COVALENT PROBES FOR FUNCTIONAL AND STRUCTURAL CHARACTERIZATION OF GLYCOSIDE HYDROLASES

by

Namrata Jain

M.Sc. Western University, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Chemistry)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

May 2020

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

COVALENT PROBES FOR FUNCTIONAL AND STRUCTURAL CHARACTERIZATION OF GLYCOSIDE HYDROLASES

submitted by Namrata Jain in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

in Chemistry

Examining Committee:

Prof. Harry Brumer, Professor, Chemistry, UBC
Supervisor

Prof. Reinhard Jetter, Professor, Chemistry, UBC
Supervisory Committee Member

Prof. David Perrin, Professor, Chemistry, UBC
University Examiner

Prof. Steven Hallam, Associate Professor, Microbiology and Immunology, UBC
University Examiner

Additional Supervisory Committee Members:

Prof. Katherine S. Ryan, Associate Professor, Chemistry, UBC
Supervisory Committee Member

Prof. Stephen G. Withers, Professor, Chemistry, UBC
Supervisory Committee Member

Abstract

Carbohydrates are ubiquitous in Nature and fundamental to the sustenance of organisms across all domains of life. Carbohydrates serve as sources and reserves of metabolic energy, participate in various cellular communication events, and provide structural support to plant and animal cells. Highly specific enzymes have evolved over several millennia to bind and manipulate carbohydrate substrates. Glycoside hydrolases (GHs) are a class of carbohydrate-active enzymes that cleave glycosidic linkages in complex carbohydrates. Organisms across all domains of life dedicate part of their genome to the production of GHs. New GHs are continually discovered through genome sequencing, while their structural and functional characterization, particularly in complex native environments, poses a persistent challenge to the dynamic field of GH characterization. One of the fundamental ways of ascribing protein function is the exploration of protein active sites, which can be used to deduce important details regarding substrate-enzyme interactions.

The work presented in this thesis describes the development of six new probes targeting a variety of GHs by irreversible covalent inhibition. These probes, developed on oligosaccharide scaffolds, feature either an *N*-bromoacetylglycosylamine electrophilic warhead or 2', 4' dinitrophenyl 2-deoxy-2-fluoro substitutions, facilitating irreversible inhibition of the target GH. The analysis presented in this work reveals key information about enzyme-inhibitor interactions through enzyme kinetic analyses, intact-protein mass spectrometry, and inhibitor-bound protein X-ray crystallography. Enzymes of diverse GH families and substrate preferences including e*ndo*-xyloglucanases, mixed-linkage glucanases, and β-(1,3) glucanases are featured to demonstrate the potency of this library of inhibitors. This small-molecule inhibitor toolkit targeting specific GH

enzymes has the potential to enhance our knowledge of the structural and functional characteristics of GHs and to provide a platform for activity-based enzyme profiling.

Lay Summary

Complex carbohydrates from diverse sources account for about 75% of the total biomass on earth. To be used as an energy source, carbohydrates are often required to be broken down into smaller parts. This task is accomplished in Nature by millions of highly specific carbohydrate-degrading enzymes (glycoside hydrolases or GHs). These enzymes have many applications such as in the bioethanol industry for producing fermentable sugar.

GHs have a variety of structures, specializations, and speed of carbohydrate digestion. Such properties can be evaluated with the help of biochemical tools that can predictably tune the activity of GHs. The work in this thesis describes the development of new GH inhibitors designed to quantitatively kill GH activity, facilitating the comparison of speed across various GHs and the prediction of their functions. Collectively, these inhibitors add to the growing field of biochemical tools that assist in the discovery of new industrially applicable enzymes.

Preface

Chapter 2: Synthesis and application of a highly branched, mechanism-based 2-deoxy-2-fluoro-oligosaccharide inhibitor of *endo*-xyloglucanases. This chapter was written by the author (Namrata Jain), who also performed the chemo-enzymatic synthesis of the inhibitor as well as all kinetics and active site labelling experiments. Dr. Mohamed Attia performed the recombinant protein production and site-directed mutagenesis of the enzyme. Wendy Offen solved the crystal structures, which were analyzed under the supervision of Prof. Gideon Davies. Prof. Harry Brumer (Ph.D. supervisor) reviewed and revised the manuscript. A version of this chapter has been published: **Jain, N., Attia, M. A., Offen, W. A., Davies, G. J., & Brumer, H.** (2018). Synthesis and application of a highly branched, mechanism-based 2-deoxy-2-fluoro-oligosaccharide inhibitor of *endo*-xyloglucanases. *Organic & Biomolecular Chemistry*, **16**(45), 8732–8741 (with a correction published as **Jain, N., Attia, M. A., Offen, W. A., Davies, G. J., & Brumer, H.** (2019). Correction: Synthesis and application of a highly branched, mechanism-based 2-deoxy-2-fluoro-oligosaccharide inhibitor of *endo*-xyloglucanases. *Organic & Biomolecular Chemistry*, **17**(2), 398-398.)

Chapter 3: This chapter was written by the author (Namrata Jain), who also performed the chemo-enzymatic synthesis of the inhibitors as well as all kinetics and active site labelling experiments. VvEG16 (ΔV152) was provided by Dr. Nicholas McGregor. Kazune Tamura performed recombinant protein production of BoGH16 and crystal structure determination, under the supervision of Prof. Filip van Petegem and Prof. Harry Brumer. Prof. Harry Brumer (Ph.D. supervisor) reviewed and revised the manuscript. A version of this chapter will be submitted for publication.

Chapter 4: This chapter was written by the author (Namrata Jain), who also performed the chemo-enzymatic synthesis of the inhibitors and β-2'-chloro-4'-nitrophenyl laminaribioside, the inactivation kinetics, and the intact-protein mass spectrometry experiments. The BuGH158 and BuGH16 proteins were produced by Dr. Guillaume Déjean, and BoGH16 was produced by Kazune Tamura. Prof. Harry Brumer (Ph.D. supervisor) reviewed and revised the chapter.

β-2'-chloro-4'-nitrophenyl laminaribioside, the synthesis of which is described in this chapter, was used for the kinetic characterization of the *Bacteroides ovatus* GH16 mixed-linkage *endo*-glucanase in the publication: **Tamura, K. *et al.*** (2017) Molecular mechanism by which prominent human gut Bacteroidetes utilize mixed-linkage beta- glucans, major health-promoting cereal polysaccharides. *Cell Reports*, **21**, 417–430.

The determination of the reaction stereoselectivity of GH158 by NMR, performed by the author together with Kazune Tamura, has been published in **Déjean, G. *et al.*** (2020) Synergy between cell-surface glycosidases and glycan-binding proteins dictates the utilization of specific beta (1,3)- glucans by human gut Bacteroides. *mBio*, **11**, e00095-20.

A version of this chapter will be submitted for publication.

Table of Contents

List of Figures

List of Schemes

List of Abbreviations

2,4-DNFB: 2, 4 dinitrofluorobenzene

2F-DNP: 2', 4'-dinitrophenyl-2-deoxy-2-fluoroglycoside

ABP: Activity-based probe

ABPP: Activity-based protein profiling

Ac: Acetate

ATT: 6-Aza-2-thiothymine

BODIPY: Boron-dipyrromethene

BSA: Bovine serum albumin

CAZy: Carbohydrate-active enzyme

CAZymes: Carbohydrate-active enzymes

CNP: 2-Chloro-4-nitrophenyl

DABCO: 1,4-Diazabicyclo [2.2.2] octane

DCM: Dichloromethane

DMF: *N, N*-dimethylformamide

DNP^-: Dinitrophenolate

DNP: 2,4-Dinitrophenyl

EC: Enzyme Commission

ESCI: Electrospray chemical ionization

ESI: Electrospray ionization

Gal: Galactose

GHs: Glycosyl hydrolases

Glc: Glucose

GlcA: Glucuronic acid

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HPAEC-PAD: High-performance anion exchange chromatography- pulsed amperometric detector

HPLC: High-performance liquid chromatography

HPSEC-UV: High- performance size-exclusion chromatography with UV detection

HRMS: High-resolution mass spectrometry

LB: Lysogeny broth

LC-MS: Liquid chromatography coupled to mass spectrometry

Ld: *Laminaria digitata*

m/z: Mass-to-charge ratio

MALDI-TOF: Matrix-assisted laser desorption ionization-time of flight

Man: Mannose

MBI: Mechanism-based inhibitor

MLG: Mixed-linkage glucan

MS: Mass spectrometry

NaCNP: 2-chloro-4-nitrophenol

NMR: Nuclear magnetic resonance

PCR: Polymerase chain reaction

PDB: Protein Data Bank

PEG: Polyethylene glycol

Q-TOF: Quadrupole time-of-flight

SDS-PAGE: Sodium dodecyl sulfate–polyacrylamide gel electrophoresis

TLC: Thin-layer chromatography

UPLC: Ultra performance liquid chromatography

UV: Ultraviolet

XDS: X-ray Detector Software

XyG: Xyloglucan

Xyl: Xylose

Acknowledgements

I would like to express my sincere gratitude to my parents, Dr. Dinesh Jain and Sarita Jain, who have always been supportive of my endeavours and given me the independence and confidence to achieve all I strive to. My sustenance is driven by their continued love and trust from afar. My brother Dr. Vinamra Jain deserves my immense gratitude for inspiring me with his genius and leading the way to success for me.

Many thanks to my Ph.D. supervisor Harry, who has provided valuable guidance and encouragement throughout the past years, and taught me not only how-to science but also how to navigate through life using the skills I have developed as a graduate student. I am utterly inspired by his patience, time- and personnel- management skills, and scientific writing. I also thank him for creating a supportive and friendly lab environment, as well as setting a high standard of academic training.

I appreciate the time I have spent with the past and present members of the Brumer lab. A big thanks to Mohamed, Nick, Guillaume, and Greg for providing me with valuable mentorship during my first couple of years in the lab, and each being a trustworthy friend and confidant. Shaheen, in the role of our lab manager, has been the most reliable, supportive, and thoughtful person and a constant source of positivity for me. I want to thank Chang for all his help with my chemistry-related queries, and his warm and friendly nature.

I want to extend my sincere appreciation to all the Brumer lab ladies. Maria, Julie, Stephanie, and Laleh are people I have learnt from a lot. Hila has been my lifeline over the past few years inside and outside the lab, and I appreciate our constant check-ins and hugs. I cherish all the great conversations I have had with these women and appreciate that we uplift each other without fail.

A shoutout to all the current and former members of GreenChem UBC. Starting and leading this student group was a wonderful supplement to my graduate experience and I thank everyone who has helped keep the ship afloat. I also thank the UBC Sustainability Initiative for letting me be a part of the Sustainability Ambassadors program, through which I have met many folks who share my passion for sustainability.

I would also like to extend gratitude to Jason Rogalski for his help in intact-MS experiments, Dr. Maria Ezhova for assistance with NMR experiments, and Dr. Yun Ling for help in MALDI and mass spectrometry studies. I want to acknowledge Dr. Hongming Chen (Withers Lab, UBC) for helpful carbohydrate-chemistry related discussions when I was still learning the ropes. I am also grateful to Dr. Zachary Armstrong for providing the glycosynthase proteins for some preliminary experiments. Many thanks to Dr. Gideon Davies and Wendy Offen for their work on crystallography.

I am grateful to the members of my supervisory committee Dr. Katherine Ryan, Dr. Reinhard Jetter, and Dr. Stephen Withers, who have provided valuable feedback over the years and helped me shape my project into a logical scientific story.

Last but foremost, I want to thank my (soon-to-be) husband Nihar, who has been my rock throughout the past year. His patience, affection, and encouragement got me through the most challenging aspects of my graduate student life. I am eternally grateful to have a partner who fuels my motivation and cheers me on to a path of success.

Dedication

I would like to dedicate this thesis to my younger brother Dr. Vinamra Jain, who inspires me to embrace my fears and limitations so they don't get in the way of accomplishments.

Chapter 1: Introduction

1.1 The Role of Complex Polysaccharides

Carbohydrates are one of the most important classes of biomolecules and serve a variety of essential functions in Nature. They serve as sources and reserves of metabolic energy, participate in various intracellular transport and recognition processes, provide structural support to plant and animal cells, and are essential components of many biologically important natural products such as glycolipids and glycoproteins[1–6]. However, exploration and utilization of carbohydrates are heavily restricted by their extensive structural complexity. Despite being composed of a handful of elements, carbohydrates display a uniquely high structural diversity through tremendous variation in their structure, molecular weight, stereochemical configuration, and linkages. A hexasaccharide, for example, is calculated to have 10^{12} possible linear and branched forms[7], although the actual diversity of the structures of naturally occurring glycans is likely to be only a small fraction of that number[8].

1.1.1 Plant Cell Wall Polysaccharides

Plant cell walls are rich in carbohydrates and account for one of the largest stores of organic carbon on earth. They are crucial in providing rigidity, flexibility, and shape to plants, storing metabolic energy, protecting against pathogens as well as playing a vital role in recognition and signalling events[9]. Carbohydrates are abundantly found in the thin, flexible, and highly permeable primary cell walls that surround growing cells, as well as in thicker secondary cell walls that develop in some cells after expansion to provide additional strength and rigidity[6]. In the complex network of glycans within the plant cell walls, paracrystalline cellulose microfibrils are woven into

a matrix of hemicelluloses, lignins, and pectins[6] **(Figure 1.1)**. Some important plant cell wall polysaccharides are highlighted below.

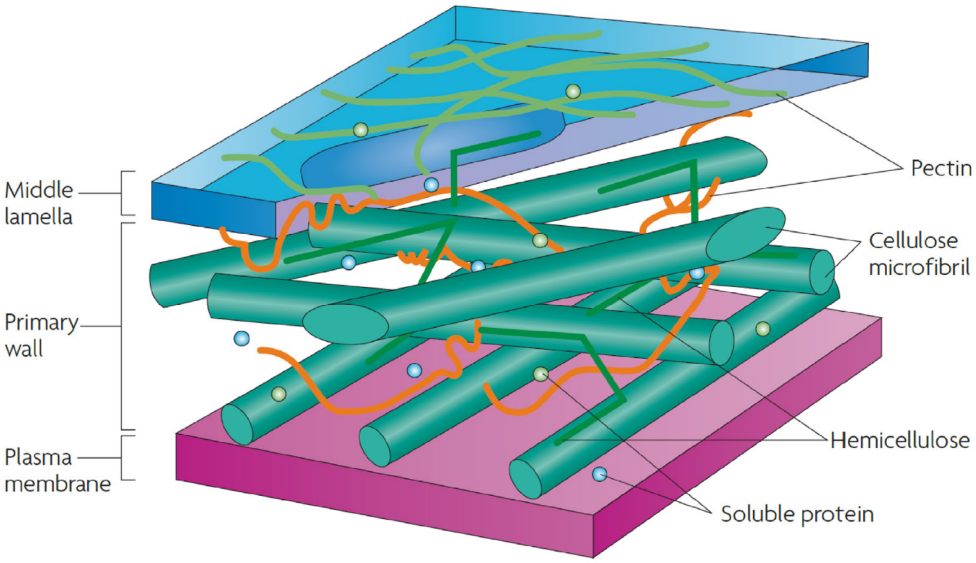


Figure 1.1: Schematic model for plant cell wall structure (reused with permission from Sticklen, 2008[10])

1.1.1.1 Cellulose

Cellulose is the most abundant biopolymer on earth and an immense resource of renewable biomass. It is an important structural constituent of plant cell walls and makes up 35-50% of dry plant matter in land plants[11]. Cellulose is a linear homopolysaccharide made up of glucose residues linked through β-(1,4) glycosidic linkages where every other glucose is rotated 180° along the chain axis. **(Figure 1.2)**. The molecular weight of cellulose ranges from 100 to 20,000[12]. The cellulose chains are held together via intermolecular hydrogen bonds and van der Waals forces, and form microfibrils with a crystalline or semi-crystalline lattice which can extend up to a few micrometres in length and 2-10 nm wide[13].

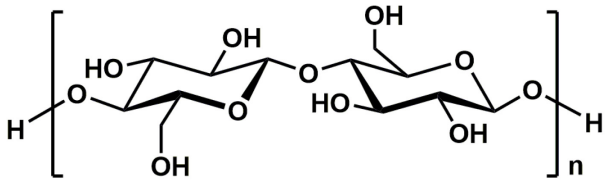


Figure 1.2: Representative structure of cellulose

1.1.1.2 Hemicelluloses

Hemicelluloses are a broad category of polysaccharides that strengthen the plant cell wall by cross-linking cellulose microfibrils via non-covalent interactions[14]. In some cases, hemicelluloses also serve as seed storage carbohydrates[15]. Their composition, abundance, and structures vary widely between different species and cell types, but they are typically amorphous polysaccharides composed of highly branched and variably linked uncharged monosaccharide units[16]. Two important hemicelluloses of relevance to this thesis are xyloglucan and mixed-linkage β-glucan.

1.1.1.2.1 Xyloglucan

Xyloglucan (XyG) is one of the major constituents of the plant cell wall hemicelluloses. It is present in all terrestrial plants in variable amounts[14,17] and is the most abundant hemicellulose in dicots and non-commelinoid monocots, constituting 20-25% of their dry cell wall weight[18]. XyG plays a distinct role in various plants in the form of structural and storage polysaccharides[15]. The primary role of XyG in the plant cell wall is to form a sheath around cellulose microfibrils to prevent them from aggregating[19], as well as to support cell enlargement during growth[20]. In many primary cell walls, the cellulose–xyloglucan network is the major load-bearing structure[21]. XyG is also potentially an important feedstock for second-generation bioethanol production[22], as well as an additive in the papermaking process to facilitate increased adhesion and lowered friction between cellulose fibers[23].

The canonical structure of XyG consists of a linear chain of β-(1,4) linked D-glucose residues which are heavily substituted by α-(1,6) D-xylosyl side groups. This moiety can be further substituted with D- and L-galactosyl, L-fucosyl, D-galacturonosyl, L-arabinopyranosyl, and L-arabinofuranosyl moieties. Acetyl groups are often connected to the backbone as well as the

branching residues[16,24]. The substitutions on xyloglucan are highly variable based on the source of the polysaccharide[25]. Owing to the highly complex structure of this polysaccharide, a shorthand nomenclature is commonly used to represent individual glucosyl residues of XyG, wherein G denotes an unbranched glucose residue, X (Xylp-α(1, 6)-β-Glcp) denotes an α-(1,6) D-xylose substituted glucose, and L (Galp-β(1, 2)-Xylp-α(1, 6)-β-Glcp) represents further β-(1,2) D-galactose substitution on this xylosyl group[26] **(Figure 1.3)**.

Predominant forms of XyG contain a repeating unit consisting of three α-(1,6) D-xylose-substituted glucoses followed by one unsubstituted glucose XXXG[27]. When further substituted by galactosyl and fucosyl residues, it results in fucogalactoxyloglucan which is found in lettuce, carrot, and soybean[25] **(Figure 1.3)**.

Xyloglucan depolymerization is highly important in plant cell wall remodelling during various stages of plant growth and is necessary for the production of fermentable monosaccharides from XyG[28,29]. Hence, there are numerous examples of xyloglucan-degrading enzymes (xyloglucanases) of prokaryotic and eukaryotic origin[20,28,30–34].

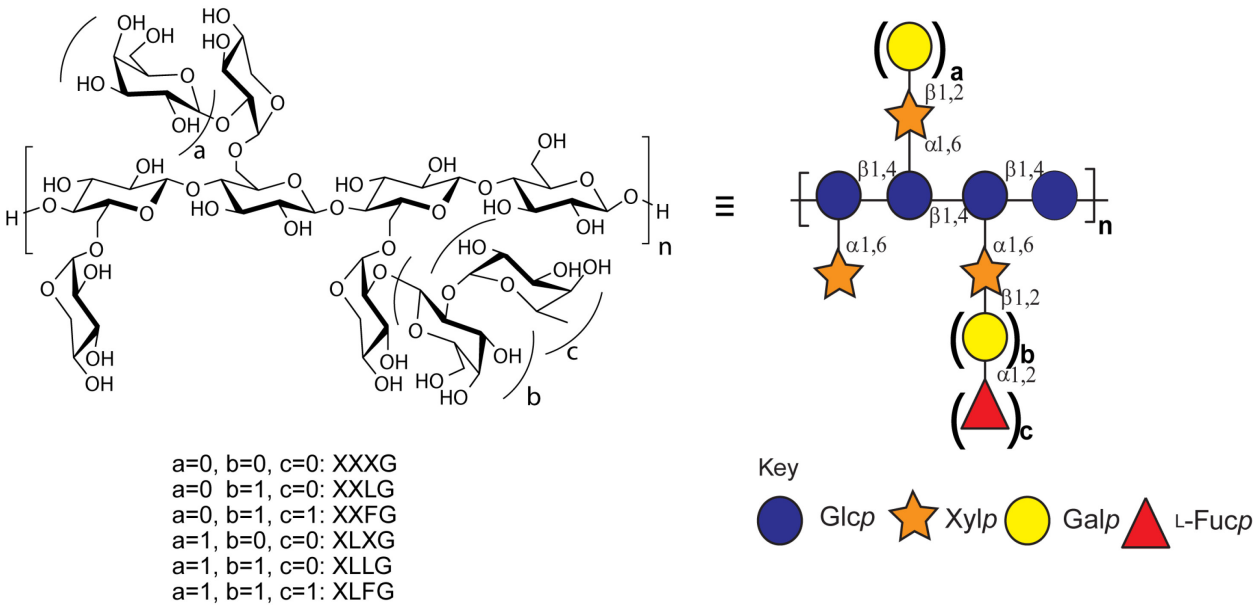


Figure 1.3: Representative structure of dicot xyloglucan. Substructure nomenclature is according to Tuomivaara, 2015 and symbols are according to the Consortium for Functional Glycomics (reused from Jain, 2018)

1.1.1.2.2 Mixed-linkage β-glucan

Mixed-linkage β-glucans (MLGs) are a major component of the cell walls of commelinoid monocots such as oat and barley[35], as well as plants of the genus *Equisetum*[36]. MLGs are typically more abundant in the primary cell walls than in the lignified secondary cell wall[37]. As important components of dietary fibre, the intake and modulation of MLGs is a critical factor in the overall digestive health, innate immunity, and disease outcome in mammalian systems[38–41].

Structurally, MLG is composed of glucosyl units connected with β-(1,4) glycosidic bonds which are regularly interspersed with β-(1,3) linkages[42,43]. The nomenclature of MLG denotes β-(1,4) linked glucosyl residues as G and β-(1,3) linked glucosyl residues as G3[43–45]. MLG is dominated by cellotriosyl and cellotetrasyl units linked by β-(1,3) linkages, but longer β-(1,4)-linked segments also occur[46] **(Figure 1.4).** Typically, about 25-30% of the total linkages are β-(1,3)[42], giving the polysaccharide a kinked shape and leading to the formation of a gel-like matrix

that facilitates wall elongation and plays an important role in supporting plant cell wall growth[47]. Mixed-linkage β-glucan can be depolymerized in nature by enzymes such as mixed-linkage glucanases (MLGases)[43,48,49]. MLGases are found across all domains of life and serve essential functions such as assist mammalian gut bacteria in the digestion of dietary MLG[44].

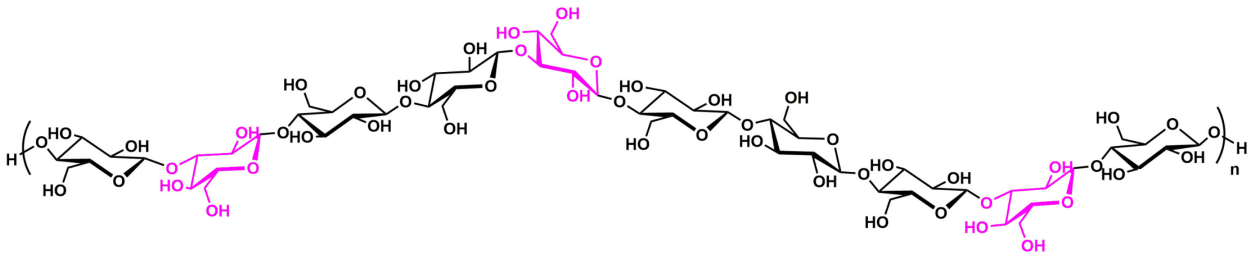


Figure 1.4: Representative structure of mixed-linkage β-glucan (MLG): The glucosyl residues in purple are substituted at the C3 position with a β-(1,3) linkage, other residues are linked β-(1,4).

1.1.2 β-(1,3) Glucans

β-(1,3) glucans are linear polysaccharides made up of chains of contiguous β-(1,3) glucosyl residues, which are often substituted with β-(1,6) glucoses. The degree of polymerization and branching varies with the specific plants, algal, fungal, and bacterial sources[50,51].

An example of a higher plant linear β-(1,3) glucan is callose. Callose has many important roles in various stages of plant development, such as separation of developing pollen grains to prevent their underlying walls from fusing, and are a major constituent of pollen tubes[52,53]. Similarly, fungal β-(1,3) glucan pachyman is a major component of the sclerotia of *Poria cocos,* a medicinal fungus with myriad pharmaceutical properties[54]. Pachyman is made up of a β-(1,6) -branched β-(1,3) glucose main chain[55], with some internal β-(1,6) bonds also reportedly present[55,56].

Bacterial polysaccharide curdlan is an unbranched β-(1,3) glucan known for its high commercial applicability due to its thermal gelation properties[57,58]. It is produced by the

fermentation of *Agrobacterium* strains from soil[59]. Both pachyman and curdlan have a high degree of polymerization (approximately 250 and 450 respectively)[60].

Algal β-(1,3) glucans, such as laminarin, are essential components of brown algae, microalgae and phytoplankton[57,61,62] and are one of the most abundant carbohydrates in the marine ecosystem[62]. Laminarin is a relatively low molecular weight storage polysaccharide typically containing 20 to 30 glucose residues[63]. This polysaccharide is composed of β-(1,3) linked glucosyl residues which are sparsely substituted with β-(1,6) glucosyl branching[63] with some chains terminated by mannitol end-groups[61,62,64] **(Figure 1.5)**. There are variations in the degree of branching, polymerization and the ratio of different linkages, according to the algae species[65]. Laminarin has anti-apoptotic[66] and immunostimulatory[38,67] properties and is an excellent source of renewable feedstock for the biofuel industry due to the low percentage of lignin in macroalgae such as seaweeds[68,69].

Due to the diverse structures and origins of β-(1,3) glucans, many enzymes have evolved to catalytically depolymerize them[70–72]. β-(1,3) glucanase enzymes have applications in yeast extract production[73] and are important in many plant functions such as seed development, germination, and defence against pathogens[74].

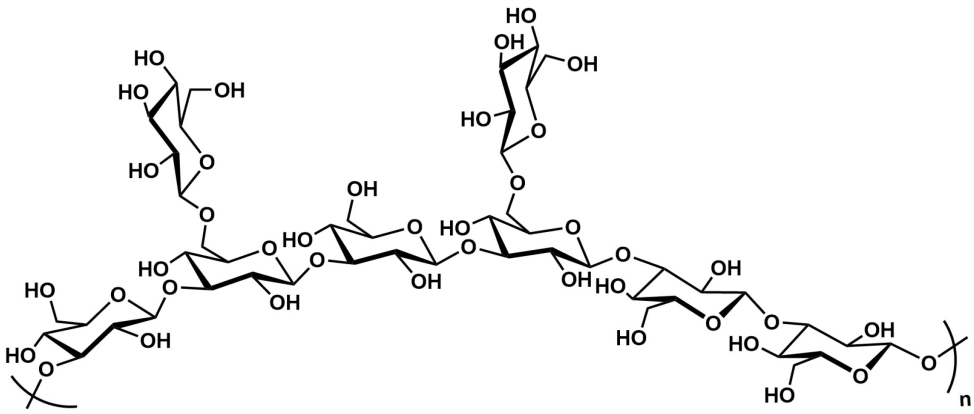


Figure 1.5: Representative structure of laminarin from brown algae *Laminaria digitate*

1.2 Glycoside Hydrolases as Carbohydrate-Active Enzymes

Carbohydrate-active enzymes (CAZymes) are highly specific proteins that have evolved over several millennia across Kingdoms to perform the assembly, degradation, isomerization, binding, and reshuffling of carbohydrates in an extraordinarily efficient manner. Some CAZymes are known to accelerate the rate of carbohydrate hydrolysis by as much as 10^{17} fold as compared to spontaneous hydrolysis with highly exquisite substrate specificity[75,76]. These proteins have diverse structures, stability, functions, specificities, and mechanisms of action. To facilitate their analysis, they have been classified in various ways, for example, based on specificity (EC classification[77]) and amino acid sequences (CAZy database[78]).

1.2.1 CAZy Database

The Carbohydrate-Active Enzyme database (CAZy database) is an actively curated resource for CAZyme discovery and organization that groups CAZymes into various classes and families based on sequence similarity[78]. Since there is a direct correlation between amino-acid sequence and overall fold, this classification enables the prediction of structural features of an enzyme, as well as the mechanism of action and evolutionary pathway.

The CAZy database classifies CAZymes into five main classes: glycosyl hydrolases[79], glycosyltransferases[80], polysaccharide lyases[81,82], carbohydrate esterases[83] and auxiliary activity[84], with an additional class for the non-catalytic carbohydrate-binding modules[85]. Each class is further subdivided into multiple families (and sometimes subfamilies[82,86–89]) based on sequence and structural similarities. The fold, active site topology, and mechanism of action is typically shared between the members of the same family[90], making the prediction of catalytic residues and structural information among uncharacterized members easier via sequence alignment with a characterized member[91].

The largest class of CAZymes is the glycoside hydrolases (GH), which catalyze the cleavage of glycosidic bonds either between saccharides or between a saccharide and a non-sugar molecule (aglycone). This cleavage is followed by hydrolysis or transglycosylation, in a wide range of glycoconjugates. GHs are classified into over 160 distinct families based on many thousands of known sequences and folding similarities[78]. Many of the sequence-based families are polyspecific, i.e. they contain enzymes of different substrate specificities, and often CAZymes with similar specificities are found in different families[79]. Typically about 1-3% of the genome of any organism encodes GHs[92]. Families with multiple activities are further divided into subfamilies, providing finer details of the evolution and substrate specificity of those families[88,89].

GHs can be further classified based on their mode of action as *endo*- or *exo*-acting. *Endo*-acting GHs cleave non-terminal glycosidic linkages on conjugated or free oligo- or polysaccharides, whereas *exo*-acting GHs cleave terminal linkages. These modes of actions are reflected in the surface topologies of GHs, whereby *exo*-acting GHs often display a deeper binding pocket to be able to accommodate the terminal end of the substrate, while *endo*-acting GHs contain a shallow groove or tunnel at the surface[90] **(Figure 1.6)**. *Endo*-acting GHs may be either *endo*-dissociative or *endo*-processive. *Endo*-dissociative GHs detach from the polysaccharide chain after each hydrolytic cleavage, and hence the enzyme produces a random mixture of mid-range molecular weight products at the beginning of the reaction. In contrast, *endo*-processive GHs bind the polysaccharide chain, catalyze the hydrolysis and slide along the polysaccharide chain without detaching to continue the hydrolysis[93].

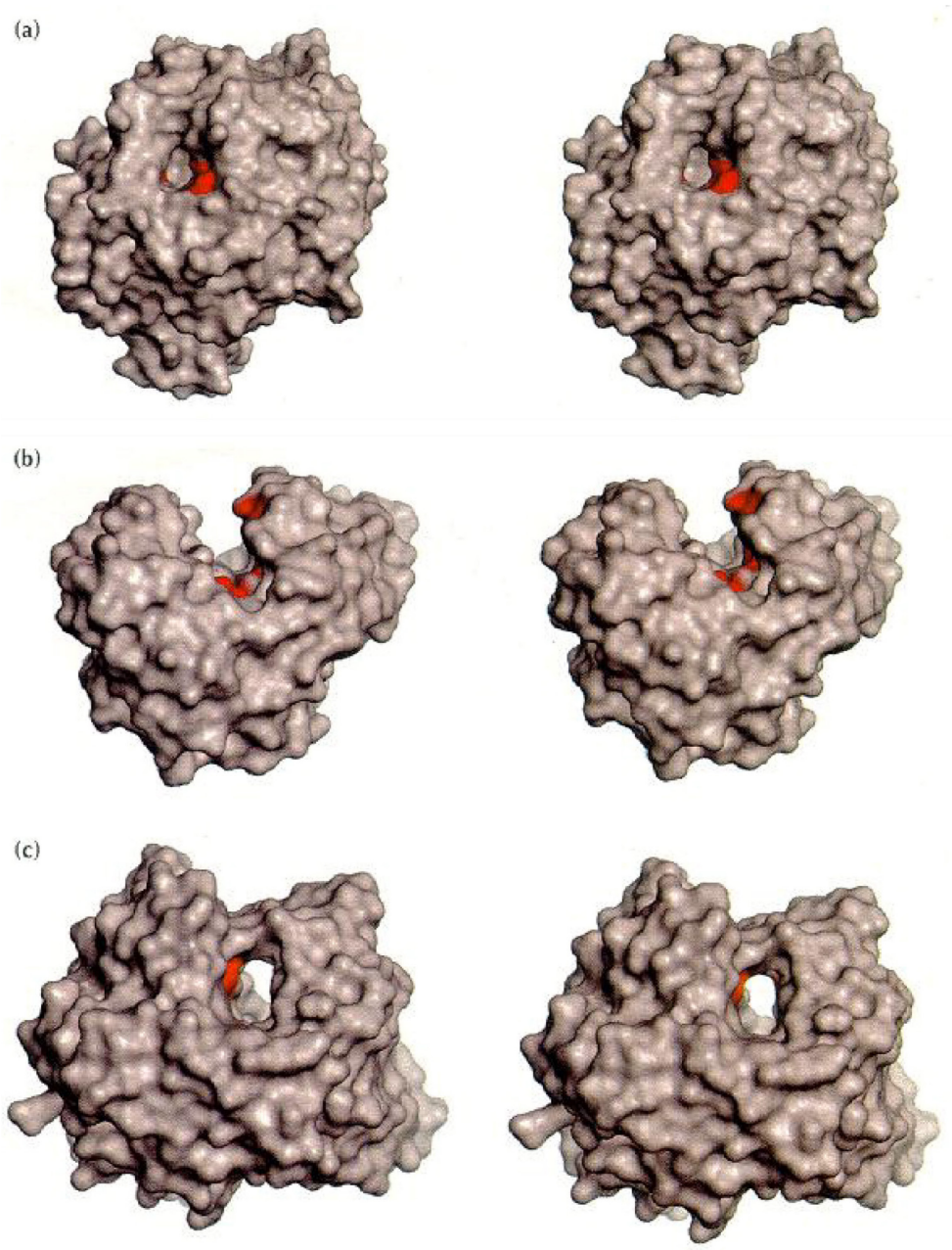


Figure 1.6: Types of active sites found in glycoside hydrolases : (a) The pocket (b) The cleft (c) The tunnel. In all panels, the proposed catalytic residues are indicated in red (reused with permission from Davies and Henrissat, 1995[90])

1.2.2 Mechanism of GH Catalysis

GHs can also be classified based on the stereospecificity of the catalysis product relative to the substrate. GHs generally act on their substrate via one of two mechanisms, inverting or

retaining, with a net inversion or retention of configuration at the anomeric carbon with respect to their substrate, respectively[90,94] **(Figure 1.7)**. Catalysis is made possible by two active site carboxylic acid/carboxylate residues. Several variations in these mechanisms have been found including substrate-assisted catalysis[95], alternative nucleophiles[96], absent key residues[97], and use of NAD-cofactor[98], but such variations are rare and out of the scope of this thesis.

Inverting enzymes progress through a single step, direct displacement mechanism where a carboxylate residue acts as a general base, deprotonating a water molecule that attacks the anomeric centre. The departure of the leaving group is assisted by protonation from a second carboxylic acid group which acts as a general acid **(Figure 1.7a)**. This reaction proceeds via an oxocarbenium ion-like transition state[76] with net inversion of anomeric stereoconfiguration. The distance between the two carboxylic acid residues is typically 6-12 Å.

Retaining GHs proceed through a two-step, double displacement mechanism involving a covalent glycosyl-enzyme intermediate **(Figure 1.7b)**. As for the inverting mechanism, both steps proceed via an oxocarbenium ion-like transition state[76]. In the retaining mechanism, the anomeric centre of the substrate is attacked by a deprotonated carboxylate amino acid (the nucleophilic residue), leading to a covalently-linked substrate-enzyme intermediate with inversion of configuration at the anomeric centre. The departing group is protonated by a second carboxylic amino acid (the catalytic acid/base residue). In the next step, the acid/base residue acts to deprotonate an incoming nucleophile (usually water or a hydroxyl group from another saccharide molecule). This deprotonated nucleophile, in turn, cleaves the covalently-linked intermediate to complete the catalytic cycle with net retention of configuration at the anomeric centre. The two carboxylic acid residues are separated by approximately 5 Å.

Even though the active site carboxylic acid residues play a key role in catalysis, it is believed that the acceleration of reaction rate is achieved by the stabilizing non-covalent interactions between the enzyme and the substrate in the oxocarbenium ion-like transition states[99].

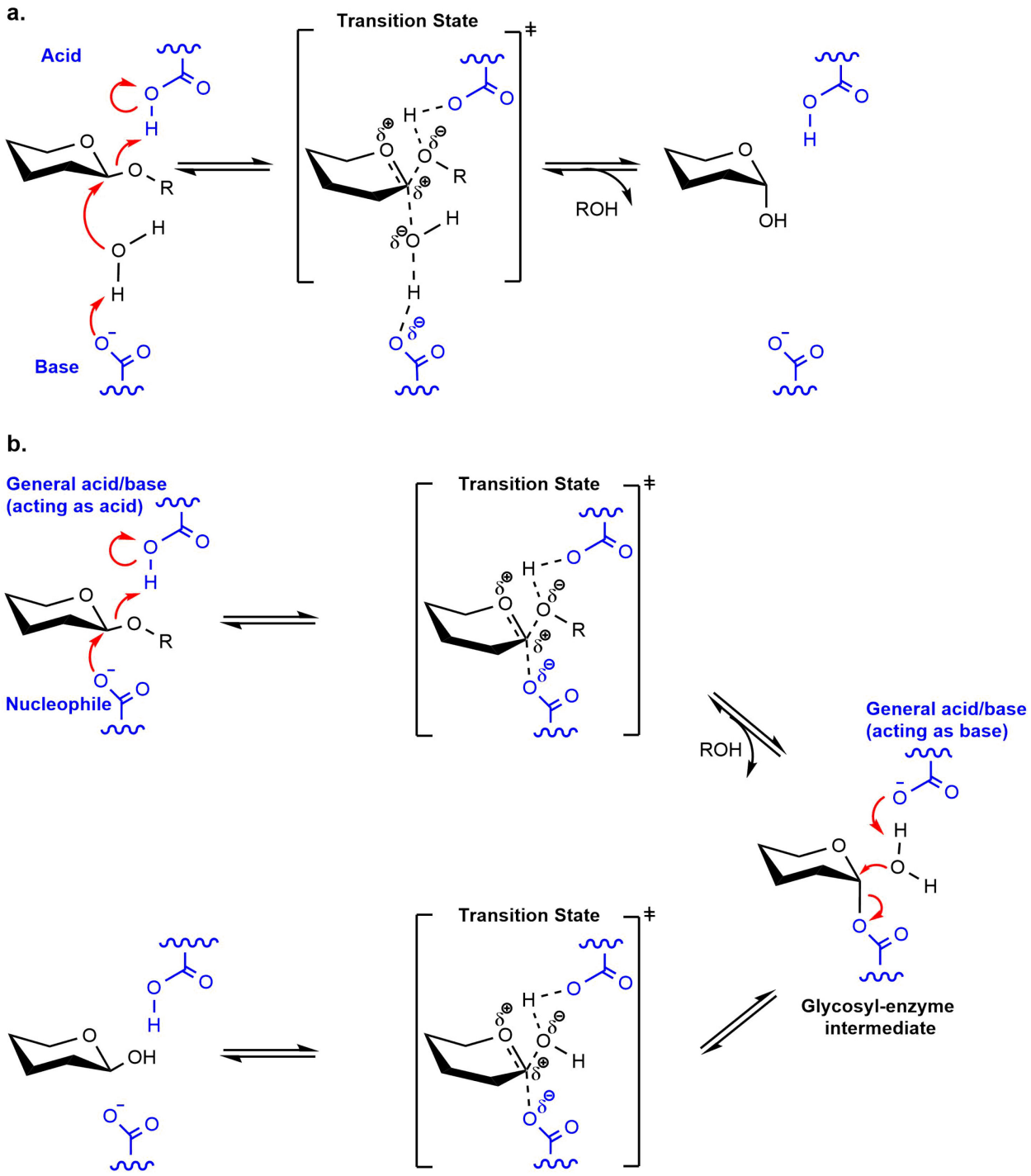


Figure 1.7: Classical Koshland mechanism of action of GHs: (a) Inverting GHs (b) Retaining GHs

During catalysis by GHs, individual substrate-binding sites on the enzyme are referred to as subsites (n) and are numbered from the point of bond cleavage according to a -n to +n system[100]. The subsites towards the non-reducing end of the substrate are labelled -1, -2, -3, etc. while those toward the reducing end are labelled as +1, +2, +3. Thus, glycosidic bond cleavage occurs between the −1 and +1 subsites[100].

1.2.3 Industrial Importance of Glycoside Hydrolases

Due to their excellent capacity to depolymerize complex carbohydrates, GHs are widely applied in many industries. For example, GHs involved in cellulose and hemicellulose biodegradation are widely used to overcome the recalcitrance of natural lignocellulosic materials for producing sustainable biobased products and biofuels[101–103]. As well, some GHs have applications in animal feed production, wherein they break down various hemicelluloses in the ingredients of the feed to reduce the viscosity of the raw material[104], in the pulp and paper industry for bleaching of cellulose pulp[105], and as key ingredients in some laundry detergents[106,107]. Other hemicellulose-degrading enzymes, such as mixed-linkage glucanases (MLGases), are used to alleviate problems such as hazing in the brewery industry due to the viscosity of their substrates[108].

Laminarinases have been reported to have usage for commercial yeast extract production[73], in malting and brewing processes[109], and in fungal and yeast cell wall degradation[110]. Other β-(1,3) glucanases have important applications in bioethanol production[111], against fungal infections[112], and to produce high-value oligosaccharides[113,114].

1.2.4 Polysaccharide Utilization in Bacteria

One of the most important roles of GHs is their capability to assist carbohydrate degradation in microbes, offering some of the greatest catalytic rate enhancements among

enzymes. In particular, microbes that reside in the lower digestive tract of mammals are of considerable importance to human nutrition[115,116], as the intrinsic capacity of mammals to absorb and utilize complex dietary polysaccharides is highly limited. Hence, there is significant research interest in the mechanisms of complex dietary polysaccharides metabolization by bacteria and their impact on human health[117].

Within the human distal colon, *Prevotella* and *Bacteroides* are important genera of the ubiquitous bacterial Phylum Bacteroidetes[118]. In particular, *Bacteroides thetaiotaomicron*[119] has served as a model species in understanding polysaccharide digestion in the human gut and its proteomic analysis has greatly added to our knowledge of the mechanism of action of *Bacteroides*[120]. Similarly, *B. ovatus* and *B. uniformis* have recently been the focus of similar research, owing to their capacity to utilize a multitude of polysaccharides[32,44,121–123]. Polysaccharide-degrading bacteria residing in the soil are also prolific polysaccharide-degraders, such as the Gram-negative soil saprophyte *Cellvibrio japonicus*, whose complete genome sequencing and genetic methods development showed immense potential for complex carbohydrate degradation, making it an important organism for CAZyme discovery[124,125].

1.3 Small Molecule Inhibitors of GH

Despite a large number of curated sequences of GHs in the CAZy database, only a small percentage have been functionally and structurally assigned, underpinning a pressing need for new tools targeting the investigation of GHs. Indeed, only about 2% of the enzymes in the CAZy database are functionally characterized, and the number of CAZymes with crystallographic information available is less than 0.25%[126,127]. Hence, novel approaches to characterize and discover new functional GH enzymes are highly important.

Over the past few decades, the structural, functional, specificity, and mechanistic characterization of GHs has been a major area of research in the field of carbohydrate enzymology. However, there is a widening gap between sequence prediction and functional analysis[127] (**Figure 1.8**). As such, there is an increasing need for the development of enzyme-binding probes that can measure the proficiency of GHs by quantifiably altering their natural function. Highly potent probes could provide insights into the active site of newly discovered GHs. Such insights could assist in the identification of important catalytic residues and determination of the mechanisms of action, preferred substrate configurations, and catalytic efficiencies of these enzymes.

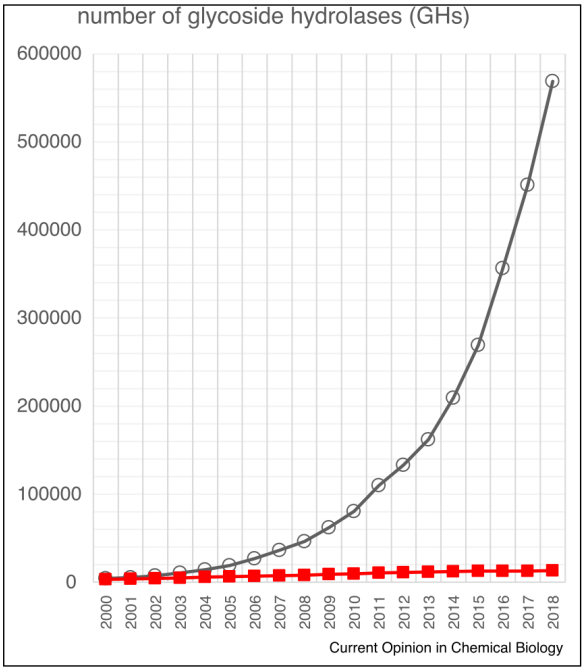


Figure 1.8: Growth of the number of GHs in the CAZy database. Black: number of sequences; red: number of biochemically characterized entries (reused with permission from Garron and Henrissat, 2019[127])

Enzyme inhibitors are a class of such probes that are capable of reducing enzyme activity. Inhibitors can be designed to slow or kill the activity of GHs and thus assist in enzyme kinetic analysis. The data obtained through inhibition kinetics help in understanding the specificity of GHs, as well as in identifying enzyme residues directly involved in catalysis, and therefore shed

light on the details of enzyme mechanisms. Often, GH inhibitors are widely used as molecular tools in chemical biology and drug discovery and many drugs are indeed enzyme inhibitors[128,129]. In general, inhibitors are classified as either reversible or irreversible, according to their binding modes and the nature of their interaction with the enzyme.

1.3.1 Reversible Inhibitors of GHs

Reversible inhibitors of GHs bind the enzyme non-covalently, primarily through hydrogen and ionic bonds, and hydrophobic interactions[130]. Various classes of reversible inhibitors of GHs are based on the incorporation of a nitrogen atom to substitute the internal ring oxygen, resulting in the iminosugar type inhibitors (such as Nojirimycin[131], deoxynojirimycin[132] and castanospermine[133]). Alternatively, the nitrogen can substitute the anomeric carbon and the ring oxygen can be replaced by a carbon atom (1-azasugar type inhibitors such as isofagomine[134]) **(Figure 1.9).**

Another class of reversible GH inhibitors is based on glycosylamine derivatives such as valienamines[135] and β-D-glycosylamidines[136]. Some other effective reversible inhibitors of GHs are mannostatin A[137], acarbose[138] and validamycin A[139], thiosugars[140,141] and glucoimidazole-type inhibitors[142]. Reversible inhibitors have a multitude of applications for therapeutic and research purposes[139,143–145], and as chemical probes for the characterization of GHs[146].

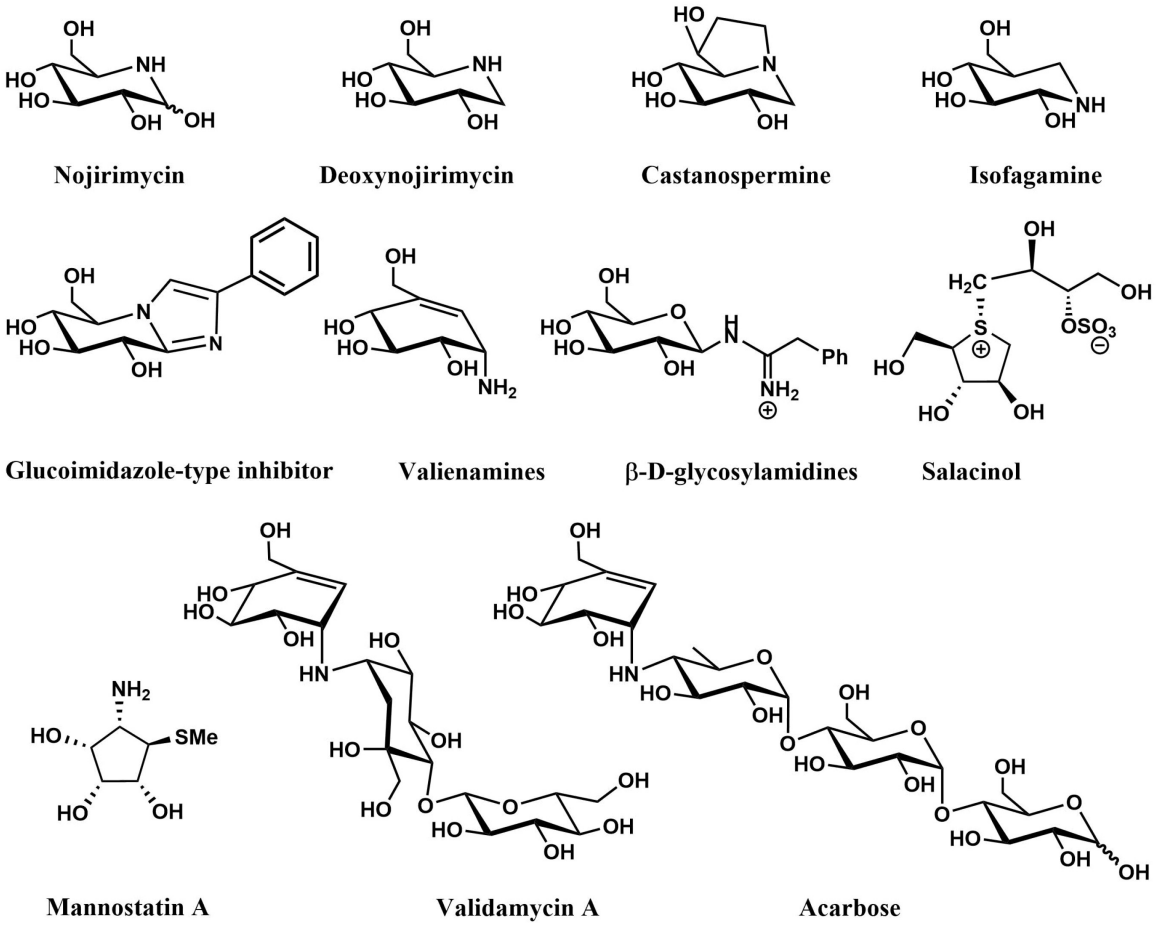


Figure 1.9: Examples of reversible inhibitors of GHs

1.3.2 Irreversible Inhibitors of GHs

Irreversible inhibitors facilitate GH inactivation via the formation of a covalent bond with the enzyme[147,148]. Due to their glycomimetic backbone, irreversible inhibitors have an affinity towards the active site of the target GH, and the loss of activity is incurred either by steric crowding of the active site or by chemical modification of important catalytic residues[147]. Irreversible inhibitors of GHs can be classified into two broad categories, as discussed below.

1.3.2.1 Active-site Affinity-based Inhibitors of GHs

Active-site affinity-based inhibitors (sometimes referred to as affinity labels) are structurally similar to the natural enzyme-substrate and are capable of covalently modifying active

site residues[149]. Active-site affinity-based inhibitors can be further subdivided into probes that require an external trigger for covalent cross-linking (such as photoaffinity labels[150]), and those that react spontaneously due to the presence of intrinsically reactive functional groups[147]. The former consists of inhibitors that have photoreactive groups such as diazirine or aryl azide situated on a substrate motif[151,152]. On activation, these groups release highly reactive nitrenes or carbenes, which in turn non-specifically react with the enzyme residues[150].

Other examples of active-site affinity-based inhibitors are C-glycosides with a highly reactive diazomethyl group[153,154], glucosylthio-hydroquinones[155], and glucosyl-isothiocyanate[156], all of which are minimally potent towards their target GH and hence have not seen wide use as glycosidase-labelling agents **(Figure 1.10).**

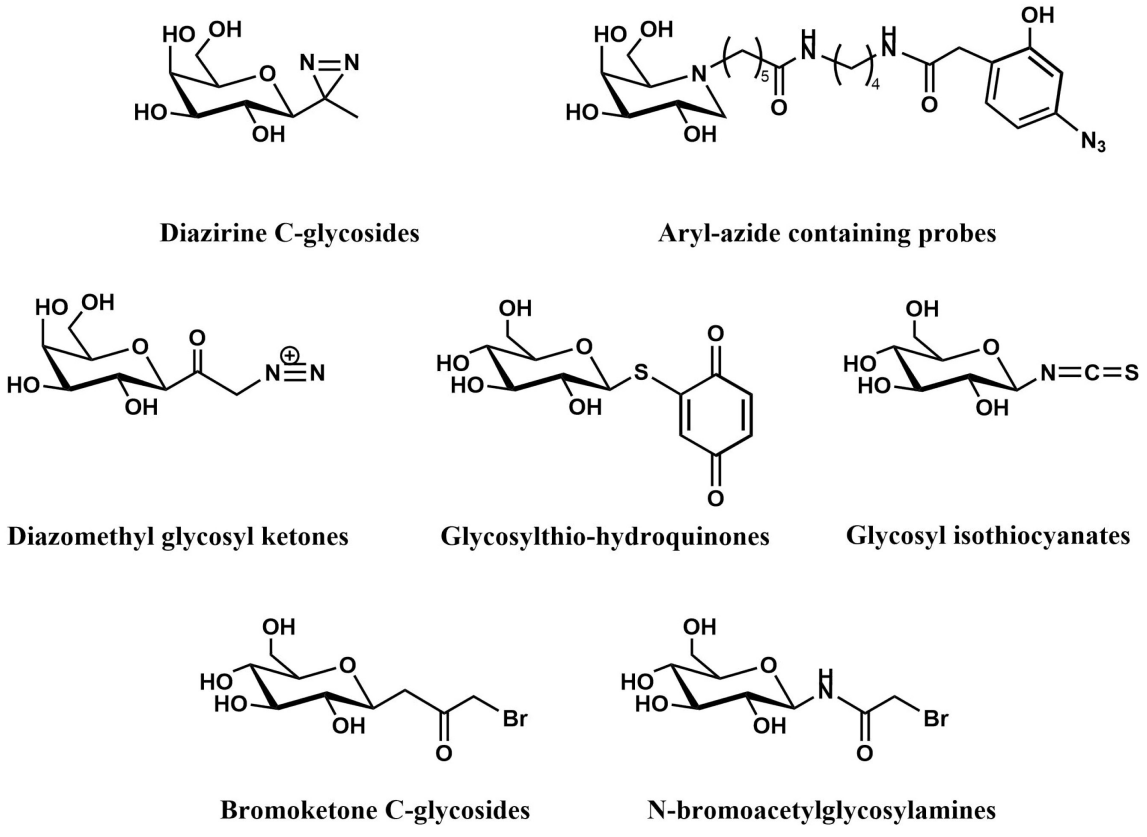


Figure 1.10: Examples of active-site affinity-based inhibitors of GHs

So far, the most effective inhibitors in this category have been the *N*-bromoacetylglycosylamine[157] and bromoketone C-glycoside[158] based affinity labels. They provide significant benefits over the previously mentioned probes as a result of their stability towards spontaneous decomposition and pH variation, as well as their ability to perform without prior stimulation[157,159]. Such inhibitors consist of an electrophilic "warhead" which reacts rapidly with the catalytic residue through the displacement of the bromo-functional group[160]. They can be synthesized through simple synthetic routes from free oligosaccharides without the need for the multiple protection and deprotection steps that are often necessary for synthetic glycochemistry.

N-bromoacetyl glycosylamines have thus far been developed on various mono- and oligosaccharide backbones such as, glucose[157,161–165], galactose[162,166], xylose[167–169], fucose[166], cellobiose[157,170], lactose[170], XXXG[170] and XLLG[170] (for the structure of XLLG, see **Figure 1.3**). Hence, *N*-bromoacetyl glycosylamine inhibitors have been used to inactivate various GHs, including: *endo*-xyloglucanases[170,171], mixed-linkage β-glucanase/xyloglucanase[43], β-galactosidase[166], *exo*glycanase[160], β-xylosidases[167–169], and β-glucosidases[161,163–165]. Analogous *N*-iodoacetylglycosylamine congeners have also been previously developed[172]; however, their application is limited by their instability, as previously discussed[170,173].

1.3.2.2 Mechanism-based Inhibitors of GHs

Contrary to active-site affinity-based inhibitors, mechanism-based inhibitors (MBIs) require activation by the catalytic machinery. During the catalytic cycle, inhibition occurs due to a covalent attachment of the enzyme catalytic amino acid residue to the anomeric carbon of the inhibitor, or to a reactive species formed during catalysis[147,149,174]. MBIs are typically substrate or transition-state analogues that take advantage of the catalytic machinery of the target GHs and render them inactive via a 'suicide' mechanism[175].

A variety of MBIs have been developed over the past few decades, for example, glycosylmethyl aryl triazenes[153,176], o- and p- difluoromethylaryl β-glucosides[177], conduritol β-epoxide[174,178], cyclophellitol[179], conduritol β-aziridines[180], N-aminoaziridines[181], cyclosulfates[182], cyclohexene carbasugar analogs[183], and bicyclo[4.1.0]heptyl galactose analogues[184] **(Figure 1.11)**.

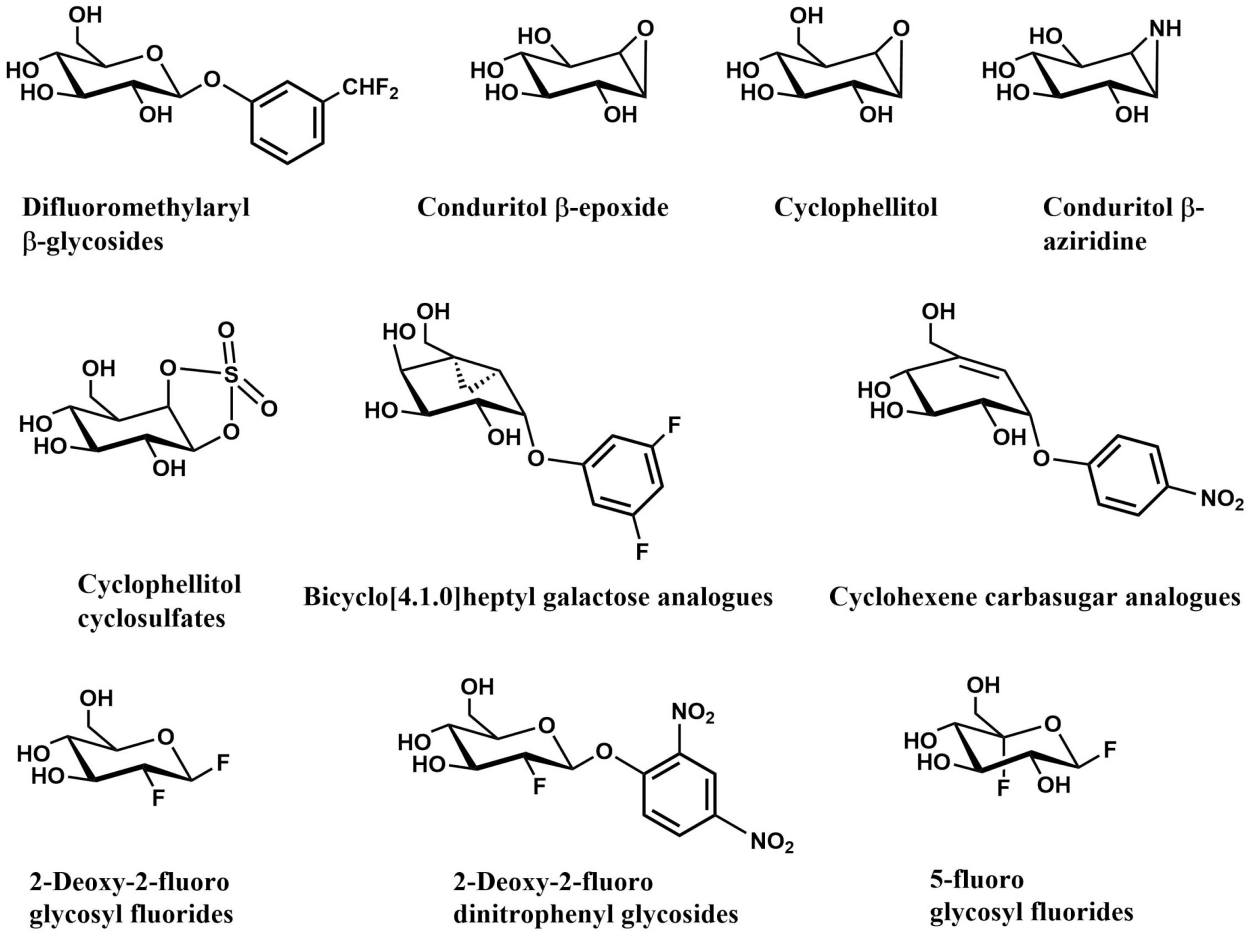


Figure 1.11: Examples of mechanism-based inhibitors of GHs

The incorporation of fluorine[185] and H[186] atoms to replace hydroxyl groups within the pyranose glycone forms the basis of a subclass of MBIs. The mechanism of this subclass is based on the destabilization of the transition state and disruption of key catalytic interactions between the hydroxyl group and enzymic residues. Since the conception of this idea, several mono- and

difluoroglycoside-type probes have since been developed[187–192]. For example, 2', 4'-dinitrophenyl-2-deoxy-2-fluoroglycoside (2F-DNP) inhibitors are targeted towards retaining GHs, taking advantage of their double-displacement mechanism.

In 2F-DNP inhibitors, the C2 hydroxyl group is substituted with the electron-withdrawing fluorine group which, during catalysis, causes a significant destabilization of the transition states and increases the activation energy. This leads to a decrease in the rates of both reaction steps. A good leaving group, such as the chromogenic dinitrophenyl group, is substituted at the anomeric position to increase the rate of the first (glycosylation) step. These substitutions, hence, facilitate the accumulation of the glycosyl-enzyme intermediate, leading to a significant reduction of the turnover rate and in most cases causing complete enzyme inhibition[185,193] **(Figure 1.12)**.

2-deoxy-2-fluoro and 2-deoxy-2, 2-difluoro inhibitors have previously been developed on various carbohydrate backbones including glucose[185], galactose[194], mannose[194], lyxose[191,192], maltose[188,191], maltotriose[191], xylobiose[195], cellobiose[196], laminaribiose (and related higher β(1,3) and mixed-linkage oligosaccharides)[197,198], and 2-acetamido-Glc-β(1, 4)-Glc[199,200]. 2-deoxy-2-fluoro glycosides have been highly successful in trapping a variety of β-retaining GHs including xylanases[195,201], galactosidases[202,203], glucocerebrosidases[204], *endo*-glucanases[196,205], and glucosidases[206,207]. The additional electron withdrawing effect of a second fluorine in 2-deoxy-2, 2-difluoro glycosides is typically required for trapping α-retaining GHs[189,190,192,208].

Similarly, 5-fluoro-glycosyl sugar congeners[209,210], which have been found to be suitable for trapping both α-retaining GHs[210–212] and β-retaining GHs[209,213], cause the destabilization of the emerging positive charge on the endocyclic oxygen during catalysis, and with an appropriate leaving group, lead to the accumulation of glycosyl-enzyme intermediate.

Many such probes have been further functionalized with reporter groups for developing activity-based probes[214–216]. However, due to the limited number of highly specific GH probes based on oligosaccharide scaffolds, there is further scope for the development of novel inhibitors targeting enzymes such as xyloglucanases and MLGases.

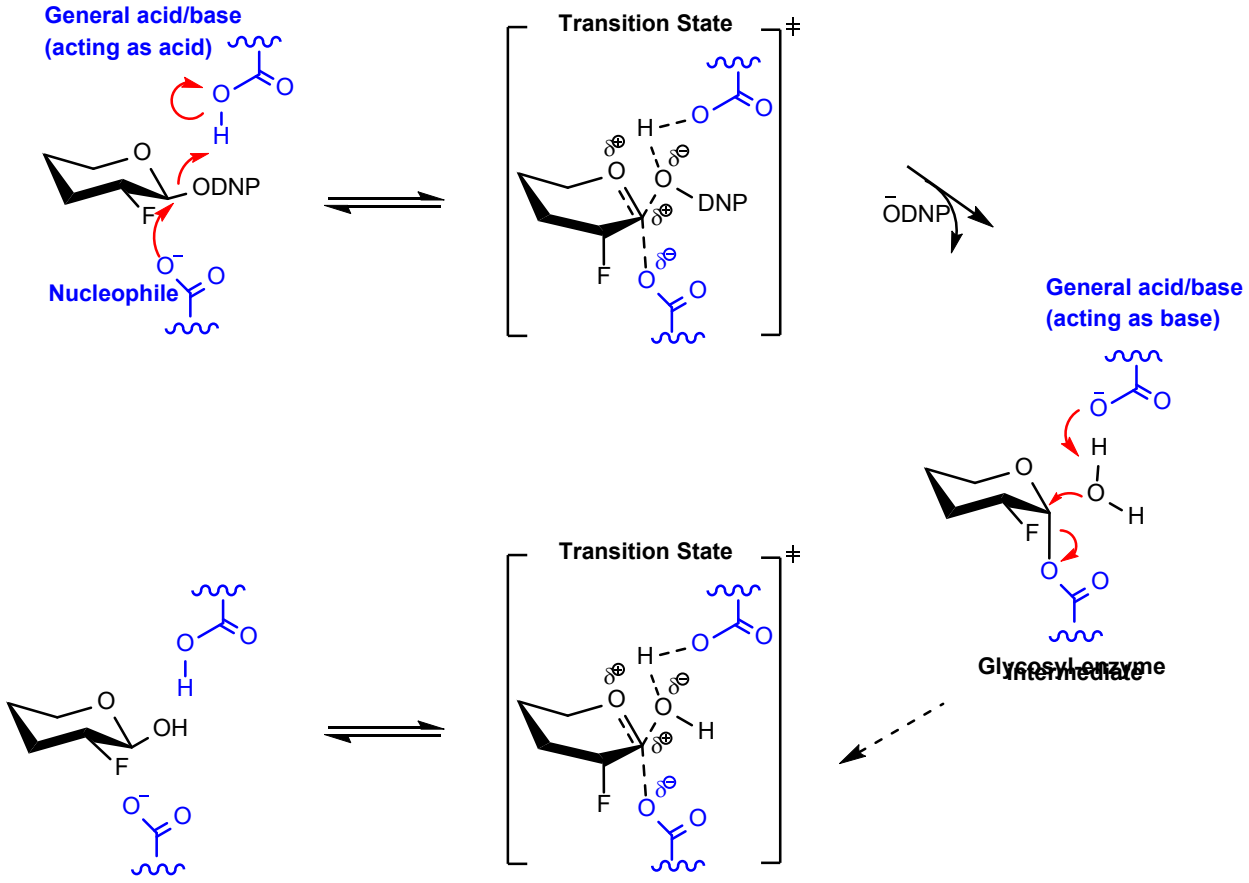


Figure 1.12: Mechanism of inhibition of GHs by 2',4'-dinitrophenyl-2-deoxy-2-fluoroglycoside inhibitors

1.4 Evaluating GH Activity using Inhibitors

Many analytical techniques can be used to test the activity of covalent inhibitors on target GHs to facilitate their structural and functional characterization, including protein crystallography, reaction product analysis using chromatography, protein mass measurement using mass

spectrometry techniques, and biochemical techniques such as enzyme kinetics analyses using chromogenic substrates and inhibition kinetics measurement.

1.4.1 Enzyme Kinetics

Enzyme kinetics help in deducing the reaction rate of an enzyme-catalyzed reaction and the effects of external conditions. A set of principles first developed by Michaelis and Menten in 1913[217] and further developed by Briggs and Haldane[218] captures the kinetic information necessary to map the interactions between enzymes and their substrates.

In the simplest model of an enzymatic reaction for retaining GHs (following a two-step mechanism) described below **(Scheme 1.1),** E stands for the enzyme, S stands for the substrate, E.S stands for enzyme-substrate complex (also known as the Michaelis complex), E-S is the covalent glycosyl-enzyme intermediate, and P is the reaction product. The reaction rate (V) is defined in terms of substrate concentration and maximum reaction rate V_{max} (**Equation 1.1**).

The variables k_2 and k_3 are the rate constants of glycosylation and deglycosylation reactions, respectively. Two important variables in this model are the Michaelis-Menten constant (K_m), defined as the substrate concentration needed to achieve a half-maximum enzyme velocity (**Equation 1.2**) and the turnover number (k_{cat}) which is the first-order rate constant for the conversion of the Michaelis complex to free enzyme and product (**Equation 1.3**). When the deglycosylation step is rate-limiting ($k_2 >> k_3$), $k_{cat} = k_3$. The ratio k_{cat}/K_m (the specificity, or performance, constant[219,220]) is an important indicator of the catalytic efficiency of enzymes and often used to compare the proficiency of enzyme catalysis across various parameters.

$$\mathrm{E} + \mathrm{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{E.S} \xrightarrow{k_2} \textit{E-S} \xrightarrow{k_3} \mathrm{E} + \mathrm{P}$$

Scheme 1.1: Kinetic model of enzyme kinetics

$$V = \frac{V_{max}\,[\mathrm{S}]}{K_m + [\mathrm{S}]} \tag{1.1}$$

$$K_m = \frac{k_3\,(k_{-1} + k_2)}{k_1\,(k_2 + k_3)} \tag{1.2}$$

$$k_{cat} = \frac{k_2 k_3}{k_2 + k_3} \tag{1.3}$$

The progress curve (i.e. depletion of substrate or accumulation of product over time) has multiple phases, including the early linear phase or steady-state, where the product is formed at an 'initial rate'. Following this, the reaction rate continues to drop and the curve eventually reaches a plateau. During the steady-state, the concentration of the enzymatic intermediate (E.S) is essentially constant. Typically, the measurement of reaction rates is assumed to be made in the steady-state at substrate concentrations far higher than enzyme concentrations ([S]>>[E]).

Further details into the enzyme kinetics can be obtained by analyzing the rate of the reaction as soon as the enzyme is added to the substrate mixture. Known as the pre-steady state phase (or burst phase), this initial period sees the rapid accumulation of enzyme-substrate complex (E.S), with one equivalent of leaving group released per enzyme catalytic site. The amount of released leaving group is often used to quantify the concentration of active catalytic sites in a given enzyme solution[186,188,206]. Several chromophores, e.g. phenolates, can be used as leaving group to study the pre-steady state kinetics of GHs since their release can be monitored continuously through spectroscopy. The burst phase is followed by the slower, linear phase (steady-state) described before.

1.4.2 Inhibition Kinetics

Enzyme inhibition kinetics quantify the potency of inhibitors through determination of inhibition rate constants. In the case of covalent inhibitors, the inhibition mechanism (**Scheme 1.2)** includes enzyme E binding with the inhibitor I with a dissociation constant of K_i to give a

Michaelis-type complex E.I. The Michaelis-type complex then proceeds via a first-order reaction to give the covalent enzyme-inhibitor complex E-I with an associated first-order rate constant of k_i. The term K_i, referred to as the dissociation constant, represents the concentration of the inactivator that gives half-maximal inactivation, and k_i is the maximal rate constant for inactivation or the turnover number. Kitz and Wilson (1962)[221] have shown that such reactions will obey **Equations 1.4** and **1.5**.

$$E + I \overset{K_i}{\rightleftharpoons} E.I \xrightarrow{k_i} E\text{-}I$$

Scheme 1.2: Kinetic model of irreversible enzyme inhibition kinetics

A method to study the extent of inactivation includes the incubation of a solution of the enzyme with various concentrations of the inhibitor. At various times, small aliquots of this incubate are removed. Residual enzymatic activity (ν) at the steady-state phase is determined and plotted to reveal the degree of inactivation as a function of time (**Equation 1.4, Figure 1.13a**). This plot is used to calculate the apparent rate constant (k_{app}) for each inhibitor concentration. Thereafter, k_{app} can be plotted as a function of inhibitor concentration (**Equation 1.5, Figure 1.13b**), yielding the values of the kinetic parameters k_i and K_i.

$$\nu = \nu_0 e^{-k_{app} t} \tag{1.4}$$

$$k_{app} = \frac{k_i[I]}{K_i + [I]} \tag{1.5}$$

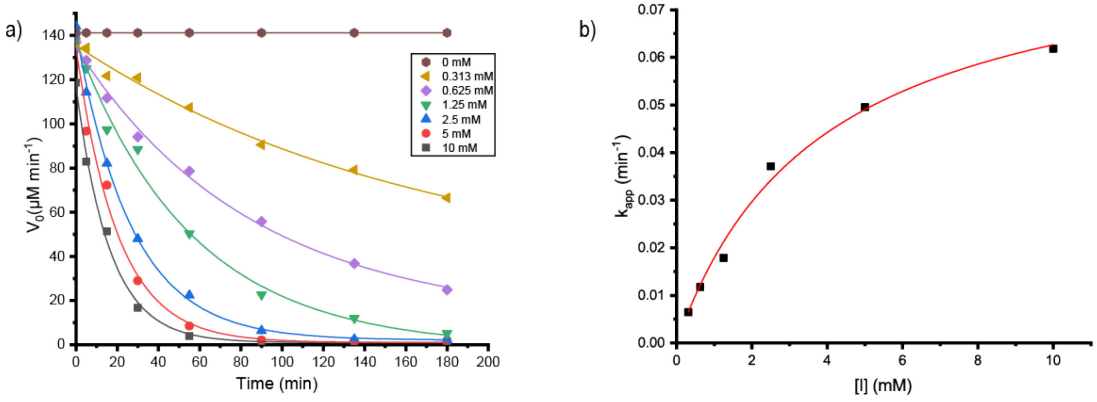


Figure 1.13: Illustrative curves for inhibition kinetics fitted using equations 1.4 and 1.5. (a) Plot of residual activity v versus time. (b) Plot of k_{app} versus inhibitor concentration [I]

1.4.3 Intact-protein Mass Spectrometry

Intact-protein mass spectrometry has been extensively used to analyze proteins in the past decades[222–224]. In particular, soft ionization techniques such as ESI have been very successful in the precise measurement of proteins and protein complexes with masses over a million daltons[225]. Such techniques have regularly been used to characterize GHs and can be used to provide evidence of covalent binding of the inhibitor to the enzyme by comparing the mass of enzyme before and after the addition of the inhibitor[226–229].

1.4.4 Inhibitor-bound Enzyme Crystallography

Three-dimensional enzyme crystallography is one of the most important and definitive techniques for GH characterization. This technique facilitates the understanding of atomic-level details of the enzymes and provides valuable information regarding the catalytic mechanism and substrate interactions. Crystal structures in complex with ligands delineate interactions between the enzyme and its substrate or inhibitor and are an invaluable tool in the structural biology of GHs[230,231]. The first enzyme structure solved by crystallography was chicken egg lysozyme in 1965[232,233]. Structures of several substrate- or inhibitor-bound GH complexes have now been solved[234–238]. The two methods of crystallography of enzyme-ligand complex are co-crystallization, in which a ligand is added to the crystallization solution, and soaking, in which the enzyme crystal is formed before the introduction of the ligand.

1.5 Aim of Investigation

The central aim of this thesis is to develop new biochemical tools for the characterization of polysaccharide-active GH enzymes. In particular, the focus of this work is on the GHs degrading

XyG, MLG, and β-(1,3) glucan, due to the myriad applications of the glucans as well as their respective GHs in various commercial sectors. This overarching aim was achieved via the chemo-enzymatic synthesis of stable small-molecule inhibitors by the functionalization of complex oligosaccharides to install reactive groups.

Two distinct modes of inhibitory action are featured in this thesis, in the form of mechanism-based and active-site affinity-based inhibitors[147]. The compound targets for this thesis are described in **Figure 1.14**. The 2-deoxy-2-fluoro mechanism-based inhibitors described in Chapters 2 and 4 target the nucleophilic residue of retaining GHs by participating in their two-step catalytic cycle, and trapping the reaction intermediate by kinetic manipulation of the transition state. The *N*-bromoacetylglycosylamine active-site affinity-based inhibitors, as described in Chapters 3, are general electrophilic reagents capable of labelling catalytically important carboxylic acid residues[43,167,170].

This work is built upon the previous development of a xyloglucan-oligosaccharide *N*-bromoacetylglycosylamine (XXXG-NHCOCH$_2$Br) inhibitor by the Brumer group (**Figure 1.14a**)[43,170]. In the first target (described in Chapter 2), the functional group at the reducing end of XyG oligosaccharide XXXG is varied to develop a mechanism-based inhibitor (XXXG(2F) β-DNP **Figure 1.14b**). Secondly, the *N*-bromoacetylglycosylamine is installed on to a mixed-linkage oligosaccharide backbone to generate active site inhibitors of MLGases (GG3G, GGG3G and G3GGG NHCOCH$_2$Br, Chapter 3, **Figure 1.14c**). Lastly, the backbone, as well as functional group, is varied to develop di-and tri-saccharide inhibitors incorporating a β-(1,3) linkage and 2-deoxy-2-fluoro dinitrophenyl moiety at the reducing end yielding G3G and GG3G (2F) β-DNP inhibitors (Chapter 4, **Figure 1.14d)**

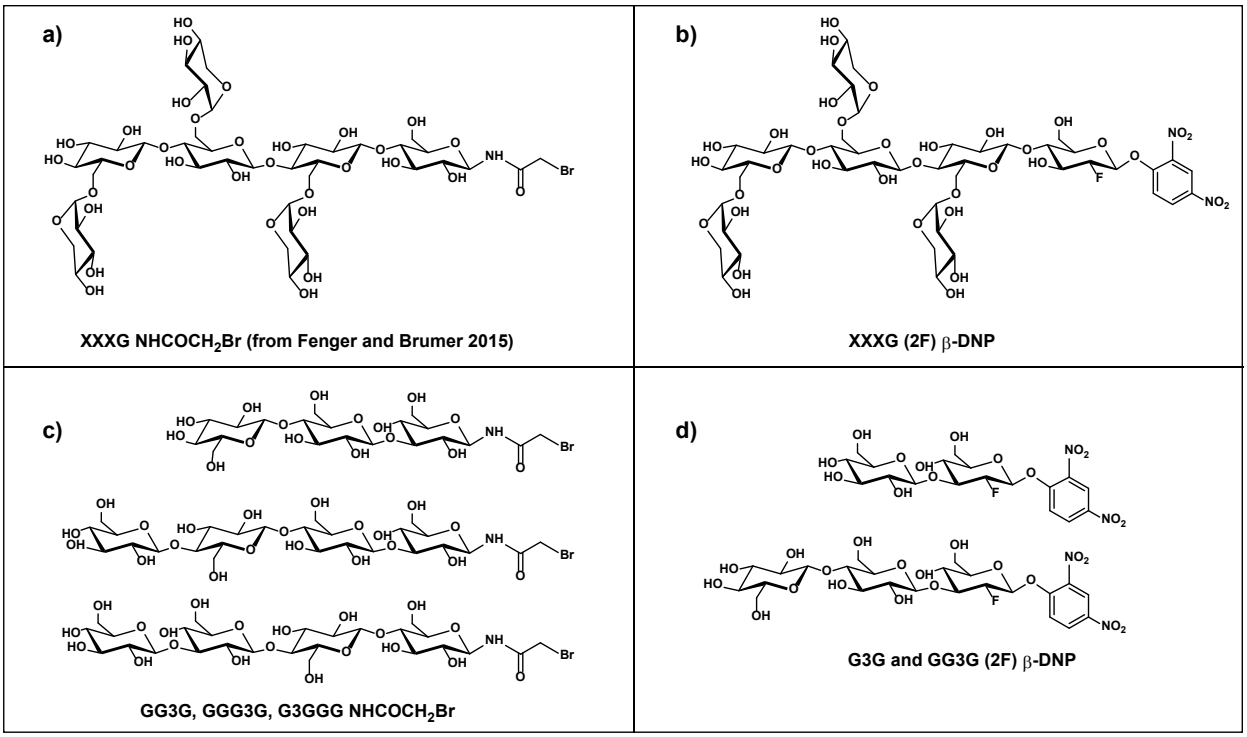


Figure 1.14: Design of inhibitors presented in this thesis, inspired by the previously synthesized (top left panel) XXXG NHCOCH$_2$Br inhibitor of *endo*-xyloglucanases[170].

The versatility in the design of the target inhibitors above is aimed to increase their utility in characterizing newly discovered GHs. Additionally, further functionalization of these active site GH inhibitors at the non-reducing end with reporter groups such as biotin or fluorogenic species like BODIPY can assist in the activity-based protein profiling of complex GH systems in native and non-native environment[239,240], significantly adding to the repertoire of covalent glycan probes.

Chapter 2: Synthesis and Application of a Highly Branched, Mechanism-based 2-deoxy-2-fluoro-oligosaccharide Inhibitor of *Endo*-xyloglucanases[1]

2.1 Introduction

Terrestrial and marine biomass represents a large and currently underutilized source of carbon in the transition from fossil petroleum to sustainable fuels, chemicals, and materials[241,242]. However, the complexity and recalcitrance of terrestrial plant cell walls, in particular, significantly impede the extraction and (bio)chemical transformation of individual component polymers, including cellulose, diverse matrix hetero-polysaccharides (e.g. hemicelluloses and pectins), and lignin. To help overcome inherent limitations in strictly chemical processes, including a lack of precise molecular control, there has been a widespread and sustained interest in the discovery and characterization of efficient enzymes for biomass transformation in the food, feed, and bioproducts industries[243,244]. In particular, Glycoside Hydrolases (GHs), which currently comprise over 150 structurally related families in the CAZyme database[78], has a particularly rich history of fundamental and applied research. Detailed mechanistic and structural studies have been essential in delineating CAZyme specificity determinants, which in turn underpin the surgical application of individual enzymes and enzyme cocktails in bioprocesses[90,245,246].

The use of small molecule glycomimetic inhibitors has been central to GH structure-function analyses in glycobiology[143,174,247]. In particular, irreversible, covalent inhibitors of GHs have been extensively used to identify key substrate-binding and catalytic residues through analytical biochemistry and enzyme crystallography and to screen for new GHs. A wide array of

[1] Adapted from: Jain, N., Attia, M. A., Offen, W. A., Davies, G. J. & Brumer, H. Synthesis and application of a highly branched, mechanism-based 2-deoxy-2-fluoro-oligosaccharide inhibitor of endo-xyloglucanases. Org. Biomol. Chem. 16, 8732–8741 (2018), with a correction published in Jain, N., Attia, M. A., Offen, W. A., Davies, G. J. & Brumer, H. Correction: Synthesis and application of a highly branched, mechanism-based 2-deoxy-2-fluoro-oligosaccharide inhibitor of endo-xyloglucanases. Org. Biomol. Chem. 17, 398-398 (2019

generally reactive, photo-activatable, or mechanism-based inhibitors based on bespoke glycan specificity motifs has been deployed, including carbasugar-epoxides (e.g. conduritol-β-epoxide, cyclophellitol), cyclopropylcarbasugars, epoxyalkyl glycosides, *N*-bromoacetylglycosylamines, bromoketone C-glycosides, glucosylthio-hydroquinones, aziridines, cyclosulfates, glycosylmethyl triazenes, activated phenylmethyl glycosides, various photoaffinity labels, and glycosides fluorinated at the 2- or 5-position (see comprehensive reviews and references therein[147,248,249], and recent primary literature[182–184,250,251].

Since their introduction by Withers and coworkers 30 years ago[194], 2-deoxy-2-fluoroglycosides bearing activated aglycones (and the related 2-deoxy-2,2-difluoroglycosides[188,191,192]) have been widely used for mechanistic and structural studies of diverse GHs[147]. By virtue of their mechanism-based inhibition and conservative steric substitution, this class of compounds has been exceptionally useful for identifying the catalytic nucleophile by protein mass spectrometry and/or crystallography in GH families that utilize a double-displacement, anomeric configuration-retaining mechanism (**Figure 1.12**)[149]. During catalysis, the presence of the fluorine group at C-2 destabilizes the transition state of both chemical steps, while the incorporation of a good nucleofuge, such as fluorine or 2',4'-dinitrophenol (2,4-DNP), increases the rate of leaving group departure sufficiently to enable the accumulation of the 2-deoxyfluoroglycosyl-enzyme intermediate. For some GHs, inhibition is essentially complete, while in others, these inhibitors act as "slow substrates" due to demonstrable turnover to release the free enzyme[206], especially in the presence of sugars as alternate glycosyl acceptor substrates[191,194,195,200,202,207,252,253].

The xyloglucans (XyGs) comprise a family of complex heteropolysaccharides, whose members are ubiquitous in land plants[254], where they can constitute up to one-quarter of the dry

weight of the primary cell wall[14]. A central structural feature of XyGs is a linear β-(1,4)-glucan backbone that is regularly substituted with α-(1,6)-xylosyl residues, which can be further extended by various other saccharide residues depending on the source tissue[25]. In dicot XyGs, the core repeating unit comprises a heptasaccharide motif (Xyl_3Glc_4), in which three of four contiguous backbone glucosyl units are branched (**Figure 1.3**). Consequently, such XyGs are referred to as "XXXG-type"[27] in the standard shorthand, in which "G" represents and unbranched β-(1,4)-linked backbone glucosyl residue, and "X" represents the disaccharide motif comprising a β-(1,4)-linked backbone glucosyl residue bearing an α-(1,6)-linked xylosyl sidechain[26].

Inspired by the success of 2-deoxy-2-fluoroglycosides as mechanism-base inhibitors and motivated by a long-standing interest in the enzymology of xyloglucan metabolism by plants and microorganisms[31,32,255,256], the chemo-enzymatic synthesis of XXXG(2F)-β-DNP as a specific mechanism-based inhibitor of *endo*-(xylo)glucanases (EC 3.2.1.151) is presented. In particular, the application of this compound is demonstrated for the covalent labelling of exemplar configuration-retaining *endo*-(xylo)glucanases by protein mass spectrometry and crystallography.

2.2 Materials and Methods

2.2.1 General Synthetic and Analytic Techniques

All reagents and solvents were analytical or HPLC grade and were purchased from Sigma-Aldrich, Alfa-Aesar or ACROS Organics. For anhydrous reactions, glassware was dried overnight in a 100-150 °C oven and purged with argon prior to use. Solvents were dried by stirring with activated 4 Å molecular sieves overnight under argon.

Thin-layer chromatography (TLC) was performed using aluminum sheet TLC plates (0.25 mm) pre-coated with Merck silica gel 60 F254, using ethyl acetate: hexanes or water: isopropanol: ethyl acetate as solvent systems (particular solvent ratios specified below), and visualized by a UV

lamp and/or 10% sulfuric acid in water with charring using a heat gun. Flash chromatography was performed using Merck silica gel 60 with ethyl acetate: hexanes or water: isopropanol: ethyl acetate as mobile phases. Fractions were analyzed by TLC, and those with the desired compounds were pooled together and evaporated under reduced pressure.

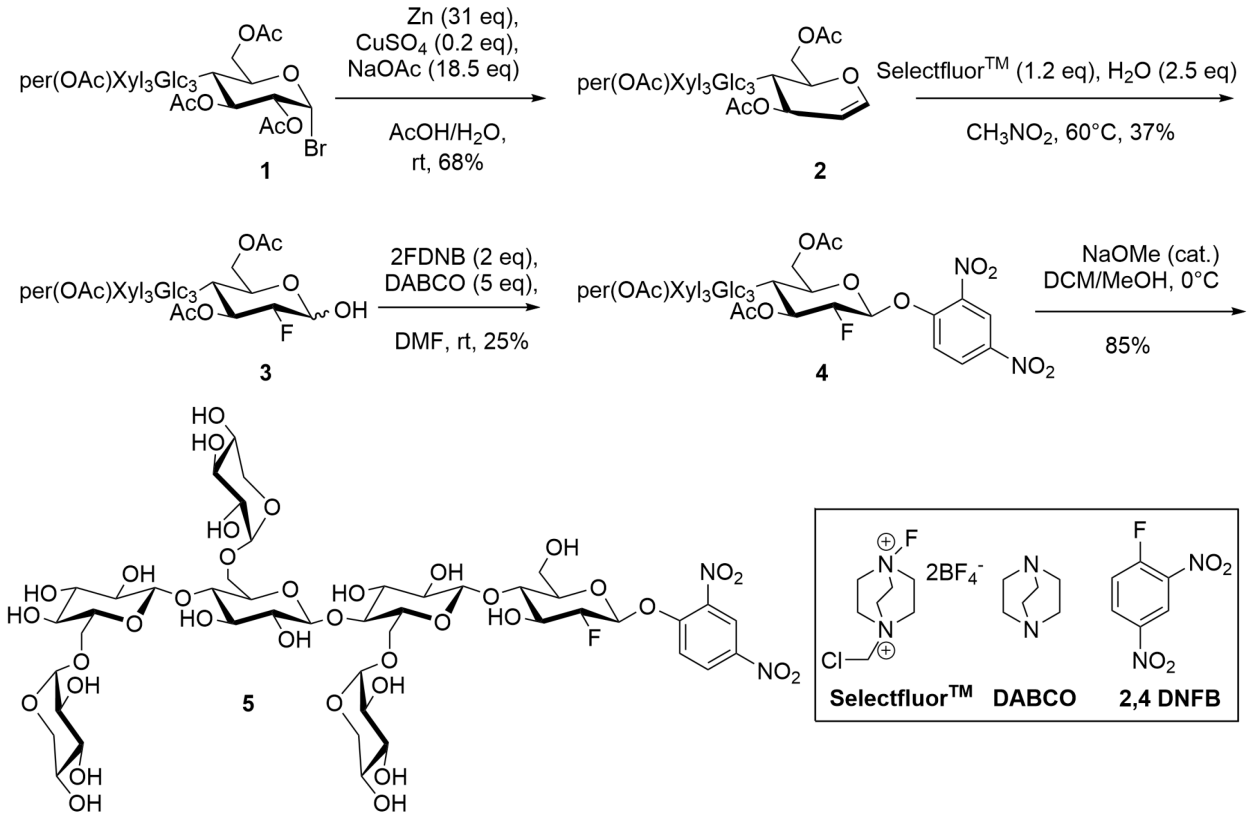
All ^{19}F-, ^{13}C- and ^{1}H-NMR data were collected on a Bruker Avance 400 MHz spectrometer at room temperature (100.6 MHz and 376.5 MHz for ^{13}C- and ^{19}F-, respectively). The NMR spectra were referenced to the solvent as follows[257]: HOD= 4.79 ppm, $CHCl_3$= 7.27 ppm. MALDI-MS data were collected on a Bruker Autoflex instrument in reflectron mode over m/z 700-3500 using 6-Aza-2-thiothymine (ATT) as the matrix. HRMS data were obtained using either a Waters Xevo G2-S Q-TOF or Waters/Micromass LCT TOF mass spectrometer in positive-ion mode, via direct infusion through an electrospray ion source.

2.2.2 Synthesis of XXXG(2F)-β-DNP

XXXG heptasaccharide ([α-D-Xylp-(1,6)]-β-D-Glcp-(1,4)- [α-D-Xylp-(1,6)]-β-D-Glcp-(1,4)- [α-D-Xylp-(1,6)]-β-D-Glcp-(1,4)-D-Glcp; abbreviated nomenclature according to Tuomivaara et al.[26]) was prepared enzymatically from tamarind kernel powder, per-*O*-acetylated, and converted to the corresponding α-glycosyl bromide (**Scheme 2.1**) as per our previously established procedure[258].

The synthesis of per-*O*-acetylated XXXG glycal (**2**) was performed by adapting the method of Xu et al.[259]. (per-*O*Ac) XXXG α-bromide (**1,** 1 g) was dissolved in 30 mL in acetic acid. A solution of Zn (31 eq.), NaOAc (18.5 eq.), and $CuSO_4.5H_2O$ (0.2 eq.) was suspended in 20 mL of water and stirred for 5 min. The solution of **1** in acetic acid was slowly added to the suspension. The solution was stirred for 3 h at room temperature, and filtered through a Celite pad. The solvent was concentrated under reduced pressure, re-dissolved in CH_2Cl_2, and washed with $NaHCO_3$ (3×)

and brine (1×). The organic layer was concentrated, and flash chromatography (mobile phase ethyl acetate/hexanes 2:1, R_f = 0.3 in the same solvent) followed by evaporation of the solvent under reduced pressure was used to isolate a white solid (0.63 g, 68% yield). ^{1}H-NMR (**Figure A1**, 400 MHz, $CDCl_3$): δ 6.41 (d, J = 6.14 Hz, 1H, H1), 5.40–5.36 (m), 5.19–4.60 (m), 4.29–3.65 (m), 3.35 (d), 2.13–1.93 (54H, $COCH_3$). ^{13}C-NMR (**Figure A2,** 100.6 MHz, $CDCl_3$): δ 170.34–168.64 (18 × CO), 145.58 (C1), 101.07–95.94 (C2, 6 × C1), 75.66–67.17 (6 × C2, 7 × C3, 7 × C4, 7 × C5), 61.77, 59.22, 59.06, 58.85 (4 × C6), 20.80–20.65 (18 × CH_3). Monoisotopic m/z calculated for $C_{75}H_{100}O_{49}K^+$: 1823.4975; MALDI-TOF MS found: 1823.3. m/z calculated for $C_{75}H_{100}O_{49}Na^+$: 1807.5236; ESI-HRMS found: 1807.5282.



Scheme 2.1: Synthesis of XXXG(2F)-β-DNP (5)

The synthesis of 1-hydroxy-2-deoxy-2-fluoro-(per-*O*Ac)XXXG (**3**) was adapted from previously reported methods[260,261]. Compound **2** (0.63 g) was dissolved in dry nitromethane (20 mL) and 1.2 eq. of Selectfluor^{TM} was added to the solution to create a suspension, which was stirred at room temperature for 6 h, at which time water (2.4 eq.) was added. The solution was then heated to 60°C and stirred at that temperature for 16 h. The reaction was filtered over Celite to remove the Selectfluor^{TM} by-product and the filtrate was concentrated by rotary evaporation. Elution of the crude product through a silica column (ethyl acetate/hexanes, 3.5:1) to remove polar impurities and remaining by-products yielded a mixture of anomers (yield 0.24 g, 37%), which was used directly in the next step without further separation. ^{19}F-NMR (**Figure A3,** 376.5MHz, $CDCl_3$) gluco β-anomer: -197.91 (ddd, $J_{H2\text{-}F2}$=51.1 Hz, $J_{H3\text{-}F2}$=13.7 Hz, $J_{H1\text{-}F2}$=1.6 Hz, F2), gluco α-anomer:-199.30 (dd, $J_{H2\text{-}F2}$=49.5 Hz, $J_{H3\text{-}F2}$=12.0 Hz, $J_{H1\text{-}F2}$=0 Hz, F2), consistent with values for the corresponding cellobioside[191,262]. Monoisotopic m/z calculated for $C_{75}H_{101}FO_{50}Na^+$: 1843.5248; MALDI-TOF MS found: 1843.3; ESI-HRMS found: 1843.5494

The synthesis of the per-*O*-acetylated 2'4'-dinitrophenyl β-glycoside of 2-deoxy-2-fluoro-XXXG (**4**) was adapted from a previously reported glycosylation method[191]**.** A solution of 2,4 dinitrofluorobenzene (2,4-DNFB, 2 eq.) was dissolved in dry dimethylformamide (DMF) and stirred over activated 4 Å molecular sieves overnight. Crude **3** (0.24 g) was dissolved in dry DMF (10 mL) and under dry conditions, 5 eq. of 1,4-diazabicyclo [2.2.2] octane (DABCO) was added to the solution as a solid powder. The solution was stirred for 15 min, after which it was added to the solution of 2,4-DNFB through a syringe with an oven-dried steel needle. The reaction was stirred for 3.5 h, after which time the molecular sieves were filtered away and the DMF in the filtrate was evaporated under reduced pressure. The solution was re-dissolved in CH_2Cl_2 and washed with $NaHCO_3$ (3×) and brine (1×). Flash chromatography (mobile phase ethyl

acetate/hexanes 2.5:1) was used to isolate the pure compound (TLC R_f= 0.35 in the same solvent mixture) with a yield of 66 mg, 25%. ^{1}H-NMR (**Figure A4,** 400 MHz, $CDCl_3$): δ 8.75 (d, 1H, H′3), 8.45 (dd, 1H, H′5), 7.42 (d, 1H, H′6), 5.56 (d, 1H, H1), 5.43–5.32 (m), 5.17–4.60 (m), 4.15–3.68 (m), 3.43 (d), 2.16–1.97 (54H, $COCH_3$). ^{13}C-NMR (**Figure A5,** 100.6 MHz, $CDCl_3$): δ 170.23–168.56 (18 × CO), 153.22 (C′3), 142.00 (C′4), 140.01 (C′5), 128.65 (C′1), 121.54 (C′2), 117.49 (C′6), 100.38–95.91 (7 × C1), 88.09 ($J_{C2–F2}$ = 189.0 Hz, C2), 75.22–67.32 (6 × C2, 7 × C3, 7 × C4, 7 × C5), 60.40, 59.09, 58.94, 58.74 (4 × C6), 21.05–20.58 (18 × CH_3). ^{19}F-NMR (**Figure A6,** 376.5MHz, $CDCl_3$ δ −194.75 (ddd, $J_{H2–F2}$ = 47.7 Hz, $J_{H3–F2}$ = 15.7 Hz, $J_{H1–F2}$ = 2.5 Hz, F2). Monoisotopic m/z calculated for $C_{81}H_{103}FN_2O_{54}Na^+$: 2009.5257; MALDI-TOF MS found: 2009.4; ESI-HRMS found: 2009.5463.

The 2’4’-dinitrophenyl β-glycoside of 2-deoxy-2-fluoro-XXXG (XXXG(2F)-β-DNP (**5**) was produced by Zemplen deprotection of **4** (66 mg) in 10 mL methanol/CH_2Cl_2 9:1, to which 0.5 equivalents of NaOMe (25% in MeOH) was added. The reaction was stirred at 4°C and monitored by TLC overnight. The product was purified by flash chromatography using water/isopropanol/ethyl acetate (1:3:4) as the mobile phase. The purified product was re-dissolved in water and freeze-dried to give a pale, fluffy powder in 85% yield (34 mg). ^{1}H-NMR (**Figure A7,** 400 MHz, $CDCl_3$): δ 8.75 (d, 1H, H′3), 8.51 (dd, 1H, H′5), 7.67 (d, 1H, H′6), 5.66 (d, 1H, H1), 4.84–4.80 (m), 4.61–4.39 (m), 3.98–3.31 (m). ^{13}C-NMR (**Figure A8,** 100.6 MHz, $CDCl_3$): δ 153.37 (C′3), 141.19 (C′4), 138.42 (C′5), 129.35 (C′1), 121.72 (C′2), 117.32 (C′6), 102.35–96.97 (7 × C1), 90.43 ($J_{C2–F2}$ = 186.8 Hz, C2), 78.81–65.33 (6 × C2, 7 × C3, 7 × C4, 7 × C5), 61.93, 61.05, 61.00, 60.68 (4 × C6). Small amounts of CD_3OD (sep, 47.52) and CH_3COONa (181.02, 22.74) were also detected. ^{19}F-NMR (**Figure A9,** 376.5MHz, $CDCl_3$) −199.69 (ddd, $J_{H2–F2}$ = 51.4

Hz, $J_{H3–F2}$ = 15.2 Hz, $J_{H1–F2}$ = 2.5 Hz, F2). Monoisotopic m/z calculated for $C_{45}H_{67}FN_2O_{36}Na^+$:1253.3360; MALDI-TOF MS found: 1253.3; ESI-HRMS found: 1253.3354.

2.2.3 *Endo*-xyloglucanase Production

The *endo*-(xylo)glucanases *Cellvibrio japonicus* GH5D (CjGH5D)[171], *Cellvibrio japonicus* GH74 (CjGH74)[263], *Bacteroides ovatus* GH5 (BoGH5)[32], and *Prevotella bryantii* GH5 (PbGH5)[43] were produced recombinantly and purified as previously described. The purity of the recombinant proteins was confirmed by SDS-PAGE[32,43,171,263].

The catalytic acid/base mutant CjGH5D(E255A) was generated using the PCR-based QuickChange II Site-Directed Mutagenesis Kit (Agilent, USA) following the manufacturer's protocol. PCR amplification was conducted using the forward primer 5'-TTTGCCGGCACTAACGCCCCCAATGCGGAAAAT-3' and the reverse primer 5'-ATTTTCCGCATTGGGGGCGTTAGTGCCGGCAAA-3', utilizing pET28a::CjGH5D [171] as the template. The resulting plasmid pET28a:: CjGH5D(E255A) was sequenced to confirm the desired mutation.

Chemically competent *E. coli* Rosetta DE3 cells were transformed with the plasmid pET28a:: CjGH5D(E255A). The resulting colonies were grown on LB solid media containing kanamycin (50 μg mL^{-1}) and chloramphenicol (30 μg mL^{-1}). Gene overexpression and recombinant protein purification were then performed as described for the wild-type enzyme[171]. The concentration of the purified recombinant CjGH5D(E255A) was determined using the Epoch Micro-Volume Spectrophotometer System (BioTek®, USA) at 280 nm. The presence of the desired mutation was confirmed by intact protein mass spectrometry[264]. The purified protein was aliquoted and stored at -80 °C until needed.

2.2.4 Screening for Active Site Labelling

Individual 20 µL solutions of CjGH5D, PbGH5, BoGH5, CjGH74, and CjGH5D(E255A) (2.3 µM) in 0.5 mM sodium phosphate buffer, pH 7.0, containing 2.5 mM XXXG(2F)-β-DNP were incubated for 12 h. Intact protein masses were determined on a Waters Xevo LC-ESI-MS Q-TOF with a NanoAcuity UPLC system essentially as previously described[264], and analyzed using the software Masslynx 4.0.

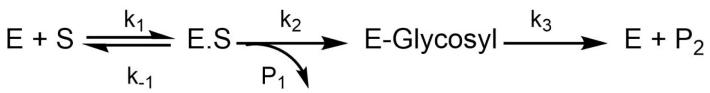
2.2.5 Inhibition Kinetics Measurements

All kinetic data were obtained using an Agilent Cary 60 UV–Vis Spectrophotometer equipped with a Peltier temperature-controlled cell holder. The data were fit by linear least-squares analysis using the bundled Cary Kinetics software or nonlinear regression using Origin 8 (OriginLab, Northampton, MA, USA).

To measure time-dependent enzyme inactivation, 2.5 mM of XXXG(2F)-β-DNP (**5**) was incubated with 20 µM CjGH5D in 50 mM sodium phosphate buffer, pH 7.5, in a total volume of 100 µL (also including 0.1 mg mL^{-1} bovine serum albumin (BSA) to prevent non-specific loss of activity) for a period of 420 min at 40°C. A control experiment was run in parallel, in which the inhibitor was omitted from the buffered enzyme/BSA solution to ensure the stability of the enzyme over the incubation time. Periodically, 10 µL of each solution was withdrawn and diluted 1:100 in 50 mM sodium phosphate buffer, pH 7.5, and 100 µL of the diluted solution was added to 100 µL of 0.4 mM XXXG-β-CNP[258], which was previously dissolved in water and preincubated at 40°C. Linear initial-rate kinetics of 2-chloro-4-nitrophenolate release was measured at 405 nm (ε= 17.74 mM^{-1} cm^{-1}) over 1 min in a quartz cuvette (l = 1 cm) maintained at 40°C, essentially as previously described[171].

Prior to measuring burst phase kinetics, the background hydrolysis rate of 235 μL of 5 mM XXXG(2F)-β DNP (**5**) in 50 mM phosphate buffer, pH 7.5, was measured at 40°C for 75 min (ε_{405} = 11.17 mM^{-1} cm^{-1}, determined using a standard curve of absorbance as a function of dinitrophenolate concentration at the specified buffer concentration and pH conditions. To make the standard solution, 2,4-dinitrophenol was desiccated overnight in presence of phosphorus pentoxide, and thereafter dissolved in the specified buffer). Thereafter, CjGH5D (15 μL, 0.45 mM), which had been preincubated at 40°C in the same buffer, was added to the cuvette and mixed rapidly (final solution of volume 250 μL, final enzyme concentration 27 μM). The rate of dinitrophenolate (DNP^-) release of the final solution was monitored for an additional 700 min and **Equation 2.1**[206] was fit to the data using Origin 8 software, according to the inhibition mechanism shown in **Scheme 2.2**. The cuvette was covered with parafilm to limit evaporation during the entire course of the experiment.

$$[\mathrm{DNP^-}] = [\mathrm{DNP^-}]_0(1 - \mathrm{e}^{-k_2 t}) + k_3 t \tag{2.1}$$



Scheme 2.2: Kinetic scheme of anomeric-configuration-retaining glycoside hydrolases. In the case of 2-deoxy-2-fluoroglycoside mechanism-based inhibitors, the rate of k_3 is reduced relative to k_2, resulting in the accumulation of the covalent glycosyl-enzyme intermediate

2.2.6 X-ray Crystallography and Structure Solution

CjGH5D (E255A) was crystallized in 20 mM sodium phosphate buffer, pH 7.0, containing 10% (v/v) glycerol using a protein concentration of 20 mg mL^{-1}, by employing similar conditions to those used to crystallize the wild-type protein[171]. A crystal grown from a protein stock solution supplemented with 5 mM inhibitor was harvested directly into liquid nitrogen using a CryoLoop™

(Hampton Research, Aliso Viejo, CA, USA). In addition, an unliganded crystal was soaked with 2 mM inhibitor for 45 min, and harvested similarly.

Data were collected at Diamond beamline IO3 and processed using XDS[265] for the co-crystallized complex and collected at IO4-1 and processed using DIALS[266] for the soaked complex. Both datasets were put through the data reduction pipeline in ccp4i2[267] which uses POINTLESS, AIMLESS and CTRUNCATE[268], and cut-off at resolutions of 1.7 and 2.0 Å (judged by R_{pim}<0.60 and $CC_{1/2}$ > 0.50 in outer bin), respectively (**Table A1**). Both structures were solved using the unliganded structure of the wild-type enzyme (PDB ID 5OYC) as the model for refinement with REFMAC[269]. After refinement with water molecules added, the ligands were built into difference electron density in the weighted $2F_o$-F_c maps using COOT[270], and validated using PRIVATEER[271] prior to deposition at the Protein Data Bank with accession codes 6HAA and 6HA9.

2.3 Results and Discussion

2.3.1 Inhibitor Synthesis

The synthesis of the target compound XXXG(2F)-β-DNP (**5**) was accomplished using a linear synthetic scheme starting from the heptasaccharide XXXG (**Scheme 2.1**). XXXG was produced via sequential *endo*-(xylo)glucanase and β-galactosidase hydrolysis of tamarind kernel XyG, per-*O*-acetylated, and converted to the corresponding α-bromide as per our established procedure[258]. Transformation of per-*O*-acetyl XXXG α-bromide to the corresponding glycal via a reductive elimination was straightforward using Zn metal and $CuSO_4$ in the presence of NaOAc/HOAc in water[259].

The crucial step of installing fluorine at the 2-position of the protected XXXG glycal was achieved using the electrophilic fluorinating agent Selectfluor™, followed by hydrolytic work-up

to obtain the corresponding protected 1-hydroxy-2-deoxy-2-fluoro-oligosaccharide. Selectfluor™ has been widely used for the regiospecific 2-fluorination of a number of glycals. [272] In the case of D-glucal, an equimolar mixture of *gluco* and *manno* 2F-epimers was obtained, while in the case of the corresponding disaccharides, e.g. cellobiose, the amount of the 2F-*gluco* epimer reached 80% in polar solvents[261,262]. In our case, only the desired *gluco* configured product was obtained in dry nitromethane[260], albeit as a mixture of α/β anomers, as indicated by ^{19}F NMR (trace amounts of *manno* epimers were perhaps also present[262], **Figure A3**). As this represents the only example, to our knowledge, of the application of Selectfluor™ to such a large oligosaccharide glucal, the origins of this high stereoselectivity are not fully clear but may be based in increased local steric congestion or altered access to boat-conformer addition products.[262]

The reaction of the anomeric mixture with 2,4-dinitrofluorobenzene (2,4-DNFB) in presence of DABCO yielded exclusively the desired β-configured, kinetic product **4**[273,274], as ascertained by the large H1-H2 coupling of 7.6 Hz and F2-H1 coupling of 3.0 Hz. With the key stereochemical configuration set, careful Zemplen deprotection followed by column chromatography yielded pure **5**.

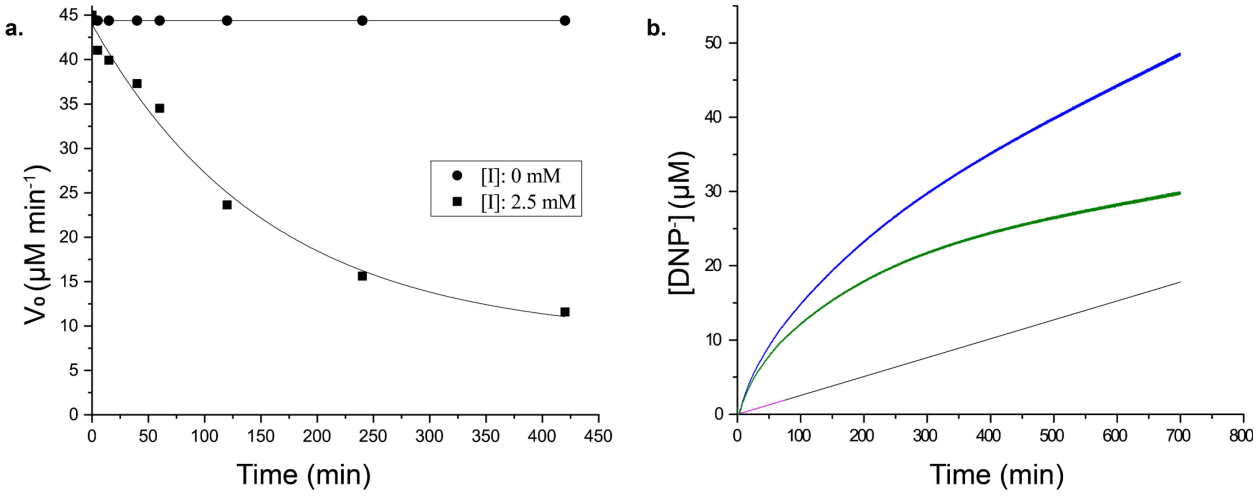


Figure 2.1: Inhibition of CjGH5D with XXXG(2F)-β-DNP: (a) Time-dependent loss of activity upon incubation with 5. (b) Real-time burst phase kinetics. Pink line: spontaneous hydrolysis of 5 over 75 min, with

linear extrapolation shown in black. Blue line: 2,4-dinitrophenylate release after the addition of CjGH5D. Green line: enzymic release of 2,4-dinitrophenylate obtained after subtraction of spontaneous hydrolysis

2.3.2 Active Site Labelling of *Endo*-xyloglucanases

To examine the efficacy of XXXG(2F)-β-DNP (**5**) as an inhibitor, a high concentration (2.5 mM) of the compound was incubated with three configuration-retaining GH5 *endo*-(xylo)glucanases with differing specificity profiles: BoGH5 from the human gut bacterium *Bacteroides ovatus*[32], PbGH5 from the ruminant bacterium *Prevotella bryantii*[43], and CjGH5D from the soil saprophyte *Cellvibrio japonicus*[171]. The configuration-inverting GH74 *endo*-xyloglucanase from *Cellvibrio japonicus*, CjGH74[263], which is not expected to be covalently inhibited due to a single step, direct water-attack mechanism[275], was also examined as a negative control. As anticipated, time-dependent, 1:1 stoichiometric labelling of all three configuration-retaining GH5 enzymes was observed by intact protein mass spectrometry, while CjGH74 remained entirely unmodified during a 12 h incubation (**Figure A11**). On the basis of this semi-quantitative analysis, the strict *endo*-(xylo)glucanase BoGH5 was completely labelled by 1 h (**Figure A11a**) and the promiscuous mixed-linkage β-(1-3)/β-(1-4)-*endo*-glucanase/(xylo)glucanase PbGH5 showed near-complete labelling by 9 h (**Figure A11b**). In contrast, the predominant *endo*-(xylo)glucanase CjGH5D was only partially labelled after a 12 h incubation (**Figure A11c**). Further analysis, detailed below, revealed that this was due to the turnover of the fluoroglycosyl-enzyme intermediate.

2.3.3 Kinetics of Inhibition of CjGH5D with XXXG(2F)-β-DNP

To understand better the mechanism of inhibition of CjGH5D with XXXG(2F)-β-DNP, kinetics measurements at the same high concentration of compound used in initial intact protein MS analyses were performed. As shown in **Figure 2.1a,** a time-dependent loss of activity toward

the chromogenic substrate XXXG-β-2'-chloro-4'-nitrophenol is observed following incubation with 2.5 mM XXXG(2F)-β-DNP. However, it was consistently observed that the activity of the enzyme was not completely lost, even after extended incubation (e.g. 420 min, **Figure 2.1a**). Similar results were observed in preliminary screening experiments using lower and higher inhibitor concentrations (up to 10 mM) with lengthy incubation times (e.g. 18 h, **Figure A12**). These observations suggested the kinetically relevant turnover of the fluoroglycosyl-enzyme intermediate, as has been reported previously for other combinations of 2-deoxy-2-fluoroglycoside inhibitors and enzymes[191,194,195,200,202,207,252,253]. The inhibition kinetics constants of CjGH5D using an active-site affinity-based covalent *N*-bromoacetylglycosylamine inhibitor synthesized on the XXXG backbone were determined (**Figure 1.14a**), the details of which can be found in Appendix A (**Figure A10**)[171].

Further evidence to suggest turnover of the glycosyl enzyme was obtained by the observation of apparent pre-steady-state burst kinetics, by monitoring the release of 2,4-dinitrophenolate over time (**Figure 2.1b**). The biphasic data were fit by **Equation 2.1**, which describes an initial pseudo-first-order exponential accumulation of the fluoroglycosyl-enzyme, followed by a linear phase dominated by the steady-state turnover of the covalent intermediate (**Figure 2.1b**)[206]. After accounting for an effectively linear background hydrolysis rate of **5** in the buffer (0.28×10^{-3} min^{-1}), values for k_2 and k_3 of 7.83×10^{-3} min^{-1} and 0.16×10^{-3} min^{-1}, respectively, were obtained. The concentration of enzyme active sites, $[E_o]$ =17.60 μM, obtained from the magnitude of the burst (A_{405} = 0.20, DNP^- ε_{405} = 11.17 $mM^{-1}cm^{-1}$) suggested that the CjGH5D preparation was not fully active ([P] = 27 μM in the assay).

2.3.4 Crystallography of the XXXG(2F)-CjGH5D(E255A) Covalent Complex

The turnover of the fluoroglycosyl-enzyme (**Figure 2.1b**) precluded the observation of a covalent complex by protein crystallography, despite MS evidence of its accumulation (**Figure A11**). To overcome this problem,[200] a site-directed mutation of the general acid/base residue Glu255 to alanine was produced in order to further reduce the rate of hydrolysis of the fluoroglycosyl-enzyme intermediate (**Figure A11d**). Our first attempt to capture this intermediate was by soaking of E255A crystals with the inhibitor, which surprisingly led to an unreacted complex; with a -3 to +1 binding mode of the glycone with the aryl glycosidic bond still intact (**Figure A13**). For this reason, co-crystallization was instead used to access the reacted tertiary structure of XXXG(2F)-CjGH5D(E255A) at 1.7 Å resolution by X-ray crystallography (**Figure 2.2**). The overall structures corresponded well to that of the wild-type enzyme in free, or "apo", form (PBD ID 5OYC), which has a classic $(\beta/\alpha)_8$-barrel fold.[171] Crystallographic data collection and refinement statistics are given in **Table A1**. Privateer results showing validation for Glc (BGC) and Xyl (XYS) residues in XXXG(2F)-CjGH5D(E255A) covalent and XXXG(2F)-β-DNP are given in **Table A2**

In the co-crystallized complex, electron density corresponding to the near-complete glycone was observed spanning the -4 to -1 subsites[100] within the active site cleft, with the exception of the non-reducing-terminal xylosyl residue that was not modelled. In keeping with the mechanism-based design of the inhibitor, the electron density clearly revealed the covalent attachment of C1 of the 2-fluoroglucosyl residue in subsite -1 with $O\varepsilon_1$ of the catalytic nucleophile Glu390 (distance ~1.4 Å). The 2-fluoroglucosyl residue was in a relaxed $^4C_1$ conformation, while the sidechain of Glu390 was rotated from its position in the apo enzyme structure, including an average 1.1 Å translation of $O\varepsilon_1$ toward the sugar ring (**Figure 2.3**).

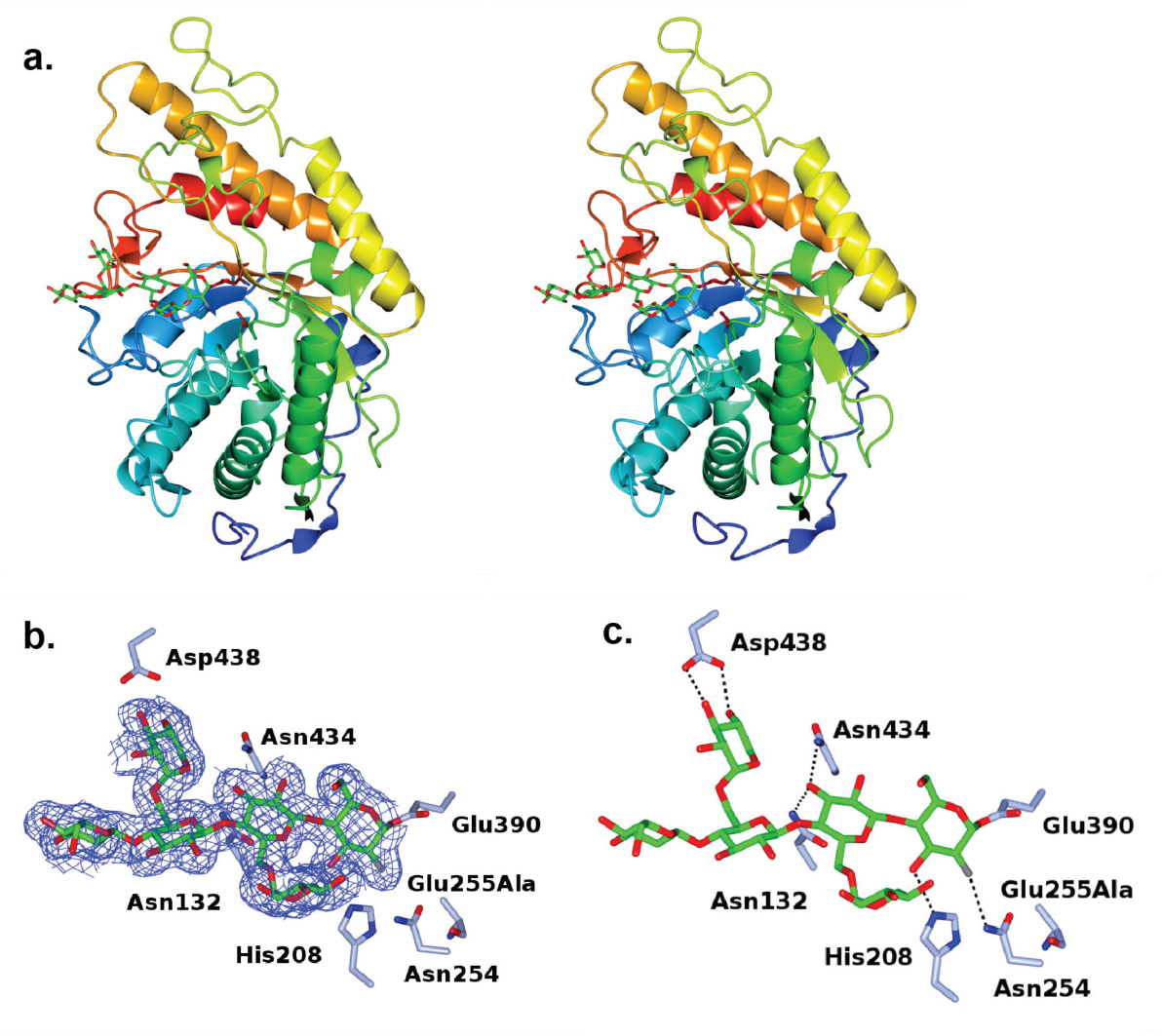


Figure 2.2: Tertiary structure of the covalent fluoroglycosyl-enzyme intermediate formed by cocrystallization of CjGH5D(E255A) with XXXG(2F)-β-DNP (a) Divergent (wall-eyed) stereo cartoon representation of the secondary structure, colour ramped from the N-terminus (blue) to the C-terminus (red), with the inhibitor represented as sticks with C atoms in green, and the catalytic residues with C atoms in tan. (b) Maximum-likelihood/σA weighted 2Fobs − Fcalc electron density map contoured at an r.m.s.d. level of 1 σ for the ligand XXXG-2F. (c) Hydrogen bonding interactions with the ligand. In panels b and c, side chains of interacting residues are shown in ice blue and hydrogen bonds are shown as dashed lines

The structure of a covalent complex of wild-type CjGH5D with the *N*-bromoacetylglycosylamine inhibitor XXXG-NHCOCH$_2$Br was determined previously, in which the catalytic acid-base residue was labelled after displacement of the bromide (PDB ID 5OYD)[171].

Superposition of XXXG-NHAc-CjGH5D with the XXXG(2F)-CjGH5D(E255A) glycosyl-enzyme complex revealed the basis for the distinct specificity of the two inhibitor types (**Figure 2.3**). Whereas the anomeric carbon of Glc-1 in the XXXG(2F)-β-DNP mechanism-based inhibitor was situated directly above the catalytic nucleophile (Glu390), the three-atom extension of the *N*-bromoacetyl group placed the α-carbon of the "warhead" in a position suitable for side-on attack by the catalytic acid/base (Glu255), and effectively out of reach of Glu390. Moreover, this latter reaction required the displacement of the sidechain of Glu255 (Cδ) by an additional 2 Å away from the ring atoms of Glc-1 to accommodate the reactive moiety. These steric and sidechain plasticity requirements are likely to underlie the current and previous observations that *N*-bromoacetylglycosylamine reagents tend to label catalytic acid/base residues in glycosidases (except in the case of off-active-site labelling[157]).

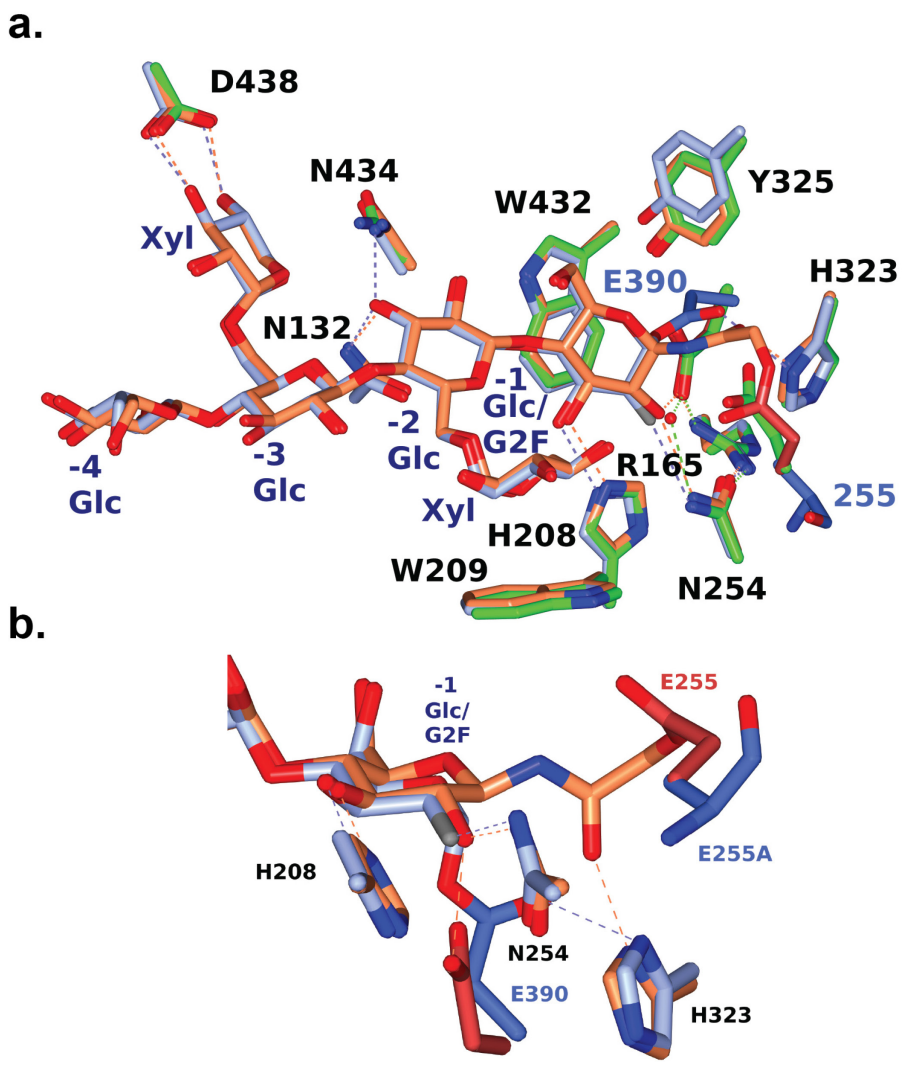


Figure 2.3: Active-site superposition of CjGH5D and corresponding covalent inhibitor complexes (a) Full view (b) Close-up of the −1 subsite. Side chains are shown with carbon atoms coloured as follows: XXXG-NHAc-CjGH5D covalent complex (PDB ID: 5OYD) in coral and XXXG(2F)-CjGH5D(E255A) covalent complex (PDB ID: 6HAA) in ice blue, and with catalytic residue carbon atoms shown in pale crimson and light blue respectively. Main chain atoms are shown for E255A for clarity

2.4 Conclusions

The chemo-enzymatic synthesis of the heptasaccharide XXXG(2F)-β-DNP has enabled the production of one of the most structurally complex 2-deoxy-2-fluorosugar mechanism-based

inactivators to date. Possessing active site-residue specificity advantages over analogous xyloglucan oligosaccharide bromoketone C-glycoside and *N*-bromoacetylglycosylamine inhibitors[170], XXXG(2F)-β-DNP is anticipated to find continued use in structure-function analyses of *endo*-(xylo)glucanases from diverse GH families. In particular, *N*-bromoacetylglycosylamines generally appear to react with catalytic acid-base residues in configuration-retaining GHs[43,160,171], contrasting the exclusive catalytic nucleophile specificity of the Withers-type fluorosugar inhibitors[147].

Chapter 3: *N*-bromoacetylglycosylamine-based Irreversible Inhibitors of Mixed-linkage Glucanases

3.1 Introduction

Glycoside hydrolases (GHs) form the largest class of Carbohydrate-Active Enzymes (CAZymes)[78]. Both reversible and irreversible inhibitors have aided in the biochemical and structural characterization of GHs, enabling the elucidation of important details of active site-specificity and identification of key catalytic residues[174,276,277]. In conjunction with the free enzyme structure, inhibitor-bound enzyme crystallographic structures provide high-resolution molecular insight into interactions of GHs with their substrates during catalysis. Irreversible inhibitors, in particular, have been used in a number of studies to define the mechanism and specificity of GH action (see reviews[147,148,187,239,248,249,278] and recent primary literature[170,182–184,206,250,251,279–281]). This knowledge can be further utilized to guide mutational analysis, and direct enzyme engineering to produce GHs with altered biochemical properties.

β(1,3)/ β (1,4)-Glucans (mixed-linkage β-glucans, MLGs) are hemicellulosic polysaccharides found in the plant cell walls of grasses and cereals such as rye, sorghum, rice, and wheat, as well as lower vascular plants such as the horsetails[36], algae[282], and some fungi[283]. MLGs are particularly abundant in oats (3%–8% dry weight) and barley (2%– 20% dry weight), comprising about 65-75% of the soluble, non-starch glycan content of the barley endosperm[284]. MLGs play a major role in maintaining plant cell wall structure and providing a store of readily mobilized glucose for the developing seedling[283]. The importance of MLG is underscored by their importance in the agri-food industry and their association with human health[285]. MLG is an important component of dietary fibre and has been known to be linked to reducing postprandial blood glucose and insulin levels, as well as lowering cholesterol levels in humans[286]. Cereal dietary

fibre, with the aid of the population of symbiotic gut microbiota[44], can provide as much as 10% of the daily caloric requirements in humans[287]. MLGs have also been of interest as feedstocks in large scale bioethanol production for second-generation biofuels[47].

Structurally, MLGs are composed predominantly of β-(1,4)-glucopyranosyl linkages that are interspersed with β-(1,3) linkages, resulting in a kinked polysaccharide chain. The ratio of (1,4)-β-D-glucopyranosyl to (1,3)-β-D-glucopyranosyl residues generally ranges from 2:1 to 3:1[288]. Although a small percentage of longer β-(1,4)-linked stretches are present, the β-(1,3) linkages typically occur after every two or three β-(1,4)-linked residues (**Figure 1.4**). Thus, hydrolysis with specific MLGases results in a limit-digest comprised of GG3G and GGG3G, where "3" represents the position of the β(1,3) linkage in the gluco-oligosaccharide[43–45].

Due to the ubiquity of MLGs in nature, a broad diversity of GHs has evolved to hydrolyze this family of polysaccharides. In addition to specific MLGases (EC 3.2.1.73; formally, (1,3)/(1,4)-β-D-glucan 4-glucanohydrolase, also known as licheninase), MLG can also be depolymerised by *endo* β-(1,4)-glucanases[289,290] (classic cellulases, EC 3.2.1.4) and *endo* β-(1,3)-glucanases (laminarinases, EC 3.2.1.6)[291]. MLG-active enzymes are found in diverse organisms, including archaea, bacteria, fungi, and plants. Specific MLGases are most commonly associated with Glycoside Hydrolase Family 16 (GH16; bacterial, fungal, and plant representatives)[86,292] and GH17 (plant representatives)[288]. MLGase activity has also been found in members of GH5[43], GH6, GH7, GH8, GH9[293,294], GH11, GH12[295], GH26[296,297], and GH51 (for an overview, see [78], http://www.cazy.org/Glycoside-Hydrolases.html), comprising both anomeric-configuration-retaining and –inverting enzymes[94,275]. In addition to their biological importance in carbohydrate metabolism[44,47,292,298,299], MLGases have numerous biotechnological applications in biomass

conversion to high-value products such as monosaccharides and oligosaccharides[300], animal feed treatment to improve digestibility[301], and in brewing to reduce mash viscosity[302].

Despite a long history of enzymological characterization[292], including through the use of bespoke aryl glycoside substrates[303–305], there are few specific inhibitors for MLGases. These are essentially limited to non-hydrolyzable thio-oligosaccharide analogs as competitive (i.e. non-covalent, reversible) active site inhibitors[296,306]. On the other hand, epoxyalkyl-glucosides, -cellobiosides, -laminaribiosides, and -mixed-linkage oligosaccharides have been synthesized and tested as covalent inhibitors of MLGases, in some cases enabling the solution of enzyme-ligand complexes by X-ray crystallography[280,307–309]. Hence, there is significant scope for the development of novel covalent inhibitors of MLGases based on MLG oligosaccharide scaffolds.

Among the classes of irreversible, active site-directed inhibitors, *N*-bromoacetylglycosylamines which comprise a highly electrophilic "warhead" attached to a carbohydrate as a specificity determinant, have been of particular interest[157,167,168,170]. Compared with mechanism-based inactivators of GHs[147,187], the *N*-bromoacetylglycosylamine functional group is intrinsically reactive and does not require activation by the catalytic machinery of the target enzyme. As such, these inhibitors are in principle agnostic to the mechanism of catalysis, thus potentially making them useful probes both for retaining and inverting GHs. The strong covalent nature of the bond formed with nucleophilic residues such as active site carboxylates makes *N*-bromoacetylglycosylamine powerful probes to label and identify important catalytic residues[160] (in contrast, for example, C-1 glycosyl-enzyme intermediates from mechanism-based inhibitors are often susceptible to slow turnover[159,281]). Early examples of the use of *N*-bromoacetylglycosylamines to inactivate *exo*-glycosidases and *endo*-glycanases include *Cellulomonas fimi exoglycanase*[160], cyanogenic β-glucosidase (linamarase) from cassava[164],

Escherichia coli β-galactosidase[166], *Thermoanaerobacterium saccharolyticum* β-xylosidase[167], *Flavobacterium meningosepticum*[163] and *Aspergillus wentii*[165] β-glucosidases, *Aspergillus carbonarius* β-xylosidase[168] and β-glucosidase[161], and *Thermobifida fusca* β-xylosidase[169].

More recently, our group has reported the straightforward synthesis of *N*-bromoacetylglycosylamine inhibitors, based on highly branched xyloglucan oligosaccharides, with remarkable potency toward *endo*-xyloglucanases from diverse GH familes[170]. In turn, the corresponding heptasaccharide inhibitor enabled the determination of inhibitor-bound crystal structures of two bacterial GH5 enzymes[43,171]. Here, this facile synthetic protocol is extended to mixed-linkage oligosaccharide scaffolds to generate three novel *N*-bromoacetylglycosylamine based inhibitors comprising one trisaccharide and two tetrasaccharide motifs. These specific motifs were selected because the corresponding oligosaccharides can be enzymatically produced as limit-digest products in large quantities using previously established protocols[44,310]. The potency of these inhibitors was demonstrated with a vanguard GH16 MLGase from the human gut bacterium *Bacteroides ovatus* by inhibition kinetics, intact protein mass spectrometry, and protein crystallography.

3.2 Materials and Methods

3.2.1 General Synthetic and Analytic Techniques

All solvents and other chemicals used were of reagent grade or better and purchased from Sigma Aldrich unless mentioned otherwise. Ultrapure water (18.2 MΩ cm^{-1}) was used for all kinetic experiments. Thin-layer chromatography (TLC) was carried out on aluminium plates coated with silica gel 60; analytes were visualized by dipping in an H_2SO_4 solution (10%) and charring using a heat gun. The Amberlite® IR120H^+ resin was washed with a copious amount of methanol before use. Bromoacetic anhydride was purchased from Acros Organics. All ^{13}C- and

^{1}H-NMR data were collected on a Bruker Avance 400 MHz spectrometer at room temperature. NMR spectra were referenced to $CHCl_3$ at 7.27 ppm or HOD at 4.79 ppm[257]. Small molecule mass spectrometry was performed on a Waters LC-MS system including Waters 2695 HPLC and Waters ZQ mass spectrometer equipped with ESCI ion source. HPAEC-PAD was performed on a Dionex ICS-5000 system equipped with an AS-AP autosampler and temperature-controlled sample tray, run in a sequential injection configuration using CHROMELEON 7 software, as described previously[310]. The chromogenic substrate 2'-chloro-4'-nitrophenyl β-glycoside of GG3G (GG3G-β-CNP) was purchased from Megazyme International (Ireland), product code: O-CNPBG3. Oat β-glucan powder (70%) was obtained from Garuda International (California, USA).

The enzyme VvEG16 (ΔV152) was produced by Dr. Nicholas McGregor as per an established protocol[310]. Recombinant BoGH16 was produced in *E. coli* BL21 (DE3) and purified by nickel affinity chromatography as described previously[44] with the exception that HEPES buffers were used instead of sodium phosphate (binding buffer: 20 mM HEPES pH 7.4, 500 mM sodium chloride, 20 mM imidazole; elution buffer: 20 mM HEPES pH 7.4, 500 mM sodium chloride, 500 mM imidazole). The protein for use in crystallography was further purified by size exclusion chromatography using Superdex 75 (GE Healthcare) packed in XK 16/100 column (GE Healthcare) run with 10 mM HEPES pH 7.0. The fractions were inspected for purity by SDS-PAGE and pure fractions were pooled and concentrated in Vivaspin centrifugal filters (GE Healthcare). Final protein concentration was determined by spectrophotometry at 280 nm in an Epoch Microplate Spectrophotometer (BioTek) using the extinction coefficient 54890 $M^{-1}cm^{-1}$, which was calculated from the primary sequence using ProtParam tool from the ExPASy Bioinformatics Resource Portal[44].

3.2.2 Chemoenzymatic Synthesis and Characterization of Target Compounds

Preparation of GG3G and GGG3G oligosaccharides

Oat β-glucan powder (30 g, 70%) was suspended in 25 mM sodium citrate buffer, pH 6.5 (1 L), heated to 50 °C and stirred for 20 min. 100 µL of 0.45 mM *Bo*GH16 enzyme was added to this suspension and the reaction was stirred overnight at 37 °C, following which the mixture was concentrated under reduced pressure. The remaining syrup was then freeze-dried under vacuum to yield a pale powder, confirmed by HPAEC-PAD chromatography to contain a mixture of trisaccharide Glcβ-(1,4)-Glcβ-(1,3)-Glc (GG3G) and tetrasaccharide Glcβ-(1,4)-Glc β-(1,4)-Glcβ-(1,3)-Glc (GGG3G) (**Figure B1**) as the only oligosaccharides[44].

20 g of this crude oligosaccharide mixture was per-*O*-acetylated at 60 °C using pyridine (150 mL) and acetic anhydride (100 mL) with *N,N*-dimethylaminopyridine (0.5 g) as a catalyst. After 3 h, the reaction mixture was concentrated under reduced pressure to near dryness, and subsequently dissolved in 7 mL of dichloromethane (DCM) and loaded on to a silica flash column directly. The peracetylated tri- and tetrasaccharide were separated using flash chromatography with isocratic toluene/acetone (6.5:1) as the solvent system ((per-*O*Ac)GG3G: R_f= 0.3, yield= 3g, (per-*O*Ac)GGG3G: R_f= 0.25, yield= 1 g). The per-*O*-acetylated oligosaccharides were stored at room temperature and were subsequently deprotected using Zemplen conditions, as needed. The characterization of the resulting deprotected oligosaccharides, as well as their corresponding peracetylated congeners, was in accordance with the previous literature[304,311,312].

Preparation of G3GGG oligosaccharide

G3GGG was prepared according to a previously described protocol,[310] with modifications. For large scale preparation, 5 g of oat β-glucan powder (70%) was suspended in 25mM sodium

citrate buffer, pH 6 (1 L) with 100 µL of 600 µM VvEG16 (ΔV152) and stirred for 48 h at 37 °C. Upon concentration under reduced pressure and subsequent lyophilization, a pale powder was obtained. This powder was confirmed by HPAEC-PAD to contain a mixture of glucose, cellobiose and Glcβ-(1,3)-Glcβ-(1,4)-Glcβ-(1,4)-Glc (G3GGG) (**Figure B2**). Isolation of pure acetylated oligosaccharide was achieved using peracetylation and flash chromatography, using conditions as described in the previous section (per-*O*Ac)G3GGG: R_f= 0.25, yield= 700 mg). Deprotection using Zemplen conditions yielded pure deprotected oligosaccharide in near quantitative yield.

(per-*O*Ac)G3GGG: R_f= 0.25 (toluene/acetone: 6.5:1). ^{1}H-NMR ($CDCl_3$, 400 MHz, **Figure B3**): δ 6.21 (H1α, d, J_{H1-H2}= 3.69 Hz, 0.45 H), 5.62 (H1β, d, J_{H1-H2}= 8.39 Hz, 0.54H), 5.38 (t, J= 9.78, 0.5H), 5.19 (dd, J= 9.27, 8.95, 0.5H), 5.08- 4.78 (m, 7H), 4.52-3.55 (m, 19H), 2.14-1.94 (s, CH_3). ^{13}C-NMR: ($CDCl_3$, 100.6 MHz, **Figure B4**): δ 168.56- 170.69 (CO), 100.89, 100.73, 100.37 ($C1^{II}$, $C1^{III}$, $C1^{IV}$), 91.64 (C1β), 89.02 (C1α), 78.65, 76.33, 76.04, 75.8, 75.69, 73.65, 73.02, 72.92, 72.86, 72.83, 72.66, 72.6, 72.4, 72.26, 71.84, 71.76, 71.11, 70.85, 70.63, 69.53, 69.47, 68.01, 67.89, 62.37, 61.86, 61.70, 61.32, 20.95-20.44 (CO*C*H_3); Monoisotopic m/z calculated for $C_{52}H_{70}O_{35}Na^+$: 1277.36; LC-MS found: 1277.6; ESI-HRMS found: 1277.3602

G3GGG: ^{1}H-NMR ($CDCl_3$, 400 MHz, **Figure B5**): δ 5.24 (H1α, d, J_{H1-H2} = 3.72 Hz 0.40H), 4.76 ($H1^{II}$, d, J_{H1-H2} = 7.95 Hz, 1H), 4.67 (H1β, d, J_{H1-H2} = 7.95 Hz, 0.63H), 4.55 ($H1^{III}$, $H1^{IV}$, d, d, J= 8.14 Hz, 7.95 Hz, 2H), 4.01- 3.27 (m, 23H); ^{13}C-NMR (D_2O, 100.6 MHz, **Figure B6**): δ 102.98, 102.76, 102.52 ($C1^{II}$, $C1^{III}$, $C1^{IV}$), 95.96 (C1β), 92.02 (C1α), 84.14, 78.83, 78.68, 78.59, 78.57, 76.19, 75.79 , 75.74, 75.68, 75.01, 74.45, 74.26, 74.1, 73.63, 73.19, 73.15, 71.5, 71.43, 70.32, 69.78, 69.65, 68.19, 60.89, 60.76, 60.21, 60.09; Monoisotopic m/z calculated for $C_{24}H_{42}O_{21}Na^+$: 689.21; LC-MS found: 689.2; ESI-HRMS found: 689.2115

General synthesis of *N*-bromoacetylglycosylamines:

The final products were obtained using previously established methods with modifications[157,170] **(Scheme 3.1)**. 150 mg of the corresponding oligosaccharide was dissolved in 15 mL of 30% ammonium hydroxide, and ammonium bicarbonate (2 eq.) was added to the mixture. The mixture was then stirred overnight (20 h) at 42 °C, and thereafter concentrated under reduced pressure to a powder before lyophilization to yield a pale amorphous solid

GG3G-β-NH$_2$: ^{1}H-NMR (D_2O, 400 MHz, **Figure B7**): δ 4.73 ($H1^{II}$, d, J= 9.28 Hz, 1H), 4.51 ($H1^{III}$, d, J= 8.27 Hz, 1H), 4.13-3.29 (m, 18 H)

GGG3G-β-NH$_2$: ^{1}H-NMR (D_2O, 400 MHz, **Figure B8**): δ 4.70 ($H1^{II}$, d, J= 9.15 Hz, 1H), 4.50 ($H1^{III}$, d, J= 8.60 Hz, 1H), 4.48 ($H1^{IV}$, d, J= 7.96 Hz, 1H), 4.10-3.15 (m, 24H)

G3GGG-β-NH$_2$: ^{1}H-NMR (D_2O, 400 MHz, **Figure B9**): δ 4.74 ($H1^{II}$, d, 1H), 4.55 ($H1^{III}$, d, J= 7.85 Hz, 1H), 4.53 ($H1^{IV}$, d, J= 8.07 Hz, 1H), 4.13-3.18 (m, 24H)

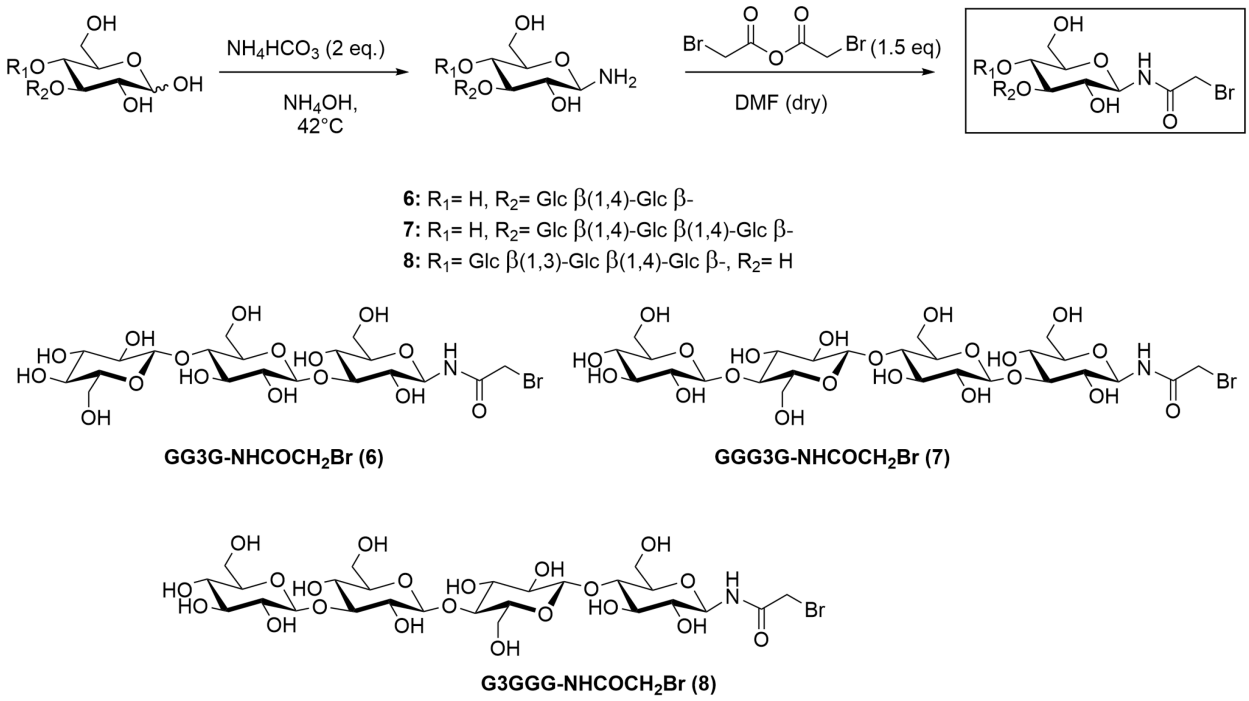
The aminated sugar (150 mg) was subsequently dissolved in 7 mL of dry *N*,*N*-dimethylformamide (DMF) under an inert atmosphere, and stirred for 10 min to form a pale-yellow suspension. Bromoacetic anhydride (1.5 eq.) was added to the reaction mixture as a solution in 0.5 mL dry DMF solution through a syringe. Thereafter, the solution turned into a homogenous dark yellow solution. The reaction was monitored via TLC. Generally, the reaction time was between 1.5 to 2 h. The reaction mixture was concentrated in the presence of silica (7 g) and dry-loaded on to a flash column. The product was eluted using a water/isopropanol/ethyl acetate (1:3:6) solvent system. Fractions containing the product were pooled together, concentrated, and dried to a pale powder. Final yields were 100 mg (54%) for **6**, 44 mg (25%) for **7**, and 68 mg (39%) for **8**.

GG3G-NHCOCH$_2$Br (6): R_f= 0.25 (water/isopropanol/ethyl acetate 1:3:6); ^{1}H-NMR (D_2O, 400 MHz, **Figure B10**): δ 5.05 (H1β, d, J= 9.30 Hz, 1H), 4.84 ($H1^{II}$, d, J= 8.06 Hz, 1H),

4.57 ($H1^{III}$, d, J= 7.82 Hz, 1H), 4.03 (s, $COCH_2Br$, 2H), 3.98-3.37 (m, 18H); ^{13}C-NMR: (D_2O, 100.6 MHz, **Figure B11**): δ 171.31 (C=O), 102.77, 102.74, 84.92, 79.76, 78.79, 77.58, 76.73, 76.18, 75.69, 75.05, 74.34, 73.38, 71.65, 69.69, 67.97, 60.83, 60.26, 60.07, 28.16 (*C*H_2Br); Monoisotopic m/z calculated for $C_{20}H_{34}BrNO_{16}Na^+$: 646.10; LC-MS found: 646.1; ESI-HRMS found: 646.0948.

GGG3G-NHCOCH$_2$Br (7): R_f= 0.25 (water/isopropanol/ethyl acetate 1:3:6); ^{1}H-NMR (D_2O, 400 MHz, **Figure B12**): δ 5.03 (H1β, d, J= 9.31 Hz, 1H), 4.82 ($H1^{II}$, d, 1H), 4.54 ($H1^{III}$, $H1^{IV}$, d, d, J=8.36 Hz, J= 8.42 Hz, 2H), 4.00 (s, $COCH_2Br$, 2H), 4.02-3.33 (m, 24H); ^{13}C-NMR: (D_2O, 100.6 MHz, **Figure B13**): δ 171.34 (C=O), 102.77, 102.5, 79.74, 78.61, 78.05, 76.72, 76.19, 75.69, 75.1, 75.05, 74.24, 73.35, 73.14, 71.75, 69.66, 60.78, 60.12, 59.95, 27.98 (*C*H_2Br); Monoisotopic m/z calculated for $C_{26}H_{44}BrNO_{21}Na^+$: 808.15; LC-MS found: 808.5; ESI-HRMS found: 808.1477

G3GGG-NHCOCH$_2$Br (8): R_f= 0.25 (water/isopropanol/ethyl acetate 1:3:6); ^{1}H-NMR (D_2O, 400 MHz, **Figure B14**): δ 5.02 (H1β, d, J= 9.19 Hz, 1H), 4.76 ($H1^{II}$, d, 1H) 4.56 ($H1^{III}$, d, J= 8.19 Hz, 1H), 4.53 ($H1^{IV}$, d, J= 8.09 Hz, 1H), 4.00 (s, $COCH_2Br$, 2H), 3.95-3.31 (m, 24H); ^{13}C-NMR: (D_2O, 100.6 MHz, **Figure B15**): δ 171.32 (C=O), 102.99, 102.77, 102.50, 84.18, 79.75, 78.62, 78.08, 76.72, 76.19, 75.69, 75.04, 74.24, 73.24, 73.64, 73.36, 73.14, 71.75, 69.80, 69.67, 68.21, 60.97, 60.79, 60.13, 59.97, 28.04 (*C*H_2Br); Monoisotopic m/z calculated for $C_{26}H_{44}BrNO_{21}Na^+$: 808.15; LC-MS found: 808.1; ESI-HRMS found: 808.1483



Scheme 3.1: Synthesis of inhibitors 6, 7 and 8

3.2.3 Inhibition Kinetics Measurement

The determination of time-dependent enzyme inhibition kinetics was performed essentially as previously described[170,171]. Briefly, 100 µL solutions of BoGH16 (8.9 µM) were incubated with a range of concentrations of inhibitor (0.019 to 1.25 mM) at 37 °C in 50 mM sodium citrate buffer pH 6.5 (with 0.1 mg mL^{-1} bovine serum albumin to prevent non-specific loss of activity). A control experiment was run in parallel, in which no inhibitor was added to the buffered enzyme/BSA solution. Periodically, 10 µL of this solution was withdrawn and diluted 1:100 in 50 mM sodium citrate pH 6.5 buffer, and 100 µL of the diluted incubate was added to 100 µL of the pre-incubated solution of chromogenic substrate GG3G-β-CNP in ultrapure water at 37 °C (0.2 mM final substrate concentration in the assay).

Residual activities were determined by measuring the rate of release of the chromophore 2-chloro-4-nitrophenolate over 1–2 min in a 1 cm quartz cuvette maintained at 37 °C at 405 nm in

an Agilent Cary 60 UV–vis spectrophotometer (ε= 16.57 mM^{-1} cm^{-1}). The inhibition data were fitted according to the Kitz-Wilson model[221] (**Equation 1.4**), as shown in **Figures 3.1a, 3.2a**, and the inactivation rate constants (k_{app}) were measured. Individual K_i and k_i values were determined (unless otherwise stated) by fitting **Equation 1.5** to plots of k_{app} versus inhibitor concentrations [I] (**Figures 3.1b, 3.2b**) by nonlinear regression using Origin Pro graphing software, as previously described[170].

3.2.4 Intact-protein Mass Spectrometry Labelling Studies

A 20 μL solution of BoGH16 (3.27 μM in 50 mM pH 6.5 sodium citrate buffer) containing 2.5 mM inhibitor was incubated for 3 h (for **6**, **7**) or 7 h (for **8**) at 37 °C. A negative control experiment omitting the inhibitor from the incubate was run in parallel. Intact-protein masses were determined on a Waters Xevo LC-ESI-MS Q-TOF with a NanoAcuity UPLC system as previously described[264], and analyzed using Masslynx 4.0 software.

3.2.5 Protein Crystallography

To reproduce previous crystallography[44], crystals of unliganded recombinant BoGH16 were obtained by first screening sitting drops in 96-well format at room temperature (set-up using a Phoenix robot, Art Robbin). A hit observed in the JCSG+ screen (Qiagen) condition G7 was pursued for optimization by screening around this condition, varying the PEG concentration in one dimension and the buffer pH in the other (24-well format hanging drops, set up by hand). Large, orthorhombic crystals were readily reproduced by mixing 27.2 mg mL^{-1} protein solution one-to-one with reservoir solution comprised of 0.1 M succinic acid pH 7.2 and 15% (w/v) PEG3350. To obtain the inhibitor complex structure, crystals were soaked for one hour in mother liquor containing inhibitor at 2.5 mM, after which time they were cryoprotected in mother liquor supplemented with 20% ethylene glycol. X-ray data from these crystals were collected at

Advanced Photon Source beamline 23 ID-B. Datasets were indexed and integrated using XDS[265]. The structure was determined by molecular replacement using PHASER from the CCP4i2 software suite[313] with the original unliganded BoGH16 structure (PDB ID: 5NBO)[44] as the search model. After density modification using PARROT[314], iterative rounds of manual model building and refinement were performed in COOT[315] and REFMAC5[269], respectively. The quality of the model was monitored using MOLPROBITY[316] and sugar conformations were validated using PRIVATEER[271].

3.3 Results and Discussion

3.3.1 Chemoenzymatic Synthesis and Characterization of Target Compounds

Oat β-glucan powder is abundant in mixed-linkage β-glucan (70% by mass)[286] and therefore constitutes an inexpensive agricultural source of the polysaccharide. Taking advantage of this, large-scale preparation of the mixed-linkage gluco-oligosaccharides GG3G and GGG3G was achieved by digestion with a specific mixed-linkage *endo*-glucanase previously characterized in our laboratory, *Bacteroides ovatus*, GH16 (*Bo*GH16, **Figure B1)**. Similarly, a mixed-function *Vitis vinifera* GH16 *endo*-glucanase variant (*Vv*EG16-ΔV152) characterized by our group[310] was used to produce G3GGG via limit-digestion of oat β-glucan (**Figure B2**). To assist purification on a multi-gram scale, per-*O*- acetylation, flash chromatography on silica gel, and Zemplen deprotection was employed.

Subsequently, an established,[157,168,170,317] two-step synthetic protocol (**Scheme 3.1**) was applied and further optimized to install the electrophilic "warhead". Conveniently, both steps were performed on the unprotected oligosaccharides, thereby avoiding additional, and potentially tedious, protection and deprotection steps. Thus, GG3G, GGG3G, and G3GGG were quantitatively aminated via dissolution in ammonium hydroxide and addition of ammonium

bicarbonate at 42 °C. Here, a 20 h reaction time was sufficient, versus two-day reactions reported previously[170,318]. Following removal of the solvent and reagents under reduced pressure, amide formation using bromoacetic anhydride in DMF was facile in all cases. The resulting *N*-bromoacetyl glycosamines were amenable to flash chromatography on silica gel to produce pure GG3G-NHCOCH$_2$Br (**6**), GGG3G-NHCOCH$_2$Br (**7**), and G3GGG-NHCOCH$_2$Br (**8**) in moderate yields.

3.3.2 Evaluation of Inhibitor Potency with BoGH16 as a Model MLGase

To investigate their potency as inhibitors, compounds **6**, **7**, and **8** were incubated with the well-characterized representative MLGase, BoGH16[44]. A time- and concentration-dependent inactivation of the enzyme (8.9 μM) was observed upon incubation with GG3G-NHCOCH$_2$Br (compound **6**, **Figure 3.1**) and GGG3G-NHCOCH$_2$Br (compound **7**, **Figure 3.2**) at concentrations of 0.019 mM to 1.25 mM. At high concentrations of inhibitors, the decline in enzyme activity is particularly rapid. Indeed, complete inhibition of the enzyme is apparent within the first 20 min of incubation of BoGH16 with ≥ 0.156 mM inhibitor **7**.

The covalent nature of the inhibition was demonstrated by intact-protein mass spectrometry after incubation of BoGH16 with 2.5 mM of inhibitors **6** and **7**. In both cases, a 1:1 labelling stoichiometry is observed with no signs of over-labelling, as has been observed with *N*-bromoacetylglycosylamines in some cases[157,164,170,317]. Under identical concentrations of enzyme and G3GGG-NHCOCH$_2$Br (compound **8)**, there was no indication of labelling, even with extended incubation time (**Figure 3.3**). These results are consistent with the known specificity of BoGH16, which requires a β-(1,3) linkage spanning the -2 and -1 enzyme subsites[44] (subsite nomenclature according to ref. [100]). However, this compound could see future use in inhibiting MLGases preferring substrates with alternative linkage profiles[310].

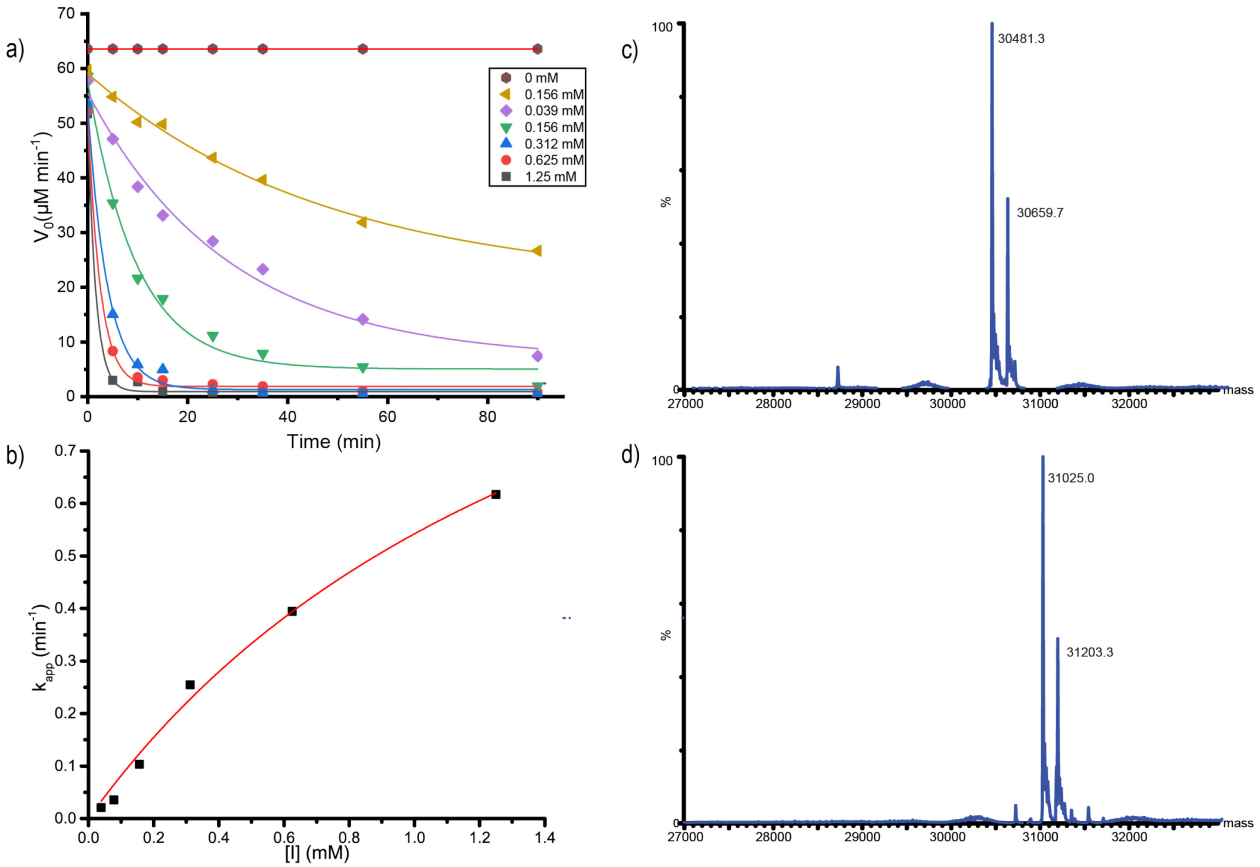


Figure 3.1: Inhibition kinetics and intact-protein MS of BoGH16 by inhibitor 6: a) Plot of initial-rate enzyme activity versus time (single determinations) at inhibitor concentrations indicated. b) Plot of pseudo-first-order rate constants (k_{app}), obtained from the fitted curves shown in panel a, versus inhibitor concentration. c) wild type BoGH16. Expected mass: 30480.8 Da; a significant phosphogluconoylation peak is observed at +178 Da. d) BoGH16 incubated with 2.5 mM GG3G-NHCOCH$_2$Br (624.1 Da) at 37 °C for 3 h. Expected mass: 31025.0 Da; the phosphogluconoylated proteins are equally labelled.

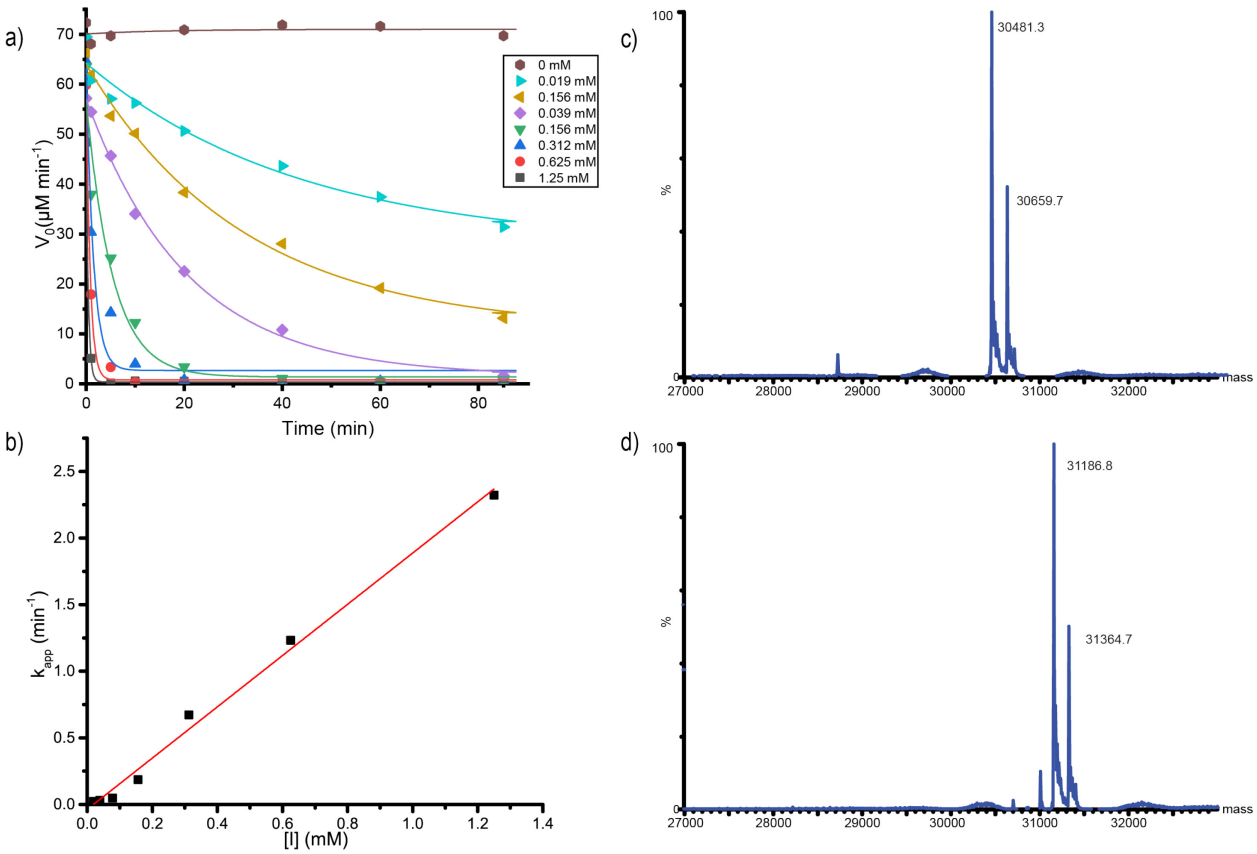


Figure 3.2: Inhibition kinetics and intact-protein MS of BoGH16 by inhibitor 7: a) Plot of initial-rate enzyme activity versus time (single determinations) at inhibitor concentrations indicated b) Plot of pseudo-first-order rate constants (k_{app}), obtained from the fitted curves shown in panel a, versus inhibitor concentration. c) wild type BoGH16. Expected mass: 30480.8 Da; a significant phosphogluconoylation peak is observed at +178 Da. d) BoGH16 incubated with 2.5 mM GGG3G-NHCOCH$_2$Br (785.2 Da) at 37 °C for 3 h. Expected mass: 31186.1 Da; the phosphogluconoylated proteins are equally labelled.

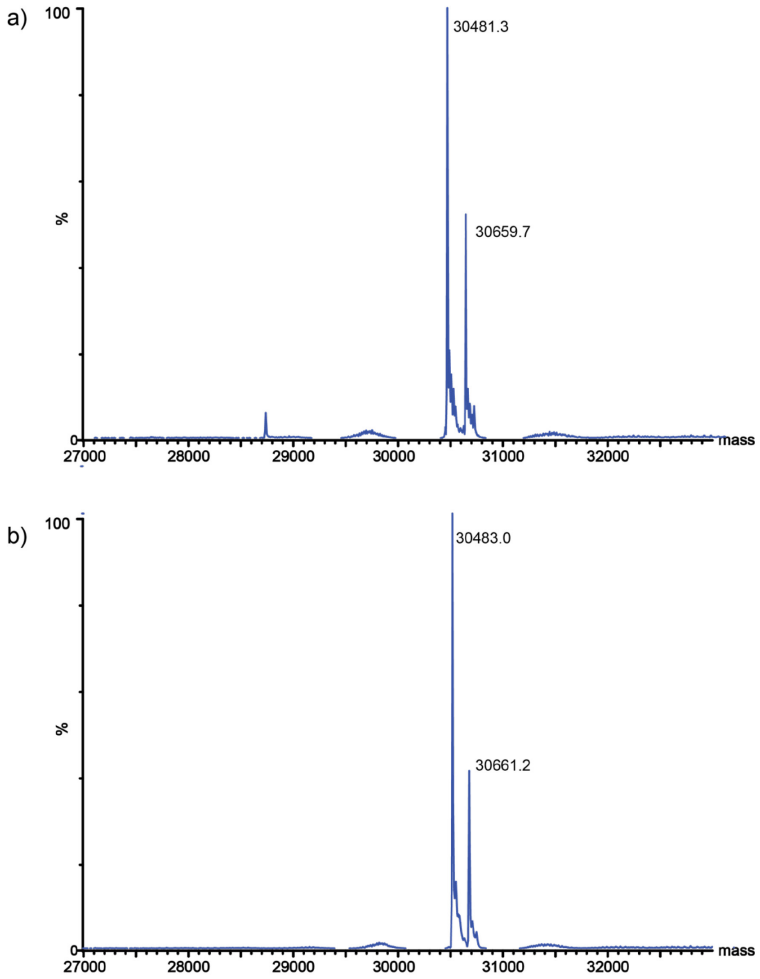


Figure 3.3: Intact-protein MS of BoGH16 in presence of 8: a) wild type BoGH16. Expected mass: 30480.8 Da; a significant phosphogluconoylation peak is observed at +178 Da. b) BoGH16 incubated with 2.5 mM G3GGG-NHCOCH$_2$Br (785.2 Da) at 37 °C for 7 h. Expected mass: 30480.8 Da; the phosphogluconoylated proteins are equally labelled.

Fitting of the Kitz-Wilson model[221] (**Equations 1.4** and **1.5**) to the time- and concentration-dependent inhibition data obtained for GG3G-NHCOCH$_2$Br enabled the calculation of the specific inhibition constants with BoGH16: K_i 1.68 ± 0.47 mM, k_i 1.45 ± 0.27 min^{-1} (k_i/K_i 0.86 mM^{-1} min^{-1}). These values compare favourably to those of other *N*-bromoacetylglycosylamine inhibitors, which likewise exhibit K_i values in the millimolar range[43,157,161–163,168–170,174]. However, the inactivation constants, k_i, for other *N*-bromoacetylglycosylamine inhibitors are reportedly 2-3

orders of magnitude lower[43,157,161–163,168–170]. The large error on the K_i value, in this case, reflects an apparently poor ability to saturate the enzyme and a correspondingly flat k_{app} vs. [I] curve (**Figure 3.1**). Likewise, saturation inhibition kinetics were not achieved with the longer inhibitor GGG3G-NHCOCH$_2$Br (compound **7**). Thus, a linear fit to the data was used to obtain a k_i/K_i value of 1.92 $mM^{-1}min^{-1}$, which is comparable to that of GG3G-NHCOCH$_2$Br. As discussed further below, these results are concordant with the three known negative subsites of BoGH16[44].

3.3.3 Structural Analysis of an Inhibitor-bound Enzyme Complex

To provide molecular insight into the mechanism of inhibition, the complex formed by the incubation of GGG3G-NHCOCH$_2$Br with BoGH16 was solved to 2.13 Å resolution by X-ray crystallography (**Table B1**). The space group, unit cell dimensions, and presence of two chains in the asymmetric unit were identical to the previously determined unliganded and GGG3G product-complex structures of BoGH16 (PDB IDs: 5NBO and 5NBP, respectively[44]) despite crystallization in a different condition. Likewise, the overall tertiary structure was virtually identical to that of these previous structures, comprising a canonical β-jelly roll fold typical of GH16[86] (**Figure 3.4a**). The inhibitor-bound complex structure was successfully obtained by soaking crystals in a solution containing 2.5 mM inhibitor (the same concentration used for intact protein MS experiments and two-fold greater than the highest concentration used for kinetics experiments). Notably, the observed electron density corresponding to the inhibitor was far superior in chain A, with the lower occupancy observed in chain B attributing to the proximity of its active site to a protein chain from a neighbouring asymmetric unit. Nevertheless, the near-complete glycone was modelled into the active site of chain A, spanning subsites -1 to -3, including the *N*-bromoacetylated glucose in subsite -1. The observed electron density unambiguously revealed that the catalytic acid/base residue, Glu148, was labelled by nucleophilic substitution of the bromide (**Figure 3.4a, 3.4b,**

3.4c). The density of the non-reducing-terminal glucosyl residue corresponding to subsite -4 was too poor to model, reminiscent of the relatively high disorder of this residue in the BoGH16:GGG3G product complex (**Figure 3.4d**)[44].

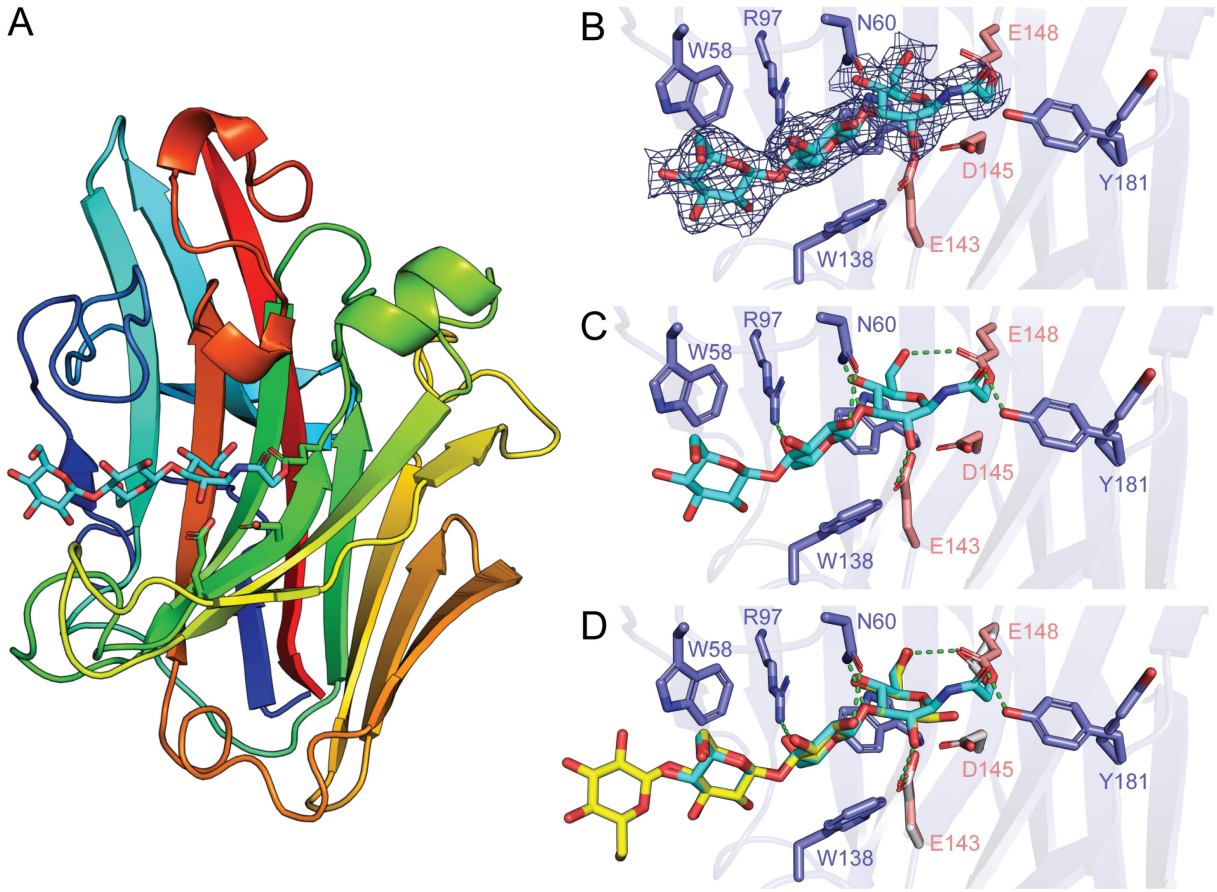


Figure 3.4: Crystal structure of the BoGH16-GGG3GNHCOCH$_2$Br inhibitor covalent complex: a) Overall structure with BoGH16 shown in cartoon representation and colour ramped from blue (N-terminus) to red (C-terminus), the inhibitor shown as cyan sticks, and the key catalytic residues shown as green sticks. b) Inhibitor-bound active site showing the F_o-F_c omit density map (generated by Privateer[271]) of the inhibitor contoured at 3.0 σ. Catalytic residues are shown as salmon sticks and other important active site residues are shown as slate sticks. c) Inhibitor-bound active site showing hydrogen bonding interactions (3.0 Å or shorter) with the ligand as green dashed lines. d) Superposition of the inhibitor complex and the GGG3G complex (PDB ID: 5NBP) structures. The tetrasaccharide is shown as yellow sticks and catalytic residues from 5NBP are shown as white sticks.

The β orientation of the *N*-bromoacetyl group extended the electrophilic warhead away from the BoGH16 catalytic nucleophile Glu143, which is positioned directly below the anomeric carbon. Instead, the *N*-bromoacetyl group filled the space that typically accommodates a water molecule in the glycosyl-enzyme intermediate of the normal catalytic cycle, ideally positioning the electrophile for nucleophilic attack by Glu148. The natural role of this residue is to deprotonate a water molecule for turnover of the glycosyl-enzyme intermediate in the second step of the canonical double displacement GH mechanism[94]. That the catalytic acid/base residue is labelled, rather than the catalytic nucleophile, is similar to previous observations with *N*-bromoacetylglycosylamine inhibitors[160,163,164,167,169,171].

Notably, the Glu148 sidechain underwent a minor rotation and shift in orientation to accommodate the covalent linkage between Oε and *N*-acetyl group of the inhibitor (**Figure 3.4d**). The *N*-acetylated glucose in subsite -1 assumed a ^{4}E envelope conformation (**Table B2**), distinct from the favoured $^{4}C_{1}$ chair conformation observed for the subsite -1 glucose in the BoGH16:GGG3G product complex (PDB ID: 5NBP). As such, it appears that the anomeric carbon is pulled up by the *N*-acetyl linkage to Glu148 (**Figure 3.4d**), suggesting an overall rigidity of the protein structure itself. The rest of the glycone overlaid very well with the GGG3G oligosaccharide, retaining the hydrogen bonding and aromatic stacking interactions observed for the natural product complex (**Figure 3.4d**). Interestingly, the same "swung in" rotamer of Tyr181 observed in the GGG3G-complex[44] was seen here as well, but instead made a hydrogen bond with the carbonyl oxygen of the *N*-acetyl group of the inhibitor, rather than the anomeric (C1) hydroxyl of the glucosyl residue in subsite -1 (**Figure 3.4d**). Considering these observations together, the ^{4}E envelope conformation can be rationalized by the energy required to bring the sugar out of the

relaxed $^4C_1$ chair conformation being lower than the energy of compromising the numerous interactions between the glycone and the various sidechains of the enzyme active site.

3.4 Conclusions

In this chapter, the chemoenzymatic synthesis of three novel *N*-bromoacetylglycosylamine-based covalent inhibitors with tri- and tetrasaccharide backbones of mixed β-(1,3)/(1,4) linkage is presented. The potential of these inhibitors to inactivate mixed-linkage glucanases active on β-(1,3)/(1,4)-glucans is demonstrated through protein-labelling and inhibition kinetics studies, as well as inhibitor-bound enzyme crystallography. Tri- and tetrasaccharide inhibitors with β-(1,3) linkage between the reducing end glucosyl residues showed potency towards a canonical GH16 MLGase. A 2.13 Å crystal structure of this enzyme was obtained in complex with the tetrasaccharide inhibitor covalently bound to the catalytic acid/base residue of the enzyme. The production of these compounds with an abundant natural source and a two-step synthetic protocol with minimal chemical modifications is anticipated to aid in the biochemical, kinetic, and structural characterization of novel mixed-linkage glucanases.

Chapter 4: Synthesis and Application of Mechanism-based 2-deoxy-2-fluoro Oligosaccharide Inhibitors of β-(1,3) and β-(1,3)/(1,4) Glucanases

4.1 Introduction

β-(1,3) glucans are widely distributed in nature and serve various biological functions, e.g. as storage polysaccharides in brown algae (laminarin), in the cell walls of fungi and yeast for structural support (pachyman), and as exopolysaccharides in bacteria (curdlan) enabling pathogenicity and symbiotic interactions with hosts[319]. The structure of β-(1,3) glucans varies with the source of polysaccharide, but they are typically composed of linear β-(1,3) linked glucose chains, which may be substituted with β-(1,6) glucose-linked side chains and intrachain β-(1,6) glucosyl linkages. The ratio between β-(1,3)and β-(1,6) linkages varies significantly[320]. The algal β-(1,3) storage polysaccharide laminarin is the most abundant polysaccharide in marine algae, constituting up to 25-35% of their dry cell wall weight[320,321]. Laminarin has a relatively low molecular weight of about 5000 Da, with a degree of polymerization between 20 and 30[319].

β-(1,3)/(1,4) glucan or MLG is major plant cell wall hemicellulose of commelinoid monocots such as oats, barley, sorghum and other cereals and grasses, as well as in horsetails[36], algae[282], and some fungi[283]. MLG is made up of chains of β-(1,4) and β-(1,3) linked glucosyl residues with the number of β-(1,3) linkages ranging between 25-30%. Typically, cellotriose and cellotetraose units are linked by β-(1,3) glycosidic bonds, though longer units are also present.

β-(1,3) glucanases are enzymes that catalyze the depolymerization of β-(1,3) glucan via *exo* (EC 3.2.1.58) or *endo* (EC 3.2.1.39) mechanisms, among which, the *endo* β-(1,3) glucanases typically fall within the GH families 5, 16, 17, 55, 64, 81, 128, 152, 157 and 158[78]. They have been isolated from various sources and organisms such as fungi[322,323], plants[324], bacteria[325], archaea[326] and marine invertebrates[327].

β-(1,3) glucanases have found important applications in bioethanol production[111], against fungal infections[112], in yeast extract production[73] and to produce high-value oligosaccharides[113,114]. β-(1,3)/(1,4) glucanases (MLGases, EC 3.2.1.73) break down MLG and are of high commercial importance due to their numerous biotechnological applications in animal feed treatment[301], in laundry detergents to remove food stains[328], biomass conversion to high-value products[300], and in brewing to reduce mash viscosity and increase extract yield[302]. They are also valuable for the production of prebiotic oligosaccharides[329].

Retaining GHs employ a Koshland double displacement mechanism to catalyze the hydrolysis of carbohydrates[94]. Upon substrate binding, these enzymes utilize two active site carboxylic amino acid residues (Glu and/or Asp) located about 5–6 Å apart to perform a two-step hydrolysis wherein the nucleophile enables the formation of a glycosyl-enzyme intermediate via substitution at the substrate, which gets hydrolyzed with assistance from the general acid/base residue acting as a base. Both the catalytic steps proceed through an oxocarbenium-ion like transition state (**Figure 1.7b**).

2-deoxy-2-fluoro glycoside inhibitors are well-established as potent inactivators of retaining glycoside hydrolases[185,194]. They are structural mimics of the GH natural substrates capable of covalently and irreversibly binding and inactivating a GH enzyme by kinetically trapping the glycosyl-enzyme intermediate (**Figure 1.12**). The 2-deoxy-2-fluoro substitution inductively destabilizes the transition states, while also eliminating key hydrogen-bonding interactions of the C2-hydroxyl group with enzyme active site residues, and decreases the reaction rate of both the aforementioned catalytic steps, while fluorine (pK_a of HF = 3.2) or 2',4' dinitrophenyl substitution (pK_a of dinitrophenol = 4.1) at the anomeric carbon facilitates departure of the aglycone thereby increasing the rate of enzyme glycosylation. These inhibitors are

essentially “slow-substrates” with a highly variable lifetime of the glycosyl-enzyme intermediate, depending on the GH studied. Where observed, kinetic analysis indicates that incomplete inhibition of retaining GHs by 2F DNP inhibitors is due to significant rates turnover by hydrolysis[206,281] or transglycosylation[195,207,252].

In this chapter, the chemoenzymatic synthesis of two 2F DNP mechanism-based inhibitors of β-(1,3)-glucan-active enzymes based on β-(1,3) linked diglucoside (laminaribiose) and β-(1,3)/(1,4) linked triglucoside scaffolds is reported. The potency of these inhibitors was evaluated using a putative β-(1,3) glucanase using enzyme inhibition kinetic analysis and intact-protein mass spectrometry.

4.2 Materials and Methods

4.2.1 General Synthetic and Analytic Techniques

All reagents were analytical or HPLC grade and were purchased from Sigma-Aldrich, Alfa-Aesar or ACROS Organics. For anhydrous reactions, glassware was dried overnight in a 100-150 °C oven and purged with argon before use. Solvents were dried by stirring with activated 4 Å molecular sieves overnight under argon. TLC was performed using aluminum sheet TLC plates (0.25 mm) pre-coated with Merck silica gel 60 F254, using ethyl acetate: hexanes, toluene: acetone or water: methanol: ethyl acetate as solvent systems (particular solvent ratios specified below), and visualized by a UV lamp and/or 10% sulfuric acid in water with charring using a heat gun. Flash chromatography was performed using Merck silica gel 60 with ethyl acetate: hexanes, toluene: acetone or water: methanol: ethyl acetate as mobile phases. All ^{19}F-, ^{13}C- and ^{1}H-NMR data were collected on a Bruker Avance 400 MHz spectrometer at room temperature (100.6 MHz and 376.5 MHz for ^{13}C- and ^{19}F-, respectively). The NMR spectra were referenced to the solvent as follows[257]: HOD= 4.79 ppm, $CHCl_3$= 7.27 ppm. MALDI-TOF MS data were collected on a

Bruker Autoflex instrument in reflectron mode over *m/z* 700-3500 using 6-Aza-2-thiothymine (ATT) as the matrix. HRMS data were obtained using either a Waters Xevo G2-S Q-TOF or Waters/Micromass LCT TOF mass spectrometer in positive-ion mode, via direct infusion through an electrospray ion source. All kinetic data were obtained using an Agilent Cary 60 UV-Vis spectrophotometer equipped with a Peltier temperature-controlled cell holder. The data were fit with **Equation 1.4 or 1.5** by linear least-squares analysis using nonlinear regression using OriginPro 8 (OriginLab, Northampton, MA, USA). The BoGH16 enzyme was prepared using the method described in Chapter 3 (see section 3.2.1). The BuGH158 and BuGH16 enzymes were obtained as per an established protocol and their purities were confirmed by SDS-PAGE (Supplementary Figure 16: http://doi.org/10.14288/1.0388792)[123].

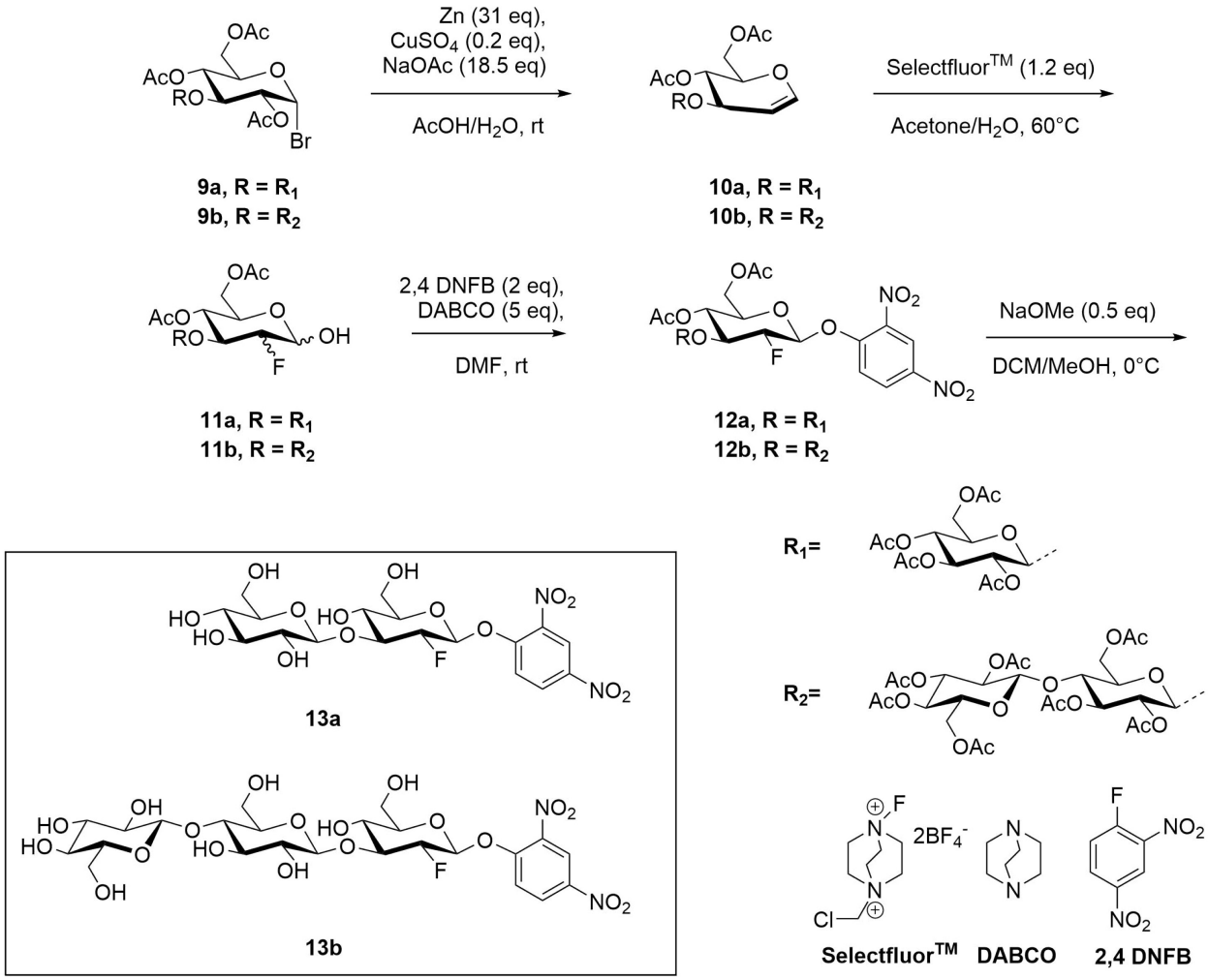
4.2.2 Preparation of Oligosaccharides G3G and GG3G

Laminaribiose (G3G) was prepared enzymatically from 3 g of the commercially available polysaccharide laminarin from *Laminaria digitate* as per an established protocol[123] which yielded a mixture of glucose, G3G, and a small amount of a mixed-linkage trisaccharide. The crude mixture was acetylated by dissolving in pyridine (60 mL) and acetic anhydride (40 mL) and heated at 60 °C for 3 h. Following the workup, the mixture was purified using column chromatography with toluene/acetone (6.5:1) as the eluent to yield 3.5 g of pure (per-*O*Ac)G3G (R_f= 0.33). The characterization data was in accordance with literature values[330]. Similarly, (per-*O*Ac)GG3G (2.8 g) was obtained using a previously established method from oat β-glucan (see Chapter 3 Materials and Methods).

4.2.3 General Synthesis of (2F)-β-DNP Inhibitors

The synthesis of target inhibitors proceeded through a five-step synthetic protocol (**Scheme 4.1**) starting from the per-*O*-acetylated oligosaccharides, (per-*O*Ac)G3G (1.2 g) and (per-

*O*Ac)GG3G (1 g), which were dissolved in 30 mL dichloromethane and converted to their respective α-glycosyl bromides **9a** and **9b** as per previously established protocols[304,330] using 33% hydrogen bromide in glacial acetic acid. The characterization data of compounds **9a** (81% yield) and **9b** (89% yield) were in accordance with literature values[304,330].



Scheme 4.1: Synthesis of target inhibitors (13a and 13b)

Per-*O*-acetylated G3G glycal (**10a**) and GG3G glycal (**10b**): Compounds **10a** and **10b** were synthesized by adapting previously established methods[259,281]. Briefly, glycosyl α-bromide (1000 mg **9a** or 900 mg **9b**) was dissolved in 30 mL of acetic acid and slowly added to a suspension of Zn (31 eq.), NaOAc (18.5 eq.), and $CuSO_4.5H_2O$ (0.2 eq.) in 20 mL of water which had been

stirring for 5 min. The solution was stirred vigorously for 3 h at room temperature and subsequently filtered through Celite. The mixture was concentrated under reduced pressure to remove the solvent, re-dissolved in DCM (100 mL), and washed with $NaHCO_3$ (3×) and brine (1×). The DCM layer was concentrated, and the resulting syrup was purified using flash chromatography with isocratic ethyl acetate/hexanes (1:1) as the mobile phase (for **10a**: R_f = 0.3, for **10b**: R_f = 0.25). The evaporation of solvents under reduced pressure yielded **10a** (150 mg, 19% yield) and **10b** (160 mg, 21% yield) as white solids.

10a: ^{1}H-NMR (**Figure C1**, 400 MHz, $CDCl_3$): δ 6.39 (d, $J_{(H1-H2)}$ = 6.26 Hz, 1H, H1), 5.18-5.11 (m, 2H, $H1^{II}$, $H1^{III}$), 5.00 (t, J= 9.62 Hz, 1H), 4.88 (dd, J= 8.37 Hz, J= 9.20 Hz, 1H), 4.79 (t, J= 5.10 Hz, 1H), 4.65 (d, J= 7.97 Hz, 1H), 4.33-4.00 (m, 6H), 3.69-3.65 (m, 1H), 1.92-2.01 (6s, 18H, $COCH_3$). ^{13}C-NMR (**Figure C2**, 100.6 MHz, $CDCl_3$): δ 171.15-169.31 (6×CO), 145.30 (C1), 98.85 ($C1^{II}$), 97.38 (C2), 73.93, 72.93, 72.00, 71.37, 69.96, 68.37, 67.92, 61.90 (C6 or $C6^{II}$), 61.56 (C6 or $C6^{II}$), 21.06-20.62 (6×CH_3); Monoisotopic m/z calculated for $C_{24}H_{32}O_{15}Na^+$: 583.16; MALDI-TOF MS found: 583.1; ESI-HRMS found: 583.1636

10b: ^{1}H-NMR (**Figure C3**, 400 MHz, $CDCl_3$): δ 6.45 (d, $J_{(H1-H2)}$ = 6.26 Hz, 1H, H1), 5.25-5.23 (m, 1H), 5.19-4.82 (m, 6H), 4.66 (d, J= 8.05 Hz, 1H), 4.58 (dd, J= 12.01 Hz, J= 1.83 Hz, 1H), 4.52 (d, J= 7.97 Hz, 1H), 4.39-4.33 (m, 3H), 4.13-4.03 (m, 4H), 3.78 (t, J= 9.47 Hz, 1H), 3.68-3.60 (m, 2H), 2.13-1.98 (9s, 27H, $COCH_3$). ^{13}C-NMR (**Figure C4**, 100.6 MHz, $CDCl_3$): δ 170.64-169.17 (9×CO), 145.26 (C1), 100.92 ($C1^{III}$), 99.06 ($C1^{II}$), 97.44 (C2), 76.43, 74.03, 73.07, 73.00, 72.68, 72.13, 71.73, 71.66, 70.05, 68.00, 67.90, 61.79 (C6/$C6^{II}$/$C6^{III}$), 61.74 (C6/$C6^{II}$/$C6^{III}$), 61.67 (C6/$C6^{II}$/$C6^{III}$), 20.99-20.67 (9×CH_3); Monoisotopic m/z calculated for $C_{36}H_{48}O_{23}Na^+$: 871.25; MALDI-TOF MS found: 871.3; ESI-HRMS found: 871.2483.

1-hydroxy 2-deoxy-2-fluoro-(per-*O*-acetylated)glycosides (**11a, 11b**): Utilizing established synthetic protocols[260,261,281] with modifications (see Results and Discussion), the corresponding glycal **10a** (150 mg) or **10b** (160 mg) was dissolved in minimum amount of acetone and added to a stirring suspension of Selectfluor^{TM} (1.2 eq.) in 18 mL of acetone/water (5:1). The solution was stirred first at room temperature for 16 hr, and then at 60 °C for 3 hr. It was subsequently filtered over a Celite plug and concentrated under reduced pressure. The resulting syrup was re-dissolved in minimum DCM and loaded to a silica column for partial purification using flash chromatography using isocratic ethyl acetate/hexanes (1.2:1) as eluent (R_f= 0.25-3.00 for **11a** and **11b**). This reaction yielded a mixture of diastereomers of **11a** (60 mg, 38% yield) and **11b** (90 mg, 54% yield) respectively, due to the addition of 2-fluoro and 1-hydroxyl functional groups either equatorially or axially. These mixtures of isomers were characterized using ^{19}F-NMR and high-resolution mass spectrometry and used in the next step directly without further purification.

11a: ^{19}F-NMR (**Figure 4.2**, 376.5 MHz, $CDCl_3$): δ -199.17 (ddd, $J_{(F2-H2)}$ = 50.96 Hz, $J_{(F2-H3)}$ = 14.47 Hz, $J_{(F2-H1)}$ = 1.55 Hz, 0.13F, *gluco* β-anomer), -200.60 (dd, $J_{(F2-H2)}$ = 49.39 Hz, $J_{(F2-H3)}$ = 12.66 Hz, 0.59F, *gluco* α-anomer), -205.68 (ddd, $J_{(F2-H2)}$ = 49.60 Hz, $J_{(F2-H3)}$ = 27.70 Hz, $J_{(F2-H1)}$ = 6.18 Hz, 1F, *manno* α-anomer); Monoisotopic m/z calculated for $C_{24}H_{33}O_{16}FNa^+$: 619.17; MALDI-TOF MS found: 619.1; ESI-HRMS found: 619.1650

11b: ^{19}F-NMR (**Figure 4.3**, 376.5 MHz, $CDCl_3$): δ -198.84 (ddd, $J_{(F2-H2)}$ = 51.11 Hz, $J_{(F2-H3)}$ = 15.11 Hz, $J_{(F2-H1)}$ = 1.91 Hz, 0.06F, *gluco* β-anomer), -200.40 (dd, $J_{(F2-H2)}$ = 49.51 Hz, $J_{(F2-H3)}$ = 12.63 Hz, 0.39F, *gluco* α-anomer), -205.82 (ddd, $J_{(F2-H2)}$ = 49.33 Hz, $J_{(F2-H3)}$ = 27.53 Hz, $J_{(F2-H1)}$ = 6.70 Hz, 1F, *manno* α-anomer); Monoisotopic m/z calculated for $C_{36}H_{49}O_{24}FNa^+$: 907.25; MALDI-TOF MS found: 907.1; ESI-HRMS found: 907.2496

β-2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated)glycosides (**12a, 12b**): The synthesis for these compounds followed previously reported glycosylation methods[191,281] with modifications. The diastereomeric mixture of **11a** (60 mg) or **11b** (90 mg) was dissolved in 5 mL of dry DMF containing 4 Å activated molecular sieves and anhydrous DABCO (5 eq.). A solution of 2 eq. of 2,4-DNFB in 1 mL dry DMF was added to the reaction mixture through an oven-dried steel needle. The reaction was stirred for 3.5 h, concentrated using rotary evaporator, re-dissolved in DCM (100 mL) and washed with $NaHCO_3$ (3×) and brine (1×). TLC analysis of the product mixtures using toluene/acetone (2.5:1) indicated 3 spots each when visualized under UV lamp (indicating the presence of aromatic group), as well as on charring after dipping in 10% sulfuric acid solution (**Figure 4.4**, R_f = 0.40, 0.47, 0.55 for diastereomeric mixture of **12a**, R_f = 0.34, 0.38, 0.47 for diastereomeric mixture of **12b**), of which, spots A2 and B2 were separated by column chromatography using isocratic toluene/acetone (6.5:1) and identified to be the desired β-gluco diastereomers **12a** and **12b**. The spots A1/B1 and A3/B3 were identified to be the *manno* α and *gluco* α diastereomers respectively (for ^{19}F NMR spectra and assignments of A1 and A3, see **Figures C5** and **C6** respectively). The column fractions indicating the desired product were pooled together and concentrated using rotary evaporation to yield diastereomerically pure **12a** (24 mg, 33%) and **12b** (26 mg, 24%).

12a: ^1H-NMR (**Figure C7**, 400 MHz, $CDCl_3$): δ 8.78 (d, $J_{(H'3\text{-}H'5)}$ = 2.70 Hz, 1H, H′3), 8.44 (dd, $J_{(H'5\text{-}H'6)}$ = 9.16 Hz, $J_{(H'5\text{-}H'3)}$ = 2.70 Hz, 1H, H′5), 7.39 (d, $J_{(H'6\text{-}H'5)}$ = 9.16 Hz, 1H, H′6), 5.29 (dd, $J_{(H1\text{-}H2)}$ = 7.47 Hz, $J_{(H1\text{-}H3)}$ = 3.43 Hz, 1H, H1), 5.25-4.96 (m, 4H), 4.74 (d, $J_{(H1II\text{-}H2II)}$ = 7.86 Hz, 1H, $H1^{II}$), 4.69-4.64 (m, 1H), 4.34 (dd, J= 12.51 Hz, J= 4.23 Hz, 1H), 4.22 (d, J= 3.76 Hz, 2H), 4.14-4.07 (m, 2H), 3.94-3.89 (m, 1H), 3.72-3.68 (m, 1H), 2.10-2.01 (6s, 18H, $COCH_3$). ^{13}C-NMR (**Figure C8**, 100.6 MHz, $CDCl_3$): δ 170.46-169.05 (6×CO), 153.39 (C′3), 142.43 (C′4), 140.38

(C′5), 128.56 (C′1), 121.72 (C′2), 118.39(C′6), 101.39 ($C1^{II}$), 98.59 ($J_{(C1-F2)}$ = 24.93 Hz, C1), 91.07 ($J_{(C2-F2)}$ = 190.19 Hz, C2), 79.47, ($J_{(C3-F2)}$ = 18.13 Hz, C3), 72.85, 72.56, 71.88, 71.33, 67.97 ($C4^{II}$), 67.37 ($J_{(C4-F2)}$ = 8.31 Hz, C4), 61.70 (C6, $C6^{II}$), 20.69-20.33 (6×CH_3). ^{19}F-NMR (**Figures C9, 4.5**, 376.5 MHz, $CDCl_3$): δ -198.81 (ddd, $J_{(F2-H2)}$ = 50.51 Hz, $J_{(F2-H3)}$ = 15.27 Hz, $J_{(F2-H1)}$ = 3.17 Hz, *gluco* β-anomer); Monoisotopic m/z calculated for $C_{30}H_{35}O_{20}N_2FNa^+$: 785.17; MALDI-TOF MS found: 785.2; ESI-HRMS found: 785.1667

12b: ^{1}H-NMR (**Figure C10**, 400 MHz, $CDCl_3$): δ 8.76 (d, $J_{(H'3-H'5)}$ = 2.38 Hz, 1H, H′3), 8.43 (d, $J_{(H'5-H'6)}$ = 9.16 Hz, $J_{(H'5-H'3)}$ = 2.38 Hz, 1H, H′5), 7.38 (d, $J_{(H'6-H'5)}$ = 9.16 Hz, 1H, H′6), 5.28 (dd, $J_{(H1-H2)}$ = 7.36 Hz, J= 3.97 Hz, 1H, H1), 5.23-5.00 (m, 5H), 4.91 (t, J= 8.07 Hz, 2H), 4.69 (d, $J_{(H1II-H2II)}$ = 7.68 Hz, 1H, $H1^{II}$), 4.53-3.59 (m, 12H), 2.13-1.98 (9s, 27H, $COCH_3$). ^{13}C-NMR (**Figures 4.6, C11**, 100.6 MHz, $CDCl_3$): δ 170.54-169.00 (9×CO), 153.41 (C′3), 142.36 (C′4), 140.28 (C′5), 128.60 (C′1), 121.78 (C′2), 118.22 (C′6), 101.39 ($C1^{III}$), 100.59 ($C1^{II}$), 98.47 ($J_{(C1-F2)}$ = 25.15 Hz, C1), 91.01 ($J_{(C2-F2)}$ = 191.14 Hz, C2), 79.58 ($J_{(C3-F2)}$ = 18.11 Hz, C3), 75.96, 72.89, 72.83, 72.76, 72.16, 72.04, 71.63, 71.50, 67.72, 67.29 ($J_{(C4-F2)}$ = 7.04 Hz, C4), 61.80 ($J_{(C6-F2)}$ = 23.14 Hz, C6), 61.50 ($C6^{II}$, $C6^{III}$), 20.82-20.43 (9×CH_3). ^{19}F-NMR (**Figure C12**, 376.5 MHz, $CDCl_3$): δ -198.32 (ddd, $J_{(F2-H2)}$ = 50.23 Hz, $J_{(F2-H3)}$ = 15.28 Hz, $J_{(F2-H1)}$ = 3.54 Hz, *gluco* β-anomer); Monoisotopic m/z calculated for $C_{42}H_{51}O_{28}N_2FNa^+$: 1073.25; MALDI-TOF MS found: 1073.6; ESI-HRMS found: 1073.2513

β-2’4’-dinitrophenyl 2-deoxy-2-fluoroglycosides (**13a, 13b**): The deprotected glycosides were obtained by Zemplen deprotection of **12a** (24 mg) or **12b** (26 mg) in 10 mL methanol/DCM (9:1) containing 0.5 eq. of NaOMe (25% in MeOH). The solution was stirred at 4°C for 16 h and monitored by TLC. Upon completion, the reaction was neutralized by adding 3 g of silica and concentrated to dryness under reduced pressure to yield a pale silica powder to which the product

was adsorbed. This powder was dry loaded on a flash column and pure product was eluted using water/methanol/ethyl acetate (1:2:18) as the mobile phase. The fractions containing the product were pooled together and concentrated to give a syrup, which was re-dissolved in water and freeze-dried to give a white, fluffy powder **13a** (5.1 mg, 30%) and **13b** (6.6 mg, 40%).

13a: ^{1}H-NMR (**Figure C13**, 400 MHz, D_2O): δ 8.93 (d, $J_{(H'3\text{-}H'5)}$ = 2.65 Hz, 1H, H′3), 8.57 (dd, $J_{(H'5\text{-}H'6)}$ = 9.31 Hz, $J_{(H'5\text{-}H'3)}$ = 2.65 Hz, 1H, H′5), 7.64 (d, $J_{(H'6\text{-}H'5)}$ = 9.31 Hz, 1H, H′6), 5.78 (dd, $J_{(H1\text{-}H2)}$ = 7.62 Hz, J= 2.90 Hz, 1H, H1), 4.74 (d, $J_{(H1II\text{-}H2II)}$ = 7.71 Hz, 1H, $H1^{II}$), 4.30-4.22 (m, 1H), 4.01-3.32 (m, 11H). ^{13}C-NMR (**Figure C14**, 100.6 MHz, D_2O): δ 154.08 (C′3), 141.97 (C′4), 139.22 (C′5), 130.04 (C′1), 122.40 (C′2), 118.10 (C′6), 102.66 ($C1^{II}$), 97.91 ($J_{(C1\text{-}F2)}$ = 23.14 Hz, C1), 91.39 ($J_{(C2\text{-}F2)}$ = 188.12 Hz, C2), 81.43 ($J_{(C3\text{-}F2)}$ = 17.10 Hz, C3), 76.49, 76.18, 75.72, 73.38 (C5), 69.79 ($C4^{II}$), 67.40 ($J_{(C4\text{-}F2)}$ = 7.92 Hz, C4), 60.91 ($C6^{II}$), 60.21 ($J_{(C6\text{-}F2)}$ = 20.12 Hz, C6). ^{19}F-NMR (**Figure C15**, 376.5 MHz, D_2O): δ -199.39 (ddd, $J_{(F2\text{-}H2)}$ = 50.61 Hz, $J_{(F2\text{-}H3)}$ = 15.05 Hz, $J_{(F2\text{-}H1)}$ = 2.39 Hz, *gluco* β-anomer); Monoisotopic m/z calculated for $C_{18}H_{23}O_{14}N_2FNa^+$: 533.10; MALDI-TOF MS found: 533.1; ESI-HRMS found: 533.1018

13b: ^{1}H-NMR (**Figure C16**, 400 MHz, D_2O): δ 8.90 (d, $J_{(H'3\text{-}H'5)}$ = 2.50 Hz, 1H, H′3), 8.55 (dd, $J_{(H'5\text{-}H'6)}$ = 9.33 Hz, $J_{(H'5\text{-}H'3)}$ = 2.50 Hz, 1H, H′5), 7.62 (d, $J_{(H'6\text{-}H'5)}$ = 9.33 Hz, 1H, H′6), 5.76 (dd, $J_{(H1\text{-}H2)}$ = 7.36 Hz, J= 2.21 Hz, 1H, H1), 4.71 (d, $J_{(H1II\text{-}H2II)}$ = 7.40 Hz, 1H, $H1^{II}$), 4.51 (d, $J_{(H1III\text{-}H2III)}$ = 7.33 Hz, 1H, $H1^{III}$), 4.29-3.29 (m, 18H). ^{13}C-NMR (**Figure C17**, 100.6 MHz, D_2O): δ 153.86 (C′3), 141.72 (C′4), 138.94 (C′5), 129.86 (C′1), 122.19 (C′2), 117.89 (C′6), 102.54, ($C1^{III}$), 102.21 ($C1^{II}$), 97.65 ($J_{(C1\text{-}F2)}$ = 24.14 Hz, C1), 91.18 ($J_{(C2\text{-}F2)}$ = 188.12 Hz, C2), 81.11 ($J_{(C3\text{-}F2)}$ = 16.09 Hz, C3), 78.53, 76.24, 75.94, 75.44, 74.77, 74.05, 73.12, 72.94, 67.50 ($C4^{II/III}$), 67.14 ($J_{(C4\text{-}F2)}$ = 8.05 Hz, C4), 60.54 ($C6^{II}$, $C6^{III}$), 60.03 ($J_{(C6\text{-}F2)}$ = 11.06 Hz, C6). ^{19}F-NMR (**Figure C18**, 376.5 MHz, D_2O): δ -199.24 (ddd, $J_{(F2\text{-}H2)}$ = 50.88 Hz, $J_{(F2\text{-}H3)}$ = 15.02 Hz, $J_{(F2\text{-}H1)}$ = 2.42 Hz, *gluco* β-

anomer); Monoisotopic m/z calculated for $C_{24}H_{33}O_{19}N_2FNa^+$: 695.16; MALDI-TOF MS found: 695.1; ESI-HRMS found: 695.1552

4.2.4 Synthesis of β-CNP Laminaribioside

β-2'-chloro-4'-nitrophenyl-(per-*O*-acetylated) laminaribioside (**14**) was synthesized from α-laminaribiosyl bromide **9a** according to a previously established protocol[258] with some modifications. Traditional phase-transfer conditions[331], although generally high-yielding, can result in partially de-*O*-acetylated side-products, which necessitates an extra re-acetylation step to obtain maximum yields[258]. To circumvent this, the sodium salt of 2-chloro-4-nitrophenol (NaCNP) was prepared first by the addition of NaOH (400 mg) into an aqueous solution (50 mL) of 2-chloro-4-nitrophenol (CNP, 1.74 g). The solution was stirred for 10 min and subsequently acetone (200 mL) was added to precipitate the salt. The precipitated salt was filtered, washed with acetone several times, and stored as a dry powder (1.92 g, 98%) at 4°C until further needed. A solution of the sodium phenolate salt (1.5 g, 4.5 eq.) in water (3 mL) was added to a solution of α-laminaribiosyl bromide[304] (**9a,** 1.2 g) and benzyl tri-n-butylammonium chloride (536 mg, 1 eq.) in DCM (120 mL). The reaction mixture was stirred at room temperature for 48 h and monitored via TLC analysis. Subsequently, it was diluted with more DCM (120 mL), washed with water (2×), dried over $MgSO_4$, and concentrated in vacuo. The pure compound (1.1 g, 81%) was isolated via flash chromatography using ethyl acetate/hexanes (3:4) as eluent (R_f= 0.25).

^{1}H-NMR (**Figure C19**, 400 MHz, $CDCl_3$): δ 8.30 (d, $J_{(H'3\text{-}H'5)}$ = 2.63 Hz, 1H, H′3), 8.12 (dd, $J_{(H'\text{-}H'6)}$ = 9.05 Hz, $J_{(H'5\text{-}H'3)}$ = 2.63 Hz, 1H, H′5), 7.22 (d, $J_{(H'6\text{-}H'5)}$ = 9.05 Hz, 1H, H′6), 5.40 (d, $J_{(H1\text{-}H2)}$ = 8.18 Hz, 1H, H1), 5.19-4.90 (m, 5H), 4.65 (d, $J_{(H2\text{-}H1)}$ = 8.18 Hz, 1H, H2), 4.39 (dd, J= 12.31 Hz, J= 4.32 Hz, 1H), 4.25-4.00 (m, 4H), 3.93- 3.88 (m, 1H), 3.72-3.68 (m, 1H), 2.18-2.00 (7s, 21H, $COCH_3$). ^{13}C-NMR (**Figure C20**, 100.6 MHz, $CDCl_3$): δ 170.60-168.79 (7×CO), 157.45

(C′3), 143.38 (C′4), 126.31 (C′5), 125.15 (C′1), 123.66 (C′2), 116.77 (C′6), 101.08 ($C1^{II}$), 99.52 (C1), 78.34 (C3), 72.97, 72.73, 71.92, 71.18, 68.12 (C2), 62.07 (C6 or $C6^{II}$), 61.82 (C6 or $C6^{II}$), 21.01-20.49 (7×CH_3); Monoisotopic m/z calculated for $C_{32}H_{38}O_{20}NClNa^+$: 814.16; MALDI-TOF MS found: 814.3; ESI-HRMS found: 814.1571

β-2'-chloro-4'-nitrophenyl laminaribioside (**15**): Compound **14** (1.1 g) was dissolved in MeOH/DCM (9:1) (30 mL) containing 0.5 eq. of NaOMe (25% in MeOH) and stirred at 4°C overnight. The reaction mixture was then neutralized with Amberlite® IR120 hydrogen foam and purified with flash chromatography (water/methanol/ethyl acetate (1:2:9) to afford **15** (394 mg, 57%). ^{1}H-NMR (**Figure C21**, 400 MHz, D_2O): δ 8.33 (d, $J_{(H'3\text{-}H'5)}$ = 2.75 Hz, 1H, H′3), 8.21 (dd, $J_{(H'5\text{-}H'6)}$ = 9.19 Hz, $J_{(H'5\text{-}H'3)}$ = 2.75 Hz, 1H, H′5), 7.46 (d, $J_{(H'6\text{-}H'5)}$ = 9.19 Hz, 1H, H′6), 5.27 (d, $J_{(H1\text{-}H2)}$ = 7.55 Hz, 1H, H1), 4.65 (d, $J_{(H1II\text{-}H2II)}$ = 7.93 Hz, 1H, $H1^{II}$), 3.95-3.32 (m, 12H). ^{13}C-NMR (**Figure C22**, 100.6 MHz, D_2O): δ 157.46 (C′3),142.67 (C′4), 126.39 (C′5), 124.49 (C′1), 123.56 (C′2), 115.82 (C′6), 102.98 ($C1^{II}$), 99.94 (C1), 83.77 (C3), 76.25, 76.21, 75.74, 73.65, 72.61, 69.80, 67.92, 60.91 (C6 or $C6^{II}$), 60.56 (C6 or $C6^{II}$); Monoisotopic m/z calculated for $C_{18}H_{24}O_{13}NClNa^+$: 520.08; MALDI-TOF MS found: 520.0; ESI-HRMS found: 520.0831

4.2.5 Inhibition Kinetics Measurement

The inhibition of a laminarinase from *Bacteroides uniformis* (BuGH158) in the presence of compounds **13a** and **13b** was measured using a previously described method[170,171]. The chromogenic substrate β-2'-chloro-4'-nitrophenyl laminaribioside (**15**) was used to measure residual activity. The Michaelis-Menten kinetic parameters for **15** were determined in duplicates at 50 °C in 50 mM sodium citrate buffer at pH 6 using a non-linear fit of the Michaelis-Menten equation with concentrations ranging from 0.078-7.5 mM (**Figure C24**).

Inhibition kinetics were performed using a 50 mM sodium citrate buffer at pH 6 at 50 °C (optimum temperature of BuGH158 activity[123]). The range of concentrations of each inhibitor used was 10 mM to 0.31 mM. 100 µL solution of 0.13 mM enzyme in buffer was incubated with the inhibitor (with 0.1 mg mL^{-1} bovine serum albumin added to prevent non-specific loss of activity). At various times up to 180 min, a 10 µL aliquot of this incubate was withdrawn and diluted 1:100 in the same buffer and 100 µL of this diluted enzyme- inhibitor incubate was mixed with 100 µL solution of the chromogenic substrate **15** (0.25 mM after dilution). This solution was added to a 1 cm quartz cuvette prewarmed to 50°C and residual activity of the enzyme at that time interval was determined by measuring the rate of release of 2-chloro-4-nitrophenolate at 405 nm in an Agilent Cary 60 UV–vis spectrophotometer (ε= 15.8 mM^{-1} cm^{-1}). Inhibition kinetics parameters K_i and k_i were determined according to the Kitz-Wilson kinetics model[221] using **Equations 1.4** and **1.5** as described previously[170,171].

4.2.6 Intact-protein Mass Spectrometry Labelling Studies

A 20 µL solution of the enzyme BuGH158 (4.13 µM), BoGH16 (2.60 µM) or BuGH16 (4.24 µM) in 10 mM sodium citrate buffer at pH 6 containing 2.5 mM inhibitor was incubated at 50 °C for 2 h. A control experiment with no inhibitor was run in parallel. Thereafter, intact-protein masses were determined on a Waters Xevo LC-ESI-MS Q-TOF with a NanoAcuity UPLC system as previously described[264], and analyzed using Masslynx 4.0 software.

4.3 Results and Discussion

4.3.1 Chemoenzymatic Synthesis of Target Inhibitors and Substrate

To produce the target inhibitors, the polysaccharides laminarin and MLG were enzymatically digested to yield the oligosaccharides G3G and GG3G respectively, which were then used as starting materials for chemical syntheses. G3G was obtained from commercially

purchased laminarin, which was subjected to a limit digest by a laminarinase from *Bacteroides uniformis* (BuGH158)[123]. This yielded a mixture of oligosaccharides, which was easily separable via column chromatography upon per-*O*-acetylation. Per-*O*-acetylated GG3G was obtained from oat β-glucan MLG through digestion by the proficient MLGase from *Bacteroides ovatus*, BoGH16[44] acetylation, and purification, as described in Chapter 3. Facile preparation of both BuGH158 and BoGH16 enabled multi-gram-scale production of both oligosaccharide starting materials.

The peracetylated oligosaccharides were each converted to their respective glycosyl bromides using established protocols with excellent yields[304,330]. The conversion of glycosyl bromides **9a** and **9b** to the respective glycals **10a** and **10b** was accomplished through a method established previously[281], with minor modification of the mobile phase used for the column chromatography.

The incorporation of the C2-fluoro group in 2F DNP glycosides is typically achieved by electrophilic fluorination of the corresponding per-*O*-acetylated glycal. Previously, highly reactive reagents such as acetyl or trifluoromethyl hypofluorite[332,333], xenon fluoride[334] or elemental fluorine[335] were used for this conversion. More recently, these reagents have been superseded by Selectfluor™ (1-chloromethyl-4-fluoro-1,4-diazoniabicyclo[2.2.2]octane bis(tetrafluoroborate))[336,337], due to its inexpensive commercial availability, ease of handling, high yields, and simultaneous one-pot introduction of an anomeric functional group[336]. However, the stereoselectivity of fluorine and hydroxyl group addition upon reaction with Selectfluor™ depends upon various factors including the specific configuration of the glycal, the nature of the hydroxyl protecting groups, and the choice of solvent and temperature[260]. In numerous previous examples,

all four possible stereoisomers were obtained, yielding difficult-to-separate mixtures[199,261,262,338,339] and additional anomeric acetylation or selective crystallization are often required[191,340].

To overcome this problem, various approaches to achieve diastereomerically pure products have previously been devised. For example, individual functionalized monosaccharide moieties were assembled using trichloroacetimidate chemistry to obtain the desired stereochemical configuration, such as in the synthesis of the disaccharides 2,4-dinitrophenyl 4'-azido-2,4'-dideoxy-2-fluoro-β-xylobioside[215] and 2,4-dinitrophenyl 2-acetamido-2-deoxy- β-D-glucopyranosyl-(1,4)-2-deoxy-2-fluoro-β-D-glucopyranoside[199]. In another approach, specifically engineered glycosynthase enzymes were used to add activated glycosyl donors to the non-reducing end of monosaccharide inhibitors to yield the corresponding di-and tri-saccharide inhibitors[198]. Yet, the use of Selectfluor™ on glycals based on larger (>1) oligosaccharides, particularly those containing β-(1,3) glycosidic linkages has not yet been explored.

Therefore, in an attempt to optimize the reaction to obtain exclusively the *gluco* epimer, small amounts (20 mg each) of **10a** and **10b** were subjected to electrophilic fluorination using Selectfluor™ using several solvents (DMF, acetonitrile, nitromethane/H_2O, and acetone/H_2O mixtures). However, no notable preference for the *gluco* epimer was observed in any case (**Figure 4.1**). Reactions using solvents with $MeNO_2$ or acetone yielded the *gluco* (α and β anomer) and *manno* (α anomer) epimers approximately in the same ratio. Reactions performed in DMF or acetonitrile yielded additional products, that were not further investigated, but were likely the result of solvent participation, as has been observed previously[260,272]. The absence of the *manno* β- anomer in all cases can be rationalized by the anomeric effect as well as the *syn*- addition of Selectfluor™, with the subsequent addition of the nucleophile proceeding with an inversion of configuration, consistent with the previous reports[339–341]. Notably, in our previous work (described

in Chapter 2)[281], the reaction of a highly branched xyloglucan heptasaccharide glycal afforded exclusively the *gluco* epimer with an anomeric α/β mixture (as confirmed by ^{19}F-NMR spectroscopy).

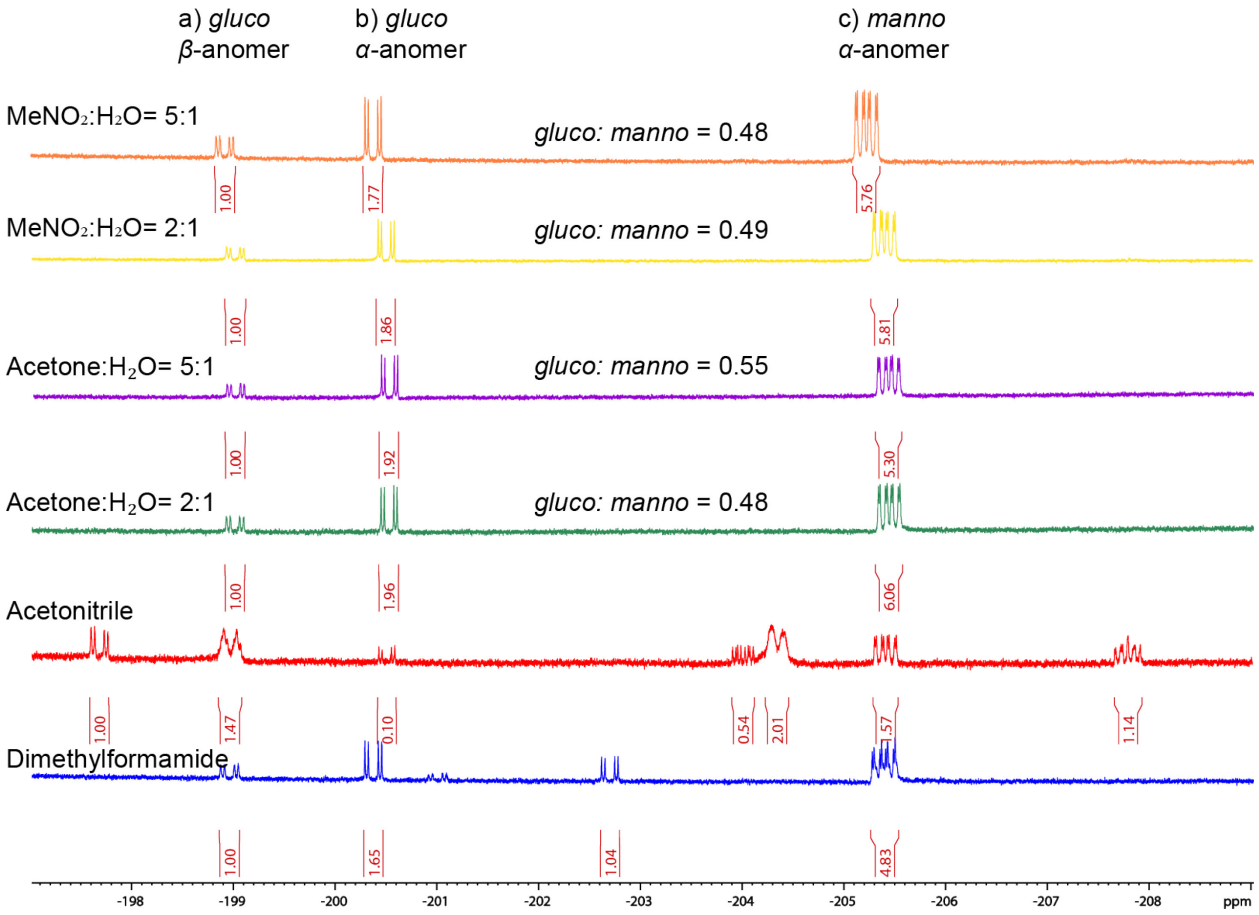


Figure 4.1: Overlay of 19F NMR spectra of crude product mixture 11a obtained by reaction of glycal 10a with Selectfluor™ in various solvents, as indicated with their epimeric product ratio. The diastereomers are identified using fluorine-proton coupling constant values. gluco β-anomer: ddd, J(F2-H2) = 51.78 Hz, J(F2-H3) = 14.35 Hz, J(F2-H1) = 1.89 Hz, gluco α-anomer: dd, J(F2-H2) = 49.96 Hz, J(F2-H3) = 13.06 Hz, manno α-anomer: ddd, J(F2-H2) = 49.48 Hz, J(F2-H3) = 26.97 Hz, J(F2-H1) = 6.35 Hz

Based on these results, the reactions of glycals **10a** and **10b** with Selectfluor™ were scaled up using acetone/water (5:1) as the solvent system. The product mixtures **11a** and **11b** respectively were obtained in moderate yields. The ^{19}F-NMR spectra of the compounds **11a** (**Figure 4.2**) and

11b (**Figure 4.3**) indicated a mixture of diastereomers similar to that observed in the small-scale reaction of glycal **10a** described above.

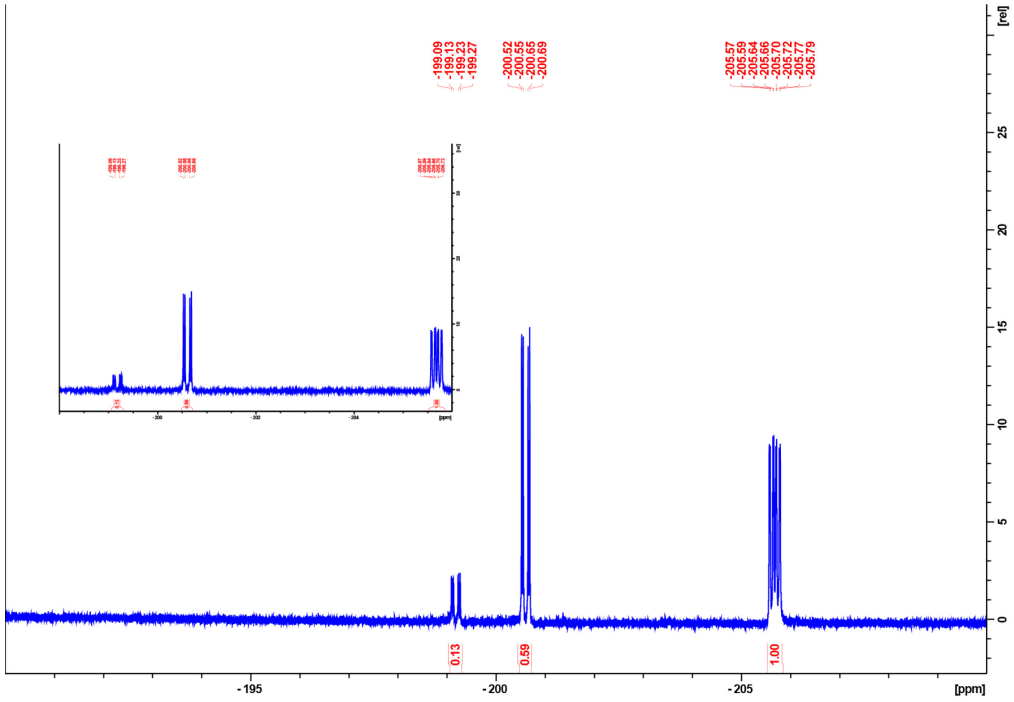


Figure 4.2: ^{19}F NMR spectrum of 11a (1-hydroxyl 2-deoxy-2-fluoro per-*O*-acetylated G3G)

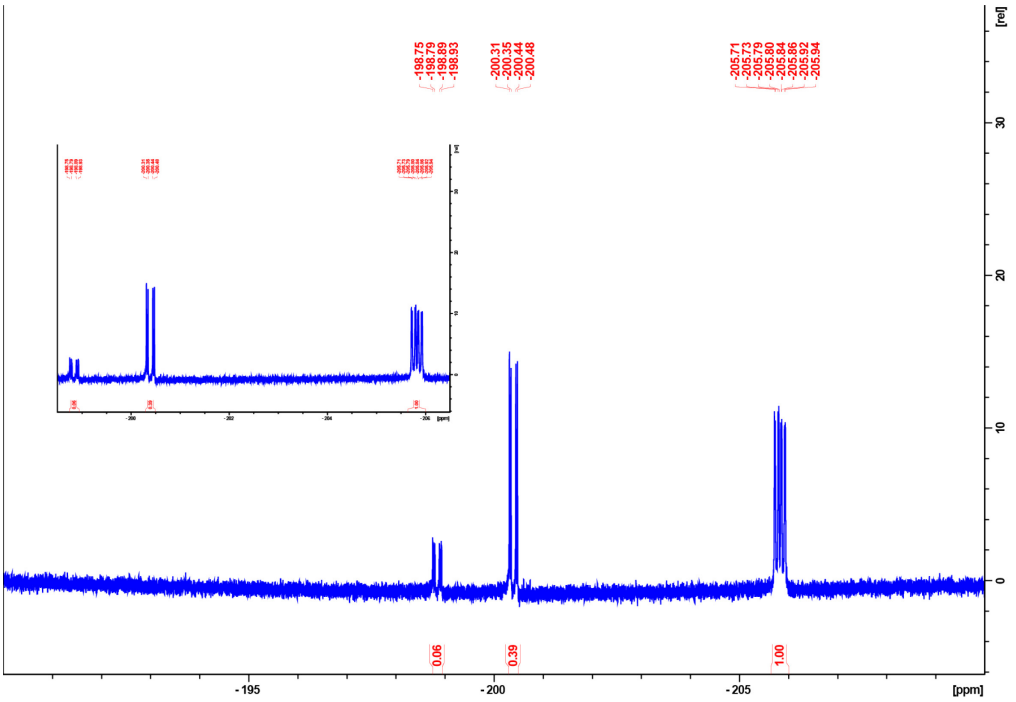


Figure 4.3: ^{19}F NMR spectrum of 11b (1-hydroxyl 2-deoxy-2-fluoro per-*O*-acetylated GG3G)

Due to the difficulty in separating the three diastereomers as free sugars, the mixture was directly subjected to reaction with 2,4 dinitrofluorobenzene in the presence of DABCO. Upon workup, TLC of the crude product mixture showed three individual spots in each case (**Figure 4.4**). Purification using flash chromatography (toluene/acetone, 6.5:1) enabled the isolation of fractions corresponding to spots A2 and B2 as the desired *gluco*-β diastereomers, **12a** and **12b**, respectively, as ascertained by the ^{19}F NMR chemical shift and spin coupling constant values of fluorine with adjacent protons (**Figures 4.5, 4.6**). Spots A1 and B1 correspond to *manno* α diastereomers while spots A3 and B3 correspond to the *gluco* α diastereomers, respectively, as also concluded by ^{19}F NMR analysis (for ^{19}F NMR spectra and assignments of A1 and A3, see **Figures C5** and **C6** respectively). Synthesis of a mixed-linkage tetrasaccharide based GGG3G (2F)-β-DNP was attempted but subsequently abandoned due to an inability to chromatographically separate of a similar mixture of diastereomers of this longer oligosaccharide. Subsequent

deprotection of **12a** and **12b** using Zemplen conditions and purification using flash chromatography yielded the target compounds **13a** and **13b**.

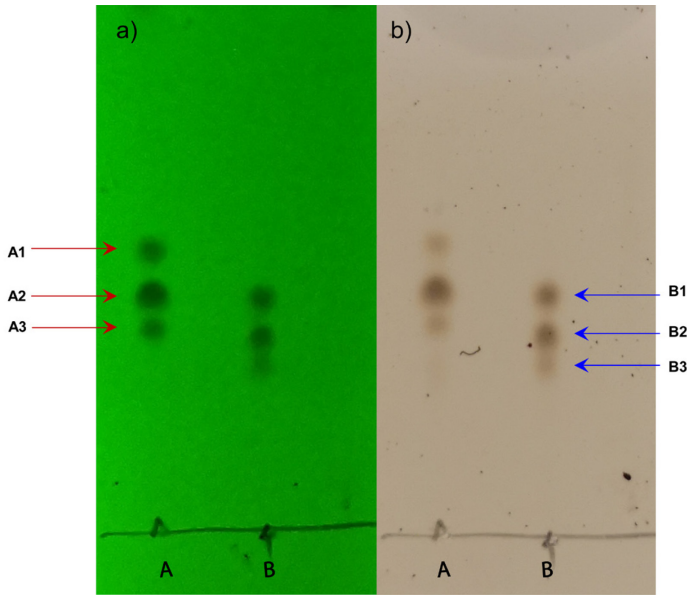


Figure 4.4: Thin-layer chromatography of crude product mixture of 12a and 12b 2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated) glycosides in toluene/acetone (2.5:1) depicting a mixture of diastereomers. First lane (A) is spotted with crude 12a and second lane (B) is spotted with crude 12b. (a) TLC plate visualized under a UV lamp. (b) TLC plate on charring after dipping in 10% sulfuric acid solution. A2 and B2 are the desired β-glycoside diastereomers at R_f= 0.47 (12a) and 0.38 (12b) respectively.

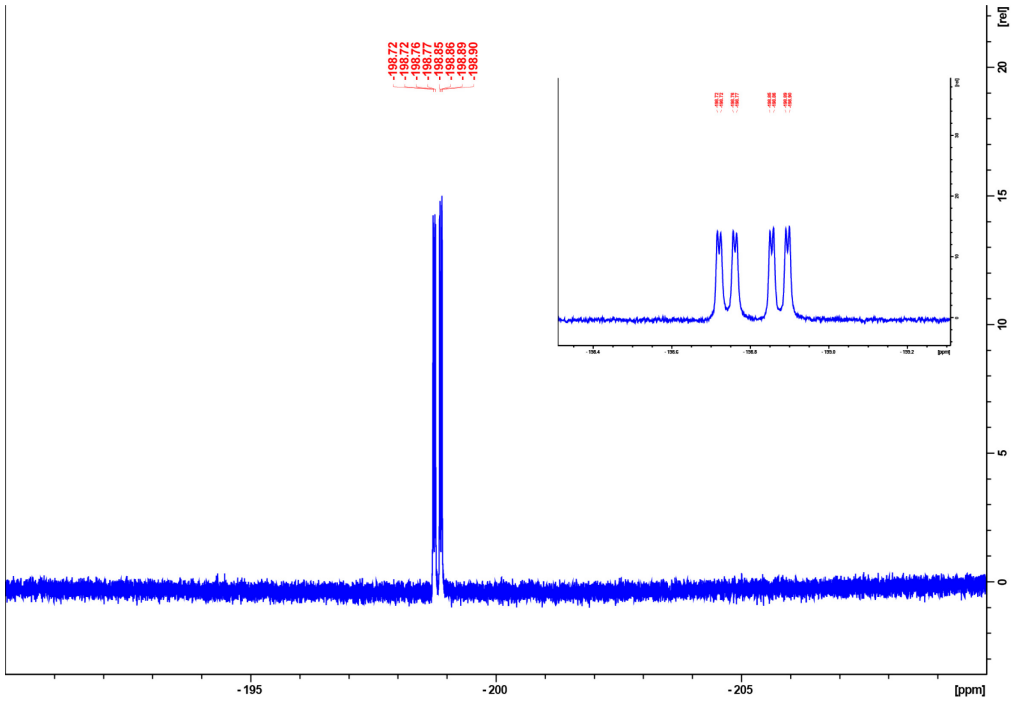


Figure 4.5: ^{19}F NMR spectrum of 12a (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated) G3G)

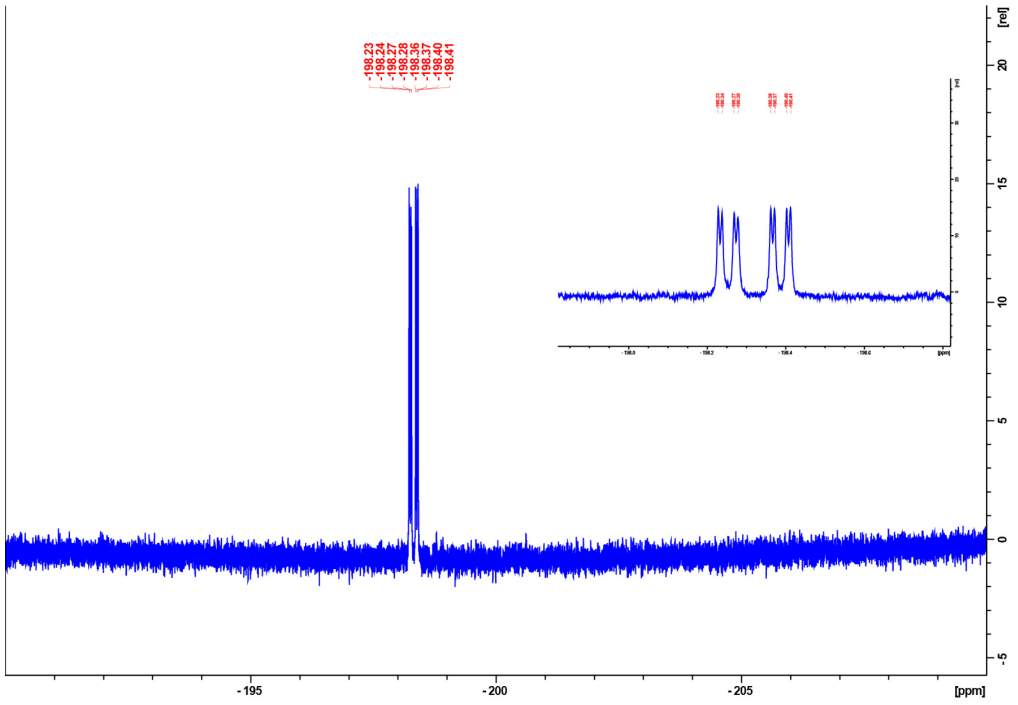


Figure 4.6: ^{19}F NMR spectrum of 12b (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated) GG3G)

4.3.2 Evaluation of Inhibitory Potential Toward *Endo*-β(1,3) Glucanases

Recently, our group elucidated the mechanism of laminarin digestion by a prominent member of the human gut microbiota *Bacteroides uniformis*[123], with a particular focus on the structural and biochemical characterization of the GHs BuGH158 and BuGH16 encoded within a complex Polysaccharide Utilization Locus[342]. BuGH158 was confirmed to be a highly specific *endo*-β-(1,3) glucanase with high activity on laminarin from *Laminaria digitata* (*Ld*) (brown algae) and limited activity on barley MLG (about 140 × lower). BuGH16 was found to be a broad-specificity *endo*-β(1,3)/β(1,4)-glucanase with activity on *Ld* laminarin about 4 times higher than barley MLG[123].

Prior to this work, the stereochemical outcome of catalysis by GH158 members was unknown. Hence, compound **15** was used to reveal the catalytic mechanism of BuGH158 as configuration-retaining by ^{1}H-NMR spectroscopy (**Figure C23**, published in ref.[123]). ^{1}H-NMR spectra at various time intervals were recorded upon the addition of BuGH158 (20 μM final concentration) to **15** (10 mM final concentration) in D_2O. The anomeric configuration of the first-formed laminaribiose product (5 min) was observed to be β, which subsequently underwent mutarotation over 24 h.[123].

As a representative retaining laminarinase/MLGase, BuGH158 was then used to study the potency of **13a** and **13b** as potential inhibitors, using the chromogenic substrate β-CNP laminaribioside (**15**) to measure residual activity (K_m= 12.6 ± 1.15 mM, k_{cat}= 1918 ± 154 min^{-1}, **Figure C24**). At concentrations ranging from 0.31 to 10 mM, both **13a** and **13b** exhibited a time- and concentration-dependent inactivation of BuGH158 over 3 h, with rapid and complete inactivation by **13b** being particularly apparent (**Figures 4.7a, 4.8a**). The values of the apparent inactivation rate constant k_{app} were calculated by fitting **Equation 1.4** to the data. Subsequent

fitting of **Equation 1.5** to the plot of individual k_{app} vs inhibitor concentration (**Figures 4.7b, 4.8b**) allowed the calculation of inhibition constants for **13a** (k_i= 0.087 ± 0.006 min^{-1}, K_i= 3.88 ± 0.624 mM, k_i /K_i= 0.022 min^{-1} mM^{-1}) and **13b** (k_i= 0.374 ± 0.018 min^{-1}, K_i= 3.12 ± 0.373 mM, k_i /K_i = 0.119 min^{-1} mM^{-1}). The k_i values reported here are either in a similar range[99,191,196,205,343,344] or 1-2 orders of magnitude lower[194,195,202,206,345] than similar 2-deoxy-2-fluoro-dinitrophenyl or 2-deoxy-2-fluoro-glycosyl fluoride inhibitors. The K_i values observed here are in a comparable millimolar range[99,191,194,205,206,344,345].

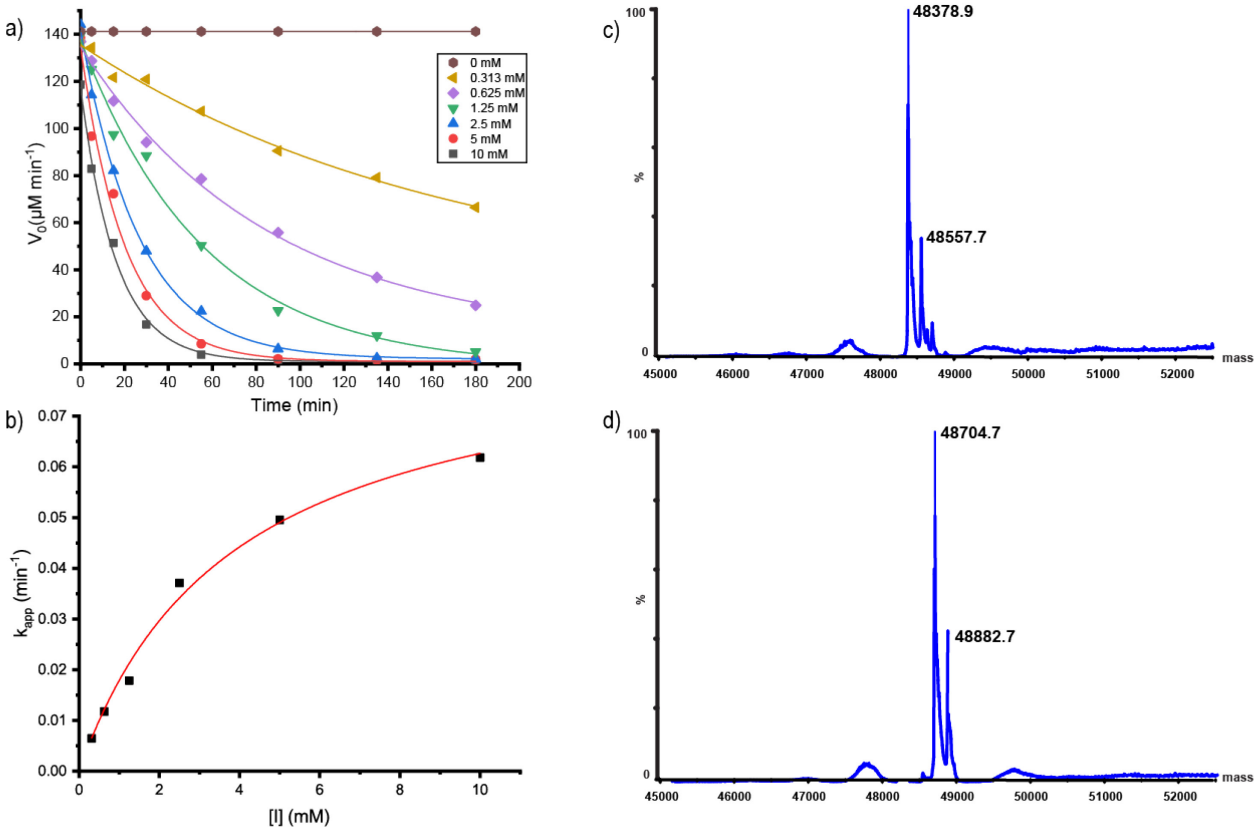


Figure 4.7: Inhibition kinetics and intact-protein MS of BuGH158 by inhibitor 13a. (a) Plot of initial-rate enzyme activity versus time (single determinations) at inhibitor concentrations indicated. (b) Plot of pseudo-first-order rate constants (k_{app}), obtained from the fitted curves shown in panel a, versus inhibitor concentration. (c) wild type BuGH158. Expected mass: 48377.5 Da; a significant phosphogluconoylation peak is observed at +178 Da. (d) BuGH158 incubated with 2.5 mM 13a (510.1 Da) at 50 °C for 2 h. Expected mass: 48704.6 Da; the phosphogluconoylated proteins are equally labelled.

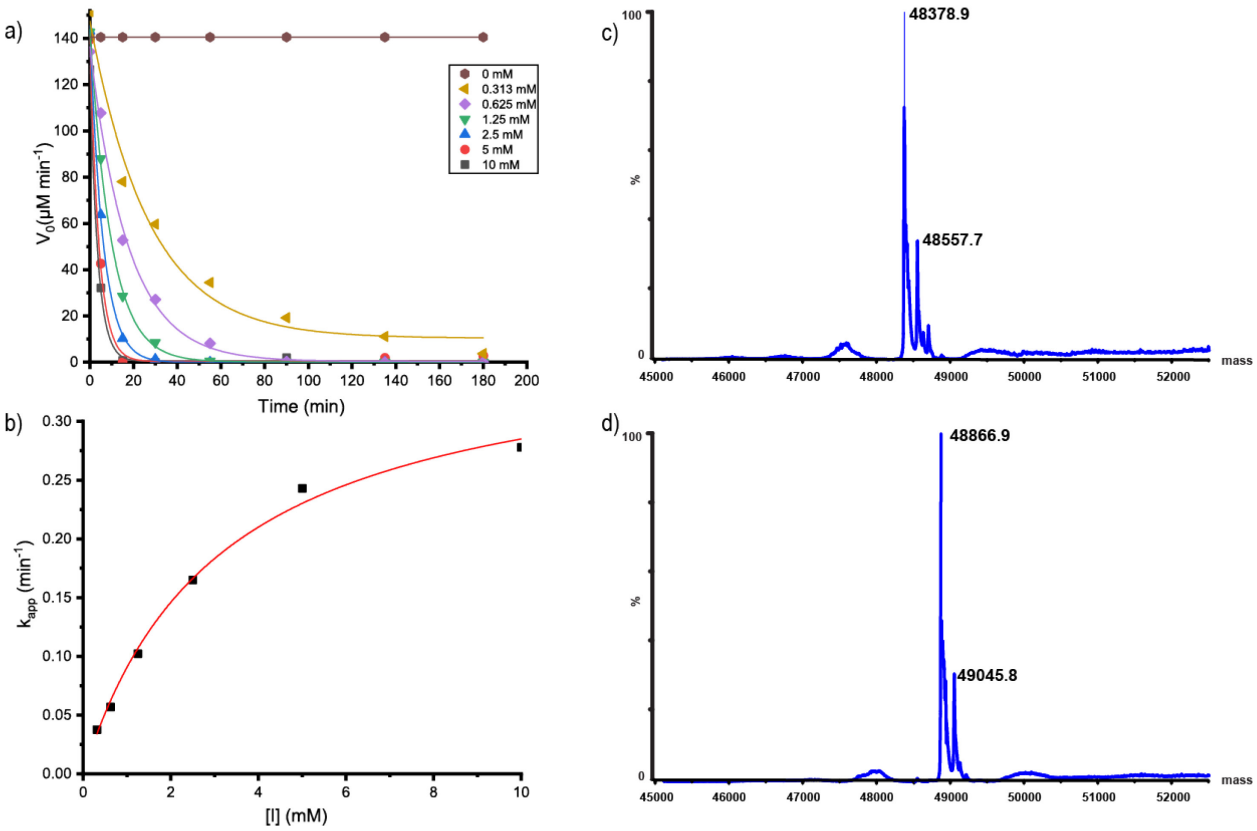


Figure 4.8: Inhibition kinetics and intact-protein MS of BuGH158 by inhibitor 13b. (a) Plot of initial-rate enzyme activity versus time (single determinations) at inhibitor concentrations indicated. (b) Plot of pseudo-first-order rate constants (k_{app}), obtained from the fitted curves shown in panel a, versus inhibitor concentration. (c) wild type BuGH158. Expected mass: 48377.5 Da; a significant phosphogluconoylation peak is observed at +178 Da. (d) BuGH158 incubated with 2.5 mM 13b (672.2 Da) at 50 °C for 2 h. Expected mass: 48866.7 Da; the phosphogluconoylated proteins are equally labelled.

Intact-protein mass spectrometry of the enzyme upon incubation with 2.5 mM of **13a** or **13b** revealed a stoichiometric 1:1 labelling of the enzyme after 2 h and the covalent nature of inhibition (**Figures 4.7c, 4.7d and 4.8c, 4.8d)**. An increase in protein molecular weight of 325.8 Da and 488.0 Da, respectively, were observed for **13a** and **13b** upon incubation with the enzyme. These results are consistent with the addition of a single 2-deoxy-2-fluoro G3G moiety and a single 2-deoxy-2-fluoro GG3G moiety, respectively. No trace of the free enzyme was observed for either inhibitor, indicative of relatively stable glycosyl-enzyme intermediates with limited

turnover[191,195,200,202]. For comparison, the retaining MLGase BuGH16[123] was labelled by **13a (Figure C25a)** and **13b (Figure C25b)**. However, labelling with **13a** was incomplete, suggesting either slow inhibition or slow turnover. Complete labelling of the MLGase BoGH16[44] was observed using **13b (Figure C25c)**.

4.4 Conclusion

In summary, two mechanism-based oligosaccharide inhibitors of retaining GHs were chemoenzymatically synthesized in moderate overall yields, starting from readily available per-*O*-acetylated oligosaccharides. To the best of our knowledge, this is the first report of the synthesis and isolation of β-2'4'dinitrophenyl-2-deoxy-2-fluoro glucosides with β-(1,3) glycosidic linkages. Through covalent labelling and inhibition kinetics, the utility of these inhibitors was demonstrated for the biochemical characterization of β-(1,3) glucanases and MLGases. This work contributes to the developing field of small-molecule covalent probes for GH characterization and aims to enhance our understanding of the functional and mechanistic aspects of GH catalysis[147,148,346].

Chapter 5: General Conclusions and Future Directions

5.1 General Conclusions

The CAZy database is a dynamic collection of carbohydrate-processing enzymes with an exponentially increasing number of entries[78,127]. GHs acting on myriad substrates form the largest class of CAZy database and see the bulk of this input, with extensive research conducted to delineate various aspects of their utility. The industrial applications of GHs in diverse biotechnological fields motivate the profiling of GHs to obtain information about their homology, atomic structure details, mechanism of catalytic action, substrate specificity, and product configuration. Careful examination of the active site of GHs elicits this information and facilitates GH application.

Complex carbohydrate substrates of GHs found in the plant, algal and bacterial cell walls are highly useful biomolecules and key components of dietary fibre. Such carbohydrates serve important roles in influencing bacterial pathogenicity, provide renewable feedstocks for the biofuel industry, and are used as additive materials for the food industry. Highly potent covalent inhibitors of GHs, along with crystallographic analysis of in-complex enzyme-structures, have a precedent of dissecting transition-state structures and elucidating minute mechanistic details. The work described in this thesis focuses on utilizing the affinity of small-molecule covalent inhibitors towards GH active site to develop strategies for the exploration and analysis of the latter.

In chapter 2, a highly branched mechanism-based inhibitor XXXG (2F)-β-DNP targeting *endo*-xyloglucanases was developed using a multistep synthetic strategy starting from a natural xyloglucan source tamarind kernel powder. The synthetic methodology yielded one of the most complex oligosaccharide-based 2-deoxy-2-fluoro inhibitors to date. An inhibitor-bound protein crystal structure of a representative GH5 *endo*-xyloglucanase enzyme, along with inhibition

kinetics and protein mass spectrometry data confirmed the inhibitory potential of XXXG (2F)-β-DNP, which turned out to be a "slow-substrate" of the CjGH5D enzyme with relatively high turnover rate.

In chapter 3, the synthetic strategy from a previously developed active-site affinity-based inhibitor of *endo*-xyloglucanases (XXXG-NHCOCH$_2$Br[43,170]) was adopted for the incorporation of an electrophilic warhead on a mixed-linkage oligosaccharidic backbone to target mixed-linkage β-(1,3)/(1,4)-glucanases. Two (GG3G-NHCOCH$_2$Br and GGG3G-NHCOCH$_2$Br) of the three inhibitors proved to be highly potent inactivators of a vanguard MLGase, BoGH16, with the lack of inhibition by the third (G3GGG-NHCOCH$_2$Br) reflecting the linkage preference of the enzyme. The applicability of this group of inhibitors extends beyond BoGH16, with the variable linkage providing versatility across MLGases with alternative substrate specificities[310].

Chapter 4 describes a combination of the chemoenzymatic synthetic strategies from Chapters 2 and 3, with the use of β-(1,3) and β-(1,3)/(1,4) linked disaccharide and trisaccharide to synthesize a dinitrophenyl 2-deoxy-2-fluoroglycoside mechanism-based inhibitor targeting β-(1,3) glucanases and MLGases. These inhibitors showed potency towards the laminarinase BuGH158, with intact-protein mass spectrometric experiments suggesting successful inhibition of MLGases (BuGH16 and BoGH16) by the 2F-DNP inhibitors. A similar synthetic protocol to synthesize the corresponding tetrasaccharide GGG3G (2F)-β-DNP inhibitor was met with failure due to difficulty in chromatographic separation of the diastereomeric mixture obtained after the installation of the 2,4 dinitrophenyl group.

Collectively, the work described in this thesis further enhances the knowledge of GH active site make-up by identification of key active site residues via covalent labelling. Inexpensive and abundantly available natural sources of the polysaccharide starting materials facilitate large-scale

production of novel, highly versatile inhibitors with proven potency towards vanguard GHs. With the aid of such rationally designed chemical tools, the roles of GHs and glycans can be better understood, and new approaches towards examining GH functions and structures can be developed. Additionally, the synthetic methodologies described in this thesis could be extended for the discovery of novel covalent inhibitors targeting GHs of diverse substrate specificities.

5.2 Future Directions

Taking advantage of the affinity and specificity of covalent inhibitors towards GHs of interest, novel chemical probes derived by further chemical derivatization of these inhibitors can be designed[347]. This strategy has found direct application in the field of activity-based protein profiling (ABPP)[184,348,349]. ABPP is a proteomics technique that facilitates the study of protein function *in vivo* and/or *in vitro* using chemical probes (activity-based probes or ABPs) targeting certain proteins or protein classes. The common features of such probes are a reactive group (or warhead), a recognition element with affinity towards the target protein(s), and a reporter entity (or a ligation handle for facile attachment of the reporter group)[350] (**Figure 5.1**). The reporter group aids the imaging, purification, or enrichment of the labelled proteins.

The ABPP technique relies on the development of enzyme labels that bind to the target enzyme and facilitate the quantifiable detection of the enzyme with the help of reporter groups such as fluorophores, chromophores, radioactive labels, and biotin[350–354]. This technique can also be used to discover new enzymes with unknown functions[148,355,356] as well as for targeted imaging of subcellular function and localization[357,358].

A well-known method to attach a reporter group towards the non-reducing end of inhibitors is via ligation handles such as azido, amino, or alkyne groups. Such handles can be installed using

chemically functionalized donor sugars with the help of glycosynthases of GHs[237,305,359–361], resulting in the development of various ABPs.

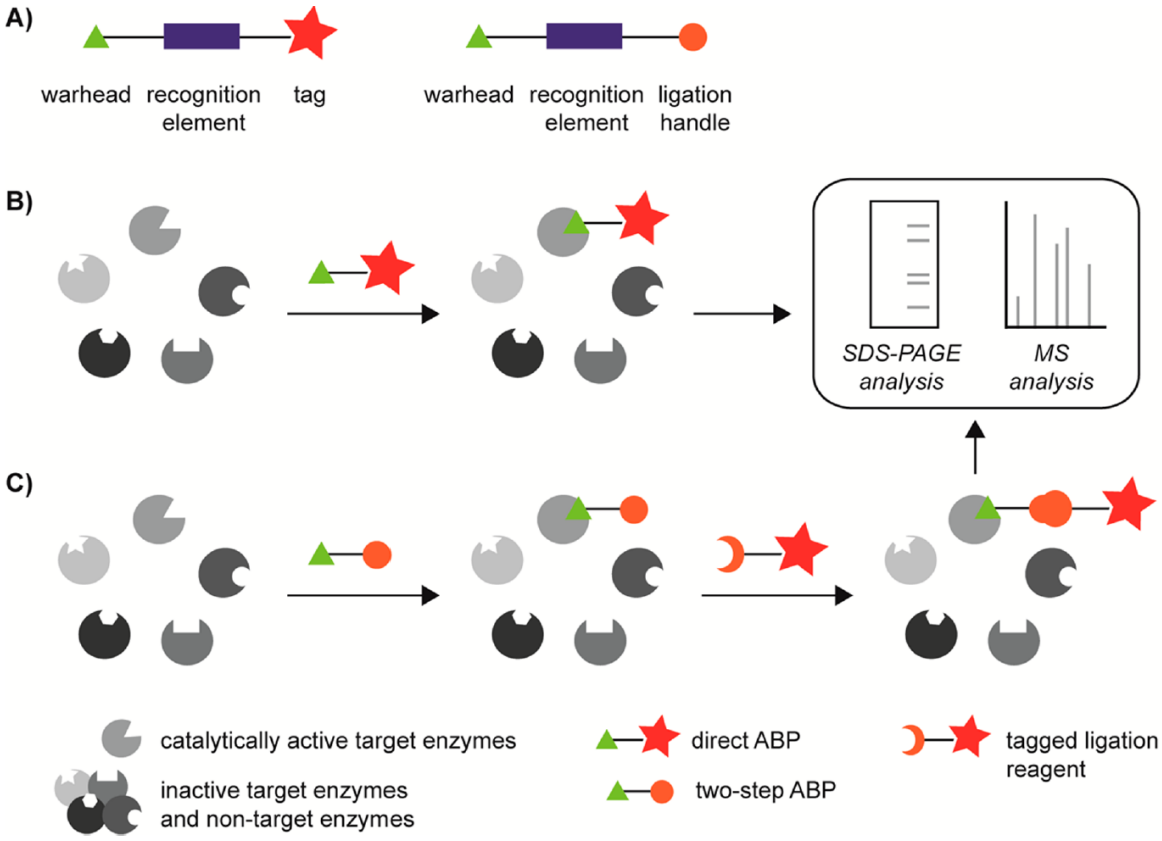
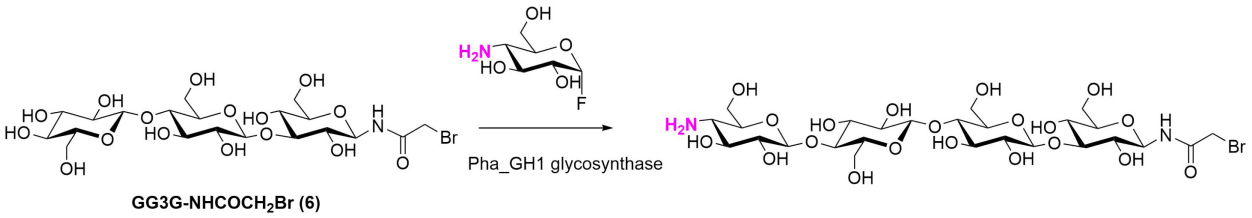


Figure 5.1: Schematic representations of one- and two-step ABPP strategies. (A) General composition of a direct ABP (left) and a two-step ABP (right). (B, C) Labeling of enzymatic activity by a direct ABP (B) or a two-step ABP (C). (Reused with permission from Willems, Overkleeft, and van Kasteren[350] (Copyright (2014) American Chemical Society))

There have been multiple reports of ABPs based on GH inhibitors[173,239,240,362–366]. For example, the mechanism-based inhibitor cyclophellitol[367,368] was used to develop a fluorescently-labelled probe capable of labelling and detecting human enzyme glucocerebrosidase in cell extracts and cultured cells[369,370]. Similarly, 2-deoxy-2-fluoroglycosides have been used as ABPP probes for retaining glycosidases[198,371]. The covalent inhibitors described in Chapters 2, 3, and 4

could be functionalized by the addition of azide or amino group-bearing α-fluoro glucose donors using straightforward glycosynthase chemistry.

To investigate the scope of ABPs based on the inhibitors described in this thesis, preliminary experiments using a GH1 glycosynthase from *Phanerochaete chrysosporium* (Pha_GH1)[372,373] (using a previously reported protocol[373]) were conducted. This glycosynthase was generously provided by Dr. Zachary Armstrong from Withers lab, UBC. The glycosynthase Pha_GH1 was used to attach 4-amino glucose at the non-reducing end of GG3G $NHCOCH_2Br$ (**Scheme 5.1**) using 4-amino-α-fluoro glucose as the donor. The reaction, as monitored by MALDI-TOF spectrometry, indicated full conversion. Further confirmation of the identity of the product would be needed by NMR spectroscopy experiments. Incomplete conversions were also recorded by MALDI-TOF spectrometry for the addition of the same donor to GG3G-β-CNP and GGG3G $NHCOCH_2Br$, suggesting promising results in this field and setting the stage for future investigation into complex molecular probes for activity-based protein profiling of GHs.



Scheme 5.1: Glycosynthase-mediated addition of 4-amino glucose to GG3G-$NHCOCH_2Br$

References

1. Rauvala, H. & Finne, J. Structural similarity of the terminal carbohydrate sequences of glycoproteins and glycolipids. *FEBS Lett.* **97,** 1–8 (1979).
2. Smets, L. A., Enninga, I. C. & Rooy, H. Van. Cell communication reduced by changes in cell surface carbohydrates. *Exp. Cell Res.* **139,** 181–189 (1982).
3. Bidlack, J., Malone, M. & Benson, R. Molecular structure and component integration of secondary cell walls in plants. *Proc. Oklahoma Acad. Sci.* **72,** 51–56 (1992).
4. Jéquier, E. Carbohydrates as a source of energy. *Am. J. Clin. Nutr.* **59,** 682S-685S (1994).
5. Muramatsu, T. Essential roles of carbohydrate signals in development, immune response and tissue functions, as revealed by gene targeting. *J. Biochem.* **127,** 171–176 (2000).
6. Cosgrove, D. J. Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.* **6,** 850–861 (2005).
7. Laine, R. A. A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05 × 1012 structures for a reducing hexasaccharide: The Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* **4,** 759–767 (1994).
8. Lapébie, P., Lombard, V., Drula, E., Terrapon, N. & Henrissat, B. Bacteroidetes use thousands of enzyme combinations to break down glycans. *Nat. Commun.* **10,** (2019).
9. Carpita, N. C. & Mccann, M. C. in *Biochem. Mol. Biol. Plants* (eds. Buchanan, B. B., Gruissem, W. & Jones, R. L.) 55–108 (John Wiley & Sons, Inc., 2000).
10. Sticklen, M. B. Plant genetic engineering for biofuel production: Towards affordable cellulosic ethanol. *Nat. Rev. Genet.* **9,** 433–443 (2008).
11. Lynd, L. R., Paul J. Weimer, Zyl, W. H. van & Pretorius, I. S. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* **66,** 506–77 (2004).
12. O'Sullivan, A. C. Cellulose: The structure slowly unravels. *Cellulose* **4,** 173–207 (1997).
13. Carvalho, K. C. C. de, Montoro, S. R., Cioffi, M. O. H. & Voorwald, H. J. C. in *Des. Appl. Nanostructured Polym. Blends Nanocomposite Syst.* (eds. Thomas, S., Shanks, R. & Chandrasekharakurup, S.) 261–285 (Elsevier Inc., 2016).
14. Scheller, H. V. & Ulvskov, P. Hemicelluloses. *Annu. Rev. Plant Biol.* **61,** 263–289 (2010).
15. Reid, J. S. G. in *Adv. Bot. Res.* (eds. Callow, J. A. & Woolhouse, H. W.) **11,** 125–155 (Academic Press, 1985).

16. Sandhu, A. P. S., Randhawa, G. S. & Dhugga, K. S. Plant cell wall matrix polysaccharide biosynthesis. *Mol. Plant* **2,** 840–850 (2009).
17. Carpita, N. C. & Gibeaut, D. M. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* **3,** 1–30 (1993).
18. McDougall, G. J., Morrison, I. M., Stewart, D. & Hillman, J. R. Plant Cell Walls as Dietary Fibre: Range, Structure, Processing and Function. *J. Sci. Food Agric.* **70,** 133–150 (1996).
19. Anderson, C. T., Carroll, A., Akhmetova, L. & Somerville, C. Real-time imaging of cellulose reorientation during cell wall expansion in Arabidopsis roots. *Plant Physiol.* **152,** 787–796 (2010).
20. Hayashi, T. Xyloglucans in the Primary Cell Wall. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40,** 139–168 (1989).
21. Pauly, M., Albersheim, P., Darvill, A. & York, W. S. Molecular domains of the cellulose/xyloglucan network in the cell walls of higher plants. *Plant J.* **20,** 629–639 (1999).
22. de Souza, A. P., Leite, D. C. C., Pattathil, S., Hahn, M. G. & Buckeridge, M. S. Composition and Structure of Sugarcane Cell Wall Polysaccharides: Implications for Second-Generation Bioethanol Production. *Bioenergy Res.* **6,** 564–579 (2013).
23. Christiemin, M., Henriksson, G., Lindström, M. E., Brumer, H., Teeri, T. T., Lindström, T. & Laine, J. The effects of xyloglucan on the properties of paper made from bleached kraft pulp. *Nord. Pulp Pap. Res. J.* **18,** 182–187 (2003).
24. Hoffman, M., Jia, Z., Peña, M. J., Cash, M., Harper, A., Blackburn, A. R., Darvill, A. & York, W. S. Structural analysis of xyloglucans in the primary cell walls of plants in the subclass Asteridae. *Carbohydr. Res.* **340,** 1826–1840 (2005).
25. Pauly, M. & Keegstra, K. Biosynthesis of the Plant Cell Wall Matrix Polysaccharide Xyloglucan. *Annu. Rev. Plant Biol.* **67,** 235–259 (2016).
26. Tuomivaara, S. T., Yaoi, K., O'Neill, M. A. & York, W. S. Generation and structural validation of a library of diverse xyloglucan-derived oligosaccharides, including an update on xyloglucan nomenclature. *Carbohydr. Res.* **402,** 56–66 (2015).
27. Vincken, J.-P., York, W. S., Beldman, G. & Voragen, A. G. J. Two general branching

patterns of xyloglucan, XXXG and XXGG. *Plant Physiol.* **114,** 9–13 (1997).

28. Fry, S. C. Polysaccharide-Modifying Enzymes in the Plant Cell Wall. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **46,** 497–520 (2003).
29. Rashmi, R. & Siddalingamurthy, K. R. Microbial xyloglucanases : a comprehensive review. *Biocatal. Biotransformation* **36,** 280–295 (2018).
30. Edwards, M., Dea, I. C. M., Bulpin, P. V. & Reid, J. S. G. Purification and properties of a novel xyloglucan-specific endo-(1→4)-β-D-glucanase from germinated nasturtium seeds (Tropaeolum majus L.). *J. Biol. Chem.* **261,** 9489–9494 (1986).
31. Eklöf, J. M. & Brumer, H. The XTH Gene Family: An Update on Enzyme Structure, Function, and Phylogeny in Xyloglucan Remodeling. *Plant Physiol.* **153,** 456–466 (2010).
32. Larsbrink, J., Rogers, T. E., Hemsworth, G. R., McKee, L. S., Tauzin, A. S., Spadiut, O., Klinter, S., Pudlo, N. A., Urs, K., Koropatkin, N. M., Creagh, A. L., Haynes, C. A., Kelly, A. G., Cederholm, S. N., Davies, G. J., Martens, E. C. & Brumer, H. A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes. *Nature* **506,** 498–502 (2014).
33. Gloster, T. M., Ibatullin, F. M., Macauley, K., Eklöf, J. M., Roberts, S., Turkenburg, J. P., Bjørnvad, M. E., Jørgensen, P. L., Danielsen, S., Johansen, K. S., Borchert, T. V., Wilson, K. S., Brumer, H. & Davies, G. J. Characterization and three-dimensional structures of two distinct bacterial xyloglucanases from families GH5 and GH12. *J. Biol. Chem.* **282,** 19177–19189 (2007).
34. Ariza, A., Eklöf, J. M., Spadiut, O., Offen, W. A., Roberts, S. M., Besenmatter, W., Friis, E. P., Skjøt, M., Wilson, K. S., Brumer, H. & Davies, G. Structure and activity of paenibacillus polymyxa xyloglucanase from glycoside hydrolase family 44. *J. Biol. Chem.* **286,** 33890–33900 (2011).
35. McCleary, B. V & Codd, R. Measurement of (1 → 3),(1 → 4)-β-D-glucan in barley and oats: A streamlined enzymic procedure. *J. Sci. Food Agric.* **55,** 303–312 (1991).
36. Fry, S. C., Nesselrode, B. H. W. A., Miller, J. G. & Mewburn, B. R. Mixed-linkage (1→3,1→4)-β-D-glucan is a major hemicellulose of Equisetum (horsetail) cell walls. *New Phytol.* **179,** 104–115 (2008).
37. Vega-Sánchez, M. E., Verhertbruggen, Y., Scheller, H. V. & Ronald, P. C. Abundance of mixed linkage glucan in mature tissues and secondary cell walls of grasses. *Plant Signal.*

Behav. **8,** 8–11 (2013).

38. Bohn, J. A. & BeMiller, J. N. (1→3)-β-d-Glucans as biological response modifiers: a review of structure-functional activity relationships. *Carbohydr. Polym.* **28,** 3–14 (1995).
39. Chen, J. & Seviour, R. Medicinal importance of fungal β-(1→3), (1→6)-glucans. *Mycol. Res.* **111,** 635–652 (2007).
40. Williams, D. L. Overview of (1-3)-b-D-glucan immunobiology. *Mediators Inflamm.* **6,** 247–250 (1997).
41. Keogh, G. F., Cooper, G. J. S., Mulvey, T. B., McArdle, B. H., Coles, G. D., Monro, J. A. & Poppitt, S. D. Randomized controlled crossover study of the effect of a highly β-glucan-enriched barley on cardiovascular disease risk factors in mildly hypercholesterolemic men. *Am. J. Clin. Nutr.* **78,** 711–718 (2003).
42. Johansson, L., Virkki, L., Maunu, S., Lehto, M., Ekholm, P. & Varo, P. Structural characterization of water soluble β-glucan of oat bran. *Carbohydr. Polym.* **42,** 143–148 (2000).
43. McGregor, N., Morar, M., Fenger, T. H., Stogios, P., Lenfant, N., Yin, V., Xu, X., Evdokimova, E., Cui, H., Henrissat, B., Savchenko, A. & Brumer, H. Structure-function analysis of a mixed-linkage β-glucanase/xyloglucanase from the key ruminal bacteroidetes prevotella bryantii B14. *J. Biol. Chem.* **291,** 1175–1197 (2016).
44. Tamura, K., Hemsworth, G. R., Dejean, G., Rogers, T. E., Pudlo, N. A., Urs, K., Jain, N., Davies, G. J., Martens, E. C. & Brumer, H. Molecular Mechanism by which Prominent Human Gut Bacteroidetes Utilize Mixed-Linkage Beta- Glucans, Major Health-Promoting Cereal Polysaccharides. *Cell Rep.* **21,** 417–430 (2017).
45. Desmet, T., Cantaert, T., Gualfetti, P., Nerinckx, W., Gross, L., Mitchinson, C. & Piens, K. An investigation of the substrate specificity of the xyloglucanase Cel74A from Hypocrea jecorina. *FEBS J.* **274,** 356–363 (2007).
46. Stone, B. A. & Clarke, A. E. *Chemistry and Biology of 1, 3-β-Glucans*. (Intl Specialized Book Service Inc, 1992).
47. Fincher, G. B. Exploring the evolution of (1,3;1,4)-β-D-glucans in plant cell walls: comparative genomics can help! *Curr. Opin. Plant Biol.* **12,** 140–147 (2009).
48. Hahn, M., Pons, J., Planas, A., Querol, E. & Heinemann, U. Crystal structure of Bacillus licheniformis 1,3-1,4-β-D-glucan 4-glucanohydrolase at 1.8 A resolution. *FEBS Lett.* **374,**

221–224 (1995).

49. Planas, A., Juncosa, M., Cayetano, A. & Querol, E. Studies on Bacillus licheniformis endo-β-l,3-1,4-D-glucanase: characterization and kinetic analysis. *Appl. Microbiol. Biotechnol.* **37,** 583–589 (1992).
50. Wade Abbott, D. & Boraston, A. B. in *Chem. Biochem. Biol. 1-3 Beta Glucans Relat. Polysaccharides* (eds. Bacic, A., Fincher, G. B. & Stone, B. A.) **16,** 171–199 (Elsevier Inc., 2009).
51. Stone, B. A. *Chemistry of β -Glucans*. *Chem. Biochem. Biol. (1→3)-β-Glucans Relat. Polysaccharides* (Elsevier Inc., 2009).
52. Nishikawa, S. I., Zinkl, G. M., Swanson, R. J., Maruyama, D. & Preuss, D. Callose (β-1,3 glucan) is essential for Arabidopsis pollen wall patterning, but not tube growth. *BMC Plant Biol.* **5,** 1–9 (2005).
53. Chen, X. Y. & Kim, J. Y. Callose synthesis in higher plants. *Plant Signal. Behav.* **4,** 489–492 (2009).
54. Li, X., He, Y., Zeng, P., Liu, Y., Zhang, M., Hao, C., Wang, H., Lv, Z. & Zhang, L. Molecular basis for Poria cocos mushroom polysaccharide used as an antitumour drug in China. *J. Cell. Mol. Med.* **23,** 4–20 (2019).
55. Hoffmann, G. C., Simon, B. W. & Timell, T. E. Structure and molecular size of pachyman. *Carbohydr. Res.* **20,** 185–188 (1971).
56. Wang, Y., Zhang, M., Ruan, D., Shashkov, A. S., Kilcoyne, M., Savage, A. V. & Zhang, L. Chemical components and molecular mass of six polysaccharides isolated from the sclerotium of *Poria cocos*. *Carbohydr. Res.* **339,** 327–334 (2004).
57. McIntosh, M., Stone, B. A. & Stanisich, V. A. Curdlan and other bacterial (1→3)-β-D-glucans. *Appl. Microbiol. Biotechnol.* **68,** 163–173 (2005).
58. Harada, T., Misaki, A. & Saito, H. Curdlan: A Bacterial Gel-Forming β-1,3-Glucan. *Arch. Biochem. Biophys.* **124,** 292–298 (1968).
59. McArdle, R. & Hamill, R. in *Process. Meats Improv. Safety, Nutr. Qual.* (eds. Kerry, J. P. & Kerry, J. F.) 243–269 (Woodhead Publishing Limited, 2011).
60. Saito, H., Misaki, A. & Harada, T. A Comparison of the Structure of Curdlan and Pachyman. *Agr. Biol. Chem.* **32,** 1261–1269 (1968).
61. Lewin, R. A. *Physiology and biochemistry of algae*. (Academic Press, 1962).

62. Percival, E. & Mcdowell, R. H. in *Methods Plant Biochem.* (ed. Dey, P. M.) **2,** 523–547 (Academic Press, 1990).
63. Dunstan, D. E. & Goodall, D. G. Terraced self assembled nano-structures from laminarin. *Int. J. Biol. Macromol.* **40,** 362–366 (2007).
64. Read, S. M., Currie, G. & Bacic, A. Analysis of the structural heterogeneity of laminarin by electrospray-ionisation-mass spectrometry. *Carbohydr. Res.* **281,** 187–201 (1996).
65. Chizhov, A. O., Dell, A., Morris, H. R., Reason, A. J., Haslam, S. M., McDowell, R. A., Chizhov, O. S. & Usov, A. I. Structural analysis of laminarans by MALDI and FAB mass spectrometry. *Carbohydr. Res.* **310,** 203–210 (1998).
66. Kim, K. H., Kim, Y. W., Kim, H. B., Lee, B. J. & Lee, D. S. Anti-apoptotic activity of laminarin polysaccharides and their enzymatically hydrolyzed oligosaccharides from *Laminaria japonica*. *Biotechnol. Lett.* **28,** 439–446 (2006).
67. Neyrinck, A. M., Mouson, A. & Delzenne, N. M. Dietary supplementation with laminarin, a fermentable marine β (1–3) glucan, protects against hepatotoxicity induced by LPS in rat by modulating immune response in the hepatic tissue. *Int. Immunopharmacol.* **7,** 1497–1506 (2007).
68. Wargacki, A. J., Leonard, E., Win, M. N., Regitsky, D. D., Santos, C. N. S., Kim, P. B., Cooper, S. R., Raisner, R. M., Herman, A., Sivitz, A. B., Lakshmanaswamy, A. & Kashiyama, Y. An Engineered Microbial Platform for Direct Biofuel Production from Brown Macroalgae. *Science* **335,** 308–314 (2012).
69. Al Abdallah, Q., Nixon, B. T. & Fortwendel, J. R. The enzymatic conversion of major algal and cyanobacterial carbohydrates to bioethanol. *Front. Energy Res.* **4,** 1–15 (2016).
70. Mrsa, V., Klebl, F. & Tanner, W. Purification and characterization of the *Saccharomyces cerevisiae BGL2* gene product, a cell wall endo-β-1,3-glucanase. *J. Bacteriol.* **175,** 2102–2106 (1993).
71. Shi, P., Yao, G., Yang, P., Li, N., Luo, H., Bai, Y., Wang, Y. & Yao, B. Cloning, characterization, and antifungal activity of an endo-1,3-β-d-glucanase from *Streptomyces* sp. S27. *Appl. Microbiol. Biotechnol.* **85,** 1483–1490 (2010).
72. Balasubramanian, V., Vashisht, D., Cletus, J. & Sakthivel, N. Plant β -1,3-glucanases : Their biological functions and transgenic expression against phytopathogenic fungi. *Biotechnol Lett* **34,** 1983–1990 (2012).

73. Ryan, E. M. & Ward, O. P. Study of the effect of β-1,3-glucanase from Basidiomycete QM 806 on yeast extract production. *Biotechnol. Lett.* **7,** 409–412 (1985).
74. Branco, A. T., Ferreira, B. dos S., Lourenco, G. F., Marques, V. C. L., Machado, O. L. T., Pereira, M. G. & Filho, J. C. A. A. and G. A. de S. Induction of β-1,3-Glucanase in Seeds of Maize Defective-Kernel Mutant (827Kpro1). *Protein Pept. Lett.* **18,** 651–657 (2011).
75. Wolfenden, R., Lu, X. & Young, G. Spontaneous hydrolysis of glycosides. *J. Am. Chem. Soc.* **120,** 6814–6815 (1998).
76. Zechel, D. L. & Withers, S. G. Glycosidase mechanisms: Anatomy of a finely tuned catalyst. *Acc. Chem. Res.* **33,** 11–18 (2000).
77. *Enzyme Nomenclature Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes*. (Academic Press, 1992). at <http://www.sciencedirect.com/science/article/pii/B9780122271649500017>
78. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42,** D490–D495 (2014).
79. Henrissat, B. & Davies, G. Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* **7,** 637–644 (1997).
80. Coutinho, P. M., Deleury, E., Davies, G. J. & Henrissat, B. An evolving hierarchical family classification for glycosyltransferases. *J. Mol. Biol.* **328,** 307–317 (2003).
81. Sutherland, I. W. Polysaccharide lyases. *FEMS Microbiol. Rev.* **16,** 323–347 (1995).
82. Lombard, V., Bernard, T., Rancurel, C., Brumer, H., Coutinho, P. M. & Henrissat, B. A hierarchical classification of polysaccharide lyases for glycogenomics. *Biochem J* **432,** 437–444 (2010).
83. Cantarel, B. I. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res.* **37,** 233–238 (2009).
84. Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M. & Henrissat, B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol. Biofuels* **6,** 41 (2013).
85. Boraston, A. B., Bolam, D. N., Gilbert, H. J. & Davies, G. J. Carbohydrate-binding

modules: fine-tuning polysaccharide recognition. *Biochem. J.* **382,** 769–781 (2004).

86. Viborg, A. H., Terrapon, N., Lombard, V., Michel, G., Czjzek, M., Henrissat, B. & Brumer, H. A subfamily roadmap of the evolutionarily diverse glycoside hydrolase family 16 (GH16). *J. Biol. Chem.* **294,** 15973–15986 (2019).
87. St John, F. J., González, J. M. & Pozharski, E. Consolidation of glycosyl hydrolase family 30: A dual domain 4/7 hydrolase family consisting of two structurally distinct groups. *FEBS Lett.* **584,** 4435–4441 (2010).
88. Aspeborg, H., Coutinho, P. M., Wang, Y., Brumer, H. & Henrissat, B. Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). *BMC Evol. Biol.* **12,** (2012).
89. Mewis, K., Lenfant, N., Lombard, V. & Henrissat, B. Dividing the large glycoside hydrolase family 43 into subfamilies: A motivation for detailed enzyme characterization. *Appl. Environ. Microbiol.* **82,** 1686–1692 (2016).
90. Davies, G. J. & Henrissat, B. Structures and mechanisms of glycosyl hydrolases. *Structure* **3,** 853–859 (1995).
91. Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J. P. & Davies, G. Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc. Natl. Acad. Sci. U. S. A.* **92,** 7090–7094 (1995).
92. Davies, G. J., Gloster, T. M. & Henrissat, B. Recent structural insights into the expanding world of carbohydrate-active enzymes. *Curr. Opin. Struct. Biol.* **15,** 637–645 (2005).
93. Beckham, G. T., Ståhlberg, J., Knott, B. C., Himmel, M. E., Crowley, M. F., Sandgren, M., Sørlie, M. & Payne, C. M. Towards a molecular-level theory of carbohydrate processivity in glycoside hydrolases. *Curr. Opin. Biotechnol.* **27,** 96–106 (2014).
94. Koshland, D. E. Stereochemistry and the Mechanism of Enzymatic Reactions. *Biol. Rev.* **28,** 416–436 (1953).
95. Mark, B. L., Vocadlo, D. J., Knapp, S., Triggs-Raine, B. L., Withers, S. G. & James, M. N. G. Crystallographic Evidence for Substrate-assisted Catalysis in a Bacterial β-Hexosaminidase. *J. Biol. Chem.* **276,** 10330–10337 (2001).
96. Vocadlo, D. J. & Withers, S. G. Detailed comparative analysis of the catalytic mechanisms of β-N-acetylglucosaminidases from families 3 and 20 of glycoside hydrolases. *Biochemistry* **44,** 12809–12818 (2005).

97. Burmeister, W. P., Cottaz, S., Rollin, P., Vasella, A. & Henrissat, B. High resolution x-ray crystallography shows that ascorbate is a cofactor for myrosinase and substitutes for the function of the catalytic base. *J. Biol. Chem.* **275,** 39385–39393 (2000).

98. Rajan, S. S., Yang, X., Collart, F., Yip, V. L. Y., Withers, S. G., Varrot, A., Thompson, J., Davies, G. J. & Anderson, W. F. Novel catalytic mechanism of glycoside hydrolysis based on the structure of an NAD+/Mn2+-dependent phospho-α-glucosidase from Bacillus subtilis. *Structure* **12,** 1619–1629 (2004).

99. McCarter, J. D., Adam, Mi. J. & Withers, S. G. Binding Energy and Catalysis. Fluorinated and deoxygenated glycosides as mechanistic probes of Escherichia coli (lacZ) β-galactosidase. *Biochem. J.* **286,** 721–727 (1992).

100. Davies, G. J., Wilson, K. S. & Henrissat, B. Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* **321,** 557–559 (1997).

101. Percival Zhang, Y. H., Himmel, M. E. & Mielenz, J. R. Outlook for cellulase improvement: Screening and selection strategies. *Biotechnol. Adv.* **24,** 452–481 (2006).

102. Hasunuma, T., Okazaki, F., Okai, N., Hara, K. Y., Ishii, J. & Kondo, A. A review of enzymes and microbes for lignocellulosic biorefinery and the possibility of their application to consolidated bioprocessing technology. *Bioresour. Technol.* **135,** 513–522 (2013).

103. Van Wyk, J. P. H. Biowaste as a resource for bioproduct development. *Environ. Earth Sci.* **19,** 875–883 (2011).

104. Twomey, L. N., Pluske, J. R., Rowe, J. B., Choct, M., Brown, W., McConnell, M. F. & Pethick, D. W. The effects of increasing levels of soluble non-starch polysaccharides and inclusion of feed enzymes in dog diets on faecal quality and digestibility. *Anim. Feed Sci. Technol.* **108,** 71–82 (2003).

105. Sandrim, V. C., Rizzatti, A. C. S., Terenzi, H. F., Jorge, J. A., Milagres, A. M. F. & Polizeli, M. L. T. M. Purification and biochemical characterization of two xylanases produced by Aspergillus caespitosus and their potential for kraft pulp bleaching. *Process Biochem.* **40,** 1823–1828 (2005).

106. Mitidieri, S., Souza Martinelli, A. H., Schrank, A. & Vainstein, M. H. Enzymatic detergent formulation containing amylase from Aspergillus niger: A comparative study with commercial detergent formulations. *Bioresour. Technol.* **97,** 1217–1224 (2006).

107. Tsai, C. T. & Meyer, A. S. Enzymatic cellulose hydrolysis: Enzyme reusability and visualization of β-glucosidase immobilized in calcium alginate. *Molecules* **19,** 19390–19406 (2014).
108. McCarthy, T., Hanniffy, O., Lalor, E., Savage, A. V. & Tuohy, M. G. Evaluation of three thermostable fungal endo-β-glucanases from Talaromyces emersonii for brewing and food applications. *Process Biochem.* **40,** 1741–1748 (2005).
109. Daskalova, N. & Marinova, G. Study of the application in the brewing of enzyme preparation, containing laminarinase and lichenase, produced from Trichoderma Sp. 405. *J. Chem. Technol. Metall.* **50,** 149–156 (2015).
110. Mann, J. W., Jeffries, T. W. & MacMillan, J. D. Production and ecological significance of yeast cell wall-degrading enzymes from Oerskovia. *Appl. Environ. Microbiol.* **36,** 594–605 (1978).
111. John, R. P., Anisha, G. S., Nampoothiri, K. M. & Pandey, A. Micro and macroalgal biomass: A renewable source for bioethanol. *Bioresour. Technol.* **102,** 186–193 (2011).
112. Castresana, C., Carvalho, F. De, Gheysen, G., Habets, M., Inze, D. & Montagu, M. Van. Tissue-Specific and Pathogen-lnduced Regulation of a Nicotiana plumbaginifolia β-1,3-Glucanase Gene. *Plant Cell* **2,** 1131–1143 (1990).
113. Poorna, C. A. Optimization and Application to Produce Gluco-oligosaccharides from Bacillus pumilus and their potential for hydrolysis of polysaccharides. *Ferment. Technol.* **1,** 1–5 (2011).
114. Wang, D., Kim, D. H. & Kim, K. H. Effective production of fermentable sugars from brown macroalgae biomass. *Appl. Microbiol. Biotechnol.* **100,** 9439–9450 (2016).
115. Kaoutari, A. El, Armougom, F., Gordon, J. I., Raoult, D. & Henrissat, B. The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat. Rev. Microbiol.* **11,** 497–504 (2013).
116. Tasse, L., Bercovici, J., Pizzut-Serin, S., Robe, P., Tap, J., Klopp, C., Cantarel, B. L., Coutinho, P. M., Henrissat, B., Leclerc, M., Doré, J., Monsan, P., Remaud-Simeon, M. & Potocki-Veronese, G. Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. *Genome Res.* **20,** 1605–1612 (2010).
117. Hamaker, B. R. & Tuncil, Y. E. A Perspective on the Complexity of Dietary Fiber Structures and Their Potential Effect on the Gut Microbiota. *J. Mol. Biol.* **426,** 3838–3850

(2014).

118. Wexler, H. M. Bacteroides: The good, the bad, and the nitty-gritty. *Clin. Microbiol. Rev.* **20,** 593–621 (2007).

119. Xu, J., Bjursell, M. K., Himrod, J., Deng, S., Carmichael, L. K., Chiang, H. C., Hooper, L. V & Gordon, J. I. A Genomic View of the Human: Bacteroides thetaiotaomicron Symbiosis. *Science* **299,** 2074–2076 (2003).

120. Foley, M. H., Cockburn, D. W. & Koropatkin, N. M. The Sus operon: a model system for starch uptake by the human gut Bacteroidetes. *Cell. Mol. Life Sci.* **73,** 2603–2617 (2016).

121. Bagenholm, V., Reddy, S. K., Bouraoui, H., Morrill, J., Kulcinskaja, E., Bahr, C. M., Aurelius, O., Rogers, T., Xiao, Y., Logan, D. T., Martens, E. C., Koropatkin, N. M. & Stalbrand, H. Galactomannan catabolism conferred by a polysaccharide utilization locus of Bacteroides ovatus: Enzyme synergy and crystal structure of a β-mannanase. *J. Biol. Chem.* **292,** 229–243 (2017).

122. Martens, E. C., Lowe, E. C., Chiang, H., Pudlo, N. A., Wu, M., McNulty, N. P., Abbott, D. W., Henrissat, B., Gilbert, H. J., Bolam, D. N. & Gordon, J. I. Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. *PLoS Biol.* **9,** (2011).

123. Déjean, G., Tamura, K., Cabrera, A., Jain, N., Pudlo, N. A., Pereira, G., Viborg, A. H., Petegem, F. Van, Martens, E. C. & Brumer, H. Synergy between Cell Surface Glycosidases and Glycan-Binding Proteins Dictates the Utilization of Specific Beta(1,3)-Glucans by Human Gut Bacteroides. *MBio* **11,** e00095-20 (2020).

124. DeBoy, R. T., Mongodin, E. F., Fouts, D. E., Tailford, L. E., Khouri, H., Emerson, J. B., Mohamoud, Y., Watkins, K., Henrissat, B., Gilbert, H. J. & Nelson, K. E. Insights into plant cell wall degradation from the genome sequence of the soil bacterium Cellvibrio japonicus. *J. Bacteriol.* **190,** 5455–5463 (2008).

125. Nelson, C. E. & Gardner, J. G. In-frame deletions allow functional characterization of complex cellulose degradation phenotypes in Cellvibrio japonicus. *Appl. Environ. Microbiol.* **81,** 5968–5975 (2015).

126. Brumer, H. & Gilbert, H. J. Editorial overview: Carbohydrate–protein interactions: The future is taking shape. *Curr. Opin. Struct. Biol.* **28,** v–vii (2014).

127. Garron, M. L. & Henrissat, B. The continuing expansion of CAZymes and their families.

Curr. Opin. Chem. Biol. **53,** 82–87 (2019).

128. Brown, J. R., Yang, F., Sinha, A., Ramakrishnan, B., Tor, Y., Qasba, P. K. & Esko, J. D. Deoxygenated disaccharide analogs as specific inhibitors of β1-4-galactosyltransferase 1 and selectin-mediated tumor metastasis. *J. Biol. Chem.* **284,** 4952–4959 (2009).
129. Laferté, S., Chan, N. W. C., Sujino, K., Lowary, T. L. & Palcic, M. M. Intracellular inhibition of blood group A glycosyltransferase. *Eur. J. Biochem.* **267,** 4840–4849 (2000).
130. Asano, N. Glycosidase inhibitors: Update and perspectives on practical use. *Glycobiology* **13,** 93–104 (2003).
131. Inouye, S., Tsuruoka, T., Ito, T. & Niida, T. Structure and synthesis of nojirimycin. *Tetrahedron* **24,** 2125–2144 (1968).
132. De La Fuente, A., Martin, R., Mena-Barragán, T., Verdaguer, X., García Fernández, J. M., Ortiz Mellet, C. & Riera, A. Stereoselective synthesis of 2-acetamido-1,2-dideoxyallonojirimycin (DAJNAc), a new potent hexosaminidase inhibitor. *Org. Lett.* **15,** 3638–3641 (2013).
133. Hohenschutz, L. D., Bell, E. A., Jewess, P. J., Leworthy, D. P., Pryce, R. J., Arnold, E. & Clardy, J. Castanospermine, A 1,6,7,8-tetrahydroxyoctahydroindolizine alkaloid, from seeds of Castanospermum australe. *Phytochemistry* **20,** 811–814 (1981).
134. Dong, W., Jespersen, T., Bols, M., Skrydstrup, T. & Sierks, M. R. Evaluation of isofagomine and its derivatives as potent glycosidase inhibitors. *Biochemistry* **35,** 2788–2795 (1996).
135. Chen, X., Fan, Y., Zheng, Y. & Shen, Y. Properties and production of valienamine and its related analogues. *Chem. Rev.* **103,** 1955–1977 (2003).
136. Guo, W., Hiratake, J., Ogawa, K., Yamamoto, M., Ma, S. J. & Sakata, K. β-D-Glycosylamidines: Potent, selective, and easily accessible β-glycosidase inhibitors. *Bioorg. Med. Chem. Lett.* **11,** 467–470 (2001).
137. Berecibar, A., Grandjean, C. & Siriwardena, A. Synthesis and Biological Activity of Natural Aminocyclopentitol Glycosidase Inhibitors: Mannostatins, Trehazolin, Allosamidins, and Their Analogues. *Chem. Rev.* **99,** 779–844 (1999).
138. Breuer, H. W. M. Review of acarbose therapeutic strategies in the long-term treatment and in the prevention of type 2 diabetes. *Int. J. Clin. Pharmacol. Ther.* **41,** 421 (2003).
139. Iwasa, T., Higashide, E., Yamamoto, H. & Shibata, M. Studies on validamycins, new

antibiotics. II. Production and biological properties of validamycins A and B. *J. Antibiot. (Tokyo)* **24,** 107 (1971).

140. Ghavami, A., Johnston, B. D. & Pinto, B. M. A new class of glycosidase inhibitor: Synthesis of salacinol and its stereoisomerst. *J. Org. Chem.* **66,** 2312–2317 (2001).
141. Andrews, J. S., Weimar, T., Pinto, B. M., Frandsen, T. P. & Svensson, B. Novel Disaccharides Containing Sulfur in the Ring and Nitrogen in the Interglycosidic Linkage. Conformation of Methyl 5'-Thio-4-N-α-Maltoside Bound to Glucoamylase and Its Activity as a Competitive Inhibitor. *J. Am. Chem. Soc.* **117,** 10799–10804 (1995).
142. Gloster, T. M., Roberts, S., Perugino, G., Moracci, M., Panday, N., Terinek, M., Vasella, A., Davies, G. J. & Zu, C.-. Structural, Kinetic, and Thermodynamic Analysis of Glucoimidazole-Derived Glycosidase Inhibitors. *Biochemistry* **45,** 11879–11884 (2006).
143. Gloster, T. M. & Vocadlo, D. J. Developing inhibitors of glycan processing enzymes as tools for enabling glycobiology. *Nat. Chem. Biol.* **8,** 683–694 (2012).
144. Lahiri, R., Ansari, A. A. & Vankar, Y. D. Recent developments in design and synthesis of bicyclic azasugars, carbasugars and related molecules as glycosidase inhibitors. *Chem. Soc. Rev.* **42,** 5102–5118 (2013).
145. Elbein, A. D., Pan, Y. T., Pastuszak, I. & Carroll, D. New insights on trehalose: a multifunctional molecule. *Glycobiol.* **13,** 17R-27R (2003).
146. Gloster, T. M. Development of inhibitors as research tools for carbohydrate-processing enzymes. *Biochem. Soc. Trans.* **40,** 913–928 (2012).
147. Rempel, B. P. & Withers, S. G. Covalent inhibitors of glycosidases and their applications in biochemistry and biology. *Glycobiology* **18,** 570–586 (2008).
148. Witte, M. D., Van Der Marel, G. A., Aerts, J. M. F. G. & Overkleeft, H. S. Irreversible inhibitors and activity-based probes as research tools in chemical glycobiology. *Org. Biomol. Chem.* **9,** 5908–5926 (2011).
149. Withers, S. G. & Aebersold, R. Approaches to labeling and identification of active site residues in glycosidases. *Protein Sci.* **4,** 361–372 (1995).
150. Vodovozova, E. L. Photoaffinity labeling and its application in structural biology. *Biochem.* **72,** 1–20 (2007).
151. Kuhn, C.-S., Lehmann, J., Jung, G. & Stevanović, S. Investigation of the active site of Escherichia coli β-d-galactosidase by photoaffinity labelling. *Carbohydr. Res.* **232,** 227–

233 (1992).

152. Liessem, B., Glombitza, G. J., Knoll, F., Lehmann, J., Kellermann, J., Lottspeich, F. & Sandhoff, K. Photoaffinity labeling of human lysosomal β-hexosaminidase B. Identification of Glu-355 at the substrate binding site. *J. Biol. Chem.* **270,** 23693–23699 (1995).

153. Marshall, P. J., Sinnott, M. L., Smith, P. J. & Widdows, D. Active-site-directed irreversible inhibition of glysosidases by the corresponding glycosylmethyl-(p-nitrophenyl)triazenes. *J Chem Soc Perkin Trans I* **2,** 366–376 (1981).

154. BeMiller, J. N., Gilson, R. J., Myers, R. W. & Santoro, M. M. Suicide-substrate inactivation of β-galactosidase by diazomethyl β-D-galactopyranosyl ketone. *Carbohydr. Res.* **250,** 101–112 (1993).

155. Schnabelrauch, M., Vasella, A. & Withers, S. G. Synthesis and Evaluation as Irreversible Glycosidase Inhibitors of Mono- and Oligo(glycosylthio)benzoquinones. *Helv. Chim. Acta* **77,** 778–799 (1994).

156. Shulman, M. L., Shiyan, S. D. & Khorlin, A. Y. Specific irreversible inhibition of sweet-almond β-glucosidase by some β-glycopyranosylepoxyalkanes and β-d-glucopyranosyl isothiocyanate. *Biochim. Biophys. Acta - Enzymol.* **445,** 169–181 (1976).

157. Black, T. S., Kiss, L., Tull, D. & Withers, S. G. N-Bromoacetyl-glycopyranosylamines as affinity labels for a β-glucosidase and a cellulase. *Carbohydr. Res.* **250,** 195–202 (1993).

158. Howard, S. & Withers, S. G. Bromoketone C- Glycosides, a New Class of β-Glucanase Inactivators. *J. Am. Chem. Soc.* **120,** 10326–10331 (1998).

159. Howard, S. & Withers, S. G. Labeling and identification of the postulated acid/base catalyst in the α-glucosidase from Saccharomyces cerevisiae using a novel bromoketone C-glycoside. *Biochemistry* **37,** 3858–3864 (1998).

160. Tull, D., Burgoyne, D. L., Chow, D. T., Withers, S. G. & Aebersold, R. A Mass Spectrometry-Based Approach for Probing Enzyme Active Sites: Identification of Glu 127 in Cellulomonas fimi Exoglycanase as the Residue Modified by N-Bromoacetyl Cellobiosylamine. *Anal. Biochem.* **234,** 119–125 (1996).

161. Jäger, S. & Kiss, L. Investigation of the active site of the extracellular β-D-glucosidase from Aspergillus carbonarius. *World J. Microbiol. Biotechnol.* **21,** 337–343 (2005).

162. Viratelle, M., Yon, J. M. & Yariv, J. The inactivation of β-galactosidase by N-

bromoacetyl-β-D-glucosylamine. *FEBS Lett.* **79,** 109–112 (1977).

163. Chir, J., Withers, S., Wan, C.-F. & Li, Y.-K. Identification of the two essential groups in the family 3 β-glucosidase from Flavobacterium meningosepticum by labelling and tandem mass spectrometric analysis. *Biochem J.* **365,** 857–863 (2002).
164. Keresztessy, Z., Kiss, L. & Hughes, M. A. Investigation of the active site of the cyanogenic β-D-glucosidase (linamarase) from Manihot esculenta Crantz (Cassava). II. Identification of Glu-198 as an active site carboxylate group with acid catalytic function. *Arch. Biochem. Biophys.* **315,** 323–330 (1994).
165. Legler, G., Sinnott, M. L. & Withers, S. G. Catalysis by β-Glucosidase A3 of Aspergillus wentii. *J. Chem. Soc. Perkin Trans. 2* 1376–1383 (1980).
166. Naider, F., Bohak, Z. & Yariv, J. Reversible alkylation of a methionyl residue near the active site of β-galactosidase. *Biochemistry* **11,** 3202–3208 (1972).
167. Vocadlo, D. J., Wicki, J., Rupitz, K. & Withers, S. G. A case for reverse protonation: Identification of Glu160 as an acid/base catalyst in Thermoanaerobacterium saccharolyticum β-xylosidase and detailed kinetic analysis of a site-directed mutant. *Biochemistry* **41,** 9736–9746 (2002).
168. Kiss, T., Erdei, A. & Kiss, L. Investigation of the active site of the extracellular β-D-xylosidase from Aspergillus carbonarius. *Arch. Biochem. Biophys.* **399,** 188–194 (2002).
169. Fekete, C. A. & Kiss, L. A New approach in the active site investigation of an inverting β-D-xylosidase from Thermobifida fusca TM51. *Protein J.* **32,** 97–105 (2013).
170. Fenger, T. H. & Brumer, H. Synthesis and Analysis of Specific Covalent Inhibitors of endo-Xyloglucanases. *ChemBioChem* **16,** 575–583 (2015).
171. Attia, M. A., Nelson, C. E., Offen, W. A., Jain, N., Davies, G. J., Gardner, J. G. & Brumer, H. In vitro and in vivo characterization of three Cellvibrio japonicus glycoside hydrolase family 5 members reveals potent xyloglucan backbone-cleaving functions. *Biotechnol. Biofuels* **11,** 45–61 (2018).
172. Ueda, T., Tomita, K., Notsu, Y., Ito, T., Fumoto, M., Takakura, T., Nagatome, H., Takimoto, A., Mihara, S. I., Togame, H., Kawamoto, K., Iwasaki, T., Asakura, K., Oshima, T., Hanasaki, K., Nishimura, S. I. & Kondo, H. Chemoenzymatic synthesis of glycosylated glucagon-like peptide 1: Effect of glycosylation on proteolytic resistance and in vivo blood glucose-lowering activity. *J. Am. Chem. Soc.* **131,** 6237–6245 (2009).

173. Chauvigne-Hines, L. M., Anderson, L. N., Weaver, H. M., Brown, J. N., Koech, P., Nicora, C. D., Hofstad, B. a, Smith, R. D., Wilkins, M. J., Callister, S. J. & Wright, A. T. A Suite of Activity-Based Probes for Cellulose Degrading Enzymes. *J. Am. Chem. Soc.* **134,** 20521–20532 (2012).
174. Legler, G. Glycoside Hydrolases: Mechanistic Information from Studies with Reversible and Irreversible Inhibitors. *Adv. Carbohydr. Chem. Biochem.* **48,** 319–384 (1990).
175. Walsh, C. Suicide substrates: mechanism-based enzyme inactivators. *Tetrahedron* **38,** 871–909 (1982).
176. Sinnott, M. L. & Smith, P. J. Active-site-directed irreversible inhibition of E. coliβ-galactosidase by the 'hot' carbonium ion precursor, β-D-galactopyranosylmethyl-p-nitrophenyltriazene. *J. Chem. Soc. Chem. Commun.* **6,** 223–224 (1976).
177. Halazy, S., Berges, V., Ehrhard, A. & Danzin, C. Ortho- and para-(difluoromethyl)aryl-β-d-glucosides: A new class of enzyme-activated irreversible inhibitors of β-glucosidases. *Bioorg. Chem.* **18,** 330–344 (1990).
178. Braun, H., Legler, G., Deshusses, J. & Semenza, G. Stereospecific ring opening of conduritol-B-epoxide by an active site aspartate residue of sucrase-isomaltase. *Biochim. Biophys. Acta - Enzymol.* **483,** 135–140 (1977).
179. Withers, S. G. & Umezawa, K. Cyclophellitol: A naturally occurring mechanism-based inactivator of beta-glucosidases. *Biochem. Biophys. Res. Commun.* **177,** 532–537 (1991).
180. Caron, G. & Withers, S. G. Conduritol aziridine: a new mechanism-based glucosidase inactivator. *Biochem. Biophys. Res. Commun.* **163,** 495–499 (1989).
181. Alcaide, A., Trapero, A., Pérez, Y. & Llebaria, A. Galacto configured N-aminoaziridines: a new type of irreversible inhibitor of β-galactosidases. *Org. Biomol. Chem.* **13,** 5690–7 (2015).
182. Artola, M., Wu, L., Ferraz, M. J., Kuo, C. L., Raich, L., Breen, I. Z., Offen, W. A., Codée, J. D. C., Van Der Marel, G. A., Rovira, C., Aerts, J. M. F. G., Davies, G. J. & Overkleeft, H. S. 1,6-Cyclophellitol Cyclosulfates: A New Class of Irreversible Glycosidase Inhibitor. *ACS Cent. Sci.* **3,** 784–793 (2017).
183. Ren, W., Pengelly, R., Farren-Dai, M., Shamsi Kazem Abadi, S., Oehler, V., Akintola, O., Draper, J., Meanwell, M., Chakladar, S., Świderek, K., Moliner, V., Britton, R., Gloster, T. M. & Bennet, A. J. Revealing the mechanism for covalent inhibition of glycoside

hydrolases by carbasugars at an atomic level. *Nat. Commun.* **9,** 1–12 (2018).

184. Chakladar, S., Wang, Y., Clark, T., Cheng, L., Ko, S., Vocadlo, D. J. & Bennet, A. J. A mechanism-based inactivator of glycoside hydrolases involving formation of a transient non-classical carbocation. *Nat. Commun.* **5,** 5590 (2014).
185. Withers, S. G., Street, I. P., Bird, P. & Dolphin, D. H. 2-Deoxy-2-fluoroglucosides: A Novel Class of Mechanism-Based Glucosidase Inhibitors. *J. Am. Chem. Soc.* **109,** 7530–7531 (1987).
186. Roeser, K. R. & Legler, G. Role of sugar hydroxyl groups in glycoside hydrolysis. Cleavage mechanism of deoxyglucosides and related substrates by β-glucosidase A3 from Aspergillus wentii. *BBA - Enzymol.* **657,** 321–333 (1981).
187. Kallemeijn, W. W., Witte, M. D., Wennekes, T. & Aerts, J. M. F. G. in *Adv. Carbohydr. Chem. Biochem.* (ed. Horton, D.) **71,** 297–338 (Elsevier Inc., 2014).
188. Braun, C., Brayer, G. D. & Withers, S. G. Mechanism-based inhibition of yeast α-glucosidase and human pancreatic α-amylase by a new class of inhibitors. *J. Biol. Chem.* **270,** 26778–26781 (1995).
189. Thanna, S., Lindenberger, J. J., Gaitonde, V. V., Ronning, D. R. & Sucheck, S. J. Synthesis of 2-deoxy-2,2-difluoro-α-maltosyl fluoride and its X-ray structure in complex with Streptomyces coelicolor GlgEI-V279S. *Org. Biomol. Chem.* **13,** 7542–7550 (2015).
190. Zhang, R., McCarter, J. D., Braun, C., Yeung, W., Brayer, G. D. & Withers, S. G. Synthesis and testing of 2-deoxy-2,2-dihaloglycosides as mechanism-based inhibitors of α-glycosidases. *J. Org. Chem.* **73,** 3070–3077 (2008).
191. McCarter, J. D., Adam, M. J., Braun, C., Namchuk, M., Tull, D. & Withers, S. G. Syntheses of 2-deoxy-2-fluoro mono- and oligo-saccharide glycosides from glycals and evaluation as glycosidase inhibitors. *Carbohydr. Res.* **249,** 77–90 (1993).
192. Hart, D. O., He, S., Chany, C. J., Withers, S. G., Sims, P. F. G., Sinnott, M. L. & Brumer, H. Identification of asp-130 as the catalytic nucleophile in the main α-galactosidase from Phanerochaete chrysosporium, a family 27 glycosyl hydrolase. *Biochemistry* **39,** 9826–9836 (2000).
193. Berkowitz, D. B., Karukurichi, K. R., Salud-bea, R. De, Nelson, D. L., Mccune, C. D., de la Salud-Bea, R., Nelson, D. L. & Mccune, C. D. Use of Fluorinated Functionality in Enzyme Inhibitor Development: Mechanistic and Analytical Advantages. *J. Fluor. Chem.*

129, 731–742 (2008).

194. Withers, S. G., Rupitz, K. & Street, I. P. 2-Deoxy-2-fluoro-D-glycosyl fluorides. A new class of specific mechanism-based glycosidase inhibitors. *J. Biol. Chem.* **263,** 17–20 (1988).
195. Miao, S., Ziser, L., Aebersold, R. & Withers, S. G. Identification of Glutamic Acid 78 as the Active Site Nucleophile in Bacillus subtilis Xylanase Using Electrospray Tandem Mass Spectrometry. *Biochemistry* **33,** 7027–7032 (1994).
196. Wang, Q., Tull, D., Meinke, A., Gilkes, N. R., Warren, R. A. J., Aebersold, R. & Withers, S. G. Glu280 is the nucleophile in the active site of Clostridium thermocellum CelC, a family A endo-β-1,4-glucanase. *J. Biol. Chem.* **268,** 14096–14102 (1993).
197. Sacffidi, A., Stick, R. V. & Stubbs, K. A. Synthesis of Some Glycosylated Derivatives of 2-Deoxy-2-fluoro-β-laminaribiosyl Fluoride: Another Success for Glycosynthases. *Aust. J. Chem.* **38,** 83–88 (2007).
198. Goddard-Borger, E. D., Fiege, B., Kwan, E. M. & Withers, S. G. Glycosynthase-Mediated Assembly of Xylanase Substrates and Inhibitors. *ChemBioChem* **12,** 1703–1711 (2011).
199. Vocadlo, D. J. & Withers, S. G. The chemical synthesis of 2-deoxy-2-fluorodisaccharide probes of the hen egg white lysozyme mechanism. *Carbohydr. Res.* **340,** 379–388 (2005).
200. Vocadlo, D. J., Davies, G. J., Laine, R. & Withers, S. G. Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature* **412,** 835–8 (2001).
201. Sabini, E., Sulzenbacher, G., Dauter, M., Dauter, Z., Jorgensen, P. L., Schulein, M., Dupont, C., Davies, G. J. & Wilson, K. S. Catalysis and specificity in enzymatic glycoside hydrolysis: a 2,5 B conformation for the glycosyl enzyme intermediate revealed by the structure of the Bacillus agaradhaerens family 11 xylanase. *Chem. Biol.* **6,** 483–492 (1999).
202. Gebler, J. C., Aebersold, R. & Withers, S. G. Glu-537, not Glu-461, is the nucleophile in the active site of (lac Z) β- galactosidase from Escherichia coli. *J. Biol. Chem.* **267,** 11126–11130 (1992).
203. Jochems, P., Mueller, T., Satyawali, Y., Diels, L., Dejonghe, W. & Hanefeld, U. Active site titration of immobilized beta-galactosidase for the determination of active enzymes. *Biochem. Eng. J.* **93,** 137–141 (2014).
204. Miao, S., McCarter, J. D., Grace, M. E., Grabowski, G. A., Aebersold, R. & Withers, S. G.

Identification of Glu340 as the active-site nucleophile in human glucocerebrosidase by use of electrospray tandem mass spectrometry. *J. Biol. Chem.* **269,** 10975–10978 (1994).

205. Mackenzie, L. F., Sulzenbacher, G., Divne, C., Jones, T. A., Wo, H. F., Schu, M., Withers, S. G. & Davies, G. J. Crystal structure of the family 7 endoglucanase I (Cel7B) from Humicola insolens at 2.2 Å resolution and identification of the catalytic nucleophile by trapping of the covalent glycosyl-enzyme intermediate. *Biochem. J.* **335,** 409–416 (1998).
206. Street, I. P., Kempton, J. B. & Withers, S. G. Inactivation of a β-glucosidase through the accumulation of a stable 2-deoxy-2-fluoro-α-D-glucopyranosyl-enzyme intermediate: A detailed investigation. *Biochemistry* **31,** 9970–9978 (1992).
207. Withers, S. G., Warren, A. J., Street, I. P., Rupitz, K., Kempton, J. B. & Aebersold, R. Unequivocal Demonstration of the Involvement of a Glutamate Residue as a Nucleophile in the Mechanism of a "Retaining" Glycosidase. *J. Am. Chem. Soc.* **112,** 5887–5889 (1990).
208. Tarling, C. A., He, S., Sulzenbacher, G., Bignon, C., Bourne, Y., Henrissat, B. & Withers, S. G. Identification of the Catalytic Nucleophile of the Family 29 α-L-Fucosidase from Thermotoga maritima through Trapping of a Covalent Glycosyl-Enzyme Intermediate and Mutagenesis. *J. Biol. Chem.* **278,** 47394–47399 (2003).
209. McCarter, J. D. & Withers, S. G. 5-Fluoro glycosides: A new class of mechanism-based inhibitors of both α- and β-glucosidases. *J. Am. Chem. Soc.* **118,** 241–242 (1996).
210. McCarter, J. D. & Withers, S. G. Unequivocal identification of Asp-214 as the catalytic nucleophile of Saccharomyces cerevisiae α-glucosidase using 5-fluoro glycosyl fluorides. *J. Biol. Chem.* **271,** 6889–6894 (1996).
211. Ly, H. D., Howard, S., Shum, K., He, S., Zhu, A. & Withers, S. G. The synthesis, testing and use of 5-fluoro-α-D-galactosyl fluoride to trap an intermediate on green coffee bean α-galactosidase and identify the catalytic nucleophile. *Carbohydr. Res.* **329,** 539–547 (2000).
212. Howard, S., He, S. & Withers, S. G. Identification of the active site nucleophile in jack bean α-mannosidase using 5-fluoro-β-L-gulosyl fluoride. *J. Biol. Chem.* **273,** 2067–2072 (1998).
213. Vocadlo, D. J., Mayer, C., He, S. & Withers, S. G. Mechanism of action and identification

of Asp242 as the catalytic nucleophile of Vibrio furnisii N-acetyl-β-D-glucosaminidase using 2- acetamido-2-deoxy-5-fluoro-α-L-idopyranosyl fluoride. *Biochemistry* **39,** 117–126 (2000).

214. Hekmat, O., Florizone, C., Kim, Y.-W., Eltis, L. D., Warren, R. A. J. & Withers, S. G. Specificity Fingerprinting of Retaining β-1,4-Glycanases in the Cellulomonas fimi Secretome Using Two Fluorescent Mechanism-Based Probes. *ChemBioChem* **8,** 2125–2132 (2007).
215. Williams, S. J., Hekmat, O. & Withers, S. G. Synthesis and testing of mechanism-based protein-profiling probes for retaining endoglycosidases. *ChemBioChem* **7,** 116–124 (2006).
216. Walvoort, M. T. C., Kallemeijn, W. W., Willems, L. I., Witte, M. D., Aerts, J. M. F. G., Van Der Marel, G. A., Codée, J. D. C. & Overkleeft, H. S. Tuning the leaving group in 2-deoxy-2-fluoroglucoside results in improved activity-based retaining β-glucosidase probes. *Chem. Commun.* **48,** 10386–10388 (2012).
217. Michaelis, L., Menten, M. L., Goody, R. S. & Johnson, K. A. Die Kinetik der Invertinwirkung/ The kinetics of invertase action. *Biochemistry* **49,** 333–369 (1913).
218. Briggs, G. E. & Haldane, J. B. S. A Note on the Kinetics of Enzyme Action. *Biochem. J.* **19,** 338–339 (1925).
219. Koshland, D. E. The Application and Usefulness of the ratio kcat/KM. *Bioorg. Chem.* **30,** 211–213 (2002).
220. Eisenthal, R., Danson, M. J. & Hough, D. W. Catalytic efficiency and k_{cat}/K_M: a useful comparator? *Trends Biotechnol.* **25,** 247–249 (2007).
221. Kitz, R. & Wilson, I. B. Esters of methanesulfonic acid as irreversible inhibitors of acetylcholinesterase. *J. Biol. Chem.* **237,** 3245–3249 (1962).
222. Heck, A. J. R. & Van Den Heuvel, R. H. H. Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev.* **23,** 368–389 (2004).
223. Daniel, J. M., Friess, S. D., Rajagopalan, S., Wendt, S. & Zenobi, R. Quantitative determination of noncovalent binding interactions using soft ionization mass spectrometry. *Int. J. Mass Spectrom.* **216,** 1–27 (2002).
224. Ruan, Q., Ji, Q. C., Arnold, M. E., Humphreys, W. G. & Zhu, M. Strategy and its implications of protein bioanalysis utilizing high-resolution mass spectrometric detection

of intact protein. *Anal. Chem.* **83,** 8937–8944 (2011).

225. Van Den Heuvel, R. H. H. & Heck, A. J. R. Native protein mass spectrometry: From intact oligomers to functional machineries. *Curr. Opin. Chem. Biol.* **8,** 519–526 (2004).
226. Schiano-di-Cola, C., Kołaczkowski, B., Sørensen, T. H., Christensen, S. J., Cavaleiro, A. M., Windahl, M. S., Borch, K., Morth, J. P. & Westh, P. Structural and biochemical characterization of a family 7 highly thermostable endoglucanase from the fungus Rasamsonia emersonii. *FEBS J.* **286,** 1–20 (2019).
227. Kallas, Å. M., Piens, K., Denman, S. E., Henriksson, H., Fäldt, J., Johansson, P., Brumer, H. & Teeri, T. T. Enzymatic properties of native and deglycosylated hybrid aspen (*Populus tremula x tremuloides*) xyloglucan endotransglycosylase 16A expressed in Pichia pastoris. *Biochem. J.* **390,** 105–113 (2005).
228. Yakovlieva, L. & Walvoort, M. T. C. Processivity in Bacterial Glycosyltransferases. *ACS Chem. Biol.* **15,** 3–16 (2020).
229. Bojarová, P., Kulik, N., Slámová, K., Hubálek, M., Kotik, M., Cvačka, J., Pelantová, H. & Křen, V. Selective β-N-acetylhexosaminidase from Aspergillus versicolor—a tool for producing bioactive carbohydrates. *Appl. Microbiol. Biotechnol.* **103,** 1737–1753 (2019).
230. Notenboom, V., Williams, S. J., Hoos, R., Withers, S. G. & Rose, D. R. Detailed structural analysis of glycosidase/inhibitor interactions: Complexes of cex from Cellulomonas fimi with xylobiose-derived aza-sugars. *Biochemistry* **39,** 11553–11563 (2000).
231. Varrot, A., Schülein, M., Pipelier, M., Vasella, A. & Davies, G. J. Lateral protonation of a glycosidase inhibitor. Structure of the Bacillus agaradhaerens Cel5A in complex with a cellobiose-derived imidazole at 0.97 Å resolution. *J. Am. Chem. Soc.* **121,** 2621–2622 (1999).
232. Johnson, L. N. & Phillips, D. C. Structure Of Some Crystalline Lysozyme-inhibitor Complexes Determined By X-ray Analysis At 6 A Resolution. *Nature* **206,** 757–761 (1965).
233. Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. Structure of Hen egg-white lysozyme. *Nature* **206,** 757–761 (1965).
234. Notenboom, V., Birsan, C., Warren, R. A. J., Withers, S. G. & Rose, D. R. Exploring the cellulose/xylan specificity of the β-1,4-glycanase Cex from Cellulomonas fimi through

crystallography and mutation. *Biochemistry* **37,** 4751–4758 (1998).

235. Notenboom, V., Birsan, C., Nitz, M., Rose, D. R., Warren, R. A. J. & Withers, S. G. Insights into transition state stabilization of the β-1,4-glycosidase Cex by covalent intermediate accumulation in active site mutants. *Nat. Struct. Biol.* **5,** 812–818 (1998).
236. Burmeister, W. P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S. & Henrissat, B. The crystal structures of Sinapis alba myrosinase and a covalent glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinery of an S-glycosidase. *Structure* **5,** 663–676 (1997).
237. Davies, G. J., Mackenzie, L., Varrot, A., Dauter, M., Brzozowski, A. M., Schülein, M. & Withers, S. G. Snapshots along an enzymatic reaction coordinate: Analysis of a retaining β-glycoside hydrolase. *Biochemistry* **37,** 11707–11713 (1998).
238. Sidhu, G., Withers, S. G., Nguyen, N. T., McIntosh, L. P., Ziser, L. & Brayer, G. D. Sugar ring distortion in the glycosyl-enzyme intermediate of a family G/11 xylanase. *Biochemistry* **38,** 5346–5354 (1999).
239. Witte, M. D., Walvoort, M. T. C., Li, K. Y., Kallemeijn, W. W., Donker-Koopman, W. E., Boot, R. G., Aerts, J. M. F. G., Codée, J. D. C., Van der Marel, G. A., Overkleeft, H. S., Codee, J. D., Van der Marel, G. A. & Overkleeft, H. S. Activity-Based Profiling of Retaining β-Glucosidases: A Comparative Study. *ChemBioChem* **12,** 1263–1269 (2011).
240. Wu, L., Jiang, J., Jin, Y., Kallemeijn, W. W., Kuo, C. L., Artola, M., Dai, W., Van Elk, C., Van Eijk, M., Van Der Marel, G. A., Codée, J. D. C., Florea, B. I., Aerts, J. M. F. G., Overkleeft, H. S. & Davies, G. J. Activity-based probes for functional interrogation of retaining β-glucuronidases. *Nat. Chem. Biol.* **13,** 867–877 (2017).
241. Saini, J. K., Saini, R. & Tewari, L. Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: concepts and recent developments. *3 Biotech* **5,** 337–353 (2015).
242. Popa, V. I. in *Biomass as Renewable Raw Material to Obtain Bioproducts of High-Tech Value* (eds. Popa, V. & Volf, I.) 1–37 (Elsevier, 2018).
243. Rubin, E. M., Himmel, M. E., Ding, S., Johnson, D. K. & Adney, W. S. Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Nature* **454,** 804–807 (2007).
244. Sun, J., Ding, S. & Doran-Peterson, J. in *Biol. Convers. Biomass Fuels Chem. Explor.*

from Nat. Util. Syst. (eds. Sun, J., Ding, S. & Doran-Peterson, J.) 1–13 (The Royal Society of Chemistry, 2014).

245. Willis, J. D., Mazarei, M. & Stewart, C. N. Transgenic Plant-Produced Hydrolytic Enzymes and the Potential of Insect Gut-Derived Hydrolases for Biofuels. *Front. Plant Sci.* **7,** 1–18 (2016).
246. Horn, S. J., Vaaje-Kolstad, G., Westereng, B. & Eijsink, V. G. Novel enzymes for the degradation of cellulose. *Biotechnol. Biofuels* **5,** 45 (2012).
247. Asano, N., Nash, R. J., Molyneux, R. J. & Fleet, G. W. J. Sugar-mimic glycosidase inhibitors: natural occurrence, biological activity and prospects for therapeutic application. *Tetrahedron: Asymmetry* **11,** 1645–1680 (2000).
248. Xu, Y., Uddin, N. & Wagner, G. K. in *Methods Enzymol.* (eds. Pyle, A. M. & Christianson, D. W.) **598,** 237–265 (Elsevier Inc., 2018).
249. Willems, L. I., Jiang, J., Li, K. Y., Witte, M. D., Kallemeijn, W. W., Beenakker, T. J. N., Schröder, S. P., Aerts, J. M. F. G., van der Marel, G. A., Codée, J. D. C. & Overkleeft, H. S. From covalent glycosidase inhibitors to activity-based glycosidase probes. *Chemistry* **20,** 10864–10872 (2014).
250. Jiang, J., Beenakker, T. J. M., Kallemeijn, W. W., van der Marel, G. A., van den Elst, H., Codée, J. D. C., Aerts, J. M. F. G. & Overkleeft, H. S. Comparing Cyclophellitol N-Alkyl and N-Acyl Cyclophellitol Aziridines as Activity-Based Glycosidase Probes. *Chem. A Eur. J.* **21,** 10861–10869 (2015).
251. Adamson, C., Pengelly, R. J., Shamsi Kazem Abadi, S., Chakladar, S., Draper, J., Britton, R., Gloster, T. M. & Bennet, A. J. Structural Snapshots for Mechanism-Based Inactivation of a Glycoside Hydrolase by Cyclopropyl Carbasugars. *Angew. Chem. Int. Ed.* 14978–14982 (2016).
252. Blanchard, J. E. & Withers, S. G. Rapid screening of the aglycone specificity of glycosidases: Applications to enzymatic synthesis of oligosaccharides. *Chem. Biol.* **8,** 627–633 (2001).
253. Watts, A. G., Damager, I., Amaya, M. L., Buschiazzo, A., Alzari, P., Frasch, A. C. & Withers, S. G. Trypanosoma cruzi trans-sialidase operates through a covalent sialyl-enzyme intermediate: Tyrosine is the catalytic nucleophile. *J. Am. Chem. Soc.* **125,** 7532–7533 (2003).

254. Popper, Z. A. Evolution and diversity of green plant cell walls. *Curr. Opin. Plant Biol.* **11,** 286–292 (2008).
255. Attia, M. A. & Brumer, H. Recent structural insights into the enzymology of the ubiquitous plant cell wall glycan xyloglucan. *Curr. Opin. Struct. Biol.* **40,** 43–53 (2016).
256. Eklöf, J. M., Ruda, M. C. & Brumer, H. in *Methods Enzymol.* (ed. Gilbert, H. J.) **510,** 97–120 (Elsevier Inc., 2012).
257. Gottlieb, H. E., Kotlyar, V. & Nudelman, A. NMR chemical shifts of common laboratory solvents as trace impurities. *J. Org. Chem.* **62,** 7512–7515 (1997).
258. Ibatullin, F. M., Baumann, M. J., Greffe, L. & Brumer, H. Kinetic Analyses of Retaining endo-(Xylo)glucanases from Plant and Microbial Sources Using New Chromogenic Xylogluco-Oligosaccharide Aryl Glycosides. *Biochemistry* **74,** 7762–7769 (2008).
259. Xu, X., Tan, Q. & Hayashi, M. Versatile and mild synthesis of Di- and trisaccharidic 2-enopyranosyl cyanides by cyanation of per-O-acetylglycals with trimethylsilyl cyanide catalyzed by palladium(II) acetate. *Synthesis (Stuttg).* 770–776 (2008).
260. Vincent, S. P., Burkart, M. D., Tsai, C. Y., Zhang, Z. & Wong, C. H. Electrophilic fluorination-nucleophilic addition reaction mediated by Selectfluor: Mechanistic studies and new applications. *J. Org. Chem.* **64,** 5264–5279 (1999).
261. Albert, M., Karl, D. & Ortner, J. A novel direct route to 2-deoxy-2-fluoro-aldoses and their corresponding derivatives. *Tetrahedron* **54,** 4839–4848 (1998).
262. Ortner, J., Albert, M., Weber, H. & Dax, K. Studies on the Reaction of D-Glucal and its Derivatives with 1-Chloromethyl-4-Fluoro-1,4-Diazoniabicyclo[2.2.2]Octane Salts. *J. Carbohydr. Chem.* **18,** 297–316 (1999).
263. Attia, M., Stepper, J., Davies, G. J. & Brumer, H. Functional and structural characterization of a potent GH74 endo-xyloglucanase from the soil saprophyte Cellvibrio japonicus unravels the first step of xyloglucan degradation. *FEBS J.* **283,** 1701–1719 (2016).
264. Sundqvist, G., Stenvall, M., Berglund, H., Ottosson, J. & Brumer, H. A general , robust method for the quality control of intact proteins using LC – ESI-MS. *J. Chromatogr. B* **852,** 188–194 (2007).
265. Kabsch, W. XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 125–132 (2010).
266. Waterman, D. G., Winter, G., Gildea, R. J., Parkhurst, J. M., Brewster, A. S., Sauter, N. K.

& Evans, G. Diffraction-geometry refinement in the DIALS framework. *Acta Crystallogr. Sect. D Struct. Biol.* **72,** 558–575 (2016).

267. Potterton, L., Agirre, J., Ballard, C., Cowtan, K., Dodson, E., Evans, P. R., Jenkins, H. T., Keegan, R., Krissinel, E., Stevenson, K., Lebedev, A., McNicholas, S. J., Nicholls, R. A., Noble, M., Pannu, N. S., Roth, C., Sheldrick, G., Skubak, P., Turkenburg, J., Uski, V., Von Delft, F., Waterman, D., Wilson, K., Winn, M. & Wojdyr, M. CCP4i2: The new graphical user interface to the CCP 4 program suite. *Acta Crystallogr. Sect. D Struct. Biol.* **74,** 68–84 (2018).
268. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69,** 1204–1214 (2013).
269. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67,** 355–367 (2011).
270. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 486–501 (2010).
271. Agirre, J., Iglesias-Fernández, J., Rovira, C., Davies, G. J., Wilson, K. S. & Cowtan, K. D. Privateer: Software for the conformational validation of carbohydrate structures. *Nat. Struct. Mol. Biol.* **22,** 833–834 (2015).
272. Dax, K., Albert, M., Ortner, J. & Paul, B. J. Synthesis of deoxyfluoro sugars from carbohydrate precursors. *Carbohydr. Res.* **327,** 47–86 (2000).
273. Chen, H. M. & Withers, S. G. Facile synthesis of 2,4-dinitrophenyl α-D-glycopyranosides as chromogenic substrates for α-glycosidases. *ChemBioChem* **8,** 719–722 (2007).
274. Koeners, H. J., de Kok, A. J., Romers, C. & van Boom, J. H. The use of the 2,4-dinitrophenyl group in sugar chemistry re-examined. *Recl. des Trav. Chim. des Pays-Bas* **99,** 355–362 (1980).
275. Sinnott, M. L. Catalytic Mechanisms of Enzymic Glycosyl Transfer. *Chem. Rev.* **90,** 1171–1202 (1990).
276. Gloster, T. M. & Davies, G. J. Glycosidase inhibition: assessing mimicry of the transition state. *Org. Biomol. Chem.* **8,** 305–20 (2010).
277. Borges de Melo, E., da Silveira Gomes, A. & Carvalho, I. α- and β-Glucosidase inhibitors:

chemical structure and biological activity. *Tetrahedron* **62,** 10277–10302 (2006).

278. Wu, L., Armstrong, Z., Schröder, S. P., de Boer, C., Artola, M., Aerts, J. M., Overkleeft, H. S. & Davies, G. J. An overview of activity-based probes for glycosidases. *Curr. Opin. Chem. Biol.* **53,** 25–36 (2019).
279. Beenakker, T. J. M., Wander, D. P. A., Offen, W. A., Artola, M., Raich, L., Ferraz, M. J., Li, K. Y., Houben, J. H. P. M., Van Rijssel, E. R., Hansen, T., Van Der Marel, G. A., Codée, J. D. C., Aerts, J. M. F. G., Rovira, C., Davies, G. J. & Overkleeft, H. S. Carba-cyclophellitols Are Neutral Retaining-Glucosidase Inhibitors. *J. Am. Chem. Soc.* **139,** 6534–6537 (2017).
280. Hoj, P. B., Rodriguez, E. B., Iser, J. R., Stick, R. V. & Stone, B. A. Active site-directed inhibition by optically pure epoxyalkyl cellobiosides reveals differences in active site geometry of two 1,3-1,4-β-D-glucan 4-glucanohydrolases: The importance of epoxide stereochemistry for enzyme inactivation. *J. Biol. Chem.* **266,** 11628–11631 (1991).
281. Jain, N., Attia, M. A., Offen, W. A., Davies, G. J. & Brumer, H. Synthesis and application of a highly branched, mechanism-based 2-deoxy-2-fluoro-oligosaccharide inhibitor of endo-xyloglucanases. *Org. Biomol. Chem.* **16,** 8732–8741 (2018).
282. Salmeán, A. A., Duffieux, D., Harholt, J., Qin, F., Michel, G., Czjzek, M., Willats, W. G. T. & Hervé, C. Insoluble (1 → 3), (1 → 4)-β-D-glucan is a component of cell walls in brown algae (Phaeophyceae) and is masked by alginates in tissues. *Sci. Rep.* **7,** 1–11 (2017).
283. Burton, R. A. & Fincher, G. B. (1,3;1,4)-β-D-glucans in cell walls of the poaceae, lower plants, and fungi: A tale of two linkages. *Mol. Plant* **2,** 873–882 (2009).
284. Fincher, G. B. Morphology and Chemical Composition of Barley Endosperm Cell Walls. *J. Inst. Brew.* **81,** 116–122 (1975).
285. Collins, H. M., Burton, R. A., Topping, D. L., Liao, M. L., Bacic, A. & Fincher, G. B. Variability in fine structures of noncellulosic cell wall polysaccharides from cereal grains: Potential importance in human health and nutrition. *Cereal Chem.* **87,** 272–282 (2010).
286. Wood, P. J. Cereal β-glucans in diet and health. *J. Cereal Sci.* **46,** 230–238 (2007).
287. McNeil, N. I. The contribution of the large intestine to energy supplies in man. *Am. J. Clin. Nutr.* **39,** 338–342 (1984).
288. Hrmova, M. & Fincher, G. B. Structure-function relationships of β-D-glucan endo- and

exohydrolases from higher plants. *Plant Mol. Biol.* **47,** 73–91 (2001).

289. Buchanan, M., Burton, R. A., Dhugga, K. S., Rafalski, A. J., Tingey, S. V, Shirley, N. J. & Fincher, G. B. Endo-(1,4)-β-Glucanase gene families in the grasses: temporal and spatial Co-transcription of orthologous genes. *BMC Plant Biol.* **12,** 1 (2012).
290. Yoshida, K. & Komae, K. A rice family 9 glycoside hydrolase isozyme with broad substrate specificity for hemicelluloses in type II cell walls. *Plant Cell Physiol.* **47,** 1541–1554 (2006).
291. Krah, M., Misselwitz, R., Politz, O., Thomsen, K. K., Welfle, H. & Borriss, R. The laminarinase from thermophilic eubacterium Rhodothermus marinus conformation, stability, and identification of active site carboxylic residues by site-directed mutagenesis. *Eur. J. Biochem.* **257,** 101–111 (1998).
292. Planas, A. Bacterial 1,3-1,4-β-glucanases: Structure, function and protein engineering. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1543,** 361–382 (2000).
293. Foley, M. H., Déjean, G., Hemsworth, G. R., Davies, G. J., Brumer, H. & Koropatkin, N. M. A Cell-Surface GH9 Endo-Glucanase Coordinates with Surface Glycan-Binding Proteins to Mediate Xyloglucan Uptake in the Gut Symbiont Bacteroides ovatus. *J. Mol. Biol.* **431,** 981–995 (2019).
294. Tamoi, M., Kurotaki, H. & Fukamizo, T. β-1,4-Glucanase-like protein from the cyanobacterium Synechocystis PCC6803 is a β-1,3-1,4-glucanase and functions in salt stress tolerance. *Biochem. J.* **405,** 137–146 (2007).
295. Bauer, M. W., Driskill, L. E. & Kelly, R. M. An Endoglucanase, EglA, from the Hyperthermophilic Archaeon Pyrococcus. *J. Bacteriol.* **181,** 284 (1999).
296. Money, V. A., Cartmell, A., Guerreiro, C. I. P. D., Ducros, V. M. A., Fontes, C. M. G. A., Gilbert, H. J. & Gideon J. Davies. Probing the β-1,3:1,4 glucanase, CtLic26A, with a thio-oligosaccharide and enzyme variants. *Org. Biomol. Chem.* **6,** 851–853 (2008).
297. Taylor, E. J., Goyal, A., Guerreiro, C. I. P. D., Prates, J. A. M., Money, V. A., Ferry, N., Morland, C., Planas, A., Macdonald, J. A., Stick, R. V., Gilbert, H. J., Fontes, C. M. G. A. & Davies, G. J. How family 26 glycoside hydrolases orchestrate catalysis on different polysaccharides: Structure and activity of a Clostridium thermocellum lichenase, CtLic26A. *J. Biol. Chem.* **280,** 32761–32767 (2005).
298. Chaari, F. & Chaabouni, S. E. Fungal β-1,3-1,4-glucanases: production, proprieties and

biotechnological applications. *J. Sci. Food Agric.* **99,** 2657–2664 (2019).

299. El Khoury, D., Cuda, C., Luhovyy, B. L. & Anderson, G. H. Beta glucan: Health benefits in obesity and metabolic syndrome. *J. Nutr. Metab.* **2012,** 1–28 (2012).
300. Gilleran, C. T., Hernon, A. T., Murray, P. G. & Tuohy, M. G. Induction of enzyme cocktails by low cost carbon sources for production of monosaccharide-rich syrups from plant materials. *BioResources* **5,** 634–649 (2010).
301. Von Wettstein, D., Mikhaylenko, G., Froseth, J. A. & Kannangara, C. G. Improved barley broiler feed with transgenic malt containing heat-stable (1,3-1,4)-β-glucanase. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 13512–13517 (2000).
302. Celestino, K. R. S., Cunha, R. B. & Felix, C. R. Characterization of a β-glucanase produced by Rhizopus microsporus var. microsporus, and its potential for application in the brewing industry. *BMC Biochem.* **7,** (2006).
303. Planas, A., Abel, M., Millet, O., Palasi, J., Pallares, C. & Viladot, J.-L. Synthesis of aryl 3-O-β-cellobiosyl-β-d- glucopyranosides for reactivity studies of 1,3- 1,4- β-glucanases. *Carbohydr. Res.* **310,** 53–64 (1998).
304. Malet, C., Vallés, J., Bou, J. & Planas, A. A specific chromophoric substrate for activity assays of 1,3-1,4-β-D-glucan 4-glucanohydrolases. *J. Biotechnol.* **48,** 209–219 (1996).
305. Faijes, M., Pérez, X., Pérez, O. & Planas, A. Glycosynthase Activity of Bacillus licheniformis 1,3-1,4-β-Glucanase Mutants: Specificity, Kinetics, and Mechanism. *Biochemistry* **42,** 13304–13318 (2003).
306. Moreau, V., Viladot, J. L., Samain, E., Planas, A. & Driguez, H. Design and chemoenzymatic synthesis of thiooligosaccharide inhibitors of 1,3:1,4-β-D-glucanases. *Bioorg. Med. Chem.* **4,** 1849–1855 (1996).
307. Keitel, T., Simon, O., Borriss, R. & Heinemann, U. Molecular and active-site structure of a Bacillus 1,3-1,4-β-glucanase. *Proc. Natl. Acad. Sci.* **90,** 5287–5291 (1993).
308. Hoj, P. B., Condron, R., Traeger, J. C., McAuliffe, J. C. & Stone, B. A. Identification of glutamic acid 105 at the active site of Bacillus amyloliquefaciens 1,3-1,4-β-D-glucan 4-glucanohydrolase using epoxide-based inhibitors. *J. Biol. Chem.* **267,** 25059–25066 (1992).
309. Hoj, P. B., Rodriguez, E. B., Stick, R. V. & Stone, B. A. Differences in active site structure in a family of β-glucan endohydrolases deduced from the kinetics of inactivation

by epoxyalkyl β-oligoglucosides. *J. Biol. Chem.* **264,** 4939–4947 (1989).

310. McGregor, N., Yin, V., Tung, C. C., Van Petegem, F. & Brumer, H. Crystallographic insight into the evolutionary origins of xyloglucan endotransglycosylases and endohydrolases. *Plant J.* **89,** 651–670 (2017).
311. Abel, M., Segade, A. & Planas, A. Synthesis of an aryl 2-deoxy-β-glycosyl tetrasaccharide to probe retaining endo-glycosidase mechanism. *Tetrahedron Asymmetry* **20,** 847–850 (2009).
312. Malet, C., Jimenez-Barbero, J., Bernabe, M., Brosa, C. & Planas, A. Stereochemical course and structure of the products of the enzymic action of endo-1,3-1-4-β-D-glucan 4-glucanohydrolase from Bacillus licheniformis. *Biochem. J.* **758,** 753–758 (1993).
313. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).
314. Cowtan, K. Recent developments in classical density modification. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 470–478 (2010).
315. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60,** 2126–2132 (2004).
316. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., Snoeyink, J., Richardson, J. S. & Richardson, D. C. MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* **35,** 375–383 (2007).
317. Yariv, J., Wilson, K. J., Hildesheim, J. & Blumberg, S. Labelling of the Active Site of β-galactosidase by N-bromoacetyl β-D-galactopyranosylamine. *FEBS Lett.* **15,** 24–26 (1971).
318. Xue, J., Guo, M., Gu, G. & Guo, Z. A Facile Synthesis of Nγ-Glycosyl Asparagine Conjugates and Short N-Linked Glycopeptides. *J. Carbohydr. Chem.* **31,** 105–113 (2012).
319. Laroche, C. & Michaud, P. New Developments and Prospective Applications for β (1,3) Glucans. *Recent Pat. Biotechnol.* **1,** 59–73 (2008).
320. Kadam, S. U., Tiwari, B. K. & O'Donnell, C. P. Extraction, structure and biofunctional activities of laminarin from brown algae. *Int. J. Food Sci. Technol.* **50,** 24–31 (2015).
321. O'Sullivan, L., Murphy, B., McLoughlin, P., Duggan, P., Lawlor, P. G., Hughes, H. & Gardiner, G. E. Prebiotics from marine macroalgae for human and animal health

applications. *Mar. Drugs* **8,** 2038–2064 (2010).

322. Sakamoto, Y., Nakade, K. & Konno, N. Endo-β-1,3-Glucanase GLU1, from the fruiting body of Lentinula edodes, belongs to a new glycoside hydrolase family. *Appl. Environ. Microbiol.* **77,** 8350–8354 (2011).
323. Hartl, L., Gastebois, A., Aimanianda, V. & Latgé, J. P. Characterization of the GPI-anchored endo β-1,3-glucanase Eng2 of Aspergillus fumigatus. *Fungal Genet. Biol.* **48,** 185–191 (2011).
324. Zamora, M. G. M., Bournonville, C. G., Castagnaro, A. P. & Ricci, J. C. D. Identification and characterisation of a novel class i endo -β 1,3-glucanase regulated by salicylic acid, ethylene and fungal pathogens in strawberry. *Funct. Plant Biol.* **39,** 412–420 (2012).
325. Fuchs, K. P., Zverlov, V. V., Velikodvorskaya, G. A., Lottspeich, F. & Schwarz, W. H. Lic16A of Clostridium thermocellum, a non-cellulosomal, highly complex endo-β-1,3-glucanase bound to the outer cell surface. *Microbiology* **149,** 1021–1031 (2003).
326. Ilari, A., Fiorillo, A., Angelaccio, S., Florio, R., Chiaraluce, R., Van Der Oost, J. & Consalvi, V. Crystal structure of a family 16 endoglucanase from the hyperthermophile Pyrococcus furiosus - Structural basis of substrate recognition. *FEBS J.* **276,** 1048–1058 (2009).
327. Sova, V. V & Branch, S. The Distribution Of Laminarinases In Marine Invertebrates. *Comp. Biochem. Physiol.* **32,** 459–464 (1970).
328. Kottwitz, B., Maurer, K. Detergent containing glucanase. (2002). at <http://www.freepatentsonline.com/6417152.html>
329. Barreteau, H., Delattre, C. & Michaud, P. Production of oligosaccharides as promising new food additive generation. *Food Technol. Biotechnol.* **44,** 323–333 (2006).
330. Marca, E., Valero-Gonzalez, J., Delso, I., Tejero, T., Hurtado-Guerrero, R. & Merino, P. Synthesis of O- and C-glycosides derived from β-(1,3)-D-glucans. *Carbohydr. Res.* **382,** 9–18 (2013).
331. Kleine, H. P., Weinberg, D. V., Kaufman, R. J. & Sidhu, R. S. Phase-transfer-catalyzed synthesis of 2,3,4,6-tetra-*O*-acetyl-β-d-galactopyranosides. *Carbohydr. Res.* **142,** 333–337 (1985).
332. Adam, M. J., Pate, B. D., Nesser, J. R. & Hall, L. D. A rapid, stereoselective synthesis of fluorinated carbohydrates: Addition of acetyl hypofluorite to vinyl ether derivatives of

sugars. *Carbohydr. Res.* **124,** 215–224 (1983).

333. Adamson, J. & Foster, A. B. Fluorinated Carbohydrates: 2-deoxy-2-fluoro-d-glucose and 2-deoxy-2-fluoro-d-mannose. *Carbohydr. Res.* **15,** 351–359 (1970).
334. Korytnyk, W., Valentekovic-Horvath, S. & Petrie, C. R. A convenient synthesis of 1,2-difluoro-1, 2-dideoxyhexoses using xenon fluoride. *Tetrahedron* **38,** 2547–2550 (1982).
335. Ido, T., Wan, C. N., Fowler, J. S. & Wolf, A. P. Fluorination with F2. A Convenient Synthesis of 2-Deoxy-2-fluoro-D-glucose. *J. Org. Chem.* **42,** 2341–2342 (1977).
336. Burkart, M. D., Zhang, Z., Hung, S. C. & Wong, C. H. A new method for the synthesis of fluoro-carbohydrates and glycosides using Selectfluor. *J. Am. Chem. Soc.* **119,** 11743–11746 (1997).
337. Abdul-Ghani, M., Banks, E. R., Besheesh, M. K., Sharif, I. & Syvret, R. G. N-Halogeno compounds. Part 14. 'Transfer fluorination' of quinuclidine using F-TEDA-BFc (Selectfluor^{TM} reagent): laboratory synthesis of N-fluoroquinuclidinium salts not requiring the use of elemental fluorine. *J. Fluor. Chem.* **73,** 255–257 (1995).
338. Benito, D., Matheu, M. I., Morère, A., Díaz, Y. & Castillón, S. Towards the preparation of 2″-deoxy-2″-fluoro-adenophostin A. Study of the glycosylation reaction. *Tetrahedron* **64,** 10906–10911 (2008).
339. Dohi, H., Périon, R., Durka, M., Bosco, M., Roué, Y., Moreau, F., Grizot, S., Ducruix, A., Escaich, S. & Vincent, S. P. Stereoselective glycal fluorophosphorylation: Synthesis of ADP-2-fluoroheptose, an inhibitor of the LPS Biosynthesis. *Chem. - A Eur. J.* **14,** 9530–9539 (2008).
340. Wolf, S., Berrio, R. M. & Meier, C. Synthesis of nonnatural nucleoside diphosphate sugars. *Eur. J. Org. Chem.* 6304–6313 (2011).
341. Vincent, S. P., Burkart, M. D., Tsai, C. Y., Zhang, Z. & Wong, C. H. Electrophilic fluorination-nucleophilic addition reaction mediated by Selectfluor: Mechanistic studies and new applications. *J. Org. Chem.* **64,** 5264–5279 (1999).
342. Grondin, J. M., Tamura, K., Déjean, G., Abbott, D. W. & Brumer, H. Polysaccharide utilization loci: Fueling microbial communities. *J. Bacteriol.* **199,** 1–15 (2017).
343. Ziser, L., Setyawati, I. & Withers, S. G. Syntheses and testing of substrates and mechanism-based inactivators for xylanases. *Carbohydr. Res.* **274,** 137–153 (1995).
344. Shaikh, F. A., Müllegger, J., He, S. & Withers, S. G. Identification of the catalytic

nucleophile in Family 42 β-galactosidases by intermediate trapping and peptide mapping: YesZ from Bacillus subtilis. *FEBS Lett.* **581,** 2441–2446 (2007).

345. Namchuk, M. N. & Withers, S. G. Mechanism of Agrobacterium β-Glucosidase: Kinetic Analysis of the Role of Noncovalent Enzyme/Substrate Interactions. *Biochemistry* **34,** 16194–16202 (1995).
346. Stubbs, K. A. Activity-based proteomics probes for carbohydrate-processing enzymes: Current trends and future outlook. *Carbohydr. Res.* **390,** 9–19 (2014).
347. Liu, Y., Fredrickson, J. K., Sadler, N. C., Nandhikonda, P., Smith, R. D. & Wright, A. T. Advancing understanding of microbial bioenergy conversion processes by activity-based protein profiling. *Biotechnol. Biofuels* **8,** 156 (2015).
348. Fonović, M. & Bogyo, M. Activity-based probes as a tool for functional proteomic analysis of proteases. *Expert Rev. Proteomics* **5,** 721–30 (2008).
349. Saghatelian, A. & Cravatt, B. F. Assignment of protein function in the postgenomic era. *Nat. Chem. Biol.* **1,** 130–142 (2005).
350. Willems, L. I., Overkleeft, H. S. & Van Kasteren, S. I. Current developments in activity-based protein profiling. *Bioconjug. Chem.* **25,** 1181–1191 (2014).
351. Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-Based Protein Profiling: From Enzyme Chemistry to Proteomic Chemistry. *Annu. Rev. Biochem.* **77,** 383–414 (2008).
352. Sieber, S. A. in *Bioelectrochemistry Bioenerg.* (eds. Houk, K. N., Hunter, C. A., Krische, M. J., Lehn, J. M., Ley, S. V., Olivucci, M., Thiem, J., Ventury, M., Vogel, P., Wong, C. H., Wong., H. & Yamamoto, H.) **7,** 401–402 (Springer, 1980).
353. Kohl, F., Schmitz, J., Furtmann, N., Schulz-Fincke, A.-C., Mertens, M. D., Küppers, J., Benkhoff, M., Tobiasch, E., Bartz, U., Bajorath, J., Stirnberg, M. & Gütschow, M. Design, characterization and cellular uptake studies of fluorescence-labeled prototypic cathepsin inhibitors. *Org. Biomol. Chem.* **13,** 10310–10323 (2015).
354. Tallman, K. R. & Beatty, K. E. Far-Red Fluorogenic Probes for Esterase and Lipase Detection. 70–75 (2015).
355. Stubbs, K. A. & Vocadlo, D. J. Affinity-Based Proteomics Probes Tools for Studying Carbohydrate-Processing Enzymes. *Aust. J. Chem.* **62,** 521–527 (2009).
356. Sumranjit, J. & Chung, S. J. Recent advances in target characterization and identification by photoaffinity probes. *Molecules* **18,** 10425–10451 (2013).

357. Wiedner, S. D., Anderson, L. N., Sadler, N. C., Chrisler, W. B., Vamsi, K., Smith, R. D. & Wright, A. T. Live Cell Organelle-Specific Activity-Based Protein Profiling. *Angew. Chem. Int. Ed.* **53,** 2919–2922 (2015).

358. Edgington, L. E., Verdoes, M., Ortega, A., Withana, N. P., Lee, J., Syed, S., Bachmann, M. H., Blum, G. & Bogyo, M. Functional imaging of legumain in cancer using a new quenched activity-based probe. *J. Am. Chem. Soc.* **135,** 174–182 (2013).

359. Malet, C. & Planas, A. From β-glucanase to β-glucansynthase: Glycosyl transfer to α-glycosyl fluorides catalyzed by a mutant endoglucanase lacking its catalytic nucleophile. *FEBS Lett.* **440,** 208–212 (1998).

360. Perugino, G., Trincone, A., Rossi, M. & Moracci, M. Oligosaccharide synthesis by glycosynthases. *Trends Biotechnol.* **22,** 31–37 (2004).

361. Armstrong, Z. & Withers, S. G. Synthesis of glycans and glycopolymers through engineered enzymes. *Biopolymers* **99,** 666–674 (2013).

362. Beeren, S. R., Christensen, C. E., Tanaka, H., Jensen, M. G., Donaldson, I. & Hindsgaul, O. Direct study of fluorescently-labelled barleyβ-glucan fate in an in vitro human colon digestion model. *Carbohydr. Polym.* **115,** 88–92 (2015).

363. Vissenberg, K., Martinez-Vilchez, I. M., Verbelen, J. P., Miller, J. G. & Fry, S. C. In vivo colocalization of xyloglucan endotransglycosylase activity and its donor substrate in the elongation zone of arabidopsis roots. *Plant Cell* **12,** 1229–1237 (2000).

364. Bourquin, V., Nishikubo, N., Abe, H., Brumer, H., Denman, S., Eklund, M., Christiernin, M., Teeri, T. T., Sundberg, B. & Mellerowicz, E. J. Xyloglucan Endotransglycosylases Have a Function during the Formation of Secondary Cell Walls of Vascular Tissues. *Plant Cell* **14,** 3073–3088 (2002).

365. Banasiak, A., Ibatullin, F., Brumer, H. & Mellerowicz, E. Glycoside Hydrolase Activities in Cell Walls of Sclerenchyma Cells in the Inflorescence Stems of Arabidopsis thaliana Visualized in Situ. *Plants* **3,** 513–525 (2014).

366. Ibatullin, F. M., Banasiak, A., Baumann, M. J., Greffe, L., Takahashi, J., Mellerowicz, E. J. & Brumer, H. A real-time fluorogenic assay for the visualization of glycoside hydrolase activity in planta. *Plant Physiol.* **151,** 1741–1750 (2009).

367. Kuniaki, T. Review Total synthesis and development of bioactive natural products. *Proc. Jpn. Acad., Ser. B* **84,** 87–106 (2008).

368. Atsumi, S., Umezawa, K., Iinuma, H., Naganawa, H., Nakamura, H., Iitaka, Y. & Takeuchi, T. Production, isolation and structure determination of a novel β-glucosidase inhibitor, cyclophellitol, from *Phellinus* sp. *J. Antibiot. (Tokyo)* **43,** 49–53 (1989).
369. Witte, M. D., Kallemeijn, W. W., Aten, J., Li, K., Strijland, A., Donker-koopman, W. E., Nieuwendijk, A. M. C. H. Van Den, Bleijlevens, B., Kramer, G., Florea, B. I., Hooibrink, B., Hollak, C. E. M., Ottenhoff, R., Boot, R. G., Marel, G. A. Van Der, Overkleeft, H. S. & Aerts, J. M. F. G. Ultrasensitive in situ visualization of active glucocerebrosidase molecules. *Nat. Chem. Biol.* **6,** 907–913 (2010).
370. Kallemeijn, W. W., Li, K. Y., Witte, M. D., Marques, A. R. A., Aten, J., Scheij, S., Jiang, J., Willems, L. I., Voorn-Brouwer, T. M., Van Roomen, C. P. A. A., Ottenhoff, R., Boot, R. G., Van Den Elst, H., Walvoort, M. T. C., Florea, B. I., Codée, J. D. C., Van Der Marel, G. A., Aerts, J. M. F. G. & Overkleeft, H. S. Novel activity-based probes for broad-spectrum profiling of retaining β-exoglucosidases in situ and in vivo. *Angew. Chem. Int. Ed.* **51,** 12529–12533 (2012).
371. Vocadlo, D. J. & Bertozzi, C. R. A strategy for functional proteomic analysis of glycosidase activity from cell lysates. *Angew. Chem. Int. Ed.* **43,** 5338–5342 (2004).
372. Tsukada, T., Igarashi, K., Yoshida, M. & Samejima, M. Molecular cloning and characterization of two intracellular β-glucosidases belonging to glycoside hydrolase family 1 from the basidiomycete *Phanerochaete chrysosporium*. *Appl. Microbiol. Biotechnol.* **73,** 807–814 (2006).
373. Armstrong, Z. Harnessing Natural Diversity for the Discovery of Glycoside Hydrolases and Design of New Glycosynthases. (2018).

Appendix A: Supporting Information for Chapter 2: Synthesis and Application of a Highly Branched, Mechanism-based 2-deoxy-2-fluoro-oligosaccharide Inhibitor of *Endo*-xyloglucanases

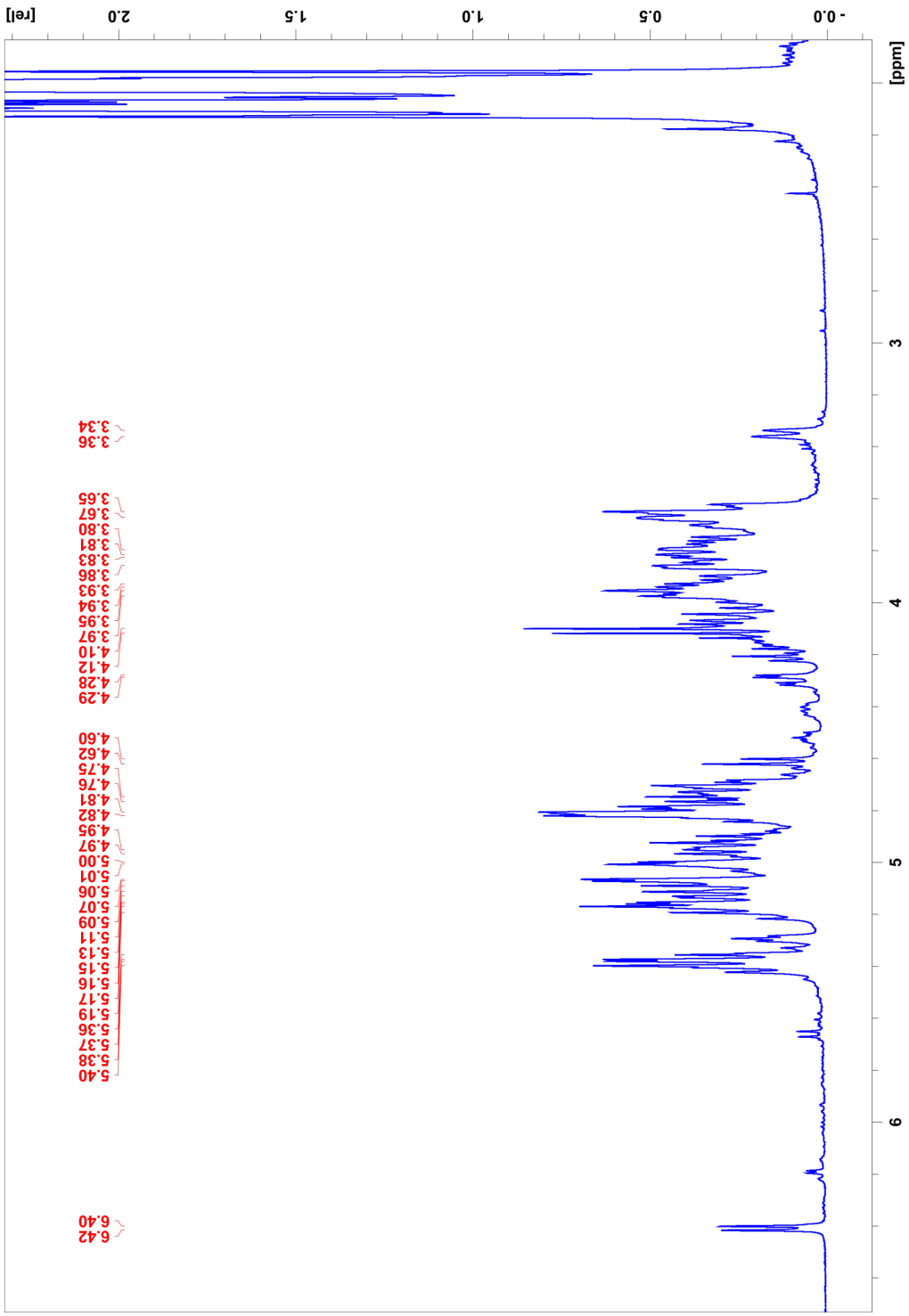


Figure A1: ^{1}H NMR spectrum of 2 (per-*O*-acetylated XXXG-Glycal)

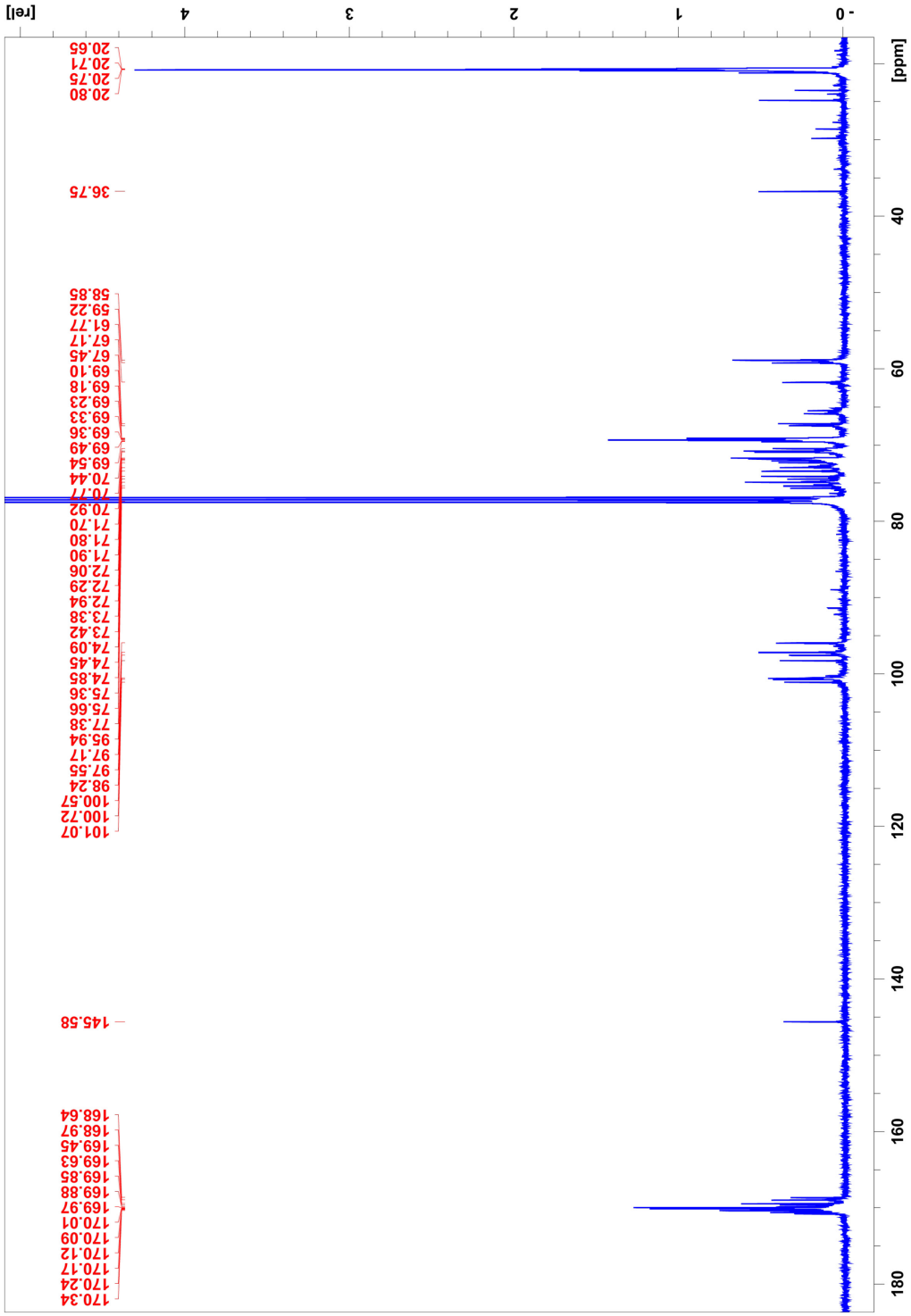


Figure A2: ^{13}C NMR spectrum of 2 (per-*O*-acetylated XXXG-Glycal)

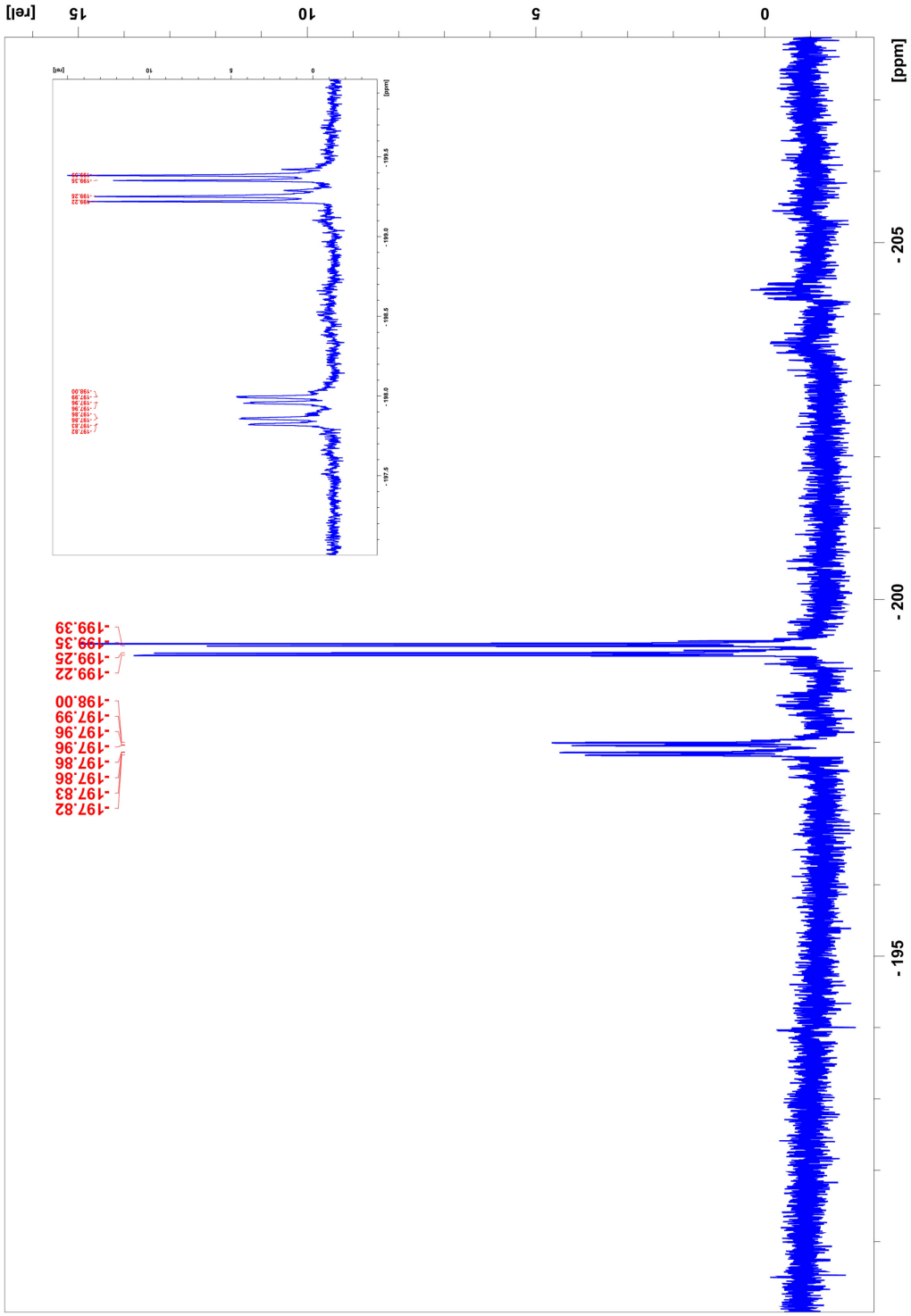


Figure A3: ^{19}F NMR spectrum of 3 (α/β hydroxyl 2-deoxy-2-fluoro per-*O*-acetylated XXXG)

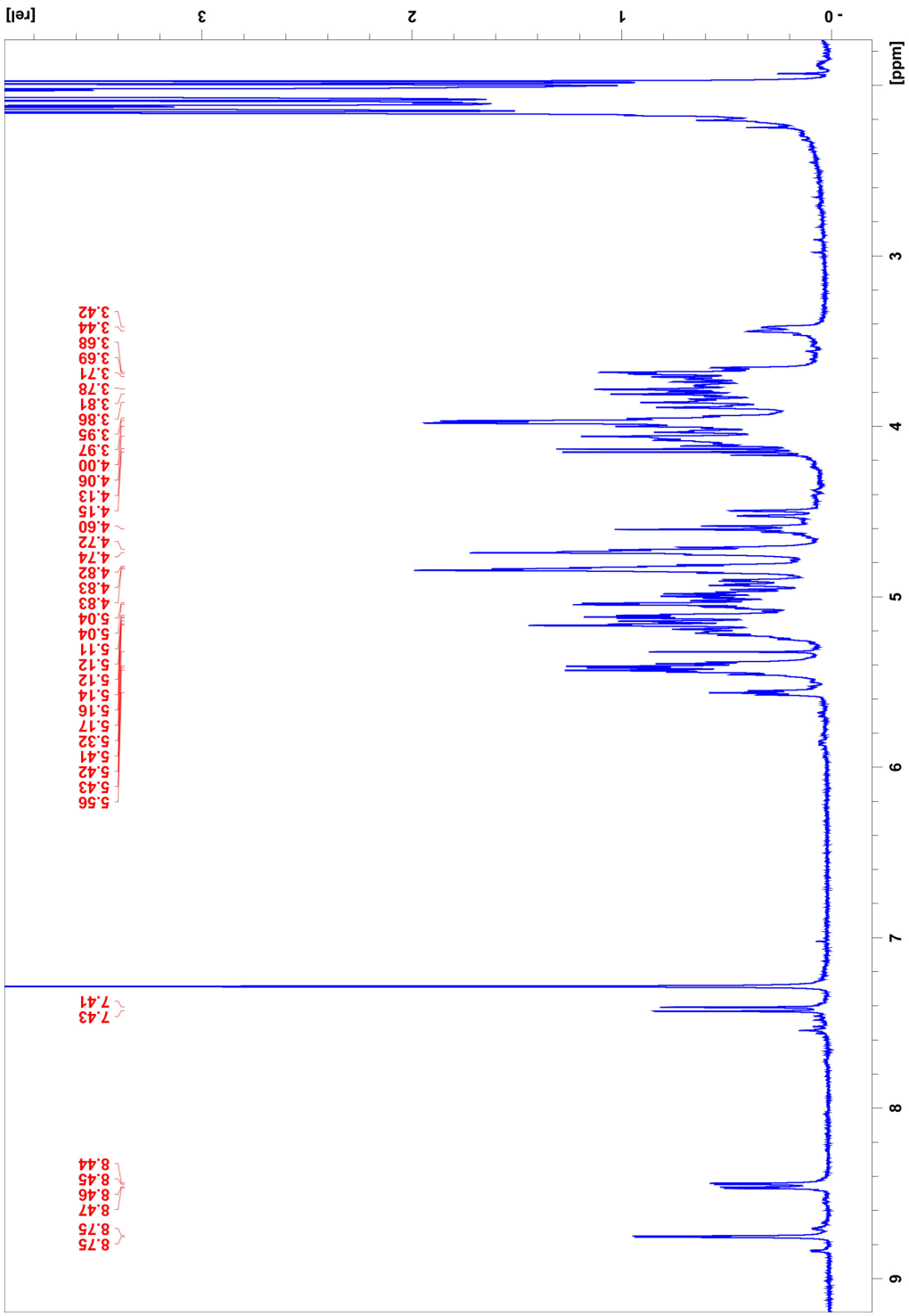


Figure A4: ^{1}H NMR spectrum of 4 (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated) XXXG)

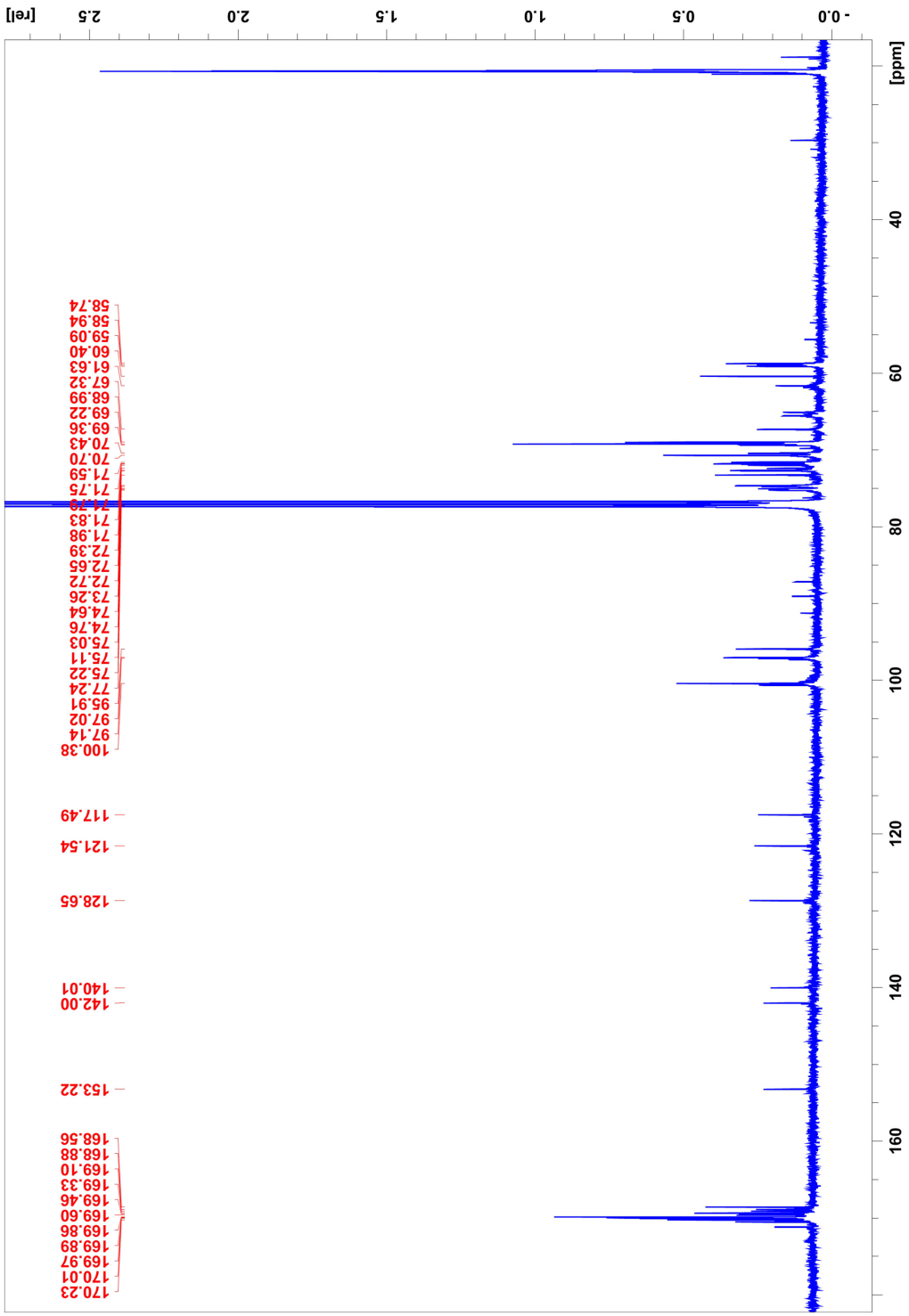


Figure A5: ^{13}C NMR spectrum of 4 (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated) XXXG)

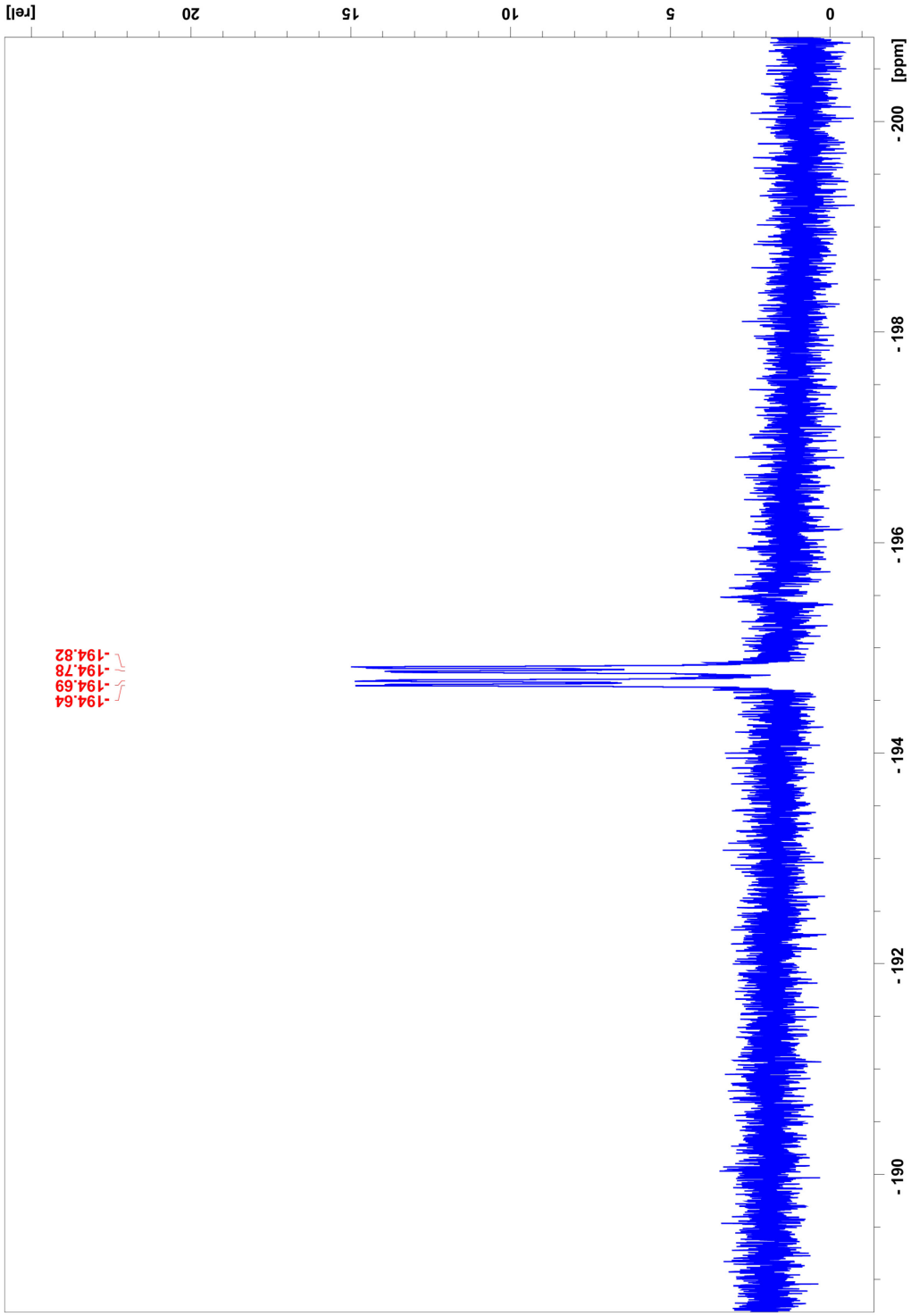


Figure A6: ^{19}F NMR spectrum of 4 (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated) XXXG)

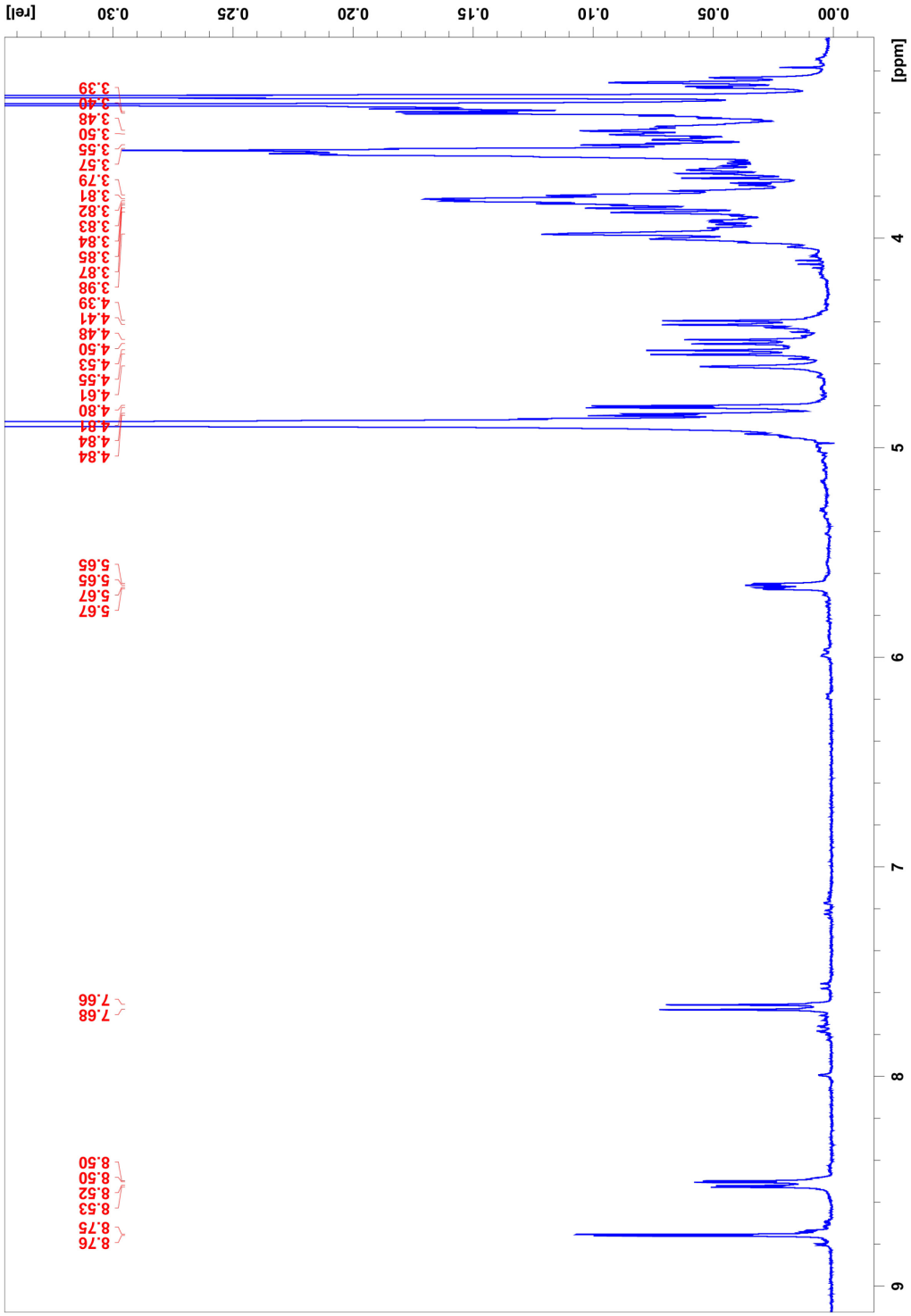


Figure A7: ^{1}H NMR spectrum of 5 (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro XXXG)

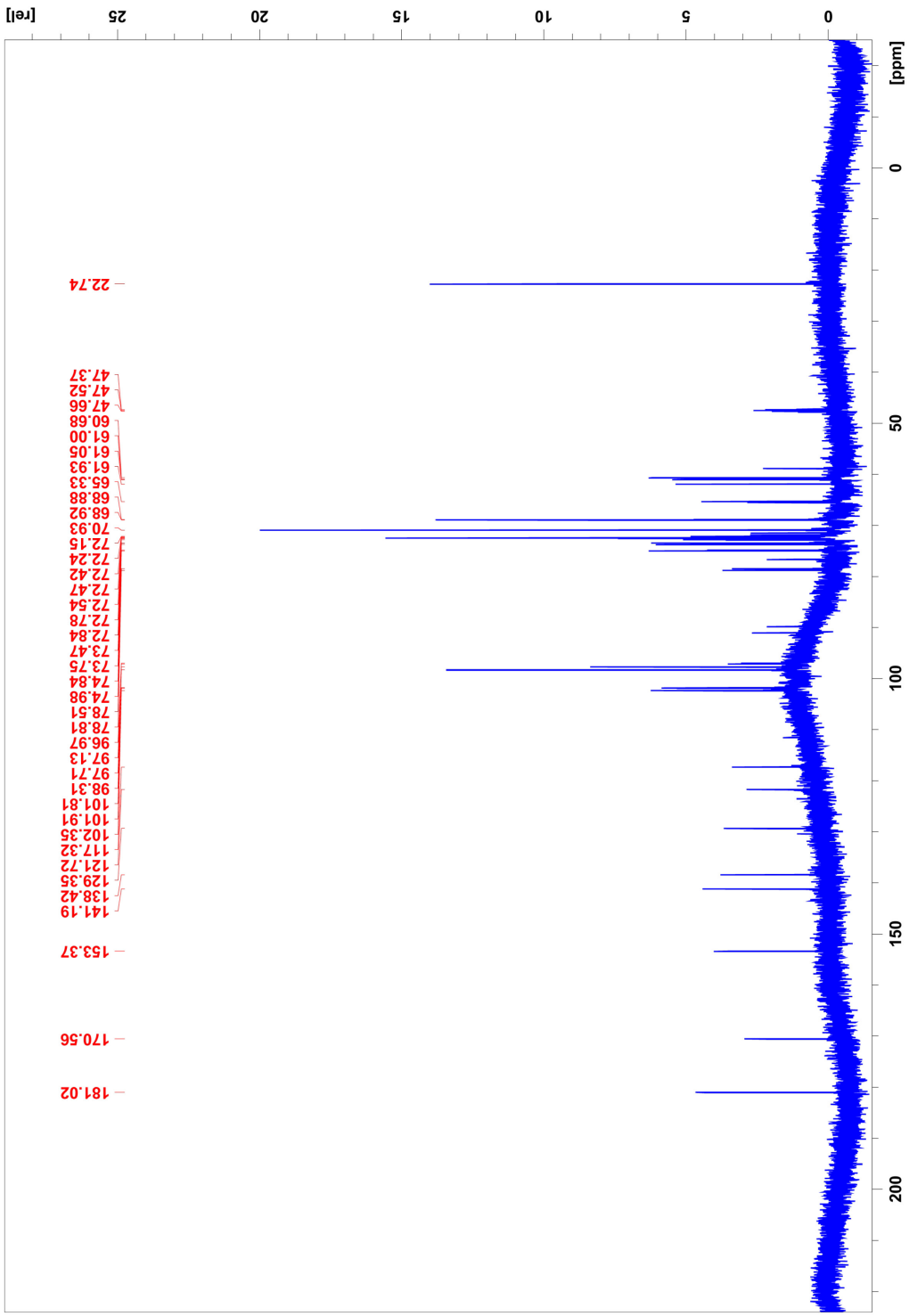


Figure A8: ^{13}C NMR spectrum of 5 (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro XXXG)

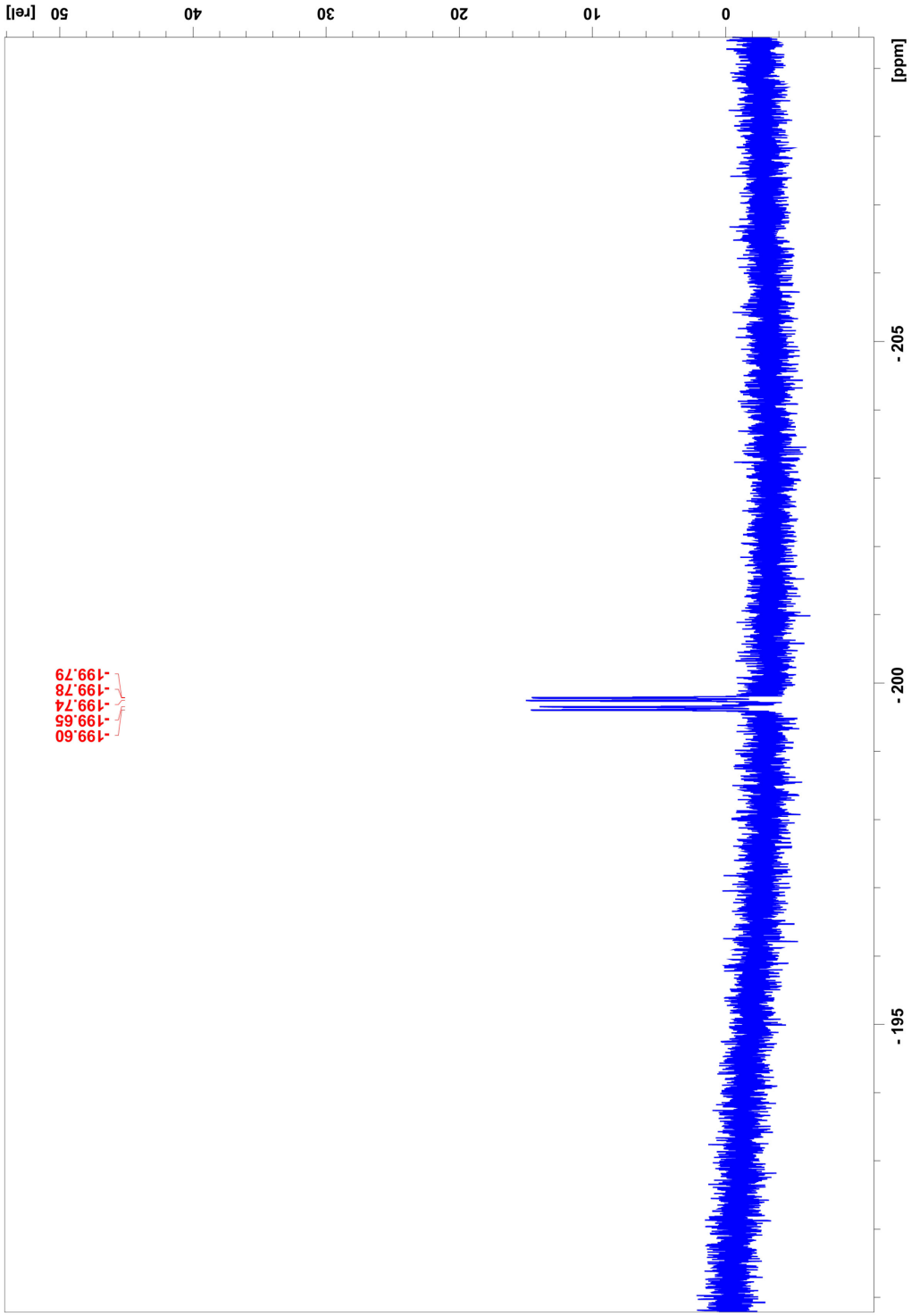


Figure A9: ^{19}F NMR spectrum of 5 (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro XXXG)

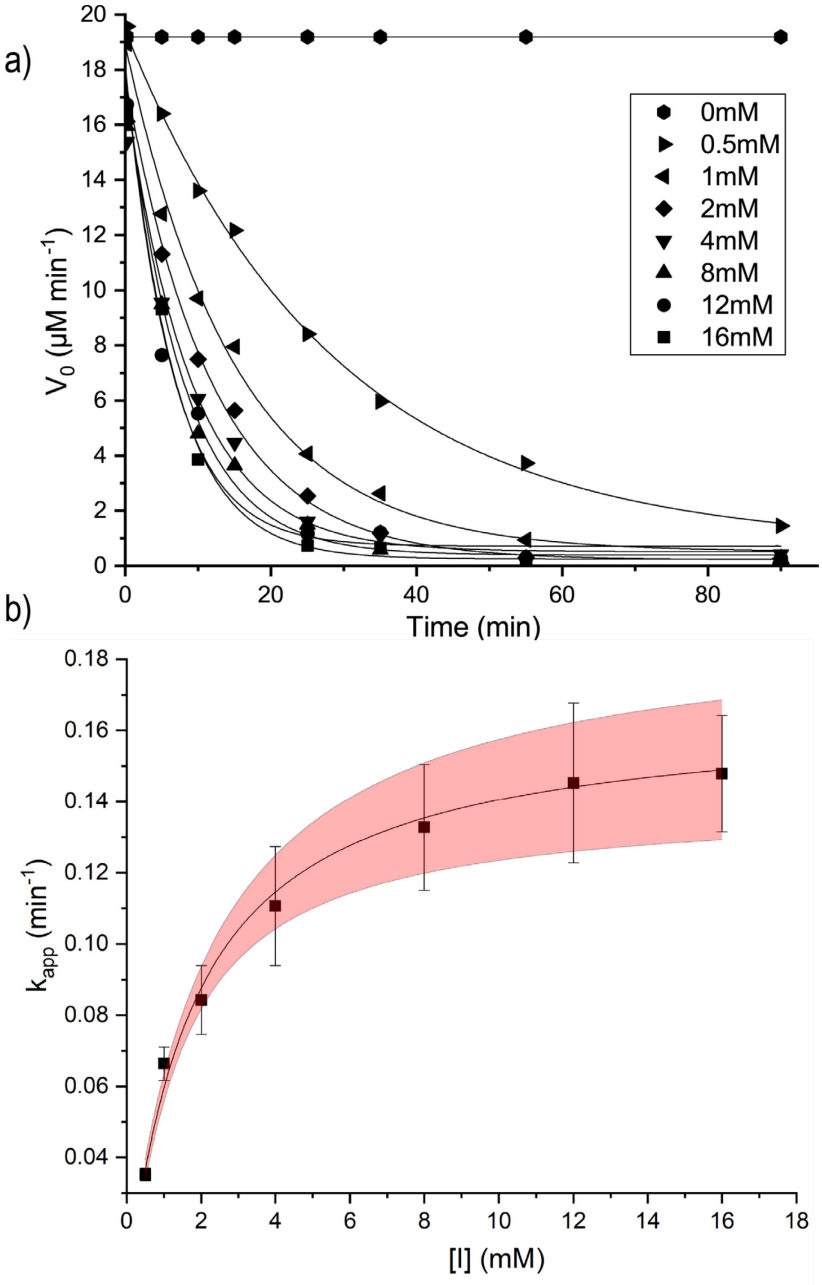


Figure A10: Inhibition of CjGH5D with XXXG $NHCOCH_2Br$. (a) Initial-rate enzyme activity over time (single determinations). (b) Pseudo-first-order rate constants (k_{app}) obtained from the fitted curves shown in a. Bars represent errors in k_{app} values from curve-fitting. The 95% confidence interval is indicated (pink band) for the fitted curve (solid line).

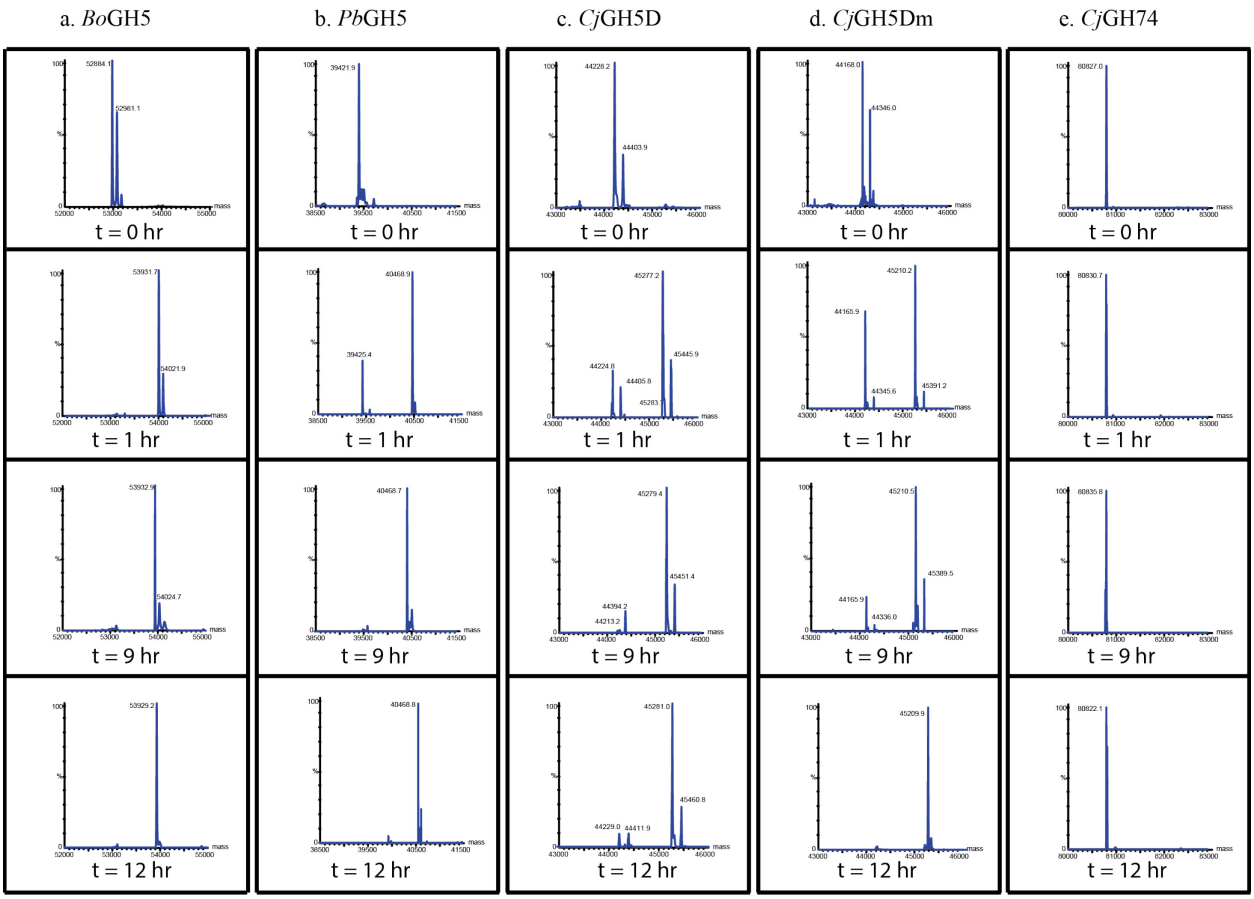


Figure A11: Intact protein mass spectrometry of endo-(xylo)glucanases incubated with 2.5 mM XXXG(2F)-β-DNP. a. *Bacteroides ovatus* GH5, b. *Prevotella bryantii* GH5, c. *Cellvibrio japonicus* GH5D, d. *Cellvibrio japonicus* GH5D(Glu255Ala) variant, e. *Cellvibrio japonicus* GH74.

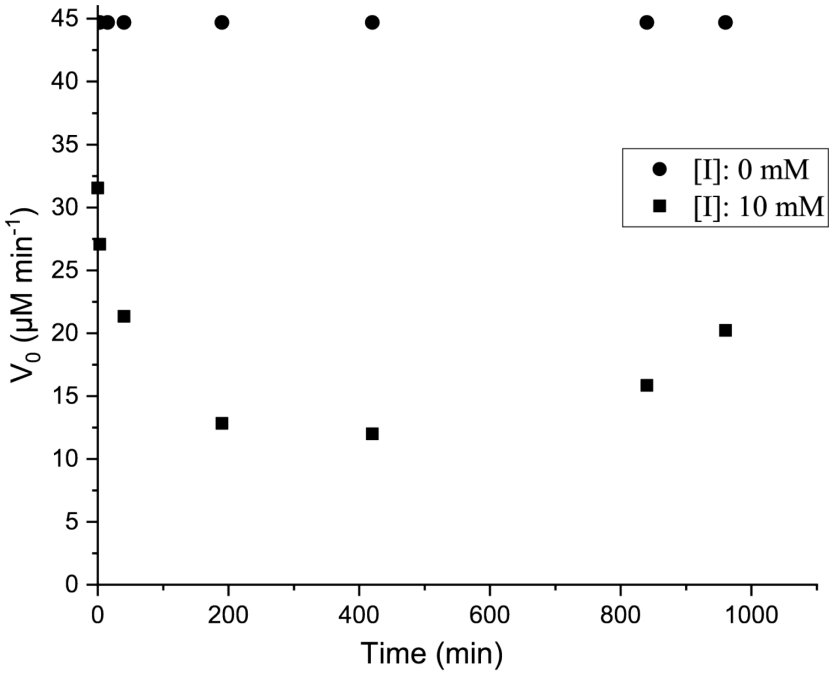


Figure A12: Inhibition of CjGH5D with 10 mM XXXG(2F)-β-DNP for up to 18 h

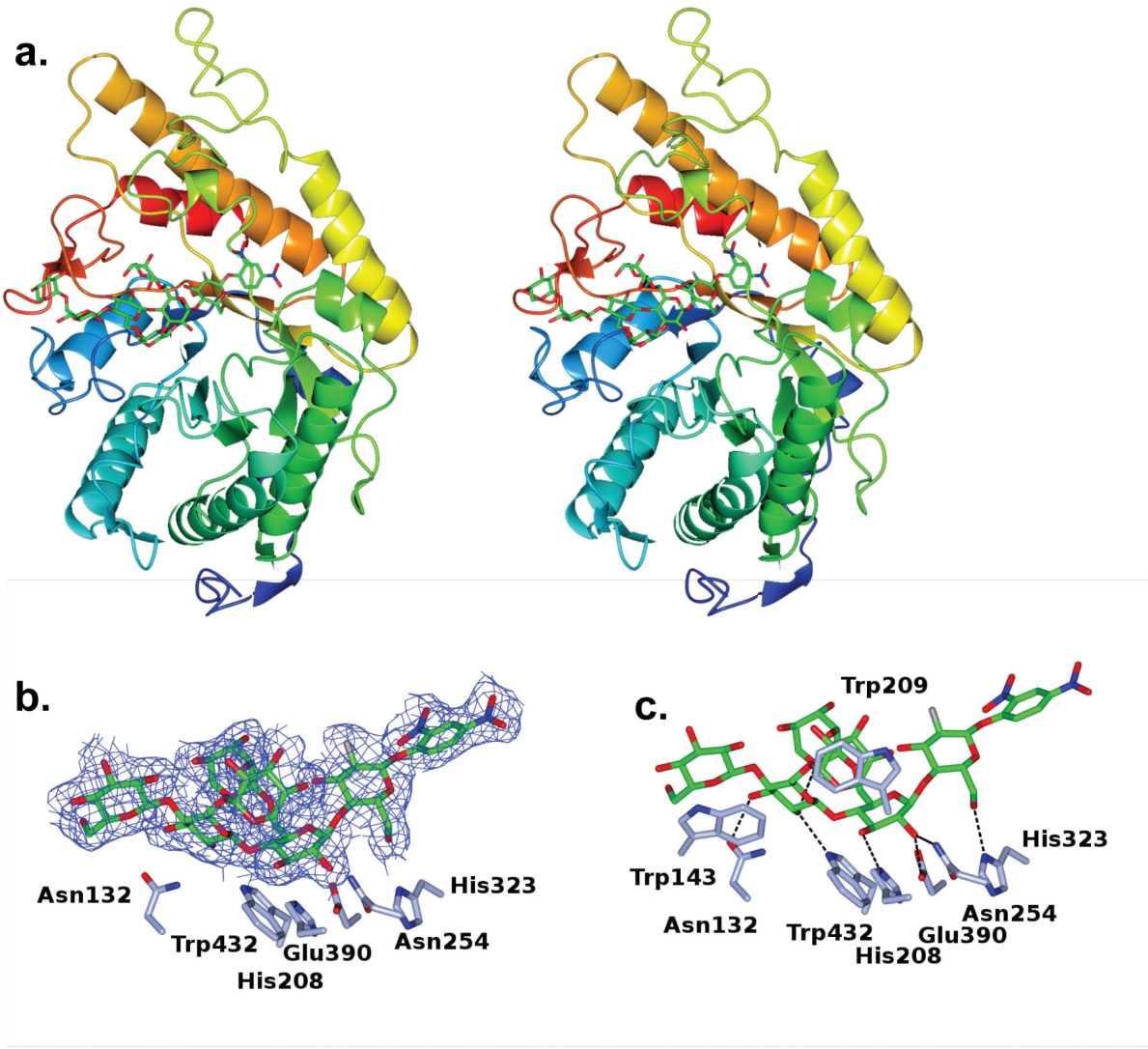


Figure A13: Tertiary structure of the non-productive complex of *Cj*GH5D(E255A) with XXXG(2F)-β-DNP. (a): Divergent (wall-eyed) stereo cartoon representation of the secondary structure of CjGH5D mutant, obtained by soaking the enzyme crystal in the inhibitor solution. Backbone colour is ramped from the N-terminus (blue) to the C-terminus (red), with the inhibitor ligand shown in green sticks. (b): Maximum-likelihood/σ_A weighted $2F_o$-F_c electron density contoured at an r.m.s.d. level of 1 σ for ligand XXXG(2F)-β-DNP (c): Hydrogen bonding and aromatic interactions. In panels b and c, side chains of interacting residues are shown in ice blue and hydrogen bonds are shown as dashed lines.

Table A1: Crystallographic data collection and refinement statistics.

	XXXG(2F)-*Cj*GH5D(E255A) covalent complex (co-crystallized)	**XXXG(2F)-β-DNP: *Cj*GH5D(E255A) non-covalent complex (soak)**
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	56.1, 97.2, 156.7	56.5, 97.8, 158.1
a, b, g (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	33.25-1.70 (1.73-1.70)	83.18-2.00 (2.05-2.00)
R_{sym} or R_{merge}	0.066 (1.325)	0.105 (1.027)
R_{pim}	0.026 (0.544)	0.05 (0.486)
$CC_{1/2}$	0.998 (0.777)	0.992 (0.630)
$I / \sigma I$	12.3 (1.2)	7.4 (1.5)
Completeness (%)	100.0 (100.0)	99.9 (99.4)
Redundancy	8.1 (7.7)	6.0 (6.3)
Refinement		
No. reflections	89231	57026
R_{work} / R_{free}	0.22/0.27	0.22/0.29
No. atoms		
Protein	5871	5870
Ligand/ion	221	212
Water	357	209
B-factors		
Protein	39.8	42.1
Ligand/ion	41.9	38.9
Water	41.1	39.4
R.m.s deviations		
Bond lengths (Å)	0.019	0.016
Bond angles (°)	1.917	1.782
Ramachandran plot residues		
In most favourable regions (%)	94.7	94.0
In allowed regions (%)	4.5	5.7
PDB code	6HAA	6HA9

Table A2: Privateer results showing validation for Glc (BGC) and Xyl (XYS) residues in XXXG(2F)-CjGH5D(E255A) covalent and XXXG(2F)-β-DNP: CjGH5D(E255A) non-covalent complexes and Glc-2-F (G2F) in the former.

Residue Name	Q[1]	Phi	Theta	Ano-mer	D/L[2]	Conform-ation	RSCC[3]	*B*-factor	Conform-ation Diagnosis[4]
XXXG(2F)-covalent									
G2F/A 1001	0.61	250.3	0.8	α	D	$^4C_1$	0.93	27.4	Ok
BGC/A 1002	0.60	81.9	1.9	β	D	$^4C_1$	0.95	26.7	Ok
BGC/A 1003	0.59	177.0	2.1	β	D	$^4C_1$	0.93	30.9	Ok
BGC/A 1004	0.57	131.8	3.6	β	D	$^4C_1$	0.84	41.3	Ok
XYS/A 1005	0.54	286.1	4.5	α	N	$^4C_1$	0.94	27.8	Ok
XYS/A 1006	0.57	340.5	1.9	α	N	$^4C_1$	0.90	37.2	Ok
G2F/B 1001	0.58	150.5	4.6	α	D	$^4C_1$	0.94	29.6	Ok
BGC/B 1002	0.61	44.0	1.7	β	D	$^4C_1$	0.94	31.5	Ok
BGC/B 1003	0.62	139.0	4.7	β	D	$^4C_1$	0.91	39.5	Ok
BGC/B 1004	0.55	222.9	10.6	β	D	$^4C_1$	0.88	54.9	Ok
XYS/B 1005	0.60	73.1	6.6	α	N	$^4C_1$	0.77	56.2	Ok
XYS/B 1006	0.60	333.9	4.7	α	N	$^4C_1$	0.92	30.4	Ok
XYS/B 1007	0.57	322.4	15.1	α	N	$^4C_1$	0.83	43.1	Ok
XXXG(2F)-β-DNP non-covalent									
BGC/A 503	0.53	250.5	25.1	β	D	^4E	0.92	31.4	*
XYS/A 504	0.58	30.9	1.0	α	N	$^4C_1$	0.94	32.5	Ok
XYS/A 505	0.46	237.6	166.0	α	N	$^1C_4$	0.84	26.1	Ok
BGC/A 506	0.60	55.2	6.3	β	D	$^4C_1$	0.93	33.4	Ok
BGC/A 507	0.57	293.5	7.3	β	D	$^4C_1$	0.84	37.7	Ok
BGC/B 506	0.49	229.1	40.6	β	D	^4E	0.93	34.8	*
XYS/B 507	0.56	55.3	8.8	α	N	$^4C_1$	0.91	35.7	Ok
XYS/B 508	0.55	149.0	22.1	α	N	$^4C_1$	0.73	44.6	Ok
BGC/B 509	0.52	151.9	5.1	β	D	$^4C_1$	0.84	47.5	Ok
XYS/B 510	0.56	136.5	167.8	α	N	$^1C_4$	0.85	28.0	Ok
BGC/B 511	0.60	5.0	5.6	β	D	$^4C_1$	0.92	36.1	Ok

[1] Q represents total puckering amplitude (in Å). [2] Handedness; N is displayed if unable to be determined based solely on the structure. [3] Real Space Correlation Coefficient, measuring

agreement between model and positive omit maximum-likelihood/σ_A weighted F_o-F_c density. [4]

Conformation either acceptable (Ok) or might be mistaken (*)

Appendix B: Supporting Information for Chapter 3: *N*-bromoacetylglycosylamine-based Irreversible Inhibitors of Mixed-linkage Glucanases

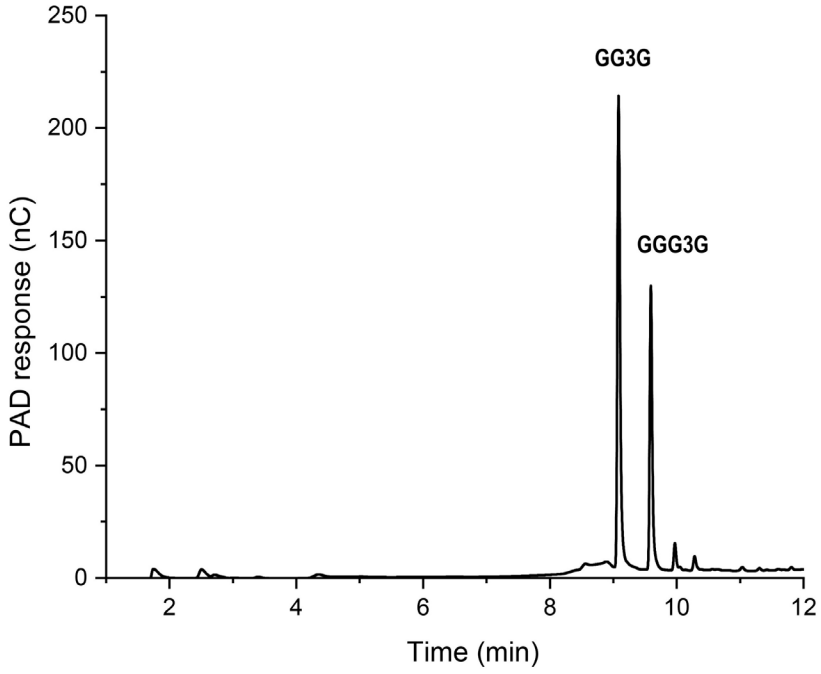


Figure B1: HPAEC-PAD chromatogram of the limit-digests products of oat β-glucan with BoGH16. GG3G: Glc β(1,4)-Glc β(1,3)-Glc; GGG3G: Glc β(1,4)-Glc β(1,4)-Glc β(1,3)-Glc

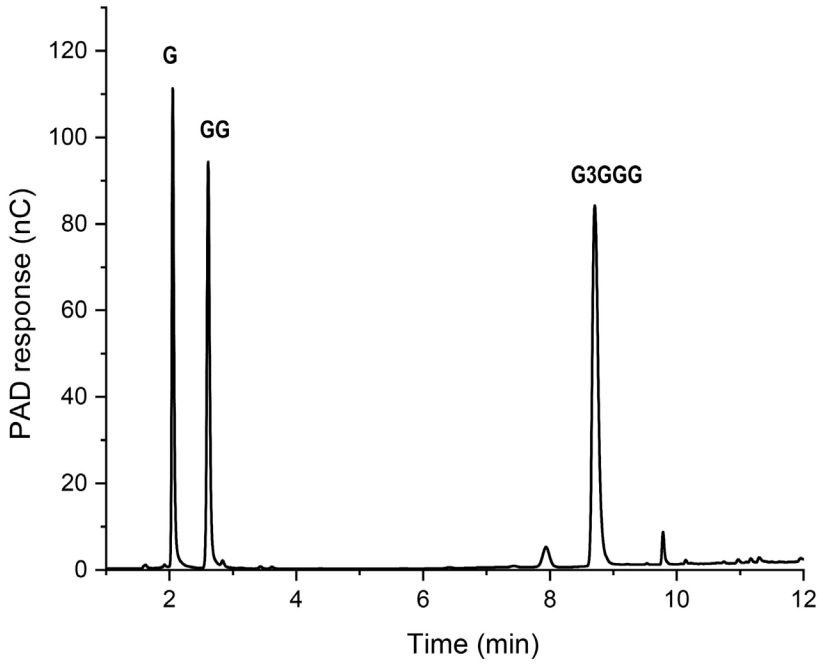


Figure B2: HPAEC-PAD chromatogram of the limit-digests products of oat β-glucan with VvEG16 (ΔV152). G: Glc, GG: Glc β(1,4)-Glc, G3GGG: Glc β(1,3)-Glc β(1,4)-Glc β(1,4)-Glc

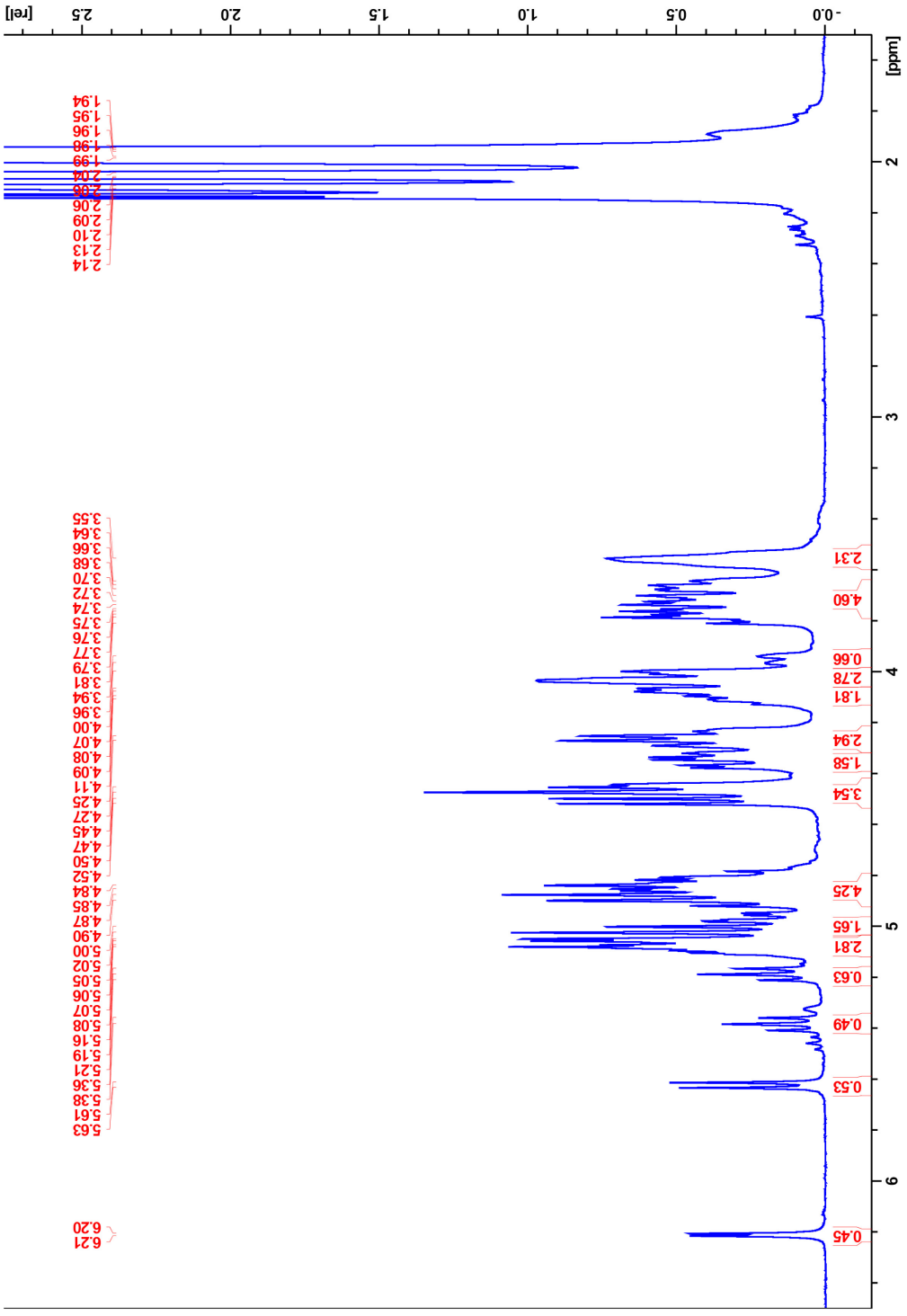


Figure B3: ^{1}H NMR of per-*O*-acetylated G3GGG

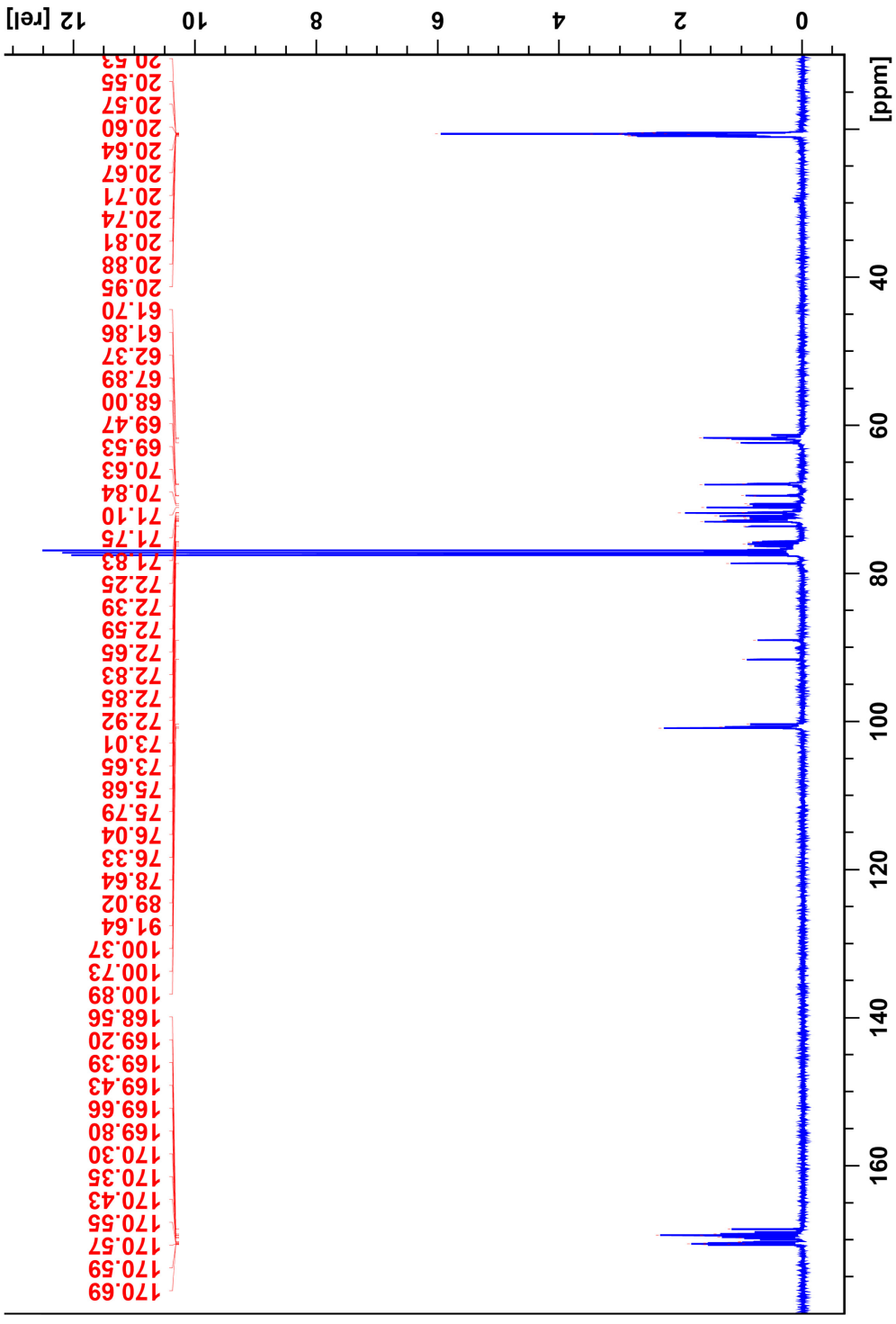


Figure B4: ^{13}C NMR of per-*O*-acetylated G3GGG

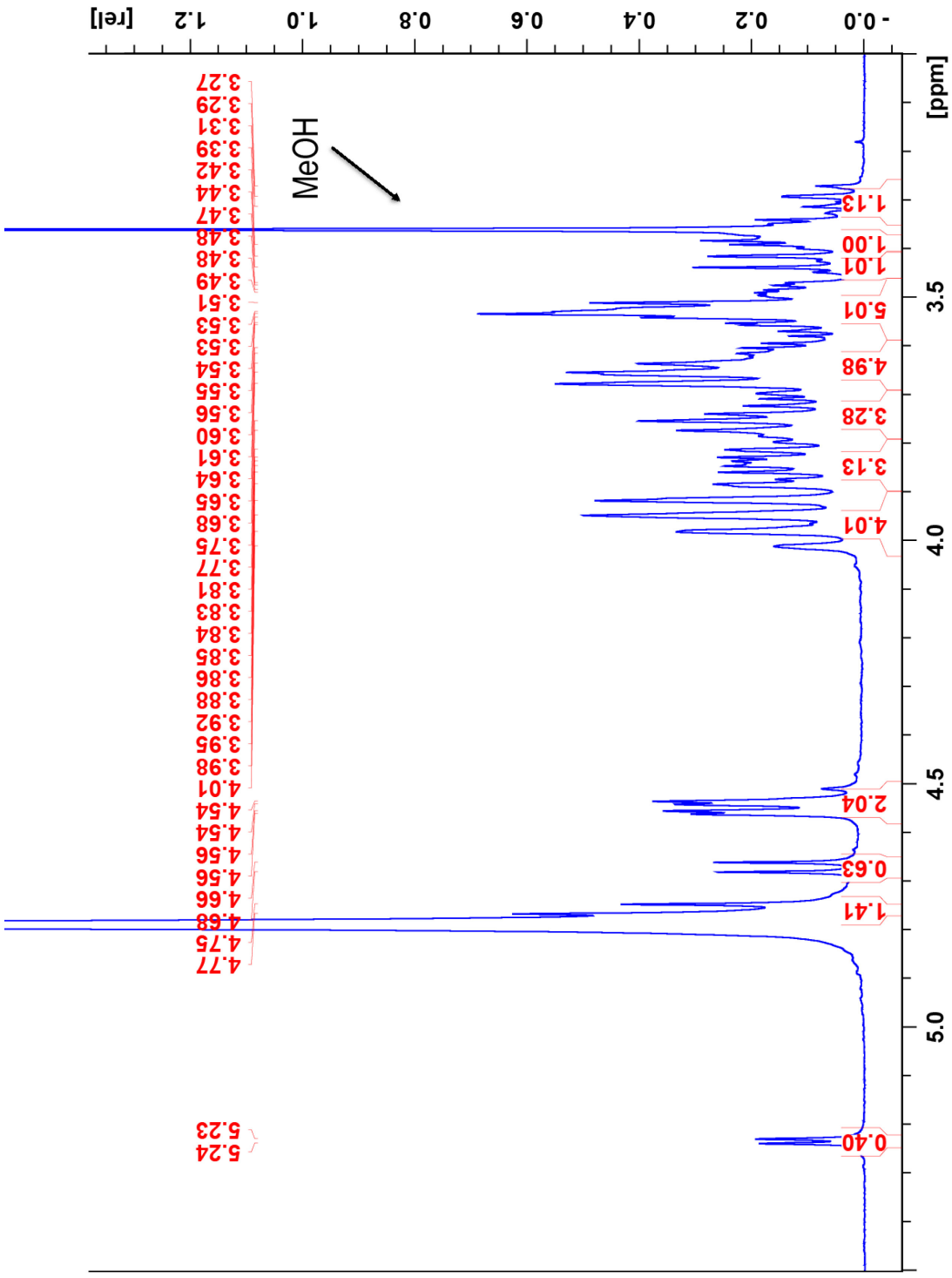


Figure B5: ^{1}H NMR of G3GGG

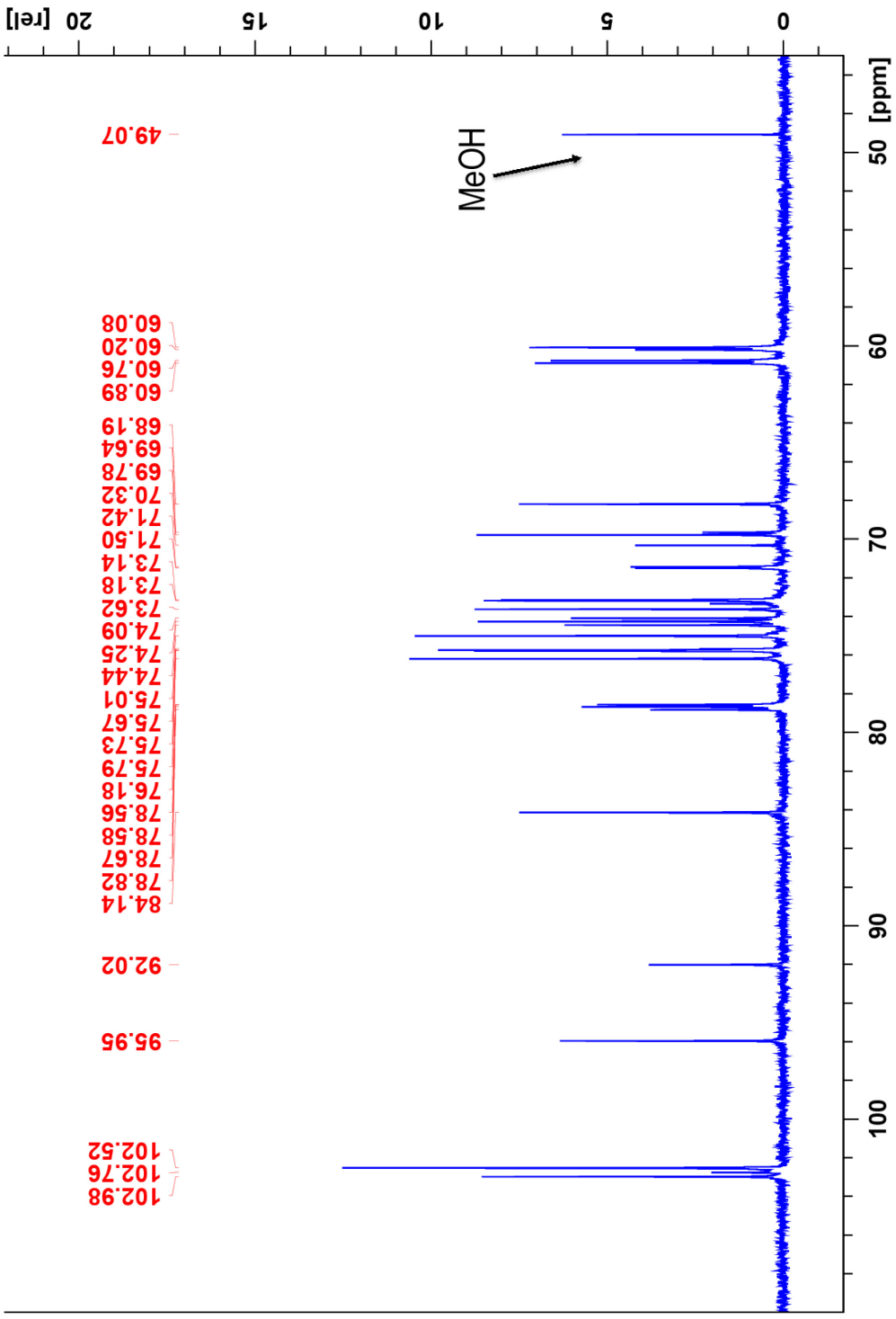


Figure B6: ^{13}C of G3GGG

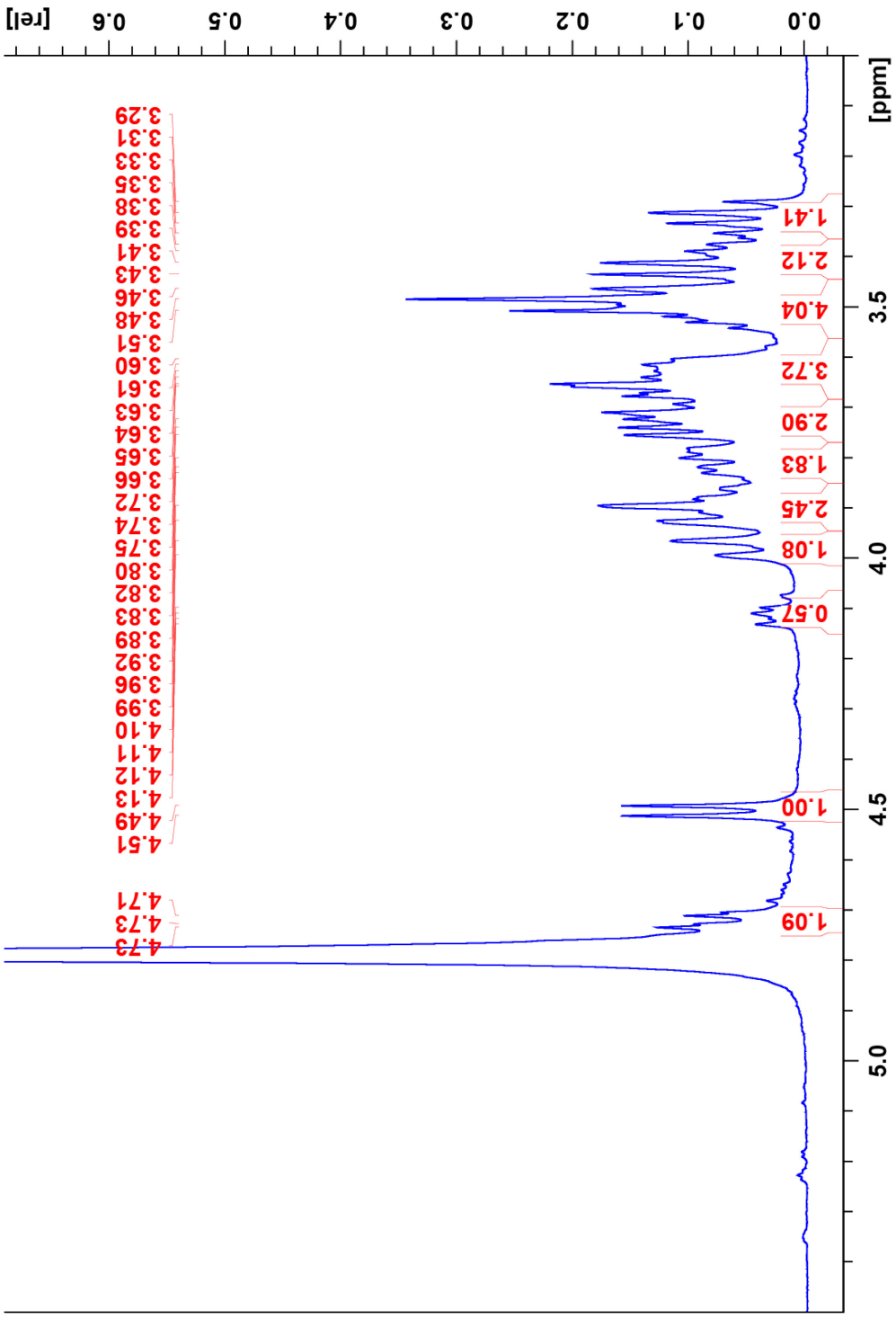


Figure B7: ^{1}H NMR of GG3G-β-NH$_2$

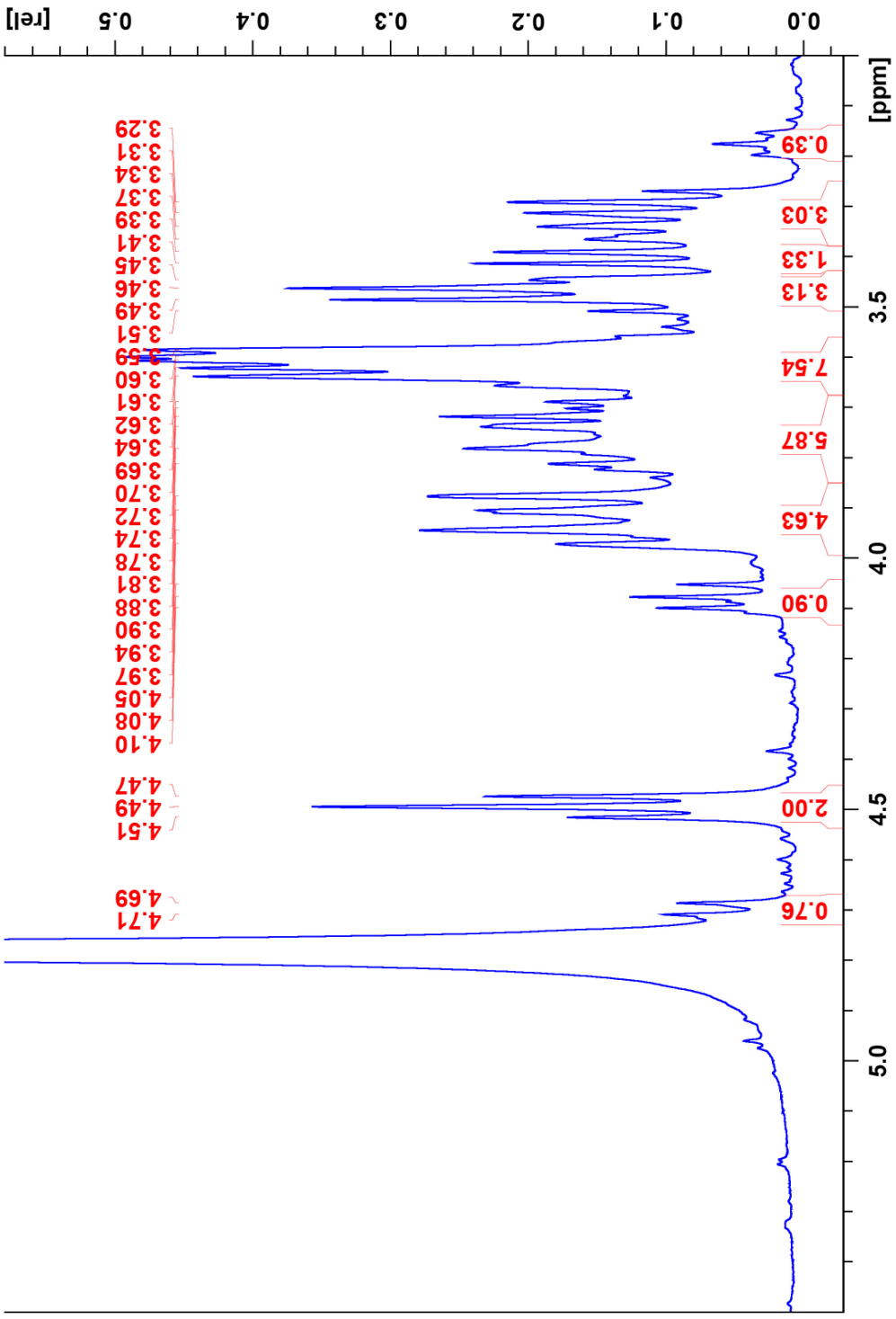


Figure B8: ^{1}H NMR of GGG3G-β-NH$_2$

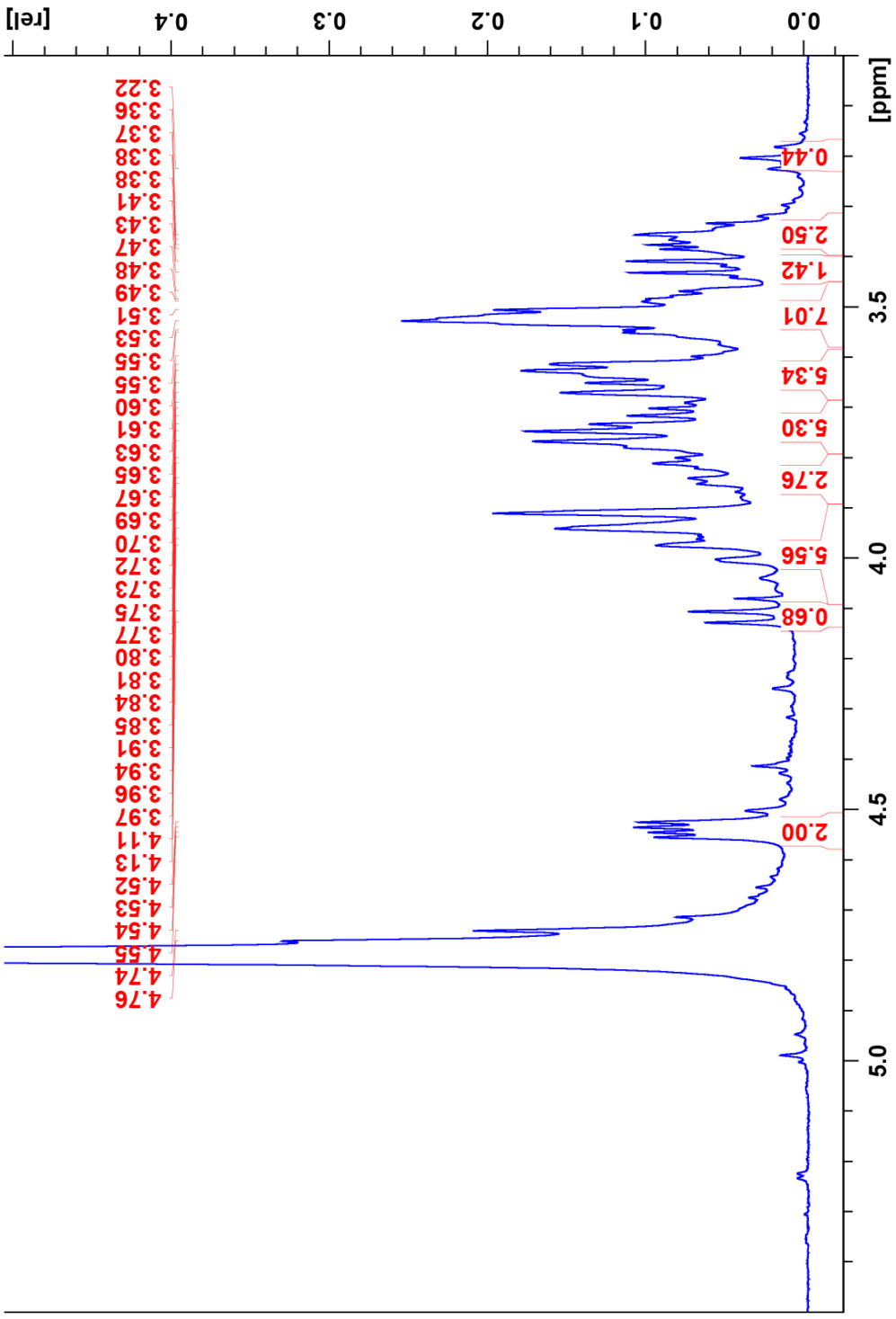


Figure B9: ^{1}H NMR of G3GGG-β-NH$_2$

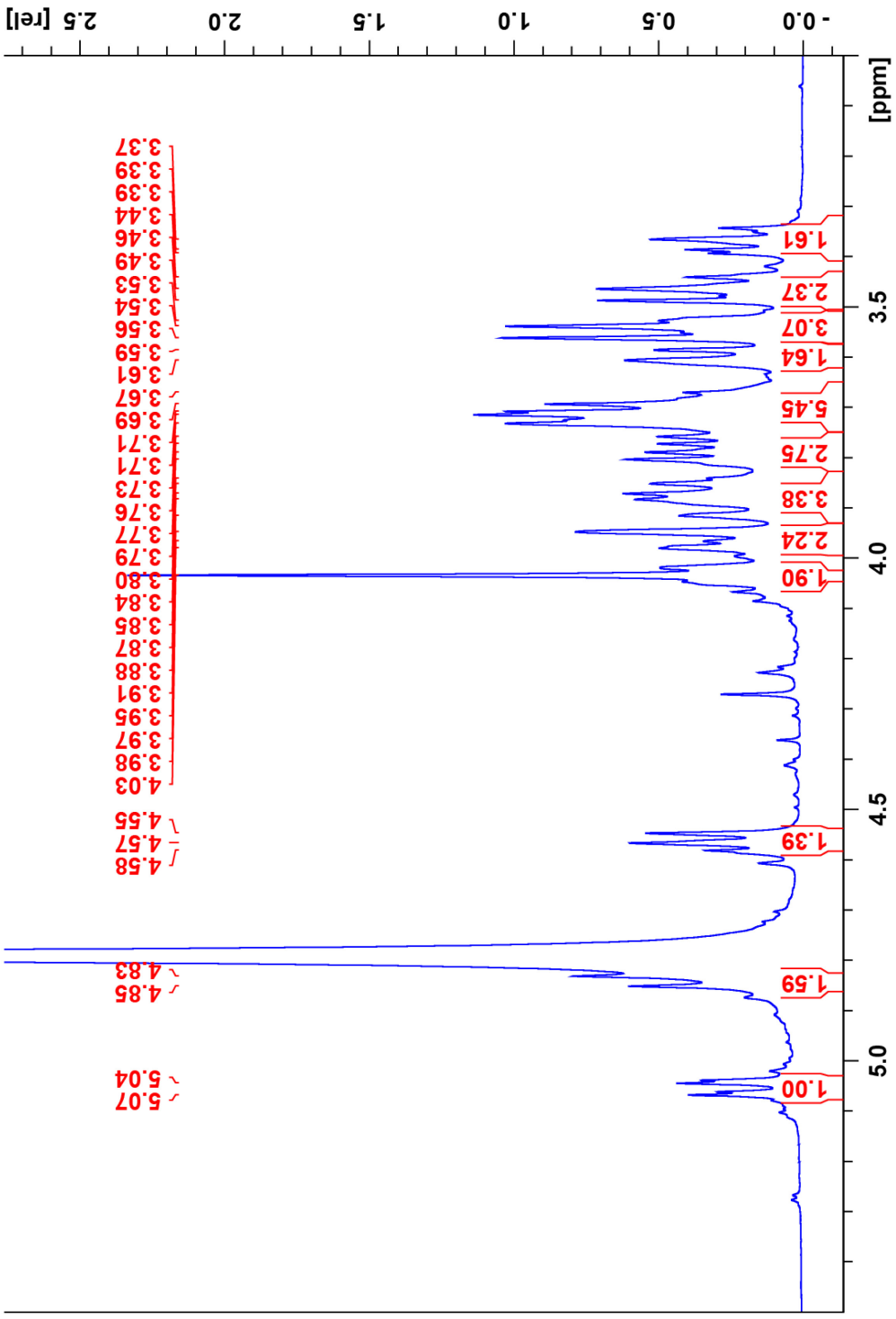


Figure B10: ^{1}H NMR of GG3G NHCOCH$_2$Br (6)

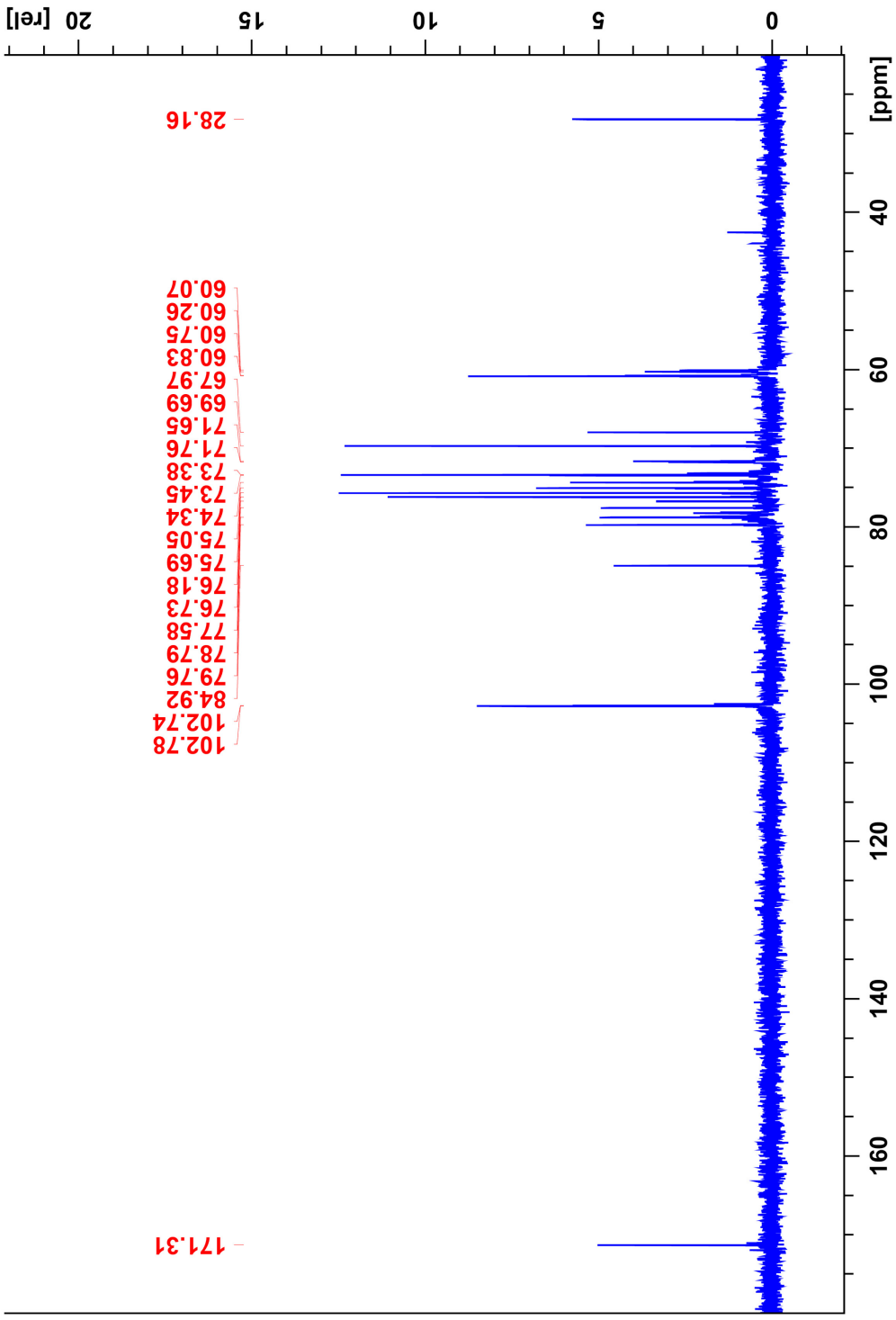


Figure B11: ^{13}C NMR of GG3G $NHCOCH_2Br$ (6)

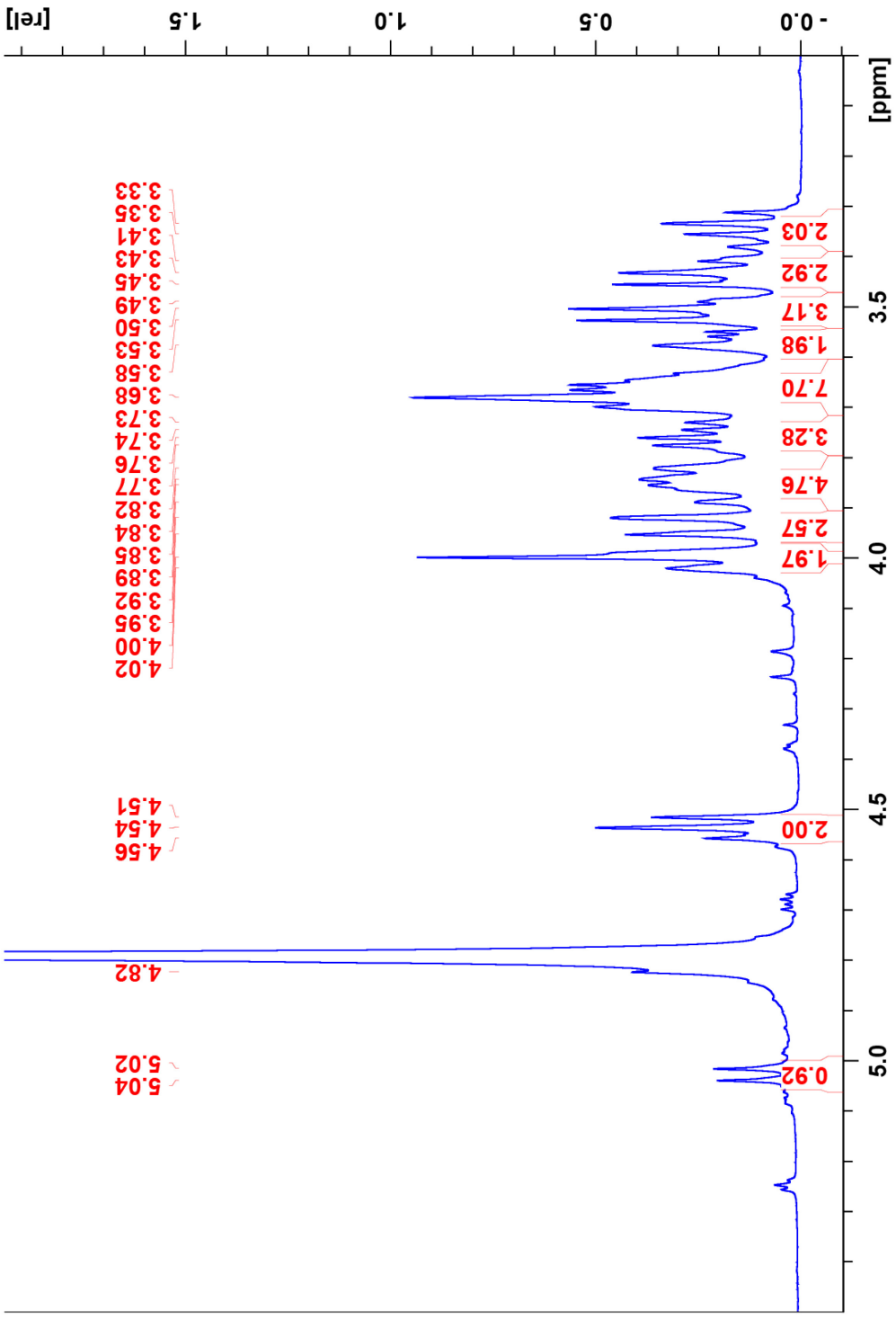


Figure B12: ^{1}H NMR of GGG3G NHCOCH$_2$Br (7)

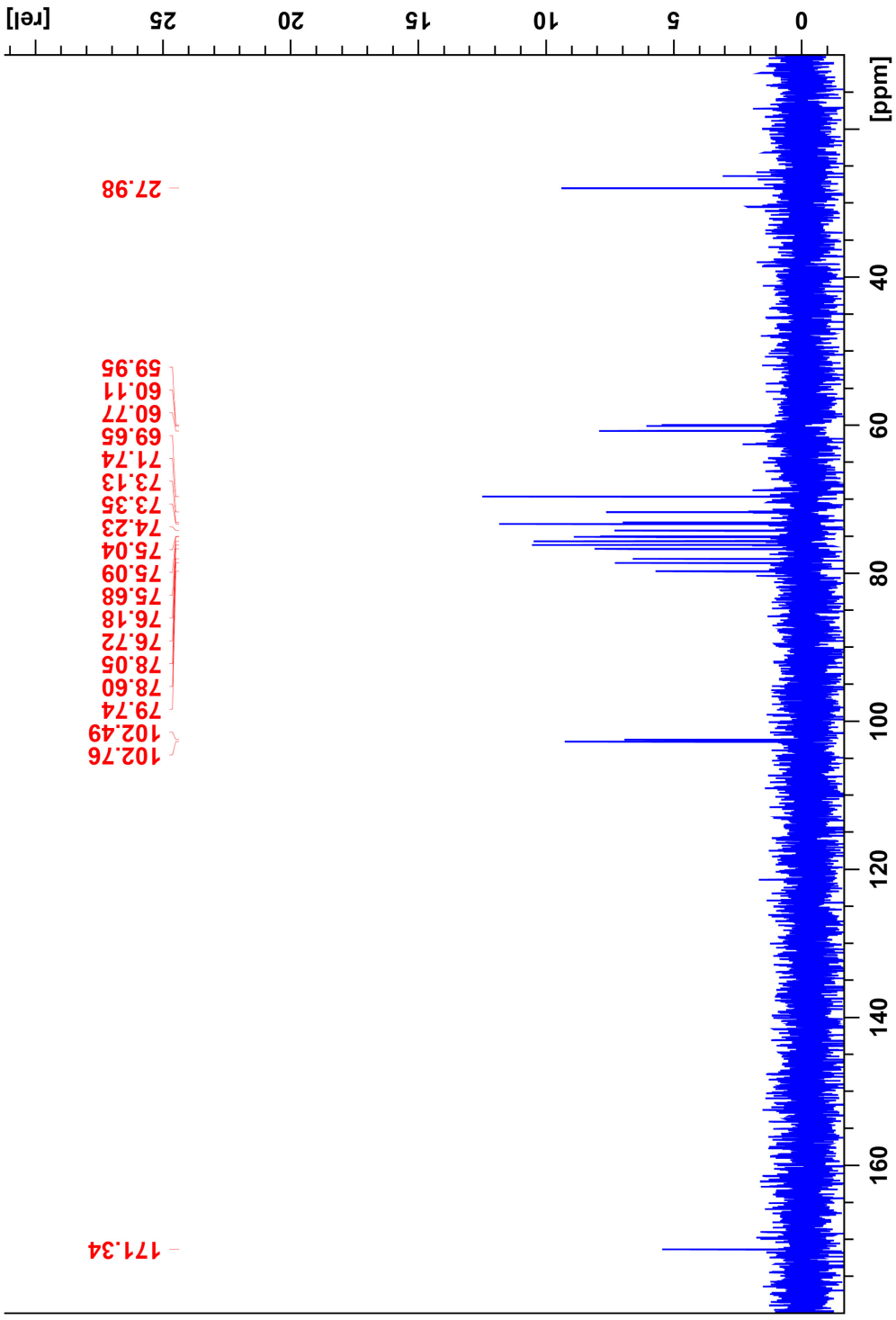


Figure B13: ^{13}C NMR of GGG3G NHCOCH$_2$Br (7)

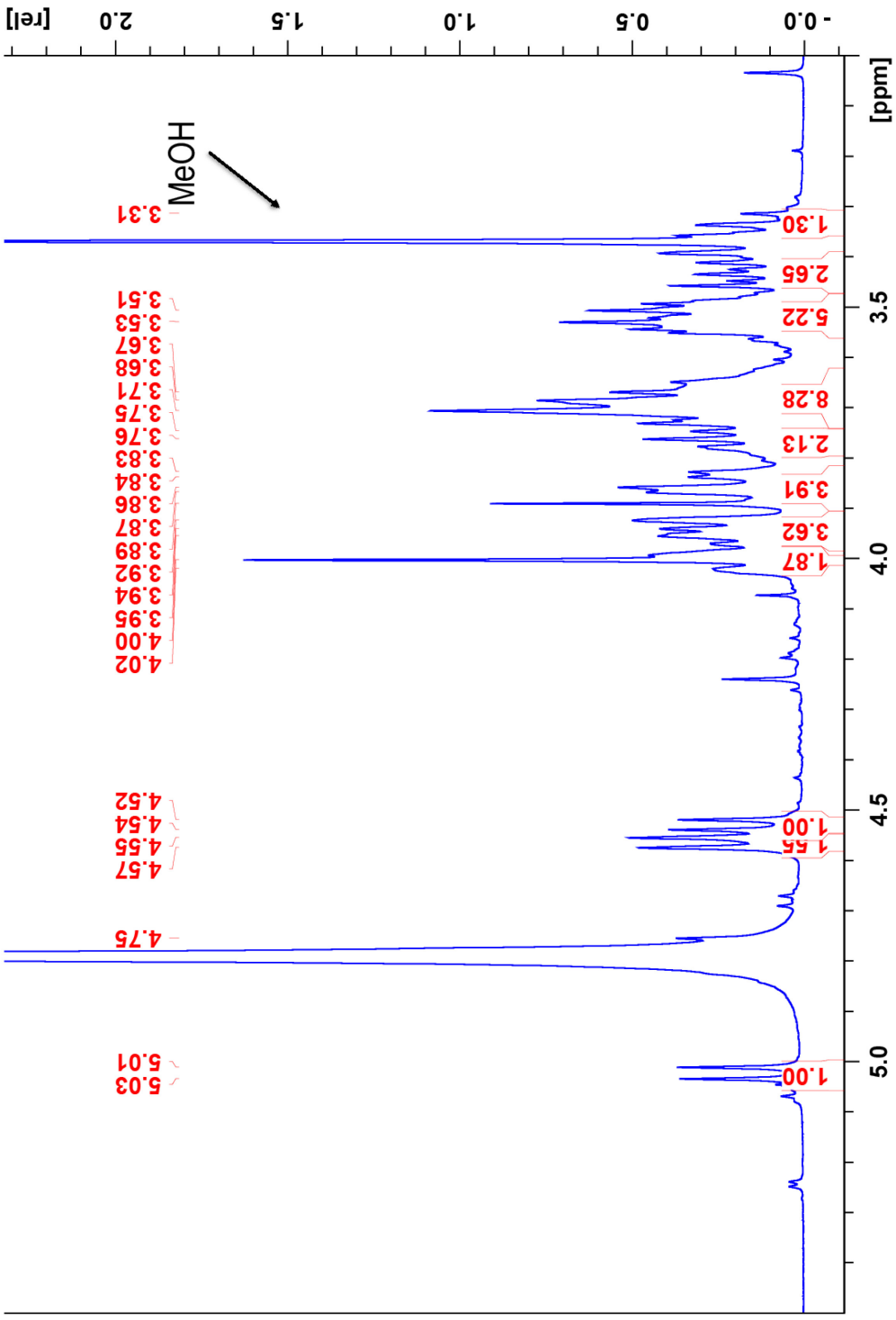


Figure B14: ^{1}H NMR of G3GGG NHCOCH$_2$Br (8)

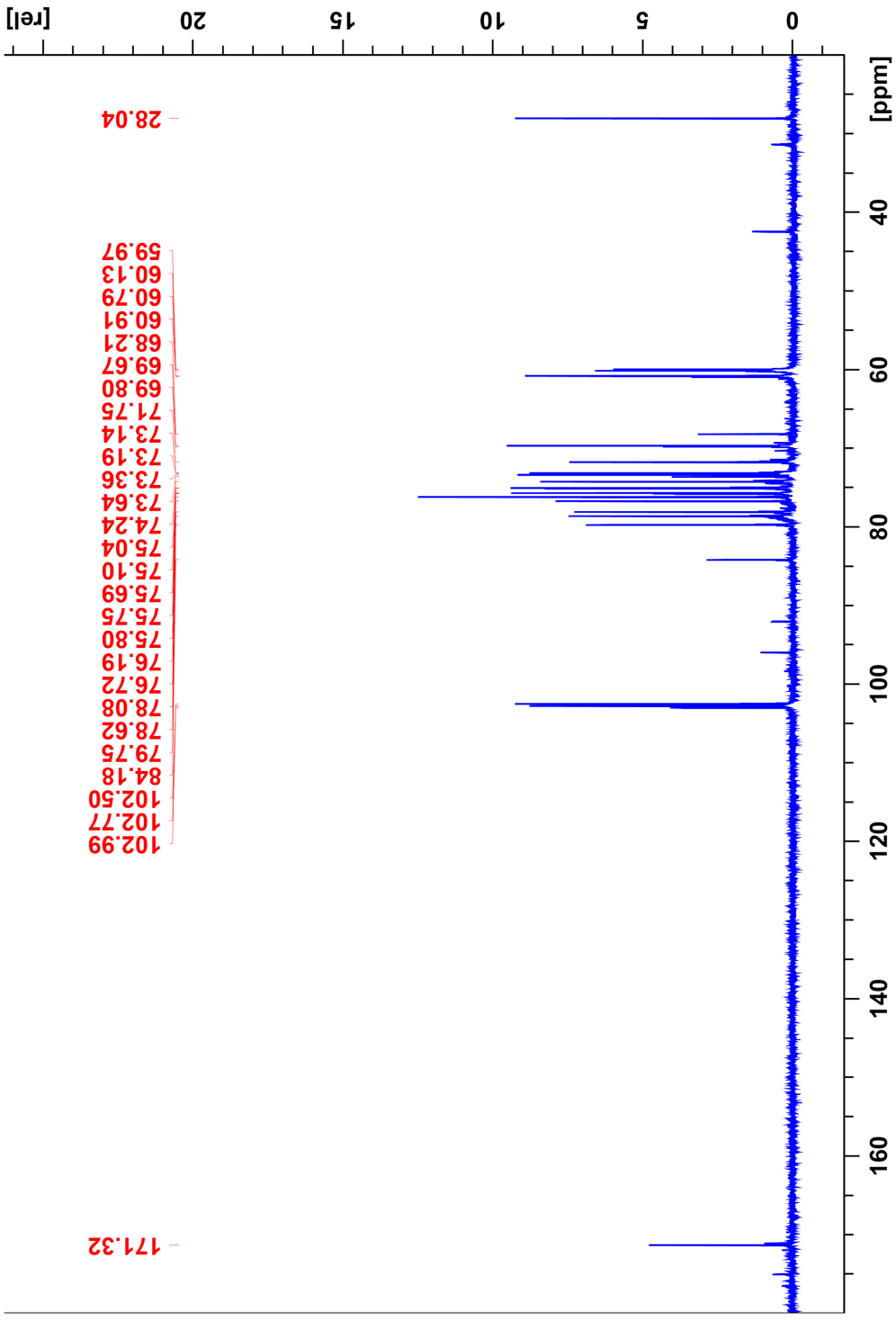


Figure B15: ^{13}C NMR of G3GGG NHCOCH$_2$Br (8)

Table B1: Data collection and refinement statistics

Data Collection	BoGH16-GGG3GNBrAc (PDB ID: 6VHO)
Beamline	APS 23ID-B
Wavelength (Å)	1.0332
Space Group	C2
Cell dimensions	
a, b, c (Å)	168.8, 61.2, 49.5
α, β, γ (°)	90.0, 93.6, 90.0
No. of reflections	
total	93228 (3800)
unique	27459 (1343)
Resolution (Å)	84.12 - 2.15 (2.19 - 2.15)
R_{meas}	0.157 (0.908)
$CC_{1/2}$	0.992 (0.569)
I/σI	6.8 (1.4)
Completeness (%)	99.5 (96.1)
Multiplicity	3.4 (2.8)
Refinement	
Resolution (Å)	84.12 - 2.15
R_{work}	0.186
R_{free}	0.206
No. of atoms	
protein	3794
ligand	37
water	199
Avg B-factor (Å^2)	
protein	31.2
ligand	39.2
water	33.1
RMS deviations	
bond length (Å)	0.0128
bond angle (°)	1.75
Ramachandran statistics	
Favoured (%)	98.72
Allowed (%)	1.28
Outliers (%)	0

Table B2: Privateer validation results.

Residue	Q[1]	Phi	Theta	Anomer	D/L[2]	Conformation	RSCC[3]	B factor	Diagnostic[4]
NBG/A 1001	0.61	239.3	52.5	beta	D	^4E	0.74	48.99	*
BGC/A 1002	0.52	254.0	4.4	beta	D	$^4C_1$	0.84	36.98	Ok
BGC/A 1002	0.62	25.1	4.5	beta	D	$^4C_1$	0.81	47.44	Ok

[1]Q is the total puckering amplitude measured in A. [2]D/L is the handedness. [3]RSCC is short for Real Space Correlation Coefficient and measures the agreement between model and positive omit density. [4]Conformation is either acceptable (ok) or might be mistaken (*).

Appendix C: Supporting Information for Chapter 4: Synthesis and Application of Mechanism-based 2-deoxy-2-fluoro Oligosaccharide Inhibitors of β-(1,3) and β-(1,3)/(1,4) Glucanases

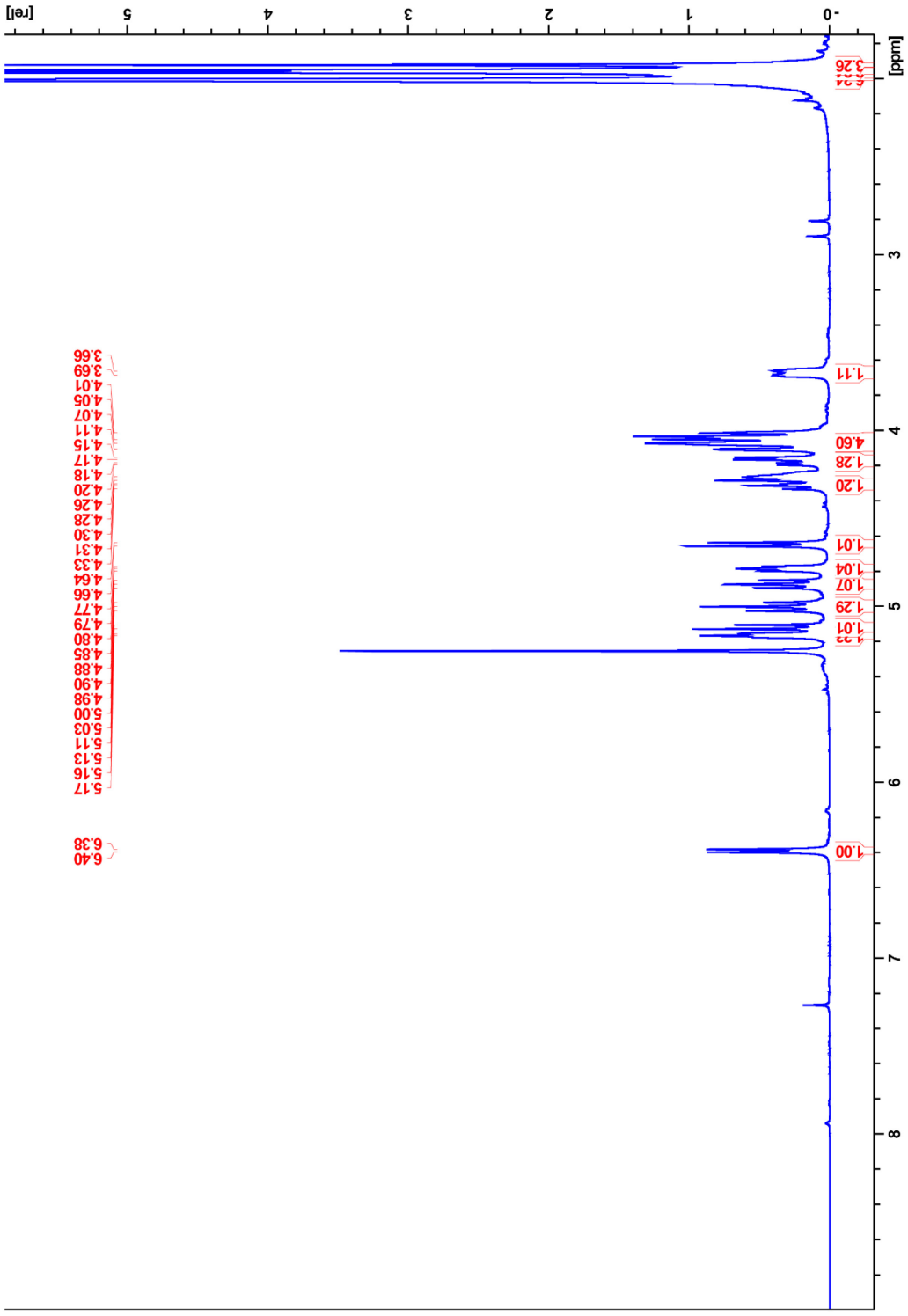


Figure C1: ^{1}H NMR spectrum of 10a (per-*O*-acetylated -G3G Glycal)

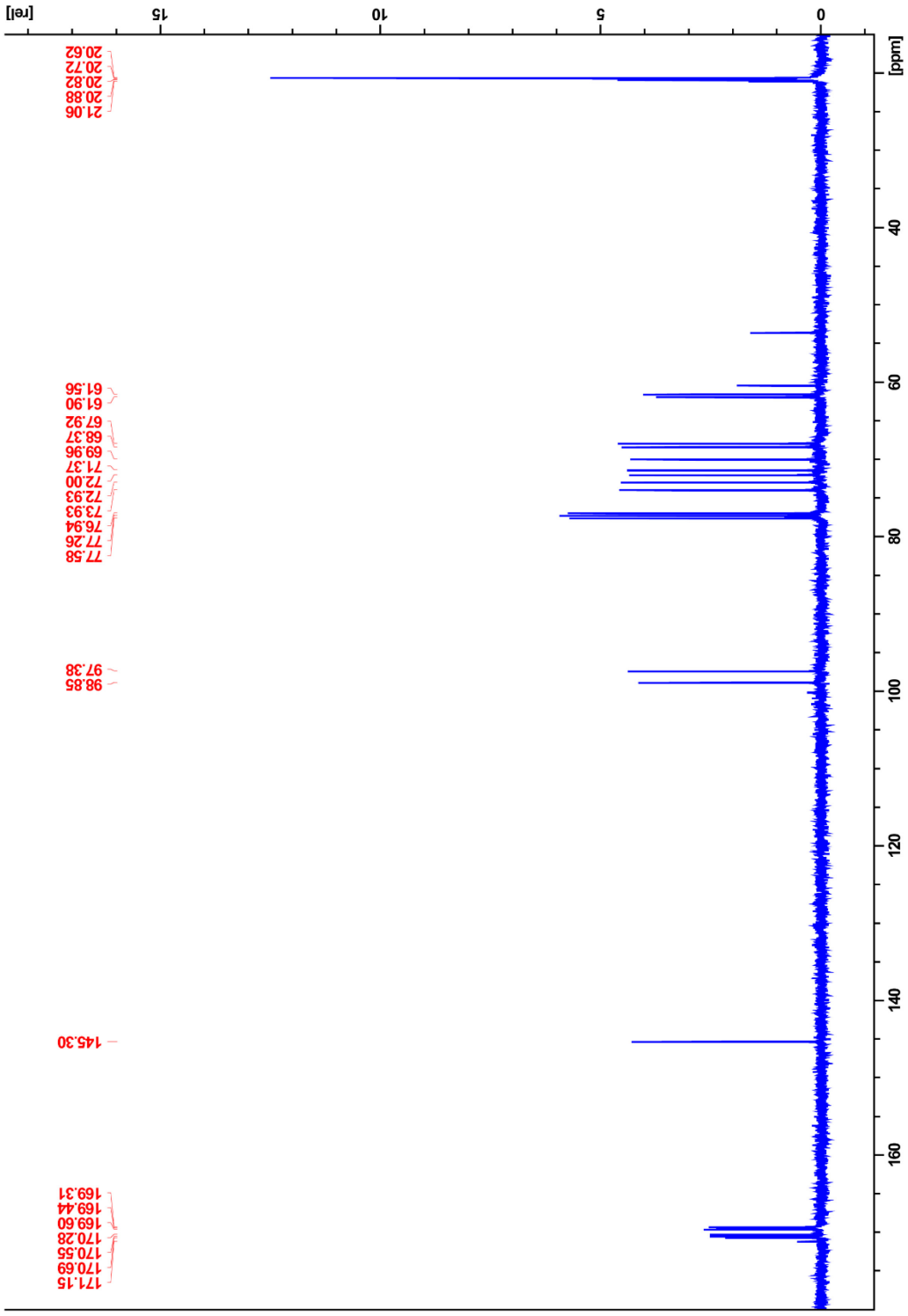


Figure C2: ^{13}C NMR spectrum of 10a (per-*O*-acetylated -G3G Glycal)

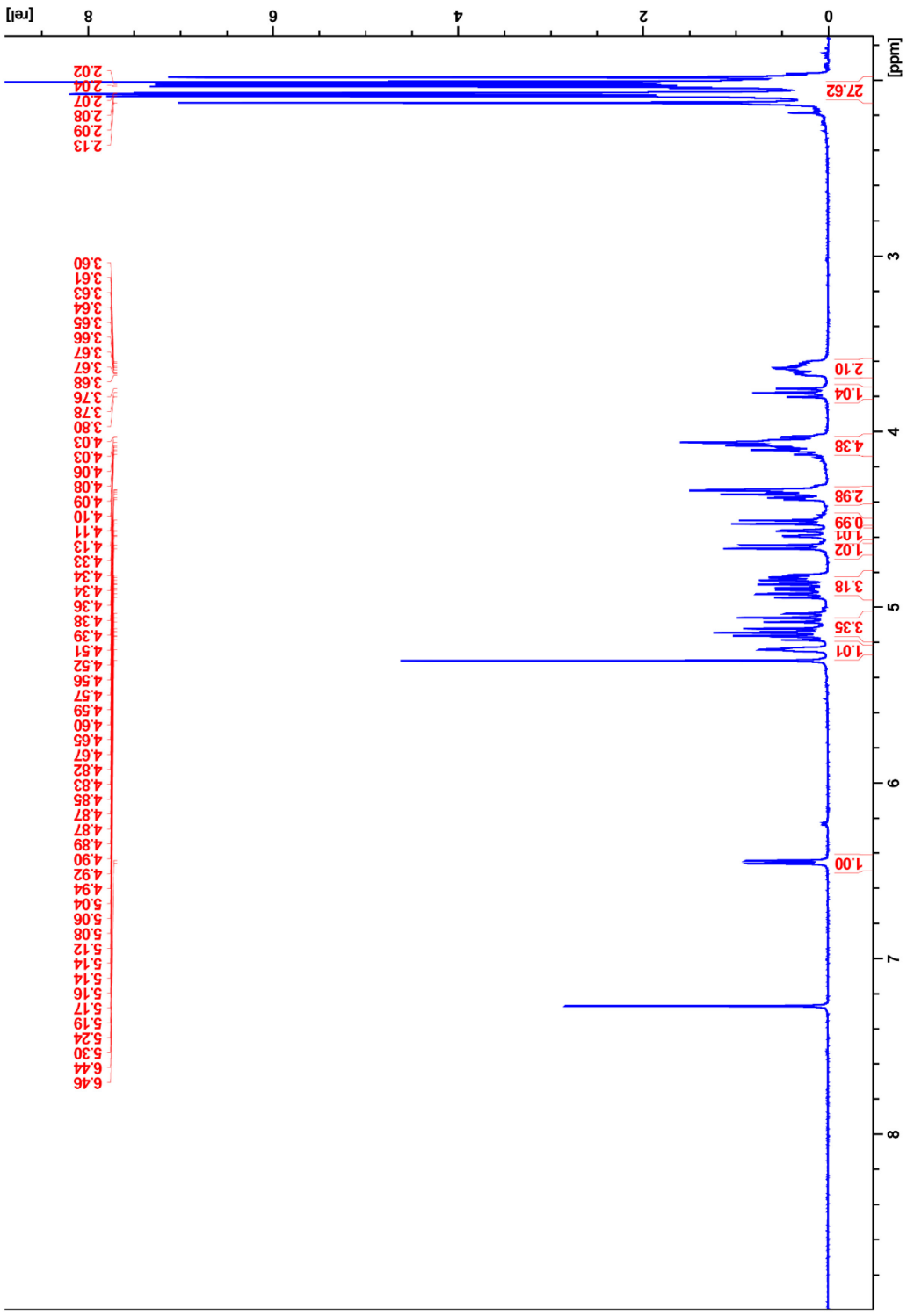


Figure C3: ^{1}H NMR spectrum of 10b (per-*O*-acetylated -GG3G Glycal)

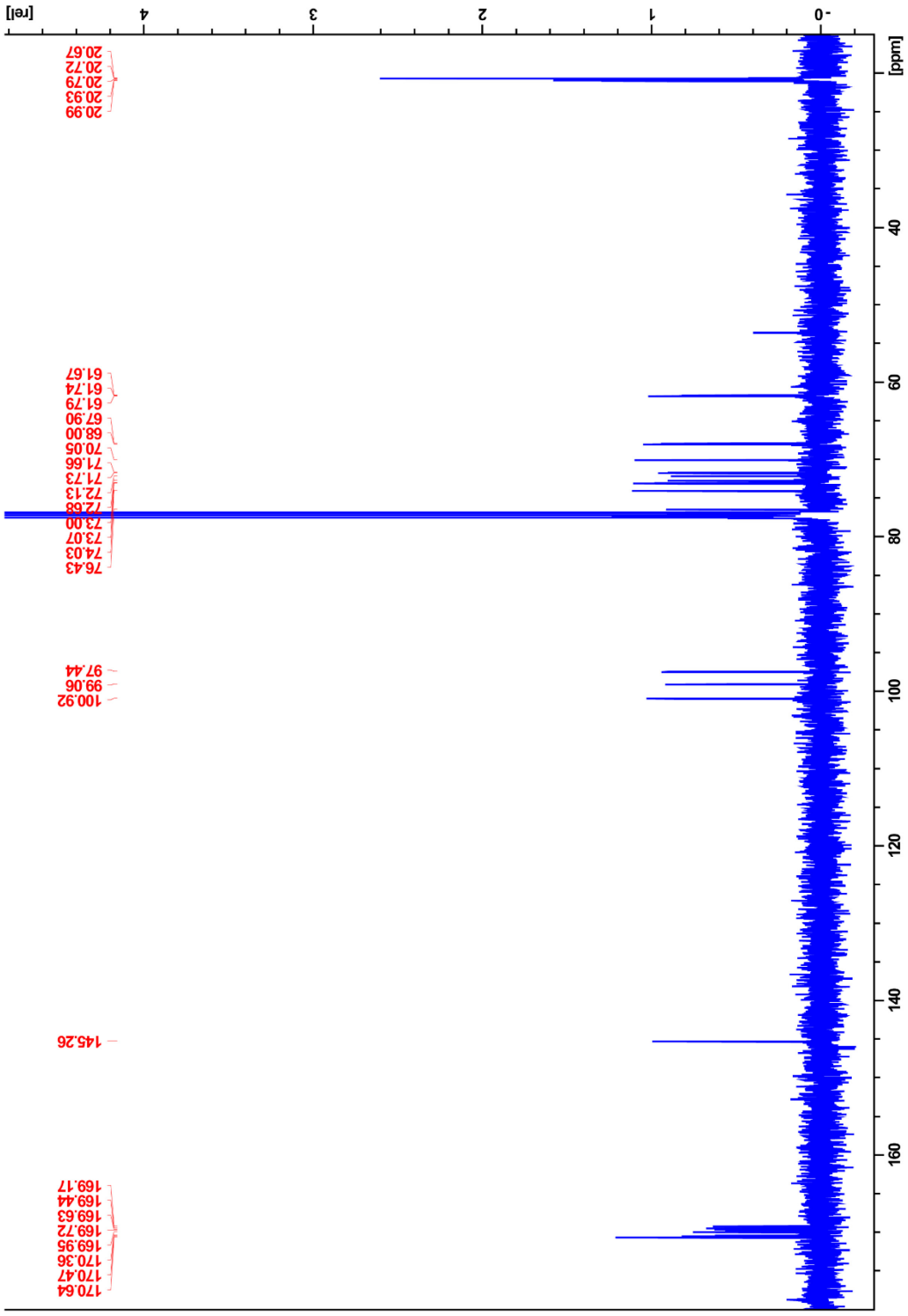


Figure C4: ^{13}C NMR spectrum of 10b (per-*O*-acetylate -GG3G Glycal)

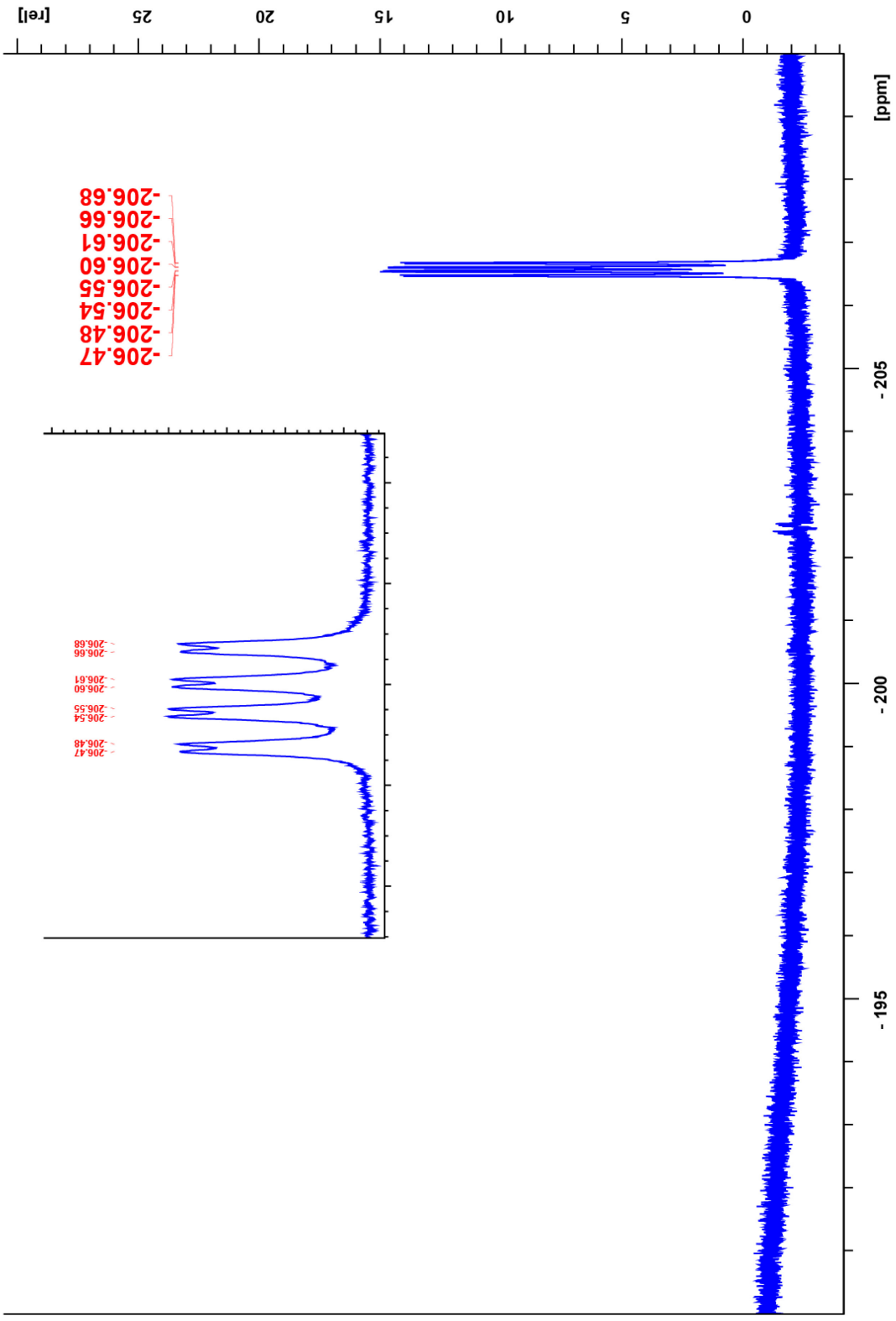


Figure C5: ^{19}F NMR spectrum of compound corresponding to spot A1 *manno* α-anomer of 12a (α-2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated) Glc β(1,3)- Man): δ -206.58 (ddd, $J_{(F2-H2)}$ = 49.67 Hz, $J_{(F2-H3)}$ = 26.49 Hz, $J_{(F2-H1)}$ = 5.54 Hz, *manno* α-anomer)

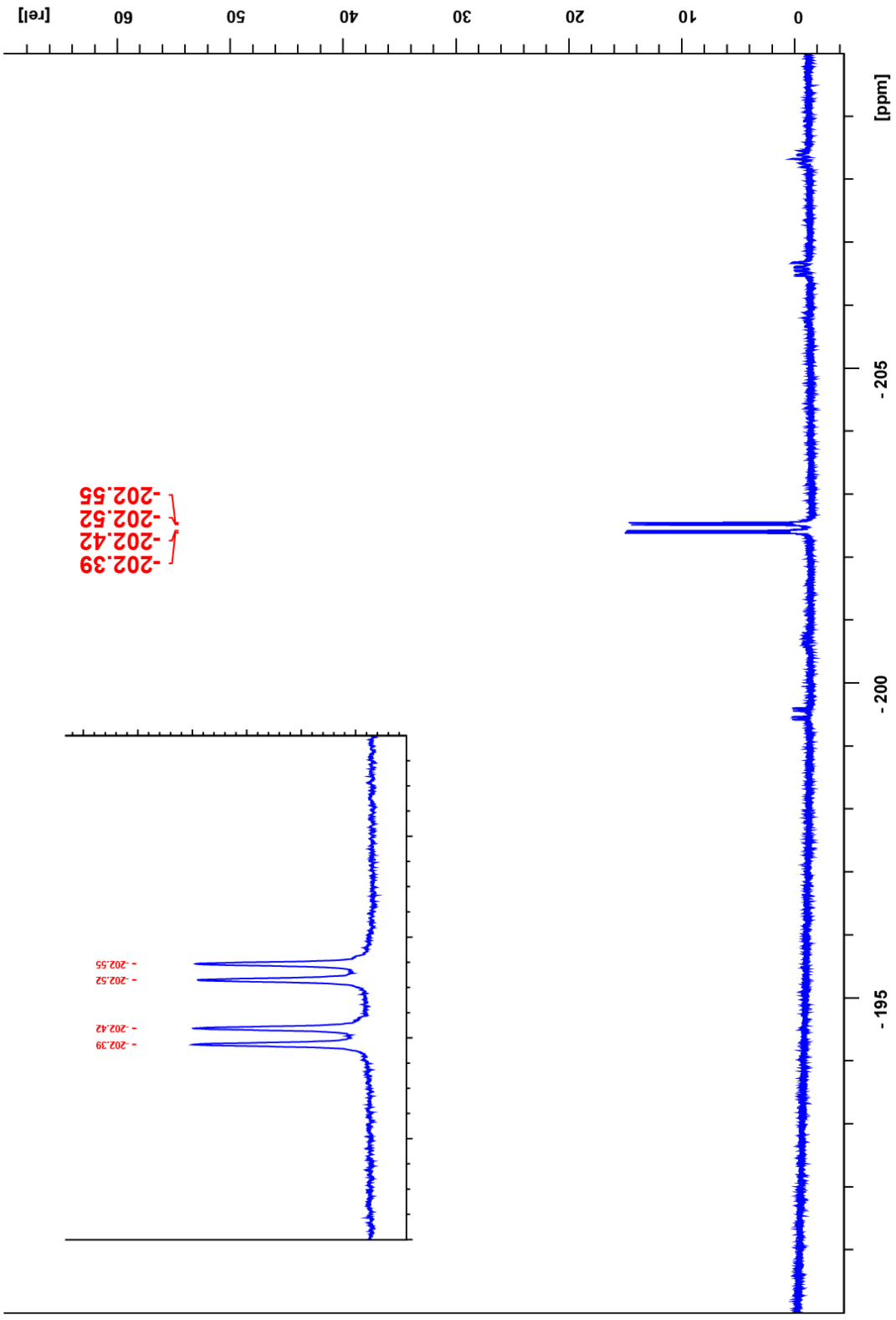


Figure C6: ^{19}F NMR spectrum of compound corresponding to spot A3 *gluco* α-anomer of 12a (α-2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated) Glc β(1,3)- Glc): δ -202.47 (dd, $J_{(F2-H2)}$ = 48.02 Hz, $J_{(F2-H3)}$ = 12.08 Hz, *gluco* α-anomer)

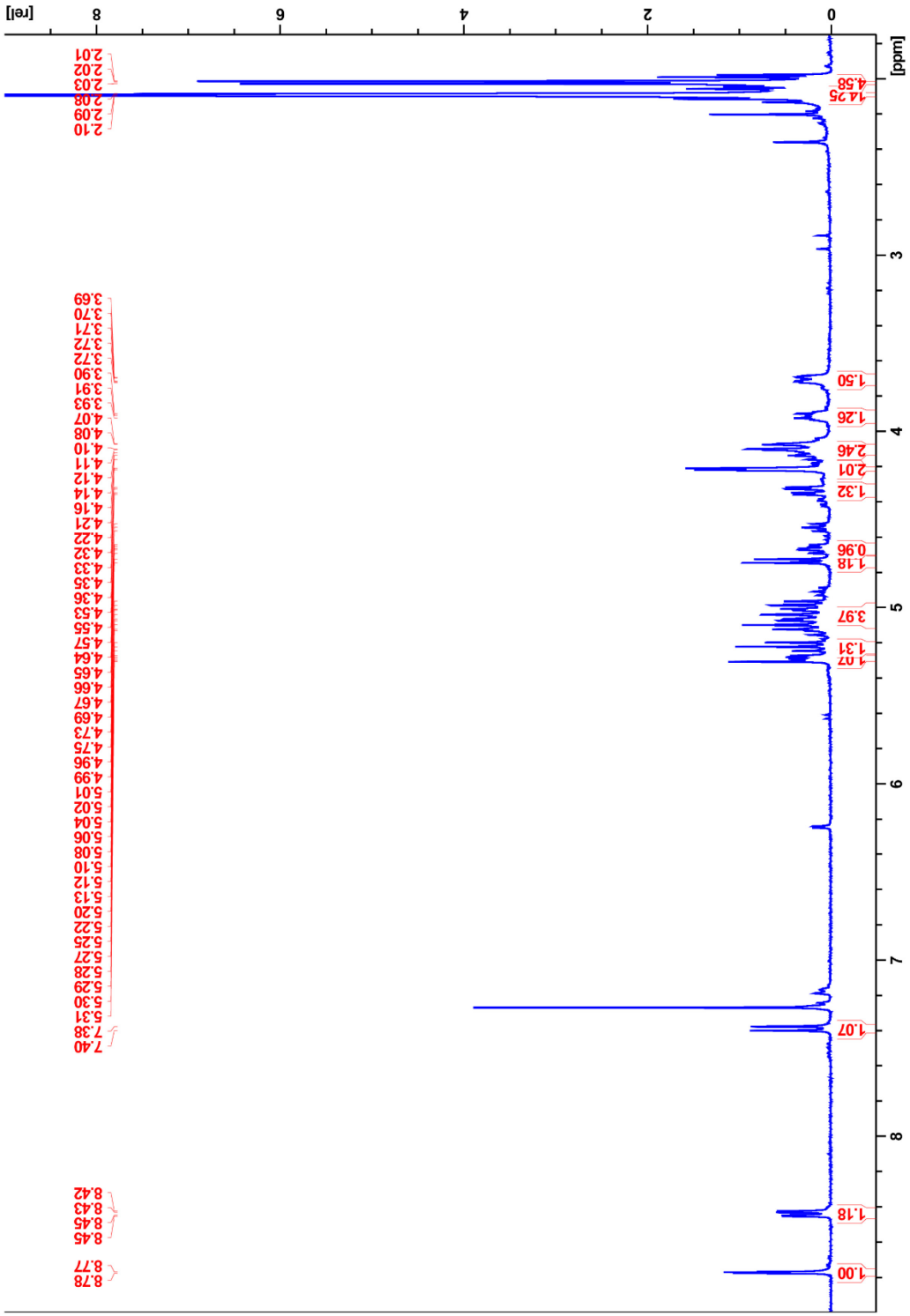


Figure C7: ^{1}H NMR spectrum of 12a (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated) G3G)

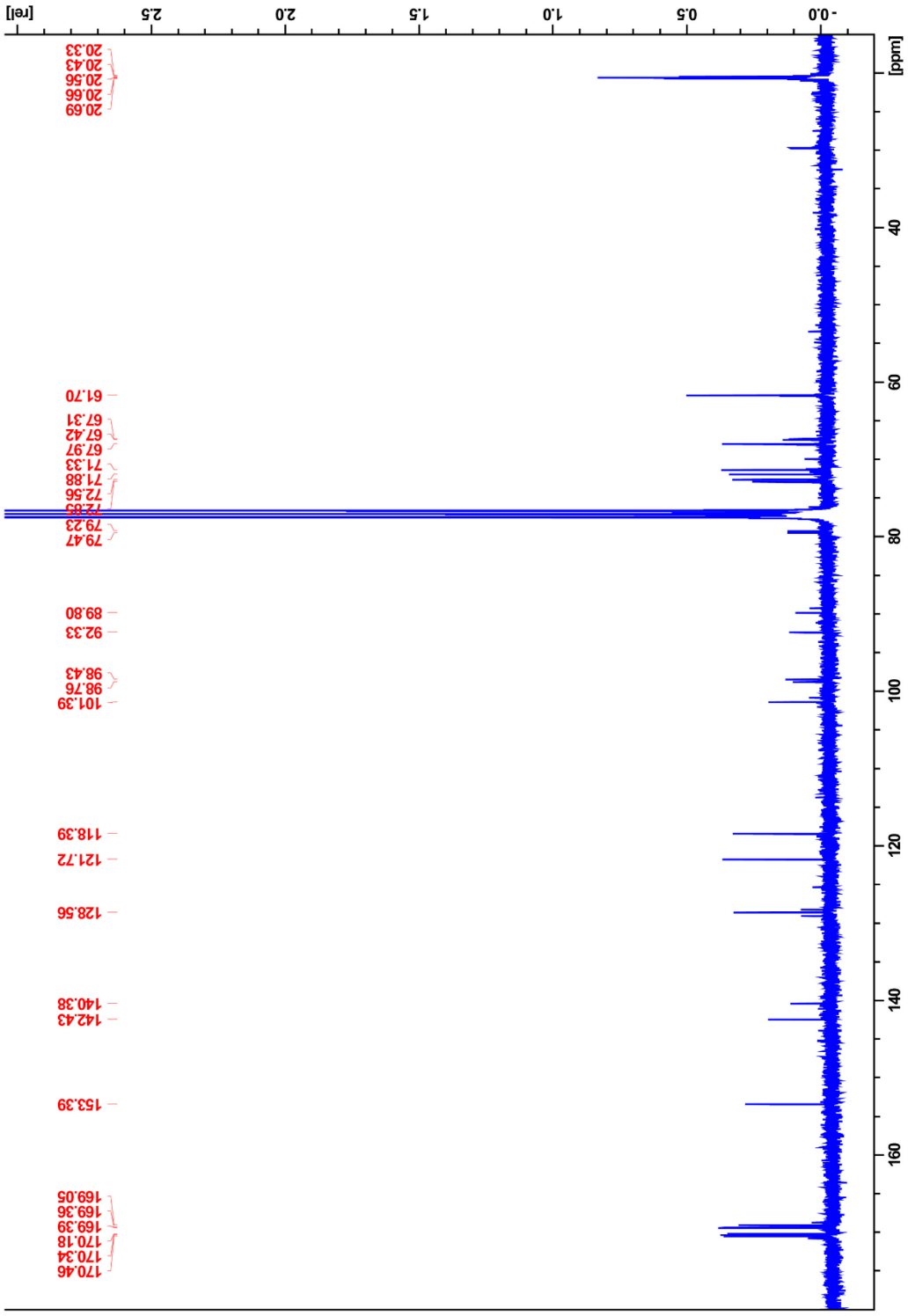


Figure C8: ^{13}C NMR spectrum of 12a (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated) G3G)

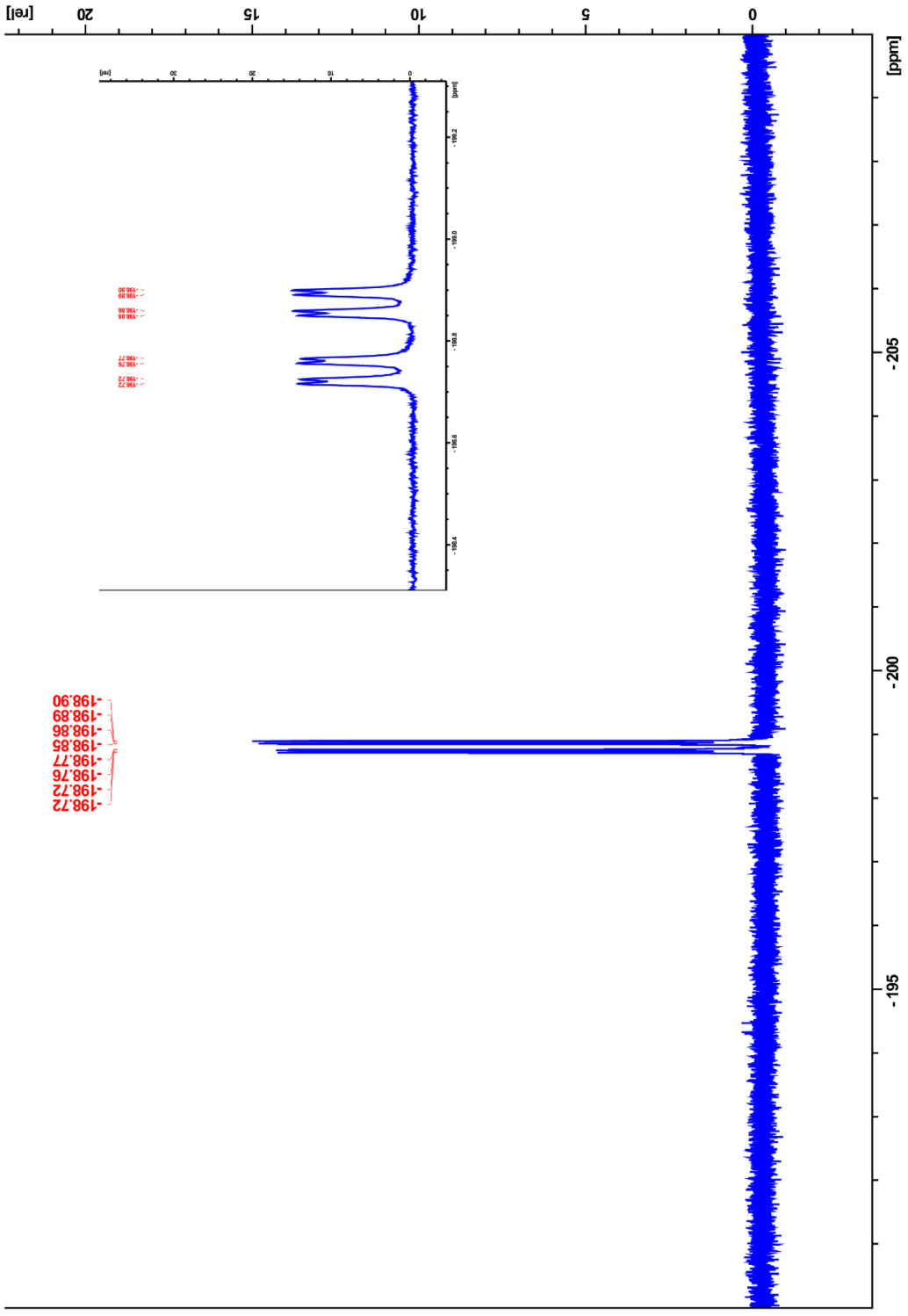


Figure C9: ^{19}F NMR spectrum of 12a (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated) G3G)

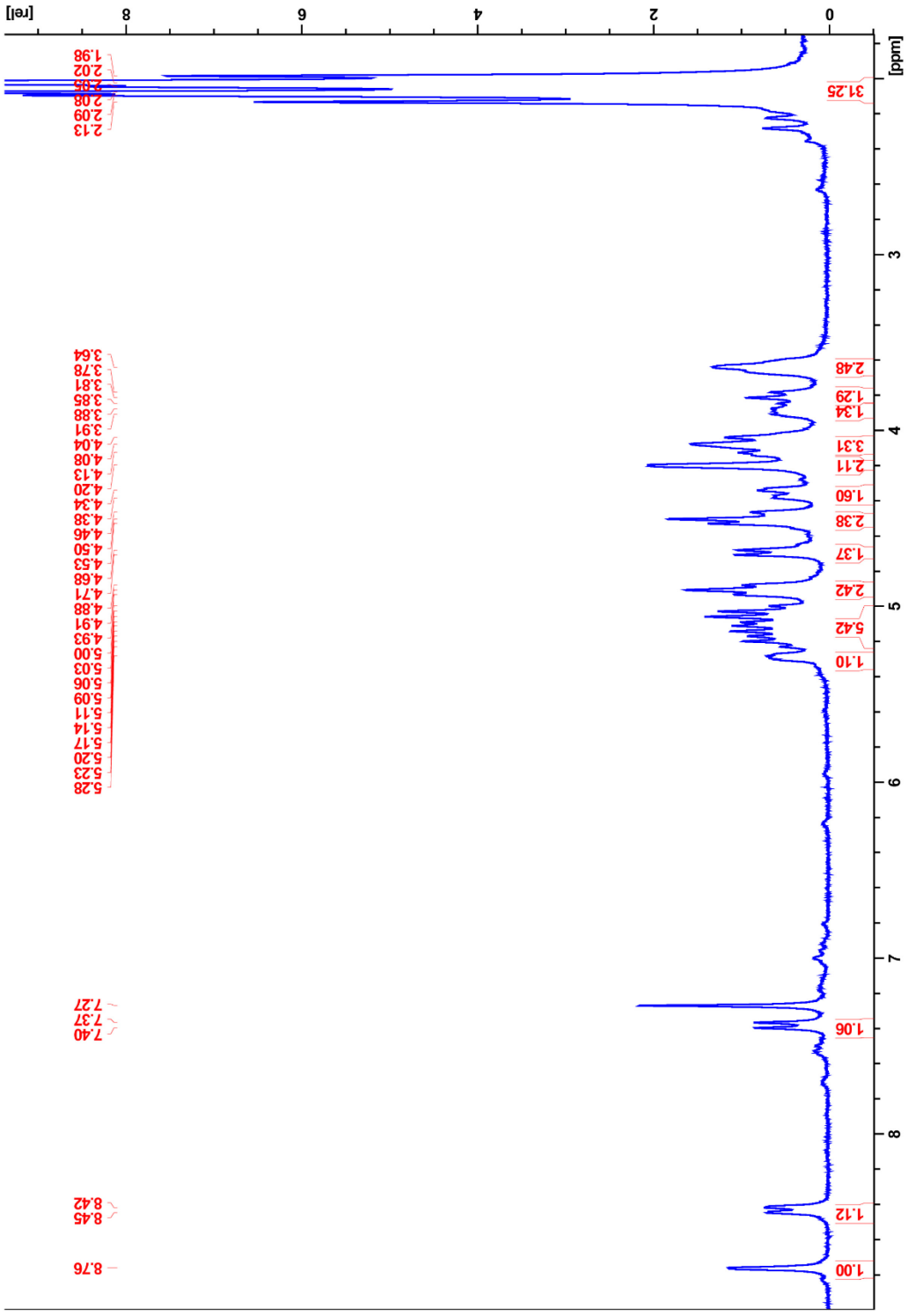


Figure C10: ^{1}H NMR spectrum of 12b (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated) GG3G)

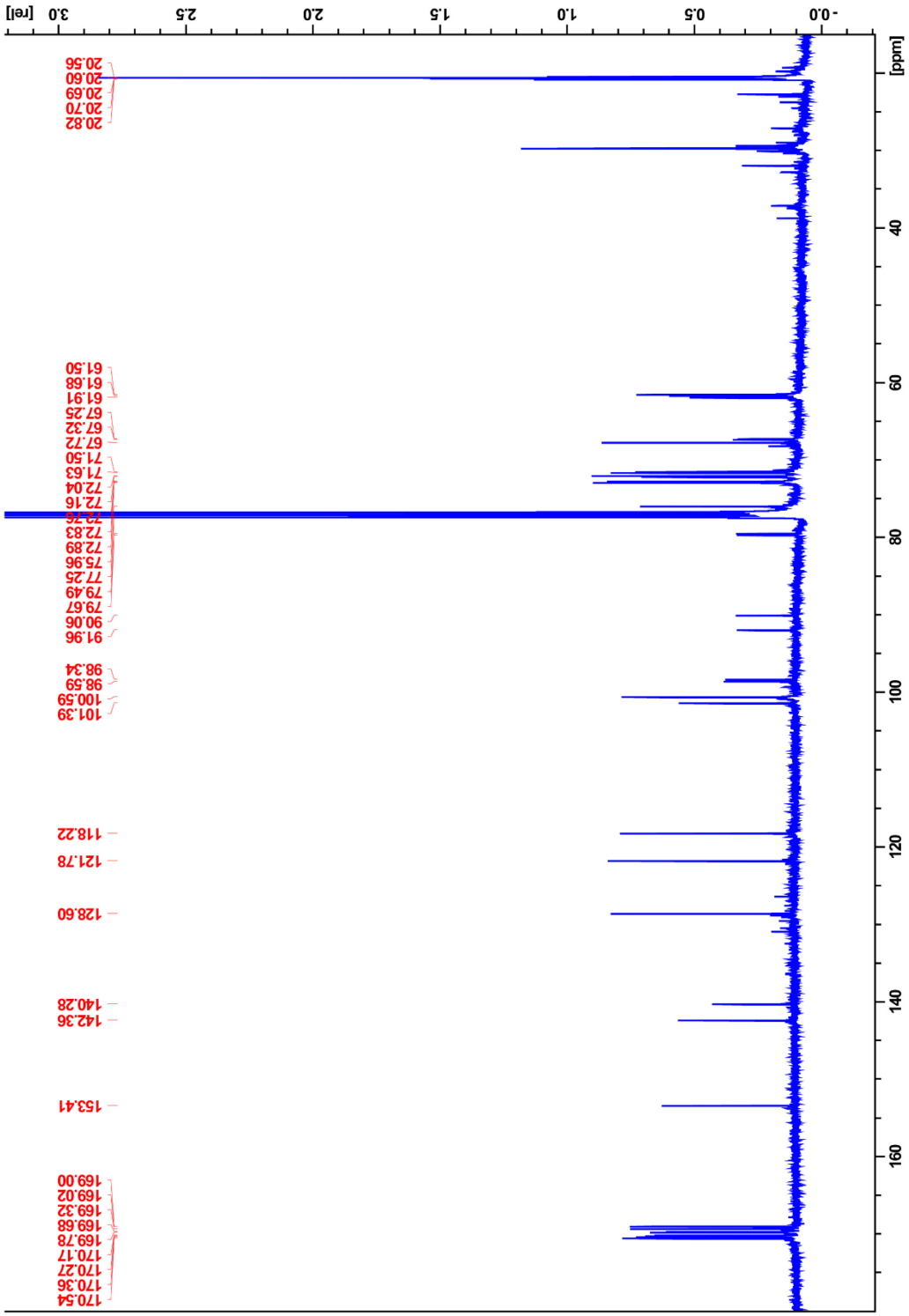


Figure C11: ^{13}C NMR spectrum of 12b (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated) GG3G)

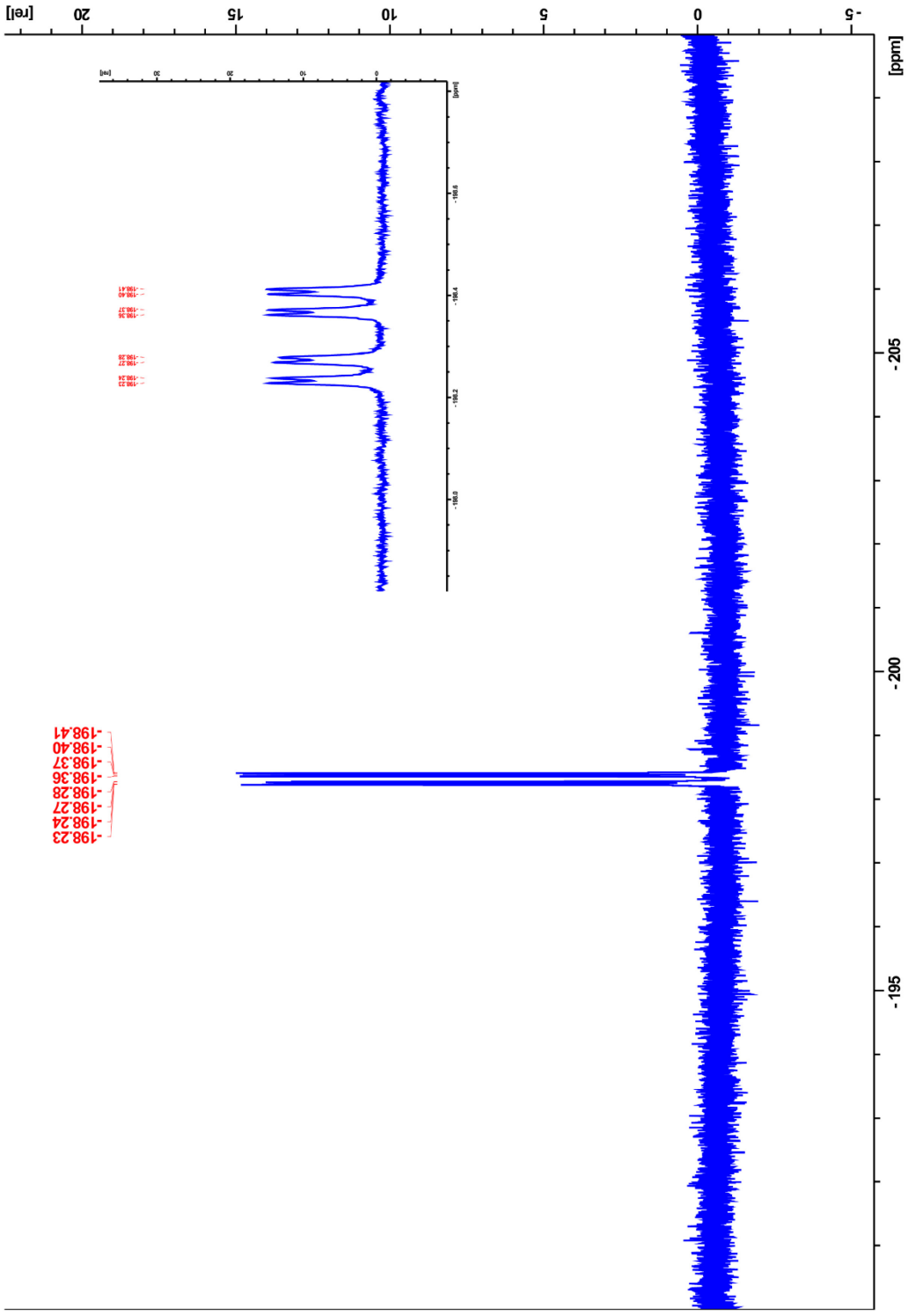


Figure C12: ^{19}F NMR spectrum of 12b (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro-(per-*O*-acetylated) GG3G)

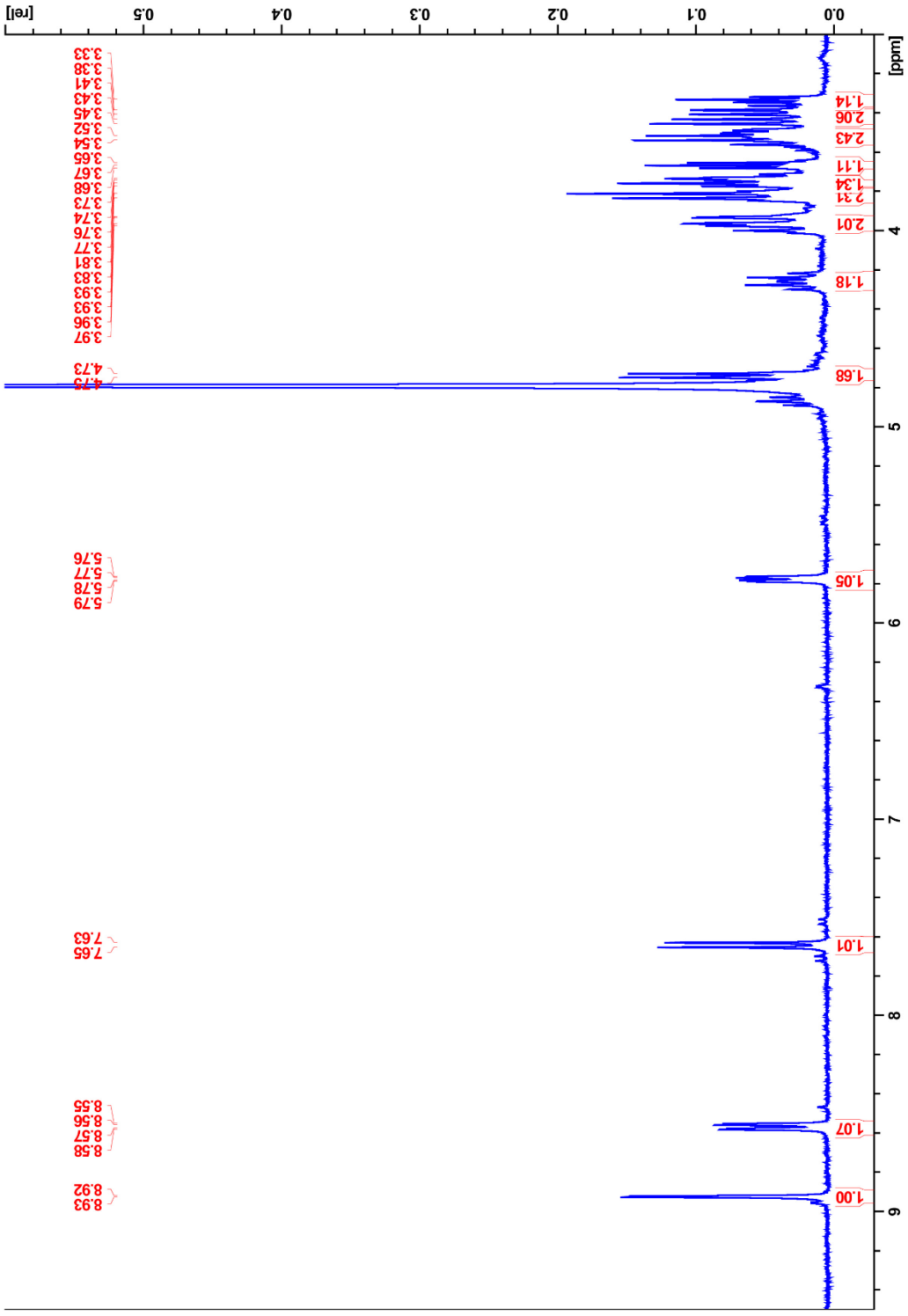


Figure C13: ^{1}H NMR spectrum of 13a (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro G3G)

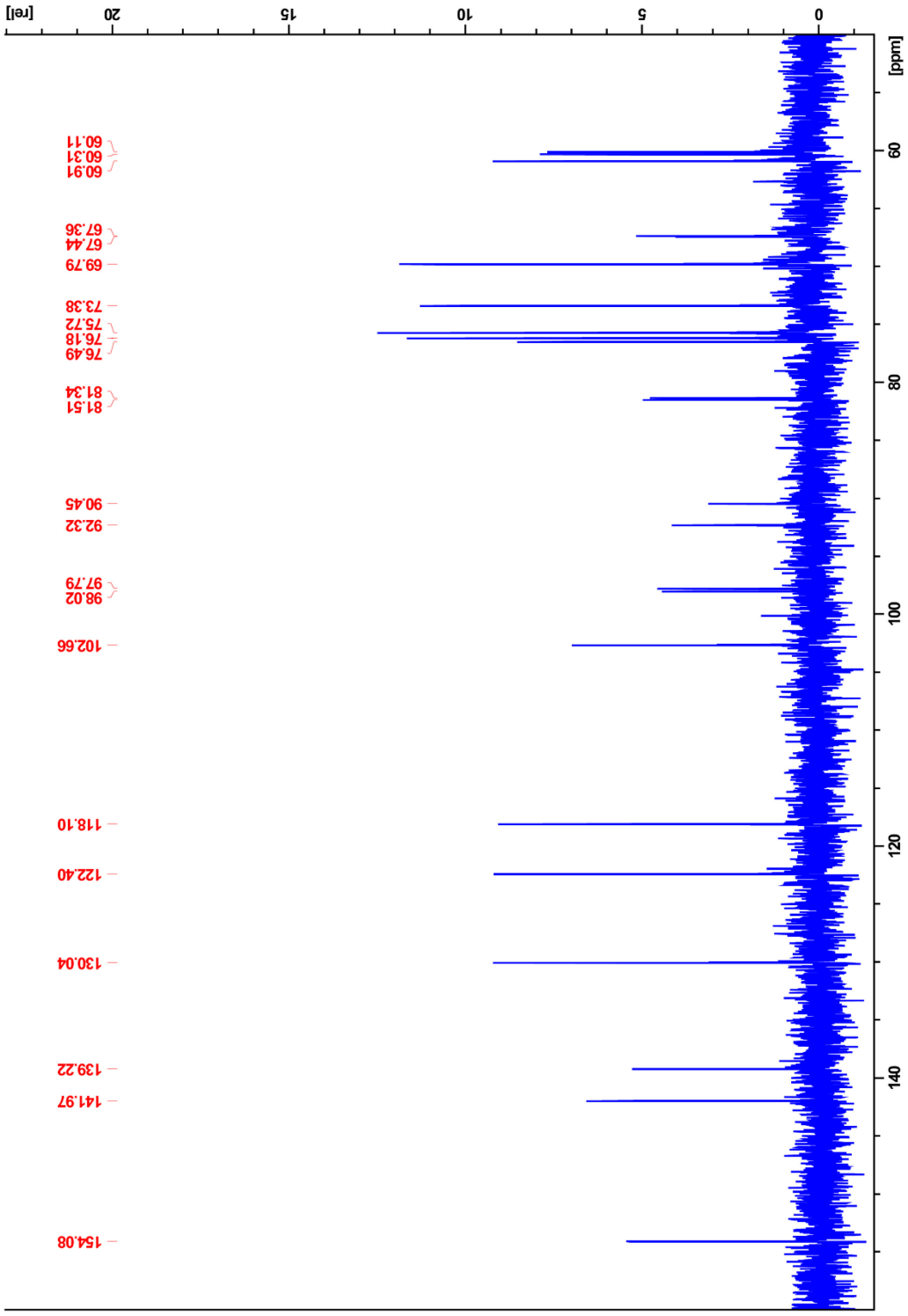


Figure C14: ^{13}C NMR spectrum of 13a (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro G3G)

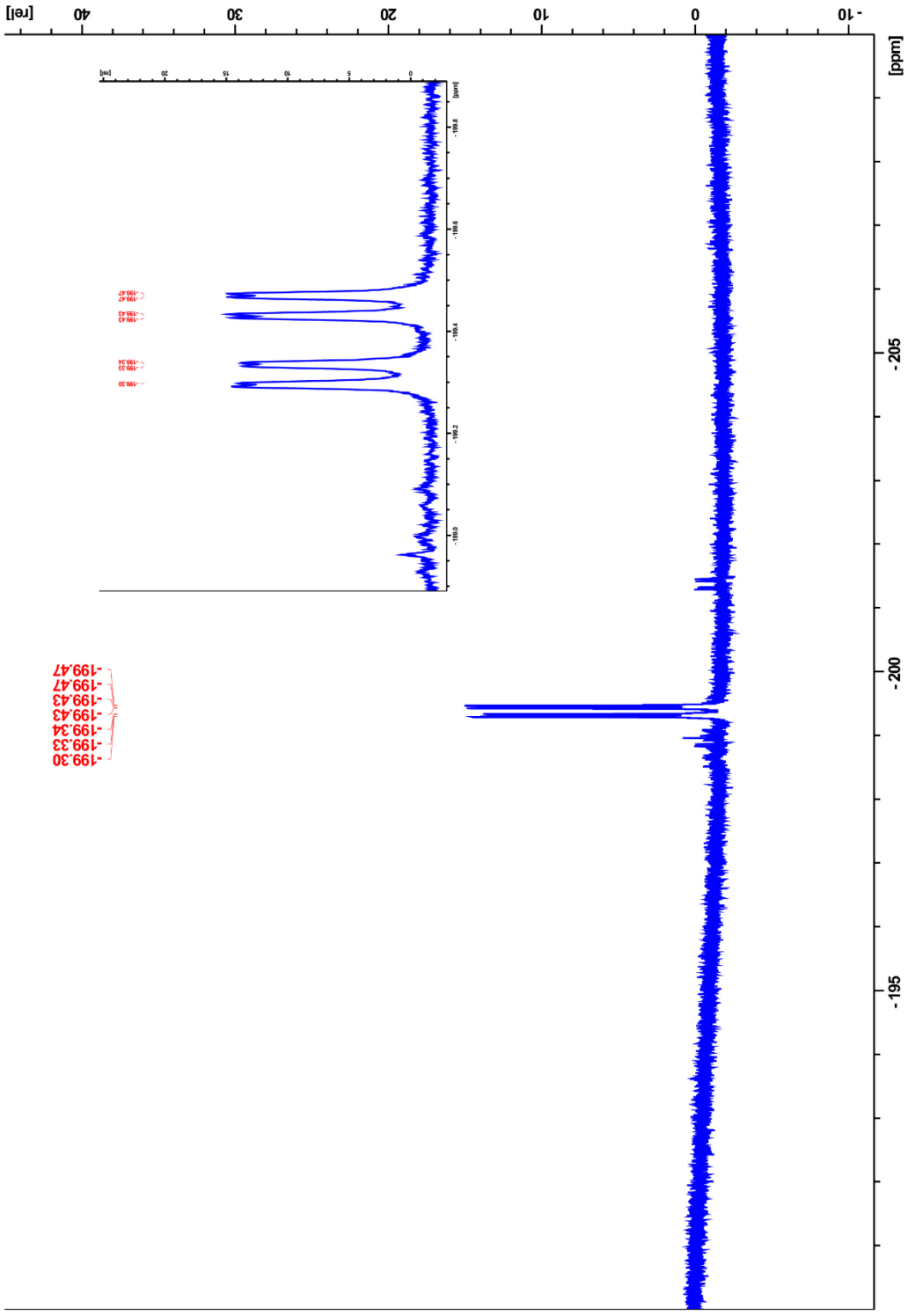


Figure C15: ^{19}F NMR spectrum of 13a (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro G3G)

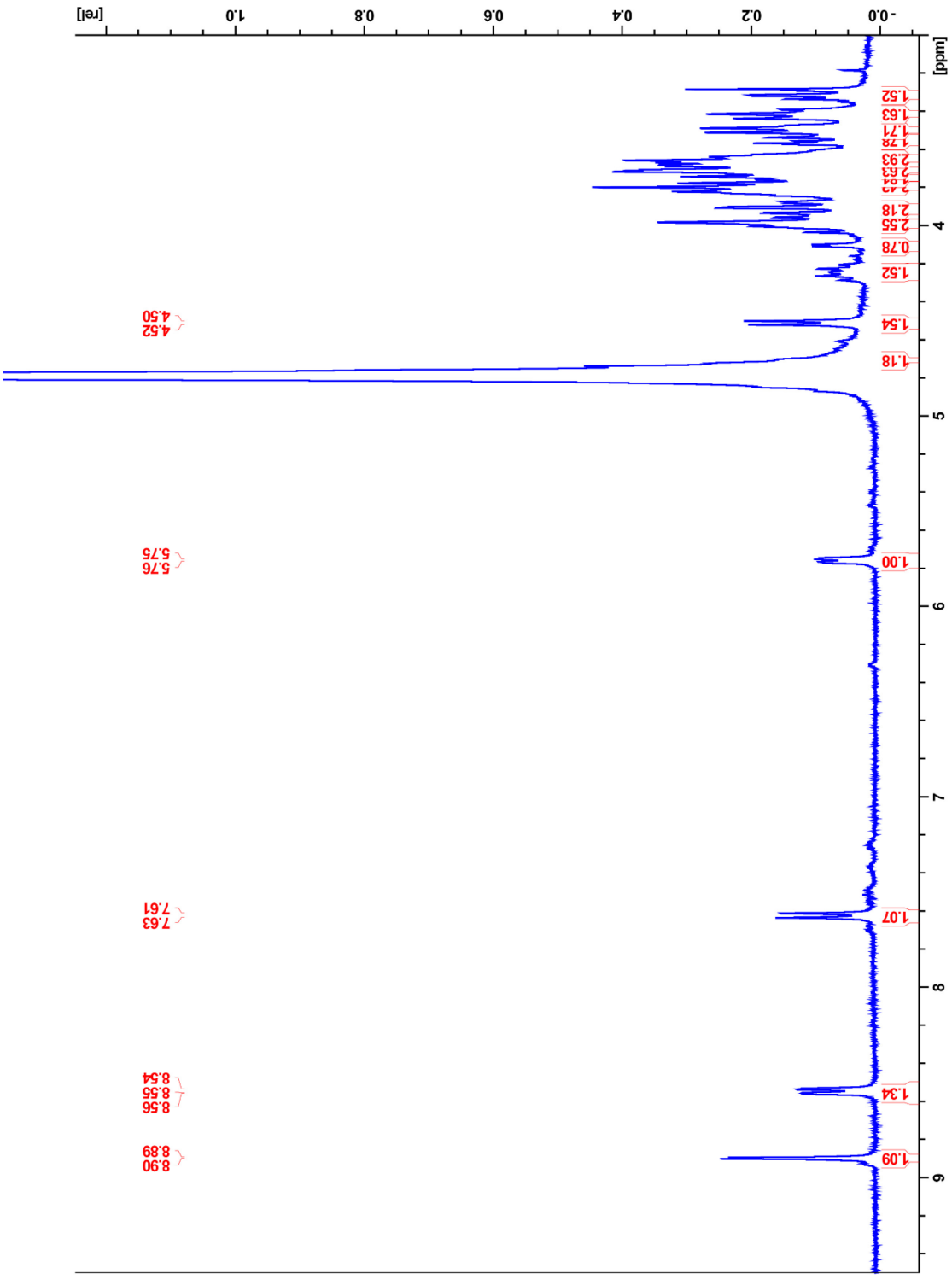


Figure C16: ^{1}H NMR spectrum of 13b (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro GG3G)

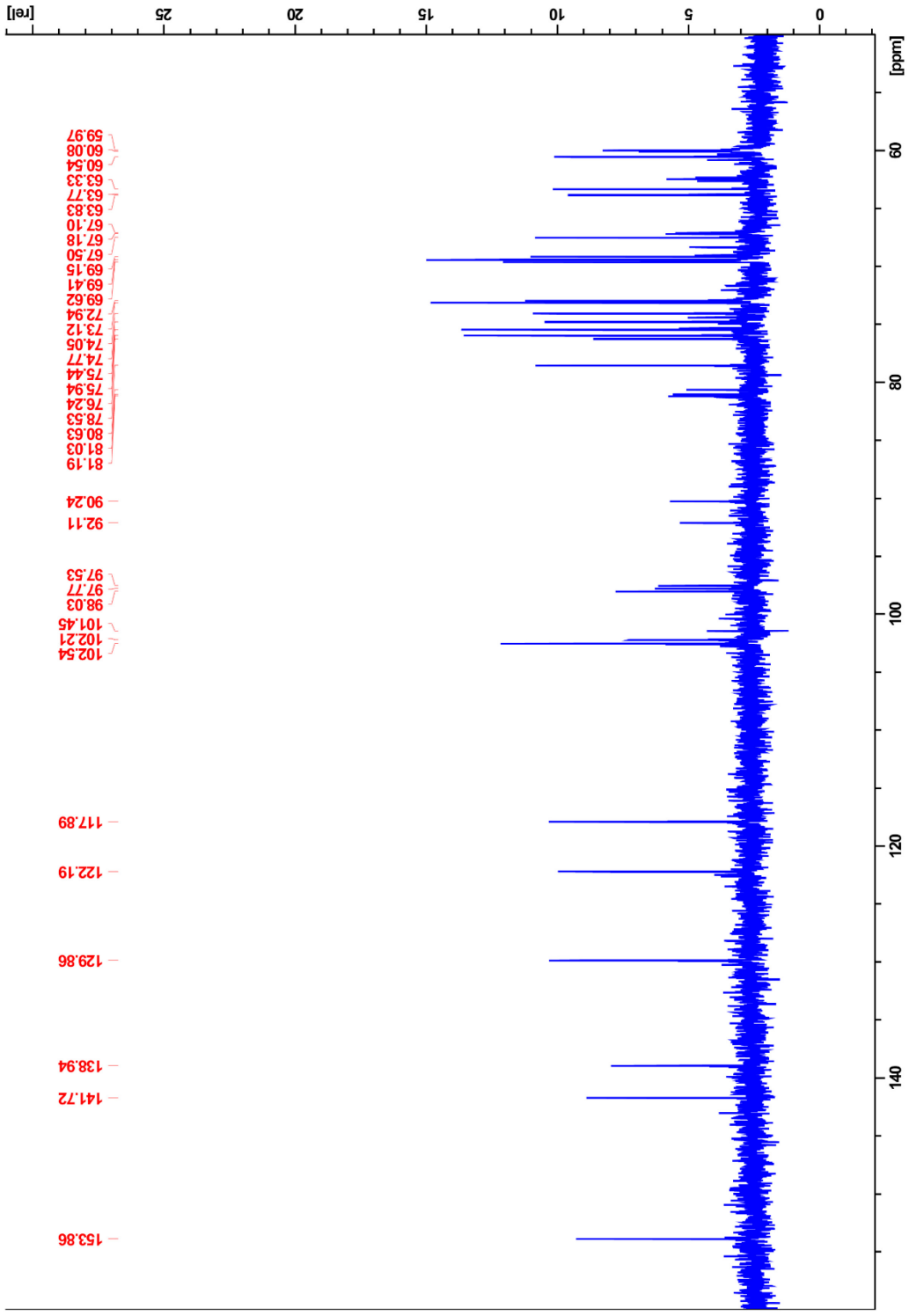


Figure C17: ^{13}C NMR spectrum of 13b (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro GG3G)

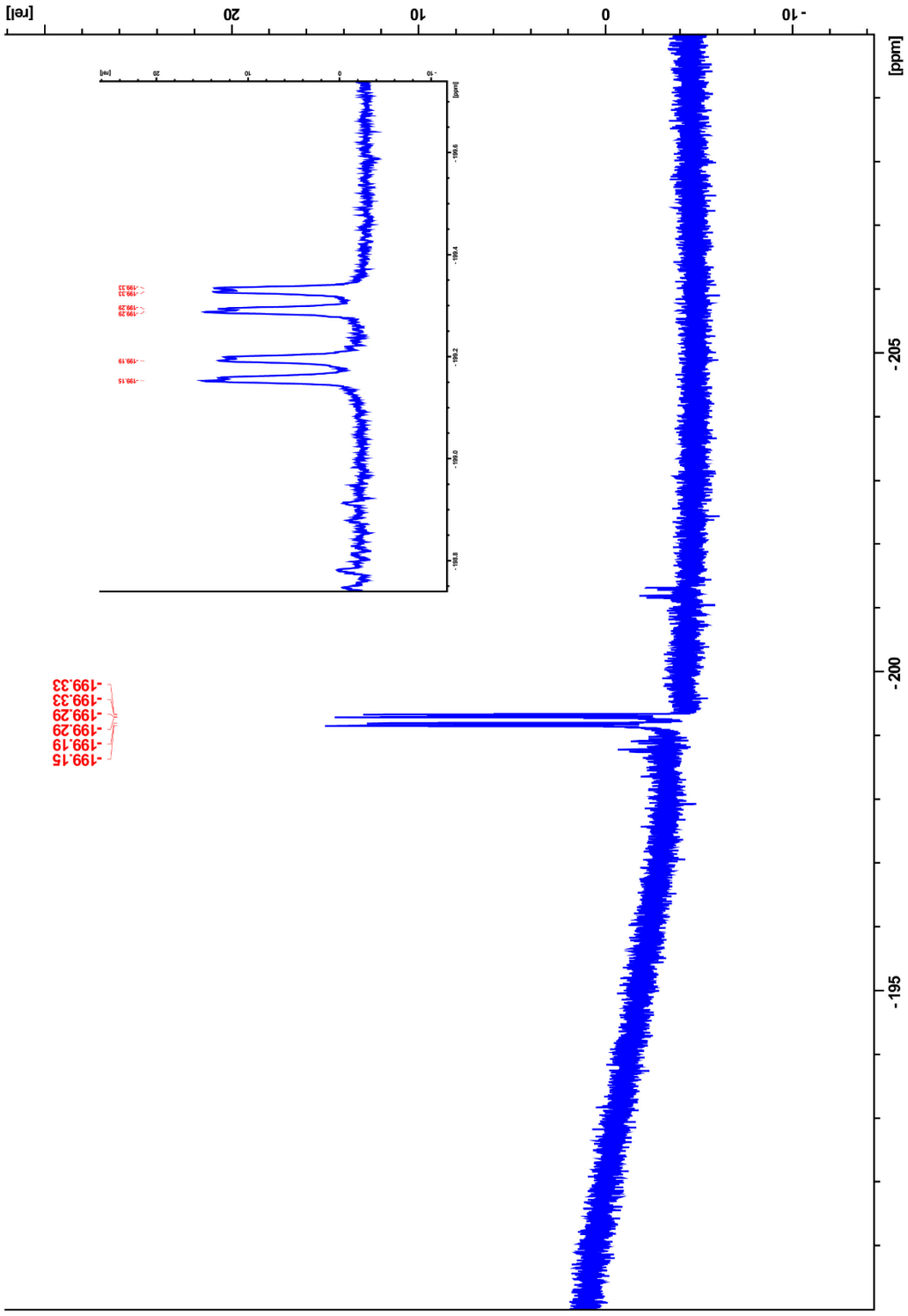


Figure C18: ^{19}F NMR spectrum of 13b (β-2'4'-dinitrophenyl 2-deoxy-2-fluoro GG3G)

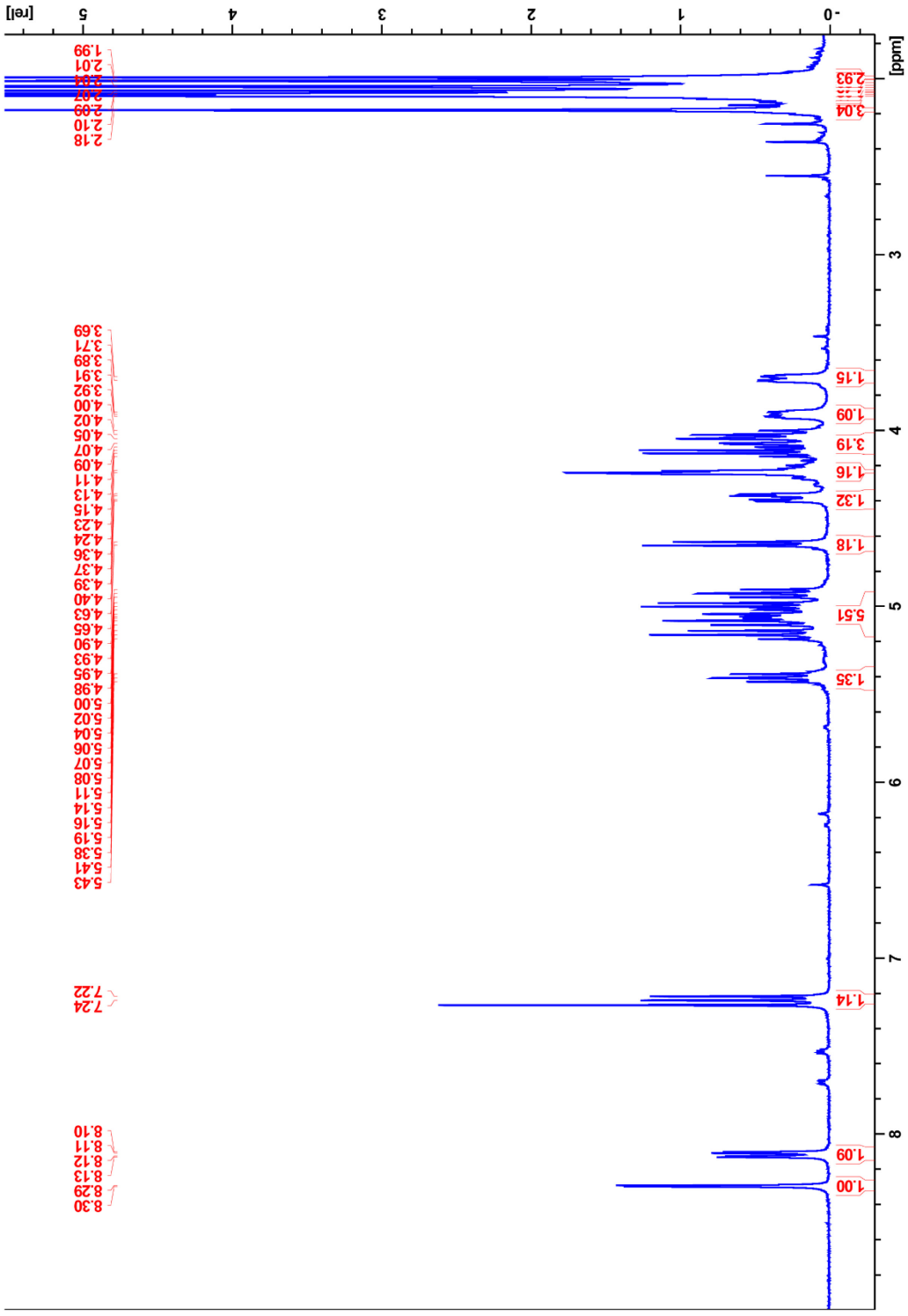


Figure C19: ^{1}H NMR spectrum of 14 (β-2'-chloro-4'-nitrophenyl-(per-*O*-acetylated) G3G)

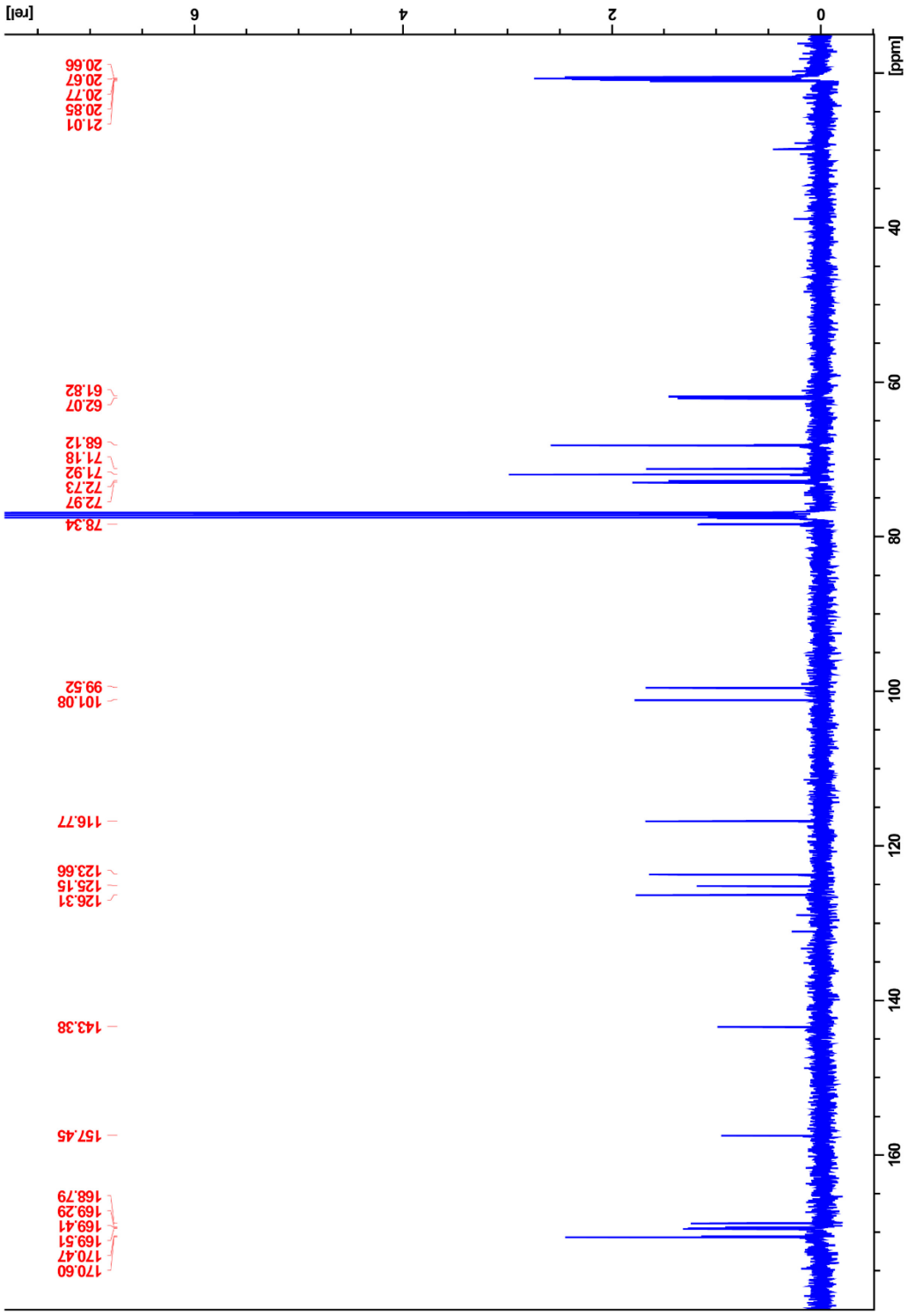


Figure C20: ^{13}C NMR spectrum of 14 (β-2'-chloro-4'-nitrophenyl-(per-*O*-acetylated) G3G)

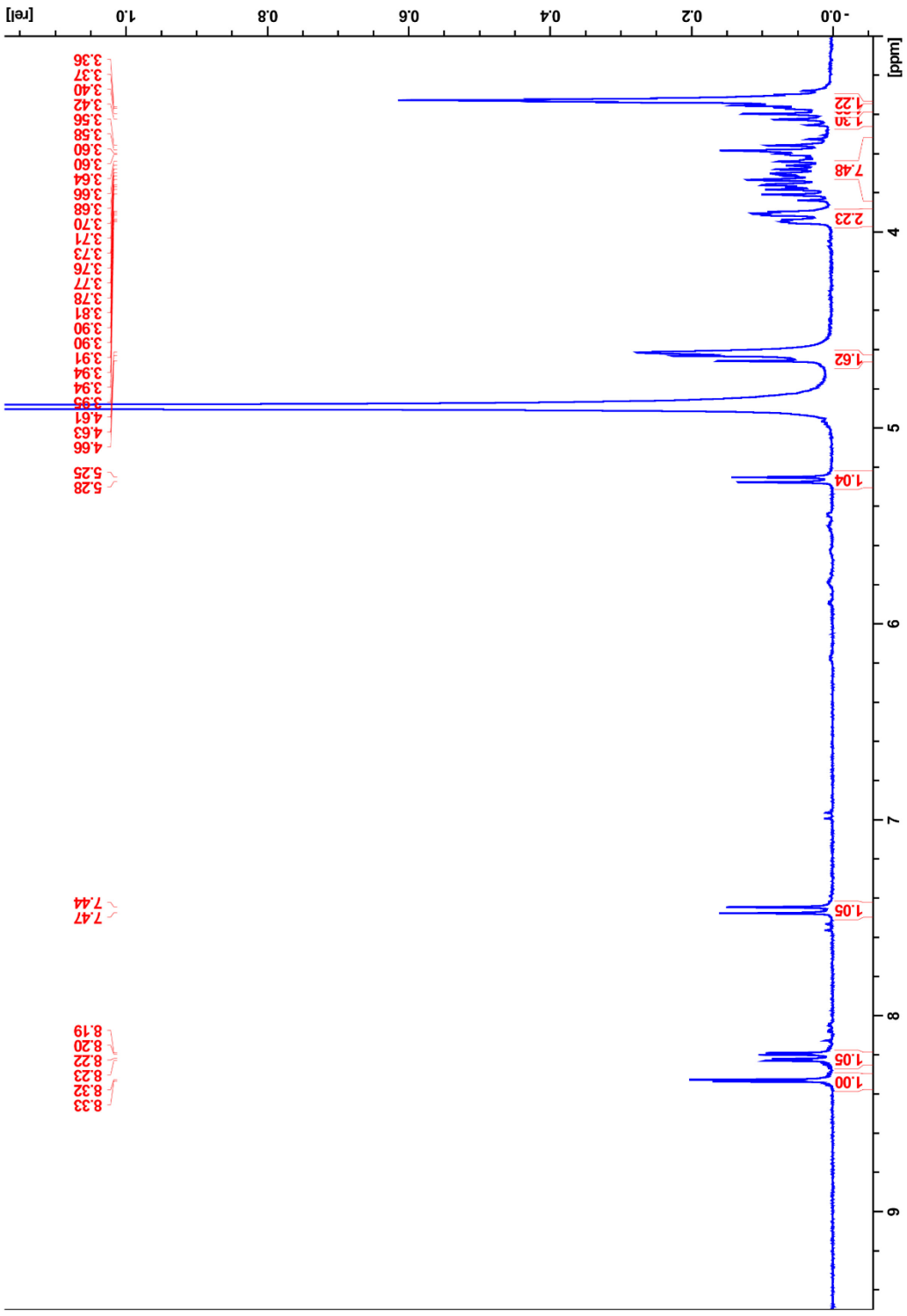


Figure C21: ^{1}H NMR spectrum of 15 (β-2'-chloro-4'-nitrophenyl G3G)

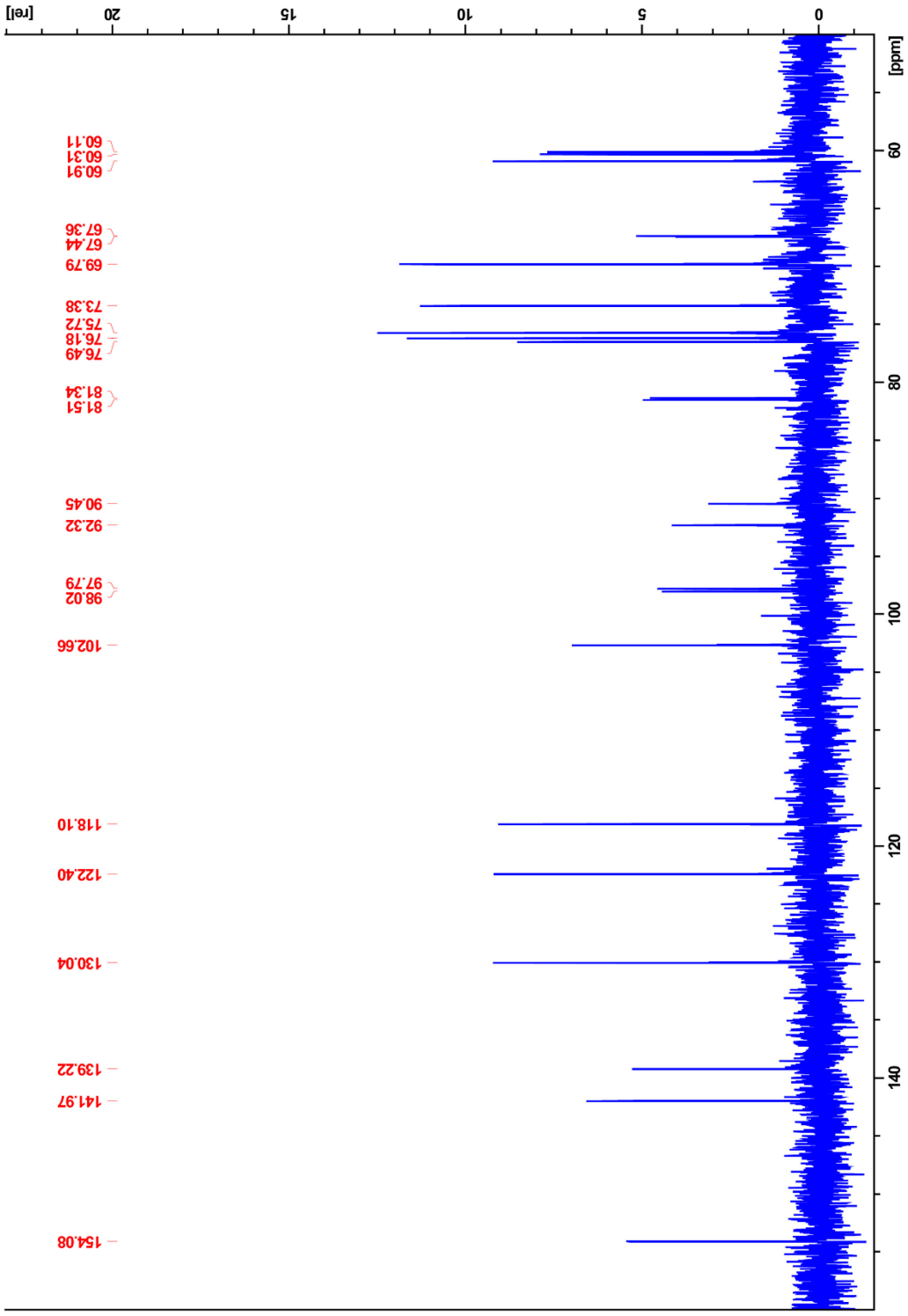


Figure C22: ^{13}C NMR spectrum of 15 (β-2'-chloro-4'-nitrophenyl G3G)

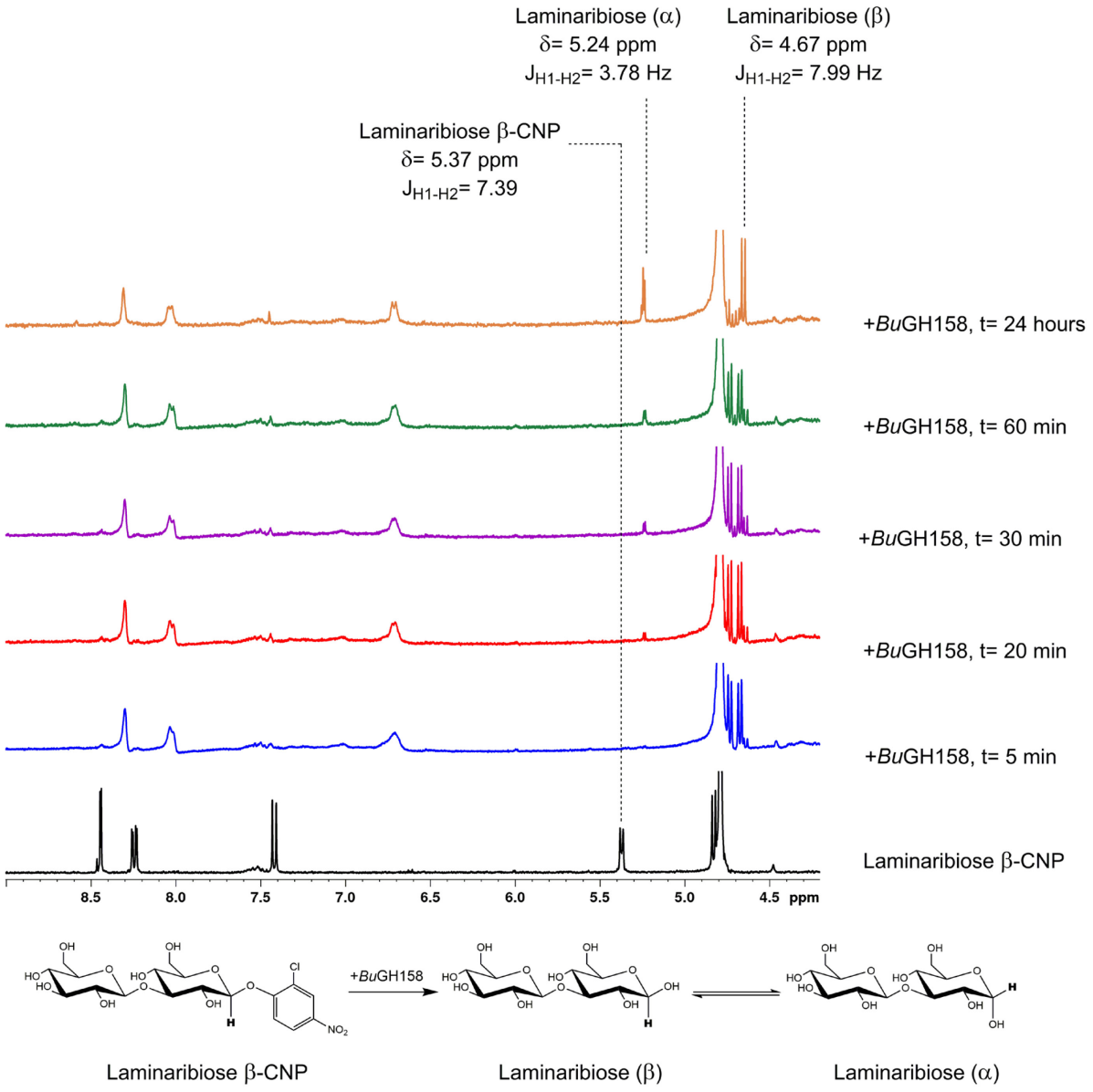


Figure C23: ^{1}H-NMR analysis of BuGH158 mechanism of hydrolysis using 15 as the substrate

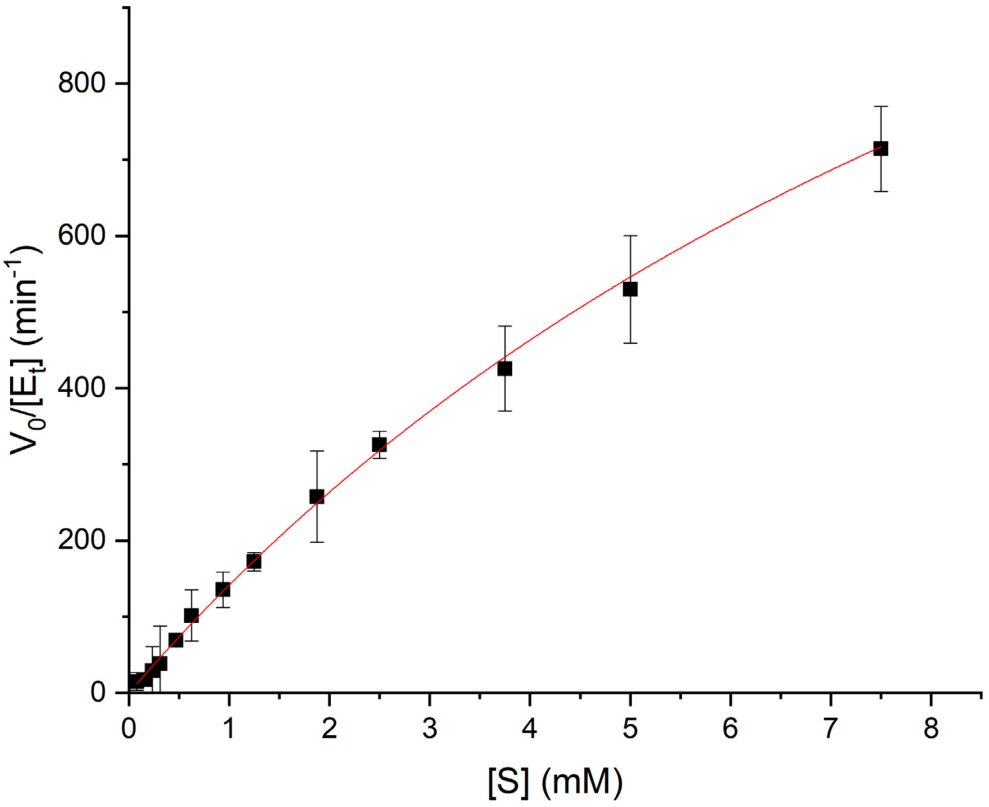


Figure C24: Michaelis-Menten plot for substrate 15 showing the relationship between substrate concentration and activity for BuGH158

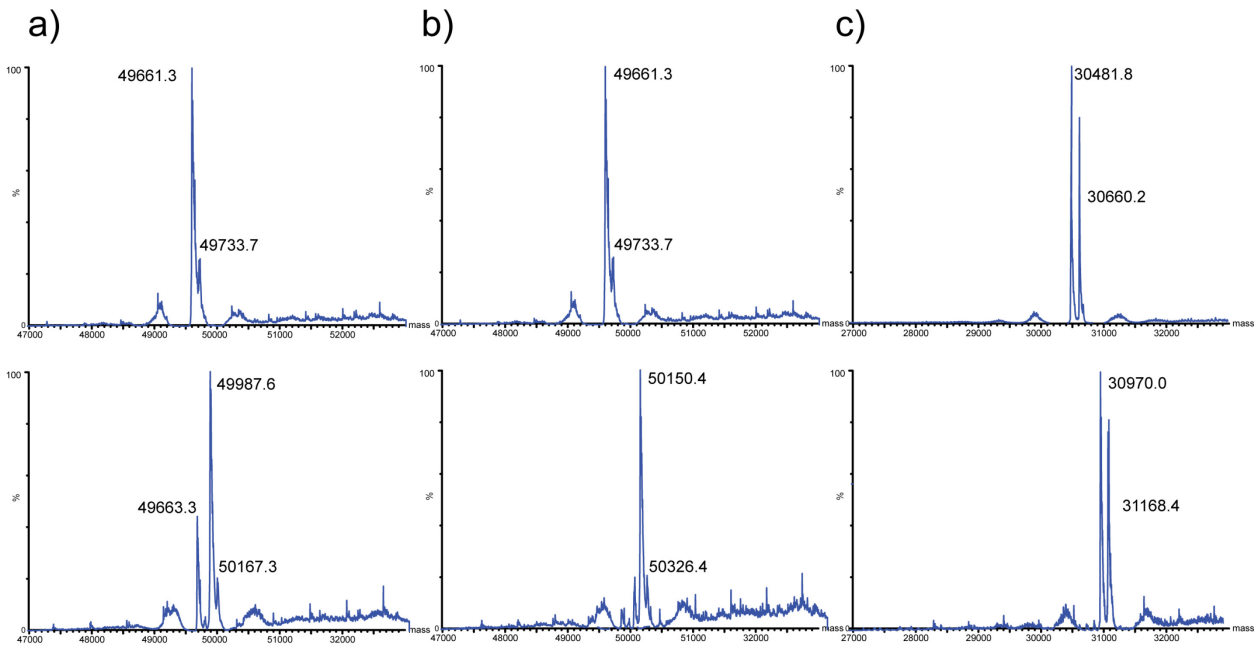


Figure C25: Intact protein mass spectrometry of β-(1,3)/(1,4) glucanases after 2 h incubation. Upper panels are enzymes with no inhibitor, lower panels are enzymes with 2.5 mM inhibitor. (a) *Bacteroides uniformis* GH16 (BuGH16) + disaccharide inhibitor 13a, (b) *Bacteroides uniformis* GH16 (BuGH16) + trisaccharide inhibitor 13b (c) *Bacteroides ovatus* GH16 (BoGH16) + trisaccharide inhibitor 13b