13}C NMR (151 MHz, Chloroform-*d*) δ 165.83, 165.47, 137.33, 136.95, 133.33, 133.27, 130.02, 129.91, 129.56, 128.53, 128.34, 128.13, 128.05, 127.84, 99.86, 75.73, 75.60, 75.12, 74.95, 72.19, 71.13, 61.78. $[\alpha]_D^{20}$ -31.39 (c 1.07 g/100 mL, $CHCl_3$). IR ν = 3447, 2927, 1725, 1272, 1094, 1069, 1027, 708, 699 cm^{-1}. (ESI-HRMS) m/z 591.200 $[M+Na]^+$ ($C_{34}H_{32}O_8Na$ requires 591.199).

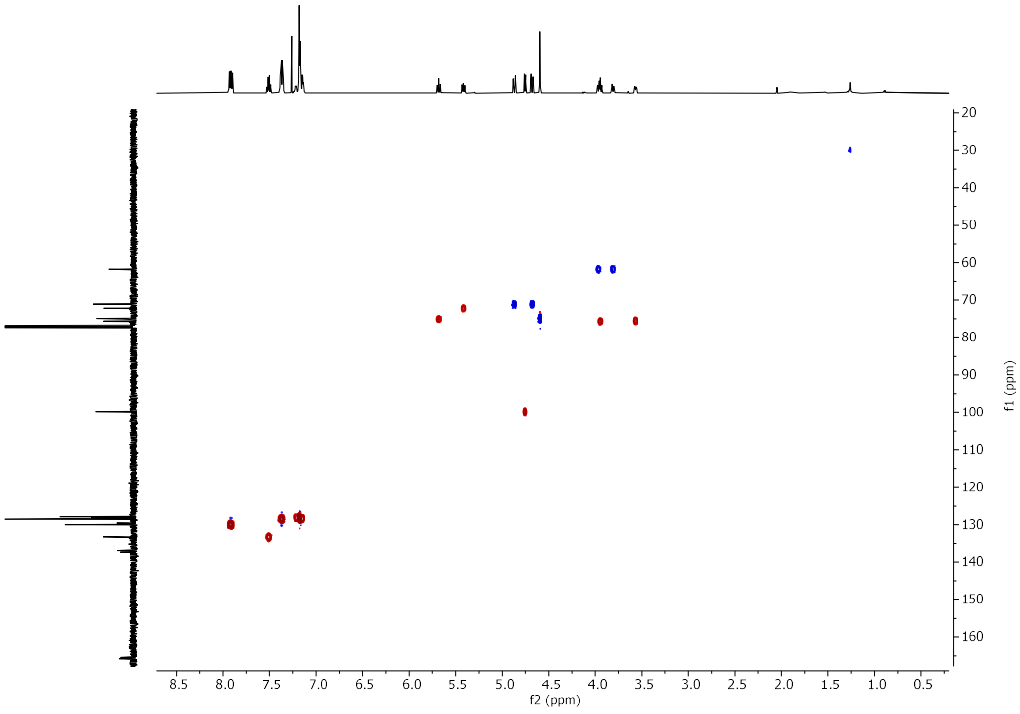
^{1}H NMR of 3.6a (600 MHz, $CDCl_3$)



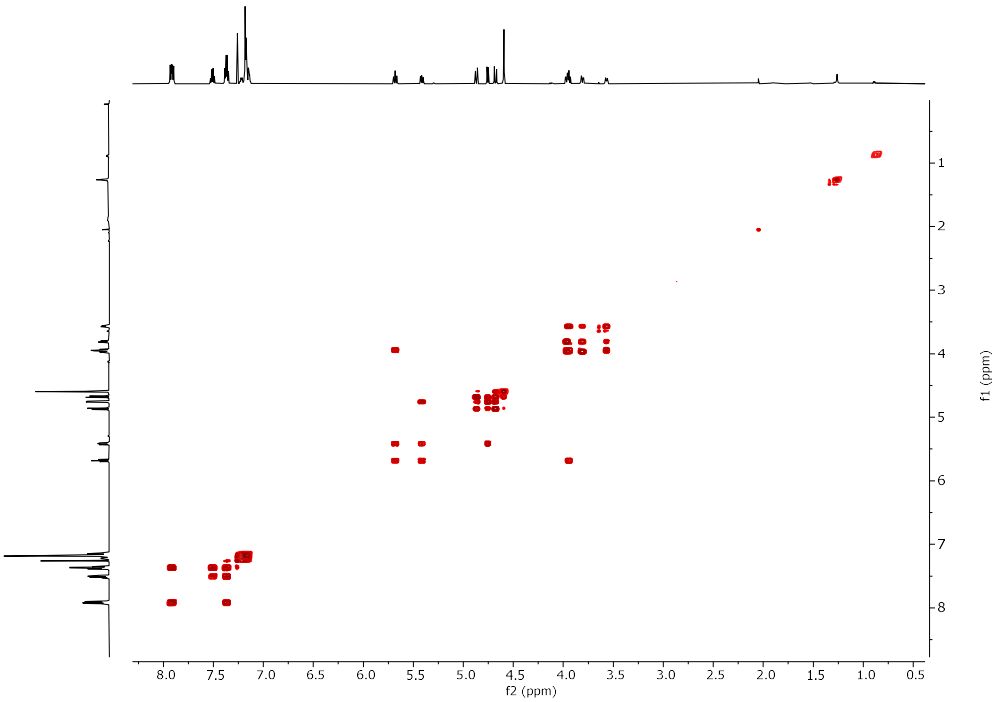
^{13}C NMR of 3.6a (151 MHz, $CDCl_3$)



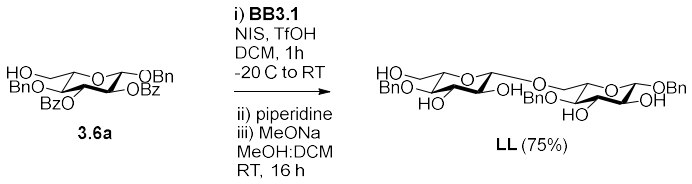
HSQC NMR of 3.6a ($CDCl_3$)



COSY NMR of 3.6a ($CDCl_3$)



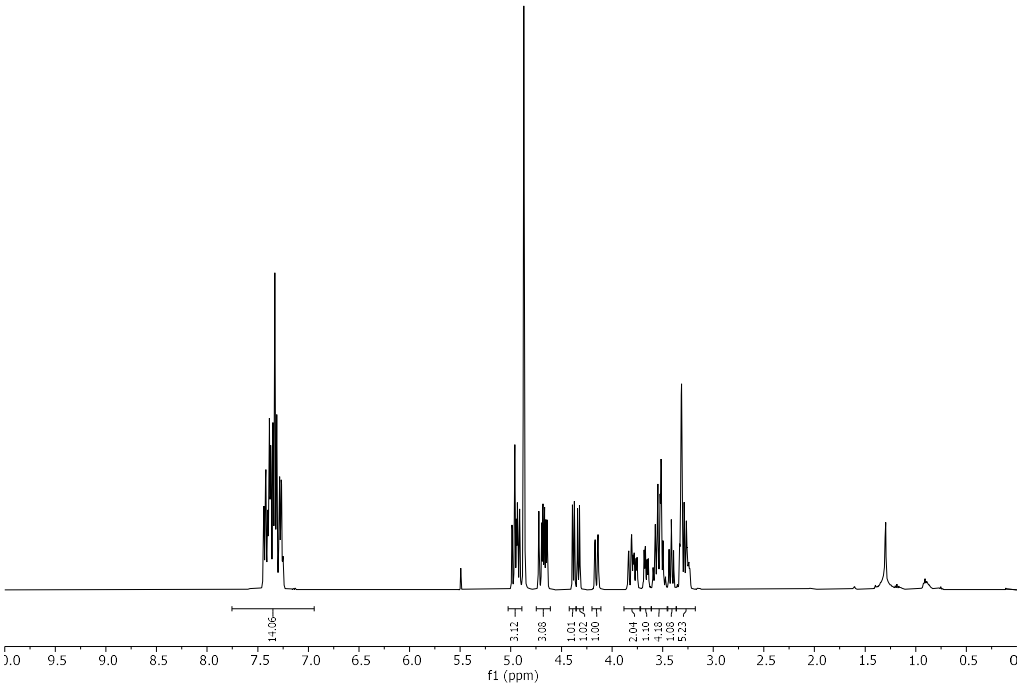
Synthesis of **LL**



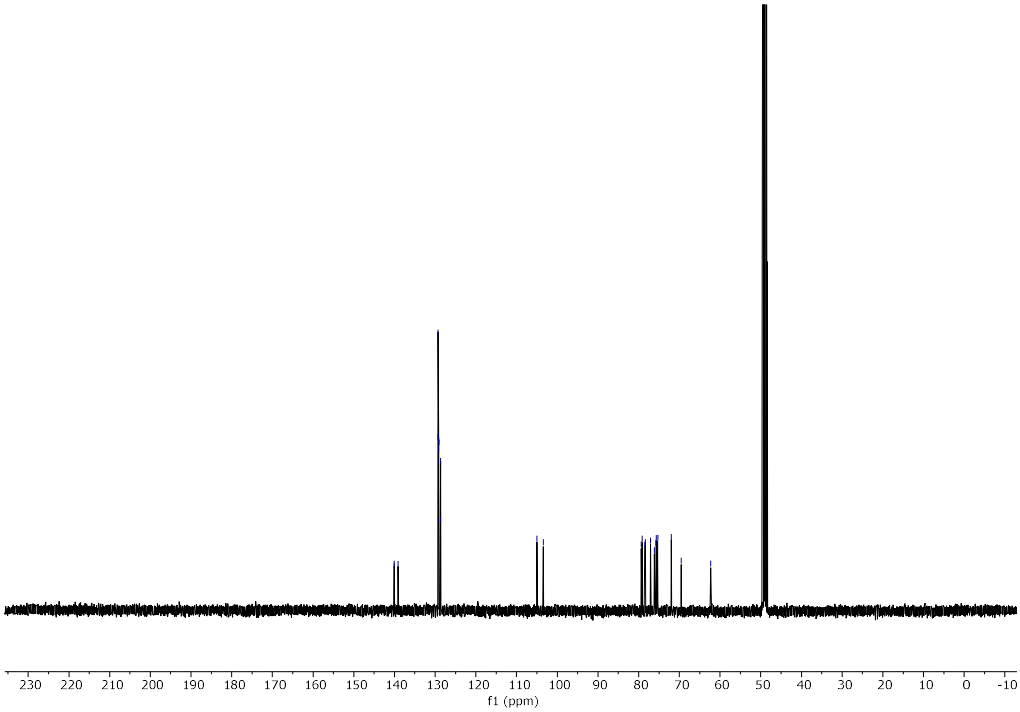
BB3.1 (69 mg, 0.093 mmol) and **3.6a** (42 mg, 0.074 mmol) were dissolved in anhydrous DCM (1 mL). The solution was stirred for 1 h with molecular sieves (4 Å) at RT under Ar atmosphere and then cooled to -20 °C. The activator solution (200 µL of a 200 mg/mL solution of NIS in DCM : Dioxane 2:1 with 1% TfOH) was added dropwise and the reaction was stirred for 5 min at -20 °C, after which time the cooling bath was removed and the reaction allowed to RT. After 1 h, ESI-MS indicated the disappearance of **3.6a**. Piperidine (0.5 mL) was added and the reaction was stirred at RT for an additional 1 h. The reaction was diluted with DCM and washed once with a 1 M aqueous solution of HCl, once with a saturated aqueous solution of $NaHCO_3$ and once with brine. The crude compound was purified by silica gel flash column chromatography (Hexane : EtOAc = 3:1→2:1→1:1→2:3). The fully protected compound (52 mg, 0.051 mmol) was dissolved in MeOH : DCM (4 mL, 1:1). MeONa in MeOH (0.5 M, 3 equiv. per benzoyl ester) was added and the solution was stirred at RT for 16 h, neutralized with Amberlite IR-120 (H^+ form) resin, filtered and concentrated *in vacuo*. The resulting yellow oil was purified by column chromatography (DCM : MeOH = 15:1) and recrystallized from DCM : Hexane to give **LL** as white solid (31 mg, 75% over 3 steps).

^{1}H NMR (400 MHz, Methanol-d_4) δ 7.51 – 7.16 (m, 15H), 5.01 – 4.89 (m, 3H), 4.75 – 4.61 (m, 3H), 4.38 (d, J = 7.8 Hz, 1H), 4.33 (d, J = 7.8 Hz, 1H), 4.15 (d, J = 10.9 Hz, 1H), 3.88 – 3.72 (m, 2H), 3.72 – 3.62 (m, 1H), 3.62 – 3.45 (m, 4H), 3.41 (t, J = 9.3 Hz, 1H), 3.36 – 3.18 (m, 3H). ^{13}C NMR (101 MHz, Methanol-d_4) δ 140.10, 140.04, 139.10, 129.31, 129.28, 129.16, 129.13, 129.08, 128.69, 128.65, 105.00, 103.42, 79.34, 79.12, 78.51, 78.37, 77.05, 76.17, 75.73, 75.67, 75.37, 75.31, 71.97, 69.56, 62.32. (ESI-HRMS) m/z 635.245 $[M+Na]^+$ ($C_{33}H_{40}O_{11}Na$ requires 635.246).

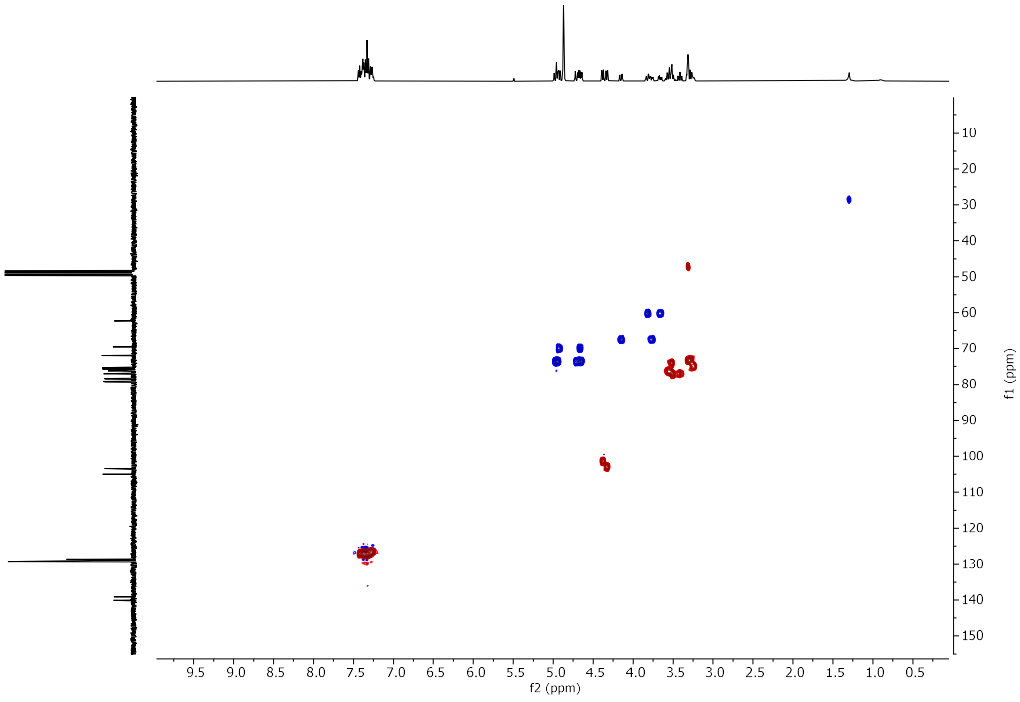
^{1}H NMR of LL (400 MHz, CD_3OD)



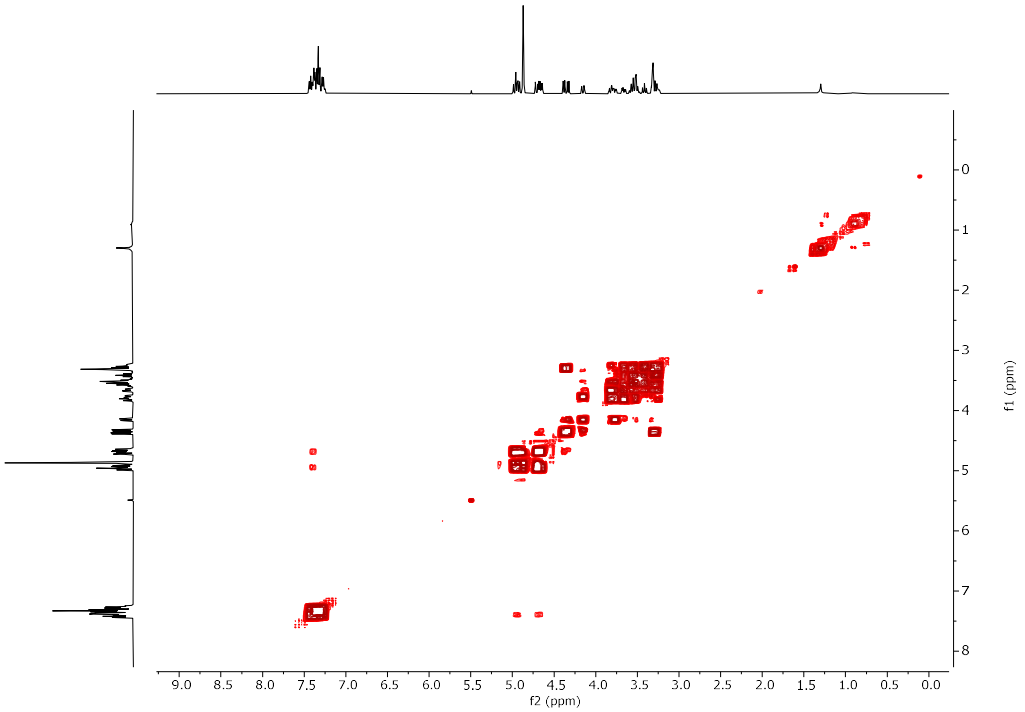
^{13}C NMR of LL (101 MHz, CD_3OD)



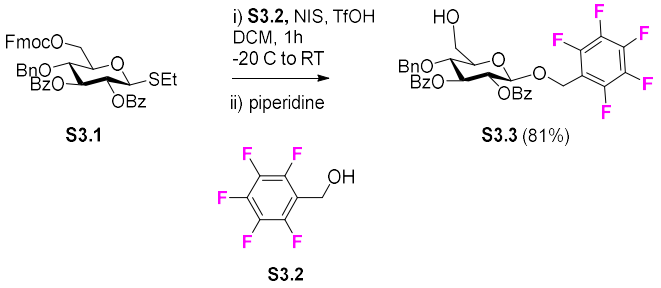
HSQC NMR of LL (400 MHz, CD_3OD)



COSY NMR of LL (400 MHz, CD_3OD)



Synthesis of **S3.3**

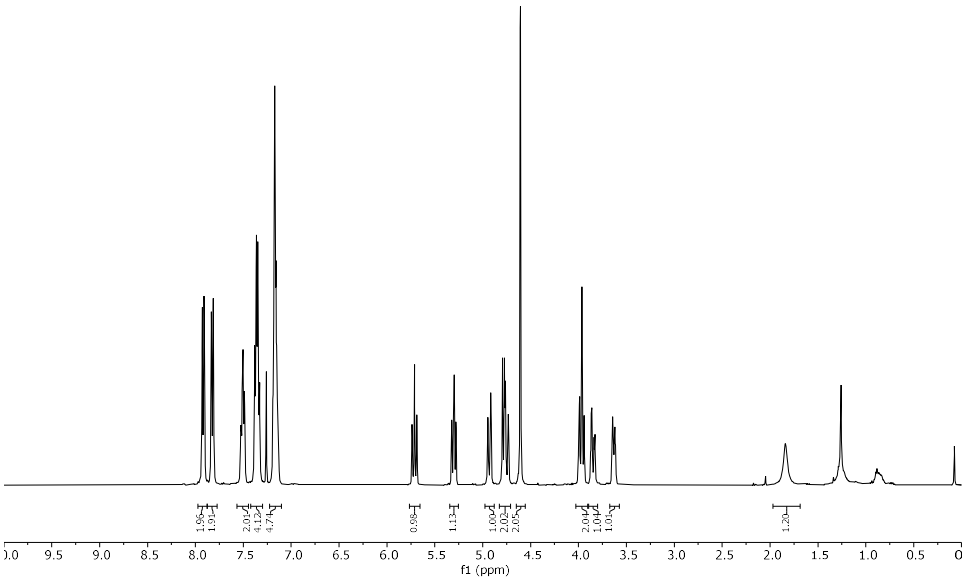


Compound **S3.1** was synthesized according to a previously reported protocol.[83]

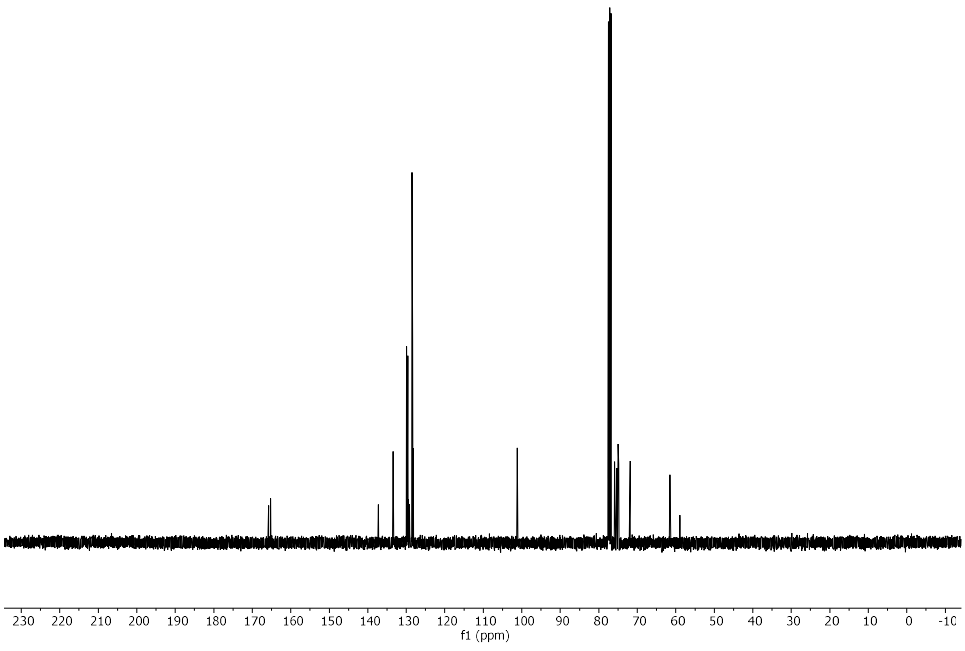
S3.1 (70 mg, 0.094 mmol) and 2,3,4,5,6-pentafluorobenzyl alcohol **S3.2** (40 mg, 0.20 mmol) were dissolved in anhydrous DCM (1 mL). The solution was stirred with molecular sieves (4 Å) for 1 h at RT under Ar atmosphere and then cooled to -20 °C. The activator solution (250 µL of a 200 mg/mL solution of NIS in DCM : Dioxane 2:1 with 1% TfOH) was added dropwise and the reaction was stirred for 5 min at -20 °C, after which time the cooling bath was removed to allow the reaction to RT. After 1 h at RT, ESI-MS indicated the disappearance of **S3.1**. Piperidine (0.5 mL) was added and the reaction was stirred at RT for additional 1 h. The reaction was diluted with DCM and washed once with a 1 M aqueous solution of HCl, once with a saturated aqueous solution of $NaHCO_3$, and once with brine. The crude was purified by silica gel flash column chromatography (Hexane : EtOAc = 3:1→2:1→1:1) to give **S3.3** as a white foam (53 mg, 81%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.95 – 7.89 (m, 2H), 7.86 – 7.79 (m, 2H), 7.55 – 7.46 (m, 2H), 7.40 – 7.31 (m, 4H), 7.21 – 7.12 (m, 5H), 5.71 (t, *J* = 9.6 Hz, 1H), 5.35 – 5.26 (m, 1H), 4.93 (d, *J* = 12.0 Hz, 1H), 4.78 (d, *J* = 8.0 Hz, 1H), 4.75 (d, *J* = 12.0 Hz, 1H), 4.01 – 3.92 (m, 2H), 3.85 (dd, *J* = 12.2, 3.6 Hz, 1H), 3.63 (dt, *J* = 9.6, 3.0 Hz, 1H), 1.84 (s, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.77, 165.22, 137.22, 133.39, 133.36, 129.89, 129.62, 129.41, 129.18, 128.54, 128.40, 128.34, 128.16, 101.17, 75.92, 75.32, 74.99, 74.83, 71.87, 61.54, 58.92. ^{19}F NMR (376 MHz, $CDCl_3$) δ -142.48 (dd, *J* = 22.2, 8.5 Hz, 2F), -152.63 (t, *J* = 20.8 Hz, 1F), -161.49 (td, *J* = 21.7, 8.5 Hz, 2F). $[\alpha]_D^{20}$ 11.41 (c 1.06 g/100 mL, $CHCl_3$). IR ν = 2927, 1729, 1507, 1273, 1096, 1071 cm^{-1}. (ESI-HRMS) m/z 681.1479 $[M+Na]^+$ ($C_{34}H_{27}F_5O_8Na$ requires 681.1518).

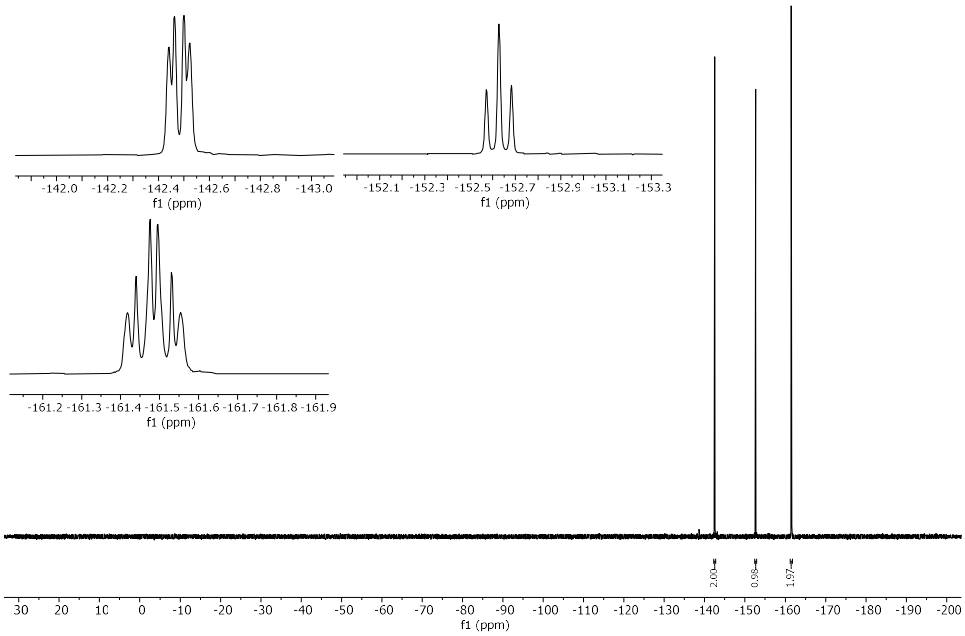
^{1}H NMR of S3.3 (400 MHz, $CDCl_3$)



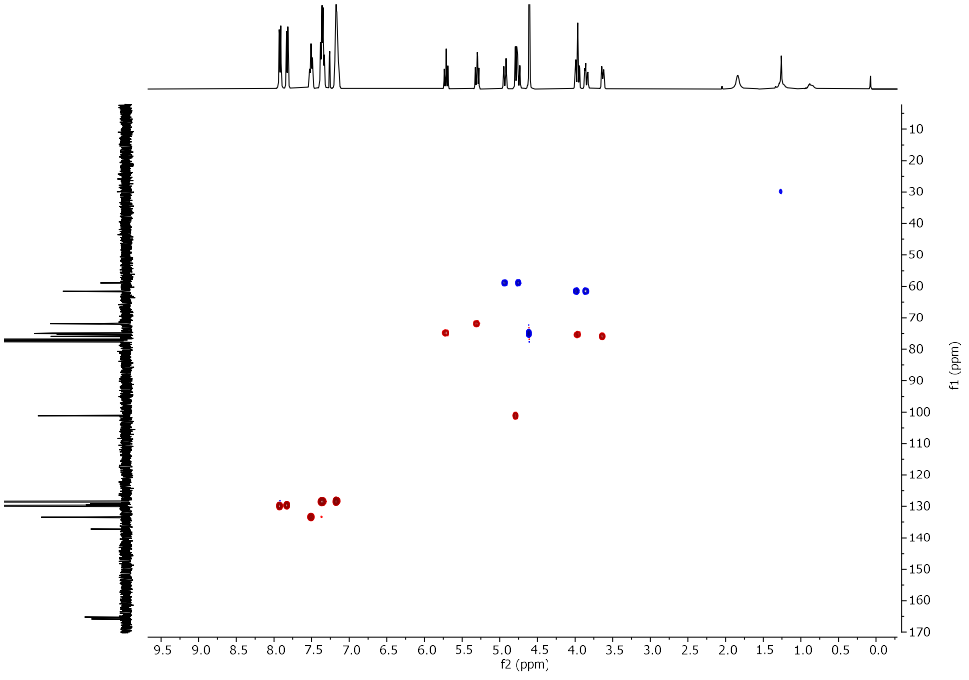
^{13}C NMR of S3.3 (101 MHz, $CDCl_3$)



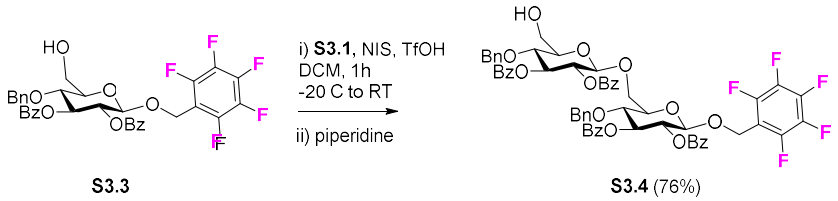
^{19}F NMR of S3.3 (376 MHz, $CDCl_3$)



HSQC NMR of S3.3 ($CDCl_3$)



Synthesis of **S3.4**

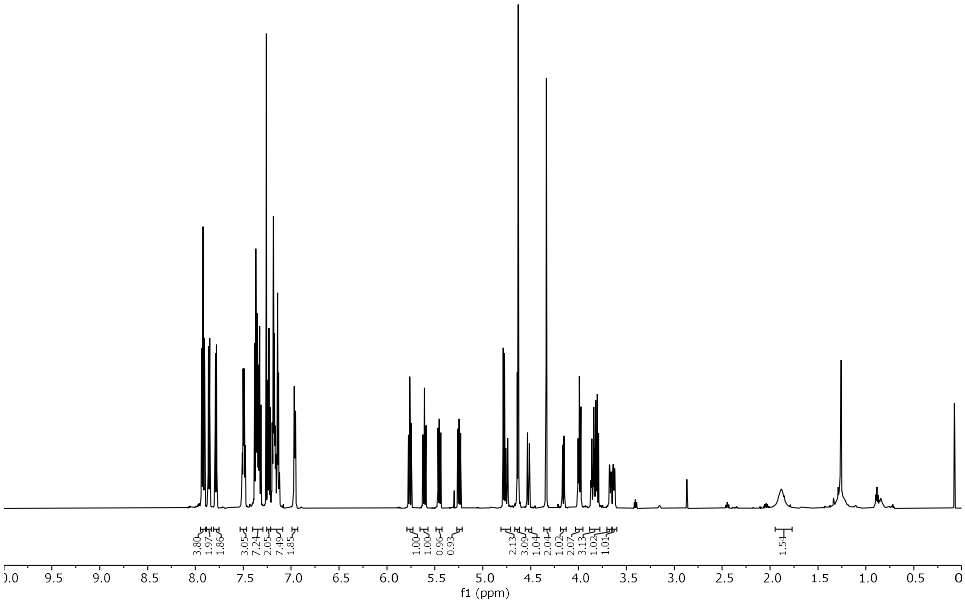


S3.3 **S3.4** (76%)

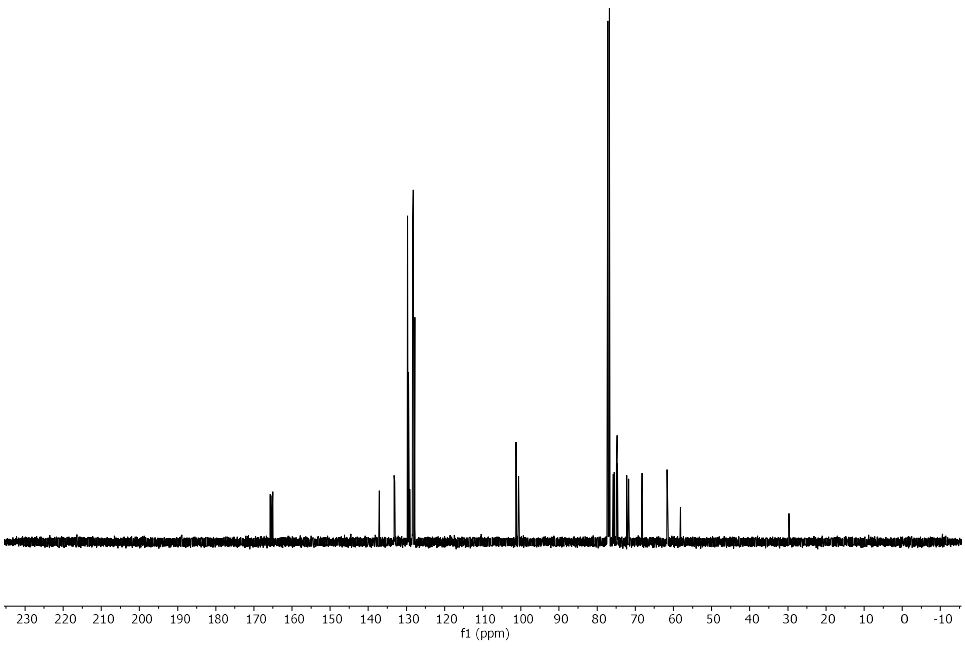
S3.3 (53 mg, 0.080 mmol) and **S3.1** (72 mg, 0.097 mmol) were dissolved in anhydrous DCM (1 mL). The solution was stirred with molecular sieves (4 Å) for 1 h at RT under Ar atmosphere and then cooled to -20 °C. The activator solution (250 µL of a 200 mg/mL solution of NIS in DCM : Dioxane 2:1 with 1% TfOH) was added dropwise and the reaction was stirred for 5 min at -20 °C, after which time the cooling bath was removed to allow the reaction to RT. After 1 h at RT, ESI-MS indicated the disappearance of **S3.3**. Piperidine (0.5 mL) was added and the reaction was stirred at RT for additional 1 h. The reaction was diluted with DCM and washed once with a 1 M aqueous solution of HCl, once with a saturated aqueous solution of $NaHCO_3$, and once with brine. The crude product was purified by column chromatography (Hexane : Acetone = 4:1→3:1→2:1) to give **S3.4** as a white solid (68 mg, 76 %).

^{1}H NMR (600 MHz, $CDCl_3$) δ 7.95 – 7.89 (m, 4H), 7.88 – 7.83 (m, 2H), 7.82 – 7.77 (m, 2H), 7.54 – 7.46 (m, 3H), 7.41 – 7.30 (m, 7H), 7.25 – 7.21 (m, 2H), 7.21 – 7.10 (m, 8H), 6.99 – 6.94 (m, 2H), 5.76 (t, *J* = 9.5 Hz, 1H), 5.61 (t, *J* = 9.4 Hz, 1H), 5.45 (dd, *J* = 9.8, 7.8 Hz, 1H), 5.24 (dd, *J* = 9.8, 7.8 Hz, 1H), 4.78 (d, *J* = 7.8 Hz, 1H), 4.75 (d, *J* = 12.0 Hz, 1H), 4.66 – 4.61 (m, 3H), 4.52 (dt, *J* = 12.1, 1.4 Hz, 1H), 4.34 (s, 2H), 4.16 (dd, *J* = 11.2, 1.7 Hz, 1H), 4.02 – 3.96 (m, 2H), 3.89 – 3.78 (m, 3H), 3.67 (ddd, *J* = 9.8, 4.8, 1.7 Hz, 1H), 3.63 (ddd, *J* = 9.6, 4.2, 2.6 Hz, 1H), 1.88 (s, 1H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 165.71, 165.51, 165.16, 165.00, 137.17, 137.12, 129.71, 129.66, 129.44, 129.33, 129.25, 129.11, 128.40, 128.33, 128.31, 128.25, 128.20, 127.99, 127.82, 127.80, 101.28, 100.62, 75.78, 75.62, 75.53, 74.93, 74.91, 74.82, 74.79, 74.68, 72.21, 71.70, 68.24, 61.64, 58.16. ^{19}F NMR (564 MHz, $CDCl_3$) δ -142.26 (dd, *J* = 22.6, 8.6 Hz, 2F), -153.08 (t, *J* = 20.8 Hz, 1F), -161.70 (td, *J* = 22.0, 8.5 Hz, 2F). (ESI-HRMS) m/z 1141.276 $[M+Na]^+$ ($C_{61}H_{51}F_5O_{15}Na$ requires 1141.304).

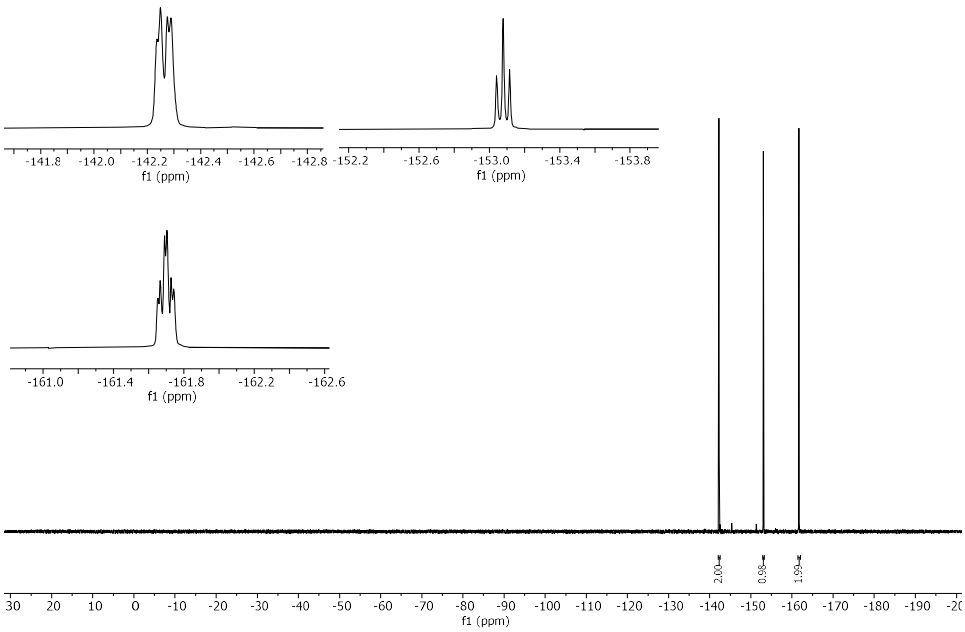
^{1}H NMR of S3.4 (600 MHz, $CDCl_3$)



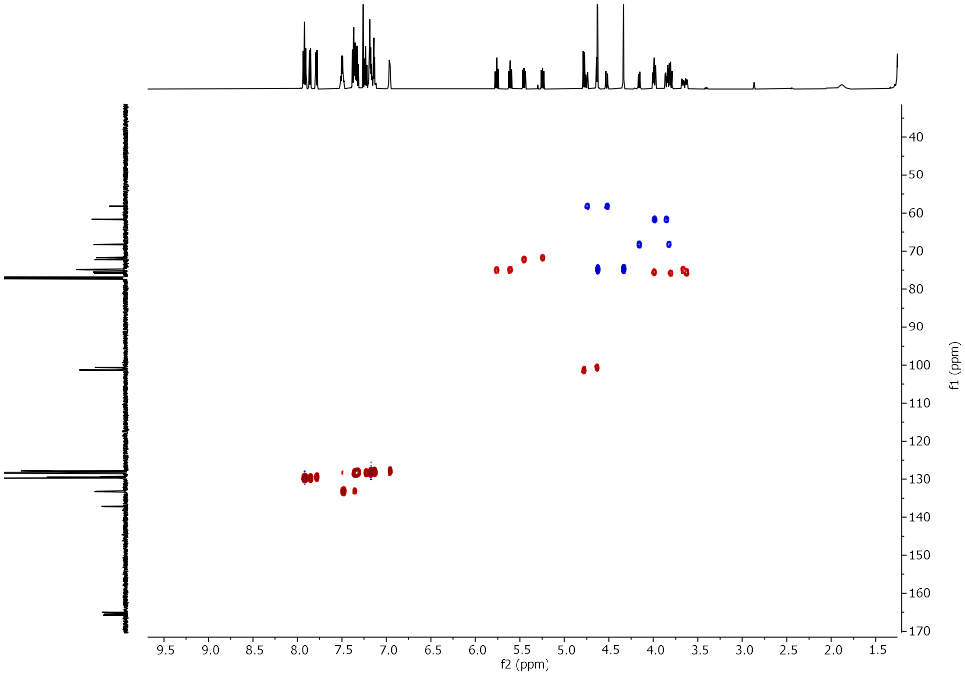
^{13}C NMR of S3.4 (151 MHz, $CDCl_3$)



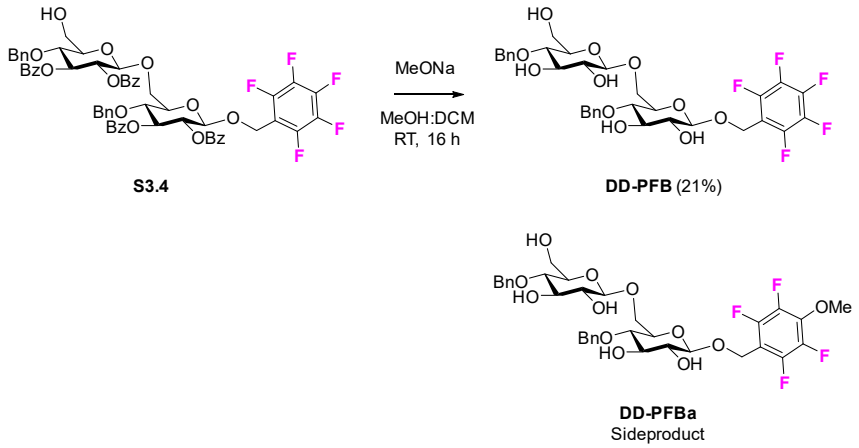
^{19}F NMR of S3.4 (564 MHz, $CDCl_3$)



HSQC NMR of S3.4 ($CDCl_3$)



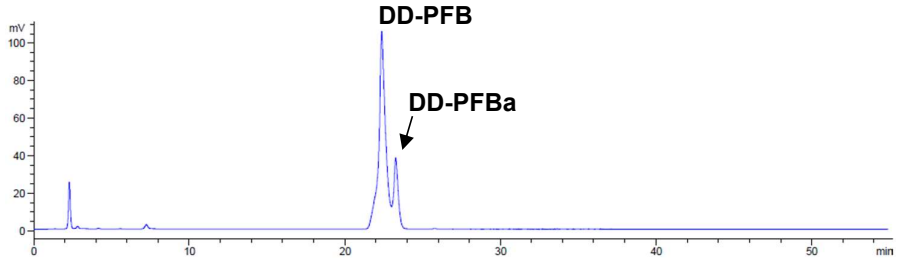
Synthesis of **DD-PFB**



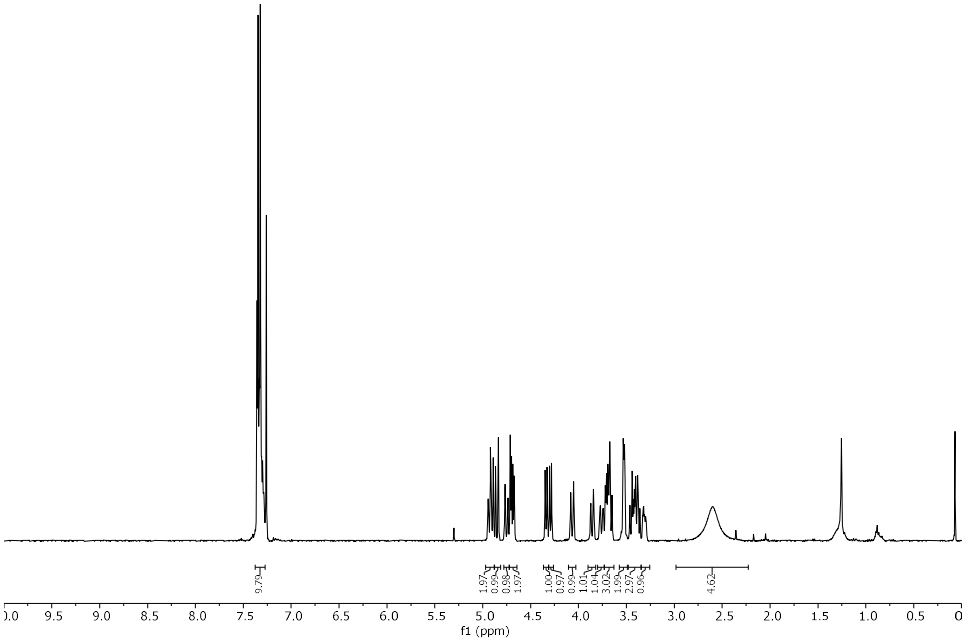
S3.a (62 mg, 0.061 mmol) was dissolved in a 1:1 mixture of MeOH : DCM (3 mL). MeONa in MeOH (0.5 M, 3 equiv. per benzoyl ester) was added and the solution was stirred at RT for 16 h, neutralized with Amberlite IR-120 (H^+ form) resin, filtered, and concentrated *in vacuo*. ^{19}F NMR and ESI-MS analysis of the crude revealed the presence of side product **DD-PFBa** (suggested in the scheme above) lacking the *para* fluorine (^{19}F NMR (564 MHz, CD_3OD) δ -146.61 (dd, J = 20.8, 8.9 Hz, 2F), -160.76 (dd, J = 21.0, 8.9 Hz, 2F) and (ESI-HRMS) m/z 737.2225 $[M+Na]^+$ ($C_{34}H_{38}F_4O_{12}Na$ requires 737.2197)). This side product **DD-PFBa** was inseparable from **DD-PFB** by TLC or silica gel flash column chromatography. The resulting yellow oil was passed through a short plug of silica gel (DCM : MeOH = 15:1), precipitated from DCM : Hexane and further purified with normal phase HPLC (Method: YMC Diol, isocratic 30% EtOAc (5 min), linear gradient to 100% EtOAc (30 min), isocratic 100% EtOAc (5min)) to yield **DD-PFB** as a colorless solid (9 mg, 21% yield).

^1H NMR (400 MHz, $CDCl_3$) δ 7.39 – 7.27 (m, 10H), 4.91 (dd, J = 11.6, 9.1 Hz, 2H), 4.85 (d, J = 11.3 Hz, 1H), 4.75 (d, J = 11.7 Hz, 1H), 4.69 (dd, J = 11.4, 5.2 Hz, 2H), 4.34 (d, J = 7.7 Hz, 1H), 4.29 (d, J = 7.8 Hz, 1H), 4.06 (d, J = 11.7 Hz, 1H), 3.86 (dd, J = 12.0, 2.6 Hz, 1H), 3.80 – 3.73 (m, 1H), 3.73 – 3.64 (m, 3H), 3.52 (d, J = 5.0 Hz, 2H), 3.48 – 3.34 (m, 3H), 3.34 – 3.27 (m, 1H).^{13}C NMR (101 MHz, $CDCl_3$) δ 138.16, 128.73, 128.71, 128.24, 128.19, 128.11, 103.38, 102.10, 77.27, 76.50, 75.52, 75.08, 74.84, 74.74, 73.71, 73.66, 68.27, 62.14, 58.27.^{19}F NMR (376 MHz, $CDCl_3$) δ -142.51 (dd, J = 22.5, 8.6 Hz, 2F), -152.60 (t, J = 20.8 Hz, 1F), -161.38 (td, J = 22.1, 8.5 Hz, 2F). (ESI-HRMS) m/z 725.2022 $[M+Na]^+$ ($C_{33}H_{35}F_5O_{11}Na$ requires 725.1992).

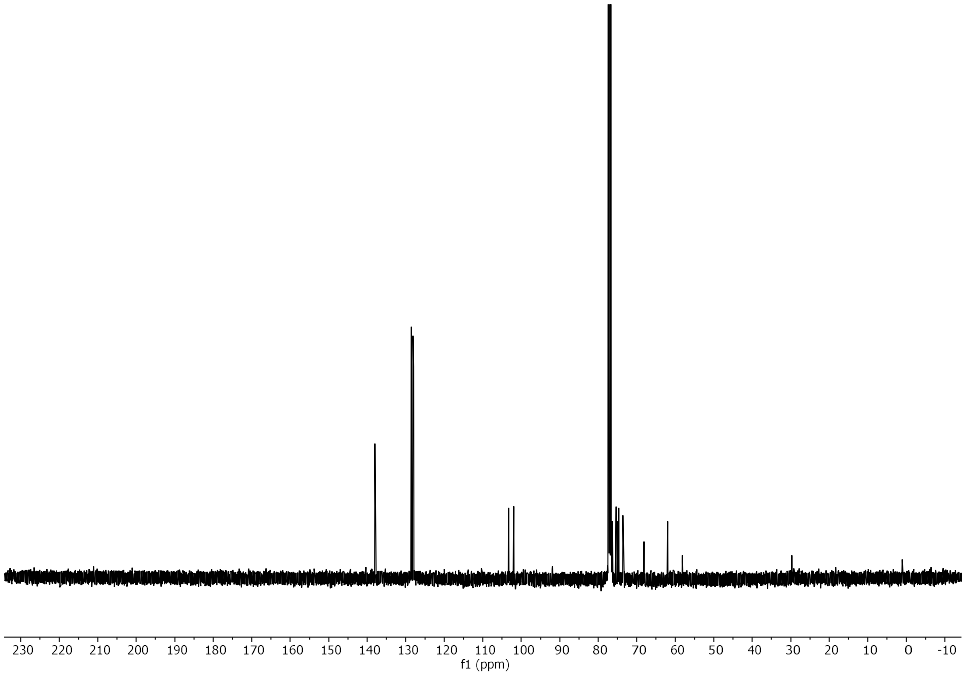
NP HPLC of crude reaction mixture (ELSD trace, t_R (6) = 22.4 min 6, t_R (6a) = 23.3 min)



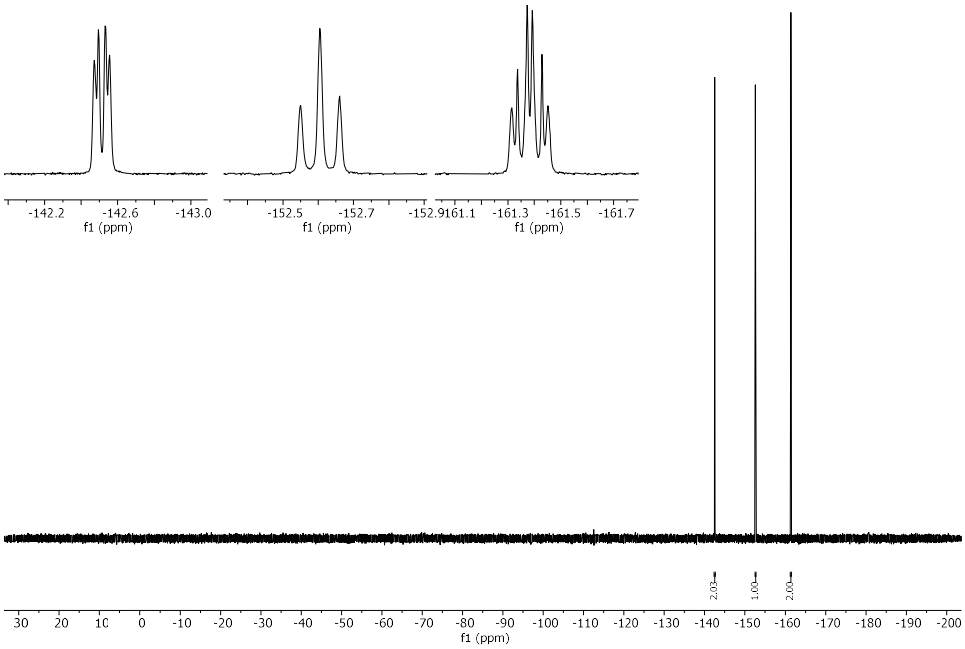
^{1}H NMR of DD-PFB (400 MHz, $CDCl_3$)



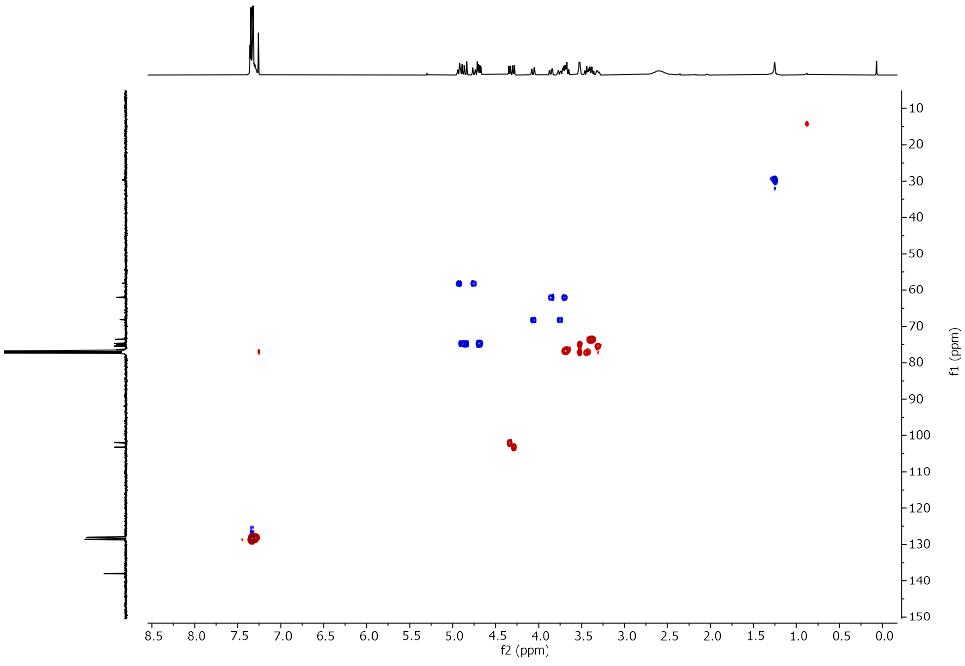
^{13}C NMR of DD-PFB (101 MHz, $CDCl_3$)



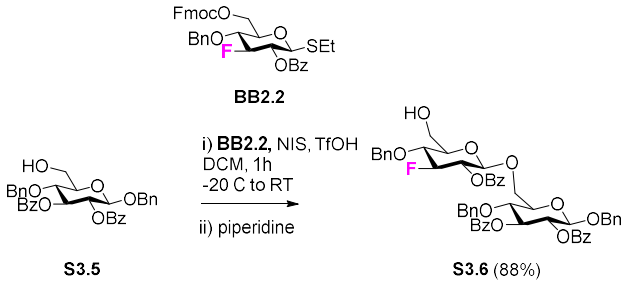
^{19}F NMR of DD-PFB (376 MHz, $CDCl_3$)



HSQC NMR of DD-PFB ($CDCl_3$)



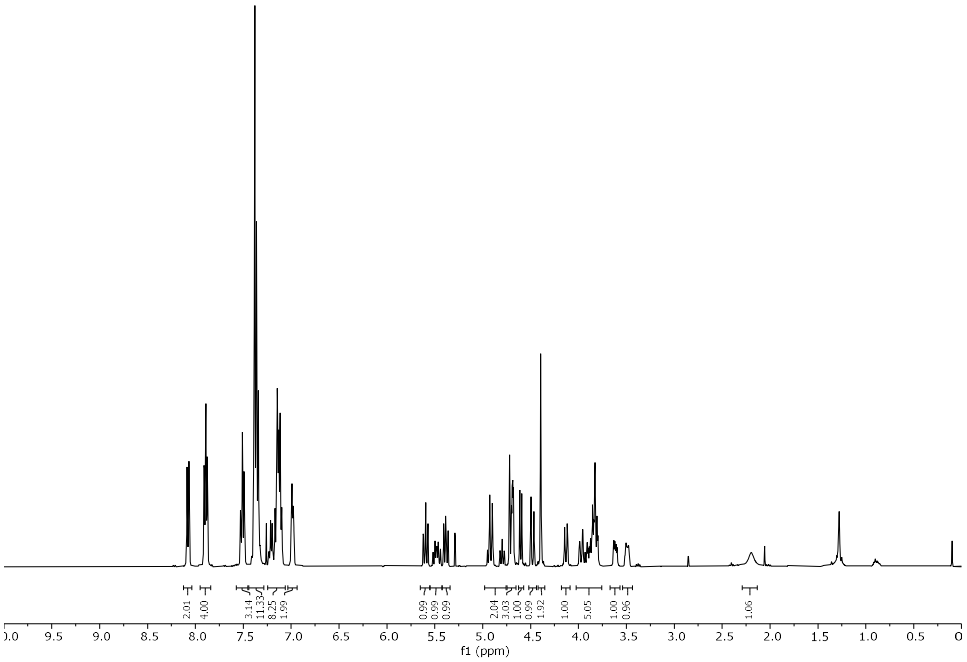
Synthesis of **S3.6**



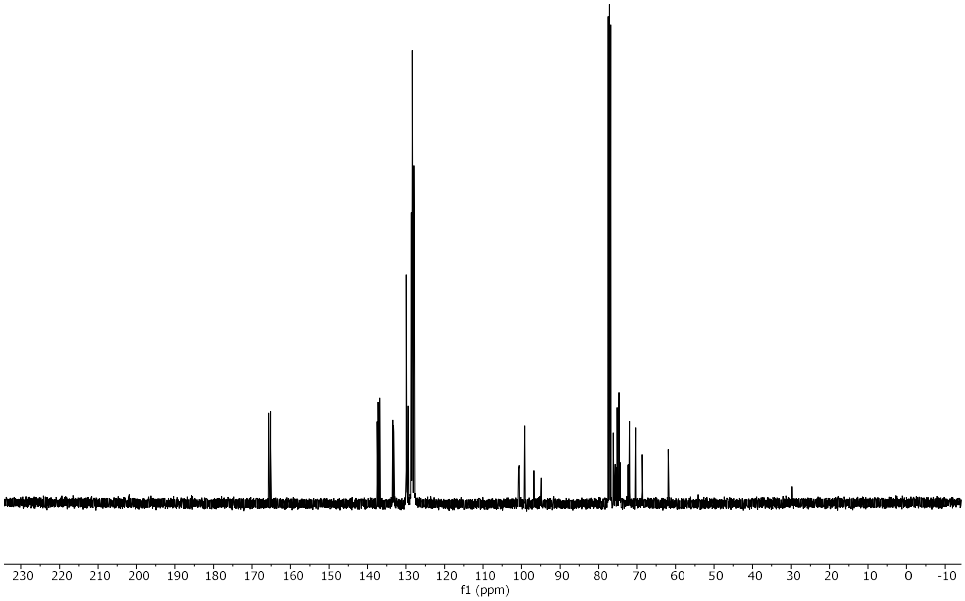
BB2.2 (110 mg, 0.171 mmol) and **S3.5** (79 mg, 0.139 mmol) were dissolved in anhydrous DCM (1.3 mL). The solution was stirred with molecular sieves (4 Å) for 1 h at RT under Ar atmosphere and then cooled to -20 °C. The activator solution (350 µL of a 133 mg/mL solution of NIS in DCM : Dioxane 2:1 with 1% TfOH) was added dropwise and the reaction was stirred for 5 min at -20 °C, after which time the cooling bath was removed to allow the reaction to RT. After 1 h stirring at RT, ESI-MS indicated the disappearance of BB**2.2**. Piperidine (0.5 mL) was added and the reaction was stirred at RT for additional 1 h. The reaction was diluted with DCM and washed once with a 1 M aqueous solution of HCl, once with a saturated aqueous solution of $NaHCO_3$ and once with brine. The crude was purified by silica gel flash column chromatography (Hexane : EtOAc = 3:1→2:1→1:1→2:3) to give **S3.6** as a white foam (114 mg, 88%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.13 – 8.02 (m, 2H), 7.96 – 7.84 (m, 4H), 7.54 – 7.47 (m, 3H), 7.41 – 7.31 (m, 11H), 7.27 – 7.09 (m, 8H), 7.02 – 6.95 (m, 2H), 5.60 (t, *J* = 9.5 Hz, 1H), 5.48 (ddd, *J* = 13.6, 9.2, 7.9 Hz, 1H), 5.38 (dd, *J* = 9.8, 7.9 Hz, 1H), 4.98 – 4.76 (m, 2H), 4.75 – 4.66 (m, 3H), 4.60 (d, *J* = 7.9 Hz, 1H), 4.48 (d, *J* = 12.7 Hz, 1H), 4.13 (dd, *J* = 11.2, 1.6 Hz, 1H), 4.02 – 3.76 (m, 5H), 3.67 – 3.57 (m, 1H), 3.53 – 3.44 (m, 1H); ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.67, 165.31, 165.18, 137.51, 137.25, 136.82, 133.48, 133.27, 133.17, 100.69 (d, *J* = 11.1 Hz), 99.22, 95.81 (d, *J* = 188.1 Hz), 76.14, 75.63 (d, *J* = 16.7 Hz), 75.17, 74.74, 74.65, 74.37 (d, *J* = 8.6 Hz), 72.31 (d, *J* = 18.8 Hz), 71.94, 70.35, 68.68, 61.84. ^{19}F NMR (376 MHz, $CDCl_3$) δ -191.14 (dt, *J* = 52.6, 13.5 Hz); (ESI-HRMS) m/z 949.3256 $[M+Na]^+$ ($C_{54}H_{51}FO_{13}Na$ requires 949.3206).

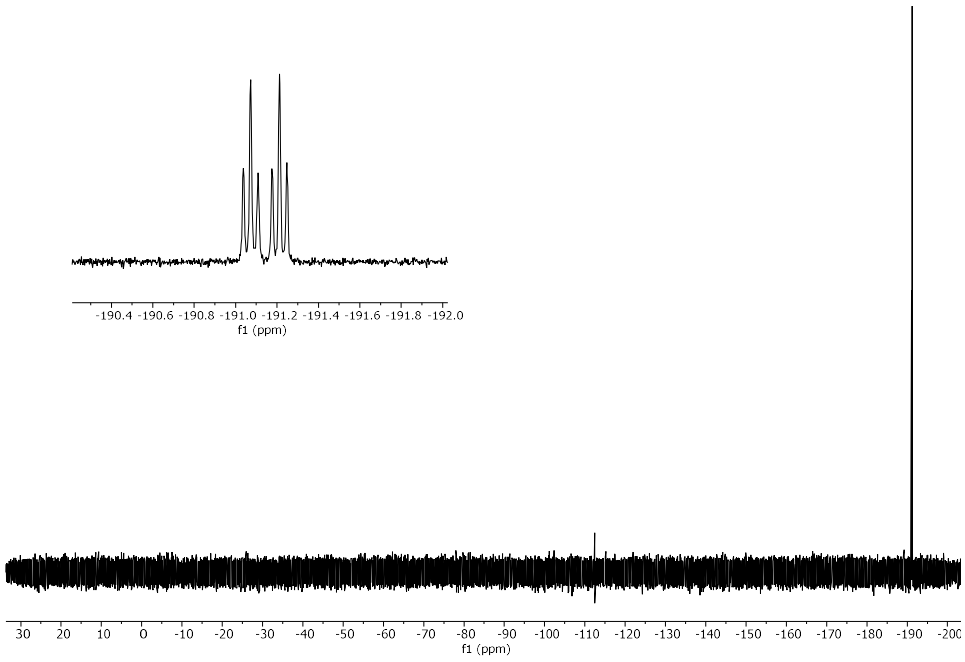
^{1}H NMR of S3.6 (400 MHz, $CDCl_3$)



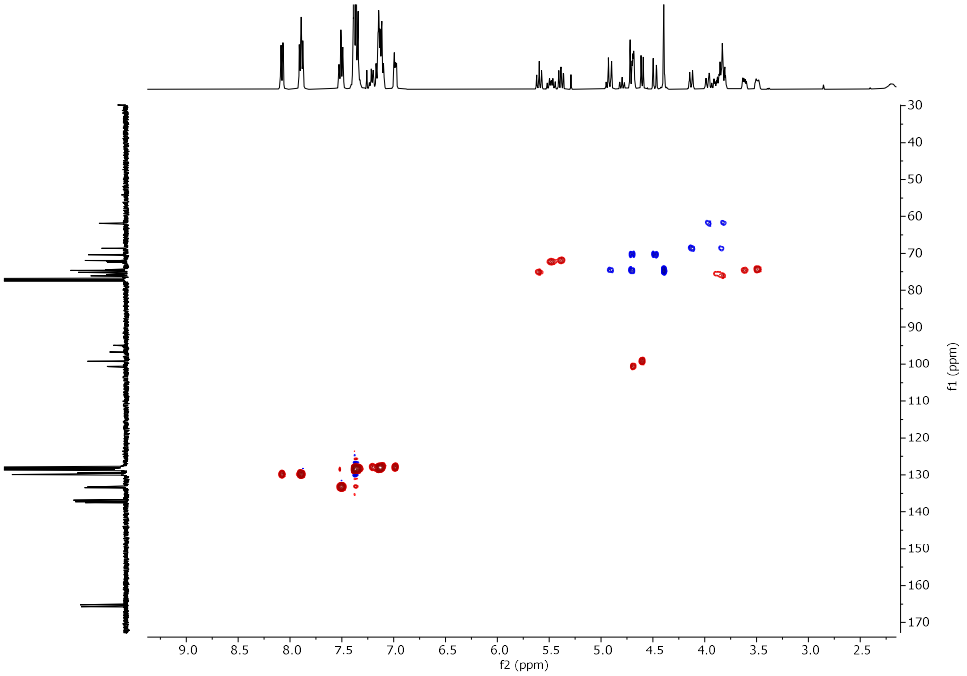
^{13}C NMR of S3.6 (101 MHz, $CDCl_3$)



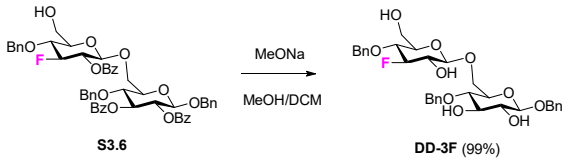
^{19}F NMR of S3.6 (376 MHz, $CDCl_3$)



HSQC NMR of S3.6 ($CDCl_3$)



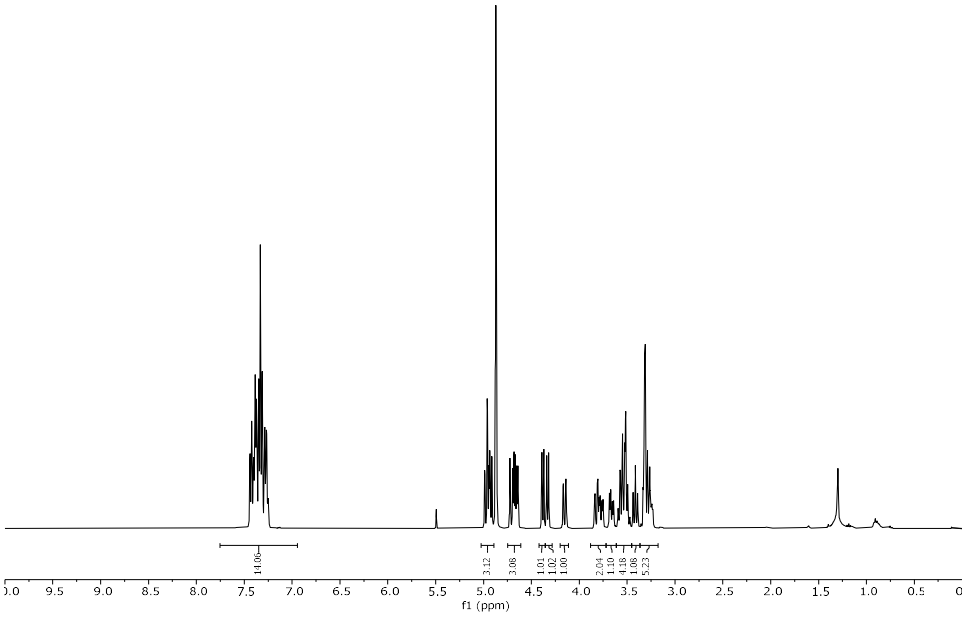
Synthesis of **DD-3F**



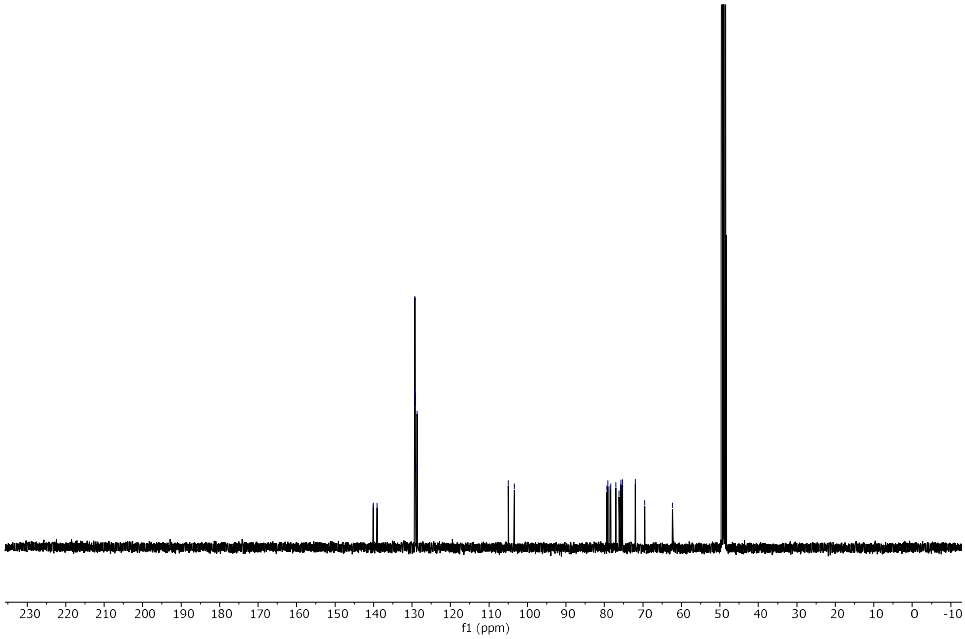
S3.6 (52 mg, 0.051 mmol) was dissolved in a 1:1 mixture of MeOH : DCM (4 mL). MeONa in MeOH (0.5 M, 3 equiv. per benzoyl ester) was added and the solution was stirred at RT for 16 h, neutralized with Amberlite IR-120 (H^+ form) resin, filtered, and concentrated *in vacuo*. The resulting yellow oil was purified by silica gel flash column chromatography (DCM : MeOH = 15:1) and recrystallized from DCM : Hexane to give **DD-3F** as white solid (31 mg, 99%).

^{1}H NMR (400 MHz, CD_3OD) δ 7.51 – 7.16 (m, 15H), 5.01 – 4.89 (m, 3H), 4.75 – 4.61 (m, 3H), 4.38 (d, *J* = 7.8 Hz, 1H), 4.33 (d, *J* = 7.8 Hz, 1H), 4.15 (d, *J* = 10.9 Hz, 1H), 3.88 – 3.72 (m, 2H), 3.72 – 3.62 (m, 1H), 3.62 – 3.45 (m, 4H), 3.41 (t, *J* = 9.3 Hz, 1H), 3.36 – 3.18 (m, 3H). ^{13}C NMR (101 MHz, CD_3OD) δ 140.10, 140.04, 139.10, 129.31, 129.28, 129.16, 129.13, 129.08, 128.69, 128.65, 105.00, 103.42, 79.34, 79.12, 78.51, 78.37, 77.05, 76.17, 75.73, 75.67, 75.37, 75.31, 71.97, 69.56, 62.32. ^{19}F NMR (376 MHz, CD_3OD) δ -191.91 (dt, *J* = 52.8, 14.2 Hz). (ESI-HRMS) m/z 637.2440 $[M+Na]^+$ ($C_{33}H_{40}O_{11}Na$ requires 637.2419).

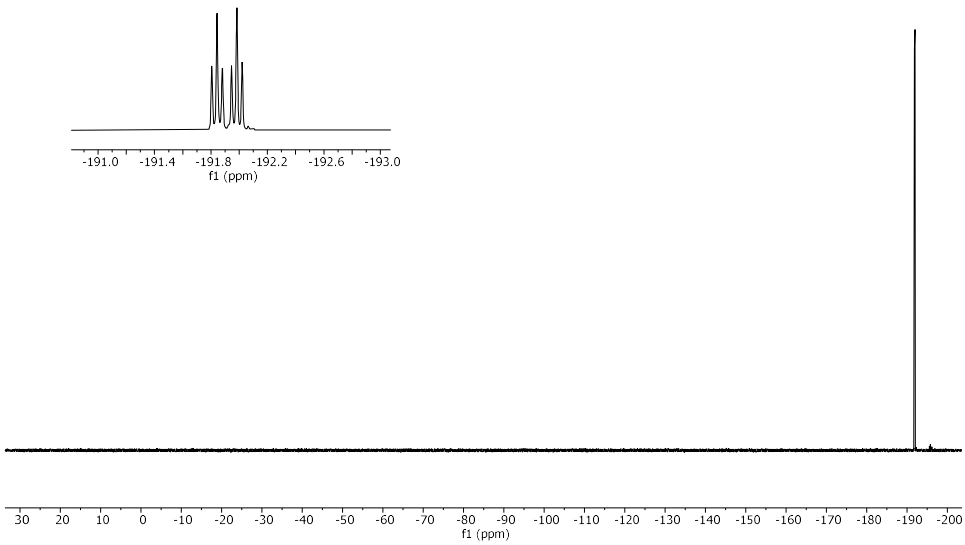
^{1}H NMR of DD-3F (400 MHz, CD_3OD)



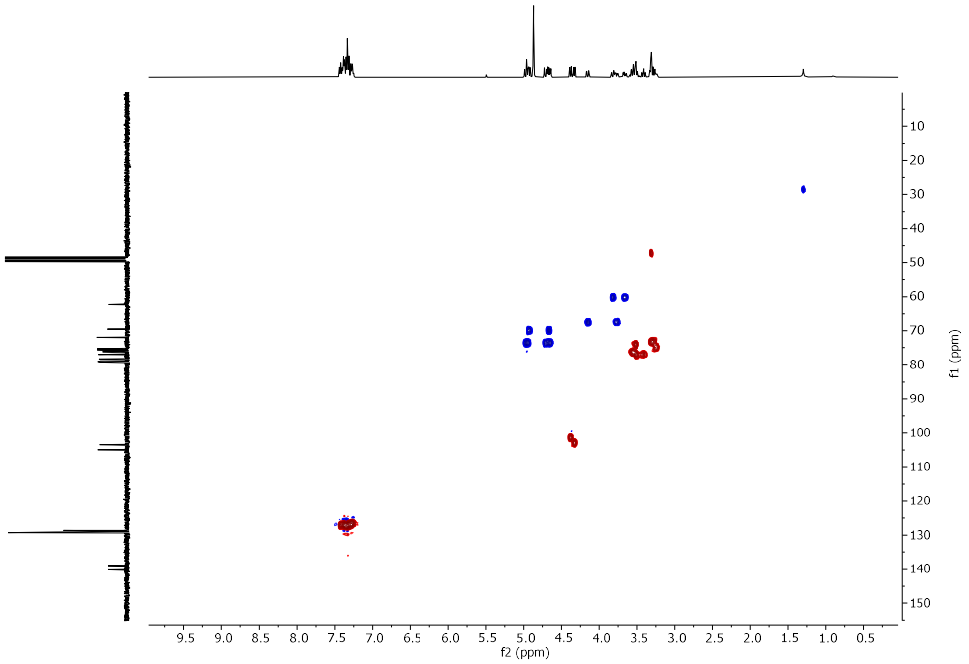
^{13}C NMR of DD-3F (101 MHz, CD_3OD)



^{19}F NMR of DD-3F (376 MHz, CD_3OD)



HSQC NMR of DD-3F (CD_3OD)



6.4.3 Oligosaccharides synthesis

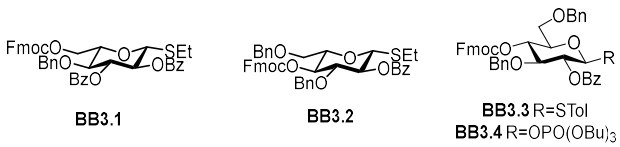
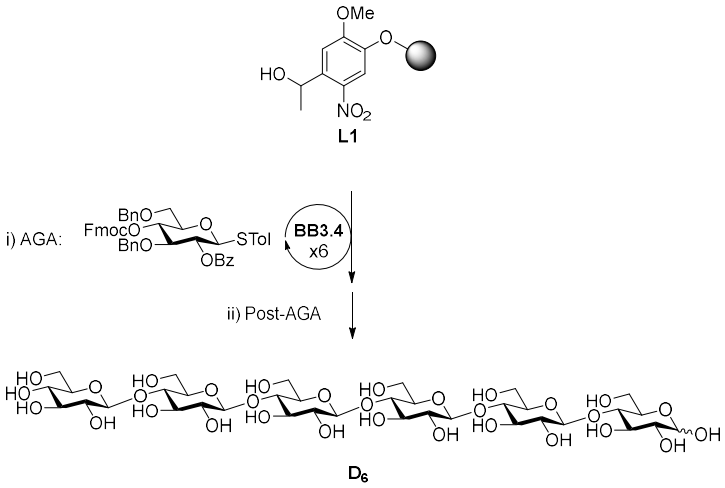


Figure S3.1 BBs used for the AGA of oligosaccharides described in Chapter 3.

Compounds **D5**, **D7**, **D8**, and **D9** were synthesized by Dr. Denisa Vargová.

Synthesis of **D6**

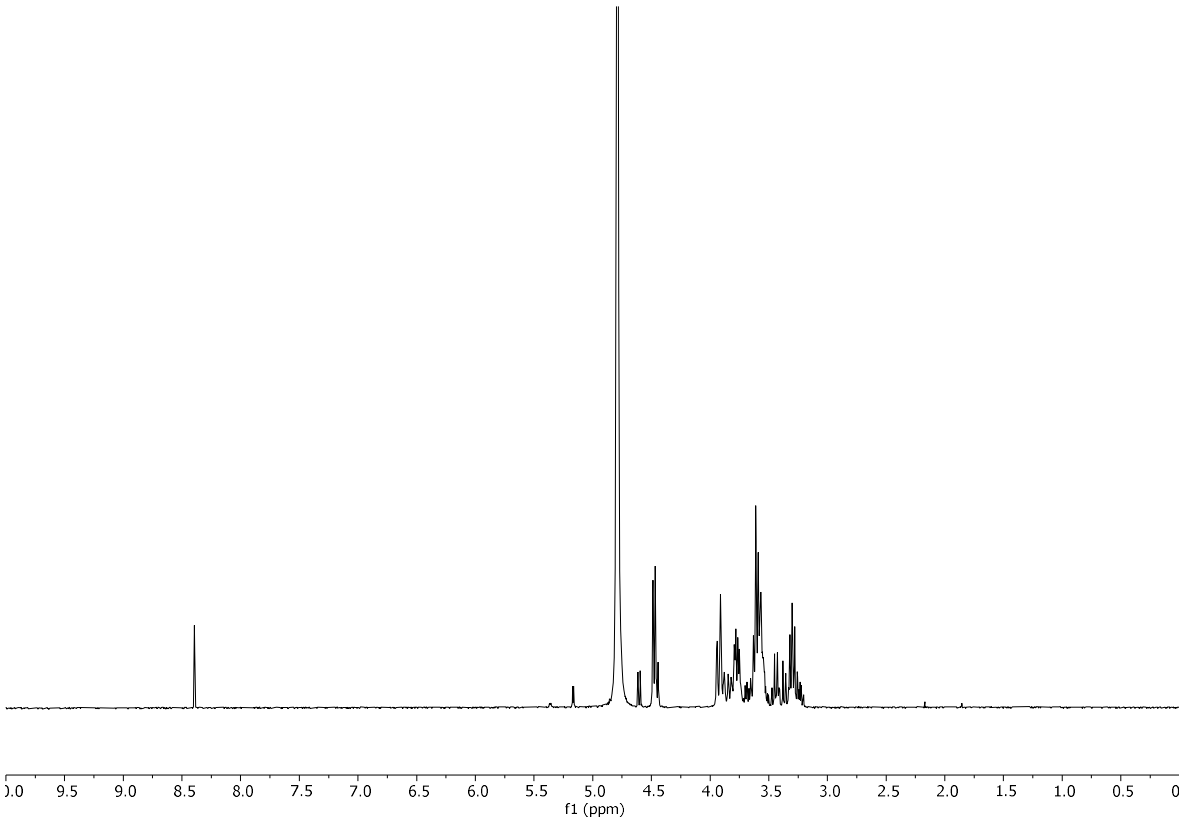


Step	Modules		Notes
AGA	-	**A**	**L1** swelling
	(**BB3.4**)x6	(**B**, **C1**, **D**, **E1**)x6	**C1**: (**BB3.4**, -20 °C for 5 min, 0 °C for 20 min)
Post-AGA	-	**F1**, **G1**, **H2**, **I**	**F1**: (3 d) **H2**: (4 h) **I**: (Method D)

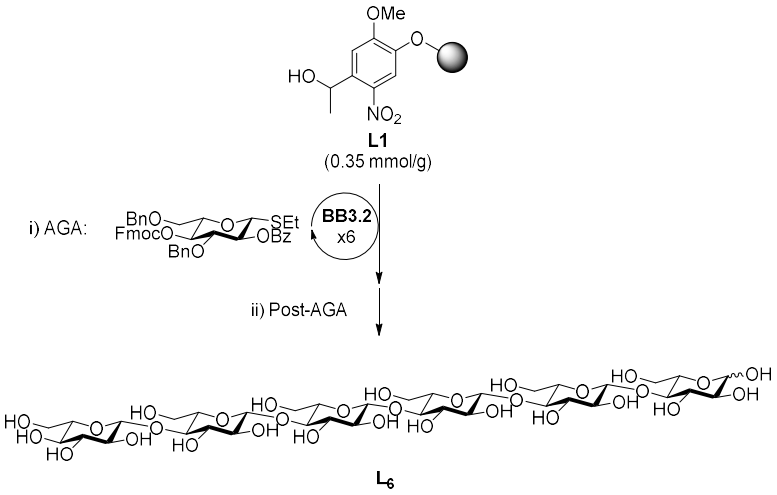
Automated synthesis, global deprotection, and purification afforded compound **D6** as a white solid (8.0 mg, 57% overall yield).

Analytical data for **D6** were in good agreement with previously reported data.[190]
^{1}H NMR (600 MHz, D_2O) δ 5.24 (d, *J* = 3.7 Hz, 0.4H, H1^{I}-α), 4.68 (d, *J* = 7.9 Hz, 0.6H, H1^{I}-β), 4.58 – 4.50 (m, 5H, H1$^{II, III, IV, V, VI}$), 4.05 – 3.90 (m, 6H), 3.90 – 3.80 (m, 6H), 3.79 – 3.57 (m, 16H), 3.55 – 3.48 (m, 2H), 3.46 – 3.41 (m, 1H), 3.41 – 3.27 (m, 5H).

^{1}H NMR of D_6 (400 MHz, D_2O)



Synthesis of $\mathbf{L_6}$

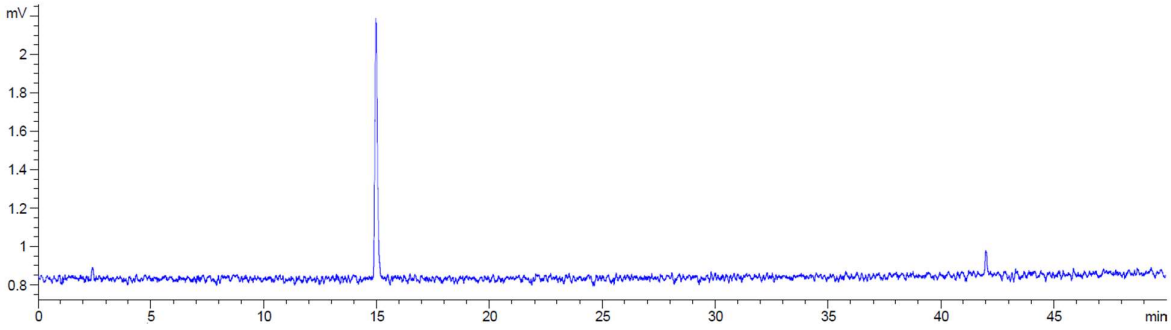


Step	Modules		Notes
AGA	-	**A**	**L1** swelling
	(**BB3.2**)x6	(**B**, **C1**, **D**, **E1**)x6	**C1**: (**BB3.2**, -20 °C for 5 min, 0 °C for 20 min)
Post-AGA	-	**F1**, **G1**, **H2**, **I**	**F1**: (3 d) **H2**: (6 h) **I**: (Method D)

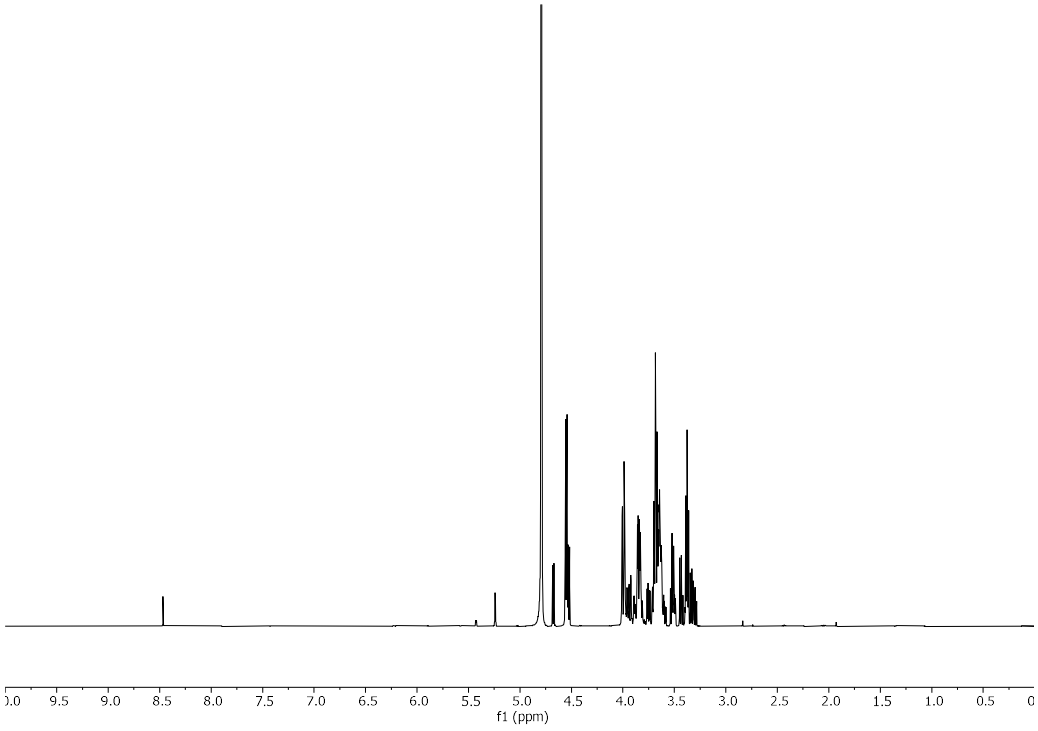
Automated synthesis, global deprotection, and purification afforded compound $\mathbf{L_6}$ as a white solid (8.4 mg, 60% overall yield).

^{1}H NMR (600 MHz, D_2O) δ 5.24 (d, *J* = 3.8 Hz, 0.4H, H1^{I}-α), 4.68 (d, *J* = 7.9 Hz, 0.6H, H1^{I}-β), 4.59 – 4.50 (m, 5H, H1$^{II, III, IV, V, VI}$), 4.05 – 3.91 (m, 6H), 3.90 – 3.78 (m, 6H), 3.78 – 3.57 (m, 16H), 3.55 – 3.47 (m, 2H), 3.47 – 3.41 (m, 1H), 3.40 – 3.35 (m, 4H), 3.35 – 3.28 (m, 1H). ^{13}C NMR (151 MHz, D_2O) δ 102.48, 102.27, 95.68 (s, β-C1), 91.75 (s, α-C1), 78.56, 78.41, 78.31, 78.18, 75.91, 75.41, 74.75, 74.18, 73.96, 73.92, 73.82, 73.07, 72.85, 71.23, 71.16, 70.05, 69.37, 60.49, 59.94, 59.78. (ESI-HRMS) m/z 1013.322 $[M+Na]^+$ ($C_{36}H_{62}O_{31}Na$ requires 1013.317).

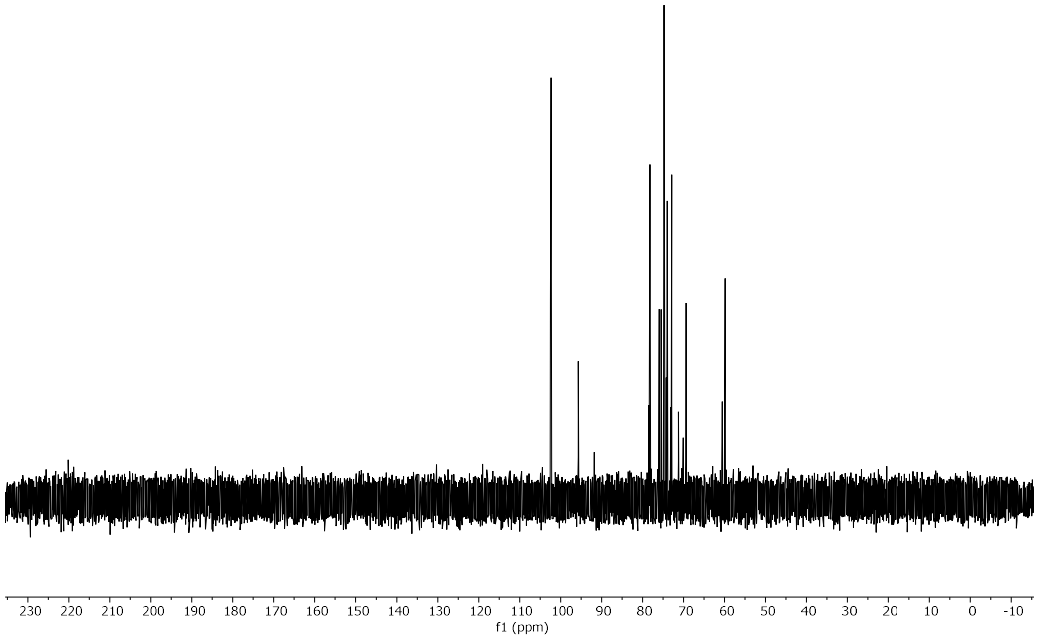
RP-HPLC of L_6 (ELSD trace, Method C1, t_R = 15.0 min)



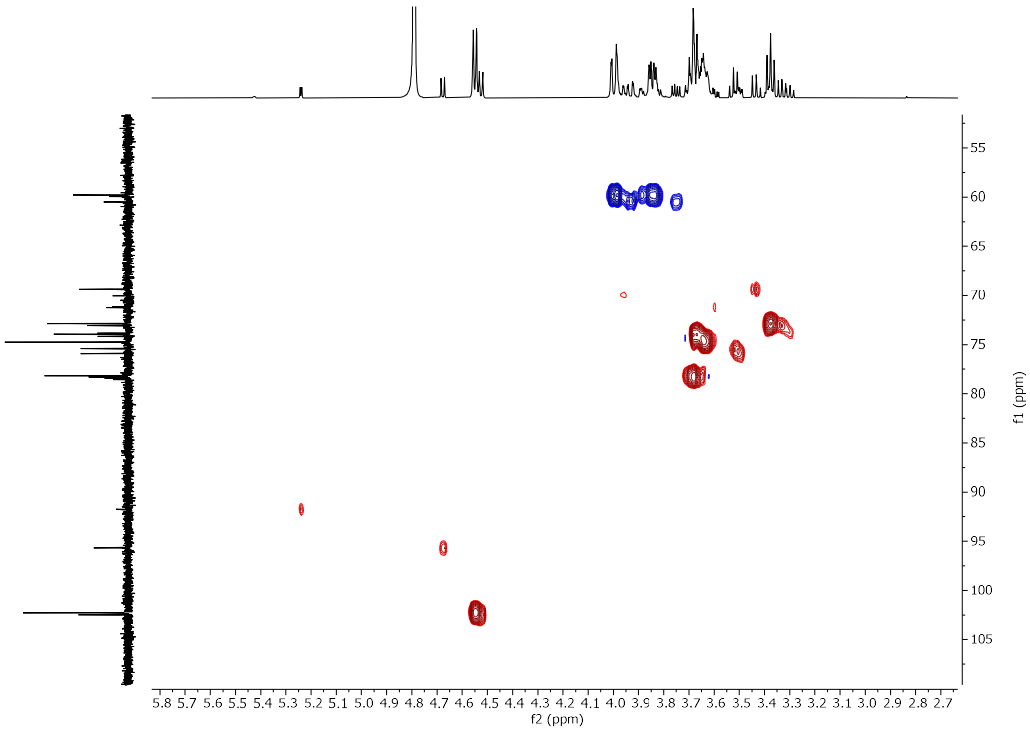
^{1}H NMR of L_6 (600 MHz, D_2O)



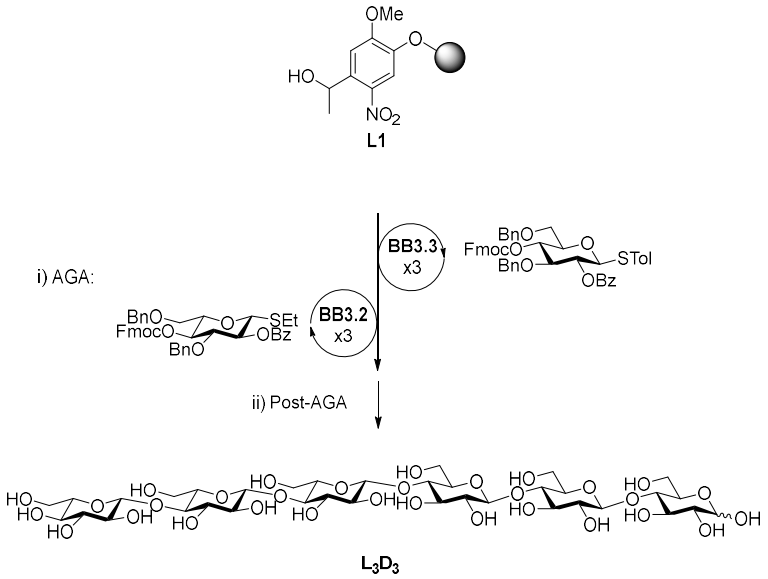
^{13}C NMR of L_6 (151 MHz, D_2O)



HSQC NMR of L_6 (D_2O)



Synthesis of **L_3D_3**

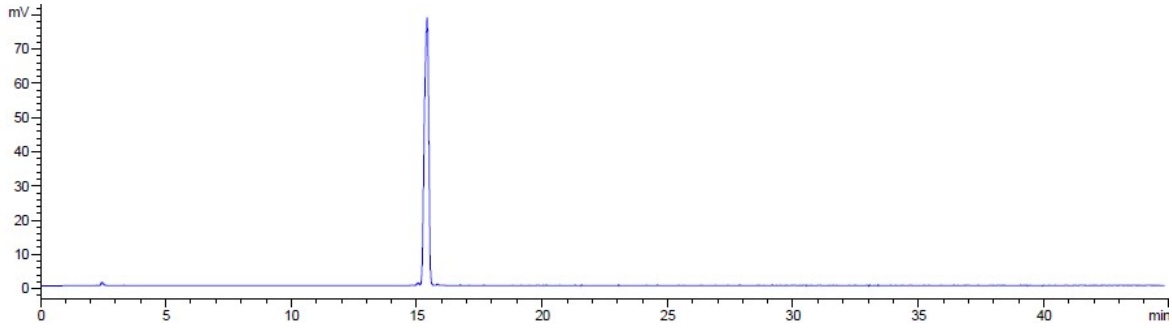


Step	Modules		Notes
AGA	**-**	**A**	**L1** swelling
	(**BB3.3**)x3	(**B**, **C1**, **D**, **E1**)x3	**C1**: (**BB3.3**, -20 °C for 5 min, 0 °C for 20 min)
	(**BB3.2**)x3	(**B**, **C1**, **D**, **E1**)x3	**C1**: (**BB3.2**, -20 °C for 5 min, 0 °C for 20 min)
Post-AGA	**-**	**F1**, **G1**, **H2**, **I**	**F1**: (16 h) **H2**: (4 h) **I**: (Method D)

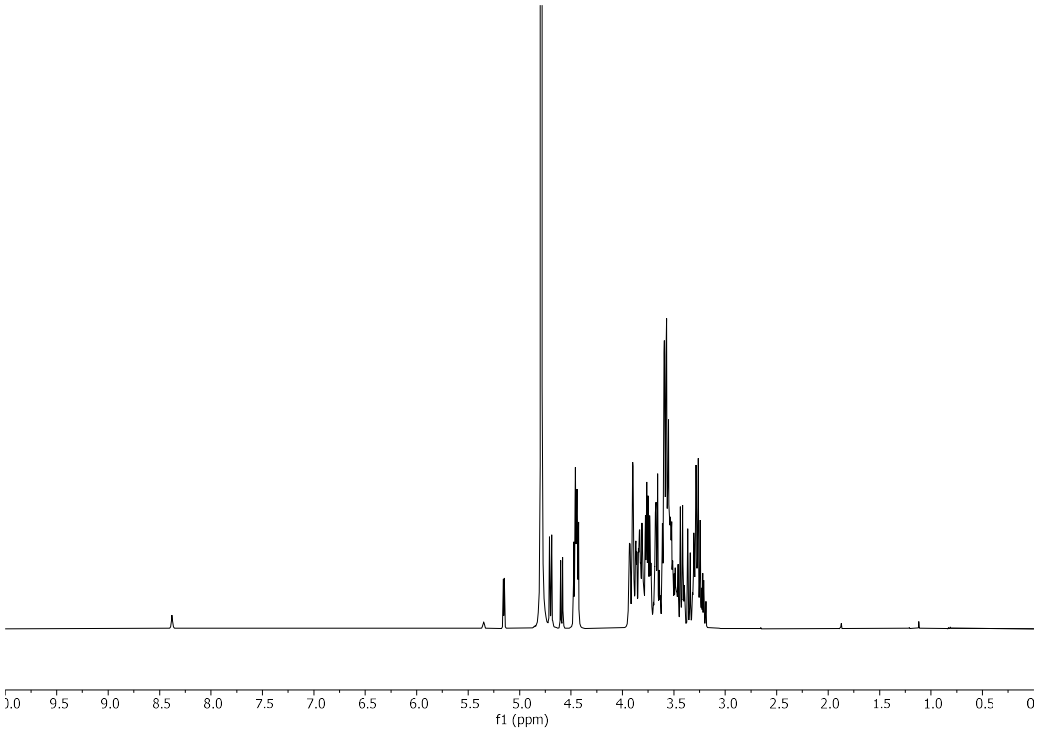
Automated synthesis, global deprotection, and purification afforded compound **L_3D_3** as a white solid (5.3 mg, 37% overall yield).

^{1}H NMR (400 MHz, D_2O) δ 5.15 (d, *J* = 3.7 Hz, 0.4H, $H1^{I}$-α), 4.70 (d, *J* = 7.9 Hz, 1H, $H1^{IV}$), 4.59 (d, *J* = 8.0 Hz, 0.6H, $H1^{I}$-β), 4.50 – 4.40 (m, 4H, $H1^{II, III, V, VI}$), 3.97 – 3.71 (m, 12H), 3.71 – 3.63 (m, 3H), 3.63 – 3.45 (m, 12H), 3.45 – 3.38 (m, 2H), 3.35 (d, *J* = 9.1 Hz, 1H), 3.33 – 3.17 (m, 6H). ^{13}C NMR (101 MHz, D_2O) δ 102.67, 102.44, 102.30, 102.22, 95.63, 91.69, 76.43, 75.85, 75.32, 75.07, 74.69, 74.51, 74.33, 74.10, 74.03, 73.88, 73.85, 73.75, 73.17, 73.00, 72.93, 72.80, 71.16, 69.99, 69.30, 60.41, 60.13, 59.89, 59.82, 59.71. (ESI-HRMS) m/z 1013.324 $[M+Na]^+$ ($C_{36}H_{62}O_{31}Na$ requires 1013.317).

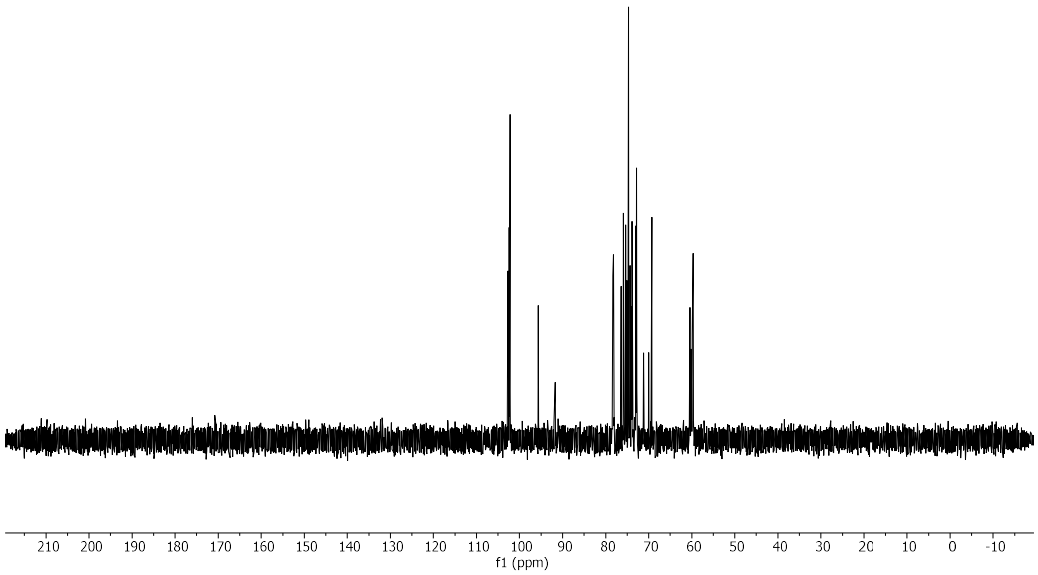
RP-HPLC of L_3D_3 (ELSD trace, Method C1, t_R = 15.4 min)



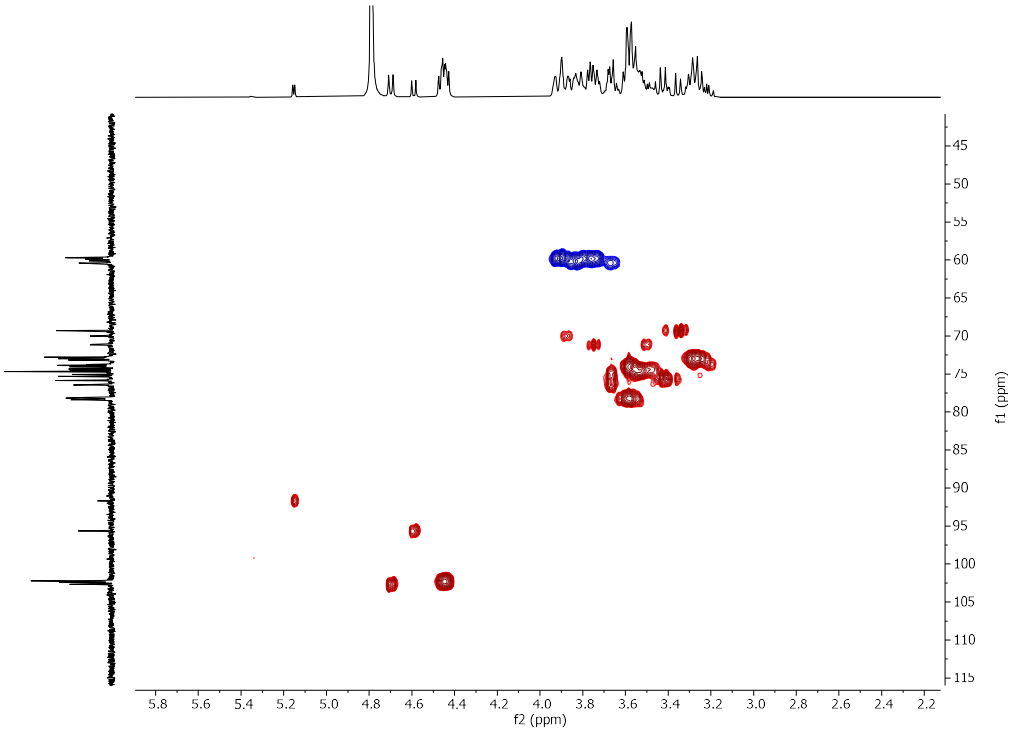
^{1}H NMR of L_3D_3 (400 MHz, D_2O)



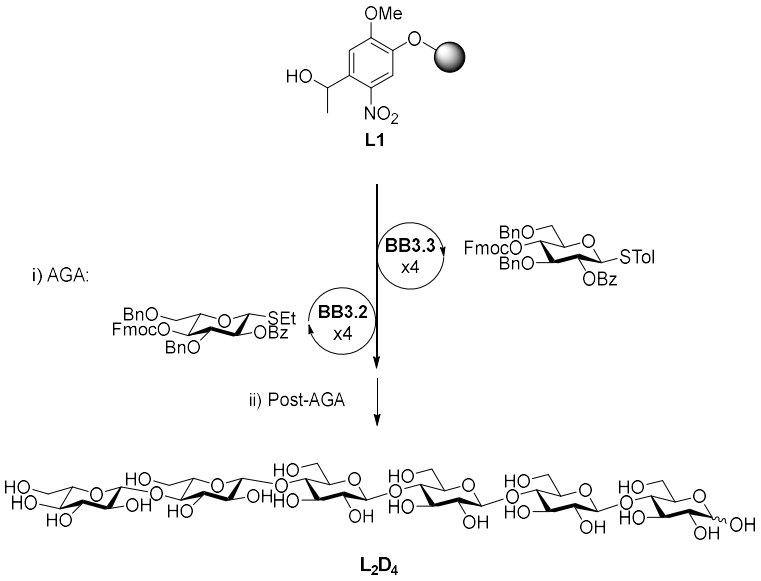
^{13}C NMR of L_3D_3 (101 MHz, D_2O)



HSQC NMR of L_3D_3 (D_2O)



Synthesis of $\mathbf{L_2D_4}$

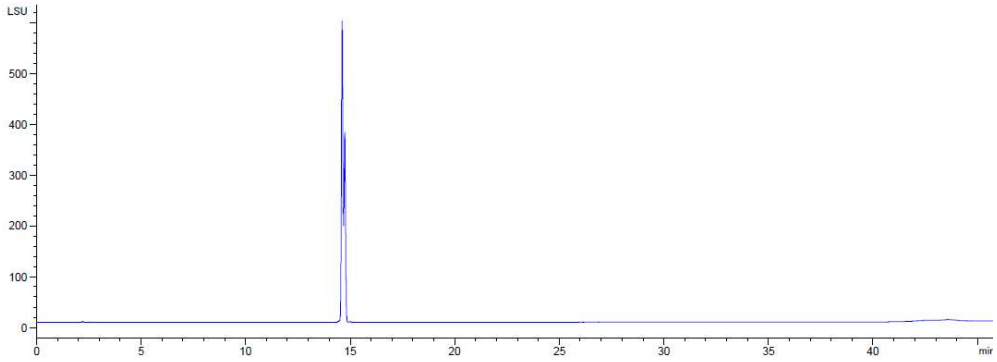


Step	Modules		Notes
AGA	-	**A**	**L1** swelling
	(**BB3.3**)x4	(**B**, **C1**, **D**, **E1**)x4	**C1**: (**BB3.3**, -20 °C for 5 min, 0 °C for 20 min)
	(**BB3.2**)x2	(**B**, **C1**, **D**, **E1**)x2	**C1**: (**BB3.2**, -20 °C for 5 min, 0 °C for 20 min)
Post-AGA	-	**F1**, **G1**, **H2**, **I**	**F1**: (16 h) **H2**: (4 h) **I**: (Method D)

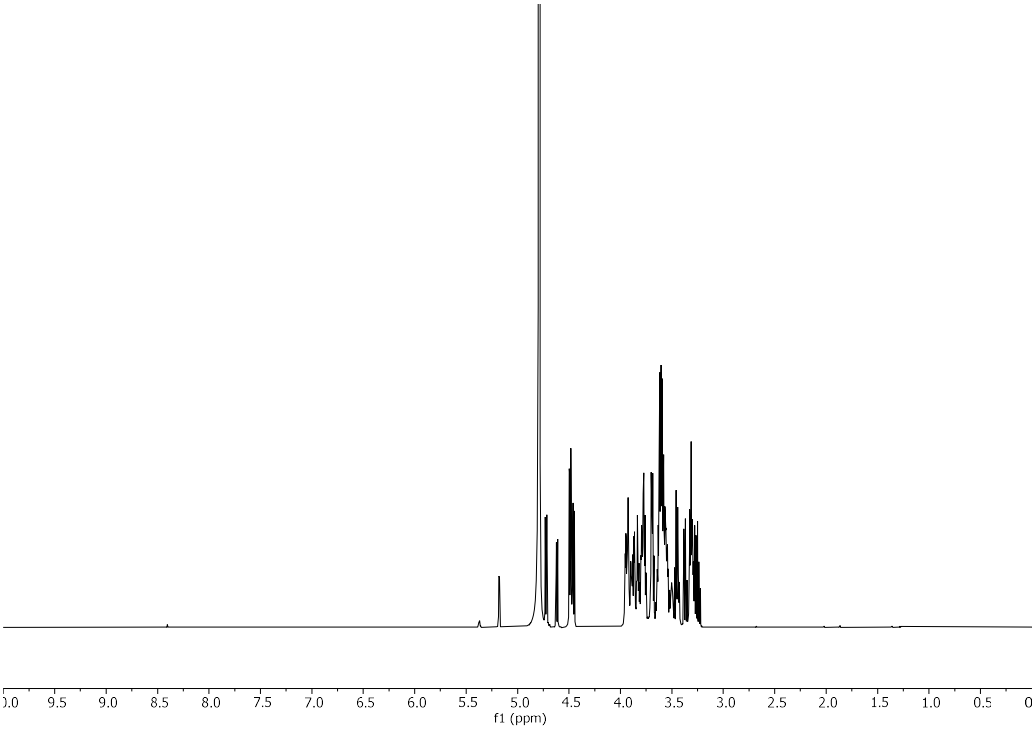
Automated synthesis, global deprotection, and purification afforded compound $\mathbf{L_3D_3}$ as a white solid (2.8 mg, 23% overall yield).

^{1}H NMR (600 MHz, D_2O) δ 5.18 (d, *J* = 3.7 Hz, 0.4H, H1^{I}-α), 4.72 (d, *J* = 7.9 Hz, 1H, H1^{V}), 4.62 (d, *J* = 8.0 Hz, 0.6H, H1^{I}-β), 4.52 – 4.44 (m, 4H, H1$^{II, III, IV, VI}$), 3.98 – 3.73 (m, 12H), 3.72 – 3.66 (m, 3H), 3.66 – 3.41 (m, 15H), 3.39 – 3.34 (m, 1H), 3.34 – 3.20 (m, 5H). ^{13}C NMR (151 MHz, D_2O) δ 102.73, 102.51, 102.36, 102.30, 95.70, 91.77, 78.61, 78.55, 78.40, 78.22, 78.17, 76.55, 75.93, 75.41, 75.16, 74.76, 74.58, 74.42, 74.19, 74.16, 73.94. 73.84, 73.25, 73.09, 73.00, 72.89, 72.87, 71.24, 71.17, 71.07, 69.39, 60.51, 60.22, 60.05, 59.94, 59.80. (ESI-HRMS) m/z 1013.323 $[M+Na]^+$ ($C_{36}H_{62}O_{31}Na$ requires 1013.317).

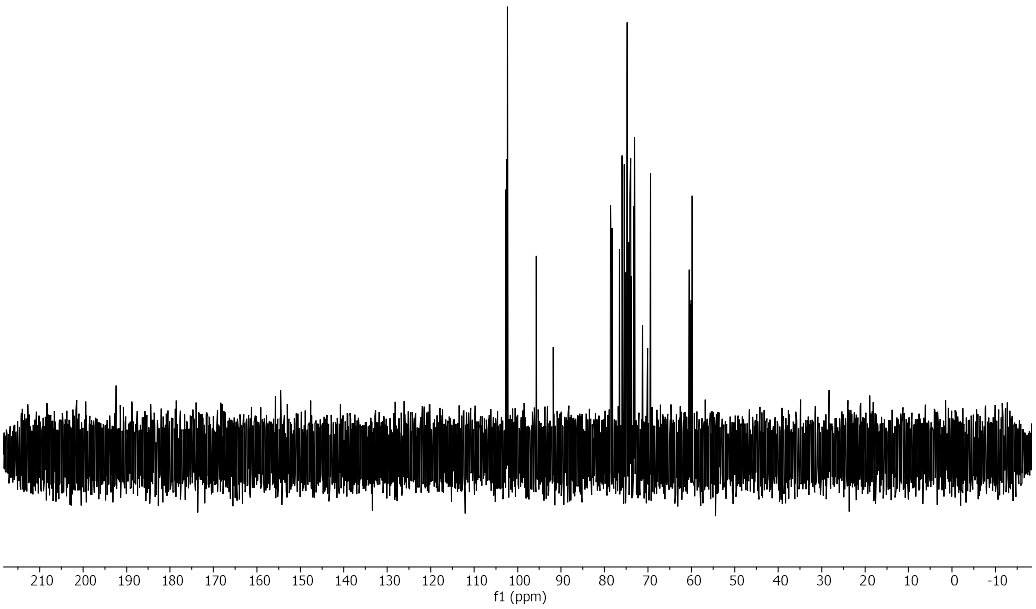
RP-HPLC of L_2D_4 (ELSD trace, Method C1, t_R = 14.6, 14.7 min)



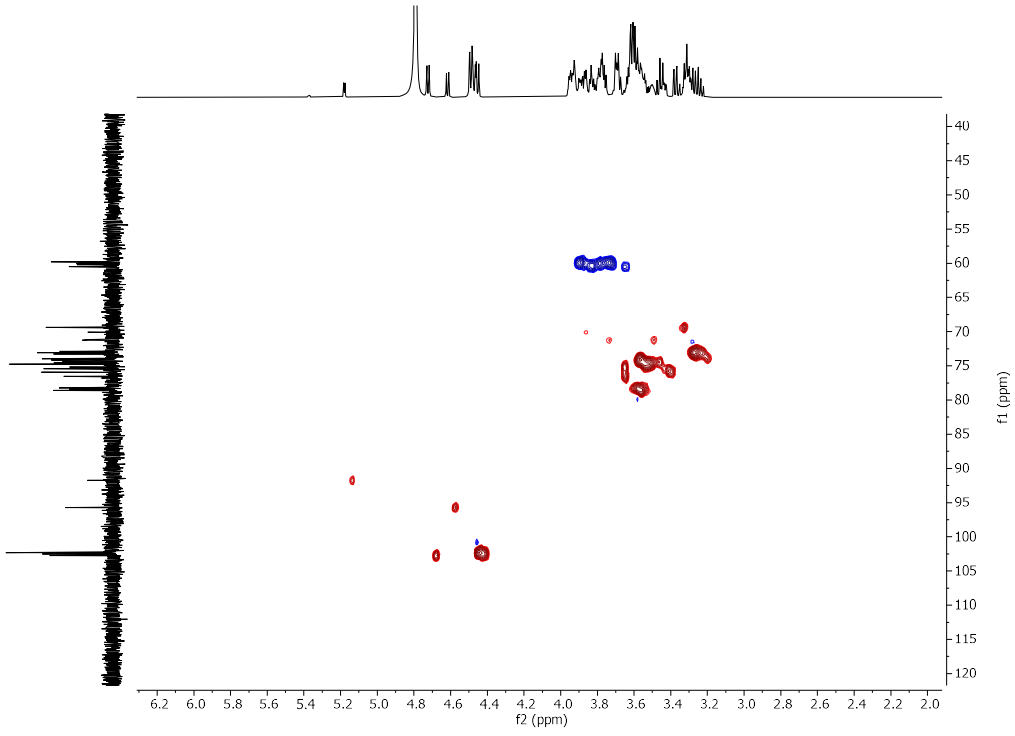
^{1}H NMR of L_2D_4 (400 MHz, D_2O)



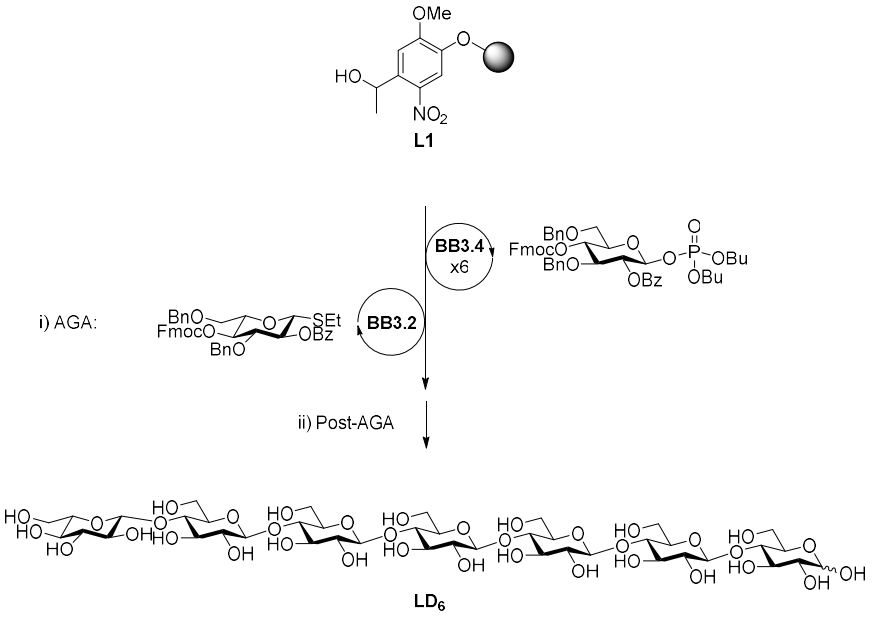
^{13}C NMR of L_2D_4 (101 MHz, D_2O)



HSQC NMR of L_2D_4 (D_2O)



Synthesis of **LD_6**

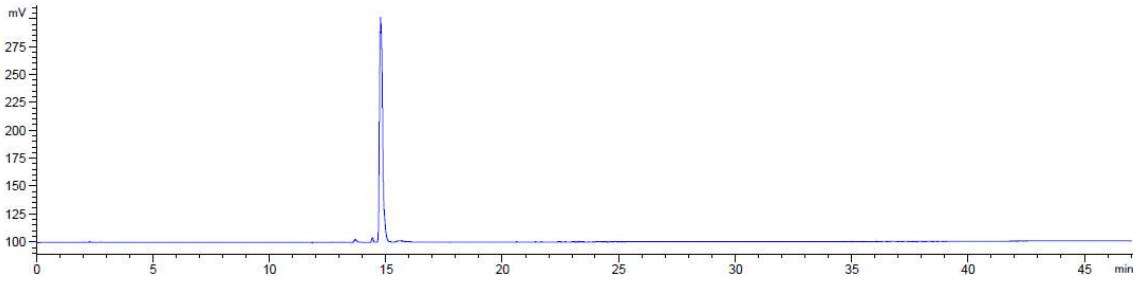


Step	Modules		Notes
AGA	**-**	**A**	**L1** swelling
	(**BB3.4**)x6	(**B**, **C2**, **D**, **E1**)x6	**C2**: (**BB3.4**, -35 °C for 5 min, -15 °C for 20 min)
	BB3.2	**B**, **C1**, **D**, **E1**	**C1**: (**BB32**, -20 °C for 5 min, 0 °C for 20 min)
Post-AGA	**-**	**F1**, **G1**, **H2**, **I**	**F**: (1 d) **H2**: (4 h) **I**: (Method D)

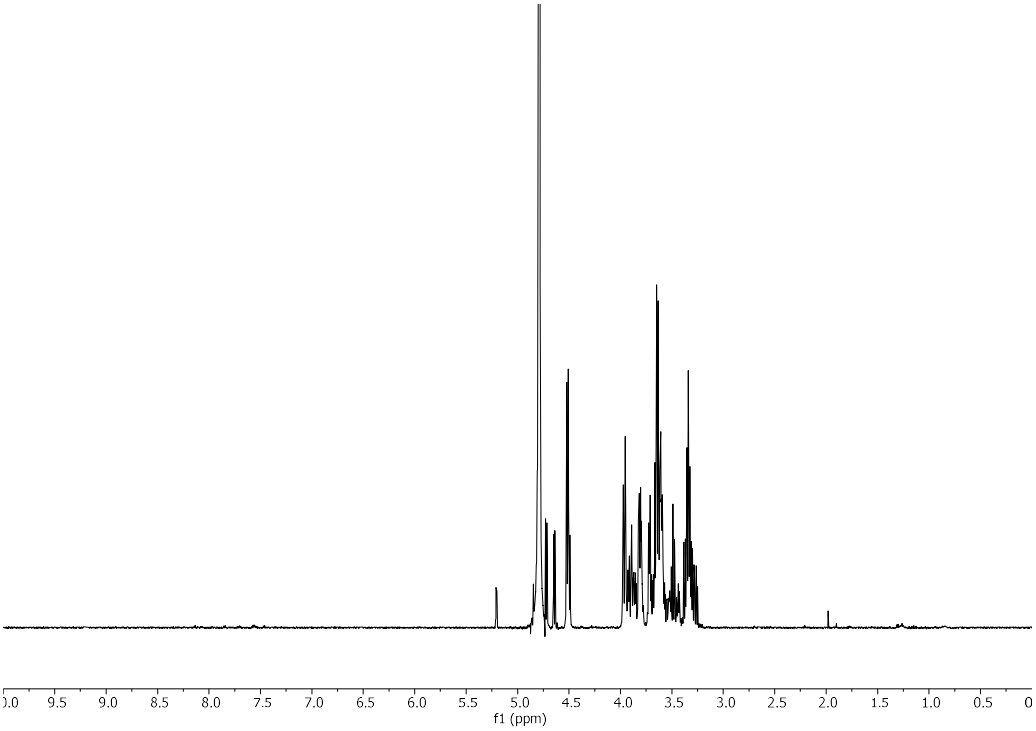
Automated synthesis, global deprotection afforded compound **LD_6** as a white solid (11.1 mg, 74% overall yield).

^{1}H NMR (600 MHz, D_2O) δ 5.12 (d, *J* = 3.8 Hz, 0.3H, H1^{I}-α), 4.63 (d, *J* = 8.0 Hz, 1H, H1^{VII}), 4.55 (d, *J* = 7.9 Hz, 0.7H, H1^{I}-β), 4.42 (dd, *J* = 10.7, 8.0 Hz, 5H, H1$^{II, III, IV, V, VI}$), 3.91 – 3.75 (m, 8H), 3.72 (dt, *J* = 15.8, 8.2 Hz, 5H), 3.66 – 3.60 (m, 3H), 3.60 – 3.47 (m, 15H), 3.45 (dd, *J* = 16.5, 3.3 Hz, 1H), 3.40 (t, *J* = 9.2 Hz, 1H), 3.35 (ddt, *J* = 12.3, 8.6, 4.2 Hz, 1H), 3.31 – 3.15 (m, 8H). ^{13}C NMR (151 MHz, D_2O) δ 102.97, 102.30, 95.70, 91.77, 78.54, 78.39, 78.18, 76.57, 75.70, 75.60, 75.16, 74.77, 74.43, 74.19, 73.93, 73.83, 73.47, 73.00, 72.87, 71.24, 71.17, 70.07, 69.66, 60.74, 60.18, 59.93, 59.78. (ESI-HRMS) m/z 1153.393 $[M+H]^+$ ($C_{42}H_{73}O_{36}$ requires 1153.388).

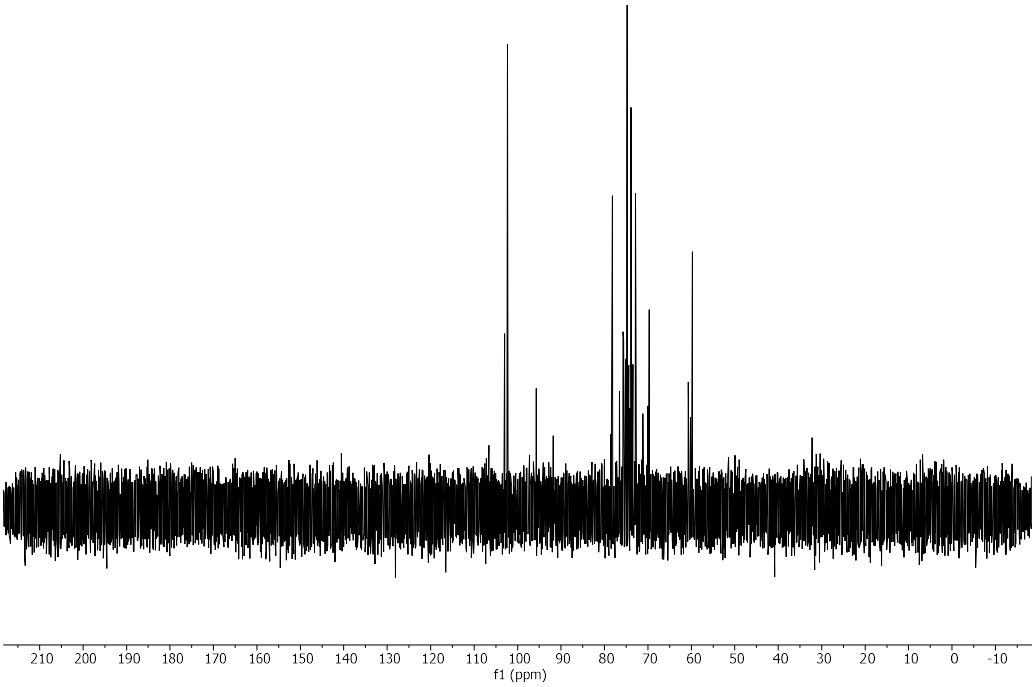
RP-HPLC of LD_6 (ELSD trace, Method C1, t_R = 14.8 min)



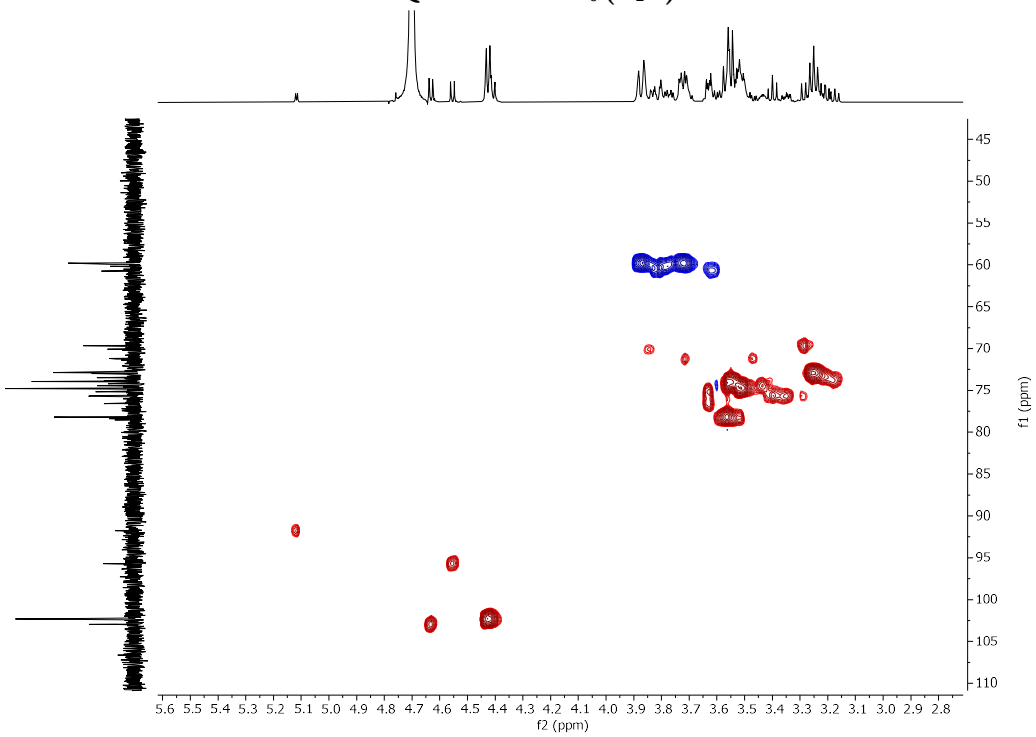
^{1}H NMR of LD_6 (600 MHz, D_2O)



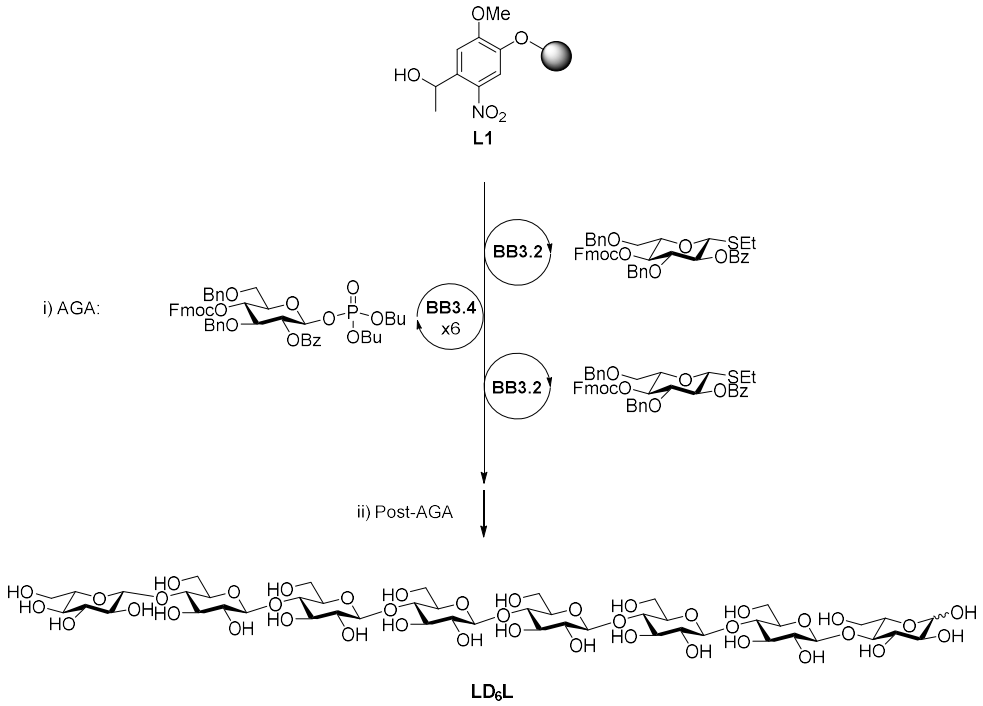
^{13}C NMR of LD_6 (D_2O)



HSQC NMR of LD_6 (D_2O)



Synthesis of **LD$_6$L**

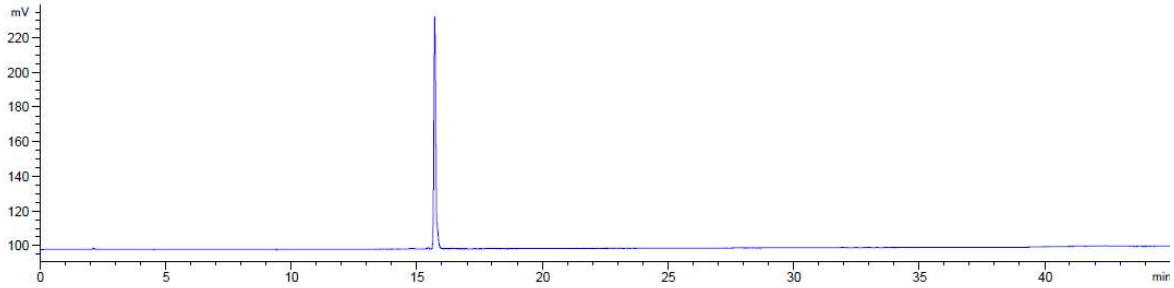


Step	Modules		Notes
AGA	**L1**	**A**	**L1** swelling
	BB3.2	**B**, **C1**, **D**, **E1**	**C1**: (**BB3.2**, -20 °C for 5 min, 0 °C for 20 min)
	(**BB3.4**)x6	(**B**, **C2**, **D**, **E1**)x6	**C2**: (**BB3.4**, -35 °C for 5 min, -15 °C for 20 min)
	BB3.2	**B**, **C1**, **D**, **E1**	**C1**: (**BB3.2**, -20 °C for 5 min, 0 °C for 20 min)
Post-AGA		**F1**, **G1**, **H2**, **I**	**F1**: (3 d) **H2**: (4 h) **I**: (Method D)

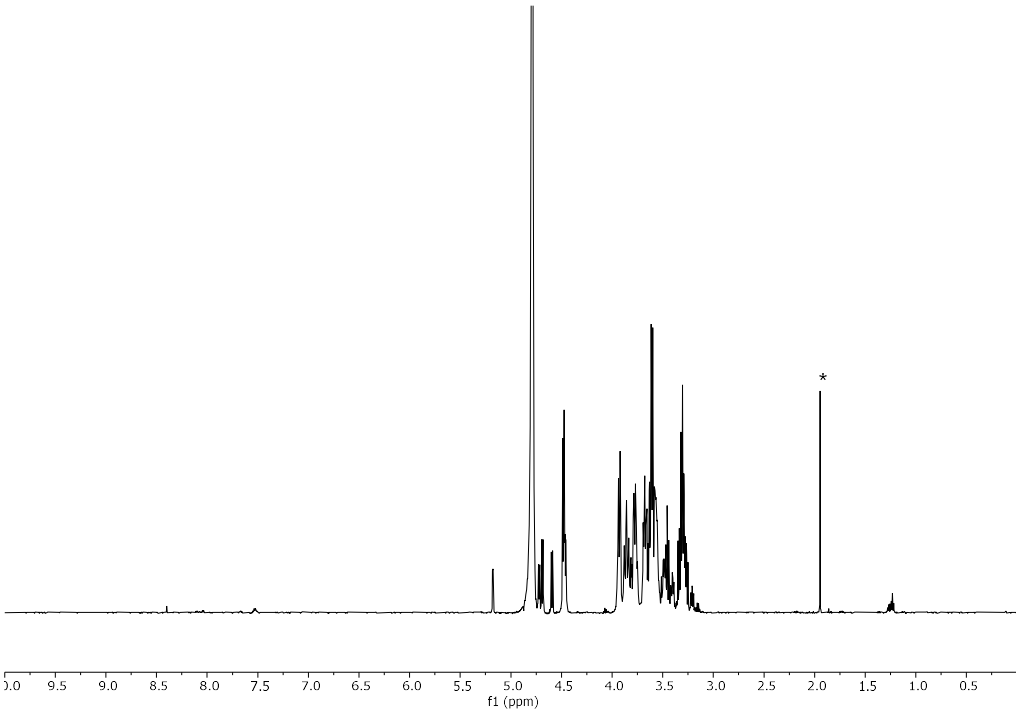
Automated synthesis, global deprotection afforded compound **LD$_6$L** as a white solid (8.5 mg, 50% overall yield).

f^{1}H NMR (600 MHz, D_2O) δ 5.18 (d, *J* = 3.7 Hz, 0.4H, H1^{I}-α), 4.72 (d, *J* = 8.0 Hz, 1H, H1^{II}), 4.69 (d, *J* = 7.9 Hz, 1H, , H1^{VIII}), 4.59 (d, *J* = 8.0 Hz, 0.6H, H1^{I}-β), 4.51 – 4.43 (m, 5H, H1$^{III, IV, V, VI, VII}$), 3.93 (d, *J* = 10.8 Hz, 5H), 3.90 – 3.73 (m, 11H), 3.73 – 3.52 (m, 19H), 3.52 – 3.37 (m, 5H), 3.37 – 3.19 (m, 8H). ^{13}C NMR (151 MHz, D_2O) δ 102.93, 102.71, 102.33, 102.26, 95.73, 91.86, 78.11, 76.66, 76.51, 75.65, 75.55, 75.40, 75.11, 74.72, 74.55, 74.53, 74.38, 74.34, 74.07, 73.88, 73.42, 73.22, 72.95, 72.83, 69.62, 60.68, 59.72. (ESI-HRMS) m/z 1337.427 $[M+Na]^+$ ($C_{48}H_{82}O_{41}Na$ requires 1337.422).

RP-HPLC of LD_6L (ELSD trace, Method C1, t_R = 15.7 min)

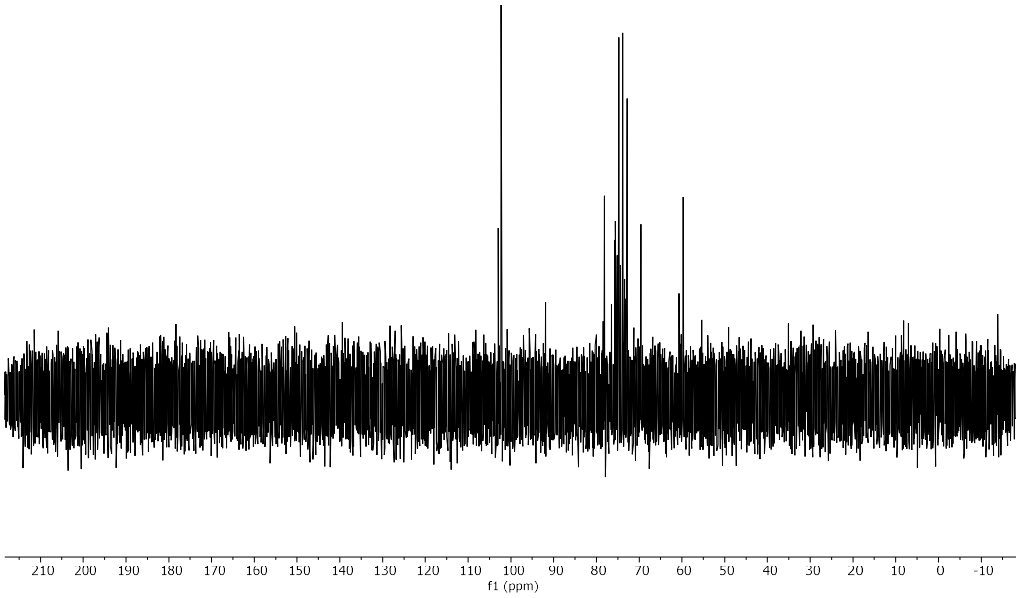


^{1}H NMR of LD_6L (400 MHz, D_2O)

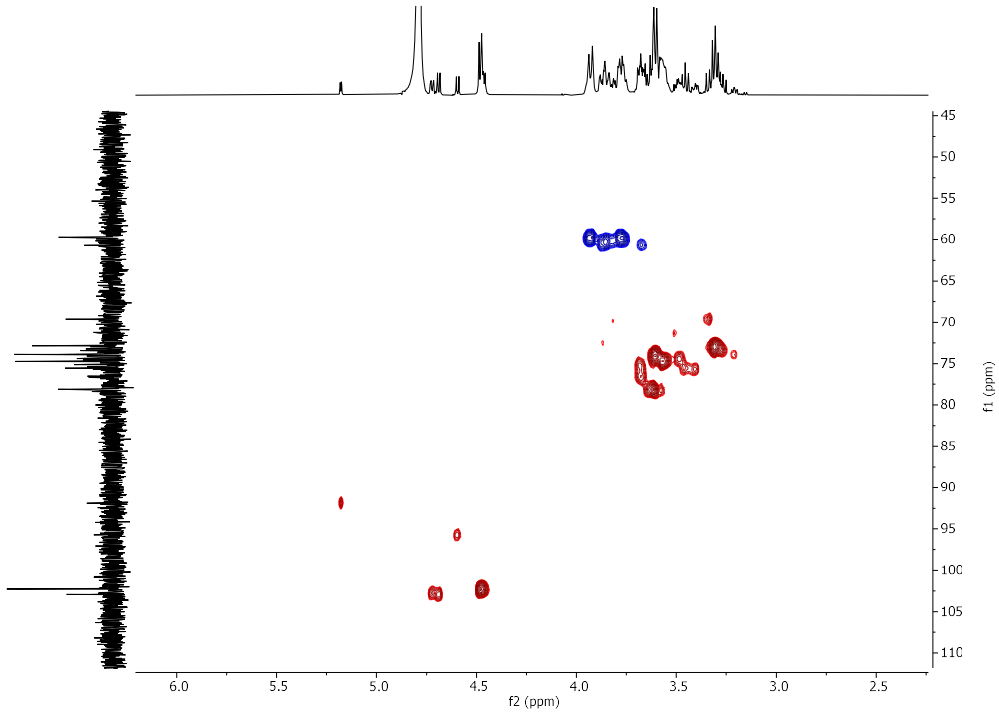


*Unidentified impurity

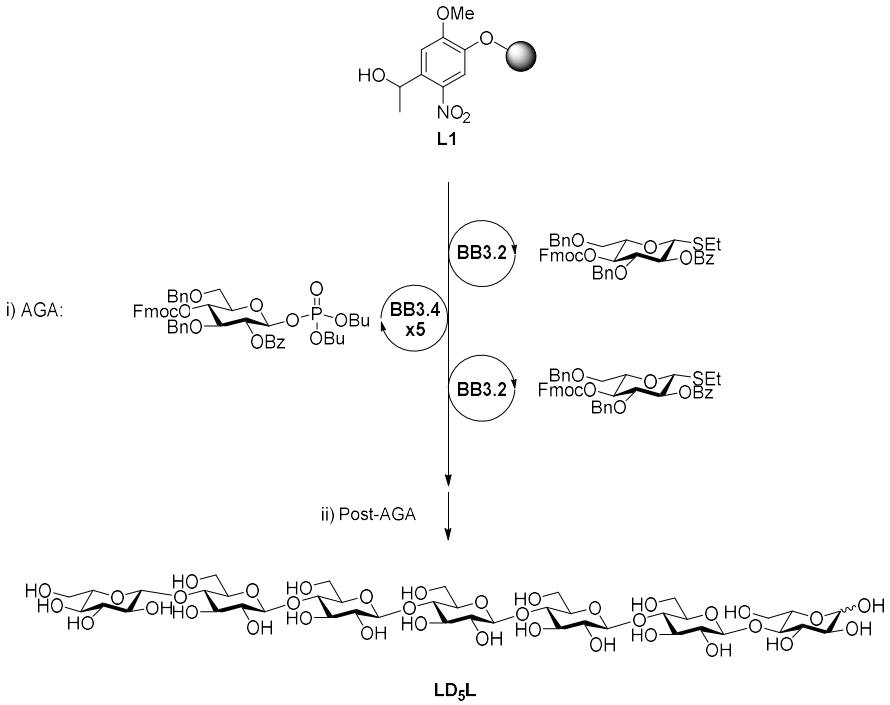
^{13}C NMR of LD$_6$L (101 MHz, D$_2$O)



HSQC NMR of LD$_6$L (D$_2$O)



Synthesis of **LD_5L**

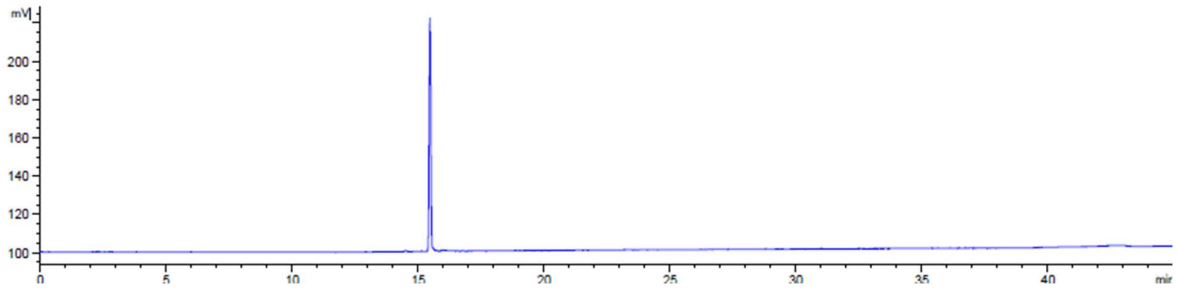


Step	Modules		Notes
AGA	**-**	**A**	**L1** swelling
	BB3.2	**B**, **C1**, **D**, **E**	**C1**: (**BB3.2**, -20 °C for 5 min, 0 °C for 20 min)
	(**BB3.4**)x5	(**B**, **C2**, **D**, **E**)x5	**C2**: (**BB3.4**, -35 °C for 5 min, -15 °C for 20 min)
	BB3.2	**B**, **C1**, **D**, **E**	**C1**: (**BB3.2**, -20 °C for 5 min, 0 °C for 20 min)
Post-AGA	**-**	**F1**, **G1**, **H2**, **I**	**F1**: (2 d) **H2**: (5 h) **I**: (Method D)

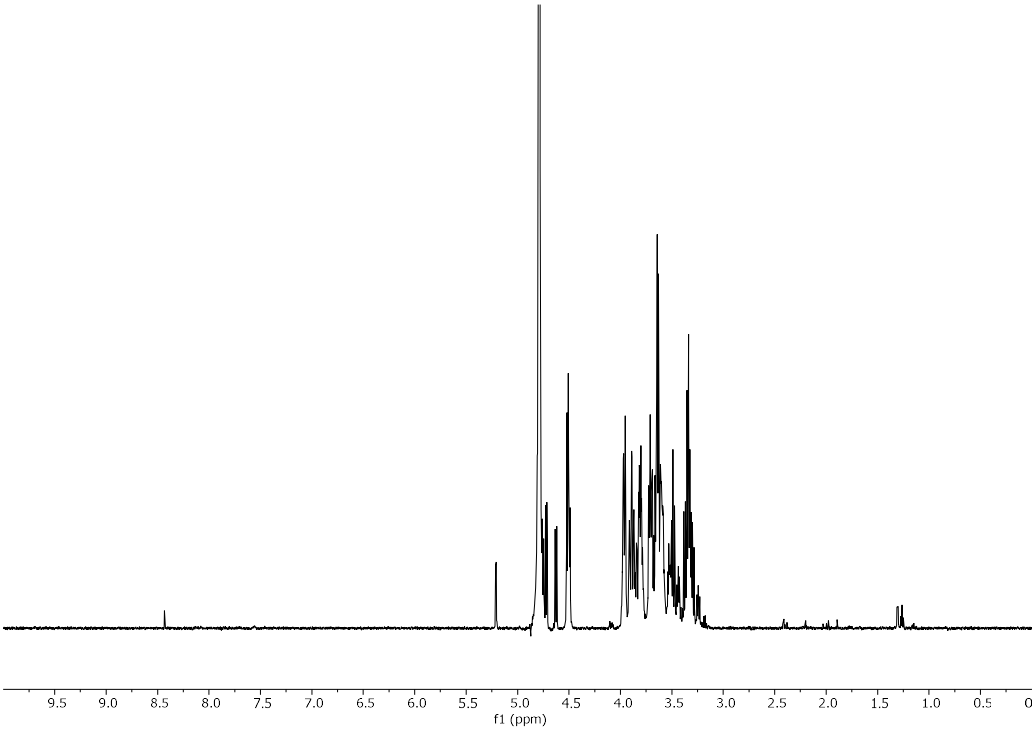
Automated synthesis, global deprotection afforded compound **LD_5L** as a white solid (5.1 mg, 35% overall yield).

^{1}H NMR (600 MHz, D_2O) δ 5.21 (d, *J* = 3.7 Hz, 0.4H, $H1^{I}$-α), 4.75 (m, 1H, $H1^{II}$), 4.72 (d, *J* = 7.9 Hz, 1H, $H1^{VII}$), 4.63 (d, *J* = 8.0 Hz, 0.6H, $H1^{I}$-β), 4.54 – 4.47 (m, 4H, $H1^{III, IV, V, VI}$), 3.96 (d, *J* = 10.8 Hz, 4H), 3.93 – 3.76 (m, 10H), 3.76 – 3.55 (m, 16H), 3.55 – 3.40 (m, 4H), 3.40 – 3.27 (m, 7H), 3.27 – 3.20 (m, 1H). ^{13}C NMR (151 MHz, D_2O) δ 102.97, 102.76, 102.37, 102.30, 95.76, 91.84, 78.23, 78.18, 76.57, 75.70, 75.60, 75.45, 75.16, 74.76, 74.59, 74.43, 74.11, 73.93, 73.47, 73.27, 73.00, 72.87, 72.52, 71.27, 69.85, 69.66, 60.74, 60.38, 60.19, 59.98, 59.79. (ESI-HRMS) m/z 1175.381 $[M+Na]^+$ ($C_{42}H_{72}O_{36}Na$ requires 1175.370).

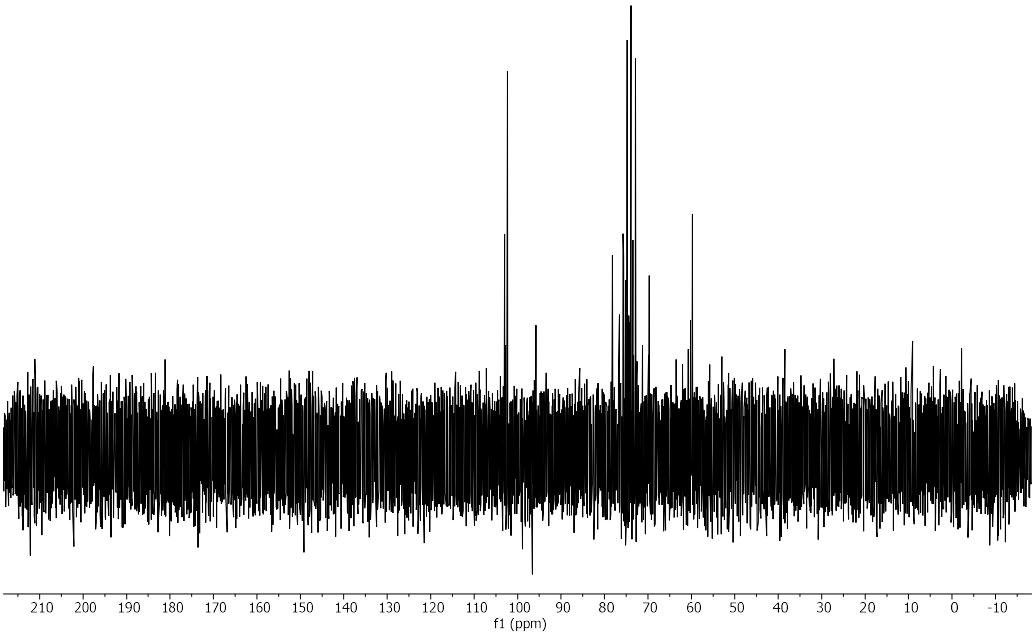
RP-HPLC of LD_5L (ELSD trace, Method C1, t_R = 15.5 min)



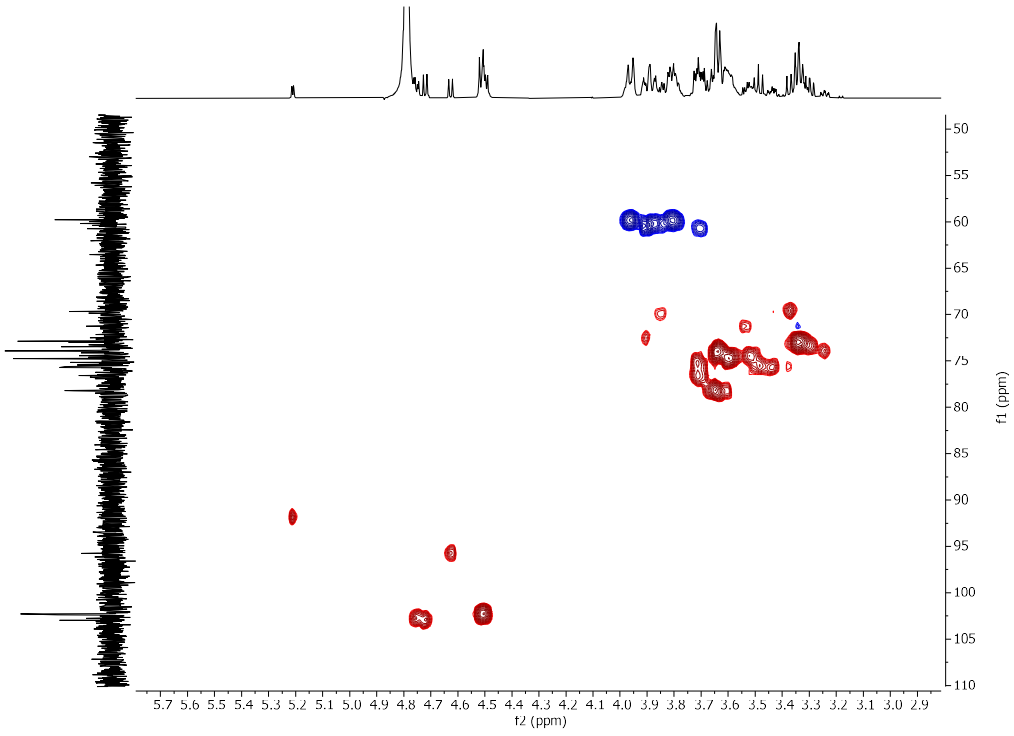
^{1}H NMR of LD_5L (400 MHz, D_2O)



^{13}C NMR of LD_5L (101 MHz, D_2O)



HSQC NMR of LD_5L (D_2O)



6.4.4 Dimers self-assembly

The self-assembly, SEM, and XRD analysis was performed by Dr. Soeun Gim. The electron diffraction analysis was performed by Dr. Yu Ogawa.

Solvent switch method. Stock solutions of the disaccharides (20, 40 and 100 mg/mL) in HFIP were diluted with water to reach a final concentration of 2 mg/mL with different ratio of water and HFIP. The samples were incubated for 3 days without agitation before the measurement.

Details on solid state NMR (ssNMR) and MicroED analysis of **DD** and **DD-6F** are reported in the original publications.[358,359]

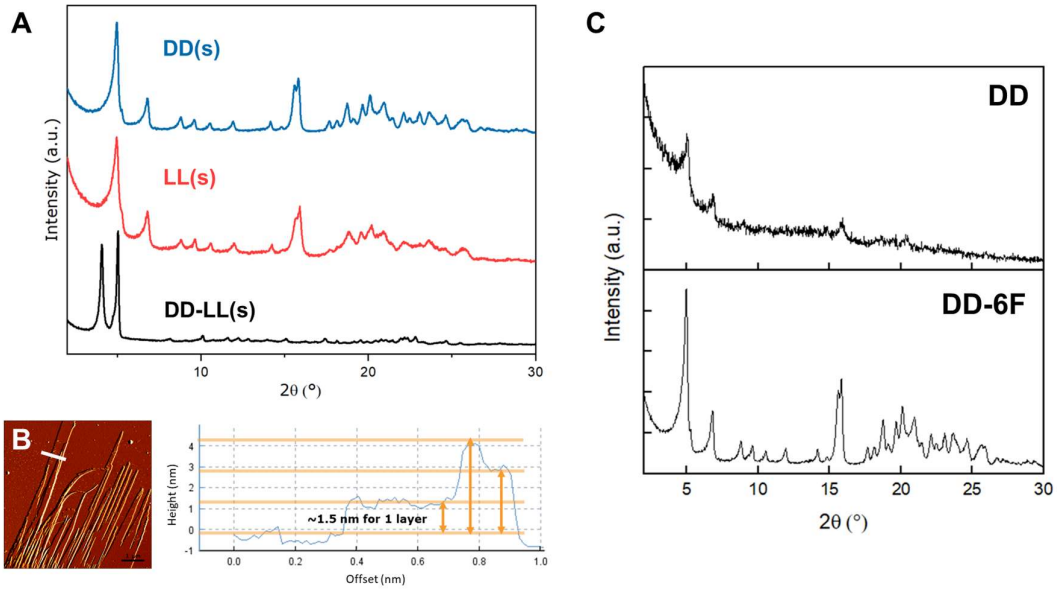


Figure S3.2 A) Powder XRD profiles for **DD** (*blue*), **LL** (*red*), and **DD-LL** (*black*). B) AFM image and cross-section analysis of **DD-LL**. C) Comparison between powder XRD patterns of **DD** and **DD-6F**.

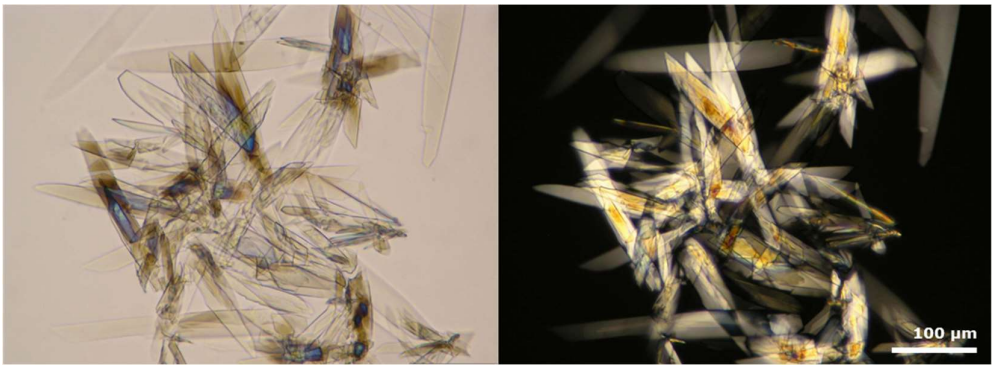


Figure S3.3 Polarized optical microscope (POM) image of **DD-LL** taken between parallel (*left*) and crossed (*right*) polarizers.

6.4.5 Oligosaccharides self-assembly

Solubility measurement. The lyophilized powder was weighed, MilliQ water was added in portions, and the mixture was bubbled with N_2 for 30 s. Upon visual disappearance of the precipitate a range of solubility was calculated. Water addition was stopped when the calculated solubility was < 0.5 mg/mL.

Compound	Mass (mg)	Volume upon dissolution (μL)	Solubility (mg/mL)
D_5	0.8	60	≥ 13.3
D_6	1.1	1000	1
D_7	0.9	-	< 0.5
L_6	1.1	1000	≥ 1
L_3D_3	1.0	20	≥ 50
L_2D_4	1.0	20	≥ 50
LD_6	1.0	200	≥ 5
LD_6L	1.1	2000	≥ 0.6
LD_5L	1.0	600	≥ 1.7

Table S3.1 Solubility measurement.

Recrystallization. The lyophilized compounds were dissolved in DMSO (10 mg/mL) under sonication in a heated bath at 40 °C until a clear solution was obtained. Crystallization was induced by the addition of MeOH (5 times the volume of DMSO used) to the DMSO solution. After 16 h the suspension was centrifuged, the supernatant removed and the solid washed three times with MeOH.

Compound	Dissolved in DMSO?	Precipitation rate
D_6	Yes	Fast
D_6 + L_6 (1:1)	Yes	Slow
L_3D_3	Yes	Slow
L_2D_4	Yes	Slow
LD_6	Yes	Fast
LD_6L	Yes	Fast
LD_5L	Yes	Intermediate

Table S3.2 Fast: within 5 min. Intermediate: within 2-3 h. Slow: more than 16 h.

XRD analysis. X-ray diffraction experiments were carried out using a D8 Avance diffractometer (Bruker) in reflection mode with monochromatic Cu Kα radiation (λ= 1.5418 Å) generated at 40 kV and 40 mA (Siemens X-ray tube KFL CU 2K). The scans were performed in the scattering angle range between 4°and 40° with a step of 0.02°and an accumulation time of 6 or 10 s. Raw XRD profiles were corrected by subtraction of the sample holder signal, smoothing and baseline correction. The oligosaccharide samples were lyophilized prior to XRD measurement.

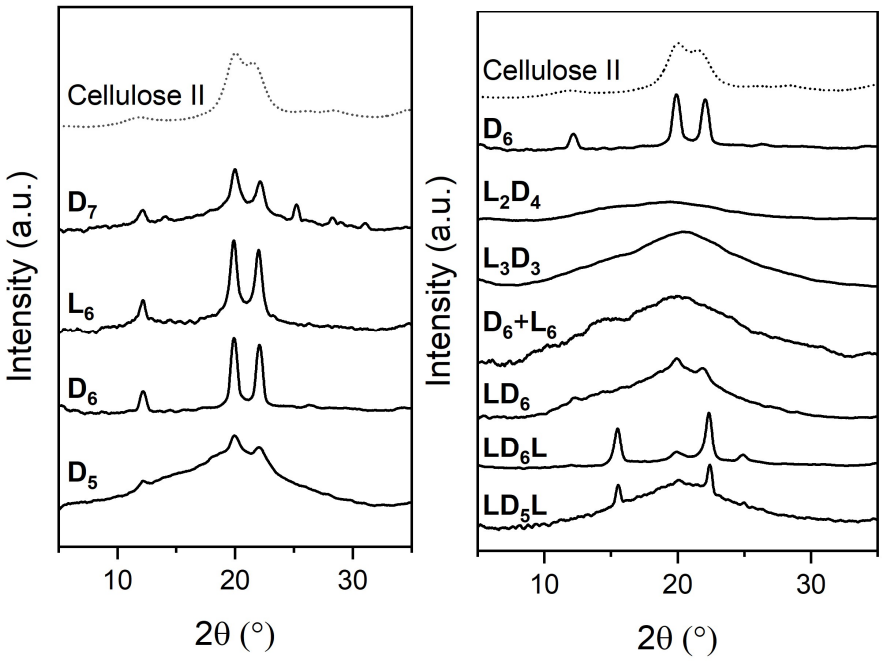


Figure S3.4 XRD profiles obtained from lyophilized powder after synthesis.

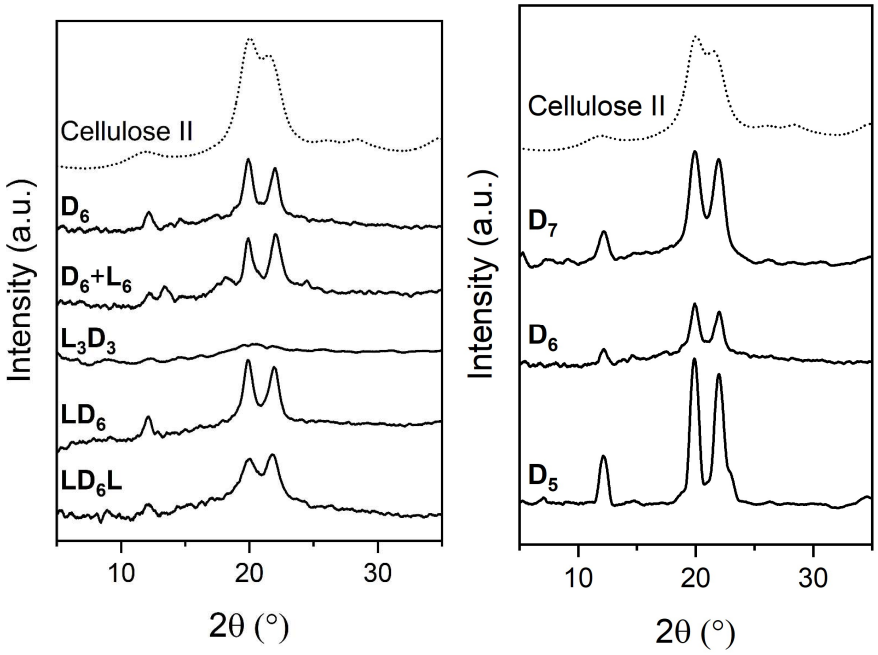


Figure S3.5 XRD profiles obtained from lyophilized powders after recrystallization.

TEM imaging. Transmission electron microscopy was performed using a JEM 2100Plus transmission electron microscope (Jeol, Japan) operated at an accelerating voltage of 200 kV or a FEI Talos L120C (Thermo Fisher, USA) operated at an accelerating voltage of 120 kV (LaB_6 cathode). Drops (3-4 μL) of aqueous suspensions (or MeOH suspensions) of crystallites were deposited on glow-discharged carbon-coated copper grids. Negative staining (2% uranyl acetate aqueous solution) was applied (3 μL drop) to the grid after the sample was deposited, allowed to settle for approximately 1 min, and then carefully blotted away with filter paper. Negative staining of the TEM grids has been used to enhance contrast only where specified in the images reported below.[256]

D_6

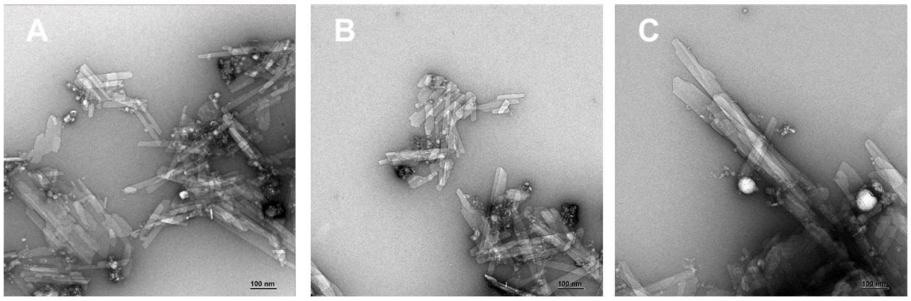


Figure S3.6 Representative TEM images of **D_6** obtained from aqueous suspension (1 mg/mL) using negative staining.

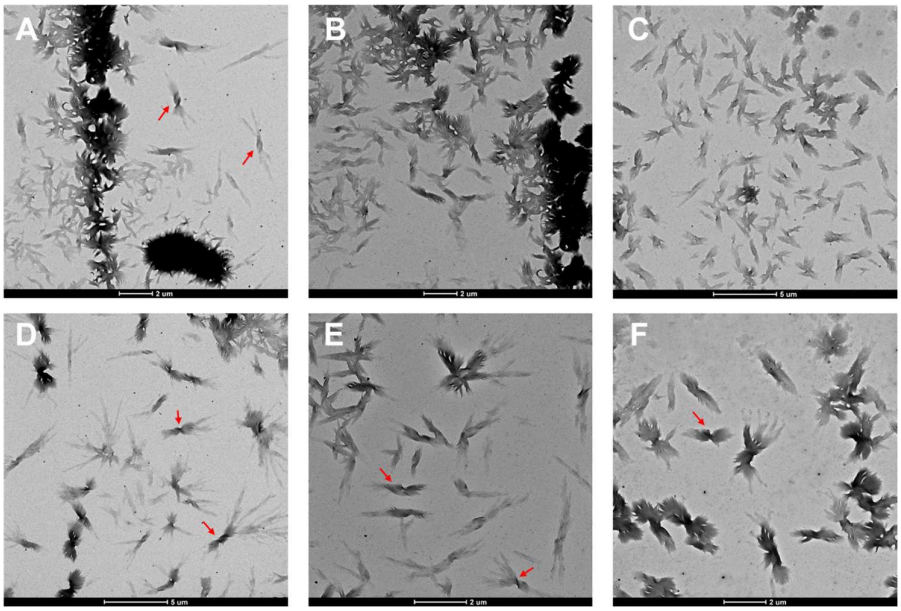


Figure S3.7 Representative TEM images of **D_6** obtained from aqueous suspension (1 mg/mL). The red arrows indicate the assemblies where the features indicating a twisted morphology are more evident.

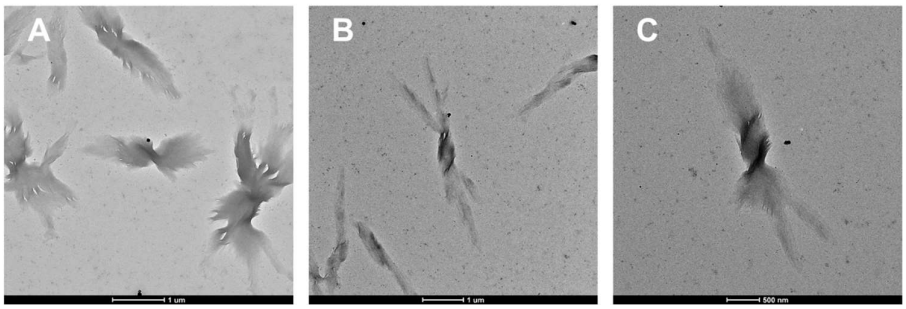


Figure S3.8 Representative TEM images of $\mathbf{D_6}$ twisted assemblies at high magnifications obtained from aqueous suspension (1 mg/mL).

L_6

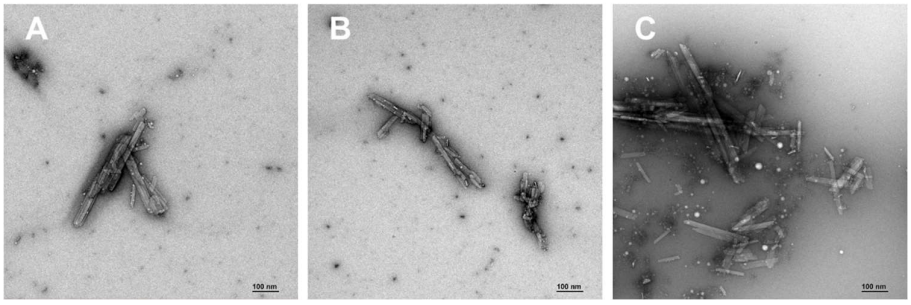


Figure S3.9 Representative TEM images of $\mathbf{L_6}$ obtained from aqueous suspension (1 mg/mL) using negative staining.

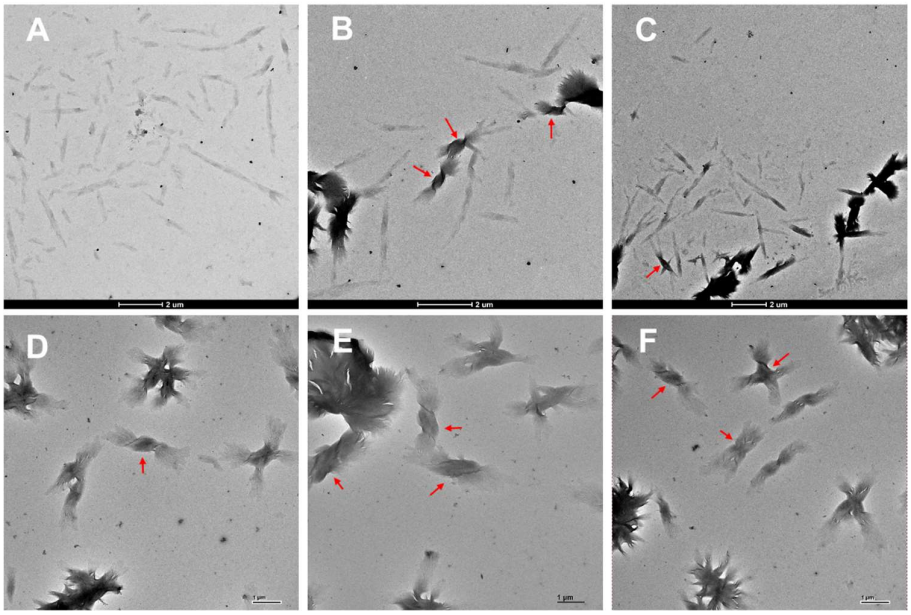


Figure S3.10 Representative TEM images of $\mathbf{L_6}$ obtained from aqueous suspension (1 mg/mL). The red arrows indicate the assemblies where the features indicating a twisted morphology are more evident.

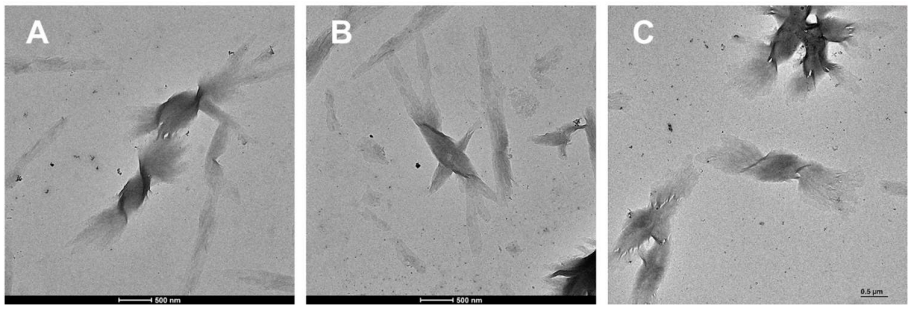


Figure S3.11 Representative TEM images of $\mathbf{L_6}$ twisted assemblies at high magnifications obtained from aqueous suspension (1 mg/mL).

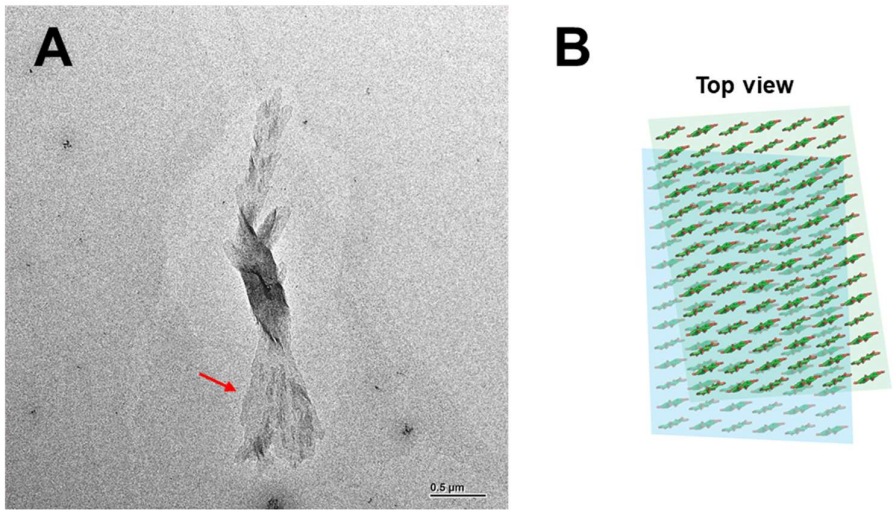


Figure S3.12 A) Representative TEM images of $\mathbf{L_6}$ obtained from aqueous suspension (1 mg/mL) showing a fan-like arrangement of the stacking platelets (*red arrow*). B) The fan-like arrangement of the stacking platelets was interpreted as a rotation between the (001) planes.

$\mathbf{D_7}$

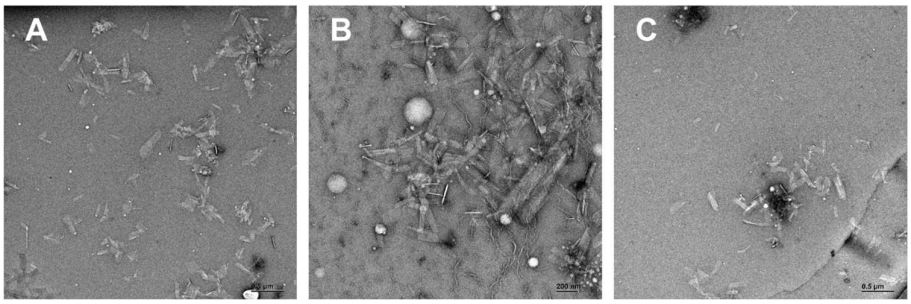


Figure S3.13 Representative TEM images of $\mathbf{D_7}$ obtained from aqueous suspension (1 mg/mL) using negative staining.

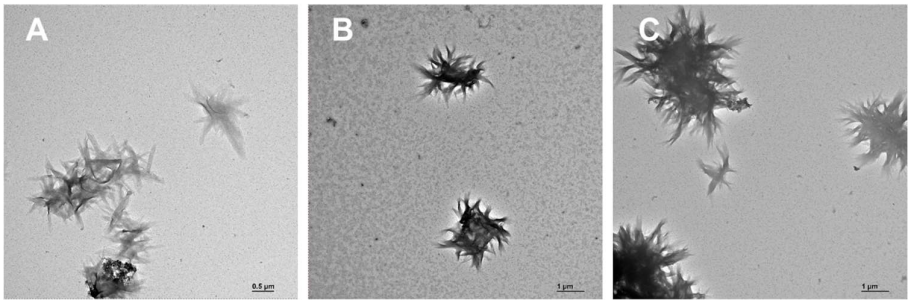


Figure S3.14 Representative TEM images of $\mathbf{D_7}$ obtained from aqueous suspension (1 mg/mL).

$\mathbf{D_8}$

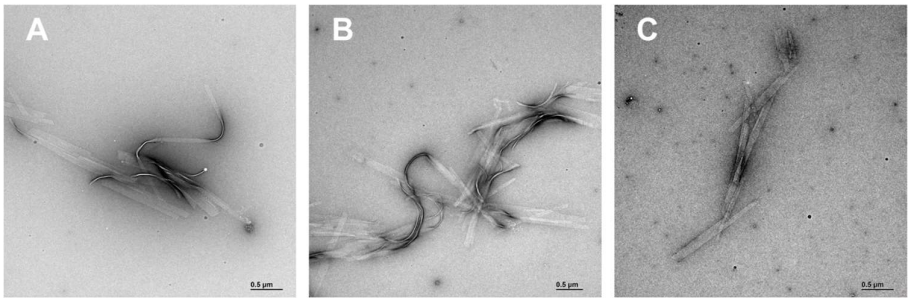


Figure S3.15 Representative TEM images of $\mathbf{D_8}$ obtained from aqueous suspension (1 mg/mL) using negative staining.

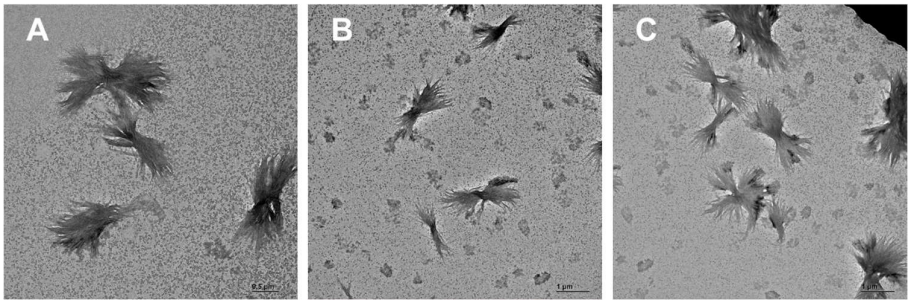


Figure S3.16 Representative TEM images of $\mathbf{D_8}$ obtained from aqueous suspension (1 mg/mL).

$\mathbf{D_6 + L_6}$

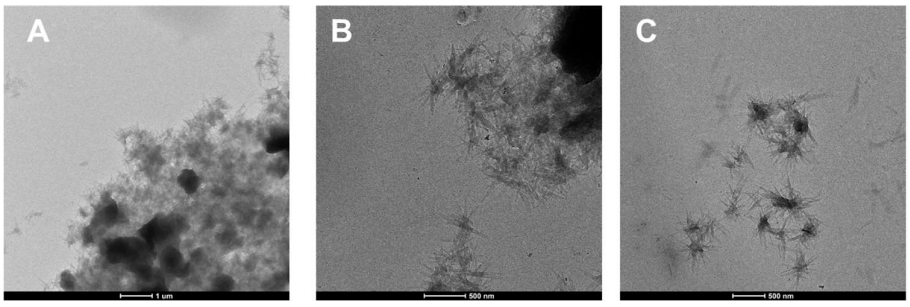


Figure S3.17 Representative TEM images of $\mathbf{D_6 + L_6}$ (1:1) obtained from MeOH suspension after recrystallization.

LD_6

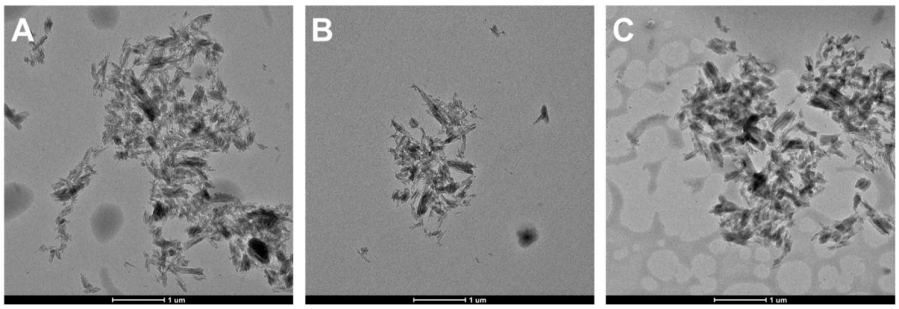


Figure S3.18 Representative TEM images of **LD_6** platelets obtained from MeOH suspension after recrystallization.

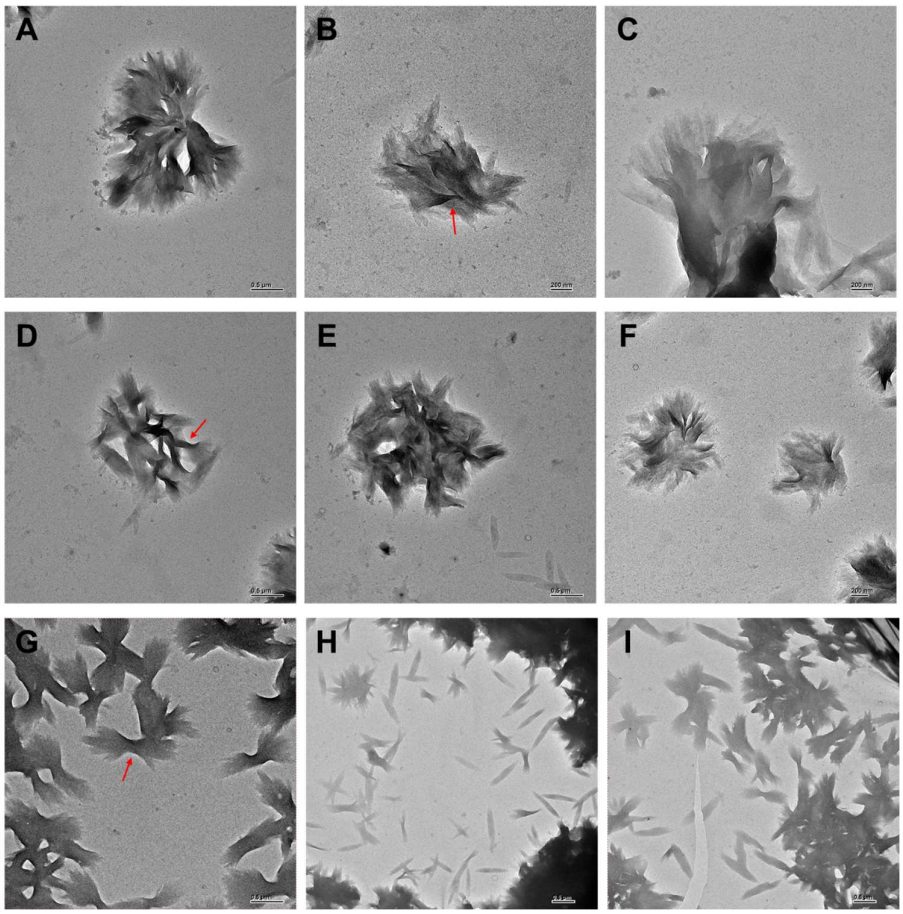


Figure S3.19 Representative TEM images of **LD_6** bundles obtained from aqueous suspension (1 mg/mL).

LD_6L

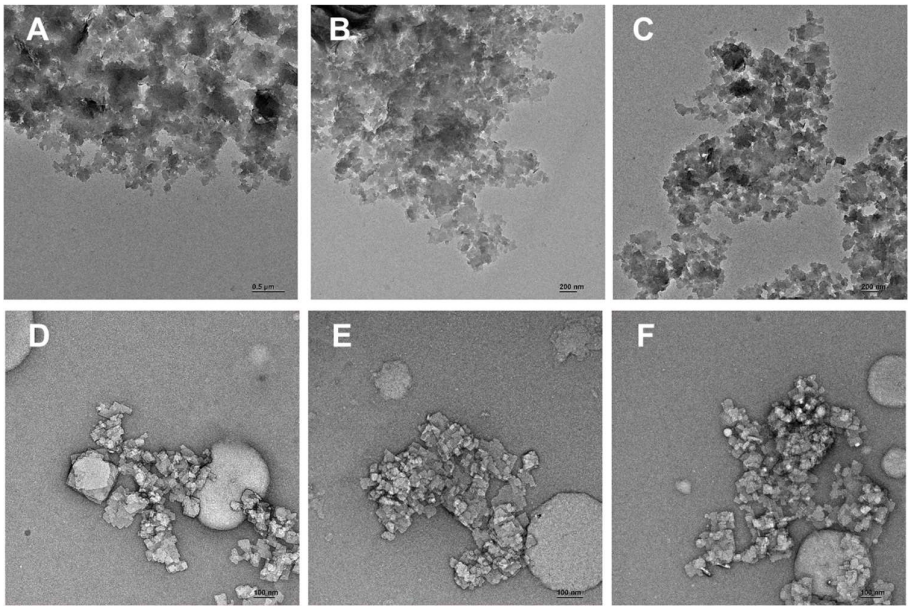


Figure S3.20 A-C) Representative TEM images of **LD_6L** obtained from aqueous suspension (1 mg/mL). D-F) Representative TEM images of **LD_6L** obtained from aqueous suspension (1 mg/mL) using negative staining.

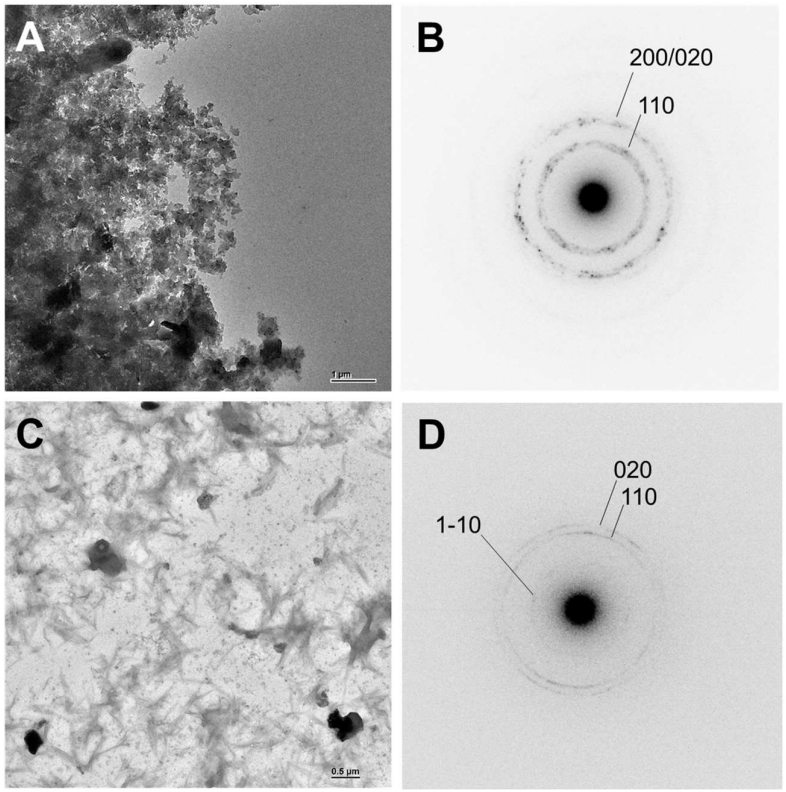


Figure S3.21 A, B) Electron diffraction analysis of **LD_6L** obtained from aqueous suspension (1 mg/mL). The pattern was assigned to the cellulose IV_{II} allomorph. C, D) Electron diffraction analysis of **LD_6L** obtained from aqueous suspension after recrystallization. The pattern was assigned to the cellulose II allomorph.

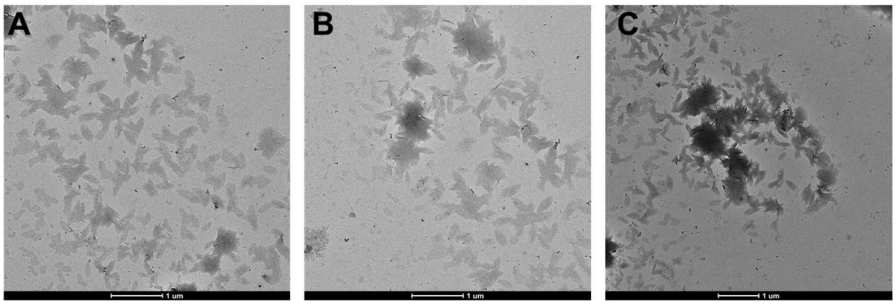


Figure S3.22 Representative TEM images of **LD_6L** platelets obtained from MeOH suspension after recrystallization.

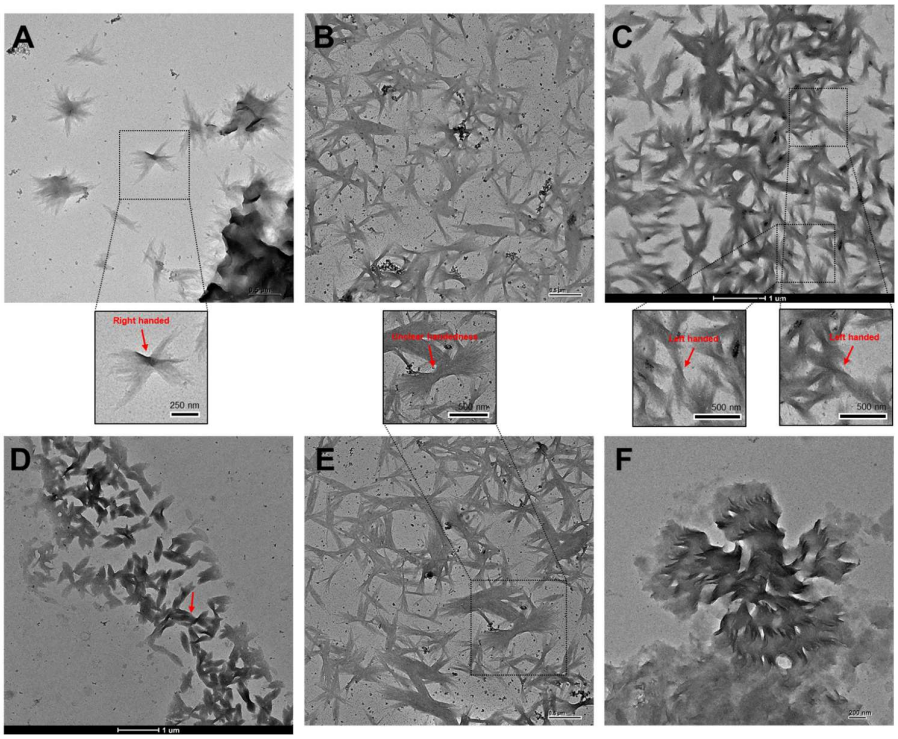


Figure S3.23 Representative TEM images of **LD_6L** bundles obtained from MeOH suspension after recrystallization.

LD_5L

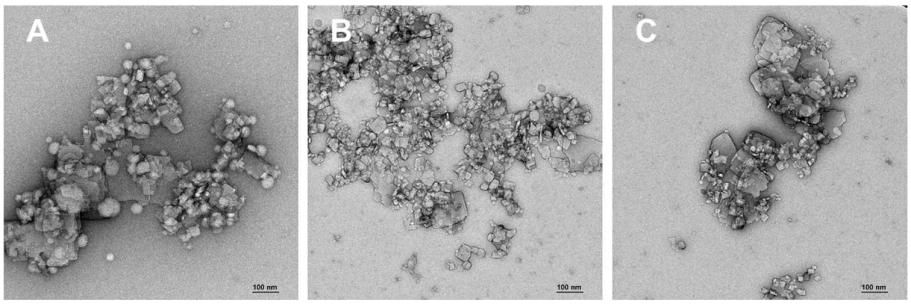


Figure S3.24 Representative TEM images of **LD_5L** obtained from aqueous suspension (1 mg/mL) using negative staining.

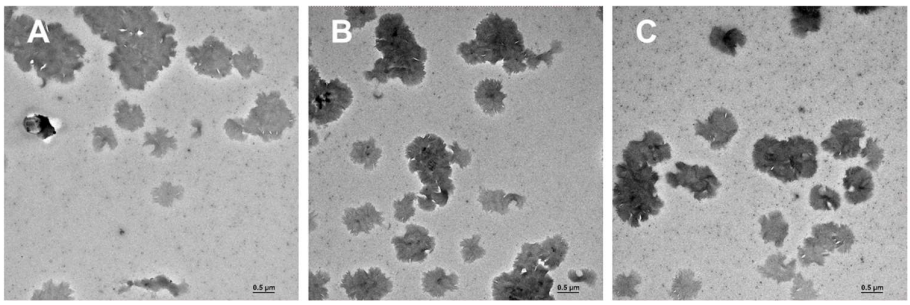


Figure S3.25 Representative TEM images of **LD$_5$L** obtained from MeOH suspension after recrystallization.

AFM imaging. Atomic force microscopy was performed with a JPK NanoWizard 4 AFM in tapping mode (AC mode) or a Dimension ICON instrument (Bruker) in pulse force (PeakForce) mode using SNL-10 A tip (0.35 N/m, 65 kHz, Bruker) or Arrow NCR tip (42 N/m, 285 kHz, Nano World). The samples were prepared as follows: approximately 0.1 mg of the lyophilized powder was weighed (in a glass or plastic vial), diluted with MilliQ water to reach the concentration of 1 mg/mL. For AFM imaging, the solution was further diluted with MilliQ water to reach the concentration of 0.1 mg/mL. Drops of aqueous suspensions were deposited on freshly cleaved mica or on glow-discharged (0.8 mbar, 30 mA for 20 s using air) silicon wafer and dried at room temperature. AFM images were collected with 1024 x 1024 pixels/frame and analyzed with the JPK Data Processing software.

The AFM analysis of **D$_7$** and **D$_8$** was performed by Dr. Denisa Vargová.

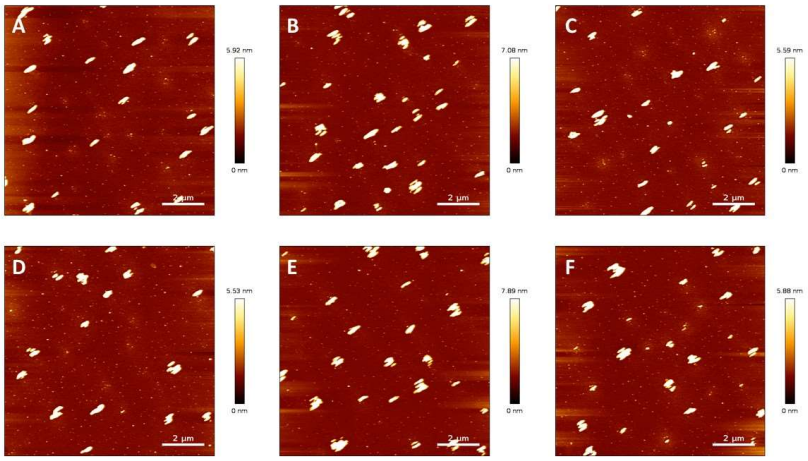


Figure S3.26 Representative AFM images of **D$_6$** obtained from aqueous suspension (0.1 mg/mL) drop casted on freshly cleaved mica.

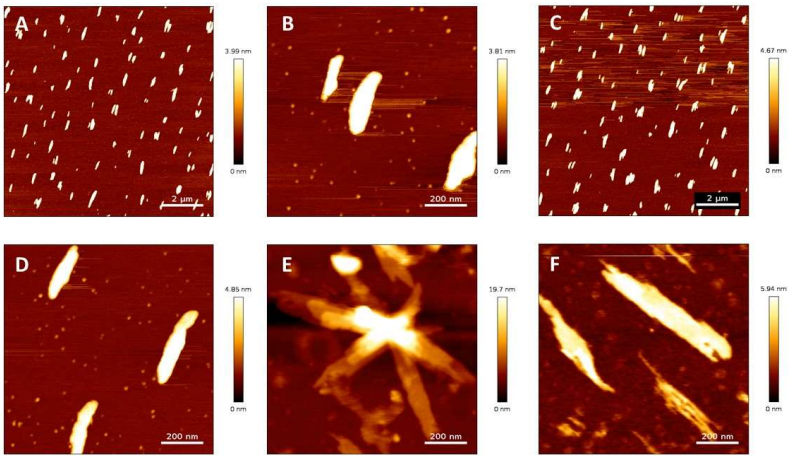


Figure S3.27 Representative AFM images of $\mathbf{D_7}$ obtained from aqueous suspension (0.1 mg/mL) drop casted on freshly cleaved mica.

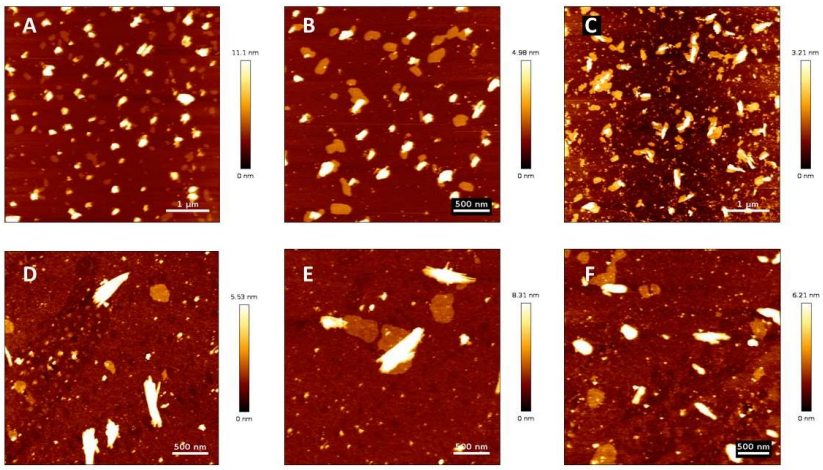


Figure S3.28 Representative AFM images of $\mathbf{D_8}$ obtained from aqueous suspension (approx. 0.1 mg/mL) drop casted on freshly cleaved mica (A-C) and silicon wafer (D-F).

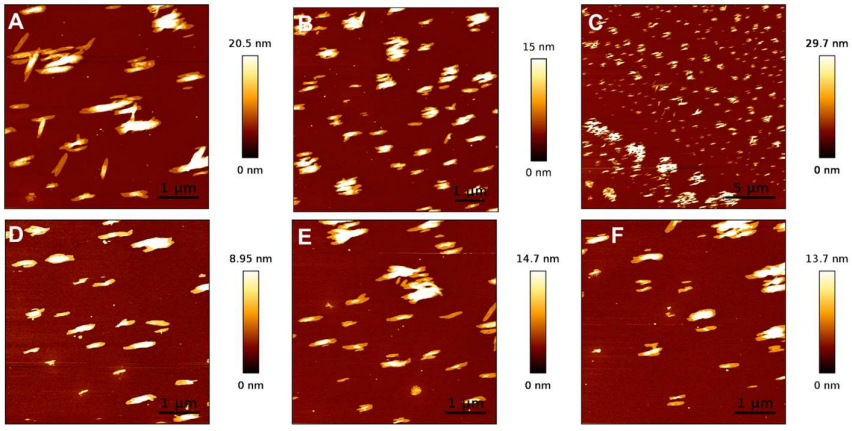


Figure S3.29 Representative AFM images of $\mathbf{L_6}$ obtained from aqueous suspension (0.1 mg/mL) drop casted on freshly cleaved mica.

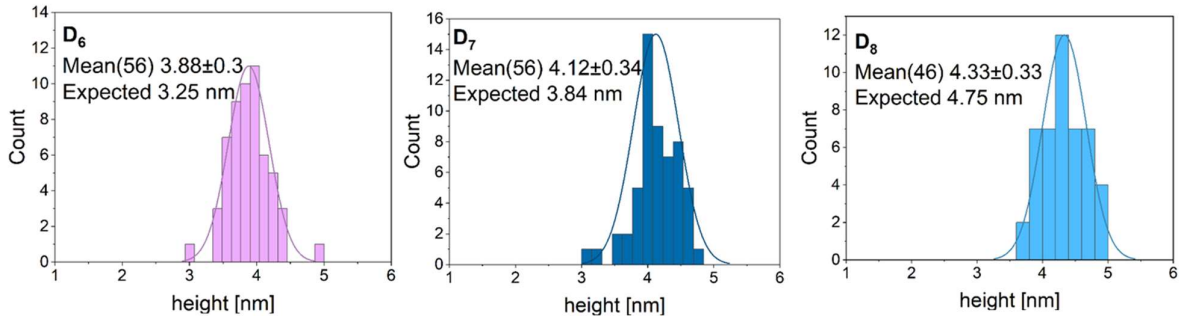


Figure S3.30 AFM height histograms for single-layer platelets. The height of the platlets was measured manually to avoid artifacts caused by edges or overlays.

Chirality analysis

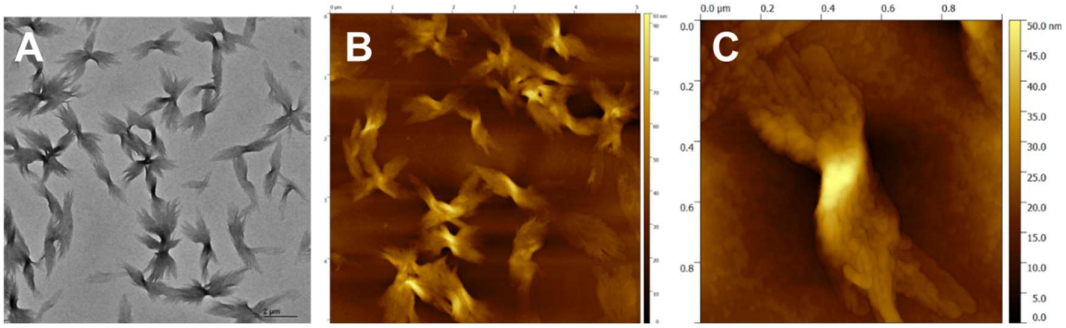


Figure S3.31 A) Representative TEM image obtained from $\mathbf{D_6}$ aqueous suspension (1 mg/mL). B-C) AFM height image obtained from the same TEM grid.

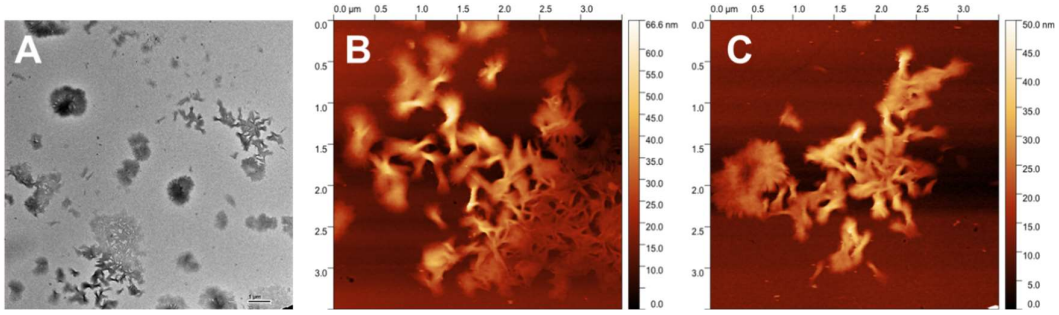
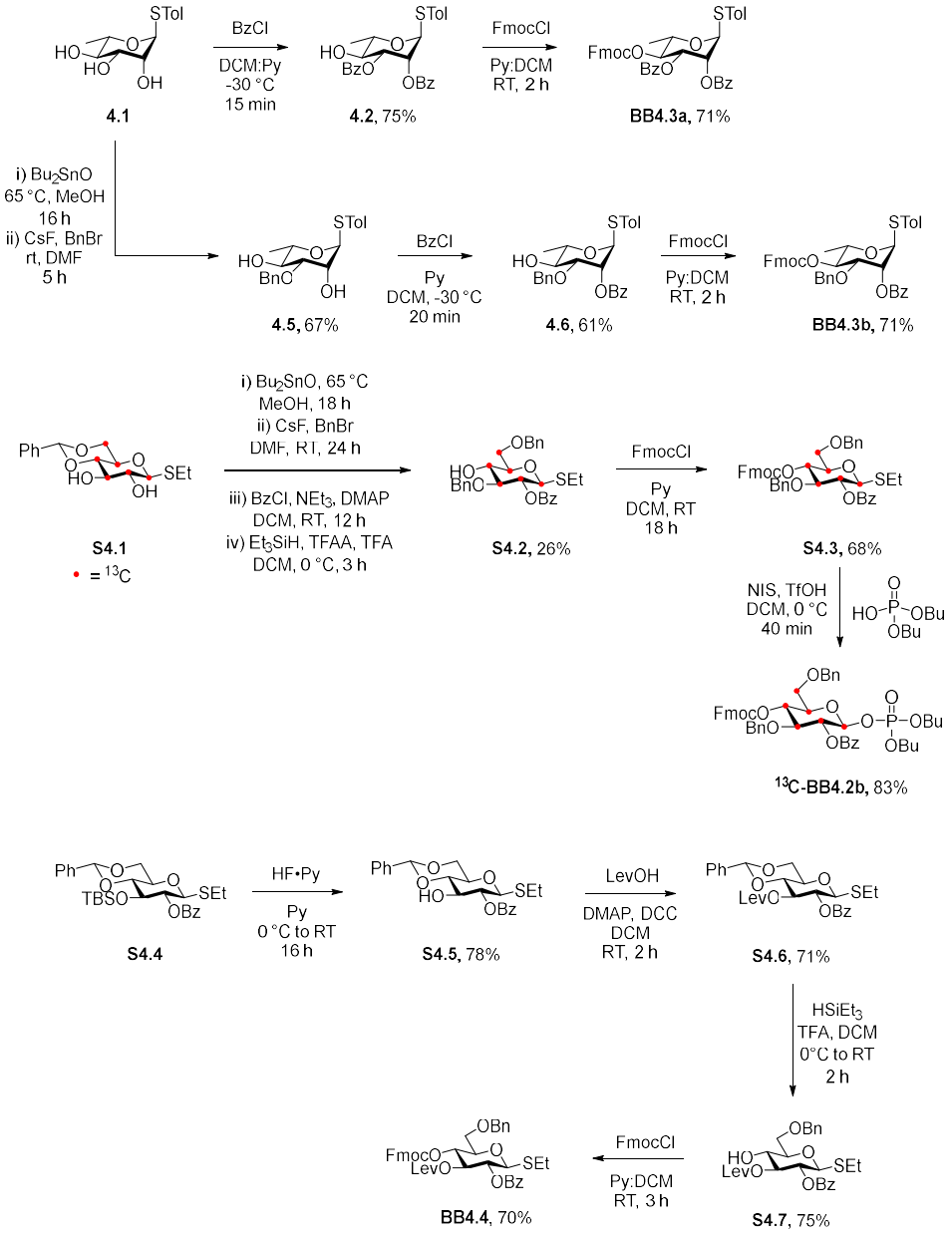


Figure S3.32 A) Representative TEM image obtained from $\mathbf{LD_6L}$ MeOH suspension. B-C) AFM height image obtained from the same TEM grid.

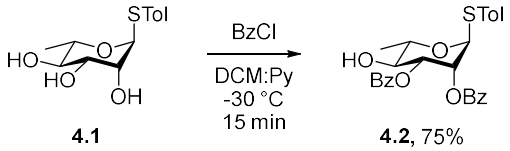
6.5 Experimental section for Chapter 4

6.5.1 Synthesis of building blocks



Scheme 4.1 Synthesis of **BB3a**, **BB3b**, **^{13}C-BB2b**, and **BB4**.

Synthesis of **4.2**



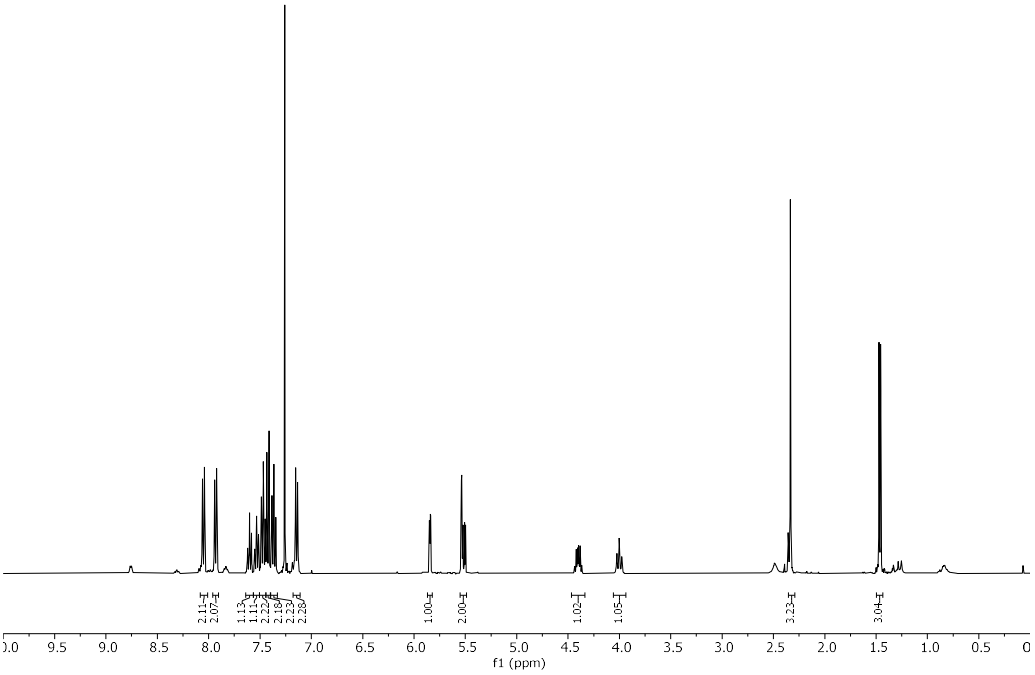
4.1 **4.2**, 75%

4.1 was synthesized according to previously reported procedures.[360]

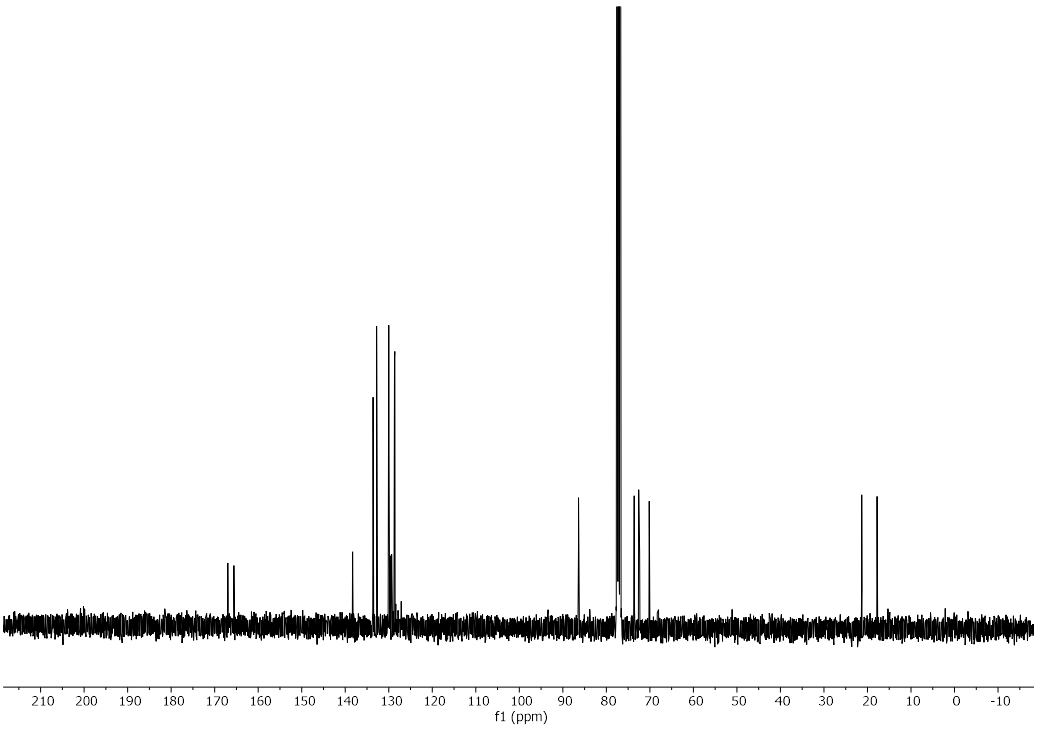
4.1 (2.27 mg, 8.4 mmol) was dissolved in a 1:1 mixture of DCM:Py anhydrous (35 mL) and cooled to -35 °C (ACN:dry ice bath) under Ar atmosphere. A solution of BzCl (2.2 mL in 6.5 mL of DCM, 19.0 mmol) was added dropwise over 5 min while stirring. The reaction was kept between -35 and -30 °C and the reaction progress was monitored every 3 min by TLC. Upon complete consumption of **4.1** (ca. 15 min), the reaction was quenched with a saturated aqueous solution of $NaHCO_3$. The reaction mixture was diluted with DCM, washed three times with saturated aqueous solution of $NaHCO_3$ and once with brine. The organic layer was dried with Na_2SO_4 and concentrated under reduce pressure. The crude product was purified by silica gel flash column chromatography (Hexane : Acetone, 4:1→3:1→2:1) to yield **4.2** as a sticky colorless solid (3.03 g, 75%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.05 (dd, *J* = 8.3, 1.2 Hz, 2H), 7.93 (dd, *J* = 8.3, 1.2 Hz, 2H), 7.63 – 7.57 (m, 1H), 7.56 – 7.50 (m, 1H), 7.47 (t, *J* = 7.7 Hz, 2H), 7.42 (d, *J* = 8.1 Hz, 2H), 7.36 (t, *J* = 7.8 Hz, 2H), 7.14 (d, *J* = 7.9 Hz, 2H), 5.84 (dd, *J* = 3.3, 1.6 Hz, 1H), 5.55 – 5.49 (m, 2H), 4.40 (dq, *J* = 9.4, 6.2 Hz, 1H), 4.00 (t, *J* = 9.6 Hz, 1H), 2.33 (s, 3H), 1.46 (d, *J* = 6.2 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.96, 165.56, 138.32, 133.63, 132.76, 130.11, 130.01, 129.96, 128.69, 128.58, 86.41, 73.59, 72.54, 72.36, 70.13, 21.30, 17.76. $[\alpha]_D^{20}$ -41.37 (c = 1.07 g/100mL, $CHCl_3$). IR (film) ν = 3495, 1726, 1280, 1095, 1071 cm^{-1}. R_f = 0.43 (Hexane : Acetone 2:1). ESI-HRMS m/z 501.1401 $[M+Na]^+$ ($C_{27}H_{26}O_6SNa$ requires 501.1342).

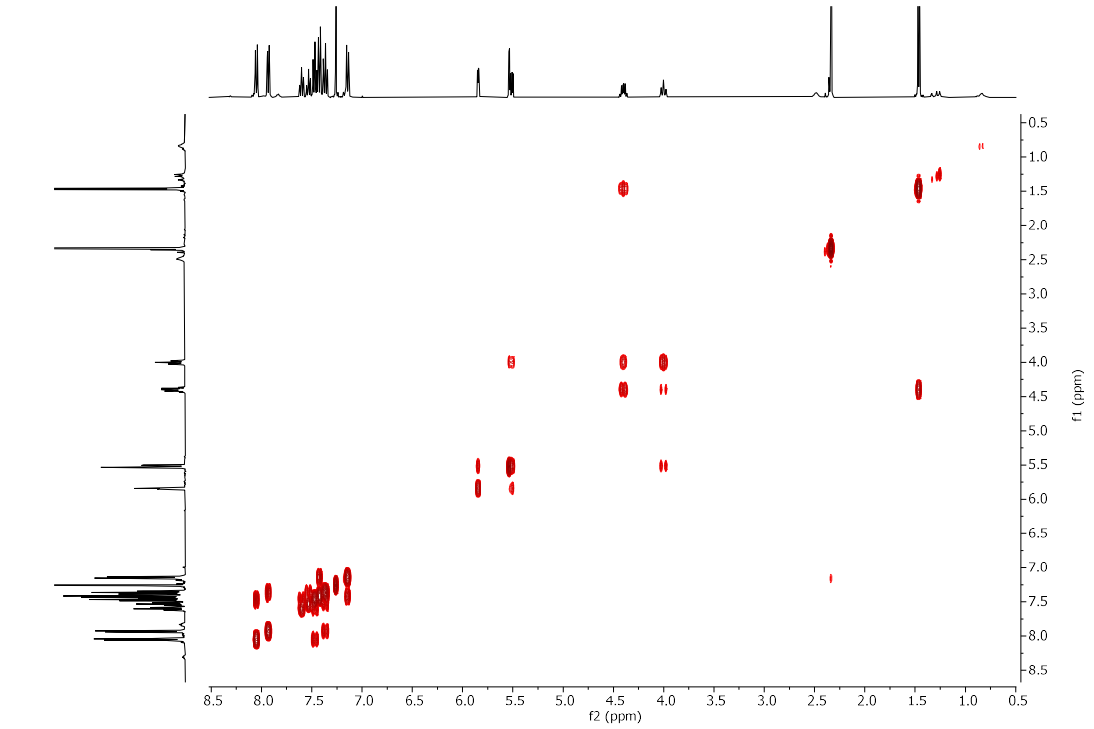
^{1}H NMR of 4.2 (400 MHz, $CDCl_3$)



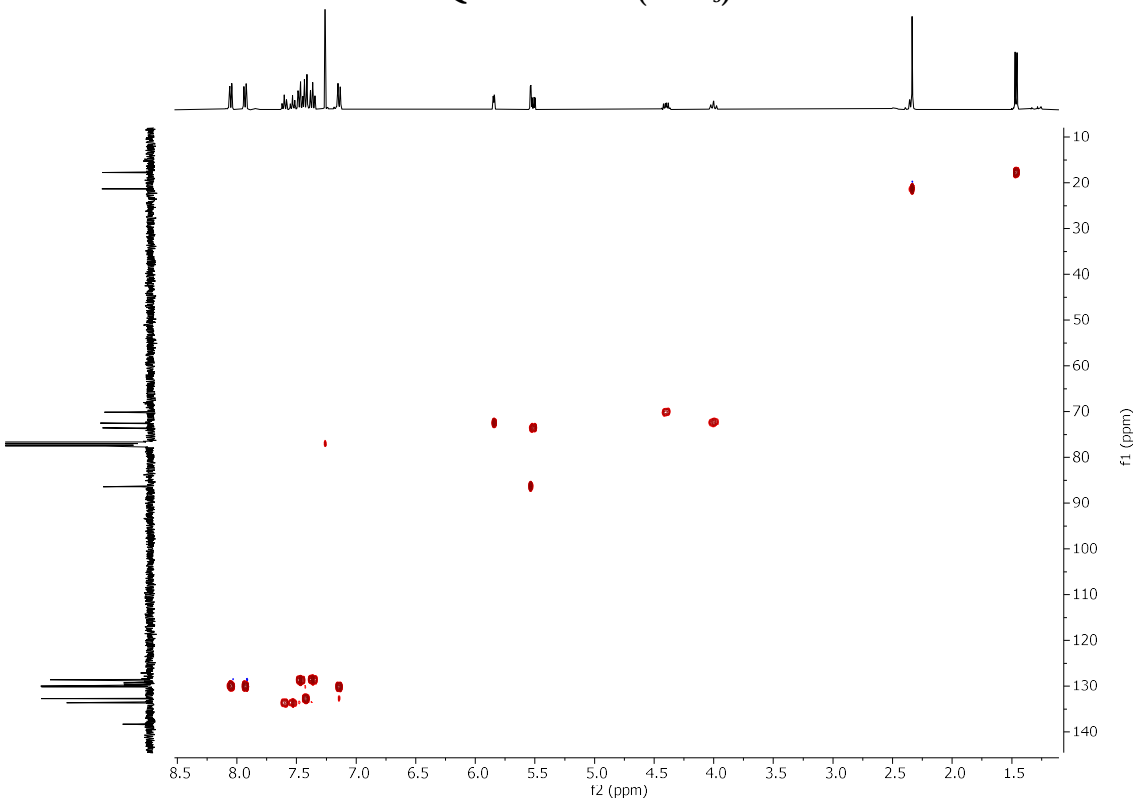
^{13}C NMR of 4.2 (101 MHz, $CDCl_3$)



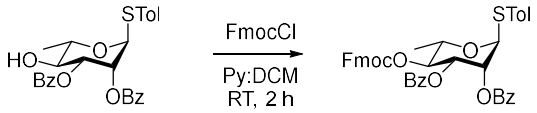
COSY NMR of 4.2 ($CDCl_3$)



HSQC NMR of 4.2 ($CDCl_3$)



Synthesis of **BB4.3a**

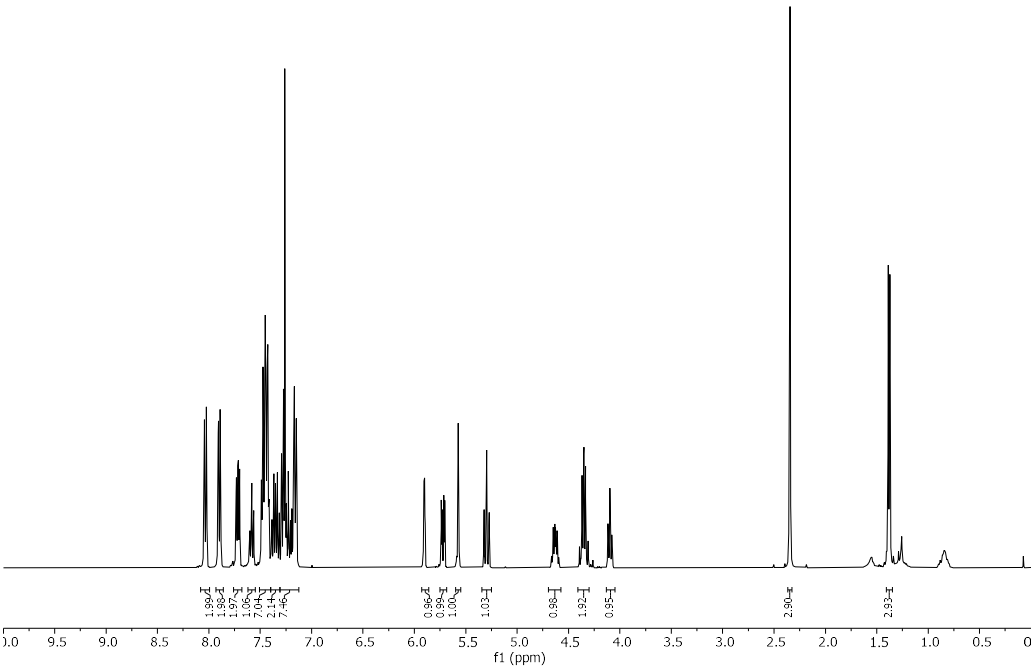


4.2 **BB4.3a**, 71%

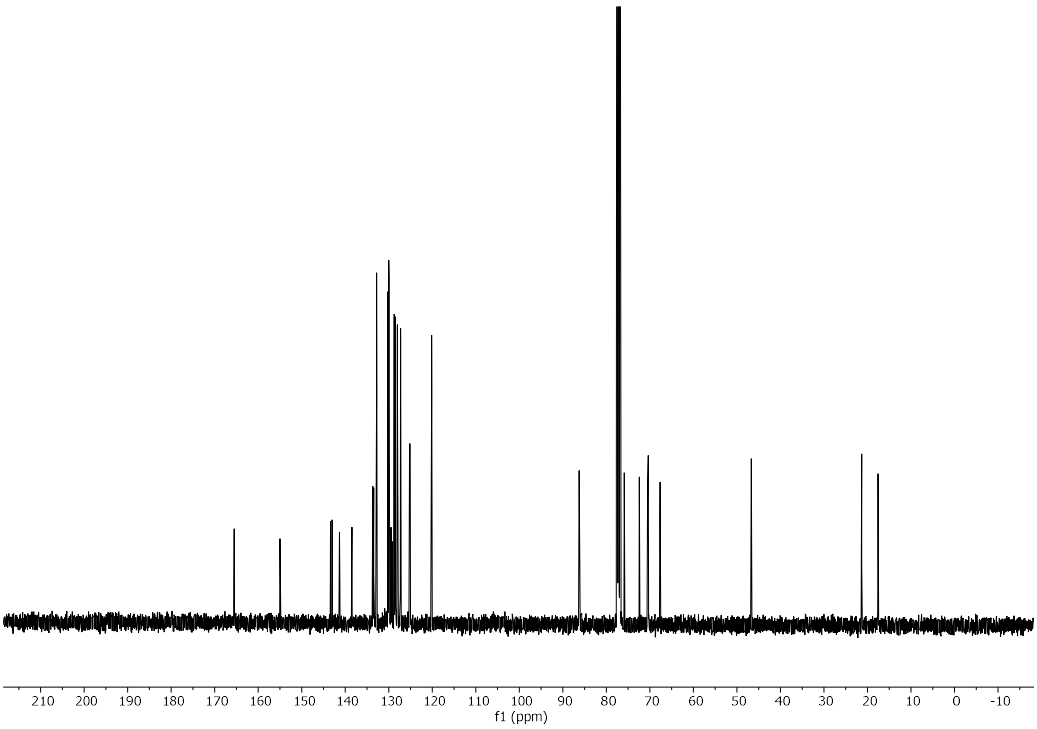
4.2 (3.03 g, 6.3 mmol) was dissolved in DCM:Py anhydrous (3:1, 30 mL) and cooled to 0 °C under Ar atmosphere. FmocCl (3.2 g, 12.4 mmol) was dissolved in DCM (5 mL) and added dropwise to the stirred reaction mixture. The solution was stirred for 30 min at 0 °C and then allowed to RT and stirred for additional 2 h, after which time the reaction was quenched with an aqueous solution of citric acid (1 M). The reaction mixture was diluted with DCM, washed twice with an aqueous solution of citric acid (1 M) and once with brine. The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by silica gel flash column chromatography (Hexane : Acetone, 4:1→3:1) and recrystallized from DCM : Hexane to yield **BB4.3a** as a sticky colorless solid (3.18 g, 71%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.06 – 8.01 (m, 2H), 7.92 – 7.87 (m, 2H), 7.72 (dd, *J* = 7.5, 5.0 Hz, 2H), 7.58 (t, *J* = 7.4 Hz, 1H), 7.50 – 7.40 (m, 7H), 7.35 (dt, *J* = 14.3, 7.5 Hz, 2H), 7.30 – 7.13 (m, 6H), 5.90 (dd, *J* = 3.3, 1.5 Hz, 1H), 5.72 (dd, *J* = 10.1, 3.3 Hz, 1H), 5.57 (d, *J* = 1.1 Hz, 1H), 5.30 (t, *J* = 9.9 Hz, 1H), 4.68 – 4.58 (m, 1H), 4.41 – 4.30 (m, 2H), 4.10 (t, *J* = 7.3 Hz, 1H), 2.35 (s, 3H), 1.38 (d, *J* = 6.2 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.54, 165.49, 154.97, 143.37, 142.99, 141.38, 141.30, 138.47, 133.66, 133.47, 132.78, 130.19, 130.01, 129.96, 129.51, 129.45, 129.19, 128.73, 128.50, 128.01, 127.30, 125.20, 125.06, 120.15, 86.25, 75.91, 72.42, 70.52, 70.36, 67.67, 46.71, 21.33, 17.56. $[\alpha]_D^{20}$ 13.48 (c = 0.78 g/100mL, $CHCl_3$). IR (film) ν = 1754, 1730, 1282, 1248, 1103, 711 cm^{-1}. R_f = 0.54 (Hexane : Acetone 2:1). ESI-HRMS m/z 723.2101 $[M+Na]^+$ ($C_{42}H_{36}O_8SNa$ requires 723.2023).

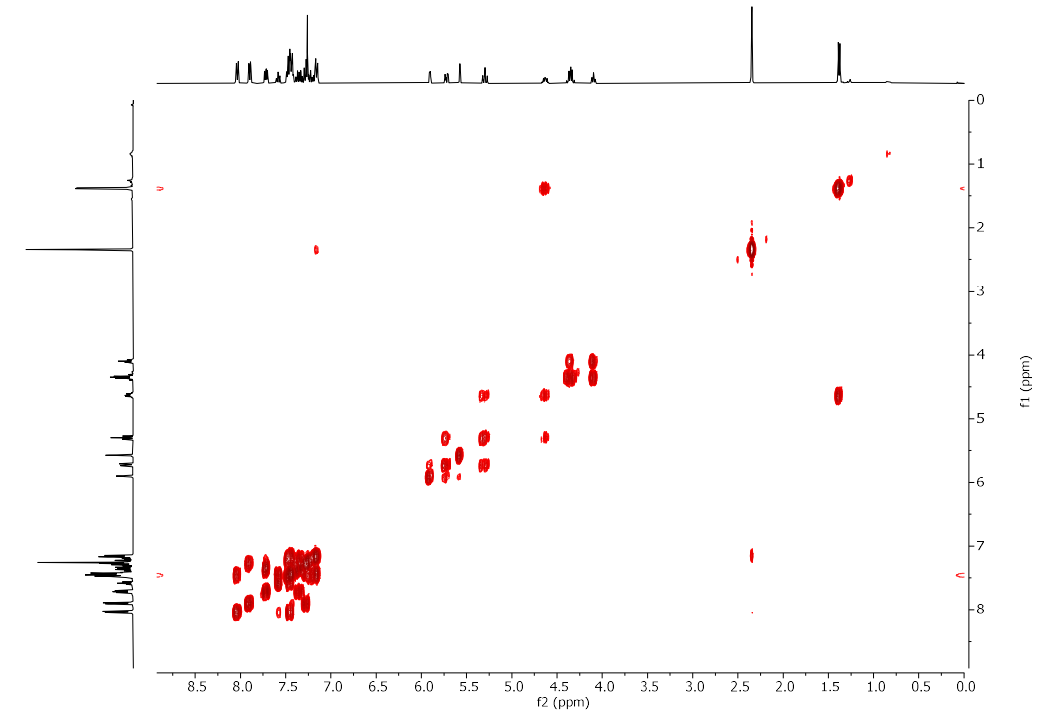
^{1}H NMR of BB4.3a (400 MHz, $CDCl_3$)



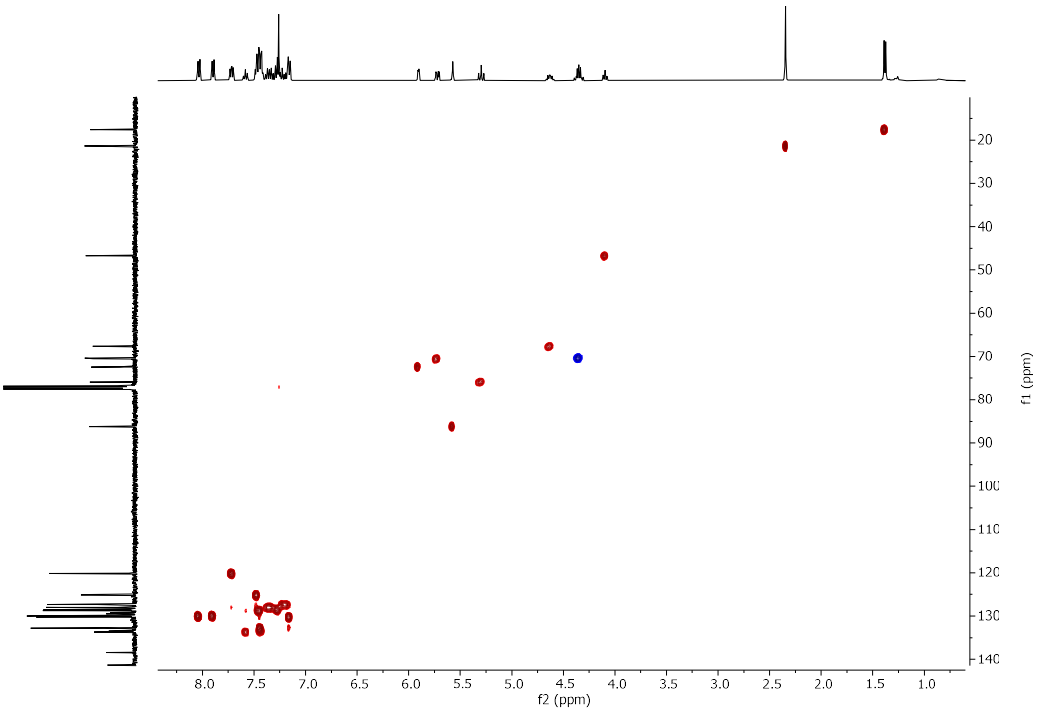
^{13}C NMR of BB4.3a (101 MHz, $CDCl_3$)



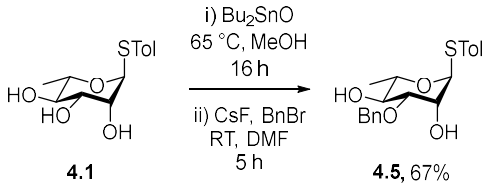
COSY NMR of BB4.3a ($CDCl_3$)



HSQC NMR of BB4.3a ($CDCl_3$)



Synthesis of **4.5**

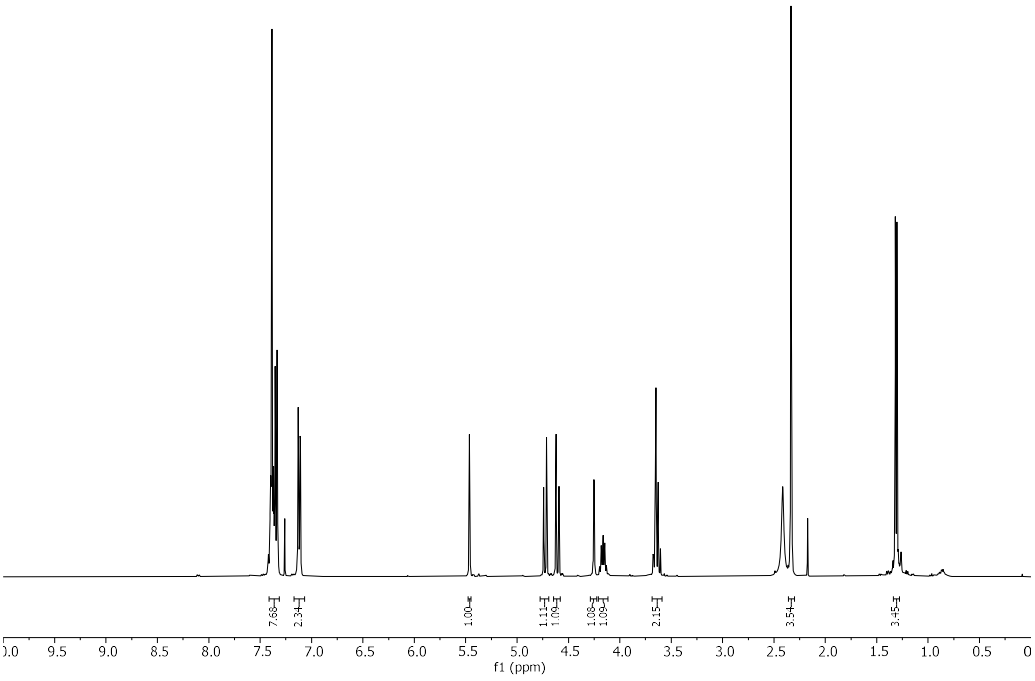


4.1 was synthesized according to previously reported procedures.[360]

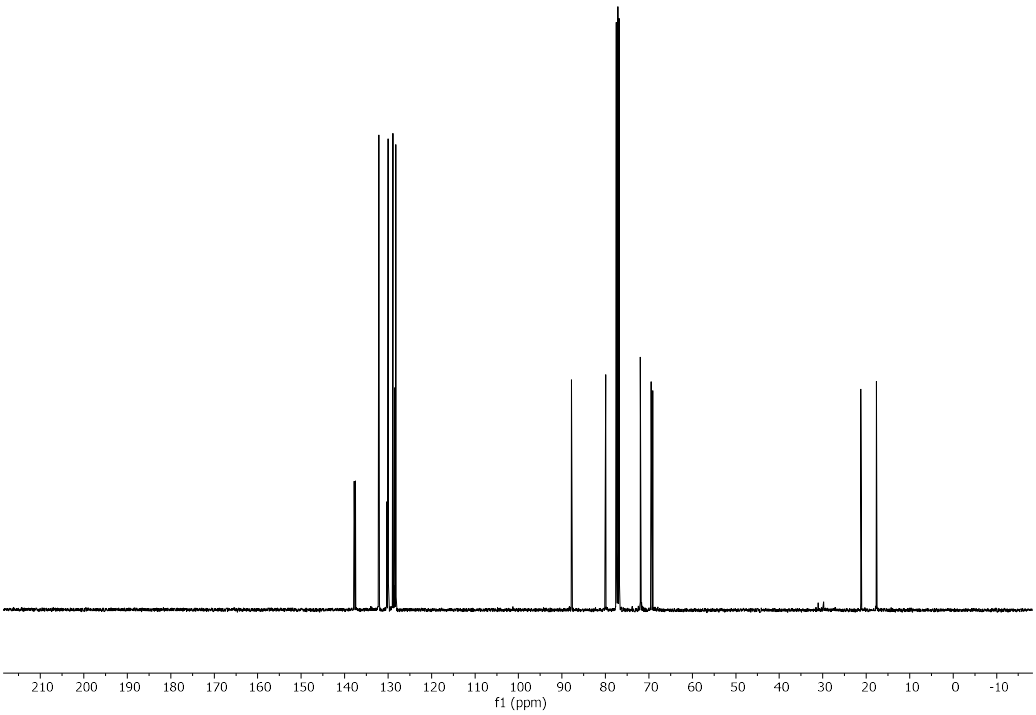
4.1 (2.98 g, 11.0 mmol) was dissolved in MeOH (100 mL), di-*n*-butyltin oxide (3.28 g, 13.2 mmol) was added and the reaction mixture (white suspension) heated at reflux at 65 °C while vigorously stirring for 16 h. The reaction mixture (clear solution) was then cooled, concentrated under reduced pressure and the crude product was used in the next step without further purification. The crude product was dissolved in DMF (50 mL), benzyl bromide (1.6 mL, 13.5 mmol) and cesium(I) fluoride (2.2 g, 14.5 mmol) were added and the clear solution stirred at RT for 5 h under Ar atmosphere. The reaction mixture (cloudy) was diluted with EtOAc and the organic layer was passed through a short plug of silica gel and dried under reduced pressure. The crude product was then diluted with EtOAc and the organic layer washed once with an aqueous solution of KF (1 M), dried with Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by silica gel flash column chromatography (Hexane : Acetone, 4:1→3:1→2:1→1:1) to yield **4.5** as a sticky colorless solid (2.67 g, 67%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.42 – 7.30 (m, 7H), 7.12 (d, *J* = 8.0 Hz, 2H), 5.49 – 5.44 (m, 1H), 4.73 (d, *J* = 11.5 Hz, 1H), 4.61 (d, *J* = 11.5 Hz, 1H), 4.27 – 4.23 (m, 1H), 4.16 (quint, *J* = 6.2 Hz, 1H), 3.70 – 3.59 (m, 2H), 2.33 (s, 3H), 1.31 (d, *J* = 6.2 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 137.81, 137.47, 132.15, 130.28, 129.98, 128.92, 128.50, 128.21, 87.78, 79.95, 71.95, 71.92, 69.47, 69.12, 21.23, 17.66. $[\alpha]_D^{20}$ -158.13 (c = 1.00 g/100mL, $CHCl_3$). IR (film) ν = 3437, 2923, 1494, 1100, 1061 cm^{-1}. R_f = 0.46 (Hexane : Acetone 2:1). ESI-HRMS m/z 383.1288 $[M+Na]^+$ ($C_{20}H_{24}O_4SNa$ requires 383.1287).

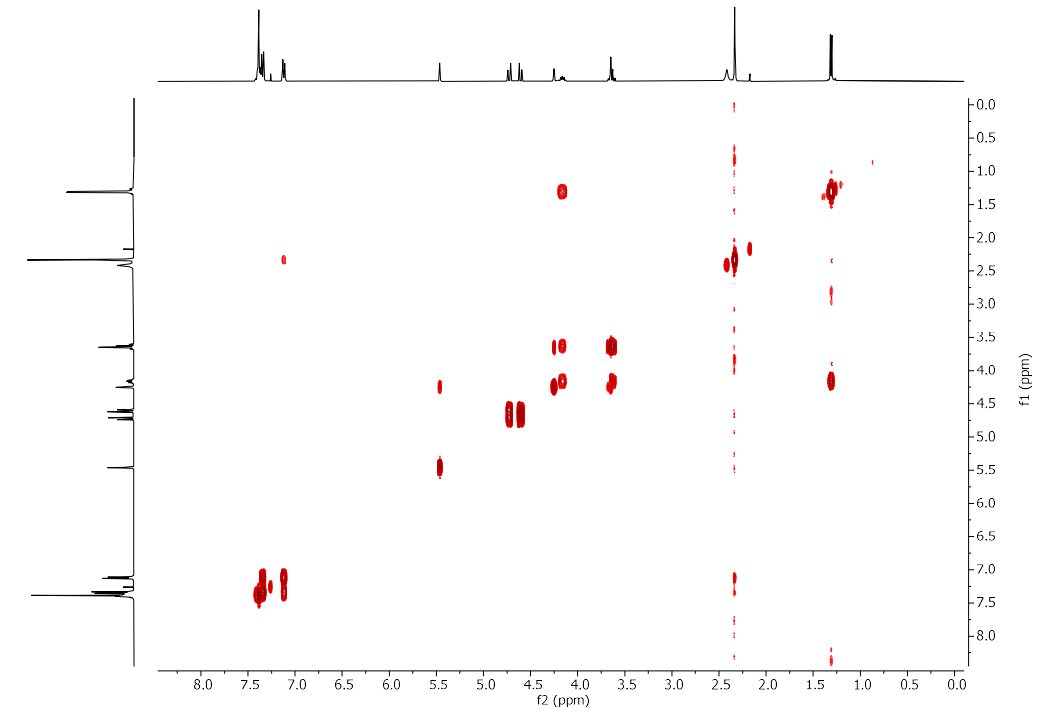
^{1}H NMR of 4.5 (400 MHz, $CDCl_3$)



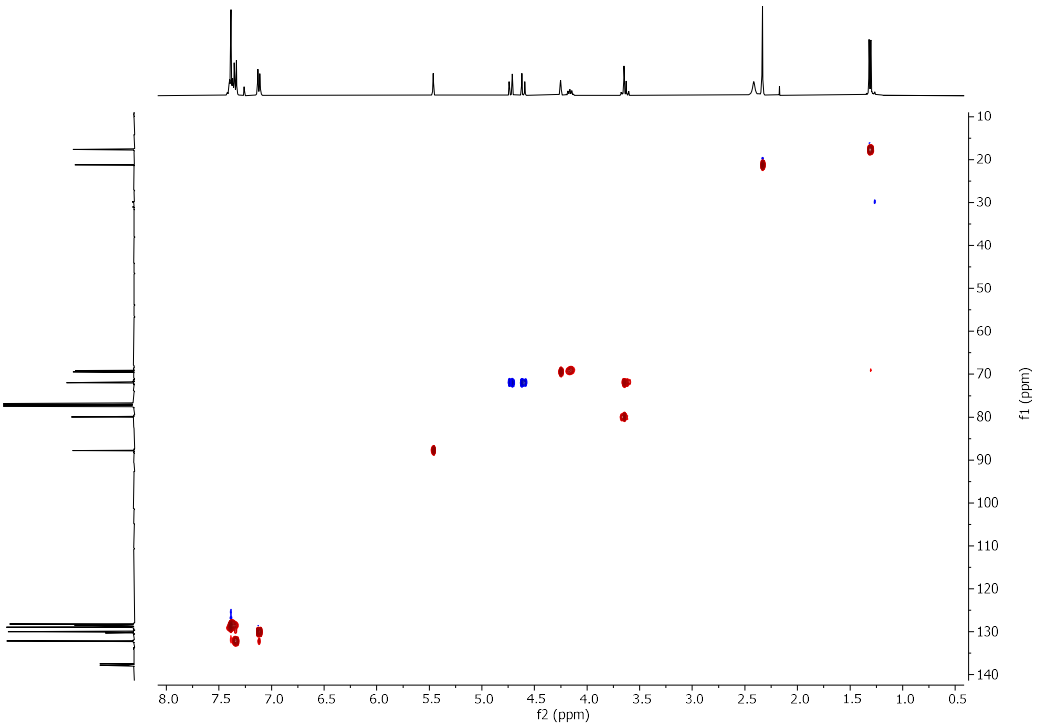
^{13}C NMR of 4.5 (101 MHz, $CDCl_3$)



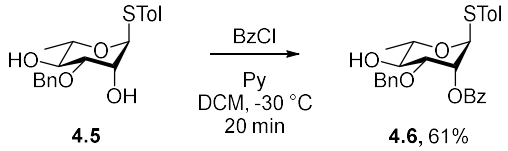
COSY NMR of 4.5 ($CDCl_3$)



HSQC NMR of 4.5 ($CDCl_3$)



Synthesis of **4.6**

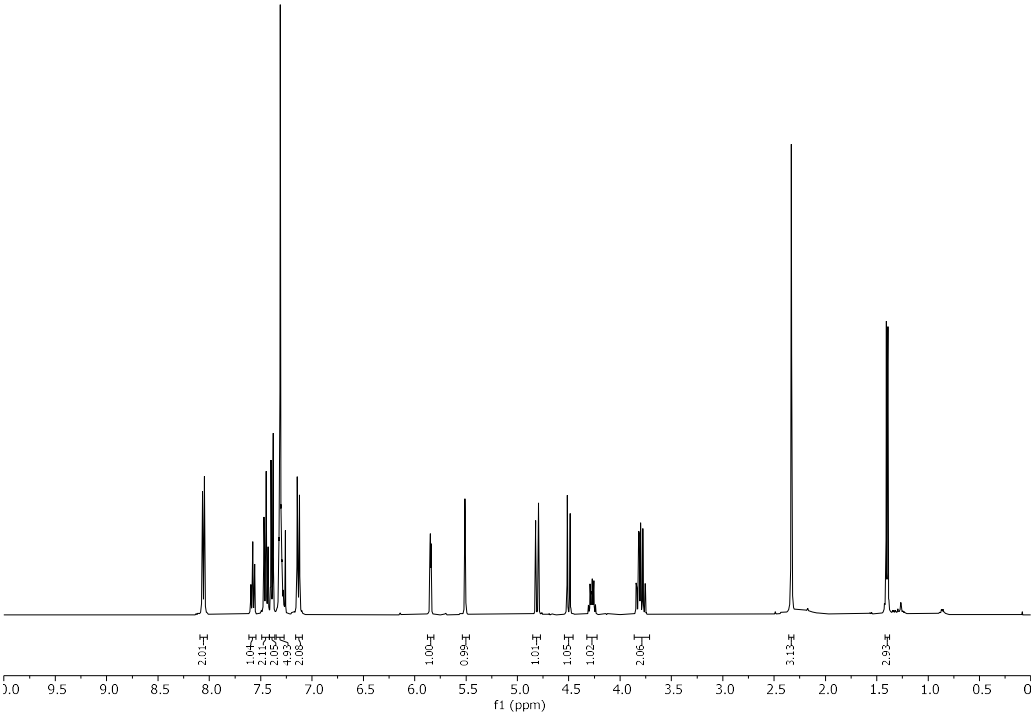


4.5 **4.6**, 61%

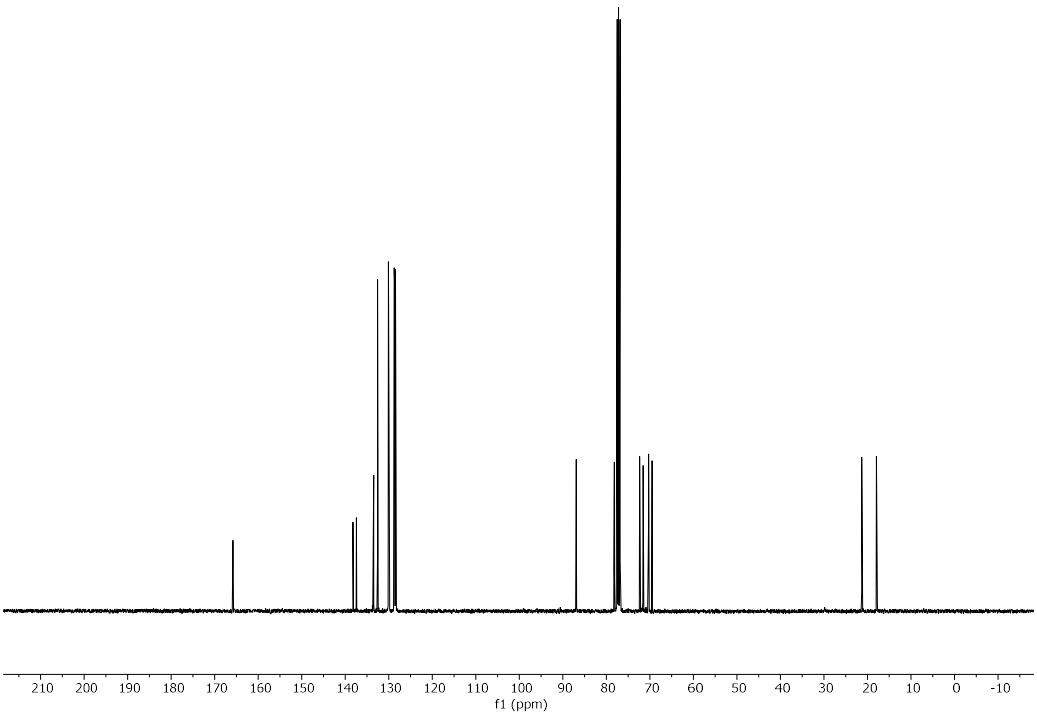
4.5 (2.07 mg, 5.7 mmol) was dissolved in a 1:1 mixture of DCM:Py anhydrous (30 mL) and cooled to -35 °C (ACN:dry ice bath) under Ar atmosphere. A solution of BzCl (790 μL in 4 mL of DCM, 6.8 mmol) was added dropwise over 5 min while stirring. The reaction was kept between -35 and -30 °C and the reaction progress was monitored every 3 min by TLC. Upon complete consumption of **4.5** (ca. 20 min), the reaction was quenched with a saturated aqueous solution of $NaHCO_3$. The reaction mixture was diluted with DCM, washed three times with saturated aqueous solution of $NaHCO_3$ and once with brine. The organic layer was dried over Na_2SO_4 and concentrated under reduce pressure. The crude product was purified by silica gel flash column chromatography (Hexane : Acetone, 4:1→3:1→2:1) and recrystallized from DCM : Hexane to yield **4.6** as a sticky colorless solid (1.61 g, 61%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.09 – 8.02 (m, 2H), 7.58 (tt, *J* = 7.0, 1.3 Hz, 1H), 7.49 – 7.42 (m, 2H), 7.42 – 7.36 (m, 2H), 7.35 – 7.27 (m, 5H), 7.13 (d, *J* = 7.9 Hz, 2H), 5.85 (dd, *J* = 2.8, 1.7 Hz, 1H), 5.51 (d, *J* = 1.3 Hz, 1H), 4.81 (d, *J* = 11.2 Hz, 1H), 4.50 (d, *J* = 11.2 Hz, 1H), 4.27 (dq, *J* = 8.6, 6.2 Hz, 1H), 3.85 – 3.73 (m, 2H), 2.33 (s, 3H), 1.40 (d, *J* = 6.2 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.77, 138.17, 137.40, 133.46, 132.54, 130.10, 130.05, 129.99, 128.73, 128.57, 128.39, 128.26, 86.87, 78.12, 72.30, 71.54, 70.27, 69.46, 21.26, 17.90. $[\alpha]_D^{20}$ -36.20 (c = 1.00 g/100mL, $CHCl_3$). IR (film) ν = 3496, 1721, 1268, 1109, 1070, 711 cm^{-1}. R_f = 0.37 (Hexane : Acetone = 3:1). ESI-HRMS m/z 487.1548 $[M+Na]^+$ ($C_{27}H_{28}O_5SNa$ requires 487.1549).

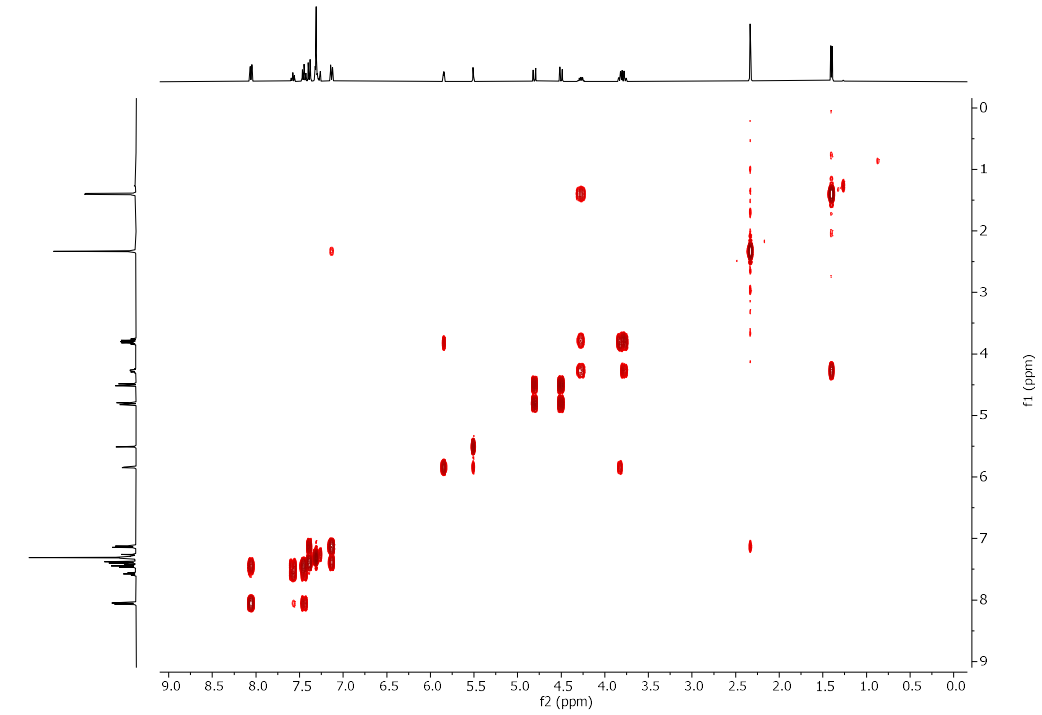
^{1}H NMR of 4.6 (400 MHz, $CDCl_3$)



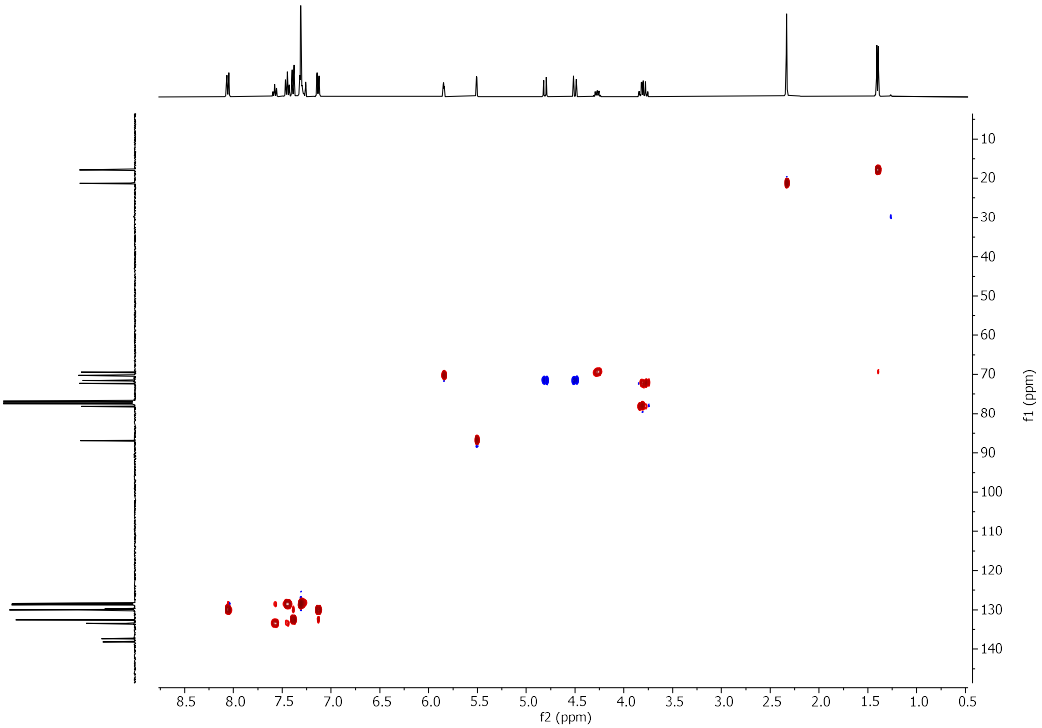
^{13}C NMR of 4.6 (101 MHz, $CDCl_3$)



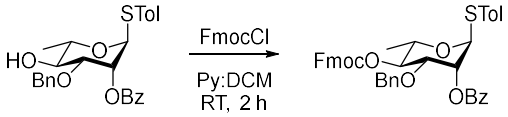
COSY NMR of 4.6 ($CDCl_3$)



HSQC NMR of 4.6 ($CDCl_3$)



Synthesis of **BB4.3b**

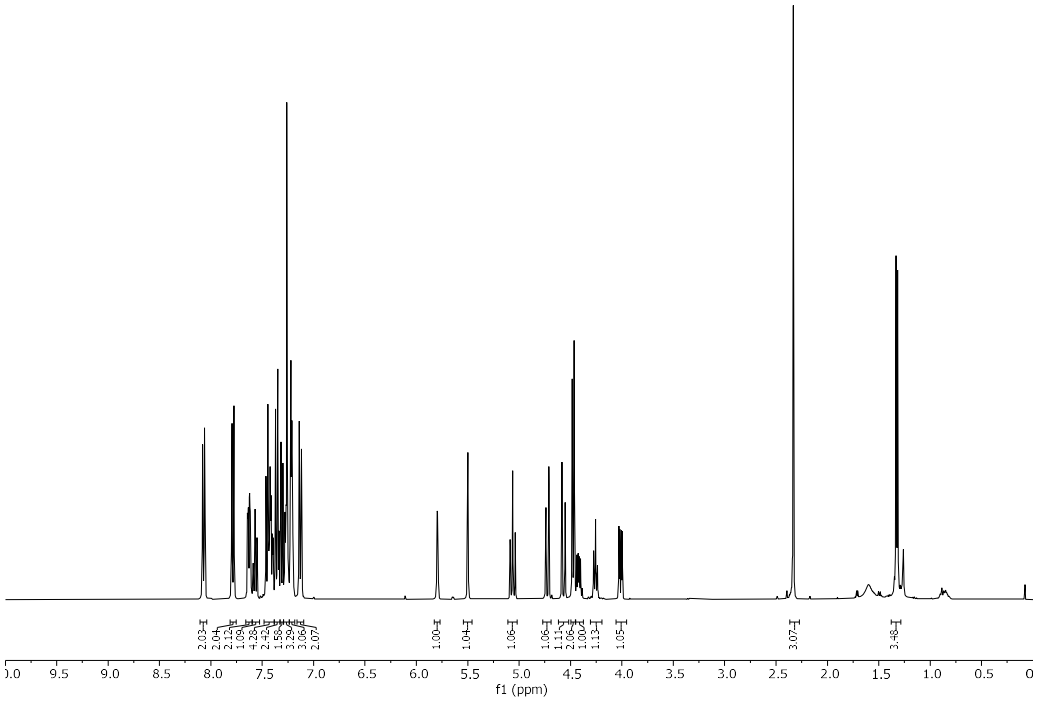


4.6 **BB4.3b**, 71%

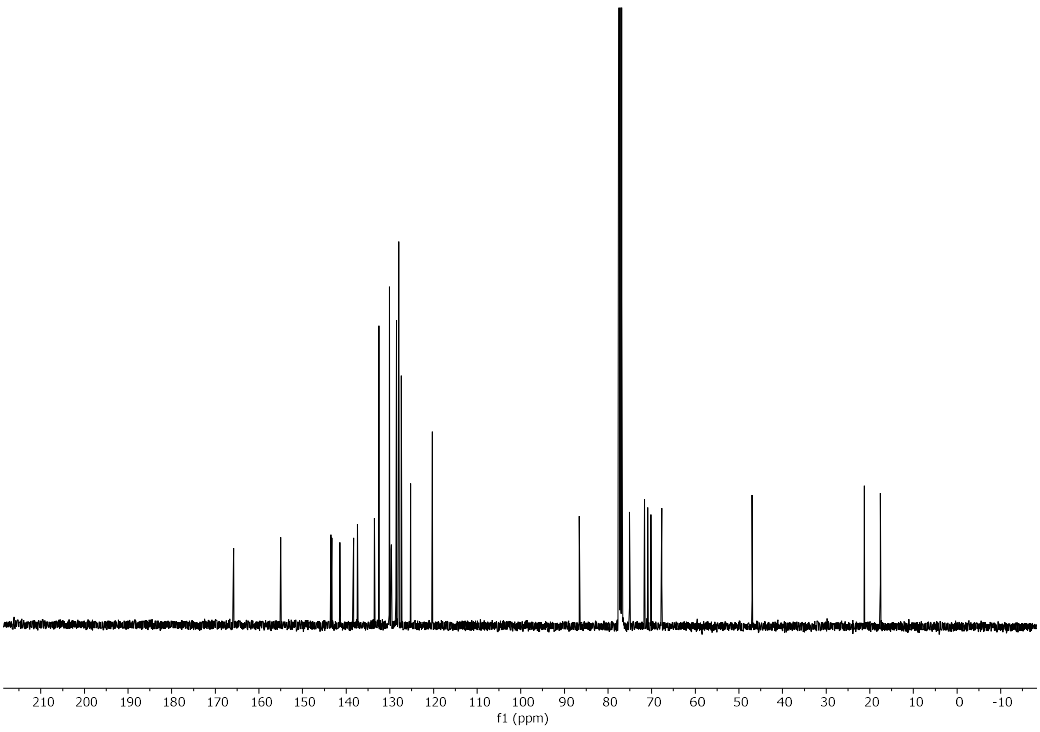
4.6 (1.61 g, 3.5 mmol) was dissolved in DCM:Py anhydrous (3:1, 20 mL) and cooled to 0 °C under Ar atmosphere. FmocCl (1.79 g, 6.9 mmol) was dissolved in DCM (6 mL) and added dropwise to the stirred reaction mixture. The solution was stirred for 15 min at 0 °C and then allowed to RT and stirred for additional 2 h, after which time the reaction was quenched with an aqueous solution of citric acid (1 M). The reaction mixture was diluted with DCM, washed once with an aqueous solution of citric acid (1 M), and once with brine. The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified using automated silica gel flash column chromatography (Hexane : Acetone, linear gradient of Acetone from 5% to 30%) to yield **BB4.3b** as a sticky colorless solid (1.70 g, 71%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.11 – 8.03 (m, 2H), 7.79 (d, *J* = 7.5 Hz, 2H), 7.63 (dd, *J* = 7.4, 2.2 Hz, 2H), 7.60 – 7.54 (m, 1H), 7.50 – 7.38 (m, 4H), 7.38 – 7.33 (m, 2H), 7.31 (dd, *J* = 7.5, 0.9 Hz, 2H), 7.29 – 7.24 (m, 2H), 7.24 – 7.19 (m, 3H), 7.13 (d, *J* = 8.0 Hz, 2H), 5.80 (dd, *J* = 3.1, 1.7 Hz, 1H), 5.50 (d, *J* = 1.4 Hz, 1H), 5.06 (t, *J* = 9.8 Hz, 1H), 4.72 (d, *J* = 12.1 Hz, 1H), 4.57 (d, *J* = 12.1 Hz, 1H), 4.47 (d, *J* = 7.2 Hz, 2H), 4.42 (dd, *J* = 9.8, 6.2 Hz, 1H), 4.26 (t, *J* = 7.1 Hz, 1H), 4.01 (dd, *J* = 9.7, 3.2 Hz, 1H), 2.33 (s, 3H), 1.32 (d, *J* = 6.2 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.81, 155.04, 143.54, 143.31, 141.47, 141.44, 138.31, 137.45, 133.51, 132.50, 130.11, 130.08, 129.75, 129.65, 128.62, 128.51, 128.06, 127.95, 127.34, 125.24, 125.22, 120.24, 86.57, 75.06, 71.60, 70.91, 70.16, 46.98, 21.28, 17.58. $[\alpha]_D^{20}$ -34.10 (c = 1.00 g/100mL, $CHCl_3$). IR (film) ν = 2928, 1754, 1723, 1452, 1249, 1104, 758, 743, 711 cm^{-1}. R_f = 0.40 (Hexane : Acetone 3:1). ESI-HRMS m/z 709.2239 $[M+Na]^+$ ($C_{42}H_{38}O_7SNa$ requires 709.2230).

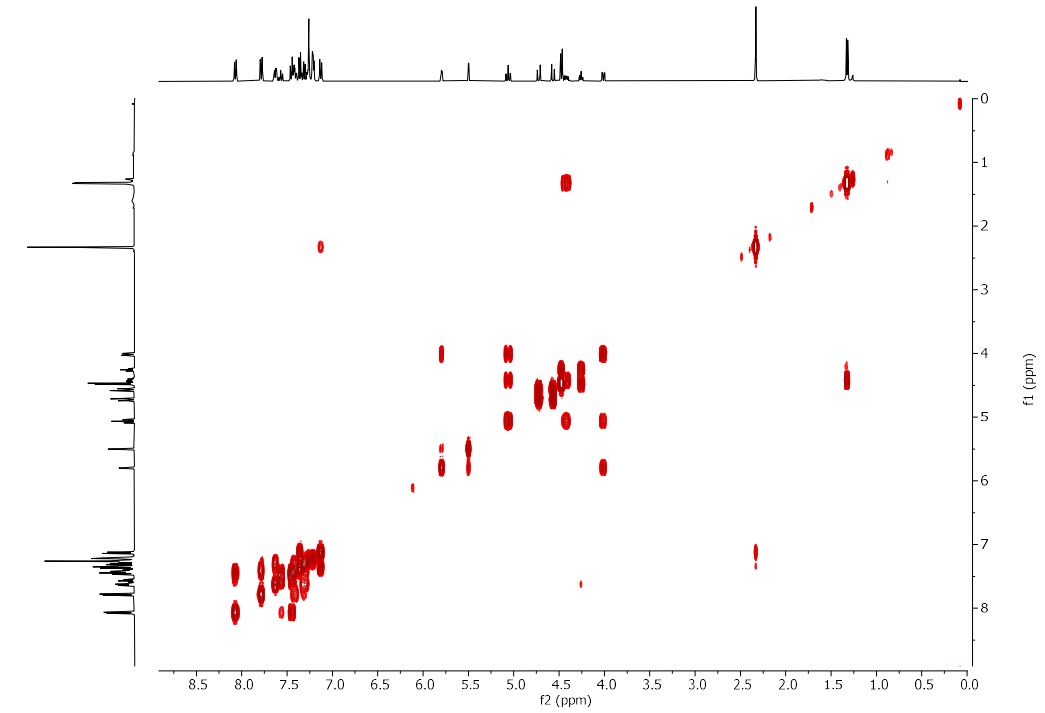
^{1}H NMR of BB4.3b (400 MHz, $CDCl_3$)



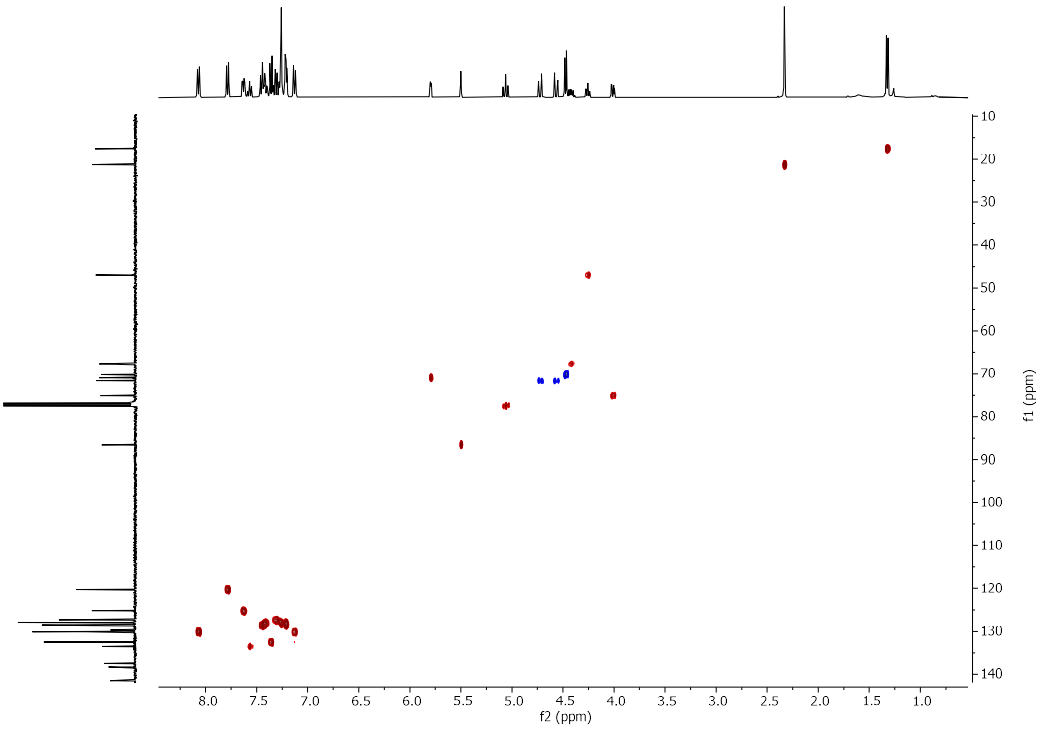
^{13}C NMR of BB4.3b (101 MHz, $CDCl_3$)



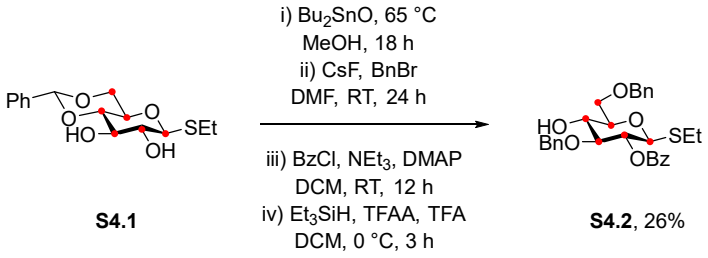
COSY NMR of BB4.3b ($CDCl_3$)



HSQC NMR of BB4.3b ($CDCl_3$)



Synthesis of **S4.2**



S4.1 **S4.2**, 26%

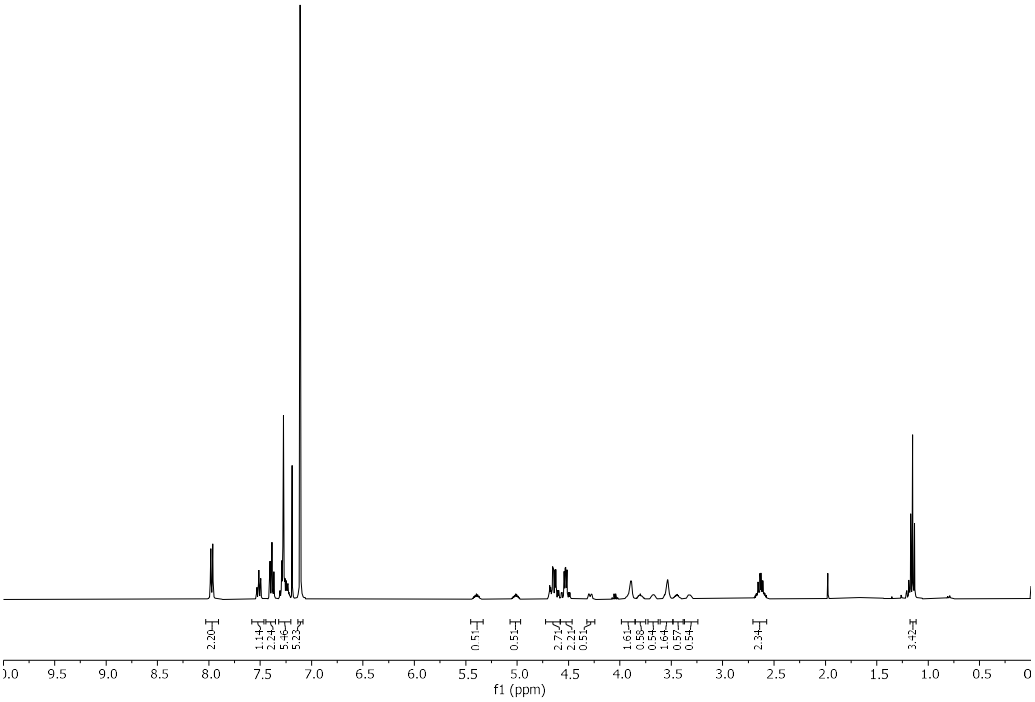
• = ^{13}C

S4.1 was synthesized according to previously reported procedure.[83]

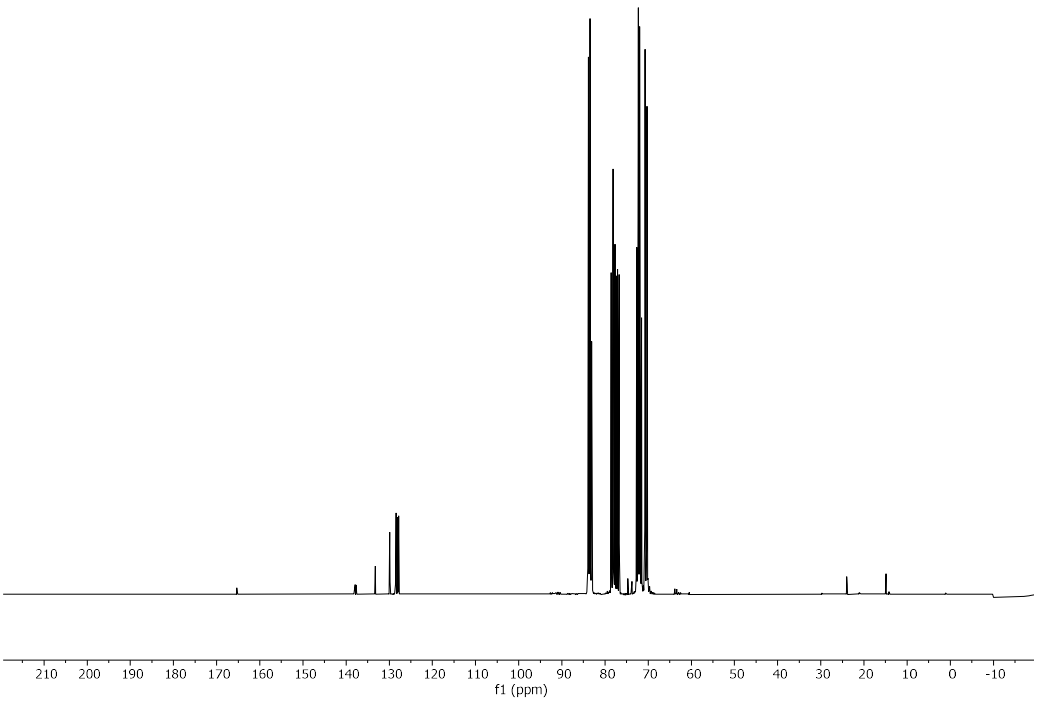
S4.1 (4.9 g, 18 mmol) was dissolved in MeOH (100 mL), di-*n*-butyltin oxide (5.5 g, 22 mmol) was added and the reaction mixture heated to 65 °C for 18 h. The reaction mixture was then cooled, concentrated *in vacuo* and the crude product was used in the next step without further purification. The crude product was dissolved in DMF (100 mL), benzyl bromide (3.76 g, 22 mmol) and cesium(I) fluoride (3.34 g, 22 mmol) were added and the mixture stirred at RT for 24 h under Ar atmosphere. The reaction mixture was concentrated *in vacuo* and the residue dissolved in DCM (120 mL). The organic layer was washed with aqueous potassium fluoride (100 mL, 1 M), dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude product was recrystallized from DCM : Hexane to give a white solid. The while solid was then dissolved in DCM (100 mL), benzoyl chloride (4.6 g, 33 mmol), DMAP (439 mg, 0.2 eq), and TEA (15 mL, 108 mmol) were added sequentially at 0 °C under Ar atmosphere. The reaction mixture was stirred at RT for 12 h, diluted with DCM and washed once with a saturated aqueous solution of $NaHCO_3$ and once with brine. The organic layer was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The crude product was recrystallized from DCM : Hexane to give a white solid. The white solid was dissolved in anhydrous DCM (50 mL) and triethylsilane (14.3 mL, 90 mmol) was added and the solution was cooled to 0 °C under Ar atmosphere. Trifluoroacetic acid (TFA) (6.9 mL, 90 mmol) and trifluoroacetic anhydride (TFAA) (1.2 mL, 9.0 mmol) were added sequentially. The solution was stirred at 0 °C for 3 h, after which time the reaction was diluted with DCM (50 mL) and washed once with a saturated aqueous solution of $NaHCO_3$ and once with brine. The organic layer was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The crude product was purified by silica gel flash column chromatography (Hexane : EtOAc, 3:1) to give **S4.2** as a white solid (2.5 g, 26% over 4 steps).

^1H NMR (400 MHz, $CDCl_3$) δ 8.01 – 7.94 (m, 2H), 7.56 – 7.47 (m, 1H), 7.43 – 7.34 (m, 2H), 7.31 – 7.17 (m, 5H), 7.11 (s, 5H), 5.21 (dtt, *J* = 152.7, 9.5, 4.8 Hz, 1H), 4.49 (dm, *J* = 157.2 Hz, 1H), 3.72 (dm, *J* = 143.5 Hz, 3H), 3.63 (dm, *J* = 143.8 Hz, 1H), 3.51 (dm, *J* = 141.0 Hz, 1H), 2.73 (quintd, *J* = 7.4, 3.8 Hz, 2H), 1.25 (t, *J* = 7.5 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.33, 137.92, 137.90, 137.69, 137.66, 133.27, 129.90, 129.80, 128.52, 128.47, 128.44, 128.04, 127.91, 127.85, 127.82, 84.24 – 82.82 (m), 78.79 – 77.61 (m), 77.49 – 76.62 (m), 74.72, 73.80, 72.89 – 71.39 (m), 70.50 (dt, *J* = 44.7, 4.2 Hz), 24.00, 14.92; $[\alpha]_D^{20}$ -8.36 (c 1.00 g/100mL, $CHCl_3$). IR (film) ν = 1727, 1269, 1052 cm^{-1}. R_f = 0.39 (Hexane : EtOAc = 3:1). ESI-HRMS m/z 537.1986 $[M+Na]^+$ ($C_{23}{}^{13}C_6H_{32}O_6SNa$ requires 537.2013).

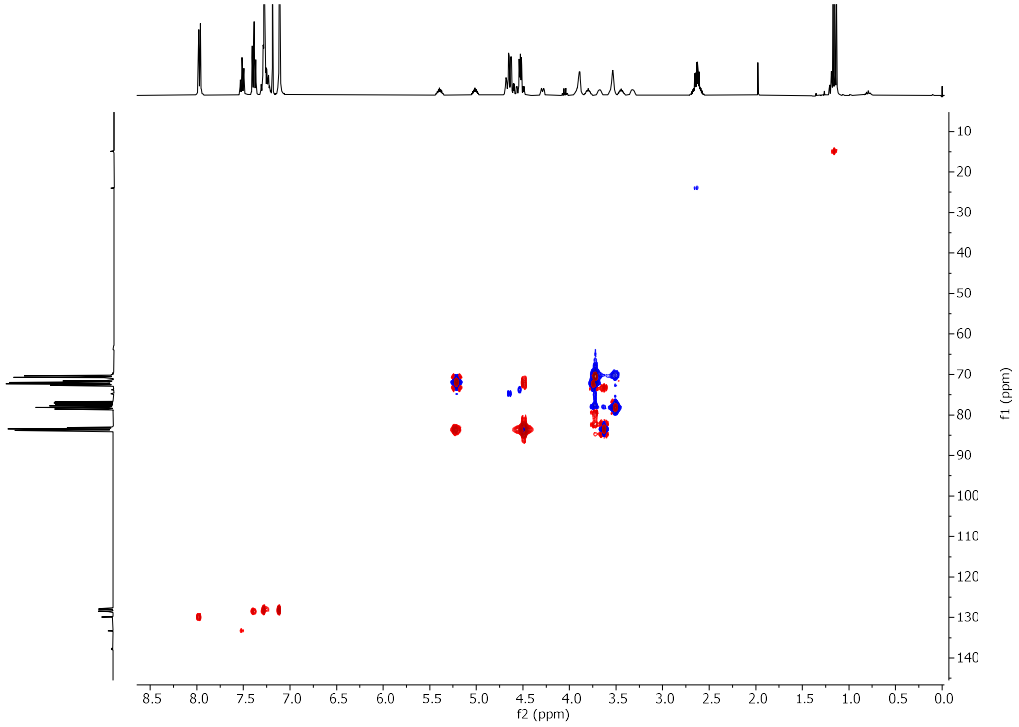
^{1}H NMR of S4.2 (400 MHz, $CDCl_3$)



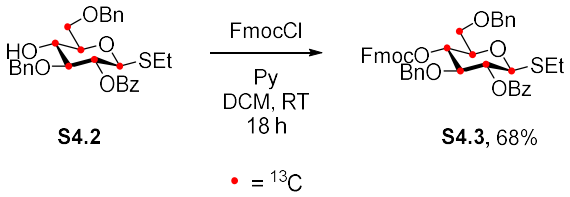
^{13}C NMR of S4.2 (101 MHz, $CDCl_3$)



HSQC NMR of S4.2 ($CDCl_3$)



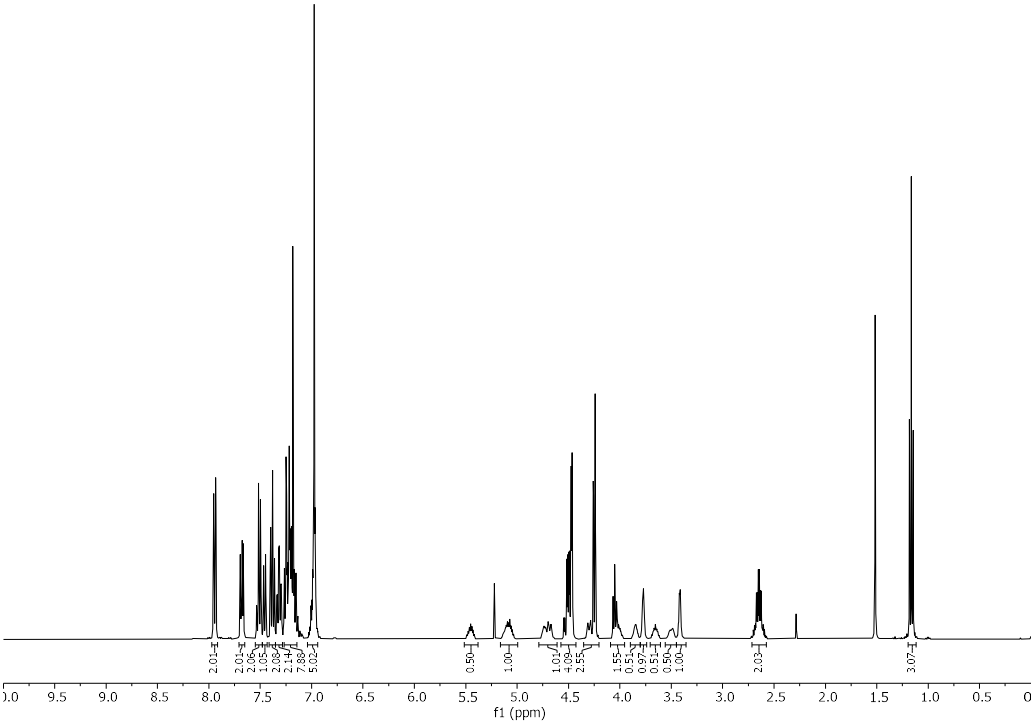
Synthesis of **S4.3**



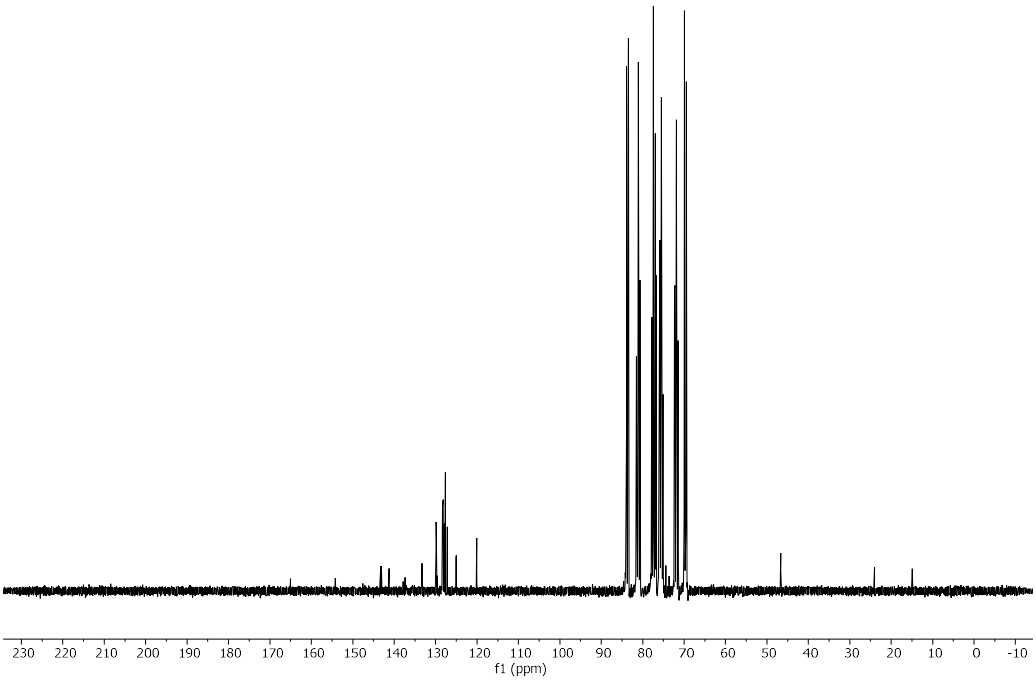
S4.2 (2.5 g, 4.9 mmol) was dissolved in DCM (50 mL) under Ar atmosphere. Pyridine (1.2 mL, 15 mmol) and FmocCl (2.5 g, 9.8 mmol) were added to the stirred reaction mixture. The yellow solution was stirred for 3 h and then quenched with an aqueous solution of HCl (1 M). The organic layer was washed once with an aqueous solution of HCl (1 M), once with a saturated aqueous solution of $NaHCO_3$, and once with brine. The crude product was purified with silica gel flash column chromatography (Hexane : EtOAc, 5:1) to give compound **S4.3** as white solid (2.9 g, 78%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.98 – 7.91 (m, 2H), 7.68 (ddt, *J* = 7.5, 4.3, 0.9 Hz, 2H), 7.51 (d, *J* = 7.4 Hz, 2H), 7.46 (dt, *J* = 7.4, 1.0 Hz, 1H), 7.38 (t, *J* = 7.8 Hz, 2H), 7.34 – 7.29 (m, 2H), 7.27 – 7.14 (m, 7H), 7.03 – 6.91 (m, 5H), 5.26 (dm, *J* = 153.0 Hz, 1H), 4.92 (dm, *J* = 153.2 Hz, 1H), 4.49 (dm, *J* = 151.9 Hz, 1H), 4.25 (d, *J* = 6.8 Hz, 2H), 4.05 (t, *J* = 7.1 Hz, 1H), 3.83 (dm, *J* = 142.7 Hz, 1H), 3.67 (dm, *J* = 144.7 Hz, 1H), 3.60 (dm, *J* = 142.6 Hz, 2H), 2.65 (td, *J* = 7.6, 3.9 Hz, 2H), 1.16 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, cdcl$_3$) δ 165.04, 143.28, 143.10, 141.30, 141.27, 137.36, 133.27, 129.89, 129.64, 128.42, 128.33, 128.15, 127.91, 127.83, 127.65, 127.63, 127.17, 125.10, 125.01, 120.08, 83.67 (dd, *J* = 42.1, 4.0 Hz), 81.07 (t, *J* = 40.3 Hz), 78.08 – 76.48 (m), 75.50 (td, *J* = 40.8, 4.5 Hz), 74.38, 73.60, 71.84 (td, *J* = 41.5, 40.7, 4.6 Hz), 70.02, 69.67 (dt, *J* = 44.7, 3.9 Hz), 46.68, 24.06, 14.89. $[\alpha]_D^{20}$ +35.87 (c 1.00 g/100mL, $CHCl_3$); IR (film) ν = 1754, 1452, 1253 cm^{-1}. R_f = 0.25 (Hexane : Acetone = 3:1). ESI-HRMS m/z 759.2678 $[M+Na]^+$ ($C_{38}{}^{13}C_6H_{42}O_8SNa$ requires 759.2694).

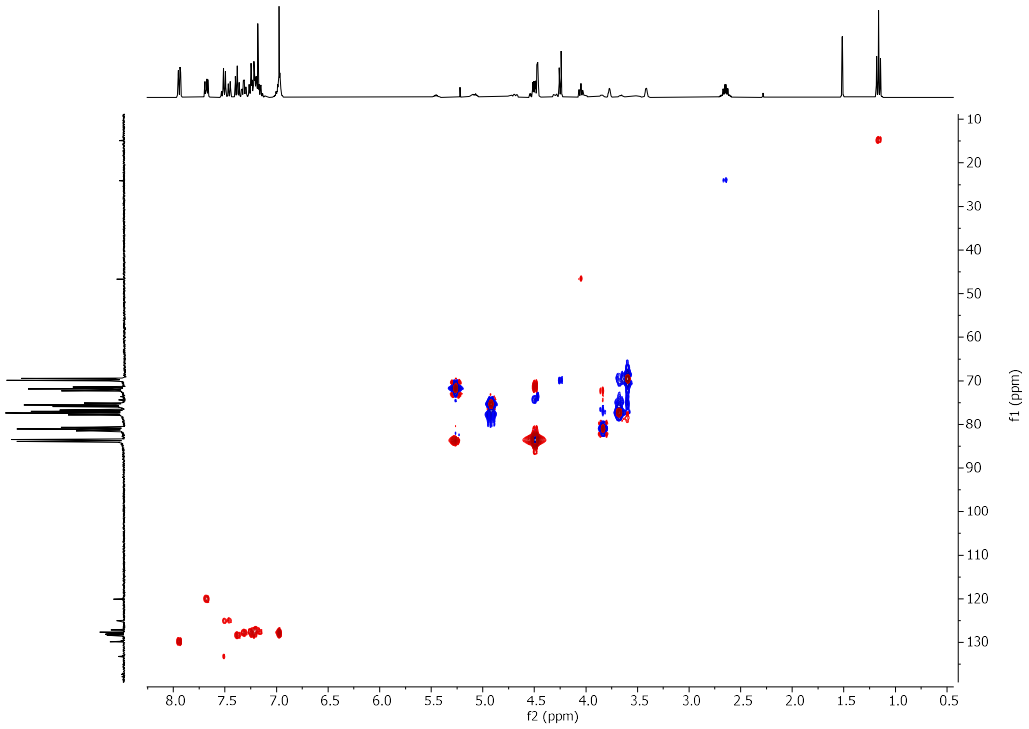
^{1}H NMR of S4.3 (400 MHz, $CDCl_3$)



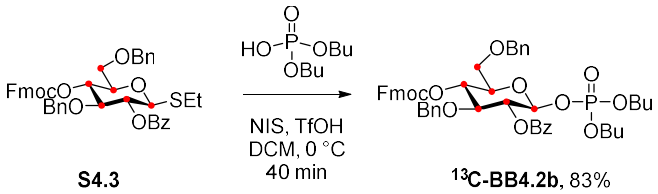
^{13}C NMR of S4.3 (101 MHz, $CDCl_3$)



HSQC NMR of S4.3 ($CDCl_3$)



Synthesis of **^{13}C-BB4.2b**

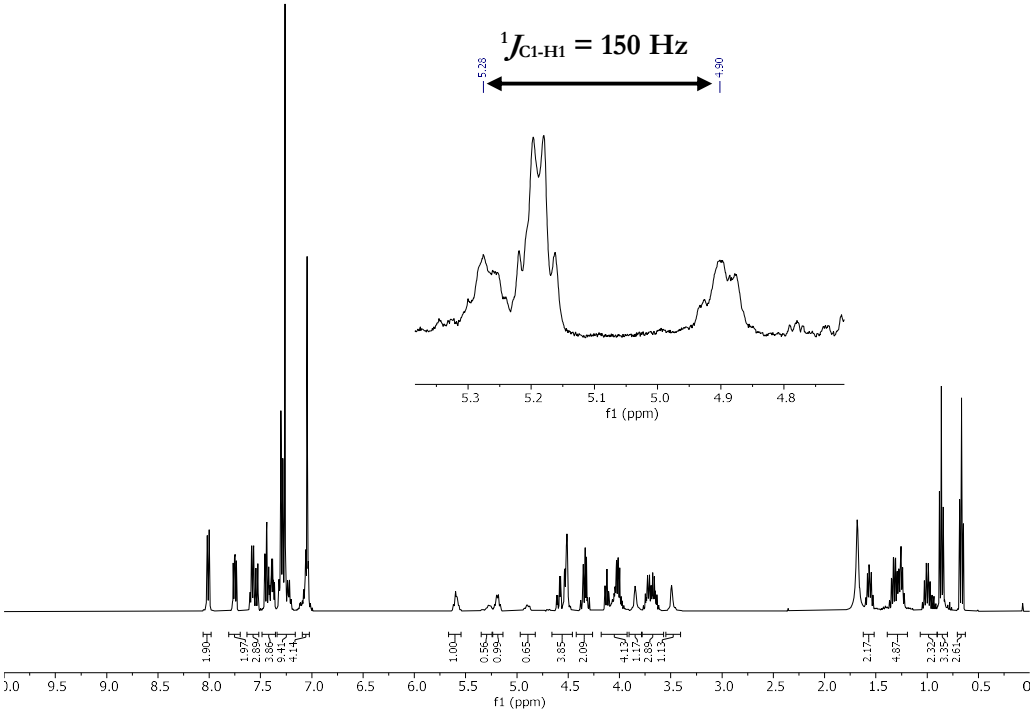


• = ^{13}C

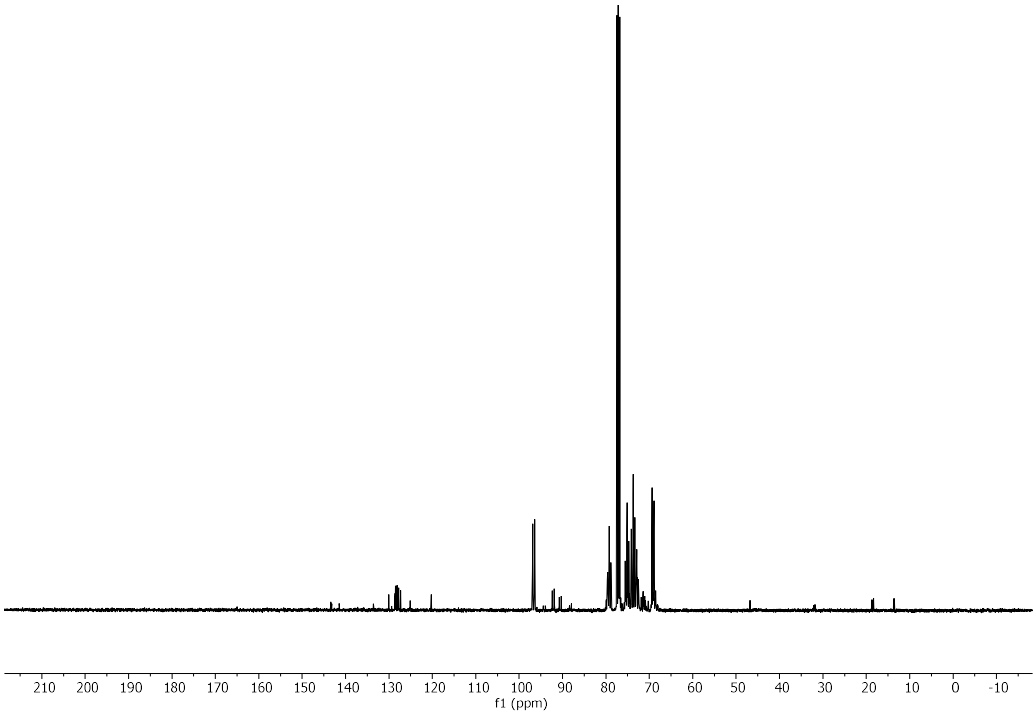
S4.3 (338 mg, 0.46 mmol) was co-evaporated three times with toluene, dissolved in anhydrous DCM (1.6 mL) and stirred in the presence of molecular sieves (4 Å, powder, flame dried) for 1 h under Ar atmoshpere. Di-*n*-butyl phosphate (375 µL) was dissolved in anhydrous DCM (2 mL) and stirred in the presence of flame dried 4 Å molecular sieves for 1 h (Solution A). Solution A (1.1 mL, 1.0 mmol of di-*n*-butyl phosphate): was added to the reaction mixture and cooled to 0 °C. NIS (124 mg, 0.55 mmol) was added in one portion followed by a dropwise addition of TfOH (3 µL, 0.03 mmol) and the reaction was stirred for 40 min at 0 °C, after which time it was quenched by addition of Py (1 drop). The reaction was diluted with DCM, filtered over celite, washed once with a saturated aqueous solution of $Na_2S_2O_3$ and once with brine. The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by silica gel flash column chromatography (Hexane : Acetone, 5:1→4:1→3:1→2:1) to yield **^{13}C-BB4.2b** as a sticky colorless solid (338 mg, 83%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.01 (d, *J* = 7.3 Hz, 2H), 7.81 – 7.71 (m, 2H), 7.64 – 7.50 (m, 3H), 7.49 – 7.35 (m, 4H), 7.34 – 7.18 (m, 8H), 7.16 – 6.98 (m, 4H), 5.66 – 5.53 (m, 1H), 5.34 – 5.23 (m, 0.5H), 5.23 – 5.11 (m, 1H), 4.99 – 4.82 (m, 0.5H), 4.64 – 4.45 (m, 4H), 4.42 – 4.27 (m, 2H), 4.19 – 3.93 (m, 4H), 3.91 – 3.80 (m, 1H), 3.79 – 3.58 (m, 3H), 3.55 – 3.43 (m, 1H), 1.56 (dt, *J* = 14.6, 6.7 Hz, 2H), 1.39 – 1.21 (m, 4H), 1.06 – 0.93 (m, 2H), 0.86 (t, *J* = 7.4 Hz, 3H), 0.67 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 143.40, 143.20, 141.42, 133.58, 130.06, 129.40, 128.60, 128.47, 128.34, 128.11, 128.07, 127.87, 127.81, 127.75, 127.33, 125.22, 125.14, 120.23, 96.91, 96.86, 96.43, 96.39, 92.36, 91.90, 90.77, 90.31, 88.40, 87.97, 79.93, 79.64, 79.24, 78.84, 77.48, 77.16, 76.84, 75.56, 75.52, 75.15, 75.11, 74.75, 74.71, 74.14, 73.73, 73.31, 72.96, 72.92, 72.52, 71.89, 71.74, 71.46, 71.28, 71.05, 70.90, 70.23, 69.38, 68.94, 68.55, 46.82, 32.16, 32.08, 31.91, 31.84, 18.78, 18.68, 18.35, 13.69, 13.50. ^{31}P NMR (162 MHz, $CDCl_3$) δ -2.75. $[\alpha]_D^{20}$ 56.57 (c = 1.00 g/100mL, $CHCl_3$). IR (film) ν = 2962, 1756, 1734, 1453, 1250, 1070, 1028, 740 cm^{-1}. R_f = 0.46 (Hexane : Acetone 2:1). ESI-HRMS m/z 907.3526 $[M+Na]^+$ ($C_{44}{}^{13}C_6H_{55}O_{12}PNa$ requires 907.3524).

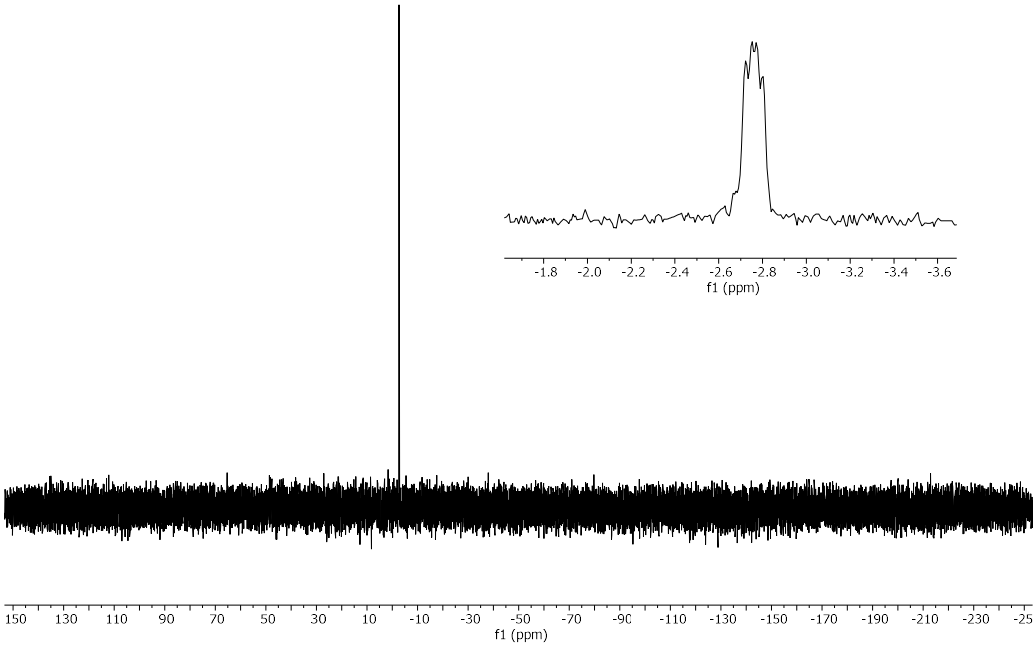
^{1}H NMR of ^{13}C-BB4.2b (400 MHz, $CDCl_3$)



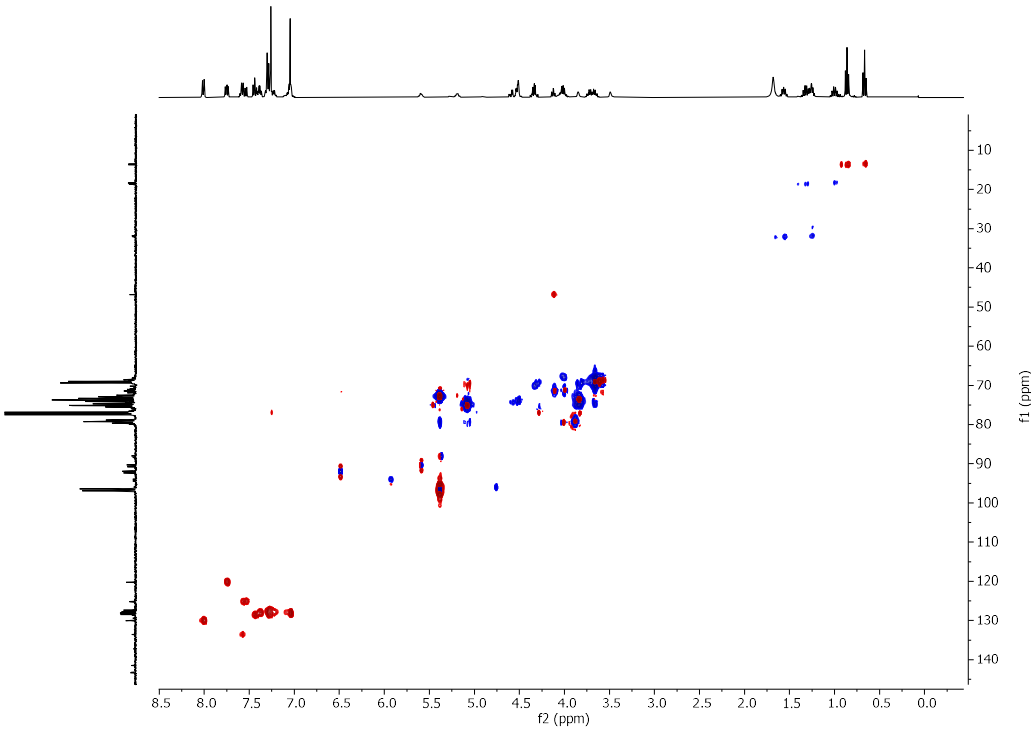
^{13}C NMR of ^{13}C-BB4.2b (101 MHz, $CDCl_3$)



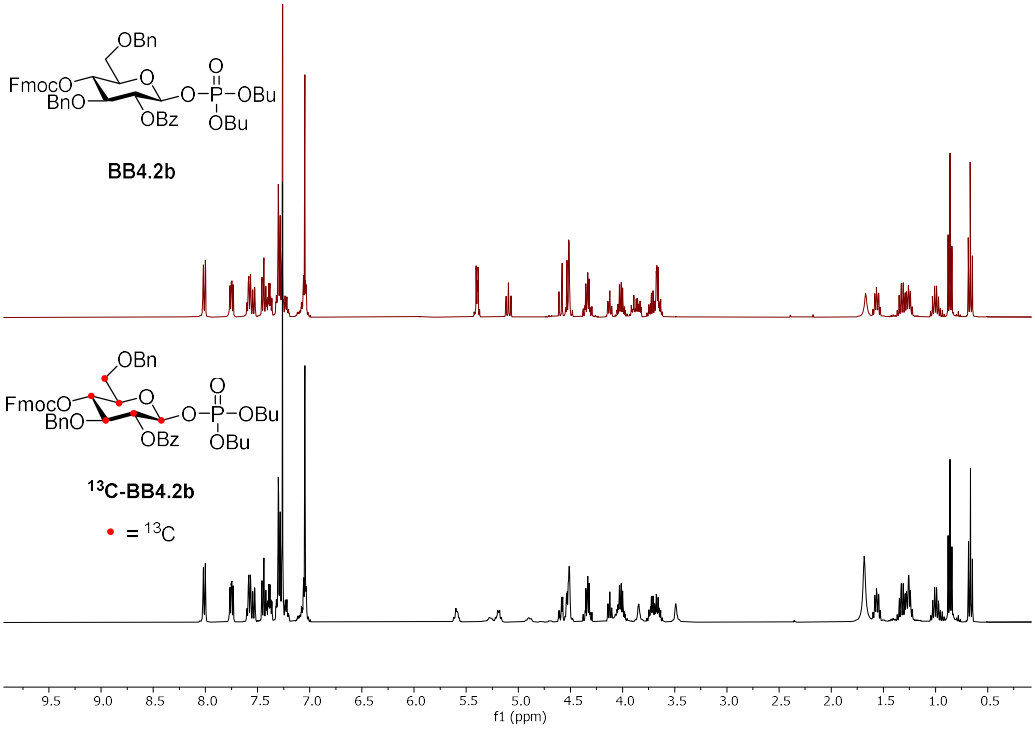
^{31}P NMR of^{13}C-BB4.2b (162 MHz, $CDCl_3$)



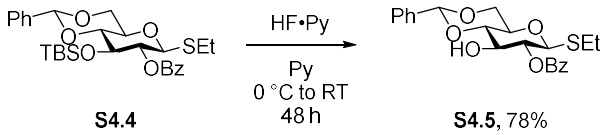
HSQC NMR of ^{13}C-BB4.2b ($CDCl_3$)



Comparison of ^{1}H NMR of BB2b (top) and ^{13}C-BB2b (bottom) ($CDCl_3$)



Synthesis of **S4.5**



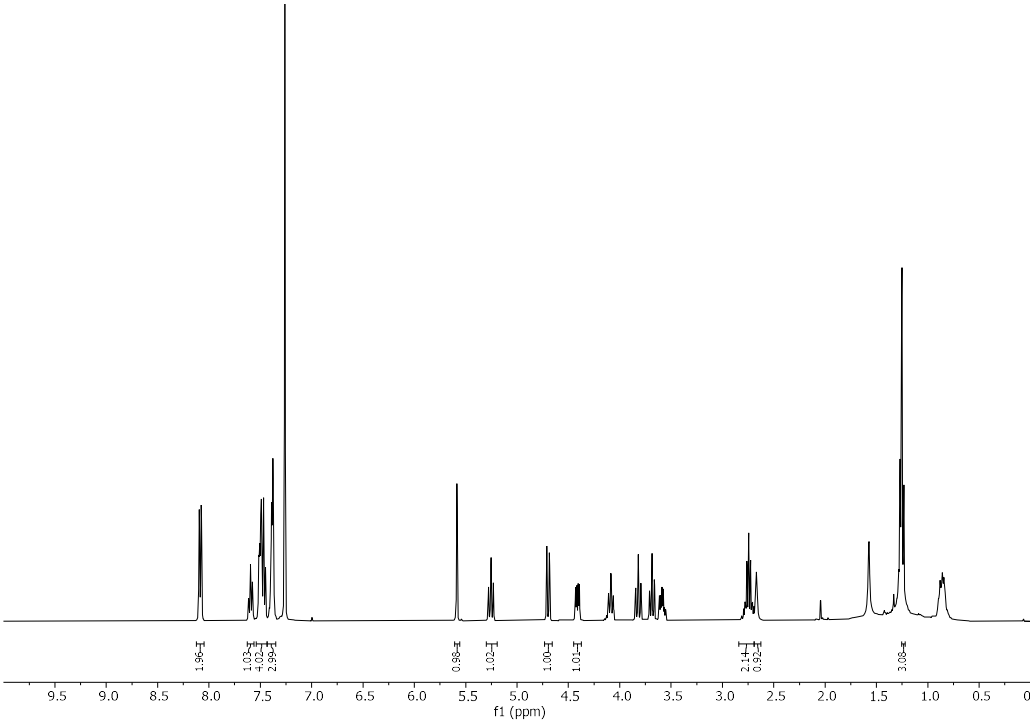
S4.4 **S4.5**, 78%

S4.4 was synthesized according to previously reported procedure.[361]

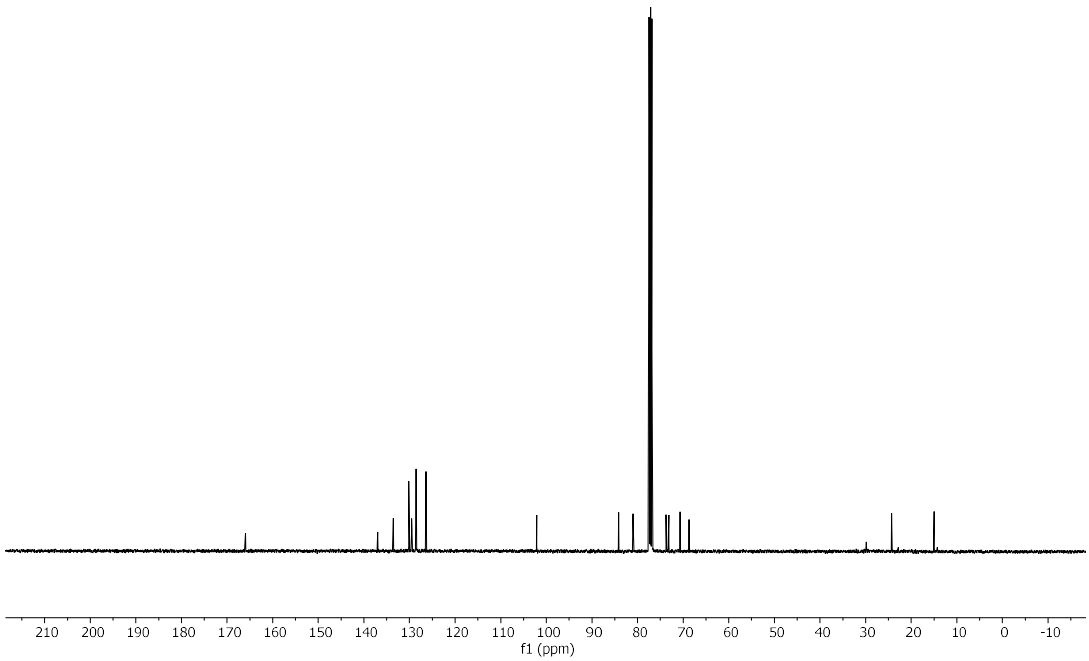
S4.4 (514 mg, 0.97 mmol) was dissolved in pyridine (Py, 7 mL) and cooled to 0 °C. HF•Py (1.25 mL, 70%$_w$ HF) was added dropwise to the stirred reaction mixture. After 1 h, the reaction was allowed to RT and stirred for 48 h, after which time the reaction was cooled to 0 °C and quenched with a saturated aqueous solution of $NaHCO_3$. The crude product was diluted with DCM and washed twice with saturated aqueous solution of $NaHCO_3$. The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by silica gel flash column chromatography (Hexane : EtOAc, 4:1→3:1→2:1) to yield **S4.5** as a colorless foam (320 mg, 78%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.09 (d, *J* = 7.8 Hz, 2H), 7.59 (t, *J* = 7.4 Hz, 1H), 7.53 – 7.43 (m, 4H), 7.38 (d, *J* = 4.9 Hz, 3H), 5.58 (s, 1H), 5.25 (t, *J* = 9.4 Hz, 1H), 4.70 (d, *J* = 10.0 Hz, 1H), 4.41 (dd, *J* = 10.5, 4.8 Hz, 1H), 4.10 (q, *J* = 8.3 Hz, 1H), 3.82 (t, *J* = 10.2 Hz, 1H), 3.69 (t, *J* = 9.3 Hz, 1H), 3.58 (td, *J* = 9.6, 4.9 Hz, 1H), 2.74 (quint, *J* = 7.1 Hz, 2H), 2.67 (s, 1H), 1.25 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 165.96, 136.96, 133.54, 130.15, 129.66, 129.52, 128.59, 128.53, 126.42, 102.11, 84.18, 80.97, 73.74, 73.14, 70.67, 68.72, 24.28, 15.00. $[\alpha]_D^{20}$ -24.01 (c = 0.80 g/100mL, $CHCl_3$). IR (film) ν = 3475, 2927, 1727, 1270, 1095, 711 cm^{-1}. R_f = 0.24 (Hexane : EtOAc = 4:1). ESI-HRMS m/z 439.1208 $[M+Na]^+$ ($C_{22}H_{24}O_6SNa$ requires 439.1186).

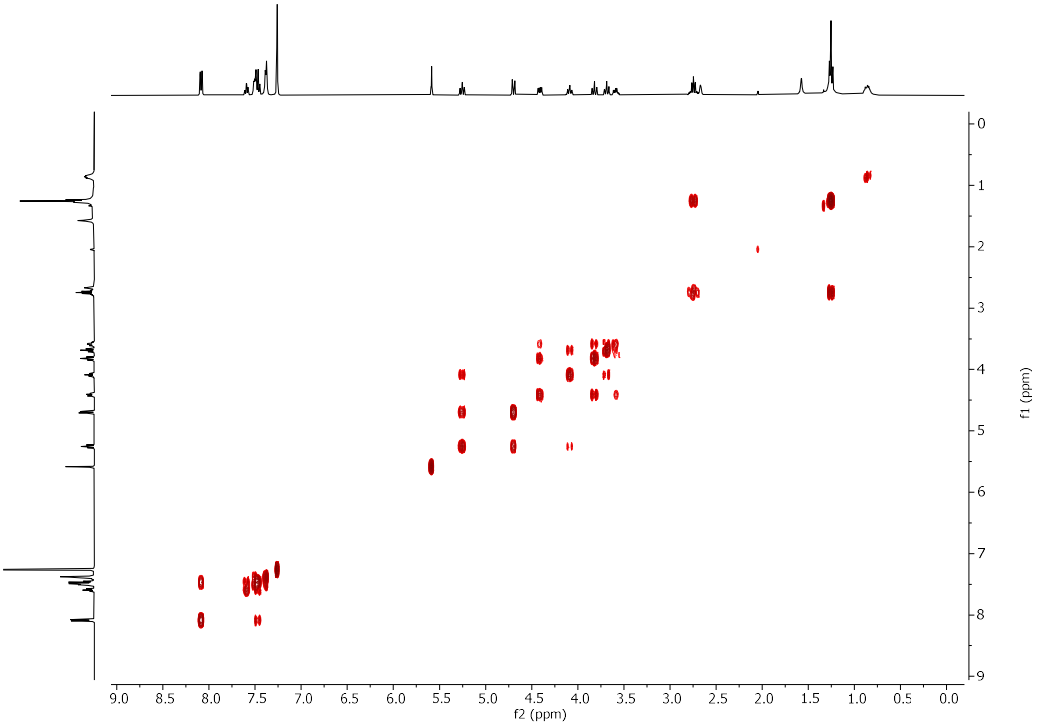
^{1}H NMR of S4.5 (400 MHz, $CDCl_3$)



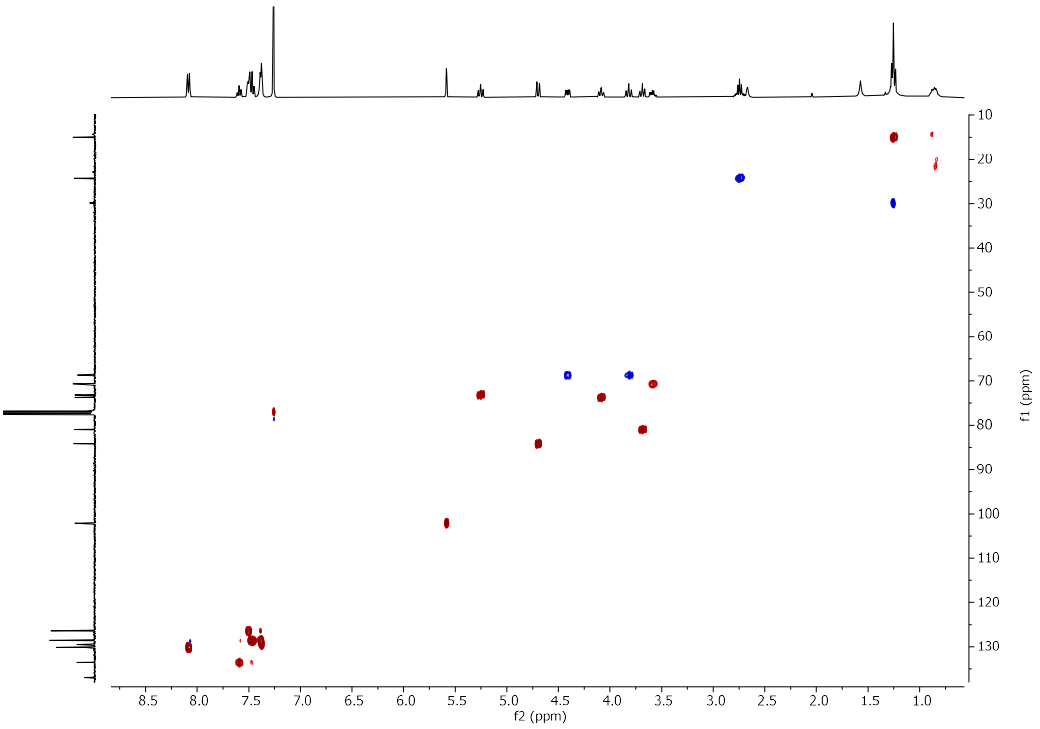
^{13}C NMR of S4.5 (101 MHz, $CDCl_3$)



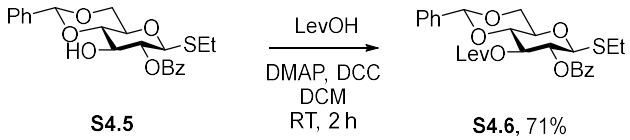
COSY NMR of S4.5 ($CDCl_3$)



HSQC NMR of S4.5 ($CDCl_3$)



Synthesis of **S4.6**

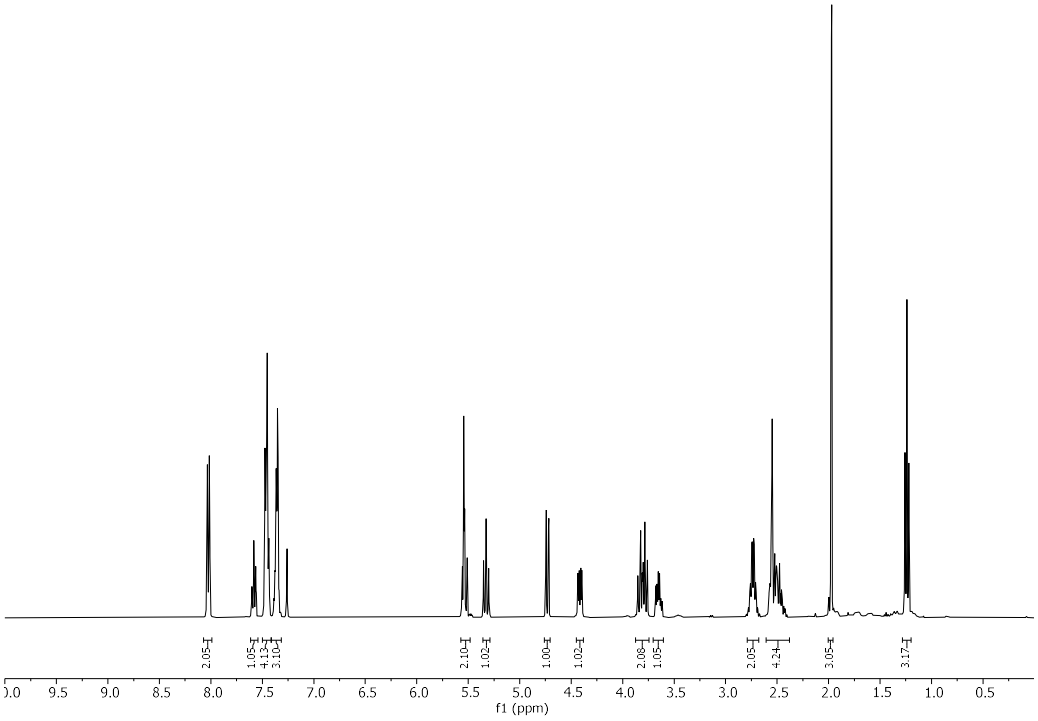


S4.5 **S4.6**, 71%

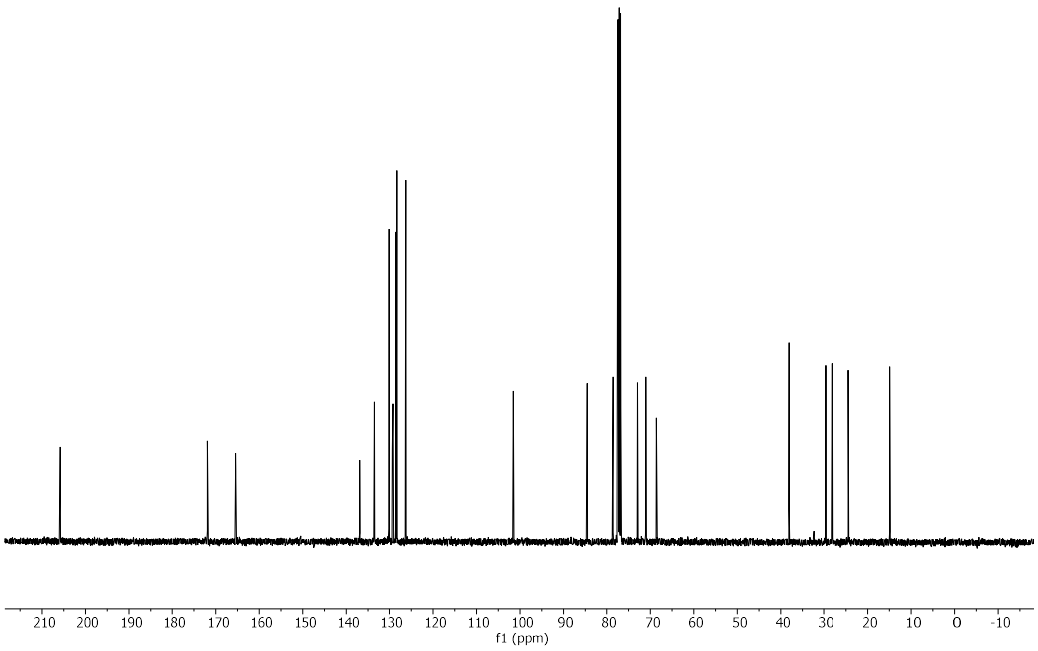
S4.5 (311 mg, 0.75 mmol) was dissolved in anhydrous DCM (4 mL) under Ar atmosphere. Levulinic acid (LevOH, 100 µL, 1.5 mmol), *N*,*N*'-dicyclohexylcarbodiimide (DCC, 280 mg, 1.35 mmol), and 4-dimethylaminopyridine (DMAP, 20 mg, 0.15 mmol) were added to the stirred reaction mixture at RT. After 2 h, the reaction was passed through a short plug of silica gel using DCM as eluent. The crude reaction mixture was concentrated under vacuum and recrystallized from DCM : Hexane to yield **S4.6** as a colorless crystalline solid (275 mg, 71%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.02 (d, *J* = 7.9 Hz, 2H), 7.58 (t, *J* = 7.3 Hz, 1H), 7.49 – 7.42 (m, 4H), 7.41 – 7.33 (m, 3H), 5.57 – 5.48 (m, 2H), 5.33 (t, *J* = 9.6 Hz, 1H), 4.73 (d, *J* = 10.0 Hz, 1H), 4.41 (dd, *J* = 10.5, 4.8 Hz, 1H), 3.88 – 3.73 (m, 2H), 3.65 (td, *J* = 9.6, 4.9 Hz, 1H), 2.74 (qt, *J* = 7.3, 4.0 Hz, 2H), 2.60 – 2.39 (m, 4H), 1.97 (s, 3H), 1.24 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 205.81, 171.93, 165.42, 136.88, 133.51, 130.10, 129.39, 129.22, 128.57, 128.36, 126.27, 101.56, 84.51, 78.57, 72.97, 71.10, 71.04, 68.64, 38.05, 29.62, 28.13, 24.51, 14.94. $[\alpha]_D^{20}$ -23.81 (c = 0.70 g/100mL, $CHCl_3$). IR (film) ν = 2930, 1724, 1269, 1097, 712 cm^{-1}. R_f = 0.42 (Hexane : EtOAc = 2:1). ESI-HRMS m/z 537.1572 $[M+Na]^+$ ($C_{27}H_{30}O_8SNa$ requires 537.1553).

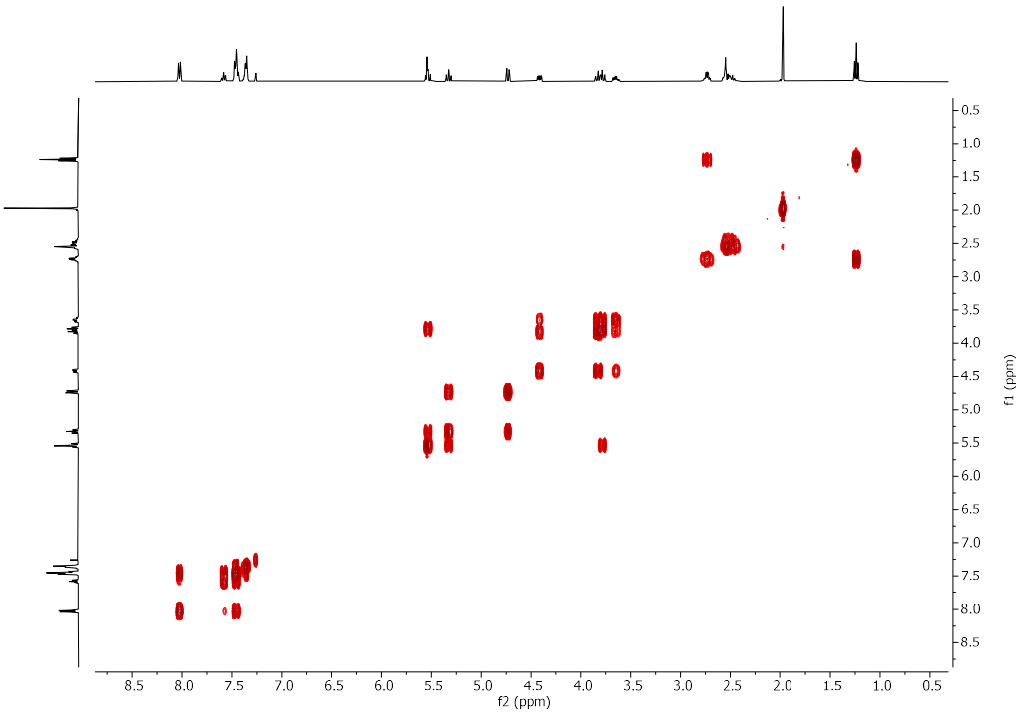
^{1}H NMR of S4.6 (400 MHz, $CDCl_3$)



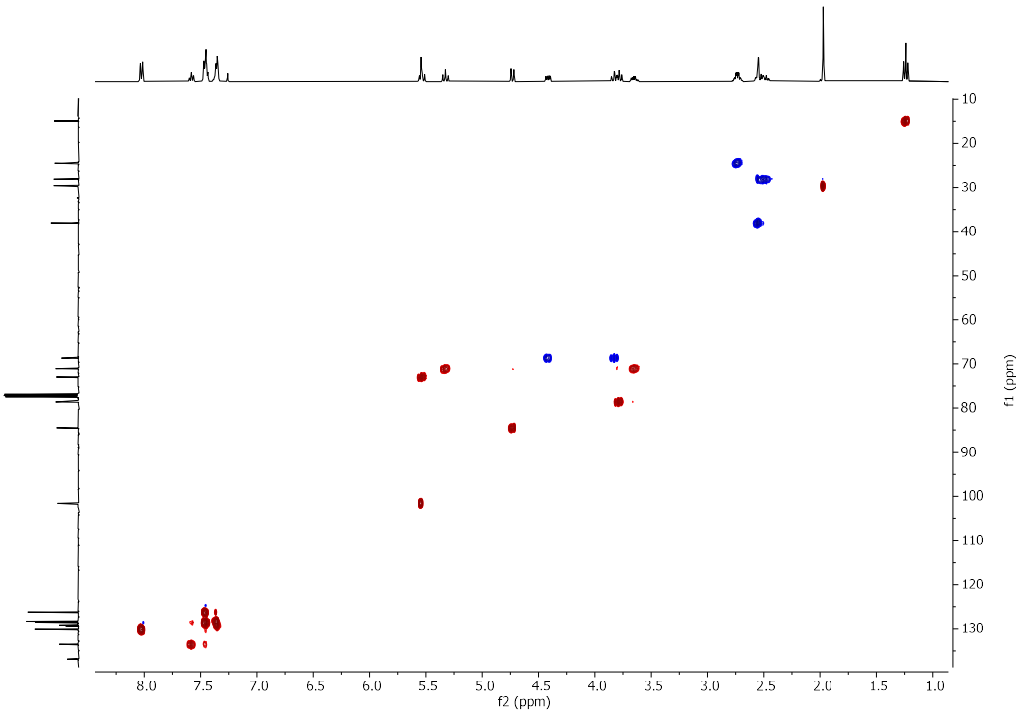
^{13}C NMR of S4.6 (101 MHz, $CDCl_3$)



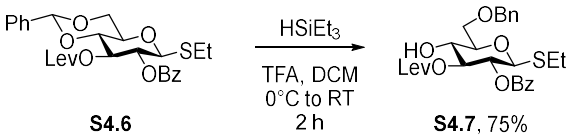
COSY NMR of S4.6 ($CDCl_3$)



HSQC NMR of S4.6 ($CDCl_3$)



Synthesis of **S4.7**

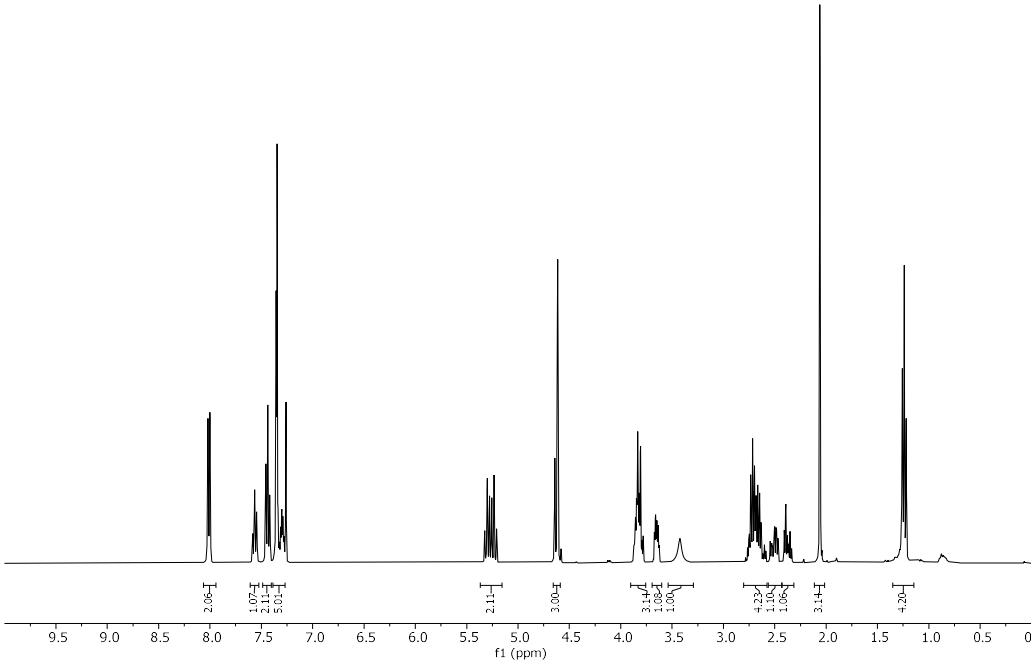


S4.6 **S4.7**, 75%

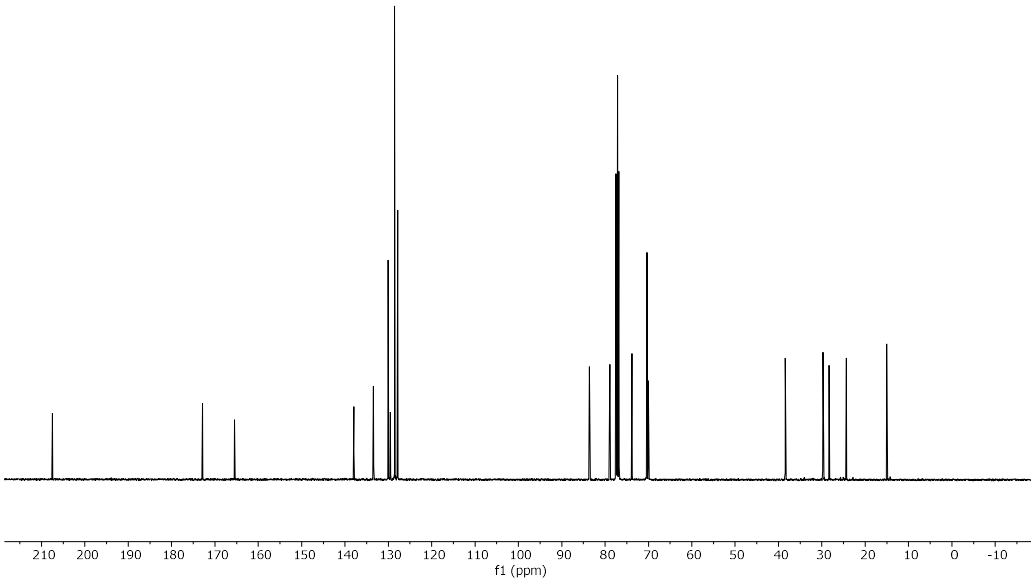
S4.6 (275 mg, 0.53 mmol) was dissolved in anhydrous DCM (10 mL) and stirred in the presence of molecular sieves (4 Å, pellet, flame dried) under Ar atmosphere. The stirred reaction mixture was cooled to 0 °C and $HSiEt_3$ (500 µL, 3.2 mmol) and TFA (235 µL, 3.2 mmol) were added. After 20 min, the reaction was allowed to RT and stirred for 3 h, after which time it was quenched by addition of a saturated aqueous solution of $NaHCO_3$. The crude product was diluted with DCM and washed once with a saturated aqueous solution of $NaHCO_3$, and once with brine. The organic layer was dried over Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by silica gel flash column chromatography (Hexane : EtOAc, 2:1→3:2→1:1→2:3) to yield **S4.7** as a colorless viscous oil (205 mg, 75%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.01 (d, *J* = 7.9 Hz, 2H), 7.57 (t, *J* = 7.3 Hz, 1H), 7.44 (t, *J* = 7.6 Hz, 2H), 7.37 – 7.27 (m, 5H), 5.30 (t, *J* = 9.2 Hz, 1H), 5.23 (t, *J* = 9.5 Hz, 1H), 4.67 – 4.59 (m, 3H), 3.91 – 3.75 (m, 3H), 3.68 – 3.60 (m, 1H), 3.42 (s, 1H), 2.81 – 2.58 (m, 4H), 2.56 – 2.44 (m, 1H), 2.43 – 2.32 (m, 1H), 2.06 (s, 3H), 1.24 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 207.50, 172.85, 165.47, 137.98, 133.42, 130.03, 129.50, 128.54, 127.87, 127.81, 83.65, 78.87, 77.16, 73.78, 70.37, 70.02, 38.43, 29.71, 28.31, 24.37, 15.01. $[\alpha]_D^{20}$ -1.19 (c = 1.00 g/100mL, $CHCl_3$). IR (film) ν = 3483, 2929,m 1725, 1271, 1070, 713 cm^{-1}. R_f = 0.16 (Hexane : EtOAc = 4:1). ESI-HRMS m/z 539.1705 $[M+Na]^+$ ($C_{27}H_{32}O_8SNa$ requires 539.1710).

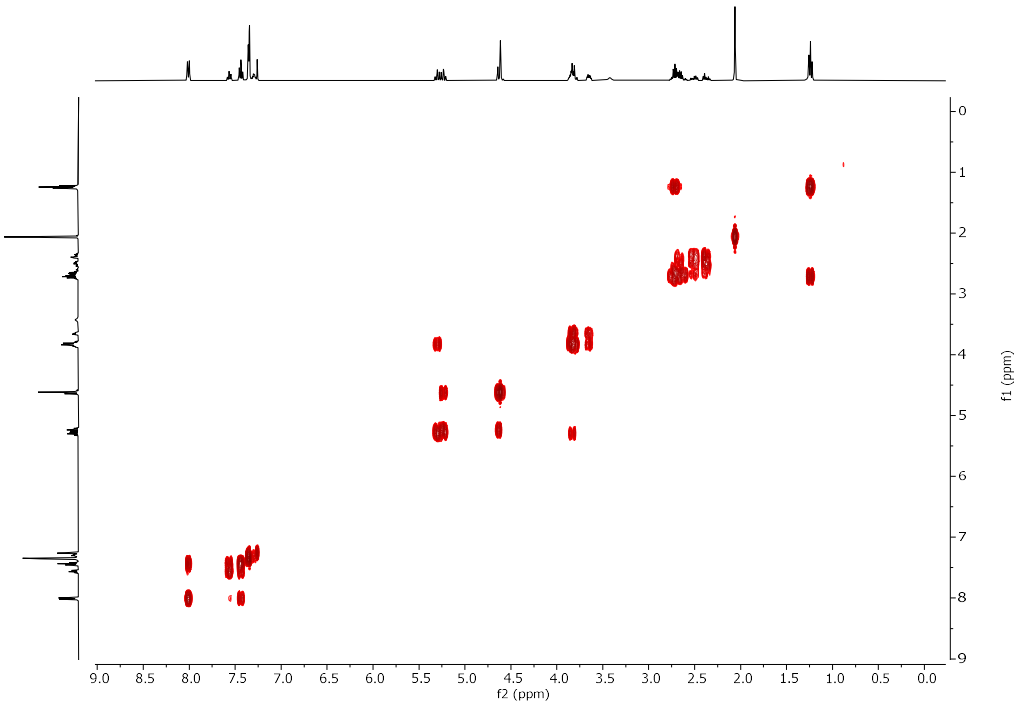
^{1}H NMR of S4.7 (400 MHz, $CDCl_3$)



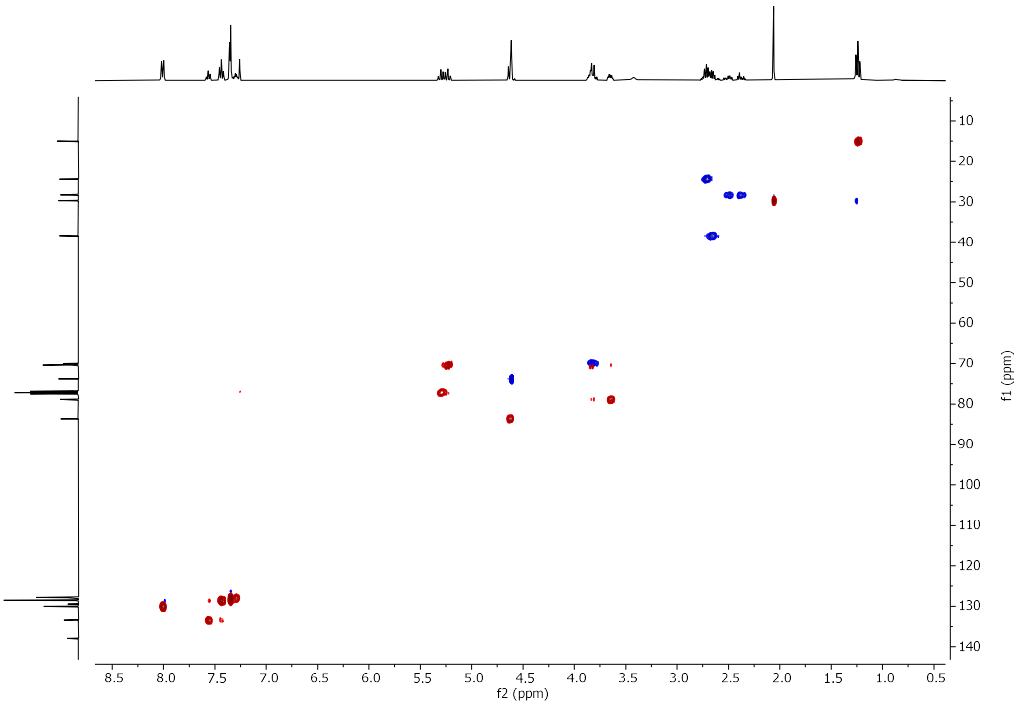
^{13}C NMR of S4.7 (101 MHz, $CDCl_3$)



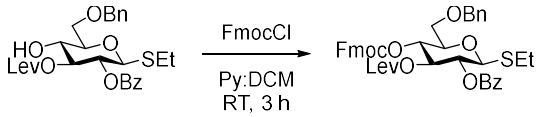
COSY NMR of S4.7 ($CDCl_3$)



HSQC NMR of S4.7 ($CDCl_3$)



Synthesis of **BB4.4**

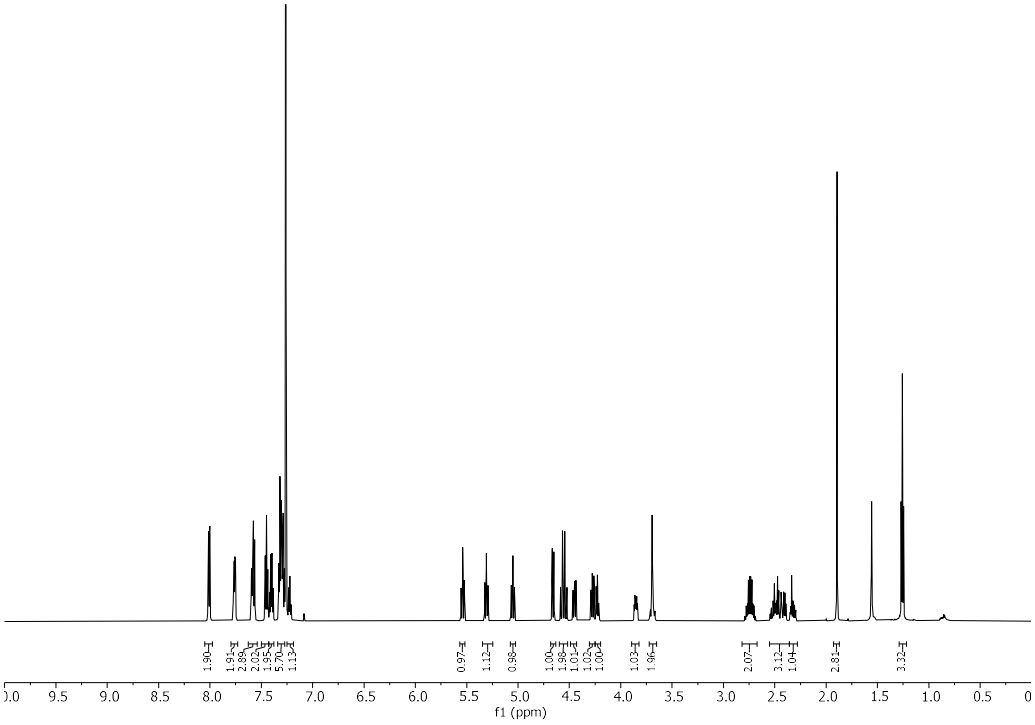


S4.7 **BB4.4**, 70%

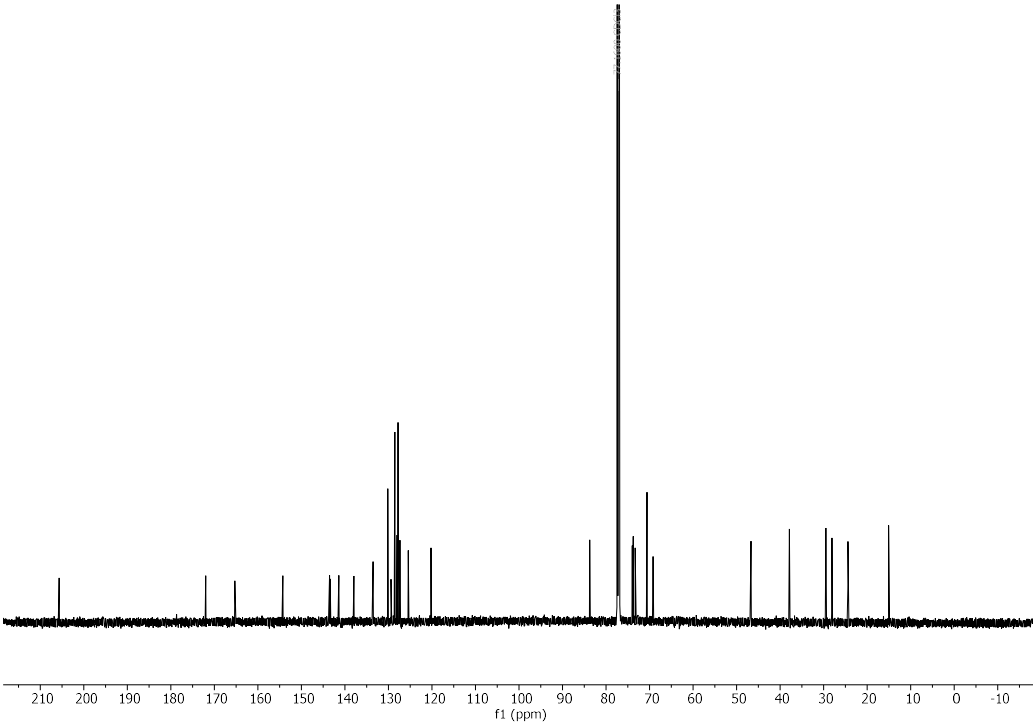
S4.7 (205 mg, 0.4 mmol) was dissolved in anhydrous DCM (6 mL) and cooled to 0 °C under Ar atmosphere. FmocCl (205 mg, 0.8 mmol) was dissolved in DCM (0.7 mL) and added dropwise to the stirred reaction mixture. The solution was stirred for 15 min at 0 °C and then allowed to RT and stirred for additional 5 h, after which time the reaction was quenched by addition of a saturated aqueous solution of $NaHCO_3$. The reaction mixture was diluted with DCM, washed once with an aqueous solution of citric acid (0.5 M), and once with brine. The organic layer was dried with Na_2SO_4 and concentrated under reduced pressure. The crude product was purified using silica gel flash column chromatography (Hexane : EtOAc, 3:1→2:1→1:1) to yield **BB4.4** as a colorless foam (208 mg, 70%).

^{1}H NMR (600 MHz, $CDCl_3$) δ 8.01 (d, *J* = 7.3 Hz, 2H), 7.76 (dd, *J* = 7.5, 3.5 Hz, 2H), 7.61 – 7.55 (m, 3H), 7.45 (t, *J* = 7.8 Hz, 2H), 7.40 (q, *J* = 7.1 Hz, 2H), 7.34 – 7.27 (m, 6H), 7.22 (t, *J* = 7.2 Hz, 1H), 5.54 (t, *J* = 9.5 Hz, 1H), 5.31 (t, *J* = 9.7 Hz, 1H), 5.05 (t, *J* = 9.7 Hz, 1H), 4.66 (d, *J* = 10.0 Hz, 1H), 4.58 (d, *J* = 12.0 Hz, 1H), 4.53 (d, *J* = 12.0 Hz, 1H), 4.45 (dd, *J* = 10.1, 7.3 Hz, 1H), 4.28 (dd, *J* = 10.1, 7.8 Hz, 1H), 4.23 (t, *J* = 7.4 Hz, 1H), 3.88 – 3.82 (m, 1H), 3.72 – 3.64 (m, 2H), 2.81 – 2.68 (m, 2H), 2.56 – 2.37 (m, 3H), 2.36 – 2.29 (m, 1H), 1.89 (s, 3H), 1.26 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 205.68, 171.95, 165.30, 154.24, 143.54, 143.41, 141.41, 141.38, 137.90, 133.53, 130.11, 129.38, 128.49, 128.03, 127.82, 127.81, 127.34, 127.32, 125.43, 125.40, 120.18, 120.15, 83.75, 77.44 (d, *J* = 10.6 Hz), 73.96, 73.76, 73.26, 70.58, 69.17, 46.70, 37.87, 29.49, 28.09, 24.40, 15.02. $[\alpha]_D^{20}$ +18.81 (c = 1.00 g/100mL, $CHCl_3$). IR (film) ν = 2929, 1756, 1725, 1253 cm^{-1}. R_f = 0.35 (Hexane : EtOAc 2:1). ESI-HRMS m/z 761.2391 $[M+Na]^+$ ($C_{42}H_{42}O_{10}SNa$ requires 761.2391).

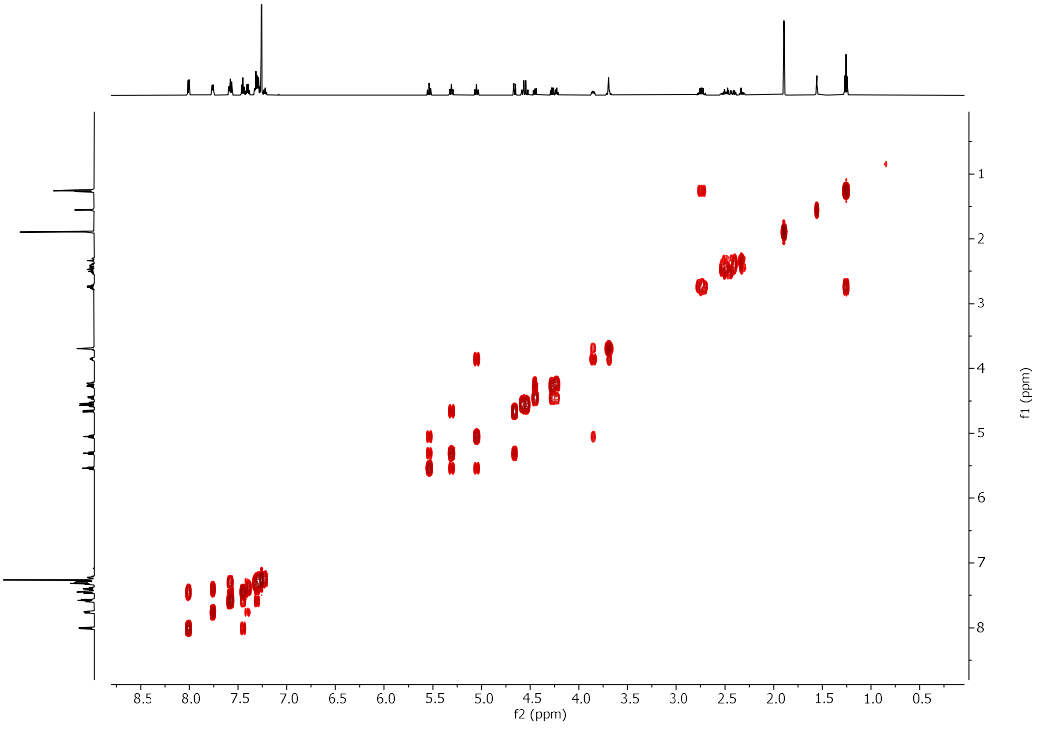
^{1}H NMR of BB4.4 (600 MHz, $CDCl_3$)



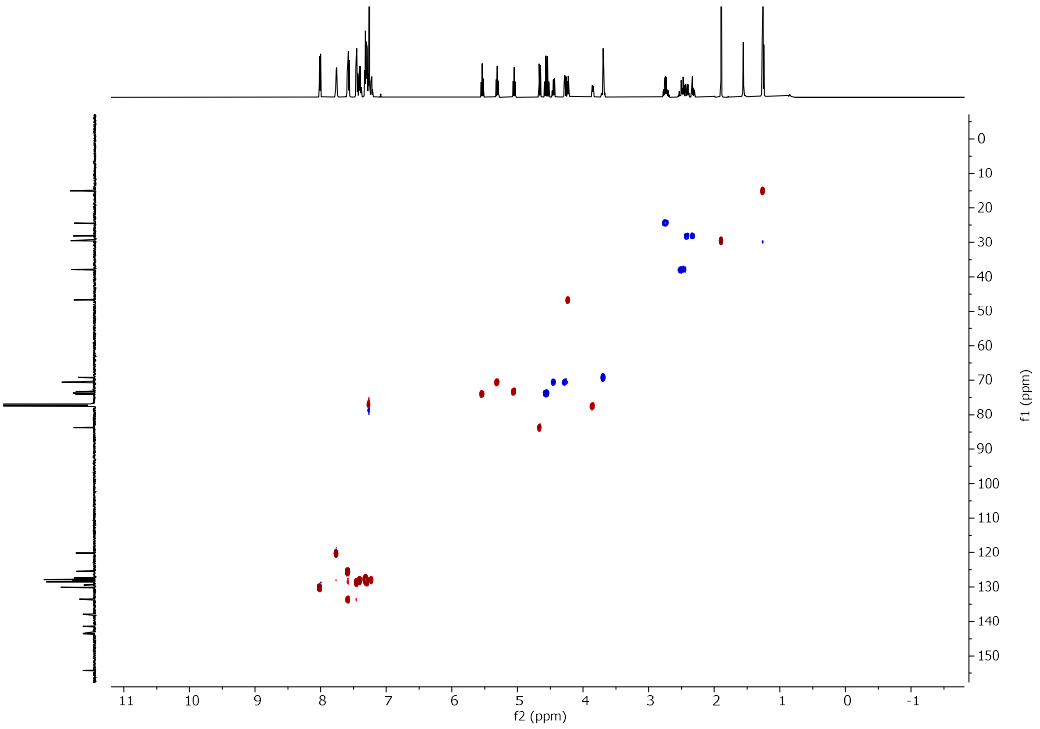
^{13}C NMR of BB4.4 (151 MHz, $CDCl_3$)



COSY NMR of BB4.4 ($CDCl_3$)



HSQC NMR of BB4.4 ($CDCl_3$)



6.5.2 Oligosaccharide synthesis

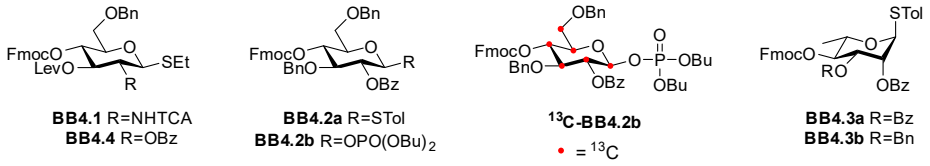


Figure S4.1 BBs used for the AGA of oligosaccharides described in Chapter 4.

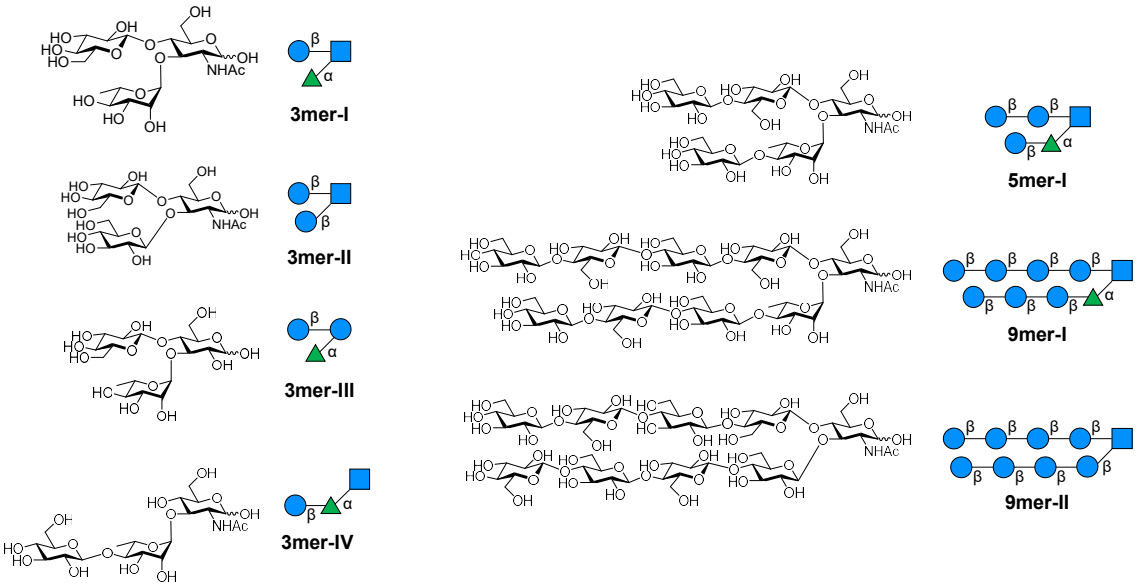


Figure S4.2 Oligosaccharides synthesized by AGA described in Chapter 4.

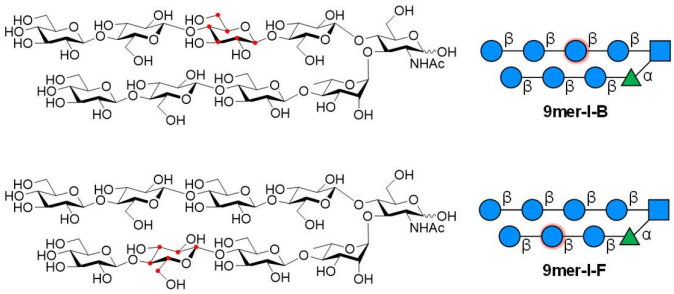
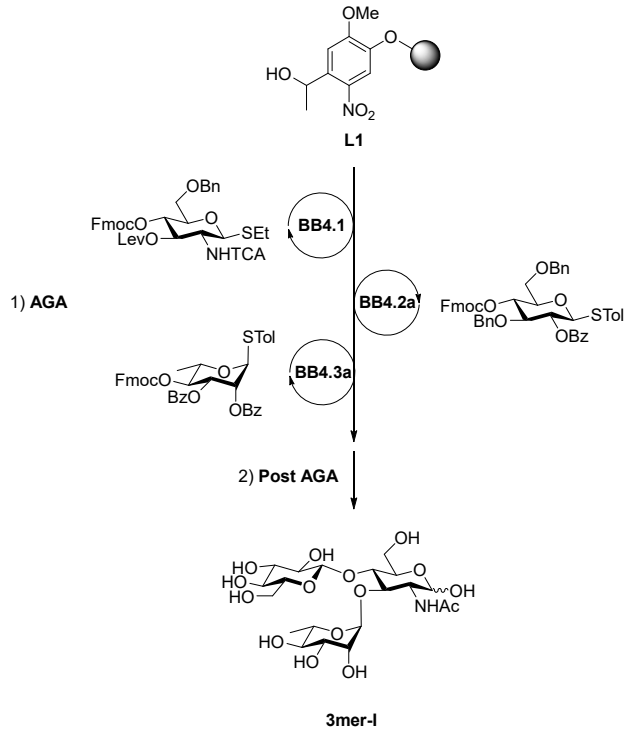


Figure S4.3 ^{13}C-labelled oligosaccharides synthesized by AGA described in Chapter 4.

Synthesis of **3mer-I**

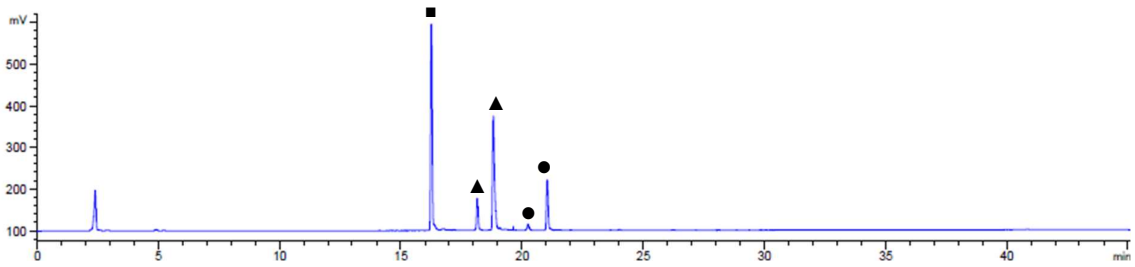


Step	BB	Modules	Notes
AGA	-	**A**	**L1** swelling
	BB4.1	**B**, **C1***, **D**, **E1**	**C1**: (**BB4.1**, -20 °C for 5 min, 0 °C for 40 min) *Double cycle
	BB4.2a	**B**, **C1***, **D**, **E1**	**C1**: (**BB4.2a**, -20 °C for 5 min, 0 °C for 20 min) *Double cycle
	-	**D**, **E2**	-
	BB4.2a	**B**, **C1**, **D**, **E1**	**C1**: (**BB4.2a**, -20 °C for 5 min, 0 °C for 20 min)
Post-AGA	-	**F1**, **G1**, **H2**, **I**	**F1**: (2 h) **H2**: (8 h) **I**: (Method A2: 19.1 min)

Automated synthesis, global deprotection, and purification afforded **3mer-II** as a white solid (0.7 mg, 11% overall yield).

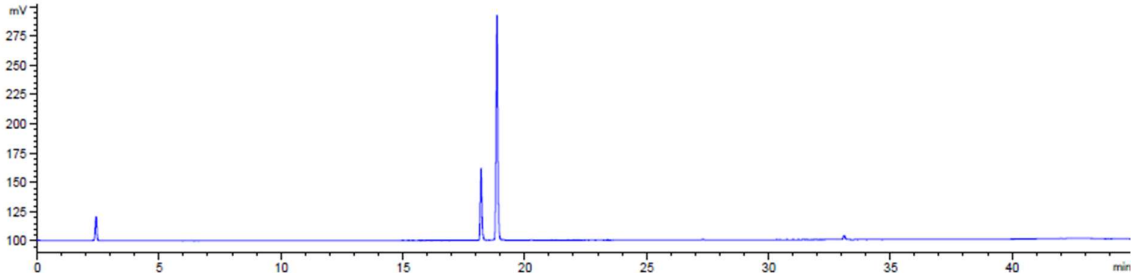
^{1}H NMR (600 MHz, D_2O) ^{1}H NMR (600 MHz, D_2O) δ 5.09 (d, *J* = 3.5 Hz, 0.6H, H-1α GlcNAc), 4.62 (d, *J* = 8.4 Hz, 0.4H, H-1β GlcNAc), 4.61 – 4.53 (m, 2H, H-1 Glcx2), 4.21 – 4.15 (m, 0.6H), 4.05 – 4.00 (m, 0.4H), 3.98 (dd, *J* = 10.4, 3.5 Hz, 0.6H), 3.93 – 3.75 (m, 6H), 3.67 – 3.60 (m, 2H), 3.51 (s, 0.4H), 3.44 – 3.22 (m, 8H), 1.95 (s, 1.8H, CH_3 Ac GlcNAc α), 1.95 (s, 1.2H, CH_3 Ac GlcNAc β). ^{13}C NMR (176 MHz, D_2O) δ 101.14 (C-1 Glc), 100.96 (C-1 Glc), 94.80 (C-1β GlcNAc), 90.63 (C-1α GlcNAc), 76.22, 76.09, 75.68, 75.43, 75.40, 75.35, 73.11, 72.96, 72.89, 72.57, 71.36, 69.37, 60.66, 59.97, 53.71, 22.06. ESI-HRMS m/z 568.1872 $[M+Na]^+$ ($C_{20}H_{35}NO_{16}Na$ requires 568.1848).

RP-HPLC of crude 3mer-II (ELSD trace, Method A1)

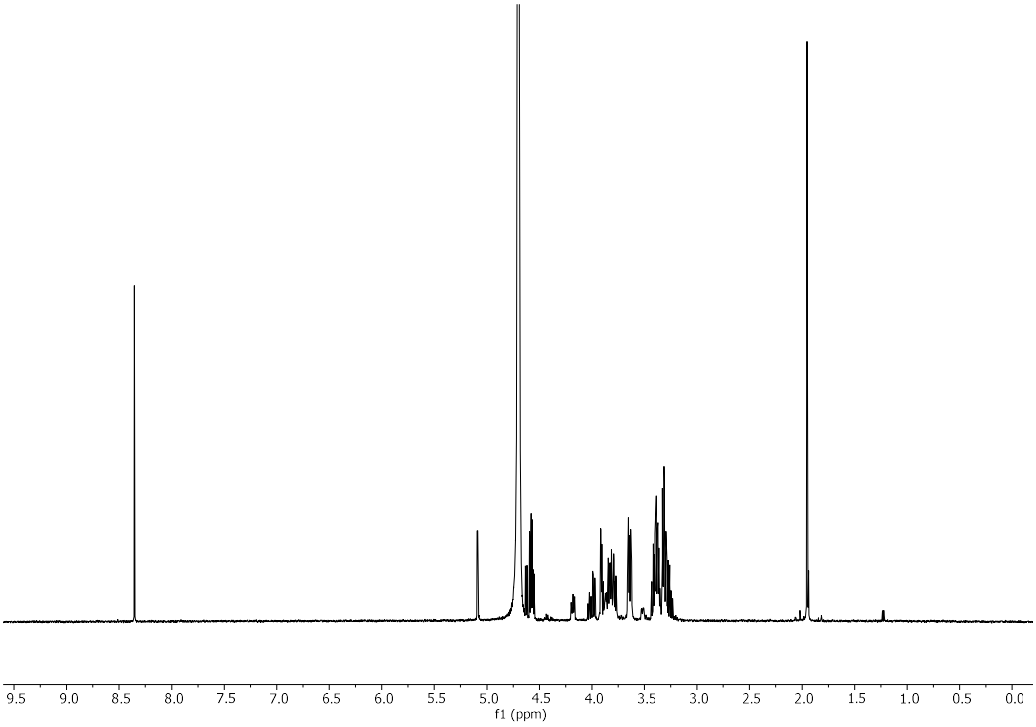


▲ **3mer-II** (α and β anomers, m/z 568 $[M+Na]^+$). ■ Sideproduct with GlcNAc reducing end hydrogenated (m/z 570 $[M+Na]^+$) formed during hydrogenolysis (Module I2). ● 2mer deletion sideproduct (m/z 406 $[M+Na]^+$).

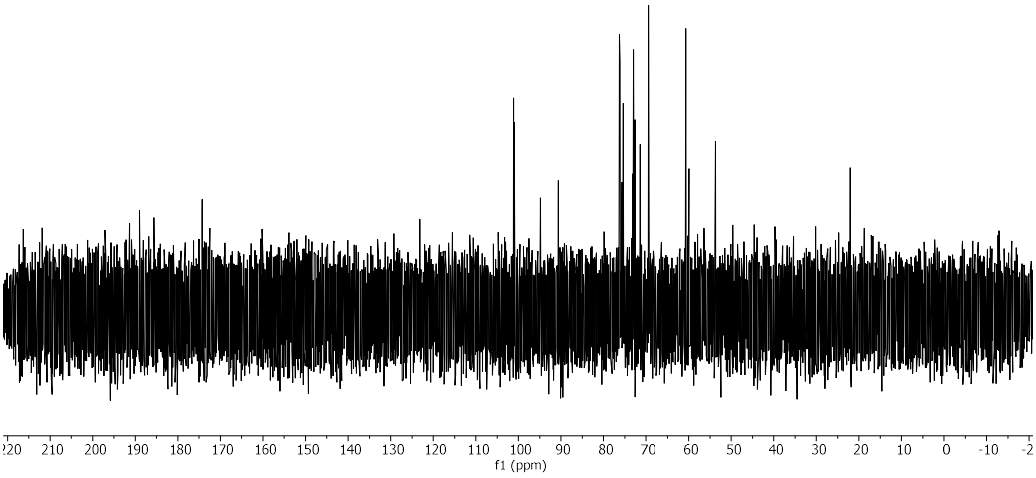
RP-HPLC of 3mer-II (ELSD trace, Method A1, t_R = 18.2, 18.9 min)



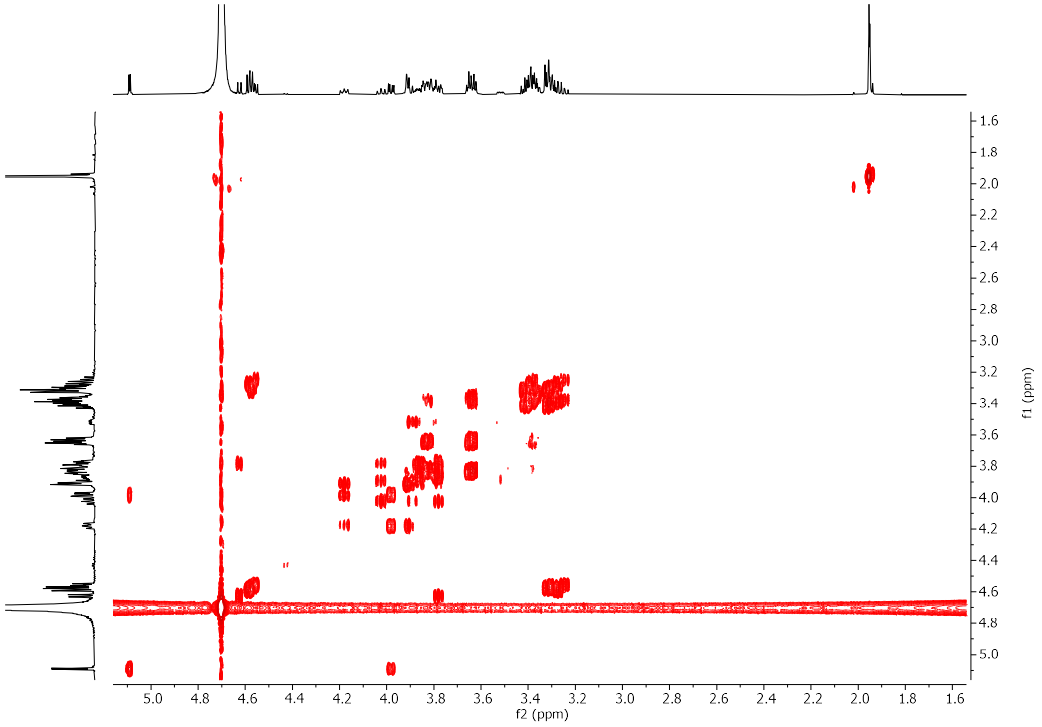
^{1}H NMR of 3mer-II (600 MHz, D_2O)



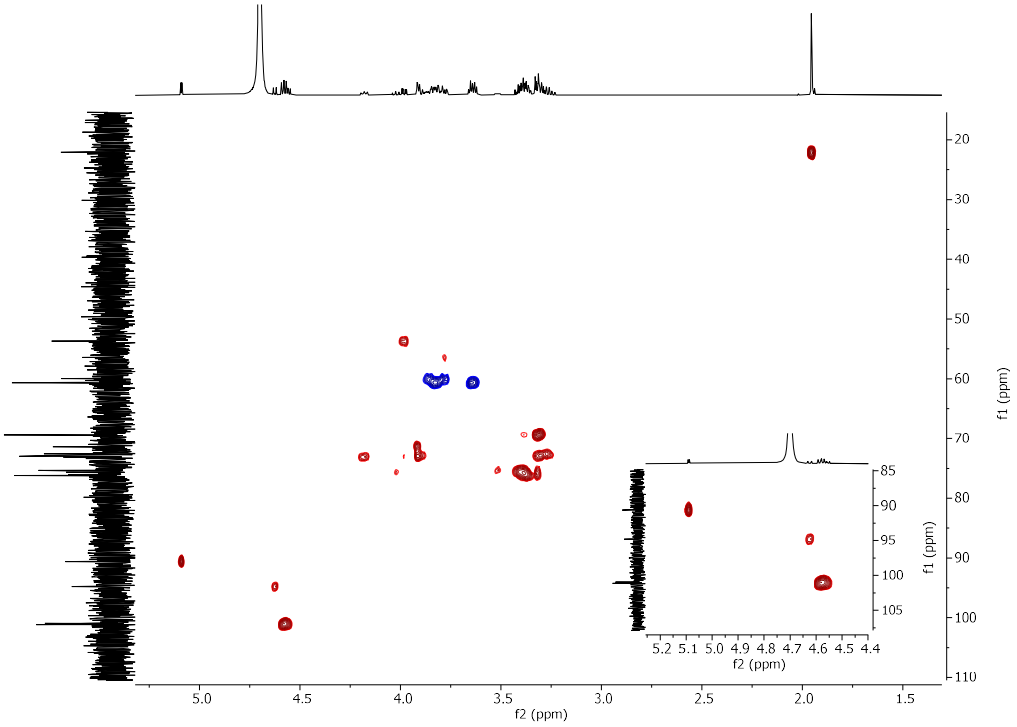
^{13}C NMR of 3mer-II (151 MHz, D_2O)



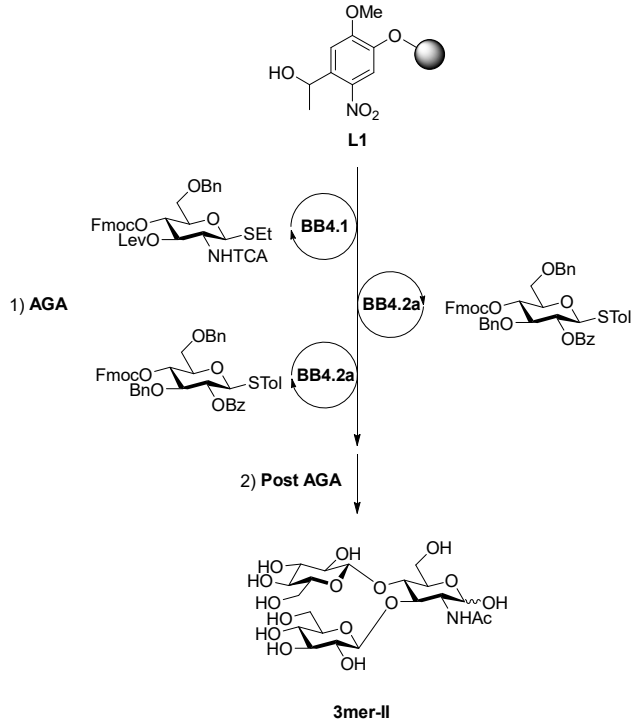
COSY NMR of 3mer-II (D_2O)



HSQC NMR of 3mer-II (D_2O)



Synthesis of **3mer-II**

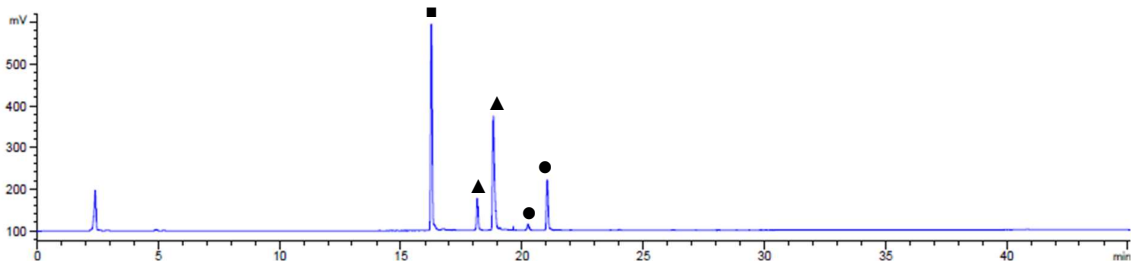


Step	BB	Modules	Notes
AGA	-	**A**	**L1** swelling
	BB4.1	**B**, **C1***, **D**, **E1**	**C1**: (**BB4.1**, -20 °C for 5 min, 0 °C for 40 min) *Double cycle
	BB4.2a	**B**, **C1***, **D**, **E1**	**C1**: (**BB4.2a**, -20 °C for 5 min, 0 °C for 20 min) *Double cycle
	-	**D**, **E2**	-
	BB4.2a	**B**, **C1**, **D**, **E1**	**C1**: (**BB4.2a**, -20 °C for 5 min, 0 °C for 20 min)
Post-AGA	-	**F1**, **G1**, **H2**, **I**	**F1**: (2 h) **H2**: (8 h) **I**: (Method A2: 19.1 min)

Automated synthesis, global deprotection, and purification afforded **3mer-II** as a white solid (0.7 mg, 11% overall yield).

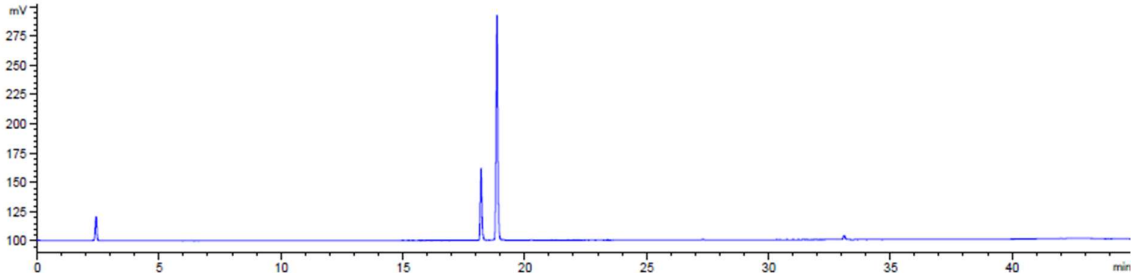
^{1}H NMR (600 MHz, D_2O) ^{1}H NMR (600 MHz, D_2O) δ 5.09 (d, *J* = 3.5 Hz, 0.6H, H-1α GlcNAc), 4.62 (d, *J* = 8.4 Hz, 0.4H, H-1β GlcNAc), 4.61 – 4.53 (m, 2H, H-1 Glcx2), 4.21 – 4.15 (m, 0.6H), 4.05 – 4.00 (m, 0.4H), 3.98 (dd, *J* = 10.4, 3.5 Hz, 0.6H), 3.93 – 3.75 (m, 6H), 3.67 – 3.60 (m, 2H), 3.51 (s, 0.4H), 3.44 – 3.22 (m, 8H), 1.95 (s, 1.8H, CH_3 Ac GlcNAc α), 1.95 (s, 1.2H, CH_3 Ac GlcNAc β). ^{13}C NMR (176 MHz, D_2O) δ 101.14 (C-1 Glc), 100.96 (C-1 Glc), 94.80 (C-1β GlcNAc), 90.63 (C-1α GlcNAc), 76.22, 76.09, 75.68, 75.43, 75.40, 75.35, 73.11, 72.96, 72.89, 72.57, 71.36, 69.37, 60.66, 59.97, 53.71, 22.06. ESI-HRMS m/z 568.1872 $[M+Na]^+$ ($C_{20}H_{35}NO_{16}Na$ requires 568.1848).

RP-HPLC of crude 3mer-II (ELSD trace, Method A1)

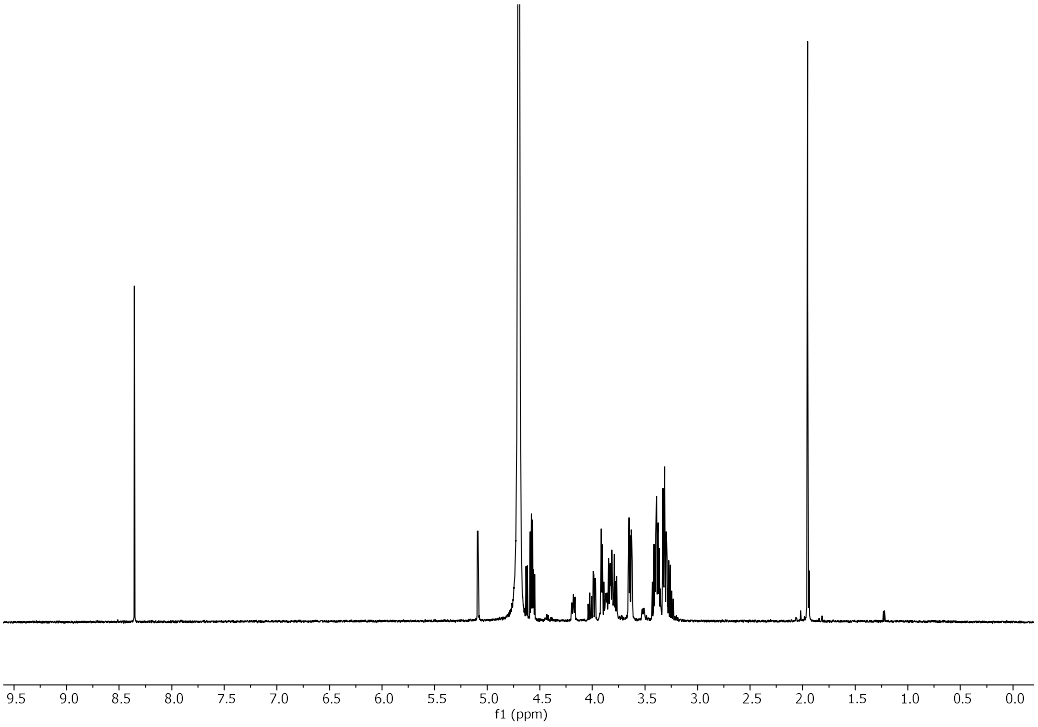


▲ **3mer-II** (α and β anomers, m/z 568 $[M+Na]^+$). ■ Sideproduct with GlcNAc reducing end hydrogenated (m/z 570 $[M+Na]^+$) formed during hydrogenolysis (Module I2). ● 2mer deletion sideproduct (m/z 406 $[M+Na]^+$).

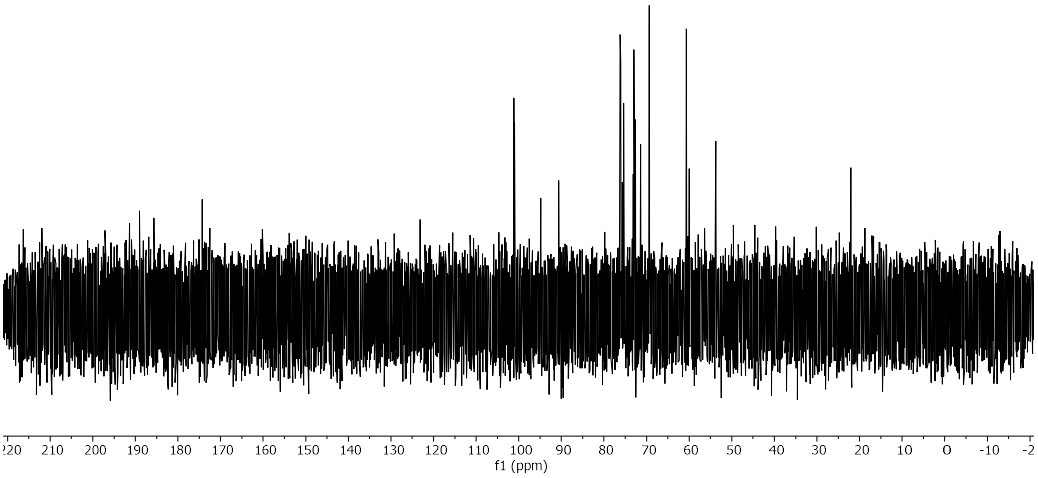
RP-HPLC of 3mer-II (ELSD trace, Method A1, t_R = 18.2, 18.9 min)



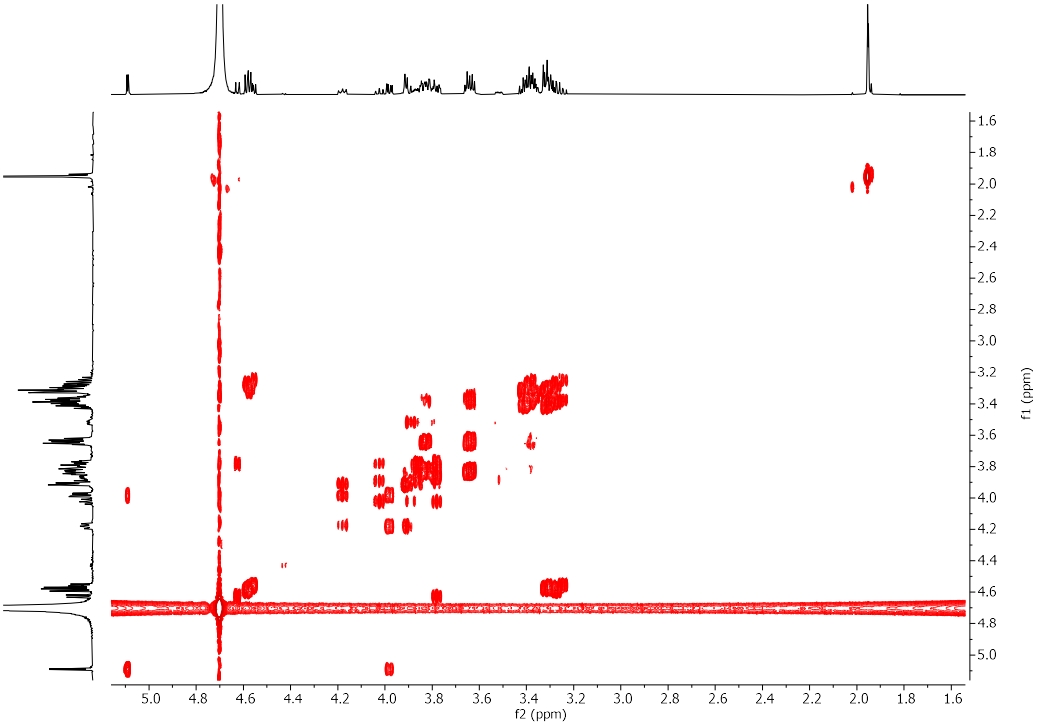
^{1}H NMR of 3mer-II (600 MHz, D_2O)



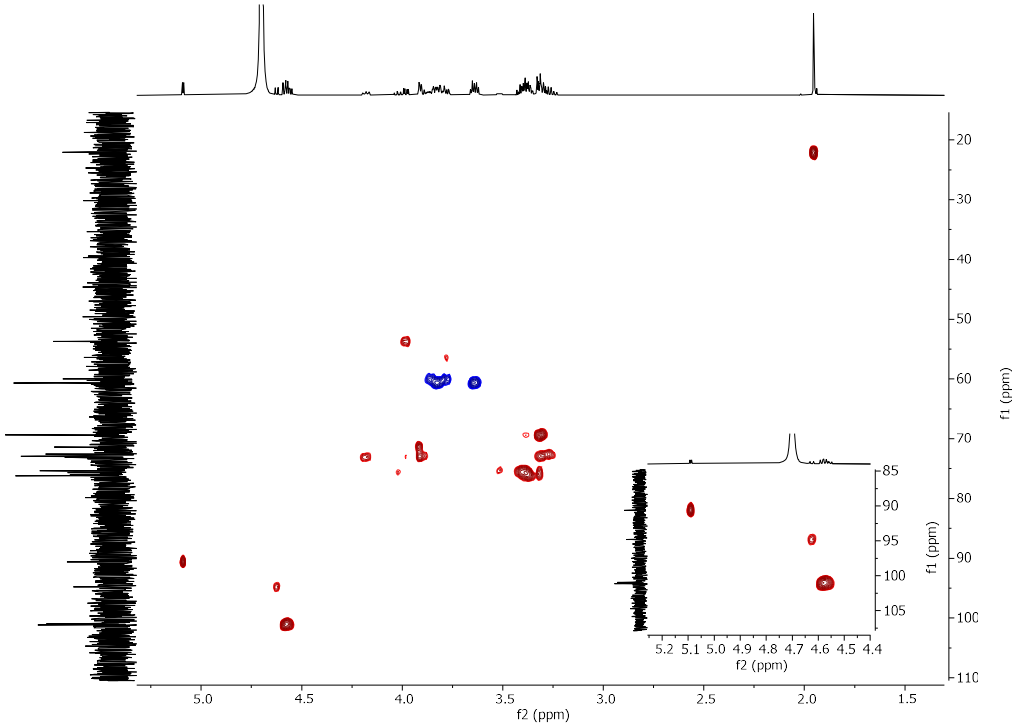
^{13}C NMR of 3mer-II (151 MHz, D_2O)



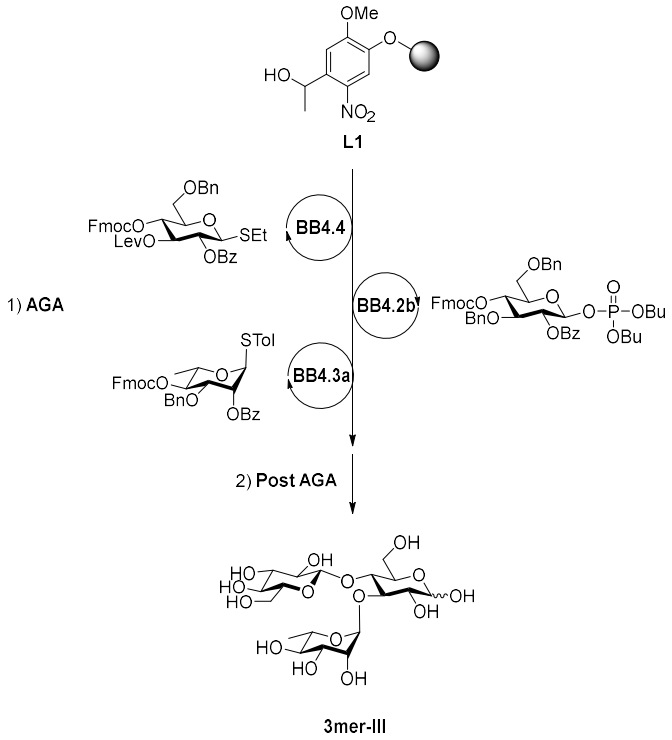
COSY NMR of 3mer-II (D_2O)



HSQC NMR of 3mer-II (D_2O)



Synthesis of **3mer-III**

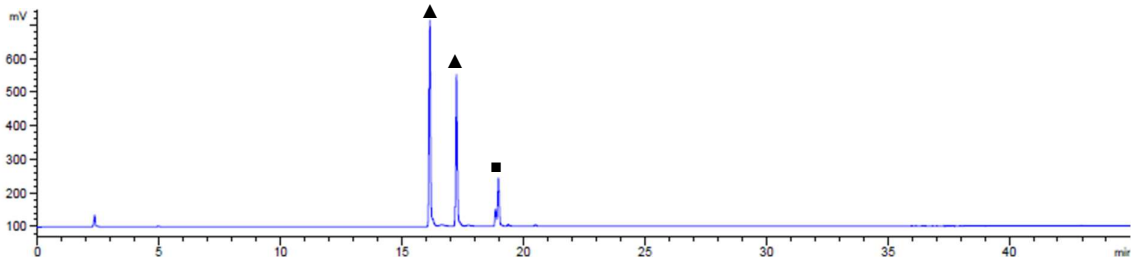


Step	BB	Modules	Notes
AGA	-	**A**	**L1** swelling
	BB4.4	**B**, **C1**, **D**, **E1**	**C1**: (**BB4.4**, -20 °C for 5 min, 0 °C for 20 min)
	BB4.2b	**B**, **C1***, **D**, **E1**,	**C1***: (**BB4.2b**, -30 °C for 5 min, -10 °C for 40 min) *Double cycle
	-	**D**, **E2**	-
	BB4.3a	**B**, **C1**, **D**, **E1**	**C1**: (**BB4.3a**, -20 °C for 5 min, 0 °C for 20 min)
Post-AGA	-	**F1**, **G1**, **H2**, **I**	**F1**: (4 h) **H2**: (3 h) **I**: (Method A2: 18.1 min)

Automated synthesis, global deprotection, and purification afforded **3mer-III** as a white solid (1.7 mg, 27% overall yield).

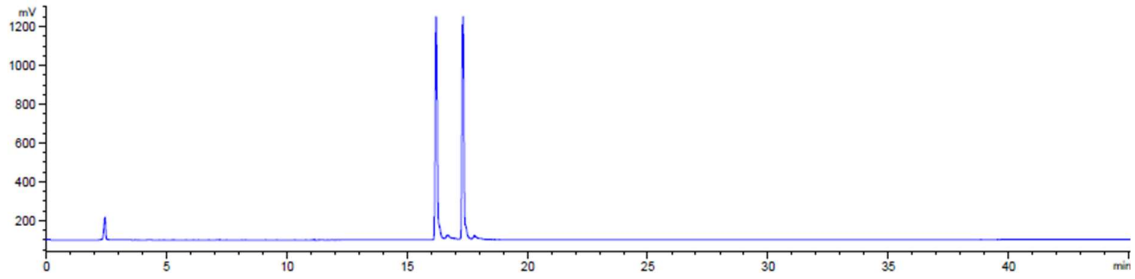
^{1}H NMR (400 MHz, D_2O) δ 5.27 (d, *J* = 1.6 Hz, 0.6H, H-1 Rha), 5.22 (d, *J* = 1.5 Hz, 0.4H, H-1 Rha), 5.20 (d, *J* = 3.8 Hz, 0.4H, H-1α Glc), 4.67 (d, *J* = 8.0 Hz, 0.6H, H-1β Glc), 4.50 (d, *J* = 7.9 Hz, 1H, H-1 Glc), 4.42 (dt, *J* = 13.5, 6.5 Hz, 1H, H-5 Rha), 4.07 – 4.03 (m, 1H), 3.96 (t, *J* = 11.1 Hz, 2.4H), 3.91 – 3.74 (m, 5H), 3.69 (dd, *J* = 9.9, 3.8 Hz, 0.4H), 3.59 (s, 0.6H), 3.55 – 3.33 (m, 4.6H), 3.25 (t, *J* = 8.7 Hz, 1H), 1.26 (d, *J* = 5.6 Hz, 3H, CH_3-6 Rha). ^{13}C NMR (101 MHz, D_2O) δ 101.49 (C-1 Glc), 100.75 (C-1 Rha), 100.62 (C-1 Rha), 95.75 (C-1α Glc), 92.01 (C-1β Glc), 78.30, 76.13, 75.94, 75.48, 75.30, 75.26, 73.57, 72.89, 72.85, 72.43, 71.89, 70.90, 70.10, 69.95, 69.66, 68.39 (C-5 Rha), 61.07, 59.56, 16.45 (CH_3-6 Rha). ESI-HRMS m/z 511.1666 $[M+Na]^+$ ($C_{18}H_{32}O_{15}Na$ requires 511.1633).

RP-HPLC of crude 3mer-III (ELSD trace, Method A1)

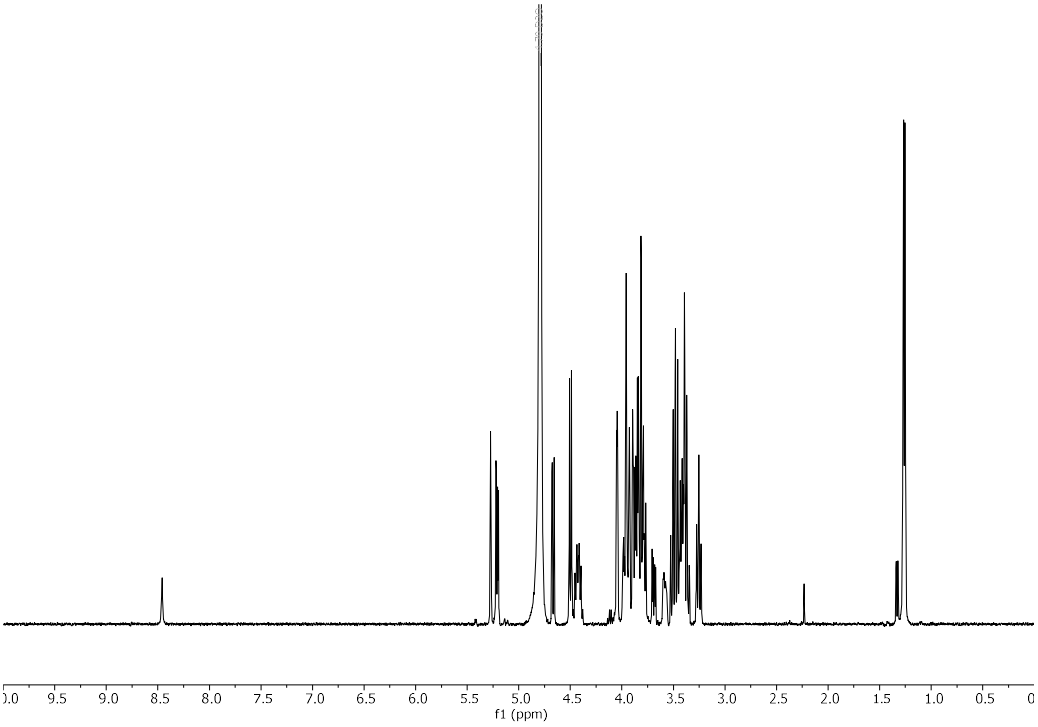


▲ **3mer-III** (α and β anomers, m/z 568 $[M+Na]^+$). ■ 2mer deletion sideproduct (-Rha, m/z 365 $[M+Na]^+$).

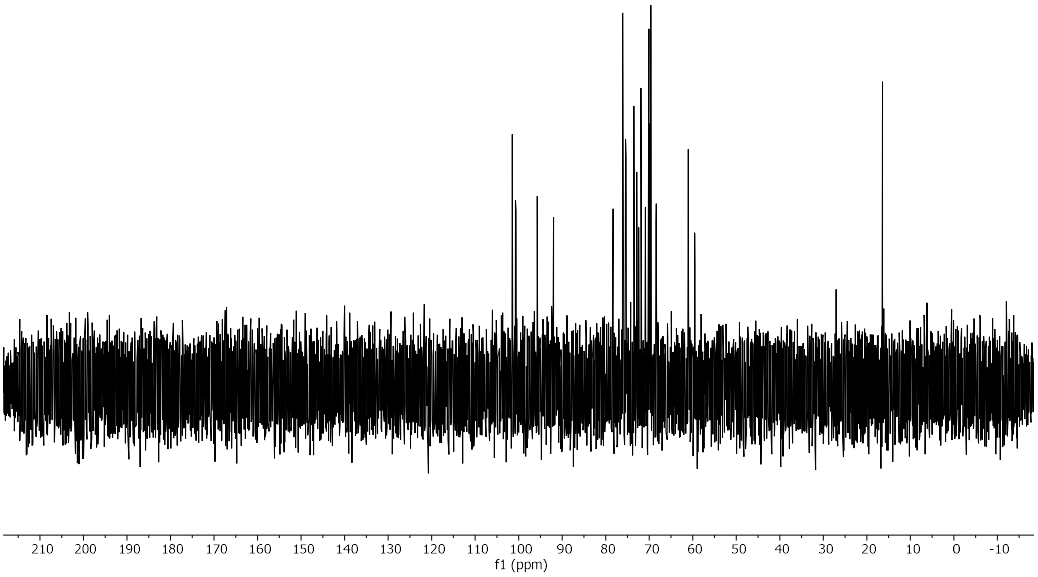
RP-HPLC of 3mer-III (ELSD trace, Method A1, t_R = 16.2, 17.3 min)



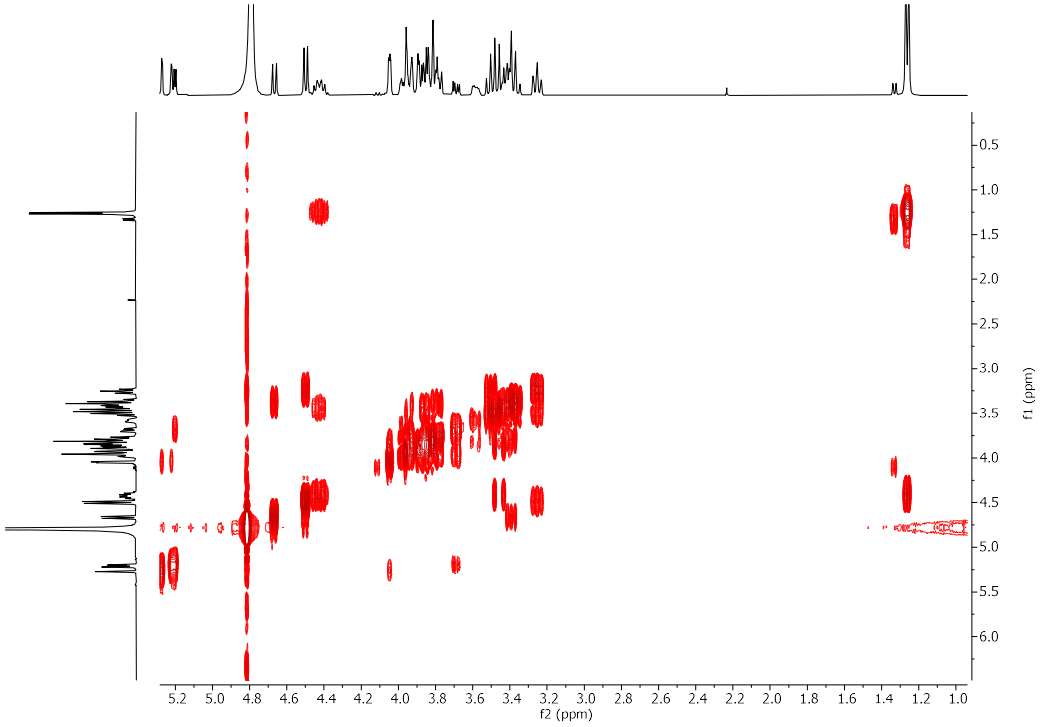
^{1}H NMR of 3mer-III (400 MHz, D_2O)



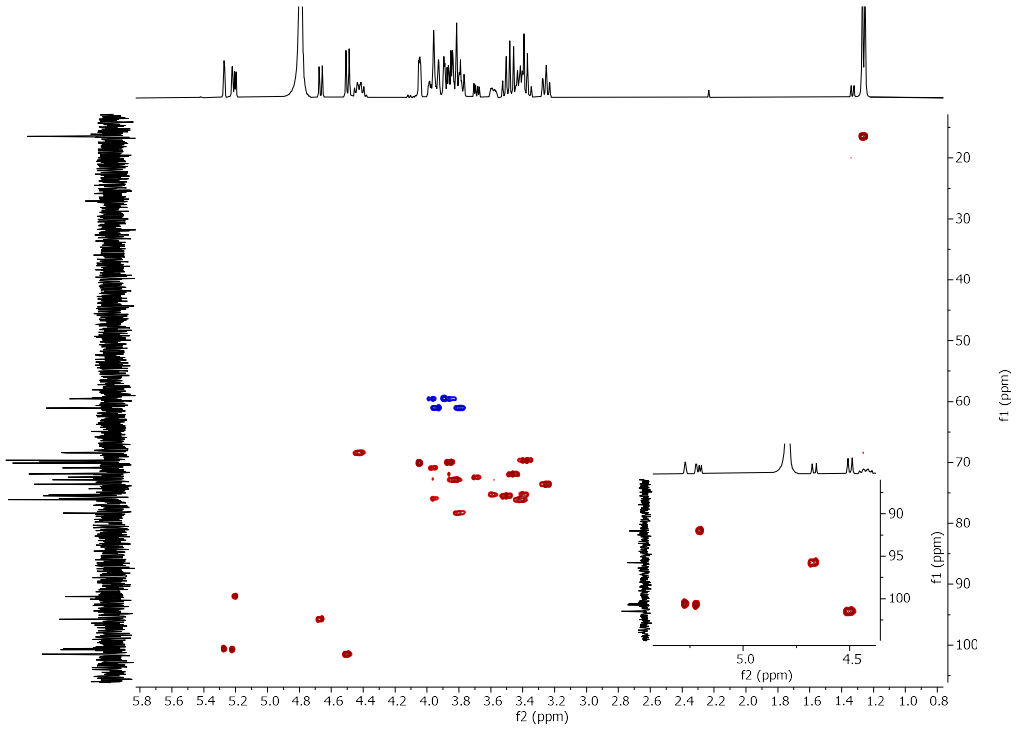
^{13}C NMR of 3mer-III (151 MHz, D_2O)



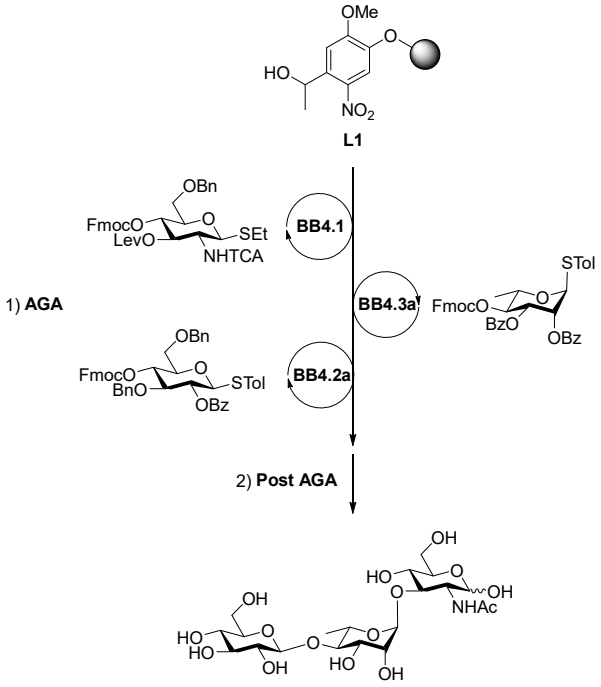
COSY NMR of 3mer-III (D_2O)



HSQC NMR of 3mer-III (D_2O)



Synthesis of **3mer-IV**

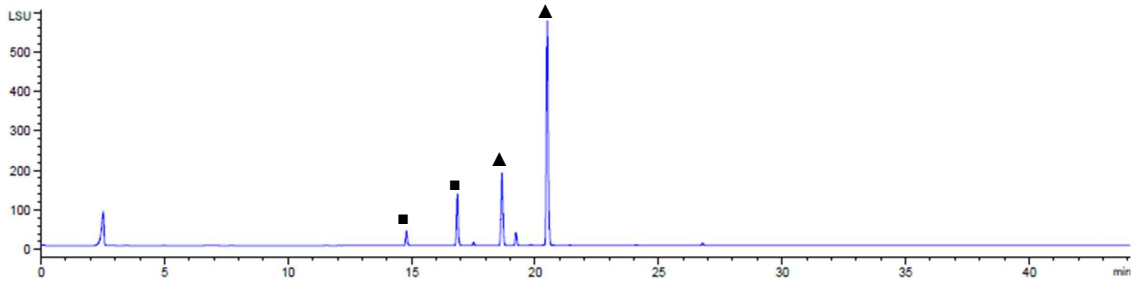


Step	BB	Modules	Notes
AGA	-	**A**	**L1** swelling
	BB4.1	**B**, **C1***, **D**, **E1**	**C1***: (**BB4.1**, -20 °C for 5 min, 0 °C for 40 min) *Double cycle
	-	**D**, **E2**	-
	BB4.3a	**B**, **C1**, **D**, **E1**	**C1**: (**BB4.3a**, -20 °C for 5 min, 0 °C for 20 min)
	BB4.2a	**B**, **C1**, **D**, **E1**	**C1**: (**BB4.2a**, -20 °C for 5 min, 0 °C for 20 min)
Post-AGA	-	**F1**, **G1**, **H2**, **I**	**F1**: (24 h) **H2**: (4 h) **I**: (Method A2: 20.5 min)

Automated synthesis, global deprotection, and purification afforded **3mer-IV** as a white solid (0.7 mg, 11% overall yield).

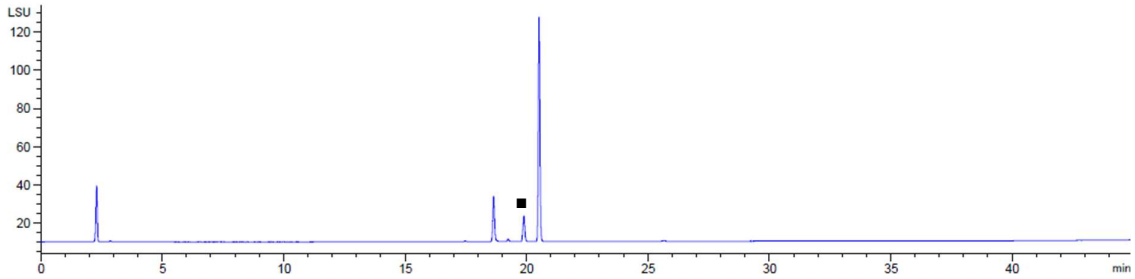
^{1}H NMR (600 MHz, D_2O) δ 5.14 (d, *J* = 3.5 Hz, 0.6H, H-1α GlcNAc), 4.87 – 4.82 (m, 1H, H-1 Rha), 4.75 (d, *J* = 8.7 Hz, 0.4H, H-1β GlcNAc), 4.71 (d, *J* = 8.0 Hz, 1H, H-1 Glc), 4.06 (dt, *J* = 9.8, 6.4 Hz, 1H, H-5 Rha), 4.02 (dd, *J* = 10.5, 3.5 Hz, 0.6H), 3.95 (dd, *J* = 9.5, 3.3 Hz, 1H), 3.93 – 3.70 (m, 6.4H), 3.67 (t, *J* = 9.6 Hz, 1H), 3.61 – 3.42 (m, 4H), 3.41 – 3.36 (m, 1H), 3.33 – 3.27 (m, 1H), 2.06 (s, 3H, CH_3 Ac GlcNAc), 1.31 – 1.27 (m, 3H, CH_3-6 Rha). ^{13}C NMR (151 MHz, D_2O) δ 103.21 (C-1 Glc), 101.08 (C-1 Rha), 94.16 (C-1β GlcNAc), 90.94 (C-1α GlcNAc), 81.76, 80.99, 79.31, 75.90, 75.69, 73.86, 71.74, 70.73, 70.30, 69.47, 68.37, 67.35 (C-5 Rha), 60.53, 56.28, 53.55, 21.88 (CH_3 Ac GlcNAc), 16.63 (C-6 Rha). ESI-HRMS m/z 552.1901 $[M+Na]^+$ ($C_{20}H_{35}NO_{15}Na$ requires 552.1899).

RP-HPLC of crude 3mer-IV (ELSD trace, Method A1)



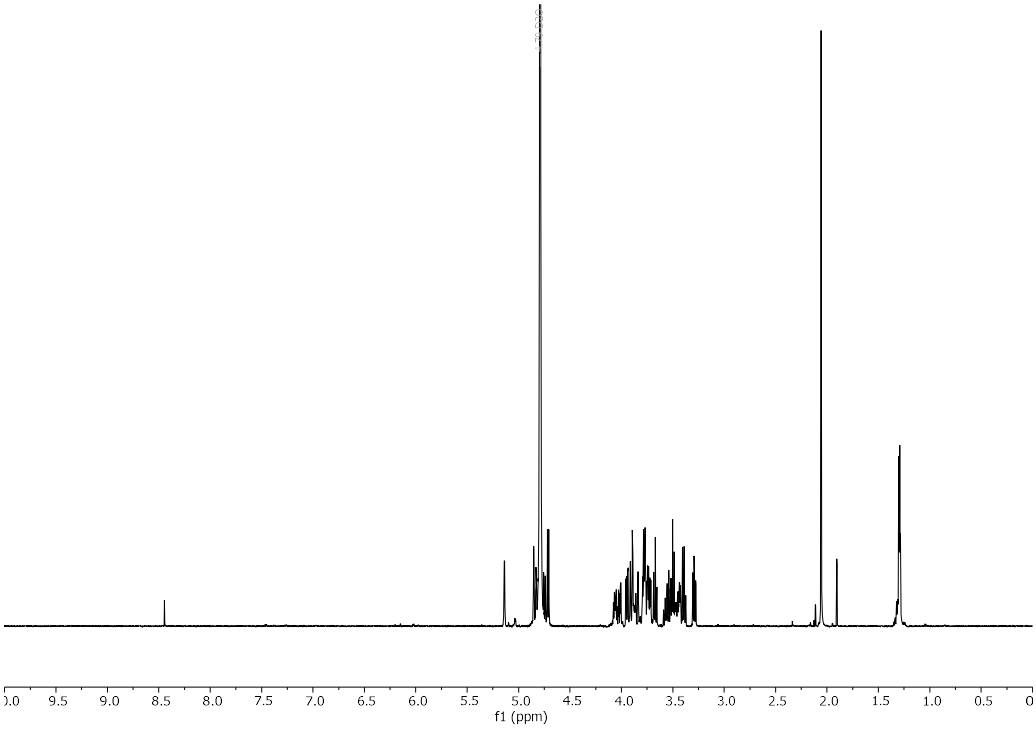
▲**3mer-IV** (α and β anomers, m/z 552 $[M+Na]^+$). ■ 2mer deletion sideproduct (-Glc, m/z 390 $[M+Na]^+$).

RP-HPLC of 3mer-IV (ELSD trace, Method A1, t_R = 19.9, 20.5 min)

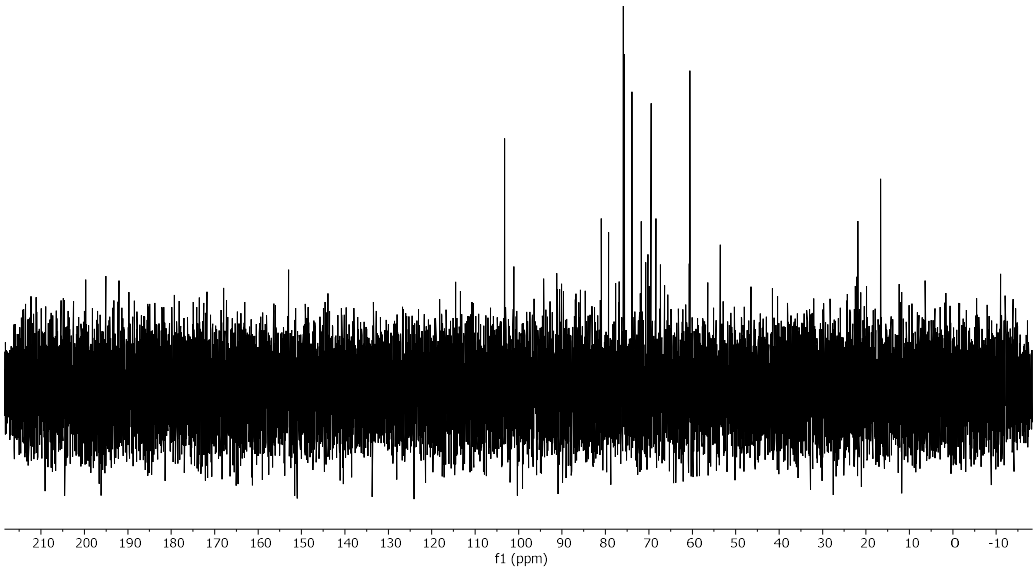


■ Unidentified impurity.

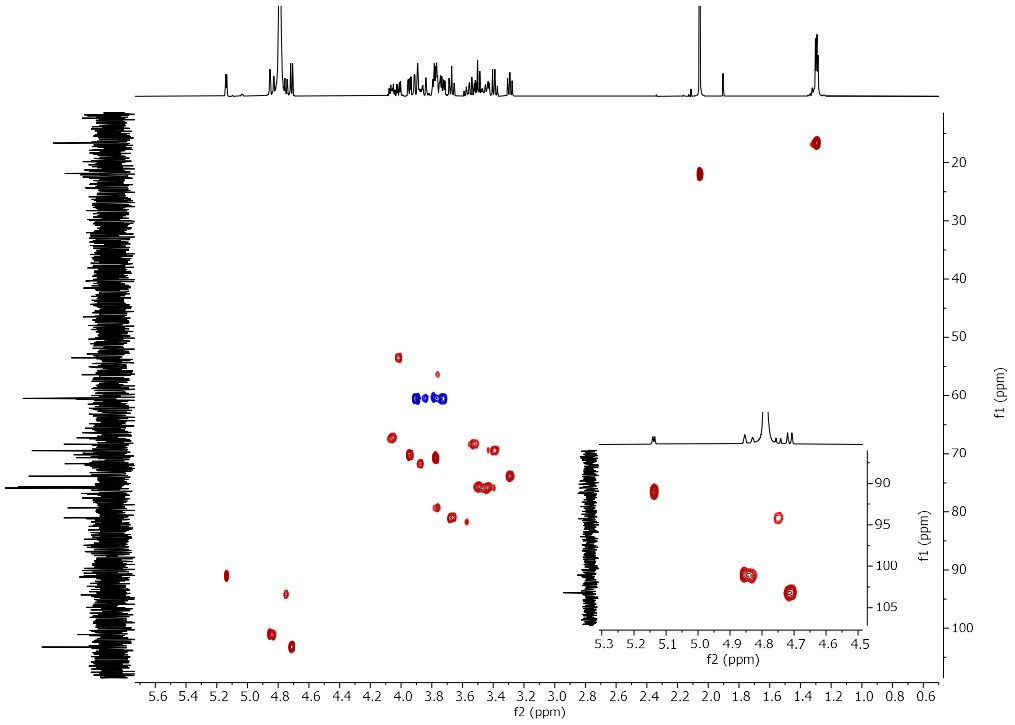
^{1}H NMR of 3mer-IV (600 MHz, D_2O)



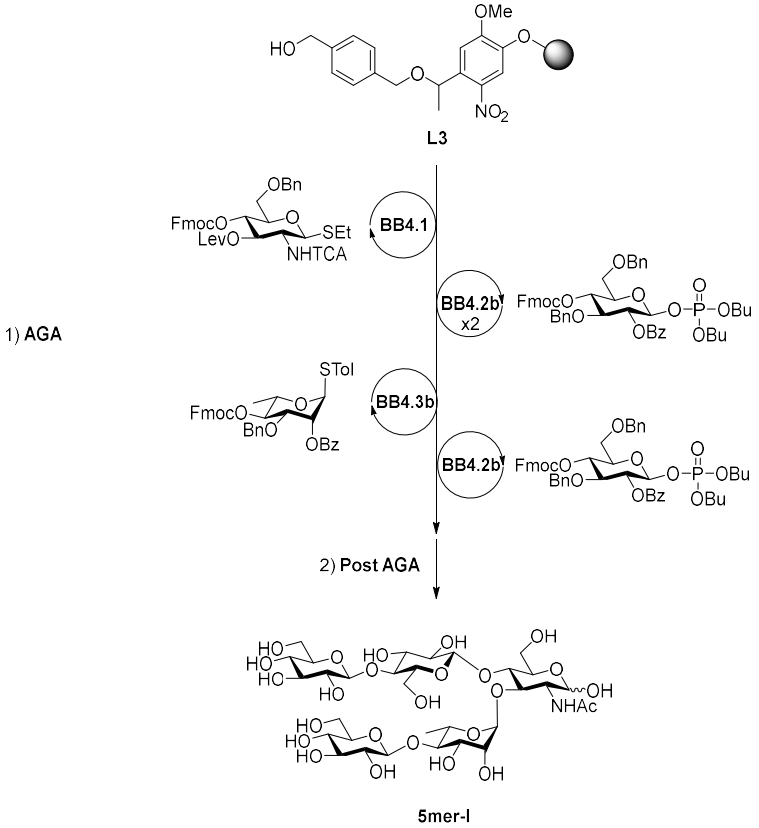
^{13}C NMR of 3mer-IV (151 MHz, D_2O)



HSQC NMR of 3mer-IV (D_2O)


f1 (ppm)
f2 (ppm)
f1 (ppm)
f2 (ppm)

Synthesis of **5mer-I**

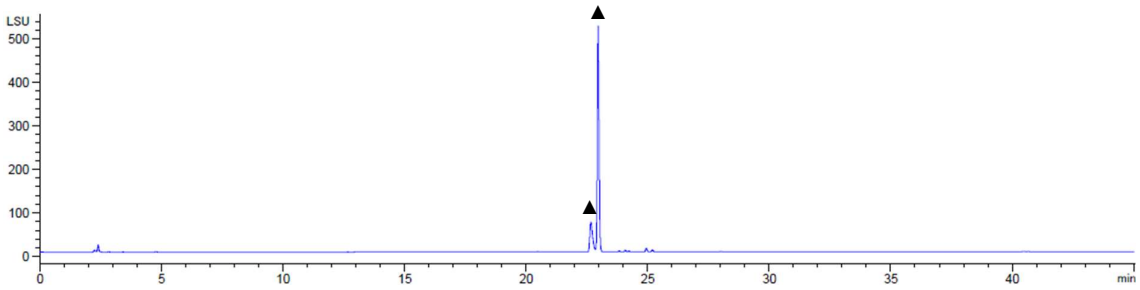


Step	BB	Modules	Notes
AGA	-	**A**	**L2** swelling
	BB4.1	**B**, **C1***, **D**, **E1**	**C1***: (**BB4.1**, -20 °C for 5 min, 0 °C for 40 min) *Double cycle
	BB4.2b	**B**, **C2***, **D**, **E1**	**C2***: (**BB4.2b**, -30 °C for 5 min, -10 °C for 40 min) *Double cycle
	BB4.2b	**B**, **C2**, **D**, **E1**	**C2**: (**BB4.2b**, -30 °C for 5 min, -10 °C for 40 min)
	-	**D**, **E2**	-
	BB4.3b	**B**, **C1**, **D**, **E1**	**C1**: (**BB4.3b**, -20 °C for 5 min, 0 °C for 20 min)
	BB4.2b	**B**, **C2**, **D**, **E1**	**C2**: (**BB4.2b**, -30 °C for 5 min, -10 °C for 40 min)
Post-AGA	-	**F1**, **G1**, **H2**, **I**	**F1**: (48 h) **H2**: (3 h) **I**: (Method A2: 23.1 min)

Automated synthesis, global deprotection, and purification afforded **5mer-I** as a white solid (2.1 mg, 19% overall yield).

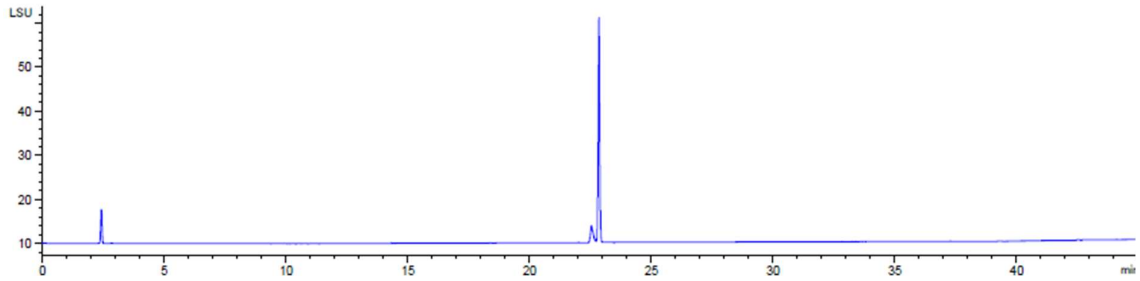
^{1}H NMR (600 MHz, D_2O) δ 5.10 (d, *J* = 3.5 Hz, 0.6H, H-1α GlcNAc), 4.93 – 4.90 (m, 1H, H-1 Rha), 4.73 (d, *J* = 2.9 Hz, 0.4H, H-1β GlcNAc), 4.65 (d, *J* = 7.9 Hz, 1H, H-1 Glc), 4.54 – 4.51 (m, 1H, H-1 Glc), 4.48 – 4.41 (m, 1H, H-5 Rha), 4.41 (d, *J* = 7.9 Hz, 1H, H-1 Glc), 4.10 (dd, *J* = 10.2, 3.6 Hz, 0.6H), 4.06 – 4.02 (m, 1H), 4.02 – 3.94 (m, 2.4H), 3.94 – 3.85 (m, 4.4H), 3.84 – 3.76 (m, 3H), 3.74 – 3.67 (m, 2H), 3.65 – 3.59 (m, 2H), 3.58 – 3.53 (m, 1.6H), 3.52 – 3.47 (m, 3H), 3.46 – 3.42 (m, 1H), 3.42 – 3.37 (m, 2H), 3.37 – 3.33 (m, 1H), 3.32 – 3.23 (m, 3H), 2.04 – 2.02 (m, 3H, CH_3 Ac GlcNAc), 1.31 – 1.27 (m, 3H, CH_3-6 Rha). ^{13}C NMR (151 MHz, D_2O) δ 103.66 (C-1 Glc), 102.86 (C-1 Glc), 101.04 (C-1 Glc), 100.15 (C-1 Rha), 94.26 (C-1β GlcNAc), 90.97 (C-1α GlcNAc), 81.29, 80.09, 75.99, 75.89, 75.31, 75.13, 74.53, 74.29, 74.27, 74.22, 73.15, 73.10, 71.09, 70.37, 70.31, 70.18, 70.13, 69.60, 69.26, 67.05 (C-5 Rha), 60.93, 60.63, 60.49, 56.84, 53.90, 21.86 (CH_3 Ac GlcNAc), 16.74 (C-6 Rha). ESI-HRMS m/z 876.2943 $[M+Na]^+$ ($C_{32}H_{55}NO_{25}Na$ requires 876.2955).

RP-HPLC of crude 5mer-I (ELSD trace, Method A1)

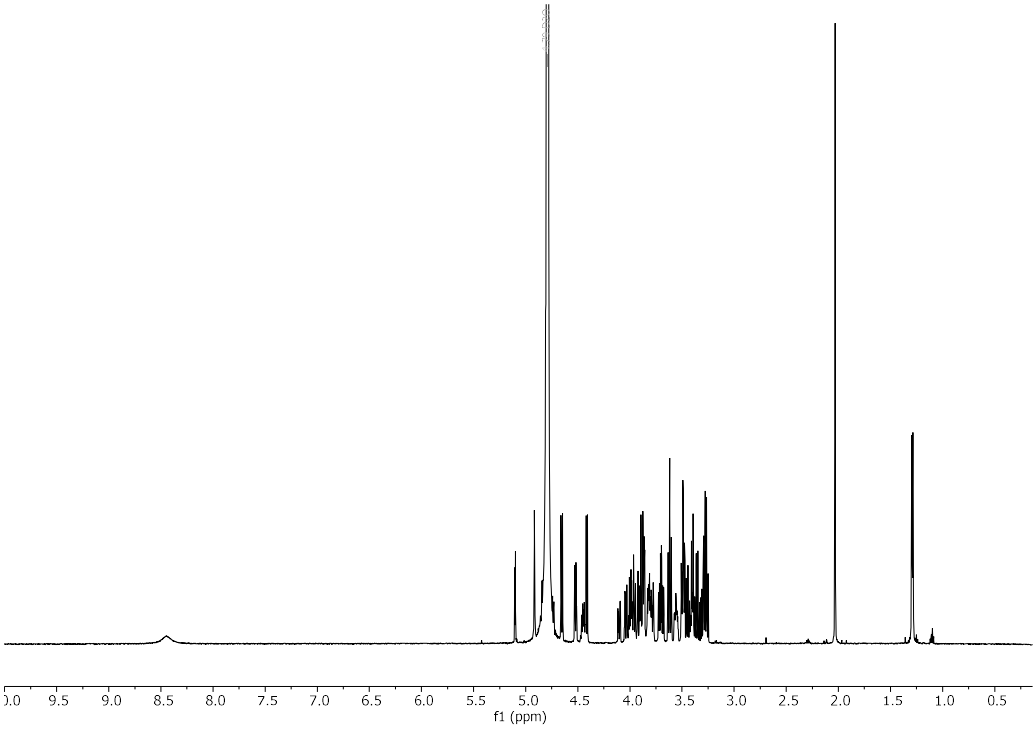


▲**5mer-I** (α and β anomers, m/z 876 $[M+Na]^+$).

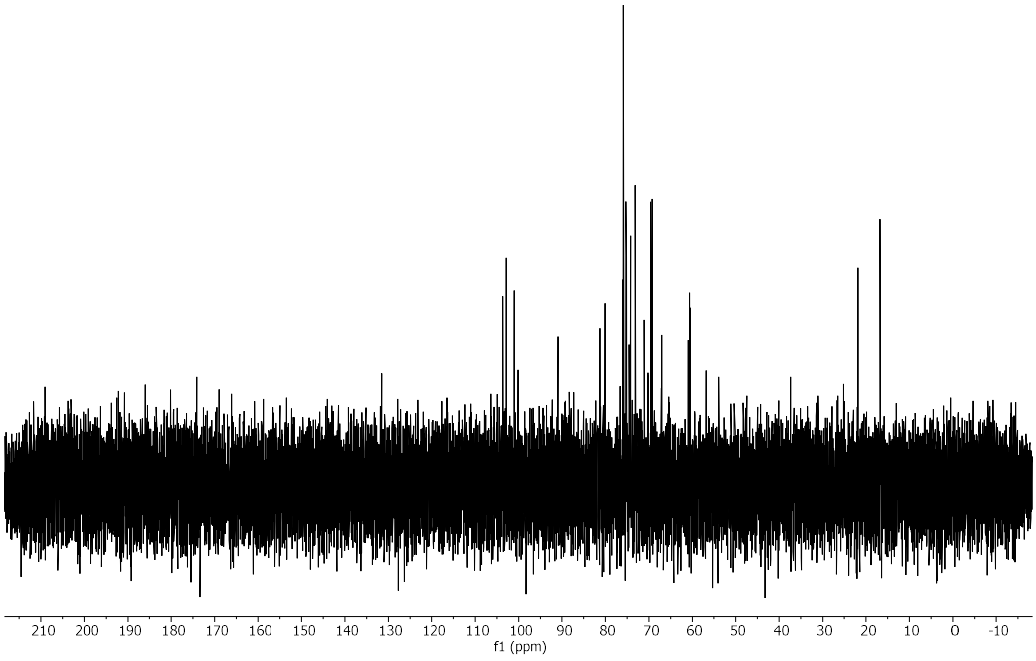
RP-HPLC of 5mer-I (ELSD trace, Method A1, t_R = 22.5, 22.8 min)



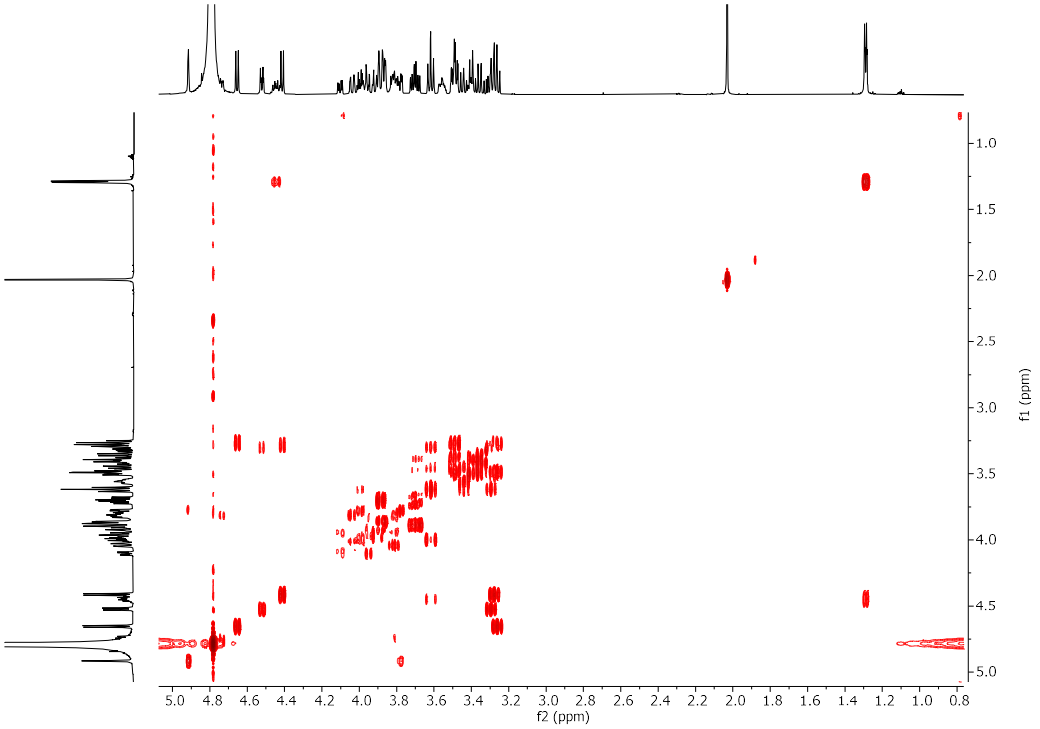
^{1}H NMR of 5mer-I (600 MHz, D_2O)



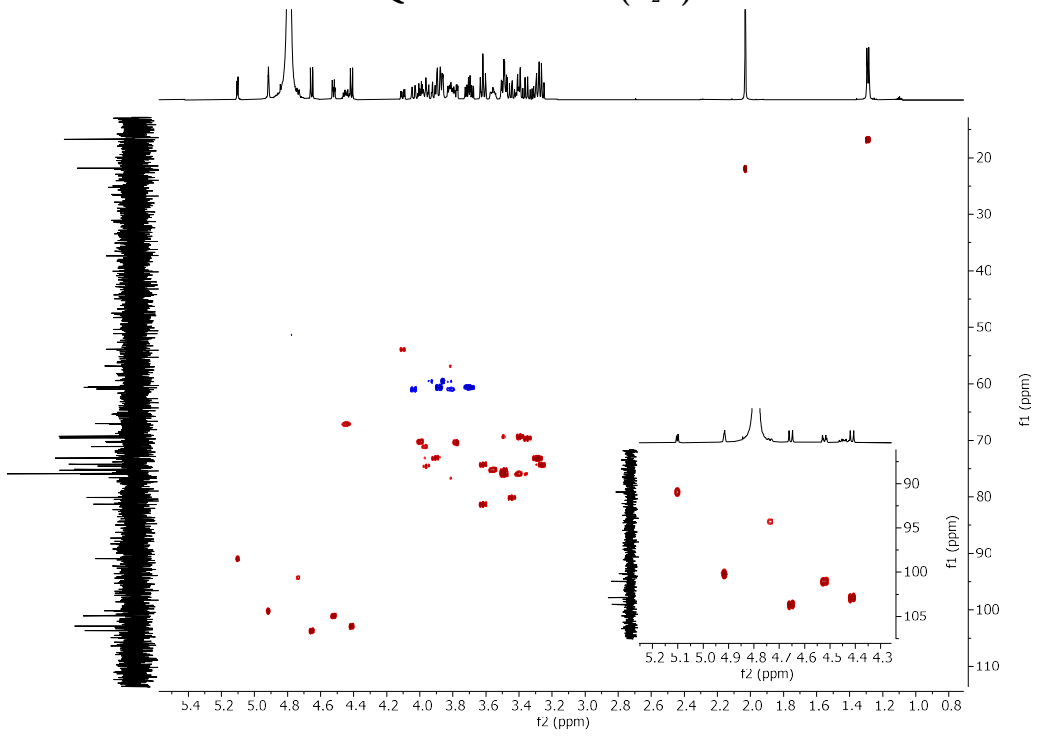
^{13}C NMR of 5mer-I (151 MHz, D_2O)



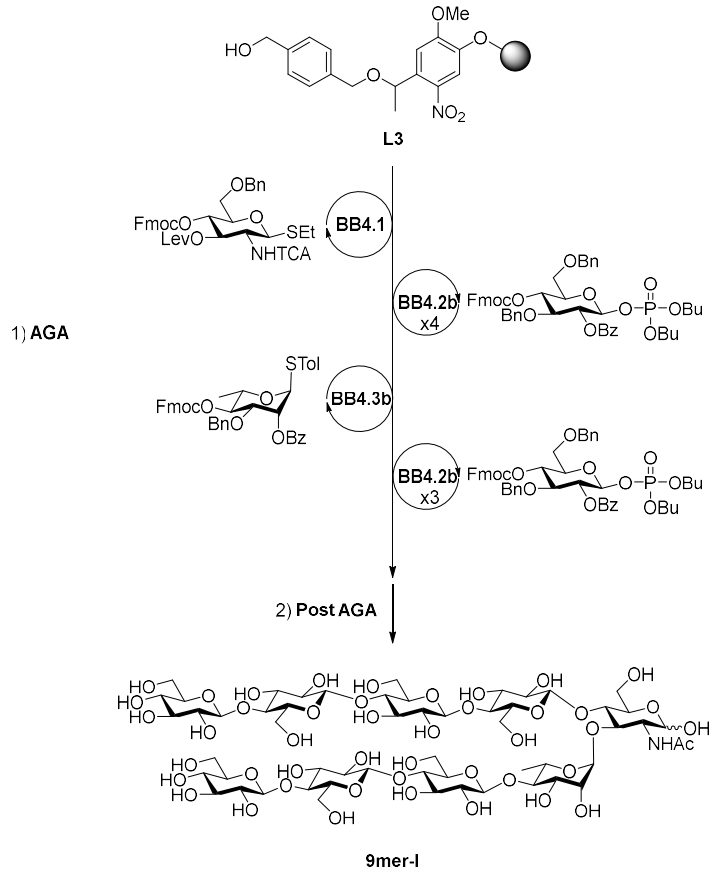
COSY NMR of 5mer-I (D_2O)



HSQC NMR of 5mer-I (D_2O)



Synthesis of **9mer-I**

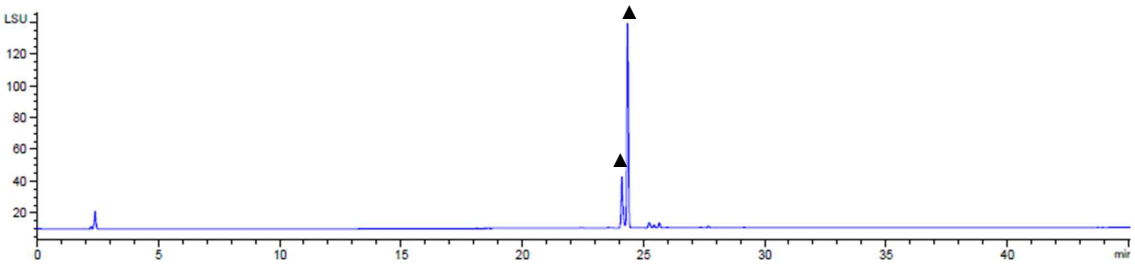


Step	BB	Modules	Notes
AGA	-	**A**	**L2** swelling
	BB4.1	**B**, **C1***, **D**, **E1**	**C1***: (**BB4.1**, -20 °C for 5 min, 0 °C for 40 min) *Double cycle
	BB4.2b	**B**, **C2***, **D**, **E1**	**C2***: (**BB4.2b**, -30 °C for 5 min, -10 °C for 40 min) *Double cycle
	(**BB4.2b**)x3	(**B**, **C2**, **D**, **E1**)x3	**C2**: (**BB4.2b**, -30 °C for 5 min, -10 °C for 40 min)
	-	**D**, **E2**	-
	BB4.3b	**B**, **C1**, **D**, **E1**	**C1**: (**BB4.3b**, -20 °C for 5 min, 0 °C for 20 min)
	(**BB4.2b**)x3	(**B**, **C2**, **D**, **E1**)x3	**C2**: (**BB4.2b**, -30 °C for 5 min, -10 °C for 40 min)
Post-AGA	-	**F1**, **G1**, **H2**, **I**	**F1**: (24 h) **H2**: (10 h) **I**: (Method B2: 26.9 min)

Automated synthesis, global deprotection, and purification afforded **9mer-I** as a white solid (2.3 mg, 12% overall yield).

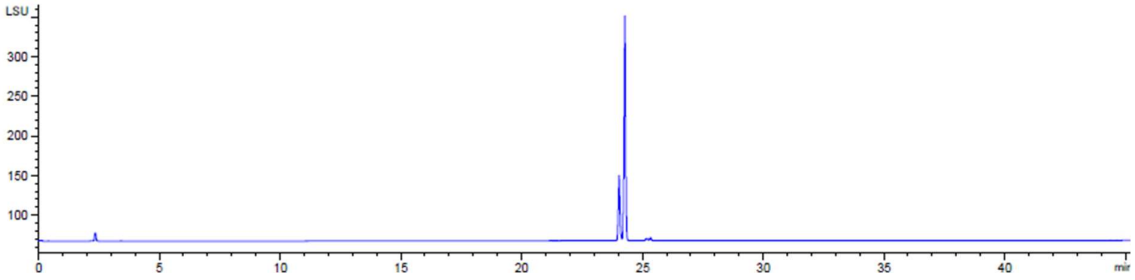
^{1}H NMR (600 MHz, D_2O) δ 5.10 (d, *J* = 3.6 Hz, 0.6H, H-1α GlcNAc), 4.91 (s, 1H, H-1 Rha), 4.73 (d, *J* = 8.0 Hz, 0.4H, H-1β GlcNAc), 4.67 (d, *J* = 7.7 Hz, 1H, H-1 Glc), 4.54 – 4.45 (m, 6H, 5xH-1 Glc, H-5 Rha), 4.44 (d, *J* = 7.9 Hz, 1H, H-1 Glc), 4.12 – 4.09 (m, 0.6H), 4.05 (d, *J* = 11.2 Hz, 1H), 4.03 – 3.92 (m, 7.4H), 3.92 – 3.85 (m, 4H), 3.85 – 3.76 (m, 7H), 3.74 – 3.52 (m, 17H), 3.47 (q, *J* = 8.6 Hz, 5H), 3.39 (t, *J* = 9.1 Hz, 2H), 3.36 – 3.26 (m, 7H), 2.03 (d, *J* = 1.2 Hz, 3H, CH_3 Ac GlcNAc), 1.33 – 1.26 (m, 3H, CH_3-6 Rha). ^{13}C NMR (151 MHz, D_2O) δ 103.58 (C-1 Glc), 102.75 (C-1 Glc), 102.57 (C-1 Glc), 102.51 (C-1 Glc), 102.37 (C-1 Glc), 100.95(C-1 Glc), 100.01 (C-1 Rha), 94.17 (C-1β GlcNAc), 90.88 (C-1α GlcNAc), 81.57, 80.31, 78.76, 78.16, 75.89, 75.45, 75.37, 75.09, 74.65, 74.44, 74.25, 74.04, 73.95, 73.06, 72.98, 72.79, 69.35, 66.96 (C-5 Rha), 60.46, 59.71, 56.82, 53.93, 21.82 (CH_3 Ac GlcNAc), 16.71 (C-6 Rha). ESI-HRMS m/z 1524.5063 $[M+Na]^+$ ($C_{56}H_{95}NO_{45}Na$ requires 1524.5068).

RP-HPLC of crude 9mer-I (ELSD trace, Method B1)

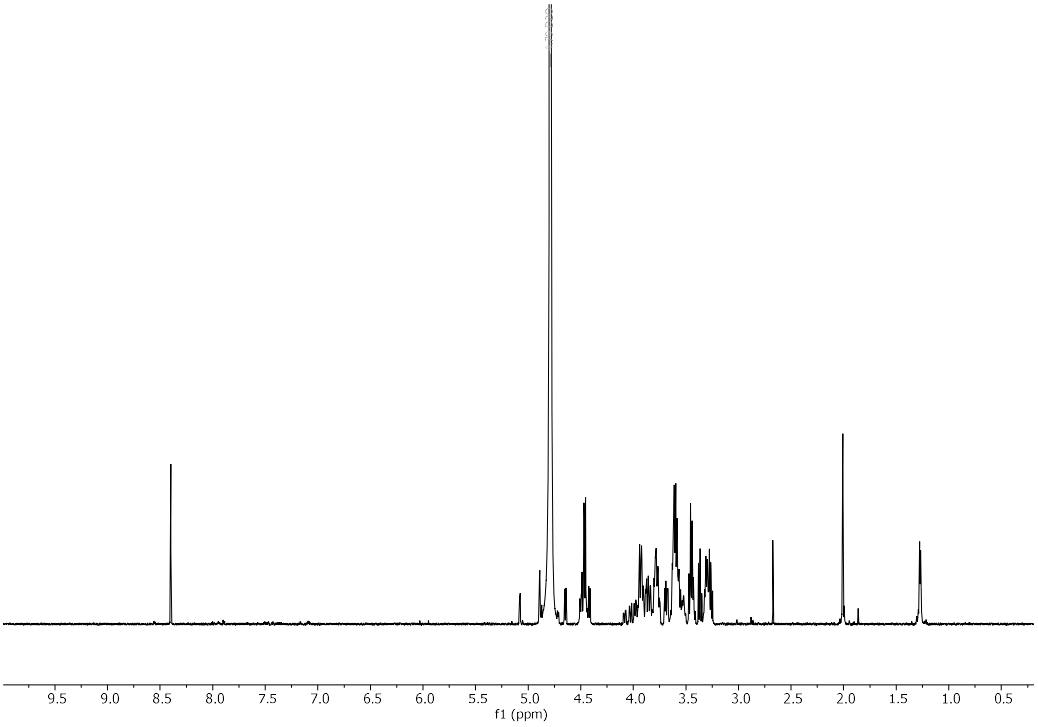


▲**9mer-I** (α and β anomers, m/z 1524 $[M+Na]^+$).

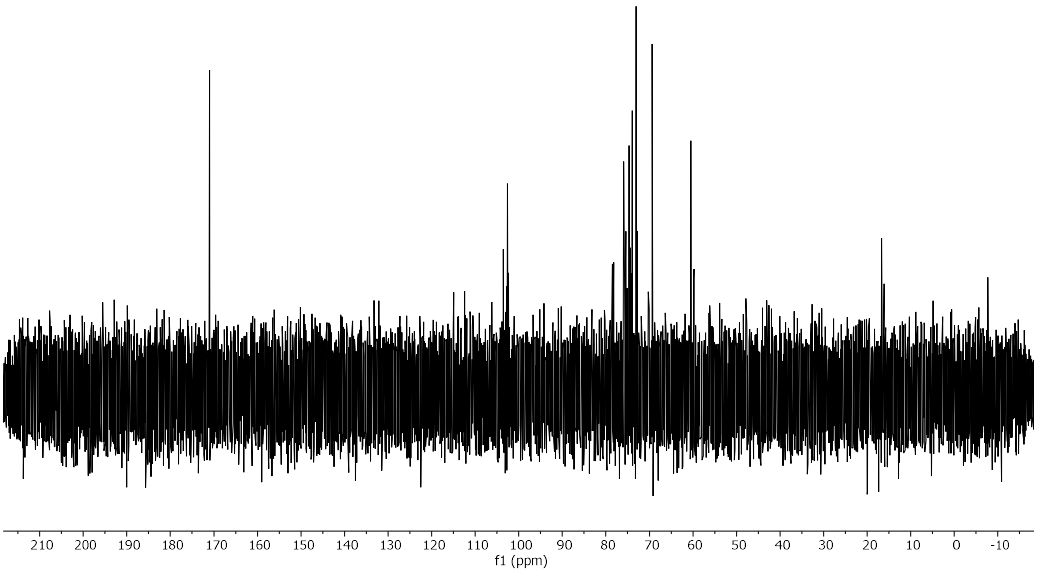
RP-HPLC of 9mer-I (ELSD trace, Method B1, t_R = 24.0, 24.3 min)



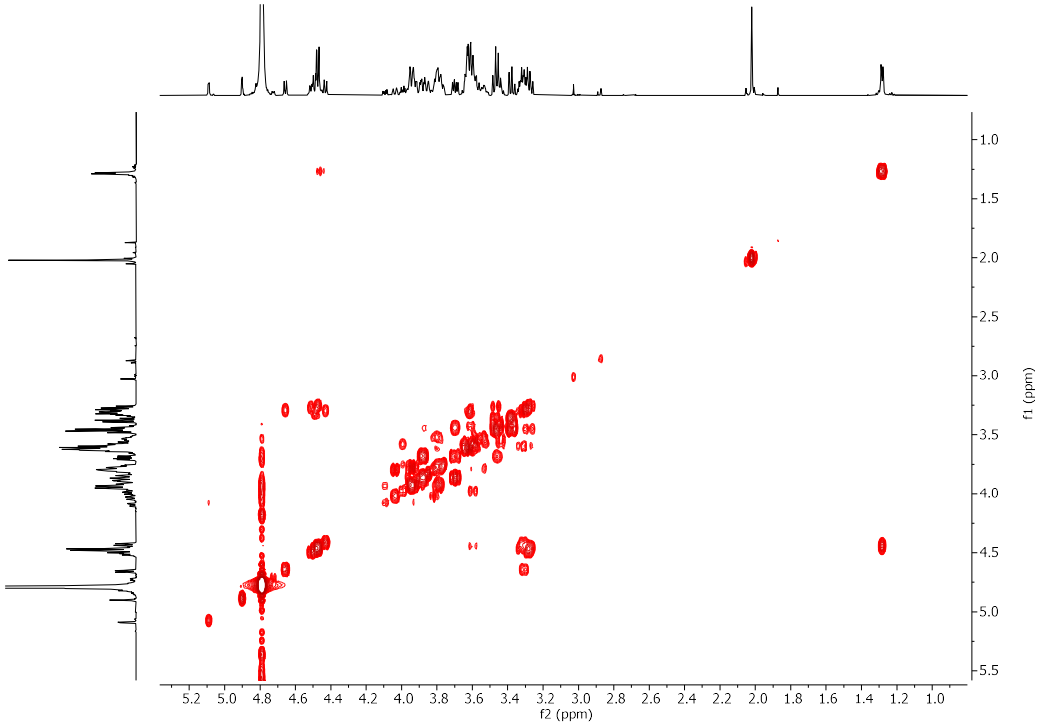
^{1}H NMR of 9mer-I (600 MHz, D_2O)



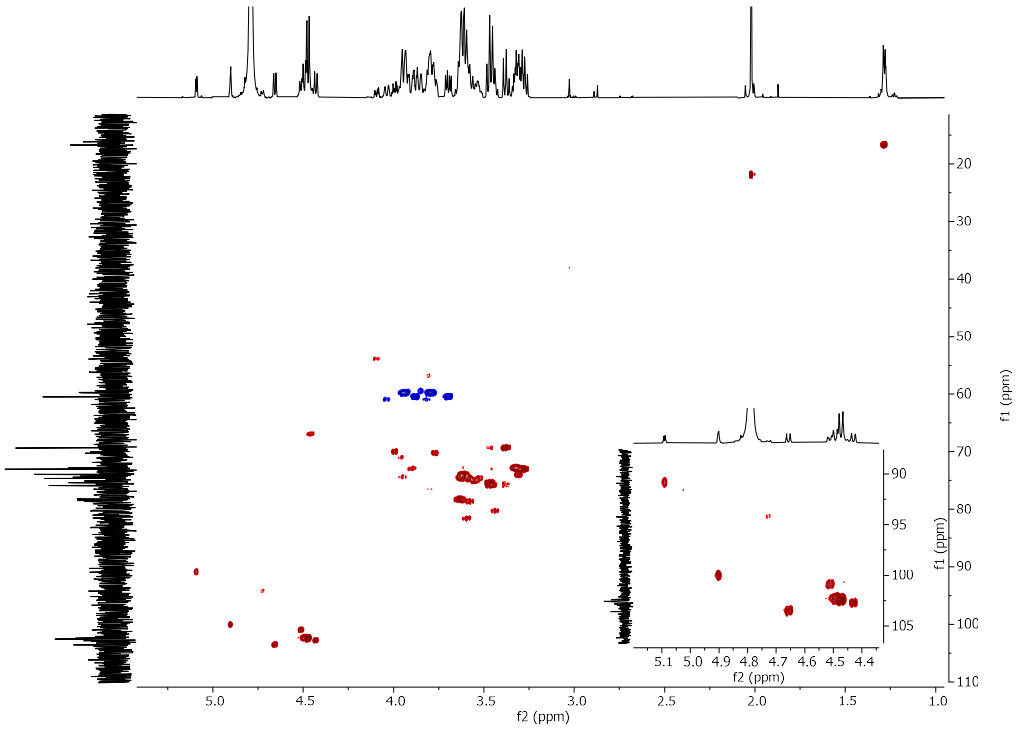
^{13}C NMR of 9mer-I (151 MHz, D_2O)



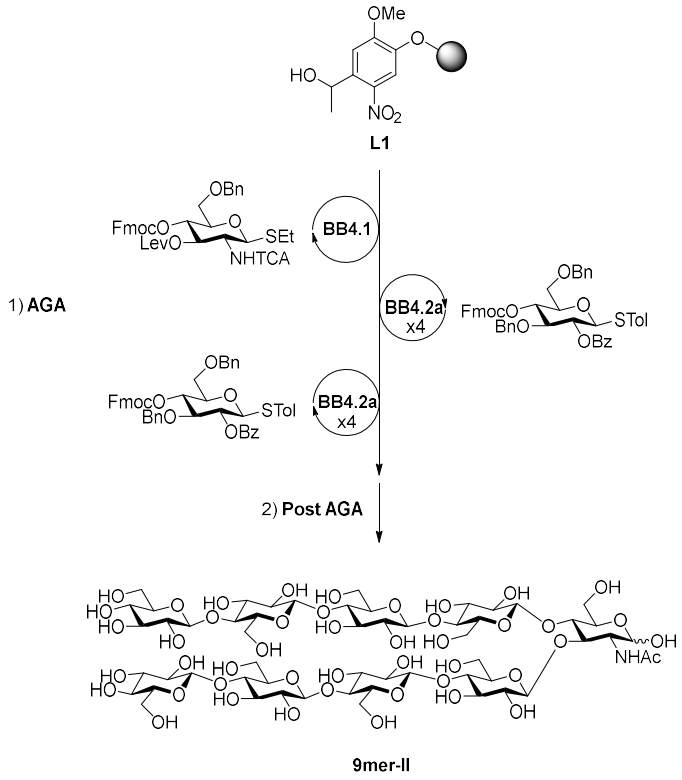
COSY NMR of 9mer-I (D_2O)



HSQC NMR of 9mer-I (D_2O)



Synthesis of **9mer-II**

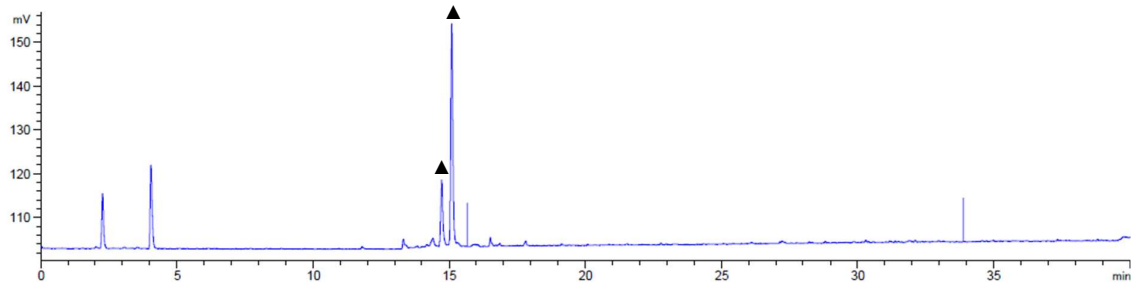


Step	BB	Modules	Notes
AGA	-	**A**	**L2** swelling
	BB4.1	**B**, **C1***, **D**, **E1**	**C1***: (**BB4.1**, -20 °C for 5 min, 0 °C for 40 min) *Double cycle
	BB4.2b	**B**, **C2***, **D**, **E1**	**C2***: (**BB4.2b**, -20 °C for 5 min, 0 °C for 20 min) *Double cycle
	(**BB4.2b**)x3	(**B**, **C2**, **D**, **E1**)x3	**C2**: (**BB4.2b**, -30 °C for 5 min, -10 °C for 40 min)
	-	**D**, **E2**	-
	(**BB4.2b**)x4	(**B**, **C2**, **D**, **E1**)x4	**C2**: (**BB4.2b**, -20 °C for 5 min, 0 °C for 20 min)
Post-AGA	-	**F1**, **G1**, **H2**, **I**	**F1**: (18 h) **H2**: (8 h) **I**: (Method C2: 15.7, 16.1 min)

Automated synthesis, global deprotection, and purification afforded **9mer-II** as a white solid (3.7 mg, 18% overall yield).

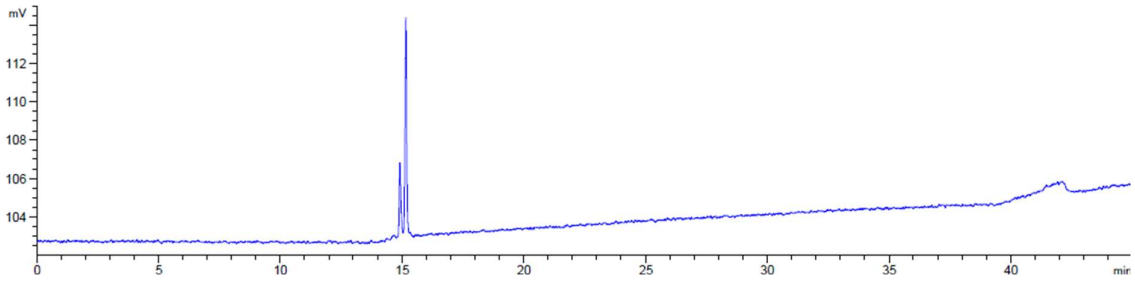
^{1}H NMR (600 MHz, D_2O) δ 5.13 (d, *J* = 3.4 Hz, 0.6H, H-1α GlcNAc), 4.69 – 4.61 (m, 2.4H, H-1β GlcNAc, 2xH-1 Glc), 4.52 – 4.44 (m, 6H, 6xH-1 Glc), 4.26 – 4.18 (m, 0.6H), 4.10 – 4.02 (m, 1H), 4.01 – 3.72 (m, 17.4H), 3.69 (dd, *J* = 12.4, 5.8 Hz, 2H), 3.66 – 3.50 (m, 16H), 3.50 – 3.23 (m, 14H), 2.00 (d, *J* = 1.4 Hz, 3H, CH_3 Ac GlcNAc). ^{13}C NMR (101 MHz, D_2O) δ 174.19, 170.96, 102.46 (C-1 Glc), 102.28(C-1 Glc), 102.24(C-1 Glc), 100.93 (C-1 Glc), 100.66 (C-1 Glc), 94.58 (C-1β GlcNAc), 90.59 (C-1α GlcNAc), 78.20, 78.08, 75.87, 75.34, 74.71, 73.90, 73.02, 72.81, 69.32, 60.44, 59.74, 53.63, 21.93 (CH_3 Ac GlcNAc). ESI-HRMS m/z 1540.5099 $[M+Na]^+$ ($C_{56}H_{95}NO_{46}Na$ requires 1540.5017).

RP-HPLC of crude 9mer-II (ELSD trace, Method C1)

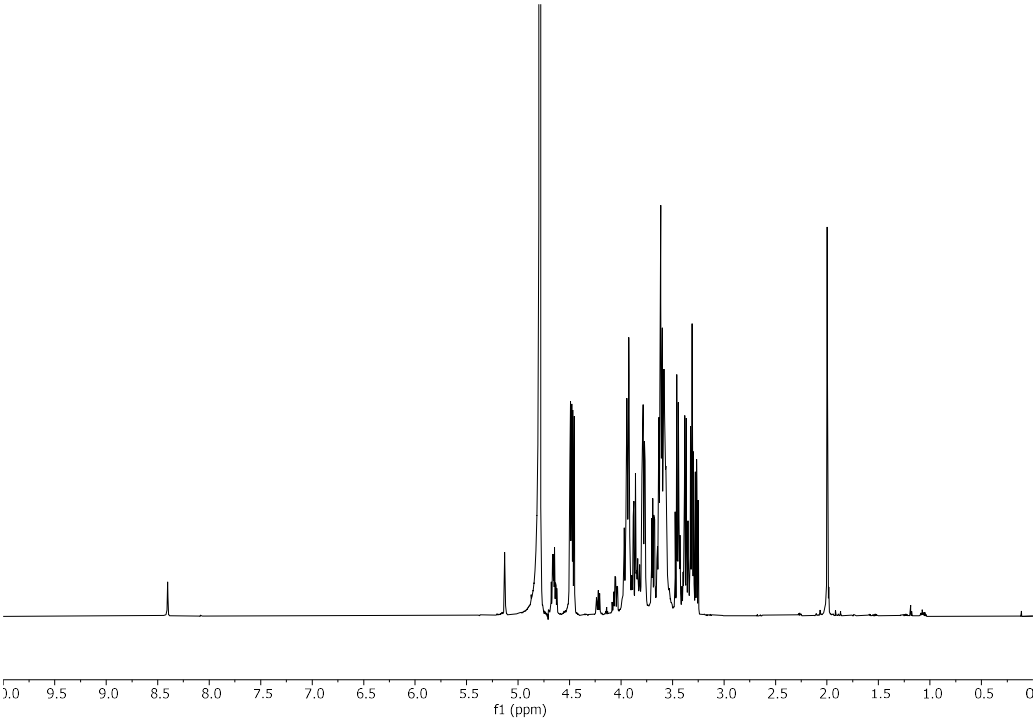


▲**9mer-II** (α and β anomers, m/z 1540 $[M+Na]^+$).

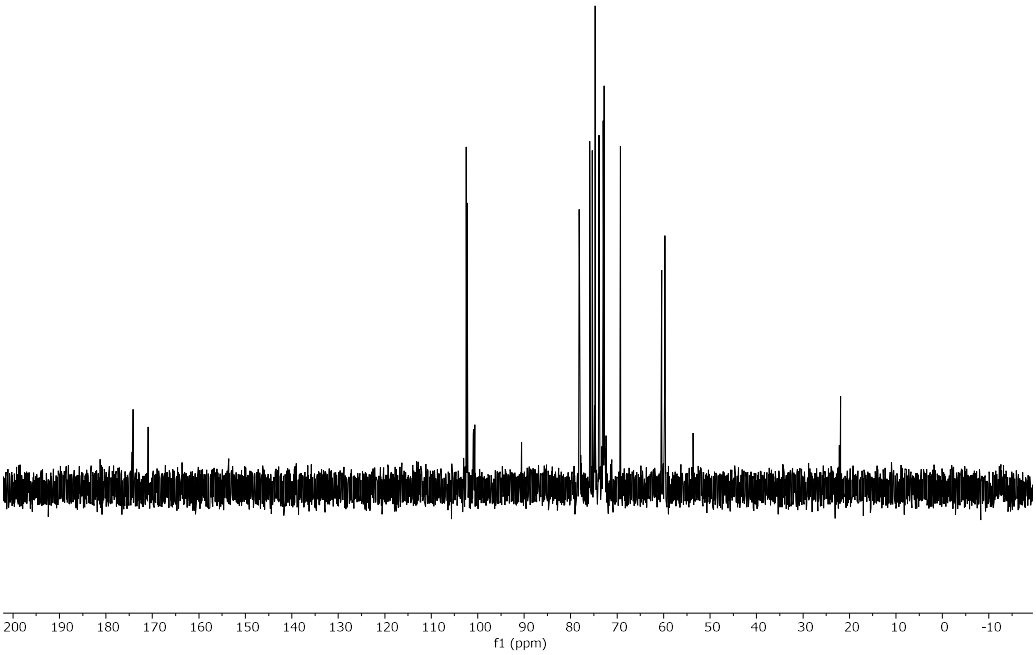
RP-HPLC of 9mer-II (ELSD trace, Method C1, t_R = 14.9, 15.1 min)



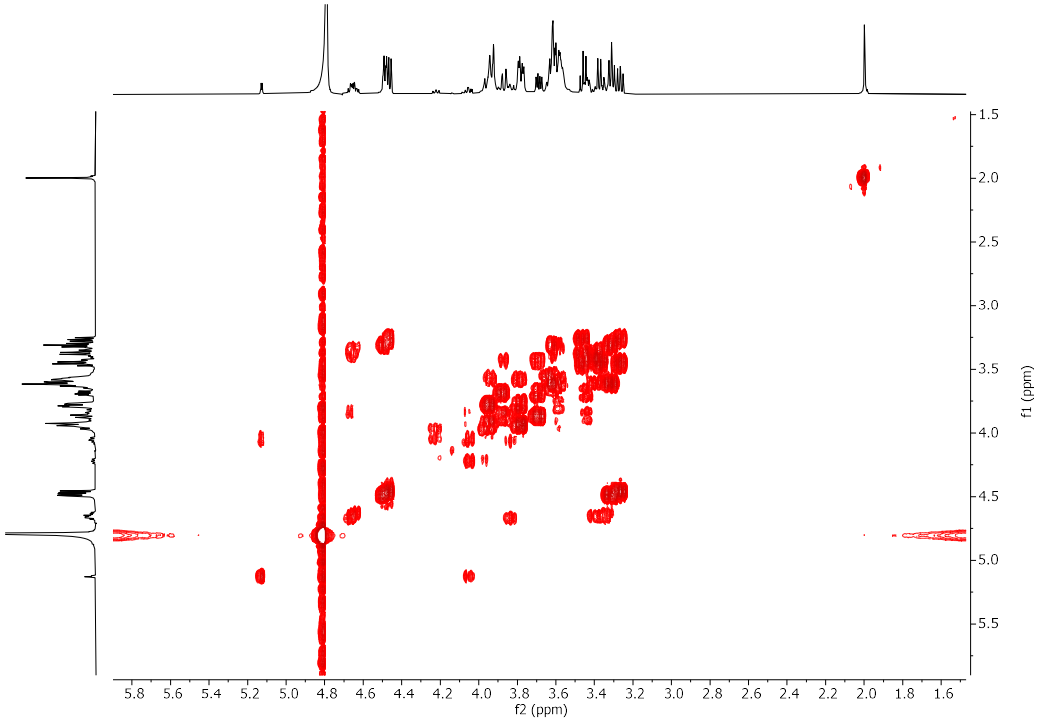
^{1}H NMR of 9mer-II (600 MHz, D_2O)



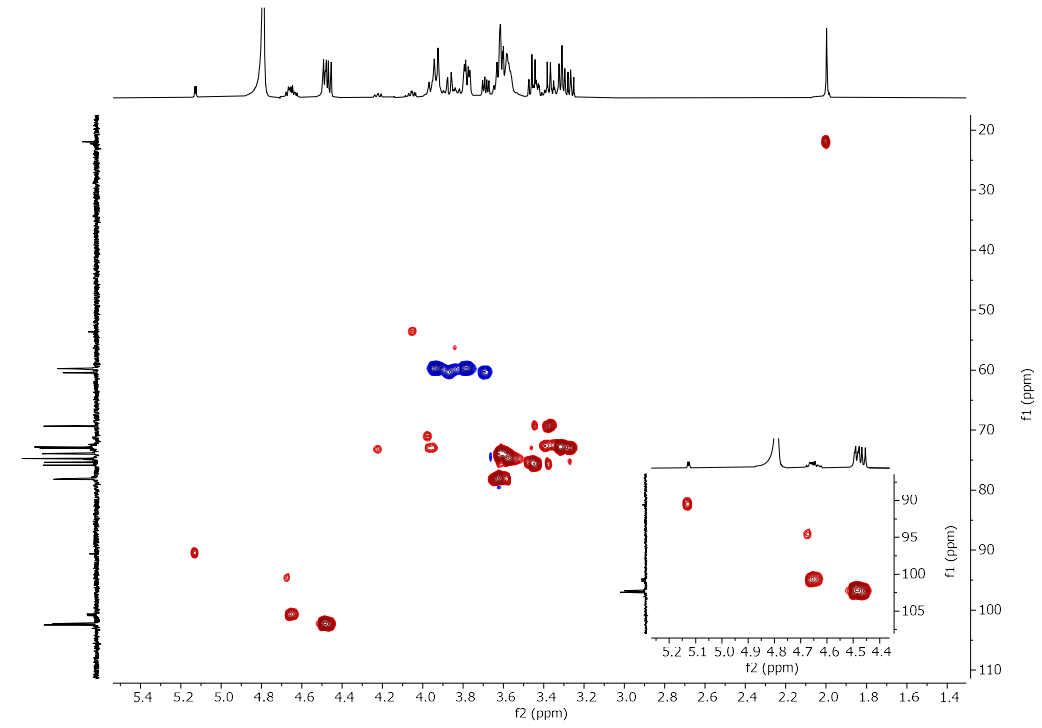
^{13}C NMR of 9mer-II (151 MHz, D_2O)



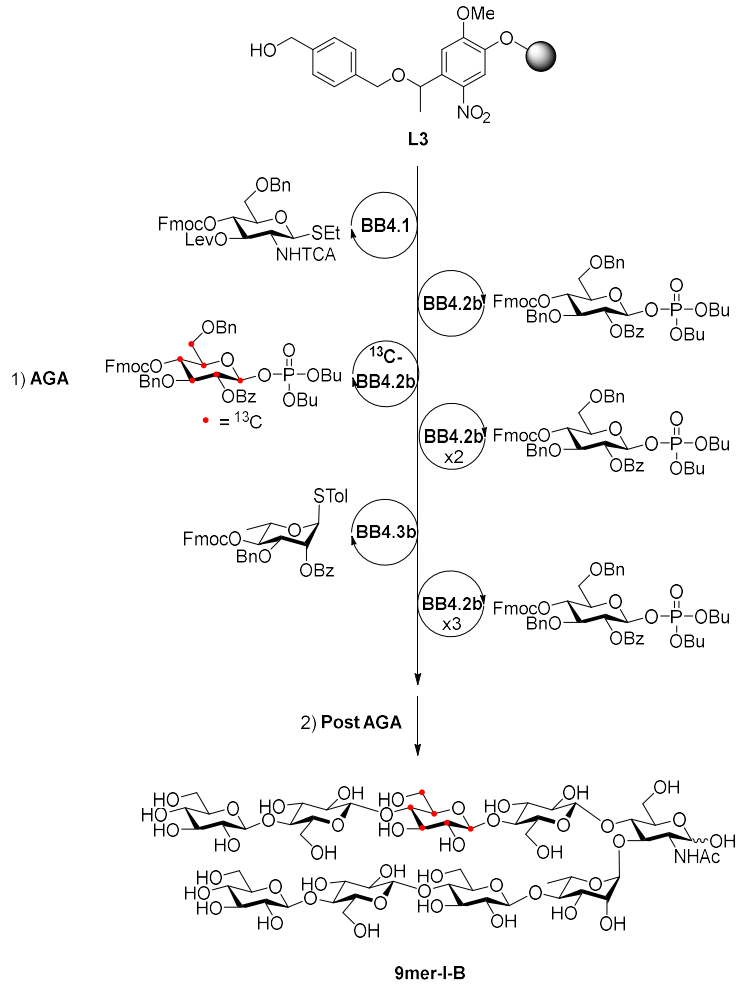
COSY NMR of 9mer-II (D_2O)



HSQC NMR of 9mer-II (D_2O)



Synthesis of **9mer-I-B**

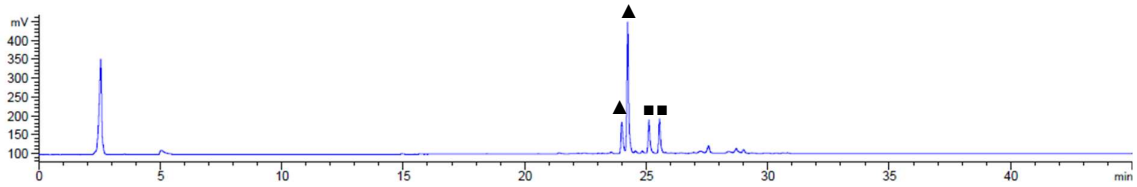


Step	BB	Modules	Notes
AGA	-	**A**	**L2** swelling
	BB1	**B**, **C1***, **D**, **E1**	**C1***: (**BB1**, -20 °C for 5 min, 0 °C for 40 min) *Double cycle
	BB2b	**B**, **C2***, **D**, **E1**	**C2***: (**BB2b**, -30 °C for 5 min, -10 °C for 40 min) *Double cycle
	^{13}C-BB2b	**B**, **C2**, **D**, **E1**	**C2**: (**^{13}C-BB2b**, -30 °C for 5 min, -10 °C for 40 min)
	(**BB2b**)x2	(**B**, **C2**, **D**, **E1**)x2	**C2**: (**BB2b**, -30 °C for 5 min, -10 °C for 40 min)
	-	**D**, **E2**	-
	BB3b	**B**, **C1**, **D**, **E1**	**C1**: (**BB3b**, -20 °C for 5 min, 0 °C for 20 min)
	(**BB2b**)x3	(**B**, **C2**, **D**, **E1**)x3	**C2**: (**BB3b**, -30 °C for 5 min, -10 °C for 40 min)
Post-AGA	-	**F1**, **G1**, **H2**, **I**	**F1**: (24 h) **H2**: (10 h) **I**: (Method B2: 26.9 min)

Automated synthesis, global deprotection, and purification afforded **9mer-I-B** as a white solid (2.0 mg, 10% overall yield).

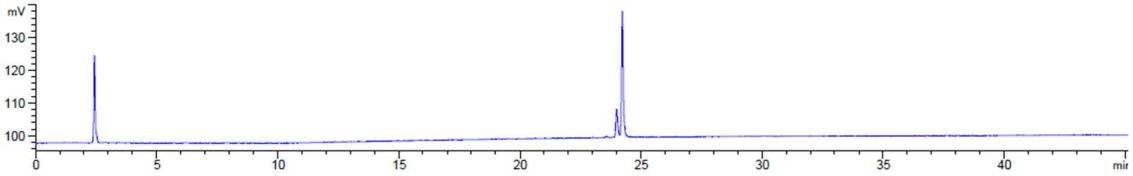
^{1}H NMR (700 MHz, D_2O) δ 5.08 (d, *J* = 3.5 Hz, 0.6H, H-1α GlcNAc), 4.89 (s, 1H, H-1 Rha), 4.72 (d, *J* = 7.5 Hz, 0.4H, H-1β GlcNAc), 4.65 (d, *J* = 7.9 Hz, 1H, H-1 Glc), 4.53 (d, *J* = 8.0 Hz, 0.5H, H-1 ^{13}C-Glc), 4.52 – 4.43 (m, 6H, 5x H-1 Glc, H-5 Rha), 4.30 (d, *J* = 7.2 Hz, 0.5H, H-1 ^{13}C-Glc), 4.12 – 4.07 (m, 0.6H), 4.07 – 4.00 (m, 2.4H), 4.00 – 3.40 (m, 36.5H), 3.37 (t, *J* = 9.5 Hz, 3H), 3.34 – 3.24 (m, 8H), 3.20 (s, 0.5H), 2.01 (s, 3H, CH_3 Ac GlcNAc), 1.28 (d, *J* = 6.9 Hz, 3H, CH_3-6 Rha). ^{13}C NMR (176 MHz, D_2O) δ 103.03, 102.76, 102.47, 102.21, 78.57, 78.48, 78.34, 78.25, 78.12, 78.03, 75.98, 75.03, 74.96, 74.80, 74.72, 74.48, 74.28, 74.06, 73.83, 73.28, 73.15, 73.01, 72.93, 72.79, 72.68, 69.45, 60.05, 59.98, 59.81, 59.72. ESI-HRMS m/z 1530.532 $[M+Na]^+$ ($C_{50}{}^{13}C_6H_{95}NO_{45}Na$ requires 1530.5269).

RP-HPLC of crude 9mer-I-B (ELSD trace, Method B1)

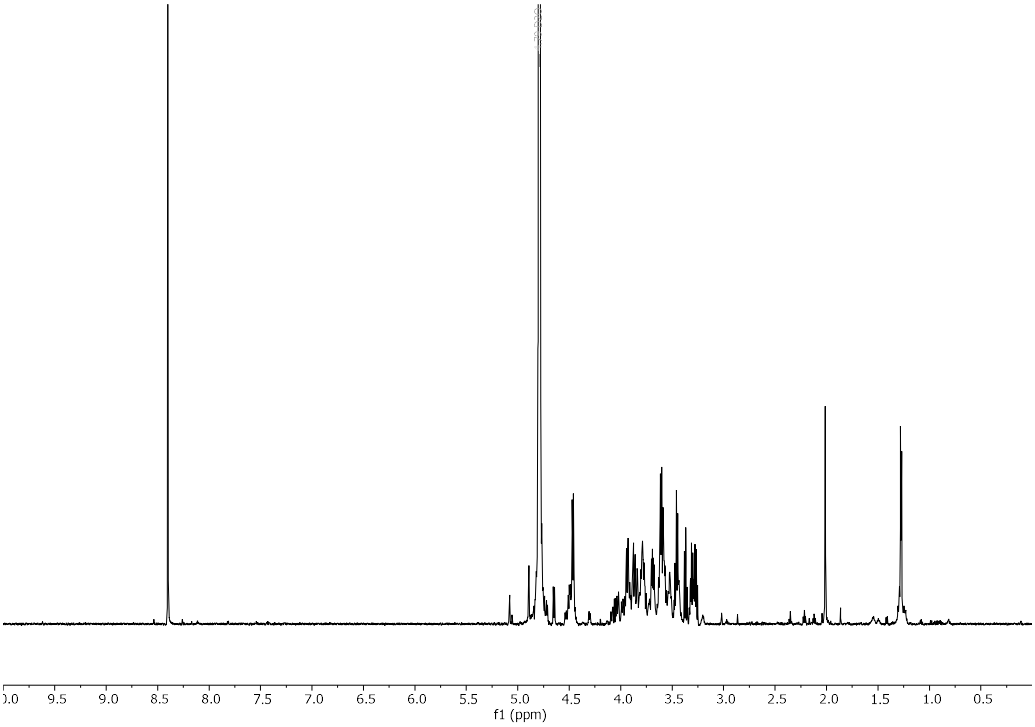


▲**9mer-I-B** (α and β anomers, m/z 1530 $[M+Na]^+$). ■ Incomplete methanolysis side-products.

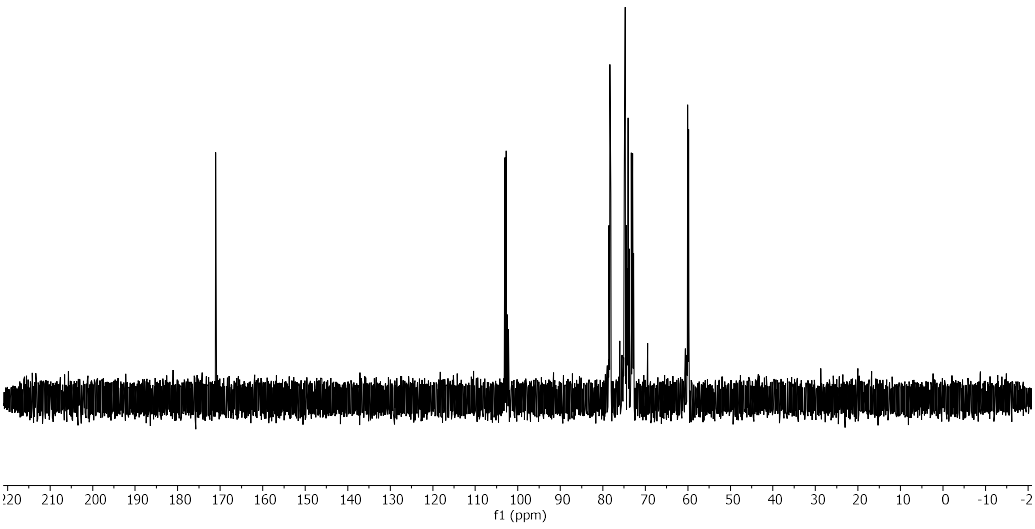
RP-HPLC of 9mer-I-B (ELSD trace, Method B1, t_R = 24.0, 24.3 min)



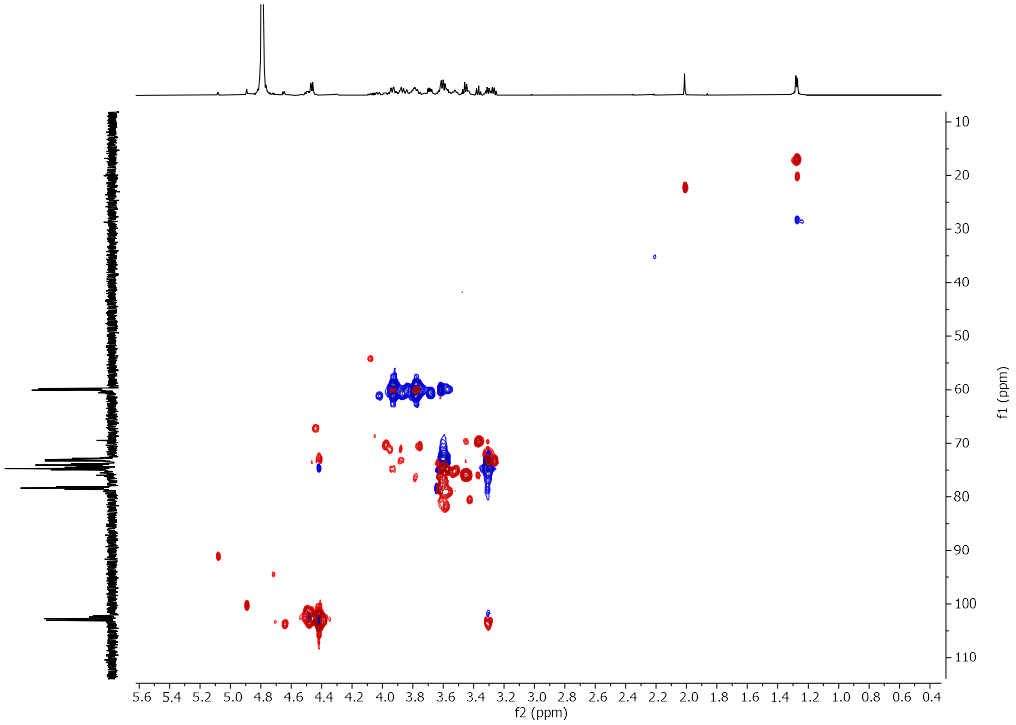
^{1}H NMR of 9mer-I-B (700 MHz, D_2O)



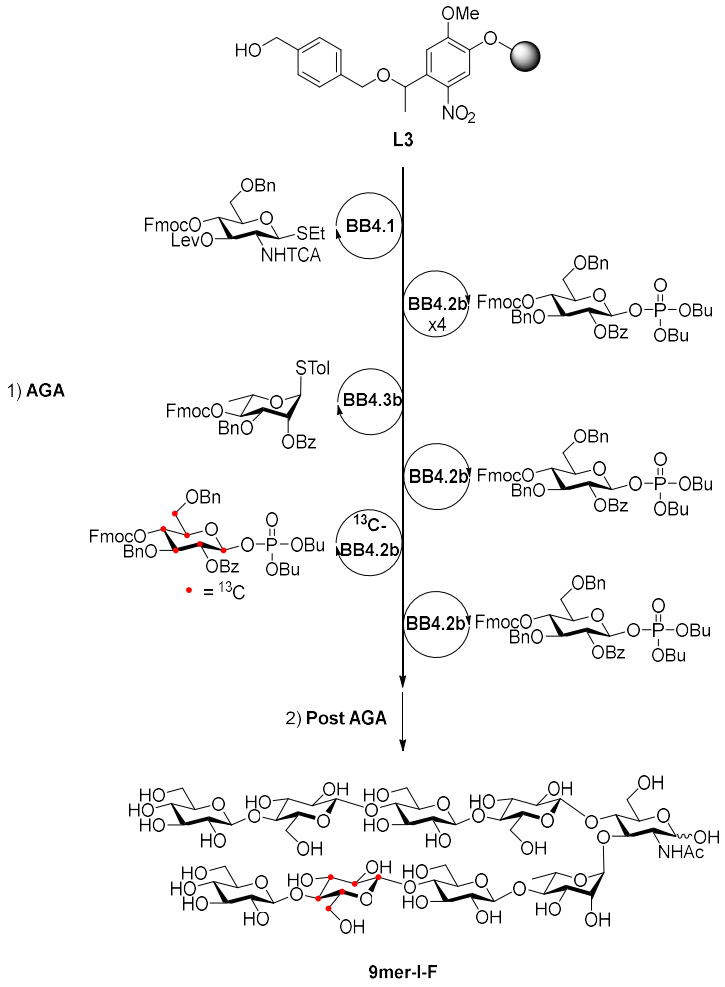
^{13}C NMR of 9mer-I-B (176 MHz, D_2O)



HSQC NMR of 9mer-I-B (D_2O)



Synthesis of **9mer-I-F**

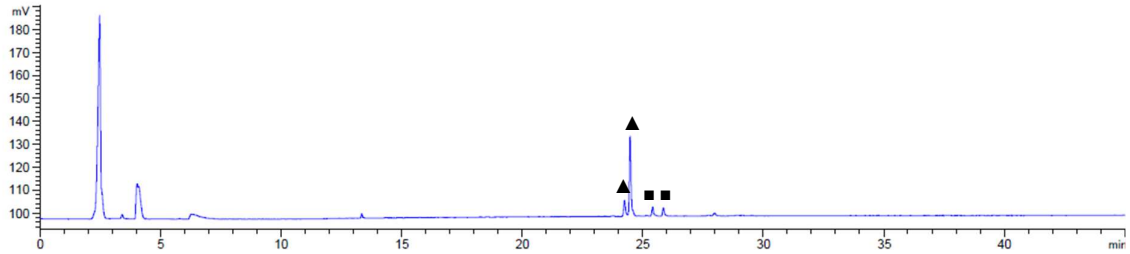


Step	BB	Modules	Notes
AGA	-	**A**	**L2** swelling
	BB4.1	**B**, **C1***, **D**, **E1**	**C1***: (**BB4.1**, -20 °C for 5 min, 0 °C for 40 min) *Double cycle
	BB4.2b	**B**, **C2***, **D**, **E1**	**C2***: (**BB4.2b**, -30 °C for 5 min, -10 °C for 40 min) *Double cycle
	(**BB4.2b**)x3	(**B**, **C2**, **D**, **E1**)x3	**C2**: (**BB4.2b**, -30 °C for 5 min, -10 °C for 40 min)
	-	**D**, **E2**	-
	BB4.3b	**B**, **C1**, **D**, **E1**	**C1**: (**BB4.3b**, -20 °C for 5 min, 0 °C for 20 min)
	BB4.2b	**B**, **C1**, **D**, **E1**	**C2**: (**BB4.2b**, -30 °C for 5 min, -10 °C for 40 min)
	^{13}C-BB4.2b	**B**, **C1**, **D**, **E1**	**C2**: (**^{13}C-BB4.2b**, -30 °C for 5 min, -10 °C for 40 min)
	BB4.2b	**B**, **C1**, **D**, **E1**	**C2**: (**BB4.2b**, -30 °C for 5 min, -10 °C for 40 min)
Post-AGA	-	**F1**, **G1**, **H2**, **I**	**F1**: (24 h) **H2**: (10 h) **I**: (Method B2: 26.9 min)

Automated synthesis, global deprotection, and purification afforded **9mer-I-F** as a white solid (1.9 mg, 10% overall yield).

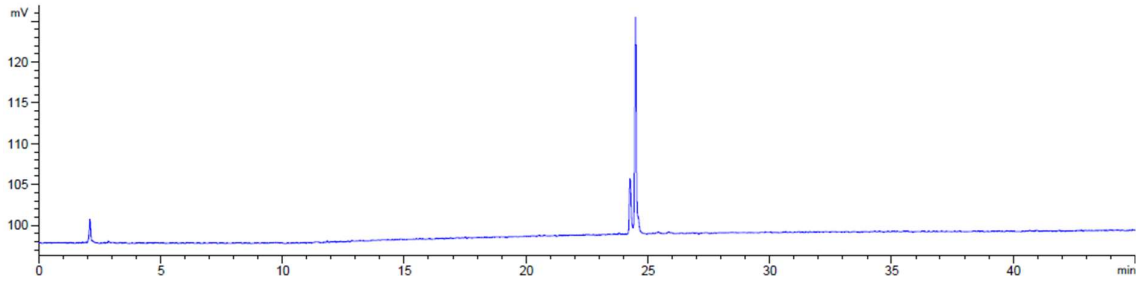
^{1}H NMR (600 MHz, D_2O) δ 5.08 (d, *J* = 3.5 Hz, 0.6H, H-1α GlcNAc), 4.89 (s, 1H, H-1 Rha), 4.71 (d, *J* = 9.2 Hz, 0.4H, H-1β GlcNAc), 4.65 (d, *J* = 8.0 Hz, 1H, H-1 Glc), 4.59 (d, *J* = 7.3 Hz, 0.5H, H-1 ^{13}C-Glc), 4.53 – 4.43 (m, 5H, 4x H-1 Glc, H-5 Rha), 4.42 (d, *J* = 7.7 Hz, 1H, H-1 Glc), 4.32 (d, *J* = 7.6 Hz, 0.5H, H-1 ^{13}C-Glc), 4.11 – 4.01 (m, 1.5H), 4.00 – 3.73 (m, 18H), 3.72 – 3.66 (m, 3H), 3.65 – 3.49 (m, 13H), 3.48 – 3.40 (m, 6H), 3.36 (t, *J* = 9.2 Hz, 2H), 3.33 – 3.24 (m, 7H), 2.01 (s, 3H, CH_3 Ac GlcNAc), 1.27 (d, *J* = 6.3 Hz, 3H, CH_3-6 Rha). ^{13}C NMR (151 MHz, D_2O) δ 102.51, 102.20, 78.64, 78.38, 78.23, 78.12, 74.90, 74.69, 74.61, 74.33, 74.20, 73.93, 73.68, 73.04, 72.74, 72.46, 59.98, 59.71. ESI-HRMS m/z 1530.531 $[M+Na]^+$ ($C_{50}{}^{13}C_6H_{95}NO_{45}Na$ requires 1530.52).

RP-HPLC of crude 9mer-I-F (ELSD trace, Method B1)

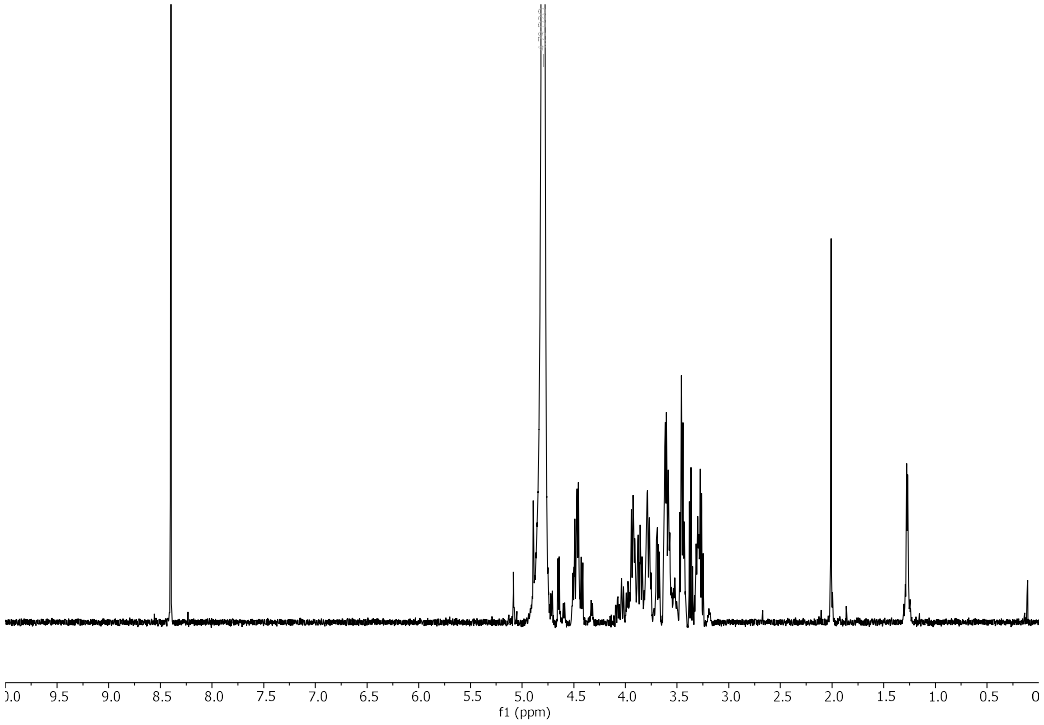


▲**9mer-I-F** (α and β anomers, m/z 1530 $[M+Na]^+$). ■ Incomplete methanolysis side-products.

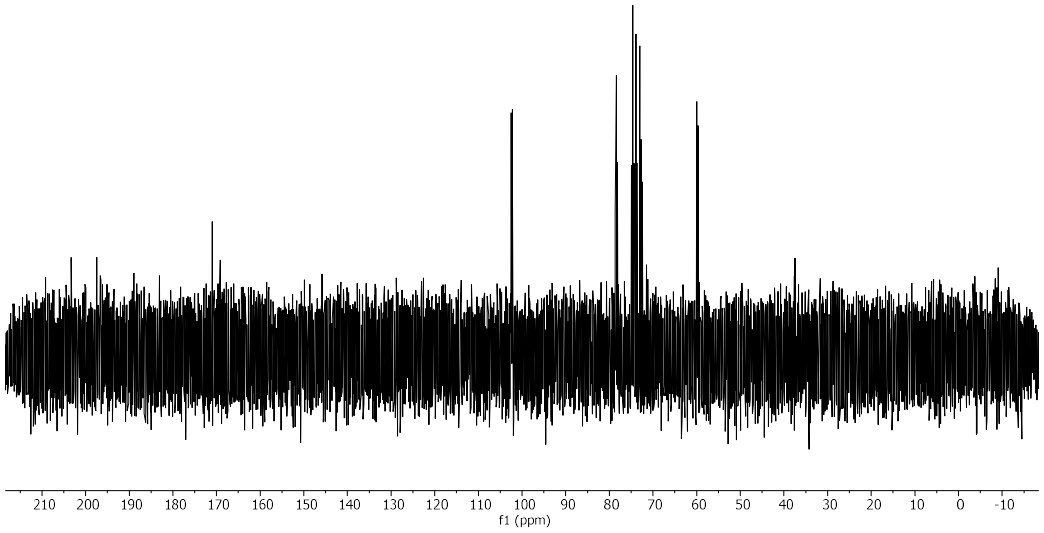
RP-HPLC of 9mer-I-F (ELSD trace, Method B1, t_R = 24.3, 24.5 min)



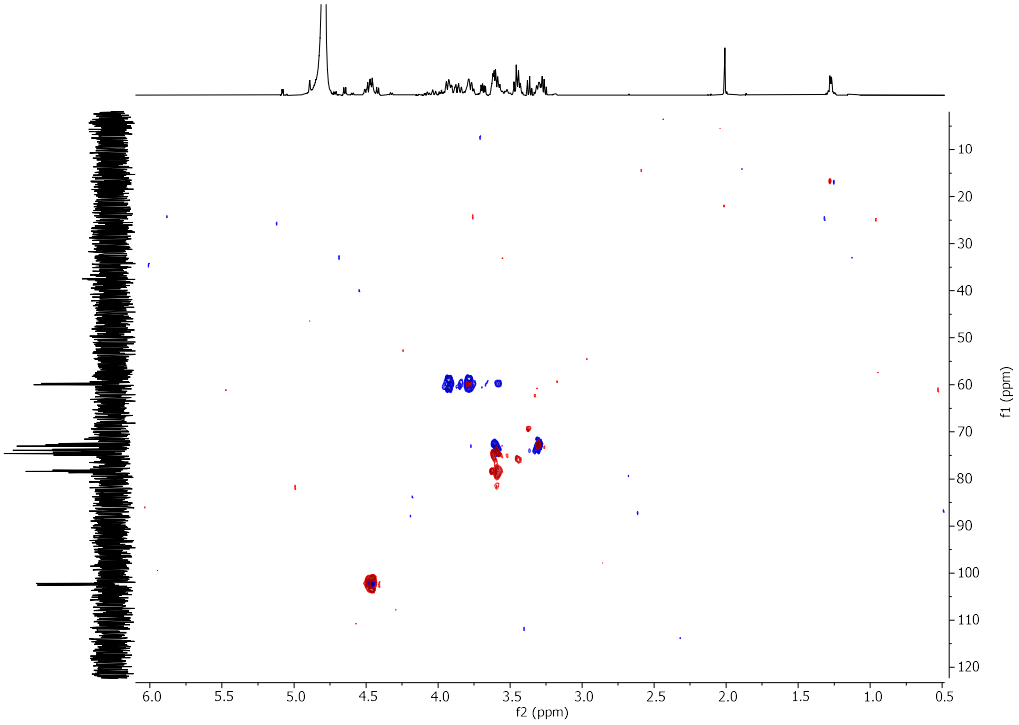
^{1}H NMR of 9mer-I-F (600 MHz, D_2O)



^{13}C NMR of 9mer-I-F (151 MHz, D_2O)



HSQC NMR of 9mer-I-F (D_2O)



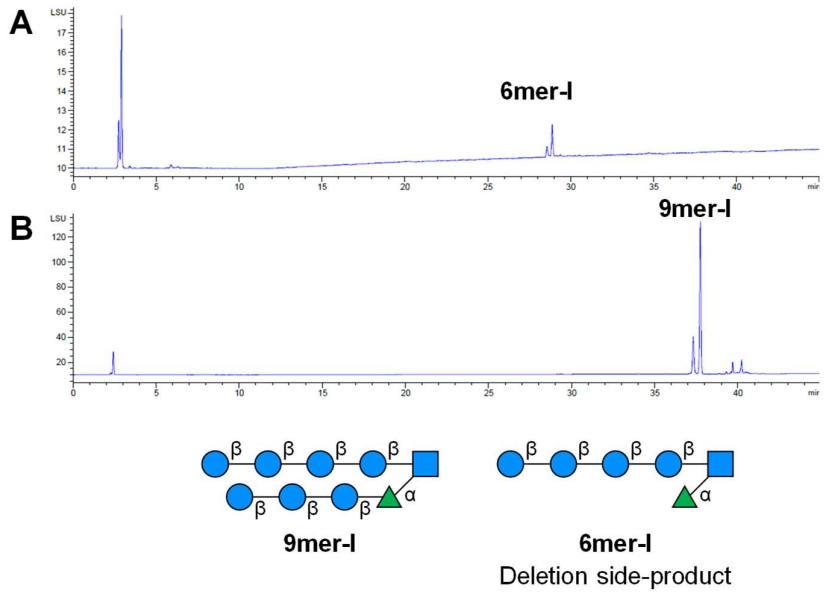


Figure S4.4 Comparison of crude RP HPLC traces (ELSD trace, Method A1) of **9mer-I** synthesized using **BB4.3a** (A) or **BB4.3b** (B).

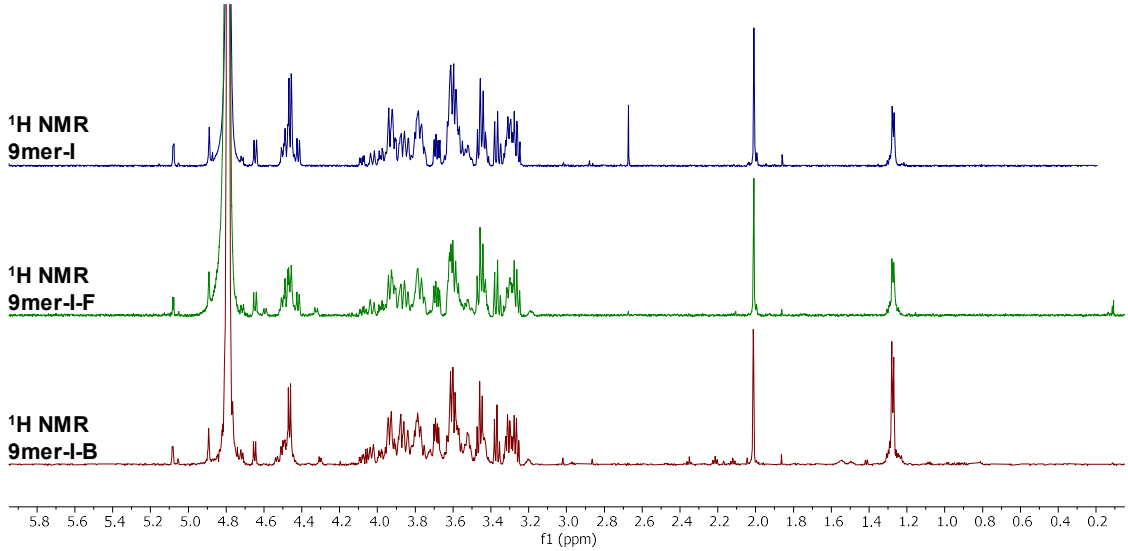


Figure S4.5 Comparison of ^{1}H NMR of ^{13}C-labelled compounds **9mer-I**, **9mer-I-B**, and **9mer-I-F**.

6.5.3 Molecular dynamics simulations

The MD simulations were performed by Dr. Theodore Tyrikos-Ergas.

General materials and methods All-atom molecular dynamics (MD) simulations were performed using gromacs 5.1.2.[362] The oligosaccharides were modeled using a modified version of GLYCAM06$_{OSMO,r14}$ force field,[205,206] and the system was solvated with TIP5P[320] water molecules to avoid excessive interactions between the monomers. The topology was converted to gromacs format using the glycam2gmx.pl script and solvated with 2100 water molecules using gromacs tools. The systems were kept at a constant temperature of 303 K using a Nosé-Hoover thermostat[363,364] and at constant pressure of 1 bar with the Parrinello-Rahman barostat.[365,366] Non-bonded interactions were cut-off at 1.4 nm, long range electrostatics were calculated using the particle mesh Ewald method.[367] Bonds involving hydrogens were constrained using the LINCS[368] to allow a 2 fs time step algorithm; water molecules were kept rigid with SETTLE.[369]

After energy minimization (steepest descent algorithm) and before the production run, the systems were equilibrated at 300 K for 50 ns in a canonical (NVT) ensemble (constant number of particles, volume and temperature) and subsequently at 300 K and 1 atm for 50 ns in an isothermal-isobaric (NPT) ensemble. All the modelled structures were simulated for 500 ns.

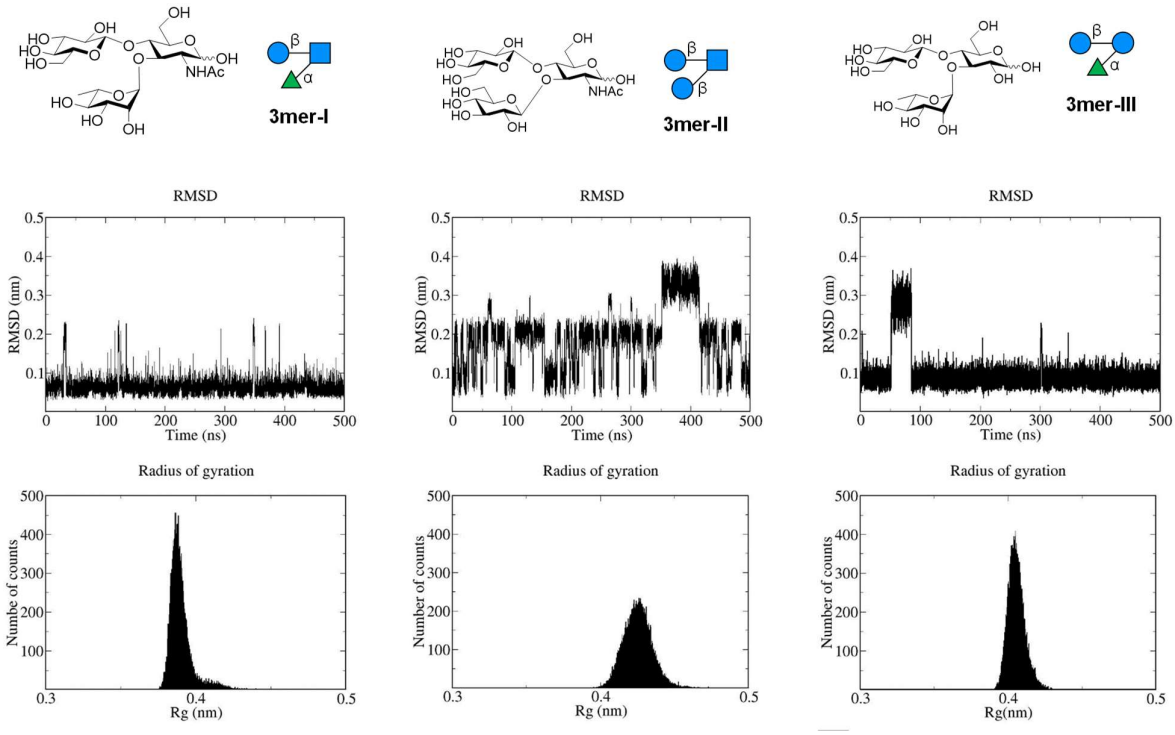


Figure S4.6 RMSD and R_g plots for **3mer-I**, **3mer-II**, and **3mer-III**.

Glycosidic torsional angles were defined as follows: $\Phi = O_5\text{-}C_1\text{-}O\text{-}C_n$ and $\Psi = C_1\text{-}O\text{-}C_n\text{-}C_{(n-1)}$.

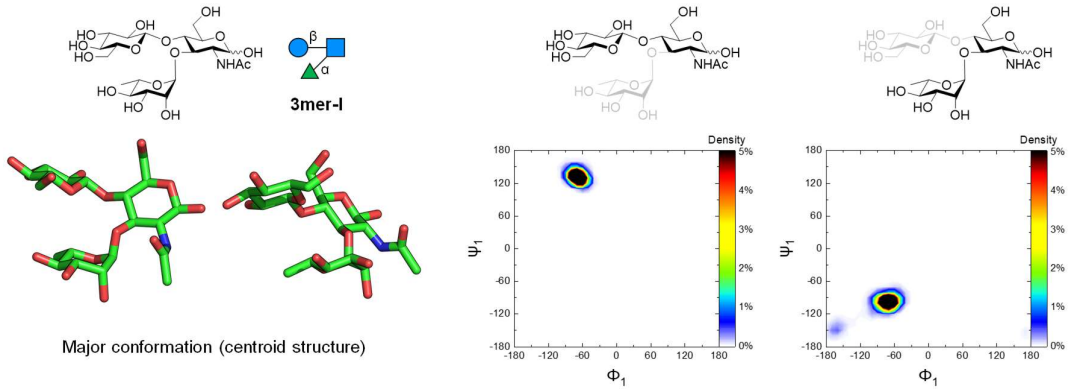


Figure S4.7 Ramachandran plots for the two glycosidic linkages of **3mer-I**.

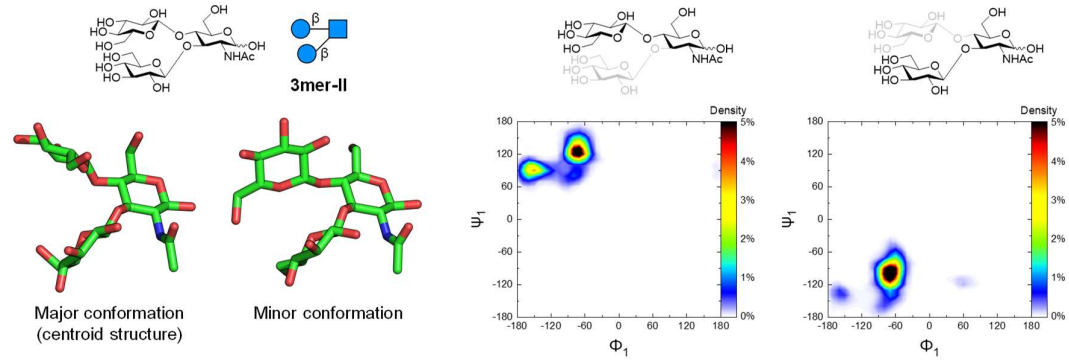


Figure S4.8 Ramachandran plots for the two glycosidic linkages of **3mer-II**.

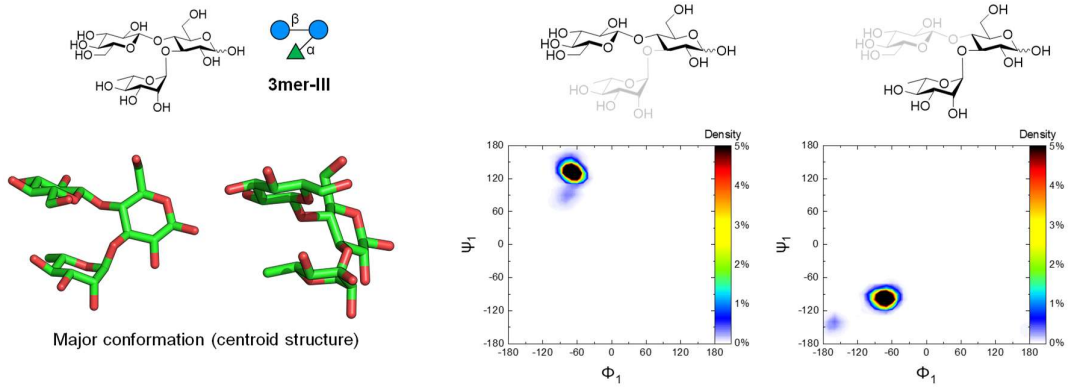


Figure S4.9 Ramachandran plots for the two glycosidic linkages of **3mer-III**.

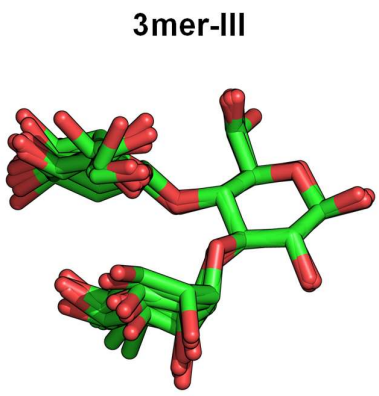


Figure S4.10 Overimposition of seven snapshot extracted from the MD simulation of **3mer-III**.

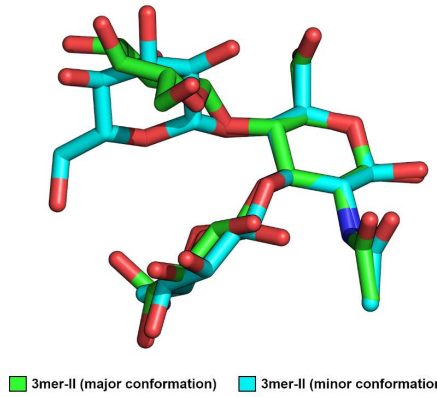


Figure S4.11 Two representative snapshot of the major and minor conformations of **3mer-II** overimposed.

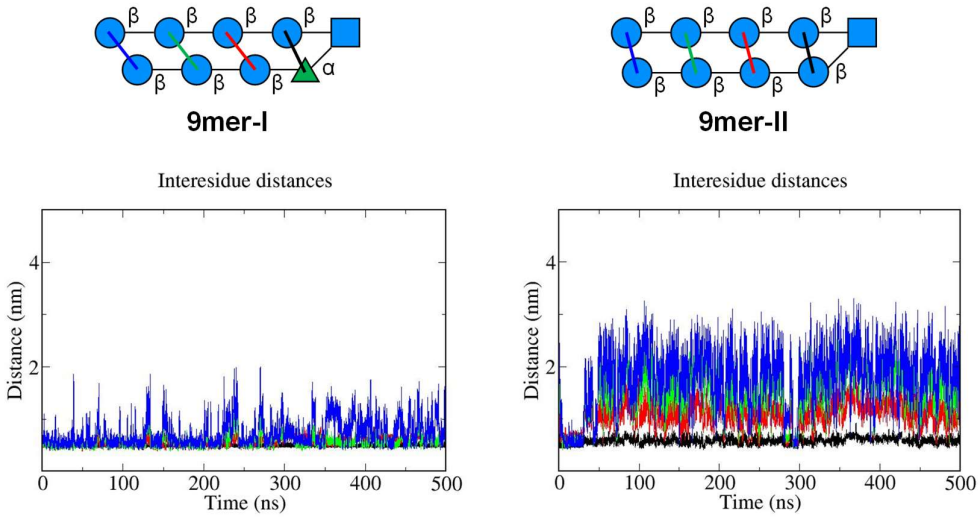


Figure S4.12 Inter-residue distance plots extracted from MD simulations for **9mer-I** and **9mer-II**.

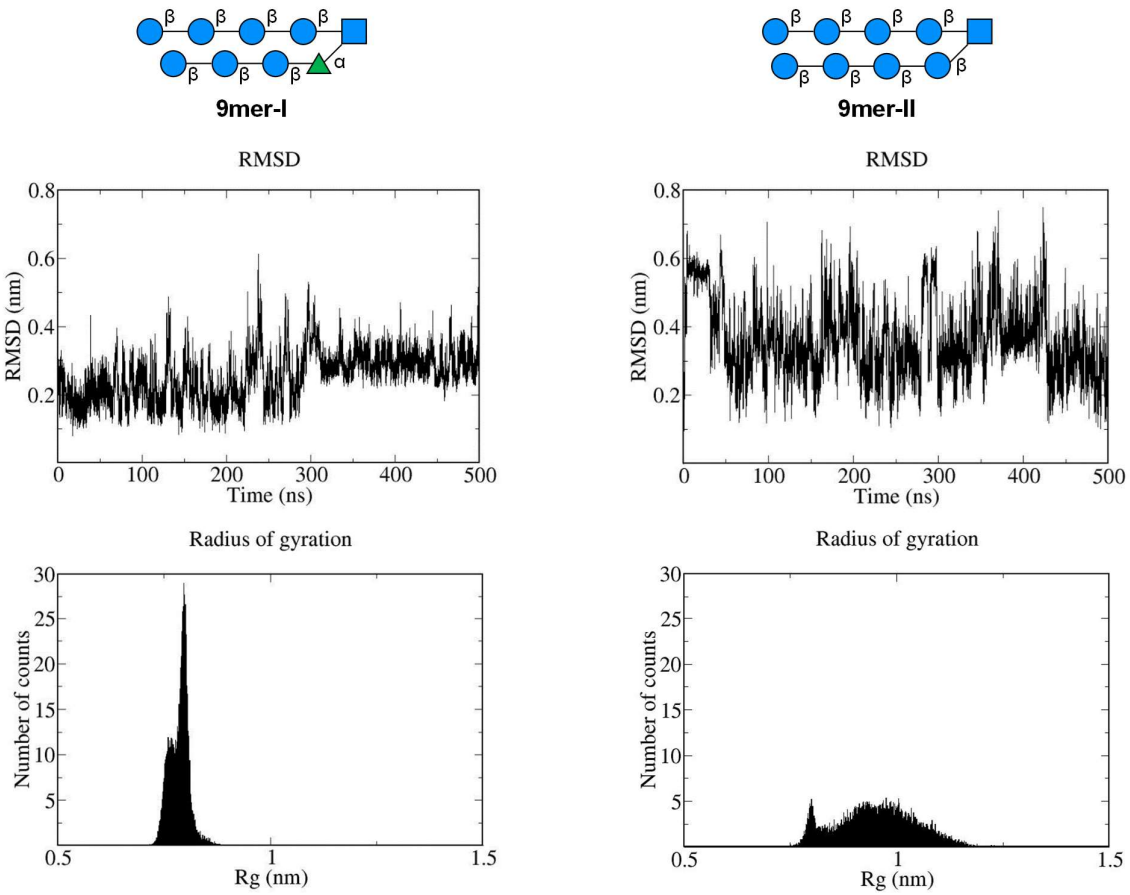


Figure S4.13 RMSD and R_g plots for **9mer-I** and **9mer-II**.

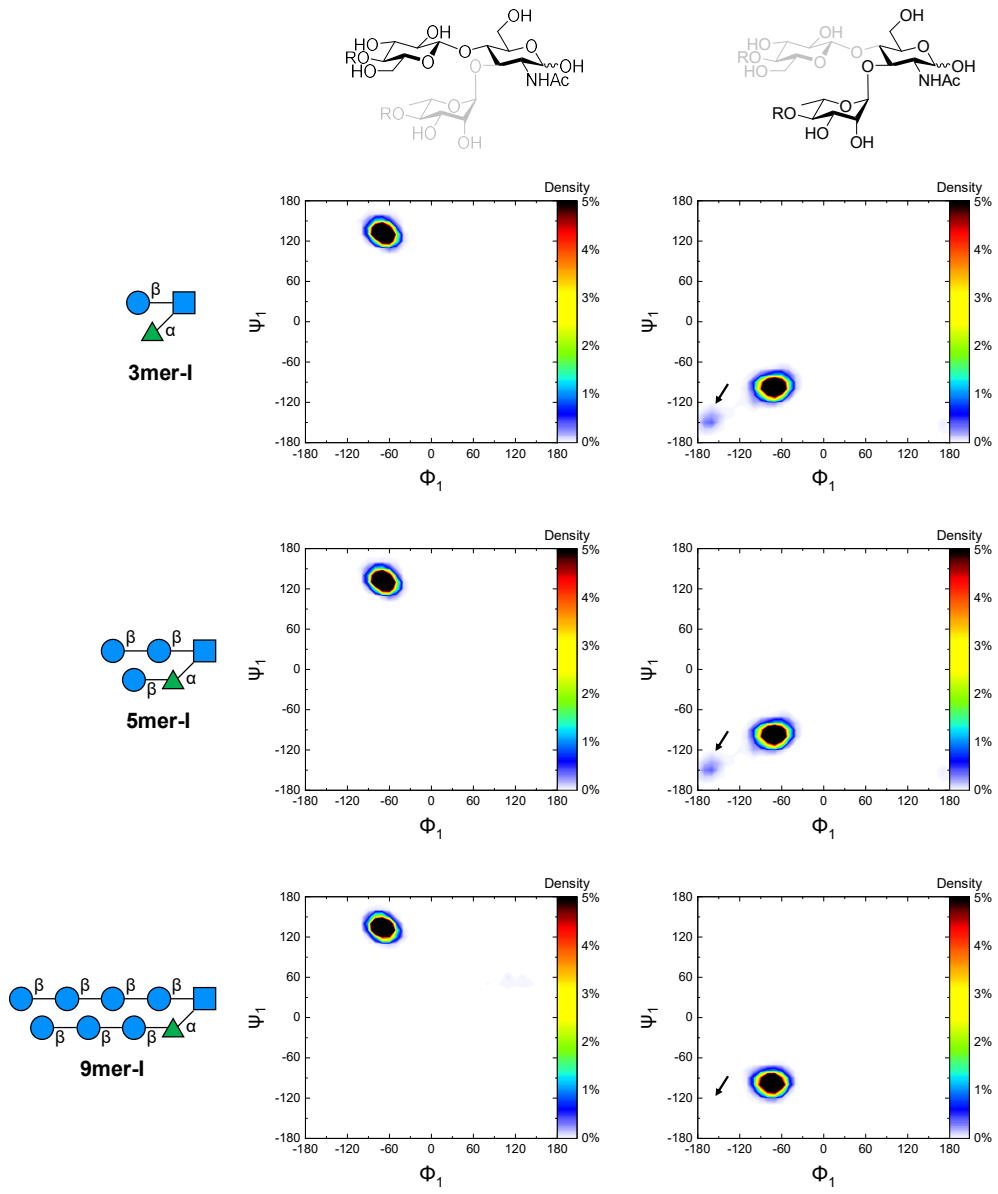


Figure S4.14 Ramachandran plots for the two glycosidic linkages of the turn unit for **3mer-I**, **5mer-I**, and **9mer-I**. Upon elongation of the hairpin strands, the population of minor conformers decreases (*black arrows*).

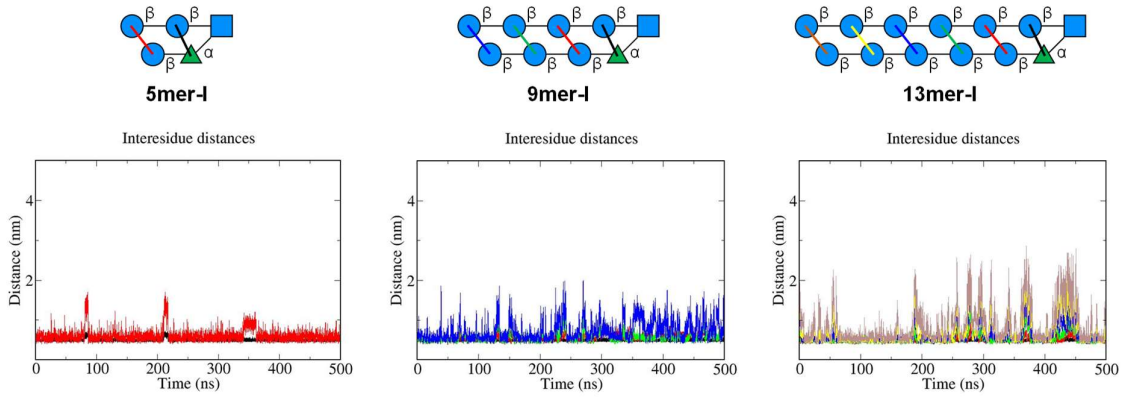


Figure S4.15 Inter-residue distance plots extracted from MD simulations for **5mer-I**, **9mer-I**, and **13mer-I**.

6.5.4 NMR studies

General materials and methods ^{1}H, ^{13}C, HSQC, 1D and 2D TOCSY, 1D and 2D ROESY, 2D NOESY NMR spectra were recorded on a Varian 400-MR (400 MHz), Varian 600-NMR (600 MHz), Bruker Biospin AVANCE700 (700 MHz) Bruker AVANCE III 800 (800 MHz) spectrometer. Samples were prepared by dissolving lyophilized samples in D_2O (concentration ≈ 1 - 6 mM). Proton resonances of the oligosaccharides were assigned using a combination of ^{1}H, 2D COSY, HSQC, 1D and 2D TOCSY. Selective 1D TOCSY (HOHAHA, pulse program: seldigpzs) spectra were recorded using different mixing times to assign all the resonances (d9 = 40, 80, 120, 160, and 200 ms). 2D TOCSY (pulse program: mlevphpp) spectra were recorded using different mixing times (d9 = 80, or 120 ms). Selective 1D t-ROESY (pulse program: selrogp.2) spectra were recorded using different mixing times (p15 = 100, 200, or 300 ms). 2D t-ROESY (pulse program: reosyph.2) and 2D NOESY (pulse program: noesygpphpp) spectra were recorded using different mixing times (p15 = 100, 200, or 300 ms for ROESY and d8 = 600, 800, or 1000 ms for NOESY). Monosaccharide were named as follows: D-glucose (Glc), D-*N*-acetyl glucosamine (GlcNAc), L-rhamnose (Rha). Labelling of protons in a monosaccharides is done as follows: e.g. proton attached to C-1 of Rha is named "Rha-1". Resonances of residues at the reducing end are additionally labelled with α or β.

Determination of the non-conventional hydrogen bond

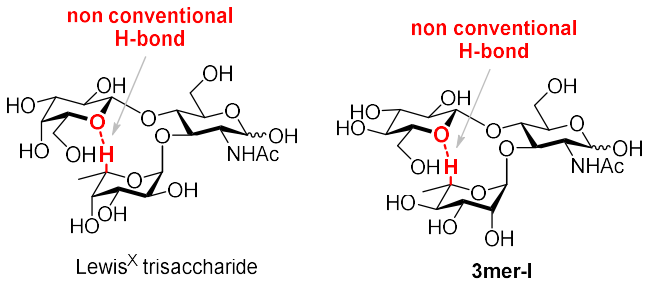


Figure S4.16 Comparison of a Lewis^{X} trisaccharide and **3mer-I** to highlight the non-conventional H-bond.[303,306]

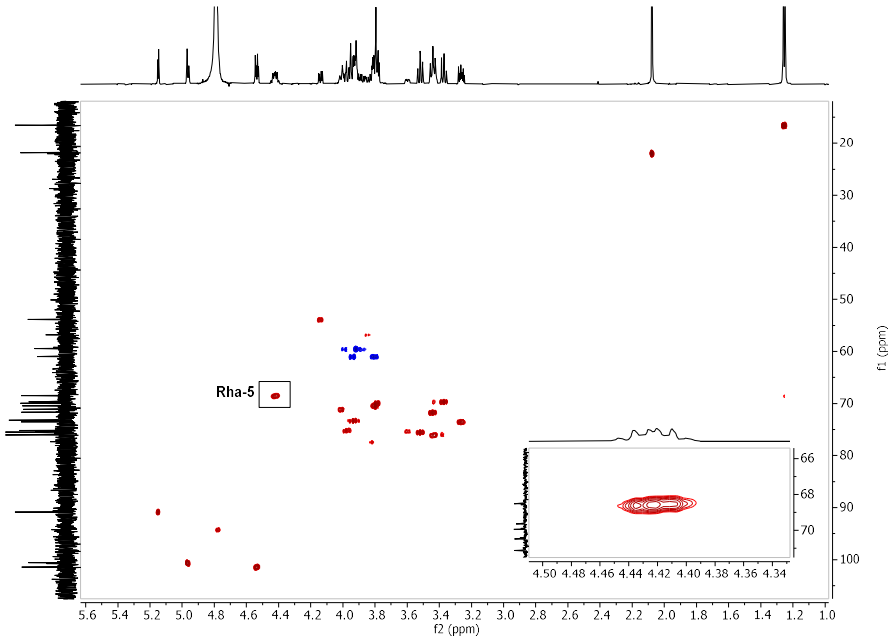


Figure S4.17 HSQC spectra of **3mer-I** with zoom-in on Rha-5 cross peak.

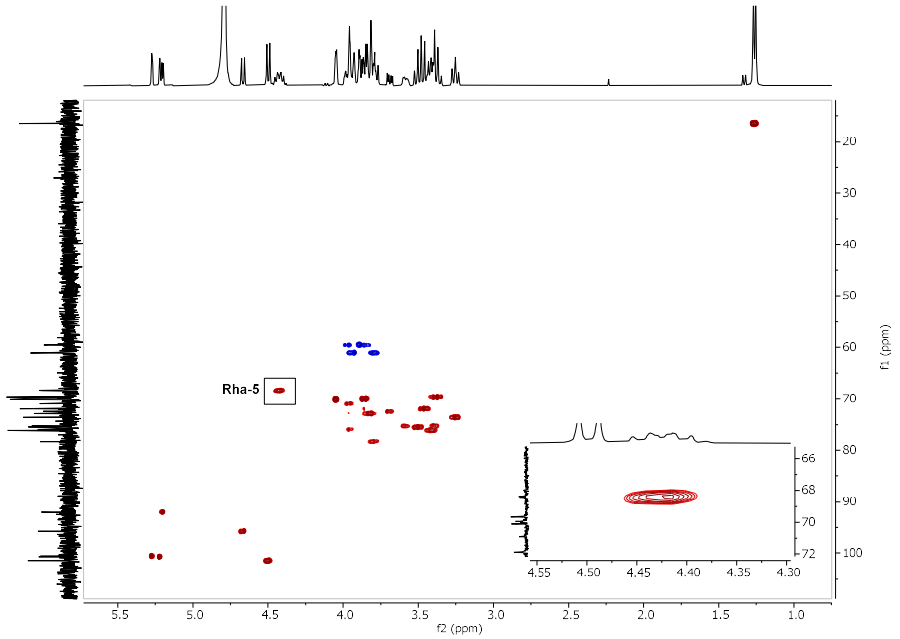


Figure S4.18 HSQC spectra of **3mer-III** with zoom-in on Rha-5 cross peak.

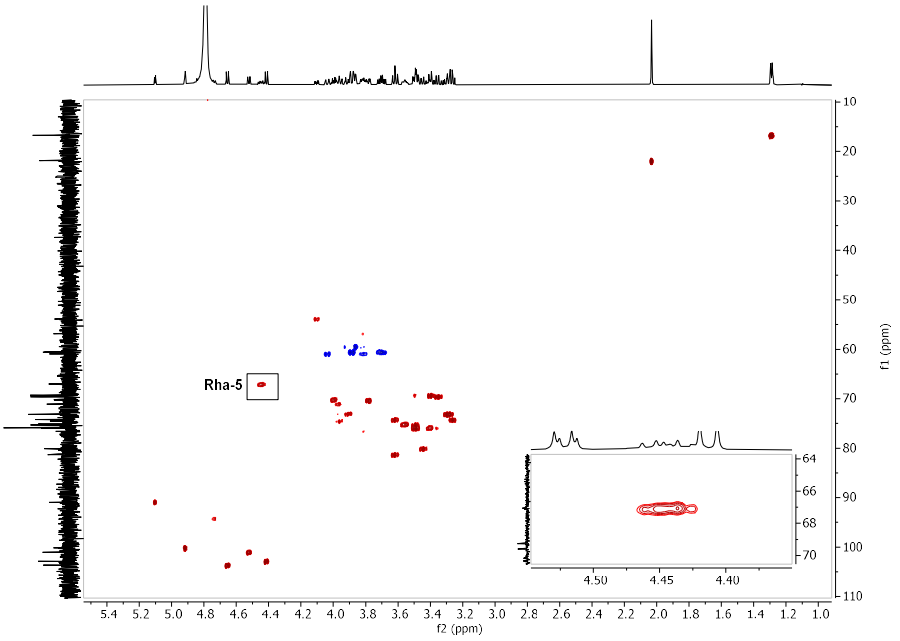


Figure S4.19 HSQC spectra of **5mer-I** with zoom-in on Rha-5 cross peak.

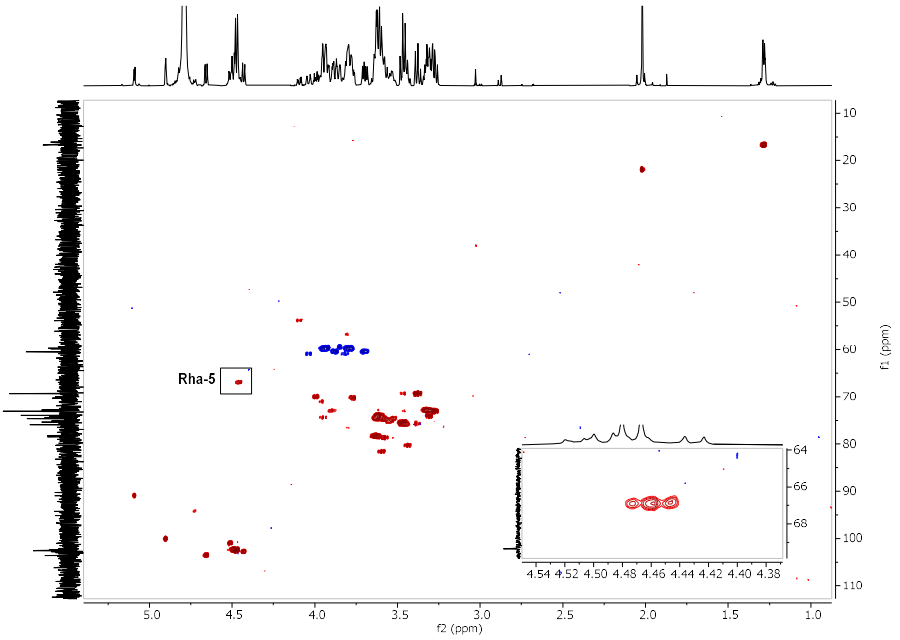


Figure S4.20 HSQC spectra of **9mer-I** with zoom-in on Rha-5 cross peak.

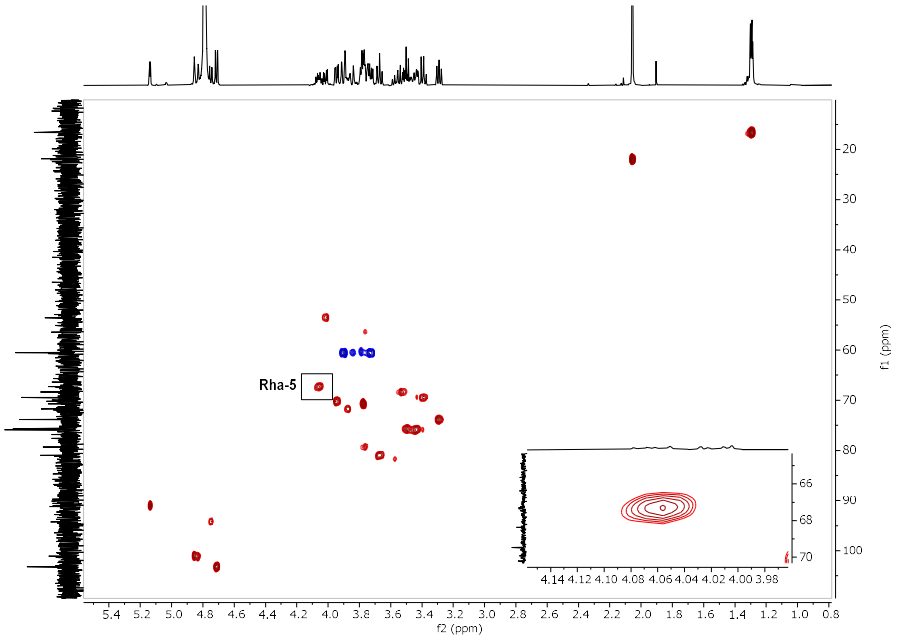


Figure S4.21 HSQC spectra of **3mer-IV** with zoom-in on Rha-5 cross peak.

Compound	δ Rha-5 (ppm)	Δδ (ppm)*
3mer-I	4.42	0.36
3mer-III	4.42	0.36
3mer-IV	4.06	0
5mer-I	4.44	0.38
9mer-I	4.46	0.40

Table S4.1 Comparison of the chemical shifts for Rha-5. *Δδ was calculated as follows Δδ = δ (Rha-5, **compound**) – δ (Rha-5, **3mer-IV**).

The chemical shift is an indicator for C-H···O hydrogen bond, as previously described for fucosylated motifs such as the Lewis^{x} trisaccharide where the H-5 of Fuc is shifted downfield (Δδ ≈ 0.45 ppm) (Figure S4.16).[304] An analogous downfield shift (Δδ ≈ 0.3-0.5 ppm) was ascribed to a non-conventional hydrogen bond in galactosaminogalactans analogues.[302] The Rha-5 ^{1}H NMR chemical shift was used to assess the presence of a non-conventional hydrogen bond (C-H···O) involving the H-5 of Rha (Rha-5) and the endocyclic O-5 of the Glc unit. A significant downfield shift (Δδ ≈ 0.3-0.4 ppm, Table S4.1) was observed for **3mer-I**, **3mer-III**, **5mer-I**, and **9mer-I** (Figures S4.17, S4.18, S4.19, and S4.20) compared to **3mer-IV** (Figure S4.21) indicating the presence of a non-conventional hydrogen bond.

NMR characterization of **3mer-I**

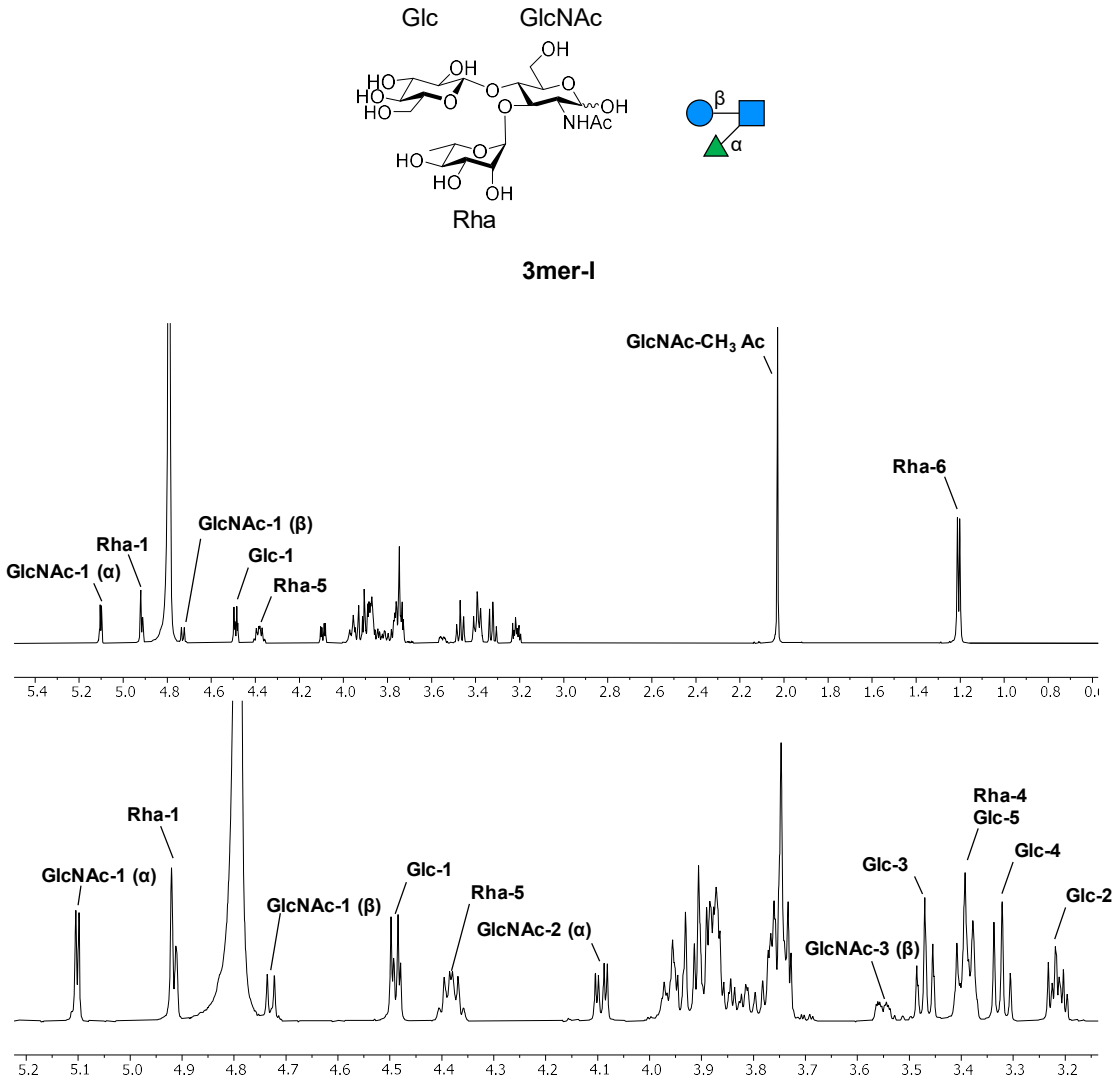


Figure S4.22 ^{1}H NMR (600 MHz, D_2O) of **3mer-I** with assignments.

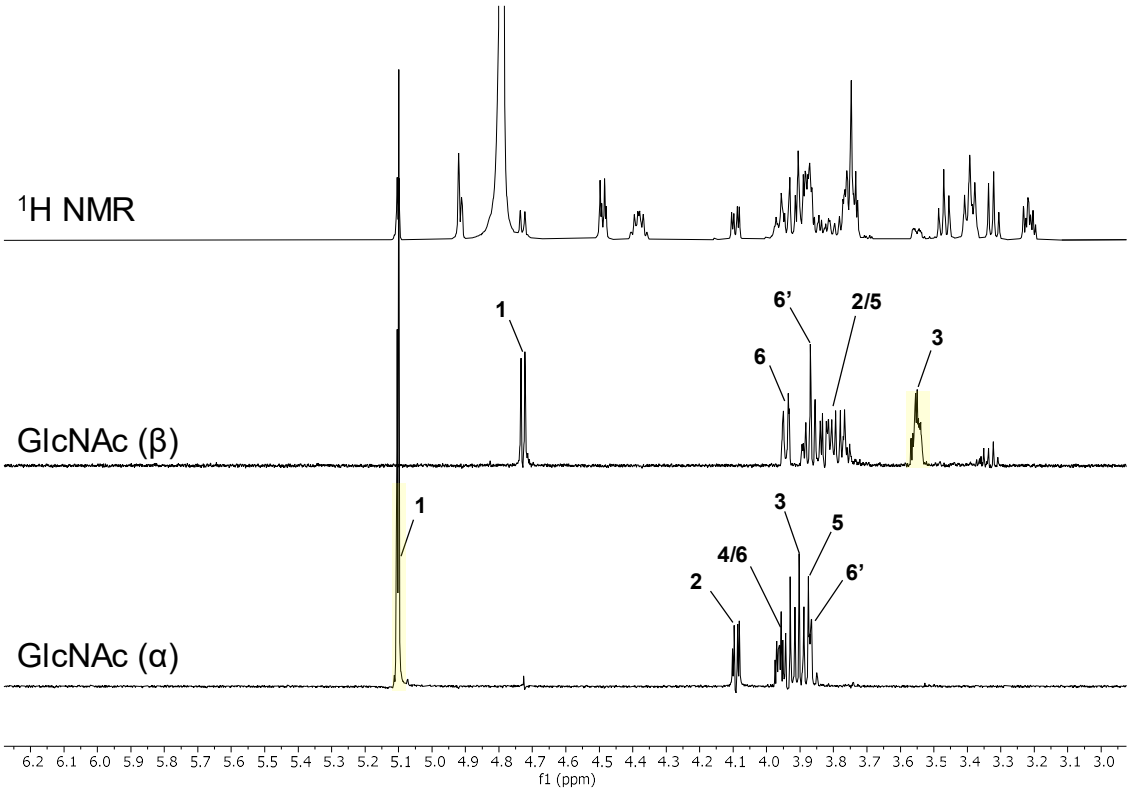


Figure S4.23 1D TOCSY (700 MHz, d9 200 ms, D_2O) of **3mer-I** with assignments. Resonances chosen for selective excitation are highlighted in yellow.

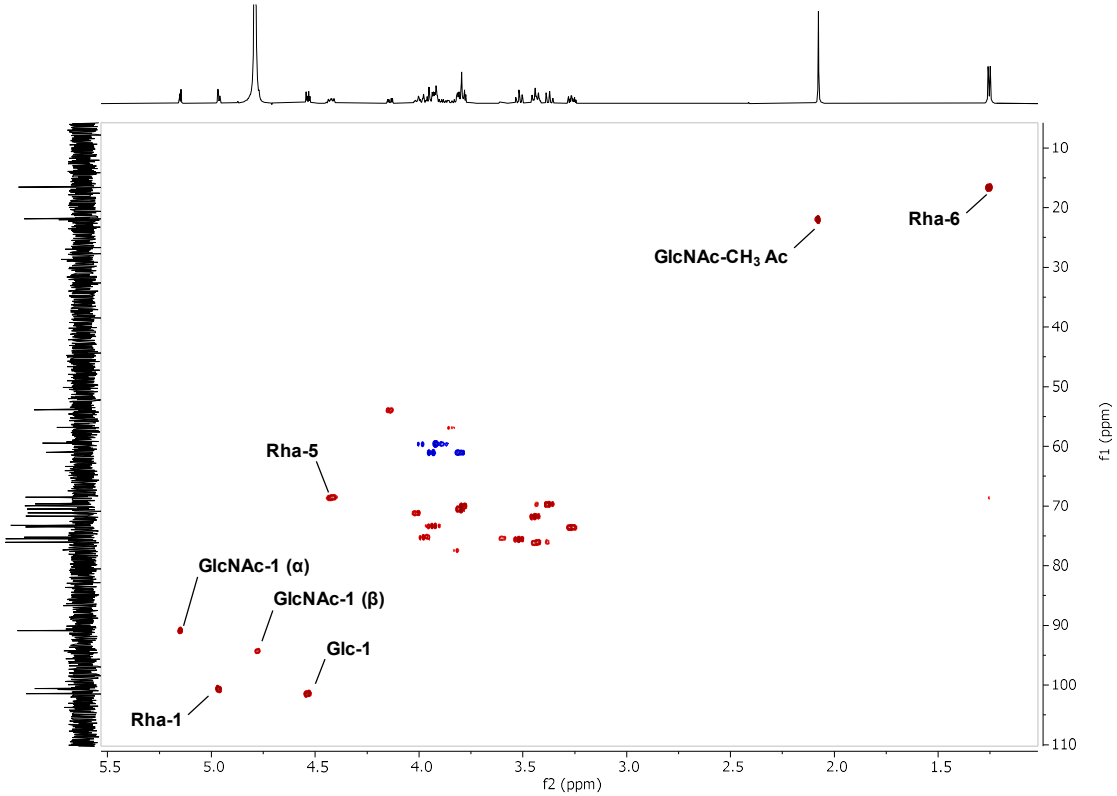


Figure S4.24 HSQC NMR (D_2O) of **3mer-I** with assignments.

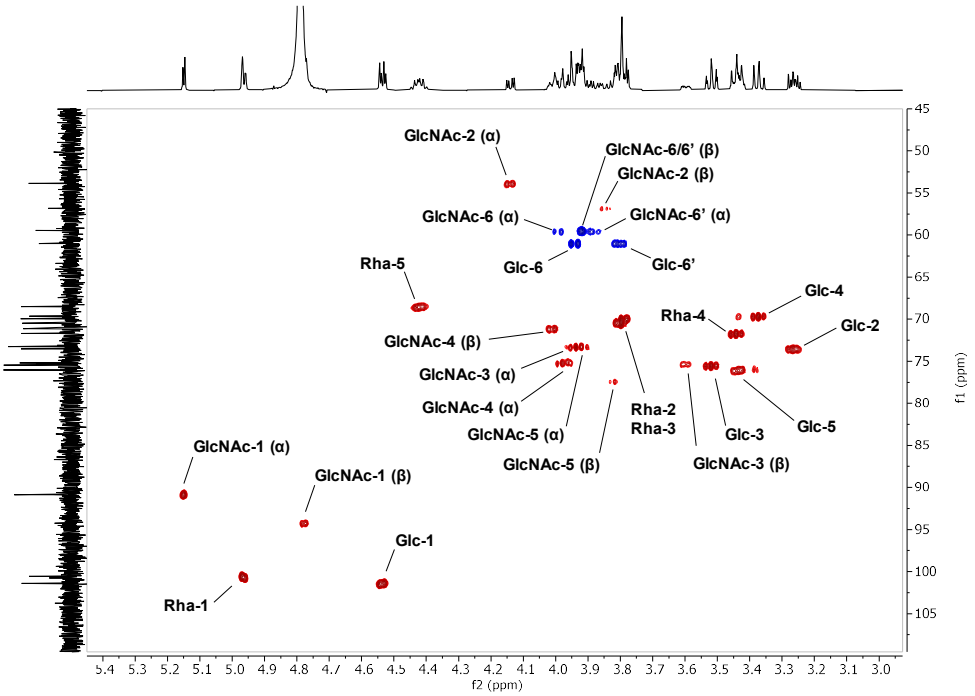


Figure S4.25 Excerpt of HSQC NMR (D_2O) of **3mer-I** with assignments.

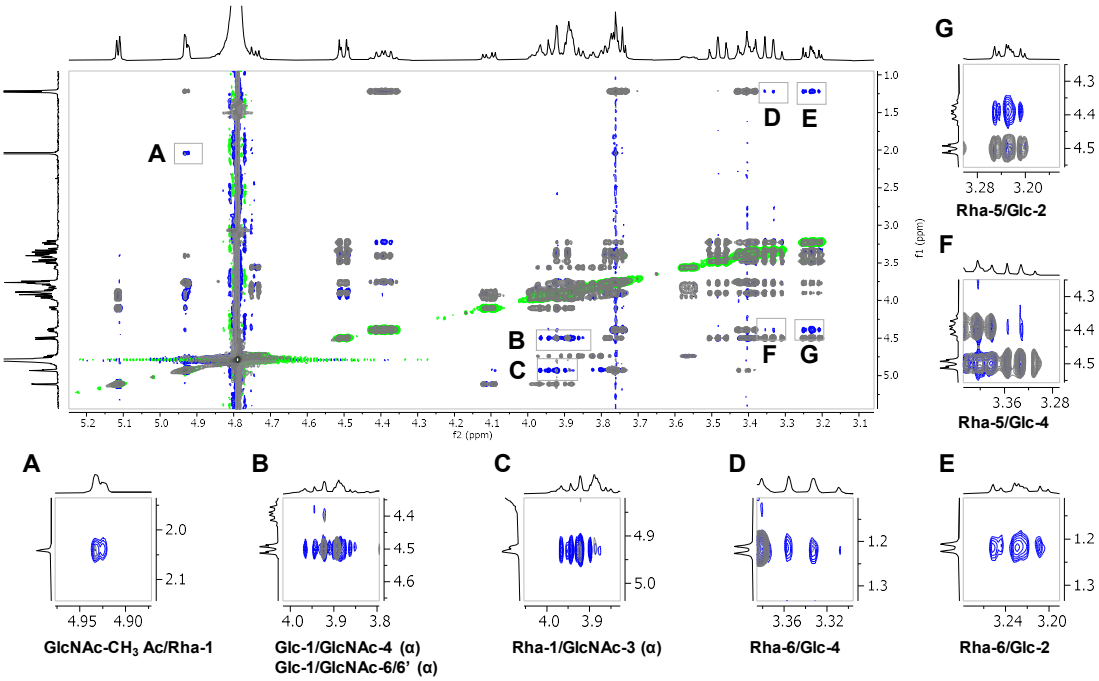


Figure S4.26 Overimposed 2D NOESY (green-blue, 400 MHz, d8 1000 ms, 295 K, D_2O) of **3mer-I** with assignments and 2D TOCSY (grey, 400 MHz, d9 150 ms, 295 K, D_2O).

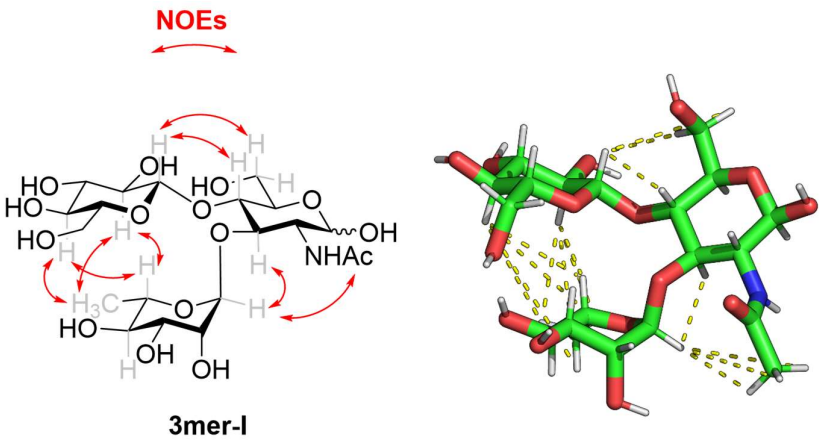


Figure S4.27 Experimentally observed NOEs (*red arrows*) and 3D model with NOEs contacts (yellow dashed line).

NMR characterization of **3mer-II**

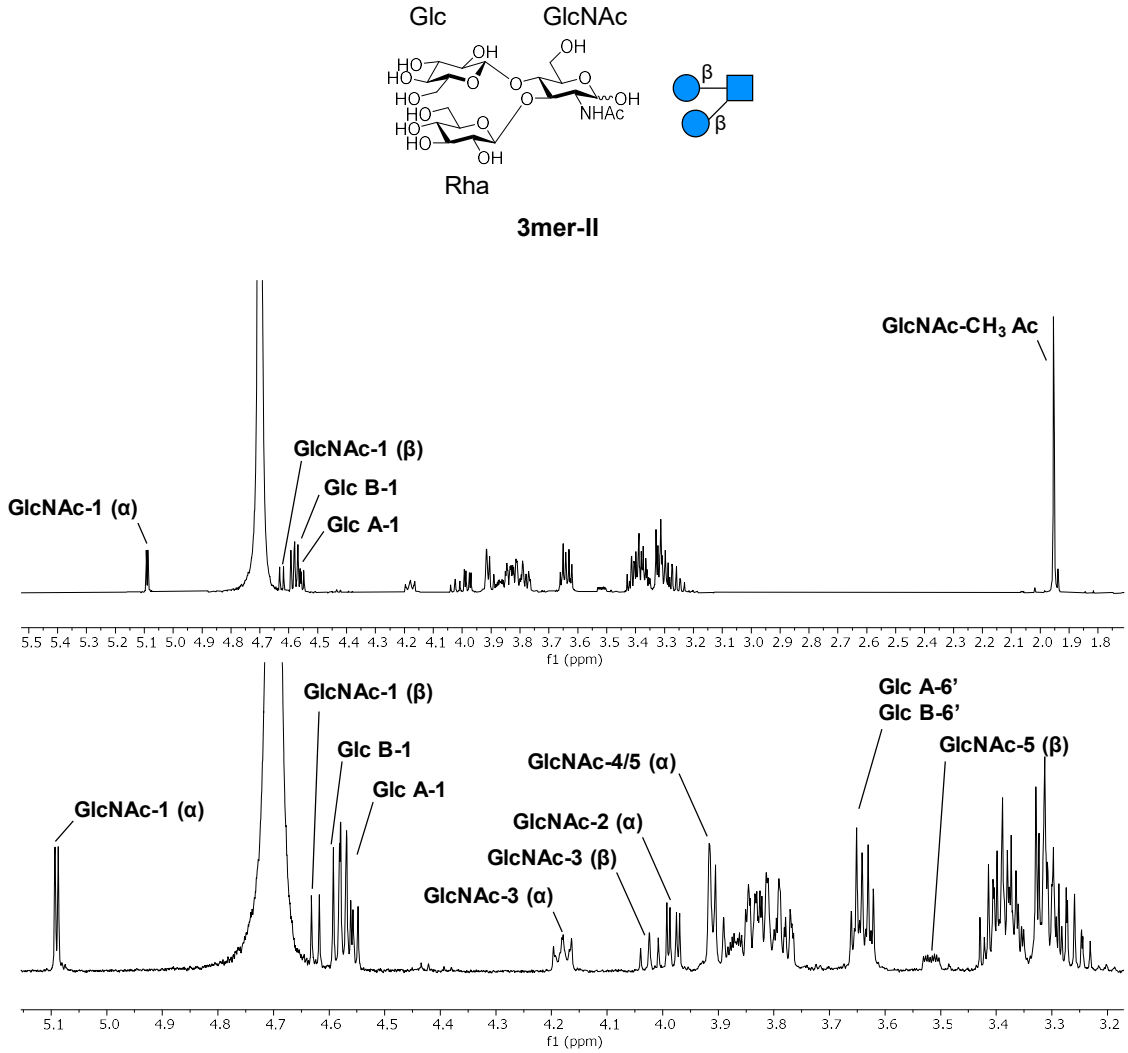


Figure S4.28 ^{1}H NMR (600 MHz, D_2O) of **3mer-II** with assignments.

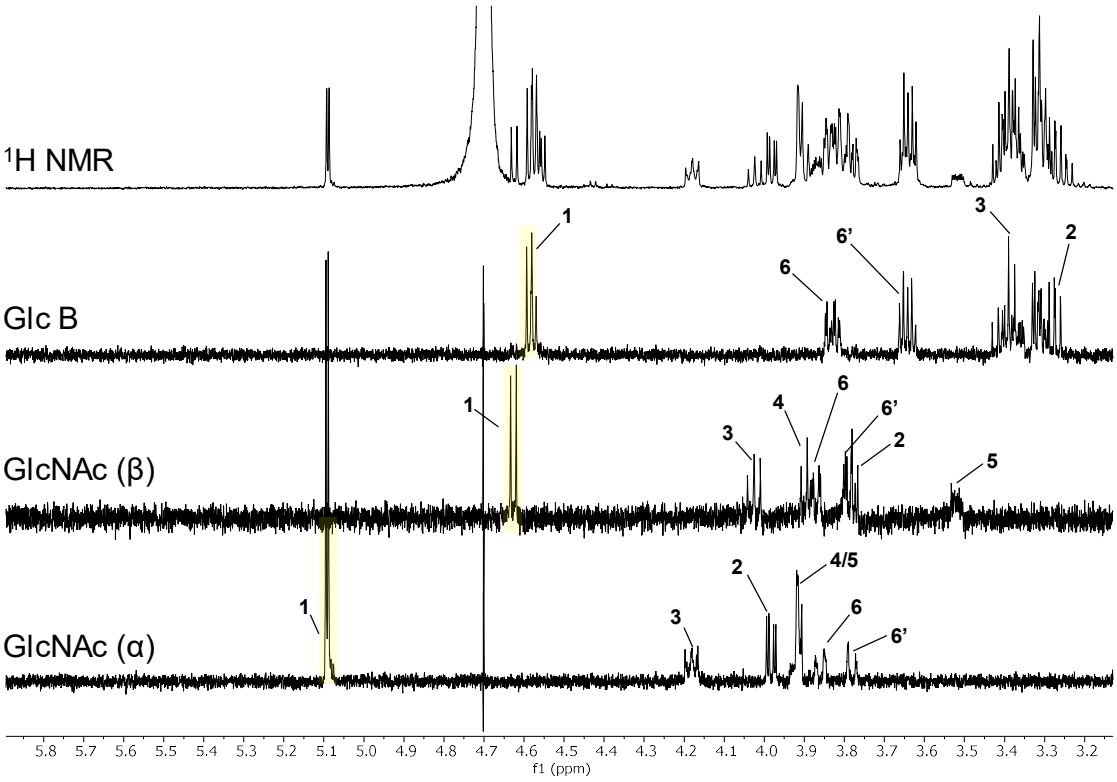


Figure S4.29 1D TOCSY (600 MHz, d9 200 ms, D_2O) of **3mer-II** with assignments. Resonances chosen for selective excitation are highlighted in yellow.

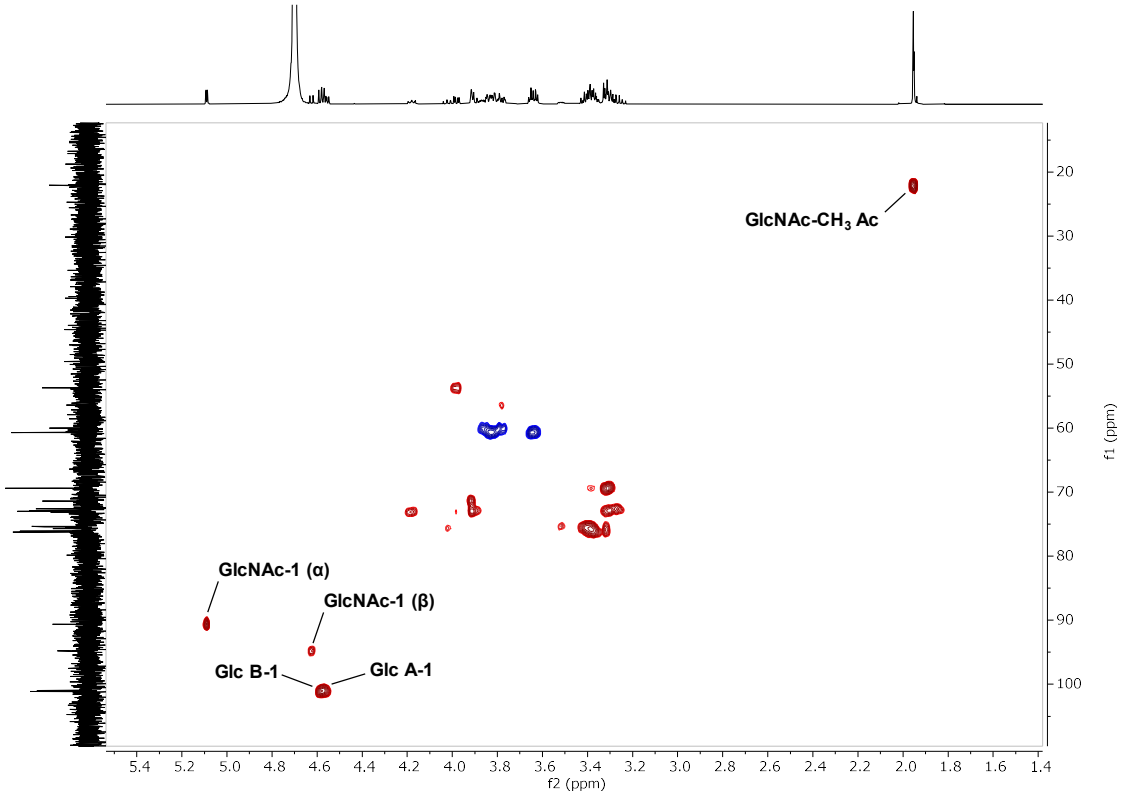


Figure S4.30 HSQC NMR (D_2O) of **3mer-II** with assignments.

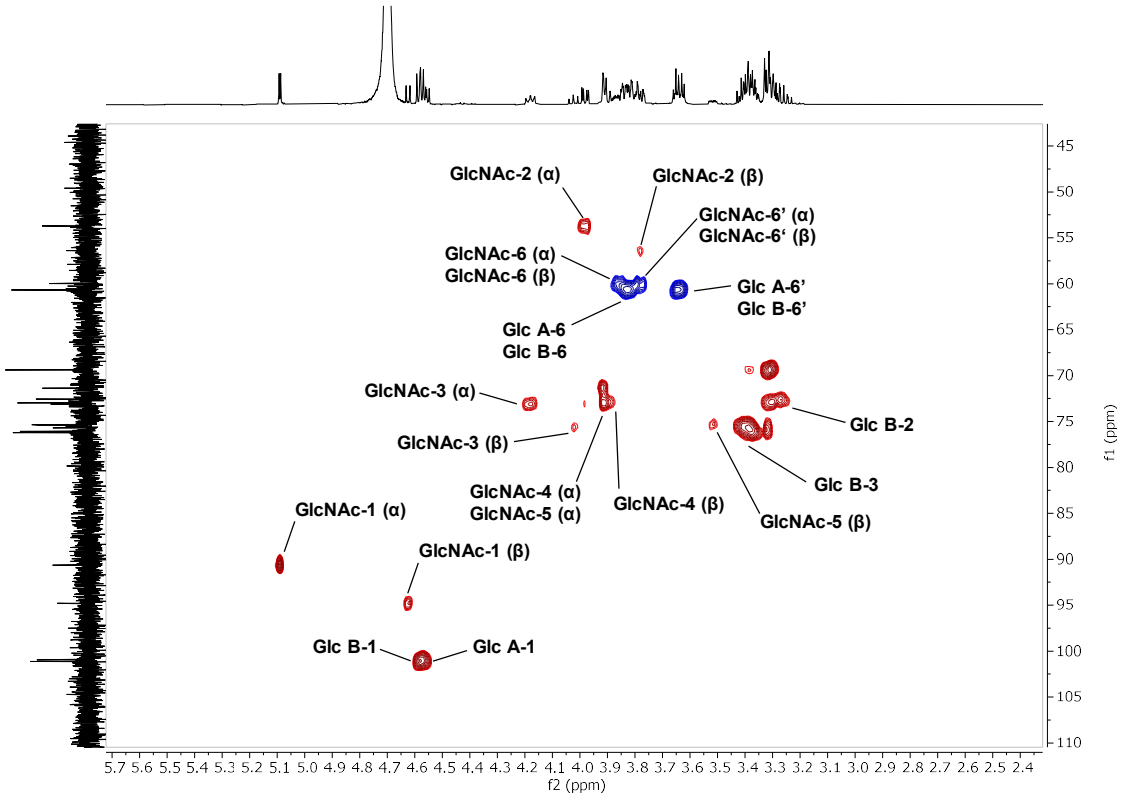


Figure S4.31 Excerpt of HSQC NMR (D_2O) of **3mer-II** with assignments.

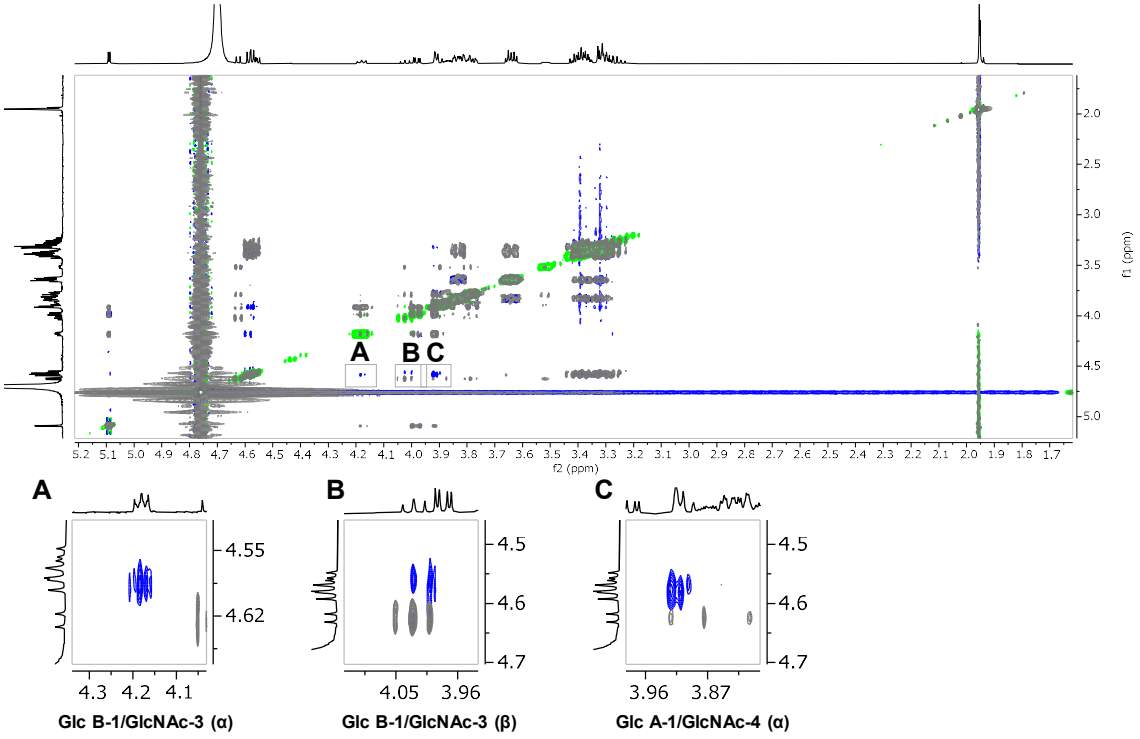


Figure S4.32 Overimposed 2D NOESY (green-blue, 400 MHz, d8 800 ms, 292 K, D_2O) of **3mer-II** with assignments and 2D TOCSY (grey, 400 MHz, d9 120 ms, 292 K, D_2O).

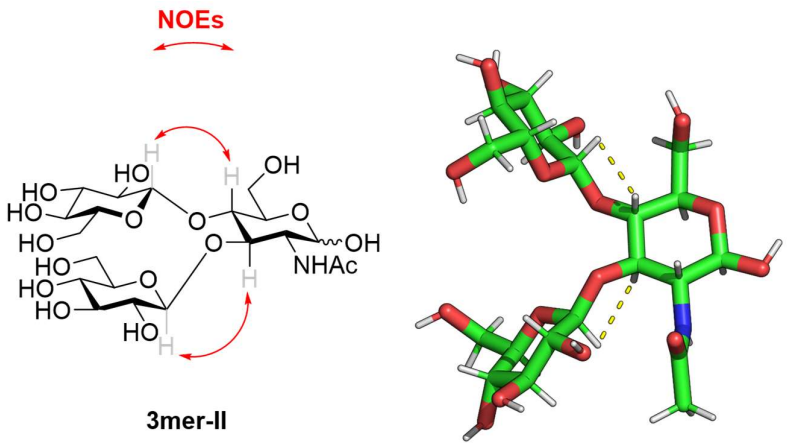


Figure S4.33 Experimentally observed NOEs (*red arrows*) and 3D model with NOEs contacts (yellow dashed line).

NMR characterization of **3mer-III**

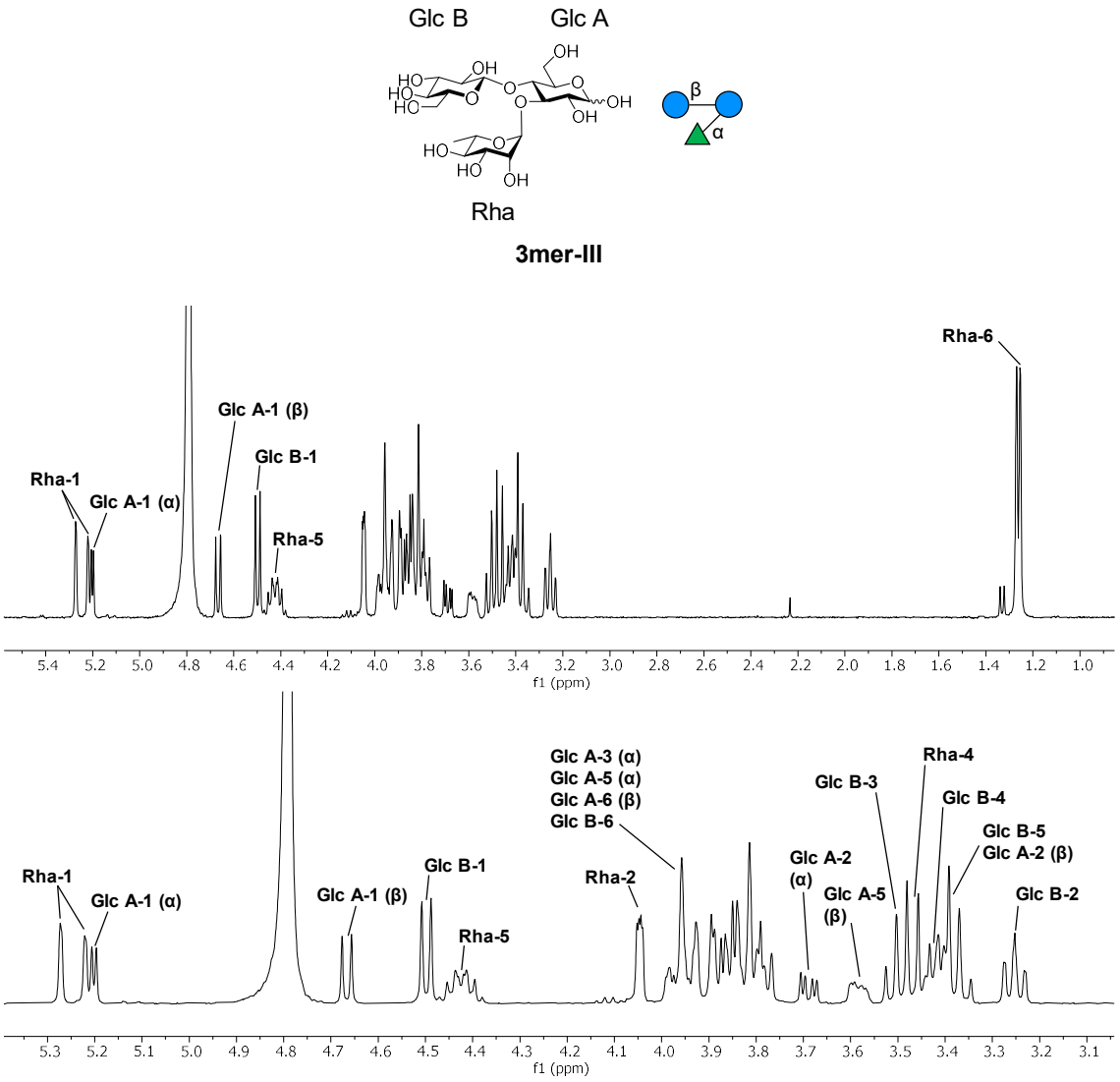


Figure S4.34 ^{1}H NMR (400 MHz, D_2O) of **3mer-III** with assignments.

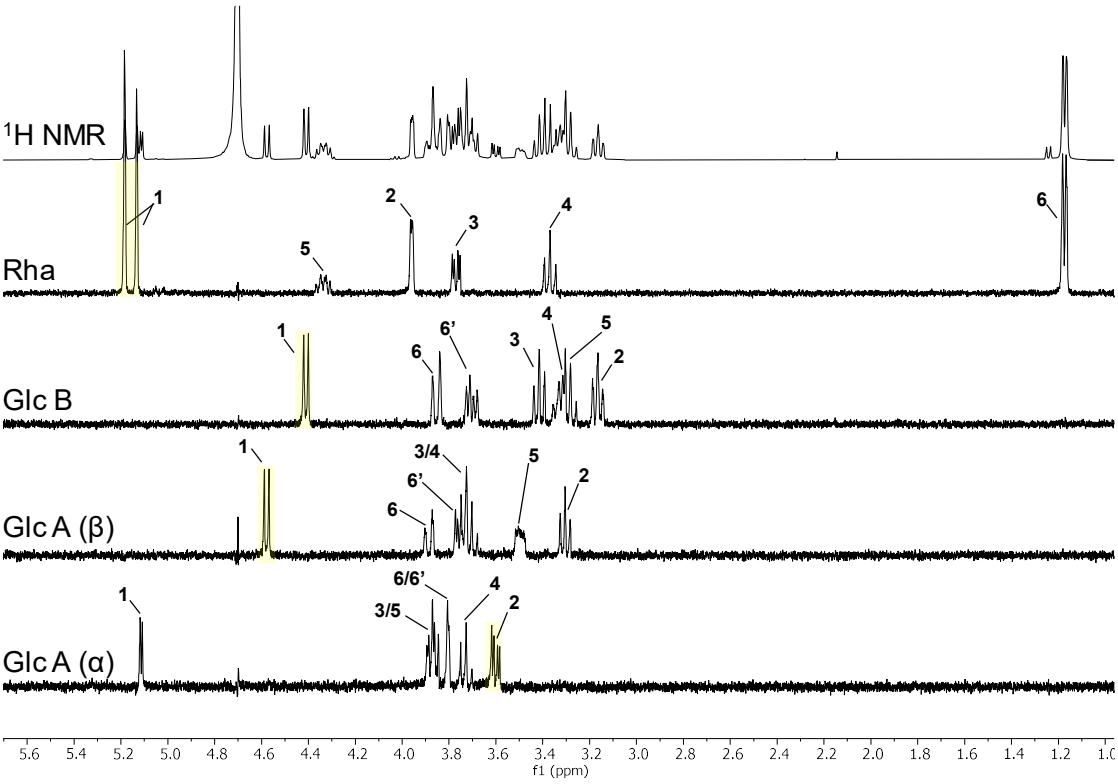


Figure S4.35 1D TOCSY (400 MHz, d9 200 ms, D_2O) of **3mer-III** with assignments. Resonances chosen for selective excitation are highlighted in yellow.

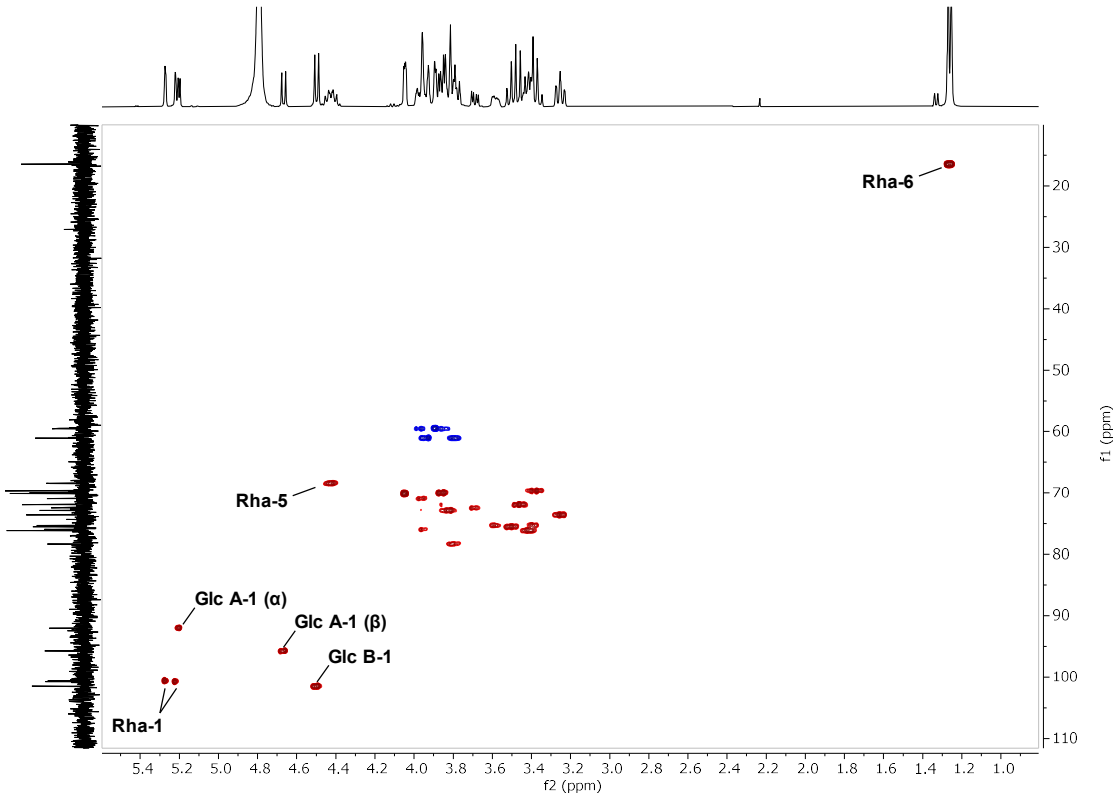


Figure S4.36 HSQC NMR (D_2O) of **3mer-III** with assignments.

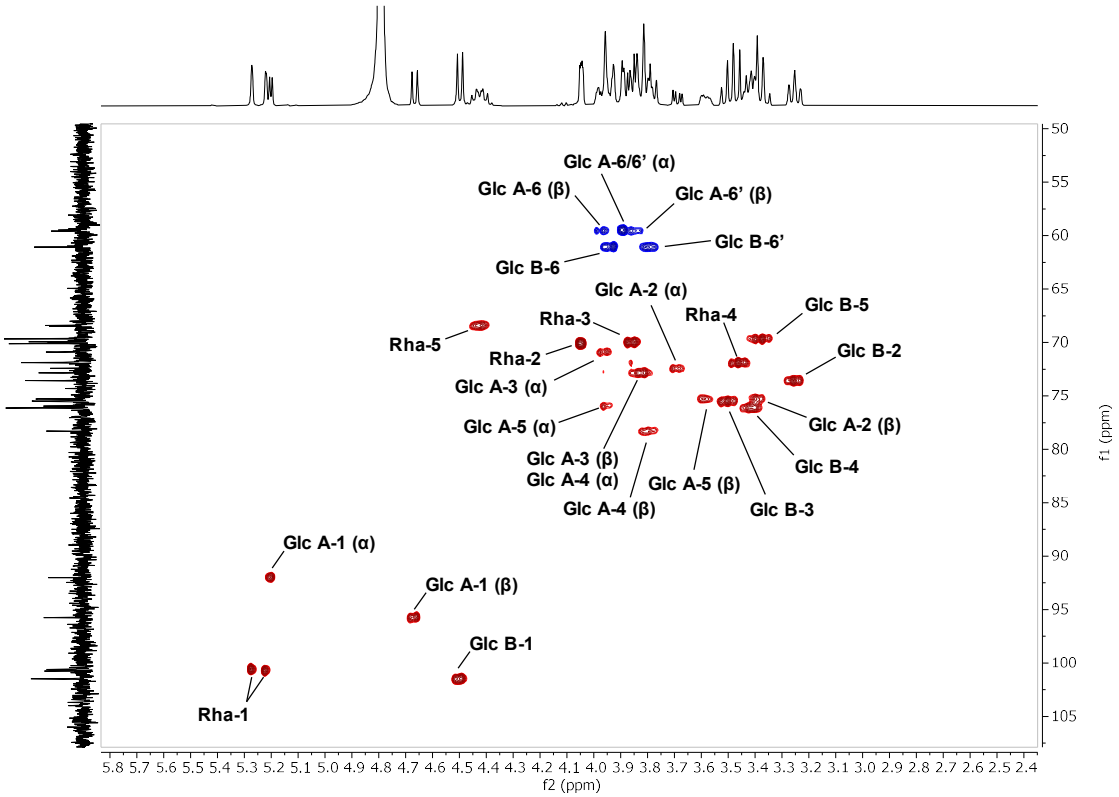


Figure S4.37 Excerpt of HSQC NMR (D_2O) of **3mer-III** with assignments.

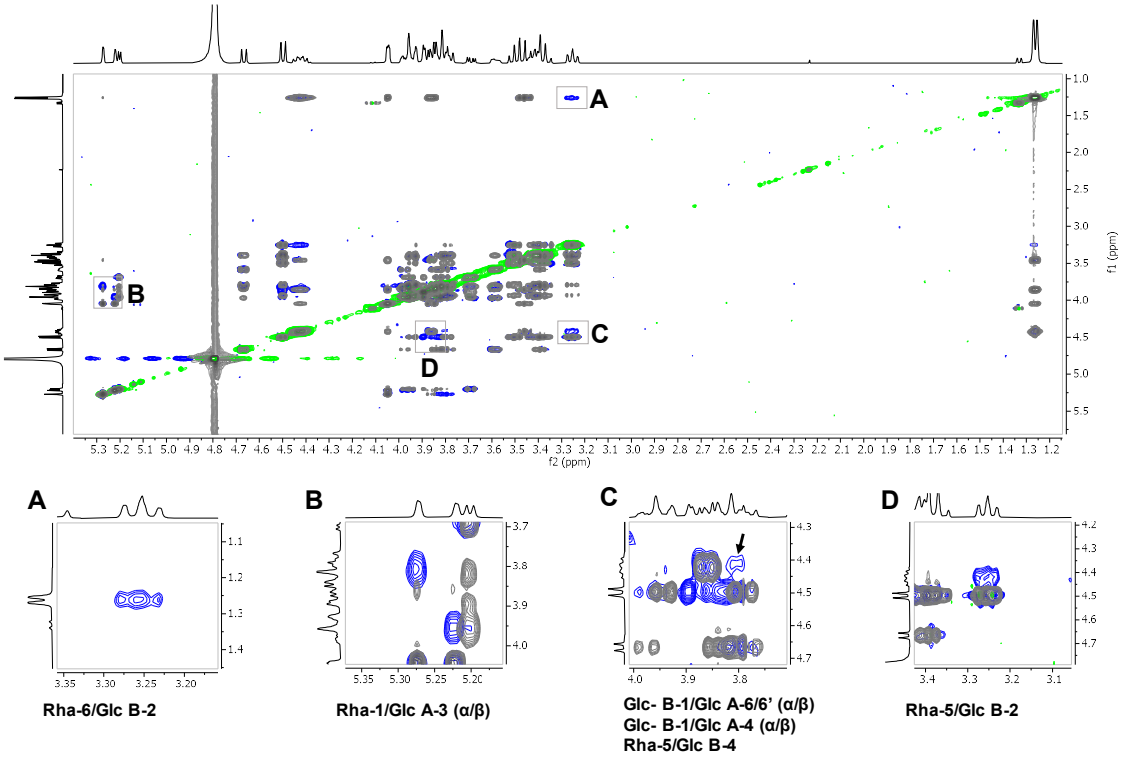


Figure S4.38 Overimposed 2D NOESY (green-blue, 400 MHz, d8 600 ms, 298 K, D_2O) of **3mer-III** with assignments and 2D TOCSY spectrum (grey, 400 MHz, d9 120 ms, 298 K).

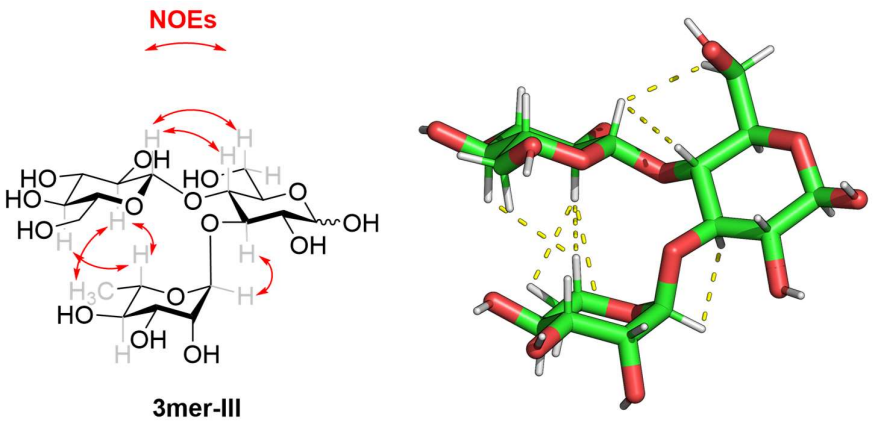


Figure S4.39 Experimentally observed NOEs (*red arrows*) and 3D model with NOE contacts (yellow dashed line).

NMR characterization of **5mer-I**

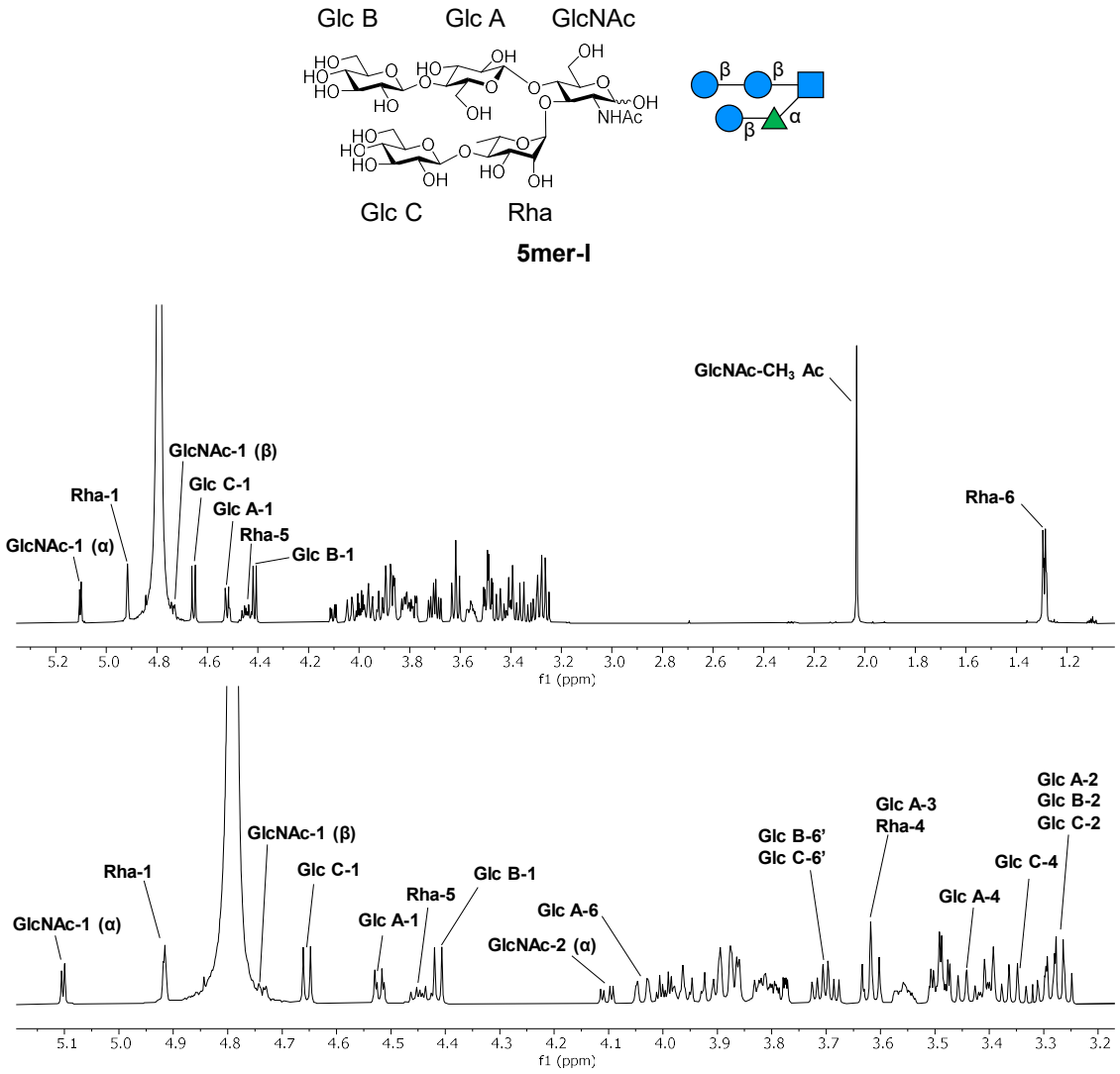


Figure S4.40 ^{1}H NMR (600 MHz, D_2O) of **5mer-I** with assignments.

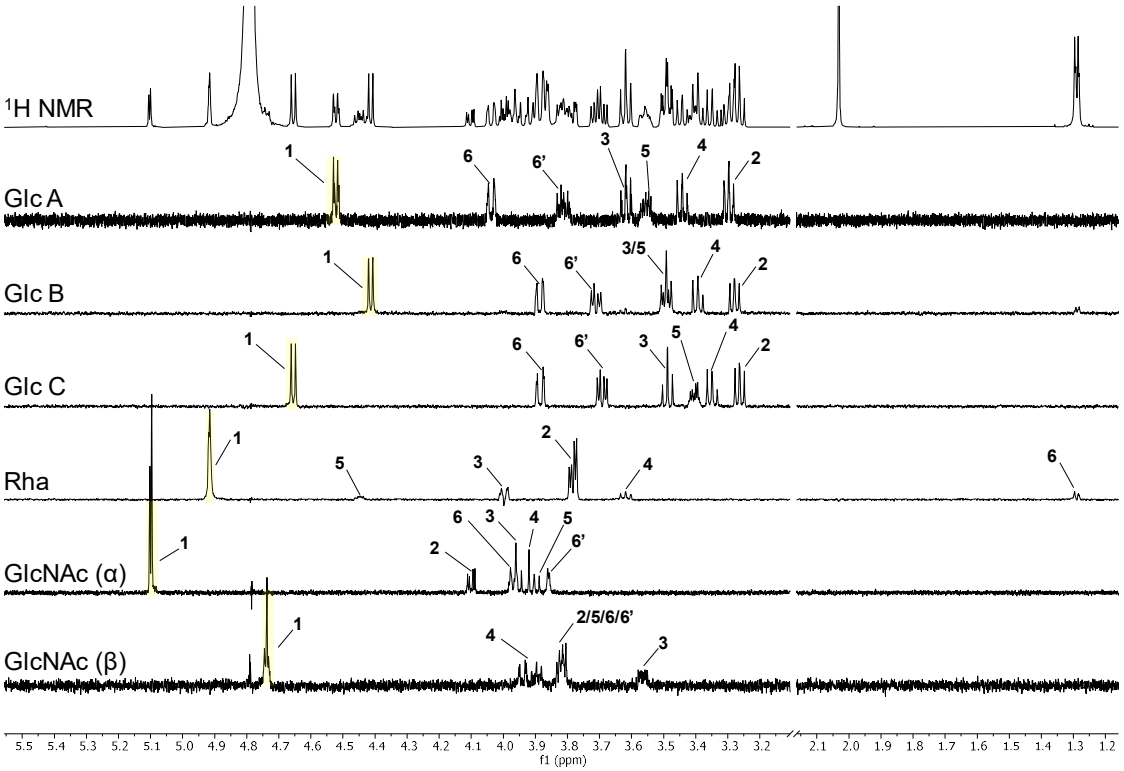


Figure S4.41 1D TOCSY (600 MHz, d9 200 ms, D_2O) of **5mer-I** with assignments. Resonances chosen for selective excitation are highlighted in yellow.

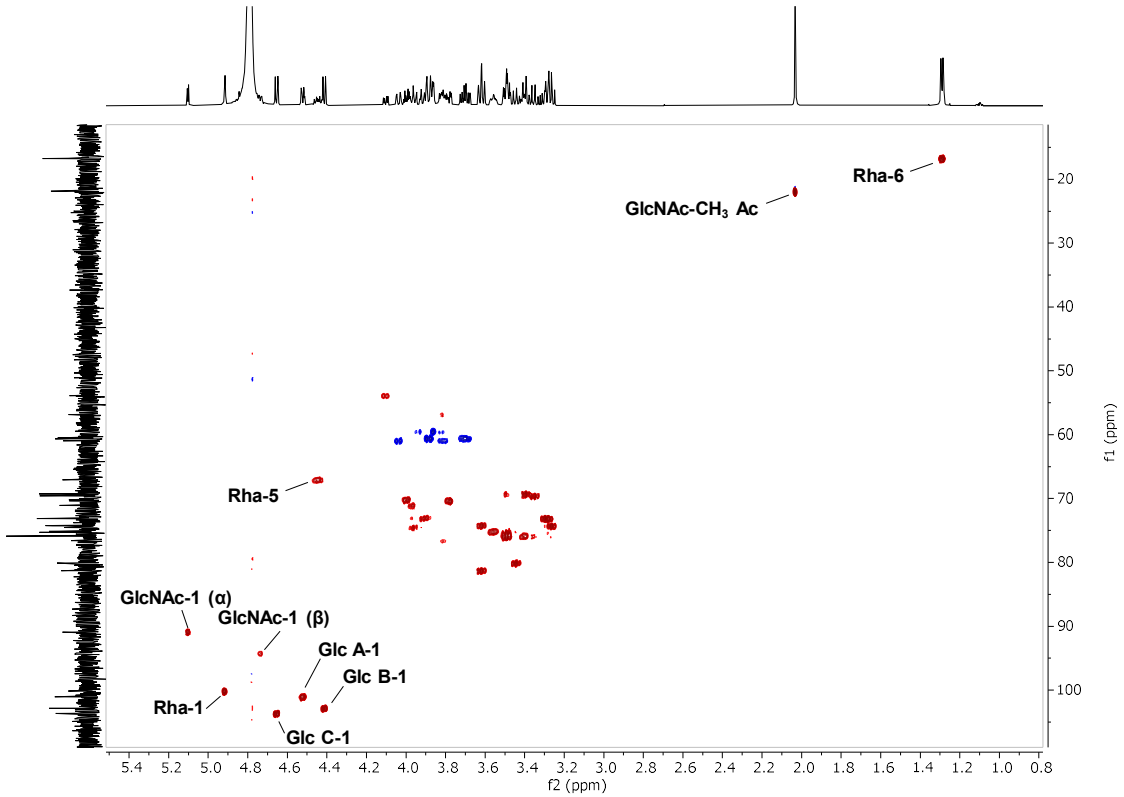


Figure S4.42 HSQC NMR (D_2O) of **5mer-I** with assignments.

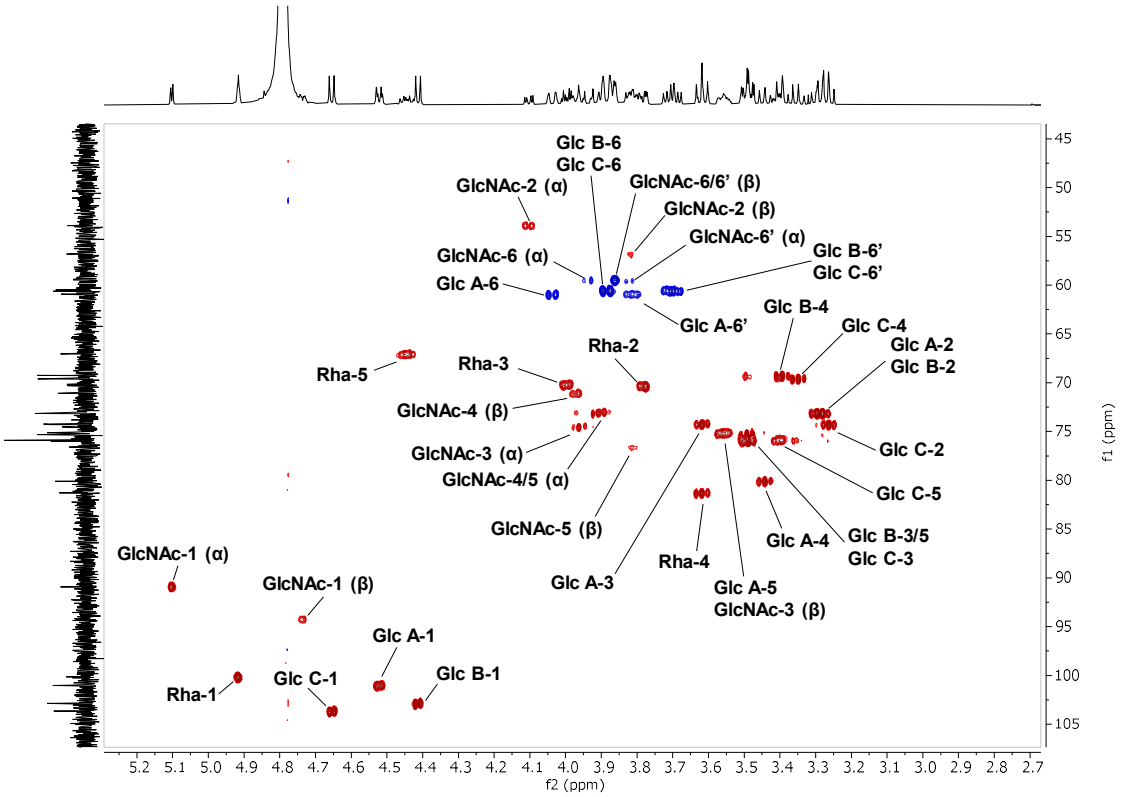


Figure S4.43 Excerpt of HSQC NMR (D_2O) of 5**mer-I** with assignments.

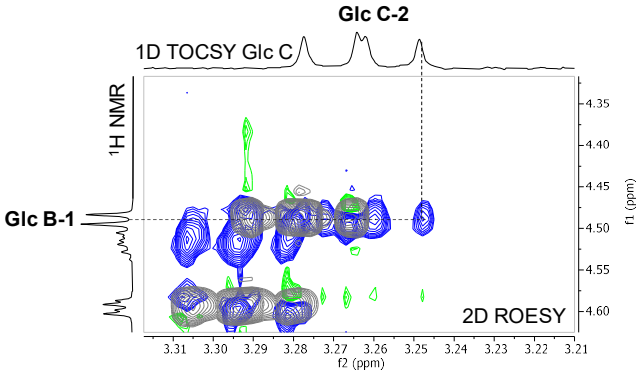


Figure S4.44 Excerpt of 2D ROESY (green-blue) and 2D TOCSY of **5mer-I** over imposed (grey). The horizontal trace presents the 1D TOCSY with selective excitation of Glc C-1. The NOE cross peak is marked with a dashed line.

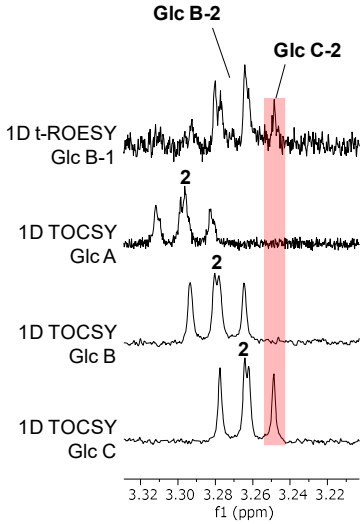


Figure S4.45 Excerpt of the overlay of 1D t-ROESY (600 MHz, p15 200 ms, 294 K, D_2O) and 1D TOCSY of **5mer-I**. The 1D t-ROESY was obtained by selective excitation of the Glc B-1 resonance.

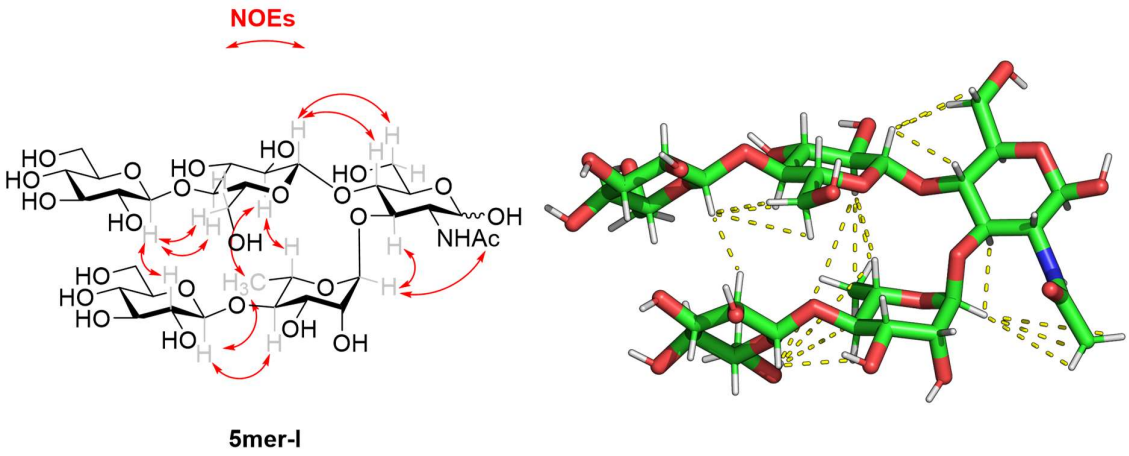


Figure S4.46 All experimentally observed NOEs (*red arrows*) and 3D model with NOEs contacts (yellow dashed lines).

NMR characterization of **9mer-I**, **9mer-I-B**, and **9mer-I-F**

The NMR characterization of **9mer-I-B** and **9mer-I-F** was performed by Dr. Ana Poveda (CIC BioGUNE).

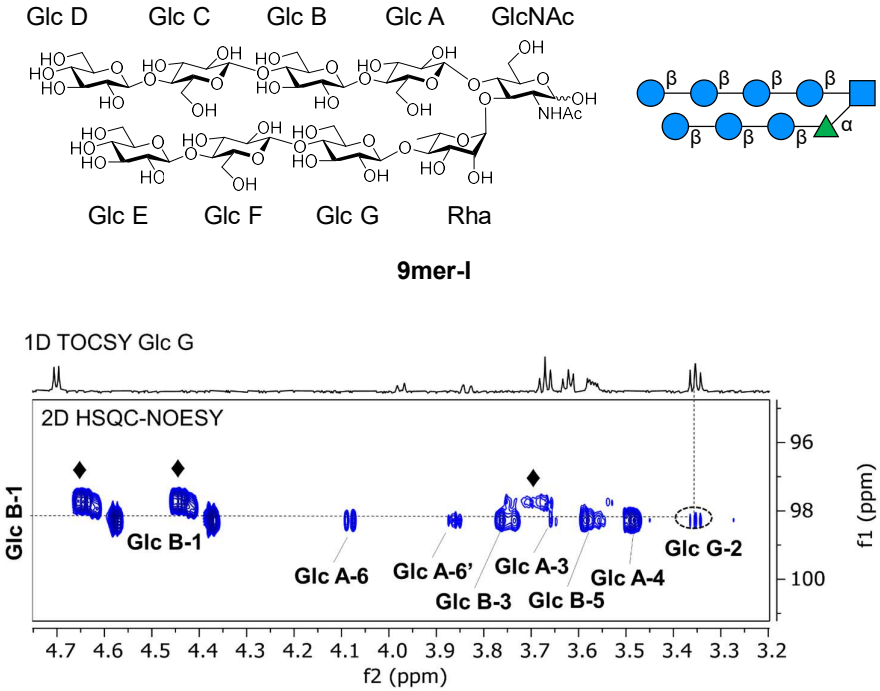


Figure S4.47 Excerpt of the 2D HSQC-NOESY NMR of **9mer-I-B** showing the expected intra-residue NOEs and the key NOE (Glc B-1/Glc G-2) between the two strands of the hairpin (298 K, mixing time 300 ms, D_2O, 800 MHz).

7 References

(1) Whitesides, G. M. Self-Assembly at All Scales. *Science* **2002**, *295* (5564), 2418–2421. https://doi.org/10.1126/science.1070821.

(2) Shoulders, M. D.; Raines, R. T. Collagen Structure and Stability. *Annu. Rev. Biochem.* **2009**, *78* (1), 929–958. https://doi.org/10.1146/annurev.biochem.77.032207.120833.

(3) Paajanen, A.; Ceccherini, S.; Maloney, T.; Ketoja, J. A. Chirality and Bound Water in the Hierarchical Cellulose Structure. *Cellulose* **2019**, *26* (10), 5877–5892. https://doi.org/10.1007/s10570-019-02525-7.

(4) Rinaudo, M. Chitin and Chitosan: Properties and Applications. *Prog. Polym. Sci.* **2006**, *31* (7), 603–632. https://doi.org/10.1016/j.progpolymsci.2006.06.001.

(5) Rest, C.; Kandanelli, R.; Fernández, G. Strategies to Create Hierarchical Self-Assembled Structures via Cooperative Non-Covalent Interactions. *Chem. Soc. Rev.* **2015**, *44* (8), 2543–2572. https://doi.org/10.1039/C4CS00497C.

(6) Sato, K.; Hendricks, M. P.; Palmer, L. C.; Stupp, S. I. Peptide Supramolecular Materials for Therapeutics. *Chem. Soc. Rev.* **2018**, *47* (20), 7539–7551. https://doi.org/10.1039/C7CS00735C.

(7) Zhang, S. Fabrication of Novel Biomaterials through Molecular Self-Assembly. *Nat. Biotechnol.* **2003**, *21* (10), 1171–1178. https://doi.org/10.1038/nbt874.

(8) Luo, Z.; Zhang, S. Designer Nanomaterials Using Chiral Self-Assembling Peptide Systems and Their Emerging Benefit for Society. *Chem. Soc. Rev.* **2012**, *41* (13), 4736. https://doi.org/10.1039/c2cs15360b.

(9) Levin, A.; Hakala, T. A.; Schnaider, L.; Bernardes, G. J. L.; Gazit, E.; Knowles, T. P. J. Biomimetic Peptide Self-Assembly for Functional Materials. *Nat. Rev. Chem.* **2020**, *4* (11), 615–634. https://doi.org/10.1038/s41570-020-0215-y.

(10) Cherny, I.; Gazit, E. Amyloids: Not Only Pathological Agents but Also Ordered Nanomaterials. *Angew. Chem., Int. Ed.* **2008**, *47* (22), 4062–4069. https://doi.org/10.1002/anie.200703133.

(11) Sinha, N. J.; Langenstein, M. G.; Pochan, D. J.; Kloxin, C. J.; Saven, J. G. Peptide Design and Self-Assembly into Targeted Nanostructure and Functional Materials. *Chem. Rev.* **2021**, *121* (22), 13915–13935. https://doi.org/10.1021/acs.chemrev.1c00712.

(12) De Santis, E.; Ryadnov, M. G. Peptide Self-Assembly for Nanomaterials: The Old New Kid on the Block. *Chem. Soc. Rev.* **2015**, *44* (22), 8288–8300. https://doi.org/10.1039/C5CS00470E.

(13) Aida, T.; Meijer, E. W.; Stupp, S. I. Functional Supramolecular Polymers. *Science* **2012**, *335* (6070), 813–817. https://doi.org/10.1126/science.1205962.

(14) Zhang, S.; Holmes, T.; Lockshin, C.; Rich, A. Spontaneous Assembly of a Self-Complementary Oligopeptide to Form a Stable Macroscopic Membrane. *Proc. Natl. Acad. Sci.* **1993**, *90* (8), 3334–3338. https://doi.org/10.1073/pnas.90.8.3334.

(15) Jiang, T.; Xu, C.; Liu, Y.; Liu, Z.; Wall, J. S.; Zuo, X.; Lian, T.; Salaita, K.; Ni, C.; Pochan, D.; Conticello, V. P. Structurally Defined Nanoscale Sheets from Self-Assembly of Collagen-Mimetic Peptides. *J. Am. Chem. Soc.* **2014**, *136* (11), 4300–4308. https://doi.org/10.1021/ja412867z.

(16) Schneider, J. P.; Pochan, D. J.; Ozbas, B.; Rajagopal, K.; Pakstis, L.; Kretsinger, J. Responsive Hydrogels from the Intramolecular Folding and Self-Assembly of a Designed Peptide. *J. Am. Chem. Soc.* **2002**, *124* (50), 15030–15037. https://doi.org/10.1021/ja027993g.

(17) Veneziano, R.; Ratanalert, S.; Zhang, K.; Zhang, F.; Yan, H.; Chiu, W.; Bathe, M. Designer Nanoscale DNA Assemblies Programmed from the Top Down. *Science* **2016**, *352* (6293), 1534–1534. https://doi.org/10.1126/science.aaf4388.

(18) Hong, F.; Zhang, F.; Liu, Y.; Yan, H. DNA Origami: Scaffolds for Creating Higher Order Structures. *Chem. Rev.* **2017**, *117* (20), 12584–12640. https://doi.org/10.1021/acs.chemrev.6b00825.

(19) Silva, G. A.; Czeisler, C.; Niece, K. L.; Beniash, E.; Harrington, D. A.; Kessler, J. A.; Stupp, S. I. Selective Differentiation of Neural Progenitor Cells by High-Epitope Density Nanofibers. *Science* **2004**, *303* (5662), 1352–1355. https://doi.org/10.1126/science.1093783.

(20) Hendricks, M. P.; Sato, K.; Palmer, L. C.; Stupp, S. I. Supramolecular Assembly of Peptide Amphiphiles. *Acc. Chem. Res.* **2017**, *50* (10), 2440–2448. https://doi.org/10.1021/acs.accounts.7b00297.

(21) Hendler, N.; Sidelman, N.; Reches, M.; Gazit, E.; Rosenberg, Y.; Richter, S. Formation of Well-Organized Self-Assembled Films from Peptide Nanotubes. *Adv. Mater.* **2007**, *19* (11), 1485–1488. https://doi.org/10.1002/adma.200602265.

(22) Ke, P. C.; Zhou, R.; Serpell, L. C.; Riek, R.; Knowles, T. P. J.; Lashuel, H. A.; Gazit, E.; Hamley, I. W.; Davis, T. P.; Fändrich, M.; Otzen, D. E.; Chapman, M. R.; Dobson, C. M.; Eisenberg, D. S.; Mezzenga, R. Half a Century of Amyloids: Past, Present and Future. *Chem. Soc. Rev.* **2020**, *49* (15), 5473–5509. https://doi.org/10.1039/C9CS00199A.

(23) Wei, G.; Su, Z.; Reynolds, N. P.; Arosio, P.; Hamley, I. W.; Gazit, E.; Mezzenga, R. Self-Assembling Peptide and Protein Amyloids: From Structure to Tailored Function in Nanotechnology. *Chem. Soc. Rev.* **2017**, *46* (15), 4661–4708. https://doi.org/10.1039/C6CS00542J.

(24) Rajagopal, K.; Schneider, J. P. Self-Assembling Peptides and Proteins for Nanotechnological Applications. *Curr. Opin. Struct. Biol.* **2004**, *14* (4), 480–486. https://doi.org/10.1016/j.sbi.2004.06.006.

(25) Lapenta, F.; Aupič, J.; Strmšek, Ž.; Jerala, R. Coiled Coil Protein Origami: From Modular Design Principles towards Biotechnological Applications. *Chem. Soc. Rev.* **2018**, *47* (10), 3530–3542. https://doi.org/10.1039/C7CS00822H.

(26) Woolfson, D. N. The Design of Coiled-Coil Structures and Assemblies; 2005; pp 79–112. https://doi.org/10.1016/S0065-3233(05)70004-8.

(27) Tanrikulu, I. C.; Forticaux, A.; Jin, S.; Raines, R. T. Peptide Tessellation Yields Micrometre-Scale Collagen Triple Helices. *Nat. Chem.* **2016**, *8* (11), 1008–1014. https://doi.org/10.1038/nchem.2556.

(28) O'Leary, L. E. R.; Fallas, J. A.; Bakota, E. L.; Kang, M. K.; Hartgerink, J. D. Multi-Hierarchical Self-Assembly of a Collagen Mimetic Peptide from Triple Helix to Nanofibre and Hydrogel. *Nat. Chem.* **2011**, *3* (10), 821–828. https://doi.org/10.1038/nchem.1123.

(29) Jiang, T.; Xu, C.; Zuo, X.; Conticello, V. P. Structurally Homogeneous Nanosheets from Self-Assembly of a Collagen-Mimetic Peptide. *Angew. Chem., Int. Ed.* **2014**, *53* (32), 8367–8371. https://doi.org/10.1002/anie.201403780.

(30) King, N. P.; Sheffler, W.; Sawaya, M. R.; Vollmar, B. S.; Sumida, J. P.; André, I.; Gonen, T.; Yeates, T. O.; Baker, D. Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy. *Science* **2012**, *336* (6085), 1171–1174. https://doi.org/10.1126/science.1219364.

(31) King, N. P.; Bale, J. B.; Sheffler, W.; McNamara, D. E.; Gonen, S.; Gonen, T.; Yeates, T. O.; Baker, D. Accurate Design of Co-Assembling Multi-Component Protein Nanomaterials. *Nature* **2014**, *510* (7503), 103–108. https://doi.org/10.1038/nature13404.

(32) Gonen, S.; DiMaio, F.; Gonen, T.; Baker, D. Design of Ordered Two-Dimensional Arrays Mediated by Noncovalent Protein-Protein Interfaces. *Science* **2015**, *348* (6241), 1365–1368. https://doi.org/10.1126/science.aaa9897.

(33) Luo, Q.; Hou, C.; Bai, Y.; Wang, R.; Liu, J. Protein Assembly: Versatile Approaches to Construct Highly Ordered Nanostructures. *Chem. Rev.* **2016**, *116* (22), 13571–13632.

https://doi.org/10.1021/acs.chemrev.6b00228.

(34) Huang, P.-S.; Boyken, S. E.; Baker, D. The Coming of Age of de Novo Protein Design. *Nature* **2016**, *537* (7620), 320–327. https://doi.org/10.1038/nature19946.

(35) Seeman, N. C. Nanomaterials Based on DNA. *Annu. Rev. Biochem.* **2010**, *79* (1), 65–87. https://doi.org/10.1146/annurev-biochem-060308-102244.

(36) Seeman, N. C. Nucleic Acid Junctions and Lattices. *J. Theor. Biol.* **1982**, *99* (2), 237–247. https://doi.org/10.1016/0022-5193(82)90002-9.

(37) Delbianco, M.; Bharate, P.; Varela-Aramburu, S.; Seeberger, P. H. Carbohydrates in Supramolecular Chemistry. *Chem. Rev.* **2016**, *116* (4), 1693–1752. https://doi.org/10.1021/acs.chemrev.5b00516.

(38) Brito, A.; Kassem, S.; Reis, R. L.; Ulijn, R. V.; Pires, R. A.; Pashkuleva, I. Carbohydrate Amphiphiles for Supramolecular Biomaterials: Design, Self-Assembly, and Applications. *Chem* **2021**, *7* (11), 2943–2964. https://doi.org/10.1016/j.chempr.2021.04.011.

(39) Diener, M.; Adamcik, J.; Sánchez-Ferrer, A.; Jaedig, F.; Schefer, L.; Mezzenga, R. Primary, Secondary, Tertiary and Quaternary Structure Levels in Linear Polysaccharides: From Random Coil, to Single Helix to Supramolecular Assembly. *Biomacromolecules* **2019**, *20* (4), 1731–1739. https://doi.org/10.1021/acs.biomac.9b00087.

(40) Schefer, L.; Adamcik, J.; Diener, M.; Mezzenga, R. Supramolecular Chiral Self-Assembly and Supercoiling Behavior of Carrageenans at Varying Salt Conditions. *Nanoscale* **2015**, *7* (39), 16182–16188. https://doi.org/10.1039/c5nr04525h.

(41) Fox, S. C.; Li, B.; Xu, D.; Edgar, K. J. Regioselective Esterification and Etherification of Cellulose: A Review. *Biomacromolecules* **2011**, *12* (6), 1956–1972. https://doi.org/10.1021/bm200260d.

(42) Habibi, Y.; Lucia, L. A.; Rojas, O. J. Cellulose Nanocrystals: Chemistry, Self-Assembly, and Applications. *Chem. Rev.* **2010**, *110* (6), 3479–3500. https://doi.org/10.1021/cr900339w.

(43) Wormald, M. R.; Petrescu, A. J.; Pao, Y.-L.; Glithero, A.; Elliott, T.; Dwek, R. A. Conformational Studies of Oligosaccharides and Glycopeptides: Complementarity of NMR, X-Ray Crystallography, and Molecular Modelling. *Chem. Rev.* **2002**, *102* (2), 371–386. https://doi.org/10.1021/cr990368i.

(44) Woods, R. J. Predicting the Structures of Glycans, Glycoproteins, and Their Complexes. *Chem. Rev.* **2018**, *118* (17), 8005–8024. https://doi.org/10.1021/acs.chemrev.8b00032.

(45) Imberty, a; Delage, M. M.; Bourne, Y.; Cambillau, C.; Pérez, S. Data Bank of Three-Dimensional Structures of Disaccharides: Part II, N-Acetyllactosaminic Type N-Glycans. Comparison with the Crystal Structure of a Biantennary Octasaccharide. *Glycoconj. J.* **1991**, *8* (6), 456–483. https://doi.org/10.1007/BF00769847.

(46) Booth, H.; Khedhair, K. A. Endo-Anomeric and Exo-Anomeric Effects in 2-Substituted Tetrahydropyrans. *J. Chem. Soc. Chem. Commun.* **1985**, *8* (8), 467. https://doi.org/10.1039/c39850000467.

(47) Perić-Hassler, L.; Hansen, H. S.; Baron, R.; Hünenberger, P. H. Conformational Properties of Glucose-Based Disaccharides Investigated Using Molecular Dynamics Simulations with Local Elevation Umbrella Sampling. *Carbohydr. Res.* **2010**, *345* (12), 1781–1801. https://doi.org/10.1016/j.carres.2010.05.026.

(48) Woods, R. J.; Pathiaseril, A.; Wormald, M. R.; Edge, C. J.; Dwek, R. A. The High Degree of Internal Flexibility Observed for an Oligomannose Oligosaccharide Does Not Alter the Overall Topology of the Molecule. *Eur. J. Biochem.* **1998**, *258* (2), 372–386. https://doi.org/10.1046/j.1432-1327.1998.2580372.x.

(49) Sletmoen, M.; Maurstad, G.; Sikorski, P.; Paulsen, B. S.; Stokke, B. T. Characterisation of Bacterial Polysaccharides: Steps towards Single-Molecular Studies. *Carbohydr. Res.* **2003**, *338* (23), 2459–2475. https://doi.org/10.1016/j.carres.2003.07.007.

(50) Rao, V. S. Raghavendra Balaji, Petety V. Chandrasekaran, R. *Conformation of Carbohydrates*; Harwood: Amsterdam, 1998.

(51) Ogawa, Y.; Putaux, J. Recent Advances in Electron Microscopy of Carbohydrate Nanoparticles. *Front. Chem.* **2022**, *10* (February), 1–8. https://doi.org/10.3389/fchem.2022.835663.

(52) Ogawa, Y. Electron Microdiffraction Reveals the Nanoscale Twist Geometry of Cellulose Nanocrystals. *Nanoscale* **2019**, *11* (45), 21767–21774. https://doi.org/10.1039/C9NR06044H.

(53) Nannenga, B. L.; Gonen, T. The Cryo-EM Method Microcrystal Electron Diffraction (MicroED). *Nat. Methods* **2019**, *16* (5), 369–379. https://doi.org/10.1038/s41592-019-0395-x.

(54) Wang, J.; Nie, S. Application of Atomic Force Microscopy in Microscopic Analysis of Polysaccharide. *Trends Food Sci. Technol.* **2018**, No. August 2017, 0–1. https://doi.org/10.1016/j.tifs.2018.02.005.

(55) Rajendran, A.; Endo, M.; Sugiyama, H. State-of-the-Art High-Speed Atomic Force Microscopy for Investigation of Single-Molecular Dynamics of Proteins. *Chem. Rev.* **2014**, *114* (2), 1493–1520. https://doi.org/10.1021/cr300253x.

(56) Iwata, K.; Yamazaki, S.; Mutombo, P.; Hapala, P.; Ondráček, M.; Jelínek, P.; Sugimoto, Y. Chemical Structure Imaging of a Single Molecule by Atomic Force Microscopy at Room Temperature. *Nat. Commun.* **2015**, *6* (1), 7766. https://doi.org/10.1038/ncomms8766.

(57) Ando, T. High-Speed Atomic Force Microscopy and Its Future Prospects. *Biophys. Rev.* **2018**, *10* (2), 285–292. https://doi.org/10.1007/s12551-017-0356-5.

(58) Camesano, T. A.; Wilkinson, K. J. Single Molecule Study of Xanthan Conformation Using Atomic Force Microscopy. *Biomacromolecules* **2001**, *2* (4), 1184–1191. https://doi.org/10.1021/bm015555g.

(59) Schefer, L.; Adamcik, J.; Mezzenga, R. Unravelling Secondary Structure Changes on Individual Anionic Polysaccharide Chains by Atomic Force Microscopy. *Angew. Chem., Int. Ed.* **2014**, *53* (21), 5376–5379. https://doi.org/10.1002/anie.201402855.

(60) Schefer, L.; Usov, I.; Mezzenga, R. Anomalous Stiffening and Ion-Induced Coil-Helix Transition of Carrageenans under Monovalent Salt Conditions. *Biomacromolecules* **2015**, *16* (3), 985–991. https://doi.org/10.1021/bm501874k.

(61) Walls, A. C.; Tortorici, M. A.; Frenz, B.; Snijder, J.; Li, W.; Rey, F. A.; DiMaio, F.; Bosch, B.-J.; Veesler, D. Glycan Shield and Epitope Masking of a Coronavirus Spike Protein Observed by Cryo-Electron Microscopy. *Nat. Struct. Mol. Biol.* **2016**, *23* (10), 899–905. https://doi.org/10.1038/nsmb.3293.

(62) Rauschenbach, S.; Ternes, M.; Harnau, L.; Kern, K. Mass Spectrometry as a Preparative Tool for the Surface Science of Large Molecules. *Annu. Rev. Anal. Chem.* **2016**, *9*, 473–498. https://doi.org/10.1146/annurev-anchem-071015-041633.

(63) Wu, X.; Delbianco, M.; Anggara, K.; Michnowicz, T.; Pardo-Vargas, A.; Bharate, P.; Sen, S.; Pristl, M.; Rauschenbach, S.; Schlickum, U.; Abb, S.; Seeberger, P. H.; Kern, K. Imaging Single Glycans. *Nature* **2020**, *582* (7812), 375–378. https://doi.org/10.1038/s41586-020-2362-1.

(64) Anggara, K.; Zhu, Y.; Delbianco, M.; Rauschenbach, S.; Abb, S.; Seeberger, P. H.; Kern, K. Exploring the Molecular Conformation Space by Soft Molecule–Surface Collision. *J. Am. Chem. Soc.* **2020**, *142* (51), 21420–21427. https://doi.org/10.1021/jacs.0c09933.

(65) Freedberg, D. I.; Kwon, J. Solution NMR Structural Studies of Glycans. *Isr. J. Chem.* **2019**, *59* (11–12), 1039–1058. https://doi.org/10.1002/ijch.201900126.

(66) Battistel, M. D.; Azurmendi, H. F.; Yu, B.; Freedberg, D. I. NMR of Glycans: Shedding New Light on Old Problems. *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *79*, 48–68. https://doi.org/10.1016/j.pnmrs.2014.01.001.

(67) Yu, Y.; Delbianco, M. Conformational Studies of Oligosaccharides. *Chem. – A Eur. J.* **2020**, *26* (44), 9814–9825. https://doi.org/10.1002/chem.202001370.

(68) Claridge, T. D. W. *High-Resolution NMR Techniques in Organic Chemistry*, 3rd ed.; Elsevier, 2016. https://doi.org/10.1016/C2015-0-04654-8.

(69) Prestegard, J. H.; Bougault, C. M.; Kishore, A. I. Residual Dipolar Couplings in Structure Determination of Biomolecules. *Chem. Rev.* **2004**, *104* (8), 3519–3540. https://doi.org/10.1021/cr030419i.

(70) Battistel, M. D.; Pendrill, R.; Widmalm, G.; Freedberg, D. I. Direct Evidence for Hydrogen Bonding in Glycans: A Combined NMR and Molecular Dynamics Study. *J. Phys. Chem. B* **2013**, *117* (17), 4860–4869. https://doi.org/10.1021/jp400402b.

(71) Battistel, M. D.; Shangold, M.; Trinh, L.; Shiloach, J.; Freedberg, D. I. Evidence for Helical Structure in a Tetramer of A2-8 Sialic Acid: Unveiling a Structural Antigen. *J. Am. Chem. Soc.* **2012**, *134* (26), 10717–10720. https://doi.org/10.1021/ja300624j.

(72) Almond, A.; Bunkenborg, J.; Franch, T.; Gotfredsen, C. H.; Duus, J. Ø. Comparison of Aqueous Molecular Dynamics with NMR Relaxation and Residual Dipolar Couplings Favors Internal Motion in a Mannose Oligosaccharide. *J. Am. Chem. Soc.* **2001**, *123* (20), 4792–4802. https://doi.org/10.1021/ja0025696.

(73) Rönnols, J.; Engström, O.; Schnupf, U.; Säwén, E.; Brady, J. W.; Widmalm, G. Interresidual Hydrogen Bonding in Carbohydrates Unraveled by NMR Spectroscopy and Molecular Dynamics Simulations. *ChemBioChem* **2019**, cbic.201900301. https://doi.org/10.1002/cbic.201900301.

(74) Canales, A.; Jiménez-Barbero, J.; Martín-Pastor, M. Review: Use of Residual Dipolar Couplings to Determine the Structure of Carbohydrates. *Magn. Reson. Chem.* **2012**, *50*, S80–S85. https://doi.org/10.1002/mrc.3888.

(75) Dupree, R.; Simmons, T. J.; Mortimer, J. C.; Patel, D.; Iuga, D.; Brown, S. P.; Dupree, P. Probing the Molecular Architecture of *Arabidopsis Thaliana* Secondary Cell Walls Using Two- and Three-Dimensional ^{13}C Solid State Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **2015**, *54* (14), 2335–2345. https://doi.org/10.1021/bi501552k.

(76) Canales, Á.; Mallagaray, Á.; Berbís, M. Á.; Navarro-Vázquez, A.; Domínguez, G.; Cañada, F. J.; André, S.; Gabius, H.-J.; Pérez-Castells, J.; Jiménez-Barbero, J. Lanthanide-Chelating Carbohydrate Conjugates Are Useful Tools To Characterize Carbohydrate Conformation in Solution and Sensitive Sensors to Detect Carbohydrate–Protein Interactions. *J. Am. Chem. Soc.* **2014**, *136* (22), 8011–8017. https://doi.org/10.1021/ja502406x.

(77) Canales, A.; Mallagaray, A.; Pérez-Castells, J.; Boos, I.; Unverzagt, C.; André, S.; Gabius, H.-J.; Cañada, F. J.; Jiménez-Barbero, J. Breaking Pseudo-Symmetry in Multiantennary Complex N-Glycans Using Lanthanide-Binding Tags and NMR Pseudo-Contact Shifts. *Angew. Chem., Int. Ed.* **2013**, *52* (51), 13789–13793. https://doi.org/10.1002/anie.201307845.

(78) Martín-Pastor, M.; Canales, A.; Corzana, F.; Asensio, J. L.; Jiménez-Barbero, J. Limited Flexibility of Lactose Detected from Residual Dipolar Couplings Using Molecular Dynamics Simulations and Steric Alignment Methods. *J. Am. Chem. Soc.* **2005**, *127* (10), 3589–3595. https://doi.org/10.1021/ja043445m.

(79) Zhang, W.; Meredith, R.; Yoon, M.-K.; Wang, X.; Woods, R. J.; Carmichael, I.; Serianni, A. S. Synthesis and O -Glycosidic Linkage Conformational Analysis of 13 C-Labeled Oligosaccharide Fragments of an Antifreeze Glycolipid. *J. Org. Chem.* **2019**, *84* (4), 1706–1724. https://doi.org/10.1021/acs.joc.8b01411.

(80) Toukach, F. V.; Ananikov, V. P. Recent Advances in Computational Predictions of NMR Parameters for the Structure Elucidation of Carbohydrates: Methods and Limitations. *Chem. Soc. Rev.* **2013**, *42* (21), 8376. https://doi.org/10.1039/c3cs60073d.

(81) Capron, I.; Brigand, G.; Muller, G. About the Native and Renatured Conformation of Xanthan Exopolysaccharide. *Polymer* **1997**, *38* (21), 5289–5295. https://doi.org/10.1016/S0032-3861(97)00079-7.

(82) Matsuo, K.; Gekko, K. Synchrotron-Radiation Vacuum-Ultraviolet Circular-Dichroism Spectroscopy for Characterizing the Structure of Saccharides; 2018; pp 101–117. https://doi.org/10.1007/978-981-13-2158-0_6.

(83) Delbianco, M.; Kononov, A.; Poveda, A.; Yu, Y.; Diercks, T.; Jiménez-Barbero, J.; Seeberger, P. H. Well-Defined Oligo- and Polysaccharides as Ideal Probes for Structural Studies. *J. Am. Chem. Soc.* **2018**, *140* (16), 5421–5426. https://doi.org/10.1021/jacs.8b00254.

(84) Armstrong, Z.; Withers, S. G. Synthesis of Glycans and Glycopolymers Through Engineered Enzymes. *Biopolymers* **2013**, *99* (10), 666–674. https://doi.org/10.1002/bip.22335.

(85) Smith, P. J.; Ortiz-Soto, M. E.; Roth, C.; Barnes, W. J.; Seibel, J.; Urbanowicz, B. R.; Pfrengle, F. Enzymatic Synthesis of Artificial Polysaccharides. *ACS Sustain. Chem. Eng.* **2020**, *8* (32), 11853–11871. https://doi.org/10.1021/acssuschemeng.0c03622.

(86) Perugino, G.; Cobucci-Ponzano, B.; Rossi, M.; Moracci, M. Recent Advances in the Oligosaccharide Synthesis Promoted by Catalytically Engineered Glycosidases. *Adv. Synth. Catal.* **2005**, *347* (7–8), 941–950. https://doi.org/10.1002/adsc.200505070.

(87) *Enzymatic Polymerization towards Green Polymer Chemistry*; Kobayashi, S., Uyama, H., Kadokawa, J., Eds.; Green Chemistry and Sustainable Technology; Springer Singapore: Singapore, 2019. https://doi.org/10.1007/978-981-13-3813-7.

(88) Faijes, M.; Planas, A. In Vitro Synthesis of Artificial Polysaccharides by Glycosidases and Glycosynthases. *Carbohydr. Res.* **2007**, *342* (12–13), 1581–1594. https://doi.org/10.1016/j.carres.2007.06.015.

(89) Shoda, S.; Uyama, H.; Kadokawa, J.; Kimura, S.; Kobayashi, S. Enzymes as Green Catalysts for Precision Macromolecular Synthesis. *Chem. Rev.* **2016**, *116* (4), 2307–2413. https://doi.org/10.1021/acs.chemrev.5b00472.

(90) Kobayashi, S.; Makino, A. Enzymatic Polymer Synthesis : An Opportunity for Green Polymer Chemistry. *Chem. Rev.* **2009**, *109*, 5288–5353. https://doi.org/10.1021/cr900165z.

(91) Kadokawa, J. Precision Polysaccharide Synthesis Catalyzed by Enzymes. *Chem. Rev.* **2011**, *111* (7), 4308–4345. https://doi.org/10.1021/cr100285v.

(92) Xiao, R.; Grinstaff, M. W. Chemical Synthesis of Polysaccharides and Polysaccharide Mimetics. *Prog. Polym. Sci.* **2017**, *74*, 78–116. https://doi.org/10.1016/J.PROGPOLYMSCI.2017.07.009.

(93) Li, W.; McArthur, J. B.; Chen, X. Strategies for Chemoenzymatic Synthesis of Carbohydrates. *Carbohydr. Res.* **2019**, *472*, 86–97. https://doi.org/10.1016/j.carres.2018.11.014.

(94) Mackenzie, L. F.; Wang, Q.; Warren, R. A. J.; Withers, S. G. Glycosynthases: Mutant Glycosidases for Oligosaccharide Synthesis. *J. Am. Chem. Soc.* **1998**, *120* (22), 5583–5584. https://doi.org/10.1021/ja980833d.

(95) Danby, P. M.; Withers, S. G. Advances in Enzymatic Glycoside Synthesis. *ACS Chem. Biol.* **2016**, *11* (7), 1784–1794. https://doi.org/10.1021/acschembio.6b00340.

(96) Hayes, M.; Pietruszka, J. Synthesis of Glycosides by Glycosynthases. *Molecules* **2017**, *22* (9), 1434. https://doi.org/10.3390/molecules22091434.

(97) Cobucci-Ponzano, B.; Moracci, M. Glycosynthases as Tools for the Production of Glycan Analogs of Natural Products. *Nat. Prod. Rep.* **2012**, *29* (6), 697. https://doi.org/10.1039/c2np20032e.

(98) Liu, F.; Chen, H.; Armstrong, Z.; Withers, S. G. Azido Groups Hamper Glycan Acceptance by Carbohydrate Processing Enzymes. *ACS Cent. Sci.* **2022**.

https://doi.org/10.1021/acscentsci.1c01172.

(99) Wang, L. X.; Huang, W. Enzymatic Transglycosylation for Glycoconjugate Synthesis. *Curr. Opin. Chem. Biol.* **2009**, *13* (5–6), 592–600. https://doi.org/10.1016/j.cbpa.2009.08.014.

(100) Tokatly, A. I.; Vinnitskiy, D. Z.; Ustuzhanina, N. E.; Nifantiev, N. E. Protecting Groups as a Factor of Stereocontrol in Glycosylation Reactions. *Russ. J. Bioorganic Chem.* **2021**, *47* (1), 53–70. https://doi.org/10.1134/s1068162021010258.

(101) Kaeothip, S.; Demchenko, A. V. Expeditious Oligosaccharide Synthesis via Selective, Semi-Orthogonal, and Orthogonal Activation. *Carbohydr. Res.* **2011**, *346* (12), 1371–1388. https://doi.org/10.1016/j.carres.2011.05.004.

(102) Jeanneret, R. A.; Johnson, S. E.; Galan, M. C. Conformationally Constrained Glycosyl Donors as Tools to Control Glycosylation Outcomes. *J. Org. Chem.* **2020**, *85* (24), acs.joc.0c02045. https://doi.org/10.1021/acs.joc.0c02045.

(103) Pardo-Vargas, A.; Delbianco, M.; Seeberger, P. H. Automated Glycan Assembly as an Enabling Technology. *Curr. Opin. Chem. Biol.* **2018**, *46*, 48–55. https://doi.org/10.1016/j.cbpa.2018.04.007.

(104) Panza, M.; Pistorio, S. G.; Stine, K. J.; Demchenko, A. V. Automated Chemical Oligosaccharide Synthesis: Novel Approach to Traditional Challenges. *Chem. Rev.* **2018**, *118* (17), 8105–8150. https://doi.org/10.1021/acs.chemrev.8b00051.

(105) Hsu, C.-H.; Hung, S.-C.; Wu, C.-Y.; Wong, C.-H. Toward Automated Oligosaccharide Synthesis. *Angew. Chem., Int. Ed.* **2011**, *50* (50), 11872–11923. https://doi.org/10.1002/anie.201100125.

(106) Panza, M.; Stine, K. J.; Demchenko, A. V. HPLC-Assisted Automated Oligosaccharide Synthesis: The Implementation of the Two-Way Split Valve as a Mode of Complete Automation. *Chem. Commun.* **2020**, *56* (9), 1333–1336. https://doi.org/10.1039/C9CC08876H.

(107) Guberman, M.; Seeberger, P. H. Automated Glycan Assembly: A Perspective. *J. Am. Chem. Soc.* **2019**, *141* (14), 5581–5592. https://doi.org/10.1021/jacs.9b00638.

(108) Seeberger, P. H. The Logic of Automated Glycan Assembly. *Acc. Chem. Res.* **2015**, *48* (5), 1450–1463. https://doi.org/10.1021/ar5004362.

(109) Hurevich, M.; Kandasamy, J.; Ponnappa, B. M.; Collot, M.; Kopetzki, D.; McQuade, D. T.; Seeberger, P. H. Continuous Photochemical Cleavage of Linkers for Solid-Phase Synthesis. *Org. Lett.* **2014**, *16* (6), 1794–1797. https://doi.org/10.1021/ol500530q.

(110) Joseph, A. A.; Pardo-Vargas, A.; Seeberger, P. H. Total Synthesis of Polysaccharides by Automated Glycan Assembly. *J. Am. Chem. Soc.* **2020**, *142* (19), 8561–8564. https://doi.org/10.1021/jacs.0c00751.

(111) Johnson, J. A.; Lu, Y. Y.; Van Deventer, J. A.; Tirrell, D. A. Residue-Specific Incorporation of Non-Canonical Amino Acids into Proteins: Recent Developments and Applications. *Curr. Opin. Chem. Biol.* **2010**, *14* (6), 774–780. https://doi.org/10.1016/j.cbpa.2010.09.013.

(112) Yoder, N. C.; Kumar, K. Fluorinated Amino Acids in Protein Design and Engineering. *Chem. Soc. Rev.* **2002**, *31* (6), 335–341. https://doi.org/10.1039/b201097f.

(113) Prescher, J. A.; Bertozzi, C. R. Chemistry in Living Systems. *Nat. Chem. Biol.* **2005**, *1* (1), 13–21. https://doi.org/10.1038/nchembio0605-13.

(114) Link, A. J.; Mock, M. L.; Tirrell, D. A. Non-Canonical Amino Acids in Protein Engineering. *Curr. Opin. Biotechnol.* **2003**, *14* (6), 603–609. https://doi.org/10.1016/j.copbio.2003.10.011.

(115) Moschner, J.; Stulberg, V.; Fernandes, R.; Huhmann, S.; Leppkes, J.; Koksch, B. Approaches to Obtaining Fluorinated α-Amino Acids. *Chem. Rev.* **2019**, *119* (18), 10718–10801. https://doi.org/10.1021/acs.chemrev.9b00024.

(116) O'hagan, D. Understanding Organofluorine Chemistry. An Introduction to the C–F Bond. *Chem. Soc. Rev.* **2008**, *37* (2), 308–319. https://doi.org/10.1039/b711844a.

(117) Salwiczek, M.; Nyakatura, E. K.; Gerling, U. I. M.; Ye, S.; Koksch, B. Fluorinated Amino Acids: Compatibility with Native Protein Structures and Effects on Protein–Protein Interactions. *Chem. Soc. Rev.* **2012**, *41* (6), 2135–2171. https://doi.org/10.1039/C1CS15241F.

(118) Marsh, E. N. G. Fluorinated Proteins: From Design and Synthesis to Structure and Stability. *Acc. Chem. Res.* **2014**, *47* (10), 2878–2886. https://doi.org/10.1021/ar500125m.

(119) Fallas, J. A.; Leary, L. E. R. O.; Hartgerink, J. D. Synthetic Collagen Mimics: Self-Assembly of Homotrimers, Heterotrimers and Higher Order Structures. *Chem. Soc. Rev.* **2010**, *39* (9), 3510–3527. https://doi.org/10.1039/b919455j.

(120) Shoulders, M. D.; Satyshur, K. A.; Forest, K. T.; Raines, R. T. Stereoelectronic and Steric Effects in Side Chains Preorganize a Protein Main Chain. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (2), 559–564. https://doi.org/10.1073/pnas.0909592107.

(121) Sloand, J. N.; Miller, M. A.; Medina, S. H. Fluorinated Peptide Biomaterials. *Pept. Sci.* **2021**, *113* (2). https://doi.org/10.1002/pep2.24184.

(122) Marsh, E. N. G.; Suzuki, Y. Using 19 F NMR to Probe Biological Interactions of Proteins and Peptides. *ACS Chem. Biol.* **2014**, *9* (6), 1242–1250. https://doi.org/10.1021/cb500111u.

(123) Chen, H.; Viel, S.; Ziarelli, F.; Peng, L. 19F NMR: A Valuable Tool for Studying Biological Events. *Chem. Soc. Rev.* **2013**, *42* (20), 7971. https://doi.org/10.1039/c3cs60129c.

(124) Smith, M. J.; Marshall, C. B.; Theillet, F.-X.; Binolfi, A.; Selenko, P.; Ikura, M. Real-Time NMR Monitoring of Biological Activities in Complex Physiological Environments. *Curr. Opin. Struct. Biol.* **2015**, *32*, 39–47. https://doi.org/10.1016/j.sbi.2015.02.003.

(125) Serber, Z.; Dötsch, V. In-Cell NMR Spectroscopy †. *Biochemistry* **2001**, *40* (48), 14317–14323. https://doi.org/10.1021/bi011751w.

(126) Rabideau, A. E.; Pentelute, B. L. A D-Amino Acid at the N-Terminus of a Protein Abrogates Its Degradation by the N-End Rule Pathway. *ACS Cent. Sci.* **2015**, *1* (8), 423–430. https://doi.org/10.1021/acscentsci.5b00308.

(127) Simon, M. D.; Maki, Y.; Vinogradov, A. A.; Zhang, C.; Yu, H.; Lin, Y.-S.; Kajihara, Y.; Pentelute, B. L. D-Amino Acid Scan of Two Small Proteins. *J. Am. Chem. Soc.* **2016**, *138* (37), 12099–12111. https://doi.org/10.1021/jacs.6b03765.

(128) Mong, S. K.; Cochran, F. V.; Yu, H.; Graziano, Z.; Lin, Y. S.; Cochran, J. R.; Pentelute, B. L. Heterochiral Knottin Protein: Folding and Solution Structure. *Biochemistry* **2017**, *56* (43), 5720–5725. https://doi.org/10.1021/acs.biochem.7b00722.

(129) Yeates, T. O.; Kent, S. B. H. Racemic Protein Crystallography. *Annu. Rev. Biophys.* **2012**, *41* (1), 41–61. https://doi.org/10.1146/annurev-biophys-050511-102333.

(130) Volpatti, L. R.; Vendruscolo, M.; Dobson, C. M.; Knowles, T. P. J. A Clear View of Polymorphism, Twist, and Chirality in Amyloid Fibril Formation. *ACS Nano* **2013**, *7* (12), 10443–10448. https://doi.org/10.1021/nn406121w.

(131) Harper, J. D.; Lieber, C. M.; Lansbury, P. T. Atomic Force Microscopic Imaging of Seeded Fibril Formation and Fibril Branching by the Alzheimer's Disease Amyloid-β Protein. *Chem. Biol.* **1997**, *4* (12), 951–959. https://doi.org/10.1016/S1074-5521(97)90303-3.

(132) Nyström, G.; Arcari, M.; Mezzenga, R. Confinement-Induced Liquid Crystalline Transitions in Amyloid Fibril Cholesteric Tactoids. *Nat. Nanotechnol.* **2018**, *13* (4), 330–336. https://doi.org/10.1038/s41565-018-0071-9.

(133) Wang, M.; Zhou, P.; Wang, J.; Zhao, Y.; Ma, H.; Lu, J. R.; Xu, H. Left or Right: How Does Amino

Acid Chirality Affect the Handedness of Nanostructures Self-Assembled from Short Amphiphilic Peptides? *J. Am. Chem. Soc.* **2017**, *139* (11), 4185–4194. https://doi.org/10.1021/jacs.7b00847.

(134) Lara, C.; Reynolds, N. P.; Berryman, J. T.; Xu, A.; Zhang, A.; Mezzenga, R. ILQINS Hexapeptide, Identified in Lysozyme Left-Handed Helical Ribbons and Nanotubes, Forms Right-Handed Helical Ribbons and Crystals. *J. Am. Chem. Soc.* **2014**, *136* (12), 4732–4739. https://doi.org/10.1021/ja500445z.

(135) Clover, T. M.; O'Neill, C. L.; Appavu, R.; Lokhande, G.; Gaharwar, A. K.; Posey, A. E.; White, M. A.; Rudra, J. S. Self-Assembly of Block Heterochiral Peptides into Helical Tapes. *J. Am. Chem. Soc.* **2020**, *142* (47), 19809–19813. https://doi.org/10.1021/jacs.9b09755.

(136) Swanekamp, R. J.; DiMaio, J. T. M.; Bowerman, C. J.; Nilsson, B. L. Coassembly of Enantiomeric Amphipathic Peptides into Amyloid-Inspired Rippled β-Sheet Fibrils. *J. Am. Chem. Soc.* **2012**, *134* (12), 5556–5559. https://doi.org/10.1021/ja301642c.

(137) Nagy-Smith, K.; Beltramo, P. J.; Moore, E.; Tycko, R.; Furst, E. M.; Schneider, J. P. Molecular, Local, and Network-Level Basis for the Enhanced Stiffness of Hydrogel Networks Formed from Coassembled Racemic Peptides: Predictions from Pauling and Corey. *ACS Cent. Sci.* **2017**, *3* (6), 586–597. https://doi.org/10.1021/acscentsci.7b00115.

(138) Nagy, K. J.; Giano, M. C.; Jin, A.; Pochan, D. J.; Schneider, J. P. Enhanced Mechanical Rigidity of Hydrogels Formed from Enantiomeric Peptide Assemblies. *J. Am. Chem. Soc.* **2011**, *133* (38), 14975–14977. https://doi.org/10.1021/ja206742m.

(139) Bera, S.; Xue, B.; Rehak, P.; Jacoby, G.; Ji, W.; Shimon, L. J. W.; Beck, R.; Král, P.; Cao, Y.; Gazit, E. Self-Assembly of Aromatic Amino Acid Enantiomers into Supramolecular Materials of High Rigidity. *ACS Nano* **2020**, *14* (2), 1694–1706. https://doi.org/10.1021/acsnano.9b07307.

(140) Tamburrini, A.; Colombo, C.; Bernardi, A. Design and Synthesis of Glycomimetics: Recent Advances. *Med. Res. Rev.* **2019**, med.21625. https://doi.org/10.1002/med.21625.

(141) Uhrig, M. L.; Lantaño, B.; Postigo, A. Synthetic Strategies for Fluorination of Carbohydrates. *Org. Biomol. Chem.* **2019**, *17* (21), 5173–5189. https://doi.org/10.1039/C9OB00405J.

(142) Lo, H.-J.; Krasnova, L.; Dey, S.; Cheng, T.; Liu, H.; Tsai, T.-I.; Wu, K. B.; Wu, C.-Y.; Wong, C.-H. Synthesis of Sialidase-Resistant Oligosaccharide and Antibody Glycoform Containing A2,6-Linked 3F Ax -Neu5Ac. *J. Am. Chem. Soc.* **2019**, *141* (16), 6484–6488. https://doi.org/10.1021/jacs.9b01991.

(143) Linclau, B.; Ardá, A.; Reichardt, N.-C.; Sollogoub, M.; Unione, L.; Vincent, S. P.; Jiménez-Barbero, J. Fluorinated Carbohydrates as Chemical Probes for Molecular Recognition Studies. Current Status and Perspectives. *Chem. Soc. Rev.* **2020**. https://doi.org/10.1039/C9CS00099B.

(144) Baumann, A.; Marchner, S.; Daum, M.; Hoffmann-Röder, A. Synthesis of Fluorinated Leishmania Cap Trisaccharides for Diagnostic Tool and Vaccine Development. *Eur. J. Org. Chem.* **2018**, *2018* (27–28), 3803–3815. https://doi.org/10.1002/ejoc.201800384.

(145) Yang, F.; Zheng, X.-J.; Huo, C.-X.; Wang, Y.; Zhang, Y.; Ye, X.-S. Enhancement of the Immunogenicity of Synthetic Carbohydrate Vaccines by Chemical Modifications of STn Antigen. *ACS Chem. Biol.* **2011**, *6* (3), 252–259. https://doi.org/10.1021/cb100287q.

(146) Martínez, J. D.; Manzano, A. I.; Calviño, E.; Diego, A. De; Rodriguez de Francisco, B.; Romanò, C.; Oscarson, S.; Millet, O.; Gabius, H.-J.; Jiménez-Barbero, J.; Cañada, F. J. Fluorinated Carbohydrates as Lectin Ligands: Simultaneous Screening of a Monosaccharide Library and Chemical Mapping by 19 F NMR Spectroscopy. *J. Org. Chem.* **2020**, *85* (24), 16072–16081. https://doi.org/10.1021/acs.joc.0c01830.

(147) Unione, L.; Alcalá, M.; Echeverria, B.; Serna, S.; Ardá, A.; Franconetti, A.; Cañada, F. J.; Diercks, T.; Reichardt, N.; Jiménez-Barbero, J. Fluoroacetamide Moieties as NMR Spectroscopy Probes for the Molecular Recognition of GlcNAc-Containing Sugars: Modulation of the CH-π Stacking

Interactions by Different Fluorination Patterns. *Chem. - A Eur. J* **2017**, *23* (16), 3957–3965. https://doi.org/10.1002/chem.201605573.

(148) Valverde, P.; Quintana, J. I.; Santos, J. I.; Ardá, A.; Jiménez-Barbero, J. Novel NMR Avenues to Explore the Conformation and Interactions of Glycans. *ACS Omega* **2019**, *4* (9), 13618–13630. https://doi.org/10.1021/acsomega.9b01901.

(149) Yuasa, H.; Hashimoto, H. Chemical Synthesis of Bioactive Oligosaccharides. Recent Advances in the Development of Unnatural Oligosaccharides. Confornration and Bioactivity. *Trends Glycosci. Glycotechnol.* **2001**, *13* (69), 31–55. https://doi.org/10.4052/tigg.13.31.

(150) Montero, E.; García-Herrero, A.; Asensio, J. L.; Hirai, K.; Ogawa, S.; Santoyo-González, F.; Cañada, F. J.; Jiménez-Barbero, J. The Conformational Behaviour of Non-Hydrolizable Lactose Analogues: The Thioglycoside, Carbaglycoside, and Carba-Iminoglycoside Cases. *Eur. J. Org. Chem.* **2000**, *2000* (10), 1945–1952. https://doi.org/10.1002/(SICI)1099-0690(200005)2000:10<1945::AID-EJOC1945>3.0.CO;2-G.

(151) Arjona, O.; Gómez, A. M.; López, J. C.; Plumet, J. Synthesis and Conformational and Biological Aspects of Carbasugars. *Chem. Rev.* **2007**, *107* (5), 1919–2036. https://doi.org/10.1021/cr0203701.

(152) Yang, Y.; Yu, B. Recent Advances in the Chemical Synthesis of C -Glycosides. *Chem. Rev.* **2017**, *117* (19), 12281–12356. https://doi.org/10.1021/acs.chemrev.7b00234.

(153) Danglad-Flores, J.; Leichnitz, S.; Sletten, E. T.; Abragam Joseph, A.; Bienert, K.; Le Mai Hoang, K.; Seeberger, P. H. Microwave-Assisted Automated Glycan Assembly. *J. Am. Chem. Soc.* **2021**, *143* (23), 8893–8901. https://doi.org/10.1021/jacs.1c03851.

(154) Li, T.; Liu, L.; Wei, N.; Yang, J.-Y. Y.; Chapla, D. G.; Moremen, K. W.; Boons, G.-J. J. An Automated Platform for the Enzyme-Mediated Assembly of Complex Oligosaccharides. *Nat. Chem.* **2019**, *11* (3), 229–236. https://doi.org/10.1038/s41557-019-0219-8.

(155) Zhu, Y.; Delbianco, M.; Seeberger, P. H. Automated Assembly of Starch and Glycogen Polysaccharides. *J. Am. Chem. Soc.* **2021**, *143* (26), 9758–9768. https://doi.org/10.1021/jacs.1c02188.

(156) Nowick, J. S.; Insaf, S. The Propensities of Amino Acids to Form Parallel β-Sheets. *J. Am. Chem. Soc.* **1997**, *119* (45), 10903–10908. https://doi.org/10.1021/ja972074p.

(157) Schenck, H. L.; Gellman, S. H. Use of a Designed Triple-Stranded Antiparallel β-Sheet To Probe β-Sheet Cooperativity in Aqueous Solution. *J. Am. Chem. Soc.* **1998**, *120* (19), 4869–4870. https://doi.org/doi.org/10.1021/ja973984+.

(158) Cordes, M. H. J.; Davidson, A. R.; Sauer, R. T. Sequence Space, Folding and Protein Design. *Curr. Opin. Struct. Biol.* **1996**, *6* (1), 3–10. https://doi.org/10.1016/S0959-440X(96)80088-1.

(159) Kortemme, T.; Ramı□ rez-Alvarado, M.; Serrano, L. Design of a 20-Amino Acid, Three-Stranded β-Sheet Protein. *Science* **1998**, *281* (5374), 253–256. https://doi.org/10.1126/science.281.5374.253.

(160) Blanco, F. J.; Rivas, G.; Serrano, L. A Short Linear Peptide That Folds into a Native Stable β-Hairpin in Aqueous Solution. *Nat. Struct. Mol. Biol.* **1994**, *1* (9), 584–590. https://doi.org/10.1038/nsb0994-584.

(161) Caruthers, M. H. Gene Synthesis Machines: DNA Chemistry and Its Uses. *Science* **1985**, *230* (4723), 281–285. https://doi.org/10.1126/science.3863253.

(162) Gellman, S. H. Foldamers: A Manifesto. *Acc. Chem. Res.* **1998**, *31* (4), 173–180. https://doi.org/10.1021/ar960298r.

(163) Guichard, G.; Huc, I. Synthetic Foldamers. *Chem. Commun.* **2011**, *47* (21), 5933. https://doi.org/10.1039/c1cc11137j.

(164) Goodman, C. M.; Choi, S.; Shandler, S.; DeGrado, W. F. Foldamers as Versatile Frameworks for

the Design and Evolution of Function. *Nat. Chem. Biol.* **2007**, *3* (5), 252–262. https://doi.org/10.1038/nchembio876.

(165) George, K. L.; Horne, W. S. Foldamer Tertiary Structure through Sequence-Guided Protein Backbone Alteration. *Acc. Chem. Res.* **2018**, *51* (5), 1220–1228. https://doi.org/10.1021/acs.accounts.8b00048.

(166) Girvin, Z. C.; Gellman, S. H. Foldamer Catalysis. *J. Am. Chem. Soc.* **2020**, *142* (41), 17211–17223. https://doi.org/10.1021/jacs.0c07347.

(167) Yoo, S. H.; Lee, H.-S. Foldectures: 3D Molecular Architectures from Self-Assembly of Peptide Foldamers. *Acc. Chem. Res.* **2017**, *50* (4), 832–841. https://doi.org/10.1021/acs.accounts.6b00545.

(168) Girvin, Z. C.; Andrews, M. K.; Liu, X.; Gellman, S. H. Foldamer-Templated Catalysis of Macrocycle Formation. *Science* **2019**, *366* (6472), 1528–1531. https://doi.org/10.1126/science.aax7344.

(169) Andrews, M. K.; Liu, X.; Gellman, S. H. Tailoring Reaction Selectivity by Modulating a Catalytic Diad on a Foldamer Scaffold. *J. Am. Chem. Soc.* **2022**, *144* (5), 2225–2232. https://doi.org/10.1021/jacs.1c11542.

(170) Purser, S.; Moore, P. R.; Swallow, S.; Gouverneur, V. Fluorine in Medicinal Chemistry. *Chem. Soc. Rev.* **2008**, *37* (2), 320–330. https://doi.org/10.1039/B610213C.

(171) Biffinger, J. C.; Kim, H. W.; DiMagno, S. G. The Polar Hydrophobicity of Fluorinated Compounds. *ChemBioChem* **2004**, *5* (5), 622–627. https://doi.org/10.1002/cbic.200300910.

(172) Zheng, H.; Gao, J. Highly Specific Heterodimerization Mediated by Quadrupole Interactions. *Angew. Chem., Int. Ed.* **2010**, *49* (46), 8635–8639. https://doi.org/10.1002/anie.201002860.

(173) Cobb, S. L.; Murphy, C. D. 19F NMR Applications in Chemical Biology. *J. Fluor. Chem.* **2009**, *130* (2), 132–143. https://doi.org/10.1016/j.jfluchem.2008.11.003.

(174) Tsukiji, S.; Miyagawa, M.; Takaoka, Y.; Tamura, T.; Hamachi, I. Ligand-Directed Tosyl Chemistry for Protein Labeling in Vivo. *Nat. Chem. Biol.* **2009**, *5* (5), 341–343. https://doi.org/10.1038/nchembio.157.

(175) Kieser, T. J.; Santschi, N.; Nowack, L.; Axer, A.; Kehr, G.; Albrecht, S.; Gilmour, R. Total Chemical Syntheses of the GM 3 and F–GM 3 Ganglioside Epitopes and Comparative Pre-Clinical Evaluation for Non-Invasive Imaging of Oligodendrocyte Differentiation. *ACS Chem. Neurosci.* **2020**, *11* (14), 2129–2136. https://doi.org/10.1021/acschemneuro.0c00319.

(176) Keenan, T.; Parmeggiani, F.; Malassis, J.; Fontenelle, C. Q.; Vendeville, J.-B.; Offen, W.; Both, P.; Huang, K.; Marchesi, A.; Heyam, A.; Young, C.; Charnock, S. J.; Davies, G. J.; Linclau, B.; Flitsch, S. L.; Fascione, M. A. Profiling Substrate Promiscuity of Wild-Type Sugar Kinases for Multi-Fluorinated Monosaccharides. *Cell Chem. Biol.* **2020**. https://doi.org/10.1016/j.chembiol.2020.06.005.

(177) Council, C. E.; Kilpin, K. J.; Gusthart, J. S.; Allman, S. A.; Linclau, B.; Lee, S. S. Enzymatic Glycosylation Involving Fluorinated Carbohydrates. *Org. Biomol. Chem.* **2020**. https://doi.org/10.1039/D0OB00436G.

(178) Matei, E.; André, S.; Glinschert, A.; Infantino, A. S.; Oscarson, S.; Gabius, H.-J.; Gronenborn, A. M. Fluorinated Carbohydrates as Lectin Ligands: Dissecting Glycan-Cyanovirin Interactions by Using 19 F NMR Spectroscopy. *Chem. - A Eur. J* **2013**, *19* (17), 5364–5374. https://doi.org/10.1002/chem.201204070.

(179) Calle, L. P.; Echeverria, B.; Franconetti, A.; Serna, S.; Fernández-Alonso, M. C.; Diercks, T.; Cañada, F. J.; Ardá, A.; Reichardt, N.-C.; Jiménez-Barbero, J. Monitoring Glycan-Protein Interactions by NMR Spectroscopic Analysis: A Simple Chemical Tag That Mimics Natural CH-π Interactions. *Chem. - A Eur. J* **2015**, *21* (32), 11408–11416. https://doi.org/10.1002/chem.201501248.

(180) Gimeno, A.; Valverde, P.; Ardá, A.; Jiménez-Barbero, J. Glycan Structures and Their Interactions

with Proteins. A NMR View. *Curr. Opin. Struct. Biol.* **2020**, *62*, 22–30. https://doi.org/10.1016/j.sbi.2019.11.004.

(181) Ardá, A.; Jiménez-Barbero, J. The Recognition of Glycans by Protein Receptors. Insights from NMR Spectroscopy. *Chem. Commun.* **2018**, *54* (38), 4761–4769. https://doi.org/10.1039/C8CC01444B.

(182) Allman, S. A.; Jensen, H. H.; Vijayakrishnan, B.; Garnett, J. A.; Leon, E.; Liu, Y.; Anthony, D. C.; Sibson, N. R.; Feizi, T.; Matthews, S.; Davis, B. G. Potent Fluoro-Oligosaccharide Probes of Adhesion in Toxoplasmosis. *ChemBioChem* **2009**, *10* (15), 2522–2529. https://doi.org/10.1002/cbic.200900425.

(183) Richards, S.-J.; Gibson, M. I.; Fascione, M. A.; Linclau, B.; Galan, M. C.; Baker, A. N.; Ledru, H.; Webster, C. S.; Council, C. E.; Walker, M.; Chidwick, H.; Wheatley, D. E.; Vendeville, J.-B.; Keenan, T.; Budhadev, D. Introducing Affinity and Selectivity into Galectin-Targeting Nanoparticles with Fluorinated Glycan Ligands. *Chem. Sci.* **2021**. https://doi.org/10.1039/D0SC05360K.

(184) Valverde, P.; Delgado, S.; Martínez, J. D.; Vendeville, J.-B.; Malassis, J.; Linclau, B.; Reichardt, N.-C.; Cañada, F. J.; Jiménez-Barbero, J.; Ardá, A. Molecular Insights into DC-SIGN Binding to Self-Antigens: The Interaction with the Blood Group A/B Antigens. *ACS Chem. Biol.* **2019**, *14* (7), 1660–1671. https://doi.org/10.1021/acschembio.9b00458.

(185) Nishiyama, Y.; Langan, P.; Chanzy, H. Crystal Structure and Hydrogen-Bonding System in Cellulose Iβ from Synchrotron X-Ray and Neutron Fiber Diffraction. *J. Am. Chem. Soc.* **2002**, *124* (31), 9074–9082. https://doi.org/10.1021/ja0257319.

(186) Langan, P.; Nishiyama, Y.; Chanzy, H. X-Ray Structure of Mercerized Cellulose II at 1 Å Resolution. *Biomacromolecules* **2001**, *2* (2), 410–416. https://doi.org/10.1021/bm005612q.

(187) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. Crystal Structure and Hydrogen Bonding System in Cellulose I α from Synchrotron X-Ray and Neutron Fiber Diffraction. *J. Am. Chem. Soc.* **2003**, *125* (47), 14300–14306. https://doi.org/10.1021/ja037055w.

(188) Medronho, B.; Romano, A.; Miguel, M. G.; Stigsson, L.; Lindman, B. Rationalizing Cellulose (in)Solubility: Reviewing Basic Physicochemical Aspects and Role of Hydrophobic Interactions. *Cellulose* **2012**, *19* (3), 581–587. https://doi.org/10.1007/s10570-011-9644-6.

(189) Glasser, W. G.; Atalla, R. H.; Blackwell, J.; Brown, M. M.; Burchard, W.; French, A. D.; Klemm, D. O.; Nishiyama, Y. About the Structure of Cellulose: Debating the Lindman Hypothesis. *Cellulose* **2012**, *19* (3), 589–598. https://doi.org/10.1007/s10570-012-9691-7.

(190) Yu, Y.; Tyrikos-Ergas, T.; Zhu, Y.; Fittolani, G.; Bordoni, V.; Singhal, A.; Fair, R. J.; Grafmüller, A.; Seeberger, P. H.; Delbianco, M. Systematic Hydrogen-Bond Manipulations To Establish Polysaccharide Structure-Property Correlations. *Angew. Chem., Int. Ed.* **2019**, *58* (37), 13127–13132. https://doi.org/10.1002/anie.201906577.

(191) Zhu, Y.; Tyrikos-Ergas, T.; Schiefelbein, K.; Grafmüller, A.; Seeberger, P. H.; Delbianco, M. Automated Access to Well-Defined Ionic Oligosaccharides. *Org. Biomol. Chem.* **2020**, *18* (7), 1349–1353. https://doi.org/10.1039/D0OB00137F.

(192) Williams, D. A.; Pradhan, K.; Paul, A.; Olin, I. R.; Tuck, O. T.; Moulton, K. D.; Kulkarni, S. S.; Dube, D. H. Metabolic Inhibitors of Bacterial Glycan Biosynthesis. *Chem. Sci.* **2020**, *11* (7), 1761–1774. https://doi.org/10.1039/C9SC05955E.

(193) Champagne, P. A.; Desroches, J.; Hamel, J. D.; Vandamme, M.; Paquin, J. F. Monofluorination of Organic Compounds: 10 Years of Innovation. *Chemical Reviews*. American Chemical Society 2015, pp 9073–9174. https://doi.org/10.1021/cr500706a.

(194) Lin, P.-C.; Adak, A. K.; Ueng, S.-H.; Huang, L.-D.; Huang, K.-T.; Ho, J. A.; Lin, C.-C. DAST-Mediated Regioselective Anomeric Group Migration in Saccharides. *J. Org. Chem.* **2009**, *74* (11), 4041–4048. https://doi.org/10.1021/jo900516r.

(195) Nicolaou, K. C.; Dolle, R. E.; Papahatjis, D. P. Practical Synthesis of Oligosaccharides. Partial Synthesis of Avermectin B1a. *J. Am. Chem. Soc.* **1984**, *106* (15), 4189–4192. https://doi.org/10.1021/ja00327a021.

(196) Suzuki, K.; Ito, Y.; Kanie, O. An Improved Method for the Synthesis of Protected Glycosyl Fluorides from Thioglycosides Using N,N-Diethylaminosulfur Trifluoride (DAST). *Carbohydr. Res.* **2012**, *359*, 81–91. https://doi.org/10.1016/j.carres.2012.07.003.

(197) Ohtsuka, I.; Ako, T.; Kato, R.; Daikoku, S.; Koroghi, S.; Kanemitsu, T.; Kanie, O. Synthesis of a Library of Fucopyranosyl-Galactopyranosides Consisting of a Complete Set of Anomeric Configurations and Linkage Positions. *Carbohydr. Res.* **2006**, *341* (10), 1476–1487. https://doi.org/10.1016/j.carres.2006.03.040.

(198) Card, P. J.; Reddy, G. S. Fluorinated Carbohydrates. 2. Selective Fluorination of Gluco- and Mannopyranosides. Use of 2-D NMR for Structural Assignments. *J. Org. Chem.* **1983**, *48* (24), 4734–4743. https://doi.org/10.1021/jo00172a054.

(199) Makino, A.; Ohmae, M.; Kobayashi, S. Synthesis of Fluorinated Chitin Derivatives via Enzymatic Polymerization. *Macromol. Biosci.* **2006**, *6* (10), 862–872. https://doi.org/10.1002/mabi.200600128.

(200) Andreana, P. R.; Crich, D. Guidelines for O -Glycoside Formation from First Principles. *ACS Cent. Sci.* **2021**, *7* (9), 1454–1462. https://doi.org/10.1021/acscentsci.1c00594.

(201) Crich, D.; Dudkin, V. Why Are the Hydroxy Groups of Partially Protected N -Acetylglucosamine Derivatives Such Poor Glycosyl Acceptors, and What Can Be Done about It? A Comparative Study of the Reactivity of N -Acetyl-, N -Phthalimido-, and 2-Azido-2-Deoxy-Glucosamine Derivativ. *J. Am. Chem. Soc.* **2001**, *123* (28), 6819–6825. https://doi.org/10.1021/ja010086b.

(202) Enugala, R.; Carvalho, L. C. R.; Dias Pires, M. J.; Marques, M. M. B. Stereoselective Glycosylation of Glucosamine: The Role of the N -Protecting Group. *Chem. – An Asian J.* **2012**, *7* (11), 2482–2501. https://doi.org/10.1002/asia.201200338.

(203) Donohoe, T. J.; Logan, J. G.; Laffan, D. D. P. Trichloro-Oxazolines as Activated Donors for Aminosugar Coupling. *Org. Lett.* **2003**, *5* (26), 4995–4998. https://doi.org/10.1021/ol0359620.

(204) Nielsen, M. K.; Ugaz, C. R.; Li, W.; Doyle, A. G. PyFluor: A Low-Cost, Stable, and Selective Deoxyfluorination Reagent. *J. Am. Chem. Soc.* **2015**, *137* (30), 9571–9574. https://doi.org/10.1021/jacs.5b06307.

(205) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: A Generalizable Biomolecular Force Field. Carbohydrates. *J. Comput. Chem.* **2008**, *29* (4), 622–655. https://doi.org/10.1002/jcc.20820.

(206) Sauter, J.; Grafmüller, A. Predicting the Chemical Potential and Osmotic Pressure of Polysaccharide Solutions by Molecular Simulations. *J. Chem. Theory Comput.* **2016**, *12* (9), 4375–4384. https://doi.org/10.1021/acs.jctc.6b00295.

(207) Denavit, V.; Lainé, D.; Bouzriba, C.; Shanina, E.; Gillon, É.; Fortin, S.; Rademacher, C.; Imberty, A.; Giguère, D. Stereoselective Synthesis of Fluorinated Galactopyranosides as Potential Molecular Probes for Galactophilic Proteins: Assessment of Monofluorogalactoside–LecA Interactions. *Chem. – A Eur. J.* **2019**, *25* (17), 4478–4490. https://doi.org/10.1002/chem.201806197.

(208) Arnold, J. R. P.; Fisher, J. Conformational Analysis; 2007; Vol. 30, pp 370–385. https://doi.org/10.1039/9781847553836-00370.

(209) Tjandra, N.; Bax, A. Direct Measurement of Distances and Angles in Biomolecules by NMR in a Dilute Liquid Crystalline Medium. *Science* **1997**, *278* (5340), 1111–1114. https://doi.org/10.1126/science.278.5340.1111.

(210) Bax, A.; Grishaev, A. Weak Alignment NMR: A Hawk-Eyed View of Biomolecular Structure. *Curr. Opin. Struct. Biol.* **2005**, *15* (5), 563–570. https://doi.org/10.1016/j.sbi.2005.08.006.

(211) Troche-Pesqueira, E.; Cid, M. M.; Navarro-Vázquez, A. Disodium Cromoglycate: Exploiting Its Properties as a NMR Weak-Aligning Medium for Small Organic Molecules. *Org. Biomol. Chem.* **2014**, *12* (12), 1957–1965. https://doi.org/10.1039/c3ob42338g.

(212) Yu, L. J.; Saupe, A. Deuteron Resonance of D 2 O of Nematic Disodium Cromoglycate-Water Systems. *Mol. Cryst. Liq. Cryst.* **1982**, *80* (1), 129–134. https://doi.org/10.1080/00268948208071026.

(213) Lydon, J. Chromonic Review. *J. Mater. Chem.* **2010**, *20* (45), 10071–10099. https://doi.org/10.1039/b926374h.

(214) Poveda, A.; Fittolani, G.; Seeberger, P. H.; Delbianco, M.; Jiménez-Barbero, J. The Flexibility of Oligosaccharides Unveiled Through Residual Dipolar Coupling Analysis. *Front. Mol. Biosci.* **2021**, *8* (November), 1–9. https://doi.org/10.3389/fmolb.2021.784318.

(215) Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A. Validation of Protein Structure from Anisotropic Carbonyl Chemical Shifts in a Dilute Liquid Crystalline Phase. *J. Am. Chem. Soc.* **1998**, *120* (27), 6836–6837. https://doi.org/10.1021/ja9812610.

(216) Rillahan, C. D.; Paulson, J. C. Glycan Microarrays for Decoding the Glycome. *Annu. Rev. Biochem.* **2011**, *80* (1), 797–823. https://doi.org/10.1146/annurev-biochem-061809-152236.

(217) Stevens, J.; Blixt, O.; Paulson, J. C.; Wilson, I. A. Glycan Microarray Technologies: Tools to Survey Host Specificity of Influenza Viruses. *Nat. Rev. Microbiol.* **2006**, *4* (11), 857–864. https://doi.org/10.1038/nrmicro1530.

(218) Mende, M.; Bordoni, V.; Tsouka, A.; Loeffler, F. F.; Delbianco, M.; Seeberger, P. H. Multivalent Glycan Arrays. *Faraday Discuss.* **2019**, *219*, 9–32. https://doi.org/10.1039/C9FD00080A.

(219) Valles, D. J.; Naeem, Y.; Rozenfeld, A. Y.; Aldasooky, R. W.; Wong, A. M.; Carbonell, C.; Mootoo, D. R.; Braunschweig, A. B. Multivalent Binding of Concanavalin A on Variable-Density Mannoside Microarrays. *Faraday Discuss.* **2019**, *219* (0), 77–89. https://doi.org/10.1039/C9FD00028C.

(220) Liang, C.-H.; Wang, S.-K.; Lin, C.-W.; Wang, C.-C.; Wong, C.-H.; Wu, C.-Y. Effects of Neighboring Glycans on Antibody-Carbohydrate Interaction. *Angew. Chemie* **2011**, *123* (7), 1646–1650. https://doi.org/10.1002/ange.201003482.

(221) Dam, T. K.; Brewer, C. F. Thermodynamic Studies of Lectin−Carbohydrate Interactions by Isothermal Titration Calorimetry. *Chem. Rev.* **2002**, *102* (2), 387–430. https://doi.org/10.1021/cr000401x.

(222) Smith, E. A.; Thomas, W. D.; Kiessling, L. L.; Corn, R. M. Surface Plasmon Resonance Imaging Studies of Protein-Carbohydrate Interactions. *J. Am. Chem. Soc.* **2003**, *125* (20), 6140–6148. https://doi.org/10.1021/ja034165u.

(223) Toone, E. J. Structure and Energetics of Protein-Carbohydrate Complexes. *Curr. Opin. Struct. Biol.* **1994**, *4* (5), 719–728. https://doi.org/10.1016/S0959-440X(94)90170-8.

(224) Dalvit, C. Ligand- and Substrate-Based 19F NMR Screening: Principles and Applications to Drug Discovery. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51* (4), 243–271. https://doi.org/10.1016/j.pnmrs.2007.07.002.

(225) Garavís, M.; López-Méndez, B.; Somoza, A.; Oyarzabal, J.; Dalvit, C.; Villasante, A.; Campos-Olivas, R.; González, C. Discovery of Selective Ligands for Telomeric RNA G-Quadruplexes (TERRA) through 19 F-NMR Based Fragment Screening. *ACS Chem. Biol.* **2014**, *9* (7), 1559–1566. https://doi.org/10.1021/cb500100z.

(226) Troelsen, N. S.; Shanina, E.; Gonzalez-Romero, D.; Danková, D.; Jensen, I. S. A.; Śniady, K. J.; Nami, F.; Zhang, H.; Rademacher, C.; Cuenda, A.; Gotfredsen, C. H.; Clausen, M. H. The 3F Library: Fluorinated Fsp 3 -Rich Fragments for Expeditious 19 F NMR Based Screening. *Angew. Chem., Int. Ed.* **2020**, *59* (6), 2204–2210. https://doi.org/10.1002/anie.201913125.

(227) Audfray, A.; Varrot, A.; Imberty, A. Bacteria Love Our Sugars: Interaction between Soluble Lectins and Human Fucosylated Glycans, Structures, Thermodynamics and Design of Competing Glycocompounds. *Comptes Rendus Chim.* **2013**, *16* (5), 482–490. https://doi.org/10.1016/j.crci.2012.11.021.

(228) Heggelund, J. E.; Varrot, A.; Imberty, A.; Krengel, U. Histo-Blood Group Antigens as Mediators of Infections. *Curr. Opin. Struct. Biol.* **2017**, *44*, 190–200. https://doi.org/10.1016/j.sbi.2017.04.001.

(229) Gimeno, A.; Delgado, S.; Valverde, P.; Bertuzzi, S.; Berbís, M. A.; Echavarren, J.; Lacetera, A.; Martín-Santamaría, S.; Surolia, A.; Cañada, F. J.; Jiménez-Barbero, J.; Ardá, A. Minimizing the Entropy Penalty for Ligand Binding: Lessons from the Molecular Recognition of the Histo Blood-Group Antigens by Human Galectin-3. *Angew. Chem., Int. Ed.* **2019**, *58* (22), 7268–7272. https://doi.org/10.1002/anie.201900723.

(230) Srivastava, A. D.; Unione, L.; Wolfert, M. A.; Valverde, P.; Arda, A.; Jiménez-Barbero, J.; Boons, G.-J. Mono- and Di-Fucosylated Glycans of the Parasitic Worm S. Mansoni Are Recognized Differently by the Innate Immune Receptor DC-SIGN. *Chem. – A Eur. J.* **2020**, chem.202002619. https://doi.org/10.1002/chem.202002619.

(231) Dalvit, C.; Invernizzi, C.; Vulpetti, A. Fluorine as a Hydrogen-Bond Acceptor: Experimental Evidence and Computational Calculations. *Chem. - A Eur. J* **2014**, *20* (35), 11058–11068. https://doi.org/10.1002/chem.201402858.

(232) Gimeno, A.; Reichardt, N.-C.; Cañada, F. J.; Perkams, L.; Unverzagt, C.; Jiménez-Barbero, J.; Ardá, A. NMR and Molecular Recognition of N-Glycans: Remote Modifications of the Saccharide Chain Modulate Binding Features. *ACS Chem. Biol.* **2017**, *12* (4), 1104–1112. https://doi.org/10.1021/acschembio.6b01116.

(233) Guberman, M.; Bräutigam, M.; Seeberger, P. H. Automated Glycan Assembly of Lewis Type I and II Oligosaccharide Antigens. *Chem. Sci.* **2019**, *10* (21), 5634–5640. https://doi.org/10.1039/C9SC00768G.

(234) Li, Q.; Jiang, W.; Guo, J.; Jaiswal, M.; Guo, Z. Synthesis of Lewis y Analogues and Their Protein Conjugates for Structure-Immunogenicity Relationship Studies of Lewis y Antigen. *J. Org. Chem.* **2019**, *84* (21), 13232–13241. https://doi.org/10.1021/acs.joc.9b00537.

(235) Guo, J.; Jiang, W.; Li, Q.; Jaiswal, M.; Guo, Z. Comparative Immunological Studies of Tumor-Associated Lewis X, Lewis Y, and KH-1 Antigens. *Carbohydr. Res.* **2020**, *492*, 107999. https://doi.org/10.1016/j.carres.2020.107999.

(236) Yu, H.; Chokhawala, H.; Karpel, R.; Yu, H.; Wu, B.; Zhang, J.; Zhang, Y.; Jia, Q.; Chen, X. A Multifunctional Pasteurella Multocida Sialyltransferase: A Powerful Tool for the Synthesis of Sialoside Libraries. *J. Am. Chem. Soc.* **2005**, *127* (50), 17618–17619. https://doi.org/10.1021/ja0561690.

(237) Kajihara, Y.; Yamamoto, T.; Nagae, H.; Nakashizuka, M.; Sakakibara, T.; Terada, I. A Novel α-2,6-Sialyltransferase: Transfer of Sialic Acid to Fucosyl and Sialyl Trisaccharides. *J. Org. Chem.* **1996**, *61* (24), 8632–8635. https://doi.org/10.1021/jo961214v.

(238) Feinberg, H.; Taylor, M. E.; Razi, N.; McBride, R.; Knirel, Y. A.; Graham, S. A.; Drickamer, K.; Weis, W. I. Structural Basis for Langerin Recognition of Diverse Pathogen and Mammalian Glycans through a Single Binding Site. *J. Mol. Biol.* **2011**, *405* (4), 1027–1039. https://doi.org/10.1016/j.jmb.2010.11.039.

(239) van Die, I.; van Vliet, S.; Kwame Nyame, A.; Cummings, R. D.; Bank, C. M. C.; Appelmelk, B.; Geijtenbeek, T. B. H.; van Kooyk, Y. The Dendritic Cell-Specific C-Type Lectin DC-SIGN Is a Receptor for Schistosoma Mansoni Egg Antigens and Recognizes the Glycan Antigen Lewis X. *Glycobiology* **2003**, *13* (6), 471–478. https://doi.org/10.1093/glycob/cwg052.

(240) van Liempt, E.; Bank, C. M. C.; Mehta, P.; García-Vallejo, J. J.; Kawar, Z. S.; Geyer, R.; Alvarez, R.

A.; Cummings, R. D.; Kooyk, Y. van; van Die, I. Specificity of DC-SIGN for Mannose- and Fucose-Containing Glycans. *FEBS Lett.* **2006**, *580* (26), 6123–6131. https://doi.org/10.1016/j.febslet.2006.10.009.

(241) García-Vallejo, J. J.; van Liempt, E.; da Costa Martins, P.; Beckers, C.; van het Hof, B.; Gringhuis, S. I.; Zwaginga, J.-J.; van Dijk, W.; Geijtenbeek, T. B. H.; van Kooyk, Y. DC-SIGN Mediates Adhesion and Rolling of Dendritic Cells on Primary Human Umbilical Vein Endothelial Cells through LewisY Antigen Expressed on ICAM-2. *Mol. Immunol.* **2008**, *45* (8), 2359–2369. https://doi.org/10.1016/j.molimm.2007.11.001.

(242) Appelmelk, B. J.; van Die, I.; van Vliet, S. J.; Vandenbroucke-Grauls, C. M. J. E.; Geijtenbeek, T. B. H.; van Kooyk, Y. Cutting Edge: Carbohydrate Profiling Identifies New Pathogens That Interact with Dendritic Cell-Specific ICAM-3-Grabbing Nonintegrin on Dendritic Cells. *J. Immunol.* **2003**, *170* (4), 1635–1639. https://doi.org/10.4049/jimmunol.170.4.1635.

(243) Guo, Y.; Feinberg, H.; Conroy, E.; Mitchell, D. A.; Alvarez, R.; Blixt, O.; Taylor, M. E.; Weis, W. I.; Drickamer, K. Structural Basis for Distinct Ligand-Binding and Targeting Properties of the Receptors DC-SIGN and DC-SIGNR. *Nat. Struct. Mol. Biol.* **2004**, *11* (7), 591–598. https://doi.org/10.1038/nsmb784.

(244) Sugiarto, G.; Lau, K.; Qu, J.; Li, Y.; Lim, S.; Mu, S.; Ames, J. B.; Fisher, A. J.; Chen, X. A Sialyltransferase Mutant with Decreased Donor Hydrolysis and Reduced Sialidase Activities for Directly Sialylating Lewis X. *ACS Chem. Biol.* **2012**, *7* (7), 1232–1240. https://doi.org/10.1021/cb300125k.

(245) Ban, L.; Pettit, N.; Li, L.; Stuparu, A. D.; Cai, L.; Chen, W.; Guan, W.; Han, W.; Wang, P. G.; Mrksich, M. Discovery of Glycosyltransferases Using Carbohydrate Arrays and Mass Spectrometry. *Nat. Chem. Biol.* **2012**, *8* (9), 769–773. https://doi.org/10.1038/nchembio.1022.

(246) Shao, W.; Sharma, R.; Clausen, M. H.; Scheller, H. V. Microscale Thermophoresis as a Powerful Tool for Screening Glycosyltransferases Involved in Cell Wall Biosynthesis. *Plant Methods* **2020**, *16* (1), 99. https://doi.org/10.1186/s13007-020-00641-1.

(247) Craven, G. R.; Steers, E.; Anfinsen, C. B. Purification, Composition, and Molecular Weight of the Beta-Galactosidase of Escherichia Coli K12. *J. Biol. Chem.* **1965**, *240* (6), 2468–2477.

(248) Wang, J.; Liu, K.; Xing, R.; Yan, X. Peptide Self-Assembly: Thermodynamics and Kinetics. *Chem. Soc. Rev.* **2016**, *45* (20), 5589–5604. https://doi.org/10.1039/C6CS00176A.

(249) Kwak, M.; Herrmann, A. Nucleic Acid Amphiphiles: Synthesis and Self-Assembled Nanostructures. *Chem. Soc. Rev.* **2011**, *40* (12), 5745. https://doi.org/10.1039/c1cs15138j.

(250) Lv, J.; Gao, X.; Han, B.; Zhu, Y.; Hou, K.; Tang, Z. Self-Assembled Inorganic Chiral Superstructures. *Nat. Rev. Chem.* **2022**, *0123456789*. https://doi.org/10.1038/s41570-021-00350-w.

(251) Brandt, J. R.; Salerno, F.; Fuchter, M. J. The Added Value of Small-Molecule Chirality in Technological Applications. *Nat. Rev. Chem.* **2017**, *1* (6), 0045. https://doi.org/10.1038/s41570-017-0045.

(252) Noyori, R. Asymmetric Catalysis: Science and Opportunities (Nobel Lecture). *Angew. Chem., Int. Ed.* **2002**, *41* (12), 2008. https://doi.org/10.1002/1521-3773(20020617)41:12<2008::AID-ANIE2008>3.0.CO;2-4.

(253) Yashima, E.; Ousaka, N.; Taura, D.; Shimomura, K.; Ikai, T.; Maeda, K. Supramolecular Helical Systems: Helical Assemblies of Small Molecules, Foldamers, and Polymers with Chiral Amplification and Their Functions. *Chem. Rev.* **2016**, *116* (22), 13752–13990. https://doi.org/10.1021/acs.chemrev.6b00354.

(254) Yu, Y.; Gim, S.; Kim, D.; Arnon, Z. A.; Gazit, E.; Seeberger, P. H.; Delbianco, M. Oligosaccharides Self-Assemble and Show Intrinsic Optical Properties. *J. Am. Chem. Soc.* **2019**, *141* (12), 4833–4838. https://doi.org/10.1021/jacs.8b11882.

(255) Kubota, R.; Tanaka, W.; Hamachi, I. Microscopic Imaging Techniques for Molecular Assemblies: Electron, Atomic Force, and Confocal Microscopies. *Chem. Rev.* **2021**, *121* (22), 14281–14347. https://doi.org/10.1021/acs.chemrev.0c01334.

(256) Ogawa, Y.; Putaux, J.-L. Transmission Electron Microscopy of Cellulose. Part 2: Technical and Practical Aspects. *Cellulose* **2019**, *26* (1), 17–34. https://doi.org/10.1007/s10570-018-2075-x.

(257) Glaeser, R. M. Limitations to Significant Information in Biological Electron Microscopy as a Result of Radiation Damage. *J. Ultrastruct. Res.* **1971**, *36* (3–4), 466–482. https://doi.org/10.1016/S0022-5320(71)80118-1.

(258) Janiak, C. A Critical Account on N-n Stacking in Metal Complexes with Aromatic Nitrogen-Containing Ligands. *J. Chem. Soc. Dalt. Trans.* **2000**, No. 21, 3885–3896. https://doi.org/10.1039/b003010o.

(259) Usov, I.; Nyström, G.; Adamcik, J.; Handschin, S.; Schütz, C.; Fall, A.; Bergström, L.; Mezzenga, R. Understanding Nanocellulose Chirality and Structure–Properties Relationship at the Single Fibril Level. *Nat. Commun.* **2015**, *6* (1), 7564. https://doi.org/10.1038/ncomms8564.

(260) Ogawa, Y.; Kimura, S.; Wada, M. Electron Diffraction and High-Resolution Imaging on Highly-Crystalline β-Chitin Microfibril. *J. Struct. Biol.* **2011**, *176* (1), 83–90. https://doi.org/10.1016/j.jsb.2011.07.001.

(261) Liu, M.; Zhang, L.; Wang, T. Supramolecular Chirality in Self-Assembled Systems. *Chem. Rev.* **2015**, *115* (15), 7304–7397. https://doi.org/10.1021/cr500671p.

(262) Lin, S.; Li, Y.; Li, B.; Yang, Y. Control of the Handedness of Self-Assemblies of Dipeptides by the Chirality of Phenylalanine and Steric Hindrance of Phenylglycine. *Langmuir* **2016**, *32* (29), 7420–7426. https://doi.org/10.1021/acs.langmuir.6b01874.

(263) Garcia, A. M.; Iglesias, D.; Parisi, E.; Styan, K. E.; Waddington, L. J.; Deganutti, C.; De Zorzi, R.; Grassi, M.; Melchionna, M.; Vargiu, A. V.; Marchesan, S. Chirality Effects on Peptide Self-Assembly Unraveled from Molecules to Materials. *Chem* **2018**, *4* (8), 1862–1876. https://doi.org/10.1016/j.chempr.2018.05.016.

(264) Raymond, D. M.; Nilsson, B. L. Multicomponent Peptide Assemblies. *Chem. Soc. Rev.* **2018**, *47* (10), 3659–3720. https://doi.org/10.1039/c8cs00115d.

(265) Lee, I. C.; Zulueta, M. M. L.; Shie, C. R.; Arco, S. D.; Hung, S. C. Deuterium-Isotope Study on the Reductive Ring Opening of Benzylidene Acetals. *Org. Biomol. Chem.* **2011**, *9* (22), 7655–7658. https://doi.org/10.1039/c1ob06056b.

(266) Pace, C. J.; Gao, J. Exploring and Exploiting Polar−π Interactions with Fluorinated Aromatic Amino Acids. *Acc. Chem. Res.* **2013**, *46* (4), 907–915. https://doi.org/10.1021/ar300086n.

(267) Vignolini, S.; Rudall, P. J.; Rowland, A. V.; Reed, A.; Moyroud, E.; Faden, R. B.; Baumberg, J. J.; Glover, B. J.; Steiner, U. Pointillist Structural Color in Pollia Fruit. *Proc. Natl. Acad. Sci.* **2012**, *109* (39), 15712–15715. https://doi.org/10.1073/pnas.1210105109.

(268) Chang, Y.; Middleton, R.; Ogawa, Y.; Gregory, T.; Steiner, L. M.; Kovalev, A.; Karanja, R. H. N.; Rudall, P. J.; Glover, B. J.; Gorb, S. N.; Vignolini, S. Cell Wall Composition Determines Handedness Reversal in Helicoidal Cellulose Architectures of Pollia Condensata Fruits. *Proc. Natl. Acad. Sci.* **2021**, *118* (51), e2111723118. https://doi.org/10.1073/pnas.2111723118.

(269) Neville, A. C. A Pipe-Cleaner Molecular Model for Morphogenesis of Helicoidal Plant Cell Walls Based on Hemicellulose Complexity. *J. Theor. Biol.* **1988**, *131* (2), 243–254. https://doi.org/10.1016/S0022-5193(88)80241-8.

(270) Parker, R. M.; Guidetti, G.; Williams, C. A.; Zhao, T.; Narkevicius, A.; Vignolini, S.; Frka-Petesic, B. The Self-Assembly of Cellulose Nanocrystals: Hierarchical Design of Visual Appearance. *Adv. Mater.* **2018**, *30* (19), 1704477. https://doi.org/10.1002/adma.201704477.

(271) Parker, R. M.; Frka-Petesic, B.; Guidetti, G.; Kamita, G.; Consani, G.; Abell, C.; Vignolini, S. Hierarchical Self-Assembly of Cellulose Nanocrystals in a Confined Geometry. *ACS Nano* **2016**, *10* (9), 8443–8449. https://doi.org/10.1021/acsnano.6b03355.

(272) Frka-Petesic, B.; Vignolini, S. So Much More than Paper. *Nat. Photonics* **2019**, *13* (6), 365–367. https://doi.org/10.1038/s41566-019-0448-9.

(273) Droguet, B. E.; Liang, H.-L.; Frka-Petesic, B.; Parker, R. M.; De Volder, M. F. L.; Baumberg, J. J.; Vignolini, S. Large-Scale Fabrication of Structurally Coloured Cellulose Nanocrystal Films and Effect Pigments. *Nat. Mater.* **2022**, *21* (3), 352–358. https://doi.org/10.1038/s41563-021-01135-8.

(274) Gonçalves, D. P. N.; Hegmann, T. Chirality Transfer from an Innately Chiral Nanocrystal Core to a Nematic Liquid Crystal: Surface-Modified Cellulose Nanocrystals. *Angew. Chem., Int. Ed.* **2021**, anie.202105357. https://doi.org/10.1002/anie.202105357.

(275) Kaushik, M.; Basu, K.; Benoit, C.; Cirtiu, C. M.; Vali, H.; Moores, A. Cellulose Nanocrystals as Chiral Inducers: Enantioselective Catalysis and Transmission Electron Microscopy 3D Characterization. *J. Am. Chem. Soc.* **2015**, *137* (19), 6124–6127. https://doi.org/10.1021/jacs.5b02034.

(276) Majoinen, J.; Hassinen, J.; Haataja, J. S.; Rekola, H. T.; Kontturi, E.; Kostiainen, M. A.; Ras, R. H. A.; Törmä, P.; Ikkala, O. Chiral Plasmonics Using Twisting along Cellulose Nanocrystals as a Template for Gold Nanoparticles. *Adv. Mater.* **2016**, *28* (26), 5262–5267. https://doi.org/10.1002/adma.201600940.

(277) Shopsowitz, K. E.; Qi, H.; Hamad, W. Y.; MacLachlan, M. J. Free-Standing Mesoporous Silica Films with Tunable Chiral Nematic Structures. *Nature* **2010**, *468* (7322), 422–425. https://doi.org/10.1038/nature09540.

(278) Ogawa, Y. Release of Internal Molecular Torque Results in Twists of Glaucocystis Cellulose Nanofibers. *Carbohydr. Polym.* **2021**, *251*, 117102. https://doi.org/10.1016/j.carbpol.2020.117102.

(279) Araki, J.; Kuga, S. Effect of Trace Electrolyte on Liquid Crystal Type of Cellulose Microcrystals. *Langmuir* **2001**, *17* (15), 4493–4496. https://doi.org/10.1021/la0102455.

(280) Chiappini, M.; Dussi, S.; Frka-Petesic, B.; Vignolini, S.; Dijkstra, M. Modeling the Cholesteric Pitch of Apolar Cellulose Nanocrystal Suspensions Using a Chiral Hard-Bundle Model. *J. Chem. Phys.* **2022**, *156* (1), 014904. https://doi.org/10.1063/5.0076123.

(281) Nyström, G.; Arcari, M.; Adamcik, J.; Usov, I.; Mezzenga, R. Nanocellulose Fragmentation Mechanisms and Inversion of Chirality from the Single Particle to the Cholesteric Phase. *ACS Nano* **2018**, *12* (6), 5141–5148. https://doi.org/10.1021/acsnano.8b00512.

(282) Bai, L.; Kämäräinen, T.; Xiang, W.; Majoinen, J.; Seitsonen, J.; Grande, R.; Huan, S.; Liu, L.; Fan, Y.; Rojas, O. J. Chirality from Cryo-Electron Tomograms of Nanocrystals Obtained by Lateral Disassembly and Surface Etching of Never-Dried Chitin. *ACS Nano* **2020**, *14* (6), 6921–6930. https://doi.org/10.1021/acsnano.0c01327.

(283) Parton, T. G.; Parker, R. M.; van de Kerkhof, G. T.; Narkevicius, A.; Haataja, J. S.; Frka-Petesic, B.; Vignolini, S. Chiral Self-Assembly of Cellulose Nanocrystals Is Driven by Crystallite Bundles. *arXiv* **2022**. https://doi.org/doi.org/10.48550/arXiv.2107.04772.

(284) Gray, D. G. Chiral Nematic Ordering of Polysaccharides. *Carbohydr. Polym.* **1994**, *25* (4), 277–284. https://doi.org/10.1016/0144-8617(94)90053-1.

(285) Conley, K.; Godbout, L.; Whitehead, M. A. (Tony); van de Ven, T. G. M. Origin of the Twist of Cellulosic Materials. *Carbohydr. Polym.* **2016**, *135*, 285–299. https://doi.org/10.1016/j.carbpol.2015.08.029.

(286) Wada, M.; Wakiya, S.; Kobayashi, K.; Kimura, S.; Kitaoka, M.; Kusumi, R.; Kimura, F.; Kimura, T. Three-Dimensional Alignment of Cellulose II Microcrystals under a Strong Magnetic Field. *Cellulose*

2021, *28* (11), 6757–6765. https://doi.org/10.1007/s10570-021-03954-z.

(287) Hirota, M.; Tamura, N.; Saito, T.; Isogai, A. Cellulose II Nanoelements Prepared from Fully Mercerized, Partially Mercerized and Regenerated Celluloses by 4-Acetamido-TEMPO/NaClO/NaClO2 Oxidation. *Cellulose* **2012**, *19* (2), 435–442. https://doi.org/10.1007/s10570-011-9642-8.

(288) Sakamoto, J.; Sugiyama, J.; Kimura, S.; Imai, T.; Itoh, T.; Watanabe, T.; Kobayashi, S. Artificial Chitin Spherulites Composed of Single Crystalline Ribbons of α-Chitin via Enzymatic Polymerization. *Macromolecules* **2000**, *33* (11), 4155–4160. https://doi.org/10.1021/ma000230y.

(289) Kobayashi, S.; Hobson, L. J.; Sakamoto, J.; Kimura, S.; Sugiyama, J.; Imai, T.; Itoh, T. Formation and Structure of Artificial Cellulose Spherulites via Enzymatic Polymerization. *Biomacromolecules* **2000**, *1* (2), 168–173. https://doi.org/10.1021/bm990010w.

(290) Aparicio, F.; Nieto-Ortega, B.; Nájera, F.; Ramírez, F. J.; López Navarrete, J. T.; Casado, J.; Sánchez, L. Inversion of Supramolecular Helicity in Oligo- p -Phenylene-Based Supramolecular Polymers: Influence of Molecular Atropisomerism . *Angew. Chemie* **2014**, *126* (5), 1397–1401. https://doi.org/10.1002/ange.201309172.

(291) Buleon, A.; Chanzy, H. Single Crystals of Cellulose IVII: Preparation and Properties. *J. Polym. Sci. Polym. Phys. Ed.* **1980**, *18* (6), 1209–1217. https://doi.org/10.1002/pol.1980.180180604.

(292) Narkevicius, A.; Steiner, L. M.; Parker, R. M.; Ogawa, Y.; Frka-Petesic, B.; Vignolini, S. Controlling the Self-Assembly Behavior of Aqueous Chitin Nanocrystal Suspensions. *Biomacromolecules* **2019**, *20* (7), 2830–2838. https://doi.org/10.1021/acs.biomac.9b00589.

(293) Venkatraman, J.; Shankaramma, S. C.; Balaram, P. Design of Folded Peptides. *Chem. Rev.* **2001**, *101* (10), 3131–3152. https://doi.org/10.1021/cr000053z.

(294) Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. A Field Guide to Foldamers. *Chem. Rev.* **2001**, *101* (12), 3893–4012. https://doi.org/10.1021/cr990120t.

(295) Ferrand, Y.; Huc, I. Designing Helical Molecular Capsules Based on Folded Aromatic Amide Oligomers. *Acc. Chem. Res.* **2018**, *51* (4), 970–977. https://doi.org/10.1021/acs.accounts.8b00075.

(296) Gellman, S. H. Minimal Model Systems for β Sheet Secondary Structure in Proteins. *Curr. Opin. Chem. Biol.* **1998**, *2* (6), 717–725. https://doi.org/10.1016/S1367-5931(98)80109-9.

(297) Fittolani, G.; Seeberger, P. H.; Delbianco, M. Helical Polysaccharides. *Pept. Sci.* **2020**, *112* (1), e24124. https://doi.org/10.1002/pep2.24124.

(298) Topin, J.; Lelimousin, M.; Arnaud, J.; Audfray, A.; Pérez, S.; Varrot, A.; Imberty, A. The Hidden Conformation of Lewis x, a Human Histo-Blood Group Antigen, Is a Determinant for Recognition by Pathogen Lectins. *ACS Chem. Biol.* **2016**, *11* (7), 2011–2020. https://doi.org/10.1021/acschembio.6b00333.

(299) Lycknert, K.; Edebrink, P.; Widmalm, G. A Conformational Carbohydrate Scaffold Is Present in the Short-Chain Lipopolysaccharides OfMoraxella Catarrhalis. *Angew. Chem., Int. Ed.* **2004**, *43* (17), 2288–2290. https://doi.org/10.1002/anie.200353581.

(300) Wyss, D. F.; Choi, J. S.; Li, J.; Knoppers, M. H.; Willis, K. J.; Arulanandam, A. R. N.; Smolyar, A.; Reinherz, E. L.; Wagner, G. Conformation and Function of the N-Linked Glycan in the Adhesion Domain of Human CD2. *Science* **1995**, *269* (5228), 1273–1278. https://doi.org/10.1126/science.7544493.

(301) Canales, A.; Boos, I.; Perkams, L.; Karst, L.; Luber, T.; Karagiannis, T.; Domínguez, G.; Cañada, F. J.; Pérez-Castells, J.; Häussinger, D.; Unverzagt, C.; Jiménez-Barbero, J. Breaking the Limits in Analyzing Carbohydrate Recognition by NMR Spectroscopy: Resolving Branch-Selective Interaction of a Tetra-Antennary N -Glycan with Lectins. *Angew. Chem., Int. Ed.* **2017**, *56* (47), 14987–14991. https://doi.org/10.1002/anie.201709130.

(302) Zhang, Y.; Gómez-Redondo, M.; Jiménez-Osés, G.; Arda, A.; Overkleeft, H. S.; Marel, G. A.; Jiménez-Barbero, J.; Codée, J. D. C. Synthesis and Structural Analysis of Aspergillus Fumigatus Galactosaminogalactans Featuring A-Galactose, A-Galactosamine and A- N -Acetyl Galactosamine Linkages. *Angew. Chem., Int. Ed.* **2020**, *59* (31), 12746–12750. https://doi.org/10.1002/anie.202003951.

(303) Zierke, M.; Smieško, M.; Rabbani, S.; Aeschbacher, T.; Cutting, B.; Allain, F. H.-T.; Schubert, M.; Ernst, B. Stabilization of Branched Oligosaccharides: Lewis x Benefits from a Nonconventional C–H···O Hydrogen Bond. *J. Am. Chem. Soc.* **2013**, *135* (36), 13464–13472. https://doi.org/10.1021/ja4054702.

(304) Aeschbacher, T.; Zierke, M.; Smieško, M.; Collot, M.; Mallet, J.; Ernst, B.; Allain, F. H. -T.; Schubert, M. A Secondary Structural Element in a Wide Range of Fucosylated Glycoepitopes. *Chem. - A Eur. J* **2017**, *23* (48), 11598–11610. https://doi.org/10.1002/chem.201701866.

(305) Zhang, Q.; Gimeno, A.; Santana, D.; Wang, Z.; Valdés-Balbin, Y.; Rodríguez-Noda, L. M.; Hansen, T.; Kong, L.; Shen, M.; Overkleeft, H. S.; Vérez-Bencomo, V.; van der Marel, G. A.; Jiménez-Barbero, J.; Chiodo, F.; Codée, J. D. C. Synthetic, Zwitterionic Sp1 Oligosaccharides Adopt a Helical Structure Crucial for Antibody Interaction. *ACS Cent. Sci.* **2019**, *5* (8), 1407–1416. https://doi.org/10.1021/acscentsci.9b00454.

(306) Battistel, M. D.; Azurmendi, H. F.; Frank, M.; Freedberg, D. I. Uncovering Nonconventional and Conventional Hydrogen Bonds in Oligosaccharides through NMR Experiments and Molecular Modeling: Application to Sialyl Lewis-X. *J. Am. Chem. Soc.* **2015**, *137* (42), 13444–13447. https://doi.org/10.1021/jacs.5b03824.

(307) Fisher, B. F.; Gellman, S. H. Impact of γ-Amino Acid Residue Preorganization on α/γ-Peptide Foldamer Helicity in Aqueous Solution. *J. Am. Chem. Soc.* **2016**, *138* (34), 10766–10769. https://doi.org/10.1021/jacs.6b06177.

(308) Freire, F.; Fisk, J. D.; Peoples, A. J.; Ivancic, M.; Guzei, I. A.; Gellman, S. H. Diacid Linkers That Promote Parallel β-Sheet Secondary Structure in Water. *J. Am. Chem. Soc.* **2008**, *130* (25), 7839–7841. https://doi.org/10.1021/ja802042c.

(309) Khakshoor, O.; Nowick, J. S. Artificial β-Sheets: Chemical Models of β-Sheets. *Curr. Opin. Chem. Biol.* **2008**, *12* (6), 722–729. https://doi.org/10.1016/j.cbpa.2008.08.009.

(310) Blanco, F. J.; Jimenez, M. A.; Herranz, J.; Rico, M.; Santoro, J.; Nieto, J. L. NMR Evidence of a Short Linear Peptide That Folds into a .Beta.-Hairpin in Aqueous Solution. *J. Am. Chem. Soc.* **1993**, *115* (13), 5887–5888. https://doi.org/10.1021/ja00066a092.

(311) Fisk, J. D.; Gellman, S. H. A Parallel β-Sheet Model System That Folds in Water. *J. Am. Chem. Soc.* **2001**, *123* (2), 343–344. https://doi.org/10.1021/ja002493d.

(312) Fisk, J. D.; Powell, D. R.; Gellman, S. H. Control of Hairpin Formation via Proline Configuration in Parallel β-Sheet Model Systems. *J. Am. Chem. Soc.* **2000**, *122* (23), 5443–5447. https://doi.org/10.1021/ja9929483.

(313) Nowick, J. S.; Smith, E. M.; Pairish, M. Artificial β-Sheets. *Chem. Soc. Rev.* **1996**, *25* (6), 401–415. https://doi.org/10.1039/CS9962500401.

(314) Samdin, T. D.; Kreutzer, A. G.; Nowick, J. S. Exploring Amyloid Oligomers with Peptide Model Systems. *Curr. Opin. Chem. Biol.* **2021**, *64*, 106–115. https://doi.org/10.1016/j.cbpa.2021.05.004.

(315) Metrano, A. J.; Chinn, A. J.; Shugrue, C. R.; Stone, E. A.; Kim, B.; Miller, S. J. Asymmetric Catalysis Mediated by Synthetic Peptides, Version 2.0: Expansion of Scope and Mechanisms. *Chem. Rev.* **2020**, *120* (20), 11479–11615. https://doi.org/10.1021/acs.chemrev.0c00523.

(316) Sebaoun, L.; Maurizot, V.; Granier, T.; Kauffmann, B.; Huc, I. Aromatic Oligoamide β-Sheet Foldamers. *J. Am. Chem. Soc.* **2014**, *136* (5), 2168–2174. https://doi.org/10.1021/ja412729s.

(317) Newberry, R. W.; Raines, R. T. Secondary Forces in Protein Folding. *ACS Chem. Biol.* **2019**, *14* (8), 1677–1686. https://doi.org/10.1021/acschembio.9b00339.

(318) Bernet, B.; Xu, J.; Vasella, A. Oligosaccharide Analogues of Polysaccharides, Part 20 , NMR Analysis of Templated Cellodextrins Possessing Two Parallel Chains: A Mimic for Cellulose I? *Helv. Chim. Acta* **2000**, *83* (9), 2072–2114. https://doi.org/10.1002/1522-2675(20000906)83:9<2072::AID-HLCA2072>3.0.CO;2-Q.

(319) Murty, K. V. S. N.; Xie, T.; Bernet, B.; Vasella, A. Oligosaccharide Analogues of Polysaccharides. Part 26. *Helv. Chim. Acta* **2006**, *89* (4), 675–730. https://doi.org/10.1002/hlca.200690068.

(320) Mahoney, M. W.; Jorgensen, W. L. A Five-Site Model for Liquid Water and the Reproduction of the Density Anomaly by Rigid, Nonpolarizable Potential Functions. *J. Chem. Phys.* **2000**, *112* (20), 8910–8922. https://doi.org/10.1063/1.481505.

(321) 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. *J. Am. Chem. Soc.* **2019**, *141* (22), 9079–9086. https://doi.org/10.1021/jacs.9b03769.

(322) Sanapala, S. R.; Kulkarni, S. S. From L-Rhamnose to Rare 6-Deoxy-L-Hexoses. *Org. Lett.* **2016**, *18* (15), 3790–3793. https://doi.org/10.1021/acs.orglett.6b01796.

(323) Wilsdorf, M.; Schmidt, D.; Bartetzko, M. P.; Dallabernardina, P.; Schuhmacher, F.; Seeberger Ab, P. H.; Pfrengle, F. A Traceless Photocleavable Linker for the Automated Glycan Assembly of Carbohydrates with Free Reducing Ends †. *Chem. Commun* **2016**, *52*, 10187. https://doi.org/10.1039/c6cc04954k.

(324) Lin, Y. H.; Ghosh, B.; Tony Mong, K.-K. In Situ Formation of β-Glycosyl Imidinium Triflate from Participating Thioglycosyl Donors: Elaboration to Disarmed–Armed Iterative Glycosylation. *Chem. Commun.* **2012**, *48* (88), 10910. https://doi.org/10.1039/c2cc35032g.

(325) Novakovic, M.; Battistel, M. D.; Azurmendi, H. F.; Concilio, M.-G.; Freedberg, D. I.; Frydman, L. The Incorporation of Labile Protons into Multidimensional NMR Analyses: Glycan Structures Revisited. *J. Am. Chem. Soc.* **2021**, *143* (23), 8935–8948. https://doi.org/10.1021/jacs.1c04512.

(326) Broecker, F.; Hanske, J.; Martin, C. E.; Baek, J. Y.; Wahlbrink, A.; Wojcik, F.; Hartmann, L.; Rademacher, C.; Anish, C.; Seeberger, P. H. Multivalent Display of Minimal Clostridium Difficile Glycan Epitopes Mimics Antigenic Properties of Larger Glycans. *Nat. Commun.* **2016**, *7* (1), 11224. https://doi.org/10.1038/ncomms11224.

(327) Martin, C. E.; Weishaupt, M. W.; Seeberger, P. H. Progress toward Developing a Carbohydrate-Conjugate Vaccine against Clostridium Difficile Ribotype 027: Synthesis of the Cell-Surface Polysaccharide PS-I Repeating Unit. *Chem. Commun.* **2011**, *47* (37), 10260. https://doi.org/10.1039/c1cc13614c.

(328) Poveda, A.; Asensio, J. L.; Martín-Pastor, M.; Jiménez-Barbero, J. Solution Conformation Dynamics of a Tetrasaccharide Related to the Lewisx Antigen Deduced by 1H NMR NOESY, ROESY, and T-ROESY Measurements. *Carbohydr. Res.* **1997**, *300* (1), 3–10. https://doi.org/10.1016/S0008-6215(97)00034-7.

(329) Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. Structure Determination of a Tetrasaccharide: Transient Nuclear Overhauser Effects in the Rotating Frame. *J. Am. Chem. Soc.* **1984**, *106* (3), 811–813. https://doi.org/10.1021/ja00315a069.

(330) Macdonald, S. S.; Pereira, J. H.; Liu, F.; Tegl, G.; DeGiovanni, A.; Wardman, J. F.; Deutsch, S.; Yoshikuni, Y.; Adams, P. D.; Withers, S. G. A Synthetic Gene Library Yields a Previously Unknown Glycoside Phosphorylase That Degrades and Assembles Poly-β-1,3-GlcNAc, Completing the Suite of β-Linked GlcNAc Polysaccharides. *ACS Cent. Sci.* **2022**, acscentsci.1c01570. https://doi.org/10.1021/acscentsci.1c01570.

(331) Speciale, I.; Di Lorenzo, F.; Gargiulo, V.; Erbs, G.; Newman, M.; Molinaro, A.; De Castro, C. Biopolymer Skeleton Produced by Rhizobium Radiobacter : Stoichiometric Alternation of Glycosidic and Amidic Bonds in the Lipopolysaccharide O-Antigen. *Angew. Chemie* **2020**, *132* (16), 6430–6436. https://doi.org/10.1002/ange.201914053.

(332) Schneider, J. P.; Kelly, J. W. Templates That Induce α-Helical, β-Sheet, and Loop Conformations. *Chem. Rev.* **1995**, *95* (6), 2169–2187. https://doi.org/10.1021/cr00038a015.

(333) Payne, C. M.; Knott, B. C.; Mayes, H. B.; Hansson, H.; Himmel, M. E.; Sandgren, M.; Ståhlberg, J.; Beckham, G. T. Fungal Cellulases. *Chem. Rev.* **2015**, *115* (3), 1308–1448. https://doi.org/10.1021/cr500351c.

(334) Basa, S.; Nampally, M.; Honorato, T.; Das, S. N.; Podile, A. R.; El Gueddari, N. E.; Moerschbacher, B. M. The Pattern of Acetylation Defines the Priming Activity of Chitosan Tetramers. *J. Am. Chem. Soc.* **2020**, *142* (4), 1975–1986. https://doi.org/10.1021/jacs.9b11466.

(335) Wattjes, J.; Niehues, A.; Cord-Landwehr, S.; Hoßbach, J.; David, L.; Delair, T.; Moerschbacher, B. M. Enzymatic Production and Enzymatic-Mass Spectrometric Fingerprinting Analysis of Chitosan Polymers with Different Nonrandom Patterns of Acetylation. *J. Am. Chem. Soc.* **2019**, *141* (7), 3137–3145. https://doi.org/10.1021/jacs.8b12561.

(336) Weikert, T.; Niehues, A.; Cord-Landwehr, S.; Hellmann, M. J.; Moerschbacher, B. M. Reassessment of Chitosanase Substrate Specificities and Classification. *Nat. Commun.* **2017**, *8* (1), 1698. https://doi.org/10.1038/s41467-017-01667-1.

(337) Van Dyken, S. J.; Locksley, R. M. Chitins and Chitinase Activity in Airway Diseases. *J. Allergy Clin. Immunol.* **2018**, *142* (2), 364–369. https://doi.org/10.1016/j.jaci.2018.06.017.

(338) Naqvi, S.; Moerschbacher, B. M. The Cell Factory Approach toward Biotechnological Production of High-Value Chitosan Oligomers and Their Derivatives: An Update. *Crit. Rev. Biotechnol.* **2017**, *37* (1), 11–25. https://doi.org/10.3109/07388551.2015.1104289.

(339) Erwig, L. P.; Gow, N. A. R. Interactions of Fungal Pathogens with Phagocytes. *Nat. Rev. Microbiol.* **2016**, *14* (3), 163–176. https://doi.org/10.1038/nrmicro.2015.21.

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

(341) Wong, S. Y. Neoglycoconjugates and Their Applications in Glycobiology. *Curr. Opin. Struct. Biol.* **1995**, *5* (5), 599–604. https://doi.org/10.1016/0959-440X(95)80050-6.

(342) Godbe, J. M.; Freeman, R.; Burbulla, L. F.; Lewis, J.; Krainc, D.; Stupp, S. I. Gelator Length Precisely Tunes Supramolecular Hydrogel Stiffness and Neuronal Phenotype in 3D Culture. *ACS Biomater. Sci. Eng.* **2020**, *6* (2), 1196–1207. https://doi.org/10.1021/acsbiomaterials.9b01585.

(343) Merg, A. D.; Touponse, G.; van Genderen, E.; Zuo, X.; Bazrafshan, A.; Blum, T.; Hughes, S.; Salaita, K.; Abrahams, J. P.; Conticello, V. P. 2D Crystal Engineering of Nanosheets Assembled from Helical Peptide Building Blocks. *Angew. Chem., Int. Ed.* **2019**, *58* (38), 13507–13512. https://doi.org/10.1002/anie.201906214.

(344) Nowick, J. S. Exploring β-Sheet Structure and Interactions with Chemical Model Systems. *Acc. Chem. Res.* **2008**, *41* (10), 1319–1330. https://doi.org/10.1021/ar800064f.

(345) Rodriguez, J.; O'Neill, S.; Walczak, M. A. Constrained Saccharides: A Review of Structure, Biology, and Synthesis. *Nat. Prod. Rep.* **2018**, *35* (3), 220–229. https://doi.org/10.1039/C7NP00050B.

(346) Xie, J.; Bogliotti, N. Synthesis and Applications of Carbohydrate-Derived Macrocyclic Compounds. *Chem. Rev.* **2014**, *114* (15), 7678–7739. https://doi.org/10.1021/cr400035j.

(347) Ciani, B.; Jourdan, M.; Searle, M. S. Stabilization of β-Hairpin Peptides by Salt Bridges: Role of

Preorganization in the Energetic Contribution of Weak Interactions. *J. Am. Chem. Soc.* **2003**, *125* (30), 9038–9047. https://doi.org/10.1021/ja030074l.

(348) Shugrue, C. R.; Miller, S. J. Applications of Nonenzymatic Catalysts to the Alteration of Natural Products. *Chem. Rev.* **2017**, *117* (18), 11894–11951. https://doi.org/10.1021/acs.chemrev.7b00022.

(349) Metrano, A. J.; Miller, S. J. Peptide-Based Catalysts Reach the Outer Sphere through Remote Desymmetrization and Atroposelectivity. *Acc. Chem. Res.* **2019**, *52* (1), 199–215. https://doi.org/10.1021/acs.accounts.8b00473.

(350) Eller, S.; Collot, M.; Yin, J.; Hahm, H. S.; Seeberger, P. H. Automated Solid-Phase Synthesis of Chondroitin Sulfate Glycosaminoglycans. *Angew. Chem., Int. Ed.* **2013**, *52* (22), 5858–5861. https://doi.org/10.1002/anie.201210132.

(351) Gude, M.; Ryf, J.; White, P. D. An Accurate Method for the Quantitation of Fmoc-Derivatized Solid Phase Supports. *Lett. Pept. Sci.* **2002**, *9* (4), 203–206. https://doi.org/doi.org/10.1023/A:1024148619149.

(352) Ohara, K.; Lin, C.-C.; Yang, P.-J.; Hung, W.-T.; Yang, W.-B.; Cheng, T.-J. R.; Fang, J.-M.; Wong, C.-H. Synthesis and Bioactivity of β-(1→4)-Linked Oligomannoses and Partially Acetylated Derivatives. *J. Org. Chem.* **2013**, *78* (13), 6390–6411. https://doi.org/10.1021/jo4005266.

(353) Zegelaar-Jaarsveld, K.; Smits, S. A. W.; Van Der Marel, G. A.; Van Boom, J. H. Synthesis of a Pentasaccharide Corresponding to the Repeating Unit of the Exopolysaccharide from Cryptococcus Neoformans Serovar D. *Bioorganic Med. Chem.* **1996**, *4* (11), 1819–1832. https://doi.org/10.1016/S0968-0896(96)00164-2.

(354) Joseph, A. A.; Verma, V. P.; Liu, X. Y.; Wu, C. H.; Dhurandhare, V. M.; Wang, C. C. TMSOTf-Catalyzed Silylation: Streamlined Regioselective One-Pot Protection and Acetylation of Carbohydrates. *Eur. J. Org. Chem.* **2012**, No. 4, 744–753. https://doi.org/10.1002/ejoc.201101267.

(355) Fittolani, G.; Shanina, E.; Guberman, M.; Seeberger, P. H.; Rademacher, C.; Delbianco, M. Automated Glycan Assembly of 19 F-labeled Glycan Probes Enables High-Throughput NMR Studies of Protein–Glycan Interactions. *Angew. Chem., Int. Ed.* **2021**, *60* (24), 13302–13309. https://doi.org/10.1002/anie.202102690.

(356) Carr, H. Y. Steady-State Free Precession in Nuclear Magnetic Resonance. *Phys. Rev.* **1958**, *112* (5), 1693–1701. https://doi.org/10.1103/PhysRev.112.1693.

(357) Meiboom, S.; Gill, D. Modified Spin-Echo Method for Measuring Nuclear Relaxation Times. *Rev. Sci. Instrum.* **1958**, *29* (8), 688–691. https://doi.org/10.1063/1.1716296.

(358) Gim, S.; Fittolani, G.; Yu, Y.; Zhu, Y.; Seeberger, P. H.; Ogawa, Y.; Delbianco, M. Targeted Chemical Modifications Identify Key Features of Carbohydrate Assemblies and Generate Tailored Carbohydrate Materials. *Chem. – A Eur. J.* **2021**, chem.202102164. https://doi.org/10.1002/chem.202102164.

(359) Gim, S.; Fittolani, G.; Nishiyama, Y.; Seeberger, P. H.; Ogawa, Y.; Delbianco, M. Supramolecular Assembly and Chirality of Synthetic Carbohydrate Materials. *Angew. Chem., Int. Ed.* **2020**, *59* (50), 22577–22583. https://doi.org/10.1002/anie.202008153.

(360) Ning, S.; Zhang, L.; Ma, J.; Chen, L.; Zeng, G.; Yang, C.; Zhou, Y.; Guo, X.; Deng, X. Modular and Scalable Synthesis of Nematode Pheromone Ascarosides: Implications in Eliciting Plant Defense Response. *Org. Biomol. Chem.* **2020**, *18* (26), 4956–4961. https://doi.org/10.1039/d0ob00652a.

(361) Mo, K.-F.; Li, H.; Mague, J. T.; Ensley, H. E. Synthesis of the β-1,3-Glucan, Laminarahexaose: NMR and Conformational Studies. *Carbohydr. Res.* **2009**, *344* (4), 439–447. https://doi.org/10.1016/j.carres.2008.12.014.

(362) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26* (16), 1701–1718.

https://doi.org/10.1002/jcc.20291.

(363) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* **1985**, *31* (3), 1695–1697. https://doi.org/10.1103/PhysRevA.31.1695.

(364) Nosé, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.* **1984**, *81* (1), 511–519. https://doi.org/10.1063/1.447334.

(365) Parrinello, M.; Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Phys. Rev. Lett.* **1980**, *45* (14), 1196–1199. https://doi.org/10.1103/PhysRevLett.45.1196.

(366) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52* (12), 7182–7190. https://doi.org/10.1063/1.328693.

(367) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N ·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092. https://doi.org/10.1063/1.464397.

(368) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463–1472. https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

(369) Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13* (8), 952–962. https://doi.org/10.1002/jcc.540130805.

8 List of publications

Bottom-up approach to understand chirality transfer across scales in cellulose assemblies
G. Fittolani, D. Vargová, P.H. Seeberger, Y. Ogawa, M. Delbianco *J. Am. Chem. Soc.*, **2022**, 144, 27, 12469-12475

Neighboring Group Participation of Benzoyl Protecting Groups in C3- and C6-Fluorinated Glucose
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**, 10.1002/ejoc.202200255

The Flexibility of Oligosaccharides Unveiled Through Residual Dipolar Coupling Analysis
A. Poveda, **G. Fittolani**, P. H. Seeberger, M. Delbianco, J. Jiménez-Berbero *Front. Mol. Biosci.*, **2021**, 8, 784318

Progress and challenges in the synthesis of sequence controlled polysaccharides
G. Fittolani, T. Tyrikos-Ergas, D. Vargová, M. A. Chaube, M. Delbianco *Beilstein J. Org. Chem.* **2021**, 17, 1981-2025

Targeted chemical modifications identify key features of carbohydrate assemblies and generate tailored carbohydrate materials
S. Gim, **G. Fittolani**, Y. Yu, Y. Zhu, P. H. Seeberger, Y. Ogawa, M. Delbianco *Chem. Eur. J.* **2021**, 27, 52, 13139-13143

Identifying the Origin of Local Flexibility in a Carbohydrate Polymer
K. Anggara, Y. Zhu, **G. Fittolani**, Y. Yu, T. Tyrikos-Ergaas, M. Delbianco, S. Rauschenbach, S. Abb, P.H. Seeberger, K. Kern *Proc. Natl. Acad. Sci.* **2021**, 118, 23, e2102168118

Automated Glycan Assembly of ^{19}F-labeled Glycan Probes Enables High-Throughput NMR studies of Protein-Glycan Interactions
G. Fittolani,[+] E. Shanina,[+] M. Guberman, P.H. Seeberger, C. Rademacher, M. Delbianco *Angew. Chem. Int. Ed.* **2021**, 60, 24, 13302-13309

Systematic Structural Characterization of Chitooligosaccharides Enabled by Automated Glycan Assembly
T. Tyrikos-Ergas, V. Bordoni, **G. Fittolani**, M. A. Chaube, A. Grafmüller, P.H. Seeberger, M. Delbianco *Chem. Eur. J.* **2021**, 27, 7, 2321-2325

Supramolecular Assembly and Chirality of Synthetic Carbohydrate Materials
S. Gim, **G. Fittolani**, Y. Nishiyama, P.H. Seeberger, M. Delbianco *Angew. Chem., Int. Ed.* **2020**, 59, 50, 2577-22583

Structural Studies Using Unnatural Oligosaccharides: Toward Sugar Foldamers
T. Tyrikos-Ergas,[+] **G. Fittolani**,[+] P.H. Seeberger, M. Delbianco, *Biomacromolecules*, **2020**, 21, 1, 18-29

Systematic Hydrogen-Bond Manipulations to Establish Polysaccharide Structure-Property Correlations
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, 37, 13127-13132

Helical Polysaccharides
G. Fittolani, P.H. Seeberger, M. Delbianco *Pept. Sci.* **2019**, 112, 1, e24124

[+]equal contribution