DEVELOPMENT AND APPLICATION OF METHODS FOR STRUCTURAL CHARACTERIZATION OF XYLOGLUCANS AND OTHER COMPLEX GLYCANS

by

SAMI TAPIO TUOMIVAARA

(Under the Direction of William S. York)

ABSTRACT

Plant cell walls are extraplasmalemmal organelles whose structural complexity and mechanical robustness render them especially challenging for detailed structural and functional scrutiny. Xyloglucans are a major class of complex polysaccharides with structural, metabolic and regulatory functions in the cell walls of land plants. The amounts and fine structures of xyloglucans vary among plant taxa, tissue and cell type, and according to the developmental stage of the cell. The range of xyloglucan functions and the functional significance of their structural diversity are currently not well understood. Technological advances are now making it possible to begin understanding these aspects of xyloglucan biology. Towards this end, I generated and purified twenty-eight xyloglucan oligosaccharides by a combination of enzymatic hydrolyses and preparative scale chromatography, and demonstrated their utility as mass spectrometry and chromatography standards and as probes in epitope characterization of thirty-two xyloglucan-binding monoclonal antibodies. Using these oligosaccharides, I first

compared the utility of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and high-pH anion exchange chromatography with pulsed amperometric detection as quantitative methods to study xyloglucan oligosaccharide mixtures. Both methods are suitable for routine analyses of xyloglucan oligosaccharide mixtures, but due to the presence of isomeric structures that leads to ambiguity in the interpretation of mass spectrometry data, and co-elution that leads to ambiguity in the interpretation of chromatography data, these methods should be used in conjunction. Immunofluorescence methods on plant tissue sections are currently one of the most powerful approaches to study the abundances, structures and localizations of xyloglucans in the plant cell walls. Despite their popularity, there has been substantial uncertainty in the interpretation of the observed binding patterns since the precise epitopes of these antibodies had not been determined. The availability of this xyloglucan oligosaccharide collection allowed me to establish the minimum epitopes of these antibodies in great chemical detail using enzyme-linked immunosorbent assays. Knowledge of the binding specificities allows structural interpretation of data from immunofluorescent and other monoclonal antibody utilizing experiments. Collectively, these contributions provide fundamental information that will facilitate our understanding of the biological roles of xyloglucans in the plant cell walls.

INDEX WORDS: Xyloglucan, plant cell wall, monoclonal antibody, enzyme-linked immunosorbent assay, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, immunofluorescence, epitope characterization, size-exclusion chromatography, high-pH anion exchange chromatography with pulsed amperometric detection, nuclear magnetic resonance spectroscopy

# DEVELOPMENT AND APPLICATION OF METHODS FOR STRUCTURAL CHARACTERIZATION OF XYLOGLUCANS AND OTHER COMPLEX GLYCANS

by

SAMI TAPIO TUOMIVAARA

B.S., University of Oulu, Finland, 2002

M.S., University of Oulu, Finland, 2004

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2013

# DEVELOPMENT AND APPLICATION OF METHODS FOR STRUCTURAL CHARACTERIZATION OF XYLOGLUCANS AND OTHER COMPLEX GLYCANS

by

SAMI TAPIO TUOMIVAARA

| | |
|---|---|
| Major Professor: | William S. York |
| Co-advisor: | Michael G. Hahn |
| Committee: | James H. Prestegard |
| | Lance Wells |

Electronic Version Approved:

Maureen Grasso

Dean of the Graduate School

The University of Georgia

December 2013

DEDICATION

To my family, who shaped me into the person I am.

To Chin Fen, who shaped me into the man I am.

ACKNOWLEDGEMENTS

I want to express my gratitude to my academic advisor, Professor William York for his support and guidance during my (lengthy) stay in his laboratory. His rigorous approach to science has had an immense influence on my thinking and maturation as a scientist. I want to thank my committee members, Professors Michael Hahn, James Prestegard, and Lance Wells as well for their help during these years. As my co-advisor and close collaborator, Professor Michael Hahn was especially instrumental in helping me with my project throughout the years. I'm grateful to current and past York lab members, colleagues and friends in Athens who have helped me during my work, especially Dr. Maria Peña, April Harper, Dr. Chenghua Deng, Katrina Saffold, Dr. Breeanna Urbanowicz, Jason Backe, Ameya Kulkarni, Christina Hopper, Dr. Ron Clay, Dr. Kyle Mohler, John Hogan, Dr. René Ranzinger, Dr. Ivana Gelineo-Albersheim, Stefan Eberhard, Dr. John Glushka, Yong Xiang, Dr. Joshua Gilmore, and everyone from ASH. Malcolm O'Neill deserves a separate mention. He has been extremely helpful during the years, giving advice whenever I needed it. Big thanks to my friends back home as well. I'm eternally in debt to my family for their support and love. Finally, the wisdom and love of Chin-Fen Teo (also a Dr.) made me everything I'm today.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

# CHAPTER 1

# INTRODUCTION AND LITERARY REVIEW

## 1.1. Plant Cell Walls

Plant cell walls are chemically and physically complex extracytoplasmic organelles whose organization and dynamics at the molecular scale are only superficially known. Classical microscopic and biochemical investigations to plant cell walls exposed a three-dimensional network of mostly high molecular weight polysaccharides immediately surrounding the plasma membrane of every plant cell. Cell walls are the main determinant of the shape and integrity of plant cells. This is evident from the observation that plant cells whose walls have been enzymatically removed are spherical and eventually rupture if placed in a hypotonic medium (Pritchard, 2007). The intracellular space (cytoplasm and the other compartments) of plant cells has a low water potential due to a large concentration of dissolved and colloidal material (Nobel, 2009). The resulting negative water potential gradient across the plasma membrane leads to an influx of water along a thermodynamic gradient to the cell from the extracellular milieu. This leads to a high osmotic pressure (up to ~1 MPa) in the cytoplasm that tends to expand and spherisize the protoplasm (everything within the plasma membrane). Naturally, the turgor pressure operates in the cells of intact plants as

well, and has many essential functions, including driving the cell growth and expansion (Pritchard, 2007). The structural rigidity of the wall protects the cell against not only this internal turgor pressure, but also the external strain conjoined by the mass of the plant tissue around and on top of it. Sizes, forms, and rheological properties of cells, tissues, organs and ultimately whole plants are largely determined by their cell walls due to the tight continuity of the extraplasmalemmal structures (collectively termed as the apoplast) - the plant stands and falls (literally!) with its cell walls.

Cumulative and combined evidence from the abovementioned as well as genetic and biophysical approaches have displaced this static view of plant cell wall as an inert mechanical restraint with a dynamic perspective, where the wall is a metabolically active and adaptive organelle with pervasive functions in plant physiology (Albersheim et al, 2011). For example, the cell wall structure is “loosened” in a regulated manner to yield to the turgor pressure, leading to the expansion and reshaping of cells undergoing growth, cell division and differentiation (Pritchard, 2007). Moreover, the confining position, adaptable porosity and gelling properties put the cell wall in charge of regulating the traffic of macromolecules and water alike between the cells and their environment. Plant cell walls are also a source of signaling molecules (Cote & Hahn, 1994; Cote et al, 1998; Fry et al, 1993a; Mohnen & Hahn, 1993; Ridley et al, 2001) and an active participant in the defense against fungal infections and other biotic challenges. It has been estimated that cell wall related genes comprise up to

15% of the genome of the model organism *Arabidopsis thaliana* (Carpita et al, 2001; Yong et al, 2005).

Nearly all types of plant cells are surrounded by a cell wall, the gametes being the most notable exception (Albersheim et al, 2011). Several types of cell walls with distinct compositions have been described in different plant taxa. The cell wall types reflect the evolutionary history of plants as well as the various strategies that plants utilize in occupying their diverse habitats (Popper, 2008; Popper et al, 2011). Furthermore, the compositions as well as chemical and physical properties of cell walls are tailored by the individual cells, not only during the initial synthesis but also throughout their lifetime. Growing plant cells synthesize primary cell walls that are biologically responsive organelles and that can accommodate the expansion of the growing cells they encase (Albersheim et al, 2011). Cells that require mechanical strength, for example to sustain water transport, synthesize additional layers of wall, collectively called the secondary cell wall. The secondary cell wall is deposited interior to the non-expanding primary cell wall and can become so thick that it eventually consumes most of the space interior to the primary cell wall, at the expense of the size of the cytoplasm. Due to the impenetrability of a typical secondary cell wall to even small solutes, these cells are metabolically inactive (dead) at maturity and only have supportive function.

The properties of many plant based products are largely determined by their cell walls (Albersheim et al, 2011). For example, the durability of construction timber arises solely from the mechanical and chemical robustness of

the secondary cell walls. Also, both the texture and shelf-life of vegetables depend largely on the integrity of the cell walls. Humans have utilized plants and thus plant cell walls throughout the history of our species. We are currently in the stage where our understanding of the genetics, biochemistry, physiology, and ecology of plants allows us to begin designing and genetically manipulating plant species for our benefit (Akpinar et al, 2013; Bock, 2013; Yoon et al, 2013).

### 1.2. Composition of Plant Cell Walls

The defining and unifying structural feature of plant cell walls is the presence of cellulose. Chemically, cellulose is a homopolymer of (1→4)-linked β-D-Glc*p* residues, but its distinctive property is the physical form it takes in the cell wall. In plants, approximately 36 individual cellulose molecules coalesce by intermolecular hydrogen bonds into dense long cylindrical microfibrils (5–15 nm wide and 1–5 μm long) as they are synthesized and extruded by the plasma membrane bound cellulose synthase complex (Albersheim et al, 2011). The high tensile strength of cellulose and the noticeable directional coherence of the microfibrils in the cell wall has led to the notion that cellulose is the main determinant of the mechanical properties of plant cell walls (Stone, 2007).

Besides cellulose, plant cell walls contain other carbohydrate and non-carbohydrate components in varying proportions (Waldron & Faulds, 2007). Pectins are a family of complex polysaccharides whose most distinguishing feature is the high proportion of anionic galacturonic acid residues in their backbone (Caffall & Mohnen, 2009). These anionic residues bind divalent cations

such as $Ca^{2+}$ that can bridge two pectin polysaccharides. An additional mode of interaction between pectins occurs via a covalent borate cross-link between two apiosyl residues (Ishii & Matsunaga, 1996; O'Neill et al, 1996). The large water-holding capacities and gelling properties of some pectins *in vitro* have been extrapolated to their function *in muro*, and it is thought that pectins determine also the porosity and hydration of the wall, hence being the main determinant of the transport properties of the cell wall (Willats *et al*., 2001).

Hemicelluloses are characterized by backbones that are composed of β(1→4)-linked D-pyranosyl residues with equatorial configuration of the hydroxyl groups at *C*-1 and *C*-4 that are involved in the backbone glycosidic bonds (Scheller & Ulvskov, 2010). Hemicelluloses are generally thought to act as cross-links, tethers or spacers between cellulose microfibrils. Several types of hemicelluloses have been described, including xylans and mannans, both of which exhibit diversity in their fine structures. For example, xylans can be substituted by arabinosyl, glucuronosyl, and 4-*O*-methylglucuronosyl residues. Mannans can be classified either as homomannans, glucomannans, galactomannans and galactoglucomannans, depending on their chemical structures. Most hemicelluloses bear *O*-acetyl groups either in the backbone or side-chain residues (Scheller & Ulvskov, 2010).

### 1.3. Xyloglucan

Xyloglucans (XyGs) are another class of hemicellulosic polysaccharides in the plant cell walls. Development of methods to study XyG structure and function is

the common theme of the research described in this dissertation. XyGs are some of the most abundant and common polysaccharides in plant cell walls (Fry, 1989). XyG structure is characterized by a β(1→4)-linked D-Glc*p* backbone with frequent side-chains initiated by α(1→6)-linked D-Xyl*p* residues. Most of the xylosyl residues are extended to create larger side-chains, whose structures and distribution depend on the plant species. The most common extension to the Xyl*p* residue is a β-D-Gal*p* residue. XyG polysaccharide from tamarind seed is composed of four major subunits whose structures are shown in Figure 1. A large number of other side-chain structures have been described, some of which are further discussed in Chapters 2 and 3.

Cumulative evidence shows that XyG is present in virtually all plant species studied so far, representing all major lineages of land plants (Embryophytes) (Popper, 2008). XyG specific enzymes, including glycosyl hydrolases, glycosyl transferases and especially endotransglycosylases, form large groups of enzymes in plants (Baumann, 2007). These observations suggest that there is a strong evolutionary pressure for maintaining XyG as a component in the cell wall. A surprising finding of a XyG-less *Arabidopsis thaliana* mutant that is viable under controlled growth conditions, has fueled speculation about the necessity and the full range of functions of XyGs in plants (Cavalier et al, 2008).

XyG analysis is hampered, as is plant cell wall polysaccharide analysis generally, by the lack of suitable standard molecules for mass spectrometry and chromatography analyses, and for the characterization of XyG-binding proteins

and XyG-processing enzymes. Generation and purification of XyG oligosaccharides for these and other analytical purposes is described in Chapter 2. Appendices B and C describe additional work utilizing mass spectrometry and other methods to study XyGs and XyG-active enzymes.

### 1.4. Ultrastructure of Plant Cell Walls

The resistance of plant cell walls to mechanical, chemical and biological degradation, or its recalcitrance, renders cell walls an especially challenging target for detailed analysis. Correlative changes in the gross compositions and mechanical properties of cell walls along evolutionary and various genetically controlled pathways, such as growth and ripening, as well as environmental conditions, such as pathogen and water stress, have been exhaustively catalogued (Percy et al, 1996). This solid understanding of the general composition and morphology of the cell wall contrasts radically the rudimentary notions we have of the mechanistic underpinnings, coordination, and cause-effect relationships during wall biogenesis, growth and function. Decades of painstaking analyses have nevertheless yielded enough information to stitch together models that can recapitulate at least some of the essential structural and functional features of plant cell walls. Models of plant cell wall architecture emphasize its chemical diversity, physical heterogeneity, and the covalent and non-covalent interactions between its various macromolecular components (Carpita & Gibeaut, 1993; Cosgrove, 2001; Keegstra et al, 1973; Somerville et al, 2004). Here, insoluble cellulose microfibrils are embedded in amorphous and hydrated matrix of hemicelluloses and pectins whose interactions with solvent

and other molecules can be tunable. Examples of such enzymatic modifications that change solubility properties of matrix components are de-methylesterification of acidic polygalacturonan and degalactosylation of XyG, that both lead to lower solubility of the corresponding polysaccharide.

The heterogenous composition of plant cell walls raises the possibility for spatially overlapping and interwoven, yet functionally independent polymer networks. Several covalent and non-covalent polymer networks have indeed been implicated *in vitro* (Mishima et al, 1998) with more-or-less convincing evidence of their functions *in muro*. The archetype of such a network is the cellulose-XyG network which has been assigned, mainly by biochemical evidence, the major load-bearing function in plants with XyG as the major hemicellulose (Whitney et al, 1999). Recent genetic evidence from viable *Arabidopsis* mutant that lack detectable XyG altogether, yet shows only minor growth phenotype in controlled environment (Cavalier et al, 2008), suggest at minimum redundancy and plasticity in the division of labor in the plant cell wall. Besides its direct role as a tether between cellulose microfibrils, XyG has other potential roles in the cell wall. From *in vitro* studies, the tendency of cellulose microfibrils to aggregate is well documented. XyG can be adsorbed to the surface of the microfibril to prevent this, in addition to providing optimal separation of the microfibrils *in muro* (Fry, 1989). Intimately linked to cell wall expansion in growing cells is the yielding of the XyG-cellulose network to turgor pressure. Peter Albersheim hypothesized the existence of XyG endotransglycosylase (XET) activity that breaks and religates XyG polymers in

the plant cell wall, leading to controlled yielding to the turgor pressure without breakdown of the network (Albersheim, 1976). The existence of such enzyme activity was experimentally confirmed later (Nishitani & Tominaga, 1992; Smith & Fry, 1991) and is understood currently to be one of the major enzymatic activities responsible for plant cell wall expansion and remodeling.

Evidence from a variety of approaches suggests two modes of non-covalent interaction between cellulose microfibrils and XyG. Parts of the XyG molecule can be either physically trapped within the nascent microfibril, or adsorbed on the surface of the assembled microfibril by hydrogen bonding (Hayashi et al, 1994). The rest of the XyG molecule is solvated and relatively free to move around, potentially tethering, or cross-linking two cellulose microfibrils. The structure and abundance of XyG varies tremendously according to the types and developmental stages of cells and tissues, as well as various levels of plant taxonomy (Fry, 1989), hinting that XyG functions might differ between plant species. Besides the XyG-cellulose network, other interpolymer networks have been proposed to contribute to the structural properties of the cell wall (Tan et al, 2013).

### 1.5. Monoclonal Antibodies as Structural and Quantitative Probes

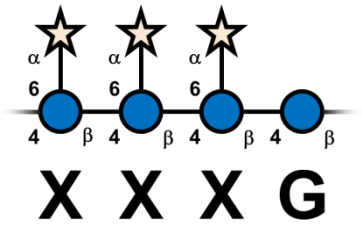
High specificity and affinity towards their epitopes make monoclonal antibodies (mAbs) a versatile probe of molecular structure and topology even in complex aggregates (Knox, 1997). The plant cell wall community has used immunolocalization to visualize structural and topological changes at the

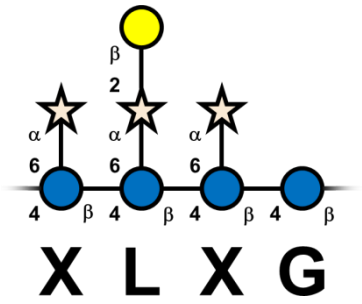
molecular level in plant cell walls with sub-cellular spatial resolution to dissect developmental pathways and responses to genetic manipulation (Knox, 2008). For instance, the CCRC-M series of mAbs, which have been generated by Michael Hahn and colleagues at the Complex Carbohydrate Research Center of the University of Georgia, contains approximately 130 mAbs against various plant cell wall polysaccharides (Pattathil et al, 2010). Some of these mAbs have been used to map changes of plant cell wall ultrastructure and composition in Arabidopsis mutants (Cavalier et al, 2008; Freshour et al, 2003). Complementary efforts for generating mAbs (for example Leeds Monoclonal (LM) and John Innes Monoclonal (JIM) series) against plant cell wall structures have also been initiated (see, for example (Smallwood et al, 1995)).

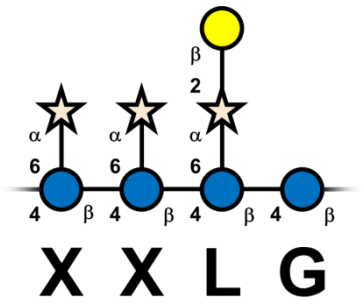
Altogether, approximately 180 mAbs have been generated against the cell wall polysaccharides. The utility of these mAbs for immunolocalization of plant cell wall polysaccharides is acknowledged, but the pace of their generation exceeds that of rigorous epitope characterization. Thus majority of the mAbs generated against plant cell wall components can only be used in a descriptive manner without derivation of more explicit structural information. Moreover, the results from immunolocalization experiments need to be interpreted in the context of the environment of the molecules under scrutiny. The plant cell wall is a complex aggregate, with the conformation, topology, and solvent availability of the epitopes differing from those of free isolated and pure molecular preparations. The deconvolution of these aspects requires additional information regarding the chemical composition of the sample. However, utilization of mAbs

as a chemical and biophysical tool in plant cell wall research remains largely unexplored.

Chapter 3 of this dissertation describes our attempts to increase the usefulness of existing mAbs by characterizing their binding epitopes in detailed manner. Our primary analytical tool for epitope characterization is the enzyme-linked immunosorbent assay (ELISA), of which several variations exist (Davies, 2005). The main approach we use is to immobilize purified and biotin-conjugated oligosaccharides on commercially available NeutrAvidin-coated plates and probe the oligosaccharides with unpurified mAbs secreted in the hybridoma cultures. We have used this approach successfully to characterize the epitopes of thirty-two mAbs. According to our experience and general knowledge of the size of antibody binding sites (Alzari et al, 1988), the minimum epitopes of mAbs are no larger than the oligosaccharides produced by XEG hydrolysis of XyG. Hence, this experiment with small oligosaccharides gives information on the minimum structure that will be recognized by the antibody in a biological sample, such as a plant cell wall. Some applications, such as isothermal titration calorimetry (Livingstone, 1996) and surface plasmon resonance (Hearty et al, 2012) require pure mAbs for their efficient utilization. Towards this end, I also explored mAb purification technologies. These efforts are described in Appendix A.


α
6
4
β
4
X X X G


β
2
α
6
4
β
4
X L X G


β
2
α
6
4
β
4
X X L G

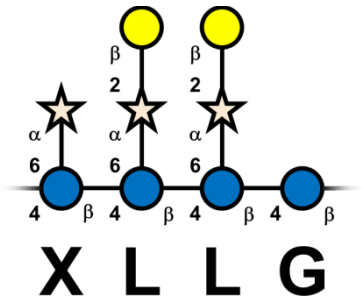

β
2
α
6
4
β
4
X L L G

**Figure 1-1.** XyG polysaccharide from tamarind seed is composed of four major subunits. The number, structures, abundances and the number of distinct subunit structures depend on the species, tissue, and developmental stage of the cell. The nomenclature of xylogucan structures was first described in (Fry et al, 1993b). The subunit structures are rendered according to Consortium of Functional Glycomics guidelines (Varki et al, 2009).

## 1.6. Literature Cited

Akpinar BA, Lucas SJ, Budak H (2013) Genomics approaches for crop improvement against abiotic stress. *TheScientificWorldJournal* **2013**: 361921.

Albersheim P (1976) The primary cell wall. *In:* Bonner J, Varner JE (*eds*) *Plant Biochemistry*, pp. 225–274. Academic Press, New York.

Albersheim P, Darvill AG, Roberts K, Sederoff R, Staehelin A (2010) *Plant Cell Walls*. Garland Science, New York.

Alzari PM, Lascombe MB, Poljak RJ (1988) Three-dimensional structure of antibodies. *Annual Review of Immunology* **6**: 555-580.

Baumann M (2007) XyG-active enzymes: properties, structures and applications (doctoral dissertation). School of Biotechnology, Royal Institute of Technology, Sweden.

Bock R (2013) Strategies for metabolic pathway engineering with multiple transgenes. *Plant Molecular Biology* **83**: 21-31.

Caffall KH, Mohnen D (2009) The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydrate Research* **344**: 1879-1900.

Carpita N, Tierney M, Campbell M (2001) Molecular biology of the plant cell wall: Searching for the genes that define structure, architecture and dynamics. *Plant Molecular Biology* **47**: 1-5.

Carpita NC, Gibeaut DM (1993) Structural models of primary cell walls in flowering plants: Consistency of molecular structure with the physical properties of the walls during growth. *The Plant Journal* **3**: 1-30.

Cavalier DM, Lerouxel O, Neumetzler L, Yamauchi K, Reinecke A, Freshour G, Zabotina OA, Hahn MG, Burgert I, Pauly M, Raikhel NV, Keegstra K (2008) Disrupting two *Arabidopsis thaliana* xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *The Plant Cell* **20**: 1519-1537.

Cosgrove DJ (2001) Wall structure and wall loosening. A look backwards and forwards. *Plant Physiology* **125**: 131-134.

Cote F, Hahn MG (1994) Oligosaccharins: Structures and signal transduction. *Plant Molecular Biology* 2**6**: 1379-1411.

Cote F, Ham KS, Hahn MG, Bergmann CW (1998) Oligosaccharide elicitors in host-pathogen interactions. Generation, perception, and signal transduction. *Sub-cellular Biochemistry* **29**: 385-432.

Davies C (2005) Principles of competitive and immunometric assays (including ELISA). *In*: Wild DG (*ed*) *The Immunoassay Handbook, 4th Edition*, pp. 29-60. Elsevier, Oxford.

Freshour G, Bonin CP, Reiter W-D, Albersheim P, Darvill AG, Hahn MG (2003) Distribution of fucose-containing xyloglucans in cell walls of the mur1 mutant of Arabidopsis. *Plant physiology* **131**: 1602-1612.

Fry SC (1989) The structure and functions of xyloglucan. *Journal of Experimental Botany* **40**: 1-11.

Fry SC, Aldington S, Hetherington PR, Aitken J (1993a) Oligosaccharides as signals and substrates in the plant cell wall. *Plant Physiology* **103**: 1-5.

Fry SC, York WS, Albersheim P, Darvill A, Hayashi T, Joseleau J-P, Kato Y, Lorences EP, Maclachlan GA, McNeil M, Mort AJ, Grant Reid JS, Seitz HU, Selvendran RR, Voragen AGJ, White AR (1993b) An unambiguous nomenclature for xyloglucan-derived oligosaccharides. *Physiologia Plantarum* **89**: 1-3.

Hayashi T, Baba KI, Ogawa K (1994) Macromolecular complexes of xyloglucan and cellulose obtained by annealing. *Plant and Cell Physiology* **35**: 219-223.

Hearty S, Leonard P, O'Kennedy R (2012) Measuring antibody-antigen binding kinetics using surface plasmon resonance. *Methods in Molecular Biology* **907**: 411-442.

Ishii T, Matsunaga T (1996) Isolation and characterization of a boron-rhamnogalacturonan-II complex from cell walls of sugar beet pulp. *Carbohydrate Research* **284**: 1-9.

Keegstra K, Talmadge KW, Bauer WD, Albersheim P (1973) The structure of plant cell walls: III. A model of the walls of suspension-cultured sycamore cells based on the interconnections of the macromolecular components. *Plant Physiology* **51**: 188-197.

Knox JP (1997) The use of antibodies to study the architecture and developmental regulation of plant cell walls. *In:* Kwang WJ (*ed*) *International Review of Cytology*, **Vol 171**, pp. 79-120. Academic Press.

Knox JP (2008) Revealing the structural and functional diversity of plant cell walls. *Current opinion in plant biology* **11**: 308-313.

Livingstone JR (1996) Antibody characterization by isothermal titration calorimetry. *Nature* **384**: 491-492.

Mishima T, Hisamatsu M, York WS, Teranishi K, Yamada T (1998) Adhesion of β-D-glucans to cellulose. *Carbohydrate Research* **308**: 389-395.

Mohnen D, Hahn MG (1993) Cell wall carbohydrates as signals in plants. *Seminars in Cell Biology* **4**: 93-102.

Nishitani K, Tominaga R (1992) Endo-xyloglucan transferase, a novel class of glycosyltransferase that catalyzes transfer of a segment of xyloglucan molecule to another xyloglucan molecule. *Journal of Biological Chemistry* **267**: 21058-21064.

Nobel PS (2009) *Physicochemical and Environmental Plant Physiology*. Elsevier, San Diego.

O'Neill MA, Warrenfeltz D, Kates K, Pellerin P, Doco T, Darvill AG, Albersheim P (1996) Rhamnogalacturonan-II, a pectic polysaccharide in the walls of growing plant cell, forms a dimer that is covalently cross-linked by a borate ester. *In vitro* conditions for the formation and hydrolysis of the dimer. *The Journal of Biological Chemistry* **271**: 22923-22930.

Pattathil S, Avci U, Baldwin D, Swennes AG, McGill JA, Popper Z, Bootten T, Albert A, Davis RH, Chennareddy C, Dong R, O'Shea B, Rossi R, Leoff C, Freshour G, Narra R, O'Neil M, York WS, Hahn MG (2010) A comprehensive toolkit of plant cell wall glycan-directed monoclonal antibodies. *Plant Physiology* **153**: 514-525.

Percy AE, O'Brien IEW, Jameson PE, Melton LD, MacRae EA, Redgwell RJ (1996) Xyloglucan endotransglycosylase activity during fruit development and ripening of apple and kiwifruit. *Physiologia Plantarum* **96**: 43-50.

Popper ZA (2008) Evolution and diversity of green plant cell walls. *Current Opinion in Plant Biology* **11**: 286-292.

Popper ZA, Michel G, Herve C, Domozych DS, Willats WG, Tuohy MG, Kloareg B, Stengel DB (2011) Evolution and diversity of plant cell walls: From algae to flowering plants. *Annual Review of Plant Biology* **62**: 567-590.

Pritchard J (2007) Turgor Pressure. *In:* Roberts K (*ed*) *Handbook of Plant Science*, pp. 148-151. John Wiley & Sons, New York.

Ridley BL, O'Neill MA, Mohnen D (2001) Pectins: Structure, biosynthesis, and oligogalacturonide-related signaling. *Phytochemistry* **57**: 929-967.

Scheller HV, Ulvskov P (2010) Hemicelluloses. *Annual Review of Plant Biology* **61**: 263-289.

Smallwood M, Martin H, Knox JP (1995) An epitope of rice threonine- and hydroxyproline-rich glycoprotein is common to cell wall and hydrophobic plasma-membrane glycoproteins. *Planta* **196**: 510-522.

Smith RC, Fry SC (1991) Endotransglycosylation of xyloglucans in plant cell suspension cultures. *Biochemical Journal* **279**: 529-535.

Somerville C, Bauer S, Brininstool G, Facette M, Hamann T, Milne J, Osborne E, Paredez A, Persson S, Raab T, Vorwerk S, Youngs H (2004) Toward a systems approach to understanding plant cell walls. *Science* **306**: 2206-2211.

Stone B (2007) Cellulose: Structure and distribution. *In:* Roberts K (*ed*) *Handbook of Plant Science*, pp. 1106-1113. John Wiley & Sons, New York.

Tan L, Eberhard S, Pattathil S, Warder C, Glushka J, Yuan C, Hao Z, Zhu X, Avci U, Miller JS, Baldwin D, Pham C, Orlando R, Darvill A, Hahn MG, Kieliszewski MJ, Mohnen D (2013) An Arabidopsis cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. *The Plant Cell* **25**: 270-287.

Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME (2009) *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Waldron KW, Faulds CB (2007) Cell wall polysaccharides: Composition and structure. *In:* Kamerling H (*ed*) *Comprehensive Glycoscience*, pp. 181-201. Elsevier, Oxford.

Whitney SE, Gothard MG, Mitchell JT, Gidley MJ (1999) Roles of cellulose and xyloglucan in determining the mechanical properties of primary plant cell walls. *Plant physiology* **121**: 657-664.

Yong W, Link B, O'Malley R, Tewari J, Hunter CT, Lu CA, Li X, Bleecker AB, Koch KE, McCann MC, McCarty DR, Patterson SE, Reiter WD, Staiger C, Thomas SR, Vermerris W, Carpita NC (2005) Genomics of plant cell wall biogenesis. *Planta* **221**: 747-751.

Yoon JM, Zhao L, Shanks JV (2013) Metabolic engineering with plants for a sustainable biobased economy. *Annual Review of Chemical and Biomolecular Engineering* **4**: 211-237.

# CHAPTER 2

# GENERATION AND STRUCTURAL VALIDATION OF A LIBRARY OF DIVERSE XYLOGLUCAN-DERIVED OLIGOSACCHARIDES FOR ANALYTICAL PURPOSES [1]

[1] Tuomivaara, S.T. and W.S. York. To be submitted to *Carbohydrate Research.*

## 2.1. Abstract

Xyloglucans are complex polysaccharide components in the cell walls of land plants with functions in cell growth and expansion, energy metabolism and signaling. Quantitation as well as elucidation of their structure-function relationships requires structurally well-defined and pure oligosaccharides. As an illustration of our laboratory's continuing efforts in structural and quantitative analysis of plant cell wall components, we have prepared and characterized a large collection of structurally diverse and homogenous xyloglucan oligosaccharides for various analytical purposes. Here, we present a comprehensive workflow for the preparation of such oligosaccharides in preparative scale using enzymatic treatments and size-exclusion chromatography. We used mass spectrometry, high-pH anion exchange chromatography and nuclear magnetic resonance spectroscopy for structural validation of these oligosaccharides. The availability of this collection of pure xyloglucan oligosaccharides enabled us to evaluate matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and high-pH anion exchange chromatography with pulsed amperometric detection in quantitative analysis of xyloglucan oligosaccharide mixtures. Whereas high-pH anion exchange chromatography with pulsed amperometric detection is suitable for absolute quantification, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry can be only used for relative quantification. Nevertheless, due to isomerism and co-elution of the oligosaccharides, unambiguous analysis requires the use both of these orthogonal methods.

## 2.2. Introduction

### 2.2.1. Plant Cell Walls and Xyloglucan

Plant cell walls are extracytoplasmic organelles that surround nearly every living plant cell (the gametes being one notable exception) and are composed mainly of polysaccharides, polylignols, and proteins in varying proportions (Albersheim et al, 2011). The structural complexity and the covalent and non-covalent interactions between these macromolecules yield a composite structure (Gibson, 2012) that is mechanically and chemically recalcitrant, thus posing formidable obstacles for analysis. Xyloglucans (XyGs), which are some of the most structurally complex and tightly integrated components in plant cell walls, embody these analytical challenges. In the generally poorly apprehended field of plant cell wall biology, XyG nevertheless stands out as one of the best characterized components. XyG has been shown to have metabolic role as a mobilizable storage polysaccharide in developing embryos of several plant species (Buckeridge, 2010). XyG oligosaccharides have also been shown to have signaling function both *in vivo* and in cell culture models (Fry et al, 1990; Vargas-Rechia et al, 1998; York et al, 1984). A large collection of plant glycosyl transferases (GTs) and glycosyl hydrolases (GHs) responsible for XyG biosynthesis (Zabotina, 2012), remodeling (Eklof & Brumer, 2010), and degradation (Iglesias et al, 2006) have been discovered. Most notably, and controversially, the archetypal plant cell wall interpolymer network formed by cellulose and XyG has been assigned load-bearing as well as regulatory functions in cell wall (Hayashi & Kaida, 2011). Further impetus for XyG

characterization has arisen from its recognized and potential applications in the food, pharmaceutical, textile, and paper industries (Mishra, 2013; Mishra & Malhotra, 2009).

### 2.2.2. The Chemical Structure of XyGs

*Side-chain Structures* The chemical structure of XyGs is characterized by a linear cellulosic backbone of β(1→4)-linked D-Glc*p* residues, with frequent and regularly distributed side-chains initiated (except for in rare instances) by an α-D-Xyl*p* residue at *O*-6. A large portion of the Xyl*p* residues are further extended to create larger side-chains containing up to four glycosyl residues. There is substantial diversity in the fine structures of XyGs; currently more than twenty side-chain structures have been described, only a small subset of which can be found in any single species. XyG structures of any length can be conveniently described in a single-letter code that was originally introduced by Fry and colleagues (Fry et al, 1993) and expanded by various researchers to accommodate newly discovered side-chain structures. In this scheme, G denotes an unsubstituted backbone Glc*p* residue, whereas X, L, and F denote Glc*p* residues substituted at *O*-6 with α-D-Xyl*p*, β-D-Gal*p*-(1→2)-α-D-Xyl*p*, and α-L-Fuc*p*-(1→2)-β-D-Gal*p*-(1→2)-α-D-Xyl*p* side-chains, respectively. An inclusive listing of relevant side-chain structures and their nomenclature is presented in Table 1. All XyGs (from wild type plants) described so far contain at minimum one additional side-chain structure (almost always L) in addition to X. However, nearly all XyGs contain additional side-chains, F being the most common.

Notably, the F side-chain is present in the structural XyGs of the model organism *Arabidopsis thaliana* (Zablackis et al, 1995) and in suspension cultured sycamore maple (*Acer pseudoplatanus*) cells (York et al, 1986).

Despite the large number of chemically distinct side-chains, there is a much smaller number of architecturally unique side-chain groups, each composed of members that share some structural features. Most of the described XyG side-chains belong in a group whose more complex side-chains are formed from the simpler ones by extension with additional glycosyl residues, as exemplified by the X, L, and F side-chains. In this regard, XyG biosynthesis represents that of several other types of complex carbohydrates, where the non-stoichiometric extension by GTs (and sometimes trimming by GHs) leads to a series of structures of varying complexity (as in the *N*- and *O*-linked glycans in proteins) or length (as in almost every type of polysaccharide) (Varki et al, 2009).

*Stereochemical Aspects of XyG Structure* Besides the structural heterogeneity arising from its non-template directed biosynthesis, XyG structures exhibit additional layer of structural complexity. Several reported side-chain structures are stereochemically identical to the common L and F side-chains despite having distinct glycosyl residue composition. For example, jojoba (*Simmondsia chinensis*) seed (Hantus et al, 1997), and notably the Arabidopsis *murus1* plant (Pauly et al, 2001a) (generated using chemical mutagen ethyl methanesulfonate and found by screening the resulting plant lines for abnormal cell wall sugar composition (Reiter et al, 1993)) and has a defect in the biosynthesis of GDP-L-Fuc (Bonin et al, 1997), have both been reported to

contain α-L-Gal*p*-(1→2)-β-D-Gal*p*-(1→2)-α-D-Xyl*p* side-chain (termed J). The chemical structure of this side-chain differs from that of the F side-chain only by the substituent on the *C*-6 of the terminal sugar residue, namely hydroxyl (OH) instead of hydrogen (H). Since L-Fuc is 6-deoxy-L-Gal, the J side-chain is stereochemically identical to the F side-chain. *Selaginella kraussiana* (a clubmoss), and *Equisetum hyemale* (a horsetail) XyGs contain side-chains with α-L-Ara*p*-(1→2)-α-D-Xyl*p* (D side-chain) and α-L-Fuc*p*-(1→2)-α-L-Ara*p*-(1→2)-α-D-Xyl*p* (E side-chain) structures (Peña et al, 2008). These side-chains are stereochemically identical to the L and F side-chains, respectively, with an α-L-Ara*p* residue replacing the β-D-Gal*p* residue. Here, the chemical difference is the substituent on the *C*-5 of the sugar residue linked to the Xyl*p* residue, namely hydrogen (H) on the α-L-Ara*p* and hydroxymethyl ($CH_2OH$) on the β-D-Gal*p*. In yet another variation of this structural theme, Peña and colleagues reported the root hair-specific anionic β-D-Gal*p*A-(1→2)-α-D-Xyl*p* (termed Y) and α-L-Fuc*p*-(1→2)-β-D-Gal*p*A-(1→2)-α-D-Xyl*p* (termed Z) side-chains in *Arabidopsis* (Peña et al, 2012). These side-chains are also stereochemically identical to the L and F side-chains, with D-Gal*p*A replacing D-Gal*p* residues. The only difference between the β-D-Gal*p*A and β-D-Gal*p* residues is the substituent on the *C*-5, namely carboxylic group (COOH) on the Gal*p*A and hydroxymethyl ($CH_2OH$) on the Gal*p* residue.

Here (and elsewhere in this dissertation), glycosyl residues (and the side-chains bearing them) are defined as stereochemically identical if they have the same spatial configuration all shared stereocenters (ignoring stereocenters that

are not present in all of the structures being compared) (IUPAC-IUBMB Joint Commission on Biochemical Nomenclature (JCBN), 1996). Thus, α-L-Gal*p* and α-L-Fuc*p* are stereochemically identical because they share the α-L-*galacto* configuration, whereas β-D-Gal*p*, β-D-Gal*p*A and α-L-Ara*p* are stereochemically identical because they share the β-D-*galacto* configuration. The stereochemical configurations of these glycosyl residues are shown in Supplemental Figure 1.

Several other families of structurally related side-chains, whose taxonomic distribution is also typically limited, have been described in the literature (Hisamatsu et al, 1992; Peña et al, 2008; York et al, 1996).

The apparent interchangeability of stereochemically identical sugar residues, for example α-L-Fuc*p*/α-L-Gal*p*, as well as β-D-Gal*p*/β-D-Gal*p*A/α-L-Ara*p* in XyGs can arise from at least two distinct reasons. First, the GTs may be able to use more than one sugar nucleotide as a donor in the transfer reaction. Second, distinct GTs dedicated to the transfer of a particular nucleotide sugar to a particular substrate may be present in the plant. The transfer and partial replacement of α-L-Fuc*p* by α-L-Gal*p* in *murus1* XyG is unlikely to be catalyzed by a dedicated L-galactosyl transferase since L-Gal is not found in the XyG of neither wild type Arabidopsis nor in the *murus2* plant that has inactive XyG-specific fucosyltransferase (AtFUT1) (Vanzin et al, 2002). Rather, it is hypothesized that in the wild type Arabidopsis, the utilization of GDP-L-Fuc by the AtFUT1 (Perrin et al, 1999) dominates over that of GDP-L-Gal (which is also a naturally occurring nucleotide sugar) due to either GDP-L-Fuc being a preferred

donor substrate, or GDP-L-Fuc having a higher concentration in the secretory pathway where the enzyme resides, or a combination of both phenomena. In contrast, dedicated GTs have been found for β-D-Gal*p* (with the underlined target positions X<u>X</u>XG (Jensen et al, 2012) and XX<u>X</u>G (Madson et al, 2003)) and β-D-Gal*p*A (Peña et al, 2012) residues in XyG. Thus, Arabidopsis utilizes at least two different strategies in XyG biosynthesis that increase the structural diversity without sacrificing any functional constraints or requirements in XyG stereochemistry. Unlike the two XyG-specific galactosyltranferases described, the XyG-specific galacturonosyltransferase (XUT1) can apparently transfer β-D-Gal*p*A to both of the underlined positions in <u>X</u>X<u>X</u>G in XyG polysaccharide. The described XyG-specific galactosyl- and galacturonosyltransferases are closely related and are classified as GT47 enzymes. It is also interesting to note that xylogucan-specific arabinofuranosyltransferases belonging to the GT47 family (Schultink et al, 2013) catalyze the transfer of an α-L-Ara*f* residue (which is not structurally similar to the pyranosyl residues with D-*galacto* configuration in L, D and Y side-chains) to the *O*-2 position of the Xyl*p* residue to form the S side-chain (York et al, 1996).

*XyG Polysaccharide is Composed of Subunits* The backbone of XyG polysaccharide is susceptible to enzymatic hydrolysis by various evolutionarily distinct endoglucanases (EGs) that differ in their site-specificity (Gilbert, 2010; Gilbert et al, 2008). Limit hydrolysis of water extractable tamarind seed XyG polymer by XyG-specific endoglucanases (XEGs) (which exclusively hydrolyze the glycosidic bond of unsubstituted backbone glucosyl residues) yields XXXG,

XLXG, XXLG, and XLLG as the most abundant XyG oligosaccharides (York et al, 1990) (see Figures 1-1 and 2-1A). The smallest possible XyG oligosaccharides generated by XEG (or other EGs) with this specificity are collectively termed the S1 fraction (for single subunit). Some structural generalities have emerged from the analysis of XyGs from a large number of plant species. The (usually) homogenous length of the XyG S1 oligosaccharides leads to an obvious conclusion that a mature XyG polymer is composed of discrete subunits. Most plants have XXXG-type subunits, where blocks of three consecutively substituted backbone residues alternate with a single unsubstituted backbone residue throughout the polymer (Hoffman et al, 2005; Hsieh & Harris, 2009; Hsieh & Harris, 2012; Vincken et al, 1997). The side-chain structures and placement within the subunit are characteristic of the XyG type. For example, in XXXG-type XyGs, both the probability of extension of the xylosyl residues and their structural complexity typically increase towards the reducing end of the subunit. Whereas the patterning of side-chains along the polymer backbone is generally restricted in a given species, their structures can vary depending on the tissue and cell types. For example, tamarind seed also contains fucosylated XXXG-type XyG that is tightly integrated to the cell walls and thus is not water extractable (see Chapter 3 for the evidence from immunofluorescence localization experiments that demonstrate this). Notable other XyG motifs include XXGG found in some Solanaceous species (York et al, 1996), and XXGGG found in Poeacea (Vincken et al, 1997). Interestingly, the XyG from Jatoba (*Hymenaea courbaril*) seeds contains both XXXG- and XXXXG-type subunits (Buckeridge et al, 1997).

Exceptions to all rules regarding XyG classification and branching patterns have been described.

### 2.2.3. XyG-active Glycosyl Hydrolases

Enzymatic hydrolyses have proven to be essential tools in the analysis of complex polysaccharides, first to generate oligosaccharides that retain at least some of the structural characteristics of the parent polysaccharide but are more amenable for analysis, and further to generate fragments with novel structural features that are not necessarily found in the parent polysaccharide (van den Brink & de Vries, 2011; Ward & Moo-Young, 1989). Several XyG backbone and side-chain hydrolyzing GHs from plants (Iglesias et al, 2006) and fungi (van den Brink & de Vries, 2011) that can be used to modify XyG *in vitro* have been characterized.

Two groups of GHs with *endo*-specificities are the most useful for generating XyG fragments. The first includes XEGs (EC 3.2.1.151) that strictly hydrolyze the glycosidic bond of an unsubstituted backbone Glc*p* residue. These enzymes are members of several Carbohydrate Active Enzymes (CAZy) families (Cantarel et al, 2009), including GH5, GH7, GH12 and GH16. This canonical specificity that strictly forms XXXG from XXXGXXXG renders these XEGs the most useful in generating XyG oligosaccharides for structural characterization since the backbone length of the formed oligosaccharides is constant (as long as the backbone length of the substrate polysaccharide subunits is homogenous). In contrast to these enzymes, the site-specificity of hydrolysis by GH74 XEG is not

as stringent, and other products, such as XXX, and GXXXG are formed as well, albeit at a lower rate (Yaoi et al, 2005). The second group of useful EGs includes two other GH74 enzymes, namely *Geotrichum* sp. M128 oligoxyloglucan reducing end cellobiohydrolase [OXG-RCBH (Yaoi & Mitsuishi, 2002)], and *Aspergillus nidulans* oligoxyloglucan reducing end-specific xyloglucanobiohydrolase [OREX (Bauer et al, 2005)]. These two enzymes (EC 3.2.1.150) release a substituted cellobiose fragment from the reducing end of XyGs if the Xyl*p* side-chain immediately on the non-reducing side of the hydrolysis site is not extended. Thus, of tamarind seed XyG S1 oligosaccharides, XLXG and XLLG are not substrates for these enzymes, whereas XXXG and XXLG are, yielding a common XX fragment, as well as XG and LG fragments, respectively (Bauer et al, 2005; Yaoi & Mitsuishi, 2002) (see Figure 3).

Isoprimeverosyl hydrolase (IPH) is another XyG-hydrolyzing GH with *endo*-activity, and like XEGs, OXG-RCBH, and OREX, IPH hydrolyzes the glycosidic bond of the backbone Glc*p* residue. The site-specificity of IPH however is unique in that it hydrolyzes only the glycosidic bond of the non-reducing end backbone Glc*p* residue and only if it is substituted by a Xyl*p* residue which itself is not extended (Kato et al, 1985). Thus, IPH hydrolyzes XXXG to X ($\alpha$-D-Xyl*p*-(1$\rightarrow$6)-D-Glc, common name isoprimeverose) and XXG fragments. IPH activity is present in some $\beta$-galactosidase preparations, as well as enzyme mixtures from *Aspergillus oryzae* [trade name Sanzyme 1000 (Kato et al, 1985)] and *Irpex lacteus* [trade name Driselase (Lorences & Fry, 1994)]. Fungal IPHs have been purified and characterized (Kato et al, 1985; Yaoi & Miyazaki, 2012).

Further diversity in XyG oligosaccharide structures can be generated with GHs with *exo*-activity. XyG active β-galactosidases (EC 3.2.1.23, GH35) (York et al, 1993), α-xylosidases (EC 3.2.1.177, GH31) (O'Neill et al, 1989), β-glucosidases (EC 3.2.1.21 GH1) (Crombie et al, 1998), and α-L-fucosidases (EC 3.2.1.63, GH95) (Leonard et al, 2008) have been described. Some of these enzymes show regiospecificity or regioselectivity towards their substrates. For example, all described XyG-active α-xylosidases only hydrolyze the xylosyl residue on the non-reducing end X. Thus, XXXG is hydrolyzed exclusively to GXXG and no other product such as XGXG is formed. At least some β-galactosidases exclusively or preferentially hydrolyze the β-D-Gal*p* residue on the "middle" xylosyl residue, thus utilizing XLXG as a substrate at higher rate compared to XXLG (York et al, 1993). In summary, GH activities against most of the common glycosyl residues in XyGs have been described, and these enzymes are an essential part of the toolbox to study XyG structures. A large part of this chapter, and dissertation as a whole, is concerned with using XyG-active enzymes to produce distinct oligosaccharides for analytical purposes.

### 2.2.4. Analytical Methods for XyG Analysis

The structural complexity of XyGs renders them challenging for detailed structural and quantitative analysis. Thus, a wide range of complementary approaches have been utilized in XyG analysis.

*Matrix-assisted Laser Desorption/Ionization Time-of-flight Mass Spectrometry (MALDI-TOF MS)* Mass spectrometry (MS) offers a powerful

combination of structural and quantitative information for oligosaccharide analysis (Bauer, 2012; Zaia, 2004). MALDI-TOF MS specifically has proven to be a convenient method for routine analysis of oligosaccharides and their mixtures. One of the main advantages of this method is its sensitivity, as little as ng of native (underivatized) neutral oligosaccharide can be detected. Furthermore, MALDI-TOF MS analysis can provide estimation on the complexity of the sample, as well as an approximation of the relative quantities of the individual oligosaccharides within a narrow mass range. Whereas single-stage MS (also termed full MS analysis) allows an unambiguous determination of the diastereomeric composition (for example, the number of hexoses, deoxyhexoses and pentoses) of oligosaccharides, isomeric aspects of structure are not accessible from this analysis. Glycan analysis by MS is hampered by both stereochemical (linkage anomericities, presence of isomeric residues) and positional (sequences as well as linkage positions of residues) isomerism. Both types of isomers exist in XyG oligosaccharides, leading to ambiguity in the structural interpretation of full MS data in the absence of other information. Two examples of XyG oligosaccharide pairs that yield identical MS spectra include the positional isomers XLXG and XXLG, as well as the XLLG and XXJG pair that exhibit both positional and stereochemical isomerism. Despite its shortcomings, full MS analysis is a staple approach in structural and quantitative analysis of XyGs. At least some of these structural ambiguities can be resolved by multiple-stage MS (or tandem MS) analysis (Dell et al, 2007). However, to take full advantage of tandem MS approach, chemical derivatization of the sample prior to

the analysis is typically required. MALDI-TOF MS is used extensively in the work described in this dissertation for routine determination of the structural identities and purities of oligosaccharide preparations, including chromatography fractions. Furthermore, we explore the capability of MALDI-TOF MS to quantitate XyG S1 oligosaccharides in complex mixtures.

*High-pH Anion Exchange Chromatography with Pulsed Amperometric Detection (HPAEC-PAD)* HPAEC-PAD provides both high resolution separation and sensitivity for oligosaccharide analysis (Corradini et al, 2012; Hardy & Rohrer, 2007). XyG polysaccharides typically yield a small number of S1 oligosaccharides whose structures are nevertheless relatively homogenous. Hence, HPAEC separation is able to separate most XyG S1 oligosaccharides from a typical XEG hydrolysis mixture. The structural information from HPAEC-PAD is limited to the comparison of elution times of the analytes to standards with known structures, and thus is prone to misinterpretation. On the other hand, PAD offers sensitive and reproducible detection of oligosaccharides without the need for derivatization. HPAEC-PAD is utilized in this dissertation for the analysis of XyG oligosaccharide purity. We also explore the utility of HPAEC-PAD to quantitate XyG oligosaccharides in mixtures.

*Nuclear Magnetic Resonance (NMR) Spectroscop*y NMR spectroscopy is the gold standard in structural determination of organic molecules (Bubb, 2003). Complete structural information that includes the sequences, linkage positions, identities, and anomeric configurations of sugar residues, non-carbohydrate decorations, as well as the solution dynamics of the molecule and

its parts can be in principle elucidated by NMR (Widmalm, 2007). So-called structural reporter concept (Vliegenthart et al, 1981) was utilized in this dissertation to verify the structural identities of the purified XyG oligosaccharides. Here, the few specific oligosaccharide signals (typically from the anomeric protons of the glycosyl residues) that are well separated from the bulk of the signals, are observed and analyzed without resorting to full structural analysis. This approach is sufficient for our purposes since all the XyGs that were used in this work have been previously structurally characterized and the chemical shifts that are the primary information content in structural analysis have been published. Besides the primary literature on the published chemical shifts, the chemical shift assignments relied heavily on the XyG $^{1}$H NMR chemical shift database (*http://www.ccrc.uga.edu/databases/index.php?db=protonnmr*). This database was established and is maintained by Dr. William York at the Complex Carbohydrate Research Center (CCRC) or the University of Georgia. Additional web-based tools that could be used in interpretation of XyG oligosaccharide NMR spectra include Computer Assisted Spectrum Evaluation of Regular Polysaccharides (CASPER, at *http://www.casper.organ.su.se/casper/*) (Lundborg & Widmalm, 2011) and GlyNest at (*http://www.glycosciences.de/database/start.php?action=form_shift_estimation*) (Lutteke et al, 2006).

Other information can come for example from biosynthetic or chemical "logic", other analytical techniques such as sugar composition and linkage analyses, liquid chromatography (either off-line, or hyphenated to MS or other

detection modalities), or interaction analysis with oligosaccharide binding proteins (Kamerling & Gerwig, 2007). Due to the individual strengths and weaknesses of each approach, complementary approaches are typically employed in conjunction for less ambiguous structural and quantitative interpretation.

## 2.3. Results and Discussion

### 2.3.1. Optimization of Size-exclusion Chromatography Conditions for XyG Oligosaccharide Purification

Our first objective was to purify the most abundant and widespread canonical XyG S1 oligosaccharides as well as some of their stereochemical analogs that are typically found only in a limited number of species or in specific tissues. We used extra-fine (particle size < 45 μm) Bio-Gel P-2 size-exclusion chromatography (SEC) polyacrylamide resin (Bio-Rad) for all of our oligosaccharide separations. This resin has an effective separation range of 100-1800 Da and has been shown to be suitable for separating neutral glycans, both native and reduced, with one to ten monosaccharide residues, using water as the eluent (Hantus et al, 1997; Hisamatsu et al, 1991; Kimura et al, 1995; York et al, 1993; York et al, 1996; York et al, 1990). Our initial attempts to resolve the tamarind XyG S1 oligosaccharide mixture using water eluent resulted in three heavily overlapping peaks as indicated by the colorimetric sulfuric acid-phenol assay chromatogram (Supplemental Figure 2A). Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) analysis

of individual fractions confirmed that all but the extreme tailing sulfuric acid-phenol positive fractions contained at least two non-isomeric oligosaccharides (data not shown). The poor separation efficiency was surprising since the same SEC resin and the same eluent conditions have been used earlier to nearly baseline separate these and other XyG oligosaccharides (see the references above).

Boric acid [$B(OH)_3$] is a Lewis acids that in aqueous solution is in equilibrium with its conjugate base borate anion [$B(OH)_4^-$] with p$K_a$ of 9.2. The borate anion can then react reversibly and covalently with *cis*-diols by a double dehydration reaction (Springsteen & Wang, 2002; Yan et al, 2004). We reasoned that this phenomenon could be utilized in XyG oligosaccharide separations since the β-D-Gal*p*, α-L-Gal*p*, α-L-Fuc*p*, and α-L-Ara*p* residues in XyG oligosaccharides contain 3,4-*cis*-diols that can complex with the borate anion. Thus, in the presence of boric acid, X[LX]G (where [LX] denotes a single L and a single X side-chain in an unspecified order) and XLLG oligosaccharides should complex with one and two borate anions, respectively. Since XXXG, X[LX]G, and XLLG form a homologus series of oligosaccharides and the number of complexed borate anions is proportional to the number of Gal*p* residues, the increase in the hydrodynamic radius should also be proportional to the number of the Gal*p* residues. Correspondingly, the SEC elution volumes of these complexes should spread to a wider range compared to the uncomplexed oligosaccharides.

Our initial efforts using tamarind XyG S1 oligosaccharides as a model system and 10 mM boric acid, 10 mM phytic acid, pH 10 (adjusted with sodium

hydroxide) were successful and this oligosaccharide mixture separated into three nearly baseline separated peaks (Supplemental Figure 2C). However, these results could not be consistently replicated. After efforts to pinpoint the reason for the irreproducibility of this effect, we concluded that the most likely difference between the initial and subsequent experiments was the quality of the Bio-Gel P-2 resin. One possible mechanistic explanation for the fortuitous effect is a small amount of negative charge in the polyacrylamide-based Bio-Gel P-2 resin resulting from deamination of the amide groups to carboxylic groups either during the manufacturing process or laboratory use. These anionic carboxylic groups would coulombically repulse the anionic borate-oligosaccharide complexes from the residue pores. Increasing the number of borate anions and thus increasing the charge of the oligosaccharide analyte would then lead to decrease in the elution volume that is proportional to the number of complexed borate anions. Due to the unavailability from the manufacturer of the batch of resin that was used in the initial experiments and that was discarded after a lengthy period of use, we were unable to confirm experimentally whether this resulted from an issue in the quality control during the manufacturing process.

We then explored other eluents that would yield consistently good separation of the oligosaccharides, and would require minimum cleanup for downstream applications. Volatile buffers that can be removed by lyophilization were deemed the most suitable for our purposes. We compared the separation efficiency of 50 mM acetic acid to water and found that the former gave enhanced separation of tamarind seed XyG S1 oligosaccharides (Supplemental

Figure 2B). In these conditions, the oligosaccharides elute in expected order according to their molecular weights. Regardless of the eluent used, XXXG and XLLG could be obtained by this single chromatography step without further separations. Again, XLXG and XXLG co-elute. We also observed that 10 mM ammonium acetate, pH 5.0 buffer gave separation comparable to 50 mM acetic acid (data not shown) but ammonium acetate buffer has low volatility and several rounds of lyophilization are needed for its complete removal. Since 50 mM acetic acid can be removed by a single round of lyophilization, we used this eluent for all subsequent separations.

### 2.3.2. Establishing A General Strategy for The Generation and Purification of XyG Oligosaccharides

The structural homology of subunits from a typical XyG polysaccharide leads to at least partial co-elution of corresponding S1 oligosaccharides in SEC, regardless of the eluent used. For example, XLXG and XXLG oligosaccharides co-elute in SEC due to their similar hydrodynamic radii (Figure 1B and Supplemental Figure 2). To purify the isomeric XLXG and XXLG oligosaccharides, we employed the complementary actions of β-galactosidase and OXG-RCBH enzymes. *Aspergillus niger* GH35 β-galactosidase preferentially hydrolyzes the β-D-Gal*p* residue on the middle Xyl*p* residue in a XyG subunit or oligosaccharide, converting XLXG to XXXG, and XLLG to XXLG at much higher rate than the conversion of XXLG to XXXG (York et al, 1993). A preparation of this enzyme (commercially available from Megazyme) contains an IPH contaminant. Besides its hydrolytic activity, IPH also has a transglycosidase

activity (York et al, 2002) which manifests at high substrate concentrations by the formation of, for example, XXXXG from XXXG. These two IPH reactions not only lower the yield of desired oligosaccharides but create unnecessary complexity in the oligosaccharide mixture as well (IPH activities and their utilization are discussed further in Section 2.3.6). In order to prevent the formation of these unwanted products, we employed a strategy whereby the β-galactosidase hydrolysis was performed on the XyG polymers, except for *Selaginella kraussiana* XyG (see below). Since IPH spcecifically hydrolyzes X units from the non-reducing end of XyGs, only a small fraction of the polysaccharide is recursively hydrolyzed until a “hanging” non-reducing end glucosyl residue (as in GXXG) or substituted X units (as in LXG) is exposed, preventing further hydrolysis. The XXXG, XLXG, XXLG, and XLLG subunit quartet in the tamarind XyG polymer was converted to a XXXG and XXLG and nearly baseline separation of the corresponding XEG-generated S1 oligosaccharides was obtained by SEC (Figure 2B). The purity and structural identity of the pooled oligosaccharide preparations were confirmed by MALDI-TOF MS (Figures 2C and 2D), as well as HPAEC-PAD and NMR spectroscopy (Supplemental Figures 3K and 3Q). The β-galactosidase hydrolysis proceeds to near completion as indicated by these analyses on the final XXLG oligosaccharide preparation.

In order to obtain the XLXG oligosaccharide, we utilized an OXG-RCBH enzyme that hydrolyzes XXXG and XXLG to a common XX oligosaccharide as well as XG, and LG oligosaccharides, respectively. The OXG-RCBH hydrolysis of tamarind XyG S1 oligosaccharides proceeds to completion as is evident from the

absence of XXXG signals in the MALDI-TOF MS spectrum (Figure 3A) and in the SEC chromatogram (Figure 3B). The resulting mixture of XG, XX, LG, XLXG, and XLLG oligosaccharides are readily separated by SEC (Figure 3B). The purity and structural identity of the pooled oligosaccharide preparation were further confirmed by MALDI-TOF MS (Figure 3B), as well as HPAEC-PAD and NMR analyses (Supplemental Figures 3G, 3H, 3O, 3P and 3S).

### 2.3.3. Neutral XyG Oligosaccharides from Other Plant Sources

XyG oligosaccharides were purified similarly from suspension cultured sycamore maple cells, as well as jojoba seeds and *Selaginella kraussiana* aerial parts. Strong alkaline conditions, which were used to extract these XyGs from the cell walls, hydrolyze the *O*-acetyl (and other ester) decorations and thus yield XyGs composed completely of glycosyl residues. Aliquots of XyG polysaccharide preparations from sycamore maple cells and jojoba seed were treated either with β-galactosidase and XEG, or XEG and OXG-RCBH to facilitate the purification of the individual oligosaccharides. XEG hydrolysis of sycamore maple XyG polysaccharide yields a mixture of oligosaccharides with XXXG, XLXG, XXLG, XXFG, XLLG, and XLFG as the major components (York et al, 1988). β-galactosidase treatment of sycamore maple XyG prior to XEG treatment simplified its subunit composition to XXXG and XXFG with traces of several rare subunits, including XFFG that can be purified to homogeneity (data not shown). (It is noteworthy that several XyG side-chains, unique to sycamore suspension culture (Hisamatsu et al, 1992; York et al, 1995), occur in the immediate vicinity of the unbranched backbone Glc*p* residues, and render the junction resistant to

XEG hydrolysis. Thus these rare side-chains are always found in S2 or larger oligosaccharides and we didn't attempt to purify them in the work described here). The SEC fractions from XEG digestion of sycamore maple XyG were assayed only by MALDI-TOF MS and colorimetric assay was not utilized (data for these MALDI-TOF MS analyses are not shown). The XXFG and XFFG oligosaccharides could be purified to homogeneity using SEC as confirmed by MALDI-TOF MS, HPAEC-PAD and NMR analyses (Supplemental Figure 3Y and 3AB). OXG-RCBH hydrolysis of sycamore maple XyG yielded XG, XX, LG, FG, XLXG, XLFG, and XFFG oligosaccharides that were subjected to SEC separation. All resulting oligosaccharides could be obtained in their pure forms and their identities and purities of FG and XLFG were confirmed by MALDI-TOF MS, HPAEC-PAD and NMR (Supplemental Figures 3W and 3AA).

Jojoba seed XyG polymer is composed of XXXG, XXLG, XXFG, XXJG, XLFG, and XLJG subunits (Hantus et al, 1997). Sequential β-galactosidase and XEG treatments of jojoba XyG yielded XXXG, XXLG, XXFG, and XXJG as the major oligosaccharides. The F and J side-chains are stereochemically identical and differ only by the presence of hydroxyl on the *C*-6 of the terminal pyranosyl residue (methyl group on the L-Fuc*p* in the F side-chain, and hydroxymethyl group on the L-Gal*p* in the J side-chain). Owing to their chemical similarity, these structures can be surmised to have very similar contributions to the hydrodynamic behavior of oligosaccharides bearing them. Nevertheless, it has been shown previously that 6-deoxyhexoses elute slightly earlier in Bio-Gel P-2 chromatography compared to their hexose counterparts using water as an eluent

(Yamada et al, 1980). Indeed, the XXJG/XXFG oligosaccharide pair separates partially in our hands using the same resin. XXJG can be purified by iteratively re-chromatographing the fractions enriched in XXJG (Supplemental Figure 3AD). The SEC fractions for jojoba XyG purification were assayed only by MALDI-TOF MS and colorimetric assay was not utilized (data for these MALDI-TOF MS analyses are not shown). XEG and OXG-RCBH hydrolysis of jojoba seed XyG polysaccharide yielded XG, LG, FG, JG, XXXG, XLFG, and XLJG as the major products that were separated by SEC. The FG/JG and XLFG/XLJG oligosaccharide pairs again severely overlap but could be purified by several rounds of re-chromatography. The purities and structural identities of the JG and XLJG oligosaccharides were validated by as described above (Supplemental Figures 3AC and 3AE). XyG active α-L-fucosidase (GH95) is commercially available from Megazyme. This enzyme doesn't hydrolyze the α-L-Gal*p* residues and thus should prove useful in the future efforts to purify J side-chain containing XyG oligosaccharides.

Due to the simultaneous presence and apparent metabolic interchangeability of the stereochemically identical β-D-Gal*p* and α-L-Ara*p* residues, XEG hydrolysis of *Selaginella* XyG polysaccharide yields a complex mixture of S1 oligosaccharides with L, D, F, and E side-chains (Peña et al, 2008). The 4 M potassium hydroxide extract is a complex mixture of several polysaccharides and other components and the β-galactosidase treatment of this extract was inefficient in our hands, as indicated by MALDI-TOF MS analysis after the β-galactosidase and XEG treatments (data not shown). Since the β-

galactosidase preparation contains an IPH contaminant and the Selaginella XyG has relatively little of the L side-chain containing oligosaccharides, the β-galactosidase treatments of these oligosaccharides and polysaccharides was omitted. We obtained nearly pure XDXG, and XDDG oligosaccharides from OXG-RCBH hydrolysis of Selaginella XyG S1 oligosaccharide mixture by SEC (Supplemental Figures 3U and 3V). Analogous oligosaccharides with F and E side-chains (for example, FG and EG) co-elute in SEC (data not shown) and we did not attempt to purify the E side-chain containing oligosaccharides to homogeneity.

A summary of all purified oligosaccharides is presented in Table 2. The analytical data validating the structures and purities of all oligosaccharide preparations is presented in Supplemental Figure 3. Supplemental Tables 1 and 2 tabulate the analytical data collected by MALDI-TOF MS, HPAEC-PAD and NMR for all oligosaccharides.

### 2.3.4. α-xylosidase Modification of XyG Oligosaccharides

We treated pure XXXG, XLLG and XXFG oligosaccharides with α-xylosidase to form GXXG, GLLG, and GXFG oligosaccharides. The MALDI-TOF MS, HPAEC-PAD, and NMR analyses show that the reactions proceeded to completion in all cases and no other products were formed (Supplemental Figures 3J, 3R, and 3X).

### 2.3.5. Reduction of XyG oligosaccharides to Corresponding Oligosaccharide Alditols

We reduced pure XXXG, XLLG and XXFG oligosaccharides to the corresponding alditols, XXXGol, XLLGol, and XXFGol by sodium borohydride treatment. The MALDI-TOF MS, HPAEC-PAD, and NMR analyses show that the reactions proceeded to completion in all cases (Supplemental Figures 3L, 3T, and 3Z).

### 2.3.6. Isoprimeverosyl-series XyG Oligosaccharides

XyG S1 oligosaccharides that have fewer or more than four glucosyl residues in the backbone have been reported in the literature. XXG subunits have been described as a minor component in XyGs from plants that have XXXG-type XyG branching pattern (Guillen et al, 1995; Hilz et al, 2007; Hsieh & Harris, 2009; Pauly et al, 2001b). XyG polymer from cotyledons of Jatoba (*Hymenaea courbaril* L.) contains subunits with the XXXXG branching motif (Buckeridge et al, 1997). The presence of subunits with different backbone lengths (either four or five glucosyl residues) and varying number of β-D-Gal*p* residues (up to two on the XXXG core structure and up to three on the XXXXG core structure) results in a complex mixture of oligosaccharides from this source that severely co-elute in SEC (data not shown).

Since these XyG oligosaccharides with unusual backbone length could be used in various ways to provide information on the enzymatic synthesis and degradation of XyGs, we were motivated to produce them *in vitro*. York and colleagues have demonstrated previously that IPH has significant hydrolase and

transglycosidase activities *in vitro* (York et al, 2002). Since both hydrolase and transglycosidase activities occur at appreciable rates, a series of oligosaccharides with different number of isoprimeverosyl residues is formed from pure preparations of XXXG as a starting material. For example, the hydrolytic activity hydrolyzes XXXG sequentially to XXG, XG, and finally G, with accumulation of free X units. At a rate comparable to the hydrolysis, a transglycosylation takes place with isoprimeverosyl (X) transfer to XXXG leading to XXXXG, and further to XXXXG leading to XXXXXG. This simple scenario is however complicated by lack of regiospecificity of the transglycosylation reaction. Here, the transglycosylation reaction by IPH using XXXG as an acceptor leads to mixed products, the canonical XXXXG, and unexpected $X^3$XXXG, where the X unit is transferred to *O*-3 of the reducing end glucosyl residue of the acceptor (York et al, 2002). According to MALDI-TOF MS analysis, which cannot differentiate the isomeric transglycosylation products, the major components have the masses corresponding to X, XG, XXG, XXXG, XXXXG, and XXXXXG (Supplemental Figure 4A). Since the oligosaccharides in this series differ in size by a disaccharyl unit, an excellent baseline separation was achieved in SEC (Supplemental Figure 4B), except for the two largest oligosaccharides which were only partially resolved. Very pure fractions were nevertheless obtained for all oligosaccharides, and X (Supplemental Figure 3F), XG, XXG (Supplemental Figure 3I), and XXXG could be purified to homogeneity by SEC. Analysis of the XXXXG and XXXXXG fractions by HPAEC-PAD and NMR mirror the previous findings of (York et al, 2002). HPAEC-PAD analysis confirms the presence of two

main species in the XXXXG preparation (Supplemental Figure 3M) and four major species in the XXXXXG preparation (Supplemental Figure 3N). 2D-NMR characterization also indicates the presence these unusual structures (data not shown). Further separation and structural characterization of the isomeric oligosaccharides was not pursued in this work.

### 2.3.7. Generation and Purification of Anionic XyG Poly- and Oligosaccharides

Two recent reports have shown that some XyG polysaccharides carry anionic side-chains (Peña et al, 2008; Peña et al, 2012). The anionic XyG from Arabidopsis root hairs contain Y and Z side-chains that are stereochemically identical to the common L and F side-chains (Peña et al, 2012). YXXG, YXFG, and XXZG are the major anionic subunits from this XyG polymer and smaller amounts of YXZG, YXYG, YXFG, and XLZG are found as well. Only two of the anionic subunits, namely XXZG and XLZG are topologically similar to the neutral Arabidopsis XyG oligosaccharides, in which the non-reducing end Xyl*p* residue is not extended, whereas in the other anionic subunits, this Xyl*p* residue is further extended by a β(1→2)-linked D-Gal*p*A residue. Given the low abundance of the galacturonosylated subunits in the XyG, and low total root hair mass, we were prompted to generate anionic XyG polymer and oligosaccharides from readily available starting materials. Selective oxidation of D-Gal*p* residues to D-Gal*p*A residues has been established for tamarind XyG using sequential enzymatic and chemical oxidation steps (Parikka et al, 2012).

Despite the topological dissimilarity of the abundant anionic subunits from Arabidopsis and the oxidized tamarind XyG (XYXG, XXYG, and XYYG), we generated anionic XyG that have these side-chains because they could be useful in characterizing carbohydrate active enzymes or carbohydrate binding proteins and antibodies with specificity towards these side-chains. The D-Gal*p* residues were first oxidized to 6-aldehydo-D-Gal*p* residues using galactose oxidase and two auxiliary enzymes, horseradish peroxidase (to regenerate the galactose oxidase active site) and catalase (to destroy the formed hydrogen peroxide). The 6-aldehydo-D-Gal*p* residues were subsequently chemically oxidized to D-Gal*p*A residues using the classical iodine/potassium iodide method.

Several considerations had to be made for the oxidation procedure since we were primarily interested in obtaining reducing oligosaccharides rather than polysaccharide products. Besides the enzymatically generated 6-aldehydo-D-Gal*p*, the iodine treatment also oxidizes all available reducing end glucose residues to gluconic acid (for example XXLG to XXLGon). In the case of oligosaccharides, the presence of a gluconic acid residue would have a dramatic effect on their chemical properties, and also destroy a convenient aldehyde handle that could be used for various downstream applications such as immobilization on a solid support, isotopic labeling, or conjugation to a fluorophore. To circumvent this problem, oxidation was carried out on tamarind XyG polysaccharide. Our ethanol precipitated tamarind XyG polymer preparation contains polysaccharides with sizes ranging from S3 upward (data not shown), and thus may contain non-negligible amount of reducing end subunits that can

pose a problem of their own. Here, the oxidation of reducing polysaccharide would lead to the formation of small amounts of gluconic acid residue-containing oligosaccharides (including XXXGon) that, owing to their similar sizes and charges, co-elute with the desired oligosaccharides such as XYXG and XXYG in SEC. On the other hand, performing the oxidation with sodium borohydride reduced polysaccharide would yield corresponding amount of anionic oligosaccharide alditols (including XXYGol) that again co-elute with their reducing analogs in SEC, and perhaps in other separation modalities as well. We performed all polysaccharide oxidations on reducing polysaccharides and showed that the small amount of gluconic acid containing XyG oligosaccharides can be separated from the desired oligosaccharides by SEC.

We used tamarind XyG polysaccharide as a substrate for the chemoenzymatic oxidation. In our hands, the enzymatic oxidation step was the bottleneck for obtaining highly oxidized tamarind XyG polysaccharide (results not shown). We increased the amounts of all three enzymes in the cocktail by five-fold compared to the amounts used by Parikka and colleagues (Parikka et al, 2012) to yield tamarind XyG polymer with approximately 90% oxidation, according to MALDI-TOF MS (data not shown). This oxidation level is comparable to the approximately 85% oxidation level as previously reported by (Parikka et al, 2012). The iodine oxidation step proceeded to completion under the reported conditions since MALDI-TOF MS analysis of the XEG hydrolysis products of this polysaccharide does not reveal any 6-aldehydo-D-Gal*p* residue bearing oligosaccharides (data not shown). The incomplete enzymatic oxidation

of the XyG polysaccharide results in small amounts of “mixed” side-chain subunits, namely XYLG and XLYG. This is evident from the MALDI-TOF MS analysis of the XEG hydrolyzed polysaccharide (data not shown). Since these oligosaccharides are anionic, they interfere with the SEC purification of the desired XYXG, XXYG, and XYYG oligosaccharides. We attempted to circumvent this problem by removing the remaining D-Gal*p* residues on the oxidized XyG polysaccharide by β-galactosidase treatment prior to XEG hydrolysis. To our surprise, the β-galactosidase did not have any measurable effect on the subunit structure of the oxidized XyGs despite using five-fold higher concentration of the enzyme compared to the concentration used for neutral XyG polysaccharides (data not shown). We also observed that the XXYG oligosaccharide is a poor substrate for OXG-RCBH compared to XXLG (as demonstrated by MALDI-TOF MS) and only partial hydrolysis is achieved in the same conditions that fully hydrolyze XXXG (data not shown). To the contrary, the chemoenzymatically oxidized tamarind XyG polysaccharide is effectively hydrolyzed by XEG (data not shown). These results suggest that the charged residues hinder the access of hydrolytic enzymes both to the backbone and the side-chains within individual subunits of XyG polysaccharide, but does not hinder access of XEG to the junction between two XyG subunits.

In SEC, anionic oligosaccharides typically elute earlier than expected based on the molecular weight alone using unbuffered or non-acidic eluents (Djordjevic et al, 1986; Knudsen et al, 1980; Peña et al, 2007; Thibault, 1980). This effect can be attributed to the coulombic repulsion of the anionic analytes by

a small amount of negative charge in the SEC resins (Bio-Gel P Polyacrylamide Gel Instruction Manual, Bio-Rad). Thus, fractionation of the XEG hydrolysis products from oxidized tamarind XyG polysaccharides to anionic and neutral oligosaccharides was conveniently performed by SEC using water as the eluent (data not shown). The anionic oligosaccharides in the void fraction were pooled and further separation of the anionic fraction to individual XyG oligosaccharides was subsequently performed in 50 mM acetic acid. In these conditions (with approximate pH 3.0), the carboxylic groups of both the oligosaccharides [p$K_a$ of Gal*p*A = 3.5, (Kohn & Kovac, 1978)] and the polyacrylamide [p$K_a$ of partially deaminated polyacrylamide = 4.5, (Michaels & Morelos, 1955)] are mostly protonated and it was expected that the hydrodynamic behavior of the these oligosaccharides would be similar to that of neutral oligosaccharides (Djordjevic et al, 1986; Thibault, 1980). Indeed, at low pH, the elution volumes of YG, XYXG, XXYG, and XYYG shifted from the void volume to the expected elution volume range of neutral oligosaccharides of similar sizes with XYXG and XXYG eluting first, followed by YG and XYYG (data not shown). The late elution of XYYG after the singly galacturonosylated oligosaccharides suggests an interaction between the oligosaccharides and the polyacrylamide resin that may or may not be charge-dependent. We did not further investigate the mechanism of this phenomenon. For this work, we purified YG and XYYG oligosaccharides from the oxidized tamarind XyG polysaccharide and characterized them structurally by MALDI-TOF MS, HPAEC-PAD, and NMR (Supplemental Figures 3AF and 3AG).

### 2.3.8. Generation and Purification of Cellodextrin Oligosaccharides

Since XyGs have a cellulosic backbone, cellodextrin oligosaccharides can serve as useful analogs of XyGs in some situations. We generated cellodextrin oligosaccharides from crystalline cellulose according to procedure reported earlier (Liebert et al, 2008) and purified them using SEC. Excellent separation was achieved for all water-soluble cellodextrins (Supplemental Figure 6B) and oligosaccharides with sizes ranging from three to seven glucosyl residues were purified to homogeneity as shown by MALDI-TOF MS and NMR spectra as well as HPAEC-PAD chromatograms (Supplemental Figures 3A to 3E).

### 2.3.9. The Validation of Structural Identities and Purities of the Oligosaccharides

MALDI-TOF MS, HPAEC-PAD and NMR spectroscopy were used to determine the structural identities and purities of the generated oligosaccharides (Supplemental Figure 3). As an example, some of the XyG oligosaccharides from suspension cultured sycamore maple cells are discussed. MALDI-TOF MS spectrum of the XEG hydrolysate of XyG from this source contains major signals at the following *m*/*z* values: 1085.3, 1247.4, 1393.4, 1555.5 and 1701.6 (data not shown). These signals have been earlier assigned to the following structures 1085.3 (XXXG), 1247.4 (X[LX]G), 1393.4 (XXFG), 1555.5 (XLFG) and 1701.6 (XFFG) using NMR. A MALDI-TOF MS spectrum collected from β-galactosidase and XEG hydrolysed XyG from this source has the following major signals (data not shown) 1085.3 (XXXG), 1247.4 (XXLG, 1393.4 (XXFG), and 1701.6 (XFFG).

Furthermore, a MALDI-TOF MS spectrum collected from XEG and OXG-RCBH hydrolysed XyG from this source has the following major signals (data not shown) 497.2 (XG), 629.3 (XX), 659.3 (LG), 805.5 (FG), 1247.4 (XLXG), and 1701.6 (XFFG). The structural assignments of the last two samples could have been made with the knowledge of the known structures of the starting material (as determined by NMR) and the known specificities of the GHs utilized. For example, the MALDI-TOF MS signal at *m*/*z* 805.3 of the XEG and OXG-RCBH treated sycamore maple XyG can only arise from an FG fragment produced from the XXFG oligosaccharide. NMR analysis of the FG oligosaccharide (Supplemental Figure 3W) confirms the structural assignment. All three analyses confirm also that the FG preparation is very pure. Based on the signal intensities of the MALDI-TOF MS, HPAEC-PAD, and NMR (Supplemental Figure 3W), the purity is estimated to be approximately 98%. Additional “hidden” structural information of the oligosaccharide structures can be extracted from the HPAEC-PAD retention time (for more detailed discussion, see Section 2.3.11), and the binding of monoclonal antibodies (for more detailed discussion, see Chapter 3).

### 2.3.10. Comparison of Mass Spectrometry and Analytical Chromatography in the Quantitation of XyG Oligosaccharides

*Relative Quantification by MALDI-TOF MS* We tested and compared the ability of MALDI-TOF MS and HPAEC-PAD to faithfully reproduce the relative signal intensities of XyG oligosaccharides in a nominally equimolar mixture analyzed at several concentrations. Six oligosaccharides, namely GXXG, XXXG, XXLG, XXFG, XLLG and XLFG were chosen to represent a typical XyG

oligosaccharide mixture from an XEG hydrolysis. The MALDI-TOF MS spectrum of a mixture with 100 pmol of each of the oligosaccharides is shown in Figure 4A for reference. The absolute integrals (and thus intensities) of MALDI-TOF MS signals can vary dramatically between samples, sample spots using the same sample, and even between experiments using the same sample spot, mostly due to uneven crystallization (Gusev et al, 1995) (data not shown). This is in contrast to the relative integral patterns of the XyG oligosaccharides in a given sample, which are very reproducible from experiment to experiment in mixtures containing from 100 to 1 pmol of each oligosaccharide (Figures 4B and 4C). The unavoidable variability in the absolute integrals among the MALDI-TOF MS spectra was eliminated by introducing a scaling factor that does not affect the integral patterns that are our primary interest. Four MS spectra were recorded (duplicate spectra from two sample spots using a given concentration oligosaccharide mixture), and one was selected arbitrarily as a standard and the remaining were multiplied by individual scaling factors to minimize the sum of the coefficients of variation of the oligosaccharide signal integrals for the four spectra.

Figures 4B and 4C show the relative responses of the XyG oligosaccharides at various dilutions, with the response of XXXG set to 1. It is clear that the integral patterns are very similar for the oligosaccharide mixture dilutions containing 1 pmol or more (up to 100 pmol) of each oligosaccharide. Correlation map of the intensity patterns between different concentrations confirms that from 100 to 1 pmol per oligosaccaride, the experiments give highly

similar relative quantifications (Supplemental Figure 7). Below 1 pmol per oligosaccharide, the intensity patterns show low correlations to each other and to the intensity patterns from experiments using higher concentration solutions. This indicates that MALDI-TOF MS yields reproducible information on the relative integrals of XyG oligosaccharides in a complex mixture with as little as 1 pmol of each oligosaccharide. However, this information doesn't reveal the relative amounts of the oligosaccharides unless their individual response factors are known.

It is evident that the MALDI-TOF MS molar response factors of the XyG oligosaccharides differ significantly from each other in our experiments (Figure 4D). The relative response factors of the six XyG oligosaccharides range (when XXXG is set to 1) from 0.76 (GXXG) to 1.78 (XXFG and XLFG) with average of 1.32 and standard deviation of 0.42 (for all relative response factors, see Supplemental Table 1). At least part of the variation could be attributed to the different glycosyl residue composition of the oligosaccharides but this aspect is not very well studied theoretically or experimentally. Our data using just six XyG oligosaccharides does not allow rigorous correlations, if they exist, to be made between the structures and the response factors. Dataset using larger number of pure oligosaccharides with diverse structures could shed light into this important issue that has wider implications in quantitative glycobiology. The nominally equimolar mixture of oligosaccharides was prepared by mixing equal volumes of gravimetrically prepared stock solutions. Only milligram quantities of oligosaccharides were available for weighing and this no doubt introduces

uncertainty in the concentration and response calculations for the oligosaccharides (this aspect is further discussed below).

Two other notable aspects that can affect the relative MALDI-TOF MS response factors both depend on the size of the XyG oligosaccharide. First, assuming the concentration of monovalent cations such as sodium in the sample-matrix mixture is very low compared to the oligosaccharide concentration (since deionized water was used in the preparation of both the sample and the matrix solutions), the sodium cations are a limiting factor in the ionization of the oligosaccharides. Due to the competition for the sodium cations among the oligosaccharides and the fact that the MALDI of oligosaccharides typically produces singly charged ions, it can be reasoned by statistical grounds that the probability of sodium coordination by an oligosaccharide, and thus its ionization, is proportional to the number of glycosyl residues in that oligosaccharide. Another contribution arises from the mass-dependent response of the multichannel plate detector. It is postulated that the response (*R*) of electron multiplier detectors, including multichannel plate detectors used in the MicroFlex LT instrument, is directly proportional to the linear momentum (*p*) of the analyte (Anderegg, 1990; Jurinke, 2005).

$$R \propto p = mv = m\frac{d}{t} \propto \frac{m}{t} \qquad \textbf{Equation 2.1}$$

Here, *m* is the mass of the analyte, *t* is the time-of-flight of the analyte, and *d* is the length of the flight tube (since *d* is a constant, it can be ignored here). Furthermore, the time-of-flight (*t*) of the analyte is proportional to the square root

of the mass of the analyte (Guilhaus, 1995) and charge (*z*) is always 1 in our experiments (and thus can be ignored here):

$$t \propto \sqrt{\frac{m}{z}} \propto \sqrt{m} \qquad \textbf{Equation 2.2}$$

Substituting Equation 2.2 to Equation 2.1 leads to

$$R \propto \frac{m}{t} \propto \frac{m}{\sqrt{m}} = \sqrt{m} \qquad \textbf{Equation 2.3}$$

Thus, the response of the analyte is proportional to the square root of the mass of the analyte. These two contributions, competition for ionizing cation and the response of the detector, have opposite effects on the response factors of oligosaccharides. There are yet other contributions that might arise, for example, from the variable interactions of the analytes with the matrix.

The poor reproducibility of the absolute responses between experiments but excellent reproducibility of the relative responses, as well as the poorly understood basis for the response factors have important consequences for quantitative oligosaccharide MALDI-TOF MS analysis. First, the changes in the relative oligosaccharide composition of two samples can be monitored without knowing anything about the response factors of the oligosaccharides and without using internal standards. On the other hand, monitoring the relative change in the total oligosaccharide amounts in two individual samples requires an internal standard. Furthermore, absolute quantification of oligosaccharides requires knowing the response factors for each oligosaccharide analyzed.

*Relative Quantification by HPAEC-PAD* Quantitation by HPAEC-PAD, whose response factors are based on physico-chemical principles that are fundamentally different than those for MALDI-TOF MS, shows much smaller variation in the relative response factors for the XyG oligosaccharides (Figure 5D). Pulsed amperometric detection (PAD) is based on the oxidation of the analyte on a gold electode in alkaline conditions (Rohrer et al, 2013). The relative response factors of the oligosaccharides (when the response of XXXG is set to 1) in HPAEC-PAD range from 0.87 (GXXG and XLLG) to 1.20 (XLFG) with an average of 1.00 and standard deviation of 0.14 (Figure 5D, Supplemental Table 1). These data indicate that the large differences between the MALDI-TOF MS responses of the oligosaccharides arise mostly from factors intrinsic to the MALDI process, and not from unequal amounts of XyG oligosaccharides in the mixture. However, there is inevitable variation in the relative amounts of the oligosaccharides that can result from variation in the weighing or pipetting. The response factors of oligosaccharides have been studied extensively by HPAEC-PAD (Corradini et al, 2012; Kazmaier et al, 1998; Kunz et al, 1996) and found to vary within similar oligosaccharides. Thus our results are in agreement with the published literature.

Figures 5B and 5C show the relative HPAEC-PAD response factors of the same six XyG oligosaccharides analyzed by MALDI-TOF MS. The relative response factors of the oligosaccharides are very reproducible across the tested dilutions down to 25 pmol per oligosaccharide. Correlation analysis of the integral patterns corroborates these results (Supplemental Figure 8). Since the PAD

integrals of oligosaccharides are very reproducible between experiments, and the PAD responses are linearly dependent on the analyte concentrations (Figure 5E) HPAEC-PAD thus can be used for absolute quantification of XyG oligosaccharides if the individual response factors are known. Supplemental Figure 9 shows the cross-correlation map of the MALDI-TOF MS and HPAEC-PAD analyses across ten different analyte concentrations. It is evident that down to 50 pmol/µl concentration, these two methods agree well with respect to the relative response factor patterns of the oligosaccharides.

### 2.3.11. HPAEC-PAD Retention Time Analysis

We also made some observations regarding the elution times of neutral oligosaccharides that could be useful for interpretation of HPAEC-PAD data. The retention time of oligosaccharides in a homologous series (cellodextrins and isoprimeverosyl series XyG oligosaccharides) increases monotonically and in a linear fashion as a function of the number of repeating units (Figures 6A and 6B). Based on these trends, the elution time of an oligosaccharide in the series can be predicted. The elution times of cellodextrins are linearly dependent on the number of glucosyl units in the oligosaccharide. Furthermore, the elution times of XG, XXG and XXXG can be used to predict the elution time of XXXXG (Figure 6B) and thus discriminate between the two HPAEC-PAD signals from the XXXXG preparation. Based on this trend, we predict that the signal at 24.7 min arises from XXXXG whereas the signal at 32.5 arises from $X^3$XXXG. Similar analysis was performed for XXXXXG preparation and we conclude that the signal at 30.6 min arises from XXXXXG. Different sugar residues in cellodextrins and

XyG oligosaccharides have distinct contributions to their elution times. Addition of either β-D-Glc*p*, α-D-Xyl*p*, β-D-Gal*p*, or α-L-Gal*p* residue increases the elution time of the oligosaccharides. This is evident when comparing any two oligosaccharides that differ in structure by one of these glycosyl residues. Figure 6C shows the chromatograms of three oligosaccharide pairs differing by the presence of a Xyl*p* residue. Interestingly, α-L-Ara*p* has only a small effect on the elution time of an oligosaccharide. Fucosylation has a more drastic effect on the elution, as all fucosylated oligosaccharides are eluted earlier compared to their non-fucosylated counterparts. Furthermore, reduction of an oligosaccharide to the corresponding alditol affects the interaction of the oligosaccharide with the resin more than losing a non-reducing end Glc*p* residue. These results are presented in Figure 6D. Since we used a multi-step gradient in our separations, extrapolation of these observations should be done with caution. However, these results demonstrate that in a given separation conditions, the elution times can be used to predict the elution behaviour or XyG oligosaccharides. The "Glycotree" concept (Takahashi & Kato, 2003) that relies on the same principle of additivity of contributions, can be used for more in-depth analysis of oligosaccharide elution times as well as response factors.

### 2.4. Conclusions

We generated and purified a total of thirty-one oligosaccharides to (near) homogeneity using enzymatic hydrolyses and preparative scale size-exclusion chromatography. Additionally, we produced mixtures of isomeric XyG oligosaccharides that contain both 1→4 and 1→3 linkages in the backbone.

Besides these *bona fide* XyG oligosaccharides containing various side-chain structures, we also purified cellodextrins that can serve as XyG analogs. Several eluent systems were tested to obtain good separation of the oligosaccharides in SEC. We conclude that volatile acetic acid buffer is an ideal eluent system to separate XyG oligosaccharides in extra fine Bio-Gel P-2 resin. The oligosaccharides were purified in their reducing forms allowing their covalent conjugation for various downstream applications. The purity and structural identity of these oligosaccharides were validated by MALDI-TOF MS, HPAEC-PAD and NMR spectroscopy. We demonstrated that this collection of pure oligosaccharides is useful as mass spectrometry and chromatography standards We showed that for rigorous quantification of XyG oligosaccharides, MALDI-TOF MS and HPAEC-PAD should be used in conjunction due to isomeric and coelution effects. We further envision that these oligosaccharides can be useful in examining the binding specificities of antibodies as well as substrate specificies of glycosyl transferases and glycosyl hydrolases.

## 2.5. Experimental

### 2.5.1. General Procedures

Removal of organic solvents and concentration of samples for efficient ethanol precipitation, dialysis or lyophilization was routinely accomplished by evaporation under reduced pressure using rotary evaporation apparatus (Büchi) with the water bath temperature set at 45 °C.

Fractionation of samples into soluble and insoluble components, in various solvent systems, was performed by centrifugation at 10000*g* for 30 min at 4 °C rather than by filtration, to minimize filtrate and retentate losses as well as potential contamination from the filter. Since native XyGs are very soluble in water, their aqueous solutions were routinely clarified by centrifugation.

Coarse fractionation of samples into high and low molecular weight material, for example into polysaccharide and oligosaccharide components, was routinely accomplished by ethanol precipitation with a final ethanol concentration of 75%. Three volumes of pre-cooled absolute ethanol was added to pre-cooled aqueous solution of the starting material, and after gentle mixing, the mixture was incubated at 4 °C for at least 2 h. Precipitated material was pelleted by centrifugation and redissolved in water.

Elimination of unwanted low-molecular weight material, including non-volatile buffer salts from aqueous polysaccharide preparations, was routinely accomplished by dialyzing two times or more against a 100-fold or larger volume of ultrapure water in a 3000 molecular weight cutoff regenerated cellulose tubing (Spectra/Por) at 4 °C with vigorous stirring. Under these conditions a single dialysis step reaches equilibrium in approximately five hours.

Enzymes were removed from reaction mixtures by spin filtering through a 10 kDa cutoff modified Nylon membrane (Nanosep MF, Pall Life Sciences).

Oligo- and polysaccharide preparations were typically stored in their lyophilized forms. Ammonium acetate and acetic acid used in enzymatic

digestions and separations are volatile, and thus can be removed from the samples by lyophilization.

### 2.5.2. Plant Cell Wall Polysaccharides

*Preparation of Alcohol-insoluble Residue (AIR) from Plant Material* The procedure used was adapted from (York et al, 1986) with some modifications. AIR was prepared from either plant tissue or suspension cultured cells by homogenization followed by extractions with organic solvents. Frozen plant material was disrupted by immersion in pre-cooled absolute ethanol (to a final concentration of 1 g tissue per 4 mL) with mixing overnight at 4 °C. The plant material was then mechanically homogenized by Polytron (Ultra-Turrax) at 4 °C until a smooth suspension was achieved. The suspension was vacuum filtered using a 50 μm Nylon net until the solvent draining stopped, and the retentate was resuspended in 75% ethanol to a final concentration of 1 g per 50 mL (based on the original plant material weight) with stirring for 1 h at room temperature, and again filtered. The overnight extraction and 1 h wash were repeated first with 1:1 (vol:vol) methanol:chloroform and then with acetone (all with volumes equal to the 75% ethanol extraction) at room temperature. The final residue was thoroughly air dried and sifted through a metal grid (850 μm openings) to yield AIR.

*Sequential Extraction of AIR* The procedure used was adapted from (York et al, 1986). The AIR was suspended in 50 mM ammonium oxalate, pH 5.0 buffer containing 0.2 mg/mL sodium azide, to a final concentration of 1 g per 50

mL and stirred overnight at room temperature. The suspension was vacuum filtered on a 50 μm Nylon net until buffer draining stopped and the retentate was resuspended in the same buffer (with volume equal to the first extraction), stirred for 1 h, and filtered again. The overnight extraction and 1 h wash were repeated sequentially with 1 M potassium hydroxide containing 1 mg/mL sodium borohydride, and 4 M potassium hydroxide solutions. The pre-cooled potassium hydroxide filtrates were stirred on ice, and glacial acetic acid was added dropwise until hydrogen gas generation stopped and a drop of the solution tested neutral on pH paper. If necessary, excessive foaming during the neutralization was quenched by few drops of octanol. The filtrates were ethanol precipitated and dialyzed to yield high molecular weight extracts.

*Tamarind Seed XyG Polysaccharide* The procedure used was adapted from (York et al, 1990). Defatted tamarind (*Tamarindus indica*) seed flour was suspended in water to a final concentration of 10 mg/mL and stirred overnight to extract water soluble material. The water insoluble material was pelleted by centrifugation and the water soluble polymeric material in the supernatant was precipitated by addition of three volumes of absolute ethanol. The precipitate was pelleted by centrifugation, dissolved in water and dialyzed. The dialysate was deionized by passing through anion and cation exchange columns to yield pure tamarind seed XyG polymer.

*Jojoba Seed XyG Polysaccharide* The procedure used was adapted from (Hantus et al, 1997). Briefly, defatted jojoba (*Simmondsia chinensis*) seed flour was extracted with 4 M potassium hydroxide solution containing 1 mg/mL sodium

borohydride. The extract was further fractionated by anion and cation exchange chromatography to yield pure neutral jojoba seed XyG polymer.

*Suspension-cultured Sycamore Maple XyG Polysaccharide* Sycamore maple (*Acer pseudoplatanus*) XyG was prepared essentially as described (York et al, 1986). The cells from a starter culture, kindly provided by Stefan Eberhard, were cultured ten days in modified Torrey and Shigemura M-6 medium (Torrey & Shigemura, 1957) as described (Talmadge et al, 1973), and harvested by vacuum filtering the cell suspension with a 50 μm Nylon net. The cell cake was washed with water, vacuum aspirated until water draining stopped, and frozen until analysis. AIR was prepared from the cells and sequentially extracted to produce cell wall polysaccharides. Polymeric material was also recovered from the spent culture medium by condensing the filtrate by rotary evaporation to approximately one fifth of the original volume, followed by ethanol precipitation and dialysis. Neutral polysaccharides were further purified from the spent culture medium and the 4 M potassium hydroxide extract of AIR by anion and cation exchange chromatographies.

*African Clubmoss XyG Polysaccharide* African clubmoss (*Selaginella kraussiana*) was collected from University of Georgia Department of Plant Biology greenhouse. Neutral polysaccharides were purified by sequential chemical extractions of the AIR prepared from aerial tissue, and purification of neutral polysaccharides from the 4 M potassium hydroxide extract.

*Ion Exchange Chromatography* Neutral polysaccharides were separated from charged polysaccharides and proteins by sequential strong anion and cation exchange chromatographies. Q Sepharose FF anion exchange resin (GE Healthcare) was packed in a borosilicate column (Bio-Rad) and conditioned by sequential washes with three bed volumes of water, 1 M imidazole, pH 7.0; and finally 10 mM imidazole, pH 7.0 buffer. XyG polysaccharide preparation was dissolved in 10 mM imidazole, pH 7.0 buffer to a final concentration of 5 mg/mL and top-loaded on the resin (up to 5 mL per 1 mL of resin). The column was then washed with three bed volumes of the same buffer. Load and wash flow-throughs were pooled, concentrated, ethanol precipitated, and dialyzed to yield anion exchanged polysaccharide preparation. The bound anionic fraction ("pectic polysaccharides") was recovered by elution with three column volumes of 1 M imidazole, pH 7.0 buffer followed by ethanol precipitation and dialysis of the eluate. SP Sepharose FF cation exchange resin (GE Healthcare) was packed in a borosilicate column and conditioned by sequential washes with water, 20 mM sodium acetate, pH 5.0 containing 0.5 M sodium chloride, and finally 20 mM sodium acetate, pH 5.0 buffer. Anion exchanged polysaccharide preparation was dissolved in 20 mM sodium acetate, pH 5.0 buffer, top-loaded on the resin, and washed in with the same buffer. Load and wash flow-throughs were pooled, concentrated, ethanol precipitated, and dialyzed to yield fully ion exchanged (neutral) polysaccharide preparation.

Removal of introduced enzymes from oxidized XyG polymer samples (see below) was performed by strong anion exchange chromatography (protocol

devised by Malcolm O'Neill at CCRC, unpublished). The Q Sepharose resin was conditioned as described above. The oxidized XyG preparation was dissolved in 10 mM imidazole, pH 7.0 buffer to a final concentration of 5 mg/mL and top-loaded on the resin (up to 2 mL per 1 mL of resin). The protein contaminants were eluted with three bed volumes of 100 mM imidazole, pH 7.0 buffer and discarded. The oxidized XyG polymer was recovered by elution with three bed volumes of 1 M imidazole, pH 7.0 buffer, followed by ethanol precipitation and dialysis of the eluate.

*Generation of Anionic XyG* Chemo-enzymatic oxidation of neutral XyG with D-Gal*p* residues to anionic XyG with D-Gal*p*A residues was performed essentially as described previously (Parikka et al, 2012), with some modifications. Tamarind XyG polymer preparation was dissolved in 100 mM sodium phosphate, pH 7.0 buffer to a final concentration of 10 mg/mL. For the enzymatic oxidation of the D-Gal*p* residues to 6-aldehydo-D-Gal*p* residues, *Dactylium dendroides* galactose oxidase, type II horseradish peroxidase, and bovine liver catalase (all enzymes from Sigma-Aldrich) were added to final concentrations of 5, 5 and 500 U per 1 mg of D-Gal*p* residues, respectively. The reaction was allowed to proceed for 48 h at room temperature with stirring. For the chemical oxidation of 6-aldehydo-D-Gal*p* residues to D-Gal*p*A residues, iodine and potassium iodide (both from Sigma-Aldrich) were added to final concentrations of 12 and 2.5 mg per 1 mg of original D-Gal*p* residues, respectively, the pH was adjusted to 9.0 by dropwise addition of 1 M sodium carbonate, and the solution was stirred in dark for five days at room temperature.

After oxidation, excess iodine was reduced to iodide by dropwise addition of 0.5 M sodium metabisulfite until the characteristic blue color of iodine-XyG complex (Gould et al, 1971) disappeared, and the reagents were removed by ethanol precipitation and dialysis. The oxidized XyG was separated from the enzymes by anion exchange chromatography (see above).

### 2.5.3. Generation and Purification of Oligosaccharides

*XyG S1 Oligosaccharides* XyG polymer preparation was dissolved in 50 mM ammonium acetate, pH 4.5 buffer to a final concentration of 10 mg/mL. XyG-specific endoglucanase [either *Aspergillus aculeatus*, Novozymes (Pauly et al, 1999) or *Paenibacillus* sp., Megazyme (Gloster et al, 2007)] was added to a final concentration of 100 mU/mL (1 U, amount of substrate-saturated XEG that creates reducing ends at a rate of 1 μmol/min) and the reaction was allowed to proceed overnight at room temperature. XEG and the residual undigested polymeric material were removed by ethanol precipitation and centrifugation. The supernatant was concentrated by rotary evaporation to remove the ethanol and lyophilized to yield XyG oligosaccharides.

*Minimal XyG Oligosaccharides* XyG S1 oligosaccharide preparation was dissolved in 50 mM ammonium acetate, pH 4.0 buffer to a final concentration of 10 mg/mL. *Geotrichum* sp. M128 oligoxyloglucan reducing end cellobiohydrolase (OXG-RCBH) (Yaoi & Mitsuishi, 2002), generously provided by Dr. Katsuro Yaoi, was added to a final concentration of 100 mU/mL and the reaction was allowed to proceed overnight at room temperature. The enzyme

was removed from the reaction mixture by spin filtering (10 kDa cutoff) and the oligosaccharides were purified by SEC.

*Isoprimeverosyl-series XyG Oligosaccharides* XXXG was dissolved in 50 mM ammonium acetate, pH 4.5 buffer to a final concentration of 100 mg/mL. *Aspergillus aculeatus* β-galactosidase preparation (Novozymes) with an isoprimeverosyl hydrolase (IPH) contaminant (York et al, 2002) was added to a final concentration of 2 mg/mL and the reaction was allowed to proceed overnight at room temperature. Enzymes were removed from the reaction mixture by spin filtering (10 kDa cutoff) and the oligosaccharides were purified by SEC.

*Modification of XyGs with Glycoside* Exo-*hydrolases* XyG preparation was dissolved in an appropriate buffer to a final concentration of 10 mg/mL, enzyme was added to a final concentration of 100 mU/mL and the reaction was allowed to proceed overnight at room temperature. The enzymes and reaction conditions used were as follows: Degalactosylation: *Aspergillus niger* β-galactosidase (Megazyme) in 50 mM ammonium acetate, pH 4.5; Dexylosylation: either *Bacillus* sp. No. 693-1 α-xylosidase (Seikagaku Corporation) in 50 mM sodium phosphate, pH 7.0. Enzymes were removed from the oligosaccharide reaction mixtures by spin filtering (10 kDa cutoff) and the oligosaccharides were purified by SEC. Enzymes were removed from the polysaccharide reaction mixtures by anion exchange chromatography. Monosaccharides and non-volatile buffer salts were removed from the polysaccharide reaction mixtures by ethanol precipitation and dialysis.

*Reduction of XyG Oligosaccharides* XyG oligosaccharide preparation was dissolved in 50 mM sodium hydroxide solution containing 10 mg/mL sodium borohydride to a final concentration of 10 mg/mL and the reaction was allowed to proceed for 2 h at room temperature. The pH of the sample was adjusted to approximately 4.5 with 1 M monobasic sodium phosphate, and purified by solid-phase extraction (see below).

*Generation and Purification of Cellodextrins* Cellodextrins were prepared as described elsewhere (Liebert et al, 2008) with minor modifications. Briefly, dried microcrystalline cellulose (Avicel PH-101, Fluka) was suspended in 85% aqueous phosphoric acid (J.T. Baker) to a final concentration of 1 g per 12 mL in a round evaporation flask and mixed by rotation in air with rotary evaporation apparatus (Büchi) for 30 minutes to form a uniform suspension. The flask was then immersed into a 55 °C water bath and rotated for 20 hours to partially hydrolyze the cellulose. The reaction mixture was cooled on ice and the cellodextrins were precipitated by adding ten volumes of ice cold absolute ethanol. The precipitate was collected by vacuum filtration through a glass microfiber filter (1.6 μm nominal pore size, GF/A, Whatman). Suspension in cold ethanol and filtering were repeated until small aliquot of water dissolved cake tested neutral on pH paper. The cake was vacuum aspirated, until ethanol draining stopped, suspended in water to a final concentration of 50 mg/mL (on a wet weight basis), and centrifuged to separate water soluble and insoluble materials. The supernatant was lyophilized to yield water soluble cellodextrins

with degree of polymerization (DP) ranging from 1 (glucose) to 7 (celloheptaose). The individual cellodextrins were purified by SEC.

*Preparative SEC of Oligosaccharides* Extra fine Bio-Gel P-2 (Bio-Rad) SEC resin was hydrated and packed into a borosilicate column (Bio-Rad) and equilibrated with at least three bed volumes of appropriate eluent prior to use. Chromatography eluents were prepared in ultrapure water and vacuum filtered through a 0.2 μm polycarbonate membrane (Nucleopore, Whatman). The following eluents were utilized: water; 10 mM ammonium acetate, pH 5.0; 50 mM acetic acid (pH approximately 3.0); and 10 mM boric acid, 10 mM phytic acid, pH 10 (adjusted with sodium hydroxide). Oligosaccharide samples were dissolved in the corresponding chromatography eluent to a final concentration of 25 to 50 mg/mL and spin filtered through a 0.2 μm modified Nylon membrane (Nanosep MF, Pall Life Sciences). Typically 10 mg of the oligosaccharide sample was top-loaded on the resin and eluted by gravity with approximately 2 m hydrostatic head, resulting in flow rate of 4 to 6 ml/h. Fractions were collected using a Foxy Jr. fraction collector (Teledyne Isco) in the drop counting mode and assayed for oligosaccharides by sulfuric acid-phenol assay as well as MALDI-TOF MS.

*Desalting of Oligosaccharides* Sephadex G-15 (GE Healthcare) size-exclusion chromatography resin was packed into a 1.5 × 50 cm borosilicate column (Bio-Rad) to an approximate bed volume of 85 mL and equilibrated in water. Oligosaccharide separation from non-volatile salts and monosaccharides, as well as fraction collection and analyses was performed as described above.

*Solid-phase Extraction* Oligosaccharide samples in non-volatile eluents were desalted with solid phase extraction (SPE) using octadecylsilane (ODS) C18 cartridges (500 mg resin, Supelco) according to (York et al, 1996). The SPE cartridges were preconditioned by washing three times with 3 ml (as were all subsequent washes and elutions) of methanol followed by three washes with water. Eluent flow was facilitated using pressurized air. The pH of the carbohydrate containing samples were adjusted to approximately 4.5 by adding an equal volume of 1 M monobasic sodium phosphate to the sample, and applied to the SPE cartridge. The resin was washed with water to remove salts and the oligosaccharides were subsequently eluted with 25% aqueous methanol. The methanol in the eluate was evaporated under a stream of air and the sample was lyophilized to yield pure oligosaccharides.

### 2.5.4. Analytical Techniques

*Sulfuric Acid-phenol Assay for Carbohydrates* Samples were analyzed for carbohydrate content using miniaturized sulfuric acid-phenol assay, with modifications from published protocols (Fox & Robyt, 1991; Masuko et al, 2005). One hundred microliters of concentrated sulfuric acid (J.T. Baker), 30 μl of ice-cooled sample, and 20 μl of 5% aqueous phenol (J.T. Baker) were transferred in this order to the wells on a 96-well polystyrene microplate (Costar 3598, Corning). The microplate was floated uncovered on a near boiling (> 90 °C) water bath for 5 min for color development and subsequently cooled to room temperature on an ice bath. The absorbances were quantified spectrophotometrically at 490 nm using a microplate reader.

*Preparation of Analytical Oligosaccharide Solutions* To prepare XyG oligosaccharide solutions with accurately known concentrations, a portion of lyophilized oligosaccharide preparation was weighed to 10 µg precision (Sartorius analytical balance) and mixed with weighed amount of ultrapure water. A nominally equimolar mixture of six common XyG oligosaccharides (GXXG, XXXG, XXLG, XXFG, XLLF and XLFG) was prepared by mixing equal volumes of these oligosaccharide solutions.

*MALDI-TOF MS* MALDI-TOF mass spectra were collected on Nitrogen laser equipped MicroFlex LT instrument (Bruker) in positive mode. Equal volumes of the oligosaccharide sample in water and the matrix solution (20 mg/mL 2,5-dihydroxybenzoic acid in 1:1 (vol:vol) methanol:water) were mixed and 1 µl of the mixture was applied and allowed to air dry on a ground steel sample plate (Bruker). Mass calibration was performed with malto-oligosacharides (Sigma-Aldrich). For routine analysis of SEC fractions, the laser power (typically < 35%) and the number of summed laser pulses (typically 50) were adjusted separately for each sample spot to yield spectra with high signal-to-noise ratios. For analytical work using the nominally equimolar mixture, the same laser power (35%) and summed laser shots (100) were used across experiments. MALDI-TOF MS signal integrals were extracted from the MS spectra using Microsoft Excel. Four spectra were included for each analysis (two repeats from two different sample spots). One of the four spectra was arbitrarily chosen as standard and the intensities of three others were scaled by an individual scaling factor. The three floating scaling factors were adjusted to

minimize the sum of the coefficients of variation of the intensities of each of the oligosaccharide among the four experiments. The GRG nonlinear algorithm in the Solver add-on in Microsoft Excel was used in this linear minimization procedure.

*HPAEC-PAD* HPAEC-PAD analysis was performed with a Dionex ICS-3000 chromatography system (Thermo Scientific) equipped with an autosampler and a pulsed amperometric detector (PAD). Pure oligosaccharides were dissolved in water to a final concentration of 0.1 mM (except for XFFG whose concentration was unknown) and 10 μl was injected to a CarboPac PA1 column (Thermo Scientific). Reconstituted mixtures from pure oligosaccharide preparations were injected in various total concentrations. Two injections were included for each analysis. Oligosaccharides were eluted at a flow-rate of 1.0 mL/min with a multi-step gradient of sodium acetate (0 to 3 min: 0 mM; 3 to 5 min: 0 to 40 mM, 5 to 27 min: 40 to 80 mM, 27 to 50 min: 80 to 430 mM, 50 to 52 min: 430 mM, 52 to 53 min: 430 to 0 mM, 53 to 60 min: 0 mM) in a constant background of 100 mM sodium hydroxide. The chromatograms were analyzed with Chromeleon software (version 6.8, Thermo Scientific).

*NMR Spectroscopy* One to three milligrams of pure oligosaccharide was lyophilized once in 500 μl of $D_2O$ (99.9%, Cambridge Isotope Laboratories) and then dissolved for analysis in 450 μl of $D_2O$ (99.96%, Cambridge Isotope Laboratories) with approximately 0.5 μl of acetone as an internal chemical shift standard. One- and two-dimensional NMR spectra were recorded at 25 °C with 5

mm cold probe equipped Varian Inova NMR spectrometer operating at 600 MHz $^{1}$H frequency. Typical one dimensional $^{1}$H NMR spectra were obtained by summing 128 transients, each with 8192 complex points and spectral width of 3612.7 Hz (0 to 6 ppm). Typical two dimensional gradient COSY spectra were collected with 4 summed transients and 800 increments in the indirectly detected dimension. The spectra were analyzed with Mnova NMR software [version 8.1.2, Mestrelab Research, (Cobas & Sardina, 2003)]. The $^{1}$H chemical shifts were referenced to internal acetone standard ($\delta$ = 2.225).

### 2.6. Acknowledgements

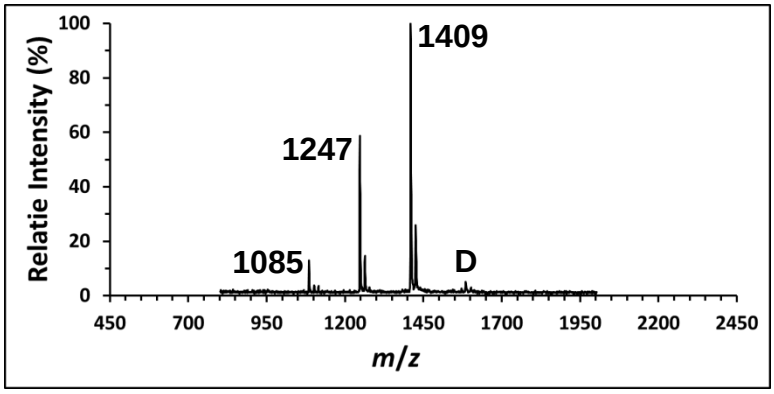
We are grateful for Dr. Maria J. Peña, Ameya Kulkarni and Dr. John Glushka for assistance on NMR analysis, Malcom O'Neill for insightful suggestions, and Dr. Chin Fen Teo for critical reading of the manuscript. We thank Dr. Katsuro Yaoi for providing the OXG-RCBH enzyme. This work was supported by the National Science Foundation Plant Genome Program (grant no. DBI–0421683).

**A** MALDI-TOF MS

**Tamarind XyG**

**XEG**

**hydrolysate**

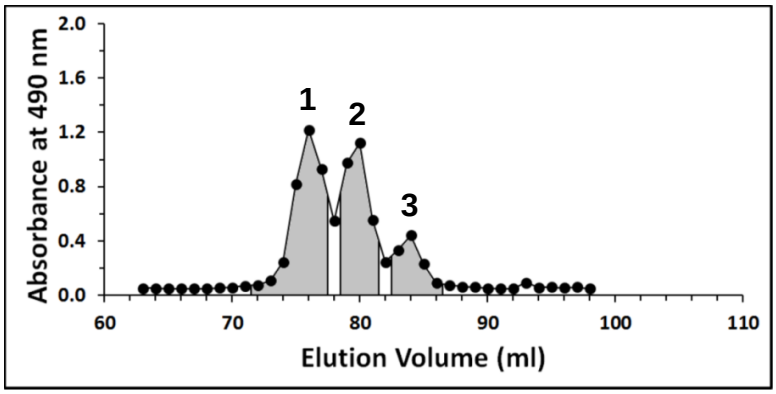



**B** SEC

**Tamarind XyG**

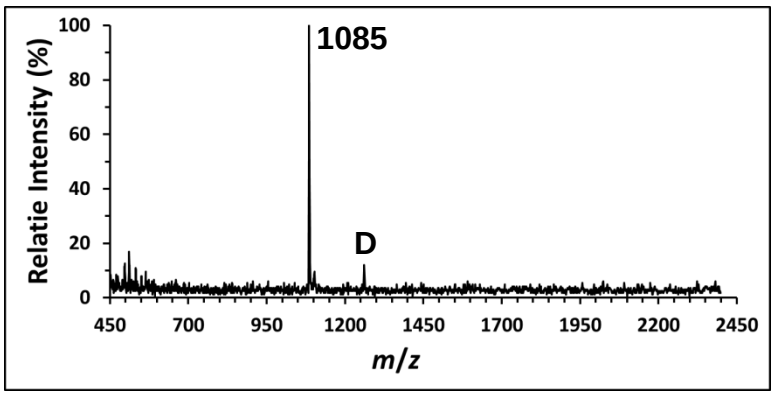
**XEG**

**hydrolysate**



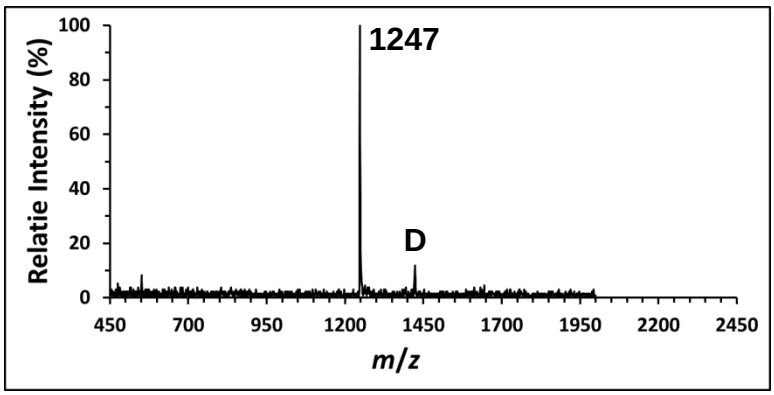

**C** MALDI-TOF MS

**Peak 3**




**D** MALDI-TOF MS

**Peak 2**



**E** MALDI-TOF MS

**Peak 1**

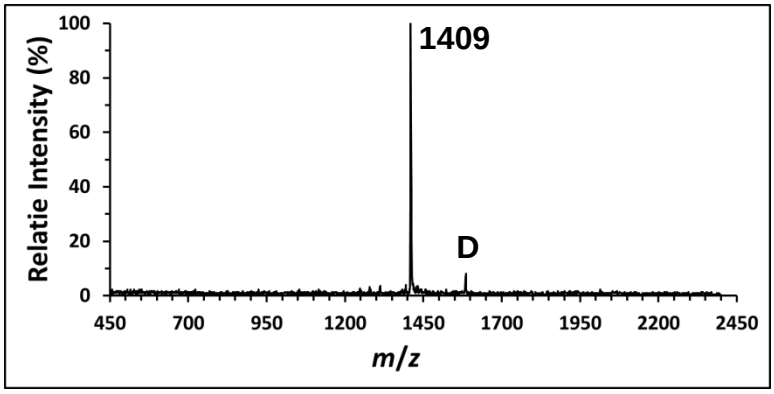
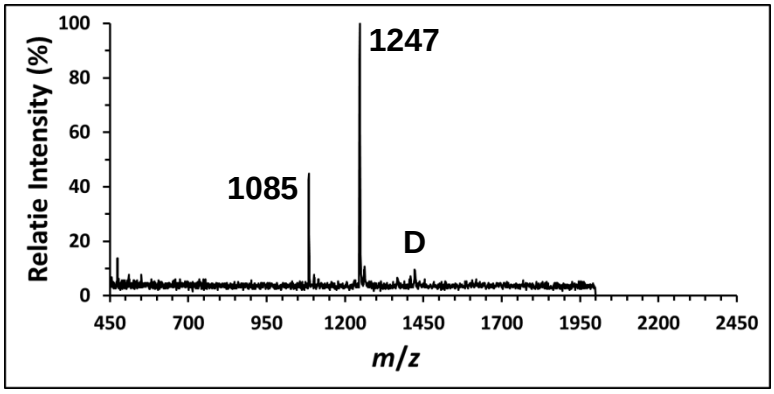



**Figure 2-1**. The generation, purification and structural validation of oligosaccharides from XEG hydrolysis of tamarind XyG polysaccharide. MALDI-TOF MS spectrum (**A**) and SEC chromatogram (**B**) of the hydrolysate. The hydrolysis yields XXXG (Peak 3, *m*/*z* 1085), XLLG (Peak 1, *m*/*z* 1409), and the isomeric XLXG and XXLG (Peak 2, *m*/*z* 1247) as the major products. The SEC fractions were analyzed by MALDI-TOF MS (data not shown) and fractions containing pure oligosaccharides (excluding Peak 2 which contains both XLXG and XXLG) were pooled as indicated by the chromatogram shading. XXXG and XLLG can be separated in their pure forms by SEC, whereas the isomeric XLXG and XXLG co-elute. MALDI-TOF MS spectra of the pure XXXG (**C**), mixture of XLXG and XXLG (**D**), and pure XLLG (**E**) are shown. The purities and structural identities of the XXXG and XLLG oligosaccharides were further established by HPAEC-PAD and NMR (see Supplemental Figures 2K and 2S). The major signal in each MALDI-TOF MS spectra arises from the sodium adduct of the oligosaccharide, $[M + Na]^+$, and is indicated by its *m*/*z* value. Other prominent, but significantly less intense signals, such as the potassium adduct $[M + K]^+$ at 16 Th higher *m*/*z* value to the parent peak, double adduct of sodium and 2,5-

dihydroxybenzoic acid (DHB) $[M + Na + DHB]^+$ at 176 Th higher *m*/*z* value (indicated by D), are observed in some spectra.

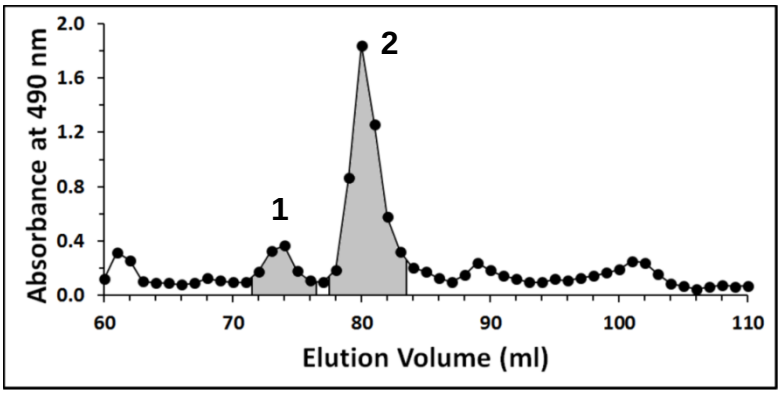
**A** MALDI-TOF MS

**Tamarind XyG β-galactosidase and XEG hydrolysate**



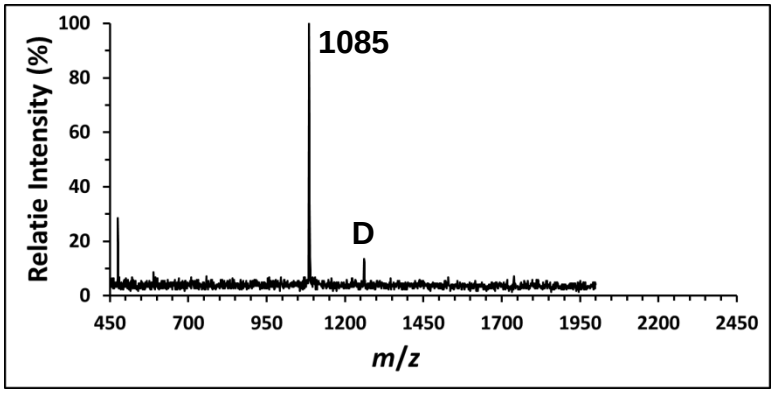

**B** SEC

**Tamarind XyG β-galactosidase and XEG hydrolysate**




**C** MALDI-TOF MS

**Peak 2**




**D** MALDI-TOF MS

**Peak 1**

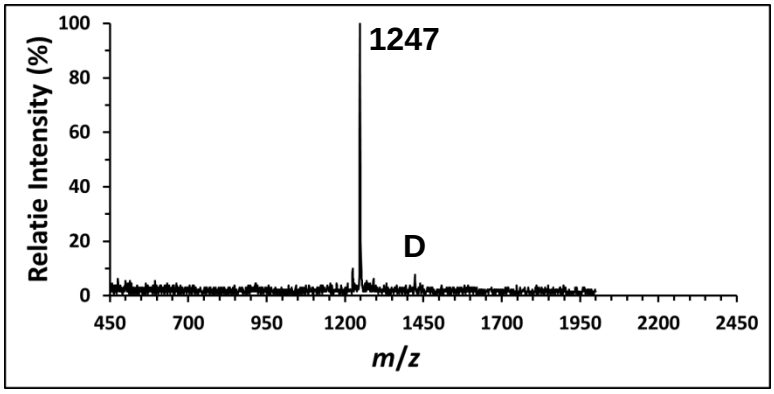
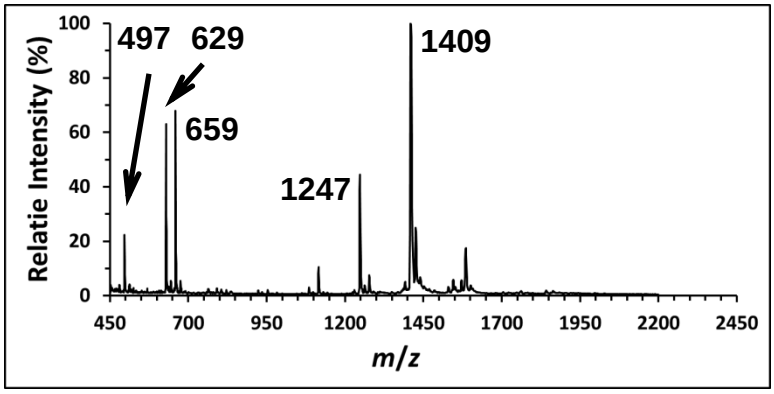


**Figure 2-2**. The generation, purification and structural validation of oligosaccharides from sequential β-galactosidase and XEG hydrolyses of tamarind XyG polysaccharide. MALDI-TOF MS spectrum (**A**) and SEC chromatogram (**B**) of the hydrolysate. The β-galactosidase hydrolyzes XLXG subunit to XXXG and XLLG subunit to XXLG. Subsequent XEG hydrolysis of this simplified polysaccharide yields a mixture of XXXG (Peak 2, *m*/*z* 1085) and XXLG (Peak 1, *m*/*z* 1247). The SEC fractions were analyzed by MALDI-TOF (data not shown) and fractions containing pure oligosaccharides were pooled as indicated by the chromatogram shading. MALDI-TOF MS spectra of the pure XXXG (**C**), and XXLG (**D**) oligosaccharides. The purity and structural identity of the XXLG oligosaccharide was further established by HPAEC-PAD and NMR (see Supplemental Figure 2). See also the legend for Figure 1.

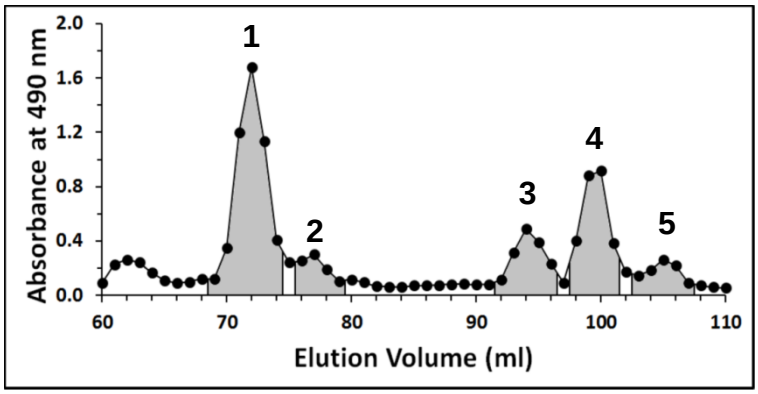
**A** MALDI-TOF MS

**Tamarind XyG XEG and OREX hydrolysate**



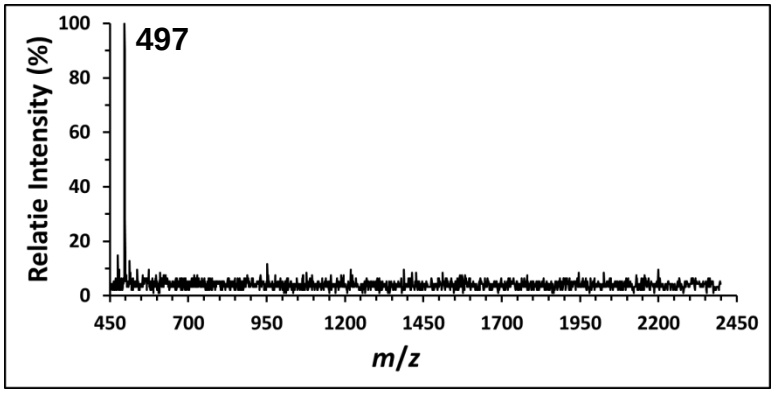

**B** SEC

**Tamarind XyG XEG and OREX hydrolysate**



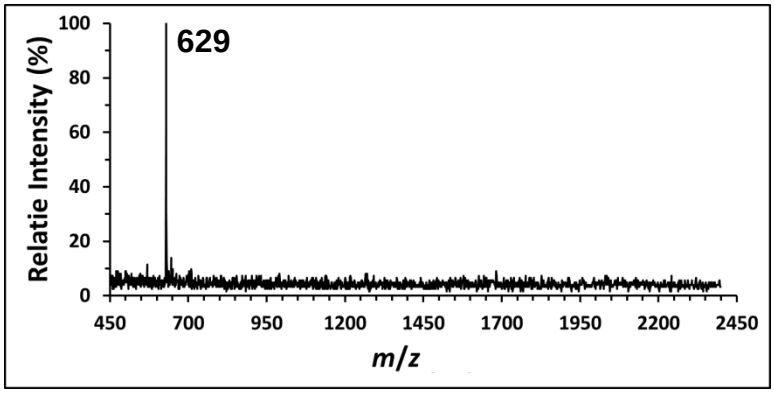

**C** MALDI-TOF MS

**Peak 5**




**D** MALDI-TOF MS

**Peak 4**



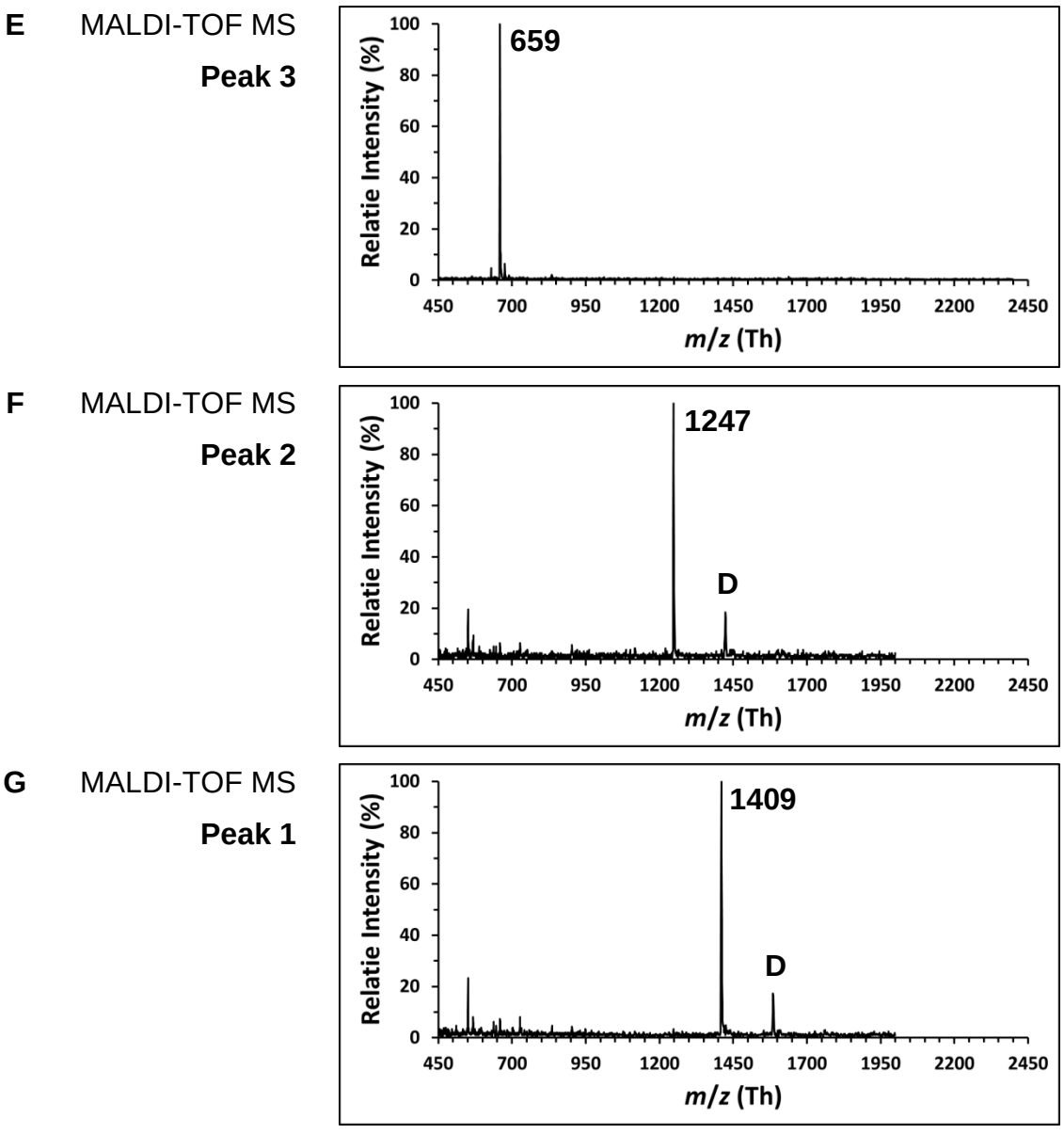
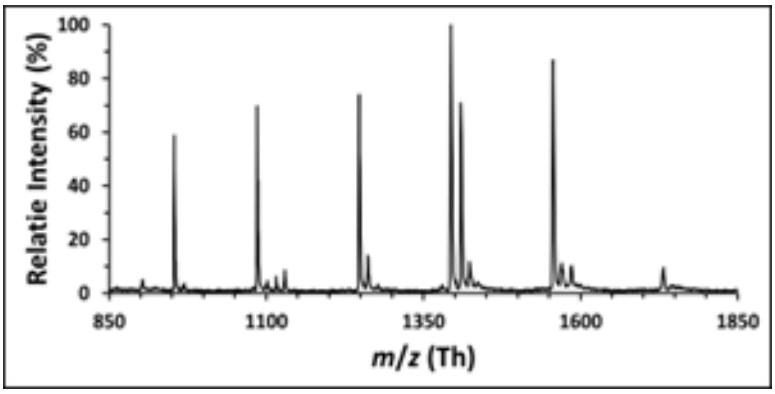


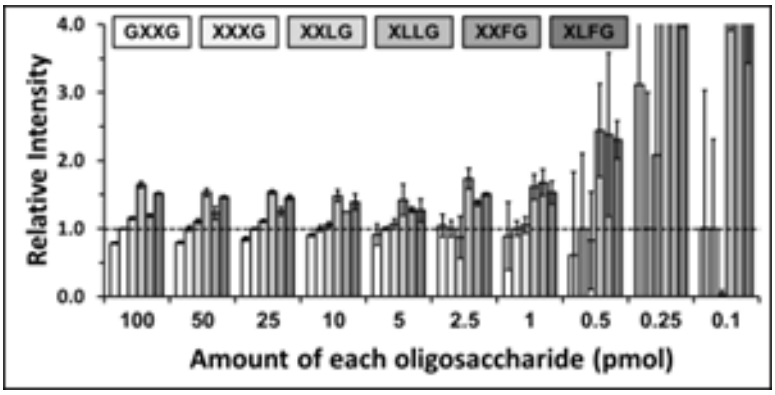

**Figure 2-3**. The generation, purification and structural validation of oligosaccharides from sequential XEG and OXG-RCBH hydrolyses of tamarind XyG polysaccharide. MALDI-TOF MS spectrum (**A**) and SEC chromatogram (**B**) of the hydrolysate. The OXG-RCBH hydrolyzes XXXG oligosaccharide to XX (Peak 4, *m*/*z* 629) and XG (Peak 5, *m*/*z* 497) fragments, and XXLG to XX and LG (Peak 3, *m*/*z* 659) fragments. XLXG (Peak 2, *m*/*z* 1247) and XLLG (Peak 1, *m*/*z*

1409) are not substrates for this enzyme and remain intact. The SEC fractions were analyzed by MALDI-TOF (data not shown) and fractions containing pure oligosaccharides were pooled as indicated by the chromatogram shading. MALDI-TOF MS spectra of the pure XG (**C**), XX (**D**), LG (**E**), XLXG (**F**), and XLLG (**G**) oligosaccharides are shown. The purities and structural identities of the XG, XX, LG, and XLXG oligosaccharides were further established by HPAEC-PAD and NMR (see Supplemental Figure 2). See also the legend for Figure 1.

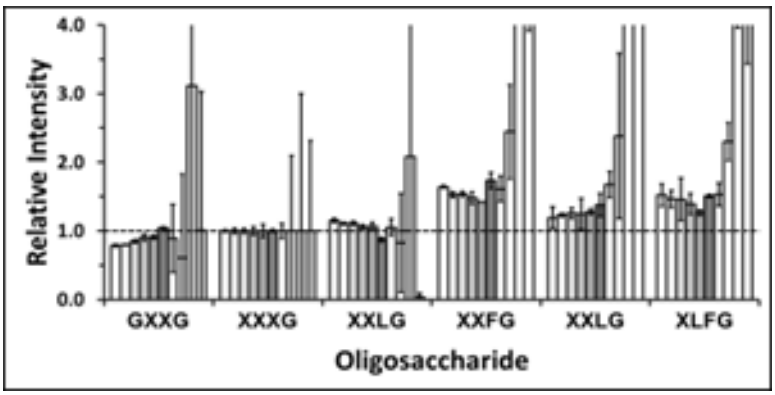
**A** MALDI-TOF MS **Mixture with 100 pmol of each of six XyG oligosaccharides**




**B** Side-by-side comparison of normalized integrals of XyG oligosaccharides




**C** Side-by-side comparison of normalized integrals of XyG oligosaccharides




**D** Relative response factors **XyG oligosaccharides**

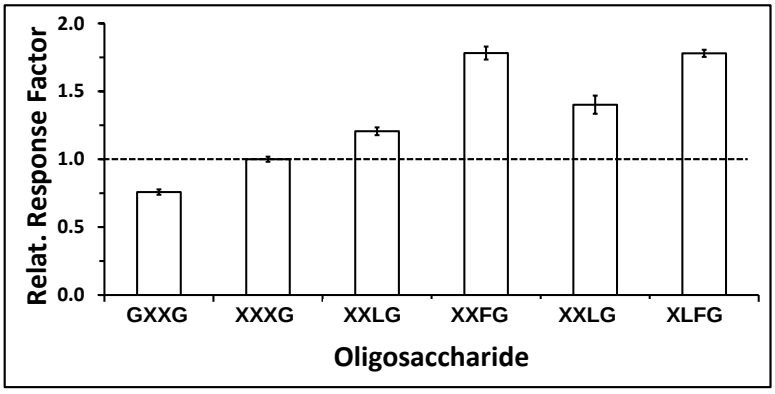
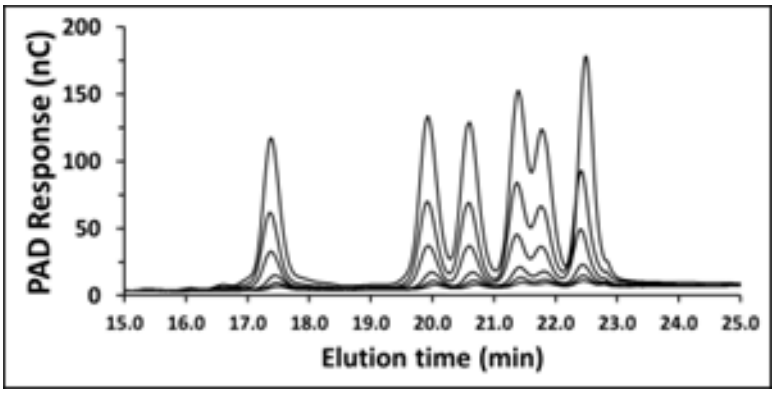
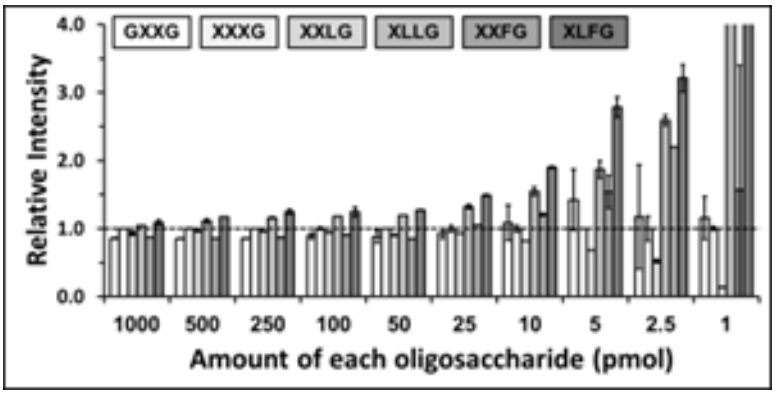


**Figure 2-4.** MALDI-TOF MS analysis of XyG oligosaccharide response factors. (**A**) MALDI-TOF MS spectrum of a mixture of GXXG, XXXG, XXLG, XXFG, XLLG and XLFG with 100 pmol of each oligosaccharide. (**B**) Comparison of the normalized signal integrals of the oligosaccharides measured for various dilutions of the mixture. Each group of bars represents a dilution and each bar in a group represents an oligosaccharide. (**C**) Comparison of the normalized signal integrals of the oligosaccharides measured for various dilutions of the mixture. Each group of bars represents an oligosaccharide and each bar in a group represents a dilution. In both B and C, the integrals are normalized to that of XXXG (set to 1). In both **B** and **C**, the reported values are averages of three measurements with the error equal to their standard deviation. (**D**) The calculated relative response factors of the XyG oligosaccharides with that of XXXG set to 1.

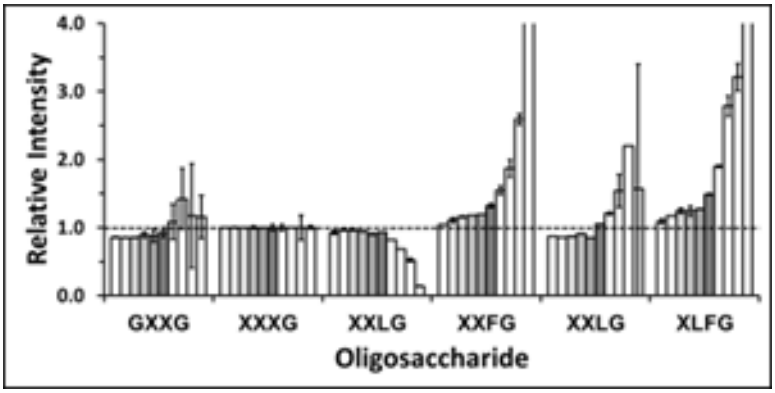
**A** HPAEC-PAD **Mixture with 100 pmol of each of six XyG oligosaccharides**



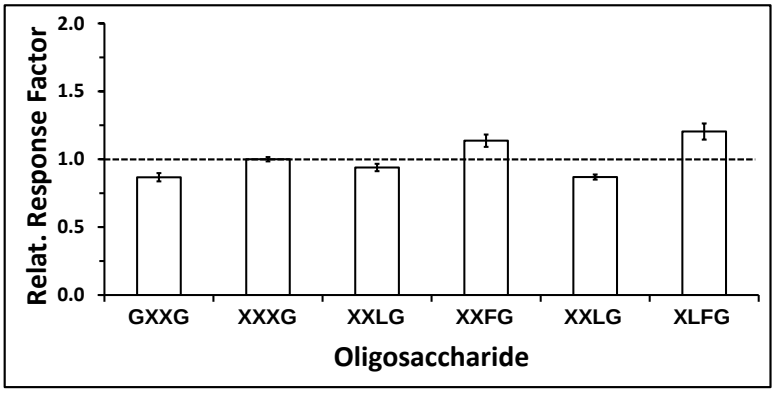

**B** Side-by-side comparison of normalized integrals of XyG oligosaccharides




**C** Side-by-side comparison of normalized integrals of XyG oligosaccharides




**D** Relative response factors **XyG oligosaccharides**



**E**

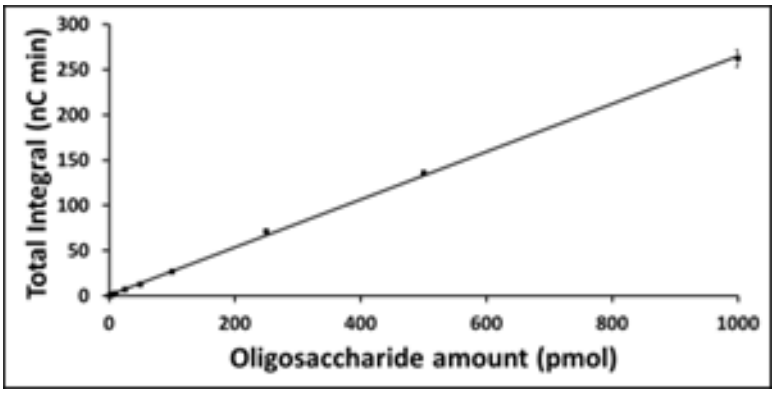
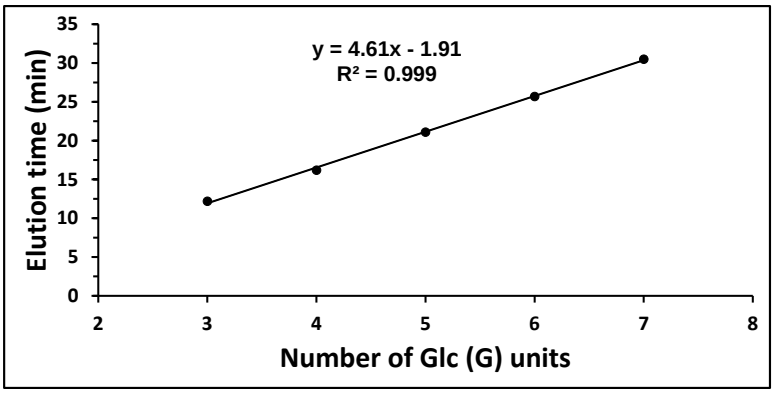
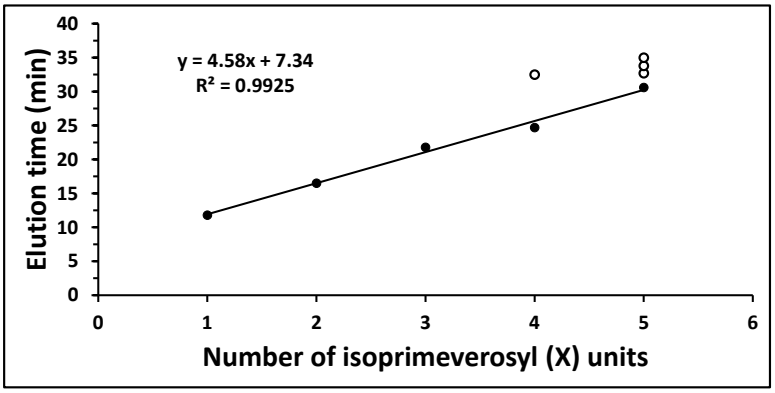



**Figure 2-5.** HPAEC-PAD analysis XyG oligosaccharide response factors. (**A**) Overlaid HPAEC-PAD chromatograms of a mixture of GXXG, XXXG, XXLG, XXFG, XLLG and XLFG with different total amounts injected (in pmol per oligosaccharide: 1000, 500, 250, 100, 50, 25, 10, 5, 2.5, and 1). (**B**) Comparison of the normalized signal integrals of the oligosaccharides measured for various dilutions of the mixture. Each group of bars represents a dilution and each bar in a group represents an oligosaccharide. (**C**) Comparison of the normalized signal integrals of the oligosaccharides measured for various dilutions of the mixture. Each group of bars represents an oligosaccharide and each bar in a group represents a dilution. In both B and C, the integrals are normalized to that of XXXG (set to 1). In both **B** and **C**, the reported values are averages of three measurements with the error equal to their standard deviation. (**D**) The calculated relative response factors of the XyG oligosaccharides with that of XXXG set to 1. (**E**) Total signal intensity (sum of all six oligosaccharides) as a function of oligosaccharide concentration.

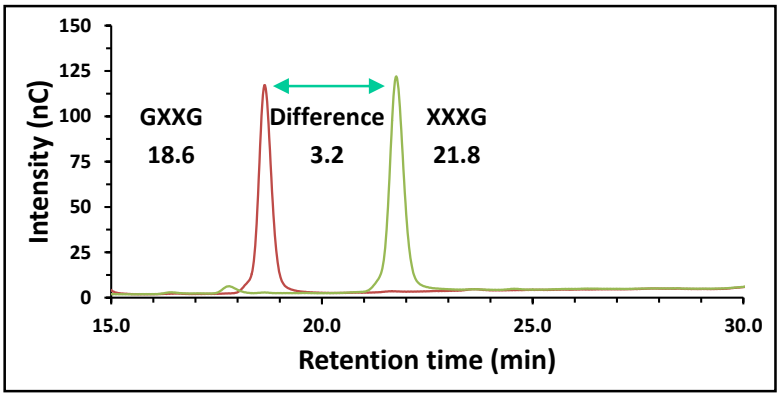
**A** **Cellodextrins**



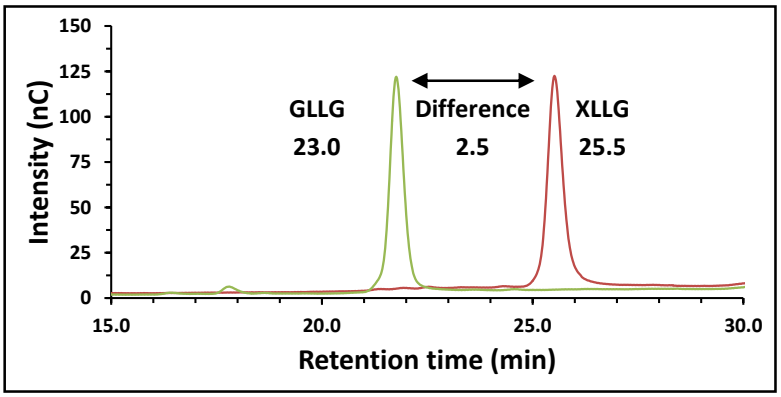

**B** **XyG isoprimeverosyl oligosaccharides**




**C** HPAEC-PAD **Overlaid GXXG and XXXG chromatograms with retention times indicated**

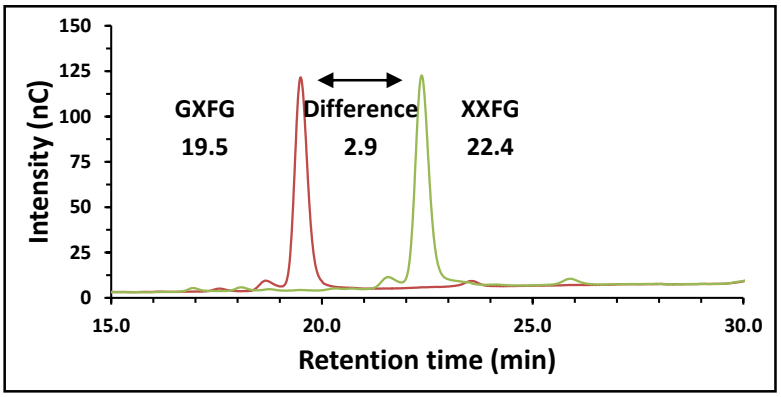



HPAEC-PAD **Overlaid GLLG and XLLG chromatograms with retention times indicated**



HPAEC-PAD

**Overlaid GXFG and XXFG chromatograms with retention times indicated**




**D** **Retention time contributions**

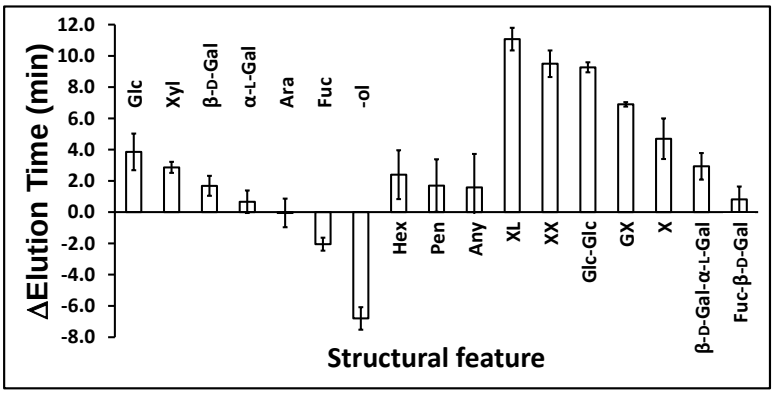



**Figure 2-6**. HPAEC-PAD analysis of oligosaccharide retention times. (**A**) Correlation of the elution time and the number of D-Glc*p* (G) units in cellodextrins (GGG, GGGG, GGGGG, GGGGGG, GGGGGGG). (**B**) Correlation of the elution time and the number of isoprimeverosyl (X) units in isoprmeverosyl XyG oligosaccharide series (XG, XXG, XXXG, XXXXG, and XXXXXG). The correlation coefficients were calculated for all eight possible combinations of signals taking into account the two signals from XXXXG and four from XXXXXG. The best fit is shown on the graph. (**C**) Comparison of the retention times of three oligosaccharide pairs that differ in the presence of Xyl*p* residue on the non-reducing end backbone Glc*p* residue. (**D**) The contributions of various structural features in cellodextrin and XyG oligosaccharides. The values were obtained by

comparing all oligosaccharide pairs that differ in structure by a specific feature. Average and standard deviation was calculated for each structural feature.

**Table 2-1.** The chemical structures of XyG side-chains relevant to this work.

| Symbol | Chemical Structure[1] | Found in | Reference |
|---|---|---|---|
| **X** | α-D-Xyl*p*-(1→6) | All XyGs | (Fry et al, 1993) |
| **L** | β-D-Gal*p*-(1→2)-α-D-Xyl*p*-(1→6) | Almost all XyGs | (Fry et al, 1993) |
| **D** | α-L-Ara*p*-(1→2)-α-D-Xyl*p*-(1→6) | *S. kraussiana*, *E. hyemale* | (Peña et al, 2008) |
| **Y** | β-D-Gal*p*A-(1→2)-α-D-Xyl*p*-(1→6) | *Arabidopsis thaliana* root hairs | (Peña et al, 2008) |
| **F** | α-L-Fuc*p*-(1→2)-β-D-Gal*p*-(1→2)-α-D-Xyl*p*-(1→6) | Majority of XyGs | (Fry et al, 1993) |
| **E** | α-L-Fuc*p*-(1→2)-α-L-Ara*p*-(1→2)-α-D-Xyl*p*-(1→6) | *S. kraussiana*, *E. hyemale* | (Peña et al, 2008) |
| **J** | α-L-Gal*p*-(1→2)-β-D-Gal*p*-(1→2)-α-D-Xyl*p*-(1→6) | Jojoba (*Simmondsia chinensis*) seed | (Hantus et al, 1997) |
| **Z** | α-L-Fuc*p*-(1→2)-β-D-Gal*p*A-(1→2)-α-D-Xyl*p*-(1→6) | *Arabidopsis thaliana* root hairs | (Peña et al, 2012) |

[1]All listed side-chains are linked to a backbone β-D-Glc*p* residue.

**Table 2-2.** Summary of the purified oligosaccharides.

| Backbone length | Structure(s) present[1] | | | | | | |
|---|---|---|---|---|---|---|---|
| | **G** | **G, X** | **G, X, L** | **G, X, <u>D</u>** | **G, X, <u>Y</u>** | **G, X, L, <u>F</u>** | **G, X, L, <u>J</u>** |
| **1** | | X | | | | | |
| **2** | | XG | LG | | YG | FG | JG |
| | | XX | | | | | |
| **3** | GGG | XXG | | | | | |
| **4** | GGGG | GXXG | XLXG | XDXG | XYYG | XXFG | XXJG |
| | | XXXG | XXLG | XDDG | | XXFG | XLJG |
| | | XXXGol | GLLG | | | XXFGol | |
| | | | XLLG | | | XLFG | |
| | | | XLLGol | | | XFFG | |
| **5** | GGGGG | <u>X</u>XXXG[2] | | | | | |
| **6** | GGGGGG | <u>XX</u>XXXG[2] | | | | | |
| **7** | GGGGGGG | | | | | | |

[1]The oligosaccharides are grouped according to the diversity and structural complexity of their side-chains. The underlined side-chains represent the characteristic side-chains that are used in the grouping.

[2]These samples contain positional isomers in which the underlined isoprimeverosyl units (X) can be either *O*-3 or *O*-4 linked.

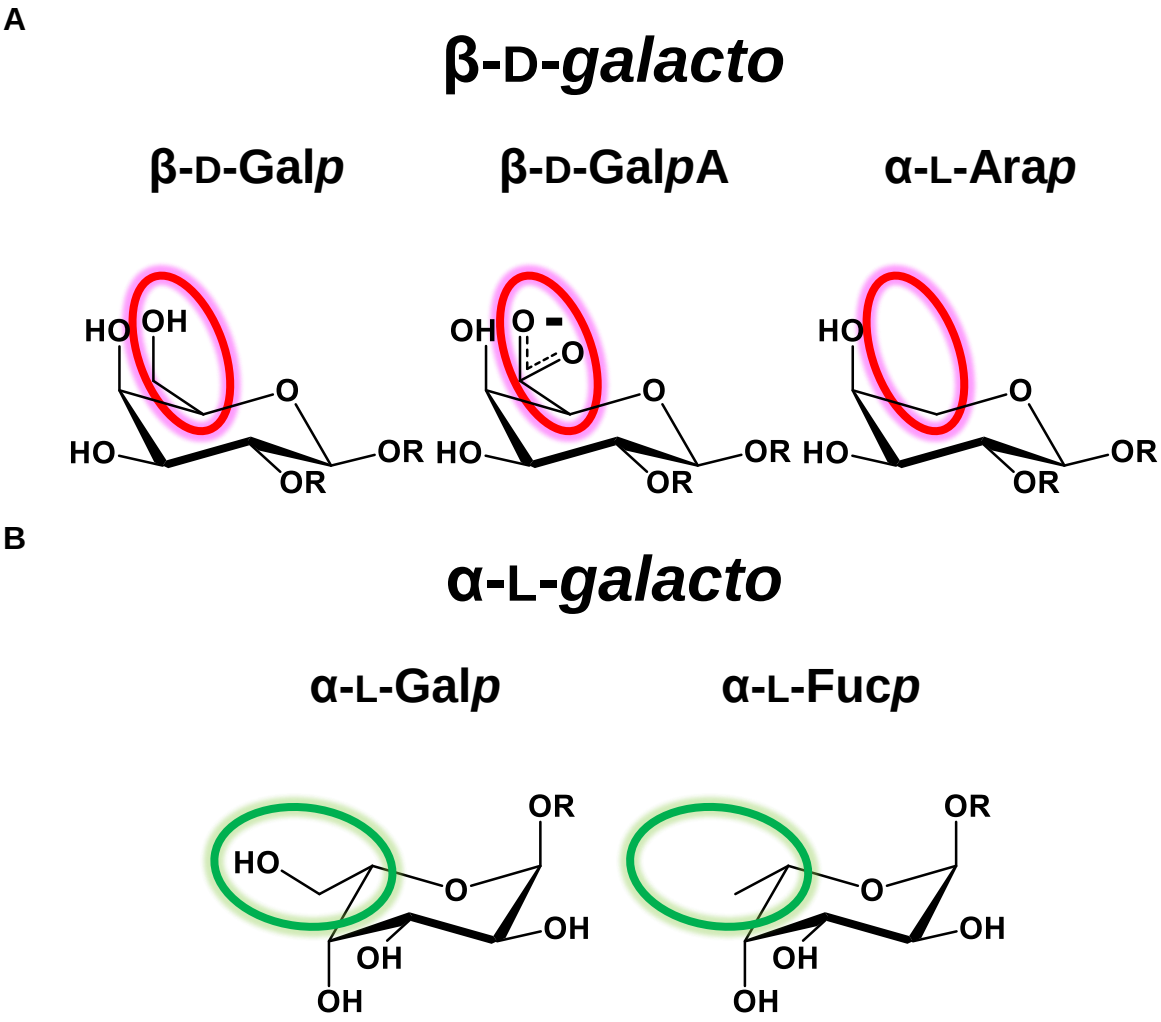



**Figure 2-S1.** Stereochemical representations of some of the glycosyl residues found in XyGs. (**A**) The β-D-Gal*p*, β-D-Gal*p*A and α-L-Ara*p* residues, that are linked to the *O*-2 position of Xyl*p* residue, all have the β-D-*galacto* configuration. (**B**) The α-L-Gal*p* (that can be linked to the *O*-2 position of β-D-Gal*p* residue) and α-L-Fuc*p* (that can be linked to the *O*-2 position of β-D-Gal*p*, β-D-Gal*p*A or α-L-Ara*p* residues) residues both have the α-L-*galacto* configuration.

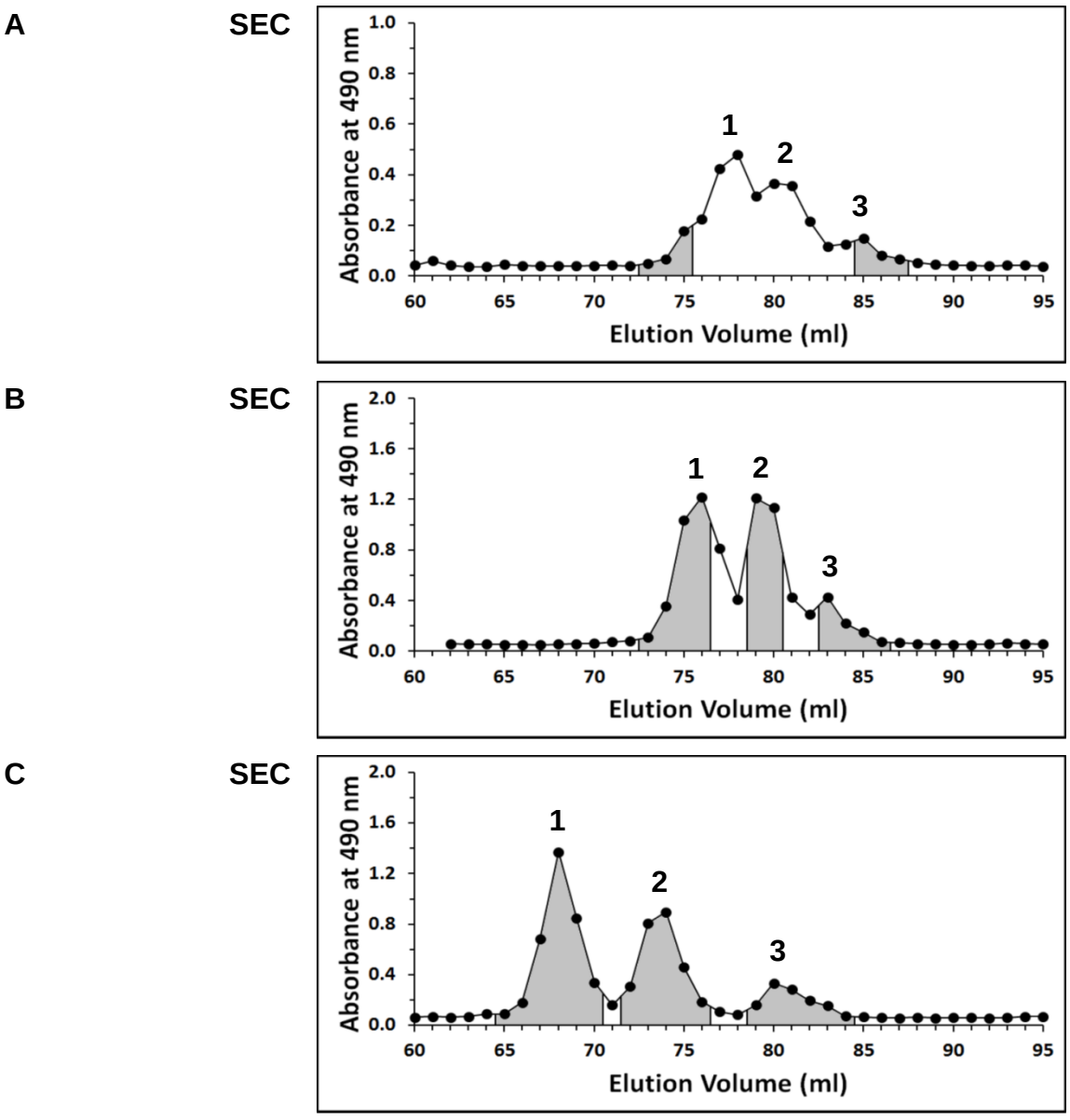



**Figure 2-S2.** Comparison of the efficiencies of eluents in the separation of tamarind XyG S1 oligosaccharides. 10 mg of tamarind XyG S1 oligosaccharides generated by XEG hydrolysis of tamarind XyG polysaccharide were loaded per experiment. (**A**) water, (**B**) 10 mM acetic acid, and (**C**) 10 mM boric acid, 10 mM phytic acid, pH 10 (adjusted with sodium hydroxide). The fractions were analyzed by MALDI-TOF MS (data not shown) and the oligosaccharide preparations were

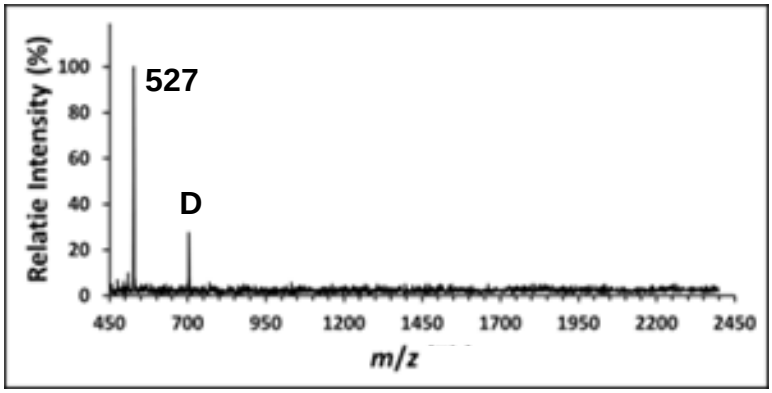
pooled as indicated by the chromatogram shading. Note that XLLG (Peak 1) and XXXG (Peak 3) oligosaccharides can be obtained in their pure forms, whereas XLXG and XXLG (Peak 2) co-elute. The oligosaccharides elute in the expected order with XLLG (9 glycosyl residues) eluting first, and XXXG (glycosyl residues) eluting last, and XLXG and XXLG (8 glycosyl residues) co-eluting in the middle.

**A** **Chemical Structure**

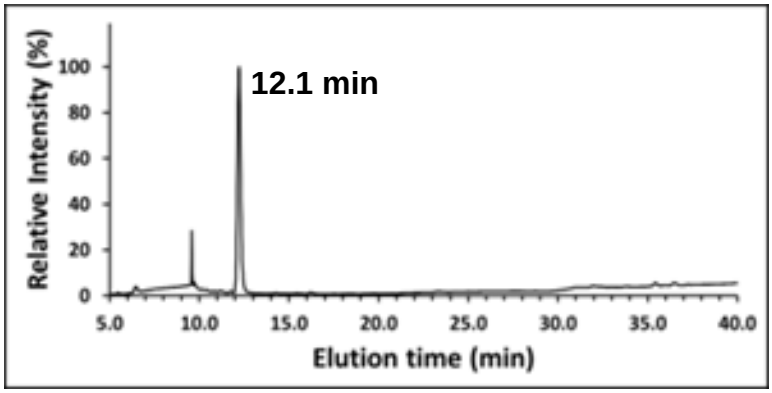

**MALDI-TOF MS**

Calc. $[M + Na]^+$

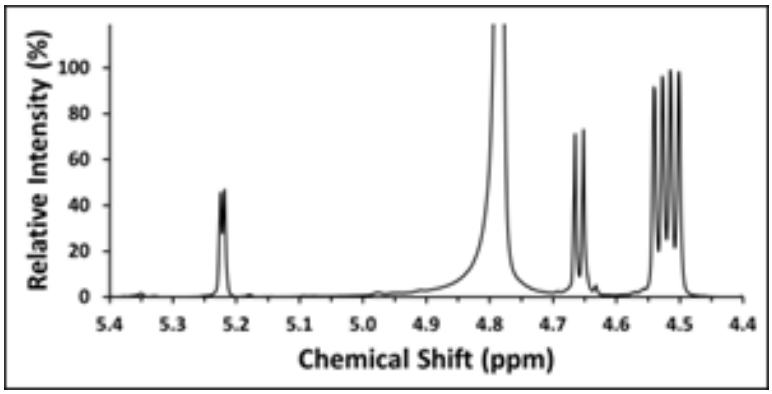
*m/z* = 527.2



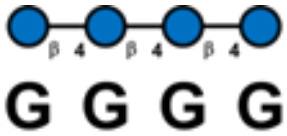

**HPAEC-PAD**




**$^1$H NMR**



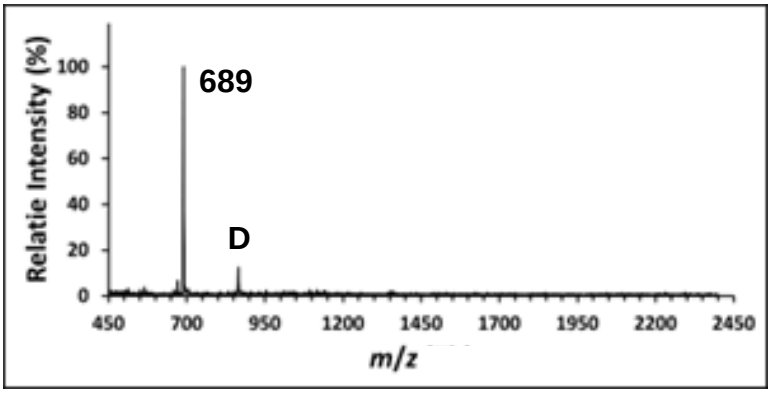
**B** **Chemical Structure**



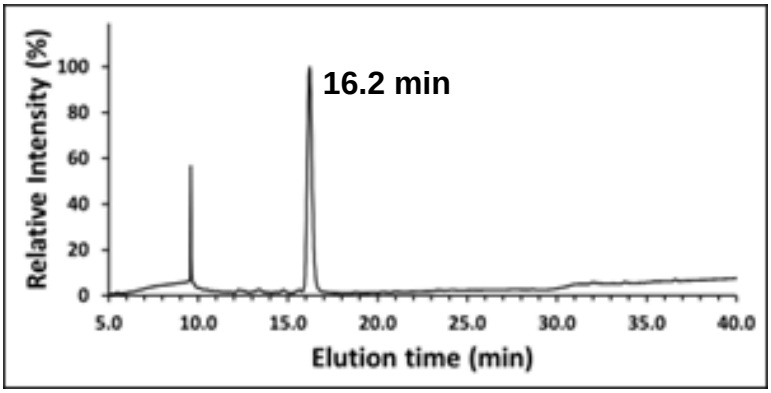

**MALDI-TOF MS**

Calc. $[M + Na]^+$

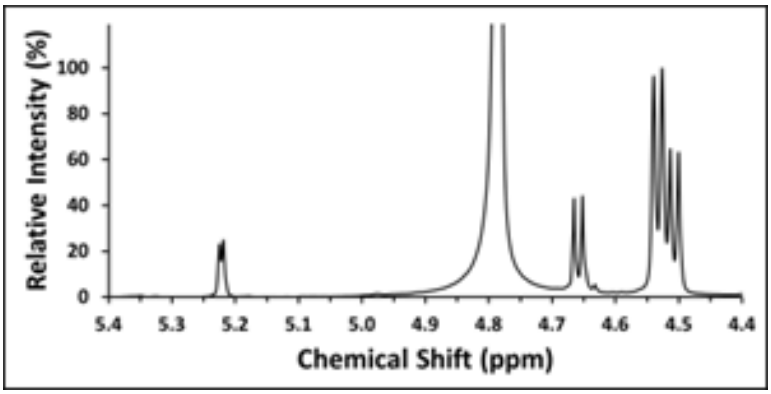
*m*/*z* = 689.2




**HPAEC-PAD**




**$^1$H NMR**

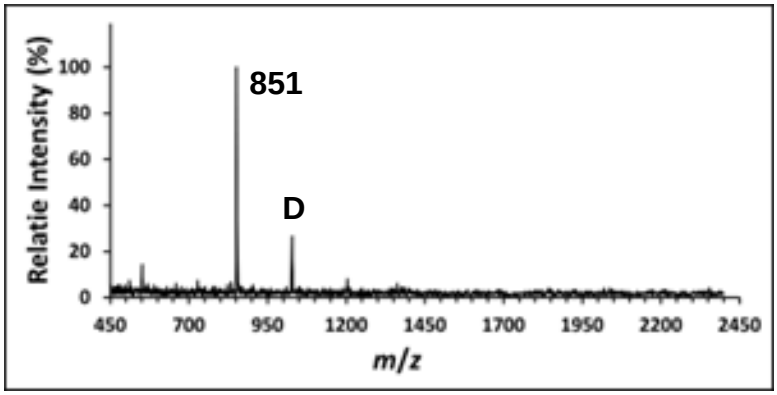


**C** **Chemical Structure**

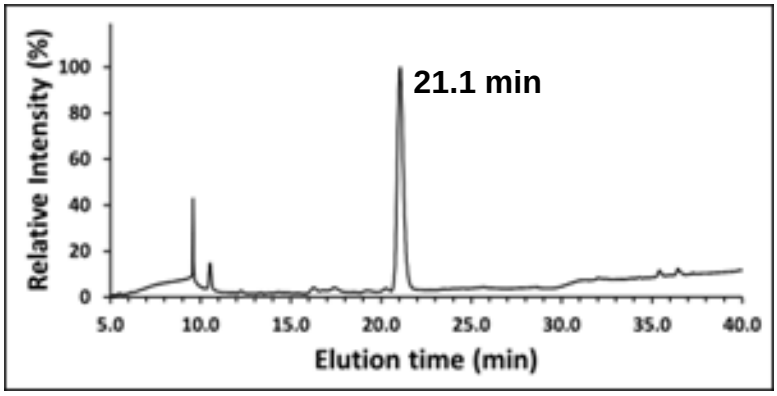

**MALDI-TOF MS**

Calc. $[M + Na]^+$

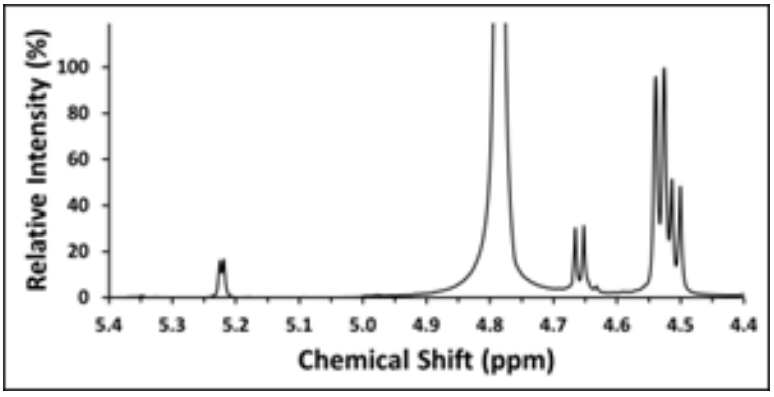
*m*/*z* = 851.3




**HPAEC-PAD**




**$^1$H NMR**



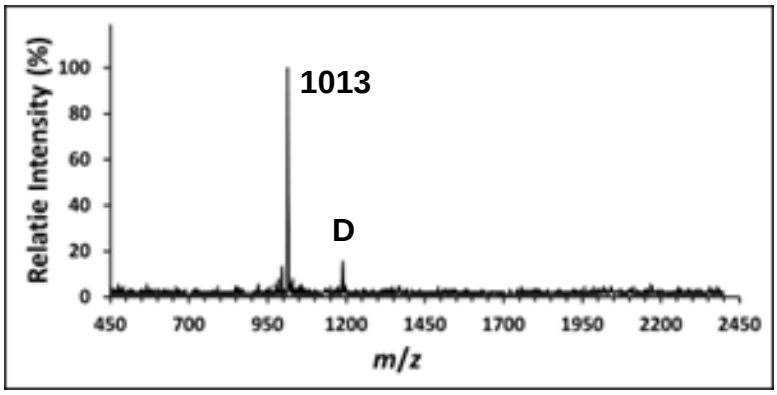
**D**

**Chemical Structure**

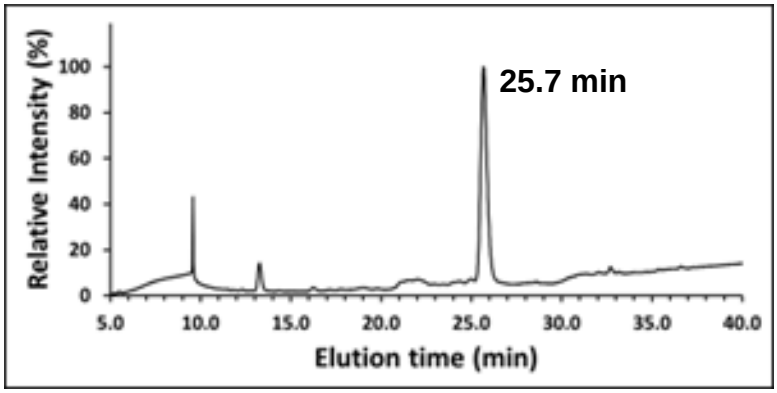

**MALDI-TOF MS**

Calc. $[M + Na]^+$

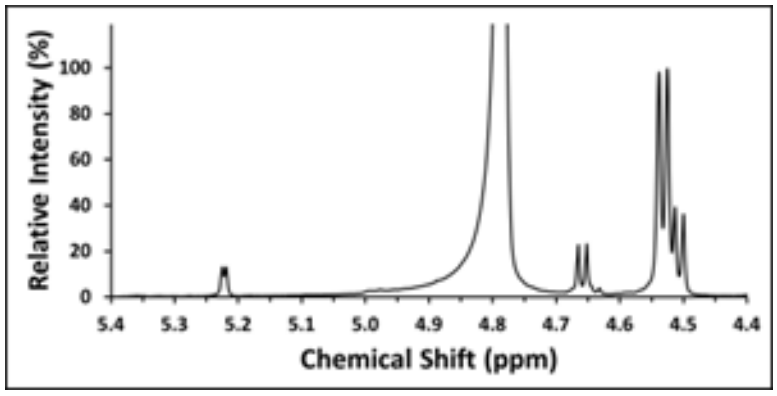
*m*/*z* = 1013.3




**HPAEC-PAD**




**$^1$H NMR**

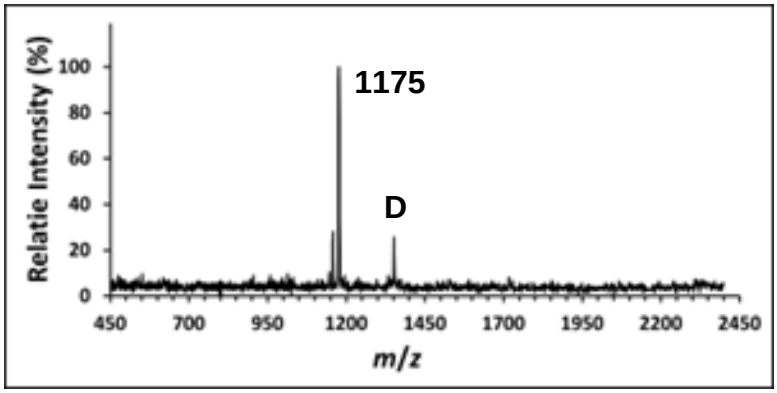


**E** **Chemical Structure**

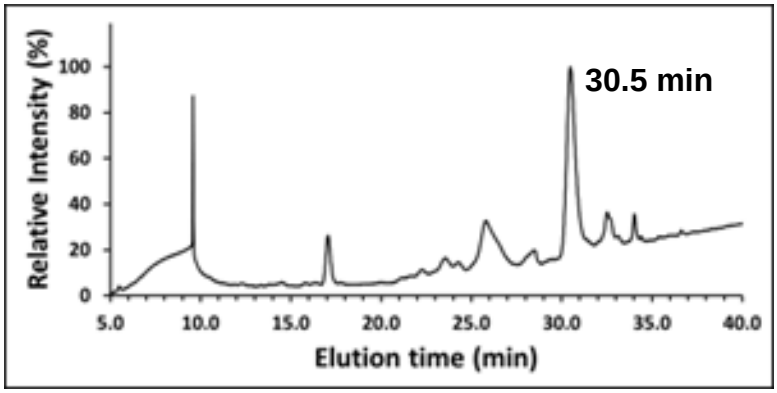

**MALDI-TOF MS**

Calc. $[M + Na]^+$

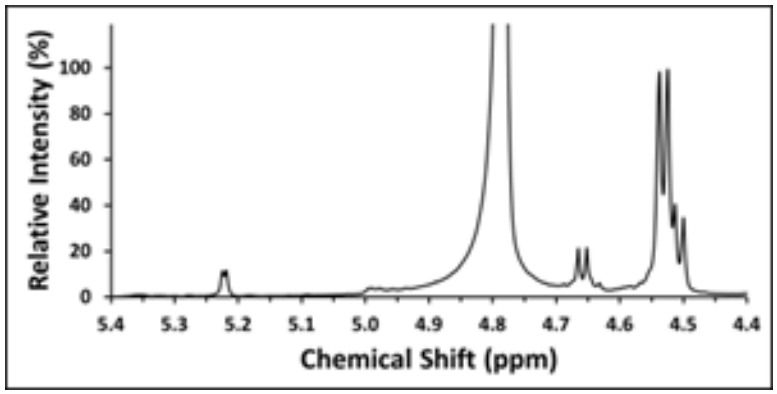
*m*/*z* = 1175.4




**HPAEC-PAD**




**$^1$H NMR**

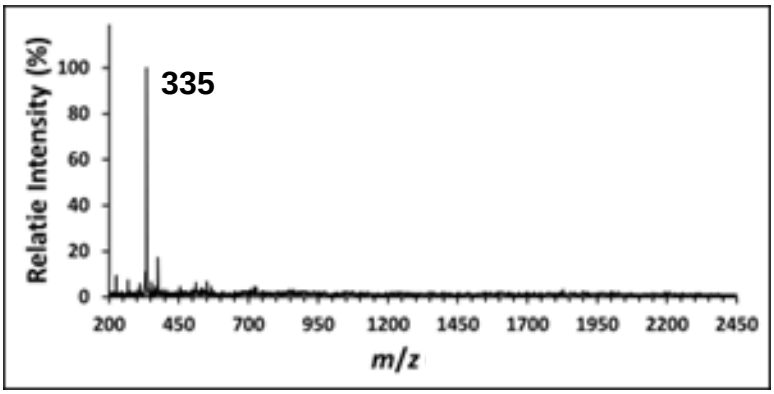


**F** **Chemical Structure**

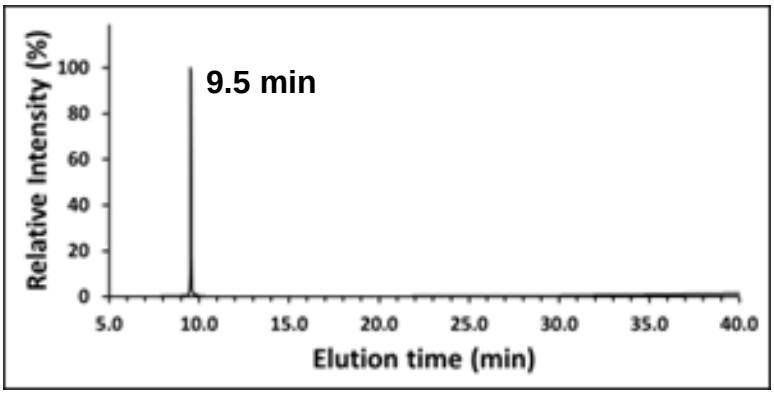

**MALDI-TOF MS**

Calc. $[M + Na]^+$

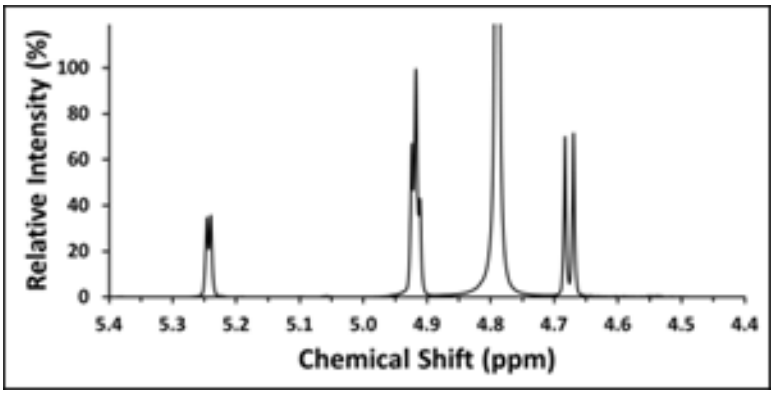
*m*/*z* = 335.1



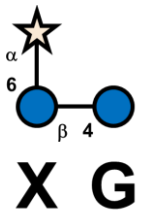

**HPAEC-PAD**




**$^1$H NMR**



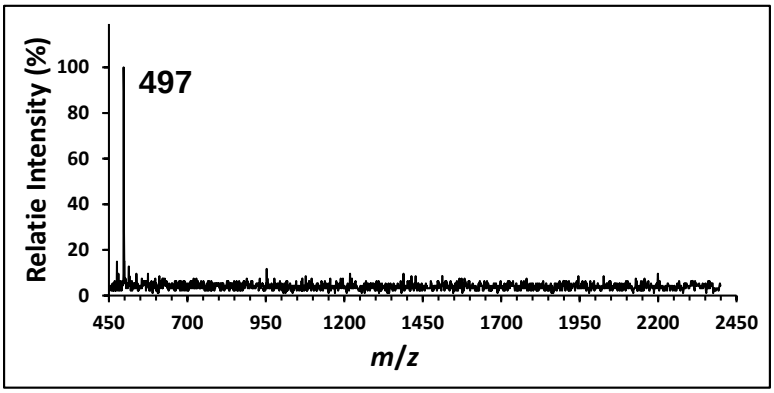
**G** **Chemical Structure**



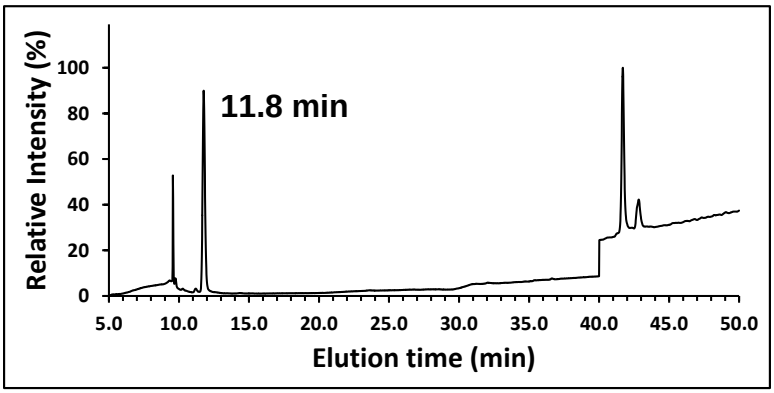

**MALDI-TOF MS**

Calc. $[M + Na]^+$

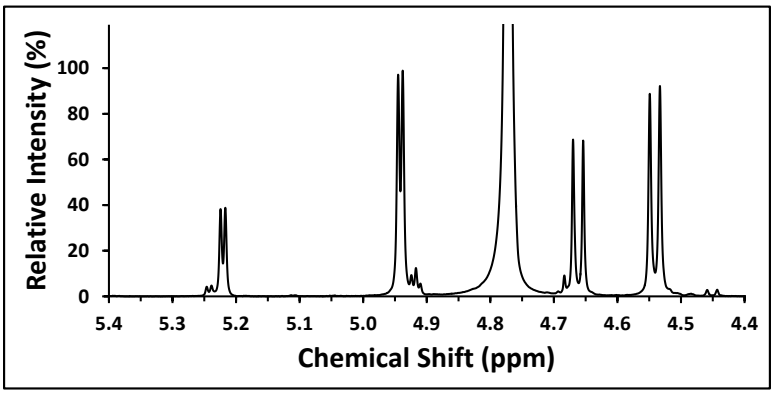
*m/z* = 497.2



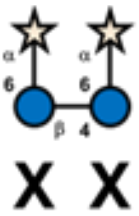

**HPAEC-PAD**




**$^1$H NMR**



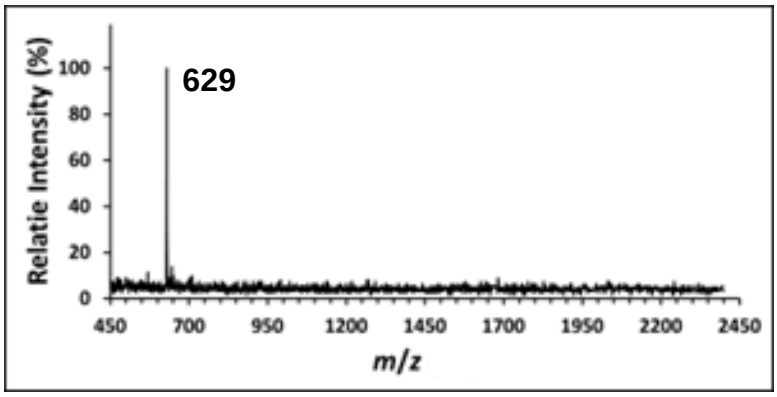
**H** **Chemical Structure**



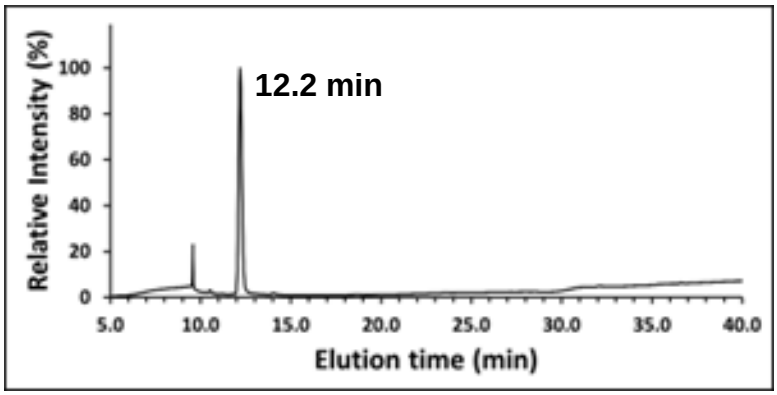

**MALDI-TOF MS**

Calc. $[M + Na]^+$

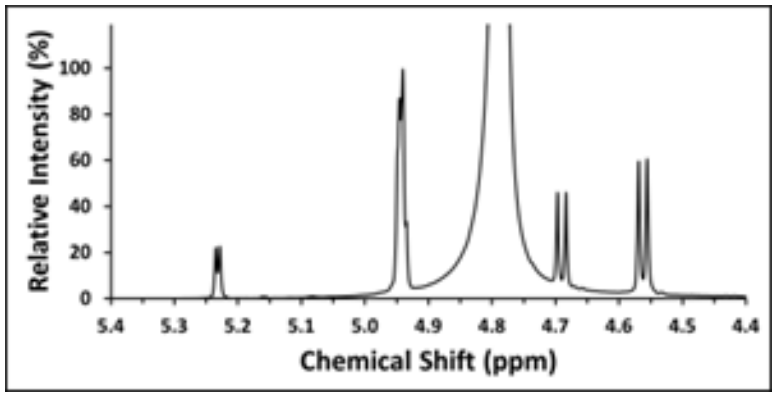
*m/z* = 629.2



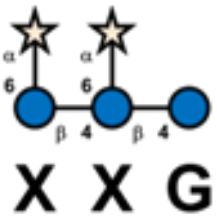

**HPAEC-PAD**




**$^1$H NMR**



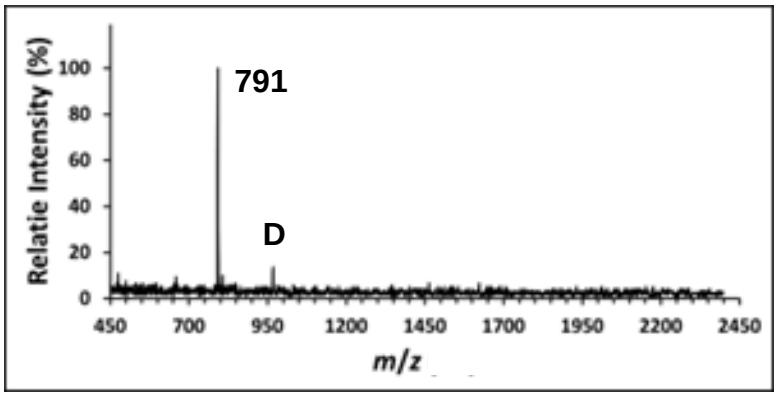
**I** **Chemical Structure**



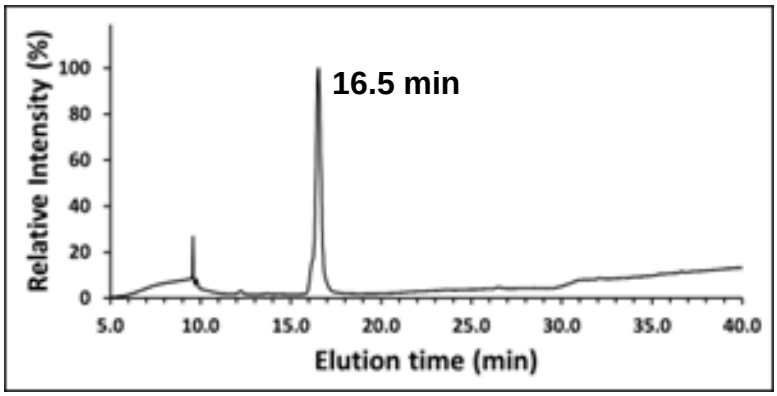

**MALDI-TOF MS**

Calc. $[M + Na]^+$

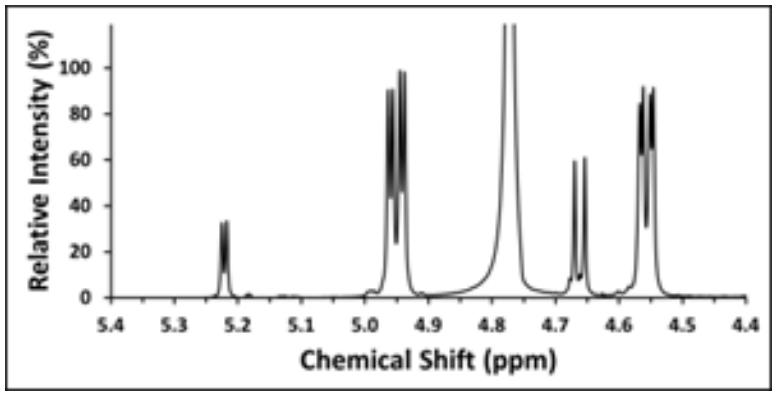
*m/z* = 791.2



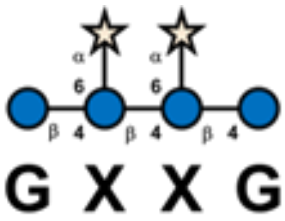

**HPAEC-PAD**




**$^1$H NMR**



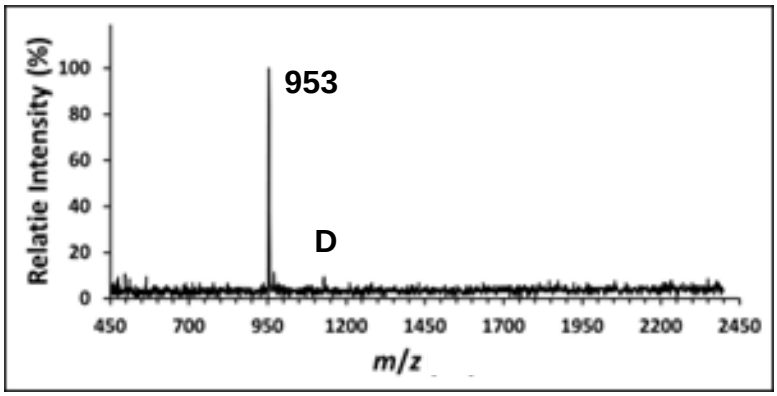
**J** **Chemical Structure**



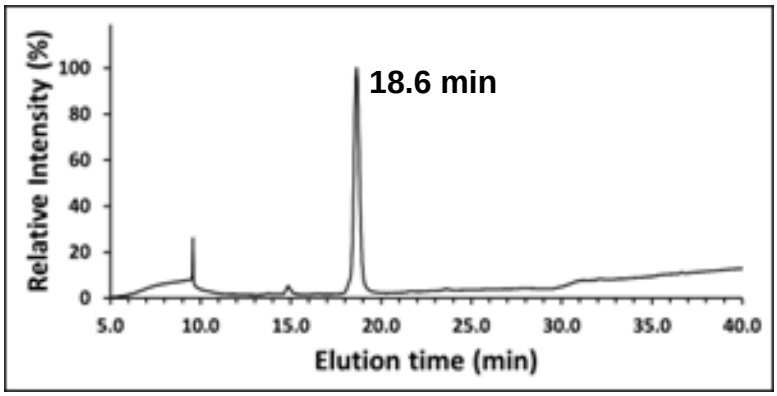

**MALDI-TOF MS**

Calc. $[M + Na]^+$

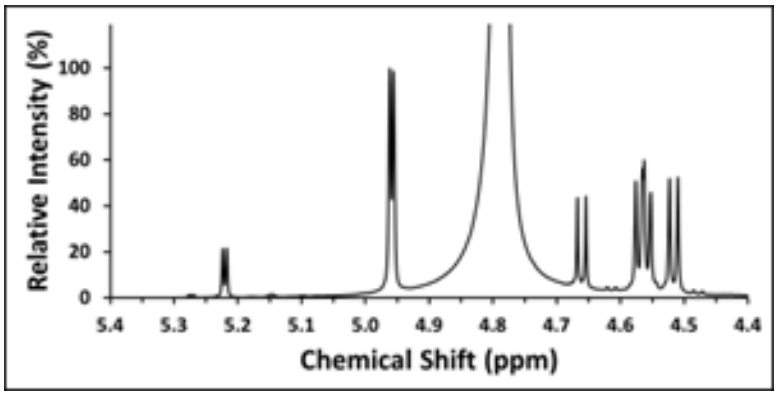
*m*/*z* = 953.3



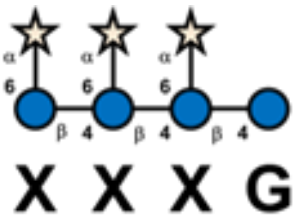

**HPAEC-PAD**




**$^1$H NMR**



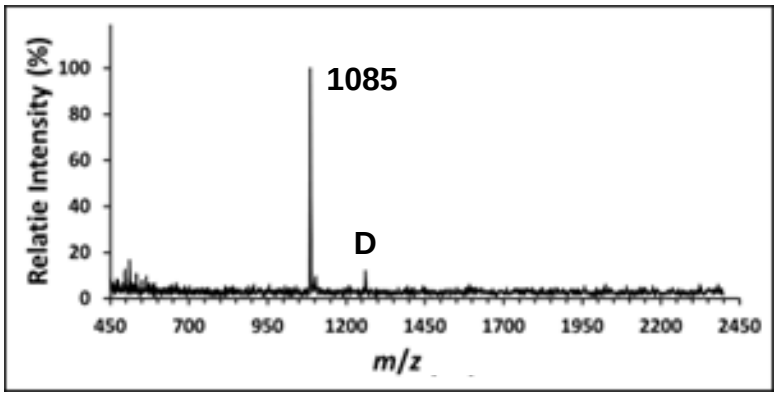
**K** **Chemical Structure**



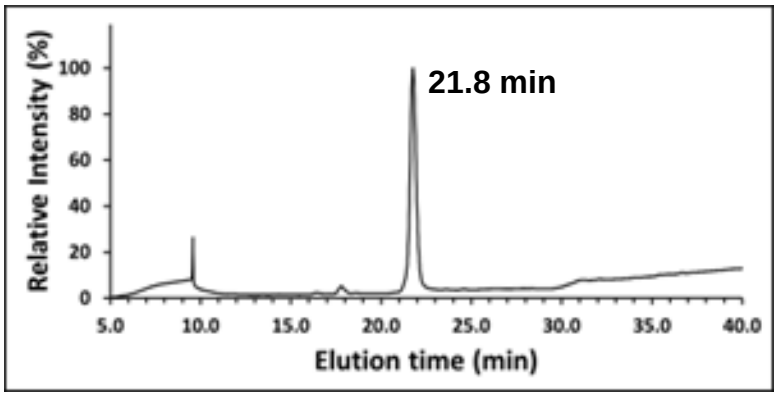

**MALDI-TOF MS**

Calc. $[M + Na]^+$

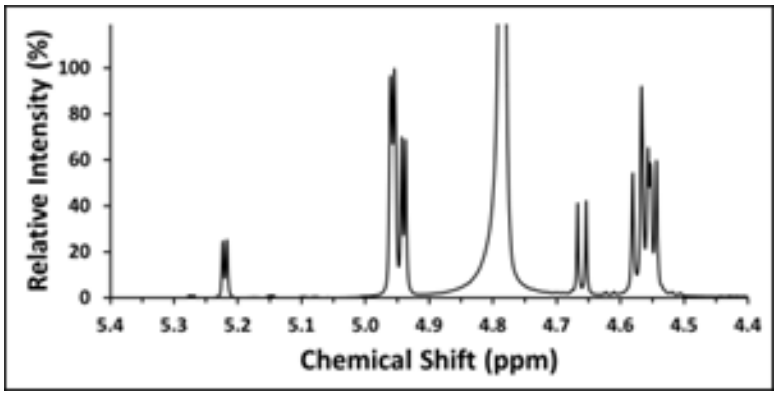
*m*/*z* = 1085.3



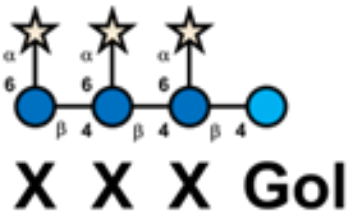

**HPAEC-PAD**




**$^1$H NMR**



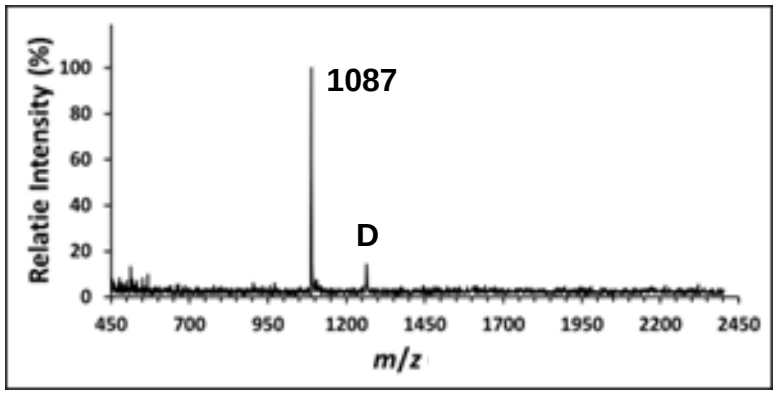
**L** **Chemical Structure**



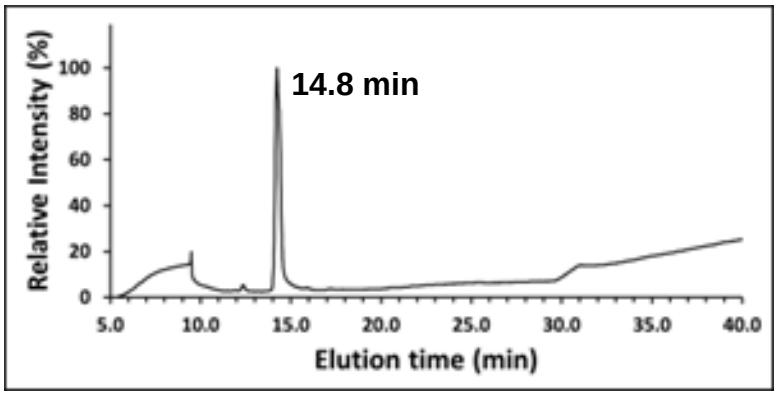

**MALDI-TOF MS**

Calc. $[M + Na]^+$

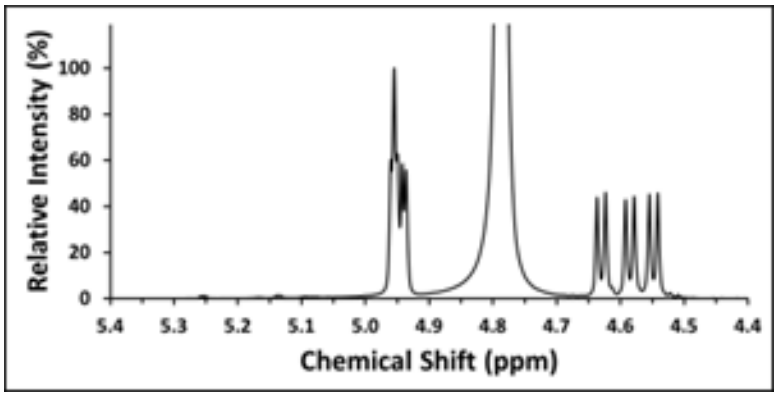
$m/z$ = 1087.4



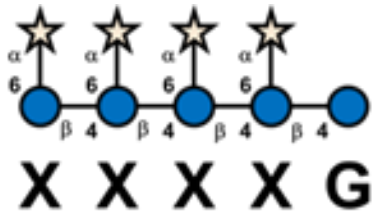

**HPAEC-PAD**




**$^1$H NMR**



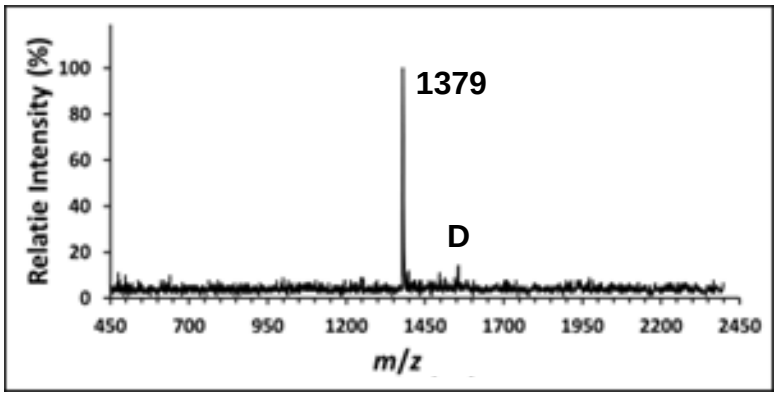
**M** **Chemical Structure**




**MALDI-TOF MS**

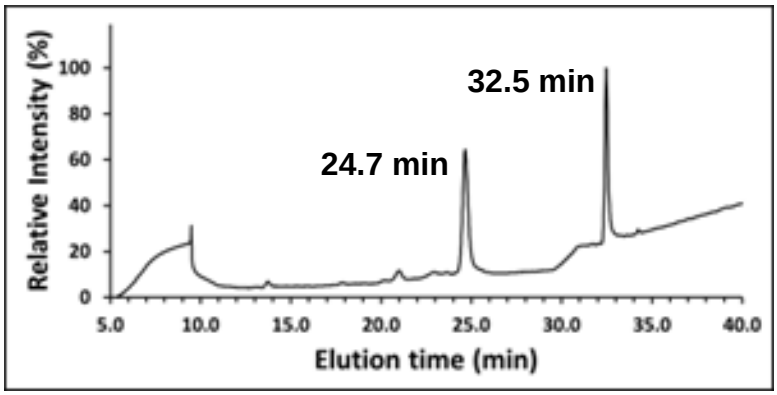
Calc. $[M + Na]^+$

*m*/*z* = 1379.4

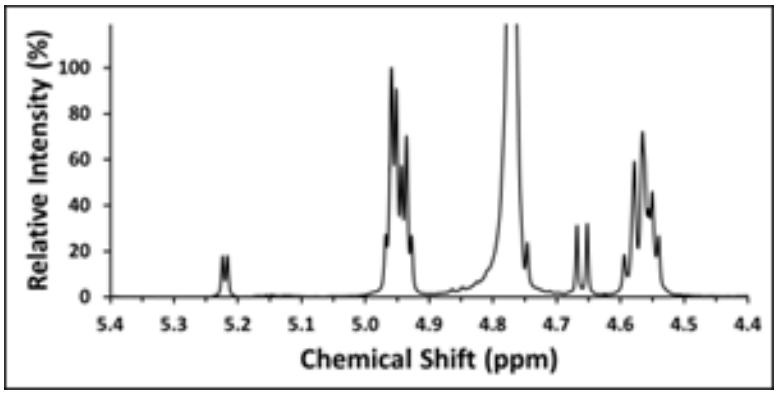



**HPAEC-PAD**

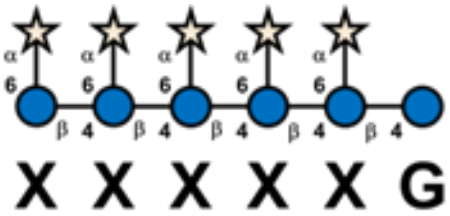
See legend for discussion of the two peaks




**$^1$H NMR**



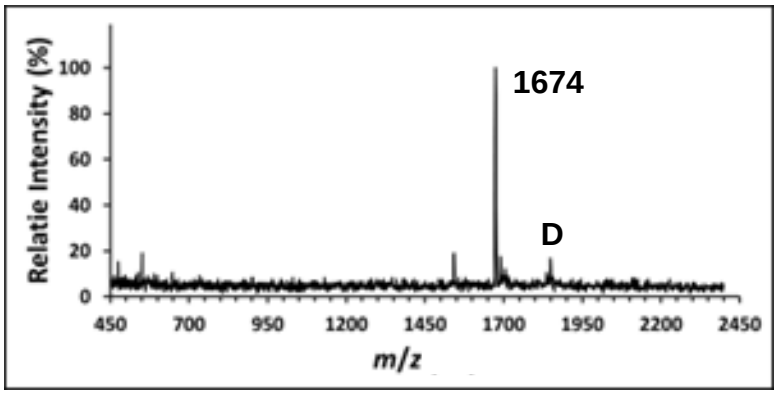
**N** **Chemical Structure**




**MALDI-TOF MS**

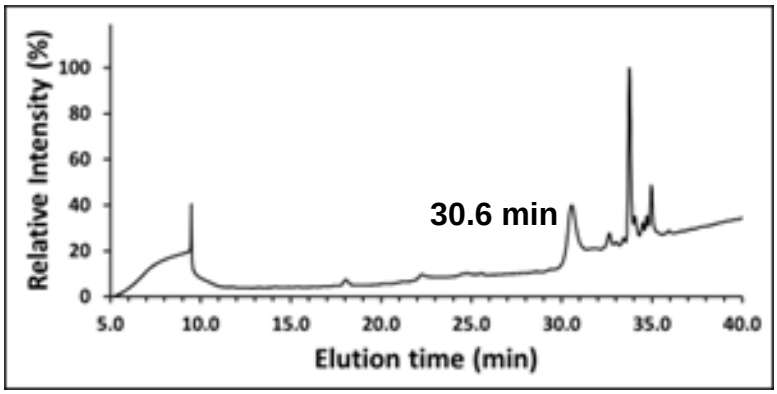
Calc. $[M + Na]^+$

*m*/*z* = 1673.5

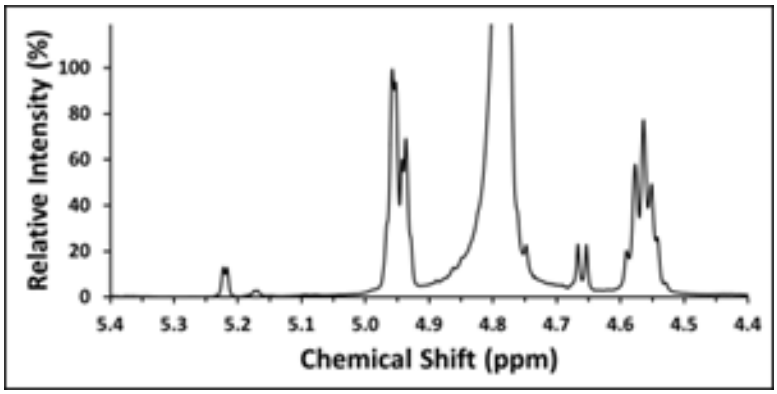



**HPAEC-PAD**

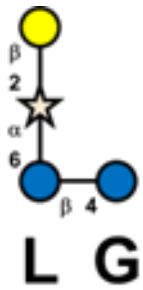
See legend for discussion of the multiple peaks




**$^1$H NMR**



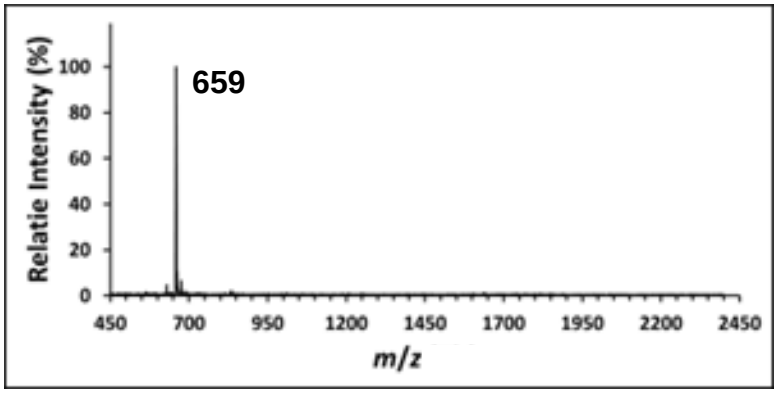
**O** **Chemical Structure**



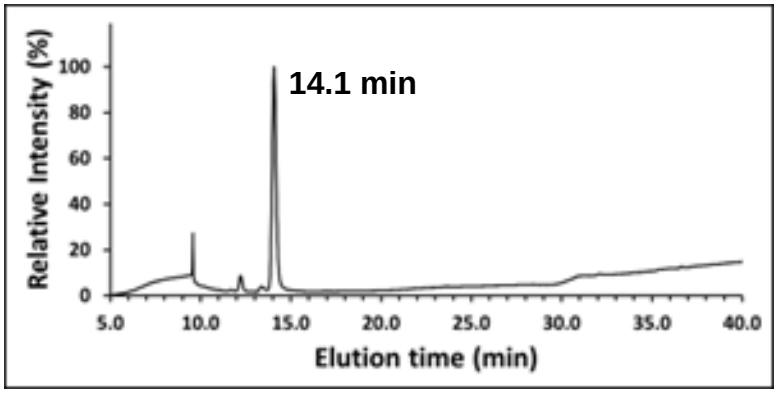

**MALDI-TOF MS**

Calc. $[M + Na]^+$

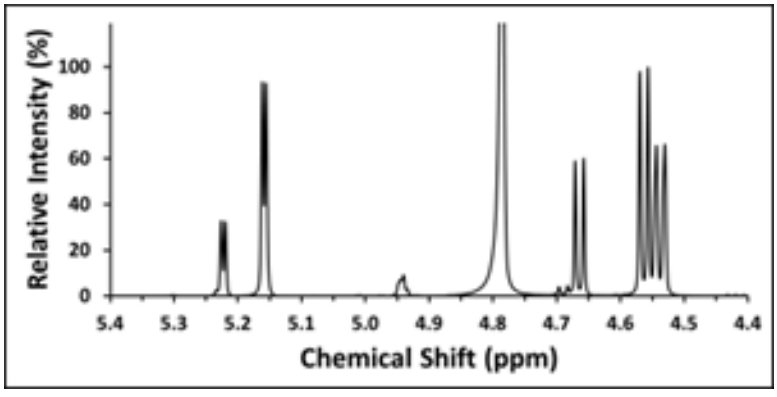
*m/z* = 659.2



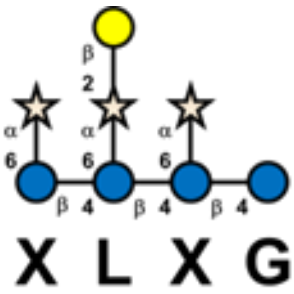

**HPAEC-PAD**




**$^1$H NMR**



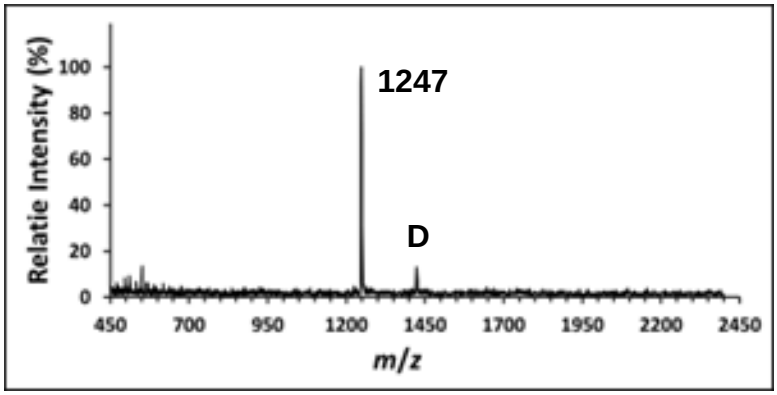
**P** **Chemical Structure**



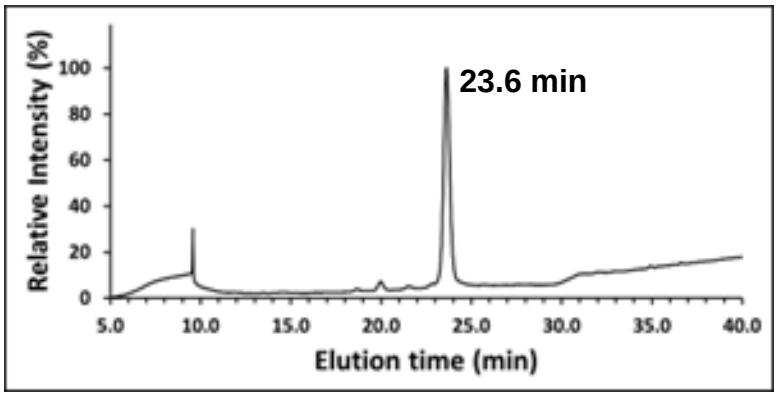

**MALDI-TOF MS**

Calc. $[M + Na]^+$

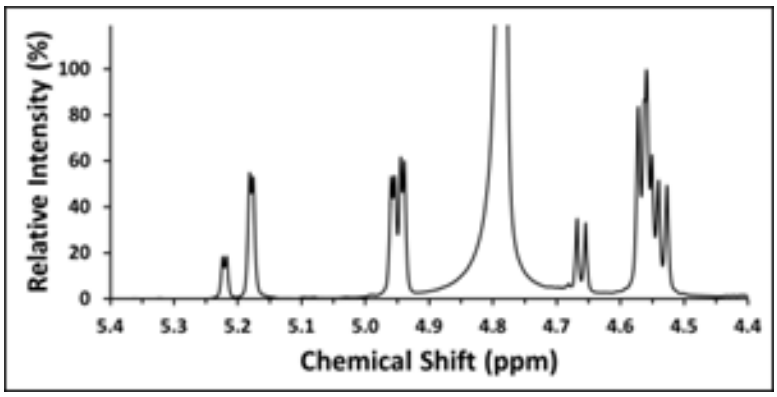
*m*/*z* = 1247.4



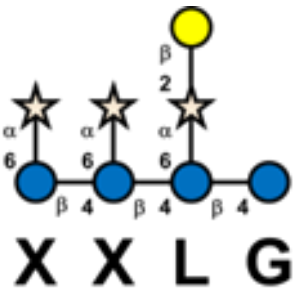

**HPAEC-PAD**




**$^1$H NMR**



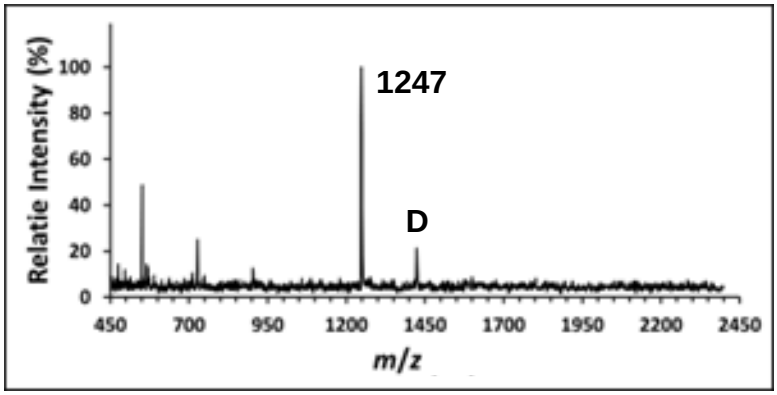
**Q** **Chemical Structure**



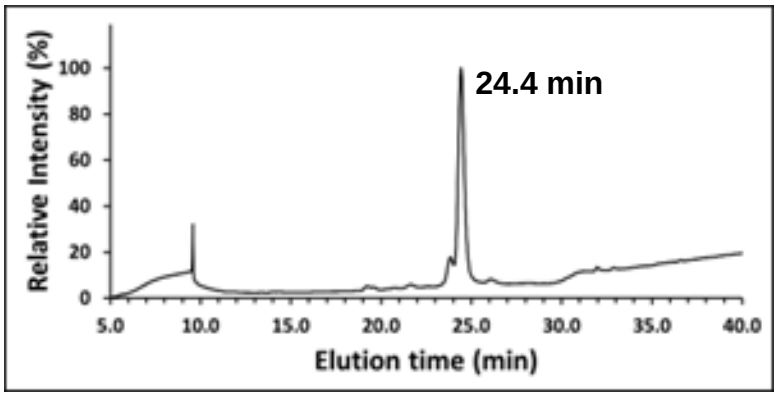

**MALDI-TOF MS**

Calc. $[M + Na]^+$

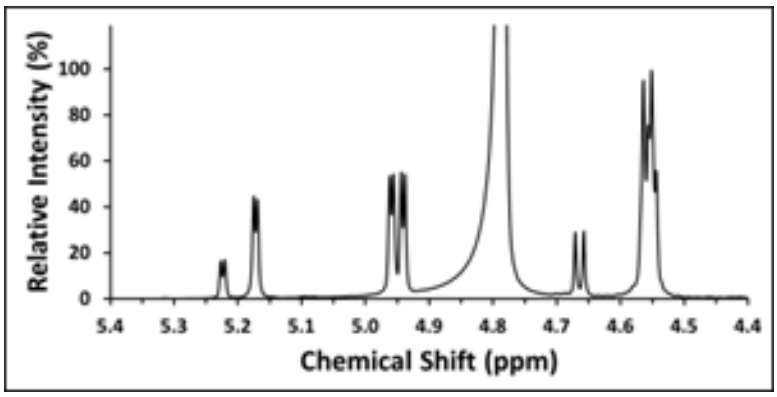
$m/z$ = 1247.4



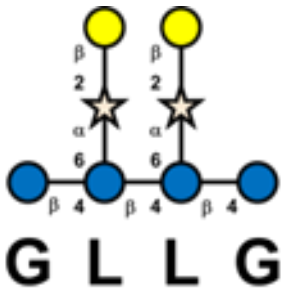

**HPAEC-PAD**




**$^1$H NMR**



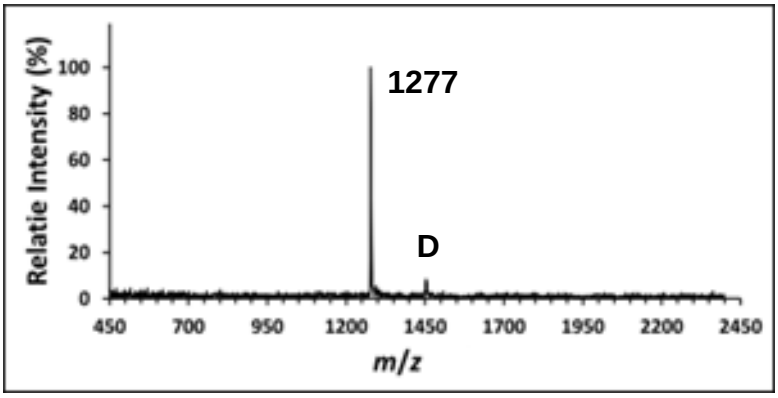
**R** **Chemical Structure**



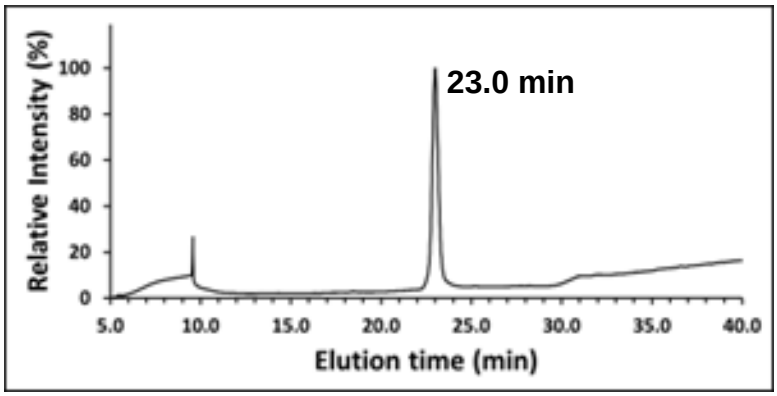

**MALDI-TOF MS**

Calc. $[M + Na]^+$

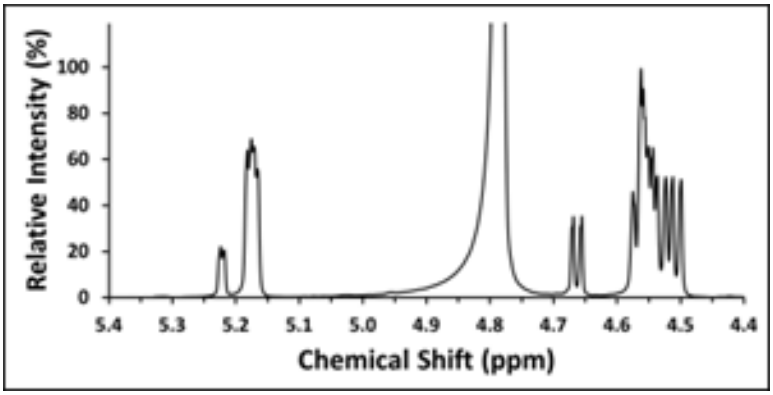
*m*/*z* = 1277.4



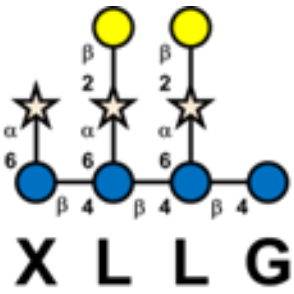

**HPAEC-PAD**




**$^1$H NMR**



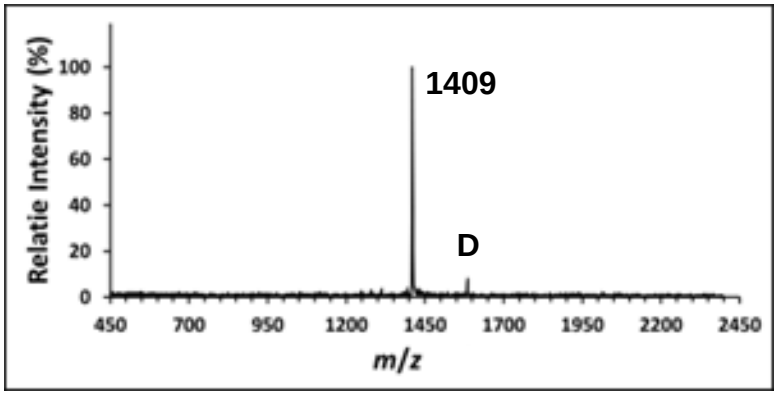
**S** **Chemical Structure**



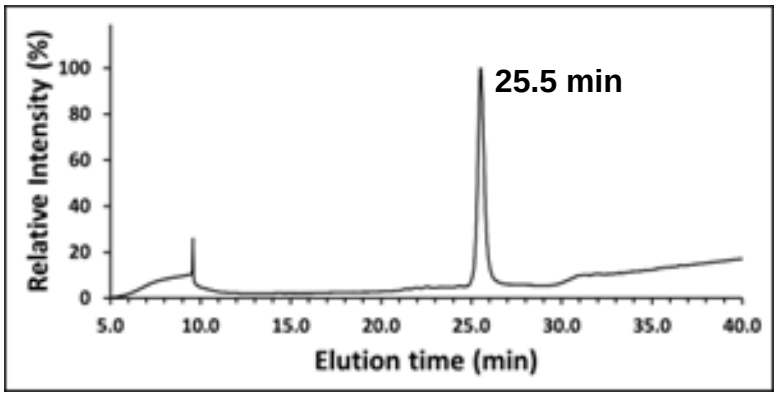

**MALDI-TOF MS**

Calc. $[M + Na]^+$

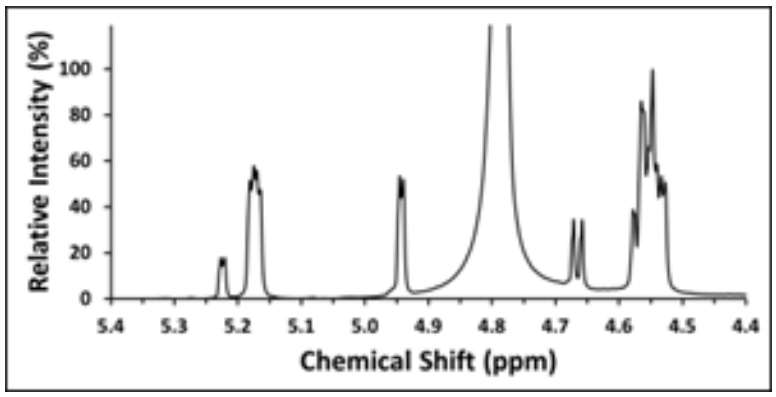
*m*/*z* = 1409.4




**HPAEC-PAD**




**$^1$H NMR**



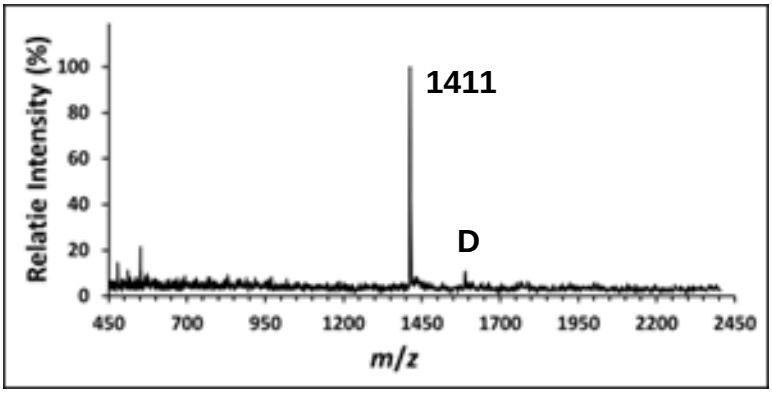
**T**

**Chemical Structure**

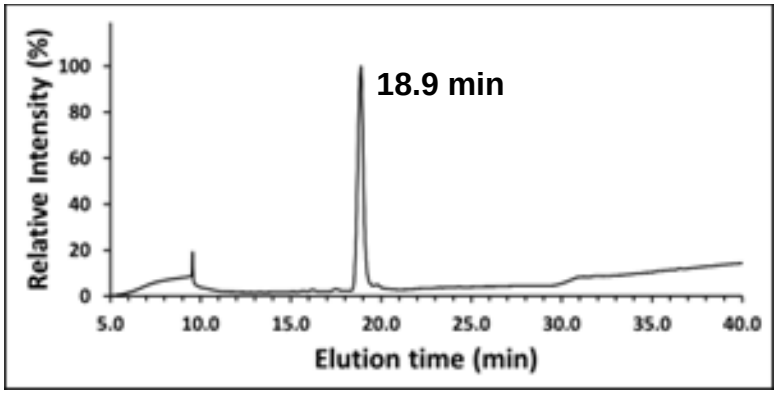

**MALDI-TOF MS**

Calc. $[M + Na]^+$

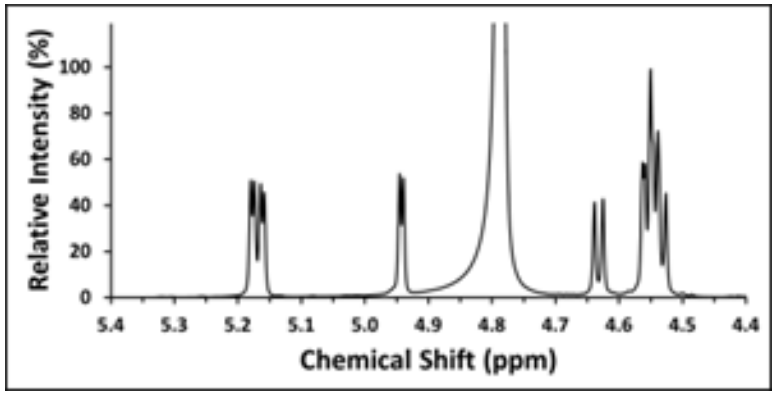
*m/z* = 1411. 5




**HPAEC-PAD**

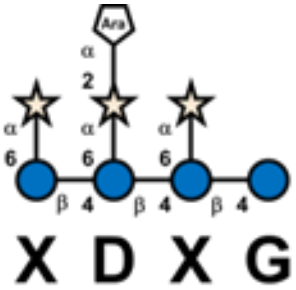



**$^1$H NMR**



**U**

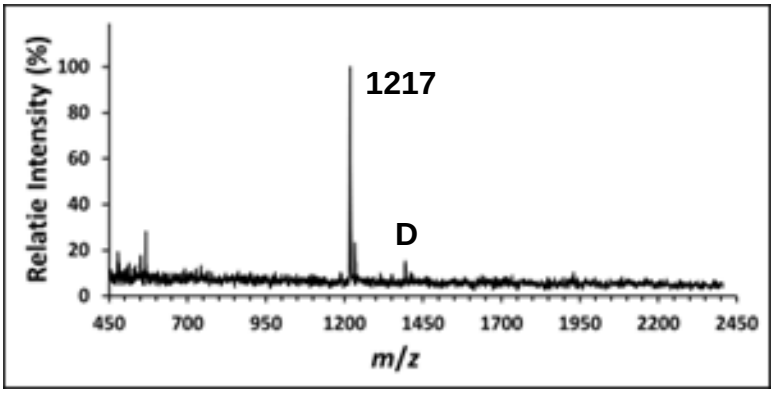
**Chemical Structure**



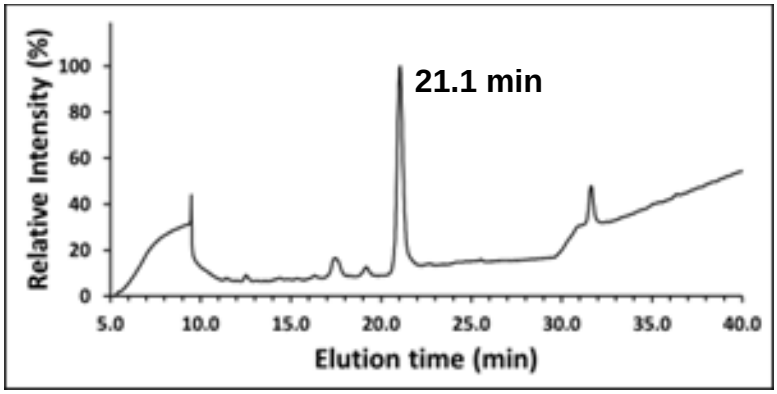

**MALDI-TOF MS**

Calc. $[M + Na]^+$

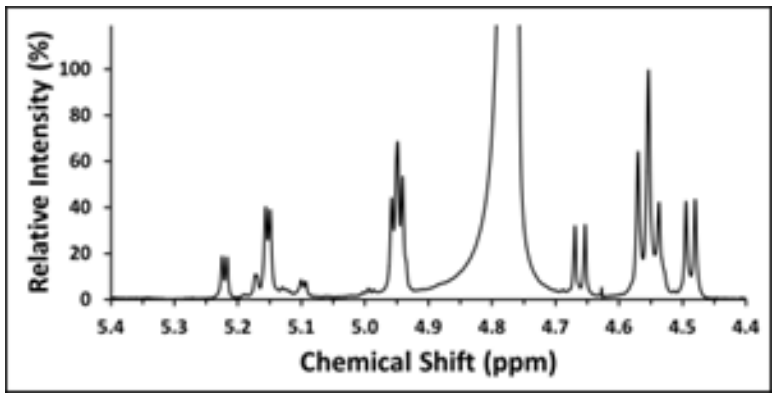
$m/z$ = 1217.4



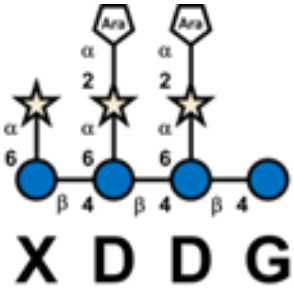

**HPAEC-PAD**




**$^1$H NMR**



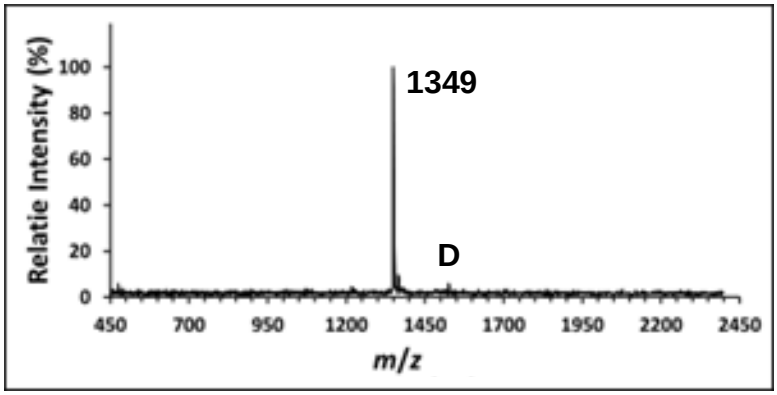
**V** **Chemical Structure**



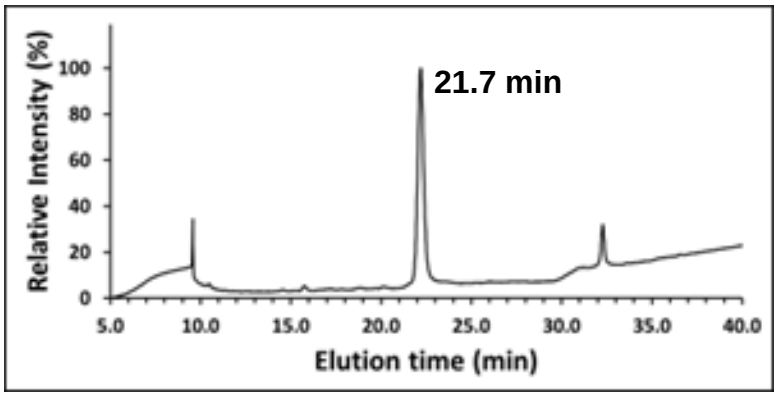

**MALDI-TOF MS**

Calc. $[M + Na]^+$

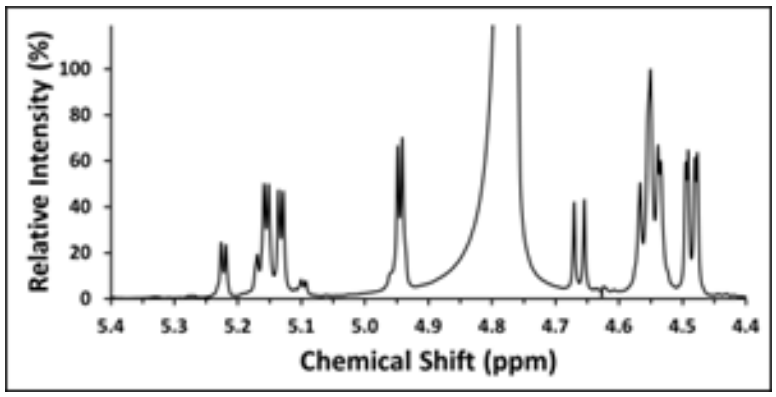
$m/z$ = 1349.4



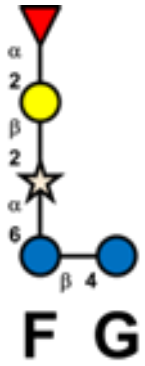

**HPAEC-PAD**




**$^1$H NMR**



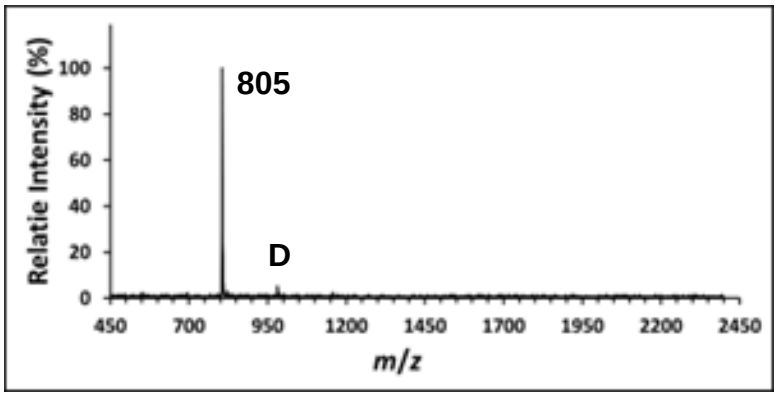
**W** **Chemical Structure**



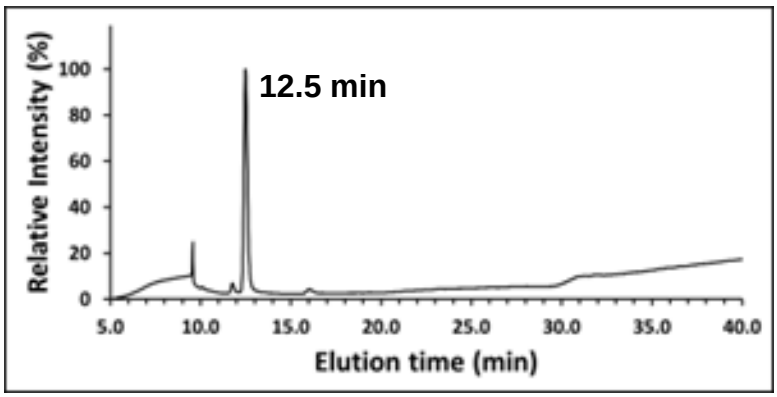

**MALDI-TOF MS**

Calc. $[M + Na]^+$

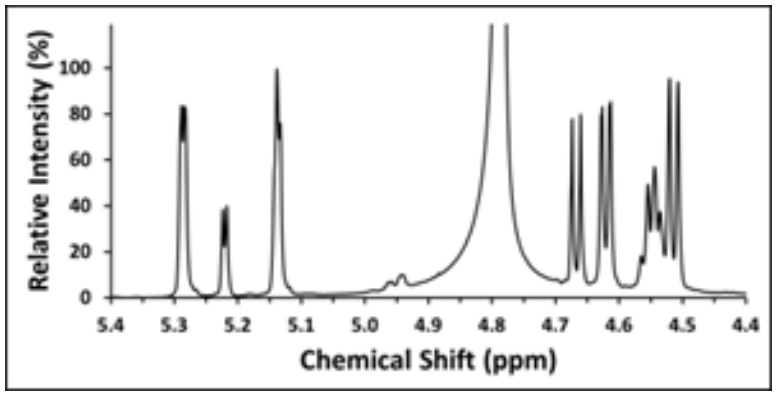
*m/z* = 805.3



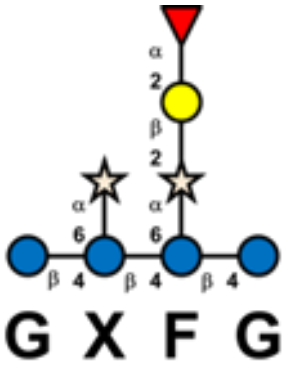

**HPAEC-PAD**




**$^1$H NMR**



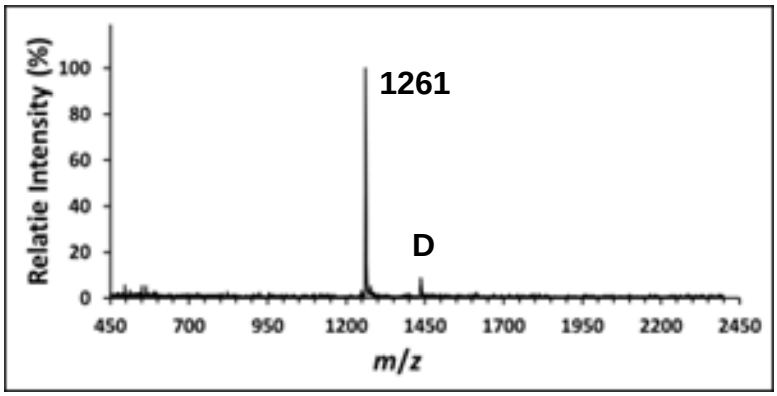
**X** **Chemical Structure**



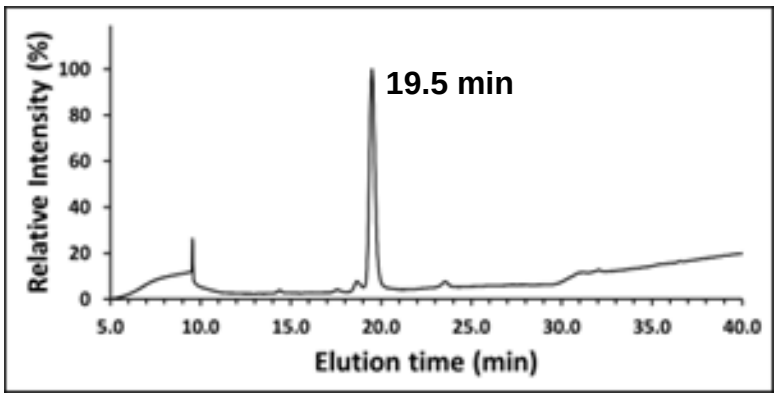

**MALDI-TOF MS**

Calc. $[M + Na]^+$

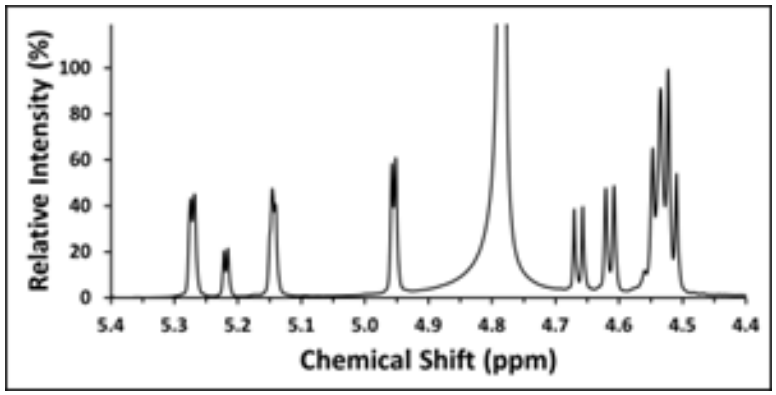
*m/z* = 1261.4




**HPAEC-PAD**

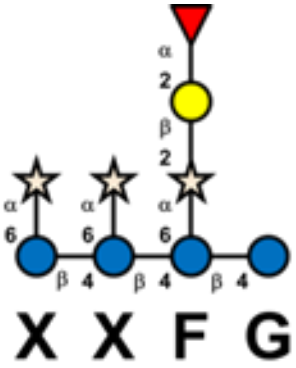



**$^1$H NMR**



**Y**

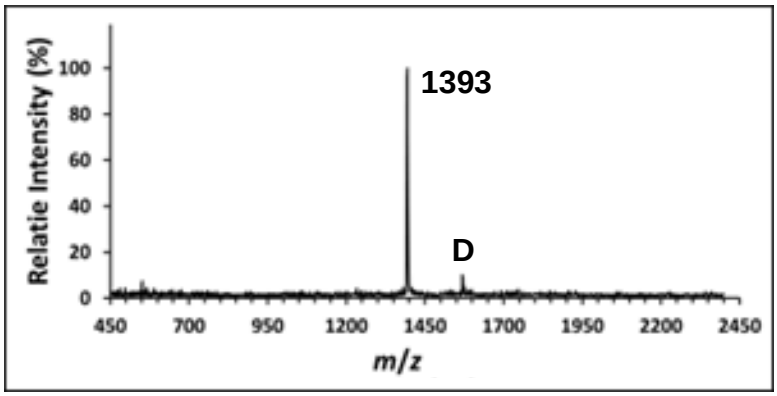
**Chemical Structure**



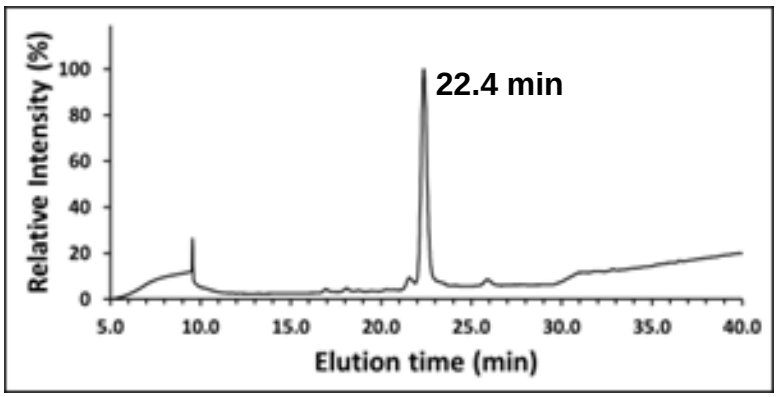

**MALDI-TOF MS**

Calc. $[M + Na]^+$

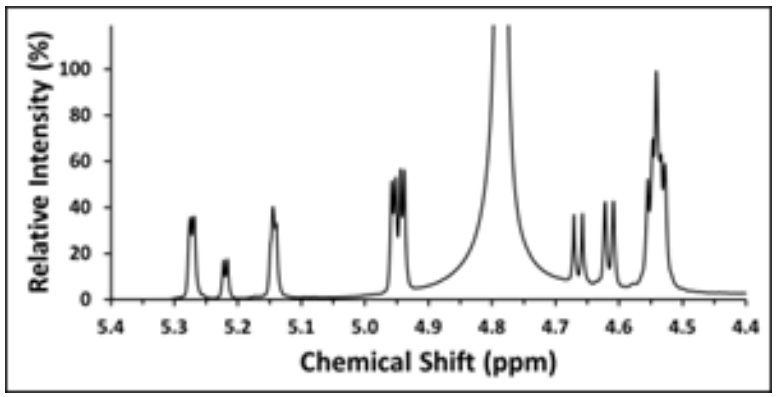
*m/z* = 1393.5




**HPAEC-PAD**




**$^1$H NMR**



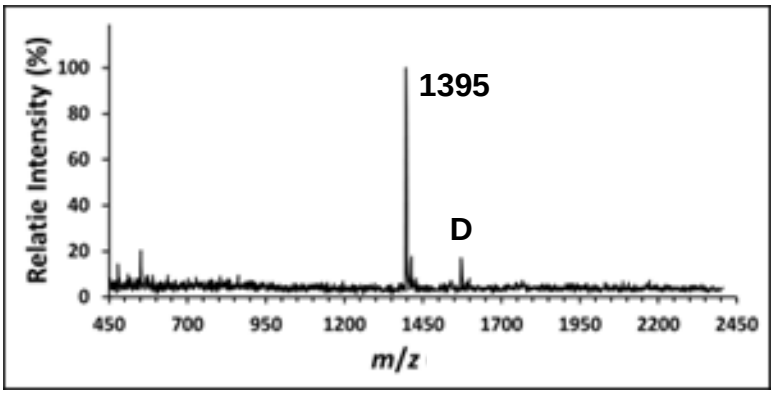
**Z**

**Chemical Structure**

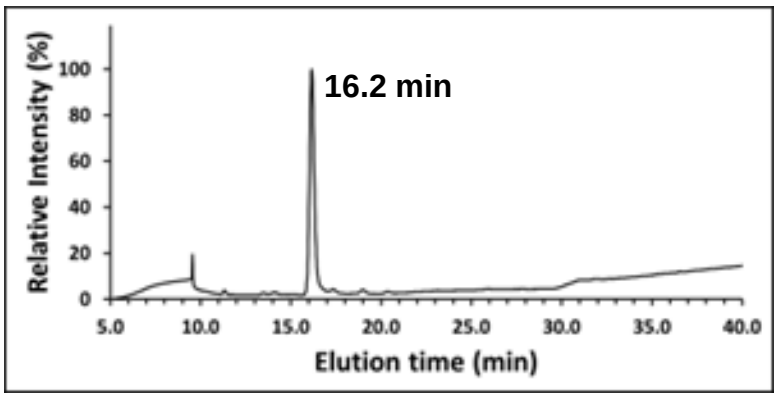

**MALDI-TOF MS**

Calc. $[M + Na]^+$

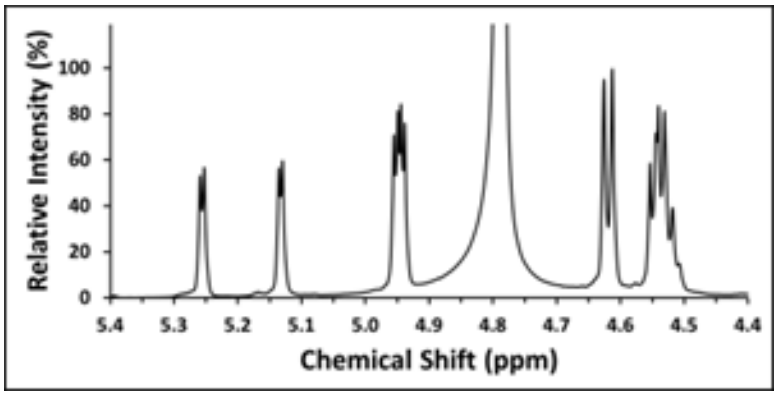
*m/z* = 1395.5



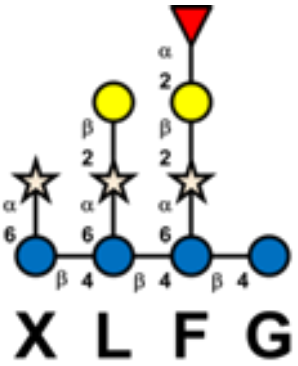

**HPAEC-PAD**




**$^1$H NMR**



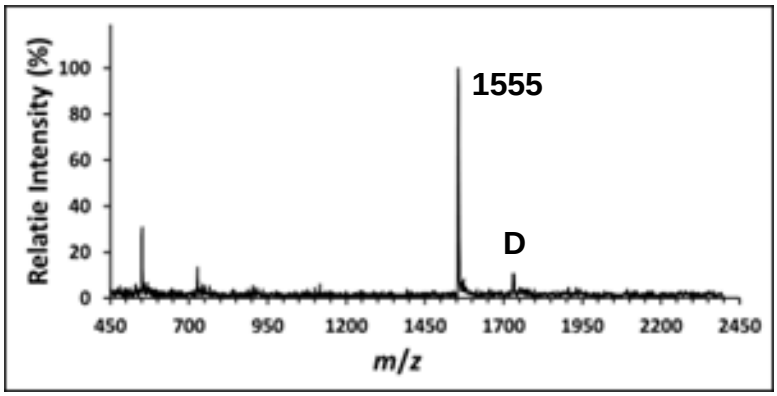
**AA** **Chemical Structure**



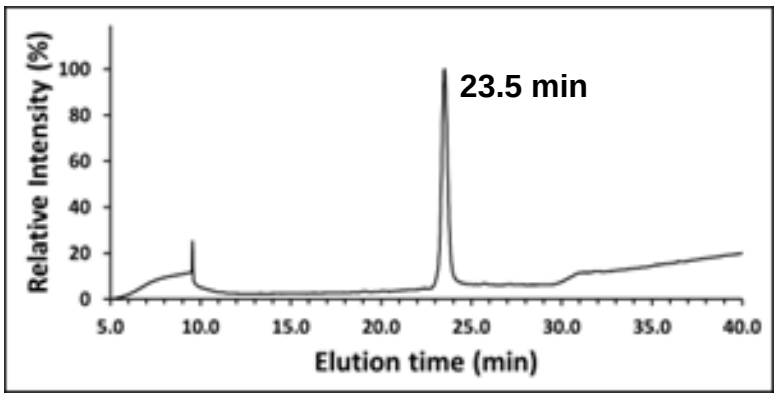

**MALDI-TOF MS**

Calc. $[M + Na]^+$

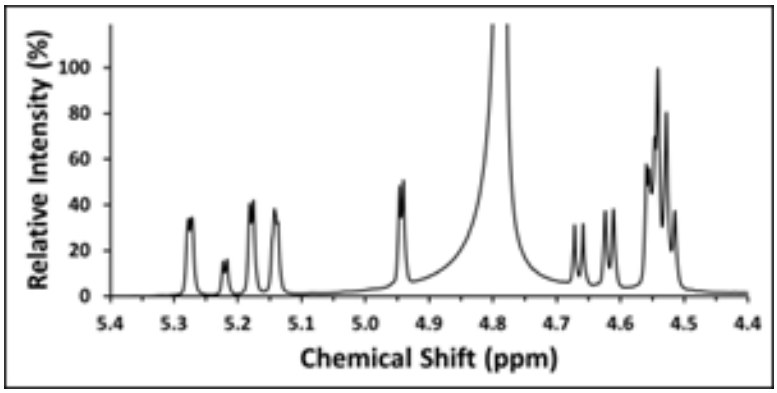
*m/z* = 1555.5



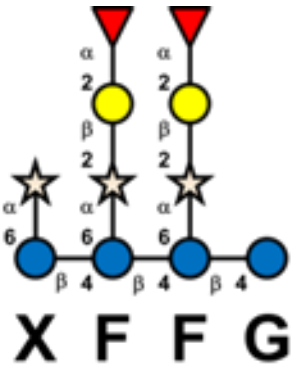

**HPAEC-PAD**




**$^1$H NMR**



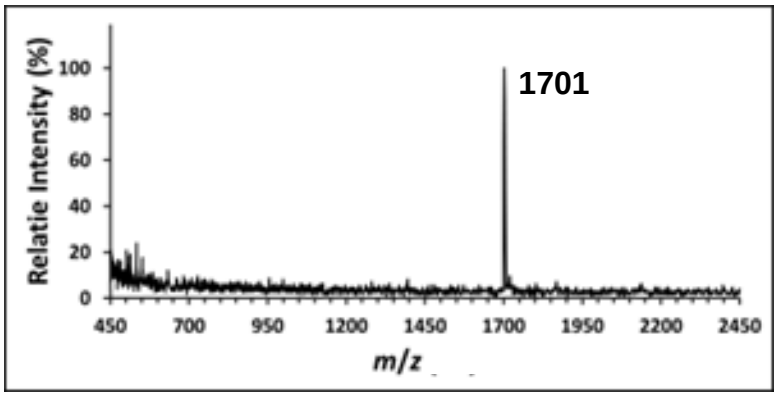
**AB** **Chemical Structure**



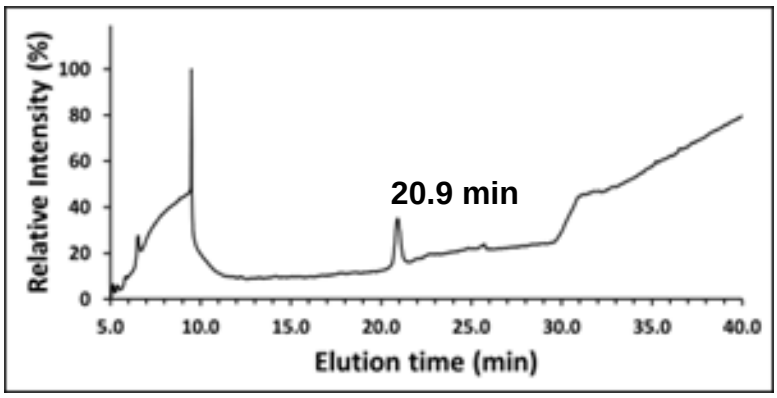

**MALDI-TOF MS**

Calc. $[M + Na]^+$

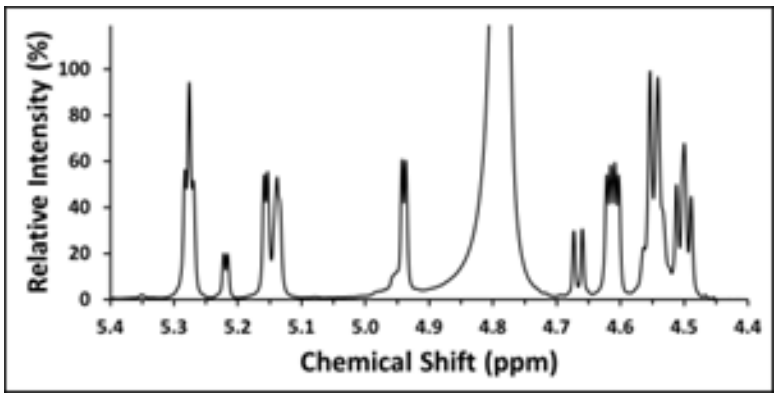
*m/z* = 1701.6



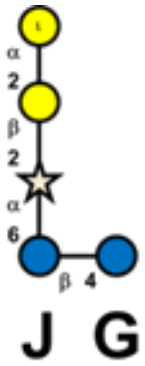

**HPAEC-PAD**




**$^1$H NMR**



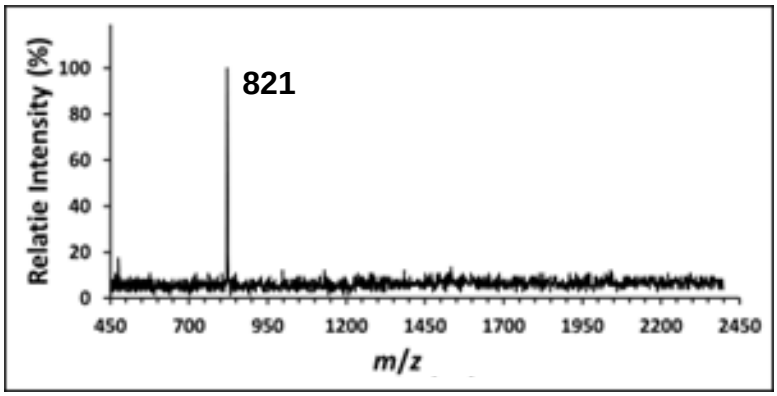
**AC** **Chemical Structure**



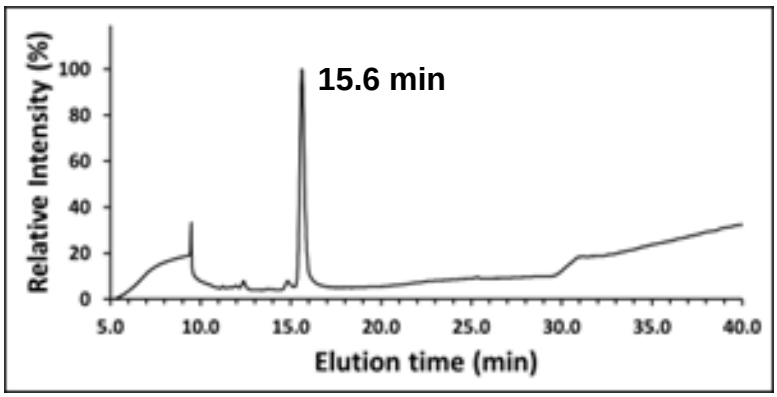

**MALDI-TOF MS**

Calc. $[M + Na]^+$

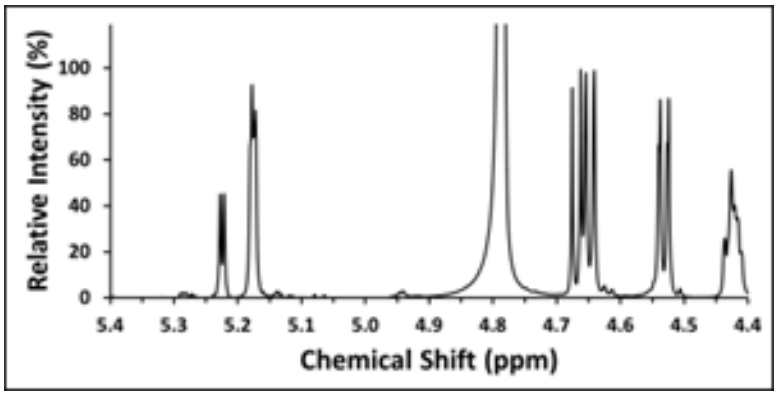
*m*/*z* = 821.3




**HPAEC-PAD**

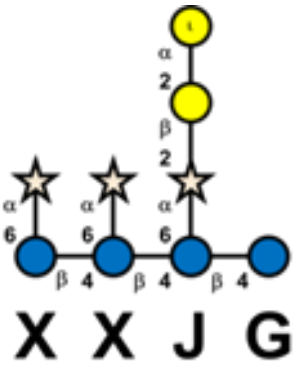



**$^1$H NMR**



**AD**

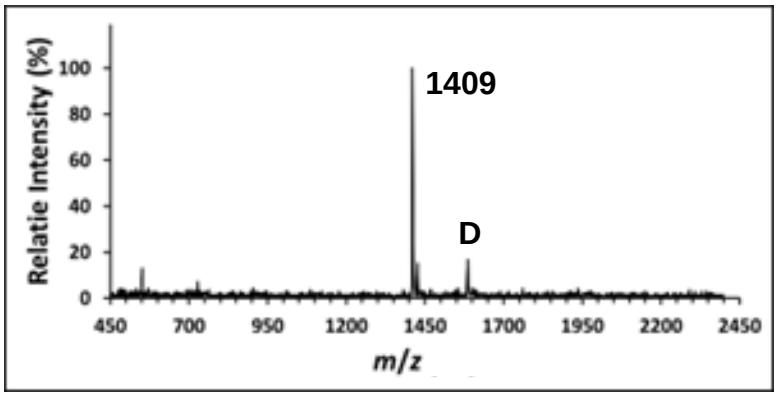
**Chemical Structure**



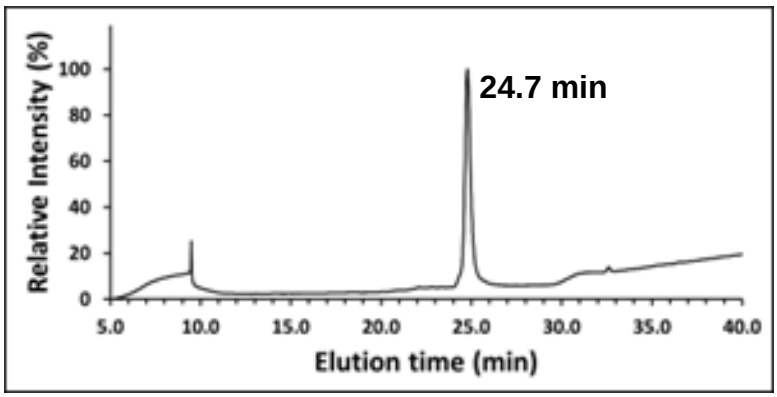

**MALDI-TOF MS**

Calc. $[M + Na]^+$

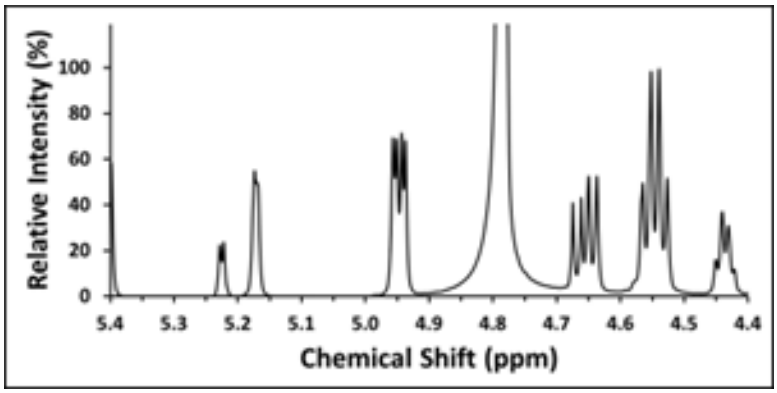
*m/z* = 1409.4




**HPAEC-PAD**

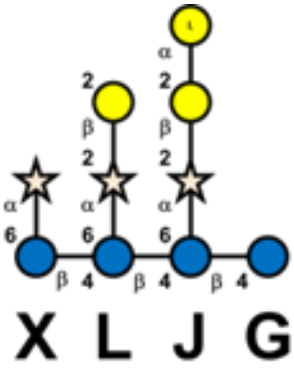



**$^1$H NMR**



**AE**

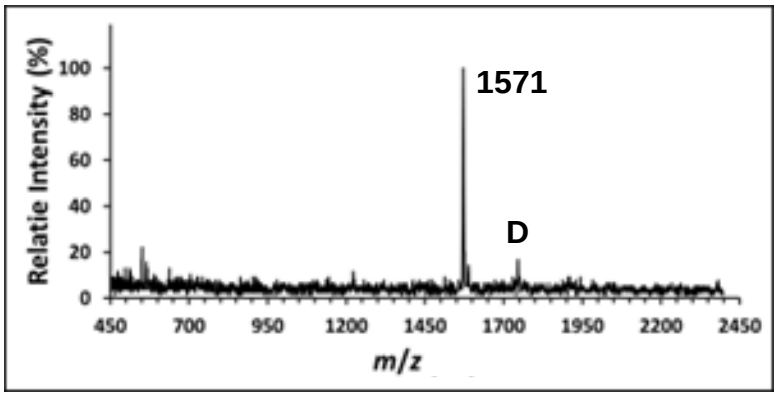
**Chemical Structure**



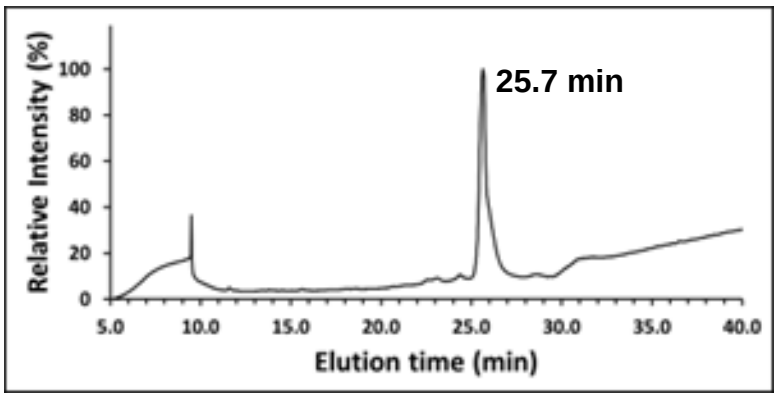

**MALDI-TOF MS**

Calc. $[M + Na]^+$

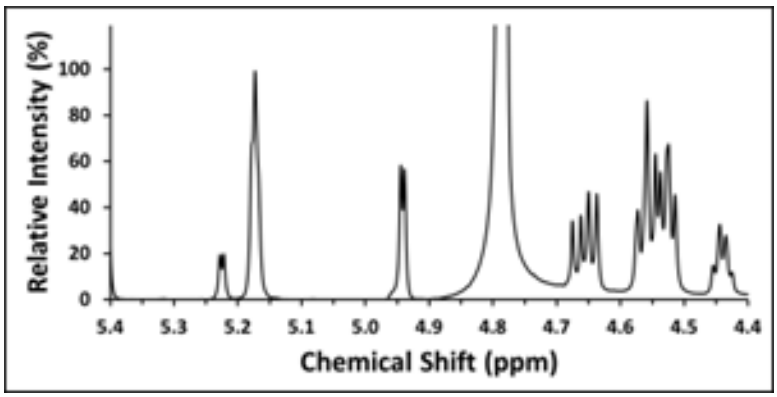
*m/z* = 1571.5



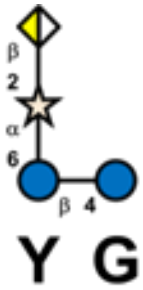

**HPAEC-PAD**




**$^1$H NMR**



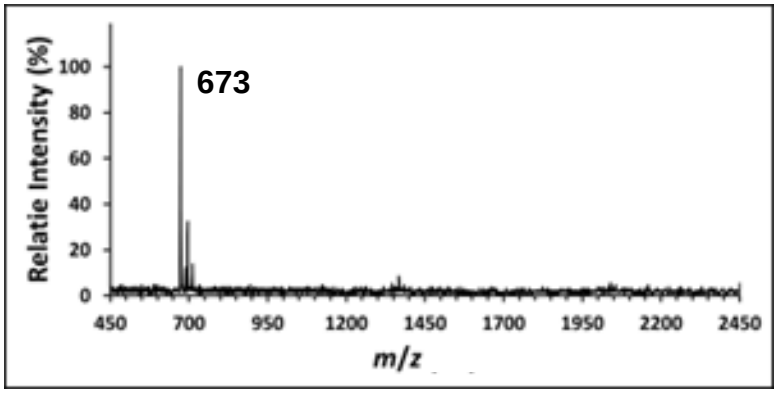
**AF** **Chemical Structure**



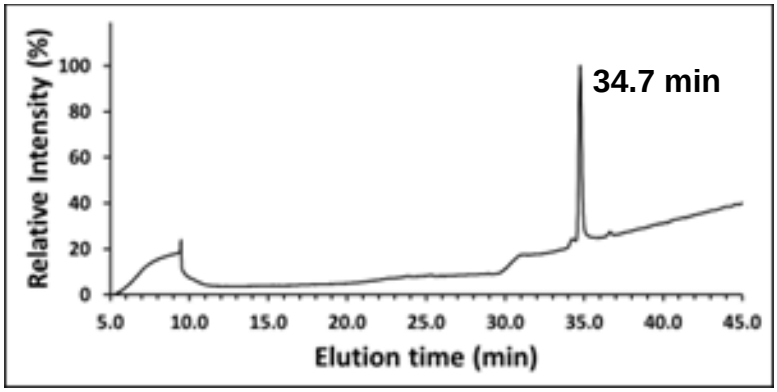

**MALDI-TOF MS**

Calc. $[M + Na]^+$

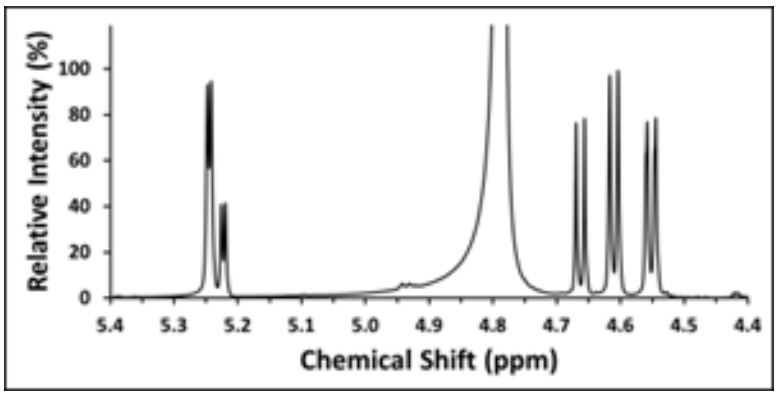
*m/z* = 673.2



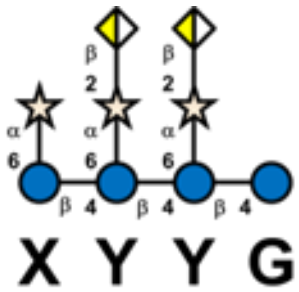

**HPAEC-PAD**




**$^1$H NMR**



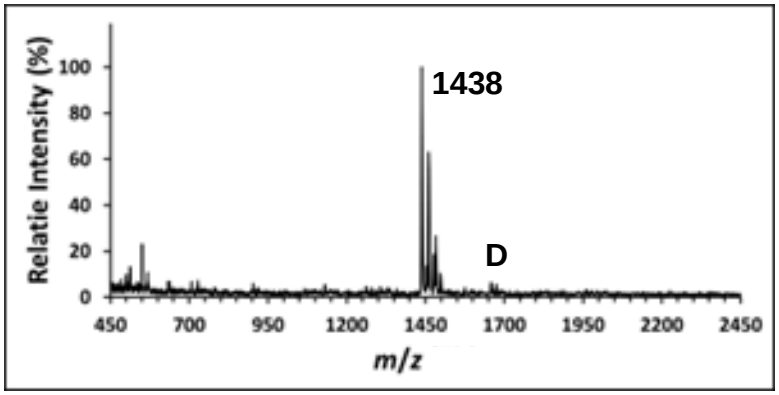
**AG** **Chemical Structure**



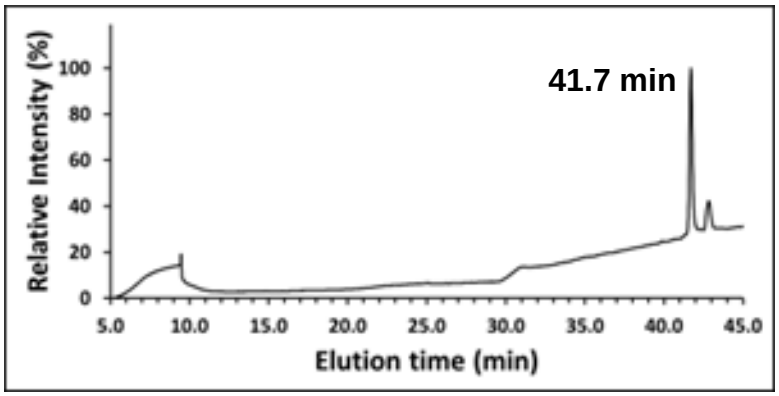

**MALDI-TOF MS**

Calc. $[M + Na]^+$

*m/z* = 1437.4




**HPAEC-PAD**




**$^1$H NMR**

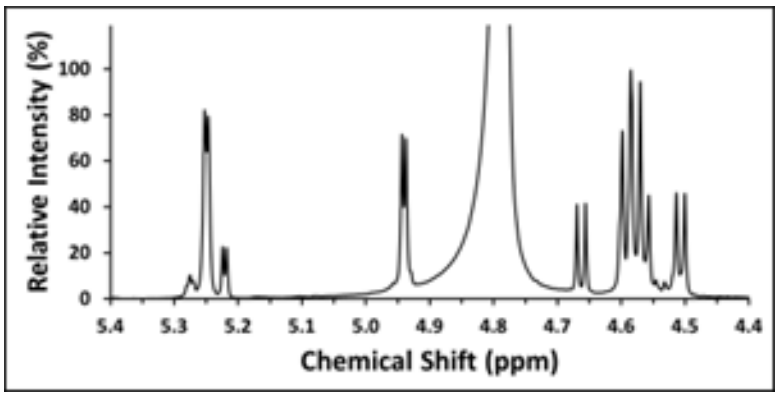



**Figure 2-S3.** Determination of the structural identities and purities of the generated oligosaccharides. The **chemical structures** of the oligosaccharides are rendered using the Consortium for Functional Glycomics symbols ((Varki et al, 2009) for key, see **AH**) and the XyG nomenclature is based on (Fry et al., 1993). In the **MALDI-TOF MS spectra**, the most intense signal always has *m*/*z* value within 1 Th of the calculated *m*/*z* value of the sodium adduct $[M + Na]^+$ of the analyte. Other prominent, but significantly less intense signals can arise from the potassium adduct $[M + K]^+$ at 16 Th higher *m*/*z* value to the parent (sodium adduct) peak, double adduct of sodium and 2,5-dihydroxybenzoic acid (DHB) $[M + Na + DHB]^+$ at 176 Th higher *m*/*z* value (indicated by D), or, dehydration products (only applicable for the cellodextrins) at 18 Th lower *m*/*z* value. For the acidic XyG oligosaccharides (YG and XYYG), additional signals can arise from the sodium or potassium salts of the carboxylic functionality at 22 or 38 Th higher *m*/*z* value, respectively. The *m*/*z* values are tabulated in Supplemental Table 1. In the **HPAEC-PAD chromatograms**, the most intense signal always corresponds to the analyte, except for XFFG (**AB**) whose intensity is lower than solvent front (at 9.5 min) due to low loaded amount of the analyte. The elution time of the oligosaccharide is indicated. Note that retention time of X (isoprimeverose, **F**) coincides with a signal from solvent front. All elution times

are presented in Supplementary Table 1. The **$^{1}$H NMR spectra** are represented by the “anomeric region” where diagnostic signals from H-1 anomeric protons reside. Few other signals in this region arise from the H-5 of the L-Fuc*p* and L-Gal*p*, as well as the H-4 of the D-Gal*p*A residues. The chemical shifts in the anomeric region are tabulated in Supplementary Table 1. XXXXG (**N**) refers to a mixture of two oligosaccharides, namely XXXXG (with canonical XyG structure) and $X^{3}$XXXG (with *O*-3 linkage in the backbone). XXXXXG (**M**) refers to a mixture as well. Here, the four potential structures are XXXXXG, $X^{3}$XXXXG, $XX^{3}$XXXG and $X^{3}X^{3}$XXXG. In both cases, the MALDI-TOF MS spectra yield only single major signal, but the heterogeneity is revealed by HPAEC-PAD chromatograms.

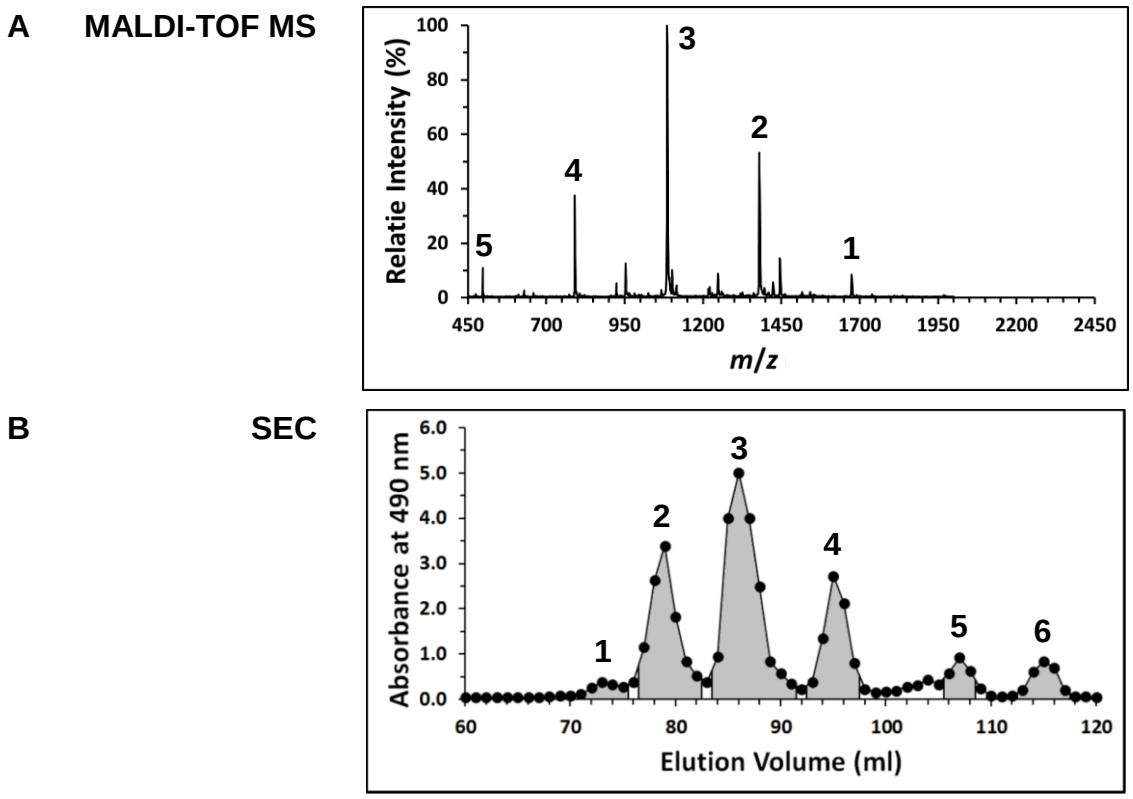



**Figure 2-S4.** The generation and purification of isoprimeverosyl-series XyG oligosaccharides. MALDI-TOF MS spectrum (**A**) and SEC chromatogram (**B**) of the oligosaccharide mixture. The SEC fractions were analyzed by MALDI-TOF (data not shown) and fractions containing pure oligosaccharides were pooled as indicated by the chromatogram shading. The structural identities of the oligosaccharides were further established by MALDI-TOF MS, HPAEC-PAD and NMR (see Supplemental Figure 2). Peaks in both MS spectra and SEC chromatogram are labeled as follows: Peak 1, XXXXXG (*m*/*z* of Na-adduct 1673.5); 2, XXXXG (*m*/*z* 1379.4); 3, XXXG (*m*/*z* 1085.3); 4, XXG (*m*/*z* 791.2); 5, XG (*m*/*z* 497.2). Isoprimeverose (Peak 6, *m*/*z* 335) was not detectable in this complex mixture by MS due to its low abundance as well as matrix interference

at low *m*/*z* values. Note that XXXXXG and XXXXG are mixtures of isomers, for structures, see Supplemental Figure 5.

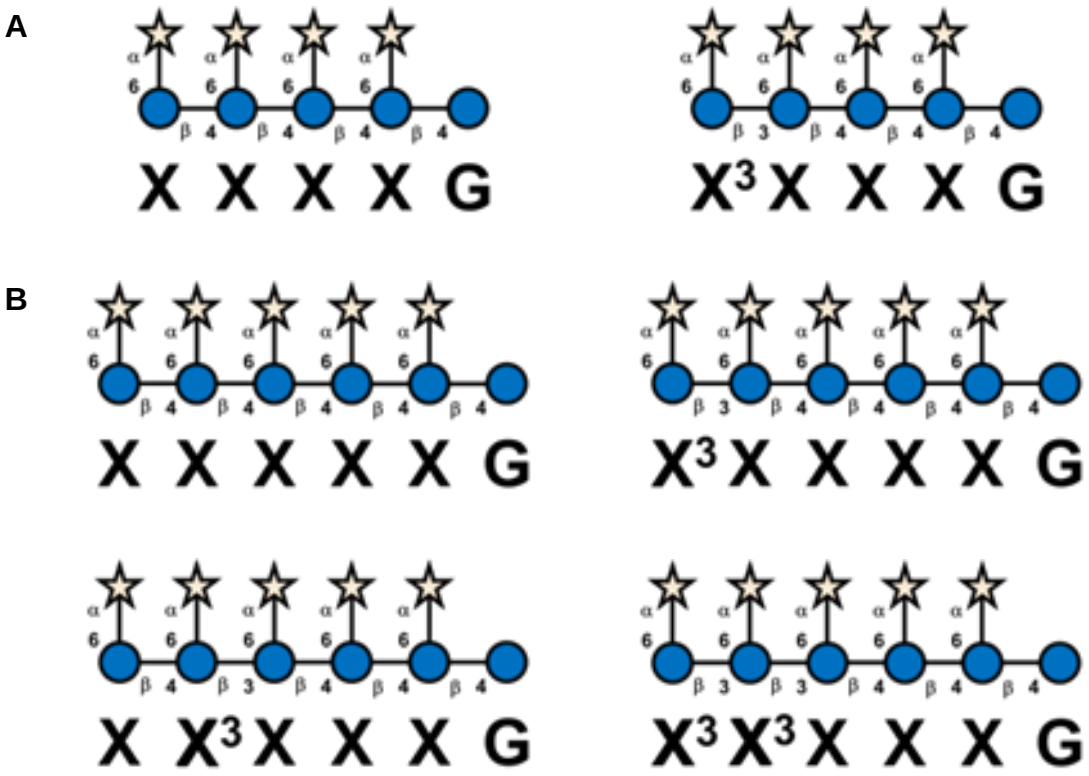



**Figure 2-S5.** Chemical structures of the transglycosylation products of isoprimeverosyl hydrolase. (**A**) XXXXG and its isomer. (**B**) XXXXXG and its isomers. Note that the only difference between the isomeric structures is the linkage position of the X units.

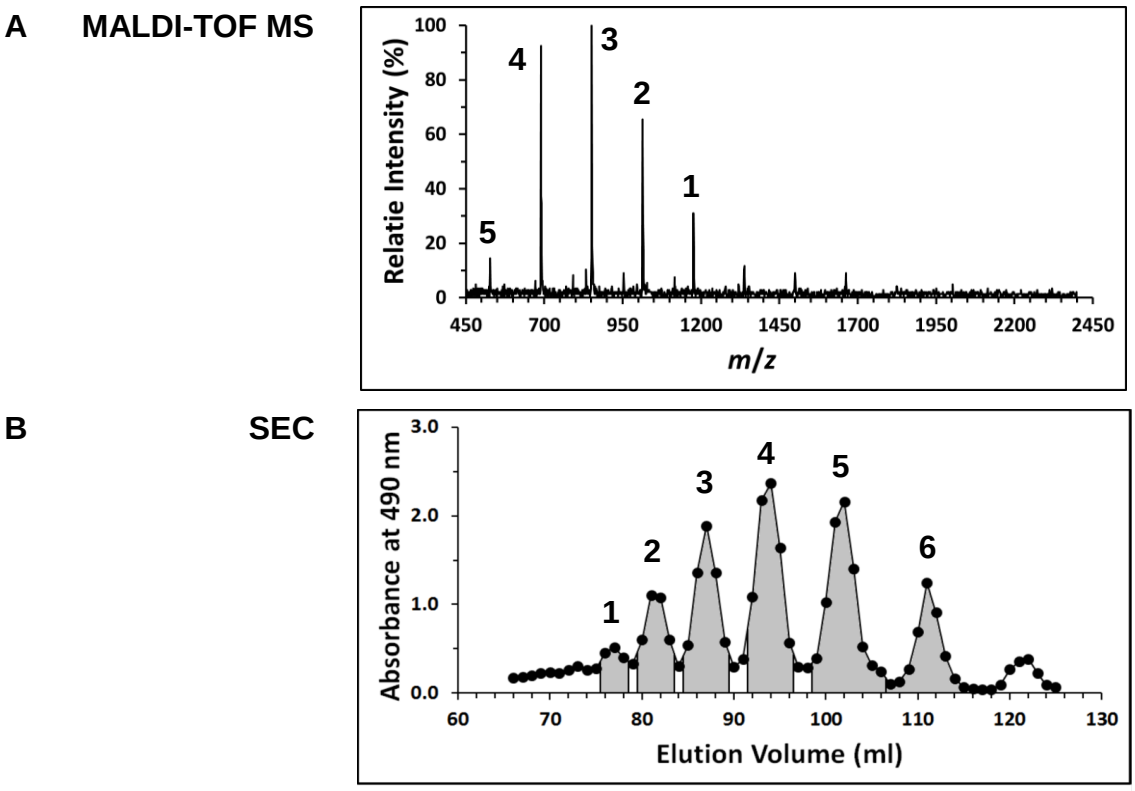



**Figure 2-S6.** The generation and purification of cellodextrin oligosaccharides. (**A**) MALDI-TOF MS spectrum and (**B**) SEC chromatogram of the water soluble cellodextrins. The fractions were analyzed by MALDI-TOF MS (data not shown) and the pure oligosaccharides were pooled as indicated by the chromatogram shading. Peaks in both MS spectra and SEC chromatogram are labeled as follows: Peak 1, celloheptaose (GGGGGGG, m/z of Na-adduct 1175.4); 2, cellohexaose (GGGGGG, m/z 1013.3); 3, cellopentaose (GGGGG, m/z 851.3); 4, cellotetraose (GGGG, m/z 689.2); and 5, cellotriose (GGG, m/z 527.2). The two remaining peaks in the SEC chromatogram correspond to cellobiose and glucose and were not analyzed in detail.

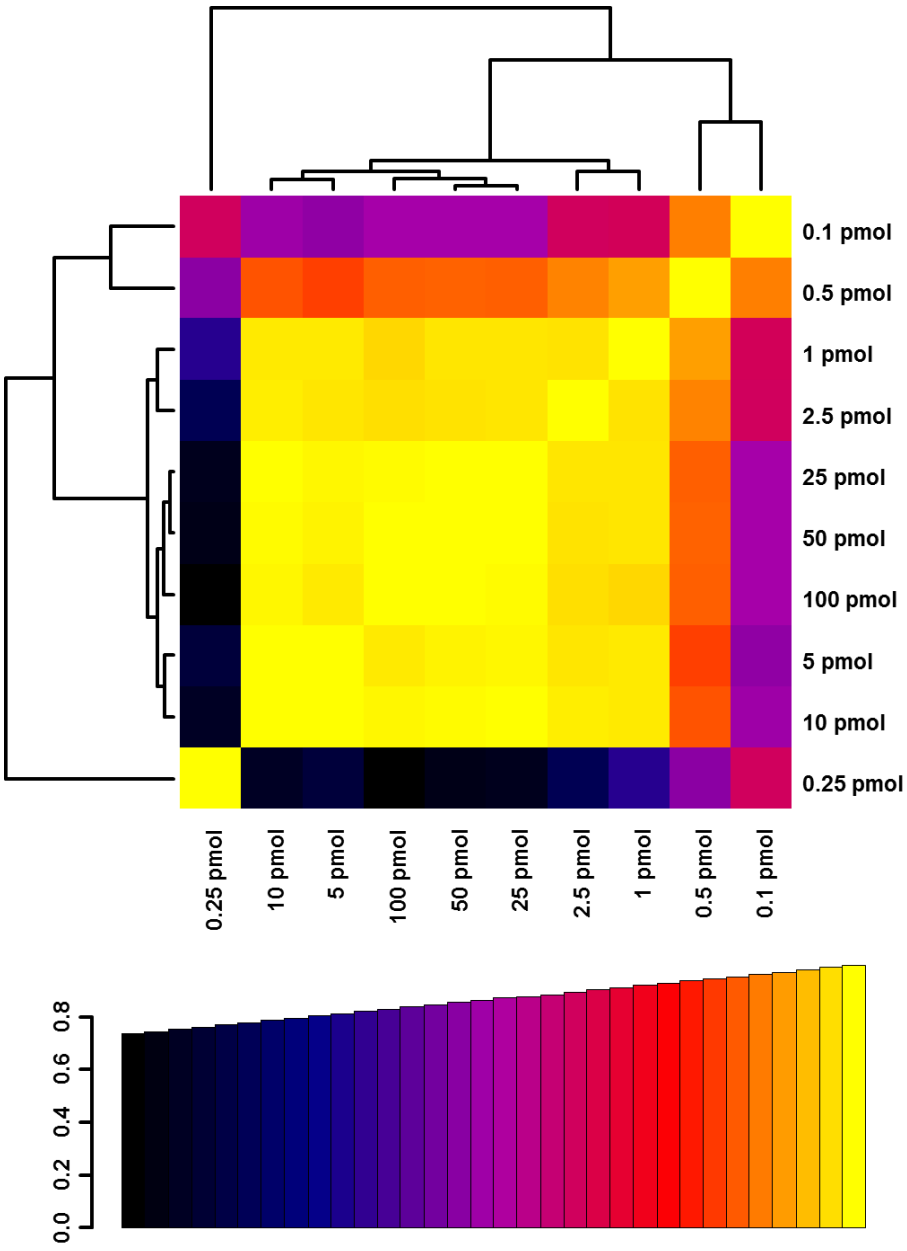



**Figure 2-S7.** Correlation heatmap of the oligosaccharide integral patterns in MALDI-TOF MS spectra between different dilutions of the oligosaccharide mixture. Palette key (**bottom**) provides the scale of the correlations that can range from zero (totally uncorrelated) to one (totally correlated).

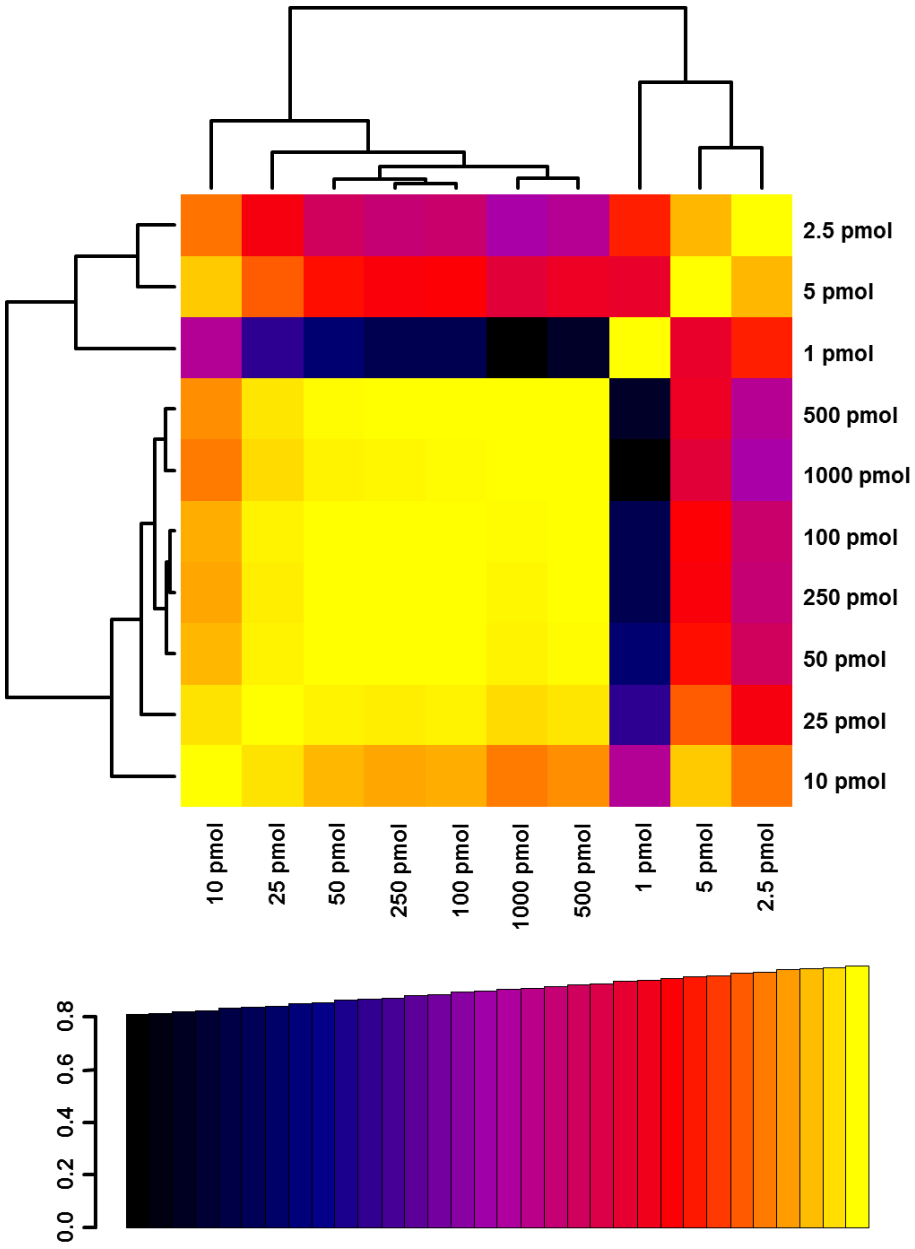



**Figure 2-S8.** Correlation heatmap of the oligosaccharide integral patterns in HPAEC-PAD chromatograms between different dilutions of the oligosaccharide mixture. Palette key (**bottom**) provides the scale of the correlations that can range from zero (totally uncorrelated) to one (totally correlated).

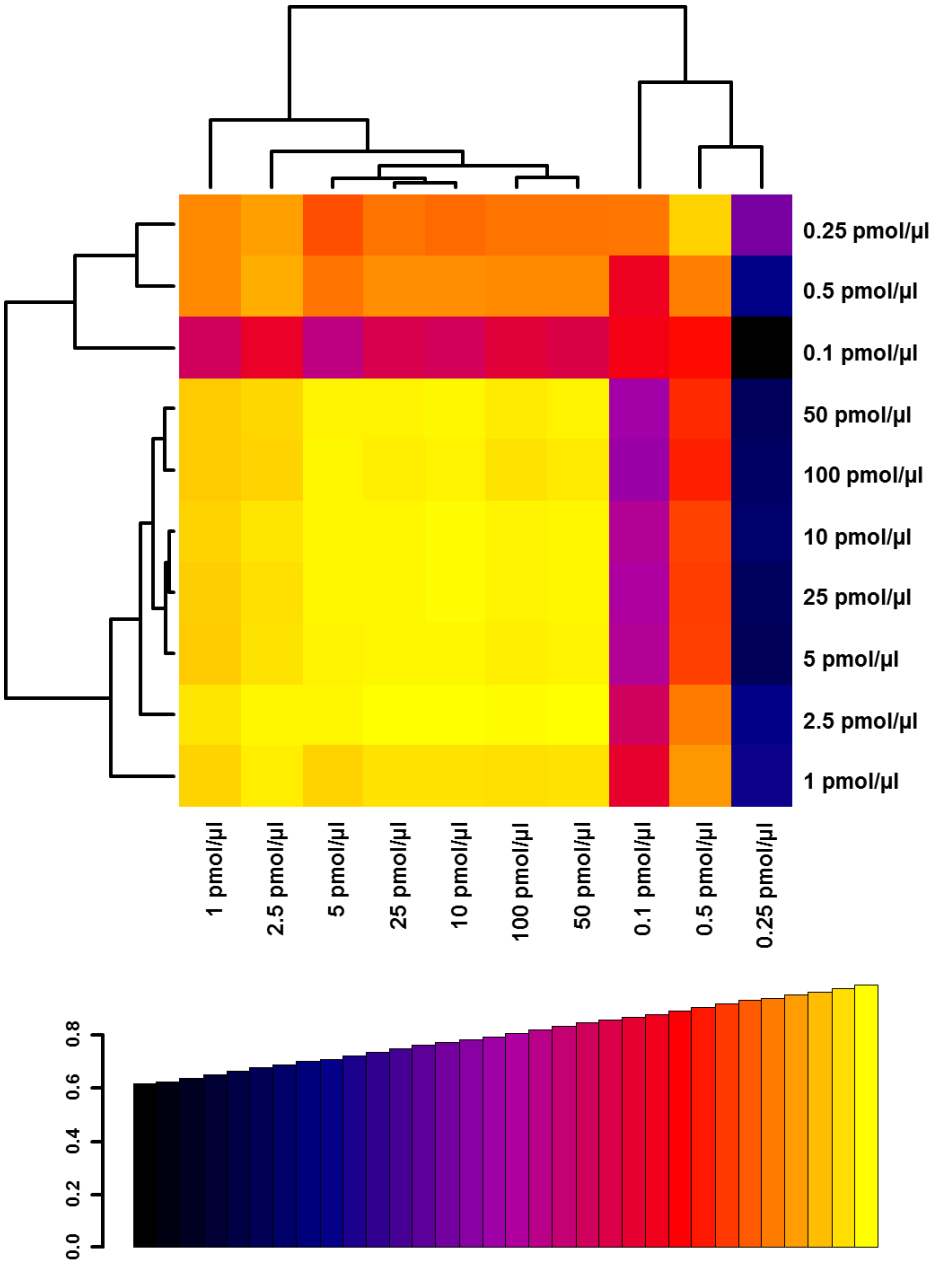



**Figure 2-S9.** Correlation heatmap of the MALDI-TOF and HPAEC-PAD integral patterns. Palette key (**bottom**) provides the scale of the correlations that can range from zero (totally uncorrelated) to one (totally correlated).

**Table 2-S1.** MALDI-TOF MS and HPAEC-PAD data of the purified oligosaccharides.

| # | Oligosaccharide | | MALDI-TOF MS | | HPAEC-PAD | |
|---|---|---|---|---|---|---|
| | Structure | Molar Mass (g/mol) | $[M + Na]^+$ *m/z* (Th) | Relative Response Factor | Elution time (min) | Relative Response Factor |
| 1 | GGG | 504.4 | 527.2 | - | 12.2 | - |
| 2 | GGGG | 666.6 | 689.2 | - | 16.2 | - |
| 3 | GGGGG | 828.7 | 851.3 | - | 21.1 | - |
| 4 | GGGGGG | 990.9 | 1013.3 | - | 25.7 | - |
| 5 | GGGGGGG | 1153.0 | 1175.4 | - | 30.5[1] | - |
| 6 | X | 312.3 | 335.1 | - | 9.5 | - |
| 7 | XG | 474.4 | 497.2 | - | 11.8 | - |
| 8 | XX | 606.5 | 629.2 | - | 12.2 | - |
| 9 | XXG | 768.7 | 791.2 | - | 16.5 | - |
| 10 | GXXG | 930.8 | 953.3 | $0.76 \pm 0.02$ | 18.6 | $0.87 \pm 0.03$ |
| 11 | XXXG | 1062.9 | 1085.3 | $1.00 \pm 0.02$ | 21.8 | $1.00 \pm 0.02$ |
| 12 | XXXGol | 1064.9 | 1087.4 | - | 14.2 | - |
| 13 | XXXXG | 1357.2 | 1379.4 | - | 24.7[2] | - |
| 14 | XXXXXG | 1651.5 | 1673.5 | - | Multiple[3] | - |
| 15 | LG | 636.6 | 659.2 | - | 14.1 | - |

**Table 2-S1.** *Continued*.

| # | Oligosaccharide | | MALDI-TOF MS | | HPAEC-PAD | |
|---|---|---|---|---|---|---|
| | Structure | Molar Mass (g/mol) | $[M + Na]^+$ *m/z* (Th) | Relative Response Factor | Elution time (min) | Relative Response Factor |
| 16 | XLXG | 1225.1 | 1247.4 | - | 23.6 | - |
| 17 | XXLG | 1225.1 | 1247.4 | 1.21 ± 0.03 | 24.4 | 0.94 ± 0.03 |
| 18 | GLLG | 1255.1 | 1277.4 | - | 23.0 | - |
| 19 | XLLG | 1387.2 | 1409.4 | 1.40 ± 0.07 | 25.5 | 0.87 ± 0.02 |
| 20 | XLLGol | 1389.2 | 1411.5 | - | 18.9 | - |
| 21 | XDXG | 1195.1 | 1217.4 | - | 21.1 | - |
| 22 | XDDG | 1327.2 | 1349.4 | - | 21.7 | - |
| 23 | FG | 782.7 | 805.3 | - | 12.5 | - |
| 24 | GXFG | 1239.1 | 1261.4 | | 19.5 | |
| 25 | XXFG | 1371.2 | 1393.5 | 1.78 ± 0.05 | 22.4 | 1.14 ± 0.05 |
| 26 | XXFGol | 1723.5 | 1395.5 | - | 16.2 | - |
| 27 | XLFG | 1533.4 | 1555.5 | 1.78 ± 0.03 | 23.5 | 1.20 ± 0.06 |
| 28 | XFFG | 1679.5 | 1701.6 | - | 20.9 | - |
| 29 | JG | 798.70 | 821.3 | - | 15.6 | - |
| 30 | XXJG | 1387.2 | 1409.4 | - | 24.7 | - |

**Table 2-S1.** *Continued*.

| | | | | | | |
|---|---|---|---|---|---|---|
| 31 | XLJG | 1549.4 | 1571.5 | - | 25.7 | - |
| 32 | YG | 650.5 | 673.2 | - | 34.7 | - |
| 33 | XYYG | 1415.2 | 1437.4 | - | 41.7 | - |

[1]Multiple signals are observed. Owing to its intensity and elution time, the signal at 30.5 min most likely corresponds to GGGGGGG. The signal at 25.8 min corresponds to GGGGGG, the origin of the remaining signals is unknown.

[2]Two signals are observed at 24.7 and 32.5 min. Owing to their elution times, the signal at 24.7 min most likely corresponds to XXXXG and that at 32.5 min to $X^3$XXXG. Here, the superscript after X indicates a non-canonical *O*-3 glycosidic linkage between the backbone Glc*p* residues.

[3]Multipe signals are observed, some of which may correspond to XXXXXG, $X^3$XXXXG, $XX^3$XXXG, and $X^3X^3$XXXG.

**Table 2-S2.** NMR chemical shifts of the purified oligosaccharides.

| # | Structure | $^1$H Chemical Shift (ppm)[1] | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Residue Type** | **D-Glc*p* H-1** | | | | | | | | | | | | | |
| | **Residue Number**[2] | **1-α** | **1-β** | **2** | **3** | **4** | **5** | **6** | **7** | | | | | | |
| 1 | **GGG** | 5.222 | 4.659 | 4.534 | 4.509 | | | | | | | | | | |
| 2 | **GGGG** | 5.222 | 4.659 | 4.534 | | 4.507 | | | | | | | | | |
| 3 | **GGGGG** | 5.222 | 4.659 | 4.533 | | | 4.507 | | | | | | | | |
| 4 | **GGGGGG** | 5.222 | 4.659 | 4.532 | | | | 4.507 | | | | | | | |
| 5 | **GGGGGGG** | 5.222 | 4.659 | 4.532 | | | | | 4.507 | | | | | | |
| | **Residue Type** | **D-Glc*p* H-1** | | | | | | | **D-Xyl*p* H-1** | | | | | | |
| | **Residue Number** | **1-α** | **1-β** | **2** | **3** | **4** | **5** | **6** | **1** | **2** | **3** | **4** | **5** | **6** | |
| 6 | **X** | 5.243 | 4.677 | -- | -- | -- | -- | -- | 4.914<br>4.921 A | -- | -- | -- | -- | -- | |
| 7 | **XG** | 5.222 | 4.662 | 4.541 | -- | -- | -- | -- | -- | 4.942 | -- | -- | -- | -- | |
| 8 | **XX** | 5.232 | 4.691 | 4.563 | -- | -- | -- | -- | 4.938<br>4.946 A | 4.944 | -- | -- | -- | -- | |
| 9 | **XXG** | 5.222 | 4.666 | 4.554, 4.559 | | -- | -- | -- | -- | 4.960 | 4.941 | -- | -- | -- | |
| 10 | **GXXG** | 5.221 | 4.661 | 4.558, 4.570 | | 4.517 | -- | -- | -- | 4.959 | | -- | -- | -- | |
| 11 | **XXXG** | 5.220 | 4.661 | 4.550, 4.560, 4.574 | | | -- | -- | -- | 4.958 | | 4.940 | -- | -- | |
| 12 | **XXXGol** | -- | | 4.631 | 4.586 | 4.549 | -- | -- | -- | 4.952, 4.958 | | 4.939 | -- | -- | |
| 13 | **XXXXG** | 5.220 | 4.660 | 4.549, 4,586, 4.559, 4.563, 4.574 | | | | -- | -- | 4.956 | | 4.962\<br>4.956 | 4.932\<br>4.940 | -- | |
| 14 | **XXXXXG** | 5.220 | 4.661 | 4.548, 4.584, 4.559, 4.570 | | | | | -- | 4.955 | | | 4.962\<br>4.955 | 4.932\<br>4.940 | |

**Table 2-S2.** *Continued*.

| | Residue Type | D-Glc*p* H-1 | | | | | D-Xyl*p* H-1 | | | D-Gal*p* H-1 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Residue Number** | **1-α** | **1-β** | **2** | **3** | **4** | **2** | **3** | **4** | **2** | **3** | | | | |
| 15 | **LG** | 5.223 | 4.665 | 4.545 | -- | -- | 5.159 | -- | -- | 4.564 | -- | | | | |
| 16 | **XLXG** | 5.220 | 4.662 | 4.534, 4.558 | | | 4.957 | 5.179 | 4.942 | -- | 4.566 | | | | |
| 17 | **XXLG** | 5.224 | 4.664 | 4.550, 4.559 | | | 5.172 | 4.959 | 4.941 | 4.559 | -- | | | | |
| 18 | **GLLG** | 5.221 | 4.665 | 4.569 | 4.532 | 4.509 | 5.171, 5.179 | | -- | 4.556 | | | | | |
| 19 | **XLLG** | 5.224 | 4.665 | 4.569 | 4.539 | 4.533 | 5.167, 5.178 | | 4.943 | 4.554 | | | | | |
| 20 | **XLLGol** | -- | | 4.632 | 4.533 | 4.544 | 5.162 | 5.177 | 4.942 | 4.549/4.560 | | | | | |
| | **Residue Type** | D-Glc*p* H-1 | | | | | D-Xyl*p* H-1 | | | L-Ara*p* H-1 | | | | | |
| | **Residue Number** | **1-α** | **1-β** | **2** | **3** | **4** | **2** | **3** | **4** | **2** | **3** | | | | |
| 21 | **XDXG** | 5.221 | 4.662 | Appr. 4.55 | | | 4.954 | 5.133 | 4.945 | -- | 4.488 | | | | |
| 22 | **XDDG** | 5.223 | 4.663 | Appr. 4.55 | | | 5.133, 5.156 | | 4.945 | 4.486 | | | | | |
| | **Residue Type** | D-Glc*p* H-1 | | | | | D-Xyl*p* H-1 | | | D-Gal*p* H-1 | | L-Fuc*p* H-1 | | L-Fuc*p* H-4 | |
| | **Residue Number** | **1-α** | **1-β** | **2** | **3** | **4** | **2** | **3** | **4** | **2** | **3** | **2** | **3** | **2** | **3** |
| 23 | **FG** | 5.220 | 4.668 | 4.550 A | -- | -- | 5.136 | -- | -- | 4.621 | -- | 5.287 | -- | 4.514 | -- |
| 24 | **GXFG** | 5.219 | 4.664 | | | | 5.144 | 4.955 | -- | 4.615 | -- | 5.272 | -- | | -- |
| 25 | **XXFG** | 5.220 | 4.665 | | | | 5.143 | 4.955 | 4.941 | 4.616 | -- | 5.272 | -- | | -- |
| 26 | **XXFGol** | -- | | | | | | 5.133 | 4.952 | 4.941 | 4.620 | -- | 5.256 | -- | |
| 27 | **XLFG** | 5.220 | 4.666 | | | | 5.140 | 5.179 | 4.944 | | | 5.275 | -- | | -- |
| 18 | **XFFG** | 5.218 | 4.667 | | | | 5.135 | 5.156 | 4.940 | | | 5.273 | 5.280 | | |

**Table 2-S2.** *Continued.*

| | Residue Type | D-Glc*p* H-1 | | | | | D-Xyl*p* H-1 | | | D-Gal*p* H-1 | | L-Gal*p* H-1 | | L-Gal*p* H-4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Residue Position** | **1-α** | **1-β** | **2** | **3** | **4** | **2** | **3** | **4** | **2** | **3** | **2** | | **2** |
| 29 | **JG** | 5.225 | 4.669 | | -- | -- | 5.178 | -- | -- | | -- | 5.420 | -- | -- |
| 30 | **XXJG** | 5.225 | 4.669 | | | | 5.172 | 4.954 | 4.940 | 4.644 | -- | 5.402 | -- | -- |
| 31 | **XLJG** | 5.225 | 4.669 | | | | 5.173 | 4.942 | | | -- | | -- | -- |
| | **Residue Type** | D-Glc*p* H-1 | | | | | D-Xyl*p* H-1 | | | D-Gal*p*A H-1 | | D-Gal*p*A H-5 | | |
| | **Residue Position** | **1-α** | **1-β** | **2** | **3** | **4** | **2** | **3** | **4** | **2** | **3** | **2** | **3** | |
| 32 | **YG** | 5.223 | 4.663 | 4.552 | -- | -- | 5.248 | -- | -- | 4.611 | -- | 4.230 | 4.198 | |
| 33 | **XYYG** | 5.221 | 4.663 | | | | 5.250 | 4.940 | | | | | | |

[1]Shaded and boxed entries of multiple nuclei indicate either that these nuclei have identical $^{1}$H chemical shifts (when a single value is reported), a corresponding number of unique resonances can be resolved but their assignments are not clear (multiple values separated by a comma), multiple resonances are observed for a single nucleus due to the anomeric effect from the reducing end (multiple values separated by a slash, with the first value associated with the α anomeric form), multiple signals are observed due to the presence of several molecules differing in structure (multiple values separated by a backslash, with the first value associated with *O*-4 linkage), or severe overlap of multiple signals prevents an accurate chemical shift assignment for that group of nuclei (when approximate value is reported).

[2]Backbone residues are numbered sequentially, with the reducing end D-Glc*p* designated as 1. In XLJG, the D-Glc*p* residue with the J side-chain attached is numbered 2, as are the other glycosyl residues (D-Xyl*p*, D-Gal*p*, L-Gal*p*) in that side-chain.

## 2.6. Literature Cited

Albersheim P, Darvill AG, Roberts K., Sederoff R, Staehelin A (2010) *Plant Cell Walls*. Garland Science, New York.

Anderegg R (1990) Mass spectrometry: An introduction. *In:* Suelter C (*ed*) *Biomedical Applications of Mass Spectrometry*, pp. 1-90. John Wiley & Sons, New York.

Bauer S (2012) Mass spectrometry for characterizing plant cell wall polysaccharides. *Frontiers in Plant Science* **3**:45-50.

Bauer S, Vasu P, Mort AJ, Somerville CR (2005) Cloning, expression, and characterization of an oligoxyloglucan reducing end-specific xyloglucanobiohydrolase from *Aspergillus nidulans*. *Carbohydrate research* **340**: 2590-2597.

Bonin CP, Potter I, Vanzin GF, Reiter WD (1997) The MUR1 gene of Arabidopsis thaliana encodes an isoform of GDP-D-mannose-4,6-dehydratase, catalyzing the first step in the de novo synthesis of GDP-L-fucose. *Proceedings of the National Academy of Sciences of the United States of America* **94**: 2085-2090.

Bubb WA (2003) NMR spectroscopy in the study of carbohydrates: Characterizing the structural complexity. *Concepts in Magnetic Resonance Part A* **19A**: 1-19.

Buckeridge MS (2010) Seed cell wall storage polysaccharides: Models to understand cell wall biosynthesis and degradation. *Plant Physiology* **154**: 1017-1023.

Buckeridge MS, Crombie HJ, Mendes CJ, Reid JS, Gidley MJ, Vieira CC (1997) A new family of oligosaccharides from the xyloglucan of *Hymenaea courbaril* L. (Leguminosae) cotyledons. *Carbohydrate Research* **303**: 233-237.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active enZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Research* **37**: D233-238.

Cobas JC, Sardina FJ (2003) Nuclear magnetic resonance data processing. MestRe-C: A software package for desktop computers. *Concepts in Magnetic Resonance Part A* **19A**: 80-96.

Corradini C, Cavazza A, Bignardi C (2012) High-performance anion-exchange chromatography coupled with pulsed electrochemical detection as a powerful tool to evaluate carbohydrates of food interest: Principles and applications. *International Journal of Carbohydrate Chemistry* **2012**: Article ID 487564.

Crombie HJ, Chengappa S, Hellyer A, Reid JS (1998) A xyloglucan oligosaccharide-active, transglycosylating beta-D-glucosidase from the cotyledons of nasturtium (*Tropaeolum majus* L) seedlings--purification, properties and characterization of a cDNA clone. *The Plant Journal* **15**: 27-38.

Dell A, Chalabi S, Hitchen PG, Jang-Lee J, Ledger V, North SJ, Pang PC, Parry S, Sutton-Smith M, Tissot B, Morris HR, Panico M, Haslam SM (2007) Mass spectrometry of glycoprotein glycans: Glycomics and glycoproteomics. *In:* Kamerling H (*ed*) Comprehensive Glycoscience, pp. 69-100. Elsevier. Oxford.

Djordjevic SP, Batley M, Redmond JW (1986) Preparative gel chromatography of acidic oligosaccharides using a volatile buffer. *Journal of Chromatography A* **354**: 507-510.

Eklof JM, Brumer H (2010) The XTH gene family: An update on enzyme structure, function, and phylogeny in xyloglucan remodeling. *Plant Physiology* **153**: 456-466.

Fox JD, Robyt JF (1991) Miniaturization of three carbohydrate analyses using a microsample plate reader. *Analytical Biochemistry* **195**: 93-96.

Fry SC, McDougall GJ, Lorences EP, Biggs KJ, Smith RC (1990) Oligosaccharins from xyloglucan and cellulose: Modulators of the action of auxin and $H^+$ on plant growth. *Symposia of the Society for Experimental Biology* **44**: 285-298.

Fry SC, York WS, Albersheim P, Darvill A, Hayashi T, Joseleau J-P, Kato Y, Lorences EP, Maclachlan GA, McNeil M, Mort AJ, Grant Reid JS, Seitz HU, Selvendran RR, Voragen AGJ, White AR (1993) An unambiguous nomenclature for xyloglucan-derived oligosaccharides. *Physiologia Plantarum* **89**: 1-3.

Gibson LJ (2012) The hierarchical structure and mechanics of plant materials. *Journal of the Royal Society Interface* **9**: 2749-2766.

Gilbert HJ (2010) The biochemistry and structural biology of plant cell wall deconstruction. *Plant Physiology* **153**: 444-455.

Gilbert HJ, Stalbrand H, Brumer H (2008) How the walls come crumbling down: Recent structural biochemistry of plant polysaccharide degradation. *Current Opinion in Plant Biology* **11**: 338-348.

Gloster TM, Ibatullin FM, Macauley K, Eklof JM, Roberts S, Turkenburg JP, Bjornvad ME, Jorgensen PL, Danielsen S, Johansen KS, Borchert TV, Wilson KS, Brumer H, Davies GJ (2007) Characterization and three-dimensional structures of two distinct bacterial xyloglucanases from families GH5 and GH12. *The Journal of Biological Chemistry* **282**: 19177-19189.

Gould SE, Rees DA, Wight NJ (1971) Polysaccharides in germination. Xyloglucans (amyloids) from the cotyledons of white mustard. *The Biochemical Journal* **124**: 47-53.

Guilhaus M (1995) Special feature: Tutorial. Principles and instrumentation in time-of-flight mass spectrometry. Physical and instrumental concepts. *Journal of Mass Spectrometry* **30**: 1519-1532.

Guillen R, York WS, Pauly M, An J, Impallomeni G, Albersheim P, Darvill AG (1995) Metabolism of xyloglucan generates xylose-deficient oligosaccharide subunits of this polysaccharide in etiolated peas. *Carbohydrate Research* **277**: 291-311.

Gusev AI, Wilkinson WR, Proctor A, Hercules DM (1995) Improvement of signal reproducibility and matrix/comatrix effects in MALDI analysis. *Analytical Chemistry* **67**: 1034-1041.

Hantus S, Pauly M, Darvill AG, Albersheim P, York WS (1997) Structural characterization of novel L-galactose-containing oligosaccharide subunits of jojoba seed xyloglucans. *Carbohydrate Research* **304**: 11-20.

Hardy MR, Rohrer JS (2007) High-pH anion-exchange chromatography (HPAEC) and pulsed amperometric detection (PAD) for carbohydrate analysis. *In:* Kamerling H (*ed*) *Comprehensive Glycoscience*, pp. 303-327. Elsevier, Oxford.

Hayashi T, Kaida R (2011) Functions of xyloglucan in plant cells. *Molecular Plant* **4**: 17-24.

Hilz H, de Jong LE, Kabel MA, Verhoef R, Schols HA, Voragen AGJ (2007) Bilberry xyloglucan—novel building blocks containing β-xylose within a complex structure. *Carbohydrate Research* **342**: 170-181.

Hisamatsu M, Impallomeni G, York WS, Albersheim P, Darvill AG (1991) A new undecasaccharide subunit of xyloglucans with two alpha-L-fucosyl residues. *Carbohydrate Research* **211**: 117-129.

Hisamatsu M, York WS, Darvill AG, Albersheim P (1992) Characterization of seven xyloglucan oligosaccharides containing from seventeen to twenty glycosyl residues. *Carbohydrate Research* **227**: 45-71.

Hoffman M, Jia Z, Peña MJ, Cash M, Harper A, Blackburn AR 2nd, Darvill A, York WS (2005) Structural analysis of xyloglucans in the primary cell walls of plants in the subclass Asteridae. *Carbohydrate Research* **340**: 1826-1840.

Hsieh YS, Harris PJ (2009) Xyloglucans of monocotyledons have diverse structures. *Molecular Plant* **2**: 943-965.

Hsieh YS, Harris PJ (2012) Structures of xyloglucans in primary cell walls of gymnosperms, monilophytes (ferns sensu lato) and lycophytes. *Phytochemistry* **79**: 87-101.

Iglesias N, Abelenda JA, Rodino M, Sampedro J, Revilla G, Zarra I (2006) Apoplastic glycosidases active against xyloglucan oligosaccharides of *Arabidopsis thaliana*. *Plant and Cell Physiology* **47**: 55-63.

IUPAC-IUBMB Joint Commission on Biochemical Nomenclature (JCBN) (1996) Nomenclature of carbohydrates. Recommendations 1996. *Pure and Applied Chemistry* **68**: 1919-2008.

Jensen JK, Schultink A, Keegstra K, Wilkerson CG, Pauly M (2012) RNA-seq analysis of developing nasturtium seeds (*Tropaeolum majus*): Identification and characterization of an additional galactosyltransferase involved in xyloglucan biosynthesis. *Molecular Plant* **5**: 984-992.

Jurinke C (2005) MALDI-TOF mass spectrometry. *In:* Fuchs J, Podda M (*eds*) *Encyclopedia of Medical Genomics and Proteomics*, pp. 103-111. Marcel Dekker, New York.

Kamerling JP, Gerwig GJ (2007) Strategies for the structural analysis of carbohydrates. *In:* Kamerling H (*ed*) *Comprehensive Glycoscience*, pp. 1-68. Elsevier, Oxford.

Kato Y, Matsushita J, Kubodera T, Matsuda K (1985) A novel enzyme producing isoprimeverose from oligoxyloglucans of *Aspergillus oryzae*. *Journal of Biochemistry* **97**: 801-810.

Kazmaier T, Roth S, Zapp J, Harding M, Kuhn R (1998) Quantitative analysis of malto-oligosaccharides by MALDI-TOF mass spectrometry, capillary electrophoresis and anion exchange chromatography. *Fresenius Journal of Analytical Chemistry* **361**: 473-478.

Kimura K, Matsumoto K, Ishihara C, Harada K, Miyagi A (1995) Structure determination of *galacto*-oligosaccharides by pyridylamination and NMR spectroscopy. *Carbohydrate Research* **270**: 33-42.

Knudsen PJ, Eriksen PB, Fenger M, Florentz K (1980) High-performance liquid chromatography of hyaluronic acid and oligosaccharides produced by bovine testes hyaluronidase. *Journal of Chromatography A* **187**: 373-379.

Kohn R, Kovac P (1978) Dissociation-constants of D-galacturonic and D-glucuronic acid and their *O*-methyl derivatives. *Chemické Zvesti* **32**: 478-485.

Kunz C, Rudloff S, Hintelmann A, Pohlentz G, Egge H (1996) High-pH anion-exchange chromatography with pulsed amperometric detection and molar response factors of human milk oligosaccharides. *Journal of Chromatography B, Biomedical Applications* **685**: 211-221.

Leonard R, Pabst M, Bondili JS, Chambat G, Veit C, Strasser R, Altmann F (2008) Identification of an Arabidopsis gene encoding a GH95 alpha1,2-fucosidase active on xyloglucan oligo- and polysaccharides. *Phytochemistry* **69**: 1983-1988.

Liebert T, Seifert M, Heinze T (2008) Efficient method for the preparation of pure, water-soluble cellodextrines. *Macromolecular Symposia* **262**: 140-149.

Lorences EP, Fry SC (1994) Sequencing of xyloglucan oligosaccharides by partial Driselase digestion: The preparation and quantitative and qualitative analysis of two new tetrasaccharides. *Carbohydrate Research* **263**: 285-293.

Lundborg M, Widmalm Gr (2011) Structural analysis of glycans by NMR chemical shift prediction. *Analytical Chemistry* **83**: 1514-1517.

Lutteke T, Bohne-Lang A, Loss A, Goetz T, Frank M, von der Lieth CW (2006) GLYCOSCIENCES.de: An internet portal to support glycomics and glycobiology research. *Glycobiology* **16**: 71R-81R.

Madson M, Dunand C, Li X, Verma R, Vanzin GF, Caplan J, Shoue DA, Carpita NC, Reiter WD (2003) The MUR3 gene of Arabidopsis encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. *The Plant Cell* **15**: 1662-1670.

Masuko T, Minami A, Iwasaki N, Majima T, Nishimura S, Lee YC (2005) Carbohydrate analysis by a phenol-sulfuric acid method in microplate format. *Analytical Biochemistry* **339**: 69-72.

Michaels AS, Morelos O (1955) Polyelectrolyte adsorption by kaolinite. *Industrial & Engineering Chemistry* **47**: 1801-1809.

Mishra A (2013) Tamarind seed xyloglucan: A food hydrocolloid for water remediation. *Journal of Biobased Materials and Bioenergy* **7**: 12-18.

Mishra A, Malhotra AV (2009) Tamarind xyloglucan: A polysaccharide with versatile application potential. *Journal of Materials Chemistry* **19**: 8528-8536.

O'Neill RA, Albersheim P, Darvill AG (1989) Purification and characterization of a xyloglucan oligosaccharide-specific xylosidase from pea seedlings. *The Journal of Biological Chemistry* **264**: 20430-20437.

Parikka K, Leppanen AS, Xu C, Pitkanen L, Eronen P, Osterberg M, Brumer H, Willfor S, Tenkanen M (2012) Functional and anionic cellulose-interacting polymers by selective chemo-enzymatic carboxylation of galactose-containing polysaccharides. *Biomacromolecules* **13**: 2418-2428.

Pauly M, Andersen LN, Kauppinen S, Kofod LV, York WS, Albersheim P, Darvill A (1999) A xyloglucan-specific *endo*-beta-1,4-glucanase from *Aspergillus aculeatus*: Expression cloning in yeast, purification and characterization of the recombinant enzyme. *Glycobiology* **9**: 93-100.

Pauly M, Eberhard S, Albersheim P, Darvill A, York WS (2001a) Effects of the mur1 mutation on xyloglucans produced by suspension-cultured *Arabidopsis thaliana* cells. *Planta* **214**: 67-74.

Pauly M, Qin Q, Greene H, Albersheim P, Darvill A, York WS (2001b) Changes in the structure of xyloglucan during cell elongation. *Planta* **212**: 842-850.

Peña MJ, Darvill AG, Eberhard S, York WS, O'Neill MA (2008) Moss and liverwort xyloglucans contain galacturonic acid and are structurally distinct from the xyloglucans synthesized by hornworts and vascular plants. *Glycobiology* **18**: 891-904.

Peña MJ, Kong Y, York WS, O'Neill MA (2012) A galacturonic acid-containing xyloglucan is involved in Arabidopsis root hair tip growth. *The Plant Cell* **24**: 4511-4524.

Peña MJ, Zhong R, Zhou G-K, Richardson EA, O'Neill MA, Darvill AG, York WS, Ye Z-H (2007) Arabidopsis irregular xylem8 and irregular xylem9: Implications for the complexity of glucuronoxylan biosynthesis. *The Plant Cell* **19**: 549-563.

Perrin RM, DeRocher AE, Bar-Peled M, Zeng W, Norambuena L, Orellana A, Raikhel NV, Keegstra K (1999) Xyloglucan fucosyltransferase, an enzyme involved in plant cell wall biosynthesis. *Science* **284**: 1976-1979.

Reiter WD, Chapple CC, Somerville CR (1993) Altered growth and cell walls in a fucose-deficient mutant of Arabidopsis. *Science* **261**: 1032-1035.

Rohrer JS, Basumallick L, Hurum D (2013) High-performance anion-exchange chromatography with pulsed amperometric detection for carbohydrate analysis of glycoproteins. *Biochemistry (Moscow)* **78**: 697-709.

Schultink A, Cheng K, Park YB, Cosgrove DJ, Pauly M (2013) The identification of two arabinosyltransferases from tomato reveals functional equivalency of xyloglucan side chain substituents. *Plant Physiology* **163**: 86-94.

Springsteen G, Wang B (2002) A detailed examination of boronic acid–diol complexation. *Tetrahedron* **58**: 5291-5300.

Takahashi N, Kato K (2003) GALAXY (Glycoanalysis by the Three Axes of MS and Chromatography): A web application that assists structural analyses of *N*-glycans. *Trends in Glycoscience and Glycotechnology* **15**: 235-251.

Talmadge KW, Keegstra K, Bauer WD, Albersheim P (1973) The Structure of plant cell walls: I. The macromolecular components of the walls of suspension-cultured sycamore cells with a detailed analysis of the pectic polysaccharides. *Plant Physiology* **51**: 158-173.

Thibault JF (1980) Separation of α-D-galacturonic acid oligomers by chromatography on polyacrylamide gel. *Journal of Chromatography A* **194**: 315-322.

Torrey JG, Shigemura Y (1957) Growth and controlled morphogenesis in pea root callus tissue grown in liquid media. *American Journal of Botany* **44**: 334-344.

van den Brink J, de Vries RP (2011) Fungal enzyme sets for plant polysaccharide degradation. *Applied Microbiology and Biotechnology* **91**: 1477-1492.

Vanzin GF, Madson M, Carpita NC, Raikhel NV, Keegstra K, Reiter WD (2002) The mur2 mutant of *Arabidopsis thaliana* lacks fucosylated xyloglucan because of a lesion in fucosyltransferase AtFUT1. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 3340-3345.

Vargas-Rechia C, Reicher F, Rita Sierakowski M, Heyraud A, Driguez H, Linart Y (1998) Xyloglucan octasaccharide XXLGol derived from the seeds of *hymenaea courbaril* acts as a signaling molecule. *Plant Physiology* **116**: 1013-1021.

Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME (2009) *Essentials of Glycobiology*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Vincken JP, York WS, Beldman G, Voragen AG (1997) Two general branching patterns of xyloglucan, XXXG and XXGG. *Plant Physiology* **114**: 9-13.

Vliegenthart JFG, van Halbeek H, Dorland L (1981) The applicability of 500-MHz high-resolution $^{1}$H-NMR spectroscopy for the structure determination of carbohydrates derived from glycoproteins. *Pure and Applied Chemistry* **53**: 45-77.

Ward OP, Moo-Young M (1989) Enzymatic degradation of cell wall and related plant polysaccharides. *Critical Reviews in Biotechnology* **8**: 237-274.

Widmalm G (2007) General NMR spectroscopy of carbohydrates and conformational analysis in solution. *In:* Kamerling H (*ed*) *Comprehensive Glycoscience*, pp. 101-132. Elsevier, Oxford.

Yamada H, Ohshima Y, Tamura K, Miyazaki T (1980) Separation of a 6-deoxyhexose and a hexose by gel filtration. *Carbohydrate Research* **83**: 377-378.

Yan J, Springsteen G, Deeter S, Wang B (2004) The relationship among p*K*a, pH, and binding constants in the interactions between boronic acids and diols—it is not as simple as it appears. *Tetrahedron* **60**: 11205-11209.

Yaoi K, Mitsuishi Y (2002) Purification, characterization, cloning, and expression of a novel xyloglucan-specific glycosidase, oligoxyloglucan reducing end-specific cellobiohydrolase. *The Journal of Biological Chemistry* **277**: 48276-48281.

Yaoi K, Miyazaki K (2012) Cloning and expression of isoprimeverose-producing oligoxyloglucan hydrolase from *Actinomycetes* species, Oerskovia sp. Y1. *Journal of Applied Glycoscienc*e **59**: 83-88.

Yaoi K, Nakai T, Kameda Y, Hiyoshi A, Mitsuishi Y (2005) Cloning and characterization of two xyloglucanases from *Paenibacillus* sp. strain KM21. *Applied and Environmental Microbiology* **71**: 7670-7678.

York WS, Darvill AG, Albersheim P (1984) Inhibition of 2,4-dichlorophenoxyacetic acid-stimulated elongation of pea stem segments by a xyloglucan oligosaccharide. *Plant Physiology* **75**: 295-297.

York WS, Darvill AG, McNeil M, Stevenson TT, Albersheim P (1986) Isolation and characterization of plant cell walls and cell wall components. *Methods in Enzymology* **118**: 3-40.

York WS, Harvey LK, Guillen R, Albersheim P, Darvill AG (1993) Structural analysis of tamarind seed xyloglucan oligosaccharides using beta-galactosidase digestion and spectroscopic methods. *Carbohydrate Research* **248**: 285-301.

York WS, Impallomeni G, Hisamatsu M, Albersheim P, Darvill AG (1995) Eleven newly characterized xyloglucan oligoglycosyl alditols: The specific effects of sidechain structure and location on $^{1}$H NMR chemical shifts. *Carbohydrate Research* **267**: 79-104.

York WS, Kumar Kolli VS, Orlando R, Albersheim P, Darvill AG (1996) The structures of arabinoxyloglucans produced by solanaceous plants. *Carbohydrate Research* **285**: 99-128.

York WS, Oates JE, van Halbeek H, Darvill AG, Albersheim P, Tiller PR, Dell A (1988) Location of the *O*-acetyl substituents on a nonasaccharide repeating unit of sycamore extracellular xyloglucan. *Carbohydrate Research* **173**: 113-132.

York WS, Pauly M, Qin Q, Jia Z, Simon JP, Albersheim P, Darvill AG (2002) The xyloglucan-cellulse network of plant cell walls: A protytype for the chemoenzymatic preparation of novel polysaccharide composites. *In:* Teeri TT, Svensson B, Gilbert HJ, Feizi T (*eds*) *Carbohydrate Bioengineering: Interdisciplinary Approaches*, pp. 143-150. The Royal Society of Chemistry.

York WS, van Halbeek H, Darvill AG, Albersheim P (1990) Structural analysis of xyloglucan oligosaccharides by $^{1}$H-n.m.r. spectroscopy and fast-atom-bombardment mass spectrometry. *Carbohydrate Research* **200**: 9-31.

Zablackis E, Huang J, Muller B, Darvill AG, Albersheim P (1995) Characterization of the cell-wall polysaccharides of *Arabidopsis thaliana* leaves. *Plant Physiology* **107**: 1129-1138.

Zabotina OA (2012) Xyloglucan and its biosynthesis. *Frontiers in Plant Science* **3**: 134.

Zaia J (2004) Mass spectrometry of oligosaccharides. *Mass Spectrometry Reviews* **23**: 161-227.

# CHAPTER 3

# EPITOPE CHARACTERIZATION OF MONOCLONAL ANTIBODIES WITH DIVERSE XYLOGLUCAN-BINDING SPECIFICITIES [1]

[1] Tuomivaara, S.T., Eberhard, S., Popper, Z., Bootten, T., Baliga, R., Boons, G.J., Hahn, M.G. and W.S. York. To be submitted to *Plant Physiology.*

## 3.1. Abstract

Monoclonal antibodies are emerging as essential tools for probing the composition of plant cell walls by offering high structural specificity and sensitivity in various experimental setups, including immunolocalization and high-throughput glycome profiling. Despite the large number of monoclonal antibodies generated against complex plant cell wall polysaccharides, the paucity of comprehensive epitope information has limited the structural interpretation of the data obtained from their use. To mitigate this shortcoming, we characterized in detail the epitopes of thirty-two monoclonal antibodies that have been previously shown to bind xyloglucans, using a sensitive enzyme-linked immunosorbent assay. Our comprehensive arsenal of structurally diverse xyloglucan derived oligosaccharides allowed us to define the binding specificities in considerable chemical detail and to discern subtle differences that would have been obscured if probed only with polysaccharides or a structurally less diverse oligosaccharide library. Specificities were established for several key xyloglucan features, namely terminal fucosyl, galactosyl, and xylosyl residues. Immunolocalization experiments using these antibodies revealed cell- and tissue-specific localizations of distinct xyloglucan epitopes in *Arabidopsis thaliana* root tissue and *Tamarindus indica* seed. Determination of the binding specificities of these antibodies allows rigorous structural interpretation of these and other immunoaffinity data and facilitates the testing of hypotheses regarding xyloglucan biosynthesis, metabolism and function in cell walls.

## 3.2. Introduction

### 3.2.1. Plant Cell Walls

Plant cell walls are complex extraplasmalemmal organelles with a wide repertoire of structural, regulatory, and metabolic functions in plant cell growth, morphogenesis, and development (Albersheim et al, 2011). Models of the plant cell wall architecture emphasize its multiphasic nature whereby insoluble cellulose microfibrils are embedded in amorphous and hydrated matrix of mostly hemicellulosic and pectic polysaccharides (Baba, 2006; Keegstra, 2010). The chemical and physical heterogeneity of plant cell walls is further compounded by the presence of (glyco)proteins and, in some walls, ossified minerals and extensively cross-linked hydrophobic polymers such as lignin. The proportions of these and other plant cell wall components vary widely and depend on the function and developmental status of the cell (Albersheim et al, 2011). The robustness of plant tissues and ultimately whole plants is largely determined by their cell walls (Gibson, 2012), suggesting a tight chemical and mechanical continuity of the apoplast. Several covalent and non-covalent interpolymer networks have indeed been implicated *in muro* by either direct structural evidence or by extrapolation from their *in vitro* properties (O'Neill et al, 1996; Tan et al, 2013). The archetype of such a network is the cellulose-xyloglucan (XyG) network which assembles *in vitro* and has thus been assigned a major load-bearing function in plants that have XyG as their major hemicellulose (Hayashi & Kaida, 2011). Furthermore, the activities of several apoplastic enzymes and other proteins that are known to remodel XyG and the cellulose-XyG network *in vitro*

have been correlated (e.g., by *in vitro* and *in vivo* stress tests) with cell wall properties in phenomena as disparate as cell wall expansion during growth and disintegration in fruit ripening, suggesting a major regulatory function for this network (Rose & Bennett, 1999). Both *in vivo* and cell culture (Fry et al, 1990; Vargas-Rechia et al, 1998) experiments have demonstrated that XyG oligosaccharides act as signaling molecules, indicating a possible feedback or defense mechanism that utilizes oligosaccharide products of XyG metabolism in the cell wall. The full spectrum of the functions as well as the essentiality of XyG are controversial, not least due to a viable *Arabidopsis thaliana* mutant that lacks XyG (Cavalier et al, 2008). Notably, XyGs have been found in the cell walls of all embryophytes (land plants) examined to date (Popper et al, 2011) and large families of enzymes and other proteins have been implicated in XyG metabolism (Rose et al, 2002), indicating a strong evolutionary pressure for maintaining XyG in the plant cell walls.

### 3.2.2. XyG Structure

XyGs have a linear cellulosic backbone of β(1→4)-linked D-Glc*p* residues that are substituted with side-chains at *O*-6. Currently nearly twenty side-chain structures have been described, but only a few can be found in any single species. The growing assortment of described side-chain structures led to the introduction of a single-letter code that can compactly describe XyG sequences of arbitrary length (Fry et al, 1993). Here, G denotes a backbone glucosyl residue without decoration at *O*-6, whereas X, L and F denote glucosyl residues with α-D-Xyl*p*, β-D-Gal*p*-(1→2)-α-D-Xyl*p*, and α-L-Fuc*p*-(1→2)-β-D-Gal*p*-(1→2)-α-D-Xyl*p*

side-chains at *O*-6, respectively. Structures and nomenclature of all side-chains relevant to this work are presented in Table 2-1.

It is noteworthy that several plant species synthesize XyGs side-chains that differ in their sugar residue composition but are stereochemically identical (for detailed discussion of XyG stereochemistry, see Chapter 2). There are several known examples of this type of structural heterogeneity in XyG. Jojoba (*Simmondsia chinensis*) seed XyG contains both the F [α-L-Fuc*p*-(1→2)-β-D-Gal*p*-(1→2)-α-D-Xyl*p*], and J [α-L-Gal*p*-(1→2)-β-D-Gal*p*-(1→2)-α-D-Xyl*p*] side-chains (Hantus et al, 1997). Other instances, whereby a particular XyG polysaccharide contains stereochemically identical side-chains differing by a single small chemical group, have been described in the literature. In a clubmoss *Selaginella kraussiana* and a horsetail *Equisetum hyemale* that were reported to contain, besides the canonical L and F side-chains, α-L-Ara*p*-(1→2)-α-D-Xyl*p* (termed D), and α-L-Fuc*p*-(1→2)-α-L-Ara*p*-(1→2)-α-D-Xyl*p* (termed E) structures (Peña et al, 2008). A root hair specific anionic XyG that contains (in addition to the X, L and F side-chains) β-D-Gal*p*A-(1→2)-α-D-Xyl*p* (termed Y) and α-L-Fuc*p*-(1→2)-β-D-Gal*p*A-(1→2)-α-D-Xyl*p* (termed Z) side-chains was reported in *Arabidopsis* (Peña et al, 2012). This aspect of glycan biosynthesis and biochemistry, where distinct glycosyl residues with the same stereochemistry are transferred to glycan structures and may confer similar functional properties to the glycan, is largely restricted to XyGs. Rhamnogalacturonan-II (RG-II) found in the plant cell walls is another complex carbohydrate in wild-type plants that has been shown to display this type of structural heterogeneity (Pabst et al, 2013).

Besides the abovementioned XyG side-chains, several others whose distribution in the plant kingdom is typically limited, have also been reported (Hilz et al, 2007; Hisamatsu et al, 1992; Peña et al, 2008; York et al, 1996).

XyG structure differs from other plant cell wall polysaccharides in other ways as well. In general, the distribution of the Xyl*p* residues on the backbone residues is not statistical (as is typical for side-chains of other plant cell wall polysaccharides) but very regular, leading to polysaccharide that is composed of apparent subunits. The regularity of the side-chains and the subunit-like structure become apparent when the XyG polysaccharide is hydrolyzed by a XyG-specific endoglucanase (XEG) that exclusively hydrolyzes the glycosidic bond of the unsubstituted backbone Glc*p* residue. For example XEG hydrolysis of XyGs from most dicotylenous plants yields oligosaccharides with XXXG core structure (where some of the Xyl*p* residues can be extended by other glycosyl residues). Several side-chain patterns have been described, that can be used for the classification of XyGs (Buckeridge et al, 1997; Hsieh & Harris, 2009; Vincken et al, 1997; York et al, 1996). Besides glycosyl residues, *O*-acetyl substitutions have been found on XyG side-chains (Hoffman et al, 2005; Lerouxel et al, 2002; York et al, 1988) and on *O*-6 of the backbone residues of XXGG-type XyGs (York et al, 1996) where they apparently replace Xyl*p* residues that are found at this position in XXXG-type XyGs. The structure and abundance of XyG varies tremendously, not only among species and higher levels of plant taxonomy, but according to the types and developmental stages of cells and tissues, and even among different regions of an individual cell wall. The functional significance of

the XyG diversity is currently not well understood, but at least in some cases the different side-chain structures seem to be functionally redundant (Schultink et al, 2013).

### 3.2.3. Monoclonal Antibodies as XyG Probes

The recalcitrance of plant cell walls poses formidable obstacles for detailed scrutiny, and XyGs, which include some of the most structurally complex and tightly integrated components of plant cell walls, epitomize these analytical challenges. Monoclonal antibodies (mAbs) can facilitate a new approach in probing the structure and abundance of XyGs and other complex polysaccharides. The principal advantage of mAbs is their ability to distinguish between unreactive and chemically similar carbohydrate structures in complex and heterogenous samples such as the plant cell wall. Additionally, mAb-based recognition can be coupled to several sensitive detection systems as well as various microscopy modalities that allow detection at sub-micrometer resolution on tissue samples. The utility of mAbs in probing cell wall polysaccharides has been demonstrated in evolutionary (Brennan & Harris, 2011; Kulkarni et al, 2012), developmental (Nishikubo et al, 2011; Vaughn et al, 2007), environmental (Balestrini et al, 1996), and biotechnological (DeMartini et al, 2013) contexts. Although the application of mAbs in providing an alternative view on the plant cell wall structure, ultrastructure, and composition is well established (Knox, 2008; Lee et al, 2011; Pattathil et al, 2010), the scarcity of detailed epitope information has thus far severely restricted the extent of structural conclusions that can be drawn from their use. Here, we report the refined epitope characterizations of

thirty-two previously described XyG-binding mAbs (Marcus et al, 2008; Pattathil et al, 2010; Pedersen et al, 2012; Puhlmann et al, 1994). Our results revealed a spectrum of complementary binding specificities among the mAbs that allow the most common structural features in XyGs to be distinguished. Since the extent of structural interpretations drawn from immunoaffinity experiments depend on the rigor of the epitope characterization, our results provide a basis for comprehensive immunoaffinity analysis of XyG structure, quantity and function. The approach we have developed provides a blueprint for epitope characterization efforts of other complex polysaccharides.

### 3.3. Results and Discussion

Thirty-two mAbs that have previously been generated using XyGs and other plant cell wall polysaccharides, or XyG oligosaccharides, as immunogens, and that are known to bind XyGs, were selected for detailed epitope characterization by enzyme-linked immunosorbent assay (ELISA). The epitopes of these mAbs were refined using XyG polysaccharides as well as a large library of XyG-derived oligosaccharides as probes. For the ELISAs, the polysaccharides were immobilized on microplate wells by drying, whereas the oligosaccharides were biotinylated and immobilized on NeutrAvidin-coated plates. Both direct and competitive ELISAs were utilized for discerning the binding patterns of the mAbs. The binding patterns were analyzed by hierarchical clustering to group the mAbs into recognition groups.

### 3.3.1. Immunogens

XyG containing immunogens were prepared by first reacting the XyG polysaccharide with 1-cyano-4-dimethylaminopyridinium tetrafluoroborate (CDAP) and then conjugating the activated XyG polysaccharide to bovine serum albumin (BSA). Immunogens containing tamarind, sycamore maple, and tomato XyGs were prepared as covalent BSA conjugates. Non-XyG polysaccharide containing immunogens were prepared as non-covalent mixtures with BSA. Whereas the non-covalent polysaccharide-BSA mixtures contain predetermined 1:1 (w/w) ratio of the components, the relative amounts of the components in covalent polysaccharide-BSA conjugates depend on several variables, including the efficacy of the conjugation reaction and the average size of the polysaccharide molecules (data not shown).

### 3.3.2. Generation and Screening of Hybridoma Lines

The diluted sera of the immunized animals were tested using ELISA for the presence of antibodies that bind underivatized and unconjugated forms of the same polysaccharide that was used for the immunization. The number of mice that generated an immune response varied depending on the immunogen (Supplemental Table 1). Hybridoma lines generated from the splenocytes of these mice were screened by ELISA using same polysaccharides previously used to test the sera. Detailed statistics of hybridoma line generation are also presented in Supplemental Table 1. These data include the number of mice that generated a strong immune response to each immunogen. These mice were

chosen for splenectomy and hybridoma line generation. Also listed is the number of hybridomas from each independent fusion that grew and tested positive by ELISA for polysaccharide binding Abs. Further passaging and cloning by limiting dilution resulted in stable hybridoma lines that continued to display a high titer of mAbs. These were used to produce large quantities of hybridoma supernatants that were stable through cryopreservation and thawing. The name, designation isotype and light-chain for each of the selected hybridoma lines are also listed in Supplemental Table 1.

### 3.3.3. mAb Binding to Immobilized Oligo- and Polysaccharides

The binding of each mAb to chemically homogeneous XyG and cellodextrin oligosaccharides (for their preparation and structural validation, see Chapter 2) or their mixtures, was determined by ELISA. The APB derivatized oligosaccharides were purified by SEC and the yields were in almost all cases over 90% (Supplemental Figures 1, 3, 4 and 5). The *m*/*z* values of the oligosaccharide-APB conjugates in MALDI-TOF MS spectra conformed to the expected values in all cases. Owing to the purity of the starting materials, as well as the mild derivatization conditions, the end products were homogenous as well, as is evident from their MALDI-TOF MS spectra. These well-defined oligosaccharides were immobilized to individual wells of streptavidin-coated ELISA plates after conjugating them to an aminooxy-$PEG_4$-biotin (APB) tag using aniline catalyst. These plates were then used to characterize each mAb by ELISA. The purity (typically > 98%) of these oligosaccharides made it possible to accurately discern the binding pattern for each of the mAbs tested.

The structural diversity of the oligosaccharide collection, and the chemical homogeneity of the individual oligosaccharides that were attached to ELISA plates, made it possible to define detailed binding specificities for each of the anti-XyG mAbs (Figure 1 and Supplemental Figures 6 and 8). The mAbs show substantial differences in their binding patterns when tested using this oligosaccharide library. That is, the mAbs are capable of distinguishing specific structural features of XyGs, including the presence and absence of structurally diverse side chains, making them useful tools for determining how these structures vary in different cells, tissues and species.

*Terminal α-L-Fuc*p *Residue-binding mAbs* Five mAbs, namely CCRC-M1, CCRC-M39, CCRC-M84, CCRC-M102, and CCRC-M106, specifically bind to oligosaccharides bearing a fucosylated side chain (either F or Z side-chain. Oligosaccharide containing the E side-chain were not available for testing). These mAbs can be further classified according to how they bind to fucosylated oligosaccharides with distinct structures. CCRC-M1 and CCRC-M102 bind to all seven prepared oligosaccharides with an F side-chain (FG, GXFG, XXFG, GLFG, XLFG, GLFG, XFFG, see Figure 1, Supplemental Figures 6A, 6Y and 8), as well as the Z oligosaccharide. Whereas CCRC-M102 binds to the Z side-chain containing oligosaccharide with nearly the same affinity as it does to its analog with the F side-chain, CCRC-M1 can distinguish between these two structures and has much higher affinity towards the F side-chain.

The smallest fucosylated oligosaccharide recognized by the CCRC-M39 and CCRC-M106 mAbs is GXFG (Figure 1, Supplemental Figures 6B and 6AB).

The affinity of these two mAbs to the other fucosylated oligosaccharides increases slightly if either or both of the terminal residues in the GX substructure are extended, as in XXFG, XLFG and GLFG. CCRC-M84 selectively binds a very rare XyG oligosaccharide that contains two adjacent F side-chains, and thus has the most stringent binding constraint of the mAbs in this clade (Figure 1, Supplemental Figure 6M). CCRC-M84 has slightly stronger binding to XFFG compared to GFFG. It is noteworthy that all five mAbs in this clade discriminate between the stereochemically identical F and J side-chains, which are both terminated by a sugar having the L-*galacto* configuration. The L-galactosyl residue at the end of the J side-chain has a hydroxyl group at *C*-6, whereas the L-fucosyl residue at the end of the F side-chain does not (with proton instead of hydroxyl group). mAbs in this clade do not bind oligosaccharides that lack an F side-chain, even if they contain a J side-chain. The absence of this hydroxyl group is thus a required structural element of the epitope recognized by these mAbs. ELISAs using the F side-chain containing oligosaccharides against dilution series of the immobilized oligosaccharides were performed as well (Supplemental Figures 6A, 6B, 6M, 6Y and 6AB). In all cases, the results mirror those from the single datapoint ELISAs. Additional experiments that would further narrow the epitope of these experiments include testing against larger Z side-chain containing XyG oligosaccharides, as well as various oligosaccharides with the E side-chain. CCRC-M39, CCRC-M106, and CCRC-M84 appear to be XyG-specific mAbs, since they require structural features for binding that are only found in XyGs.

ELISAs using oligosaccharides and polysaccharides provide complementary information on the binding specificities of mAbs. Whereas ELISAs using oligosaccharides provide detailed information on the minimum epitope, polysaccharide ELISAs, provide information on the binding to the epitope in its natural environment within the polysaccharide. Since XyG polysaccharides contain an array of epitopes, at least some of the specificity information is lost in this approach. The binding specificities of nearly all mAbs described here have previously been tested by ELISA against a large panel of plant cell wall derived polysaccharides, including tamarind, sycamore maple and tomato XyGs, as well as large number of non-XyG polysaccharides (Pattathil et al, 2010). We performed the binding assays of XyG-binding mAbs to a larger panel of XyG polysaccharides in order to identify finer differences in the binding patterns among the mAbs.

Two types of ELISAs were utilized: (1) Direct, with a single datapoint, for screening, and (2) competitive, with a dilution series of the competing polysaccharide. The information content of a single-datapoint ELISA for the purposes of epitope characterization is very limited since the absorbance reading is a function of affinity as well as concentrations of the epitope and the mAb. Thus, without a concentration series of either of the hybridoma supernatant or the immobilized polysaccharide, differences in the affinities are obscured and can be misleading. On the other hand, the single datapoint ELISA provides high-throughput platform to screen mAb-polysaccharide binding since a large number of mAb-polysaccharide combinations can be tested on a single microplate.

Supplemental Figure 7 shows the combined results from the single-datapoint ELISAs against immobilized polysaccharide in the form of a data matrix. The same data is presented in Supplemental Figure 6 for each mAb. Competitive ELISAs were performed by coating microplate wells with the same polysaccharide that was used in the immunization of a given mAb, and incubating a mAb-competitor mixture, instead of the mAb alone, in the coated microplate well. This approach can be used to test the relative affinities of competing polysaccharides.

As expected the F side-chain recognizing mAbs do not bind tamarind, tomato, or tobacco XyGs which do not contain this side-chain in measurable quantities (Supplemental Figures 6 and 7). On the contrary, sycamore maple and jojoba XyGs, which are known to contain F side-chains, bind these mAbs with high affinity. Chemical analyses of water-extracted jatoba seed XyG have not revealed the presence of the F side-chain (Buckeridge et al, 1997; Lima et al, 1995; Tiné et al, 2006). The binding of mAbs that specifically recognize the F side-chain suggests that very small amounts of the F side-chain bearing XyG is however present in this polysaccharide preparation. It is likely that the water extraction also releases small amounts of structural XyG that are typically fucosylated.

Since CCRC-M1 and CCRC-M102 bind to all tested XyG oligosaccharides with α-L-Fuc*p*-(1→2)-β-D-Gal*p* structure, these mAbs potentially bind other plant cell wall components with this structure, termed “H-antigen” (Hosoi, 2008), as well. Indeed, CCRC-M1 was shown to react with sycamore maple

rhamnogalacturonan-I (RG-I) (Puhlmann et al, 1994) that contains the α-L-Fuc*p*-(1,2)-β-D-Gal*p* epitope (Lau et al, 1987). However, CCRC-M102 does not bind XEG treated sycamore maple RG-I preparation (Supplemental Figure 6Y).

*Terminal* β-D-*Gal*p, β-D-*Gal*p*A and* α-L-*Ara*p *Residue-binding mAbs* Nineteen mAbs bind only to oligosaccharides having a side-chain (L, D, or Y) that is terminated by a glycosyl residue (D-Gal*p*, D-Ara*p* or D-Gal*p*A, respectively) with the D-*galacto* configuration (Figure 1 and Supplemental Figures 6 and 8). Most of these mAbs do not discriminate between these stereochemically identical side-chains. Nevertheless, at least four sub-specificities can be discerned among these mAbs. CCRC-M87, CCRC-M88 CCRC-M93 and CCRC-M104 bind to all tested L, D, and Y side-chain containing oligosaccharides (Supplemental Figures 6O, 6P, 6S and 6AA). CCRC-M48 is regiospecific, requiring either an L or D side-chain on the branched backbone residue closest to the reducing end of the subunit (Supplemental Figure 6C). CCRC-M51, CCRC-M99 and CCRC-M58 require two adjacent side-chains terminated with a D-*galacto* residue. That is, CCRC-M51 (Supplemental Figure 6F) and CCRC-M99 (Supplemental Figure 6V) require an LL or DD motif for binding, whereas CCRC-M58 (Supplemental Figure 6L) binds only to the DD motif.

For all of the L side-chain binding mAbs tested, the binding affinity is reduced when tamarind XyG polysaccharide is partially degalactosylated, but the degree of reduction varies. The degree of reduction appears to be proportional to the specificity of the mAb towards the structural context of the L side-chain. Thus, the reduction of the binding is highest in the mAbs that have the most stringent

requirements for binding. CCRC-M57 does not bind the partially degalactosylated XyG polysaccharide. This supports the data from the oligosaccharide ELISAs that the junction between subunits is not sufficient for binding for CCRC-M57. This mAb has high affinity towards oxidized tamarind XyG polysaccharide, indicating that it cannot discriminate effectively between the stereochemically identical L and Y side-chains.

A total of eight tested mAbs (for example CCRC-M87) that bind with high affinity to native tamarind XyG polysaccharide, show also strong binding to tamarind XyG polysaccharide whose L side-chains have been oxidized to Y side-chains with approximately 90% efficiency. The oxidize XyG thus contains mostly XXXG, XYXG, XXYG and XYYG subunits (Figure 1, Supplemental Figures 6 and 8). These mAbs are collectively called XyG-LY. These data indicate that these mAbs do not effectively discriminate between the stereochemically identical L and Y side-chains. However, ten of the tested mAbs that bind with high affinity to native tamarind XyG polysaccharide show practically no binding to the oxidized XyG polysaccharide.

*Terminal α-D-Xyl*p *Residue-binding mAbs* Three mAbs, namely CCRC-M86 (Supplemental Figure 6N), CCRC-M100 (Supplemental Figure 6W), and CCRC-M103 (Supplemental Figure 6Z) bind exclusively to the XyG S1 oligosaccharides whose Xyl*p* residues are not further extended. These two mAbs bind with the highest affinity to the GXXXG oligosaccharide. Surprisingly, these mAbs bind the oxidized tamarind XyG polysaccharide only weakly, even though the principal epitope is presumably not chemically affected by the oxidation.

According to the ELISA experiments, these three mAbs show weak binding to the XyG oligosaccharides and it is possible that the oxidation of the L side-chains perturbs the conformation of the XyG polysaccharide enough to fully prevent binding. This evidence, although not conclusive, can be interpreted that these mAbs bind to the junction between the subunits in native XyG polysaccharides.

*Larger XyG-binding mAbs* CCRC-M49, CCRC-M54, and CCRC-M57 do not bind to the tamarind XyG S1 mixture or to any of the purified XyG S1 oligosaccharides, whereas all of them bind to larger S2, S3 and S4 XyG oligosaccharides, as well as various XyG polysaccharides (Figure 1, Supplemental Figures 6D, 6I, and 6K). In all three cases the absorbance reading increases monotonically as the XyG size increases from S2 to non-hydrolyzed tamarind XyG polysaccharide (with undefined size). At least two alternative explanations are equally supported by these data. First, the minimum epitope recognized by these mAbs can flank the junction between the subunits (for example the underlined part in XXXGXXXG) that is present only in polysaccharides and oligosaccharides with two or more subunits. Second, the microscopic affinity of individual binding site of the mAb towards its epitope (for example the underlined part in XXXG) can be too weak to result in observable binding to oligosaccharide(s) with a single epitope, whereas the avidity of the mAb towards polysaccharides and oligosaccharide with two or more epitopes (for example, as in XXXGXXXG) results in tight and observable binding in ELISA. It is noteworthy that the binding of CCRC-M57 to the S2 oligosaccharides can be barely detected, indicating that the presence of the inter-subunit junction alone is

perhaps not enough for high affinity binding. The CCRC-M49, CCRC-M54, and CCRC-M57 mAbs can thus be considered a XyG-specific mAbs. Further experiments are needed to resolve the exact epitopes.

*LM-series mAbs* LM15 mAb binds with high affinity to all XyG oligosaccharides (except for GXG) that have either GX or XX at the non-reducing end (Figure 1, Supplemental Figure 6AF). GXXG is the simplest oligosaccharide with three intact Glc*p* residues in the backbone that binds LM15. On this GXXG core structure, an L side-chain on the underlined ("middle") position G<u>X</u>XG generally inhibits binding, whereas extension of the non-reducing end Glc*p* to X or reducing end Xyl*p* to L, F, or J side-chain have minor effect on the LM15 binding if the middle Glc*p* only carries a Xyl*p* extension. Notably, either an F or J side-chain on the Glc*p* adjacent to the reducing end can rescue the binding to an oligosaccharide bearing an L side-chain in the middle position.

The following affinity series (in the order of increasing binding, with underlined residues indicating structural feature that is different from GXXG) illustrates the binding requirements of LM24: GXXG (no binding), < G<u>X</u>FG < <u>X</u>XXG < <u>X</u>X<u>F</u>G < G<u>L</u>XG (Figure 1, Supplemental Figure 6AG). Thus, the minimum binding requirement for LM24 is a built on GXXG core with at least one of the terminal residues being extended, either by a single sugar in the case of G<u>L</u>XG and <u>X</u>XXG, or a disaccharyl unit in the case of G<u>X</u><u>F</u>G. Maximum binding is achieved with the G<u>L</u>XG extension alone. Two other binding determinants, for example in <u>X</u>X<u>F</u>G, combine for strong binding.

Initial ELISAs using the LM15 and LM24 with the NeutrAvidin coated and pre-blocked microplates resulted in high background absorbance in the no-oligosaccharide-APB coating control wells (data not shown). Blocking the wells for 1 h with 1 mg/mL BSA in water either before or after coating the wells, or with 10 mg/mL non-fat milk powder in TBS after coating the wells with oligosaccharide-APB conjugate was sufficient to prevent the background absorbance (Data not shown. Blocking with milk powder prior to the oligosaccharide-APB coating was not attempted since the milk powder can contain biotin). We eventually pinpointed this high background to the binding of the secondary antibodies to the NeutrAvidin coated wells. Since the CCRC-M hybridoma supernatants were used undiluted, we surmised that the bovine serum proteins in them acted as blocking agents and no background absorbance was observed. On the other hand, in the LM hybridoma supernatants were used as 1:20 dilutions in Tris-buffered saline and the final bovine serum protein concentration was not high enough to act as a blocking agent. Thus, 1 mg/mL BSA blocking prior to the oligosaccharide-APB coating was used for the ELISAs using the LM mAbs. ELISAs using the CCRC-M mAbs were performed without extra blocking.

*Effect of* O*-acetylation on mAb Binding* We performed ELISAs against both the native and de-*O*-acetylated XyG polysaccharides from sycamore maple, tomato, and tobacco, in order to determine whether the prevalent *O*-acetyl groups in these XyG polysaccharides affect the mAb binding. We observed a general decrease (globally ~15%) in the ELISA responses of the mAbs to the

chemically de-*O*-acetylated compared to the natively *O*-acetylated polysaccharide (Supplemental Figures 6 and 8). The reduction may reflect differences in the three-dimensional conformation, and thus side-chain presentation, of *O*-acetylated and non-acetylated XyGs when bound to solid surfaces. We exclude the lower adherence of the de-*O*-acetylated XyG polysaccharides to the polystyrene plate surface since the reduction in the ELISA responses varies dramatically among the mAbs, and for several mAbs the ELISA response is unchanged. Further, some mAbs showed a slight increase in the ELISA response. A single mAb, CCRC-M57, showed no binding against natively *O*-acetylated tobacco XyG, but showed high binding against de-*O*-acetylated tobacco XyG. This mAb binds strongly to the non-acetylated XyG oligo- and polysaccharides, indicating that the epitope does not contain an *O*-acetyl group. Rather this observation might be explained either by differential conformation of the *O*-acetylated and non-acetylated XyGs, or other steric reasons. The de-*O*-acetylation of the polysaccharides by alkaline treatment was ascertained by MALDI-TOF MS analysis of XEG hydrolyzed XyG samples (data not shown).

*Other Oligosaccharides* None of the tested mAbs bind to cellodextrins (Figure 1, Supplemental Figure 6), indicating that at least some canonical XyG side-chain features are required for binding. Furthermore, none of the mAbs described here bind to terminal D-Gal*p* and D-Fuc*p* residues containing oligosaccharides lacto-*N*-tetraose (L*N*T), lacto-*N*-*neo*tetraose (L*Nn*T), lacto-*N*-fucopentaose III (L*N*FPIII), and β-D-Gal*p*-(1→3)-β-D-Gal*p*NAc-(1→4)-β-D-Gal*p*-

(1→4)-D-Glc (data not shown) that are typically found on glycoproteins or glycolipids.

*Binding of mAbs to non-XyG Polysaccharide Preparations* The published ELISA experiments revealed binding of several of the XyG-directed mAbs to various polysaccharide preparations from distinct sources, compositions, and sample histories (Pattathil et al, 2010). These polysaccharide preparations include in-house prepared tomato glucomannan, sycamore pectic polysaccharide, tomato pectic polysaccharide, as well as commercially available gum guar, locust bean gum, lupin galactan, gum ghatti and gum tragacanth preparations. We treated these polysaccharide preparations with XEG to determine whether the reactivities are due to a shared epitope or the presence of contaminating XyG polysaccharide. With the exception of the sycamore pectic polysaccharide preparation, XEG treatment abolished completely or nearly completely the binding of all XyG-reactive mAbs to “non-XyG” polysaccharide preparations. Nearly all XyG binding mAbs react with the pectic polysaccharide preparation from sycamore maple. However, this binding vanishes for all but two mAbs (CCRC-M1 and LM15, see Supplemental Figures 6 and 7) if the pectin preparation is XEG treated before the ELISA. We are currently investigating the possible sources of the affinity of these two mAbs to pectin. It is possible that CCRC-M1 binds to a terminal α-L-Fuc*p*-(1→2)-β-D-Gal*p* epitope that is present in other cell wall components.

Larch arabinogalactan (Odonmažig et al, 1994) and tobacco galactoglucomannan (Eda et al, 1985; Sims et al, 1997) have been shown to

contain terminal β-D-Gal*p* residues. However, none of the XyG-binding mAbs described here bind these polysaccharides ((Pattathil et al, 2010), Supplemental Figures 6 and 7). These data collectively support the XyG specificity of the β-D-Gal*p* residue-binding mAbs characterized in this work.

In conclusion, the thirty-two mAbs can be classified in recognition groups that share an essential feature of XyG structure. The epitopes include terminal xylosyl, galactosyl, and fucosyl residues and thus span large portion of the common XyG structures.

*The Utility of ELISA Approach* It was shown by (Pattathil et al, 2010) using ELISA that both neutral and anionic polysaccharides with diverse structures adhere to the wells of polystyrene microplates upon drying from aqueous solutions. Complex polysaccharides typically contain a mixture of epitopes in variable proportions, preventing detailed epitope characterization using this approach. Additionally, polysaccharide extracts from plant cell walls may contain more than one type of polysaccharide, thus compounding the uncertainty in identifying the epitope. Epitope characterization using oligosaccharides provides potentially superior way to probe the binding specificities of mAbs compared to complex polysaccharide probes. We opted to use chemical conjugation of pure oligosaccharides to APB tag for their immobilization onto Neutravidin coated microplate wells for several reasons. First, the oligosaccharide-APB conjugates can be obtained in their pure forms by chromatographic separation. Second, Neutravidin has very strong affinity for biotin and the interaction can be considered to permanently immobilize the

oligosaccharide to the microplate surface, regardless of its size or structure. Third, using a biotinylation agent with a tetra-(ethylene glycol) ($PEG_4$) linker (approximate length of 27 Å) between the biotin moiety and the oligosaccharide allows good accessibility of mAbs to the oligosaccharides. Fourth, the microplates used have a reproducible biotin-binding capacity per well and can be used to immobilize a constant molar amount of oligosaccharides for facile comparison of their affinities to mAbs. This approach contrasts that used by Knox and colleagues whereby oligosaccharides are reductively aminated to a protein carrier and then printed on a glass slide in a microarray format (Pedersen et al, 2012). As reported by the authors, this approach suffers from variable conjugation efficiency of oligosaccharides to the protein carrier, and if uncorrected by adjusting the immobilized protein amount, can lead to non-quantiative binding patterns.

Our strategy (using polysaccharide immunogens that are chemically distinct from the immobilized oligosaccharides used for the epitope characterization) has distinct advantages over the alternative (using covalent oligosaccharide-protein conjugates for both stages (Pedersen et al, 2012)). First, the polysaccharide immunization produces mAbs with higher specificity towards side-chain patterns. This is evident from the binding patterns of the mAbs against the oligosaccharide panel. Second, the amounts of immobilized oligosaccharide can be controlled accurately for quantitative comparison of the binding affinities of the mAbs against different oligosaccharides.

Detailed annotation of the mAbs described here can be found in a database accessible on the Internet (*http://www.WallMabdb.net*). The mAbs in CCRC series are available to the cell wall research community from CarboSource (*http://www.CarboSource.net*).

### 3.3.4. Hierarchical Clustering of Oligosaccharide-ELISA Data

Examination of the ELISA data with immobilized oligosaccharides reveals the presence of groups of mAbs that have similar epitopes. Supplemental Figure 8 shows a data matrix of the ELISA data where both the oligosaccharides and mAbs are manually organized. The oligosaccharides are arranged according to the structure (first into groups with the same characteristic side-chain, and further according to their size within the groups). The mAbs are arranged according to their binding patterns against the oligosaccharides.

More quantitative analysis of the relationships in both dimensions was performed using hierarchical clustering that groups both oligosaccharides and the mAbs according to the distances between the corresponding absorbance vectors (Figure 1). Dendrograms that show the clustering relationship in hierarchical manner are included for reference. All major groups of mAbs that differ by the requirement of the side-chain structures (the X, L, and F clusters) are strongly clustered, as indicated by the higher level parting of the main branches (Figure 1). The clustering analysis provides a quantitative way to group the mAbs with similar binding patterns and allows rationality in choosing mAbs for various immunoaffinity experiments.

Equally interesting is the clustering of the oligosaccharides. The oligosaccharides form three major clades whose members might have been predicted from their chemical structures and the manual grouping (Supplemental Figure 8). For example, the first major clade (from GXG to XLLG) consists mostly of oligosaccharides with either GL or XL motif in the non-reducing end. Another homogenous main clade (from GXXXG to XXLG) is formed mostly by oligosaccharides with GX or XX motif in the non-reducing end. The clustering of oligosaccharides is informative because it indicates the degree similarity of the epitopes when probed by this set of mAbs.

### 3.3.5. Immunofluorescence Localization

*Arabidopsis root tissue* Six of the characterized mAbs were utilized to label transverse sections of wild type and a *mur3* mutant *Arabidopsis thaliana* plants (see Figure 2). Mur3 encodes a galactosyltransferase that adds the β-D-Gal*p* residue exclusively to the underlined position in XXXG to form XXLG, but not to other positions in the XyG polysaccharide (Madson et al, 2003). CCRC-M1 and CCRC-M102 which bind to F and Z side-chains in XyGs principally label the epidermal layer in the wild type Arabidopsis, with the weakest labeling in the cortical layer (Figure 2). Based on the known binding specificities of these mAbs, these data indicate that either F or Z or both side-chains are present in these tissues. The presence of both side-chains has been reported earlier in the Arabidopsis root XyG (Peña et al, 2012). Moreover, Peña and colleagues established by genetic and biochemical approaches that the Z side-chain occurs

only in the root hair of Arabidopsis, whereas the F side-chain is ubiquitously present in the root tissue.

XXFG and XLFG are the only F side-chain containing subunits observed thus far in Arabidopsis. Immunolabeling of mur3-3 plants using CCRC-M102 (whose affinity to the Z side-chain is comparable to that of the F side-chain), and CCRC-M1, selectively labels the root hairs but not the root body. Chemical analysis of the root hairs of murus3-3 plant by MALDI-TOF MS and nuclear magnetic resonance spectroscopy are in agreement with the immunofluorescence data, indicating the usefulness of mAbs in analysis of plant cell wall composition.

*Tamarind seed* Chemical analyses (glycosyl residue and mass spectrometry) of water-extractable XyG from tamarind seed have established that only non-fucosylated subunits (XXXG, XLXG, XXLG and XLLG) are present in detectable quantities. The accumulation of this type of XyG polysaccharide in the cell walls of tamarind seed is evident from the staining pattern of CCRC-M48, CCRC-M49, CCRC-M57 and CCRC-M58 that bind to non-fucosylated XyG oligo- and polysaccharides (Figure 3). CCRC-M1 that requires fucosylated side-chain for binding labels strongly the epidermal layer of the seed and to a lesser extent the cell walls within the seed. Thus, tamarind seed contains fucosylated XyG as well that has not been detected by using conventional chemical analyses.

## 3.4. Conclusions

To summarize, we have carefully characterized the epitopes of thirty-two XyG-binding mAbs using a large library of pure XyG-derived oligosaccharides as well as XyG polysaccharides as probes. The binding specificities found among the mAbs span the most common structural features found in XyG polysaccharides, ranging from the core XyG structures to specific side-chain patterns. The plant cell wall community has generated and utilized a large number of mAbs against a variety of plant cell wall polysaccharides but in most cases the exact epitopes have not been elucidated. Our results thus shed new light on the enigmatic structure-function relationships of XyGs in plant cell walls.

## 3.5. Materials and Methods

### 3.5.1. Plant Polysaccharides

Native[1], partially degalactosylated, and chemoenzymatically oxidized tamarind XyGs were generated and purified as described in Chapter 2. Jojoba seed XyG was purified as described (Hantus et al, 1997). Jatoba XyG was a generous gift from Dr. Marcos Buckeridge. Lupin galactan was obtained from Megazyme. Gum ghatti, gum, guar, gum tragacanth, and locust bean gum were obtained from Sigma-Aldrich.

[1]Native tamarind XyG polysaccharide preparation for immunization purposes was purified by Zoë Popper at CCRC.

*Sycamore Polysaccharides for Epitope Characterization* Sycamore maple (*Acer pseudoplatanus*) XyG was prepared essentially as described (York et al, 1986). The cells from a starter culture (kindly provided by Stefan Eberhard at CCRC) were cultured ten days in modified Torrey and Shigemura M-6 medium (Torrey & Shigemura, 1957) as described (Talmadge et al, 1973), and harvested by vacuum filtering the cell suspension with a 50 µm Nylon net. The cell cake was washed with small aliquot of water, vacuum aspirated until water draining stopped, and frozen until analysis (plant cell walls are a rich source of polysaccharides but the cells were not utilized in the work described here). Polymeric material ("extracellular polysaccharides") was recovered from the spent medium by concentrating the culture filtrate by rotary evaporation to approximately one fifth of the original volume, followed by sequential ethanol precipitation and dialysis steps. Ethanol precipitation was performed by gently mixing the pre-cooled sample with three volumes of pre-cooled absolute ethanol and incubating overnight at 4 °C. The precipitated material was pelleted by centrifugation (10000*g* for 30 min at 4 °C), redissolved in water and dialyzed. Dialysis was performed against water in regenerated cellulose tubing with 3000 molecular weight cutoff (Spectra/Por) at 4 °C, after which the retentate was lyophilized. Neutral and anionic fractions were purified from the lyophilizate by sequential strong anion and cation exchange chromatographies. Q Sepharose FF anion exchange resin (GE Healthcare) was packed into a column and conditioned by sequential washes with three bed volumes of water, 1 M imidazole, pH 7.0; and finally 10 mM imidazole, pH 7.0 buffer. XyG

polysaccharide preparation was dissolved in 10 mM imidazole, pH 7.0 buffer to a final concentration of 5 mg/mL and top-loaded on the resin (up to 5 mL per 1 mL of resin). The resin was then washed with three bed volumes of the same buffer. Load and wash flow-throughs were pooled, concentrated, ethanol precipitated, and dialyzed to yield anion exchanged polysaccharide preparation. The anionic fraction ("pectic polysaccharides") was recovered by elution with three column volumes of 1 M imidazole, pH 7.0 buffer, ethanol precipitation and dialysis. SP Sepharose FF cation exchange resin (GE Healthcare) was packed into a column and conditioned by sequential washes with water, 20 mM sodium acetate, pH 5.0 containing 0.5 M sodium chloride, and finally 20 mM sodium acetate, pH 5.0 buffer. Anion exchanged polysaccharide preparation was dissolved in 20 mM sodium acetate, pH 5.0 buffer, top-loaded on the resin, and washed in with the same buffer. Load and wash flow-throughs were pooled, concentrated, ethanol precipitated, and dialyzed to yield neutral polysaccharide preparation.

*Sycamore Polysaccharides for Immunization and Hybdridoma Screening*[1]

Portions of the neutral and pectic polysaccharide preparations from sycamore maple (obtained as described above) were further purified to remove the mannan polysaccharides and mannosyl-proteins that were introduced to the culture medium in the yeast extract component. The mannan polysaccharides were precipitated from the neutral polysaccharide preparation by barium hydroxide according to procedure adapted from (Fischer et al, 1996; Selvendran & O'Neill,

[1]Sycamore XyG polysaccharide preparation for immunization purposes, starting from the cell culture, was performed by Zoë Popper at CCRC.

2006). The neutral polysaccharide preparation was dissolved in water to a final concentration of 10 mg/mL and equal volume of 5% (w/v) barium hydroxide was added. The solution was incubated at room temperature for 1 h and the precipitate was pelleted by centrifugation. The supernatant was neutralized with acetic acid, dialyzed, and lyophilized. The barium hydroxide precipitation procedure was repeated to remove traces of the mannan. The mannosyl-proteins were removed from the pectic polysaccharides by lectin chromatography. Concanavalin A-Sepharose 4B resin (Sigma-Aldrich) was packed into a column to bed volume of 10 ml and conditioned with five bed volumes of 1 M sodium chloride 5 mM magnesium chloride, 5 mM manganese chloride, and 5 mM calcium chloride to saturate the Concanavalin A with these cations. The column was then equilibrated in 20 mM Tris, pH 7.4 buffer, containing 0.5 M sodium chloride, 5 mM magnesium chloride, 5 mM manganese chloride, and 5 mM calcium chloride. The pectic polysaccharide preparation was dissolved in the same buffer to a final concentration of 20 mg/mL, top-loaded to the resin, and washed with five bed volumes of the same buffer. The load and wash flow-throughs were pooled and lyophilized. The Concanavalin A purification was repeated to remove traces of the mannosyl-proteins, and the final load and wash flow-throughs were pooled, dialyzed and lyophilized.

*Tomato Polysaccharides for Epitope Characterization* Polysaccharides from tomato (*Solanum lycopersicum* L. ‘Bonny Best’) suspension cultures were purified essentially as described in (York et al, 1996). Briefly, the cells from a starter culture (kindly provided by April Harper at CCRC) were cultured for seven

days in Linsmaier and Skoog medium (Linsmaier & Skoog, 1965) supplemented with 3% sucrose and 1 mg/mL 2,4-dichlorophenoxyacetic acid, and harvested by vacuum filtering the cell suspension with a 50 μm Nylon net. The cell cake was washed with small aliquot of water, vacuum aspirated until water draining stopped, and frozen. Polymeric material was recovered from the culture filtrate as described for sycamore maple. Differential solubility in saturated ammonium sulfate of XyG and glucomannan, the two major neutral polysaccharides in tomato, was utilized for their purification. The neutral polysaccharide preparation was dissolved in water to a final concentration of 10 mg/mL, cooled and slowly stirred on ice while solid ammonium sulfate was added in small aliquots to saturation (760 mg/mL). The suspension was incubated on ice for 2 h and the ammonium sulfate insoluble material was pelleted by centrifugation (10000 *g*, 30 min, 4 °C). The pellet was dissolved in water, ethanol precipitated, and dialyzed to yield pure tomato XyG. Glucomannan was recovered from the supernatant by ethanol precipitation and dialysis.

*Tomato XyG for Immunization and Hybridoma Screening*[1] Portion of the tomato neutral polysaccharide preparation (obtained as described above) was treated with β-mannanase to hydrolyze trace amounts of glucomannan to oligosaccharides that can be separated from the remaining XyG polysaccharide material. Briefly, XyG was purified by dissolving the neutral polysaccharide preparation in ammonium formate, pH 4.5 buffer to a final concentration of 10

[1]Tomato XyG polysaccharide preparation for immunization purposes, starting from the cell culture, was performed by Zoë Popper at CCRC.

mg/mL and β-mannanase (Megazyme) was added to a final concentration of 1 U/mL. The reaction was allowed to proceed overnight at room temperature and the stopped by boiling for 5 min. The reaction mixture was ethanol precipitated and dialyzed to yield pure XyG.

*Tobacco Polysaccharides* Tobacco (*Nicotiana tabacum* L. 'Samsun') polysaccharides were purified essentially as described in (York et al, 1996). The cells were cultured for seven days in Linsmaier and Skoog medium (Linsmaier & Skoog, 1965) supplemented with 3% sucrose and 1 mg/mL 2,4-dichlorophenoxyacetic acid.[2] The XyG, galactoglucomannan, and pectic polysaccharides were purified from the spent medium as described for tomato (see above) using anion and cation exchange chromatographies and ammonium sulfate precipitation.

*XEG Treatment of Polysaccharides* Aliquots of selected non-XyG polysaccharide preparations that have shown response against XyG-binding mAbs in ELISAs (Pattathil et al, 2010) were treated with XEG. XyG polysaccharides from tamarind, sycamore maple, tomato and tobacco were treated in parallel as controls for XEG efficacy. Briefly, the XyGs, tomato glucomannan, tobacco galactoglucomannan, pectic polysaccharides, and lupin galactan preparations were dissolved in 50 mM ammonium acetate, pH 4.5 buffer to a final concentration of 10 mg/mL, XEG (*Paenibacillus* sp., Megazyme) was added to a final concentration 100 mU/mL (1 U, amount of substrate-

[2]The tobacco cell culture was performed by Stefan Eberhard at CCRC.

saturated XEG that creates reducing ends at a rate of 1 μmol/min), and the reactions were allowed to proceed 24 h at room temperature. Gum ghatti, gum guar, gum tragacanth, and locust bean gum were suspended to a final concentration of 0.2 mg/mL and dispersed by bath sonication (Model 2210, Branson Ultrasonic Corp.) for 1 h before the XEG treatment. After the reactions, the mixtures were lyophilized.

*De*-O-*acetylation of XyGs* The XyG polysaccharide preparations from sycamore maple, tomato, and tobacco were dissolved in 10 mM sodium hydroxide solution to a final concentration of 10 mg/mL and incubated overnight at room temperature. Tamarind XyG polysaccharide was treated similarly as control. The reaction mixtures were neutralized with acetic acid and lyophilized. Aliquots of the de-*O*-acetylated XyG polysaccharide preparations were XEG treated and analyzed by MALDI-TOF MS as described above to ascertain the completeness of de-*O*-acetylation. Control samples were prepared by adding sodium acetate to final concentration of 10 mM to natively *O*-acetylated XyG polysaccharide samples

### 3.5.2. Preparation of Immunogens[1]

Immunogens for the generation of mAbs were prepared by either of two strategies. In the first strategy, 1-cyano-4-dimethylaminopyridinium tetrafluoroborate (CDAP)-activated XyG polysaccharide was covalently conjugated to bovine serum albumin (BSA) (Lees et al, 1996). One hundred and

[1]All immunogen preparations were performed by Zoë Popper at CCRC.

sixty-five microliters of 100 mg/mL CDAP (Research Organics) in acetonitrile was added slowly to 22 mg of XyG polysaccharide in 1 ml of water. After 30 s, the CDAP was inactivated and reaction stopped by adding 165 µl of 0.2 M triethylamine and stirring for 90 s. The CDAP-activated XyG solution was transferred into a tube containing 2 mg of BSA (fatty acid free, Sigma-Aldrich) in 1 ml of 150 mM HEPES, pH 7.5 buffer. The reaction was incubated for 3 h at room temperature, after which it was quenched by the addition of 100 µl of 0.1 M ethanolamine in 750 mM HEPES, pH 7.5 buffer. The reaction mixture was incubated with stirring for 1 h after which the formed insoluble precipitate was pelleted by centrifugation. The supernatant was discarded and the pellet was resuspended in water. Removal of the water soluble material by resuspension and centrifugation was repeated three times. The final pellet was resuspended and dialyzed against water for 24 h at 4 °C and lyophilized. A portion of the conjugate was analyzed for sugar content by phenol-sulphuric acid assay with the appropriate XyG polysaccharide as a standard (DuBois et al, 1956) and for protein content by Bradford assay (BioRad) (Bradford, 1976) using BSA as a standard. Immunogens containing XyG polysaccharides from tamarind, sycamore maple (purified by barium hydroxide precipitation, see above) and tomato (purified by β-mannanase treatment, see above) were prepared using this strategy.

In the second strategy, the immunogen consisted of a mixture of non-XyG polysaccharide and methylated BSA (MeBSA, Sigma) dissolved in water, both in a final concentration of 1 mg/mL (Puhlmann et al, 1994). Immunogens containing

4-*O*-methylglucuronoxylan (Sigma), New Zealand flax (*Phormium tenax*) xylan (a generous gift from Industrial Research Limited) and sycamore maple pectic polysaccharide preparation (purified by Concanavalin-A chromatography, see above) were prepared using this strategy.

### 3.5.3. Immunization of Mice[1]

All immunizations were performed identically on five mice regardless of the immunogen. Five to six weeks old female BALB/c mice (Harlan-Sprague Dawley) were each given a subcutaneous injection of 100 µg of the immunogen with 50 µl of Freund's complete adjuvant in a total volume of 100 µl (Freund, 1956). The initial immunization was followed a by booster immunizations every 21 days by an intraperitoneal injection of 50 µg of the immunogen with 50 µl of Freund's incomplete adjuvant in a total volume of 100 µl (Freund, 1956). The mice were tail bled seven days after every injection and the sera (1:100, 1:300, 1:1,000, 1:3,000, 1:10,000, 1:30,000, 1:100,000, and 1:300,000 dilutions) were tested for the presence of the immunizing polysaccharide binding antibodies using ELISA. Absorbance of 0.1 or higher in ELISA (using 1:300,000 serum dilution) indicated a strong immune response against the immunogen. Four days after the positive ELISA response the mice were given a final intravenous pre-fusion booster injection of 40 µg of the immunogen in total volume of 50 µl without adjuvant. Splenectomy was performed three days after the pre-fusion booster injection.

[1]All immunizations were performed by Ruth Davis and associates at the Monoclonal antibody Facility at the UGA College of Veterinary Medicine.

The feeding, care, and euthanization of the mice followed the University Research Animal Resources guidelines.

### 3.5.4. Generation[1] and Screening[2] of Hybridoma Lines

Freshly isolated splenic lymphocytes were fused with a murine myeloma cell line Sp2/0-Ag14 (Shulman et al, 1978) and the resulting hybridomas were grown as described (Harlow & Lane, 1988; Pratt, 1984; Zola, 1987). The supernatants from initial hybridoma colonies were screened for antibodies against underivatized and unconjugated immunizing polysaccharide using ELISA. Hybridoma lines showing absorbance higher than 0.1 were passaged further and cloned by limiting dilution. The final clonal hybridoma lines were expanded and cryopreserved, and at least 1 L of their culture supernatants were collected and stored at -20 °C for further analyses. The mAb isotypes and light chain compositions were determined by ELISA against the immunizing polysaccharide using isotype-specific secondary antibodies (Fisher Scientific).

### 3.5.5. Other mAbs

The generation and characterization of CCRC-M1 mAb has been reported earlier (Puhlmann et al, 1994). The LM15 (Marcus et al, 2008) and LM24 (Pedersen et al, 2012) mAbs were a generous gift from Dr. Paul Knox.

---

[1]The generation of all hybridoma lines was performed by Ruth Davis and associates at the Monoclonal antibody Facility at the UGA College of Veterinary Medicine.

[2]The screening of the hybridoma lines was performed by Dr. Zoë Popper and David Baldwin at CCRC.

### 3.5.6. Preparation of Biotinylated Oligosaccharides

*XyG S1, minimal XyG, and cellodextrin oligosaccharides* Pure XyG S1 and minimal XyG oligosaccharides as well as cellodextrin oligosaccharide mixture (with degree of polymerization ranging from 4 to 7) were generated, purified and structurally validated as described in Chapter 2.

*XyG S*n *Oligosaccharides* Mixtures of XyG S*n* oligosaccharide with homogenous backbone length (4, 8, 12 or 16 D-Glc*p* residues for S1, S2, S3, and S4, respectively) but variable subunit composition were obtained from partial hydrolysis of tamarind seed XyG polysaccharide by XEG followed by purification by SEC. Briefly, the XyG polysaccharide was dissolved in 50 mM ammonium acetate, pH 4.5 buffer to a final concentration of 10 mg/mL. XEG from either *Aspergillus aculeatus* (Novozymes) or *Paenibacillus* sp. (Megazyme) was added to a final concentration of 10 mU/mL and the reaction was allowed to proceed at room temperature. MALDI-TOF MS was used to monitor the reaction progress. XEG was inactivated by boiling the reaction mixture for 5 min when the desired level of hydrolysis was achieved.

*Other Oligosaccharides* Lacto-*N*-tetraose (L*N*T, β-D-Gal*p*-(1→3)-β-D-Glc*p*NAc-(1→3)-β-D-Gal*p*-(1→4)-D-Glc); lacto-*N*-*neo*tetraose (L*Nn*T, β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→3)-β-D-Gal*p*-(1→4)-D-Glc); lacto-*N*-fucopentaose III (L*N*FPIII, β-D-Gal*p*-(1→4)-[α-L-Fuc*p*-(1→3)]-β-D-Glc*p*NAc-(1→3)-β-D-Gal*p*-(1→4)-D-Glc), and β-D-Gal*p*-(1→3)-β-D-Gal*p*NAc-(1→4)-β-D-Gal*p*-(1→4)-D-Glc oligosaccharides were purchased from V-Labs.

*Oligosaccharide Biotinylation* Oligosaccharides were conjugated to aminooxy-$PEG_4$-biotin (APB) tag using aniline catalyst (Thygesen et al, 2010). Briefly, typically 2 to 3 mg of oligosaccharide preparation (1.0 mg of the commercially obtained oligosaccharides) was reacted with 5 molar equivalents of APB (Thermo Scientific) in 200 μl of 100 mM ammonium acetate, 100 mM aniline, pH 4.5 buffer for 24 hours at room temperature. XyG S1 oligosaccharide-APB and cellodextrin-APB conjugates were separated from the unreacted oligosaccharides and reactants using gravity flow SEC as described in Chapter 2. The oligosaccharide-APBs were lyophilized, weighed to 0.1 mg accuracy using an analytical balance, dissolved in water to a final concentration of 1 mg/mL, and stored at -20 °C.

XyG S*n* oligosaccharide-APBs (where *n* = 1 to 4) conjugates were purified by high-pressure SEC using a Dionex UltiMate 3000 chromatography system (Thermo Scientific) equipped with a refractive index (RI) detector (R-101, Shodex). The oligosaccharide mixture was dissolved in water to a final concentration 10 mg/mL, spin filtered (0.2 μm, Nanosep MF, Pall Life Sciences), and 100 μl (total amount 1 mg) was injected to two Superdex 75 10/300 GL columns (GE Healthcare) connected in series. The oligosaccharides were eluted at 0.4 mL/min flow rate in water and individual peaks containing oligosaccharides with homogenous backbone length were collected manually. The RI chromatograms were analyzed with Chromeleon software (Thermo Scientific). The amounts of S1-, S2-, S3-, and S4-APBs were estimated from the peak areas in the RI chromatogram. The oligosaccharide-APB conjugates were dissolved in

water to a final concentration of 1.0 mg/mL, and stored at -20 °C. Small aliquot of the oligosaccharide-APB conjugates were analyzed by MALDI-TOF MS.

*Enzymatic Modification of the Biotinylated Oligosaccharides* Oligosaccharide-APB conjugates were dissolved in 50 mM sodium phosphate, pH 7.0 buffer to a final concentration of 25 mg/mL. Bacillus sp. No. 693-1 α-xylosidase (Seikagaku Corporation) was added to a final concentration of 100 mU/mL and the reaction was allowed to proceed overnight at room temperature. The enzyme was removed from the reaction mixture by spin filtering (10 kDa cutoff) and the filtrate was lyophilized. Small aliquots of the reaction mixtures were analyzed by MALDI-TOF MS.

*Chemical Synthesis of Biotinylated Z Oligosaccharide*[1] The details of the synthesis will be published elsewhere. The structure of the oligosaccharide-biotin conjugate is shown in Supplemental Figure 2.

### 3.5.7. MALDI-TOF MS

MALDI-TOF mass spectra were collected on a Nitrogen laser equipped MicroFlex LT instrument (Bruker) under the positive mode. Equal volumes of the oligosaccharide sample and the matrix solution [20 mg/mL 2,5-dihydroxybenzoic acid in 1:1 (vol:vol) methanol:water] were mixed and 1 µl of the mixture was applied and allowed to air dry on a ground steel sample plate (Bruker). Mass calibration was performed with malto-oligosacharides (Sigma-Aldrich). The laser

[1]The chemical synthesis the biotinylated Z oligosaccharide was performed by Roshan Baliga under the supervision of Professor Geert-Jan Boons at CCRC.

power and the number of summed laser pulses were adjusted separately for each sample spot to yield spectra with high signal-to-noise ratios. The spectra were analyzed with the flexAnalysis software (version 3.0, Bruker).

### 3.5.8. ELISAs

All ELISAs were performed in triplicate wells and the averages and standard deviations were calculated. Negative controls without carbohydrate coating and without hybridoma supernatant were performed in parallel.

*Direct ELISAs Against Immobilized Polysaccharides* Wells on a clear flat-bottom polystyrene 96-well microplate (Costar 3598, Corning) were coated by transferring 50 μl of 10 μg/mL polysaccharide solution in water (total amount 0.5 μg) and evaporating the solution to dryness overnight at 37 °C. The wells were blocked with 200 μl of Tris-buffered saline (TBS, 50 mM Tris, 100 mM NaCl, pH 7.4) containing 1% non-fat milk powder for 1 h at room temperature (all subsequent incubations were performed at room temperature) and the blocking solution was aspirated using microplate washer (Bio-Tek ELx405). For the preliminary single concentration ELISAs, the wells were incubated with 50 μl of hybridoma supernatant for 1 h. Neat hybridoma supernatant was used for all mAbs in the CCRC-M series, whereas mAbs in the LM series were diluted 20-fold by TBS containing 0.1% milk powder (MTBS) immediately prior to use. For the concentration series ELISAs, wells were incubated with a twelve-step dilution series of the hybridoma supernatant, prepared by successive 1:1 dilutions to MTBS. Unbound material was aspirated and washed off (3 × 5 s) with 300 μl of

MTBS. The wells were incubated with 50 μl of species and isotype-specific secondary antibody-horseradish peroxidase conjugate (goat anti-whole molecule, Sigma, 1:5000 dilution in MTBS) for 1 h and unbound material was aspirated and washed off (5 × 5 s). The wells were incubated with 50 μl of 3,3′,5,5′-tetramethylbenzidine (TMB) substrate solution (Vector Laboratories) typically for 15 to 90 min for color development. The enzyme reaction was stopped by adding 50 μl of 1 N sulfuric acid and absorbances at 490 and 655 nm were read spectrophotometrically using a microplate reader (Model 490, Bio-Rad). The absorbances were normalized to 20 min.

*Competitive ELISAs Against Immobilized Polysaccharides* Twelve aliquots of the hybridoma supernatant were mixed with equal volume of competitor dilution series, prepared by successive 1:1 dilutions of 10 mg/mL competitor in TBS. This mixture was incubated for 1 h and 50 μl was transferred to the coated wells for 1 h incubation. The procedure is otherwise identical to that described above.

*Direct ELISAs Against Immobilized Oligosaccharides* NeutrAvidin coated and pre-blocked clear polystyrene plates [standard capacity, Thermo Scientific] were used for ELISAs using immobilized oligosaccharides. The wells used for rat mAbs (LM series) were blocked with 1 mg/mL BSA (Fisher Scientific) in water for 1 h before the coating, whereas factory blocking was sufficient with wells used for mouse mAbs (CCRC-M series). For the preliminary single-concentration ELISAs, the wells were incubated with 50 μl of 200 pmol/mL oligosaccharide-APB (total

amount 10 pmol) in water for 1 h. For the concentration series ELISAs, the wells were incubated for 1 h with 50 μl of a twelve-step dilution series, prepared by successive 1:1 water dilutions of the 200 pmol/mL oligosaccharide-APB solution. Unbound material was aspirated and washed off (3 × 5 s) with 300 μl of TBS. The rest of the procedure is identical to that described above.

### 3.5.9. Hierarchical Clustering of ELISA Data

The ELISA data was hierarchically clustered using an in-house written script (Pattathil et al, 2010) in the R programming environment (R Core Team, 2013). Complete clustering method was used for the hierarchical clustering and spherical metric was used for both clustered dimensions (the mAbs as well as the oligosaccharides).

### 3.5.10. Immunofluorescence Labeling

Immunofluorescent labeling of Arabidopsis root[1] and tamarind seed[2] sections were performed as described (Avci et al, 2012).

## 3.6. Acknowledgements

We thank Dr. Chin Fen Teo for critical reading of the manuscript. We thank April Harper for providing the tomato suspension culture, Dr. Paul Knox for providing

[1]Immunofluorescence labeling of Arabidopsis root tissue (starting from the plant culture) was performed by Stefan Eberhard at CCRC.

[2]Immunofluorescence labeling of tamarind seed was performed by Dr. Glenn Freshour at CCRC.

the LM15 and LM24 hybridoma supernatants, and Industrial Research Limited (IRL) for providing the New Zealand flax (*Phormium tenax*) xylan. This work was supported by the National Science Foundation Plant Genome Program (grant no. DBI–0421683).

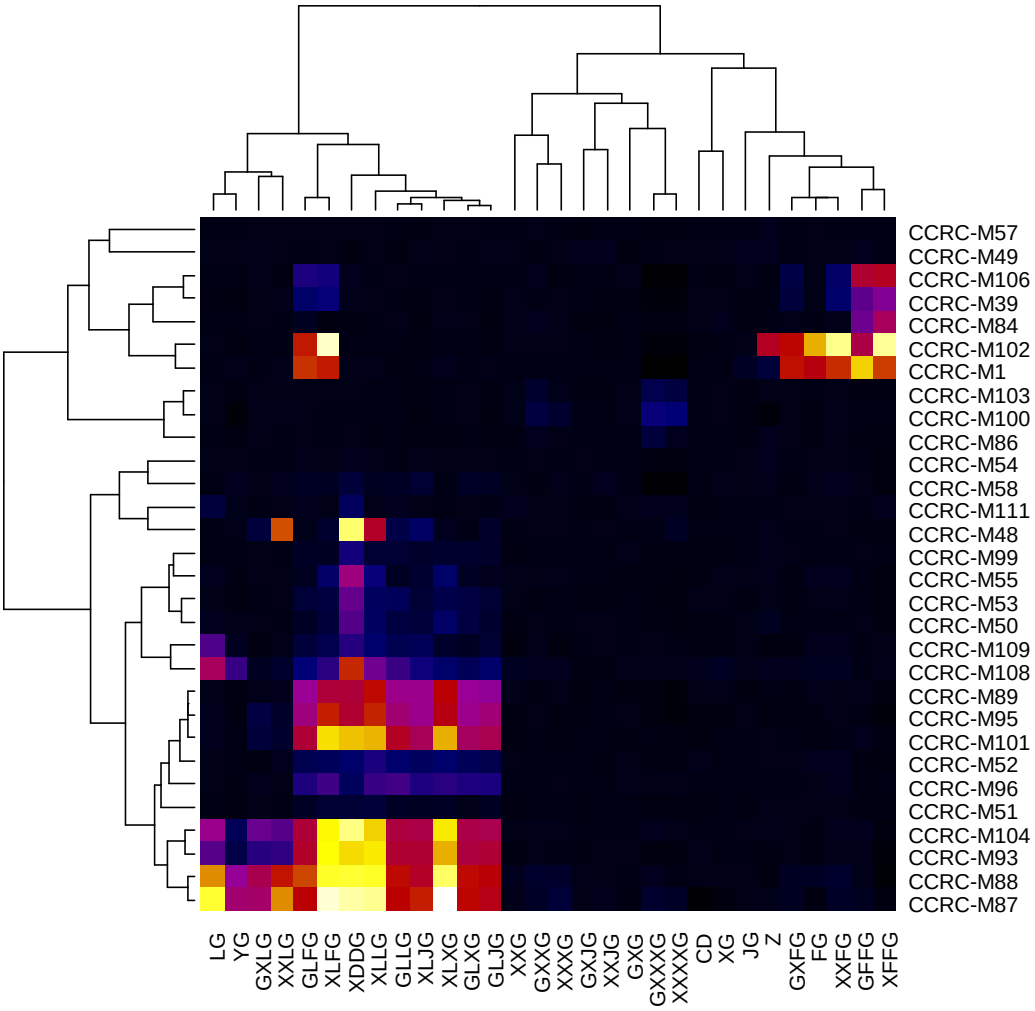



**Figure 3-1.** Heatmap of the ELISA data of the CCRC mAb binding to the immobilized oligosaccharides. Each row in the heatmap corresponds to the ELISA binding pattern of a single mAb against collection of purified and immobilized oligosaccharides. Each column corresponds to a binding pattern of a single immobilized oligosaccharide against collection of mAbs. The color of an intersection of an antibody and an oligosaccharide represents the average of three ELISA absorbance readings. The hierarchical clustering groups the mAb as well as oligosaccharides according to the similarity of their binding patterns.

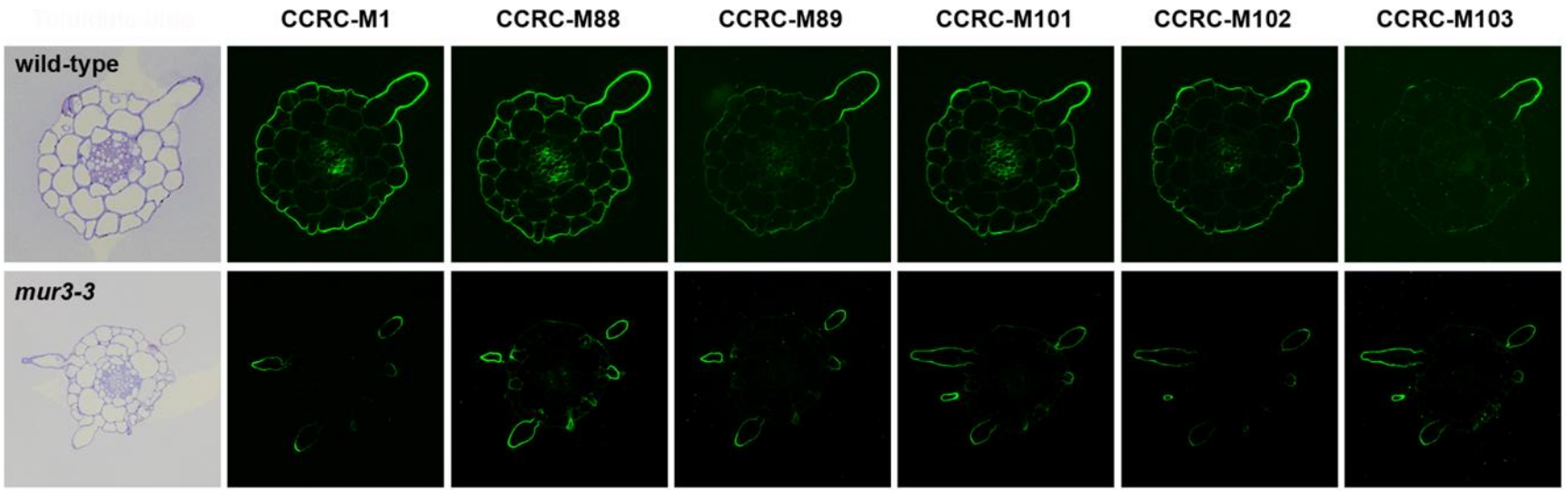

Toluidine Blue
CCRC-M1
CCRC-M88
CCRC-M89
CCRC-M101
CCRC-M102
CCRC-M103
wild-type
mur3-3

**Figure 3-2.** Immunolocalization of XyG epitopes in transverse sections of *Arabidopsis thaliana* root tissue. Toluidine blue specifically stains polymerized lignin in the cell wall and thus provides the overall root morphology, location and thickness of cell walls, and orientation of the section. Wild-type Arabidopsis (ecotype Columbia) shows variable staining by the six antibodies, ranging from CCRC-M88 that stains large number of different cell types, to CCRC-M89 and CCRC-M103 that only stain the root hairs. The labeling with all mAbs is restricted to the root hairs in *mur3-3* mutant. (Figure courtesy of Stefan Eberhard, CCRC).

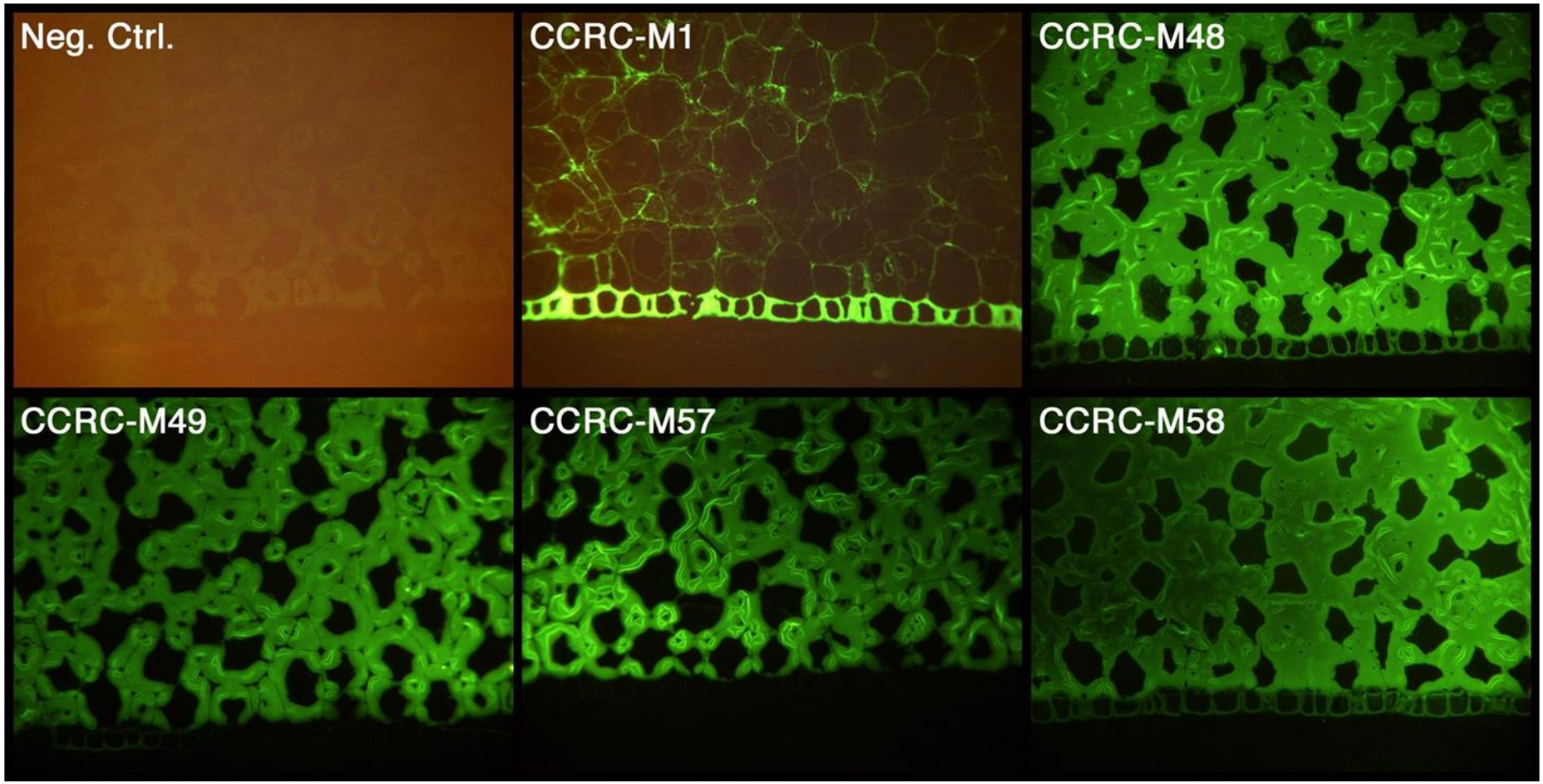

Neg. Ctrl.
CCRC-M1
CCRC-M48
CCRC-M49
CCRC-M57
CCRC-M58

**Figure 3-3.** Immunolocalization of XyG epitopes in sections of tamarind seed tissue (Figure courtesy of Dr. Glenn Freshour, CCRC).

**Table 3-S1**. Statistics of the immunizations and hybridoma line generation, and detailed information on the resulting mAbs. The fusions that resulted in at least one stable monoclonal hybridoma line are listed. The total number of tested hybridoma lines includes all fusions regardless of whether they resulted in stable hybridoma line or not.

| Immunogen | Positive Fusions | Hybridoma Lines | | mAbs[1] | |
|---|---|---|---|---|---|
| | | Tested | Stable mAb-secreting | Name | Isotype, Light chain |
| **Tamarind XyG-BSA conjugate (covalent)** | AS1 | 1222 | 8 | CCRC-M48 | $IgG_1$, κ |
| | | | | CCRC-M49 | $IgG_1$, λ |
| | | | | CCRC-M50 | $IgG_1$, λ |
| | | | | CCRC-M51 | $IgG_1$, λ |
| | | | | CCRC-M52 | $IgG_1$, λ |
| | | | | CCRC-M53 | $IgG_1$, λ |
| | | | | CCRC-M54 | $IgG_1$, λ |
| | | | | CCRC-M55 | $IgG_1$, λ |
| | AS5 | 386 | 2 | CCRC-M57 | $IgG_1$, λ |
| | | | | CCRC-M58 | $IgG_1$, κ |
| **Total** | **3** | **2146** | **10** | | |
| **Sycamore XyG-BSA conjugate (covalent)** | DB27 | 1411 | 1 | CCRC-M100 | IgM, κ |
| | DB31 | 770 | 2 | CCRC-M102 | IgM, κ |
| | | | | CCRC-M103 | IgM, κ |
| | DB33 | 277 | 1 | CCRC-M106 | $IgG_1$, κ |
| **Total** | **5** | **3635** | **4** | | |

**Table 3-S1**. *Continued*.

| Immunogen | Positive Fusions | Hybridoma Lines | | mAbs | |
|---|---|---|---|---|---|
| | | Tested | Stable mAb-secreting | Name | Isotype, Light chain |
| **Tomato XyG-BSA conjugate (covalent)** | DB16 | 293 | 4 | CCRC-M86 | IgM, κ |
| | | | | CCRC-M87 | $IgG_1$, λ |
| | | | | CCRC-M88 | $IgG_1$, λ |
| | | | | CCRC-M89 | $IgG_1$, λ |
| | DB17 | 583 | 2 | CCRC-M90 | IgM, λ |
| | | | | CCRC-M93 | $IgG_1$, λ |
| | DB24 | 608 | 3 | CCRC-M95 | $IgG_1$, λ |
| | | | | CCRC-M96 | $IgG_3$, λ |
| | | | | CCRC-M99 | $IgG_1$, λ |
| | DB30 | 343 | 2 | CCRC-M101 | $IgG_1$, λ |
| | | | | CCRC-M104 | $IgG_1$, λ |
| **Total** | **5** | **2118** | **11** | | |
| **4-*O*-methyl-glucuronoxylan-BSA mixture** | TZ1 | 1034 | 1 | CCRC-M39 | IgA, κ |
| **Total** | **2** | **2005** | **1** | | |
| **TOTAL** | **15** | **8104** | **26** | | |

[1]Other mAbs used in this work include CCRC-M84 (obtained from mouse immunized with suspension cultured sycamore maple pectic polysaccharide preparation), as well as CCRC-M108, CCRC-M109, and CCRC-M111 (obtained from mice immunized with New Zealand flax (*Phormium tenax*) xylan). The details from the hybridoma line generation will be described elsewhere.

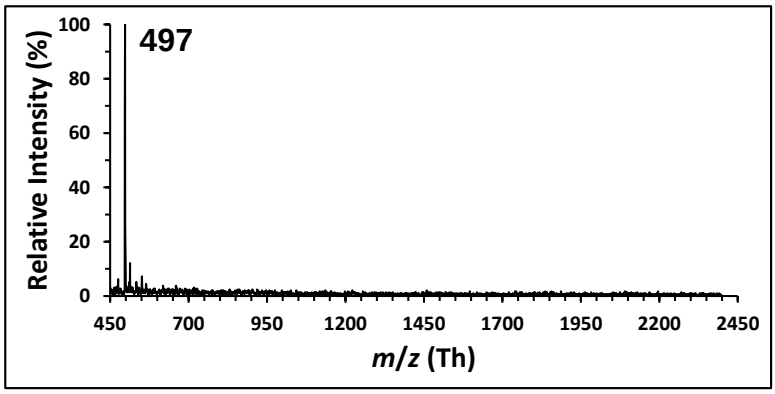
**A** **Preparation and purification of XG-APB**

MALDI-TOF MS

**XG**

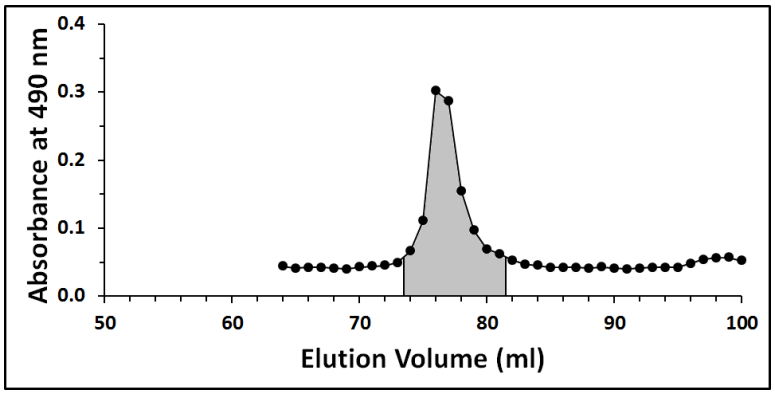
Calc. $[M + Na]^+$

$m/z$ = 497.2

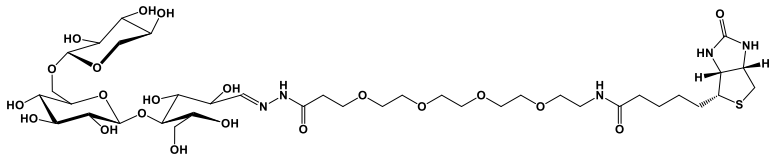



SEC

**Reaction mixture**




Structure

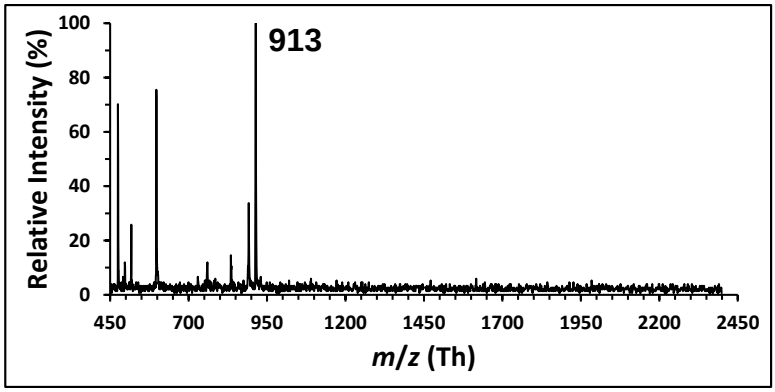


MALDI-TOF MS

**XG-APB**

Calc. $[M + Na]^+$

$m/z$ = 913.4



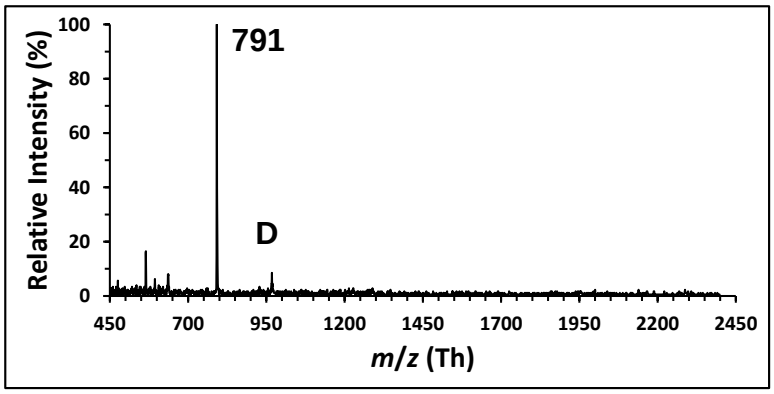
**B** **Preparation and purification of XXG-APB**

MALDI-TOF MS

**XXG**

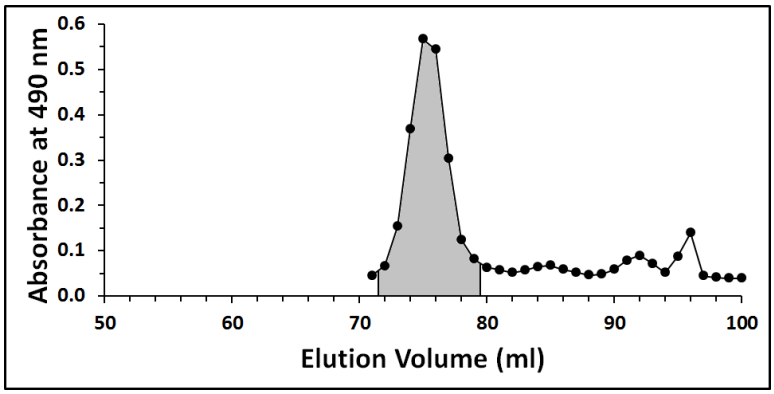
Calc. $[M + Na]^+$

*m*/*z* = 791.2




SEC

**Derivatization reaction mixture**

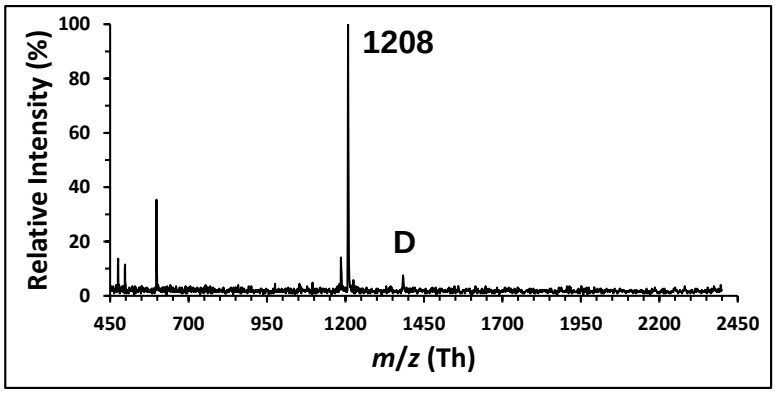



MALDI-TOF MS

**XXG-APB**

Calc. $[M + Na]^+$

*m*/*z* = 1207.5



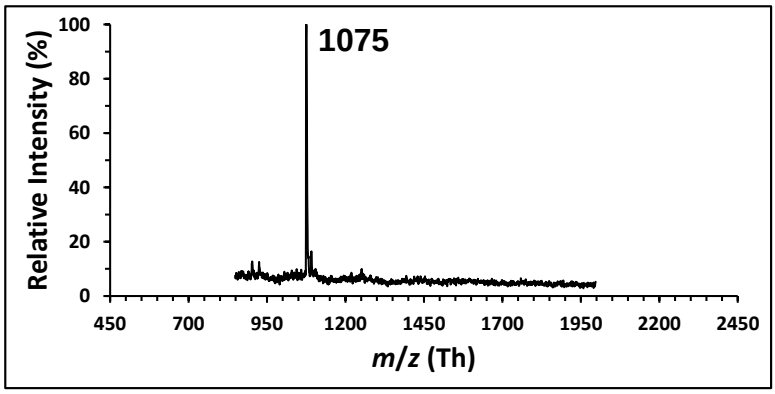
**C** **Preparation and purification of GXG-APB**

MALDI-TOF MS

**GXG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1075.4



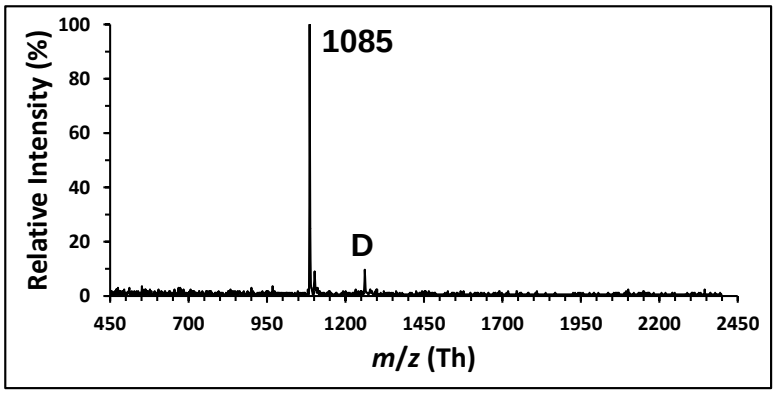
**D** **Preparation and purification of XXXG-APB**

MALDI-TOF MS

**XXXG**

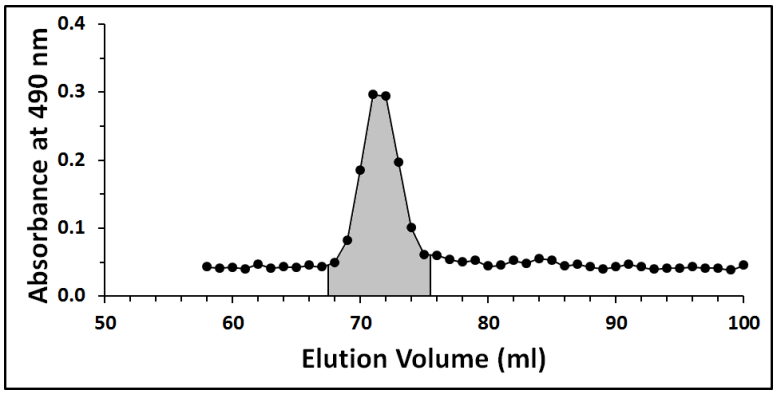
Calc. $[M + Na]^+$

*m*/*z* = 1085.3




SEC

**Derivatization reaction mixture**

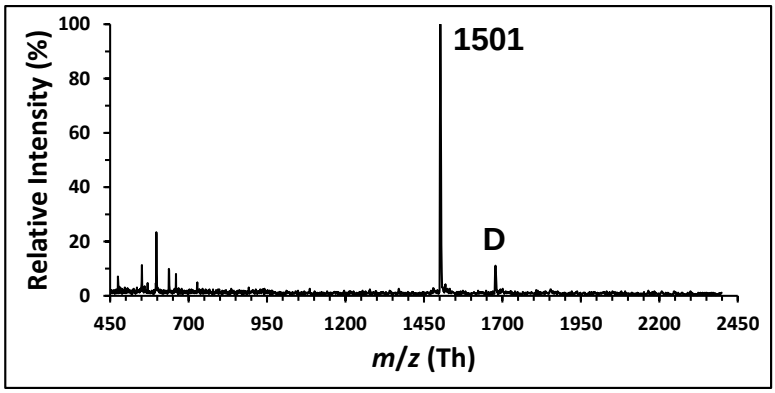



MALDI-TOF MS

**XXXG-APB**

Calc. $[M + Na]^+$

*m*/*z* = 1501.5



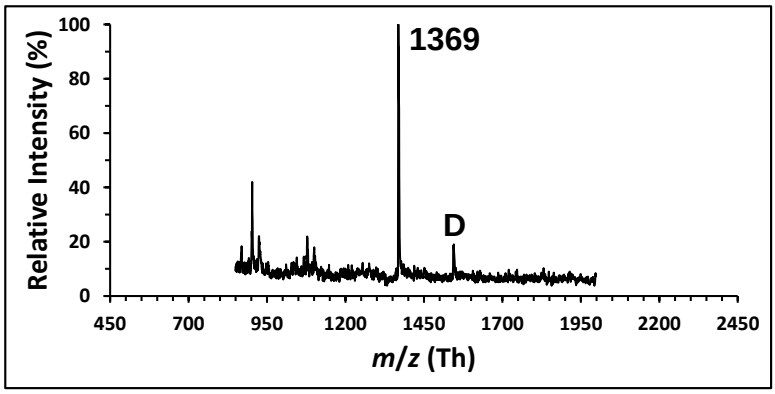
**E** **Preparation and purification of GXXG-APB**

MALDI-TOF MS

**GXXG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1369.5



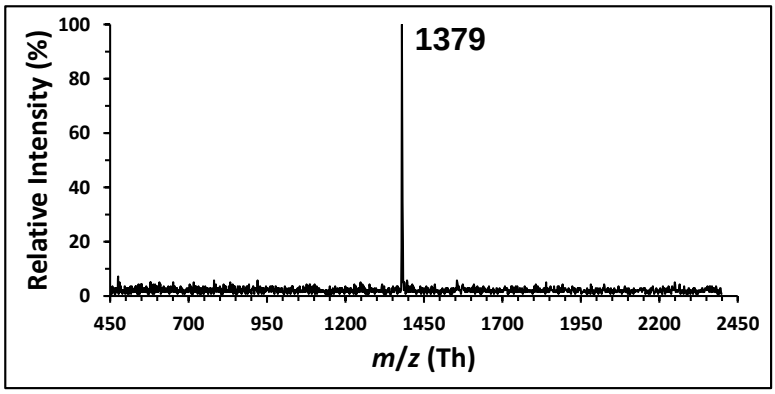
**F** **Preparation and purification of XXXXG-APB**

MALDI-TOF MS

**XXXXG**

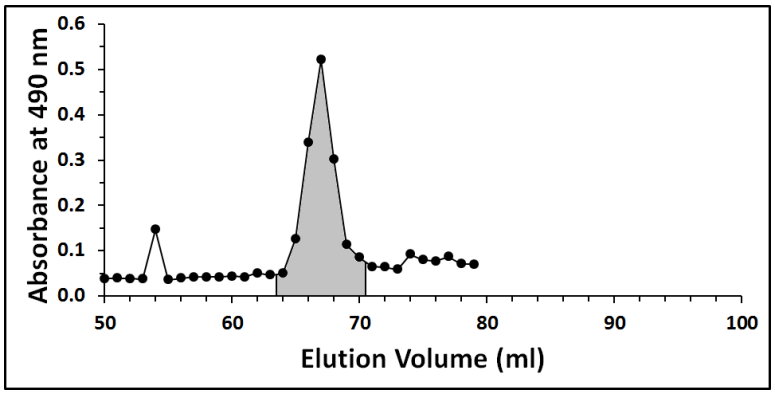
Calc. $[M + Na]^+$

*m*/*z* = 1379.4




SEC

**Derivatization reaction mixture**

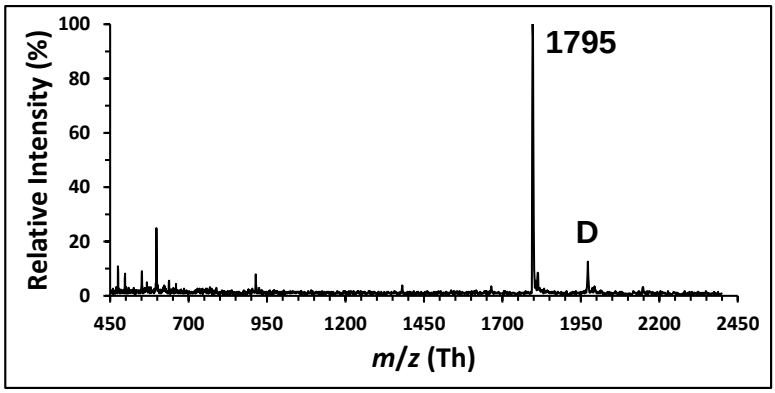



MALDI-TOF MS

**XXXXG-APB**

Calc. $[M + Na]^+$

*m*/*z* = 1795.6



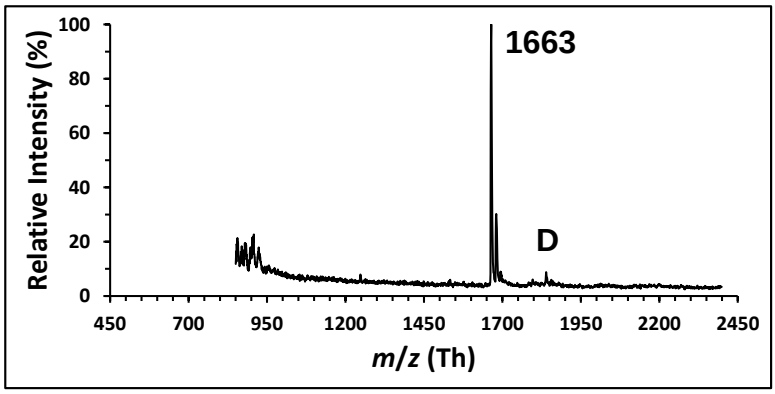
**G** **Preparation and purification of GXXXG-APB**

MALDI-TOF MS

**GXXXG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1663.6



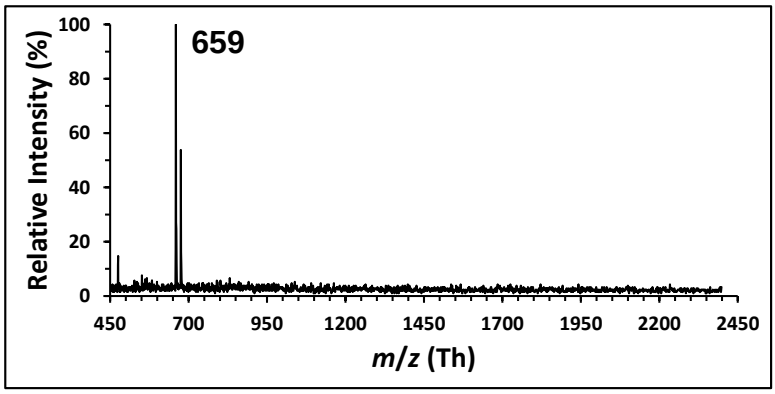
**H** **Preparation and purification of LG-APB**

MALDI-TOF MS

**LG**

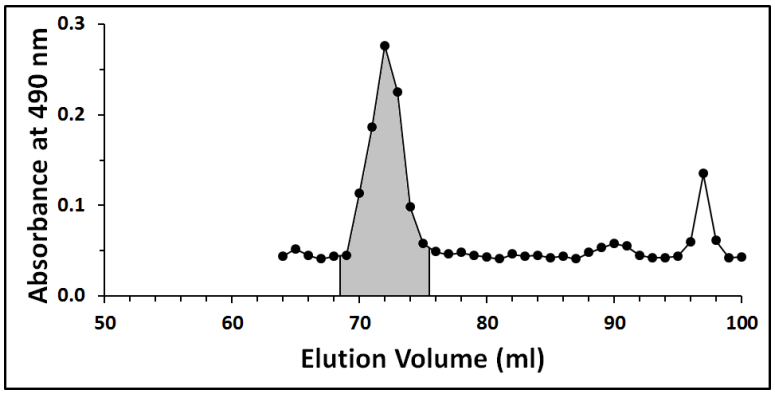
Calc. $[M + Na]^+$

*m*/*z* = 659.2




SEC

**Derivatization reaction mixture**

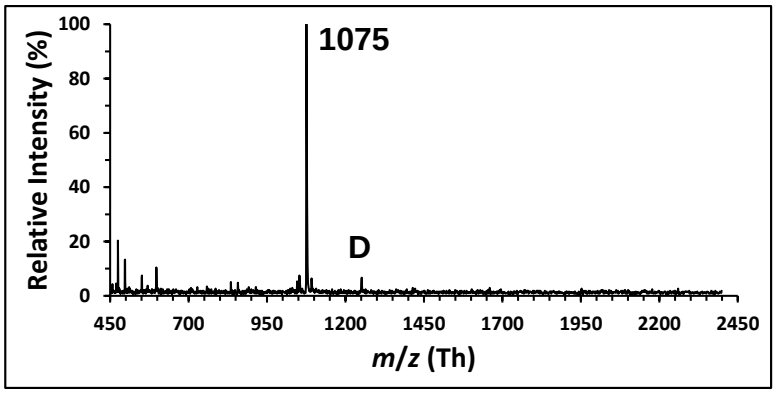



MALDI-TOF MS

**LG-APB**

Calc. $[M + Na]^+$

*m*/*z* = 1075.4



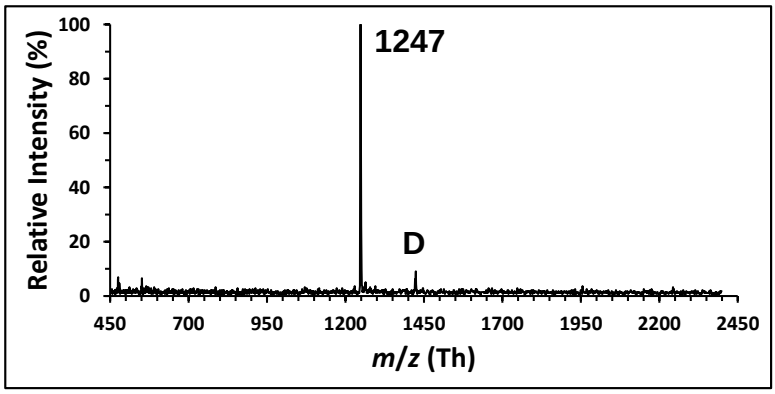
**I** **Preparation and purification of XLXG-APB**

MALDI-TOF MS

**XLXG**

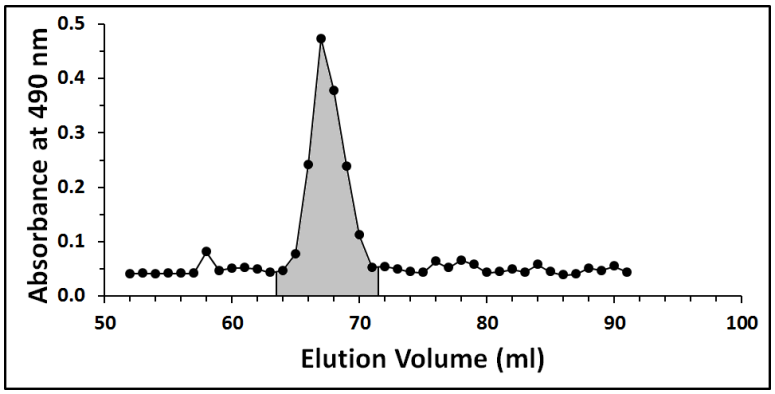
Calc. $[M + Na]^+$

$m/z$ = 1247.4




SEC

**Derivatization reaction mixture**

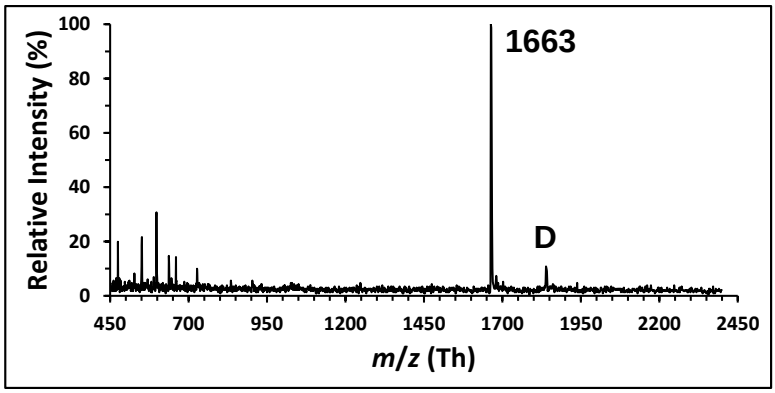



MALDI-TOF MS

**XLXG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1663.6



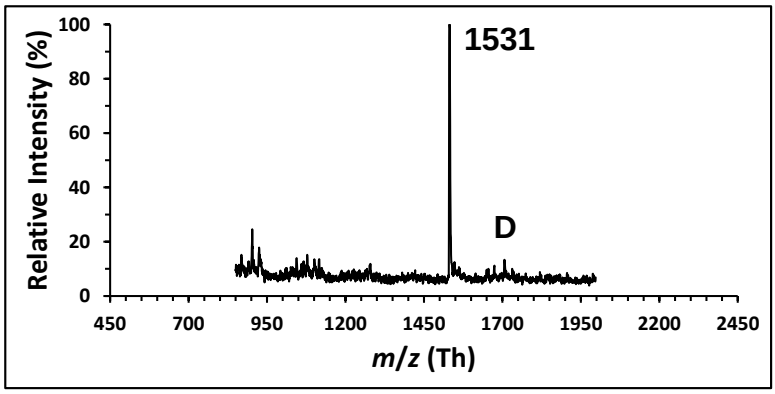
**J** **Preparation and purification of GLXG-APB**

MALDI-TOF MS

**GLXG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1531.6



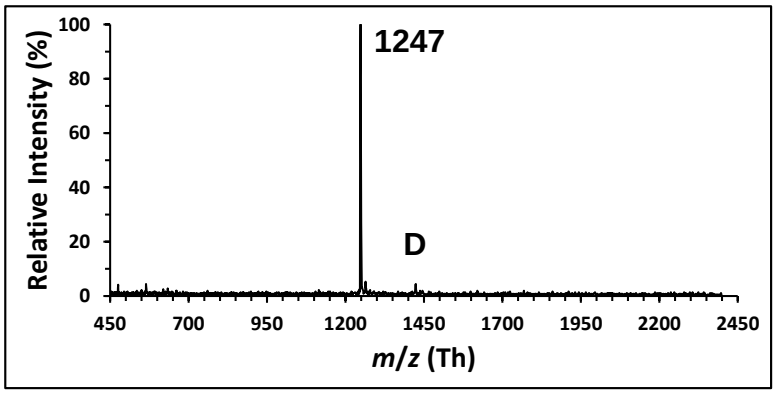
**K** **Preparation and purification of XXLG-APB**

MALDI-TOF MS

**XXLG**

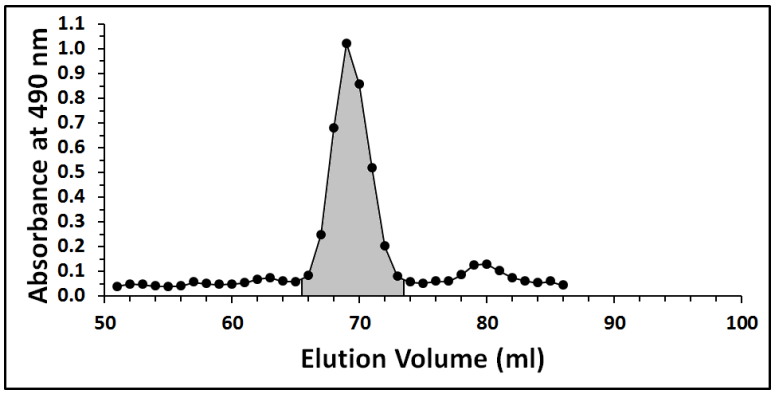
Calc. $[M + Na]^+$

$m/z$ = 1247.4




SEC

**Derivatization reaction mixture**

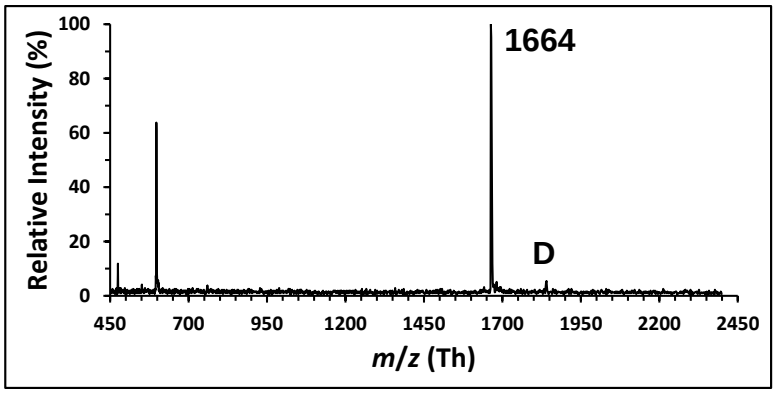



MALDI-TOF MS

**XXLG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1663.6



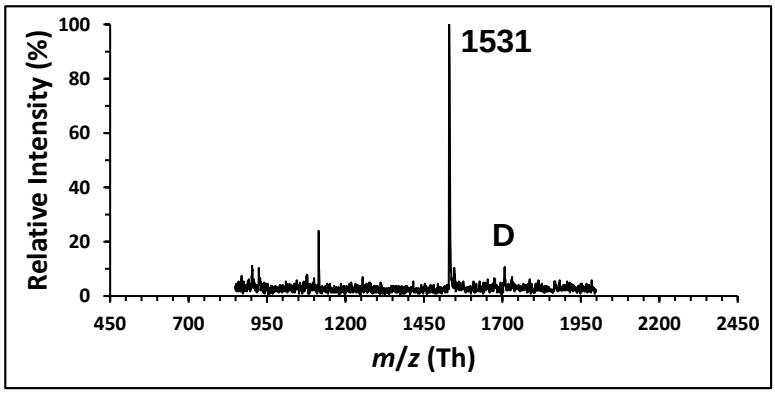
**L** **Preparation and purification of GXLG-APB**

MALDI-TOF MS

**GXLG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1531.6



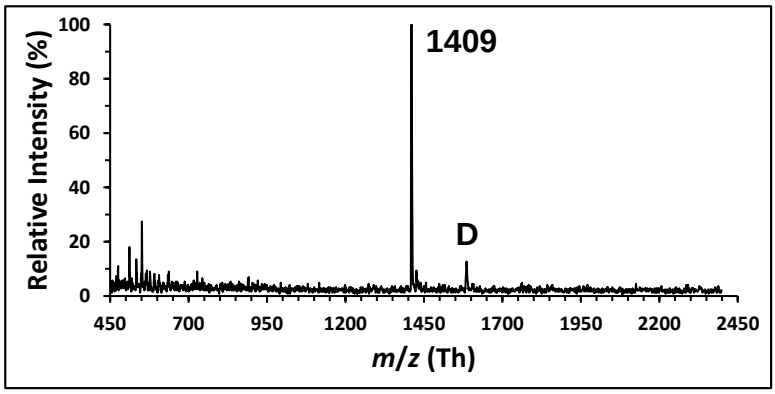
**M** **Preparation and purification of XLLG-APB**

MALDI-TOF MS

**XLLG**

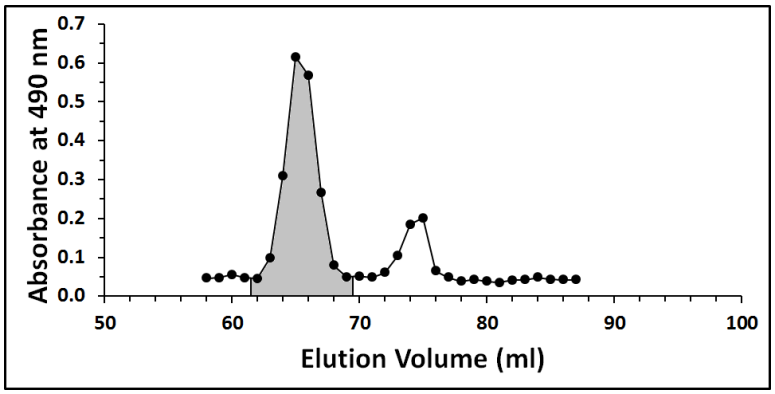
Calc. $[M + Na]^+$

*m*/*z* = 1409.4




SEC

**Derivatization reaction mixture**

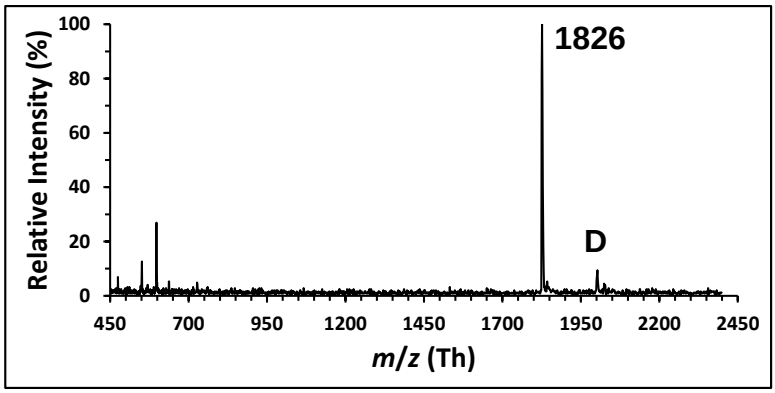



MALDI-TOF MS

**XLLG-APB**

Calc. $[M + Na]^+$

*m*/*z* = 1825.7



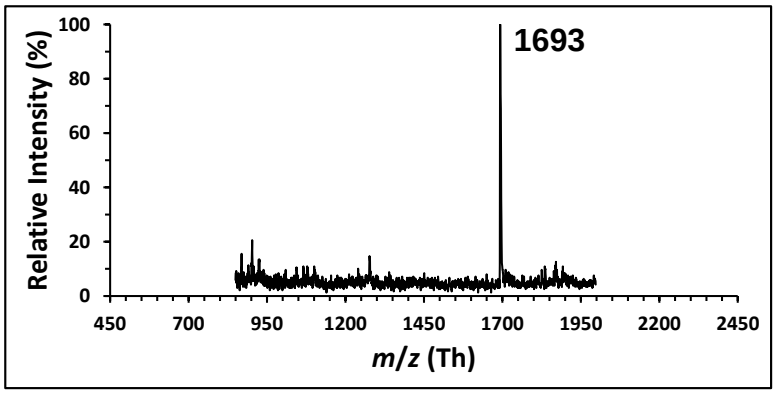
**N** **Preparation and purification of GLLG-APB**

MALDI-TOF MS

**GLLG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1693.6



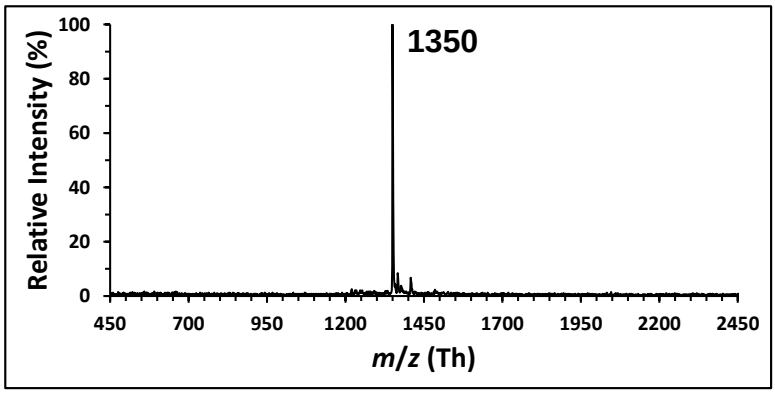
**O** **Preparation and purification of XDDG-APB**

MALDI-TOF MS

**XDDG**

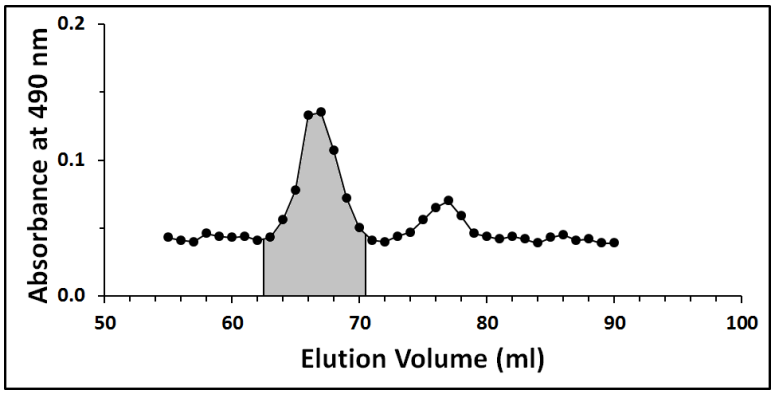
Calc. $[M + Na]^+$

$m/z$ = 1349.4




SEC

**Derivatization reaction mixture**

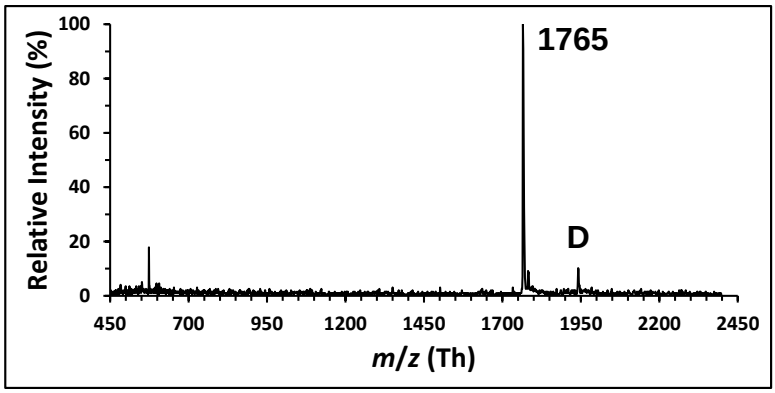



MALDI-TOF MS

**XDDG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1765.6



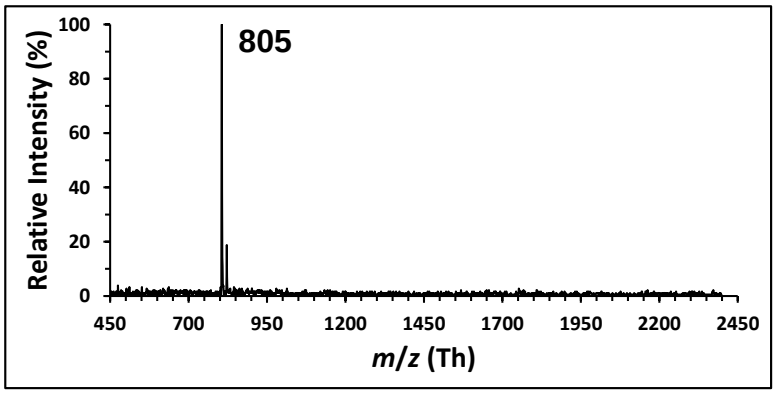
**P** **Preparation and purification of FG-APB**

MALDI-TOF MS

**FG**

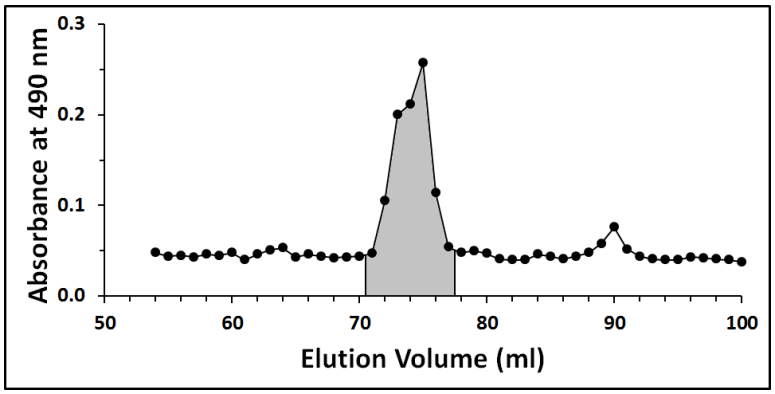
Calc. $[M + Na]^+$

$m/z$ = 805.3




SEC

**Derivatization reaction mixture**

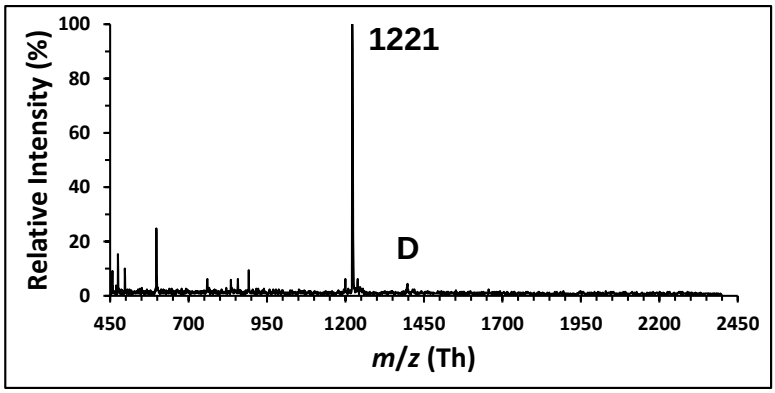



MALDI-TOF MS

**FG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1221.5



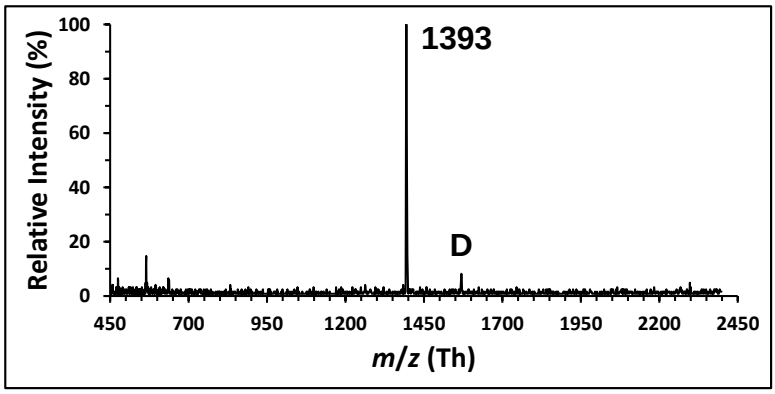
**Q** **Preparation and purification of XXFG-APB**

MALDI-TOF MS

**XXFG**

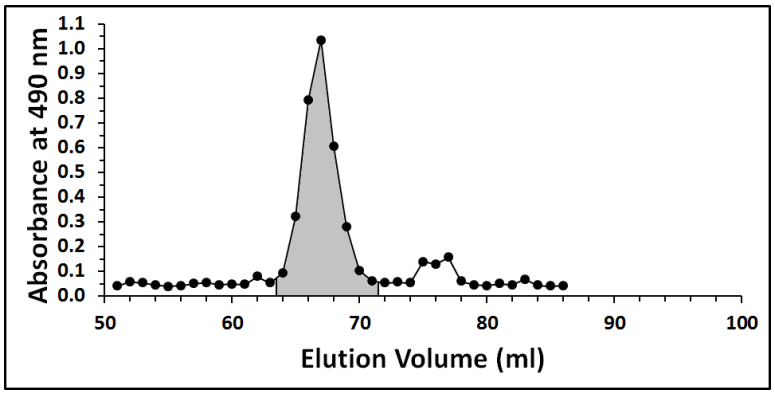
Calc. $[M + Na]^+$

$m/z$ = 1393.5




SEC

**Derivatization reaction mixture**

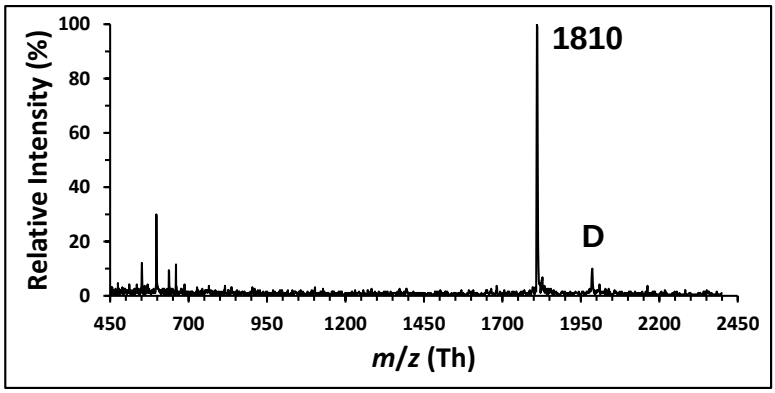



MALDI-TOF MS

**XXFG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1809.7



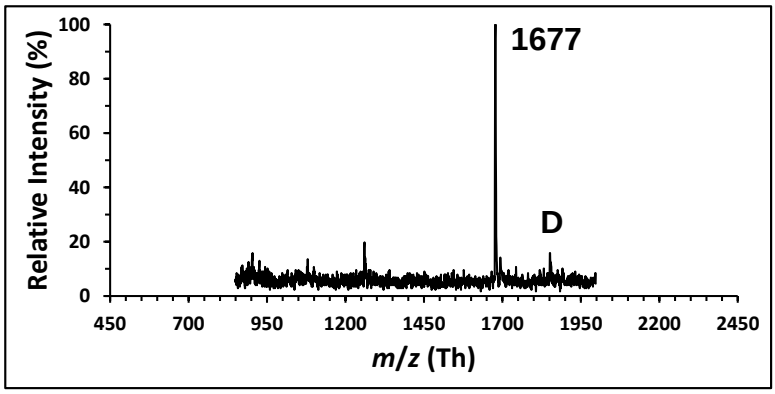
**R** **Preparation and purification of GXFG-APB**

MALDI-TOF MS

**GXFG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1677.6



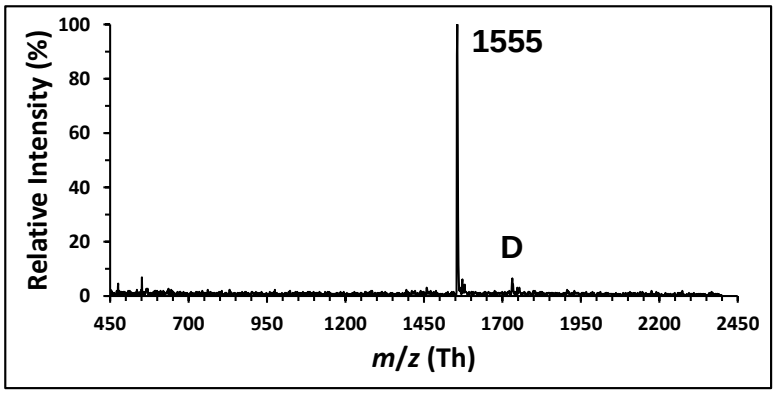
## Preparation and purification of XLFG-APB

MALDI-TOF MS

**XLFG**

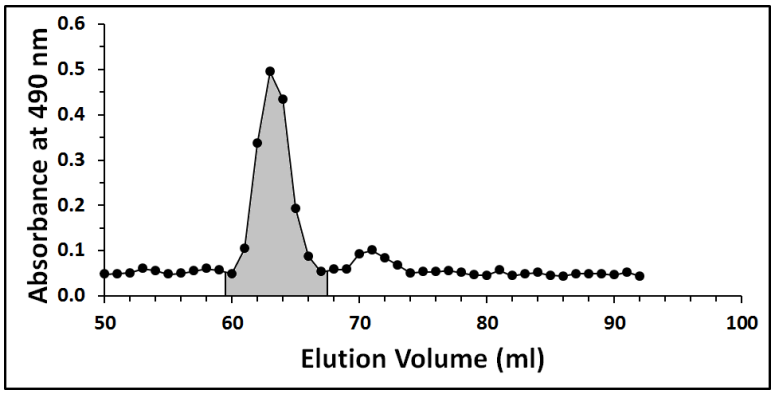
Calc. $[M + Na]^+$

$m/z$ = 1555.5




SEC

**Derivatization reaction mixture**

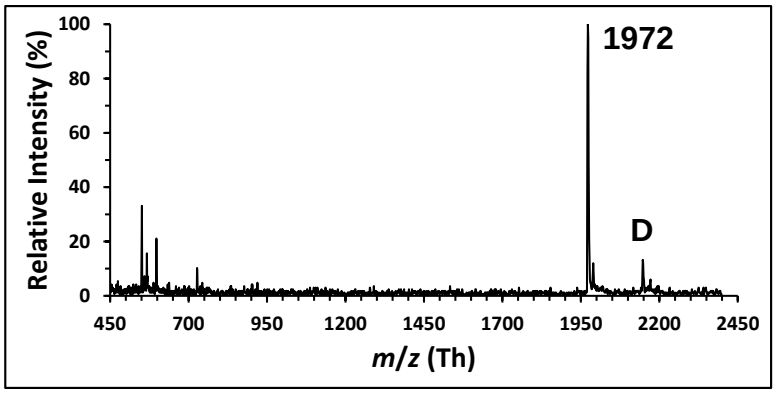



MALDI-TOF MS

**XLFG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1971.7



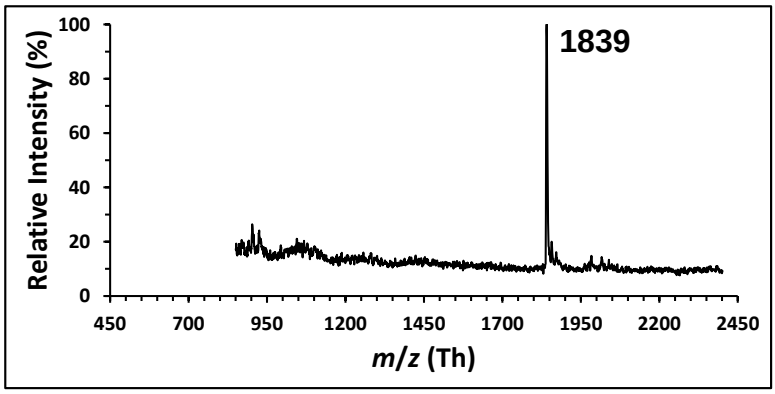
**T** **Preparation and purification of GLFG-APB**

MALDI-TOF MS

**GLFG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1839.7



**U**

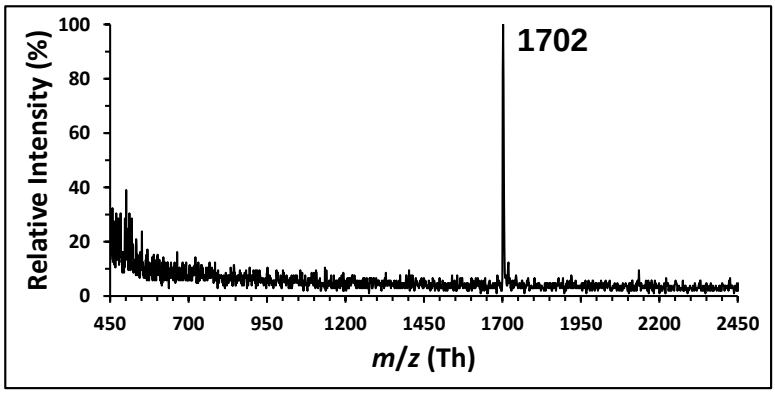
## Preparation and purification of XFFG-APB

MALDI-TOF MS

**XFFG**

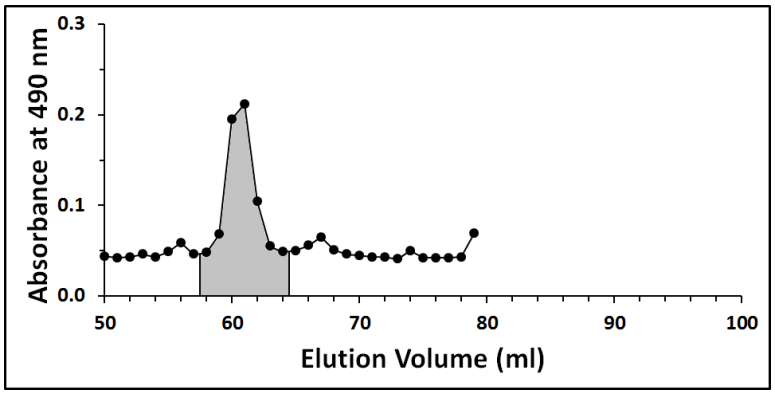
Calc. $[M + Na]^+$

$m/z$ = 1701.6




SEC

**Derivatization reaction mixture**

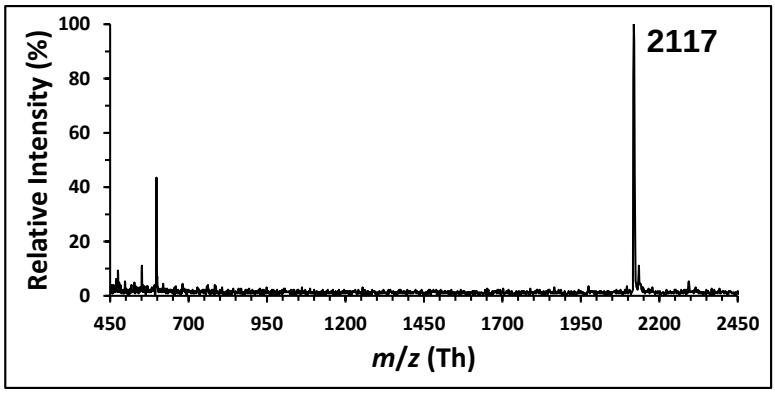



MALDI-TOF MS

**XFFG-APB**

Calc. $[M + Na]^+$

$m/z$ = 2117.8



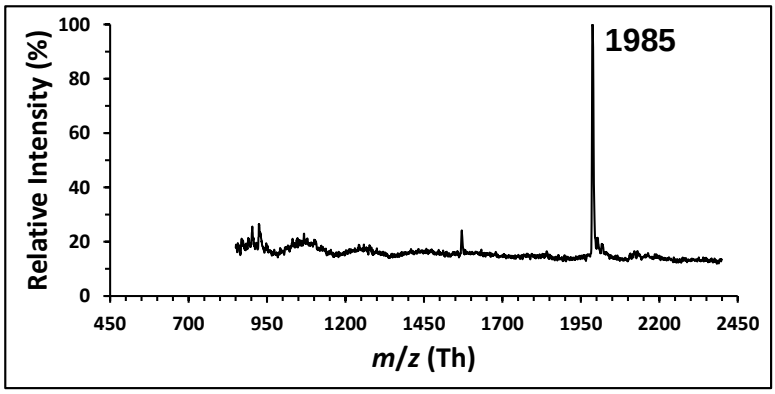
**V** **Preparation and purification of GFFG-APB**

MALDI-TOF MS

**GFFG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1985.7



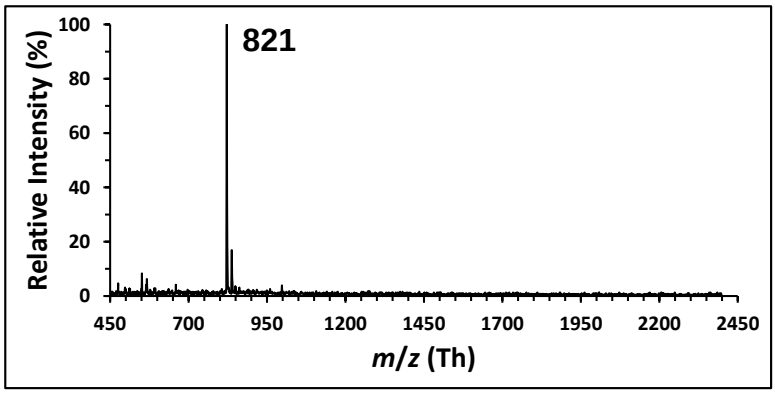
**W** **Preparation and purification of JG-APB**

MALDI-TOF MS

**JG**

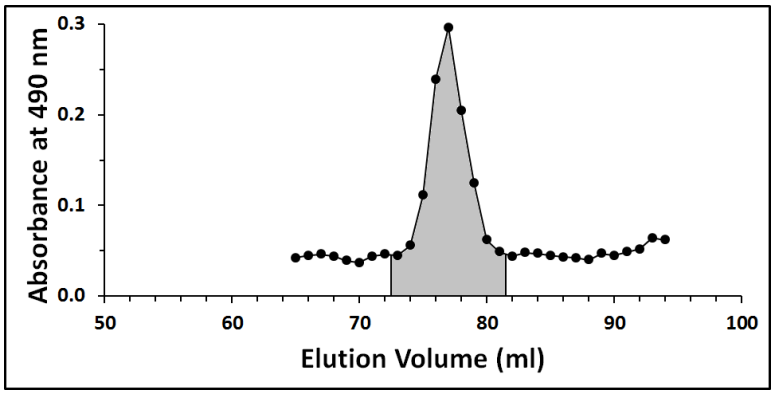
Calc. $[M + Na]^+$

$m/z$ = 821.3




SEC

**Derivatization reaction mixture**

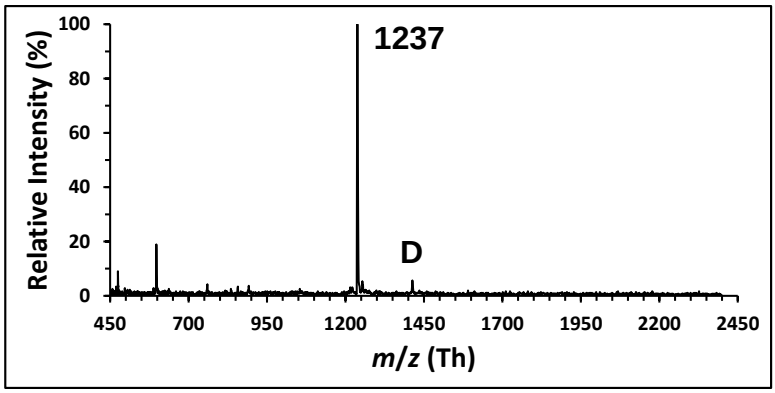



MALDI-TOF MS

**JG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1237.5



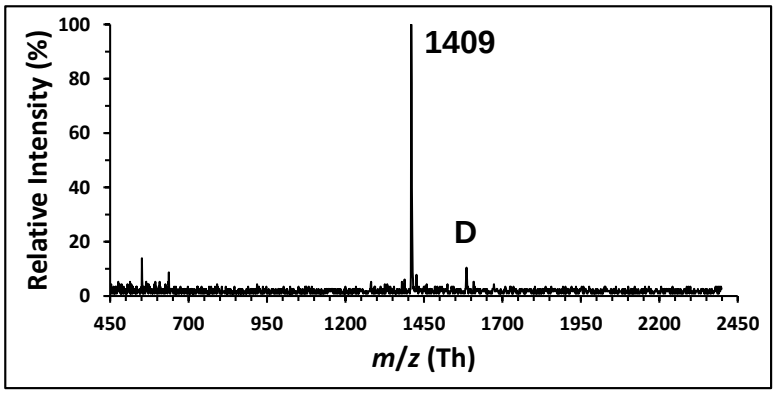
**X** **Preparation and purification of XXJG-APB**

MALDI-TOF MS

**XXJG**

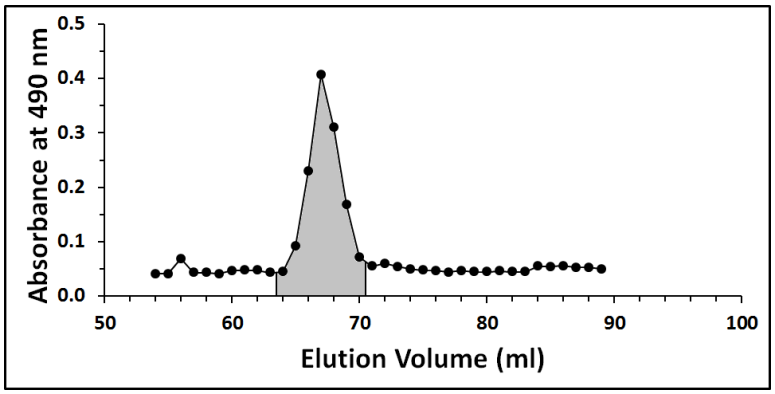
Calc. $[M + Na]^+$

*m*/*z* = 1409.4




SEC

**Derivatization reaction mixture**

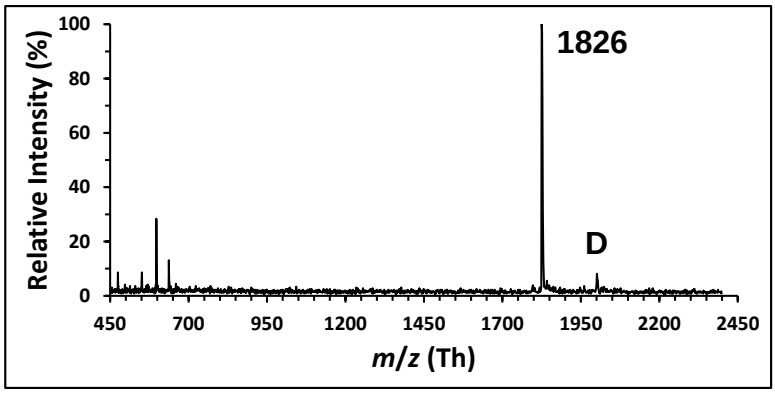



MALDI-TOF MS

**XXJG-APB**

Calc. $[M + Na]^+$

*m*/*z* = 1825. 7 Th



**Y**

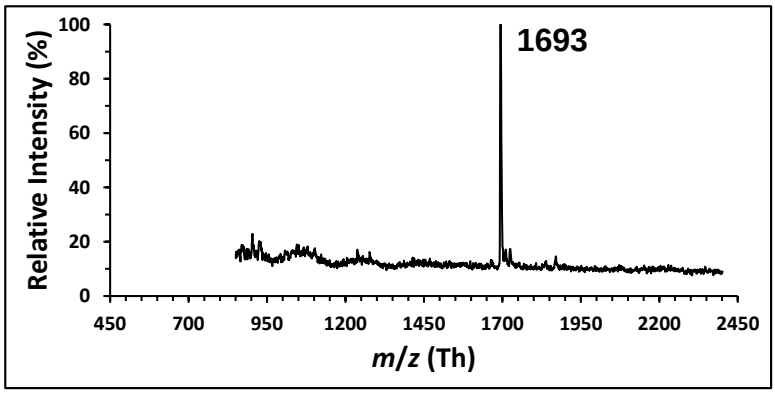
**Preparation and purification of GXJG-APB**

MALDI-TOF MS

**GXJG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1693.6



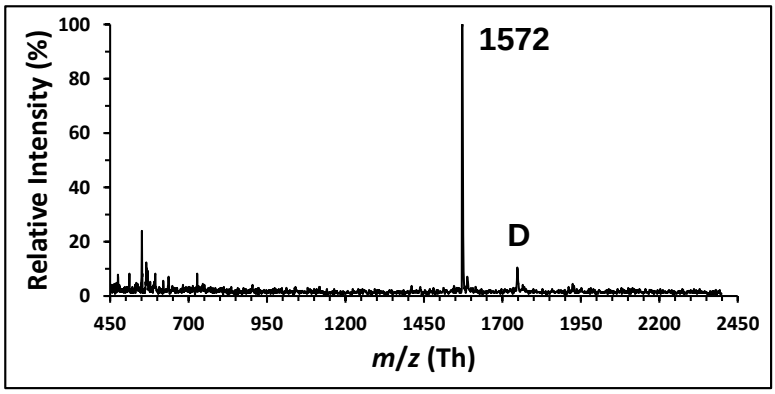
**Z** **Preparation and purification of XLJG-APB**

MALDI-TOF MS

**XLJG**

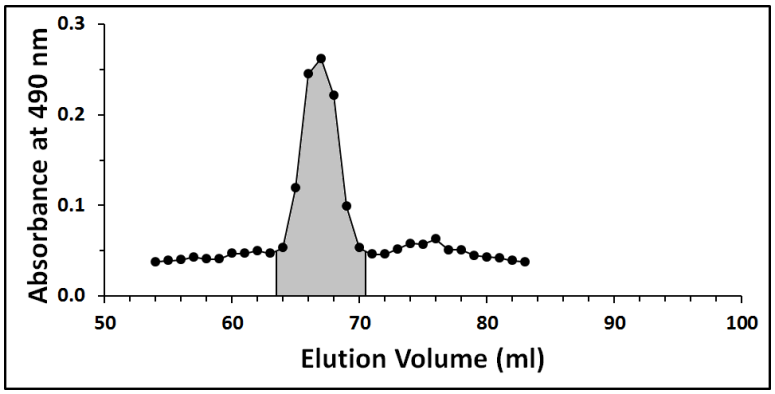
Calc. $[M + Na]^+$

*m*/*z* = 1571.5




SEC

**Derivatization reaction mixture**

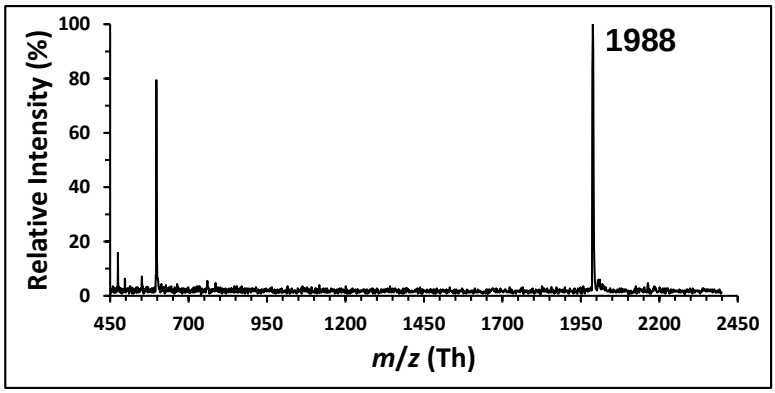



MALDI-TOF MS

**XLJG-APB**

Calc. $[M + Na]^+$

*m*/*z* = 1987.7



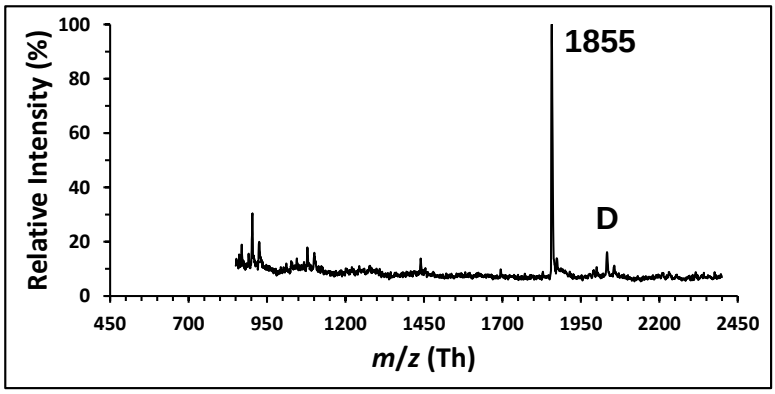
**AA** **Preparation and purification of GLJG-APB**

MALDI-TOF MS

**GLJG-APB**

Calc. $[M + Na]^+$

$m/z$ = 1855.7



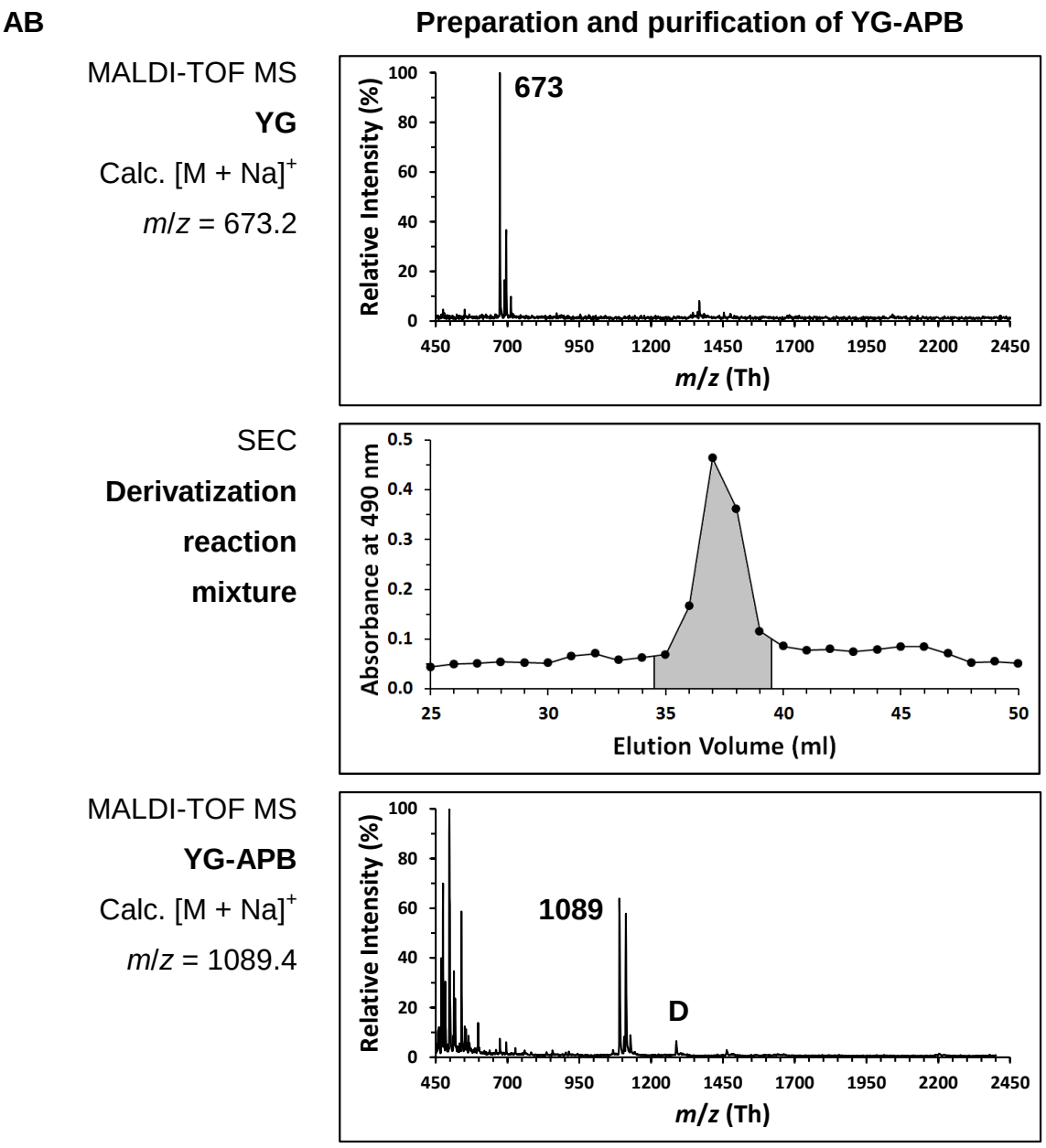



**Figure 3-S1.** Generation and purification of XyG S1 oligosaccharide-APB conjugates. For the oligosaccharide-APB conjugates that were prepared by covalent conjugation of XyG S1 oligosaccharide with the APB tag (for example XXXG-APB), the **top panel** shows a MALDI-TOF MS spectrum of the native oligosaccharide (with the measured *m*/*z* value indicated), the **middle panel**

shows a SEC chromatogram of the conjugation reaction products (with the chromatogram shading indicating the pooling of the fractions), and the **bottom panel** shows the MALDI-TOF MS spectrum of the final purified oligosaccharide-APB conjugate (with the measured *m*/*z* value indicated). For oligosaccharide-APB conjugates (for example GXXG-APB) that were generated by an α-xylosidase treatment of existing oligosaccharide-APB conjugate, only MALDI-TOF MS spectrum of the final conjugate is shown.

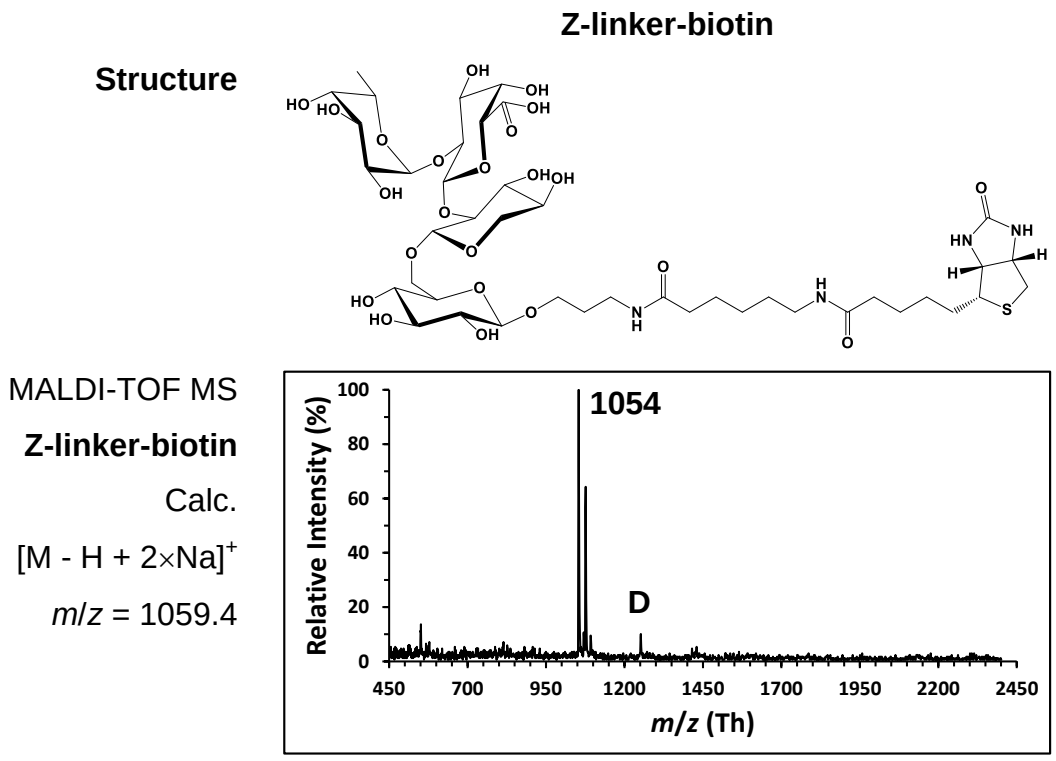


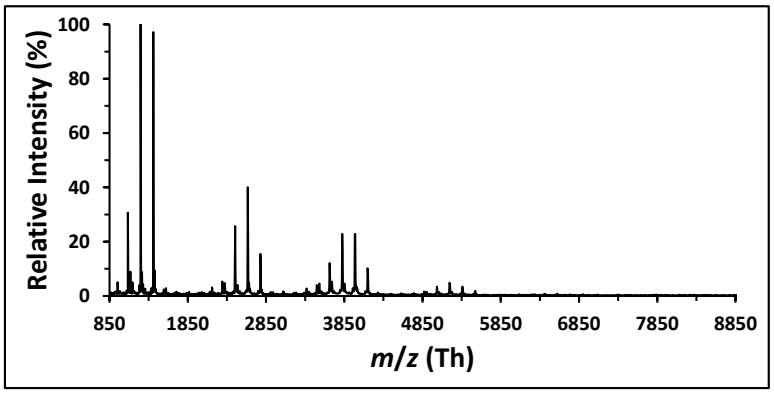

**Figure 3-S2.** Z-biotin conjugate. The chemical structure (**top panel**) and MALDI-TOF MS spectrum (**bottom panel**) of the X-biotin conjugate are shown. Note that the aglycone structure differs from that of the other biotin conjugates used in this dissertation (for comparison, see Supplemental Figure 1A). The major species detected by MALDI-TOF MS is the doubly sodiated species where one of the Na atoms replaces a proton (H) in the carboxylic group (-COOH) to form sodium carboxylate (-COONa) and the other forms an adduct with the glycoconjugate salt to increase the charge of the complex.

**A** **Preparation and Purification of XyG S*n*-APBs**

MALDI-TOF MS

**Tamarind XyG**

**S*n* mixture**

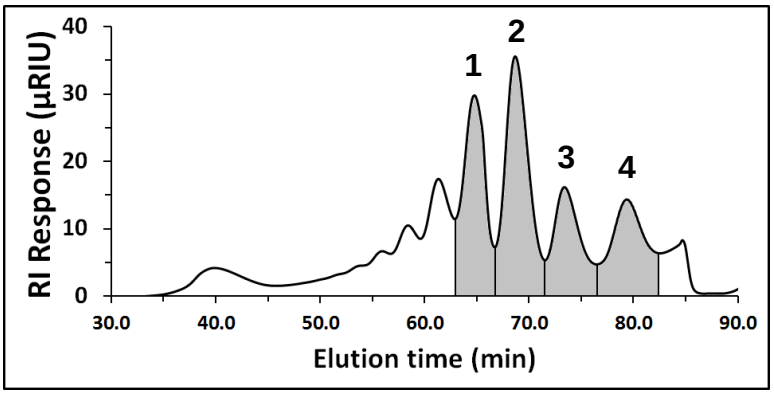



**B** SEC

**Derivatization**

**reaction**

**mixture**

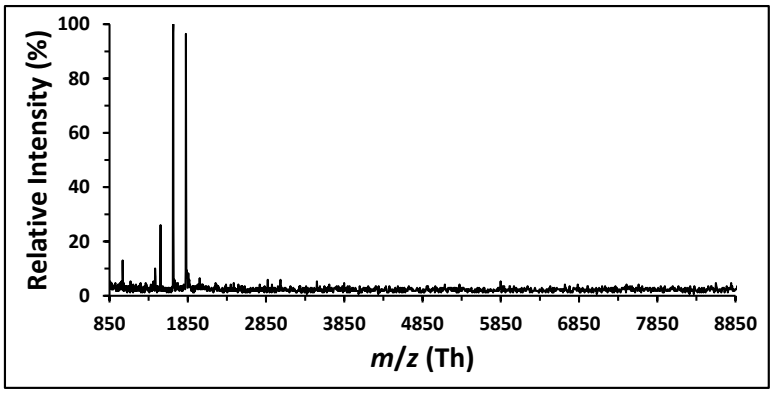



**C** MALDI-TOF MS

**Tamarind XyG**

**S1-APBs**

**(Peak 1)**

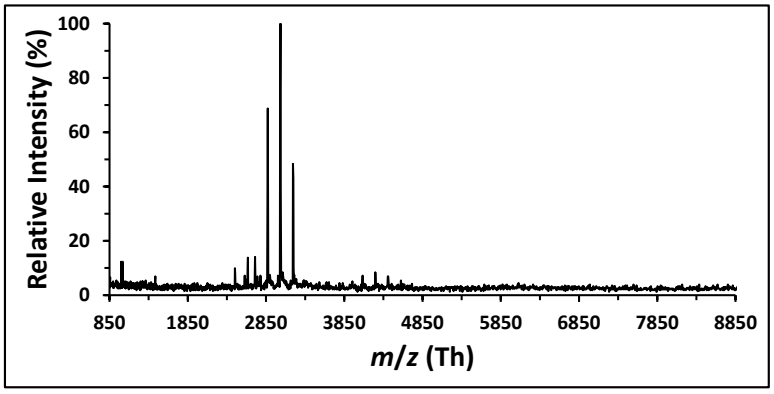



**D** MALDI-TOF MS

**Tamarind XyG**

**S2-APBs**

**(Peak 2)**



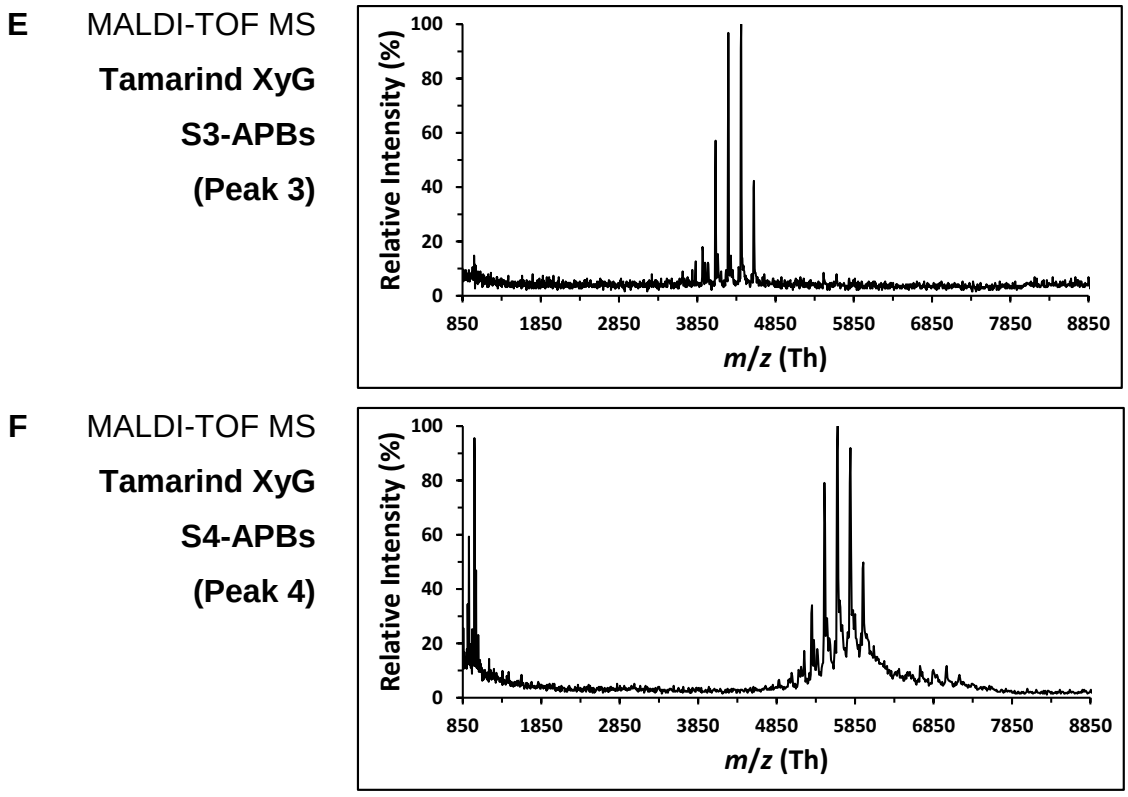



**Figure 3-S3.** Generation and purification of XyG S*n* oligosaccharide-APB conjugates. (**A**) MALDI-TOF MS spectrum of products from incomplete XEG hydrolysis of tamarind XyG polysaccharide. (**B**) SEC chromatogram of the conjugation reaction products mixture with fractions pooled as indicated by the chromatogram shading. MALDI-TOF MS spectra of the pooled tamarind XyG S*n*-APB mixtures (**C**) S1-APB, (**D**) S2-APB, (**E**) S3-APB, (**F**) S4-APB.

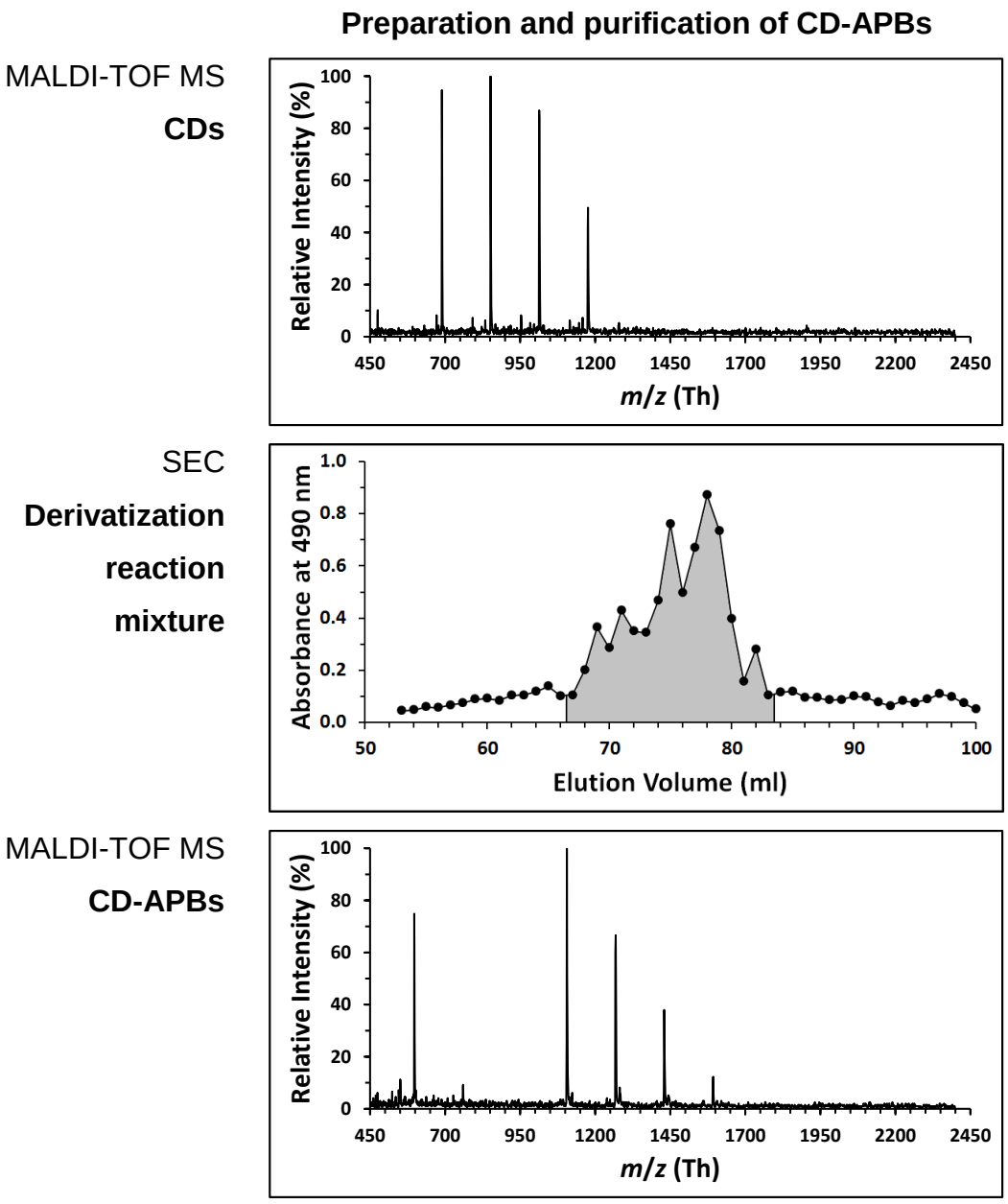



**Figure 3-S4.** Generation and purification of cellodextrin-APB conjugate mixture. (**A**) MALDI-TOF MS spectrum of the native cellodextrin mixture containging cellotetraose ($G_4$, *m*/*z* ), cellopentaose ($G_5$, *m*/*z* ), cellohexaose ($G_6$, *m*/*z* ) and celloheptaose ($G_7$, *m*/*z* ). (**B**) SEC chromatogram of the conjugation reaction products with fractions pooled as indicated by the chromatogram shading. (**C**) MALDI-TOF MS spectra of the pooled cellodextrin-APB mixture.

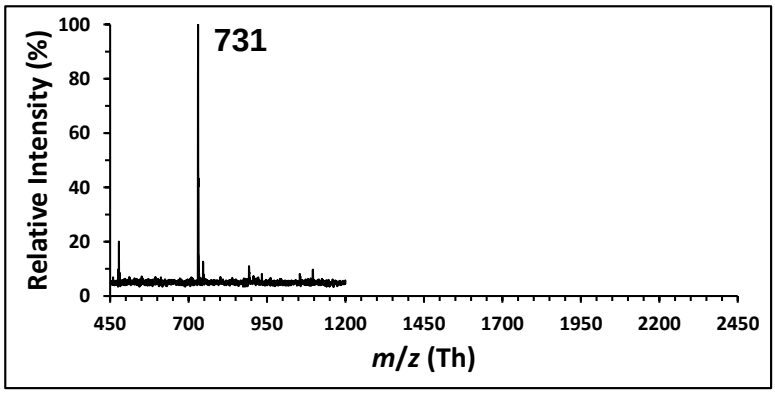
**A** **Preparation and purification of L*N*T-APB**

MALDI-TOF MS

**L*N*T**

Calc. $[M + Na]^+$

*m*/*z* = 730.6

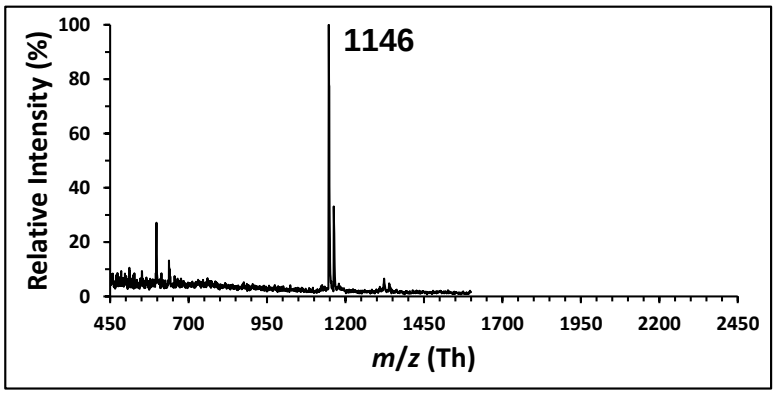



MALDI-TOF MS

**L*N*T-APB**

Calc. $[M + Na]^+$

*m*/*z* = 1146.8



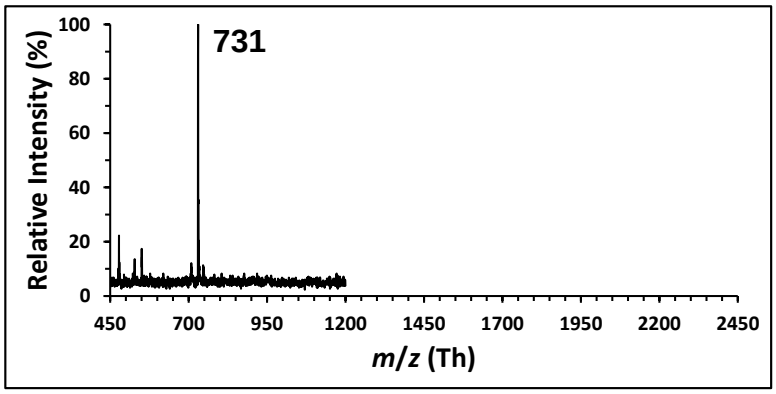
**B** **Preparation and purification of L*Nn*T-APB**

MALDI-TOF MS

**L*Nn*T**

Calc. $[M + Na]^+$

*m*/*z* = 730.8

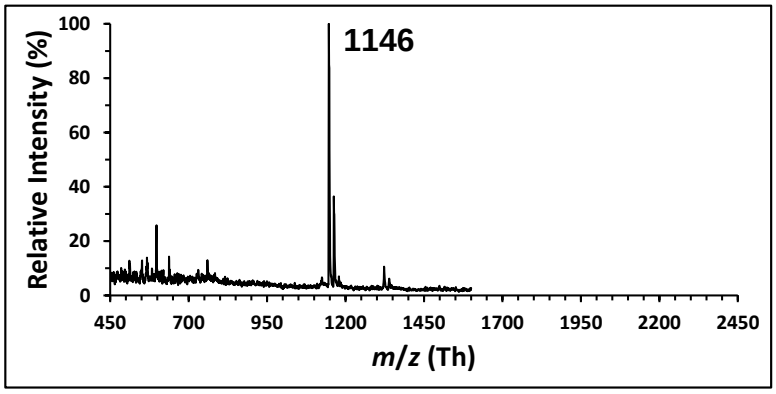



MALDI-TOF MS

**L*Nn*T-APB**

Calc. $[M + Na]^+$

*m*/*z* = 1146.8



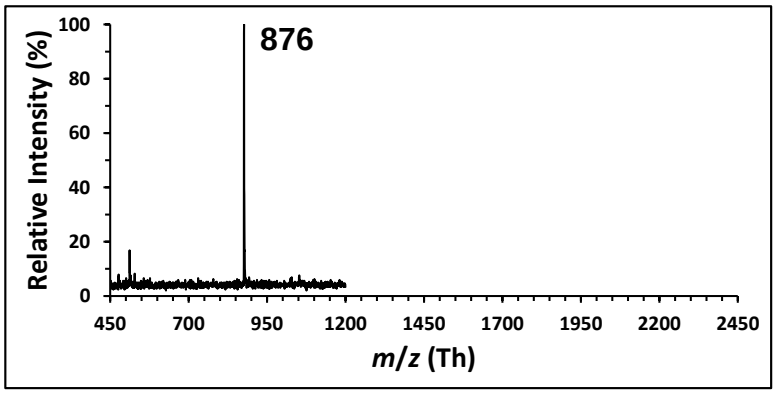
**C** **Preparation and purification of L*N*FPIII-APB**

MALDI-TOF MS

**L*N*FPIII**

Calc. $[M + Na]^+$

*m*/*z* = 876.8

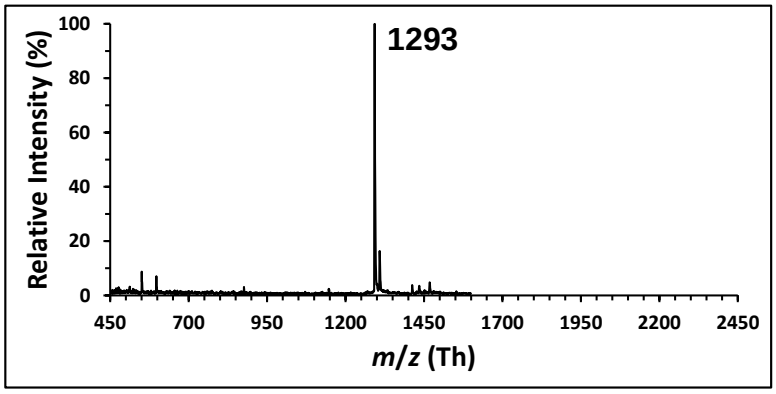



MALDI-TOF MS

**L*N*FPIII-APB**

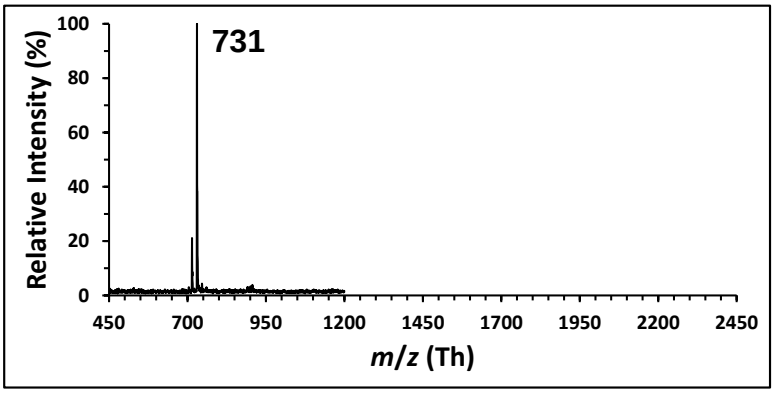
Calc. $[M + Na]^+$

*m*/*z* = 1293.0



**D** **Preparation and purification of β-D-Gal*p*-(1→3)-β-D-Gal*p*NAc-(1→4)-β-D-Gal*p*-(1→4)-D-Glc-APB**

MALDI-TOF MS
**Oligosaccharide**
Calc. $[M + Na]^+$
*m*/*z* = 730.6




MALDI-TOF MS
**Oligosaccharide-APB**
Calc. $[M + Na]^+$
*m*/*z* = 1146.8

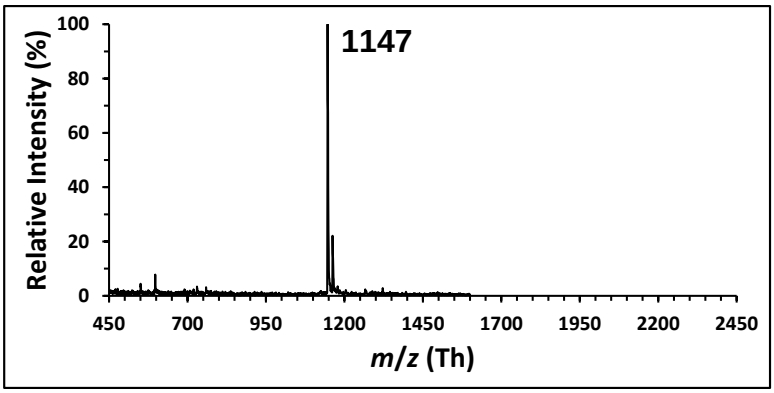
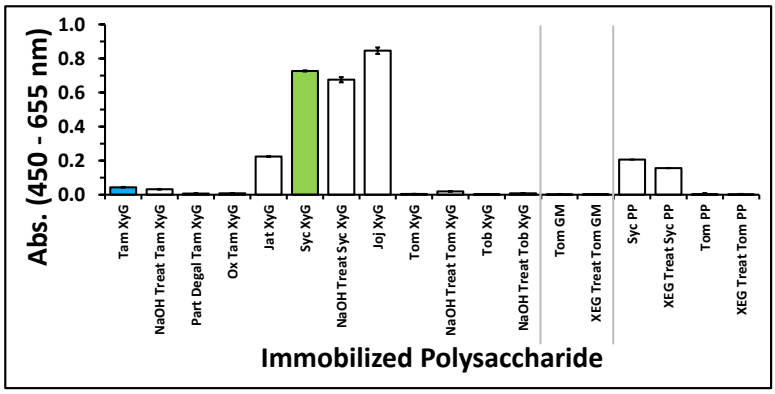



**Figure 3-S5.** Generation and purification of the APB conjugates of oligosaccharides that are typically found as glycoconjugates with proteins and lipids. MALDI-TOF MS spectra of the native oligosaccharide (**top panel**) and conjugation reaction products (**bottom panel**) are shown. (**A**) Lacto-*N*-tetraose-APB (L*N*T-APB). (**B**) Lacto-*N*-*neo*tetraose-APB (L*Nn*T-APB). (**C**) Lacto-*N*-fucopentaose III-APB (L*N*FPIII-APB). (**D**) β-D-Galp-(1→3)-β-D-GalpNAc-(1→4)-β-D-Galp-(1→4)-D-Glc-APB.

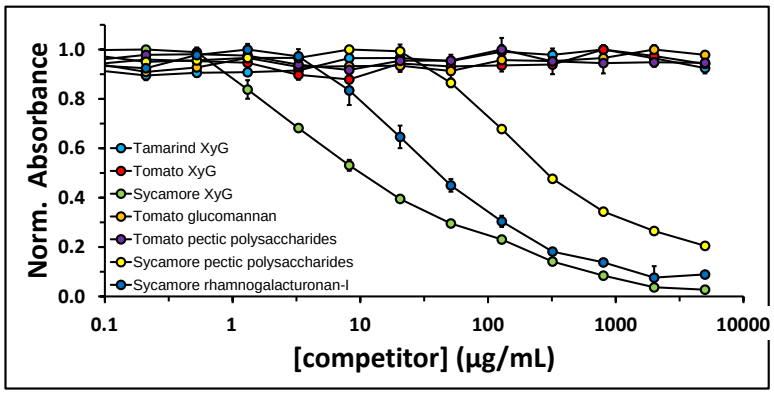
**A**

**CCRC-M1**

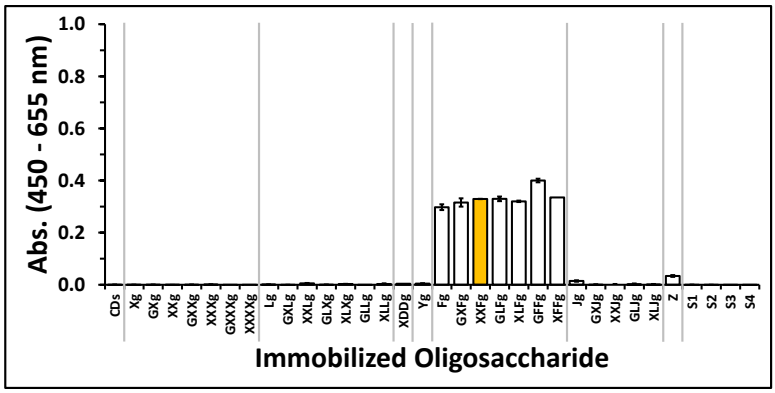
**Direct, single datapoint polysaccharide ELISA**



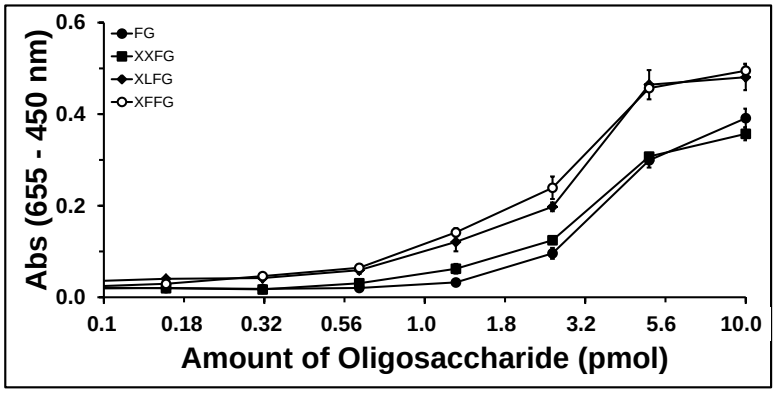

**Competitive, polysaccharide dilution series ELISA**




**Direct, single datapoint oligosaccharide ELISA**



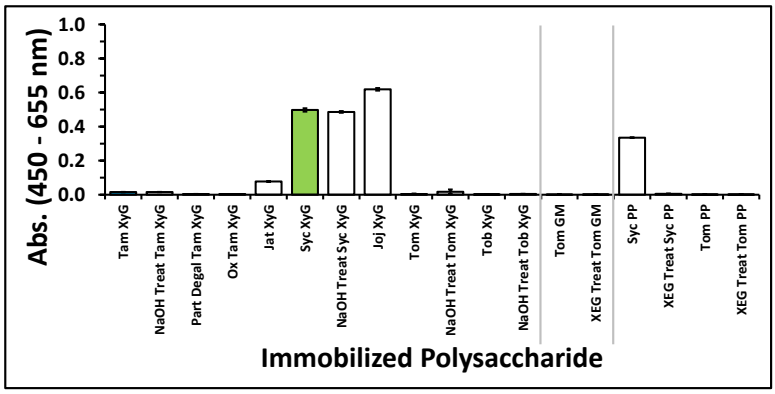

**Direct, oligosaccharide dilution series ELISA**



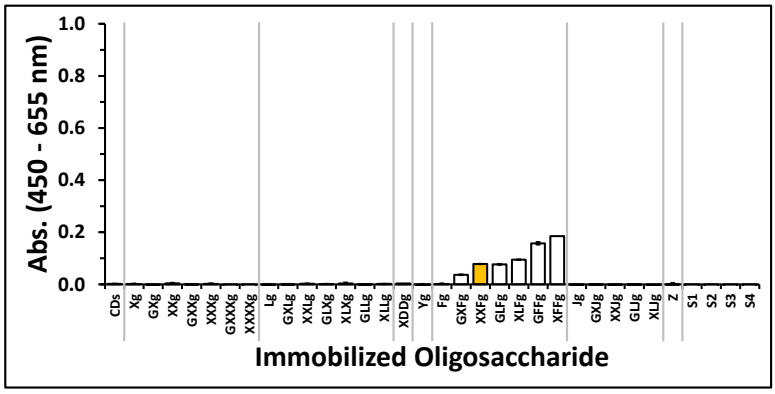
**B**

**CCRC-M39**

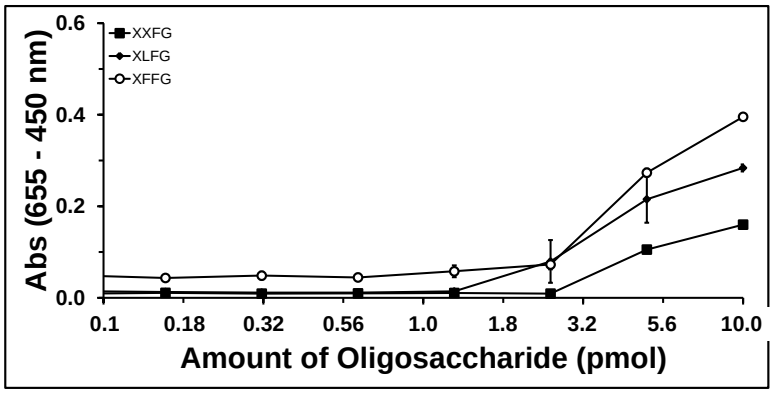
**Direct, single datapoint polysaccharide ELISA**




**Direct, single datapoint oligosaccharide ELISA**



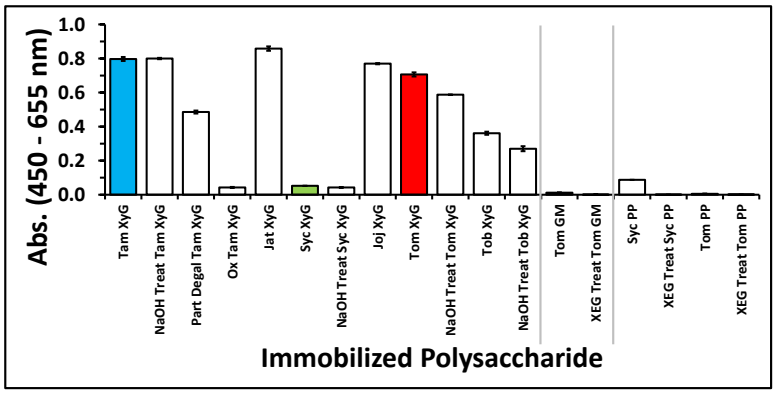

**Direct, oligosaccharide dilution series ELISA**



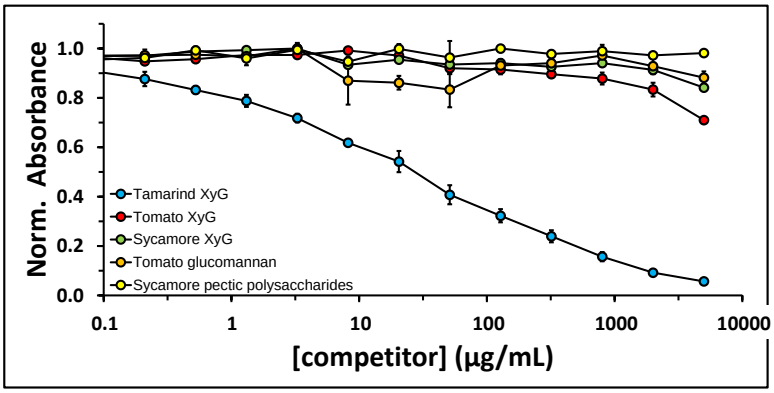
**C**

**CCRC-M48**

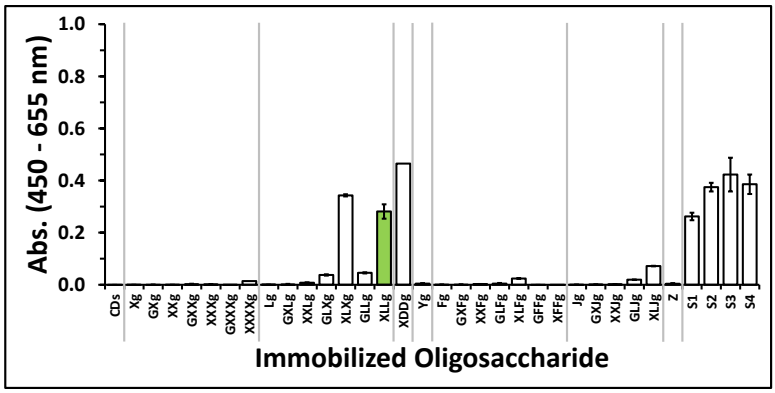
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



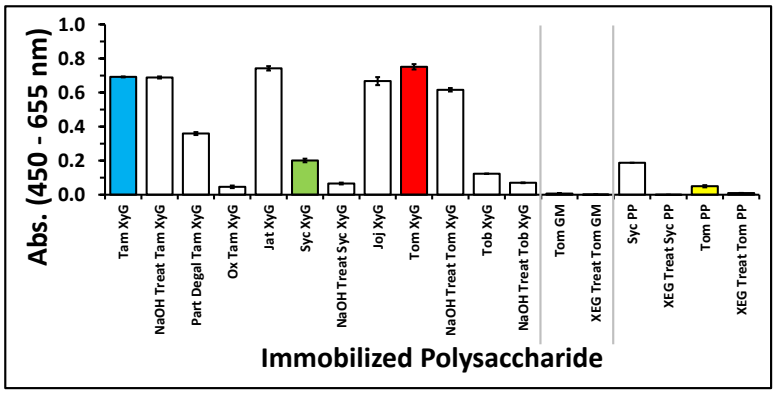

**Direct, single datapoint oligosaccharide ELISA**



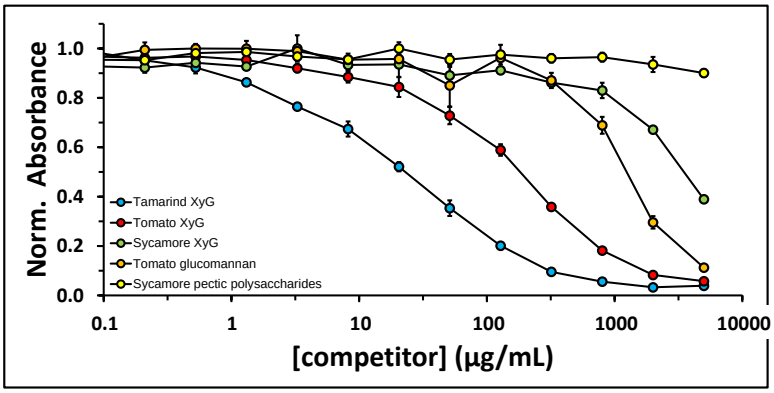
**D**

**CCRC-M49**

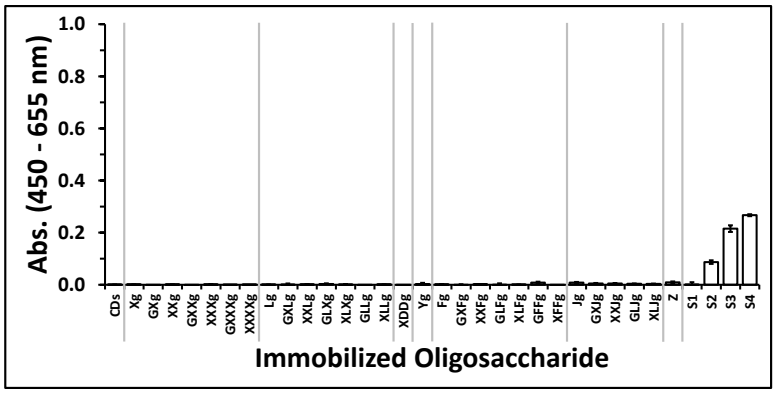
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



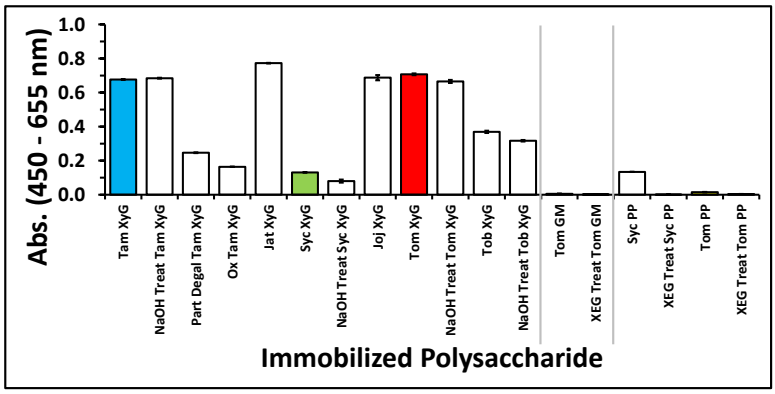

**Direct, single datapoint oligosaccharide ELISA**



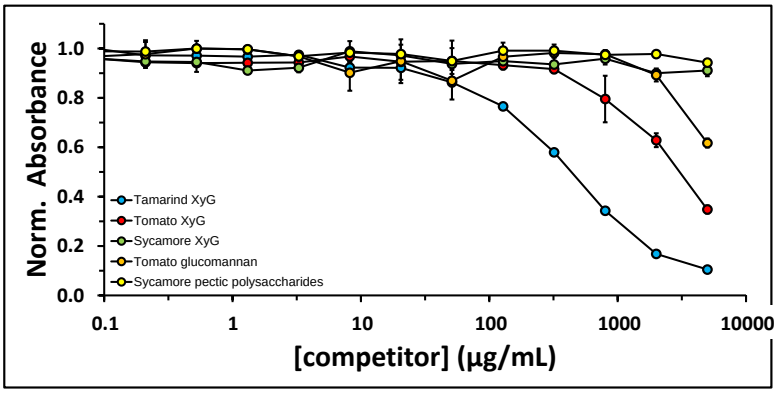
**E**

**CCRC-M50**

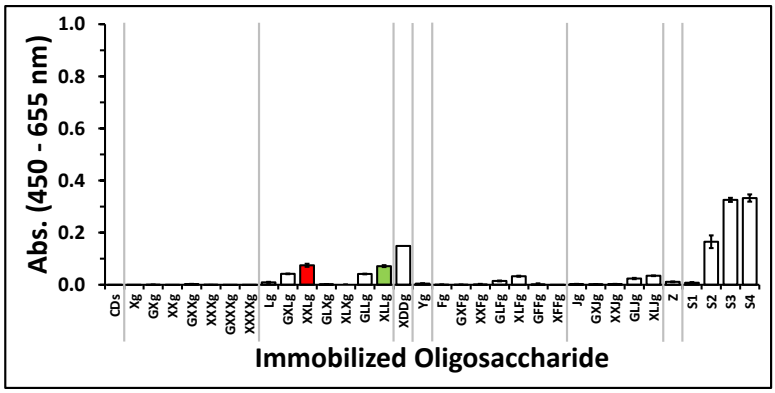
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



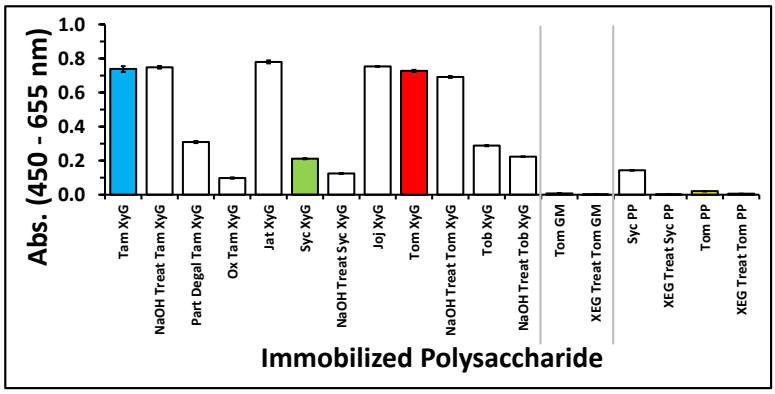

**Direct, single datapoint oligosaccharide ELISA**



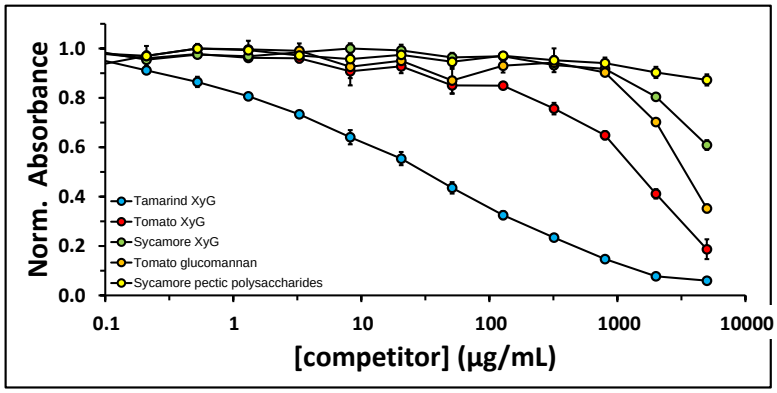
**F**

**CCRC-M51**

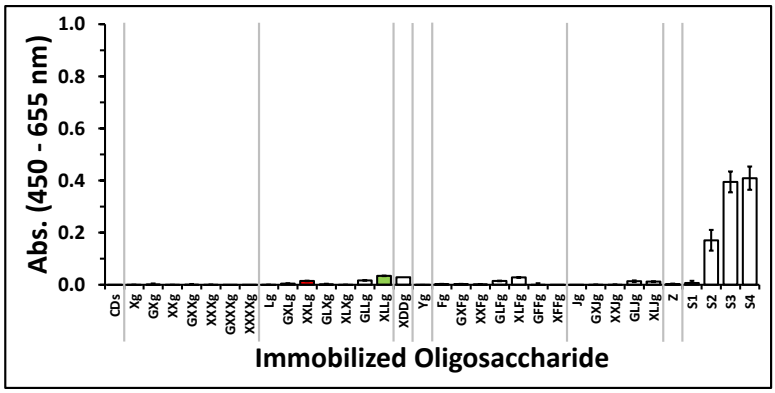
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



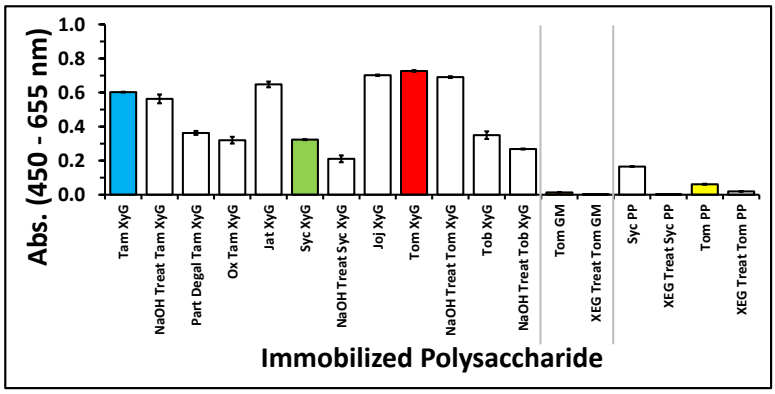

**Direct, single datapoint oligosaccharide ELISA**



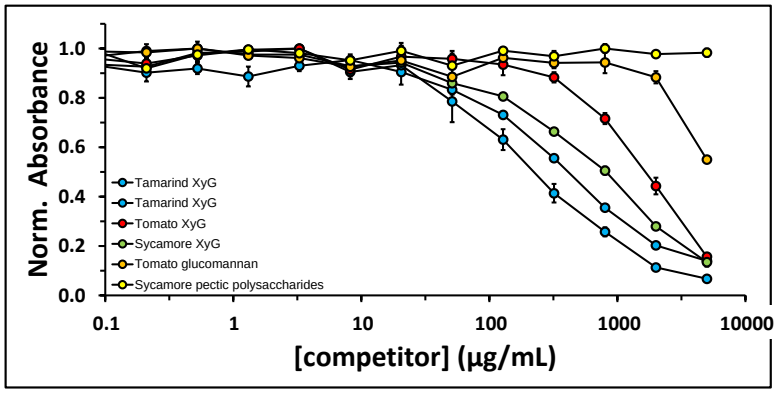
**G**

**CCRC-M52**

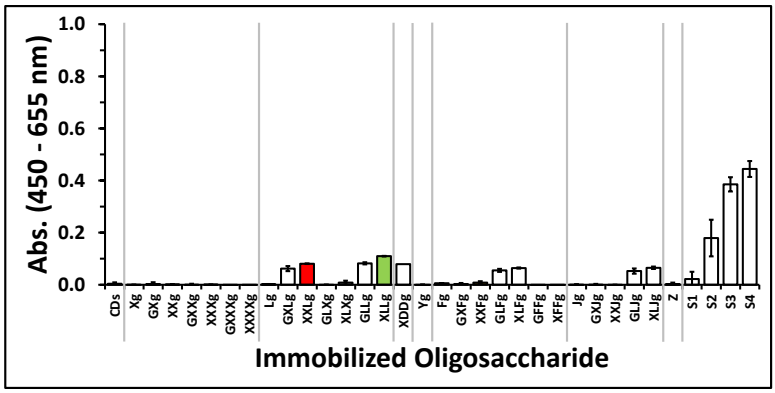
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



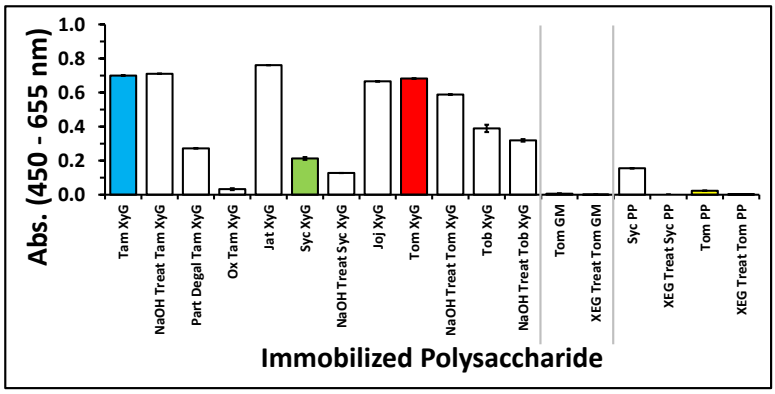

**Direct, single datapoint oligosaccharide ELISA**



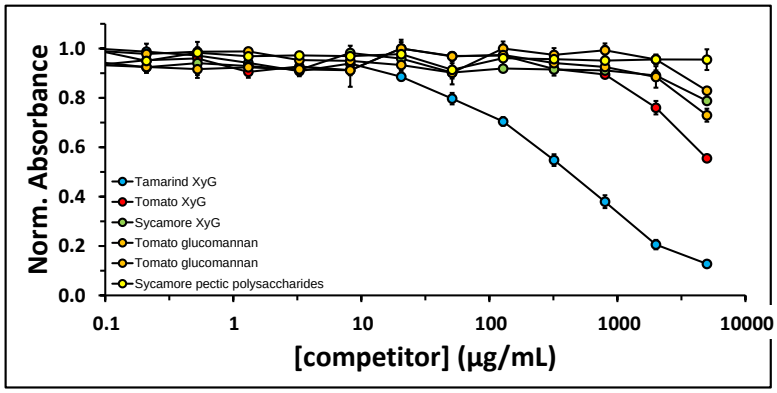
**H**

**CCRC-M53**

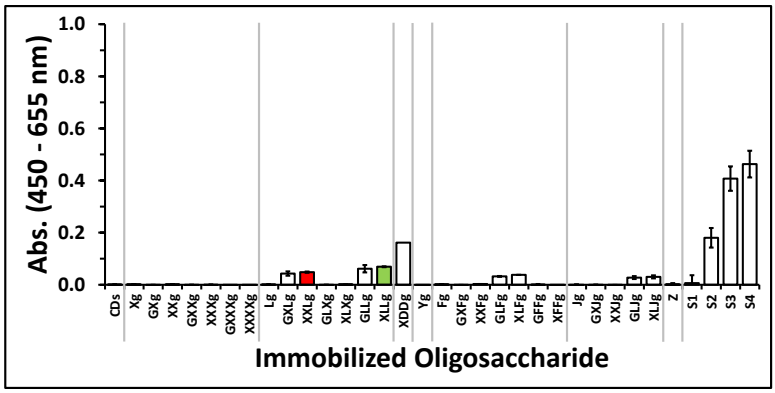
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



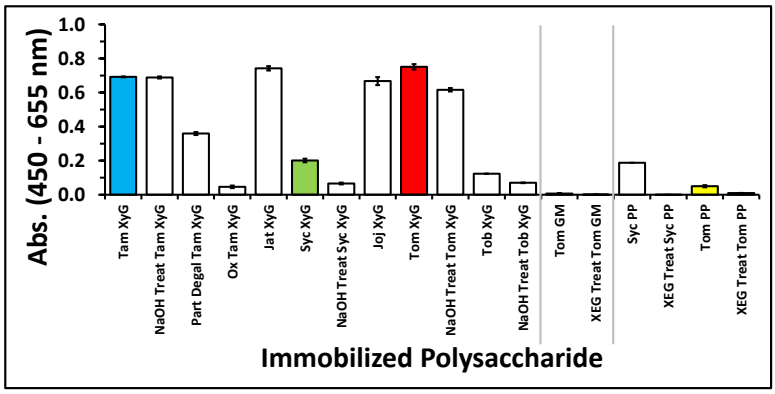

**Direct, single datapoint oligosaccharide ELISA**



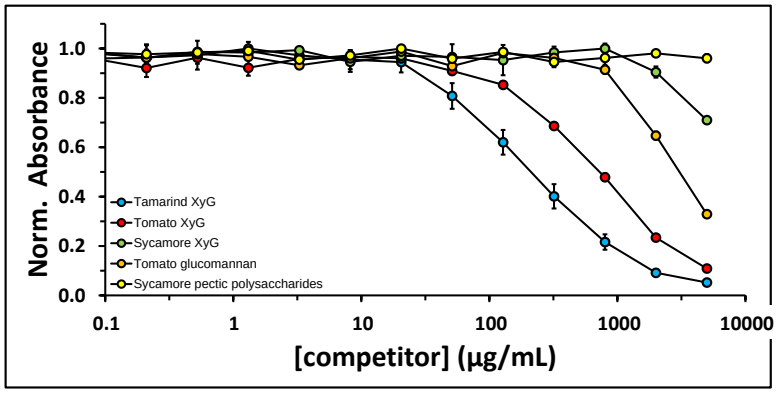
I

CCRC-M54

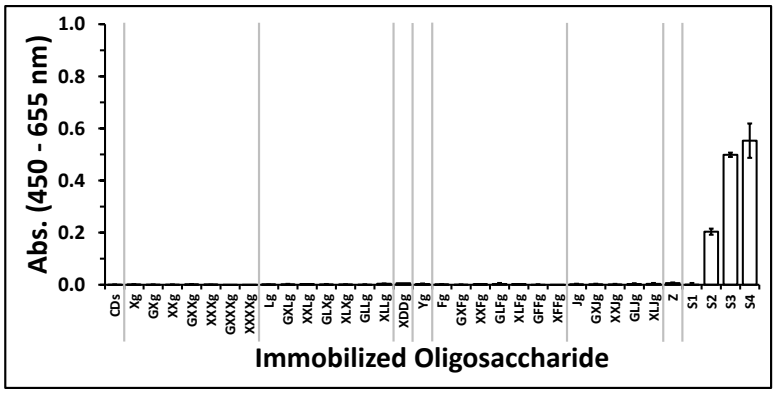
Direct, single datapoint polysaccharide ELISA




Competitive, polysaccharide dilution series ELISA



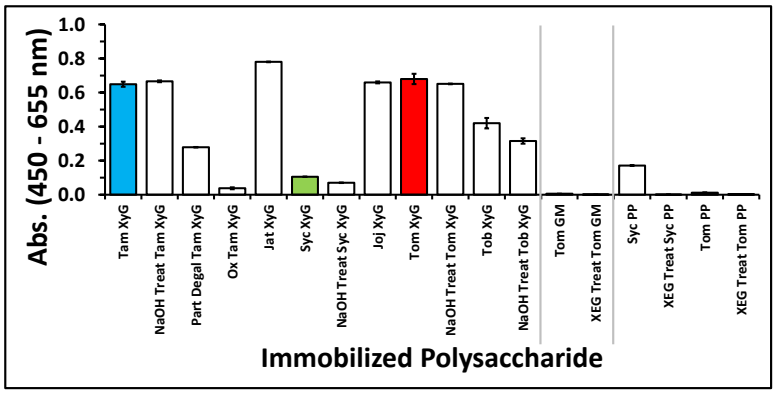

Direct, single datapoint oligosaccharide ELISA



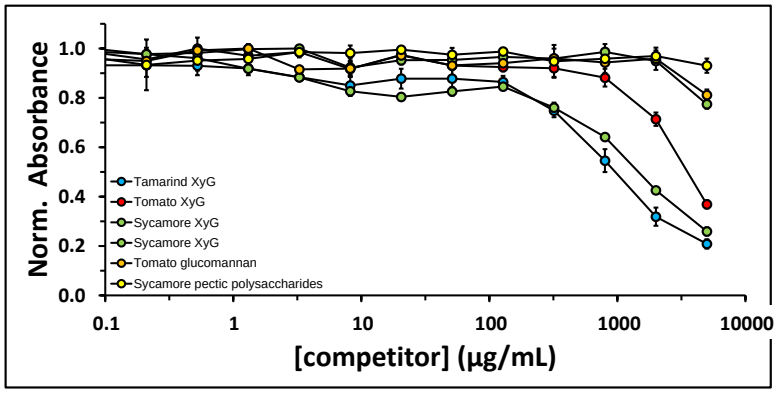
**J**

**CCRC-M55**

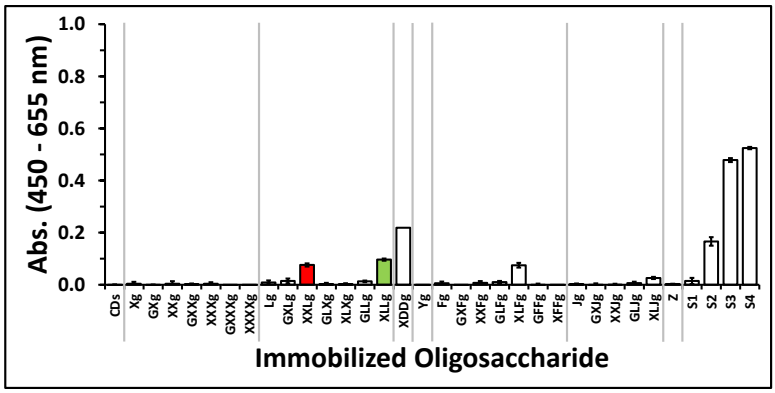
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



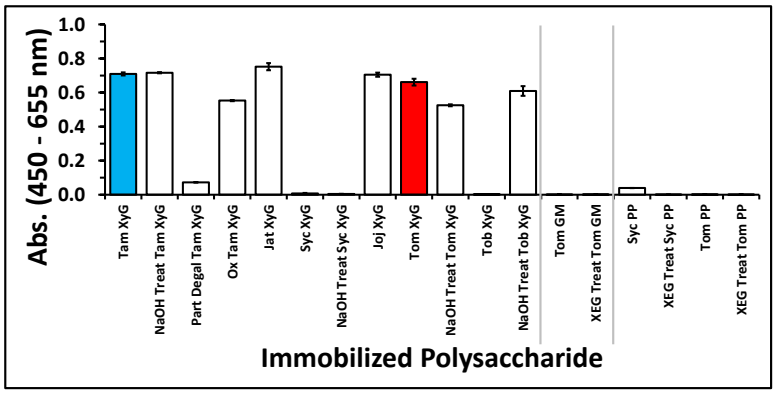

**Direct, single datapoint oligosaccharide ELISA**



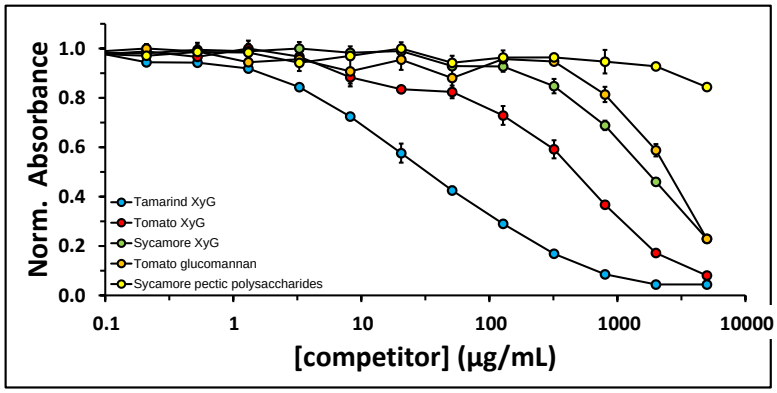
**K**

**CCRC-M57**

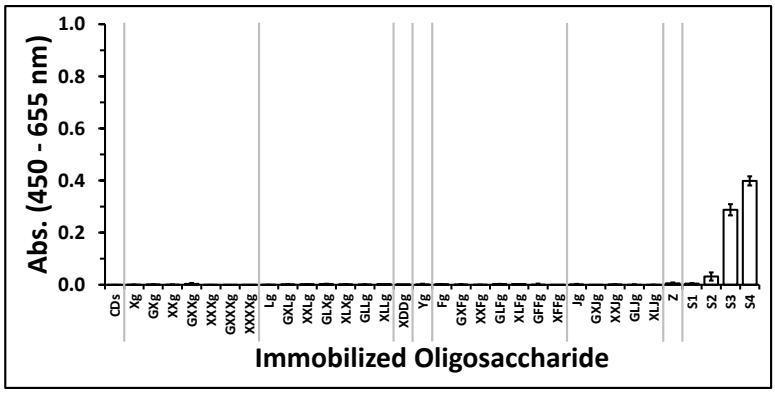
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



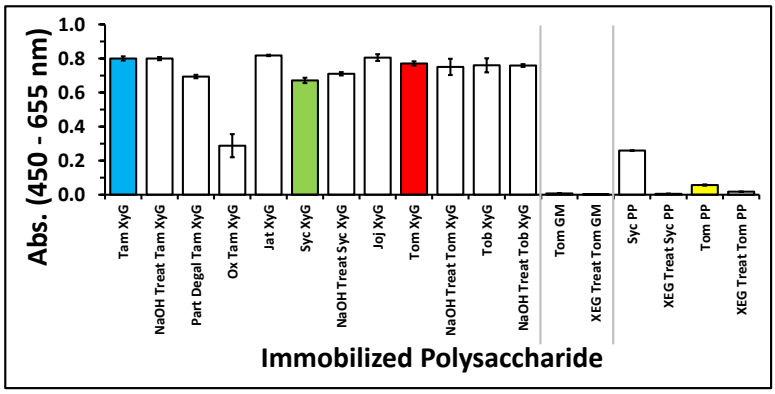

**Direct, single datapoint oligosaccharide ELISA**



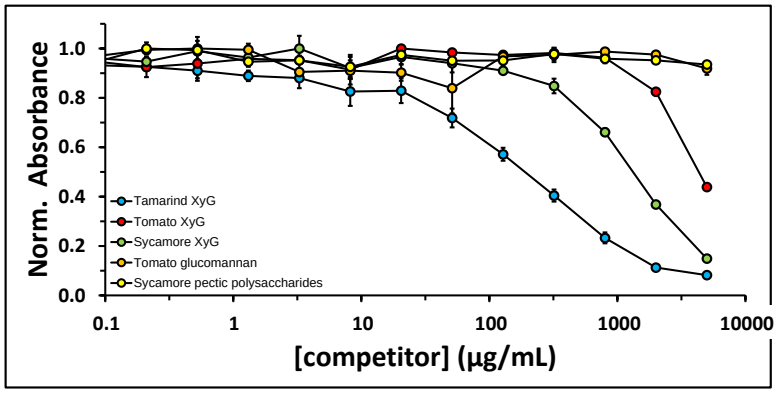
**L**

**CCRC-M58**

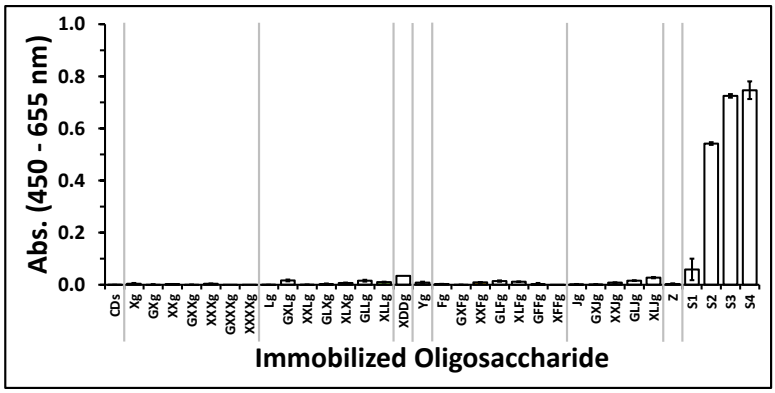
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



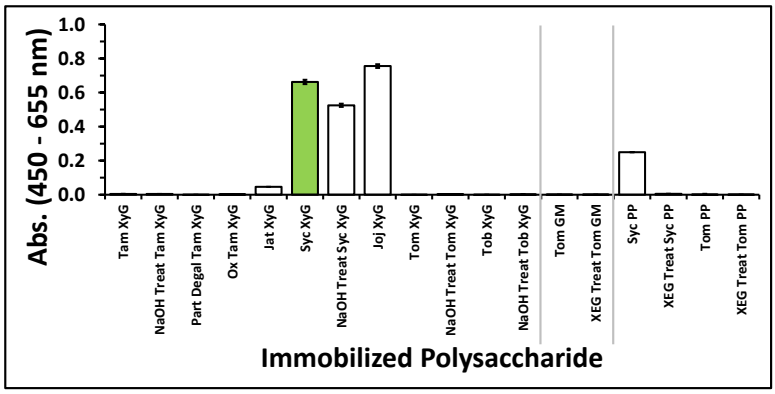

**Direct, single datapoint oligosaccharide ELISA**



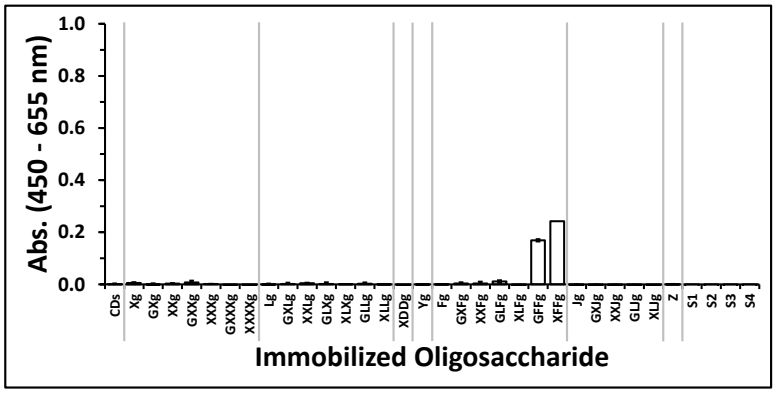
**M**

**CCRC-M84**

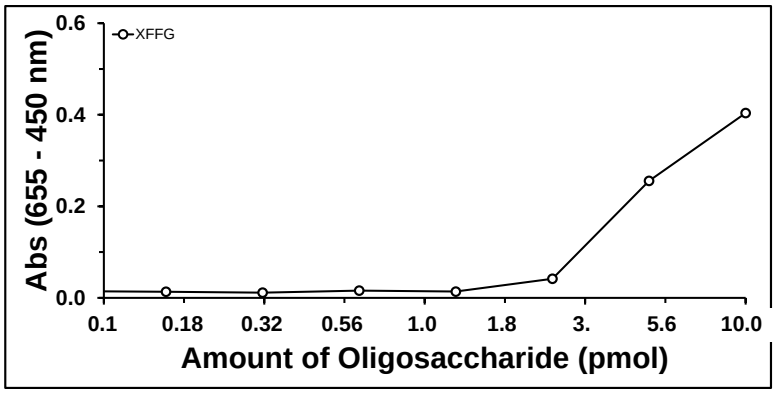
**Direct, single datapoint polysaccharide ELISA**




**Direct, single datapoint oligosaccharide ELISA**



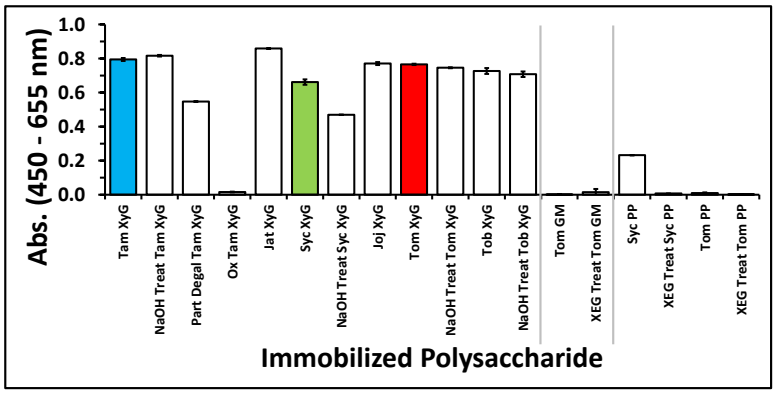

**Direct, oligosaccharide dilution series ELISA**



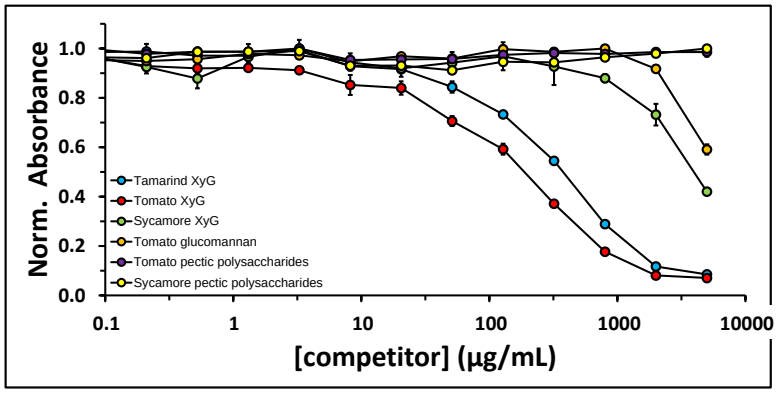
N

CCRC-M86

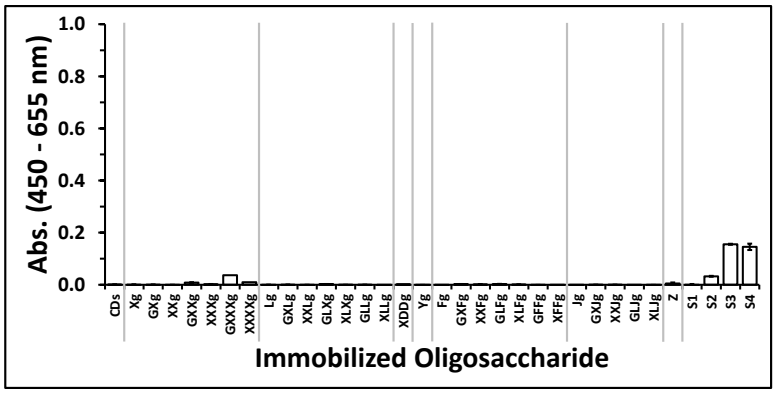
Direct, single datapoint polysaccharide ELISA




Competitive, polysaccharide dilution series ELISA



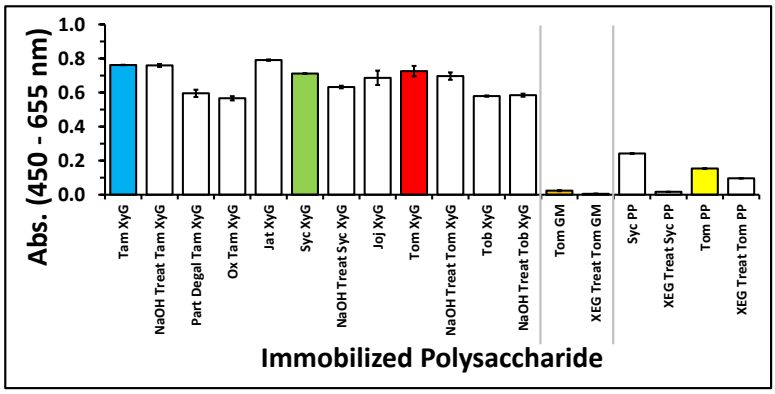

Direct, single datapoint oligosaccharide ELISA



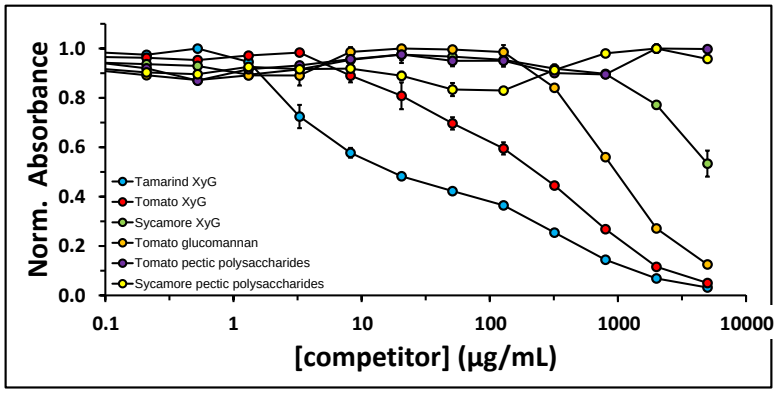
O

CCRC-M87

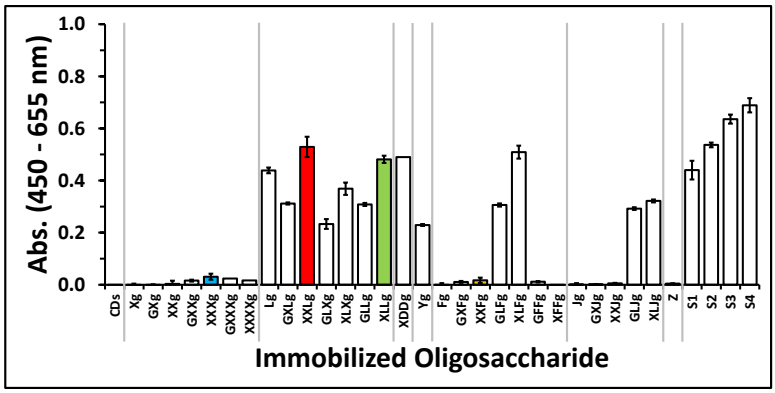
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



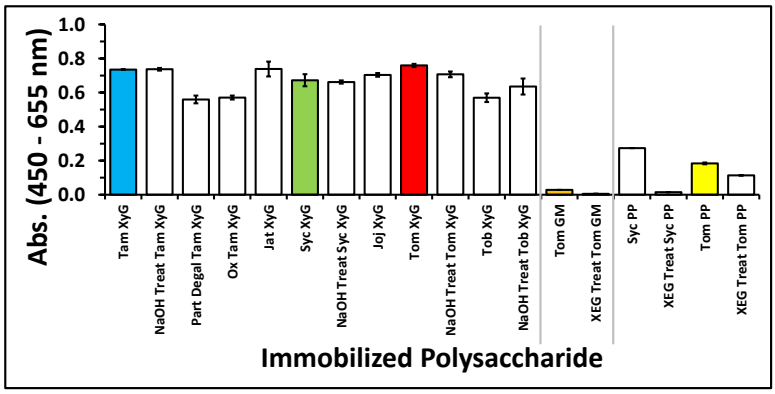

**Direct, single datapoint oligosaccharide ELISA**



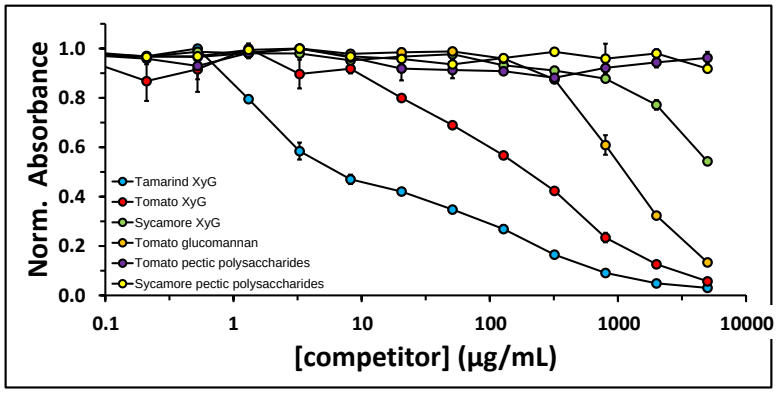
**P**

**CCRC-M88**

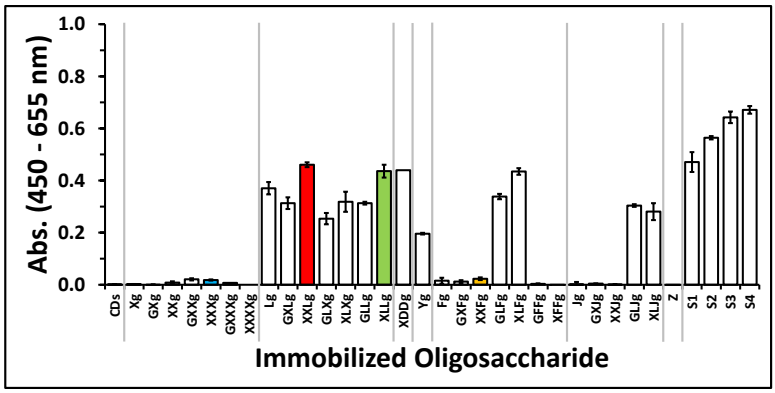
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



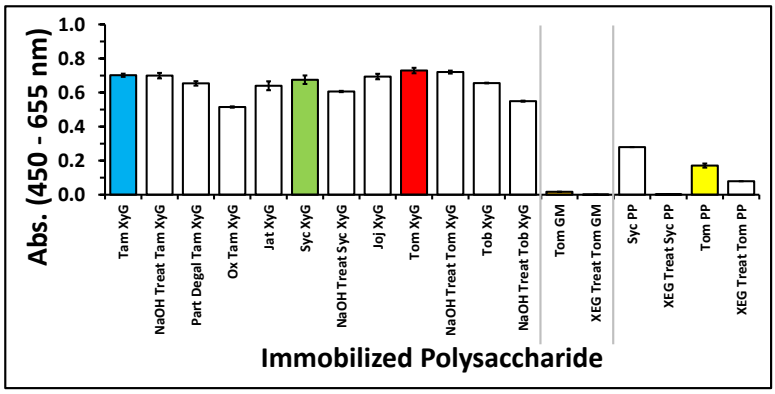

**Direct, single datapoint oligosaccharide ELISA**



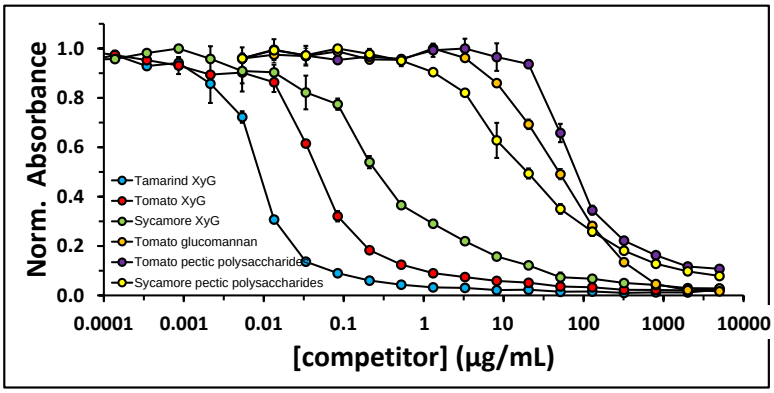
**Q**

**CCRC-M89**

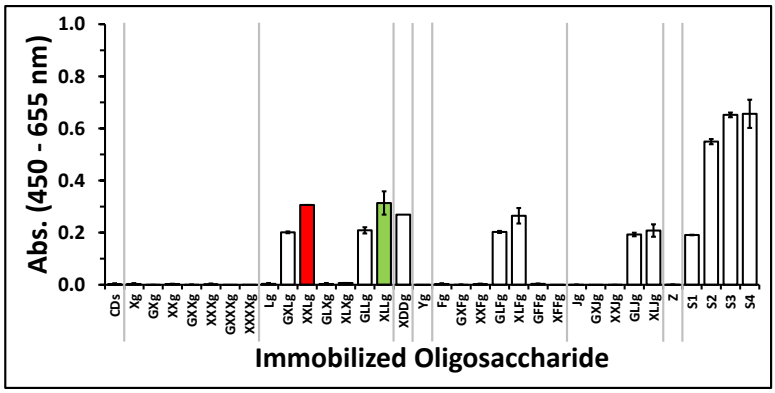
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



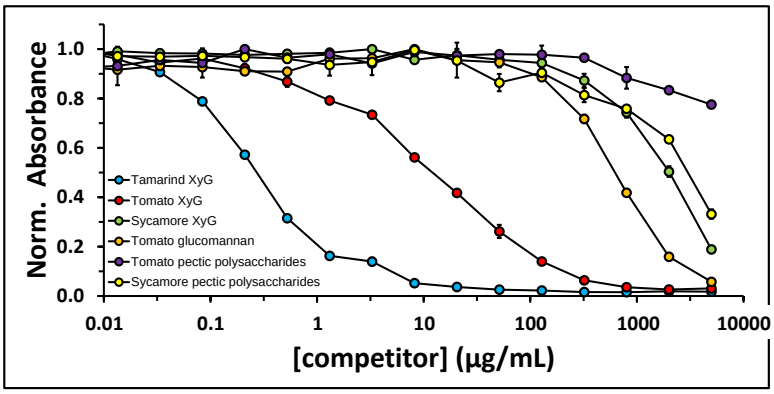

**Direct, single datapoint oligosaccharide ELISA**



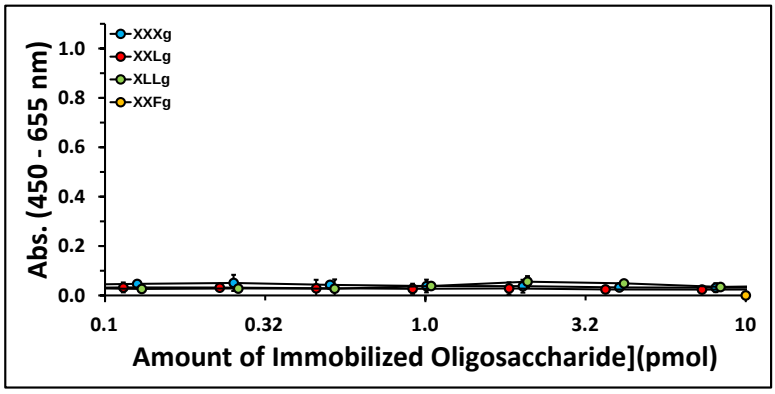
**R**

**CCRC-M90**

**Competitive, polysaccharide dilution series ELISA**



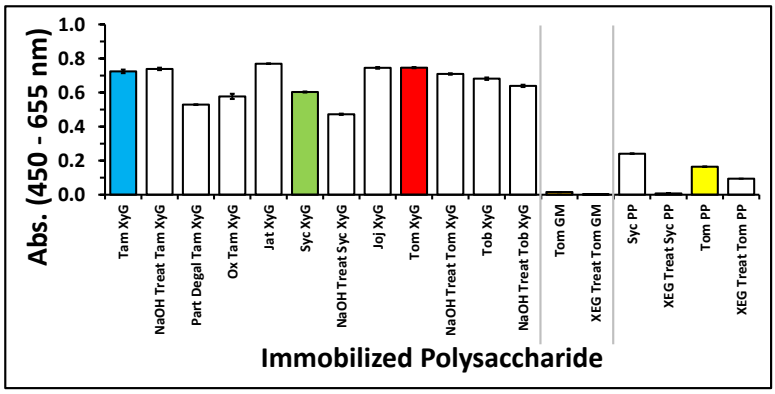

**Direct, oligosaccharide dilution series ELISA**



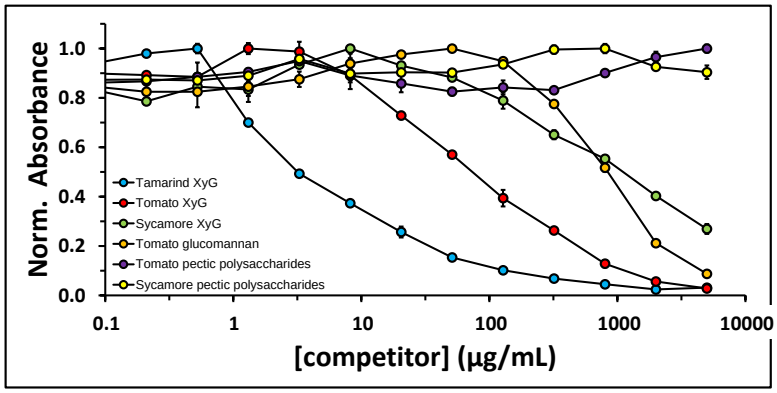
**S**

**CCRC-M93**

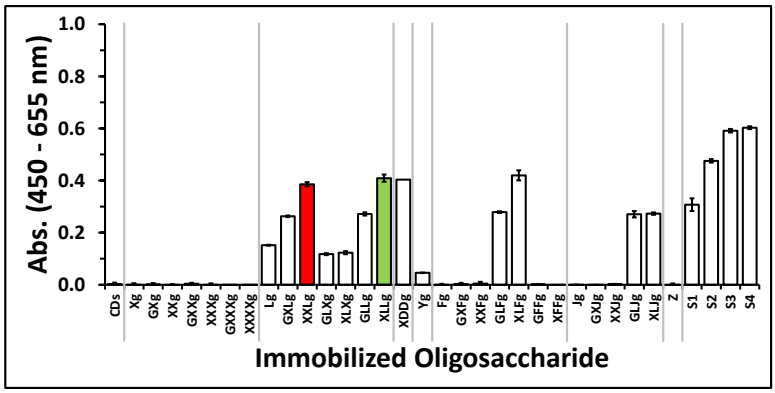
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



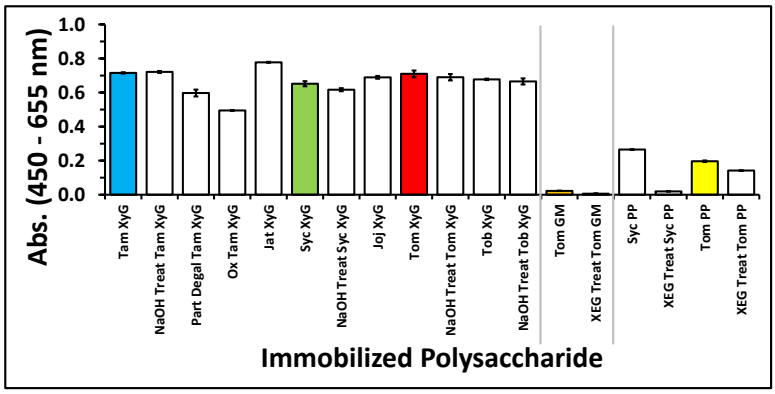

**Direct, single datapoint oligosaccharide ELISA**



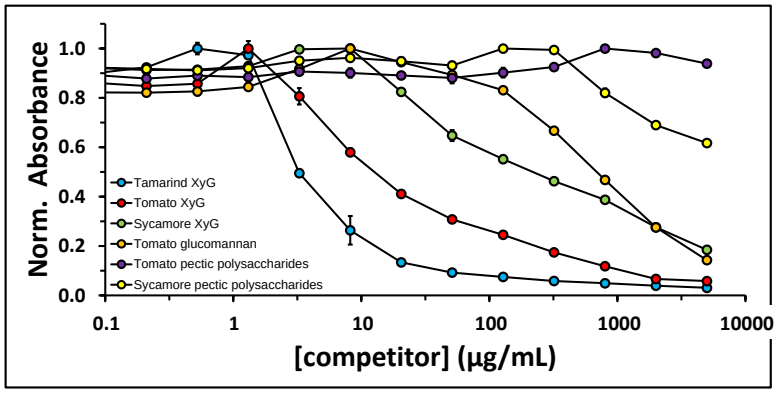
T

CCRC-M95

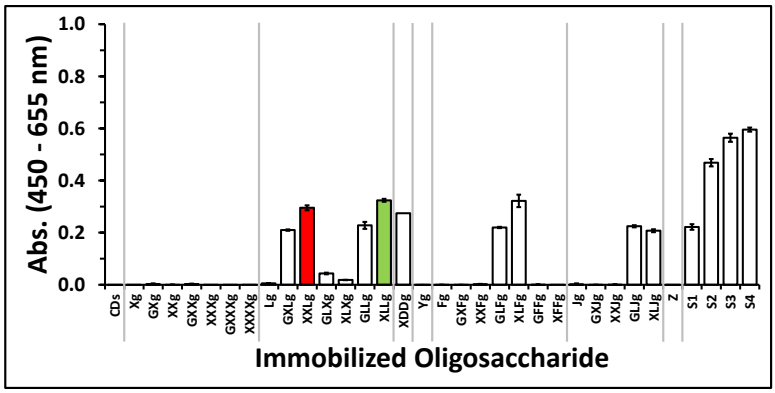
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



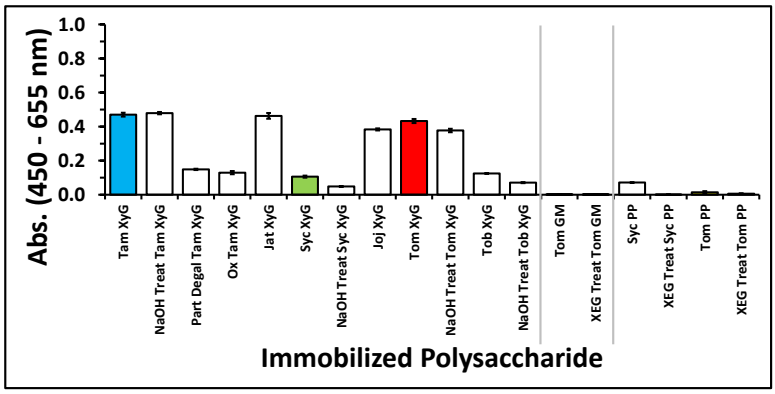

**Direct, single datapoint oligosaccharide ELISA**



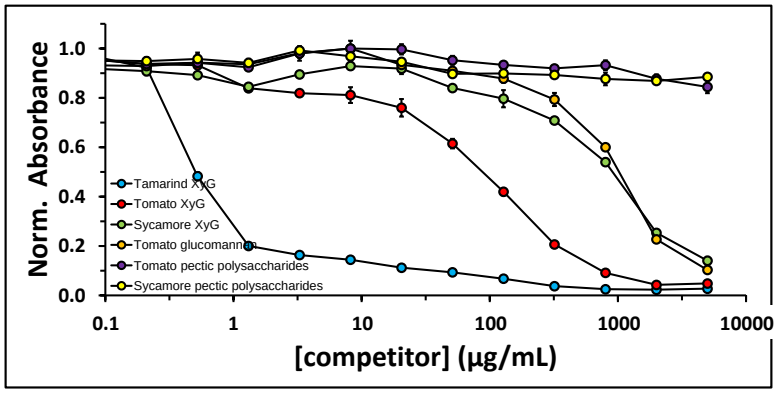
**U**

**CCRC-M96**

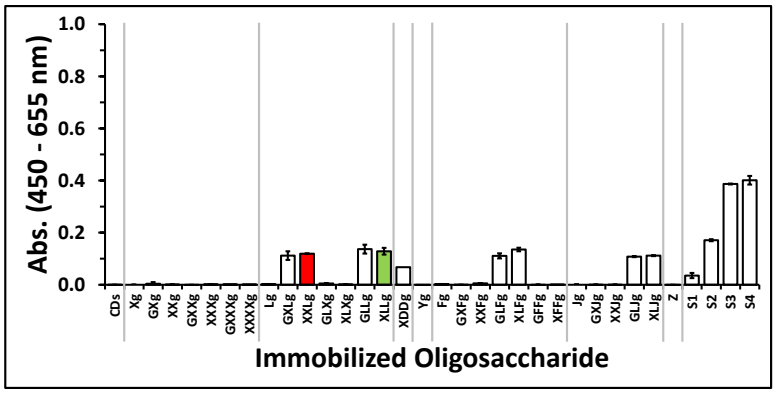
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



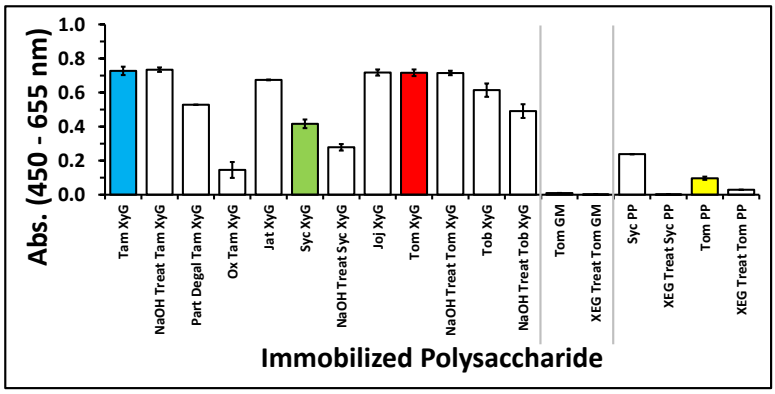

**Direct, single datapoint oligosaccharide ELISA**



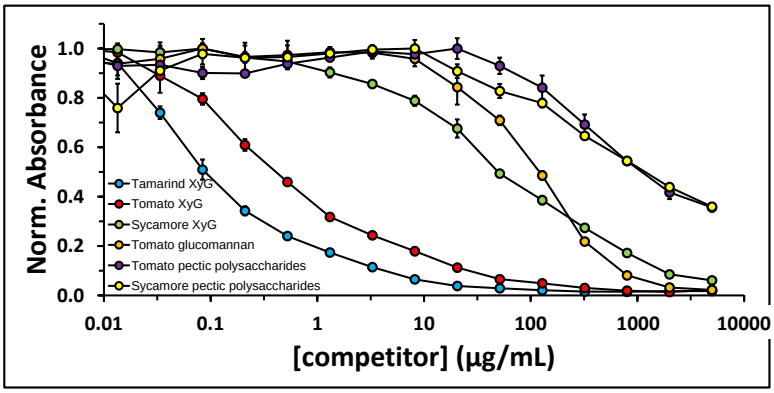
V

CCRC-M99

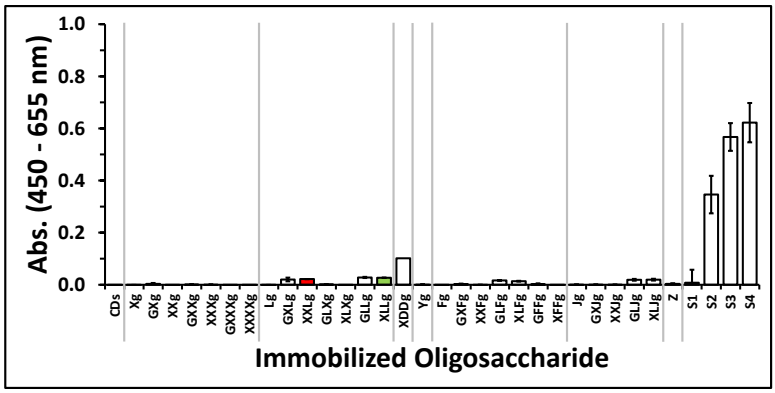
Direct, single datapoint polysaccharide ELISA




Competitive, polysaccharide dilution series ELISA



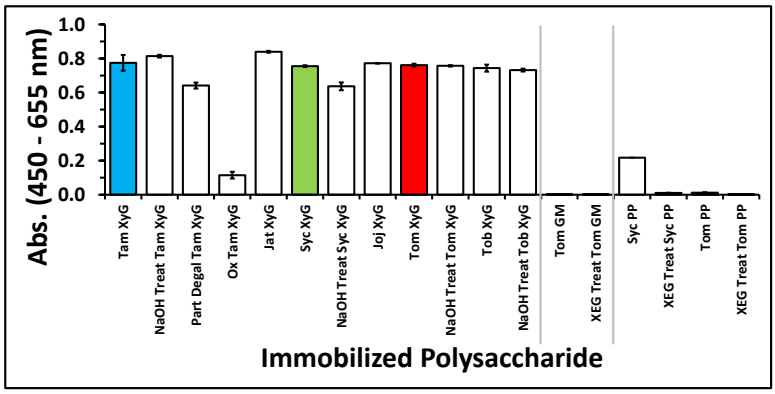

Direct, single datapoint oligosaccharide ELISA



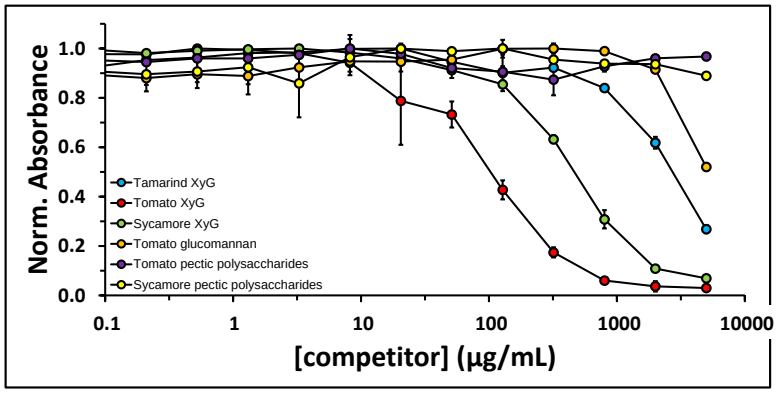
**W**

**CCRC-M100**

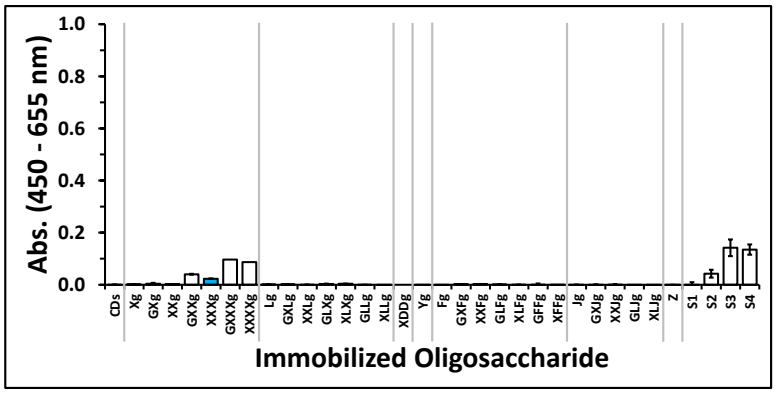
**Direct, single datapoint polysaccharide ELISA**




**Competitive, polysaccharide dilution series ELISA**



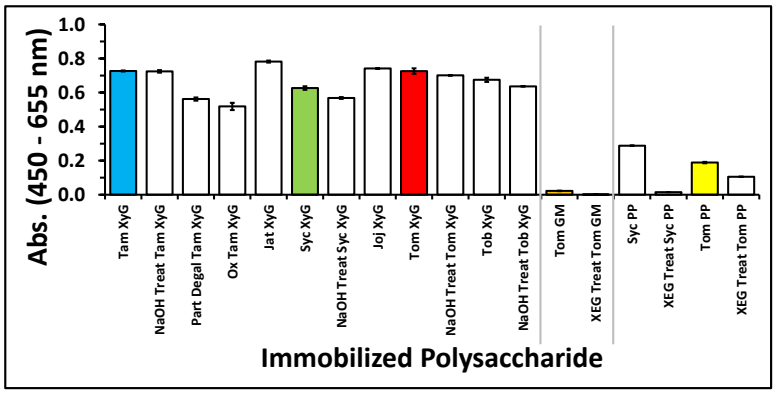

**Direct, single datapoint oligosaccharide ELISA**



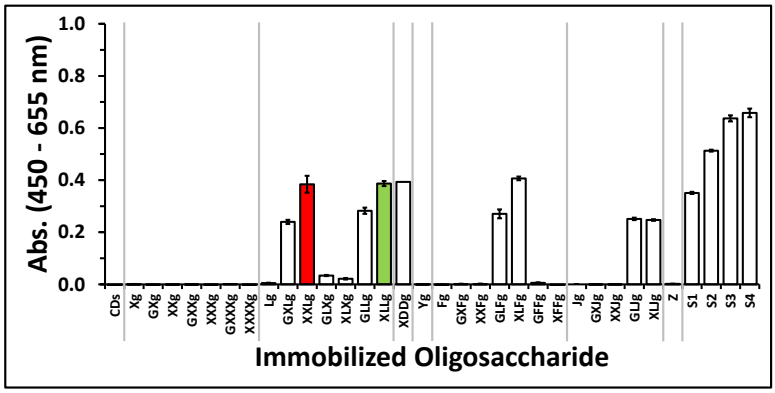
X

CCRC-M101

Direct, single datapoint polysaccharide ELISA



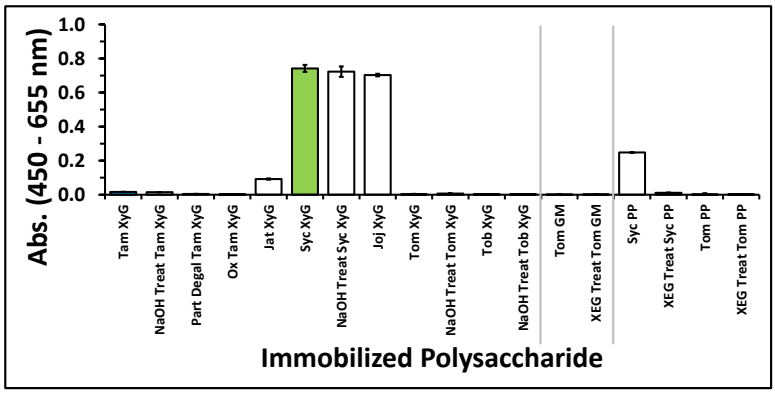

Direct, single datapoint oligosaccharide ELISA



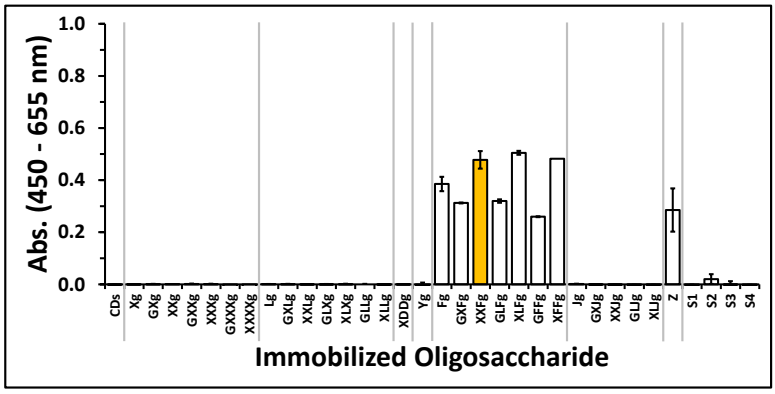
Y

CCRC-M102

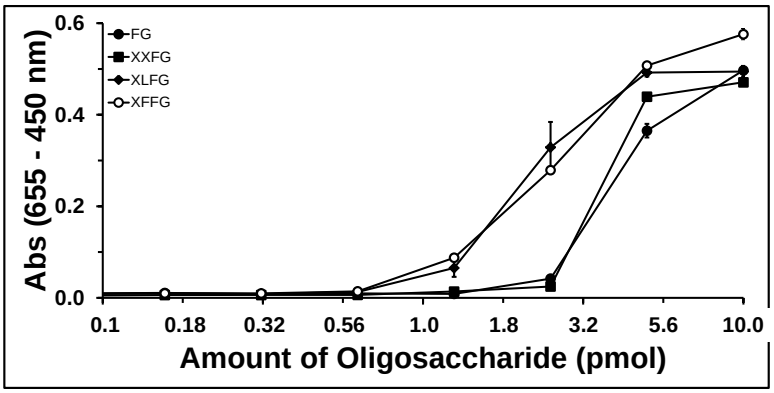
Direct, single datapoint polysaccharide ELISA




Direct, single datapoint oligosaccharide ELISA



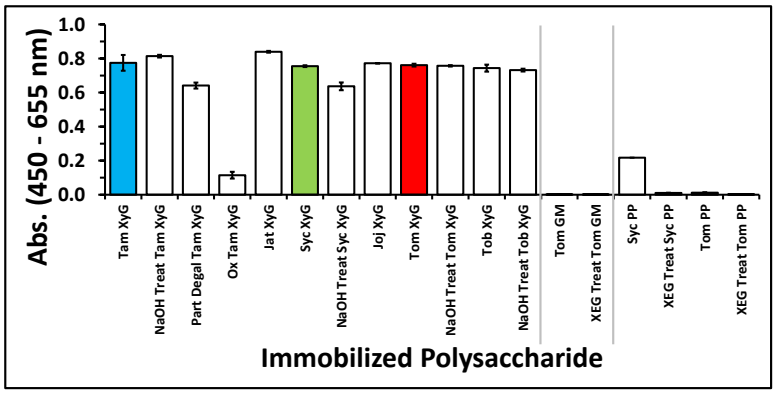

Direct, oligosaccharide dilution series ELISA



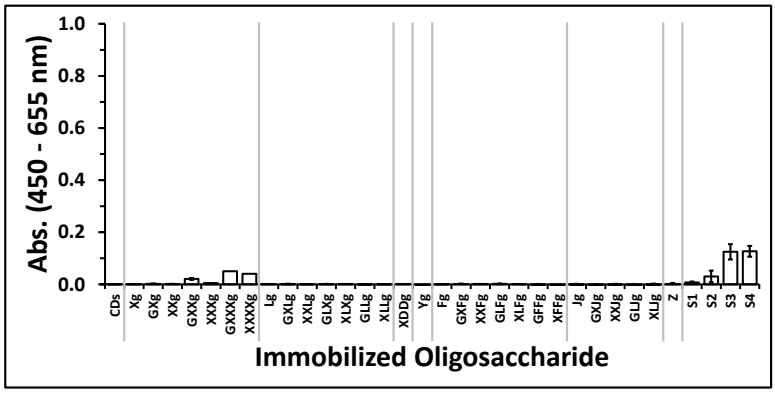
Z

CCRC-M103

Direct, single datapoint polysaccharide ELISA

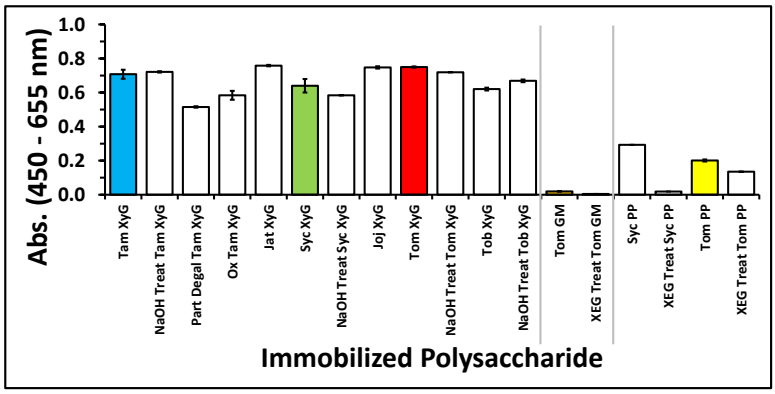



Direct, single datapoint oligosaccharide ELISA

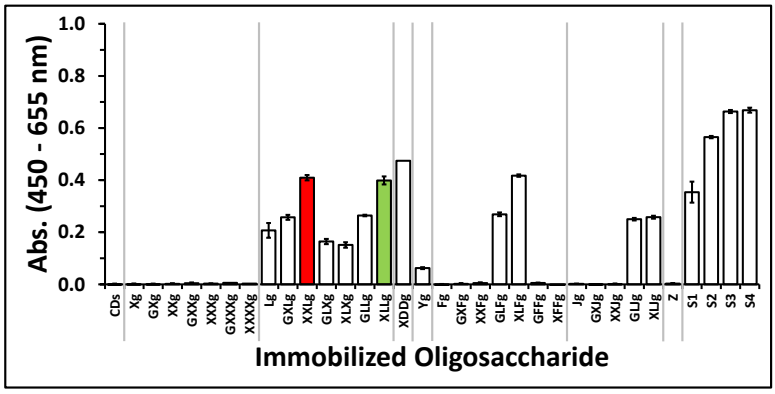


**AA** **CCRC-M104**

**Direct, single datapoint polysaccharide ELISA**

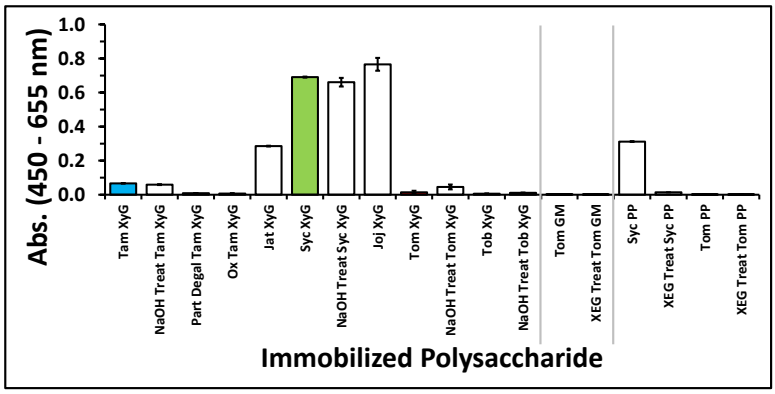



**Direct, single datapoint oligosaccharide ELISA**

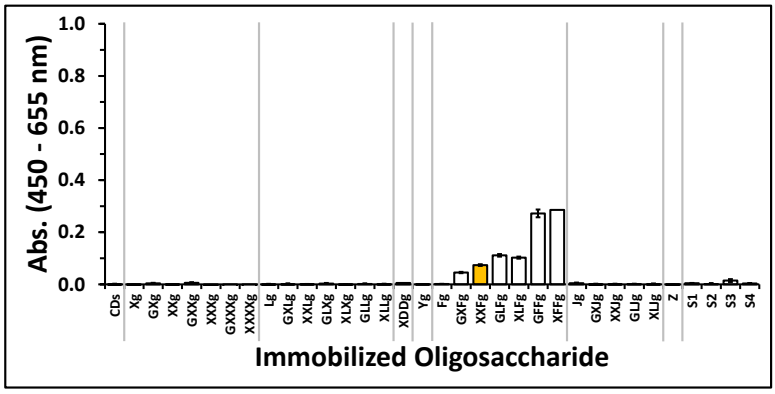


**AB** **CCRC-M106**

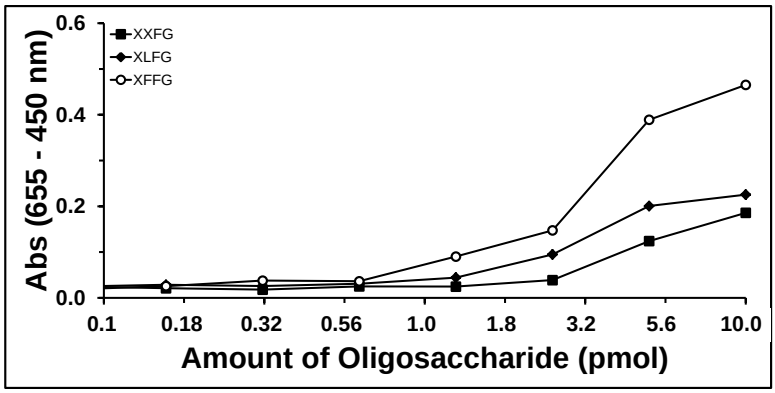
**Direct, single datapoint polysaccharide ELISA**




**Direct, single datapoint oligosaccharide ELISA**

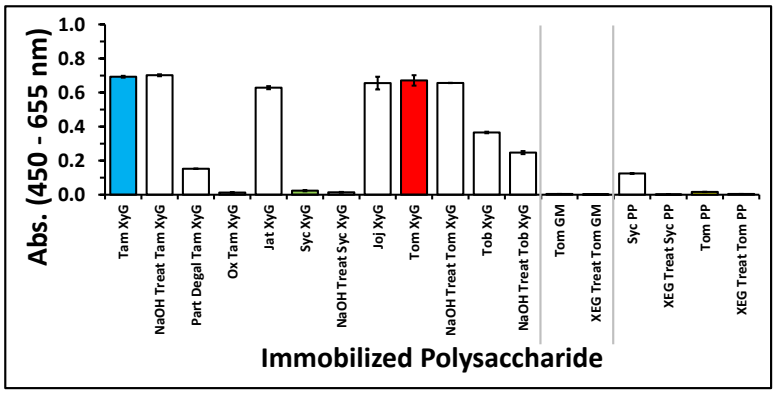



**Direct, oligosaccharide dilution series ELISA**

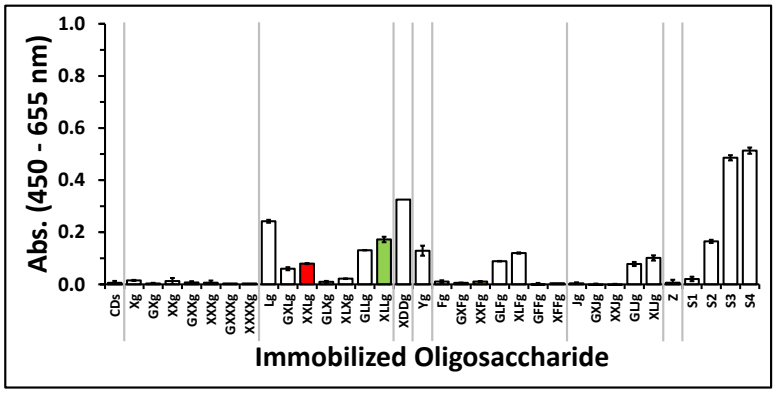


**AC** **CCRC-M108**

**Direct, single datapoint polysaccharide ELISA**



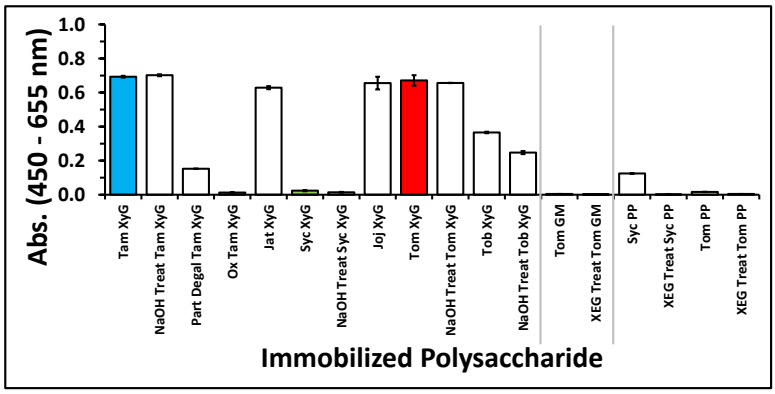

**Direct, single datapoint oligosaccharide ELISA**



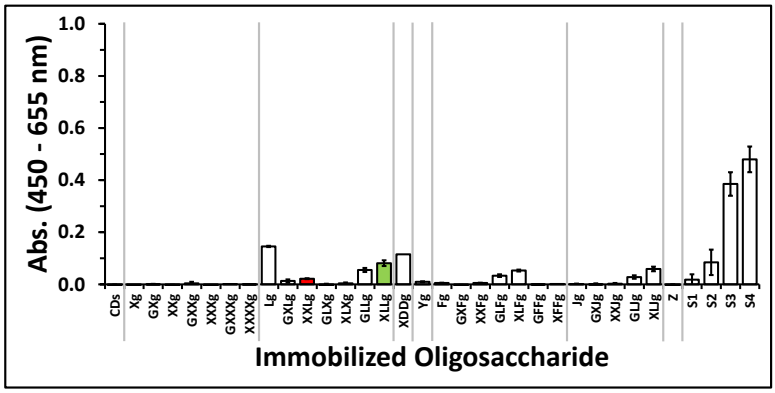
**AD**

**CCRC-M109**

**Direct, single datapoint polysaccharide ELISA**

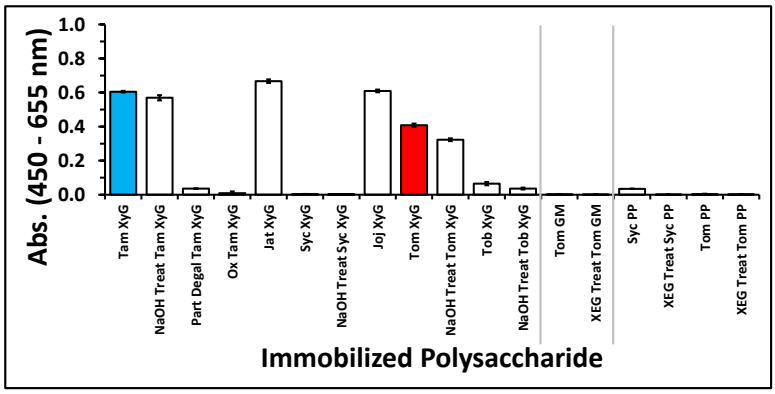



**Direct, single datapoint oligosaccharide ELISA**

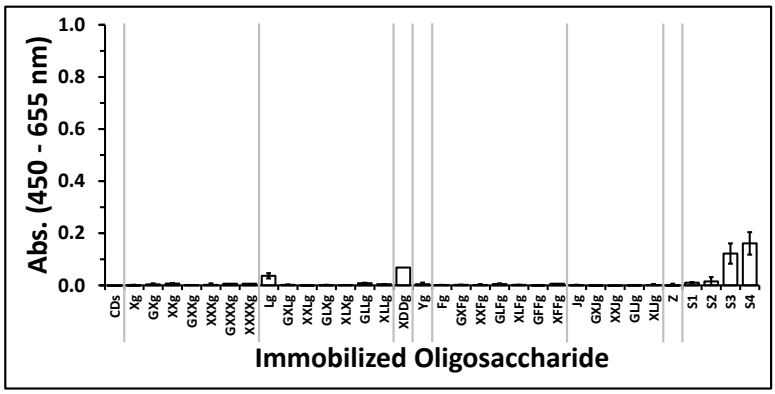


**AE** **CCRC-M111**

**Direct, single datapoint polysaccharide ELISA**

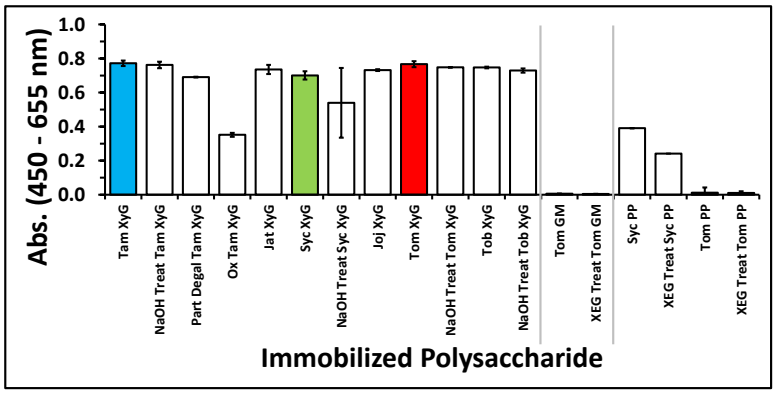



**Direct, single datapoint oligosaccharide ELISA**

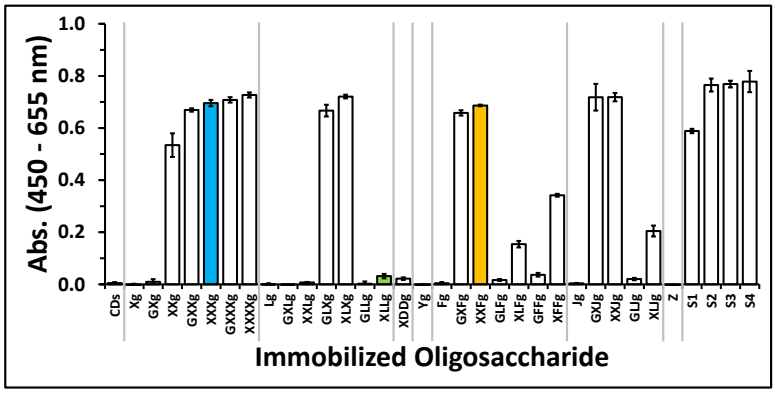


**AF** **LM15**

**Direct, single datapoint polysaccharide ELISA**




**Direct, single datapoint oligosaccharide ELISA**



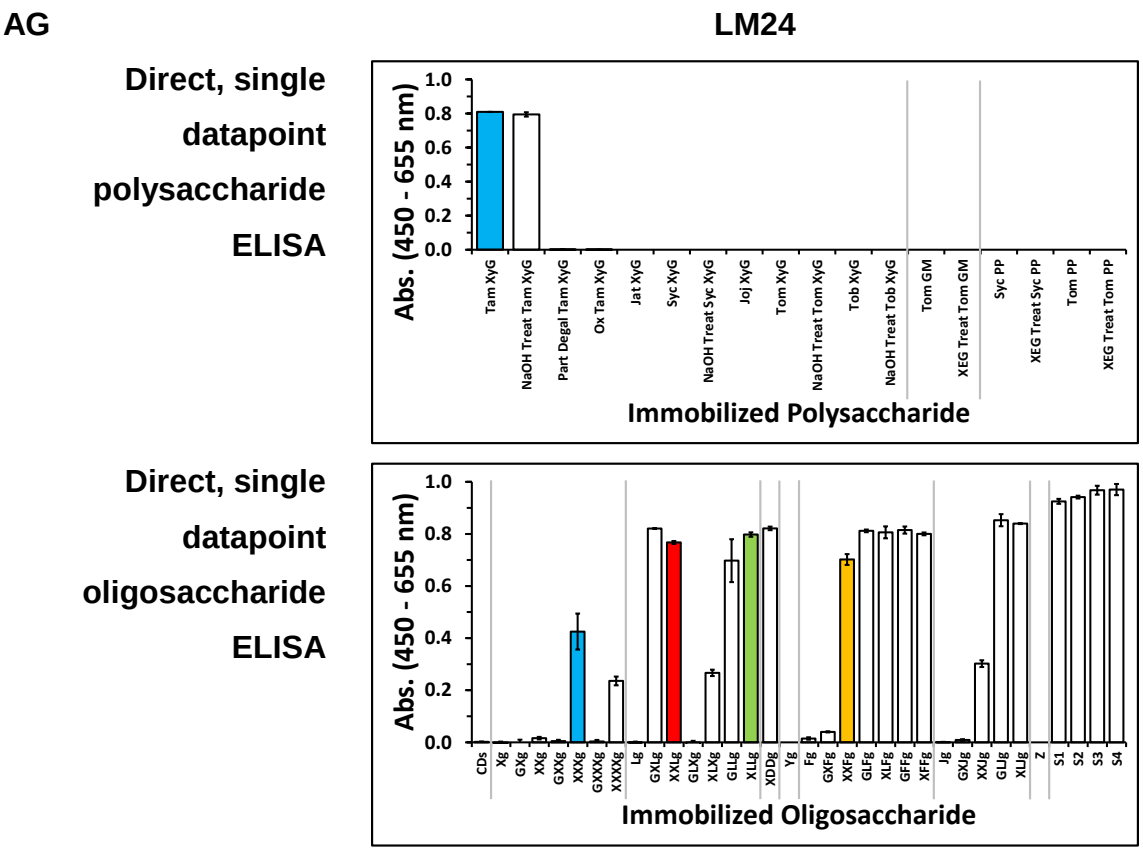



**Figure 3-S6.** ELISAs of the mAbs against immobilized poly- and oligosaccharides. Direct, single datapoint polysaccharide ELISAs were performed by incubating the hybridoma supernatant in a microplate well that was previously coated with the same polysaccharide that was used in the immunization. Competitive polysaccharide dilution series ELISAs were performed as the direct single datapoint polysaccharide ELISAs, except that the hybridoma supernatant was pre-incubated with a competitor polysaccharide prior to the incubation in the microplate well. Direct, single datapoint ELISAs against oligosaccharides were performed by incubating the hybridoma supernatant in a microplate well that was coated with oligosaccharide-APB conjugate. Direct,

oligosaccharide dilution series ELISAs were performed as the single datapoint ELISAs, except that a dilution series of the oligosaccharide-APB conjugate was immobilized on the microplate well. Note that not all experiments were carried out for all mAbs. In all cases, the reported values represent the average of three measurements and the error their standard deviation.

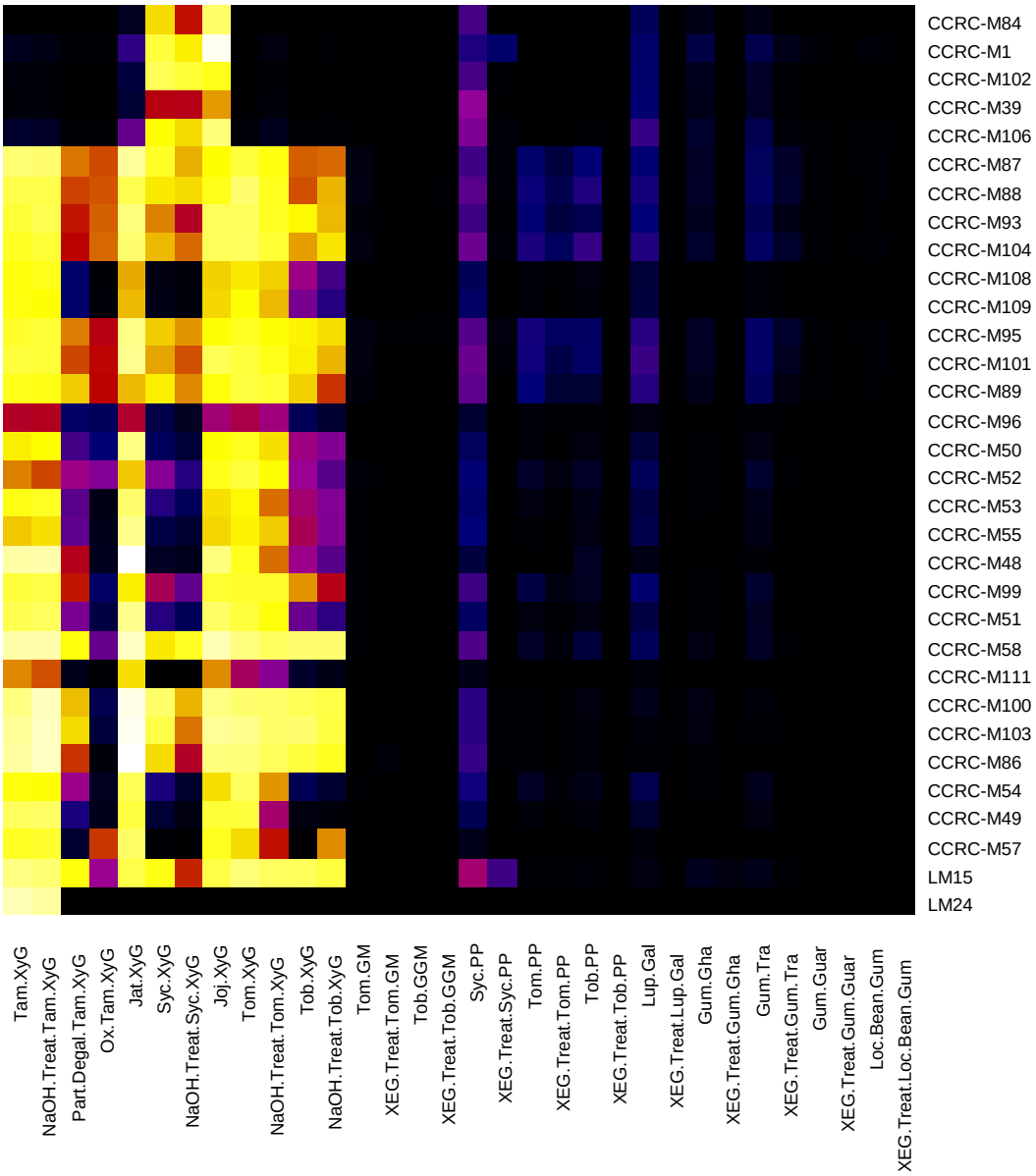



**Figure 3-S7.** Heatmap of the ELISA data of the mAb binding to the immobilized polysaccharides. Each row in the heatmap corresponds to the ELISA binding pattern of a single mAb against collection of purified and immobilized oligosaccharides. Each column corresponds to a binding pattern of a single immobilized oligosaccharide against collection of mAbs. The color of an

intersection of an mAb and an oligosaccharide represents the average of three ELISA absorbance readings. For key, see Supplemental Figure 8.

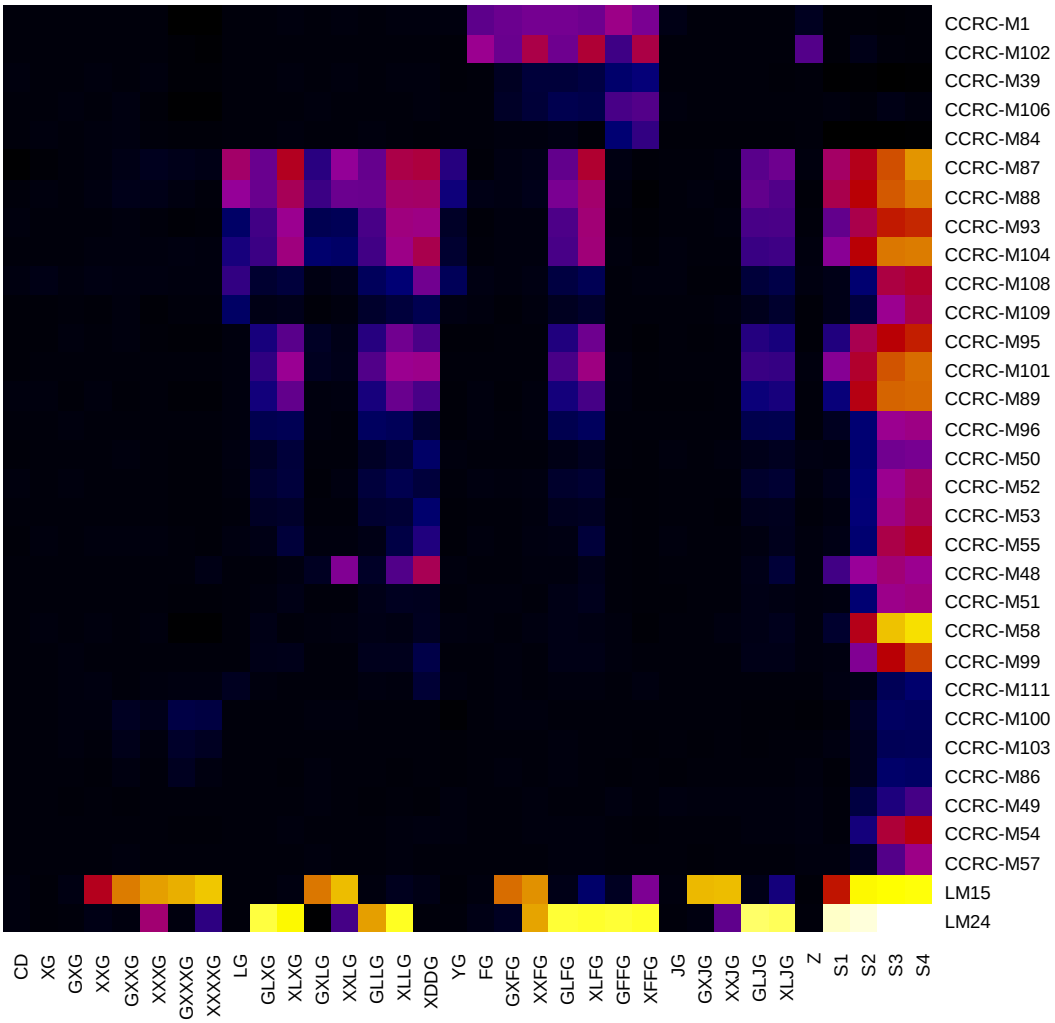



**Figure 3-S8.** Heatmap of the ELISA data of the mAb binding to the biotinylated and immobilized oligosaccharides. Each row in the heatmap corresponds to the ELISA binding pattern of a single mAb against collection of purified and immobilized oligosaccharides. Each column corresponds to a binding pattern of a single immobilized oligosaccharide against collection of mAbs. The color of an intersection of an mAb and an oligosaccharide represents the average of three ELISA absorbance readings.

### 3.8. Literature Cited

Albersheim P, Darvill AG, Roberts K, Sederoff R, Staehelin A (2010) *Plant Cell Walls*. Garland Science, New York.

Avci U, Pattathil S, Hahn M (2012) Immunological approaches to plant cell wall and biomass characterization: Immunolocalization of glycan epitopes. *In:* Himmel ME (*ed*) *Biomass Conversion* **908**, pp. 73-82. Humana Press.

Baba KI (2006) Models of plant cell walls. *In:* Hayashi T (*ed*) *The Science and Lore of the Plant Cell Wall: Biosynthesis, Structure and Function*. BrownWalker Press, Boca Raton.

Balestrini R, Hahn MG, Faccio A, Mendgen K, Bonfante P (1996) Differential localization of carbohydrate epitopes in plant cell walls in the presence and absence of arbuscular Mycorrhizal fungi. *Plant Physiology* **111**: 203-213.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**: 248-254.

Brennan M, Harris PJ (2011) Distribution of fucosylated xyloglucans among the walls of different cell types in monocotyledons determined by immunofluorescence microscopy. *Molecular Plant* **4**: 144-156.

Buckeridge MS, Crombie HJ, Mendes CJ, Reid JS, Gidley MJ, Vieira CC (1997) A new family of oligosaccharides from the xyloglucan of *Hymenaea courbaril* L. (Leguminosae) cotyledons. *Carbohydrate Research* **303**: 233-237.

Cavalier DM, Lerouxel O, Neumetzler L, Yamauchi K, Reinecke A, Freshour G, Zabotina OA, Hahn MG, Burgert I, Pauly M, Raikhel NV, Keegstra K (2008) Disrupting two *Arabidopsis thaliana* xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *The Plant Cell* **20**: 1519-1537.

DeMartini JD, Pattathil S, Miller JS, Li H, Hahn MG, Wyman CE (2013) Investigating plant cell wall components that affect biomass recalcitrance in poplar and switchgrass. *Energy & Environmental Science* **6**: 898-909.

DuBois M, Gilles KA, Hamilton JK, Rebers PA, Smith F (1956) Colorimetric method for determination of sugars and related substances. *Analytical Chemistry* **28**: 350-356.

Eda S, Akiyama Y, Katō K, Ishizu A, Nakano J (1985) A galactoglucomannan from cell walls of suspension-cultured tobacco (*Nicotiana tabacum*) cells. *Carbohydrate Research* **137**: 173-181.

Fischer M, Wegryzn TF, Hallett IC, Redgwell RJ (1996) Chemical and structural features of kiwifruit cell walls: Comparison of fruit and suspension-cultured cells. *Carbohydrate Research* **295**: 195-208.

Freund J (1956) The mode of action of immunologic adjuvants. *Advances in Tuberculosis Research* **7**: 130-148.

Fry SC, Aldington S, Hetherington PR, Aitken J (1993) Oligosaccharides as signals and substrates in the plant cell wall. *Plant Physiology* **103**: 1-5.

Fry SC, McDougall GJ, Lorences EP, Biggs KJ, Smith RC (1990) Oligosaccharins from xyloglucan and cellulose: Modulators of the action of auxin and $H^+$ on plant growth. *Symposia of the Society for Experimental Biology* **44**: 285-298.

Gibson LJ (2012) The hierarchical structure and mechanics of plant materials. *Journal of the Royal Society Interface* **9**: 2749-2766.

Hantus S, Pauly M, Darvill AG, Albersheim P, York WS (1997) Structural characterization of novel L-galactose-containing oligosaccharide subunits of jojoba seed xyloglucans. *Carbohydrate Research* **304**: 11-20.

Harlow E, Lane D (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Hayashi T, Kaida R (2011) Functions of xyloglucan in plant cells. *Molecular Plant* **4**: 17-24.

Hilz H, de Jong LE, Kabel MA, Verhoef R, Schols HA, Voragen AG (2007) Bilberry xyloglucan--novel building blocks containing beta-xylose within a complex structure. *Carbohydrate Research* **342**: 170-181.

Hisamatsu M, York WS, Darvill AG, Albersheim P (1992) Characterization of seven xyloglucan oligosaccharides containing from seventeen to twenty glycosyl residues. *Carbohydrate Research* **227**: 45-71.

Hoffman M, Jia Z, Peña MJ, Cash M, Harper A, Blackburn AR 2nd, Darvill A, York WS (2005) Structural analysis of xyloglucans in the primary cell walls of plants in the subclass Asteridae. *Carbohydrate Research* **340**: 1826-1840.

Hosoi E (2008) Biological and clinical aspects of ABO blood group system. *The Journal of Medical Investigation* **55**: 174-182.

Hsieh YS, Harris PJ (2009) Xyloglucans of monocotyledons have diverse structures. *Molecular Plant* **2**: 943-965.

Keegstra K (2010) Plant Cell Walls. *Plant Physiology* **154**: 483-486.

Knox JP (2008) Revealing the structural and functional diversity of plant cell walls. *Current Opinion in Plant Biology* **11**: 308-313.

Kulkarni AR, Peña MJ, Avci U, Mazumder K, Urbanowicz BR, Pattathil S, Yin Y, O'Neill MA, Roberts AW, Hahn MG, Xu Y, Darvill AG, York WS (2012) The ability of land plants to synthesize glucuronoxylans predates the evolution of tracheophytes. *Glycobiology* **22**: 439-451.

Lau JM, McNeil M, Darvill AG, Albersheim P (1987) Treatment of rhamnogalacturonan I with lithium in ethylenediamine. *Carbohydrate Research* **168**: 245-274.

Lee KJD, Marcus SE, Knox JP (2011) Cell wall biology: Perspectives from cell wall imaging. *Molecular Plant* **4**: 212-219.

Lees A, Nelson BL, Mond JJ (1996) Activation of soluble polysaccharides with 1-cyano-4-dimethylaminopyridinium tetrafluoroborate for use in protein—polysaccharide conjugate vaccines and immunological reagents. *Vaccine* **14**: 190-198.

Lerouxel O, Choo TS, Séveno M, Usadel B, Faye Lc, Lerouge P, Pauly M (2002) Rapid structural phenotyping of plant cell wall mutants by enzymatic oligosaccharide fingerprinting. *Plant Physiology* **130**: 1754-1763.

Lima NN, Rechia CGV, Ganter JLMS, Reicher F, Sierakowski MR (1995) Oligosaccharides derived from the xyloglucan isolated from the seeds of *Hymenaea courbaril* var. stilbocarpa. *International Journal of Biological Macromolecules* **17**: 413-415.

Linsmaier EM, Skoog F (1965) Organic growth factor requirements of tobacco tissue cultures. *Physiologia Plantarum* **18**: 100-127.

Madson M, Dunand C, Li X, Verma R, Vanzin GF, Caplan J, Shoue DA, Carpita NC, Reiter WD (2003) The MUR3 gene of Arabidopsis encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. *The Plant Cell* **15**: 1662-1670.

Marcus S, Verhertbruggen Y, Herve C, Ordaz-Ortiz J, Farkas V, Pedersen H, Willats W, Knox JP (2008) Pectic homogalacturonan masks abundant sets of xyloglucan epitopes in plant cell walls. *BMC Plant Biology* **8**: 60.

Nishikubo N, Takahashi J, Roos AA, Derba-Maceluch M, Piens K, Brumer H, Teeri TT, Stalbrand H, Mellerowicz EJ (2011) Xyloglucan endo-transglycosylase-mediated xyloglucan rearrangements in developing wood of hybrid aspen. *Plant Physiology* **155**: 399-413.

O'Neill MA, Warrenfeltz D, Kates K, Pellerin P, Doco T, Darvill AG, Albersheim P (1996) Rhamnogalacturonan-II, a pectic polysaccharide in the walls of growing plant cell, forms a dimer that is covalently cross-linked by a borate ester. *In vitro* conditions for the formation and hydrolysis of the dimer. *The Journal of Biological Chemistry* **271**: 22923-22930.

Odonmažig P, Ebringerová A, Machová E, Alföldi J (1994) Structural and molecular properties of the arabinogalactan isolated from Mongolian larchwood (*Larix dahurica* L.). *Carbohydrate Research* **252**: 317-324.

Pabst M, Fischl RM, Brecker L, Morelle W, Fauland A, Kofeler H, Altmann F, Leonard R (2013) Rhamnogalacturonan II structure shows variation in the side chains monosaccharide composition and methylation status within and across different plant species. *The Plant Journal* **76**: 61-72.

Pattathil S, Avci U, Baldwin D, Swennes AG, McGill JA, Popper Z, Bootten T, Albert A, Davis RH, Chennareddy C, Dong R, O'Shea B, Rossi R, Leoff C, Freshour G, Narra R, O'Neil M, York WS, Hahn MG (2010) A comprehensive toolkit of plant cell wall glycan-directed monoclonal antibodies. *Plant Physiology* **153**: 514-525.

Pedersen HL, Fangel JU, McCleary B, Ruzanski C, Rydahl MG, Ralet MC, Farkas V, von Schantz L, Marcus SE, Andersen MC, Field R, Ohlin M, Knox JP, Clausen MH, Willats WG (2012) Versatile high resolution oligosaccharide microarrays for plant glycobiology and cell wall research. *The Journal of Biological Chemistry* **287**: 39429-39438.

Peña MJ, Darvill AG, Eberhard S, York WS, O'Neill MA (2008) Moss and liverwort xyloglucans contain galacturonic acid and are structurally distinct from the xyloglucans synthesized by hornworts and vascular plants. *Glycobiology* **18**: 891-904.

Peña MJ, Kong Y, York WS, O'Neill MA (2012) A galacturonic acid-containing xyloglucan is involved in Arabidopsis root hair tip growth. *The Plant Cell* **24**: 4511-4524.

Popper ZA, Michel G, Herve C, Domozych DS, Willats WG, Tuohy MG, Kloareg B, Stengel DB (2011) Evolution and diversity of plant cell walls: From algae to flowering plants. *Annual Review of Plant Biology* **62**: 567-590.

Pratt LH (1984) Phytochrome immunochemistry. *In:* Smith H, Holmes MG (*eds*) *Techniques in Photomorphogenesis*, pp. 201-206. Academic Press. London.

Puhlmann J, Bucheli E, Swain MJ, Dunning N, Albersheim P, Darvill AG, Hahn MG (1994) Generation of monoclonal antibodies against plant cell-wall polysaccharides. I. Characterization of a monoclonal antibody to a terminal alpha-(1-->2)-linked fucosyl-containing epitope. *Plant Physiology* **104**: 699-710.

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rose JK, Braam J, Fry SC, Nishitani K (2002) The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: Current perspectives and a new unifying nomenclature. *Plant & Cell Physiology* **43**: 1421-1435.

Rose JKC, Bennett AB (1999) Cooperative disassembly of the cellulose–xyloglucan network of plant cell walls: Parallels between cell expansion and fruit ripening. *Trends in Plant Science* **4**: 176-183.

Schultink A, Cheng K, Park YB, Cosgrove DJ, Pauly M (2013) The identification of two arabinosyltransferases from tomato reveals functional equivalency of xyloglucan side chain substituents. *Plant Physiology* **163**: 86-94.

Selvendran RR, O'Neill MA (2006) Isolation and analysis of cell walls from plant material. *In:* Glick D (*ed*) *Methods of Biochemical Analysis*, pp. 25-153. John Wiley & Sons. London.

Shulman M, Wilde CD, Kohler G (1978) A better cell line for making hybridomas secreting specific antibodies. *Nature* **276**: 269-270.

Sims IM, Craik DJ, Bacic A (1997) Structural characterisation of galactoglucomannan secreted by suspension-cultured cells of *Nicotiana plumbaginifolia*. *Carbohydrate Research* **303**: 79-92.

Talmadge KW, Keegstra K, Bauer WD, Albersheim P (1973) The structure of plant cell walls: I. The macromolecular components of the walls of suspension-cultured sycamore cells with a detailed analysis of the pectic polysaccharides. *Plant Physiology* **51**: 158-173.

Tan L, Eberhard S, Pattathil S, Warder C, Glushka J, Yuan C, Hao Z, Zhu X, Avci U, Miller JS, Baldwin D, Pham C, Orlando R, Darvill A, Hahn MG, Kieliszewski MJ, Mohnen D (2013) An Arabidopsis cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. *The Plant Cell* **25**: 270-287.

Thygesen MB, Munch H, Sauer J, Clo E, Jorgensen MR, Hindsgaul O, Jensen KJ (2010) Nucleophilic catalysis of carbohydrate oxime formation by anilines. *The Journal of Organic Chemistry* **75**: 1752-1755.

Tiné MAS, Silva CO, Lima DUd, Carpita NC, Buckeridge MS (2006) Fine structure of a mixed-oligomer storage xyloglucan from seeds of *Hymenaea courbaril*. *Carbohydrate Polymers* **66**: 444-454.

Torrey JG, Shigemura Y (1957) Growth and controlled morphogenesis in pea root callus tissue grown in liquid media. *American Journal of Botany* **44**: 334-344.

Vargas-Rechia C, Reicher F, Rita Sierakowski M, Heyraud A, Driguez H, Linart Y (1998) Xyloglucan octasaccharide XXLGol derived from the seeds of *Hymenaea courbaril* acts as a signaling molecule. *Plant Physiology* **116**: 1013-1021.

Vaughn KC, Talbot MJ, Offler CE, McCurdy DW (2007) Wall ingrowths in epidermal transfer cells of *Vicia faba* cotyledons are modified primary walls marked by localized accumulations of arabinogalactan Proteins. *Plant and Cell Physiology* **48**: 159-168.

Vincken JP, York WS, Beldman G, Voragen AG (1997) Two general branching patterns of xyloglucan, XXXG and XXGG. *Plant Physiology* **114**: 9-13.

York WS, Darvill AG, McNeil M, Stevenson TT, Albersheim P (1986) Isolation and characterization of plant cell walls and cell wall components. *Methods in Enzymology* **118**: 3-40.

York WS, Kumar Kolli VS, Orlando R, Albersheim P, Darvill AG (1996) The structures of arabinoxyloglucans produced by solanaceous plants. *Carbohydrate Research* **285**: 99-128.

York WS, Oates JE, van Halbeek H, Darvill AG, Albersheim P, Tiller PR, Dell A (1988) Location of the *O*-acetyl substituents on a nonasaccharide repeating unit of sycamore extracellular xyloglucan. *Carbohydrate Research* **173**: 113-132.

Zola H (1987) *Monoclonal Antibodies: A Manual of Techniques*. CRC Press, Boca Raton, Florida.

# CHAPTER 4

# CONCLUSIONS

The work described in this dissertation focused on the development of methods for structural and quantitative analysis of xyloglucans (XyGs) as well as on rigorously interrogating the binding epitopes of XyG-binding monoclonal antibodies (mAbs).

In Chapter 2, I describe the purification of XyG oligosaccharides from several plant sources to obtain a structurally diverse oligosaccharide library. A comprehensive strategy was designed for the generation of twenty-eight XyG oligosaccharides using complementary enzymatic hydrolyses and subsequent separations using size-exclusion chromatography (SEC). This workflow allows the purification of milligram amounts of these oligosaccharides. The oligosaccharide collection is useful for deciphering the enzymatic and binding activities of various XyG-active proteins. The structural identities and purities of the produced oligosaccharides were verified by three analytical approaches, namely matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), high-pH anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD), and nuclear magnetic resonance (NMR) spectroscopy. MALDI-TOF and HPAEC-PAD were further compared in

their abilities to quantitate XyG oligosaccharides in complex mixtures. Both methods can be used as quantitative tools in XyG analysis, but for unequivocal interpretation, they should be used in conjunction.

In Chapter 3, I describe the utilization of the XyG oligosaccharide collection in the enzyme-linked immunosorbent assay (ELISA)-based epitope characterization of thirty-two mAbs whose detailed binding specificities had not been determined before. My results reveal various distinct mAb binding specificities, which cover the most common structures found in XyGs such as terminal xylosyl, galactosyl and fucosyl residues. These efforts provide valuable information for researchers in order to efficiently select these mAbs to probe for XyG structures and abundances in plant cell walls.

Collectively, my work contributes to the arsenal of tools that are available for the structural and functional characterization of XyGs.

# APPENDIX A

# THE PROFINIA™ PROTEIN PURIFICATION SYSTEM SIMPLIFIES ANTIBODY PURIFICATION WITH PROTEIN A [1]

[1] Berardini, M., Tuomivaara, S.T. and M.G. Hahn. 2008. *Bio-Rad Laboratories, Inc. Tech note 5712*.

Reprinted here with permission of the publisher.

## A.1 Introduction

Immobilized Protein A from *Staphylococcus aureus* has been used for many years to purify antibodies from a variety of species (Hjelm et al, 1972). The high selectivity and stability of protein A has made it a popular choice for the purification of antibodies from a wide range of sample sources, including serum, ascitic fluid, and hybridoma cell culture supernatants. M3ammalian antibodies are categorized into five major classes: IgA, IgD, IgE, IgG, and IgM. IgG is the predominant class of antibody in serum and is generated in large amounts during the secondary immune response. The IgG class of antibody is further divided into subclasses that vary depending upon the species and the properties of the heavy chain component. There are four subclasses of IgG in humans ($IgG_1$, $IgG_2$, $IgG_3$, $IgG_4$) and in mice ($IgG_1$, $IgG_{2a}$, $IgG_{2b}$, $IgG_3$). The affinity of protein A for IgG varies considerably between species and IgG subtypes and has been extensively characterized (Duhamel et al, 1979; Schwartz, 1990). In humans, protein A binds with high affinity to $IgG_1$, $IgG_2$, and $IgG_4$, but poorly to $IgG_3$. Among the four IgG subtypes in mice, protein A has the weakest affinity for $IgG_1$. The binding of antibodies to protein A is mediated, at neutral or alkaline pH values, through hydrophobic interactions involving a highly conserved histidyl residue located in the protein A binding site of IgG. The elution of IgG from immobilized protein A is commonly achieved by lowering the pH using an acidic buffer. Protein A-purified antibodies are then typically neutralized with a base, dialyzed against a neutral buffer, or desalted using a gel-filtration column to avoid acid-mediated hydrolysis and denaturation. The use of protein A affinity chromatography offers a number

of advantages to alternative separation techniques based on ion exchange, immobilized metal affinity chromatography (IMAC), and hydrophobic interaction chromatography (HIC), due to its simplicity and high recoveries of the purified product. The Profinia protein purification system is an automated chromatography instrument designed for the purification of affinity-tagged proteins and antibodies. An intuitive touch screen interface allows users to access preprogrammed chromatography methods or to customize any method for a particular application. With the Profinia system, antibodies can be purified and immediately desalted, precluding the need for neutralization, dialysis, or additional chromatography steps. The system is specifically designed for ease of use and for routine purification of 1–100 mg of IgG, using either 1 ml or 5 ml prepacked protein A cartridges. In this study, the purification, yield, and purity of IgG recovered from human, rabbit, mouse, rat, and goat sera using protein A-based separations are described. In addition, the effect of varying the elution buffer composition on the elution of IgG from protein A cartridges was investigated. Lastly, the purification of a monoclonal antibody from a hybridoma cell culture supernatant using the Profinia system and its subsequent functional analysis by an ELISA is presented.

## A.2. Literature Cited

Duhamel RC, Schur PH, Brendel K, Meezan E (1979) pH gradient elution of human IgG1, $IgG_2$ and $IgG_4$ from protein A-sepharose. *Journal of Immunological Methods* **31**: 211-217.

Hjelm H, Hjelm K, Sjoquist J (1972) Protein A from *Staphylococcus aureus*. Its isolation by affinity chromatography and its use as an immunosorbent for isolation of immunoglobulins. *FEBS Letters* **28**: 73-76.

Schwarz L (1990) Immunomodulatory properties of Protein A. *In:* Boyle M (*ed*) *Bacterial Immunoglobulin–binding Proteins*, **Vol. 2**, pp. 309-318. Academic Press, San Diego.

# APPENDIX B

# METHODS FOR STRUCTURAL CHARACTERIZATION OF THE PRODUCTS OF CELLULOSE- AND XYLOGLUCAN-HYDROLYZING ENZYMES [1]

[1] Peña, M.J., Tuomivaara, S.T., Urbanowicz, B.R., O'Neill, M.A. and W.S York. 2012. *Methods in Enzymology*. 510:121-139.

Reprinted here with permission of the publisher.

## B.1. Abstract

Structural characterization of oligosaccharide products generated by enzymatic hydrolysis of plant cell wall polysaccharides provides valuable information about the enzyme's activity and substrate specificity. In this chapter, we describe some of the chemical, chromatographic, and spectroscopic methods that we routinely use to isolate and characterize oligosaccharides formed by enzymatic fragmentation of cellulose and xyloglucan. These include techniques to determine glycosyl residue and glycosyl linkage compositions by gas chromatography and mass spectrometry. We also illustrate the use of electrospray ionization with multistage mass spectrometry, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, and nuclear magnetic resonance spectroscopy to perform detailed structural analysis of these oligosaccharides.

# APPENDIX C

# STRUCTURAL ANALYSIS OF HEMICELLULOSES BY ENZYMATIC DIGESTION AND MALDI-TOF MS [1]

[1] Tuomivaara, S.T., Kulkarni, A. and W.S. York. 2013. *GlycoPOD http://jcggdb.jp/GlycoPOD*. Web. 18.11. 2013.

Reprinted here with permission of the publisher.

## C.1. Introduction

Hemicelluloses, such as xyloglucans (XG) and xylans, are abundant polysaccharide components of plant cell walls. Typically, hemicelluloses exhibit structural microheterogeneity within the polymer, owing to a variety and (quasi-) statistical distribution of different sugar residues, and sometimes non-sugar decorations, such as acetyl and methyl groups. Some of these structural features are easier to identify if analysis is performed using oligosaccharides rather than the parent polysaccharides. Mass spectrometry (MS) offers a unique combination of sensitivity and molecular mass discrimination for oligosaccharide analysis. As little as a few ng of native (non-permethylated) oligosaccharide can be detected. Stereochemical and isomeric aspects of structure, such as sugar identities as well as linkage anomericities and positions, are not available from full MS analyses. Nevertheless full MS analysis of either a pure preparation or a mixture of oligosaccharides can typically provide the compositions of all oligosaccharide in the sample in terms of sugar types (*e.g.* hexoses and pentoses) and non-sugar decorations. Rough estimation of the relative abundances of oligosaccharides with similar molecular masses in a complex mixture is also possible. Herein, we describe a method to analyze hemicellulose structure by using a combination of enzymatic hydrolysis and matrix-assisted laser desorption-ionization MS (MALDI-TOF MS). The polysaccharide is first hydrolyzed with a glycosyl hydrolase with known specificity to form oligosaccharides, preferably containing 4 to 30 sugar residues. These oligosaccharides with relative molecular masses ranging from 500 to 4000, are easier to ionize, and the ion mass-to-charge ratios (*m*/*z*) and

intensities can be measured more accurately than for polysaccharide ions by MALDI-TOF MS. Interpretation of the spectral features can reveal some structural aspects of the parent polysaccharide. In this protocol, a detailed analysis of xyloglucan is outlined. Examples of methylglururonoxylan (MGX) and arabinoxylan (AX) are provided for reference. The procedure can be adapted to other hemicellulosic and non-hemicellulosic polysaccharides regardless of their source and structures.