Chapter 2
Chemical Composition and Structure of Natural Lignocellulose

Abstract The wide variety of natural cellulosic materials has complex and uneven components. Cellulose, hemicellulose, and lignin comprise the main composition of cell walls of plants and are important components of natural lignocellulosic materials. Cellulose molecules determine the cell wall framework, and pectin is located between the cellulose microfilaments of the cell wall. In addition, cellulosic materials contain rich cell wall protein, pigment, and ash. Understanding of the chemical composition and structure of natural lignocellulosic materials, characteristics of each component, and interrelationships between various components would contribute to the research and development regarding natural cellulose materials. This chapter mainly describes the chemical composition and structure of natural cellulosic materials.

Keywords Cellulose • Hemicellulose • Lignin • Cell wall protein • Biological properties

2.1 Main Components of Natural Lignocellulosic Materials

Cell walls of plants consist mainly of three organic compounds: cellulose, hemicellulose, and lignin. These compounds are also major components of natural lignocellulosic materials. Cellulose molecules arrange regularly, gather into bundles, and determine the framework of the cell wall. Fibers are filled with hemicellulose and lignin. The structure of the plant cell wall is compact. There is different bonding among cellulose, hemicellulose, and lignin. Cellulose and hemicellulose or lignin molecules are mainly coupled by a hydrogen bond. In addition to the hydrogen bond, there is the chemical bonding between hemicellulose and lignin, which results in the lignin, isolated from natural lignocelluloses, always contains a small amount of carbohydrates. The chemical bonds between the hemicellulose and lignin mainly refer to the chemical bonds between galactose residues, arabinose residues on the side chains of hemicellulose molecules and lignin, and carbohydrates, with this

Table 2.1 Structure and chemical composition of cellulose, hemicellulose, and lignin in cell walls of plants [2]

	Lignin	Hemicellulose	Cellulose
Subunits	Guaiacylpropane (G), syringylpropane (S), p-hydroxyphenylpropane (H)	D-Xylose, mannose, L-arabinose, galactose, glucuronic acid	D-Pyran glucose units
Bonds between the subunits	Various ether bonds and carbon-carbon bond, mainly β-O-4 ether bond	β-1,4-Glycosidic bonds in main chains; β-1.2-, β-1.3-, β-1.6-glycosidic bonds in side chains	β-1,4-Glycosidic bonds
Polymerization	4,000	Less than 200	Several hundred to tens of thousands
Polymer	G lignin, GS lignin, GSH lignin	Polyxylose, galactoglucomannan (Gal-Glu-Man), glucomannan (Glu-Man)	β-Glucan
Composition	Amorphous, inhomogeneous, nonlinear three-dimensional polymer	Three-dimensional inhomogeneous molecular with a small crystalline region	Three-dimensional linear molecular composed of the crystalline region and the amorphous region
Bonds between three components	Contain chemical bond with hemicellulose	Contains chemical bond with lignin	Without chemical bond

knowledge gained through research on the separated lignin-carbohydrate complexes (LCCs) [1, 2]. Table 2.1 shows the chemical composition and structure of cellulose, hemicellulose, and lignin.

Cell walls mainly consist of cellulose, hemicellulose, and lignin in a 4:3:3 ratio. This ratio differs from sources such as hardwood, softwood, and herbs. Besides these three components, natural lignocellulosic materials contain a small amount of pectin, nitrogenous compounds, and the secret ash. For instance, the element content of wood is about 50 % carbon, 6 % hydrogen, 44 % oxygen, and 0.05–0.4 % nitrogen.

2.2   Biological Structure of Plant Cell Walls

One of the most important components in the plant cell wall is cellulose, which determines the wall structure. Cellulose is a natural high molecular polymer composed of glucose residues, with cellobiose as the basic coupling unit. It is the most abundant renewable resource in nature, and cellulose metabolism is an

important part of the biosphere's carbon cycle [3]. Gao et al. used cotton fiber as a raw material for research on the structure of cellulose in plant cell walls. Cotton fiber is the only natural pure cellulose; its cellulose content can reach 95–97 %, and its crystallinity is about 70 %. Scanning electron microscopy showed that the diameter of fibrils is about 500 nm, so it is the largest structural unit of cellulose. A fibril is composed of entwined microfibrils, which makes cellulose stronger than steel wire of the same thickness. Microfibrils would entwine into a network as the basic framework of the cell wall; their diameter is about 10–25 nm. The microfiber is formed with elementary fibrils arranged in parallel. The diameter of the elementary fibril is approximately between 2 and 4 nm, the structural unit of which is cellulose molecules linked by β-1,4-glycosidic bonds [3]. In some regions of the microfibrils, cellulose molecules are arranged in an orderly fashion, so the cellulose has crystal properties. This regular arrangement of cellulose molecules in microfibrils is called the *micelle*. Some noncellulose molecules also exist in the network structures of cellulose, including hemicelluloses, pectin, and so on.

Another important component in the cell wall is lignin. Except for cellulose, it is the most abundant large-molecule polymer in the cell wall. Botanically, lignin encloses the bundle cells, such as wood fibers and sclerenchyma cells. From a chemistry point of view, phenylpropanoid derivatives are the basic units of the lignin; they combine into high molecular substances by ether bonds or carbon-carbon bonds. According to the physical characteristics, lignin is hard, which increases the hardness of the cell wall. Commonly, the cell wall of plants with a supporting function and mechanical action always contains a high lignin content. The lignin content is about 27–32 % in woody plants and about 14–25 % in herbaceous plants [4].

The cell wall of protective tissue usually also contains cutin, suberin, wax, and other fatty substances. For example, the cell wall surfaces of the epidermic cell are covered with cutin; the cell walls of cork cells in secondary protective tissue contain suberin, cutin, and suberin, often combined with wax. These components greatly reduce water loss from the plants.

Depending on the time of formation and chemical composition, the cell wall can be divided into the primary wall and the secondary wall. Plant cell wall formation follows after cell division; the primary cell wall is formed in the new cell plate, and intercellular layers are formed between primary cell walls. As cells differentiate, secondary cell walls are formed inside primary cell walls and outside the protoplast; with the further differentiation of the cells, the structure of the cell wall gradually adapts to the function of the cells. The intercellular layer is formed outside the primary wall, but it is difficult to identify the boundaries between them, especially after the secondary wall has been formed. The intercellular layer is mainly composed of pectic substances, which are amorphous colloids and have strong hydrophobicity and plasticity. Multicell plants rely on the pectin substances to bond neighboring cells together. Pectins are easily broken down by acids or enzymes, resulting in the isolation of cells. When the cells are lignified, the sequence of lignification degree is middle lamella (ML), primary wall, then secondary wall [5].

In the process of cell growth, the primary wall is formed from some protoplast secretions. The main components of the primary cell wall are polysaccharides, proteins (such as the expansins), and many other enzymes, glycoproteins, and some ions (such as calcium). Main polysaccharides of primary walls are cellulose, hemicellulose, and pectin. Cellulose accounts for 15–30 % dry weight of the primary cell wall. The hemicellulose interacts with celluloses, forming a network with microfibrils. Pectin accounts for about 30 % of polysaccharide of the primary cell wall [6]. Those cells with an active division property usually do not have the primary wall, which is similar to those mature cells relating to photosynthesis, respiration, and secretory action. These cells without secondary cell walls can change specialized forms and restore the ability to divide and differentiate into different cell types. Therefore, these cells that only have primary walls are relevant to callus reaction and regeneration. Usually, when the primary wall grows, it thickens unevenly. There would be a thin field in the primary wall called the primary pit field. Plasmodesmata, which connect protoplasts of adjacent cells, tend to be concentrated in this field [5]. The main function of the primary cell wall is to provide structural and mechanical support, maintain and determine cell morphology, withstand cell swelling pressure, control the rate and direction of cell growth, promote plant morphogenesis, regulate material diffusion in ML, reserve carbohydrates, maintain resistance to pathogens, resist dehydration, and activate the interaction between source signal molecules and cells [6]. In plants, many cells only have primary walls, but many others have both primary and secondary walls. When cells stop growing and the superficial area of the primary wall no longer increases, the secondary wall begins to form. Some substances produced in the metabolic process of protoplasts deposit on the inside of cell walls, then form secondary cell walls next to the plasma membrane. The secondary wall cells are formed inside the primary cell wall and have some differences in composition compared to the primary cells. In addition to containing cellulose and hemicellulose, the secondary wall contains lignin. Lignin could highly cross-link with each other to enhance mechanical support for the plants to grow upward [6]. The secondary wall is particularly important for those specialized cells that are related to mechanical reinforcement and water transportation. The secondary wall has more celluloses than the primary wall but lacks pectin. Therefore, the secondary cell wall is harder, is less extended than the primary wall, and has no enzymes and glycoprotein. The basic component of the secondary cell wall is hemicellulose. It usually can be divided into three layers: inner layer (S3), middle layer (S2), and outer layer (S1). Different layers have differences in composition, structure, microfibrillar angle, and so on. A large amount of solar energy and carbon fixed by plants is stored in secondary walls. The accumulated biomass in secondary walls accounts for the vast majority of the total plant biomass, which is the main form of biomass on Earth, and are also fiber materials and bioenergy raw materials for human life.

The primary pit field is not covered by the secondary cell wall component, resulting in the formation of many sunken areas called pits. Sometimes, the pits can also occur in the absence of a primary pit field. Pits on the cell wall are often

opposite the pits on the adjacent cell walls; the intercellular layer between the two pits and two layers of primary walls make up the pit membrane, and two opposite pits and pit membrane make up the pit pair. Pits on the secondary wall have two types: the simple pit and the bordered pit. The basic difference between them is that the secondary thickened wall uplifts toward the central part, hangs over the pit cavity, and forms a dome-shaped edge so that the pit aperture is significantly smaller, but this kind of dome-shaped edge does not exist in the simple pit.

The growth of the cell wall includes an increase in surface area and thickness; the growth process is strictly controlled by biochemical reactions in the protoplast. The growth of the cell wall should be in a relaxed state and have a high respiration rate, protein synthesis rate, and water absorption rate. Most newly synthesized microfibrils are superimposed on the original cell wall, but a few insert into the original cell wall. In those cells that grow evenly, such as marrow cells, storage cells, and culture cells, cell wall microfibrils randomly arrange in various directions and form an irregular network. In contrast, in the extended-growth cells, the deposition direction of microfibrils on the side walls makes an acute angle with the extended direction of the cells. When the surface area of the cell is increased, the external microfibrils arranged direction gradually is parallel with the long axis of the cells. Substrates (such as pectin and hemicellulose) and glycoprotein are mainly transported to the cell wall by the Golgi vesicles. The type of substrate is dependent on the development stage of the cell. For example, at the expanding stage of cells, the pectin is predominant in the matrix; otherwise, hemicelluloses predominate at the shrinking stage [5].

2.3 Cellulose

Cellulose is the most abundant renewable organic resource on Earth and is widespread in higher plants, bacteria, marine algae, and other biomass. The total annual amount of cellulose is several billion tons, revealing the huge economic value of it. Cellulose is the main component of the plant cell. Although some animals (such as tunicates) and some bacteria contain cellulose, the content of cellulose in these species is negligible when compared with plants. Cellulose was first separated by Anselme Payen (1839) from timber that was alternately treated with nitric acid and sodium hydroxide solution. It is a β-1,4-linked linear polymer of glucose units and is insoluble in water, dilute acidic solutions, and dilute alkaline solutions at normal temperatures. Although the structure and composition of the cell walls of plants vary widely, the cellulose content usually accounts for 35–50 % of dry weight and, peculiarly, almost 100 % for cotton. Study of the supramolecular structure of natural cellulose showed that the crystalline and noncrystalline phases intertwine to form the cellulose. The noncrystalline phase assumes an amorphous state when tested by X-ray diffraction because most hydroxyl groups on glucose are amorphous. However, large amounts of hydroxyl groups in the crystalline phase
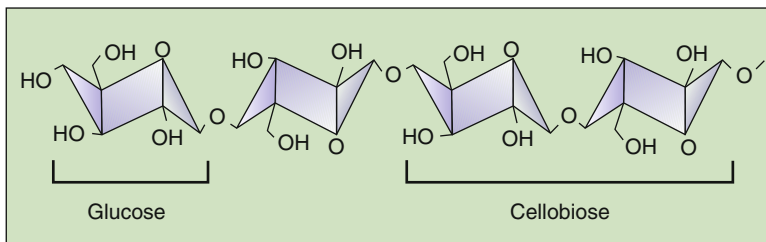
Fig. 2.1 Molecular chain structure of cellulose [8]

form many hydrogen bonds, and these hydrogen bonds construct a huge network that directly contributes the compact crystal structure [7]. In most conditions, the cellulose is wrapped by hemicellulose (dry matter accounting for 20–35 %) and lignin (dry matter accounting for 5–30 %). Cellulose has become an important raw material for the pulp and paper, textile, and fibrous chemical industries. Predictably, bioenergy generated from lignocellulosic materials will become clean energy in the future.

2.3.1 Chemical Structure of Cellulose

Cellulose is a linear homopolymer composed of D-glucopyranose units linked by β-1,4-glycosidic bonds. It mainly contains carbon (44.44 %), hydrogen (6.17 %), and oxygen (49.39 %). The chemical formula of cellulose is $(C_6H_{10}O_5)_n$; n, called the degree of polymerization (DP), represents the number of glucose groups, ranging from hundreds to thousands or even tens of thousands. In the twentieth century, it was proved that cellulose consists of pure dehydrated repeating units of D-glucoses (as shown in Fig. 2.1), and the repeating unit of the cellulose is called cellobiose.

Sodium hydroxide solution at different concentrations and different temperatures could dissolve cellulose with different DP. According to different solubilities under specific conditions, cellulose can be divided into three types: ① α-cellulose, which is dissolved in 16.5 % NaOH at 20 °C; ② β-cellulose, which is deposition extracted after neutralizing the acid solution and the remaining alkaline solution; and ③ γ-cellulose, which is the soluble remainder in the neutralized solution. Staudinger used a viscosity method to measure the DP of these three celluloses. The results indicated that the DPs of α-cellulose, β-cellulose, and γ-cellulose were more than 200, between 10 and 200, and less than 10, respectively. In industry, α-cellulose usually is used to express the purity of cellulose. Traditionally, β-cellulose and γ-cellulose are together called industrial hemicellulose. *Holocellulose* refers to all the carbohydrates in natural cellulose material, which also is the sum of cellulose and hemicellulose [2].

2.3.2   Physical Structure of Cellulose

The physical structure of cellulose refers to the spatial arrangement of different-scale structural units, including the chain structure and aggregation structure of the polymer. The chain structure, also known as the primary structure, shows the geometric arrangement of the atoms or groups in the molecular chain. The short-range structure is the first-level structure and refers to the chemical structure or stereochemical structure of one or several structural units in a single-molecule polymer. Remote structure is the second-level structure and refers to the size of a single-molecule polymer and a special structure. The aggregation structure, also called the secondary structure, refers to the inner structure of the whole polymer, including the crystal structure, noncrystal structure, orientational structure, and liquid crystal structure. The third-level structure term is used to describe how molecules in a polymeric aggregate accumulate each other, such as tangly clew structure and crystal structure formed with ordered folding chains. The chain structure of the polymer is the main structural hierarchy that reflects many characteristics of a polymer, such as melting point, density, solubility, viscosity, adhesion, and so on. The aggregation structure of a polymer is the major factor that determines the service performance of macromolecular compound products [9].

2.3.2.1   Filament Structure

A fibril is a small, stretching unit; these units aggregate and then constitute the structure of some natural and synthetic fiber materials (such as textile fibers, timbers, or fibrous protein); they also make long molecular chains gathered into bundles in one direction. Because the inequality in size of fibrillar aggregation, current terminologies include elementary fibril, microfilament, macrofilament (also called a microfilament bundle) [10]. Natural cellulose has 10,000 glucose units, and the fibril contains approximately 60–80 cellulose molecules. Hydrogen bonds are formed between adjacent molecules. In a certain range of space, hydrogen bonding can be shown in the X-ray pattern when it reaches a certain number. This space is called the crystalline region, and the rest is called the amorphous region. Microfilament is composed of elementary fibrils and is fixed in size. Macrofilament is has more than one microfilament, and its size varies with the sources or processing conditions of raw materials.

The structural model proposed by Fengel for the cell wall of timber is the representative model for the microfilament structure of each layer of the cell wall. He reported that the elementary fibril with a diameter of 3 nm is the most basic structure unit; 16 (4 × 4) elementary fibrils form a fibril with a diameter of 12 nm, then 4 fibrils form a relatively thick microfibril with a diameter of 25 nm, and more than one microfibril form macrofilament. Hemicellulose is filled between the adjacent elementary fibrils; the microfilaments are wrapped with lignin and hemicellulose. A multilayer of several hemicelluloses is filled between fibrils 12 nm

in diameter; the monolayer of hemicelluloses is filled between 3-nm elementary fibrils. Because the microfilament is formed before the lignification of the cell wall, lignin is surrounded by only microfilament with a diameter of 25 nm [10]. It is generally thought that the movement of the liquid in the cell wall occurs mainly at the elementary fibril level of hemicellulose; usually, contraction and swelling processes also mainly occur at this level.

In recent years, elementary fibrils with a diameter of 1.7 nm have been found with high-resolution electron microscopy. Because the fibrils are surrounded by hemicellulose and the microfilaments are surrounded by a large amount of lignin, the microfilament can be observed after delignification, and the elementary fibril can be found only after the hydrolysis of hemicellulose. The measurement results for elementary fibrils according to most investigators indicated that the diameter of elementary fibrils is between 30 and 35 Å, and an elementary fibril is composed of 40 cellulose macromolecular chains. The ordered region of cellulose macromolecules is the crystalline region; irregular regions form an incomplete crystalline structure. When the crystal diameter is about 3 nm, a monolayer of hemicellulose would exist around the crystal, and several cellulose crystals combine to form the cellulose crystal beam called the nanofiber. The diameter of a nanofiber is about 2–3 nm, and hemicellulose and lignin are around it. In summary, the fiber cell walls are composed of many fibers, and larger microfibrils always consist of smaller elementary fibrils.

2.3.2.2 Aggregation Structure

The aggregation state of cellulose, also called the supramolecular structure of cellulose, mainly refers to how cellulose molecules arrange to form crystal and amorphous structure, then elementary fibril, fibril, and microfibril structures. X-ray diffraction studies showed that, in the aggregates of cellulose macromolecules, molecules in crystal structure arrange regularly and display a clear X-ray pattern, so the density of cellulose in the crystalline region is high (1.588 g·cm^{-3}). Molecular chains in the amorphous region arrange irregularly and loosely, so the distance between molecules is large. The density of cellulose in the amorphous region is low, 1.500 g·cm^{-3}. However, the molecule chain is almost parallel with the main spindle of cellulose. The cellulose crystallinity, generally between 30 and 80 %, refers to the percentage of all the cellulose occupied in the crystalline region [2].

The crystallization of cellulose shows pleomorphism. There are five kinds of crystal modification in solid cellulose, whose characteristics can be reflected by characteristics of their unit cells. Under certain conditions, cellulose crystals can be converted into many crystal variants. Type I is the crystal form of the natural cellulose. Types II, III, IV, and X are those crystal forms of "artificial" cellulose under artificial processing. Now, the commonly accepted cell structure of type I is the monoclinic unit cell model introduced by Meyer and Misch in 1937 [9]. Extensive chemical treatment and heat treatment will change the crystal form; for example, ball milling can destroy crystal lattice completely. There is no distinct boundary from the crystalline region to the amorphous region. Each crystalline

region is called a microcrystal (also called a micel or micella). Since free hydroxyls at position 2, 3, and 6 of glucosyl in cellulose microcrystal regions have formed hydrogen bonds, only amorphous regions contain some free hydroxyls.

2.3.3 Physicochemical Properties of Cellulose

2.3.3.1 Chemical Properties of Cellulose

Every glucosyl ring of cellulose has three active hydroxyls: one primary hydroxyl group and two secondary hydroxyl groups. Thus, cellulose may have a series of chemical reactions related to hydroxyl. However, these hydroxyl groups also can form hydrogen bonds between molecules, which has a profound influence on the morphology and reactivity of cellulose chains, especially the intermolecular hydrogen bond formed by oxhydryl at C3 and oxygen at an adjacent molecule ring. These hydrogen bonds not only can enforce the linear integrity and rigidity of the cellulose molecule but also can make molecule chains range closely to form a highly ordered crystalline region [10]. The accessibility of cellulose refers to the difficulty reagents have in arriving at the cellulose hydroxyl. In heterogeneous reactions, the accessibility is mainly affected by the ratio of the cellulose crystalline regions to the amorphous regions. The reactivity of cellulose is the reactive capability of the primary hydroxyl and the secondary hydroxyl at the cellulose ring. Generally, because of the smallest steric hindrance, the reactivity of the primary hydroxyl groups is higher than for the secondary hydroxyl groups, so the reactivity of hydroxyl at C6 with a bulky substituting group is higher. For example, esterification of toluenesulfonyl chloride chiefly occurs in the primary hydroxyl. The reversible reaction occurs mainly in the hydroxyl group at C6, and an irreversible reaction always occurs in the hydroxyl group at C2. Thus, for the esterification of the cellulose, the reactivity of the hydroxyl group at C6 is the highest, but for the etherification, C2 is the highest [10].

The degradation of cellulose is an important reaction that can be used to produce cellulose products. Acid degradation, microbial degradation, and alkaline degradation are mainly to break the glycosidic bonds between two adjacent glucose molecules; an alkali peeling reaction and oxidation-reduction reaction of cellulose usually act on reducing ends of celluloses, and the oxidative degradation of the cellulose occurs mainly in dissociating hydroxyls at C2, C3, and C6 of the glucosyl ring. Cellulose molecule chains will form carbonyls at C2 when oxidized to some degree and then be degraded in the following alkali treatment process by the elimination reaction of β-alkoxy. After disconnecting the glycosidic bond, the reaction product is formed and then degraded to a series of organic acids [9].

Esterification and etherification reactions of cellulose act on three alcoholic hydroxyls of cellulose molecule monomer. They can greatly change the properties of cellulose, thereby producing many valuable derivatives of cellulose, such as sulfonic ester, cellulose acetate, cellulose nitrate, and cellulose ether (carboxyl

methyl cellulose, methyl cellulose, ethyl cellulose). To enhance the reactivity of ester and the ether bond of cellulose in multiphase medium and improve the quality of cellulose ester and ether, some pretreatments need to be performed. The main methods include the following: ① Preswelling treatment of celluloses can weaken the hydroxyl-binding forces between cellulose molecules to increase the reagents' diffusion velocity in cellulose, such as being immersed in concentrated caustic solution, activated by glacial acetic acid, and so on. ② The elimination of crystallinity by the ethamine can only change the DP by 20 % (usually, the concentration is higher than 1 %) when the concentration of ethamine is more than 71 %. Therefore, it was analyzed that ethylamine only enters into the microfilaments, only makes the amorphous region swell, and does not greatly change crystallization regions. ③ Cellulose derivatives with a high degree of substitution and many hydroxyl groups are substituted substantially, so the total free hydroxy declines and water absorbability decreases. So, actually some cellulose derivatives with a low degree of substitution have higher water absorbability, such as methyl, ethyl, hydroxyethyl, hydroxymethyl cellulose ether, and so on. These groups lead to the swelling of the cellulose structure and binding force decrease in macromolecules. They further result in the increase of water absorbability, degree of hydrolysis, and wrinkle resistance. The improvement of wrinkle-resistant property can be used to enhance the stiffness and moisture resistance of cardboard; also, it can improve the burst strength and the dimensional stability of paper [9].

2.3.3.2   Physical Properties of Cellulose

Free hydroxyls of cellulose have a strong attraction to many solvents and solutions, but adsorbed water only exists in the amorphous region, not the crystalline region. In the moisture sorption process, the hydrogen bonds of the amorphous region in the dry cellulose constantly could be broken; the hydrogen bonds in the cellulose molecules are replaced by the hydrogen bonds between cellulose molecules and water molecules, even forming new hydrogen bonds, and some hydrogen bonds in cellulose molecules remain. In the desorption process, because of the obstruction from inside, hydrogen bonds between cellulose molecules and water molecules cannot be broken completely and reversibly, resulting in hysteresis. Some water absorbed by cellulose enters into the amorphous region of cellulose and forms the water bound by hydrogen bonds, called bound water. Molecules of bound water attracted by hydroxyl of cellulose are arranged in a certain direction and have a high density, making swelling the cellulose and generating a heat effect. When the celluloses absorb water that reaches the fiber saturation point, water molecules continue to enter into the cell lumina and pores of cellulose to form a main layer adsorbed water or capillary water, which is called free water. No heat effect and swellability of cellulose exist when absorbing free water [9].

When solids absorb liquids, the configuration homogeneity does not change, but solids become soft with the decrease of the inner cohesive force and increased volume. This phenomenon is known as the *swellability*. Swellability of cellulose

is divided into swellability in the crystalline regions and swellability between crystalline regions. The former refers to the fact that the swelling agent can only reach the surface of crystalline and amorphous regions, and the X-ray pattern of cellulose does not change. The latter refers to the fact that the crystallization regions of microfilaments are permeated with the swelling agent and then swell to generate new crystalline lattice and display a new X-ray pattern. Unlimited swelling of the cellulose is dissolution. The hydroxyl groups in the cellulose have polarities. As a swelling agent, the greater polarity the liquid has, the greater degree of swelling the cellulose has. The metal ion in the alkali solution is usually in the form of aquo ions, which is more favorable for entering the crystallization region. Usually, 15–20 % NaOH will cause swelling within crystalline regions. If the alkali concentration is increased, the radius of aquo ions is reduced because the ion density is too high, resulting in the drop of swellability. Except for alkali, the swellability of other swelling agents, sorted from strong to weak, is as follows: phosphoric acid, water, polar organic solvents, and so on. Cellulose is saturated in a concentrated solution of NaOH to generate alkali cellulose. Although alkali cellulose is washed with water and dried, such changes cannot restore it to its original condition. Alkali cellulose may have a crystalline form of hydrate cellulose that is more stable than that of natural cellulose, which would increase its absorbability and make it easy to react with a variety of reagents. Using alkali to impregnate celluloses is also called *mercerization*. In addition, alkali cellulose is the important intermediate product for the production of viscose fibers and derivatives of cellulose ether [9].

Characteristics of polymer compounds are high molecular weight and a strong cohesive force. They have movement difficulties in the system and a poor diffusion capacity, so they cannot be dispersed in a timely manner in the solvent. The solvent dissolved with celluloses is not the real cellulose solution, but the mixed product is obtained by mixing celluloses and components in liquids. The solvents of the cellulose can be divided into two categories: aqueous and nonaqueous. Aqueous solvents include the following: ① Inorganic acids, such as H_2SO_4 (65– 80 %), HCl (40–42 %), H_3PO_4 (73–83 %), and HNO_3 (84 %) can lead to the homogeneous hydrolysis of cellulose. Concentrated HNO_3 (66 %) does not dissolve the cellulose but forms an addition compound with cellulose. ② Lewis acids, such as LiCl, $ZnCl_2$, $Be(ClO_4)_2$, thiocyanate, iodide, bromide, and others, could dissolve celluloses with a low DP. ③ Inorganic bases, such as NaOH, hydrazine and sodium zincate, NaOH, and others can only dissolve cellulose with a low DP. ④ Organic bases, such as quaternary ammonium bases $(CH_3)_4NOH$, amine oxides, and others, are also aqueous solvents. The application of amine oxide solvent to dissolve cellulose can be used to manufacture the man-made fibers. ⑤ Complexes, such as copper oxide ammonia (cuoxam), copper ethylenediamine (cuen), cobalt hydroxide ethylenediamine (cooxen), zinc ethylenediamine (zincoxen), cadmium ethylenedi-amine (cadoxen), and the iron–tartaric acid–sodium complex (EWNN, an aqueous alkaline solution of iron sodium tartrate) are included as aqueous solvents [9].

A nonaqueous solvent of cellulose refers to a nonaqueous or less-aqueous solvent that is based on the organic solvents. It consists of activators and organic liquids. The organic solvents can be used as a component of the active agent and as a

solvent of the activator, which can make the solvent have a larger polarity to dissolve cellulose. Therefore, the mechanism of cellulose dissolved in a nonaqueous solvent system cannot be easily explained by swelling theory, as in aqueous solvents. The detailed mechanism of this process can be expressed as follows: ① An oxygen atom and a hydrogen atom of cellulose hydroxyl participate in the interaction of the EDA; the oxygen atom and the hydrogen atom act as a π-electron donor and a δ-electron acceptor, respectively. ② The active agent in the solvent system has an electron donor center and an electron receiving center; the spatial location of these two centers is suitable for interaction with the oxygen atom and hydrogen atom of cellulose hydroxyl. ③ There is necessarily a suitable scope for the interaction strength of the EDA, causing the centers of the donor and acceptor to interact in polar organic solvents. When the hydroxyl charge separates to some extent, the complex of cellulose molecular chains is separated and dissolved.

Several different systems of nonaqueous solvents of cellulose exist: ① Paraformaldehyde/dimethyl sulfoxide (PF/DMSO) is an excellent new solvent system that is not biodegradable. PF resolves into formaldehyde by heating, and then formaldehyde reacts with the hydroxyl group of cellulose to generate hydroxymethylcellulose, which is dissolved in DMSO. ② Dinitrogen tetroxide/dimethylformamide (N_2O_4/DMF or DMSO) is an intermediary derivative of the reaction of N_2O_4 with cellulose to generate nitrite esters; it can be dissolved in DMF or DMSO. ③ Amine oxides directly dissolve cellulose without the intermediate derivatives. ④ Liquid ammonia/ammonium thiocyanate restricts the dissolution of the cellulose; the solvent consisting of 72.1 % (w/w) NH_4SCN, 26.5 % (w/w) NH_3, and 1.4 % (w/w) H_2O has the maximum dissolving ability. ⑤ Lithium chloride/dimethylacetamide (LiCl/DMAC) also directly dissolves cellulose without the intermediate derivatives. At room temperature, the LiCl/DMAC solution is stable and can be used for reeling off raw silk and film forming. Recently, research on nonaqueous solvents of cellulose has been active; they not only can be used to produce artificial fiber and films but also can be available for processing cellulosic materials and for the use of cellulose in homogeneous conditions to produce cellulose derivatives. The problems of cellulose solvents are the low solubility of cellulose, high price and low recovery of solvents, and environmental pollution.

Thermal decomposition of cellulose is in the narrow temperature range of 300–375 °C. Different products depend on different temperatures. Heated at a low temperature (200–280 °C), cellulose mainly dehydrates into dewatering cellulose and then forms charcoal and gas products. Heated at higher temperatures, cellulose separates into flammable volatile products (tar). The most important intermediate product of cellulose high-temperature thermal degradation is laevoglucose, which can be further degraded into low molecular products and tar-like products. Tar-like products can be polymerized into an aromatic ring structure similar to graphite structure at high temperature (400 °C or higher). Mechanical degradation of the cellulose occurs because cellulose in the mechanical process can effectively absorb

mechanical energy, causing changes of morphology and microstructure; these changes are shown as decreased DP and crystallinity and significantly increased accessibility [10].

2.3.4   Biosynthesis of Cellulose

2.3.4.1   Cellulose Synthesizing Site

Some research has already forecast that assembling of cellulose microfilaments is finished in the enzyme complex located in the extending top of the cellulose. Then, scientists hypothesized that a cellulose synthase complex was made up of many subunits, and each subunit synthesized single-chain glucose, then polymerized it to the ordered particles of cellulose. But, until 1976, through the freeze-etching technique, the complex located in the end of the cellulose microfilament was first observed in green algae. This verified the hypothesis that assembling of cellulose microfilaments is finished in the enzyme complex located in the extending top of the cellulose. The subunits of the complex are arranged linearly in three lines and form the linear enzyme complex where cellulose is synthesized [11]. The alternating self-aggregation and dispersion of the complex determine that the microfilament arrangement direction changes periodically, resulting in different levels of micro-filaments arranged perpendicular to each other. Later, similar terminal complexes were observed in bacteria, mosses, ferns, green algae, and microtubule plants, but in corns, spherical complexes were found [12].

A terminal complex like a rosette has been observed in higher plants and concentrates in the cellulose gathering place. Each six cellulose synthase sub-units of rosette synthesizes 6 glucose chains and then forms microfilaments with 36 chains. The microfilament directions are mutually different in the dif-ferent levels of the cell wall, which makes the cell wall in any direction have high mechanical strength [13]. The rosette complex not only has the function of synthase but also can bring glucose chains to the surface of cytoplasm. A complete rosette complex is essential for the synthesis of crystalline cellulose. The terminal complex would disappear or be changed when EDTA (ethylenedi-aminetetraacetic acid) is used to handle oysters or Congo red is used to han-dle banana cells, further causing the interruption or disturbance of cellulose biosynthesis. Once the cellulose synthesis recovers, the terminal complex reap-pears. The mutation of the *Ces*AI gene in *Arabidopsis* heat-sensitive mutants (RSWL) will lead to the reduction of the cellulose content, the content of antacid β-glucan, and the number of rosette complexes in the cell membrane, perhaps because the mutation of this enzyme disrupts the structure of the rosette complex. This indicates that the RSW1 (a radial swelling phenotype) maybe one component constituting the rosette complex [14]. Kimura et al. [15] used the polyclonal antibody technique on the central area in cotton *CesA* protein to verify that the plasma membrane has a rose-like structure, which was the complex of cellulose synthase, and was *Ces*A protein located in rosettes. This finding demonstrated that

the *Ces*A gene has an important role in cellulose synthesis and provided direct evidence for the hypothesis that cellulose biosynthesis takes place in the terminal complex of the rosette [11].

It is uncertain whether the rosette complex is composed of identical subunits or different subunits. Key information on assembly also still cannot be clarified. The bacteria linear terminal complex goes through the cell membrane and lipopolysaccharide layer and mainly synthesizes the 1α types of cellulose I, which is the metastable monoclinic system. But, the rosette terminal complex in plants is part of the cell membrane, which mainly synthesizes the 1β type of cellulose I, which is a stable monoclinic form. Therefore, it is generally believed that cellulose synthesis in cotton also occurs in the cellulose synthase complex connected with the plasma membrane.

In the process of cellulose biosynthesis, in addition to the terminal complex, another polypeptide with a molecular mass of 18 kDa also plays an important role. This polypeptide does not exist in the plasma membrane but is loosely connected with the plasma membrane. So, it is unlikely to be the component with catalytic activity in the cellulose synthase complexes. However, it may have regulating effect because it can combine with 2,6-dichloro-phenyl nitrile, which is the inhibitor of cellulose synthesis.

2.3.4.2 Substrate for Cellulose Synthesis

Identifying the substrate of cellulose synthesis has been difficult. Previous studies reported that callosum generated in the translating period was the substrate of cellulose synthesis. This result was derived from the fact that, with in vivo conditions, the speed of synthesizing callosum from the substrate that can supply a radioactive label is higher than the predictable accumulation level, and the conversion of callose radioactivity is consistent with the change of cellulose. The conversion may occur because of the transglycosylation of β-l,3-glucose polymerase; the discovery that the cell wall had the activity of β-l,3-glucose polymerase also supports the hypothesis mentioned. Callose is a homopolymer of β-1,3-linked glucoses. It plays an important role in the regulation of vital processes, such as metabolism of the sieve tube, the development of the gametophyte, and so on. The composition and resolution of callose are directly related to the normal growth of plant metabolic processes [16].

The precursor for the biosynthesis of cellulose is uridinediphosphate-D-glucose (UDPG). However, in the past, UDPG was thought to be obtained by the catalysis of UDPG pyrophosphorylase. Now, with research on cotton fiber development, it has been found that the catalytic reaction by the sucrose synthase could also provide UDPG. In the formation stage of the secondary wall of the cotton fibers, sucrose synthase is connected with cellulose synthase, which may be used as carbon path. Two sets of evidence support this view. It has been proved that in vitro biosynthesis of cellulose takes cellulose synthase from the cell membrane of cotton fiber as the enzyme source and UDPG as a substrate. The products are always β-1,3-glucan and
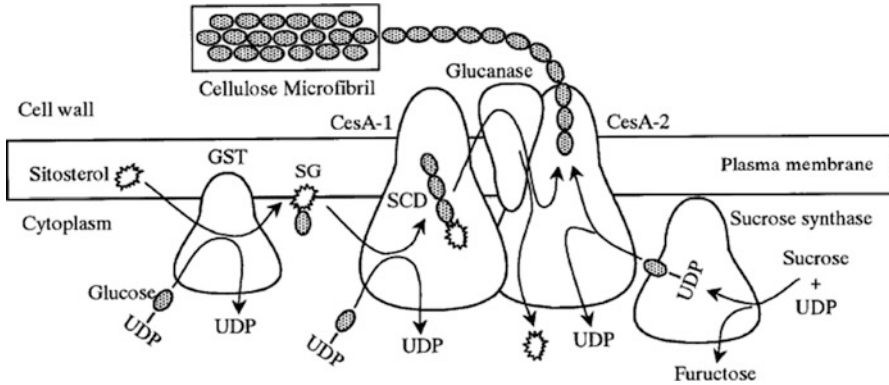
Fig. 2.2 Biosynthetic pathway of plant cellulose [18, 19]

β-1,4-glucan, a few calloses, and a small amount of cellulose. When taking sucrose as a substrate, the synthetic rate of cellulose is equal to the synthetic rate of callose, and sometimes is more than the synthetic rate of callose; the absence of Ca^{2+} is more conducive to cellulose synthesis. This shows that cellulose synthase can only use the UDGP directly from the catalytic reaction by sucrose synthase, but callose synthase can directly use the free UDPG. The other evidence is that, in the mutant of cotton without fibers, there is no gene expression of sucrose synthase in the ovule epidermal cell, but there is large gene expression of sucrose synthase in wild-type fiber cells; this shows that sucrose synthase has a close relation to the development of cotton fiber [17].

2.3.4.3 Cellulose Synthesis Process

Currently, there are different hypotheses about the mechanism of cellulose synthesis. One hypothesis suggests that the extension of the glucan chain is caused by the moving glycosyltransferase catalyzing several glucosyl residues to connect to the end of the growing cellulose chain. The synthesis of acetobacter cellulose may belong to such a mechanism. The other hypothesis claims that some short glucan could polymerize with lipid or protein and form the mature cellulose polymer. Peng et al. (2001) found that *Ces*A protein adhered to the end of noncrystalline fibers in fiber cell wall fragments treated with herbicide. Meanwhile, a small amount of the attached glucose chain was detected in these fragments, indicating that *Ces*A glycosyltransferase probably took sitosterol-β-glucoside (SG) as primers to start glucan polymerization. First, the SG and UDP-glucose is used as the substrate to generate sitosterol cellodextrin (SCD) and continue polymerization by the cellulose synthase, then enters into the crystallization process of cellulose (Fig. 2.2) [18, 19]. Schrick et al. [20] studied the relationship of biosynthesis of *Arabidopsis* sitosterol and biosynthesis of cellulose; they found that sitosterol is important in cellulose
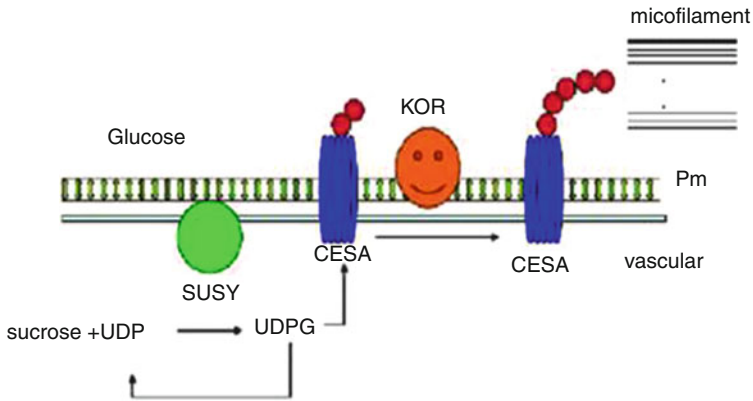
Fig. 2.3 Biosynthetic model of plant cellulose [21]

biosynthesis in *Arabidopsis*, and SG was not the only primer of cellulose synthesis. But, recent research showed that the mechanism of cellulose synthesis in higher plants may belong to the second hypothesis.

Cellulose synthesis in plants is the initiation, elongation, and termination of β-1,4-glycosidic chains. Under in vitro conditions, it is difficult to prove that UDP-Glu could directly synthesize cellulose, but it is generally considered that UDP-Glu is the substrate for it. A recent study found that SG from cotton fiber can be used as an initial extension when cellulose synthase catalyzes glucoside chains and then forms oligosaccharide linked with lipid, which is called SCDs. In the process of cellulose synthesis, cellulase (KOR) cut down SG from SCD so that the β-1,4-glucoside chain could extend more effectively. Further analysis showed that the catalytic subunit of *Ces*A is in one side of the cell membrane; the catalytic subunit of KOR is in the other side. This agrees with the following mechanism: *Ces*A accepts UDP-Glu from hyaloplasm to synthesize glucan, and then glucan crosses the plasma membrane to be further converted under the effect of cellulose. This is another significant discovery in understanding the molecular mechanisms of cellulose biosynthesis [11].

Therefore, the current anticipated process of plant cellulose biosynthesis has three steps: (1) Sucrose synthase associated with plasma membrane guides the UDP-glucose to provide a substrate for the synthesis of cellulose. (2) Hexagonal polymer is organized by the coexpressional multiple *Ces*A, and glucose monomers are polymerized to form a glucan chain; in the meantime, the discharged UDP is recycled to sucrose synthase. (3) KORRIGAN (KOR), a kind of cellulase relating to the membrane, is regarded as the editor or monitor of the cellulose microfilament and can cut defective glucan chains (Fig. 2.3). Therefore, *Ces*A, sucrose synthase, and the KOR protein interact with each other to coregulate the biosynthesis of cellulose in plants [21].

2.3.4.4 Cellulose Synthase

In 1996, by adopting complementary DNA (cDNA) random sequencing and series analysis, Pear et al. first cloned the β-1,4-glucosyltransferase gene *Cel*A, which encoded the catalytic subunit of cellulose synthase from cotton and rice [22]. Cellulose synthase has a polygene phenomenon; moreover, it constitutes a huge gene family with proteins such as cellulose synthase. Studies have shown that cellulose biosynthesis requires the participation of multiple cellulose synthase genes. Furthermore, different cellulose synthase may be related to the different tissues and cellulose synthesis of different cell wall layers.

Since the *Ces*A gene was cloned from cotton, cellulose synthase genes in many plants have also been separated and cloned in succession. The complete genomic sequence of *Arabidopsis thaliana* has been completed, so research on the *Arabidopsis Ces*A gene is the clearest. The length of the cellulose synthase gene is probably in the range of 3.5–5 kb and contains 9–13 introns. The length of the transcription product is between 3.0 and 3.5 kb. The length of the coded peptide chain is about 985–1,088 amino acids. Introns and exons of these genes are arranged conservatively; the presence of introns determines the most important differences between genes. Taking the homology comparison in the amino acid sequence of *Ces*A protein in *Arabidopsis*, corn, cotton, rice, and poplar, it has been found that each has three Asp residues and eight transmembrane domains. The three Asp residues are linked closely and together with the conserved sequence of QXXRw to form active sites used to connect with substrates on one side of the cytoplasm membrane [11]. In addition, the n-terminal of *Ces*A protein contains two conserved zinc finger domains or LIM domains (cysteine-rich zinc-binding domains) which contain four conserved and tandem cysteine residues (X3Cx2-4CX12-15CX2C). This sequence can be combined with DNA and may have a relationship with interaction between subunits of *Ces*A protein. Compared with *Ces*A protein of bacteria, there are two characteristic areas: the plant-conserved region (CR-p) and the hypervariable region (HVR) [23].

The number of amino acids in the *Ces*A protein family is approximately between 985 and 1,088; their homology 53–98 %. Studies indicated that in *Arabidopsis*, except for the identified *Ces*A gene family, there is a category of genes containing the conserved sequence of D, D, D, QXXRW similar to the structure of *Ces*A genes. However, they all belong to the glycosyltransferase family 2 (GT family 2) and do not have a specific sequence in *Ces*A genes. Meanwhile, amino acid sequences encoded by them have 7–35 % homology with the protein sequence of *Ces*A, so they are called cellulose synthase-like (Csl) protein [24]. Furthermore, they all have complete structural features of the membrane, and there are three to six transmembrane structures and one or two transmembrane domains in the C-terminal and N-terminal, respectively. So far, the function of these genes is unclear, and they may have a relationship with the synthesis of other β-polysaccharides in the cell wall.

Biosynthesis of plant cellulose requires the participation of multiple cellulose synthase genes. Analysis of gene expression showed that two or more *Ces*A proteins

are involved in the biosynthesis of cellulose in the same developmental stage of the same cell. Phenotypes of the IRX1 mutant are similar to IRX3; both appeared to have irregular xylem and a declining cellulose content; it also has been found that expression sites and periods of these two genes were identical. Ptr*Ces*A2 and Ptr*Ces*A1 of poplar are homologous with IRX1 and IRX3. Their expression sites and period are the same. They both are expressed at the secondary cell wall synthesis of xylem. So, it was suggested that these two genes may be expressed in the same cell and related to the formation of the second wall [21]. Currently, as most studies involved with cellulose biosynthetic genes are focus on the *Ces*A gene, the *Ces*A gene has been cloned from microbes and many plants. Further study showed that the mechanism of cellulose biosynthesis was complicated. Except for cellulose synthase, sitosterol glycosyltransferase, cellulase, sucrose synthase (SUSY), cytoskeletal protein, and Rac13 proteins were likely related to cellulose synthesis. Sucrose synthase is related to the supply of substrate from cellulose biosynthesis. The experiments proved that, in three different heterotrophic systems, sucrose synthase could improve the efficiency of cellulose biosynthesis. It can catalyze the reaction of sucrose and UDP to produce UDP-glucose and fructose, which can directly offer substrate to improve the biosynthesis of the cellulose. In 2004, Konishi et al. [25] further confirmed that sucrose synthase could use sucrose to synthesize UDP-glucose, which could be used directly for cellulose synthesis.

In short, cellulose biosynthesis is a highly complex biological process; thorough clarification of its mechanism will require extensive research.

2.4 Hemicellulose

Hemicellulose is another main component in plant fiber materials. In 1891, Schulz [26] thought that polysaccharides that easily separated from plant tissue were semifinished products of cellulose or precursor molecules of cellulose, so they were named hemicellulose. He also found that this component was easy to be hydrolyzed to monosaccharides in hot, dilute mineral acid or cold 5 % NaOH solution. To hemicellulose, this concept is vague in terms of both chemical structure and biological function. In recent years, people learned more about hemicellulose with improvements of polysaccharide purification as well as application of various types of chromatography, spectroscopy, nuclear magnetic resonance, mass spectrometry (MS), and electron microscopy. Aspinall in 1962 defined that hemicellulose was derived from polysaccharides of plants and included the basic chain containing residues of D-xylose, D-mannose, D-glucose, or D-galactose and other glycosyls as branched chains linked to this basic chain. The purification of hemicelluloses was conducted according to the different alkaline solubilities with cellulose. So, in 1978, Whistler thought that hemicellulose was the polysaccharide extracted by an alkali solution, except cellulose and pectin. Unlike cellulose, hemicellulose is a copolymer composed of different amounts of several saccharide molecules [2].

2.4.1 Chemical Structure of Hemicellulose

The content and structure of hemicellulose in various plants are different. The research on the chemical structure is mainly about the composition of the main chain and branched chains of glucans in hemicellulose. The main chain may consist of one or more types of glycosyls, and the connections between glycosyl are also different. Raw materials from different producing areas and different parts have different glycan compositions. Therefore, to illustrate the chemical structure, glycans must be classified first. It is generally believed that hemicellulose is the glucan in the matrix of the cell, and the main components are xylan, xyloglucan, glucomannan, manna, galactomannan, callose, and so on [27].

2.4.1.1 Chemical Structure of Xylan Hemicelluloses

Almost all plants contain xylan. D-Xylosyls are linked with each other to form homopolymer linear molecules as the main chain. Xylan hemicellulose is the glucan with a backbone of 1,4-β-D-xylopyranose and with branch chains of 4-oxymethyl-glucuronic acid.

The hemicellulose of hardwoods and gramineous forbs is mainly composed of this kind of polysaccharide. The hemicellulose of Gramineae also contains L-arabinofuranose linking to the main chain as branch chains. The number of branch chains depends on different kinds of plants. The typical molecular structure of hemicellulose of Gramineae is chiefly composed of β-D-xylopyranosyl, which is linked by β-1,4-glucosidic bonds. Branch chains consist of L-arabinofuranosyl and D-glucuronopyranosyl, respectively, on C3 and C2 of the main chain; there are also branch chains composed of xylosyl and acetyl (xylosyl acetate). The DP of hemicellulose in Gramineae is less than 100. Xylan hemicellluloses in timber are composed of linear xylans linked by β-1,4-glucosidic bonds, with some different short-branch chains linked to the main chain, similar to the Gramineae. However, average polymerization is higher than 100. In addition, hemicellulose from softwoods and hardwoods also has the distinction. Hemicellulose of hardwood is chiefly acidic xylans that have been partly acetylized; for example, the content of this hemicellulose in birch is about 35 %, while this content in *Euonymus bungenus* is only 13 % [28]. Xylan hemicellulose in softwoods is 4-O-methyl-glucuronic acid arabinose-xylan with almost no acetyl, while O-acety-L-4-O-methyl-glucuronic acid xylan is the most important hemicellulose in hardwoods [2].

2.4.1.2 Chemical Structure of Mannan Hemicellulose

Softwoods contain the highest content of mannan hemicellulose; some hardwoods also have mannan hemicellulose, but grass has little. Mannose and glucose are linked by β(1→4) bonds to form inhomogeneous polymer as the main chain.

The main chain of mannan hemicellulose in hardwoods is composed of glucose and mannose; the proportion of these two glycosyls is 1.5–2:1, and the average DP is 60–70; whether it is acetylated remains unclear. For the mannan hemicellulose of softwoods, glucose and mannose in a ratio of 3:1 arrange randomly to form the main chain, galactosyl is linked to the glucose or mannose of the main chain by α(1→6) bonds, and acetyl seems to be evenly distributed on the C2 and C3 of mannose. The average DP is more than 60; some can reach 100 [2].

2.4.1.3 Chemical Structure of Xyloglucan

The main chains of xyloglucan and cellulose are composed of D-glucopyranose linked with β-(1 → 4) bonds. The difference is that 75 % glucosyl residues are replaced by α-D-xylopyranose at *O*-6 on the main chain of the former. Xyloglucan mainly contains glucose, xylose, and galactose, and its residual ratio is approximately 4:3:1. According to differences in family, xyloglucan in plants may also contain fucose and arabinose. The main chain of xyloglucan in the dicotyledon is β-(1 → 4) glucans. α-Xylose residues are linked to the *O*-6 of the β-glucose residues. The terminal galactose is linked to the *O*-2 site of xylose residues with β bonds. If fucose is contained, it is linked to the *O*-2 site of galactose residues with α bonds. Sometimes, arabinose exists in xyloglucan, but the amount is small. The content of xyloglucan in the monocotyledon is much different; generally, terminal galactose does not exist, and the contents of xylose and galactose are lower than in the dicotyledon.

2.4.1.4 Chemical Structure of Mannan

Mannan compounds include mannan, galactomannan, glucomannan, glucuronic acid mannan, and so on. Mannose residues are connected by a β-(1 → 4) bond to form mannan, but they form galactomannan if linked to galactose residues by an α-(1 → 6) bond. The backbone of glucomannan is composed of glucose and mannose, which are linked by β-(1 → 4) bonds with the residual ratio of 1:3. Glucomannan also contains one galactose residue as a branch chain. Therefore, it is sometimes called galactoglucomannan. Furthermore, the hydroxyl group of mannose residues may also be acetylated. Glucuronic acid is prevalent in the cell wall, but its content is low. Mannose residues linked by α-(1 → 4) bonds and glucuronic acid residues linked by β-(1 → 2) bonds exist alternately in the main chain of glucuronic acid, whose side chains not only have β-(1 → 6)–linked xylose or galactose but also have 1 → 3–linked arabinose [2].

2.4.1.5 Chemical Structures of Galactan and Arabinogalactan

Galactose residues are connected by β-(1 → 4) bonds to form the backbone of galactan; galactose residues, as the side chain, are attached to the *O*-6. There are

two types of arabinogalactan: The common type has terminal galactose residues linked at *O*-3 or *O*-6, galactose residues linked at *O*-3 or *O*-6, and arabinofuranose residues linked at *O*-3 or *O*-5. Another type has galactose residues linked by *O*-4 or *O*-3 and *O*-4 bonds and arabinofuranose residues linked terminally or by an *O*-5 bond. Arabinogalactan may also be oligosaccharide constituted by several arabinose residues. Further, ferulic acid may be linked to some arabinose and galactose residues. Arabinogalactan in the cell walls may be an independent molecule or as the side chains on the polysaccharide molecules of pectin [2].

2.4.1.6   Chemical Structure of Arabinan

Arabinan entirely consists of arabinose, and α-L-arabinofuranosyl residues are linked with each other at C-5, forming the main chain. Arabinan contains many branched chains; some side chains of arabinofuranosyl are linked to *O*-2 or *O*-3 or simultaneously connected to *O*-2 and *O*-3, and some side chains are composed of arabinose.

2.4.2   Chemical Properties of Hemicellulose

Because of the low DP and few crystalline structures, hemicellulose is more easily degraded in acidic medium than cellulose. But, the category of glycosyl in hemicellulose varies, including the pyran type, furan type, α-glycoside bond-linked type, β-glycoside bond-linked type, L- configuration type, D- configuration type, and so on. The ways of linkage between glycosyls are various, such as 1-2, 1-3, 1-4 and 1-6 links [2]. Most studies showed that hydrolysis of methyl-rabopyranose is the fastest; the others are arranged in decreasing speed as follows: methyl-D-galactopyranoside, methyl-D-xylopyranoside, methyl-D-mannopyranoside, and methyl-D-galactopyranoside, which is the most stable. The β-D type of glycoside is easier to hydrolyze than the α-D type. Generally, the hydrolysis rate of the furan type is faster than that of the pyran type. The hydrolysis rate of glucuronide is 40,000; perhaps the carboxyl has positive control in the glucoside bond.

Hemicellulose is an inhomogeneous glycan composed of a variety of glycosyls, so the reducing ends have many kinds of glycosyls and some branch chains. Similar to cellulose, hemicellulose can have a peeling reaction under mild alkaline conditions. At high temperature, it would have alkaline hydrolysis. Research showed that the speed of alkaline hydrolysis of furan glycosides was many times faster than that of pyran glycosides. Hemicellulose can dissolve in both alkali solution (5 % Na_2CO_3 solution) and acid solution (2 % HCl solution). It has a relative affinity to water, which can make it form a viscous state or become a gelling agent. In rheologic studies of the viscosity of hemicellulose, this phenomenon can be well observed. For example, when the concentration of hemicelluloses in water reaches 0.5 %, the aqueous solution of hemicellulose has a certain consistency

that is the same as in human saliva; when the concentration is 2 %, the solution cannot flow because of the viscosity generated. When the concentration reaches 4 %, the solution is to be regarded as a gel. The affinity of hemicellulose is closely related to its pentose; for example, arabinose and xylose are responsible for fixing water masses on to different structures of the hemicellulose. The greatest benefit brought by this characteristic is to apply pentose in food technology. This feature also illustrates that if the percentage of the pentose in the hemicellulose is too low, the spatial organization keeps pentose away from water, resulting in low affinity of hemicellulose to water [2].

2.4.3 Biosynthesis of Hemicelluloses

Research showed that, in the plant cells obtained, the Golgi apparatus controls the biosynthesis of hemicellulose. Proteins synthesized on the ribosome of the endoplasmic reticulum of the plant cell can be transferred to the Golgi apparatus and form glycosides; the hemicellulose produced is contained in the Golgi vesicles and moved to the cell surface (moved to the cell membrane). In the cell membrane, the Golgi vesicles inosculate to the continuous plasma membrane, further causing the hemicellulose to be stuck to the cell wall. The Golgi apparatus can produce hemicellulose because it can produce the enzymes needed for its synthesis [2].

At the initial stage of polysaccharide biosynthesis in the cell wall, a certain primer is required to accept sugar residues under the effect of polysaccharide synthase. It can be speculated that the primer of polysaccharide biosynthesis in cell wall is a protein because that protein is the biosynthesis primer of many polysaccharides, such as starch, glycogen, and so on. First, the reducing end of the first sugar residue is linked to protein. Then, the protein accepts monosaccharide residues from sugar nucleotides under the effect of polysaccharide synthase, extending the polysaccharide chain. Actually, at the initial stage of the biosynthesis of glucorono-xylan in the pea and xyloglucan biosynthesis in the suspension culture of bean, it was found that some protein primers indeed participated. Inositol can be used as a primer at the initial stage of callose biosynthesis. However, regardless of the circumstance mentioned, it is not clear whether these primers have been "cut" before polysaccharide enters the cell walls or enters the cell wall together with polysaccharide. Therefore, this interesting issue should be studied further.

2.4.4 Physiological Function of Hemicelluloses

Although some data indicate that some hemicellulose, like starch, exerts a function as a storage polysaccharide, combination with lignin and cellulose molecules increases the resistance to enzymatic degradation of the cell wall and the insolubility

of components in the cell wall. Some reports also indicated that xyloglucan is related to plant morphogenesis. However, the generally accepted view is that the main function of hemicellulose is to take part in building the cell wall structure and the regulation of the cell growth process.

2.5   Lignin

Lignin is one of the most abundant organic polymers in plants, just behind cellulose. It is the exclusive chemical composition of gymnosperm and angiosperm. The content of lignin in wood and Gramineae is 20–40% and 15–20 %, respectively. Lignin is the name of a group of substances; their inhomogeneity is manifested in different species of plants, length of growing season, and different parts of the plants. Even in the different morphologies of cells of the same xylem or different cell wall layers, the structures of lignin are not the same [29].

Lignin is a complex composed of complicated phenylpropane units nonlinearly and randomly linked; three main monomers are coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol. Because of the different monomers, lignin can be divided into three types (Fig. 2.4): syringyl lignin polymerized by syringyl propane, guaiacyl lignin polymerized by guaiacyl propane, and hydroxy-phenyl lignin polymerized by hydroxy-phenyl propane. Usually, gymnosperm mainly contains guaiacyl (G) lignin; the dicotyledon mainly contains guaiacyl-syringyl (GS) lignin; the mono-cotyledon mainly contains guaiacyl-syringyl-hydroxy-phenyl (GSH) lignin [30].

At one time, lignin in plant was divided into softwood, hardwood, and grass lignins. Based on the structure of lignin, Gibbs divided lignin into G lignin and GS lignin. G lignin is chiefly formed through dehydrated oligomerization of coniferyl alcohol, and its structure is homogeneous. This kind of lignin has negative Maule reaction because less than 1.5 % of syringaldehyde and about 5 % of p-hydroxybenzaldehyde were generated when oxidized by nitrobenzene. Most lignin in softwood belongs to G lignin, which is copolymerized by guaiacyl and has a positive Maule reaction. GSH lignin is the result of the dehydrated oligomerization of coniferyl alcohol and sinapyl alcohol; the content of lignin is 17–23 %. The ratio of syringyl propane to guaiacyl propane is 0.5∼0.1; it also contains 7–12 % ester groups. p-Coumaryl alcohol in it is linked to lignin in the form of ester [10].
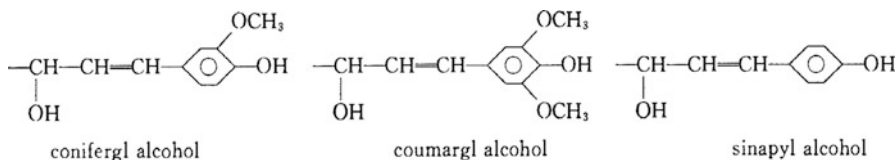

Fig. 2.4 Basic structural unit of lignin

2.5.1   Distribution of Lignin

Early studies indicated that the lignin concentration in the complex middle lamella (CML) layer is above 50 % (mass ratio), while it is about 20 % in the second wall (S layer). However, because the volume of the S layer is far greater than the volume of the ML layer, most of lignin still presents in the second wall. The lignin concentration of the cell corner of the middle lamella (CCML) layer is generally higher than that of the CML layer, even more than 70 % [31]. The lignin density of the CCML layer is about four times that of the S layer; that is, the CCML layer is full of lignins. The ultrastructure of *Salix psammophila* and the lignin distribution in various layers was researched by Xu et al. using optical (or light) microscopy (LM), transmission electron microscopy (TEM), scanning electron microscopy (SEM), electron microscopy/ energy dispersive X-ray analysis (EM-EDXA), and confocal laser scanning microscopy (CLSM). The results showed that the cell wall of *Salix psammophila* was divided into the primary (P) wall, ML, and secondary (S) wall layer; the lignin concentration ratio in the CCML, CML, and S2 was 1.96:1.33:1. CLSM pictures (530 nm) demonstrated that the lignin concentration in the vessel and ML was higher than that in fiber cells [32]. Wu et al. proposed that the composition of the lignin molecule was affected by cell type, the site of the same growth ring, vessel arrangement, and the producing area in the 25 kinds of hardwood studied [33].

Because lignin structural units differ according to timber assortments and measurement regions, the measurement of lignin distribution is uncertain and difficult. Lignin measurement methods include ultraviolet (UV) absorption spectroscopy, EM-EDXA, the interference microscopy (IM) method, CLSM, and so on. The UV method determines the distribution of lignin in the cell wall because that lignin has a typical absorption at 270–280 nm in the UV spectra. The EM-EDXA method can provide distribution information for lignin in different regions. Simultaneously using UV and EM-EDXA to analyze samples not only could provide information on lignin microdistribution but also offer information on the ratio of G-type and S-type structural units in the different microareas. The IM method measures the optical path difference first by the inference microscope and then calculates the refractive index; finally, it obtains the volume percentage and mass percentage of lignin. The CLSM method can be used quickly to measure the biological structure of samples and obtain good measurement results in the Z direction [31].

2.5.2   Structure of Lignin

Lignin is a polyphenolic polymer with a three-dimensional network. Because almost all of the delignification process includes covalent bond rupture of natural lignin, with different separation methods and separation conditions, the lignin structure

would have great differences. Therefore, a structural model was usually used to present the structure of lignin. This kind of structural model only describes a hypothetical structure inferred from the average results. Further, different plant sources, or even lignin isolated from the same plant but in different ways, would have different categories of linkages and composition of functional groups, resulting in the complicated lignin structure. Through nearly two decades of research on lignin structure, a dozen structural models have been proposed. Figure 2.4 is a structural model of lignin from softwoods. It can be seen that the lignin has a complicated structure [34].

Through the study of various types of lignin structural models, lignin is a complicated amorphous polymer with three-dimensional network, which is basically composed of phenylpropane units linked to each other by the irregular coupling of C–C and C–O. Lignin includes three basic structural monomers: p-phenyl monomer (H type) derived from coumaryl alcohol, guaiacyl monomer (G type) derived from coniferyl alcohol, and syringyl monomer (S type) derived from sinapyl alcohol (Fig. 2.4). The structural formula is as follows:

Although lignin only has three basic structures, the quantity proportions of these basic structures vary greatly in different families of plants. Lignin of hardwood includes large amounts of syringyl units. In the UV-photodegradated production from *Eucalyptus urophylla* lignin, ω (syringyl-type compounds) is 58.10 %, and ω (guaiacyl-type compounds) is 18.75 %. Compared with eucalyptus lignin, the pyrolysis products of sulfate pulp lignin are rich in guaiacol. In the sulfate pulp lignin, M (syringyl)/m (guaiacyl) is 4.3:1; in the eucalyptus lignin, this ratio is 6.4:1. Structural units of softwood lignin are mainly guaiacyl-type units; a small amount of p-hydroxyphenyl-type units remains. Wheat straw lignin mainly consists of noncondensed guaiacyl units, noncondensed syringyl units, and other condensed units; the ratio of n (noncondensed guaiacyl units) to n (noncondensed syringyl units) to n (condensed units) is 1.44:1:3.24. The content of the –OCH_3 group in bamboo lignin is similar to the content in hardwood lignin, for example, in Bai jia bamboo (ph, nidularia Mu) lignin, the ratio of n (guaiacyl units) to n (syringyl units) to n (hydroxy-phenyl units) is 1:1.15:0.54 [35].

The coupling modes between each basic unit include β-O-4, β-5, β-1, and so on. Figure 2.5 is a partial section of a softwood lignin structure.

Ether bonds in lignin include phenol-ether bonds, alkyl-ether bonds, dialkyl bonds, diaryl ether bonds, and so on. About two thirds to three quarter phenyl-propane units of lignin are linked to the adjacent structural units by ether bonds; only a small part is present in the form of free phenolic hydroxyl. Phenol-ether bonds account for 70–80 % in these groups, guaiacyl glycerol-β-aryl ethers (β-O-4) account for about half of phenol-ether bonds, followed by the guaiacyl glycerol-α-aryl ethers (α-O-4), also containing other types of ether bonds. Lignin in softwood and hardwood mainly contains aryl glycerol-β-aryl-(β-O-4) ether bonds, approximately reaching half of the lignin in softwood and more than 60 % in hardwood. In the C–C bonds of lignin, the dominant coupling type is β-5, β-β linkage, followed by β-1, β-2, 5–5, and so on [36].
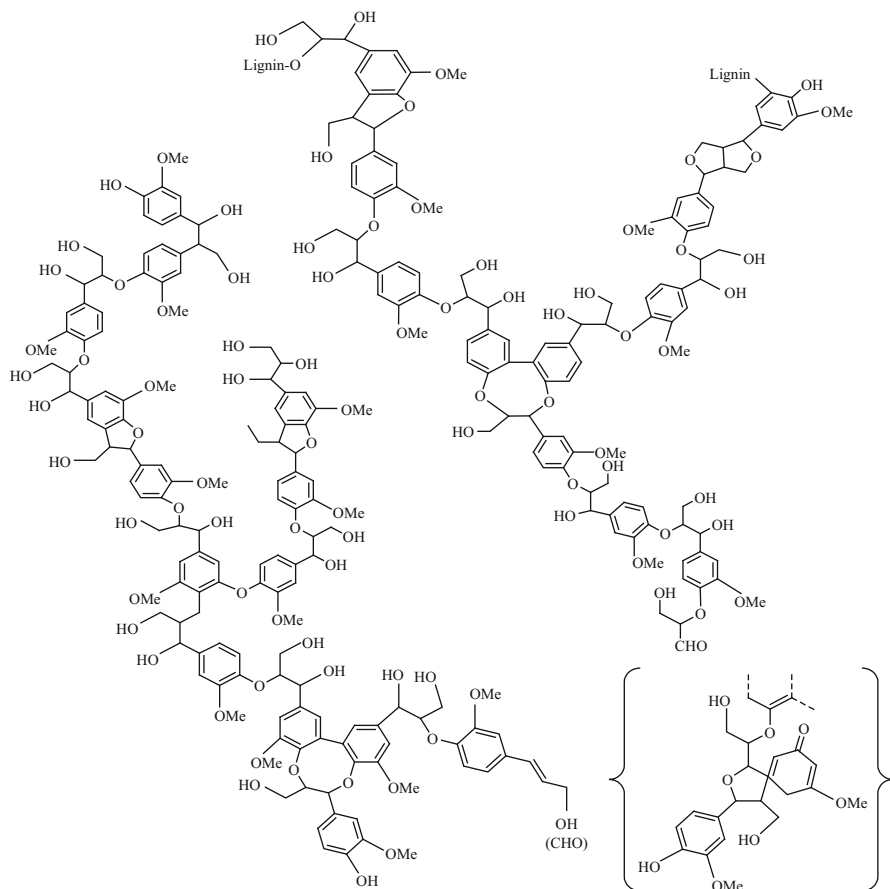
Fig. 2.5 Structural model of a section of cork lignin [34]

The main bond types of grass lignin are the same as those of lignin in wood. The main bond type of the structural unit is aromatic glycerol-β-aryl-ether bonds; these are fewer than hardwood and are similar to softwood. The proportion of carbon–carbon bonds, such as β-5 and β-β in structural units, is higher than in hardwood. In the structure of grass lignin, a considerable part of p-hydroxyphenylpropane units connects with phenylpropane units in its ester form. Taking straw as an example, 60 % p-hydroxyphenylpropane units are connected in the form of its esters. In addition, grass lignin still contains a small amount of ferulic acid esters.

Because the types and positions of functional groups in different types of lignin are different, lignin of gymnosperms has different chemical characteristics. Lignin in gymnosperm mainly contains guaiacyl lignin (G), the G-structure unit of which has a methoxy group, and is hard to remove in the papermaking process because a stable C–C linkage was formed by linkage with other monomers. Monocotyledon mainly contains GSH lignin [37].

2.5.3   Physicochemical Properties of Lignin

2.5.3.1   Chemical Properties of Lignin

The chemical properties of lignin include halogenation, nitration, and oxidation reactions on the phenyl ring; reactions on the benzyl alcohol, the aryl ether bond, and an alkyl ether bond in the side chain; lignin-modified chromogenic reaction; and so on. The chemical reactions of the lignin structural unit are divided into two major categories: nucleophilic reactions and electrophilic reactions.

(1) Chemical reactions of lignin structural unit on the side chain
Reactions on the lignin side chains are associated with pulping and lignin modification; the reaction is a nucleophilic reaction. The following reagents can conduct nucleophilic reactions with lignin:

① In alkaline medium, the effect of HO-, HS-, and S2-nucleophilic reagents leads to the cleavage of the main ether bond (e.g., α-aryl ether bond, a phenol-type α-alkoxy ether bond, and phenol-type β-aryl ether bond) and fragmentation and partial dissolution of macromolecule lignin. In alkaline medium, the phenol-type structural unit is separated into phenolate anions, and an oxygen atom that affects the benzene ring by induction and a conjugative effect, which activates their *ortho-* and *para-* positions and thereby affects the stability of the CO bond and breaks the α-aryl ether bond, then generating a methylene quinone intermediate and resulting in the aromatization of methylene quinone to generate a 1,2-diphenylethene structure.

② In neutral medium, reaction with nucleophile HSO_3^- or SO_3^{2-} leads to breaking of the ether bond and brings SO_3^{2-} groups in the degradation of lignin fragments.

③ Acidic media mainly relate to the lignin fragmentation reaction of the acidic sulfite pulping process. SO_2 aqueous solution is taken as an affinity reagent, leading to the breakage of phenol-type and nonphenolic α-aryl ether bonds, sulfonation of α-carbon, and increased lignin hydrophilicity. Phenol-type and nonphenolic α-alkoxy ether bonds may also have a similar reaction. In addition, C1, C5, and C6 on the high electron density centers of the aromatic ring could also have a condensation reaction with methylene quinone intermediates [10, 36].

(2) Chemical reaction of the aromatic ring in the lignin structure
Chemical reactions of the aromatic ring in the structural unit of lignin are closely related to the lignin-bleaching process and its modification and have been divided into electrophilic and nucleophilic reactions.

① Electrophilic substitution reaction: This mainly refers to substitution and oxidation reactions. Electrophilic reagents include chlorine, chlorine dioxide, oxygen molecule, ozone, nitro cation, nitroso cation, and so on. The electrophilic reagent replacement breaks the side chains of lignin and leads to the oxidative cleavage of β-aryl ether linkages. The aliphatic side chain is oxidized into a carboxylic acid,

and the aromatic ring is oxidized into the compound of the o-quinone structure, which will finally be oxidized into dicarboxylic acid derivatives.

② Nucleophilic reaction: Nucleophilic reagents that can react with the aromatic ring of lignin include hydroxide ions, hypochlorite ions, and hydrogen peroxide ions. These nucleophilic reagents can react with the chromophoric groups in the degraded lignin fragments, breaking the chromophoric structure to some extent [10].

(3) Lignin chromogenic reaction

① The condensation reaction of lignin with concentrated inorganic acids is mainly related to the coniferyl aldehyde structure in the lignin structure.

② Mäule chromogenic reaction: Hardwood lignin becomes reddish violet when firstly treat with $KMnO_4$ and HCl and then ammonia. The syringyl ring generates methoxy o-dihydroxybenzene under the effect of $KMnO_4$ and HCl, then generates a methoxy-o-quinone structure, which is purple after ammonia treatment.

③ Cross-Bevan reaction: Timber without extractives is treated with chlorine in the wet state; lignin will be converted into chlorinated lignin after sulfinic acid and sodium sulfite treatment, turning hardwood lignin red-purple [29].

2.5.3.2 Physical Properties of Lignin

(1) Molecular weight and polydispersity

Any type of separation method may cause some partial degradation and changes in lignin. Accordingly, the molecular weight of the original lignin is unsure. The molecular weight of the separated lignin varies with the separation method and conditions. Its molecular weight distribution can range from several hundred to several million. Under the effect of mechanical action, enzymes, or chemical reagents, the three-dimensional net structure is degraded into different size lignin fragments, which leads to the molecular weight polydispersity of lignin [29]. The molecular weight of separated lignin varies greatly; taking milled wood lignin (MWL) from spruce as an example, depending on the grinding time and extraction methods, its weight average molecular weights include 2,100, 7,100, and 11,000. The weight average molecular weight of the high molecular weight fraction of MWL from spruce and hardwood is 40,000 and 18,000, respectively, and the low molecular weight fraction is between 3,700 and 5,000. The molecular weight of lignosulfonate ranges from 10^3 to 10^5; the maximum is beyond 10^6. Kraft lignin has a much lower molecular weight [10]. Determination of lignin molecular weight includes the osmometric method, light-scattering method, supercentrifugation, gel permeation chromatography, high-pressure liquid chromatography, and so on. As for insoluble lignin, such as lignin obtained by acid hydrolysis, their molecular weights are measured depending on the linear relationship between lgM_w (the base-10 logarithm of molecular weight) and the heat-softening temperature T_s [10].

(2) Solubility

Hydroxyls and many polar groups exist in the lignin structure, resulting in strong intramolecular and intermolecular hydrogen bonds, and making the intrinsic lignin insoluble in any solvent. Condensation or degradation make the separated lignin able to be divided into soluble lignin and insoluble lignin; the former has an amorphous structure, and the latter is the morphological structure of the raw material fibers. The presence of phenolic hydroxyl and carboxyl makes the lignin able to be dissolved in alkaline solution. Separated Brauns lignin and organosolv lignin can be dissolved in dioxane, DMSO, methanol, ethanol, acetone, methyl cellosolve, and pyridine. Alkali lignin and lignosulfonate usually can be dissolved in a dilute alkali, water, and salt solution. Brauns lignin, phenol lignin, and many organosolv lignins can be completely dissolved in dioxane. Acid lignin is not soluble in any solvents. The best solvents for most separated lignin are acetyl bromide and hexafluoroisopropanol in acetic acid [29].

(3) Thermal properties

Lignin is an amorphous thermoplastic polymer. It has slight friability under high temperature and cannot form film in solution. It also has glassy transfer properties. Under the glassy transfer temperature, lignin is in the solid glass phase; it begins to move when it is above the glassy transfer temperature. The lignin is softened to become sticky and has adhesive force. The glassy transfer temperature of separated lignin varies with the raw materials, separation method, molecular weight, and water content. The softening temperature of absolutely dried lignin ranges from 127 to 129 °C, which remarkably decreased with increased water content, indicating that water acts as a plasticizer in lignin. The higher the lignin molecular weight is, the higher the softening point is. For example, for lignins with a M_W of 85,000 and 4,300, the softening points, respectively, are 176 and 127 °C [29].

(4) Relative density

The relative density of lignin is roughly between 1.35 and 1.50. Values vary with the liquid for measurement if measured by water. The relative density of sulfuric acid lignin isolated from pine is 1.451 and is 1.436 if measured by benzene. The relative density of dioxane lignin is 1.33 when measured by water at 20 °C and is 1.391 by dioxane. Lignin prepared by different methods has different relative densities, such as the relative density of pine glycol lignin is 1.362, but it is 1.348 for pine hydrochloride lignin [29].

(5) Color

Intrinsic lignin is a white or nearly colorless substance; the color of lignin we can see is the result of the separation and preparation process. For example, the color of lignin isolated by Brauns and named after him is light cream, and the colors of acid lignin, copper ammonia lignin, and periodate lignin vary from fawn to deep tan [29].
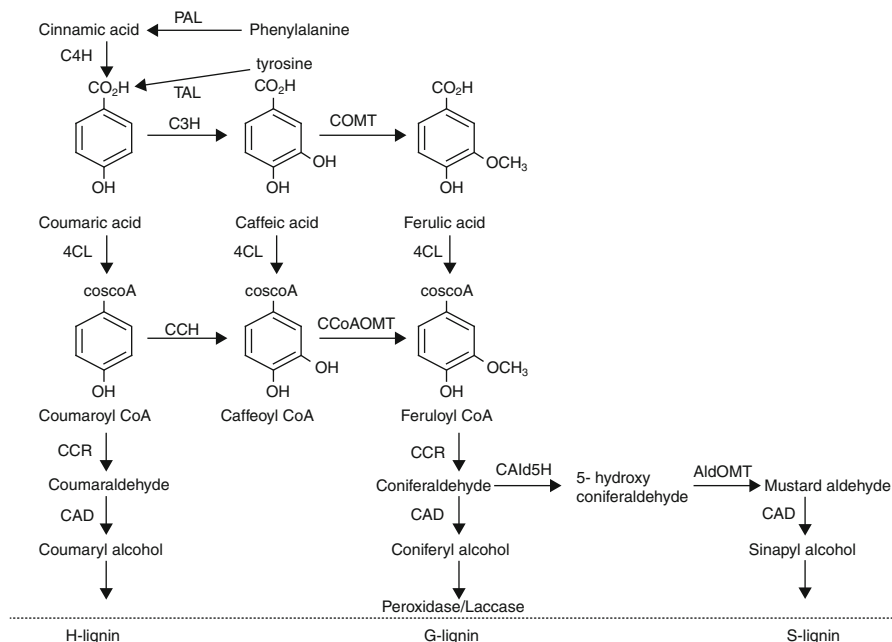
Fig. 2.6 Biosynthetic pathway of lignin monomer [38]. *PAL* phenylalanine ammonia lyase, *TAL* tyrosine deaminase, *Cald5H* coniferaldehyde-5-hydroxylase, *C4H* cinnamic acid-4-hydroxylase, *CCH* coumaroyl CoA-3-hydroxylase, *4CL* 4-coumarate CoA ligase, *CCoAOMT* caffeoyl CoA 3-O-methyltransferase, *C3H* coumarate-3-hydroxylase, *CCR* cinnamoyl CoA reductase, *AldOMT* 5-hydroxy coniferaldehyde-O-methyltransferase, *CAD* cinnamoyl alcohol dehydrogenase

2.5.4 Lignin Synthesis

2.5.4.1 Biosynthesis of Monolignols

Biosynthesis of monolignols is conducted through the phenylpropanoid pathway. In the shikimate pathway, glucose generated from the carbon dioxide by photosynthesis first is converted to the final intermediate: shikimic acid. Then, through the prephenic acid, the shikimic acid is converted into the final products of the shikimate pathway: phenylalanine and tyrosine. These two amino acids, which are widely present in plants, are starting materials for the cinnamic acid pathway. Under the effects of various enzymes, three monomers of lignin are finally synthesized after a set of reactions, such as deamination, hydroxylation, methylation, reduction, and so on (Fig. 2.6) [37, 38].

Cinnamoyl CoA reductase (CCR) catalyzes the first step of the redox reaction of lignin biosynthesis, which may be the rate-limiting step, controlling lignin synthesis and the pathway from which carbon can go into lignin biosynthesis. Cinnamyl alcohol dehydrogenase (CAD) catalyzes the redox reaction of another step in the lignin synthesis process, which may control the reduction of coniferaldehyde. Sinapyl

alcohol dehydrogenase (SAD), ferulic acid 5-hydroxylase (F5H), and bispecific caffeic acid/5-hydroxyferulic acid O-methyltransferase (COMT) are immunolocalized in cells and tissues that have S lignin. CAD is located in tissues with precipitation of G lignin. It is speculated that the last step of redox reactions of different types of lignin may be passed through different synthetic pathways and catalyzed by different enzymes [39].

Two hydroxylases are found in the phenylpropanoid pathway. C3H can catalyze coumaric acid to caffeic acid, and the other hydroxylase is F5H. *Arabidopsis fah1* mutant without F5H activity is almost full of S lignin. Now, it has been recognized that F5H can catalyze the hydroxylation of coniferaldehyde and coniferyl alcohol, converting them to monomers of S lignin.

Lignin monomer synthesis needs to go through two-step methylation reactions at the $3'$ and $5'$ site, respectively. COMT and CCoAOMT (caffeoyl CoA 3-O-methyltransferase) are methylases on two different substrate levels. 5-Hydroxyferulic acid can be converted into sinapic acid under the methylation effect of COMT to involve in the synthesis of S lignin. Much research is more inclined to support that, in lignin biosynthesis, COMT can catalyze the methylation of caffeic acid, 5-hydroxyl coniferyl aldehyde, and 5-hydroxyl coniferyl alcohol into ferulic acid, sinapic acid, and sinapic alcohol, respectively, and CCoAOMT can catalyze the methylation of caffeoyl-CoA to generate feruloyl-CoA.

2.5.4.2 Macromolecular Synthesis of Lignin

It has been proven that the polymerization of lignin macromolecules by phenylpropane units is dehydro-oligomerization. The German scientist Freudenberg first produced synthetic lignin by coniferyl alcohol and laccase collected from mushrooms under aerobic conditions. Later, peroxidase (POD) was found to catalyze this polymerization reaction effectively. But, there is no direct evidence to prove that POD and laccase could catalyze the polymerization of various lignols [38]. Hans et al. discovered that POD and manganese ion could be used to synthesize lignin in vitro, with their structures similar to natural lignin. In this process, redox shuttles of Mn^{3+}-Mn^{2+} played an important role in lignin biosynthesis. POD did not directly react with lignol to participate in the synthesis of lignin. It catalyzed Mn^{2+} oxidized to Mn^{3+}. The latter (Mn^{3+}) may contact many kinds of lignin precursors by diffusion to catalyze the formation of free radicals in terminal groups of lignin. The latter also would participate in the formation of various covalent bonds in lignin molecules [39].

There have been several different hypotheses to explain lignification; one is a random coupling model, which proposed that lignol molecules are gradually connected to the lignin polymer by oxidative coupling, and the formation of lignin is adjusted according to the categories, amount, and chemical coupling characteristics of lignol at the terminal of the lignin molecule. The random coupling hypothesis is reasonable because it explains the plasticity of lignin biosynthesis in mutants and transgenesis research. Another hypothesis involves the dirigent-like protein model,

which considered that the lignified process is strictly under the manipulation of dirigent-like protein by controlling the formation of particular chemical bonds of the lignin molecules. The hypothesis showed that the metabolism of lignin should be an orderly life process, which explains the large numbers of *O*-4 linkages in lignin molecules. Dixon et al. believed that the formation of sinapyl alcohol and coniferyl alcohol was a relatively independent process by which enzymes involved in sinapyl alcohol synthesis were associated with each other to form a multienzyme complex. This multienzyme complex was combined in the endomembrane system through C4H, C3H, C5H, and other P450 proteins, which made a series of reactions associated with mustard alcohol synthesis pathway run in an orderly fashion. The synthesis of coniferyl alcohol is conducted in the cytoplasm another way. The phenylalanine ammonia lyase (PAL), caffeoyl CoA methyltransferase (CCOMT), CCR, and CAD distribute in the cytoplasm, catalyzing the synthesis of coniferyl alcohol [39].

Peroxidase is widespread in plants with multiformity. So, how it participates in lignin monomer polymerization needs further studies. So far, it remains uncertain whether the POD or laccase catalyzes the polymerization of lignin monomers in the plant or if they work synergistically [38].

2.5.4.3   Deposition of Lignin in Plant Tissues

Lignin biosynthesis in vascular plants is controlled by morphogenesis and the environment. The lignification process is the deposition process of lignin in plant tissues. Fluorescence detection indicated that, in the normal development of *Arabidopsis* stems, lignin mainly was deposited in the fiber cell wall between the vessel cell wall and vascular bundles. Dystopia anchorage genes ELP1 were isolated from lignin deposited mutants of *Arabidopsis* pith. In mutant stems, lignin could be deposited in the usual parts and ectopically deposited in the pith of parenchyma cells. Such dystopia anchorage would exist in both young stems and mature stems. Associated with this dystopia anchorage, enzyme activity related to lignin synthesis would increase, and COMT could be ectopically expressed in the central cells of the pith. These results suggest that ELP1 genes inhibit lignin synthesis in the pith. Lignin deposition is related to coniferyl alcohol because that methylated ester generated by coniferyl alcohol can be covalently coupled with hydroxyproline-rich glycoproteins (HRGPs), resulting in the deposition of lignin in the cell wall [40].

2.5.4.4   Factors Influencing Lignin Synthesis

Different parts of plants would have a different lignin content and composition. For example, the lignin content and structure are significantly different in the node and internode of reed (*Arundo donax*); the node has a higher density than the internode because of the high content of phenolic acids (p-coumaric acid and ferulic acid). The lignin content in different parts of lettuce is different depending on the harvest-

ing time. Compared with the normal harvest, lignin in the middle of overmature harvested lettuce buds is increased most obviously, secondarily increased in the basal area, and not significantly increased at the top [40]. The reasons for this difference are varied. For instance, methyl jasmonate can significantly improve the POD activity of seedlings and the lignin content. Both jasmonic acid and gaseous methyl jasmonic acid could induce the expression of chalcone synthase (CHS), thereby increasing the lignin content. In general, overuse of nitrogen fertilizer will reduce the lignin content and thus postpone the lignification of the plant. However, nitrogen nutrition can increase the lignin content in the mulch covering of *Pinus sylvestris* L. and *Picea abies* (L.) Karst. Phosphate fertilizer will greatly limit the period of rapid growth of plants but increase the lignin content [40].

2.5.4.5 Regulation of Lignin Biosynthesis

Lignin biosynthesis may be regulated by changing the activities of different enzymes two main ways. One is to regulate the synthesis process of lignin monomers by reducing enzyme activities that participate in the synthesis process of lignin, such as those of CCR, and CAD, and by reducing the lignin content. The second way is to regulate certain particular enzyme activity to influence the composition and chemical structure of lignin. It is generally believed that the degradation of GS lignin, which is composed of guaiacyl monomers and syringyl monomers, is easier than that for G lignin, which is simply composed of guaiacyl monomers. Therefore, activate the enzyme activities of CAld5H and AldOMT for the coniferaldehyde substrate to compete with CAD, make additional syringyl lignin monomers synthesized via sinapyl aldehyde, and finally relatively reduce the synthesis of guaiacyl G monomer.

In the gene regulation of lignin biosynthesis, the search for and separation of biological enzymes related to lignin synthesis is the first task, and then the amino acid sequence of zymoprotein is analyzed to obtain the sequence of its messenger RNA (mRNA), which is the coding sequence of the functional gene exon. Gene transfer methods include agrobacterium-mediated indirect conversion and the gene gun technique. The antisense technology is also known as antisense DNA and antisense RNA. It can transfer the constructed or amplified antisense DNA or antisense RNA to the living body, then combine the exogenous antisense mRNA or antisense mRNA transcribed from antisense DNA with positive sense mRNA to form the duplexes, which could prevent the transportation of mRNA from the nucleus and combination with ribosome, consequently inhibiting the natural gene expression of endogenous mRNA and achieving the regulation of target genes. Antisense technology can be utilized for regulation of forest lignin. First, an oligonucleotide sequence is constructed that is antisense to the lignin synthase; then, the sequence is transferred into plants in direct or indirect ways, and it interacts with genes in plants to influence the translation and reduce the activity of zymoprotein. This is currently the most common and effective transgenic breeding technology in genetic engineering of lignin regulation. Finding and discovering a natural mutant in plants is also an equally effective and direct means [41].

2.5.5 Biological Function of Lignin

2.5.5.1 The Relationship Between Metabolism of Lignin and Cell Differentiation

Lignin metabolism in plants has physiological significance, which was mainly present as the close relationship between changes of its enzyme activity, the increase of intermediate and lignin contents and cell differentiation, the resistance to pathogen infection, and other physiological activities in plant development [37]. Plant cell differentiation is a core problem in developmental biology. The process of cell differentiation is essentially the changing process of cell physiology and morphology; cell differentiation and organ growth have some relation with phenylpropanoid metabolism and its products. Since the 1970s, there have been many reports about the relationship between lignin metabolism and cell differentiation; Oven found that when the plant was injured, the callus first formed lignified parenchyma cells, followed by formation of short vessels and traumatic resin canals. This process required the participation of the phenylpropanoid metabolites. Mala et al. thought that phenylalanine metabolism had a closer relationship with further development of an embryo.

The structure of cell walls from outside to inside is as follows: ML, primary wall, the first secondary wall (S1), the second secondary wall (S2), and the third secondary wall (S3). The structure of S1 and S2 of the wild-type tobacco cell wall is dense. However, the structure of the anti-CCR transgenic tobacco cell wall varies with types of cells. For example, in fiber cells, the fiber framework of S2 becomes significantly loose, but in vessel cells, the structures of S2 and S3 are changed. However, in the anti-COMT gene tobacco plants, structural changes of the cell wall are not detected. In the hybrid generation of these two transgenic plants, only the S3 structure of the vessel cell wall has slight changes. The immunocytochemistry analysis of the lignin subunit showed that the G and S units of wild-type plants concentrated on S2 in the fiber cells, G and S units of wild-type plants decreased in the hybrid generation, and cell walls became thin. In the vessel cells, G and S units in wild-type plants were unevenly distributed in the S1, S2, and S3, mostly in S2. However, G and S units of hybrid generation were distributed in S1 and S3. Corresponding with this microscopic change, plants with antisense CCR gene grow shorter and have smaller leaves. But, plants with anti-COMT gene have slight changes in height and leaf size, the height and leaf features of hybrid generation are between those of CCR gene plants and anti-COMT gene plants.

Many studies have confirmed that the formation of tracheary elements is accompanied by the synthesis of lignin. Therefore, research on the metabolism physiology of lignin is an important part in research on the differentiation physiology of tracheary elements. The results for the differentiation of wheat root tip (meristematic zone, elongation zone, maturation zone) showed that the tracheary element is mainly formed in the maturation zone. There was only preliminary differentiation and a small amount of lignin deposition in the elongation zone. Lignin is the basic factor

that causes the differentiation of totracheary elements from the apical cell of wood. In the construction of cambium and catheter, the middle structure of a nonlignified cell wall contains an irregular fine network structure, which may contain pectin and hemicellulose. After the lignification, the middle structure quickly becomes dense, fully lignified surface structure becomes more compact, and parts of surface structure are partially covered with a spherical structure [40].

2.5.5.2   Relationship Between Metabolism of Lignin and Fruit Development

Phenolics in the fruit can be mainly divided into water-soluble phenolic substances and water-insoluble phenolic substances. The former includes a variety of phenolic acids, water-soluble tannin, anthocyanin, and so on. They are soluble in both water and ethanol, but after oxidation, they cannot be redissolved in water, only in ethanol. The latter consists of condensed tannin and lignin. Lignin is the main ingredient for the formation of the secondary cell wall. The peel and pith of apple are vigorous metabolic areas of phenolics; anthocyanin is mainly synthesized in the flesh cells of peel and adjacent to peel. The large amount of lignin is in the fruit core, and the phenolics content is the lowest in the flesh.

According to reports, the content of lignin would change with the developmental stage of the apple [42]. Simple phenols and flavonoids (anthocyanin) are synthesized in the young fruit period of apple, and lignin is gradually synthesized with the beginning of the cell enlargement period. Lignin content in the peel is positively correlated with the development time of the fruit and the activity of POD [43]. In contrast, the lignin content in the yellow apple peel is slightly more than that of red apple. In addition, a low temperature has a significant impact on the lignin content in peel [40].

The harvest time also affects the lignin content. Ju et al. [44] found that the lignin content in the late-harvested Laiyang pears was the highest because many phenolics would be converted to lignin during the period of fruit development. Also, the conversion rate in flesh is higher than that in the pith tissue. Bagging also affects the lignin content. PAL is a photo-induced enzyme; its activity is positively correlated with light. However, after bagging the fruit, the activity of PAL decreases, and the lignin content in peel and pulp can both be reduced by 32 %.

2.5.5.3   Relationship Between Metabolism of Lignin and Plant Disease Resistance

Many studies have shown that, when the plant is infected or resistance is induced, the activity of enzymes related to lignin synthesis and the content of lignin would both increase, thereby enhancing the resistance of plants. Research on *Erysiphe graminis* f.sp. *tritici* showed that there was a positive correlation between the resistance of different types of wheat and the increase of lignin content that

depended on the types of wheat. Wheats different in disease resistance all have the PAL activity peak after the inoculation of powdery mildew 12 hours later. The activity peaks of enzymes related to disease resistance are high and have a long duration. The activity peak of susceptible cultivars would decline quickly and become close to the normal level after 24 h. PAL and POD activity increased sharply after bacterial infection in the leaves of mung bean. After inoculating Verticillium dahliae toresistant varieties of cotton in SBSI (Seabrook Sea Island) and susceptible varieties in Rowden, the expression of mRNA, encoding PAL and COMT, was higher than that of the reference substance. After the cotyledon, leaf, and stem of *Arabidopsis thaliana*, *Piplotaxis muralis*, *Piplotaxis tenuifolia*, and *Raphanus raphanistrum* and the cotyledon of *Sisymbrium loeselii* were inoculated with *Leptosphaeria maculans*, they all had strong resistance characteristics to tissue browning and lignin deposition.

In recent years, research on the metabolism of lignin and plant resistance has been plentiful. Pot-planting experiments of three poplar clones from *Aigeiros* sp. under waterlogging conditions showed that, in comparison to their respective controls, the activities of PAL and the contents of organic components in the basal stem phloem of all clones tested had obvious changes. When the waterlogging time was prolonged to 40 days, the metabolism within the plant was confused resulting in the abnormal contents of organic components, the PAL activity increased, and the plant growth decreased significantly. Through research on the overwintering shoots, as the temperature decreased, the PAL activities of the terminal bud, cortex, and xylem steadily increased until December, when it reached peak activity. The PAL activity of different strains in the coldest month was different [45].

In the interaction with phytopathogenic organisms, the lignification of the secondary cell wall is one of the characteristics of the disease resistance response. The lignin, as an important physical antibacterial substance in plants, can be deposited in the HRGP of cell wall as structural barriers, reinforcing the cell wall and protecting cells from infection by phytopathogenic organisms. The infection of ulcer germs can induce accumulation of HRGP and lignin deposition in the cell wall of poplar. The accumulation of different varieties of HRGP and lignin in the poplar cell wall is different and has a positive correlation with disease resistance. Treating tobacco seedlings with jasmonate acid not only can improve the ability of seedlings to be resistant to anthracnose but also can significantly increase seedling POD activity and lignin content. POD does not play a critical role in the disease resistance induced by methyl jasmonate, but lignin is closely related to the disease resistance. In addition, the lignin content in leaves of grape and cucumber has a relationship to disease resistance. Xu et al. reported that, when watermelon was infected by *Fusarium oxysporum*, lignin in seedling leaves and stems/roots, HRGP content, and POP and PAL activity would increase in different degrees with the infection in resistant varieties, and the increase of resistant varieties was significantly higher than for the susceptible varieties.

The biosynthetic pathway of lignin is a branch of the phenylpropanoid material synthesis pathway of a plant. There are reciprocal chiasmata on enzymes involved in these synthesis pathways. Therefore, in those susceptible and under waterlogging,

low-temperature, and other adverse conditions, lignin synthesis is correlated with synthesis of other phenylpropanoid substances. Chen et al. [46] reported that the total phenolic content and the pisatin and phaseoline content in the flavonoids were significantly higher than those in susceptible varieties. Zhou et al. [47] reported that, under noninfective *Verticillium* wilt, PAL activity and the content of phenol and lignin of the rootstock and grafted strains were obviously higher than those of the control strains. After infection with *Verticillium* wilt, the increasing amplitude of PAL activity, phenol content, and lignin content was higher than that of the control group. PAL activity and phenol content have some relevance to disease resistance to *Verticillium* spp. In buckwheat seedlings and different tissues, the flavonoid content always varies with the PAL activity, and this relationship is also reflected in the time sequence. As can be seen, in disease resistance and adversity resistance, nonlignin metabolism has some relevance for lignin metabolism in the phenylalanine metabolism. But, relatively little research currently is available on the degree of relevance in the different plants and the degree of relevance under normal circumstances.

2.5.5.4 Relationship Between Metabolism of Lignin and Plant Growth and Development

The growth and differentiation of plants are closely related to the activities of the cell wall, and lignin is one of the main components of plant cell walls. Since the 1970s, there have been many reports on the relationship between lignin metabolism and plant growth and development. Suzuki and Itoh [48] analyzed the lignifications of bamboo terminal buds and 2- to 14-year-old stem nodes, results of which indicated that the lignification of terminal buds was prior to the extension growth of stems; the stem node growth was closely related to the degree of lignifications. Through the study in the clones of the annual cutting alnus, they found that the sequence of the PAL activity of different clones corresponded to the sequence of the high growth of clones. Throughout the entire growth period, the PAL activity of alnus clones had a significant positive correlation with the high growth.

Lignin, as a secondary metabolite in plant growth and development, has important biological functions in the growth and development and disease resistance of plants. In the cell wall lignification process, lignin penetrates into the cell walls and fills in their framework, which increases the hardness of the cell wall, enhancing the mechanical and the compressive strengths, promoting the formation of mechanical tissues, and consolidating the plant body and water conduction. Because of the chemical characteristics of the lignin, such as the insolubility and complex phenolic polymer, lignin has hydrophobicity of the cell wall and higher resistance to pests. Clarifying the physiological effects of lignin would contribute to better understanding of many biological phenomena of the plant and provide a scientific basis for industrial and agricultural production, but further studies are needed in this field.

2.6 Ash

In plants, except for carbon, hydrogen, oxygen, and other basic elements, there are many kinds of other elements that are indispensable materials in living plant cells, such as nitrogen, sulfur, phosphorus, calcium, magnesium, iron, potassium, sodium, copper, zinc, manganese, chlorine, and more. When the sample is dried at 105 °C and further processed at 750 °C in a high temperature furnace, elements such as carbon, hydrogen, oxygen, nitrogen, sulfur, etc. disappear in the form of gaseous compounds, and the residue is the ash which contains many types of mineral elements in the form of oxide. Different types of plants and growth environment would lead to different types and contents of elements. These elements are absorbed from soil and generally exist as ions. It is well known that different parts (organs) of the plants have different functions in the process of plant growth, causing different distributions of metal ions. Residues from reed and bamboo are mainly parenchyma cells. Table 2.2 shows that, except Mn and Fe, the content of other metal elements in the residues of reed and bamboo are all significantly higher than those of wheat and bagasse. In the process of standard sieve-screening classification of bagasse, with the increase of the mesh number, the particle size becomes finer, the fiber content is reduced, the content of parenchyma cells increases (cells through the 80 mesh are mainly parenchyma cells), and the metal contents rapidly increase. The metal elements and silicon content in the joints, leaves, and ears of wheat straw are significantly higher than in the internode stem [2].

In the combustion process, a small part of the mineral elements may also be lost through gasification. The content and elemental proportion of ash vary with the varieties of raw materials, the kind of plant organs, the age of the plants, and so on. The elements in the growth environment of the plant would also have influence on the content and elemental composition of ash. Generally, wood has a low ash content, usually less than 1 % (the majority have 0.3–0.5 %, dry raw material). The

Table 2.2 Comparison of metal ion contents from several grass families [2]

Sample		Mn	Cu	Fe	Ca	Mg	Al	Na	Si	Co	Ni
Reed	Reed chip	460.2	9.9	481.4	1,569	625	763	134	9,413	<0.1	<0.2
	Reed residue	305.2	102.2	243.9	4,285	1,569	1,628	805	24,285	<0.1	12.1
Bamboo	Bamboo chip	251.3	2.5	36.4	593	440	39.1	21	8,674	<0.1	0.5
	Bamboo residue	379.8	11.1	2,181.9	2,901	1,110	353	98	17,810	0.5	4.7
Wheat	Internode stem	9.1	5.4	64	2,325	1,535	63.3	122	14,005	<0.1	1.78
	Joint	16	4.27	106	3,995	3,937	170.9	294	17,827	<0.1	1.61
	Leaf	40	4.61	329	5,061	4,376	54.1	357	4,032	<0.1	2.35
	Panicle	30	3.0	279	1,421	1,421	662.5	132	19,422	<0.1	1.21
Bagasse	>20 mesh	65	4.4	90	582	1,060	71.9	44	2,644	<0.1	0.56
	20/40 mesh	98	6.3	267	929	1,490	204.4	57	7,366	<0.1	0.65
	40/80 mesh	161	19.1	547	1,124	1,895	360.4	68	11,551	<0.1	0.78
	Over 80 mesh	312	118.1	3,004	3,084	3,075	138	209	4,408	<0.1	9.08

ash content in Gramineae and bark is high, with most having 2–5 % (straw ash content is 10–15 %) [2]. It is worthwhile to note that, compared with other plant materials, SiO_2 accounts for more than 60 % of straw ash, even more than 80 % in rice straw.

Silicon is a necessary nutrient in corn, rice, and other grass families [49, 50] There are several forms of silicon in the soil, but the silicon that can be directly used by higher plants mainly is the monomeric silicic acid [$Si(OH)_4$] in the molecular state. After $Si(OH)_4$ enters plants, the hydroxyl can form the $SiO(OH)$ units through weak intermolecular forces with a variety of the hydrophilic components; then, the units are hydrolyzed to silica sol-gel. Finally, the $SiO_n(OH)_{4-2n}$ spherical nanoparticles are assembled into the silicon in any shape [51, 52]. It is considered that, according to the Si content, the cultivated plants can be divided into three categories. The first is plants that contain a high amount of Si, with the Si content ranging from 5 to 20 %, such as in rice. The second category involves plants that contain a moderate amount of Si, with the Si content ranging from 2 to 4 %, such as in wheat, barley, oats, and the like. The third category has plants with a low amount of Si; the Si content is below 1 %, such as in legumes and dicotyledonous plants [53]. The Si content in plants often is calculated using the percentage (%) of SiO_2 to the plant's dry weight; different plant species have different Si contents. For example, Takahashi et al. analyzed components of 175 plants cultivated on the same soil; the results showed that Si has the largest difference between plant species (the difference of Si content was as high as 200 times between the top 10 high-Si-containing plants and the top 10 low-Si-containing plants) [2].

The presence of silicon in plant materials may influence their subsequent use. For example, a large amount of Na_2SiO_3, generated from silicon of raw material in the alkaline pulping process, dissolved in alkali waste liquid would increase the viscosity of the waste liquid, lower the extraction rate of black liquor in pulp washing, and bring adverse effects in the evaporation, combustion, cauterization, and lime mud recovery processes and so on [2]. Chen et al. reported that the presence of ash in the raw materials would affect enzymatic hydrolysis and fermentation of straw. Therefore, according to the distribution characteristics of SiO_2, Chen and colleagues introduced a fractionated processing technology that included steam explosion, carding, and screening (referred to as steam-exploded fractionation) to realize the high-value utilization of lignocellulose. First, the abundant nanosilica was isolated without destroying the chemical composition of lignocellulose. Then, the lignocellulose that remained was used to prepare ethanol, butanol, and other high-value-added products through biological fermentation [54, 55]. In addition, studies showed that pulps from any source would contain a complex system composed of a variety of metal ions, with part of the metal ions coming from plant materials. In these ions, transition metal ions such as Cu, Fe, or Mn will exert an adverse influence on pulp color; decompose H_2O_2, O_2, O_3, and other bleaching agents; produce some •OH radicals; degrade carbohydrates; and finally reduce the effect of bleaching. Alkaline earth metal ions represented by Ca and Mg could have a stabilization effect on bleaching agents and protect carbohydrates. However, excessive alkaline earth metal ions would stabilize the lignin, thereby reducing

brightness [2]. Therefore, some appropriate pretreatments need to be performed to improve ion distribution in the pretreated materials, further improving the utilization efficiency of the raw material.

2.7   Pectin

Pectin is a major component of the plant cell wall; it is located between the cellulose microfilaments of the cell wall, mainly condensed by galacturonic acid and its methyl ester and composed of rhamnose and arabinose. The ML of plant cells is mainly composed of pectin, which can combine adjacent cells. According to the conditions of combination and physicochemical characteristics, the pectin can be divided into pectic acid, pectin, and protopectin. Pectic acid is a straight chain linked by an α-1,4 – bond by 100 galacturonic acids. Pectic acid is soluble in water and easily acts with calcium to generate the gel of pectate calcium. Pectin is the main component of intercellular substances in the higher plants; it is a long-chain polymer compound linked by α-1,4-glycosidic bonds by galacturonic acid ester and a small amount of galacturonic acid. The molecular weight of pectin ranges from 25,000 to 50,000, with each chain containing more than 200 galacturonic acid residues. Pectin can be dissolved in water and mainly exists in the middle and primary wall; some even is present in the cytoplasm or vacuole [56]. The molecular weight of protopectin is higher than that of pectin acid and pectin. The protopectin mainly exists in the primary wall and is insoluble in water, but it can be converted to soluble pectin under the effects of dilute acid and protopectinase. Because of the formation of calcium bridges between pectin substances, pectin molecules cross-link with each other and form a network structure that plays a bonding role as the middle layer between the cells and allows water molecules to pass through freely. The gel formed by pectin substances has viscosity and elasticity. The increase of the calcium bridge will cause decreased fluidity of the cell wall matrix. Along with the increase of the esterification degree, the formation of the calcium bridge would be reduced, and the elasticity of the cell wall would increase [57].

Pectinase is a complex enzyme system consisting of pectate lyase, pectin esterase, polygalacturonase, and so on. Polygalacturonase is also known as the pectinase and includes exonuclease and endonuclease. These can catalyze the hydrolysis of α-(1,4)-polygalacturonic acid by breaking pectin molecules down into small molecules [58]. A study found that polygalacturonase was related to the maturity of fruits; its main function is to decompose the ML of the cell wall. Many mature fruits accumulated a large number of polygalacturonases. Pectin esterase could deesterify the esterified part of polygalacturonic acids and change them to acidic pectin molecules so that the calcium bridge is formed in the polygalacturonic acid chains, making the pectin firm and strengthening the cell wall structure. Therefore, the esterification and deesterification of the polygalacturonic acid may affect the expansion of the wall. Pectinase not only exists in fruit tissue but also

exists in leaves. It is related to physiological processes such as fruit ripening and softening and leaf abscission. Pectinase is also the earliest enzyme secreted when a fungus invades the plant cell wall, which plays a key role in the utilization of natural fibrous materials by microorganisms [59].

2.8   Pigment

Pigment is also an important substance of lignocellulosic materials. Wood rays exist in the cambium, which is the meristematic tissue of wood existing between timber and bark. The function of cambium is to transport horizontally and store tissues, which makes the wood have various color changes because of the susceptibility to colored bacteria. Up to 1 year of age, tissues near the center of the xylem would be gradually replaced by branches, tannins, pigments, and other substances and then degrade to dead cells with the lost of physiological activity. Many types of heartwood have special pigments; for example, ebony heartwood is black, mahogany is red, and pine is brown [2].

Plant pigment is generally generated from lignin and some lignin degradation products with benzene ring. Lignin contains the chromophore and auxochrome, but the current understanding of the lignin chromophore is not uniform. It is generally believed that the lignin chromophore mainly contains quinone structures and conjugated systems of carbonyls and double bonds. The results of studies showed that the content of the lignin quinoid structure is less than the content of catechol. However, as the optical absorption coefficient of the quinone is higher than that of α-carbonyl (680,000 and 800 $m^2 \cdot kg^{-1}$, respectively), the quinoid structure has great optical absorption; the optical absorption in the visible region accounts for 35–60 % of the total absorption. Therefore, Kringstad et al. reported that the quinoid structure was the most important chromophore in lignin. The formation and degradation of lignin are involved with formations of paraquinones and *O*-quinone. The *O*-quinone is red, and the paraquinones are yellow; they are the major chromogenic structures of lignin. Paraquinone structures are mostly dissolved in the unit form instead of condensed units form except for a few condensed units, but *O*-quinone structures mainly remain in the lignin macromolecules without dissolution [60, 61].

Plants also contain small amounts of flavonoids and anthocyanin. In the transformation of plant materials, these pigment-like substances will be extracted, affecting the further conversion of lignocellulosic materials. In the pulp and papermaking process, the blue stain of pine would exacerbate the wood color and having adverse impacts on wood products and the papermaking process. The presence of pigments would cause a certain degree of inhibition of the properties of enzymolysis and fermentation. Therefore, pretreatments of lignocellulosic feedstock usually need to remove pigments in the hydrolysate; these treatments include macroporous resin adsorption, ion exchange, electrodialysis, and precipitation.

2.9 Cell Wall Protein

In addition to the main components mentioned, the cell wall is enriched with proteins. According to physiological function, the cell wall proteins can be roughly divided into two categories: structural protein, which is the main portion of the cell wall proteins; zymoprotein, including dozens of proteins with various enzyme activities [62].

2.9.1 Structural Proteins

(1) Extensin
Extensin is the most important structural protein in the cell wall and a subtribe of HRGP. It approximately is made up of 300 amino acid residues, with the content of hydroxyproline (Hyp) is especially high, typically 30–40 %. In addition to the Hyp, the extensin contains serine, valine, threonine, histidine, tyrosine, and so on [63]. The main sugar components in extensin are arabinose and galactose, accounting for 26–65 % of the glycoprotein, and sugars connected to the amino acid play an important role in maintaining the conformation of extensin. The extensin also can cross-link with hemicelluloses and pectin, conducive to the formation of the network structure in the cell wall and consequently increasing the rigidity and strength of the cell wall [64].

(2) Proline-rich protein
Proline-rich proteins in plant cell walls contain a repeated proline-proline sequence, which can make proline-rich protein molecules cross-link themselves or with extension and interact with acidic pectin through ionic bonds.

(3) Glycine-rich protein
The characteristics of glycine-rich protein cell wall protein are a high content of glycine (up to 70 %), highly repetitive primary structure units, and a similar protein structure as the structural protein-pectin in animals and cocoon silk protein in the silkworm. Arabinogalactan protein is also found in the plant cell wall, where the main sugars are arabinose and galactose. In the corn and other grass families, threonine- and Hyp-rich glycoproteins and histidine- and Hyp-rich glycoproteins are also found. These two glycoproteins contain threonine and histidine, respectively; the composition and content of the remaining amino acids and sugars are similar to the stretching protein [65].

2.9.2 Zymoprotein

Active proteins in the cell wall mainly refer to various enzymes and lectins. The primary wall of higher plants is enriched with a variety of enzymes that coalesce to the cell wall through ionic, covalent, and hydrogen bonds and so on. To date, dozens

of enzymes have been found in the cell walls, such as cellulase. In the plant growth and development process, the degradation of cell wall polysaccharides is mainly synergistic, completed by exoglucanase, endonuclease, and β-glucosidase [66]. POD is involved in the polymerization and cross-linking of various components in the cell wall and regulates the expansion of the cell wall. Therefore, POD plays an important role in the composition and plasticity of the cell wall [67]. Extensin is a class of cell-wall-specific apoenzyme in the cucumber hypocotyl; it was discovered by Cosgrove in 1989. It can break the hydrogen bonds among the cell wall polymers, inducing the relaxation and acidic growth of the cell wall. At present, more than 100 types of extensins have been separated and identified, and they may be prevalent in the dicotyledon and monocotyledon [68]. Studies have shown that the expression of extensin genes is subject to strict regulation in time and space. Extensin is involved in almost the entire development process of the plant [69]. Enzymatic analysis indicated that extensin has no cellulase activity but has a synergistic effect with cellulase in the enzymatic hydrolysis of microcrystalline cellulose, increasing hydrolysis efficiency of microcrystalline cellulose by 50 %. Extensin also can loosen structures of the filter paper, crystalline cellulose, hemicelluloses, and so on [70].

From the analysis of the plant cell structure mentioned, a large number of structural proteins and zymoproteins were found in the plant cell wall and cell membrane, where enzymes play an important role in cell wall growth, structural changes, and other aspects. Although in the aging stage of plants the protein content would decrease, as the corn stems stay green while seeds mature, the fresh cornstalk after harvest is still rich in protein. Currently, the utilization of straw mainly focuses on the utilization of cellulose, lignin, hemicellulose, and other major components but neglects the application of cell wall proteins. After harvest, straw often is dried and stored under natural conditions, which would significantly reduce its protein content. Some pretreatments can improve the accessibility of the substrate. But, the operating conditions of most pretreatment methods are intense, such as for steam explosion, acid or alkali treatment, or organic solvent treatment, denaturing the cell wall proteins and causing loss of active ingredients.

Han and Chen [71] separated the β-glucosidase from corn straw cell wall protein; the specific activity reached 13.5 IU·mg^{-1}, and the molecular weight was 62.4 kDa, which was similar to the β-glucosidase found in corn by other organizations. ESI-MS (Electron Spray Ionization) analysis showed that the β-glucosidase from corn straw had the highest similarity to the β-glucosidase from soybean (AAD09291). However, their species source and the molecular weights were different, indicating that they were different proteins. The best reaction temperature of β-glucosidase is 37 °C, and the optimum pH is 4.8, which also indicates the important applications in simultaneous saccharification and fermentation of cellulose because simultaneous saccharification and fermentation can react at the optimum temperature of the β-glucosidase and yeast. Studies of cornstalk simultaneous saccharification and fermentation found that cornstalk β-glucosidase was more suitable for ethanol fermentation than *Aspergillus* β-glucosidase. Through ammonium sulfate precipitation, hydrophobic chromatography, and CM-Sephadex G75, β-exoglucanase with a molecular weight of 63.1 kDa can be isolated from fresh corn straw cell wall

proteins. ESI-MS analysis found that this enzyme had the highest similarity with β-(1–3,1–4)-exoglucanase (AAF79936) in maize coleoptiles. The comprehensive comparisons of species, molecular weight, enzyme activity, and other characteristics illustrated that they may be the same protein. The optimum reaction conditions for this enzyme are 40 °C and pH 6.0, and this characteristic enables potential applications in simultaneous saccharification and fermentation of cellulose because it relieves the inconsistent temperature of β-exoglucanase and yeast. β-Exoglucanase can promote the simultaneous saccharification and fermentation and enzymatic hydrolysis, especially simultaneous saccharification and fermentation.

2.9.3   Hydrophobic Protein

Fresh cornstalks contain various proteins; in addition to the structural proteins and zymoproteins, they have hydrophobic proteins. Han and Chen obtained several kinds of hydrophobic proteins from the total proteins of fresh cornstalks; these proteins had strong adsorption of lignin. The molecular weights of these proteins ranged from 15 to 35 kDa. ESI-MS analysis of protein with a molecular weight of 18.5 kDa showed that the protein and corn lipid body-associated protein L2 (P21641) had a high degree of similarity. It can be concluded that they were the same protein by comprehensive analysis of the protein characteristics, species source, and so on. Structural analysis showed that the protein had an amphiphilic structure, 46–122 amino acids consisted of the significant hydrophobic region, while 1–46 amino acids at N-terminal and 122–187 amino acids at C-terminal were hydrophilic. Therefore, it can be inferred that, in the hydrophilic-hydrophobic interface, the protein will self-assemble into amphiphilic film. Hydrophobic protein is without cellulase activity. It can be adsorbed onto lignin and steam-exploded straw but cannot be adsorbed onto microcrystalline cellulose. It has a synergistic effect with cellulase in the enzymatic hydrolysis of steam-exploded straw but has no effect on the enzymatic hydrolysis of microcrystalline cellulose. Principle analysis showed that hydrophobic protein may play a role in the enzymatic hydrolysis of steam-exploded straw through the combination with lignin, reducing the nonspecific adsorption of cellulase on the substrate, particularly exonuclease and beta-glucosidase.

In view of the plentiful proteins in plants, as well as the possible synergy between the plant zymoprotein and exogenous cellulase in the enzymatic hydrolysis of lignocellulose, Chen et al. studied the application of vegetable protein in cellulose degradation for several years and had some success. Di Lu performed some research on the change of content of the apoplast protein of corn straw during storage and the impact on *Penicillium expansum* cellulase activity; it was found that the apoplast protein and *Penicillium expansum* cellulase had obvious synergies, with the new harvest straw apoplast protein having the most obvious synergies [72]. Han and Chen [71, 73] found that in vitro autolysis producing glucose as the main product also happened in the fresh corn straw; autolysis substrates accounted for 2.56 % of the straw dry weight; these autolysis substrates were mainly cellulose and the

β-(1–3, 1–4) glucan of hemicellulose. According to this characteristic, he developed the new technology of straw enzymolysis and fermentation based on the use of cell wall proteins in fresh corn straw; glucose and ethanol production increased by 61.2 and 112.7 %, respectively, compared with the control. Fresh cornstalks were used as the substrate for enzymatic hydrolysis and fermentation; the cell wall proteins also played an additive effect on *Trichoderma viride* cellulase.

With further research and development of the cell wall proteins in the future, the application of the plant cell wall proteins in straw biotransformation will increasingly be broadened, achieving prominent and significant results.

References

1. Chen HZ. Ecological high value-added theory and application of crop straws. Beijing: Chemical Industry Press; 2006.
2. Yang SH. Plant fiber chemistry. Beijing: China Light Industry Press; 2008.
3. Zhang YZ, Liu J, Gao P. Scanning tunneling microscopy of the ultrastructure of native cellulose. Acta Biophys Sin. 1997;13(3):375–9.
4. Chen HQ, Gong Y, Fang Z. The situation and aspect of application for macromolecule (I)-cellulose, lignin and starch. Yunnan Chem Technol. 1996;11(1):41–6.
5. Li XB, Wu Q. Plant cell wall. Beijing: Peking University Press; 1993.
6. Song DL, Shen JH, Li LG. Cellulose synthesis in the cell walls of higher plants. Plant Physiol J. 2008;44(4):791–7.
7. Zhang JQ, Lin L, Sun Y, Mitchell G, Liu SJ. Advance of studies on structure and decrystallization of cellulose. Chem Ind For Prod. 2008;28(6):109–14.
8. Zugenmaier P. Conformation and packing of various crystalline cellulose fibers. Prog Polym Sci. 2001;26(9):1341–417.
9. Zhan HY. Fiber chemistry and physics. Beijing: Science Press; 2005.
10. Gao J, Tang LG. Cellulose science. Beijing: Science Press; 1996.
11. Tai FJ, Li XB. Cellulose biosynthesis in plant and the enzymes involved in it. Chin J Cell Biol. 2004;26(5):490–4.
12. Mueller SC, Brown Jr RM, Scott TK. Cellulosic microfibrils: nascent stages of synthesis in a higher plant cell. Science. 1976;194(4268):949–51.
13. Reiter WD. Biosynthesis and properties of the plant cell wall. Curr Opin Plant Biol. 2002;5(6):536–42.
14. Arioli T, Peng L, Betzner AS, Burn J, Wittke W, Herth W, Camilleri C, Höfte H, Plazinski J, Birch R. Molecular analysis of cellulose biosynthesis in *Arabidopsis*. Science. 1998;279(5351):717–20.
15. Kimura S, Laosinchai W, Itoh T, Cui X, Linder CR, Brown RM. Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant *Vigna angularis*. Plant Cell Online. 1999;11(11):2075–85.
16. Wang YJ. Callose in plants. Bull Biol. 2005;40(1):18–9.
17. Bian HY, Zhou ZG, Chen BL, Jiang GH. Biological synthesis of cellulose during cotton fiber thickening process. Cotton Sci. 2004;16(6):374–8.
18. Zhong R, Burk DH, Ye ZH. Fibers. A model for studying cell differentiation, cell elongation, and cell wall biosynthesis. Plant Physiol. 2001;126(2):477–9.
19. Kimura S, Kondo T. Recent progress in cellulose biosynthesis. J Plant Res. 2002;115:297–302.
20. Schrick K, Fujioka S, Takatsuto S, Stierhof YD, Stransky H, Yoshida S, Jürgens G. A link between sterol biosynthesis, the cell wall, and cellulose in *Arabidopsis*. Plant J. 2004;38(2):227–43.

21. Yan SP, Wang QY, Yang CP. Research advances in the plant cellulose biosynthesis. J Anhui Agric Sci. 2008;36(21):9049–51.
22. Pear JR, Kawagoe Y, Schreckengost WE, Delmer DP, Stalker DM. Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase. Proc Natl Acad Sci U S A. 1996;93(22):12637–42.
23. Doblin MS, Kurek I, Jacob-Wilk D, Delmer DP. Cellulose biosynthesis in plants: from genes to rosettes. Plant Cell Physiol. 2002;43(12):1407–20.
24. Richmond TA, Somerville CR. The cellulose synthase superfamily. Plant Physiol. 2000;124(2):495–8.
25. Konishi T, Ohmiya Y, Hayashi T. Evidence that sucrose loaded into the phloem of a poplar leaf is used directly by sucrose synthase associated with various β-glucan synthases in the stem. Plant Physiol. 2004;134(3):1146–52.
26. Zhang P, Hu HR, Shi SL. Application of hemicellulose. Tianjin Pap Mak. 2006;2:16–8.
27. Yin ZF, Fan RW. The research progress of plant cell wall. Bull Bot Res. 1999;19(4):407–14.
28. Xu F, Sun RC, Zhan HY. Progress in non-wood hemicellulose research. Trans Chin Pul Pap. 2003;18(1):145–52.
29. Jiang TD. Lignin. Beijing: Chemical Industry Press; 2001.
30. Wei JH, Song YR. Recent advances in study of lignin biosynthesis and manipulation. J Integr Plant Biol. 2001;43(8):771–9.
31. Lv WJ, Xue CY, Cao CY, Zhang Y. Lignin distribution in wood cell wall and its testing methods. J Beijing Univ. 2010;32(1):136–41.
32. Xu F, Zhong XC, Sun RC, Jones GLL. Lignin distribution and ultrastructure of *Salix psammophila*. Trans Chin Pul Pap. 2005;20(1):6–9.
33. Wu J, Fukazawa K, Ohtani J. Distribution of syringyl and guaiacyl lignins in hardwoods in relation to habitat and porosity form in wood. Holzforschung-Int J Biol Chem Phys Technol Wood. 1992;46(3):181–6.
34. Qiu WH, Chen HZ. Structure, function and higher value application of lignin. J Cellul Sci Technol. 2006;14(1):52–9.
35. Zheng DF, Qiu XQ, Lou HM. The structure of lignin and its chemical modification. Fine Chem. 2005;22(4):249–52.
36. Tao YZ, Guan YT. Study of chemical composition of lignin and its application. J Cellul Sci Technol. 2003;11(1):42–55.
37. Li W, Xiong J, Chen XY. Advances in the research of physiological significances and genetic regulation of lignin metabolism. Acta Bot Boreali-Occidentalia Sin. 2003;23(4):675–81.
38. Geng S, Xu CS, Li YC. Advance in biosynthesis of lignin and its regulation. Acta Bot Boreali-Occidentalia Sin. 2003;23(1):171–81.
39. Lin ZB, Ma QH, Xu YY. Lignin biosynthesis and its molecular regulation. Prog Nat Sci. 2003;13(5):455–61.
40. Yu MG, Yang HQ, Zhai H. Lignin and physiological function in plant. J Shandong Agric Univ. 2003;34(1):124–8.
41. Chen Y, Tan X, Clapham D. Lignin biosynthesis and genetic regulation. Acta Agric Univ Jiangxiensis. 2003;25(4):613–17.
42. Yu M. Lignin metabolism and its regulation in apple rootstock M. Hupehensis Rehd [dissertation]. Tai'an: Shandong Agricultural University; 2002.
43. Guo XF, Zhang YL, Liu H. Yearly changes of phenol content in *Danxia* apple tree. J Fruit Sci. 2004;21(6):606–8.
44. Ju ZG, Yan SP. Laiyang pear phenolic substances synthetic regulation and its effect on the quality of fruit. Sci Agric Sin. 1993;26(4):44–8.
45. Cheng SW, Tang LZ, Xiao Y, Xu X. Pal activity and organic components in basal stem phloem of poplar clones under waterlogging and flooding condition. J Nanjing For Univ. 1997;21(1):51–5.
46. Chen HM, Liu JY, Ran B, Zhou J, Li T. Dynamics of some related enzymes of tobacco infected with brown spot. J Yunnan Agric Univ. 1995;10(1):1–6.
47. Zhou BL, Lin GR, Gao YX. The resistance of grafted eggplant to vertillium wilt and its function. J Shenyang Agric Univ. 2000;31(1):57–60.

48. Suzuki K, Itoh T. The changes in cell wall architecture during lignification of bamboo, *Phyllostachys aurea* Carr. Trees-Struct Func. 2001;15(3):137–47.
49. Liu WG, Wang LQ, Bai YH. Research progress in the beneficial elements-silicon for plants. Acta Bot Boreali-Occidentalia Sin. 2003;23(12):2248–53.
50. Xu CX, Liu ZP, Liu YL. The physiological function of silicon in plants. Plant Physiol Commun. 2004;40(6):753–7.
51. Ma JF, Yamaji N. Silicon uptake and accumulation in higher plants. Trends Plant Sci. 2006;11(8):392–7.
52. Wang LJ, Wang YH. Nanostructures SiO_2 in plant body. Chin Sci Bull. 2001;46(8):625–32.
53. Ji XE, Zhang MS, Yu HQ, Bai BZ. Silicon nutrient of plants. Agric Technol. 1998;18(2):11–3.
54. Yu B, Chen HZ. Effects of steam-exploded fractionation on the structure and distribution of silicon dioxide in corn straw. Trans Chin Soc Agric Eng. 2008;24(10):190–4.
55. Chen HZ, Yu B. A preparation method of nano-silica by straw. China patent 200710062669. 2007.
56. Du JY, Bai L, Bai B. Chemical composition and basic characteristics of pectin. Agric Technol. 2002;22(5):72–6.
57. Hardell H, Leary G, Stoll M, Westermark U. Variations in lignin structure in defined morphological parts of birch [*Betula verrucosa*, middle lamella, primary wall, secondary wall, ray cells, vessels]. Svensk papperstidning. 1980;83(2):44–9.
58. Lu SM, Xi Y, Jin YF, Zhang Y. Structure and function of plant polygalacturonases. Acta Hortic Sin. 1999;26(6):369–75.
59. Xue CH, Zhang YQ, Li ZJ, Li ZJ. Recent development of pectin and pectolytic enzyme. J Food Sci Biotechnol. 2005;24(6):94–9.
60. Xi W, Li XP. An investigation to reaction properties of hydrogen peroxide with β-O-4 lignin quinoid chromophoric group. China Pulp Pap Ind. 2008;29(16):32–5.
61. Li XP, Wu S. Research development of the reaction characteristic between lignin quinonoid chromophoric group and hydrogen peroxide. J Cellul Sci Technol. 2006;14(4):52–6.
62. Han YJ, Chen HZ. Plant cell wall proteins & enzymatic hydrolysis of lignocellulose. Prog Chem. 2007;19(7/8):1153–8.
63. Li LC, Wang X, Jing JH. The existence of expansion and its properties in the hypocotyls of soybean seedlings. Acta Bot Sin. 1998;40(7):627–34.
64. Lee SJ, Saravanan RS, Damasceno C, Yamane H, Kim BD, Rose JKC. Digging deeper into the plant cell wall proteome. Plant Physiol Biochem. 2004;42(12):979–88.
65. Li XB, Yang ZH. Structure, function, crossing linking and biosythesis of extensions. Plant Physiol Commun. 1990;3:7–13.
66. Sandgren M, Ståhlberg J, Mitchinson C. Structural and biochemical studies of GH family 12 cellulases: improved thermal stability, and ligand complexes. Prog Biophys Mol Biol. 2005;89(3):246–91.
67. Jamet E, Canut H, Boudart G, Pont-Lezica RF. Cell wall proteins: a new insight through proteomics. Trends Plant Sci. 2006;11(1):33–9.
68. Tong B, Rao JP, Ren XL, Li JR. Studying progress of plant cell wall proteins expansions. Chin Agric Sci Bull. 2005;21(9):112–15.
69. Shcherban TY, Shi J, Durachko DM, Guiltinan MJ, McQueen-Mason SJ, Shieh M, Cosgrove DJ. Molecular cloning and sequence analysis of expansins—a highly conserved, multigene family of proteins that mediate cell wall extension in plants. Proc Natl Acad Sci U S A. 1995;92(20):9245–9.
70. Whitney SEC, Gidley MJ, McQueen-Mason SJ. Probing expansin action using cellulose/hemicellulose composites. Plant J. 2001;22(4):327–34.
71. Han YJ, Chen HZ. Characterization of β-glucosidase from corn stover and its application in simultaneous saccharification and fermentation. Bioresour Technol. 2008;99(14):6081–7.
72. Lu D, Chen HZ, Ma RY. Effect of straw apoplast protein on cellulase activity. Chin J Biotechnol. 2006;22(2):257–62.
73. Han YJ, Chen HZ. Synergism between corn stover protein and cellulase. Enzyme Microb Technol. 2007;41(5):638–45.