



HAL
open science

Viridiplantae and Algae

M A O'Neill, A G Darvill, M E Etzler, D Mohnen, S Perez, M Pauly, J
Mortimer

► **To cite this version:**

M A O'Neill, A G Darvill, M E Etzler, D Mohnen, S Perez, et al.. Viridiplantae and Algae. Essentials of Glycobiology, 4th edition, 2022. hal-04376359

HAL Id: hal-04376359

<https://hal.science/hal-04376359>

Submitted on 6 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Viridiplantae and Algae

M.A. O'Neill, A. G. Darvill, M.E. Etzler, D. Mohnen, S. Perez, M. Pauly & J. Mortimer

INTRODUCTION

Viridiplantae (green plants) are defined as a clade of photosynthetic organisms that contain chlorophylls a and b, produce and store their photosynthetic products inside a double-membrane-bounded chloroplast, and have cell walls that typically contain cellulose. As photoautotrophic organisms, green plants are capable of converting carbon dioxide into carbohydrates. Thus, in these organisms, carbohydrates are not limiting, and their functional utilization throughout their life cycle has expanded enormously both in functionality and structural diversity.

The Viridiplantae are comprised of two clades—the Chlorophyta and the Streptophyta. The Chlorophyta contain most of the organisms typically referred to as "green algae." The term "algae" is also used for several other groups of photosynthetic eukaryotes, including diatoms and the red, brown, golden, and yellow-green algae. The Streptophyta comprise several other lineages that are also referred to as "green algae" and the land plants. Land plants include liverworts, mosses, hornworts, lycopods, ferns, gymnosperms, and flowering plants. In this chapter, we provide an overview of the current knowledge of green plant glycan structures with an emphasis on the features that are unique to land plants.

PLANT GLYCAN DIVERSITY

Green plants synthesize diverse glycans that vary in their structural complexity and molecular size. Soluble small molecular weight compounds encompass the transport disaccharide sucrose or other glycoconjugates that contain either aromatic (e.g., phenolic glycosides) or aliphatic (glycolipids) aglycones.

Many of these compounds function in plant protection or defence e.g. to ward off herbivores. Plant polysaccharides are linear or branched polymers composed of the same or different monosaccharides. Examples of homopolymers made entirely out of glucose include the storage polymer starch and the structural polymer cellulose (Figure 24.1).

An example of a structurally vastly complex polysaccharide is the plant cell wall pectic polysaccharide, referred to as rhamnogalacturonan II (RG-II), containing 12 different monosaccharides linked together by up to 21 different glycosidic linkages (Figure 24.2). Plant proteoglycans are structurally diverse glycans in which carbohydrate generally O-linked to the protein via hydroxy amino acids accounts for up to 90% of the molecule (Figure 24.3). Plant glycoproteins typically contain 15% or less carbohydrate in the form of N-linked oligomannose, complex, hybrid,

and paucimannose oligosaccharides (see Figure 24.4). Land plants also form O-GlcNAc-modified nuclear and cytosolic proteins.

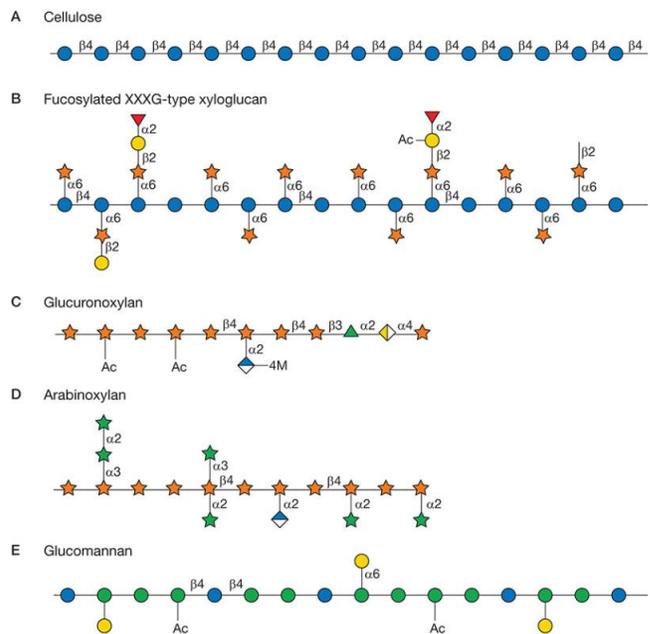


Figure 24.1 Glycosyl sequences of cellulose, selected hemicelluloses, mixed-linkage glucan.

NUCLEOTIDE SUGARS – THE BUILDING BLOCKS

Nucleotide sugars are the donors used for the synthesis of glycoconjugates and glycosylated secondary metabolites (Chapter 5). In plants, most of these activated monosaccharides exist as their nucleotide-diphosphates (e.g., UDP-Glc or GDP-Man), although at least one monosaccharide, Kdo, exists as its cytidine monophosphate derivative (CMP-Kdo). Nucleotide sugars are formed from the carbohydrate generated by photosynthesis, from the monosaccharides released by hydrolysis of sucrose and storage carbohydrates, and by recycling monosaccharides from glycans and the cell wall. Nucleotide sugars are also formed by interconverting preexisting activated monosaccharides. To date, 30 different nucleotide sugars and at least 100 genes encoding proteins involved in their formation and interconversion have been identified in plants.

PLANT GLYCOSYLTRANSFERASES AND GLYCAN MODIFYING ENZYMES

As carbohydrate-rich organisms occur, both in abundance and structural diversity, plant genomes contain numerous genes encoding proteins involved in the synthesis and modification of glycans. These proteins are spread across many enzyme classes in the Carbohydrate-Active Enzymes (CAZy) database (Table 24.1). Many of these proteins may be involved in the formation and modification of the polysaccharide-rich cell wall. Indeed, the unicellular alga *Ostreococcus tauri*, one of the few plants that does not form a cell wall, has a much smaller number of genes predicted to be involved in glycan metabolism.

Organism	GTs	Hydro-lases	Lyases	Ester-ases	CBMs	AA
<i>Arabidopsis</i>	564 (42)	426 (37)	34 (2)	85 (4)	124 (10)	49 (3)
Rice	568 (43)	429 (34)	16 (2)	53 (4)	122 (10)	(4)
Human	243 (46)	94 (28)	ni	1 (1)	39 (6)	
<i>Ostreococcus</i>	67 (25)	29 (12)	ni	2 (1)	14 (5)	

Table 1. Estimated number of genes encoding proteins involved in the synthesis and modification of glycans in plants and humans

PLANT MUTANTS PROVIDE CLUES TO GLYCAN FUNCTION

The availability of plant lines carrying mutations in specific genes has yielded considerable insight into glycan biosynthesis and function. Initially, *Arabidopsis* has been widely used as a model dicot as it is easy to grow, has a short life cycle, and its relatively small genome has been sequenced and extensively annotated, leading to the generation of chemically induced or transfer DNA (T-DNA) insertion mutant collections. These mutants have been vital in studying both the function and substrate specificity of, e.g. GTs, and for demonstrating the function or redundancy of glycans and/or glycan substituents in the life-cycle of a plant. With the advent of economically viable whole-genome sequencing, the sequence of nearly 100 plant species has been assembled (PlantGDB (<http://www.plantgdb.org/>) including rice, maize, barley, poplar, potato, etc. With this sequence and annotation information, plant mutants can be generated to test the functionality of genes related to glycans uniquely present in those plant species.

A forward genetic approach, where *Arabidopsis* seeds are randomly chemically mutagenized, and the resulting plantlets screened for structural changes in their glycans, has led to the discovery of multiple genes involved in nucleotide sugar interconversion pathway as well as plant cell wall-related GTs. Reverse genetic approaches, where based on a known gene sequence, plants carrying the corresponding loss-of-function mutations have been characterized, has led to the discovery of GTs involved in heteroxylan synthesis of primary and secondary walls, pectin synthesis, N- and O-proteoglycan synthesis, and glycolipid

synthesis. In addition, glycan modifying enzymes such as O-acetyl- and methyltransferases have been identified and characterized.

During photosynthesis, plants fix the carbon from carbon dioxide into sugars which are incorporated into complex polysaccharides. This process makes plant cell walls a terminal carbon sink and constitutes the dominant carbon sequestration system on Earth. All plant cell walls share in common the characteristic of having a complex composite structure organized on a cellulose framework. They are essentially made of polysaccharides, more than 90% in the primary walls and between 65 and 70% in the secondary walls. Their composition is based on three different categories of polysaccharides, cellulose, and the so-called matrix polysaccharides: pectic polysaccharides and hemicelluloses. Such a diversity of complex carbohydrate polymers is due to the great diversity of monosaccharide constituents allied to the diversity of inter-glycosidic linkages. This results in a great heterogeneity in the structure and function of the plant cell walls (Burton *et al.*, 2010).

More than fifteen different monosaccharides of the D and L series engaged in various modes of glycosidic linkages with α and β configurations are found in the cell wall polymers. It is interesting that most of these monosaccharides originate from the activated form of glucose, UDP- α -D-glucose, via common nucleotide sugar interconversion pathways (Feingold & Avigad, 1980; Reiter & Vanzin, 2001) catalyzed by a series of oxidation, reduction, epimerization, and/or decarboxylation reactions leading to activated sugars that can be used directly by glycosyltransferases. As many as 30 different nucleotide sugars have been identified in plants. In some cases, the free form of certain sugars can be incorporated into polysaccharides. (Bar-Bedel & O'Neill 2011)

PLANT METABOLIC CARBOHYDRATES

The initial product of photosynthesis, a triosephosphate, is utilized by plants in multiple ways. Triosephosphate can be converted to sucrose, the most abundant soluble carbohydrate in green plants. Sucrose, a disaccharide (α -D-Glcp-1-2- β -D-Fruf), does not contain a reducing end and is thus relatively chemically inert. Sucrose represents the dominant transport carbohydrate for the distribution of carbohydrates throughout the organism, in particular to non-photosynthetic organs such as roots. Other soluble carbohydrates that are almost ubiquitous in the plant kingdom include raffinose oligosaccharides (raffinose, stachyose, and verbascose). These oligosaccharides are derivatives of sucrose that contain one or more α -D-Galp residues.

The most abundant branched polysaccharide is starch, the dominant carbohydrate energy storage form in green plants. Starch represents most of the calories that humankind utilizes directly as food or indirectly as animal feed via the consumption of grains, tubers, and fruit. Starch consists entirely of glucose, consisting of α -1,4-linked D-Glcp amylose polymers, and such polymers branched via 1-6 linkages to form amylopectin. Starch polymers are arranged in insoluble, semicrystalline granules in chloroplasts or specialized plastids.

Due to a high abundance of soluble sucrose, which contains the ketose fructose, plants also generate fructose-containing polysaccharides termed fructans. The simplest fructan is linear inulin consisting of 1-2- β -D-Fruf. Other inulins contain two fructose chains on a sucrose core molecule. Levan type fructans in grasses contain linear 2-6- β -D-Fruf polymers. Fructans are often utilized by green plants as an alternative or additional storage polymers but are stored in the vacuole of plant cells. It is thought that fructans are also involved in plant protection, in particular abiotic stress such as drought, salt, or cold stress.

PLANT CELL WALLS

A substantial portion of the carbohydrate formed by photosynthesis is used to produce the polysaccharide-rich walls that surround plant cells. Primary and secondary cell walls are distinguished by their composition, architecture, and functions. A primary wall surrounds growing and dividing plant cells and nongrowing cells in the soft tissues of fruits and leaves. These walls are capable of controlled extension to allow the cell to grow and expand but are sufficiently strong to resist the cells internal turgor pressure. A much thicker and stronger secondary wall is often formed once a cell has ceased to grow. This wall is deposited between the plasma membrane and the primary wall and is composed of layers that differ in the orientation of their cellulose microfibrils and incorporation of the hydrophobic, non-carbohydrate, polyphenolic polymer lignin. For example, the secondary walls of vascular tissues involved in the movement of water and nutrients through the plant are further strengthened by the incorporation of lignin. The ability to form conducting tissues with lignified and rigid secondary walls was an indispensable event in the evolution of vascular land plants, as it facilitated the transport of water and nutrients and allowed extensive upright growth and thus a competitive advantage for the exposure and capture of sunlight. Secondary cell walls account for most of the carbohydrate in plant biomass, such as straw for animal feed and wood for the production of paper and lumber for construction. Due to its abundance, it is also considered as renewable, carbon-neutral feedstock for the production of biofuels, biomaterials and other value-added commodity chemicals (Chapter 59).

PRIMARY CELL WALL GLYCANS

Primary cell walls are composites that resemble fibre-reinforced porous, aqueous gels. The complex structures and functions of these walls result from the assembly and interactions of a limited number of structurally defined polysaccharides and proteoglycans. It is important to point out that wall structure and organization differs between plant species and tissue types within a plant, thus exhibiting a large diversity of structural features within the various glycans. Also, during cell division and differentiation, and in response to biotic and abiotic challenges, the cell responds by the differential synthesis and modification of the non-cellulosic components or by the addition of new components.

Primary walls of land plants contain cellulose, hemicellulose, and pectin in different proportions. They also contain structural

proteins/proteoglycans, enzymes, low molecular weight phenolics, and minerals. Pectin and hemicellulose are present in approximately equal amounts in the so-called type I primary walls of gymnosperms, dicots, and nongraminaceous monocots, whereas hemicellulose is far more abundant than pectin in the type II walls of the grasses. Recent more cell-type-specific structural characterization of plant cell wall glycans indicates that the breadth of wall structural diversity is enormous, and thus, one should think of the walls as a structural continuum rather than specific types.

Cellulose. Cellulose, the most abundant biopolymer in nature, is a linear polymer composed of 1-4-linked β -D-Glcp residues (Figure 24.1A). Several of these chains are hydrogen-bonded to one another to form para-crystalline microfibrils. Each microfibril is predicted to contain between 18 and 24 glucan chains. The glucan chain is synthesized by a cellulose synthase complex at the cell's plasma membrane. Three cellulose synthases, encoded by three different genes, are believed to interact to form a trimeric complex, which in turn assembles into a hexameric rosette at the plasma membrane. The catalytic site of each cellulose synthase is located in the cytosol and transfers glucose from UDP-Glc onto the elongating glucan chain. The mechanisms involved in the formation of a microfibril from individual glucan chains are not well understood, although it may involve a self-assembly process that is facilitated by numerous specific proteins. The newly formed microfibrils are deposited in the wall of a growing cell with an orientation that is transverse to the axis of elongation. This orientation may be guided in part by protein-mediated interactions between cellulose synthase proteins and cortical microtubules.

Many properties of native cellulose depend on interactions that occur at the surface of the microfibrils. The surface chains are accessible and reactive, whereas the hydroxyl groups of the internal chains in the crystal participate in extensive intra- and intermolecular hydrogen bonding. Cellulose is insoluble in water and somewhat resistant to hydrolysis by endo- and exoglucanases because of this highly packed arrangement of the glucan chains.

A number of organisms such as fungi are specialized in degrading plant biomass, including cellulose. For cellulose depolymerization, several types of enzymes, including endoglucanases, cellobiohydrolases, and β -glucosidases, are required. Many of these enzymes consist of a catalytic domain connected to a cellulose-binding module. This module facilitates the binding of the enzyme to the insoluble substrate. Some microorganisms also produce copper-dependent oxidases that render crystalline cellulose more susceptible to hydrolysis. Cellulases and other enzymes involved in cellulose hydrolysis often exist as macromolecular complexes referred to as cellulosomes. Improving the effectiveness of cellulosomes is an area of active research to increase the conversion of plant biomass to fermentable sugar.

Hemicelluloses. Hemicelluloses are branched polysaccharides with a backbone composed of 1-4-linked β -D-pyranosyl residues with an equatorial O-4 (Glc_p, Man_p, and Xyl_p). Xyloglucan, glucurono/arabinoxylan, and heteroglucomannan (Figure 24.1B-E) are included under this chemical definition of

hemicelluloses. Hemicelluloses and cellulose have structural and conformational similarities that allow them to form strong, non-covalent associations with one another in the cell wall, although the biological significance of these interactions is still a subject of debate.

Xyloglucan consists like cellulose of a backbone of 1-4-linked β -D-Glc_p residues, but unlike cellulose, contains sidechain substituents. The xyloglucan backbone is highly substituted with α -Xyl substituents at O-6 and, in some cases, with O-acetyl-substituents (Figure 24.1B). Each Xyl residue may itself be extended by the addition of one or more monosaccharides including β -D-Galp, α -L-Fuc_p, α -L-Araf and α -L-Arap, β -D-Xyl_p, β -D-GalA_p, and O-acetyl-substituents. Twenty-three structurally unique side chains have been identified to date, although only a subset of these are synthesized by a single plant species.

Early models of dicot primary walls predicted that xyloglucan polysaccharides acted as tethers between cellulose microfibrils and that controlled enzymatic cleavage or reorganization of xyloglucan by proteins facilitated wall expansion and thus plant cell growth. However, genetic engineering of xyloglucan structures in Arabidopsis plants has revealed that eliminating xyloglucan entirely from the primary wall has remarkably little effect on overall plant growth and development, challenging this notion. In contrast, removing only xyloglucan sidechain substituents beyond the initial Xyl residues is detrimental to plant growth, while replacing the extended sidechains with various glycosyl-moieties restores plant growth independent of the glycosyl-moiety added. It has now been suggested that xyloglucan acts a spacer molecule to keep cellulose microfibrils apart and/or that pectin may have a more important role in controlling wall expansion than previously thought.

Glucuronarabinoxylan (GAX, Figure 24.1C) is the predominant noncellulosic polysaccharide in the type II walls of the grasses. Its backbone is composed of 1-4-linked β -D-Xyl_p residues, many of which are substituted at O-3 with α -L-Araf residues. These Araf residues may be further substituted at O-2 with an α -L-Araf or a β -D-Xyl_p residue. A small number of the backbone residues are substituted at O-2 with α -D-Glc_pA and its 4-O-methylated counterpart (MeGlc_pA).

The presence of 1-3,1-4-linked β -glucans (also referred to as mixed-linkage glucans) in the walls of grasses was once considered to be a unique feature of these plants. However, structurally related mixed-linkage β -glucans have also been identified in the walls of Selaginella (lycopod) and Equisetum (horsetails), although the evolutionary relationship between these β -glucans is not known. Mixed linkage glucan in grasses is present mainly in young tissues, and hence it has been suggested that it acts as a carbohydrate storage molecule as also indicated by its degradation by the plant.

Pectins. Pectins are structurally complex polysaccharides that contain 1-4-linked α -Gal_pA. Three structurally distinct pectins—homogalacturonan, substituted galacturonan, and rhamnagalacturonan—have been identified in plant cell walls (Figure 24.2). Homogalacturonan, which may account for up to 65%

the pectin in a primary wall, is composed of 1-4-linked α -Gal_pA. The carboxyl group may be methyl-esterified, thus neutralizing the negative charge of this glycosyl moiety, and the uronic acid itself may be acetylated at O-2 or O-3. The extent of methyl-esterification is controlled by pectin methyl esterases present in the wall and affects the ability of homogalacturonan-containing glycans to form ionic calcium cross-links with themselves and with other pectic polymers. Such interactions alter the mechanical properties of the wall and may influence plant growth and development.

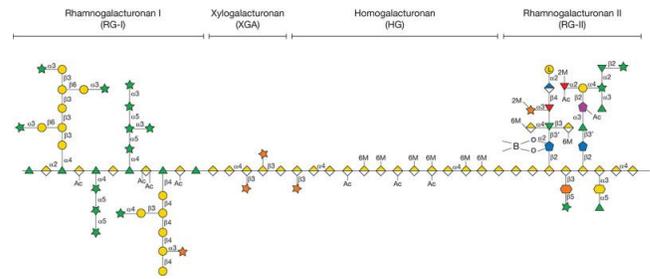


Figure 24.2. Schematic structure of pectin. The three main pectic polysaccharides are shown: rhamnagalacturonan I (RG-I), rhamnagalacturonan II (RG-II), and homogalacturonan (HG). The borate ester is formed between the apiosyl residue in side chain A of each RG-II monomer. A region of substituted galacturonan, referred to as xylogalacturonan (XGA), is also shown. XGA is not present in most cell wall pectins. The relative abundance of each pectin domain (HG, RG-I, and RG-II) is dependent on the plant species.

Rhamnagalacturonan-I is a family of polysaccharides with a backbone composed of a repeating disaccharide 4- α -D-GalpA-1-2- α -L-Rhap-1. Many of the Gal_pAs are acetylated at O-2 and/or O-3. Depending on the plant, between 20% and 80% of the Rhap residues may be substituted at O-4 with linear or branched side chains composed predominantly of Ara and Gal, together with smaller amounts of Fuc_p and Glc_pA (Figure 24.2). Little is known about the functions of these side chains and their contribution to the properties of the primary wall.

Substituted galacturonans have a backbone composed of 1-4-linked α -D-GalpA acid residues that are substituted to varying degrees with mono-, di-, or oligosaccharides. For example, xylogalacturonans contain single β -Xyl_p residues linked to O-3 of some of the backbone residues (Figure 24.2), whereas apiogalacturonans have β -D-apiose (Apif) and apiobiose linked to O-2 of some of the backbone residues. Apiogalacturonans have only been identified in the walls of duckweeds and seagrasses.

Rhamnagalacturonan-II, which accounts for between 2% and 5% of the primary cell wall, is the most structurally complex polysaccharide yet identified in nature. It is composed of 12 different monosaccharides linked together by up to 21 different glycosidic linkages (Figure 24.2). Four structurally different side chains are attached to the galacturonan backbone. Two structurally conserved disaccharides (side chains C and D) are linked to O-2 of the backbone. The A and B side chains, which contain between 7 and 9 monosaccharides, are linked to O-3 of the

backbone. Several of the monosaccharides in RG-II are O-methylated and/or O-acetylated.

Virtually all of the RG-II exists in the primary wall as a dimer cross-linked by a borate ester. The ester is formed between the Api residue in sidechain A of each RG-II monomer (Figure 24.2). The dimer forms rapidly *in vitro* when the RG-II monomer is reacted with boric acid and a divalent cation. However, the mechanism and site of dimer formation in planta has not been determined. Borate cross-linking of RG-II is likely to have substantial effects on the properties of pectin and the primary wall as RG-II is itself linked to homogalacturonan (Figure 24.2). Indeed, mutations that affect RG-II structure and cross-linking result in plants with abnormal walls and severe growth defects. Swollen primary walls and abnormal growth together with reduced RG-II cross-linking are also characteristic of boron deficient plants.

Pectin is believed to exist in the cell wall as a macromolecular complex comprised of structural domains—homogalacturonan, rhamnogalacturonan, and substituted galacturonan—that are covalently and non-covalently linked to one another. However, there is only a limited understanding of how these structural domains are organized (Figure 24.2). Molecular modelling of a pectin (~50 kDa) containing homogalacturonan (degree of polymerization ~100) and rhamnogalacturonan with arabinogalactan side chains, together with modelling of RG-II conformation, have begun to provide insights into the conformations and relative dimensions of each pectin structural domain.

The homogalacturonan region has a persistence length of approximately 20 GalpA residues, which is likely to be sufficient to stabilize junction zones formed with Ca⁺⁺. *In vitro* studies suggest that the maximum stability of such junction zones is obtained with oligomers containing approximately 15 nonesterified GalpA residues. Thus, controlling the distribution of methyl-ester groups along the homogalacturonan backbone provides a mechanism to regulate the physical properties of pectin, including its ability to form gels. For a gel to form and not to be brittle, other features, including sequences that interrupt interchain associations in the pectin macromolecule, may be important. For example, the structural diversity and the conformational flexibility of the oligosaccharide side chains of the rhamnogalacturonan domain will limit or prevent interchain pairing. The presence of 1-2-linked Rhap residues does not introduce kinks into the backbone geometry of rhamnogalacturonan and thereby limit interchain associations. Rather, it is the side chains linked to these residues that are responsible for preventing or limiting interchain associations.

The conformations of the four side chains attached to the homogalacturonan backbone may lead RG-II to adopt a "disk-like" shape. Well-defined tertiary structures are predicted for the monomer and the dimer. In the dimer, borate-ester cross-linking and Ca⁺⁺ interchain pairing further stabilizes the two disks. The apparent resistance of RG-II to wall-modifying enzymes together with the formation of a cation-stabilized RG-II dimer likely results in a structure that resists temporal changes. In contrast, homogalacturonan is continually modified by the action of wall

enzymes, and its contribution to wall architecture is, therefore time-dependent.

Increased knowledge of the physical properties of primary wall polysaccharides and proteoglycans is required to understand how modulating the amounts, and structural features of a few common polysaccharides and glycan domains lead to primary walls with diverse properties and functionalities. Further research is also needed to determine if wall structure and function results from the noncovalent interactions of polysaccharides and proteoglycans or from the formation of glycan-containing architectural units with specific structural and functional roles. The latter scenario is analogous to the organization of proteoglycans and O-linked mucins in the extracellular matrix of animal cells.

PLANT SECONDARY CELL WALL GLYCANS

The secondary walls of woody tissue and grasses are composed predominantly of cellulose, hemicellulose, and polyphenol lignin. The inclusion of lignin results in a hydrophobic composite that is a major contributor to the structural characteristics of secondary walls.

Heteroxylans are the major hemicellulosic polysaccharide present in the secondary (lignified) cell walls of flowering plants. These heteroxylans are classified according to the type and abundance of the substituents on the 1-4-linked β -D-Xylp residues of the polysaccharide backbone. Glucuronoxylans, which are major components in the secondary walls of woody and herbaceous eudicots, have an α -D-GlcpA or MeGlcpA substituent at O-2 (Figure 24.1C). Gymnosperm secondary walls contain arabinoglucuronoxylans (AGXs), which in addition to MeGlcpA substituents, have Araf residues attached to O-3 of some of the backbone residues. The glucuronoarabinoxylan in the secondary walls of grasses typically contains less Araf residues than their primary wall counterpart. Ferulic or coumaric acids are often esterified to the Araf residues of xylan in grass primary and secondary cells walls. There is some evidence that the lignin polymer is covalently linked to the secondary wall hemicelluloses.

Eudicot and gymnosperm secondary wall xylans have a well-defined glycosyl sequence 1-4- β -D-Xylp-1-3- α -L-Rhap-1-2- α -D-GalpA-1-4-D-Xylp at their reducing end (Figure 24.1C). This sequence is required for normal xylan synthesis during secondary cell wall formation and may have a role in regulating the polymers chain length. This sequence is present at the reducing end of heteroxylans of all monocots except the grasses.

HEMICELLULOSE AND PECTIN BIOSYNTHESIS

Genes that encode polysaccharide biosynthetic enzymes, including many of those required for xyloglucan, glucuronoxylan, arabinoxylan, cellulose synthesis and some of those required for pectin synthesis have been identified. This information, together with improved methods demonstrate the enzymatic activity of recombinant plant glycosyltransferases (GTs) *in vitro* and wall structural analyses of the corresponding plant mutants is providing a framework for an increased understanding of how plant cell wall polysaccharides are synthesized. In contrast to cellulose,

which is synthesized at the plasma membrane where the glucan chains are extruded into the apoplast to form cellulose microfibrils, and lignin which is polymerized in the apoplast by random non-enzymatic radical reactions, most pectins and hemicellulose are synthesized in the Golgi apparatus where they are secreted into the apoplast via vesicles. Despite advances in identifying and understanding the GTs involved in polysaccharide synthesis, we still do not know how many of the wall polymers are synthesized by Golgi-localized multienzyme complexes or if they are assembled by GTs localized in different regions of the Golgi apparatus. We also do not understand how the newly synthesized polymers are assembled into a functional cell wall in the apoplast.

PLANTS PRODUCE PROEOGLYCAN CONTAINING O-LINKED OLIGOSACCHARIDES AND O-LINKED POLYSACCHARIDES

Plants produce glycoproteins and proteoglycans that contain oligo- or polysaccharides that are linked to hydroxyproline (Hyp) and serine (Ser) albeit in low abundance. Hyp is formed post-translationally by endoplasmic reticulum (ER)-localized prolyl hydroxylases and is O-glycosylated in the ER and in the Golgi apparatus. The degree and type of Hyp glycosylation is determined to a large extent by the protein's primary sequence and the arrangement of Hyp residues. Hyp glycosylation is initiated by the addition of an Ara or a Gal residue. Contiguous Hyp residues are arabinosylated, whereas clustered but noncontiguous Hyp residues are galactosylated. Ser residues and occasionally threonine residues may also be O-glycosylated in these proteins.

Three structurally distinct plant proteoglycans containing glycosylated Hyp and Ser—the extensins, proline/hydroxyproline-rich proteoglycans, and arabinogalactan proteins—have been identified. Extensins are hydroxyproline-rich proteoglycans with Ser(Hyp)₄ repeat sequences and contain between 50% and 60% (w/w) glycan. Most of the carbohydrate exists as oligosaccharides containing one to four Ara residues linked to Hyp together with a small number of single Gal residues α -linked to Ser. The genetic modification of extensin glycosylation levels resulted in the blockage of polarized cell growth, e.g. in root hairs, underpinning the importance of extensin glycosylation in plant growth and development. The proline/hydroxyproline-rich proteoglycans, which contain from 3% to 70% (w/w) carbohydrate, are distinguished from the extensins by amino acid sequence. Both of these families of hydroxyproline-rich glycoproteins (HRGPs) likely have a structural role in the cell wall. The expression of genes involved in their synthesis is developmentally regulated and is often induced by wounding and fungal attack of plant tissues. Various plant glycopeptide signalling molecules such as *Clavata* contain arabinosylated hydroxyproline and have numerous roles in plant growth and development.

Arabinogalactan proteins (AGPs) have a glycan content of up to 90% (w/w). Chains of between 30 and 150 monosaccharides are linked to the protein by Gal-O-Ser and Gal-O-Hyp linkages. These chains have a 1-3-linked β -Galp backbone that is extensively substituted at O-6 with side chains of 1-6-linked β -Galp. These side chains are terminated with Ara, GlcpA, and Fucp residues. Some AGPs may contain homogalacturonan, RG-I and

xylan covalently linked to the arabinogalactan (Figure 24.3) thereby forming a protein-hemicellulose-pectin complex referred to as APAP1. The location of this complex in the plant and its biological function remains to be determined.

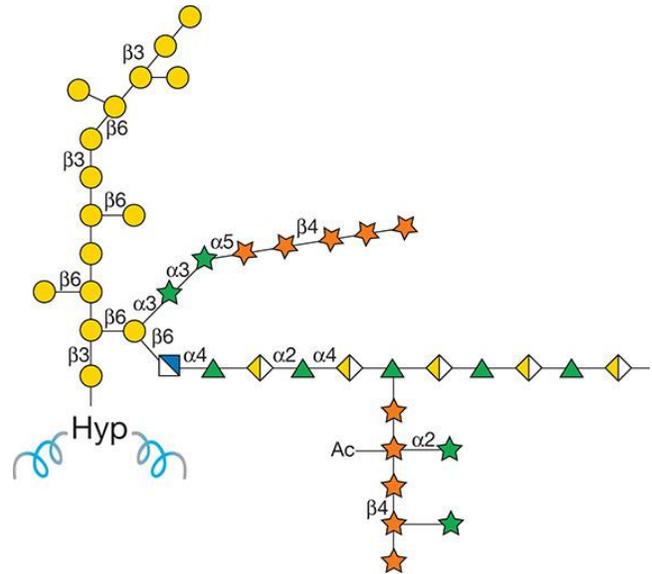


Table 24.3. Schematic structure of the proteoglycan referred to as arabinoxylan pectin arabinogalactan protein1 (APAP1).

Several AGPs are secreted into the cell wall, whereas others are linked to the plasma membrane by a glycosylphosphatidylinositol (GPI) anchor. Plant GPI anchors contain a phosphoceramide core. The glycan portion of the GPI anchor of pear cell AGP has the sequence α -D-Manp1-2 α -D-Manp-1-6 α -D-Manp-1-4-GlcpN-inositol. At least 50% of the Manp attached to the GlcpN is itself substituted at O-4 with a β -Galp, a feature that may be unique to plants. Many functions have been proposed for the AGPs, including their participation in signalling, development, cell expansion, cell proliferation, and somatic embryogenesis.

THE N-LINKED GLYCANS OF PLANT GLYCOPROTEINS HAVE UNIQUE STRUCTURES

Many of the proteins that have passed through the plant secretory system contain N-linked oligomannose, complex, hybrid, or paucimannose-type glycans (Figure 24.4). The initial stages of the synthesis of these N-glycans, including the transfer of the oligosaccharide precursor from its dolichol derivative and the control of protein folding in the ER are comparable in plants and animals. However, two modifications of N-glycans during passage through the Golgi are unique to plants.

Oligomannose-type N-glycans are often trimmed in the cis-Golgi and then modified in the medial-Golgi by N-GlcpNAc transferase I (GnT-I) catalyzed addition of GlcpNAc to the distal Man of the core. In reactions that are unique to plants, a β -Xylp is often added to O-2 of the core Manp. In the trans-Golgi, α -Fucp may be added to O-3 of the GlcpNAc residue that is itself linked to asparagine. The XylT and FucT that catalyze these reactions act independently of one another but do require at least one

terminal GlcpNAc residue for activity. The FucT is related to the Lewis FucT family, whereas the XylIT is unrelated to other known GTs.

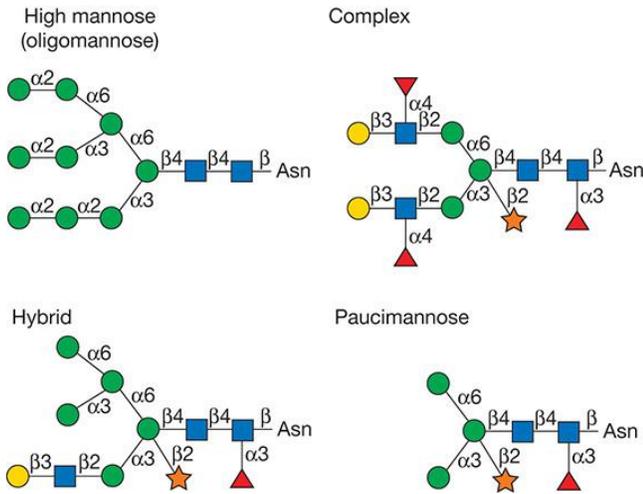


Figure 24.4 Types of N-glycans identified in plants

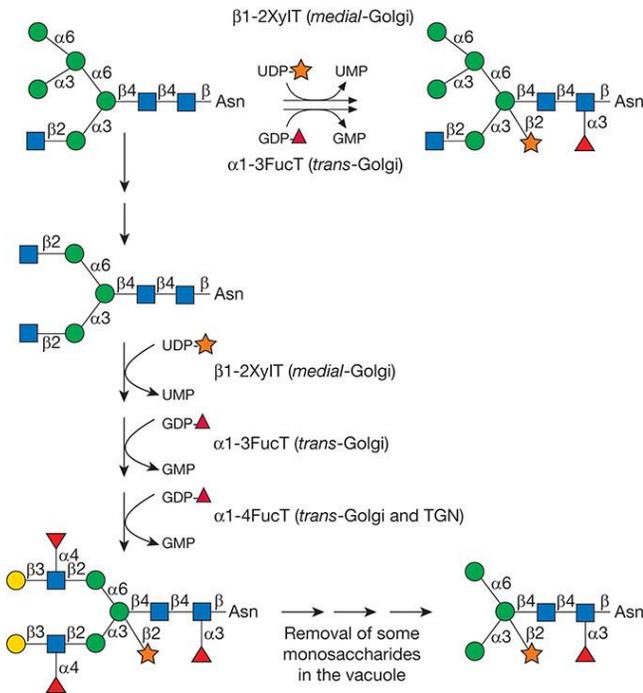


Figure 24.5 Processing of N-glycans in the plant secretory system. Only those events that are unique to plants are shown in detail.

The xylosylated and fucosylated N-glycans are often trimmed by α -mannosidase II. A second GlcNAc may then be added by GnT-II. Some plant N-glycans do not undergo further mannose trimming and proceed through the Golgi as hybrid-type N-glycans. Complex and hybrid-type N-glycans may be further modified by the addition of Galp and Fucp in the trans-Golgi. Plant glycoproteins are either secreted from the cell or transported to the vacuoles. Many of the glycoproteins present in the vacuoles contain

paucimannose type glycans, suggesting that they are trimmed by vacuolar glycosidases.

The presence of sialic acid in the N-glycans of plant glycoproteins was claimed but likely represented environmental contamination. Plants do have genes that encode proteins containing sequences similar to sialyltransferase motifs, but their functions have not been established.

Plant-specific modifications of N-glycans result in glycoproteins that are often highly immunogenic and cause allergic responses in humans. The demonstration that complex N-glycans are not essential for plant growth, initiated studies to engineering plant N-glycosylation pathways to produce glycoproteins that do not activate the mammalian immune system. Plants lacking the GTs that add Xylp and Fucp to N-linked glycans produce glycoproteins lacking immunogenic glyco-epitopes. Other glycosylation pathways involved in adding sialic acid and Gal must be introduced to fully "humanize" the glycoproteins if plants are to be used to produce recombinant therapeutic glycoproteins.

ALGAL GLYCANS

Only a few glycans of green algae in both the Chlorophyta and Streptophyta clades have been studied in detail. However, understanding their diverse polysaccharide structures could have important implications for the evolution of the more complex structures present in current land plants.

Evolutionary ancestors to land plants are the Chlorophyta. The extracellular matrix of, for example, the chlorophyte *Chlamydomonas reinhardtii* is a crystalline lattice formed from hydroxyproline rich proteins. The Streptophyta comprise land plants and the charophyte algae, including the late diverging Zygnematomyceae, which encompass the closest relatives to land plants. The cell wall of the unicellular Zygnematomyceae alga *Penium margaritaceum* has cellulose, pectic and hemicellulose-like glycans. With the emergence of whole-genome sequencing and extensive transcriptomic analyses, detailed phylogenetic trees of glycan synthesis related GTs are being established, including GTs from green algae and land plants. Such data will support the proposing of hypotheses regarding when and how different structural forms of diverse cell wall polysaccharides evolved across the Chlorophyta and Streptophyta. Detailed structural and architectural analyses of algal and high plant cell wall glycans and the substrate and acceptor specificities of the corresponding GTs will be required to resolve the evolutionary trends within the Viridiplantae.

Numerous specialized algal polysaccharides are utilized by humankind. For example, several of the polysaccharides produced by red and brown algae are used in the food industry as gelling agents, stabilizers, thickeners, and emulsifiers. They are also used in paints, cosmetics, paper, and as reagents for scientific research. These polysaccharides include agarose (agar) and carrageenan, which are sulfated galactans obtained from red seaweeds. These polysaccharides are comprised of the repeating disaccharide 3- β -D-Galp-1-4-3,6-anhydro- α -L-Galp-1 unit. Some of the D-Galp and L-Galp units are O-methylated. Pyruvate and sulfate groups may also be present in small quantities. Another example is alginate, a

linear polysaccharide composed of 1-4-linked β -D-ManpA and its C-5 epimer 1-4-linked α -L-GulpA, present in various species of brown seaweed. These monosaccharides are typically arranged in blocks of either ManpA or GulpA separated by regions comprised of 4-ManpA-1-4-GulpA-1 sequences. Brown seaweeds produce polysaccharides that have potential in the treatment of diseases. Laminaran, a linear storage polysaccharide composed of 1-3- and 1-6-linked β -D-Glcp residues. There are reports that laminaran has antiapoptotic and antitumor activities. Fucoidans are a group of sulfated polysaccharides isolated from several brown algae that have been reported to have anticoagulant, antitumor, antithrombosis, anti-inflammatory, and antiviral properties. Fucoidans have a backbone of 1-3-linked α -Fucp that is substituted at O-2 with fucose and at O-4 with sulfate or fucose. Other fucoidans have backbones of alternating 1-3- and 1-4-linked α -Fucp residues.

PLANT GLYCOLIPIDS

Glycoglycerol lipids are the most abundant glycolipids in plants. Mono- and digalactosyldiacylglycerol have been identified in all plants, whereas tri- and tetra-galactosyldiacylglycerol have a more restricted taxonomic distribution. The synthesis of these galactolipids is initiated by the formation of diacylglycerol in the ER membrane and the chloroplast membrane. Galactolipids formed in the ER membrane contain predominantly C16 fatty acids at the sn2 position and C18 fatty acids at the sn3 position. The chloroplast pathway produces C18 fatty acids at both positions. Each of these fatty acids is then desaturated to 16:3 or 18:3 acyl groups.

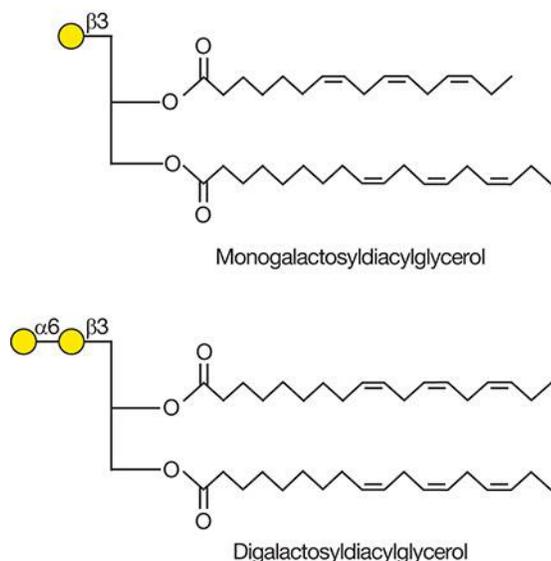


Figure 24.6. The most abundant plant galactolipids.

Monogalactosyldiacylglycerol (MGDG) is synthesized by the transfer of Gal from UDP-Gal to diacylglycerol by an MGDG synthase. Digalactosyldiacylglycerol (DGDG) is formed from MGDG by the transfer of Gal from UDP-Gal by a DGDG synthase. These reactions occur primarily in the outer chloroplast membrane. The products are then transported to the inner membrane and the thylakoid membranes of the chloroplast. The

presence and abundance of MGDG in the chloroplast thylakoid membrane are important for normal photosynthesis to occur. Sulfoquinovosyldiacylglycerol (SQDG) which is formed from diacylglycerol, is also abundant in the thylakoid membrane and may also have a role in photosynthesis.

Small amounts of MGDG and DGDG are present in the cytosolic leaflet of the plasma membrane, although the mechanism of galactolipid exchange among the membranes is not understood. The outer leaflet of the plasma membrane is instead dominated by glycosylinositolphosphorylceramides (GIPCs), lipids that are absent in animals. GIPCs were first described in the 1950s, but have remained poorly characterized partly due to their insolubility in standard membrane extraction protocols. GIPCs are estimated to comprise up to 40% of the plasma membrane, with more than 200 species described. Following synthesis in the ER, the ceramide moves to the Golgi, where an inositol phosphate is added, followed by several glycosylation steps. The first sugar added in higher plants is GlcpA, but after that, the identity of the sugar unit, the number of units, and the degree of branching depend on tissue and species. GIPCs from two (Series A) to seven (Series F) sugars have been regularly described, containing Manp, Glcp, GlcpN, GlcpNAc, Galp, and Araf, although there have been reports of GIPCs with up to twenty sugars. Further improvements to the isolation and characterization methods are required so that the full array of glycan structures can be described. No gangliosides have been identified in plants. However, there is evidence that Kdo-containing lipids with homology to bacterial lipid A are present in plant organelle membranes.

GIPC glycosylation mutants can be lethal or have severe developmental phenotypes. Altering the glycan structure can induce severe constitutive defence responses, alter plant-microbe interactions (both pathogens and beneficial microbes) and reduce cellulose content of the cell wall. Only a small number of GIPC GTs have yet been described, and much work is required to fully describe their function.

OTHER PLANT GLYCOCONJUGATES

Plants produce numerous phenolics, terpenes, steroids, and alkaloids that are collectively referred to as secondary metabolites. Many of these compounds are O-glycosylated or contain sugars linked via N, S, or C atoms. Glycosylated secondary metabolites often have important roles in a plants response to biotic and abiotic challenges and may also have value as pharmaceuticals.

In general, the addition of a single sugar or an oligosaccharide may increase water solubility, enhance chemical stability, or alter both chemical and biological activity. For example, the activity of several plant hormones may be regulated by converting them to their glucose esters or their glucosides. Digoxin and oleandrin are potent cardiac glycosides isolated from foxglove and oleander, respectively. Myrosinase-catalyzed cleavage of S-linked Glc from glucosinolates leads to the formation of pungent mustard oils when mustard and horseradish are damaged. The steviol glycosides, which are far sweeter than sucrose, are used as natural sugar substitutes. The bitter taste of citrus fruits is due to naringin, a glycosylated flavanoid.

FURTHER READING

- Painter, T. 1983. Algal polysaccharides. In *The polysaccharides* (ed. Aspinall, G), pp. 195–285. Academic, New York.
- Pérez, S, Mazeau, K, du Penhoat, CH. 2000. The three-dimensional structures of the pectic polysaccharides. *Plant Physiol Biochem* 38: 37–55.
- Gachon, CM, Langlois-Meurinne, M, Saindrenan, P. 2005. Plant secondary metabolism glycosyltransferases: The emerging functional analysis. *Trends Plant Sci* 10: 542–549.
- Hözl, G, and Dörmann, P. 2007. Structure and function of glyco-glycerolipids in plants and bacteria. *Prog Lipid Res* 46: 225–243.
- Albersheim, P, Darvill, A, Roberts, K, Sederoff, R, Staehelin, A. 2010. *Plant cell walls. From chemistry to biology.* Garland Science, New York.
- Gomord, V, Fitchette, A-C, Menu-Bouaouiche, L, Saint-Jore-Dupas, C, Plasson, C, Michaud, D, Faye, L. 2010. Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant Biotechnol J* 8: 564–587.
- Bar-Peled, M, and O'Neill, MA. 2011. Plant nucleotide sugar formation, interconversion, and salvage by sugar recycling. *Annu Rev Plant Biol* 62: 127–155.
- Kieliszewski, MJ, Lamport, D, Tan, L, Cannon, M. 2011. Hydroxyproline-rich glycoproteins: Form and function. *Annu Plant Rev* 41: 321–342.
- Popper, ZA, Michel, G, Hervé, C, Domozych, DS, Willats, WGT, Tuohy, MG, Kloareg, B, Stengel, DB. 2011. Evolution and diversity of plant cell walls: From algae to flowering plants. *Annu Rev Plant Biol* 62: 567–590.
- Atmodjo, MA, Hao, Z, Mohnen, D. 2013. Evolving views of pectin biosynthesis. *Annu Rev Plant Biol* 64: 747–779.
- Pauly, M, Gille, S, Liu, L, Mansoori, N, de souza, A, Schultink, A, Xiong, G. 2013. Hemicellulose biosynthesis. *Planta* 238: 627–642.
- Cosgrove, DJ. 2014. Re-constructing our models of cellulose and primary cell wall assembly. *Curr Opin Plant Biol* 22: 122–131.
- Knoch, E, Dilokpimol, A, Geshi, N. 2014. Arabinogalactan proteins: Focus on carbohydrate active enzymes. *Frontiers Plant Sci* 5: 198.
- Matsubayashi, Y. 2014. Posttranslationally modified small-peptide signals in plants. *Annu Rev Plant Biol* 65: 385–413.
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The Carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42:D490–D495.
- McNamara, JT, Morgan, JL, Zimmer, J. 2015. A molecular description of cellulose biosynthesis. *Annu Rev Biochem* 84: 895–921.
- Mortimer, JC, Scheller, HV. 2020 Synthesis and function of complex sphingolipid glycosylation. *Trends Plant Sci* 25: 522–524
- Höfte H, Voxeur A. 2017. Plant cell walls. *Curr Biol* 27:PR865–R870
- Jiao C, Sørensen I, Sun X, Sun H, Behar H, Alseekh S, Philippe G, Palacio Lopez K, Sun L, Reed R, Jeon S, Kiyonami R, Zhang S, Fernie AR, Brumer H, Domozych DS, Fei Z, Rose JKC. 2020. The *Penium margaritaceum* Genome: Hallmarks of the Origins of Land Plants. *Cell* 181:1097-1111