Chapter 8
Properties of Plant Proteins

Muhammad Arif and K. Peter Pauls

Chapter Highlights

- Proteins are linear assemblies of different quantities and sequences of amino acids.
- They have structural roles as building blocks of organisms, regulatory roles as transcription factors or metabolic roles as catalysts of biochemical reactions (enzymes) in the cell.
- The functional properties of a protein are defined by the physical and chemical properties of the protein.
- A number of crops including cereals grains, legumes and pulses are rich in plant proteins that can be used directly for manufacturing protein-based bioproducts.
- These plant proteins have a wide range of structures and functions, making them suitable for the manufacturing of a broad variety of bioproducts.

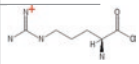
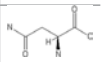
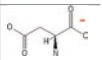
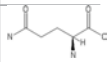
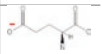
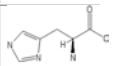
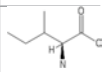
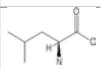
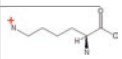
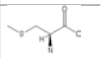
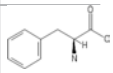
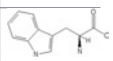
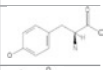
8.1 Introduction

Proteins are linear polymers of amino acids that have a great variety of functional, structural and regulatory roles in organisms. The word "protein" comes from the Greek word "prota", meaning "of primary importance" (http://www.peptideguide.com/proteins.html). Most proteins are assemblages of 20 possible standard amino acids that are directly coded for by DNA codons (Table 8.1, Fig. 2.12) and some nonstandard amino acids in some organisms. A typical amino acid contains an amino group (NH_2 –), a carboxylic group (– CO_2H) and a side chain R group (neutral, acidic or basic) that are all covalently attached to the α-carbon (Fig. 2.11).

M. Arif · K. P. Pauls (✉)
Department of Plant Agriculture, University of Guelph, Guelph, ON, Canada
e-mail: ppauls@uoguelph.ca

G. Chen et al. (eds.), *Plant Bioproducts*,
https://doi.org/10.1007/978-1-4939-8616-3_8

Table 8.1 Primary amino acids: their chemical structures and properties

Amino acid	Symbol		Side-group charge	Side-group polarity	Structure
	Three letter	One letter			
Alanine	Ala	A	Neutral	Non-polar	
Arginine	Arg	R	Positive	Polar	
Asparagine	Asn	N	Neutral	Polar	
Aspartic acid	Asp	D	Negative	Polar	
Cysteine	Cyc	C	Neutral	Non-polar	
Glutamine	Gln	Q	Neutral	Polar	
Glutamic acid	Glu	E	Negative	Polar	
Glycine	Gly	G	Neutral	Non-polar	
Histidine	His	H	Positive (10%) Neutral (90%)	Polar	
Isoleucine	Ile	I	Neutral	Non-polar	
Leucine	Leu	L	Neutral	Non-polar	
Lysine	Lys	K	Positive	Polar	
Methionine	Met	M	Neutral	Non-polar	
Phenylalanine	Phe	F	Neutral	Non-polar	
Proline	Pro	P	Neutral	Non-polar	
Serine	Ser	S	Neutral	Polar	
Threonine	Thr	T	Neutral	Polar	
Tryptophan	Trp	W	Neutral	Non-polar	
Tyrosine	Tyr	Y	Neutral	Polar	
Valine	Val	V	Neutral	Non-polar	

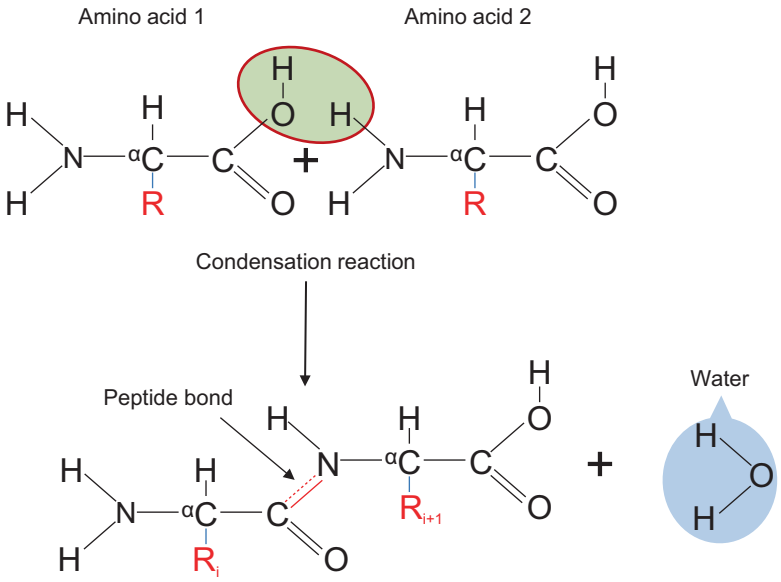


Fig. 8.1 Schematic representation of a peptide (amide) bond in a dipeptide molecule

The side chain R group contains acid, amine, amide, alcohol, alkyl or benzyl groups that determine whether an amino acid is acidic, basic or neutral.

Amino acid monomers can be assembled into a limitless number of sequences and can exhibit a wide range of interactions and chemical reactions (McMurry 1994; Stevens 1999; Pommet et al. 2003). When two amino acids are joined together during protein synthesis on a ribosome of a cell, a partial double covalent peptide (amide) bond (Fig. 2.14) is established between the carboxylic group of the nascent polypeptide or protein and the amino group of the additional amino acid, resulting in the release of a water molecule (Fig. 8.1). Typical proteins are polymers of approximately 350 amino acid residues; therefore, proteins are also called polypeptide chains. The term amino acid residue typically refers to amino acids which are bonded into a polypeptide, or a shorter segment, referred to as a peptide. The ends of polypeptide chains contain free amino and carboxylic groups, referred to as the **N-terminus** and the **C-terminus**, respectively. The ionized N-terminal amino group, C-terminal carboxyl group and groups on side chains determine the net charge, polarity and functionality of the protein polymers.

After its synthesis, the primary polypeptide chain undergoes several chemical and structural modifications that serve to protect the newly synthesized protein from breakdown by the cell's own proteolytic enzymes and functionalize it for various cellular roles. The range of modifications includes interchain folding; phosphorylation; conjugation with sugars, lipids and metals; as well as oligomerization. Proteins have structural roles as building blocks of organisms, regulatory roles as transcription factors or metabolic roles as catalysts of biochemical reactions (enzymes) in the cell.

Functional or structural proteins can be pure or conjugated or globular or filamentous, depending on their chemical and physical nature (Damodaran 1996). This chapter builds upon the information on amino acids and proteins presented in Chap. 2 extending the discussion into the realm of plant proteins, including the major sources and methods of isolation and processing. A considerable amount of the discussion that follows is based on information that can be found in a textbooks used for senior undergraduate courses in biochemistry (e.g. Heldt and Piechulla 2010; Moran et al. 2012; Nelson and Cox 2017).

8.2 Structural and Functional Properties of Proteins

The functional properties of a protein are defined by all the physical and chemical properties of the protein, which affect its behaviour during its interaction with other molecules, processing, storage, isolation, utilization and degradation. Protein structure and function depend on the type, number, order, orientation of its amino acids and interactions among them. **Protein structure** is determined by the basic structure of the peptide bond, the sequence of its amino acids, folding of the primary chain and interactions, as well as intra-protein and inter-protein cross-links among its constituent amino acid side chains (Krochta 2002). The primary polypeptide chain is organized into secondary, tertiary and quaternary structures (Fig. 2.15), based on the kind, number and sequence order of the amino acid residues and their cross-linking interactions within the chain. Hydrogen bonding, van der Waals forces and electrostatic, hydrophobic and covalent disulphide cross-linking bonds organize proteins into their final structures (Krochta 2002).

Discussions about polypeptides or proteins often refer to the terms molecular weight (MW) or relative molecular mass (M_r) which is the mass of the molecule relative to one twelfth the mass of an atom of carbon-12. MW or M_r is dimensionless and has no units. The molecular mass (i.e. absolute molecular mass, m) has the same magnitude as MW (or M_r) but instead is presented in daltons (Da) where 1 Da is an atomic mass unit.

8.2.1 Primary Structure

A protein's primary structure refers to the linear sequence of amino acid residues in the polypeptide chain. The partial double bond nature of the peptide bond, which is the result of the resonance structure created by the delocalization of the lone pair of electrons on the nitrogen, creates a slight positive charge on the amino group (N – Hδ+) and restricts the rotation of the CO–NH bond to a maximum torsion angle (ω) of 6°. This rigidity of the peptide bond keeps its six atoms in a single plane with mostly (1000:1) trans configurations, which reduces steric interactions of the

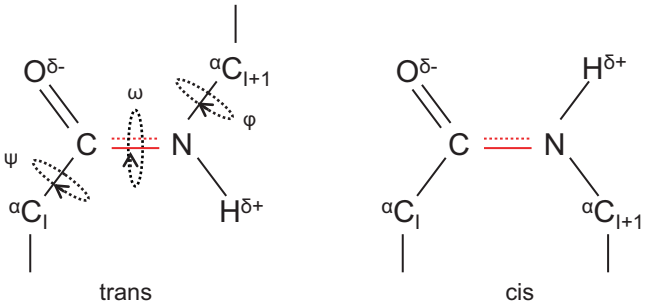


Fig. 8.2 Schematic representation of *tran-cis* orientation of amino acids in polypeptide chain. (Source: Damodaran (2008), Nelson and Cox (2017), Voet et al. (2016))

α-carbons, except for X-proline peptide groups, which have 3:1 (*trans:cis*) ratios. Rotation about the other bonds in the primary chain, namely, between N–Cα (Φ) and Cα–CO (Ψ), occurs freely but is affected by interactions among R groups (Fig. 8.2) (Damodaran 2008).

8.2.2 Secondary Structure

After synthesis, the majority of proteins fold into α-helices and β-sheets, which make up their secondary structures. **α-Helices** are right-handed spirals stabilized by hydrogen bonds between carbonyl oxygens (at positions n) and amide hydrogens of the fourth amino acid (at positions n + 4) towards the C-terminus. Every turn in the α-helix contains 3.6 amino acid residues. Within the α-helix, the side chains of the amino acid residues project outward from the axis of the helix, and its stability is affected by the properties of these side chains. For example, an alanine residue, which has a small uncharged side chain, occurs often in α-helices, whereas tyrosine and asparagine residues, which have large side chains, are less commonly found. Proline residues cannot make up α-helices because their rigid cyclic side chains disrupt the helix conformation and because they do not have hydrogen atoms on their amide nitrogens to form the hydrogen bonds that stabilize helices (Damodaran 1996).

In **β-sheets**, the primary amino acid chain is folded back on itself so that interactions can occur at the sides of the chain for some length to create a planar surface. Hydrogen bonds occur between the carbonyl oxygens of one strand and the amide hydrogens of the next strand, which lie in the sheet in roughly perpendicular positions relative to the long axis of the sheet. Side chains project above and below the plane of the sheet. β-Sheets with primary strands running antiparallel (running in the same N- to C-terminal direction) and parallel (running in opposite N- to C-terminal direction) exist, but the former structure is more stable.

8.2.3 Tertiary Structure

To optimize the intra-chain attractive forces (van der Waals, hydrophobic and electrostatic) and to reduce the free energy and interfacial area, protein secondary structures fold into more compact tertiary structures. The process of folding is complex and often brings hydrophobic side chains, such as those of phenylalanine, alanine, valine, leucine, isoleucine and methionine, into the interior of the molecule, where they are shielded from contact with water, and positions hydrophilic side chains, such as those of lysine, arginine, serine, glutamic acid, histidine and glutamine, to the outside in contact with water to create a hydrate coat. Electrostatic interactions between oppositely charged amino acids residues (such as the amino of side chain of lysine and the carboxylic group of glutamic acid) within close proximity in the three-dimensional space create salt bridges that stabilize tertiary structures. Oxidation of the thiol groups of two cysteine residues (Cys–SH) to form intra-strand disulphide bonds (Cys–S–S–Cys) is another important mechanism for stabilizing tertiary structures.

8.2.4 Quaternary Structure

Some proteins are actually aggregates of monomers arranged into a quaternary structure. These structures are stabilized by the same non-covalent interactions and disulphide bonds as the tertiary structures. For example, most of the plant seed storage proteins, including soy protein (soybean), phaseolin (bean), gluten (wheat) and zein (maize), are oligomers.

8.2.5 Protein Modifications

Post-translational modifications give proteins physical, chemical and functional properties that are used to classify them as glycoproteins (conjugated with sugar), phosphoproteins (phosphorylated), lipoproteins (conjugated with lipids) or metalloproteins (conjugated with metals). These modifications affect the functional and structural properties of proteins as well as their localization, stability and interactions with other molecules.

Glycosylation is the addition of oligosaccharide chains called glycans to proteins during or after their synthesis in the endoplasmic reticulum (ER) and/or Golgi apparatus (GA) of the cell. The oligosaccharide side chain is attached to the amide nitrogen of asparagine (Asn) residues (N-glycosylation) or to the hydroxyl group of serine (Ser), threonine (Thr) or hydroxyproline (Hyp) residues (O-glycosylation) in the protein backbone. N-glycosylation occurs during or after translation in the ER, with the transfer of a preformed lipid-linked oligosaccharide onto the polypeptide.

In contrast, O-glycosylation occurs by the transfer of individual saccharides onto folded proteins in the ER and/or the GA.

Glycosylation represents the most widespread post-translational modification found in biopharmaceutical proteins. Approximately 50% of human proteins are glycosylated, and the functional activities of therapeutic glycoproteins are frequently dependent on the presence and composition of their glycans, since they can affect their plasma half-life and tissue targeting. For foreign proteins synthesized in plant platforms, the glycosylation patterns are of particular importance because plant glycosylation patterns differ in several ways from mammalian N- and O-glycosylation patterns. For example, because plants lack N-acetylglucosaminyltransferase-III, N-acetylglucosaminyltransferase-IV and N-acetylglucosaminyltransferase-V enzymes, which produce branching in glycans, plant N-glycans are bi-antennary structures, instead of the multi-antennary structures found in animal glycoproteins. Also, plant N-glycans contain α (1, 3) fucose and/or β (1, 2) xylose linked to the core $Man_3GIucNAc_2$-Asn of glycans, in contrast to the occurrence of β (1, 4)galactose in mammalian proteins. Furthermore, most oligosaccharides of human glycoproteins are capped by the addition of sialic acid on a penultimate galactose residue, but the pathway for the addition of sialic acids is missing from plants. The absence of the sialic acid modification of galactose causes rapid removal of the protein from circulation through uptake by hepatic galactose-specific receptors. Additional protein engineering is required to produce mammalian compatible proteins in plant-based systems (Webster and Thomas 2012).

8.2.6 Protein Targeting

Higher plants have evolved the ability to accumulate large amounts of proteins in stable forms in their storage organs such as seeds, tubers and roots. The proteins are deposited in specialized protein storage vacuoles or protein bodies in the cell (Herman and Larkins 1999). The storage proteins are transported into the ER during their synthesis. Storage **albumins** and **globulins**, found in most land plants, are trafficked through the normal secretory pathway that includes the GA and accumulate in storage vacuoles of cotyledonary cells, especially in legumes (Otegui et al. 2006). In contrast, **prolamins**, which are found in cereals, form large insoluble polymers within the ER that do not proceed along the secretory pathway and result in the formation of protein bodies in endosperm cells (Herman 2008).

8.3 Physicochemical Properties of Proteins

Knowledge of the physicochemical properties of proteins, including their amino acid composition, structure, net charge, charge distribution, hydrophobicity/hydrophilicity ratio, molecular flexibility/rigidity, MW and ability to interact with other

components, is useful to evaluate their potential utility for manufacturing bioproducts such as adhesives, glues, films, fibres and pharmaceuticals. Collectively, the physicochemical properties of proteins affect their functional properties, such as solubility, hydration, oil miscibility, aroma trapping, viscosity, gelation, elasticity, emulsification and foaming characteristics, which are important parameters for protein-based products. These properties can be grouped into properties related to hydration mechanisms, protein structure and rheology and protein surfaces (Moure et al. 2006).

8.3.1 Protein Hydration Properties

The **hydration properties** of proteins are the result of their amino acid compositions and in particular the ratio of polar versus non-polar and ionic versus neutral amino acids (Table 8.1). Water is a bipolar molecule and interacts easily with other polar molecules, including polar amino acids. Proteins with high polar amino acid fractions have high water binding and holding capacities (Kuntz and Kauzmann 1974).

The locations of the **hydrophobic (non-polar)** and **hydrophilic (polar) amino acids** in protein molecules also influence their water retention properties (Damodaran 1996). To perform their different functions, proteins are usually folded in a way that exposes their hydrophilic amino acids on their surfaces and places their hydrophobic amino acid residues in the centers of their structures (Mierovich and Scheraga 1980). The ratio of hydrophobic to hydrophilic amino acid residues on the surface of the protein molecule and the charge frequency determine the solubility of the protein in various solvents. Higher ratios of surface hydrophobic amino acids increase protein-protein interactions, resulting in lower protein solubilities, while charged amino acids promote protein-solvent interactions, resulting in increased protein solubilities (Damodaran 1996; Moure et al. 2006). Protein concentration, pH, temperature, ionic strength, type of solute ion and atmospheric pressure are the external factors that affect the water-holding capacities and solubilities of proteins (Moure et al. 2006).

8.3.2 Protein Rheology Properties

Viscosity and gelation are the two most studied rheological properties of proteins. **Viscosity** is the resistance to flow of a liquid due to internal friction. A **gel** is an intermediate state between solid and liquid. The viscosity and gelation properties of proteins depend on protein-protein and protein-solvent interactions. These depend on their molecular properties, including MW, molecular size, shape, flexibility and hydration (Damodaran 2008). Proteins containing polar amino acid residues form transparent gels and proteins containing non-polar amino acid residues form opaque gels (Shimada and Matsushita 1980; Totosaus et al. 2002; Moure et al. 2006).

8.3.3 Protein Surface Properties

Proteins procured from different sources have different surface activities, including emulsion and foam properties. Differences in protein surface activities are mostly related to different conformations proteins can assume. Several interdependent factors (including polypeptide chain stability/flexibility, adaptability to environmental change, amounts and distribution patterns of hydrophilic and hydrophobic amino acid residues in their primary polypeptide chain and on the surfaces of the proteins) affect their surface properties. A balance of non-covalent interactions, including electrostatic attractions, hydrogen bonding, covalent disulphide bonding and hydrophobic interactions, determines the emulsion and foaming properties of proteins (Damodaran 1996). However, not only intrinsic factors but also extrinsic factors such as protein concentration, pH, temperature, ionic strength and type of ion and pressure can affect protein-based products (Phillips et al. 1994). A proper balance of attractive, repulsive and hydration forces is required to form strong and stable protein-based products.

The structural and functional properties of proteins make them useful for various industrial applications. For example, proteins can be made into elastic materials that can survive repeated stress-strain cycles, and they can aggregate to form films that provide barriers to gases, moisture and bacteria. Films produced from protein could be used in food packaging, in paper coatings and in bandaging materials. Protein-based materials have become a research focus because of their high performance, low cost and environmentally friendly characteristics (Kumar et al. 2008; Sasmal et al. 2008).

8.4 Plant Proteins

Agricultural crops have been the main source of food and feed for many centuries because they contain high levels and high concentrations of protein, starch and oil. Plants have the potential to function as efficient platforms for pharmaceutical and therapeutic protein manufacturing because their production costs are low, they produce high quantities of proteins and extraction protocols are relatively facile and efficient.

A number of crops including cereals grains, legumes and pulses are rich in plant proteins (Dangaran et al. 2011) that can be used directly for manufacturing protein-based bioproducts. Corn zein, wheat gluten, sorghum kafirin, rice (*Oryza sativa*) bran protein, soy protein, peanut protein and cottonseed protein have all been utilized to manufacture protein bioproducts. Because plant proteins are varied in their chemical, physical, functional and structural properties (Table 8.2), bioproducts made from them have a wide range of uses. In this section, a few common compositions of plant proteins are discussed.

Table 8.2 Amino acid composition (% of protein) of some plant proteins

Amino acids	Corn zein[a]	Wheat gluten[b]	Soybean[c]	Mung bean[d]	Red kidney bean[e]
Glycine	0.0	3.9	1.8	3.6	4.2
Alanine	10.5	3.1	1.8	5.1	4.5
Valine	4.0	5.3	2.5	6.7	4.6
Leucine	21.3	8.3	3.3	9.3	8.1
Isoleucine	5.0	4.9	1.9	5.5	3.9
Phenylalanine	7.3	6.0	2.2	6.8	5.9
Tryptophan	0.2	NR	0.5	NR	1.2
Proline	10.5	14.5	2.0	4.1	4.4
Serine	7.1	5.2	1.9	2.5	6.1
Threonine	3.5	3.5	1.6	1.8	4.5
Tyrosine	5.3	3.8	1.5	3.0	2.9
Methionine	2.4	1.7	0.6	1.3	1.4
Cysteine	0.8	2.3	0.7	–	1.0
Lysine	0.0	2.1	2.7	6.8	7.5
Arginine	4.7	4.5	3.2	7.0	6.3
Histidine	1.3	2.3	1.2	2.5	3.0
Aspartic acid	4.6	3.7	4.8	10.2	13.1
Glutamic acid	26.9	41.0	7.7	19.5	17.3

[a]Golenkov (1985)
[b]Coffman and Garcia (1977)
[c]Mundi and Aluko (2012)
[d]Mosse (1961)
[e]Kovalenko et al. (2006)

8.4.1 Corn Zein

Corn (*Zea mays*) contains 8–12% protein in its kernels (Earle 1977). Gorham first described corn zein in 1821, and Osborn classified it as a prolamin, which is a seed storage, globular, water-ethanol-soluble, protein containing large amounts of proline and glutamine and small amounts of arginine, lysine and histidine (http://www.britannica.com/EBchecked/topic/478591/prolamin), in 1924. Corn **zein** makes up 50% or more of the corn kernel proteins. It is distributed in the outer layer of corn kernels as small, compact bodies embedded in the glutelin protein matrix (Padua and Wang 2002). Because of its low solubility in water, it has a greasy and glossy appearance. It was first commercialized in 1939 for the production of adhesives, plastic films, coatings and fibre applications.

Zein is an aggregate of single polypeptide protein subunits, including α-zein (23–27 kDa), β-zein (15 kDa), γ-zein (27 kDa) and δ-zein (10 kDa). Commercial corn zein is 75–85% **α-zein**. This subunit is rich in non-polar amino acids, including glutamine (21–26%), leucine (20%), proline (10%), alanine (10%) and phenylalanine (8%), but is deficient in basic and acidic amino acids (Table 8.2; Wilson 1988). A higher proportion of non-polar amino acids in α-zein make it soluble in alcohols.

β-Zein is rich in methionine (10%) and tyrosine (8%) and constitutes 10–15% of the total corn zein. **γ-Zein** is rich in proline (25%) and histidine (8%) and accounts for 5–10% of the total corn zein. **δ-Zein** is rich in sulphur-containing amino acids (methionine and cysteine) and constitutes less than 5% of the total zein protein (Thompson and Larkins 1989; Shukla and Cheryan 2001). β-Zein, γ-zein and δ-zein are soluble in aqueous alcohol and a reducing agent (Thompson and Larkins 1989) and are classified as **glutelins**, which are cereal storage proteins soluble in dilute acid or base and not coagulated by heat (http://www.merriam-webster.com/medical/glutelin).

8.4.2 Wheat Gluten

Wheat kernels contain 8–14% protein (Delcour et al. 2012). Wheat proteins (Fig. 8.3) are classified into four groups on the basis of their solubilities and include water-soluble albumins, salt-soluble globulins, alcohol-soluble gliadins and diluted alkaline−/acid-soluble glutenins. Generally, wheat proteins are classified as

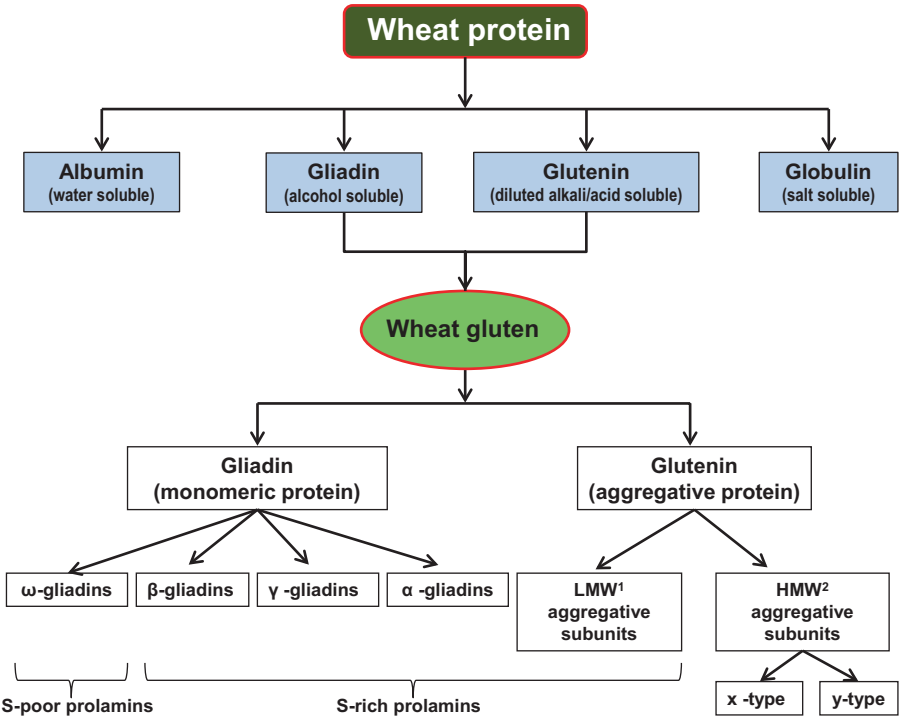


Fig. 8.3 Classification of wheat protein. [1]Lower molecular weight.[2]High molecular weight. S sulphur

glutens, comprising gliadins and glutenins, and **non-glutens**, comprising albumins and globulins. Non-gluten wheat proteins generally constitute 15 to 20% of the total wheat protein (Gupta et al. 1992). They are predominantly monomeric proteins with molecular mass lower than 25 kDa. However, a significant proportion of the chains exist as aggregated polymers with molecular mass between 60 and 70 kDa that are stabilized by intermolecular disulphide bonds. The non-gluten wheat proteins have metabolic (albumins) or structural (globulins) functions.

The addition of water to wheat flour followed by mixing results in the production of a tough, rubbery, elastic substance called gluten. At a microscopic level, gluten is a continuous network of strands composed of gliadin and glutenin proteins stabilized by intermolecular disulphide bonds and non-covalent bonds such as hydrogen bonds and hydrophobic interactions (Figs. 8.4, 8.5). The structure has important viscoelastic properties that make it ideal for bread-making (Veraverbeke and

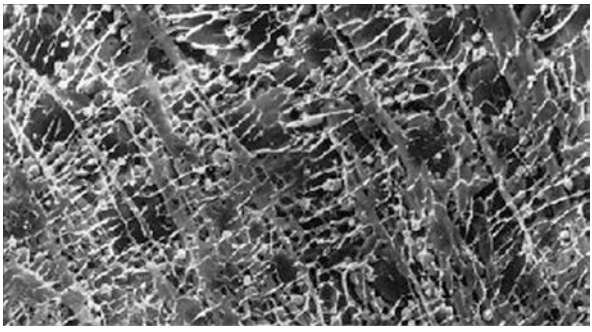


Fig. 8.4 Scanning electron micrograph of wheat gluten. (Source: http://www.bakeinfo.co.nz/Facts/Gluten)

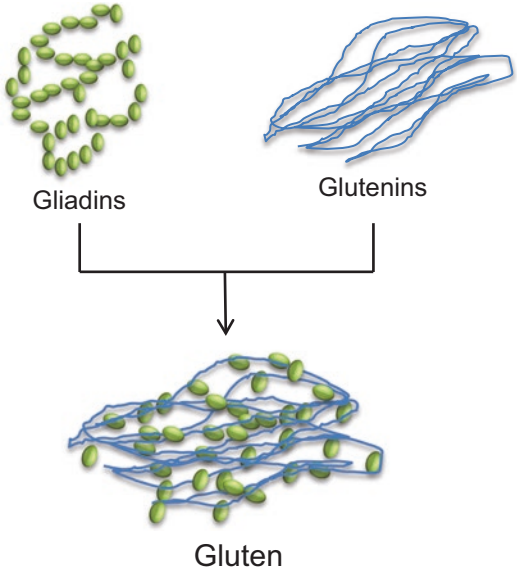


Fig. 8.5 Schematic representation of wheat gluten

Delcour 2002), including its high water absorption capacity (approximately twice its weight), stickiness, extensibility and elasticity. To extract gluten, wheat flour is mixed with water and washed extensively to remove the starch and other water-soluble proteins. The remaining gluten mass is extruded, chopped and dried for packaging. Industrial wheat gluten contains approximately 70–80% protein (33–45% alcohol-soluble gliadin and 40–50% alkaline−/acid-soluble glutenin), 10–14% polysaccharide, 6–8% lipid and 0.8–1.4% minerals. Both gliadin and glutenin are rich in non-polar amino acids, proline and glutamine, and poor in polar amino acid residues (Table 8.2), which contributes to their low solubility in water. Gliadins are compact, globular and viscous in nature (Shewry et al. 1986; Veraverbeke and Delcour 2002). They act as plasticizers in gluten and contribute to its gas retention properties in bread dough.

8.4.3 Soy Protein

Soybean contains up to 50% protein in the dry seed. Soybean protein is a by-product of the soybean oil industry (Rhim et al. 2000) and is available as soy flour, soy protein concentrates and soy protein isolates (SPI) containing 50–59%, 65–72% and ≥90% protein, respectively. Globulins are the most abundant class of proteins in soy protein isolates. Glutamic acid, aspartic acid, arginine, leucine, lysine and valine are the major constituent amino acids in soy protein isolate profiles (Table 8.2). Soy proteins are classified as 2S, 7S, 11S and 15S on the basis of their sedimentation during ultracentrifugation (Fig. 8.6) (larger **Svedberg (S) numbers** indicate smaller protein MW). They account for 22, 37, 31 and 1% of the total protein, respectively (Wolf and Briggs 1956). 11S (**glycinin**) and 7S (**β-conglycinin**) are the principal proteins and account for approximately 70% of the total seed protein in soybean (Thanh and Shibasaki 1976).

The 11S and 7S soy proteins are significantly different in their physical, chemical and functional properties (Wolf and Tamura 1968; Mori et al. 1981). For example, the 11S protein precipitates faster and forms larger aggregates than 7S, and its gels have a higher water-holding capacity than 7S gels. In addition, 11S proteins have higher tensile values and higher hardness and expand more on heating than 7S proteins. The 11S:7S ratio is therefore a predictive indicator of relative protein functional properties (Kwanyuen et al. 1998).

8.4.4 Bean Protein

Dry beans including navy (*Phaseolus vulgaris*), pinto (*Phaseolus vulgaris*), lima (*Phaseolus lunatus*), kidney (*Phaseolus vulgaris*), black (*Phaseolus vulgaris*), white (*Phaseolus vulgaris*), red (*Vigna angularis* and *Vigna umbellata*), pink (*Phaseolus vulgaris*), lentil (*Lens culinaris*, also known as *Lens esculenta)*, black

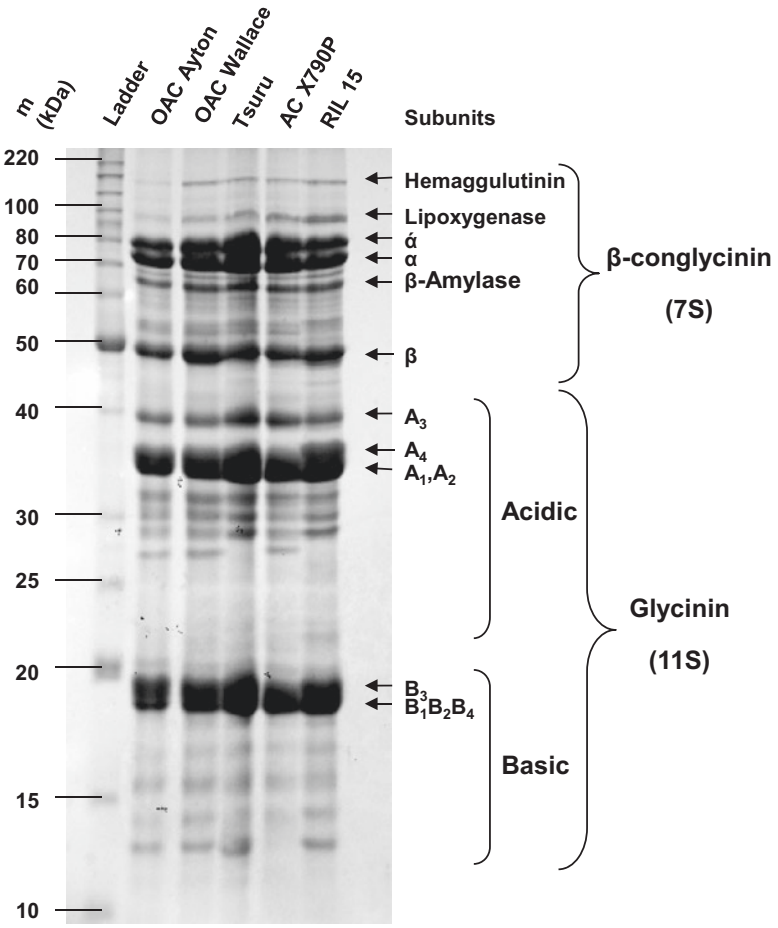


Fig. 8.6 1D SDS-polyacrylamide gel of seed storage proteins in selected soybean genotypes. Regions containing 7S and 11S proteins are shown, and some specific protein bands are identified. m molecular mass

eye (*Vigna unguiculata*), black gram (*Vigna mungo*), garden pea (*Pisum sativum*), chickpea (*Cicer arietinum*), moth bean (*Vigna aconitifolia*), jack bean (*Canavalia ensiformis*)and tepary bean (*Phaseolus acutifolius)* are important food sources in many countries. Common dry bean (*Phaseolus* spp.) seeds contain 15 to 30% protein, on a dry weight basis (Sathe 2002; Yin et al. 2008). The major classes of proteins in dry beans are water-soluble albumin and mildly alkali-soluble globulins. There are also some minor proteins, such as enzyme inhibitors and lectins. Albumin and globulin account for ~80% of the total storage protein with their molecular masses ranging from 10 to 400 kDa. The quantities and ratios of albumins to globulins depend on the species and variety of the dry bean. Generally, dry beans contain 10–30% albumins and 45–70% globulins (Sathe 2002).

Globulins are further grouped by their solubilities at different ionic strengths and by sedimentation in ultracentrifugation. They are classified into **vicilin/phaseolin** and **legumin** with coefficients of sedimentation of 7S and 11S, respectively (Danielsson 1949). Both vicilin and legumin can also be distinguished by their oligomeric organization and polypeptide structures. The 7S vicilin/phaseolin is a glycoprotein of three single peptide chains with molecular mass of 50–75 kDa, while 11S legumin is a hexameric protein linked by disulphide bond(s) between acidic α- and basic β-subunits with molecular mass of ~40 and ~60 kDa, respectively. Six polypeptides are linked together to constitute the legumin molecule with a molecular mass of 320–400 kDa (Lawrence et al. 1990; Sathe 2002). The 7S protein accounts for 50–55% of the total globulin protein content in dry beans on a dry weight basis (Sathe 2002). There is an inverse relationship between 7S and 11S concentrations (e.g. increase in the quantity of 7S causes a proportional decrease in the amount of 11S and vice versa, so the net protein concentrations remain constant) (Ogawa et al. 1989). Dry beans have small amounts of other storage proteins in addition to the globulins, including lectins (a tetrameric glycoprotein of 7S) or phytohaemagglutinin with molecular masses of 27–37 kDa, arcelin (a 2S dimer glycoprotein) with a molecular mass of 35–42 kDa, sulphur-rich protein (a 3S dimer glycoprotein) with a molecular mass of 29–32 kDa and enzyme inhibitor proteins (Sathe 2002).

8.5 Protein Isolation and Purification from Plants

Often, the first step in producing bioproducts from proteins is isolating pure fractions of proteins of interest from mixtures of several other macromolecules, including polysaccharides, lipids and other proteins. However, there is no standard procedure for isolating and purifying proteins from all sources (organisms) because the proteins may vary from source to source in amino acid composition, sequence, structure, size, shape, net charge, isoelectric point, solubility, heat stability, hydrophobicity, ligand/metal binding capacity and post-translational modification.

The level of purity, functionality and quantity of the protein required for a particular application will influence the approach used to purify it. For example, a few micrograms for enzymatic kinetic studies to several kilograms for industrial and pharmaceutical applications can be required. The highest quality and functionality are required for medical uses compared to lower stringencies for proteins used to manufacture commercial films, coatings, fibres and adhesives.

The first step in protein isolation is crushing/milling of the plant source to release the protein from the biological matrix of cells. Generally, the desired protein is released as a mixture with other macromolecules, including other proteins, polysaccharides and lipids. The next step is to separate the targeted proteins from the mixture of molecules through different procedures or a combination of more than one procedure. Protein solubility plays a crucial role in protein isolation and purification. Plant proteins including corn zein, wheat gluten, soy protein, bean protein, peanut protein and cottonseed protein differ in their solubility properties. For

Table 8.3 Protein isolation and purification process[a]

Separation process	Basis of separation
1. Precipitation	
Alkali (ammonium sulphate)	Solubility
Acetone	Solubility
Polyethyleneimine	Charge, size
Isoelectric	Solubility, pI (isoelectric point)
2. Phase partitioning	
Polyethylene glycol	Solubility
3. Chromatography	
Ion exchange (IEX)	Charge, charge distribution
Hydrophobic interaction (HIC)	Hydrophobicity
Reverse-phase HPLC	Hydrophobicity, size
Affinity	Ligand-binding site
DNA affinity	DNA-binding site
Lectin affinity	Carbohydrate content and type
Immobilized metal affinity (IMAC)	Metal binding
Immunoaffinity (IAC)	Specific antigen site
Chromatofocussing	pI (isoelectric point)
Gel filtration/size exclusion (SEC)	Size, shape
4. Electrophoresis	
Gel electrophoresis (PAGE)	Charge, size, shape
Isoelectric focussing (IEF)	pI (isoelectric point)
5. Centrifugation	Size, shape, density
6. Ultrafiltration	Size, shape

[a]Adapted from Burgess (2008)

example, corn zein is soluble in aqueous solution, soy and bean proteins in mild alkali and wheat gluten in mild acid, mild alkali and aqueous ethanol solutions.

A good isolation procedure is simple, reproducible, specific, reliable and cost-effective. Protein extraction procedures, based on their physicochemical properties, are categorized into precipitation, phase partitioning, chromatographic, centrifugation and ultrafiltration procedures (Table 8.3).

8.5.1 Corn Zein Isolation

Commercial zein is a by-product of the corn milling industry. In corn wet-milling, the fine slurry that remains after centrifugal extraction of starch from the endosperm comprises a protein-rich mass called corn gluten meal from which zein is extracted. Corn zein can also be extracted from dried distillers' grains with solubles, a by-product of corn ethanol manufacturing. The commercial method of zein extraction was patented in 1970 by Carter and Reck (Fig. 8.7). Corn gluten meal is solubilized with organic solvent such as isopropyl alcohol a few times to get highly

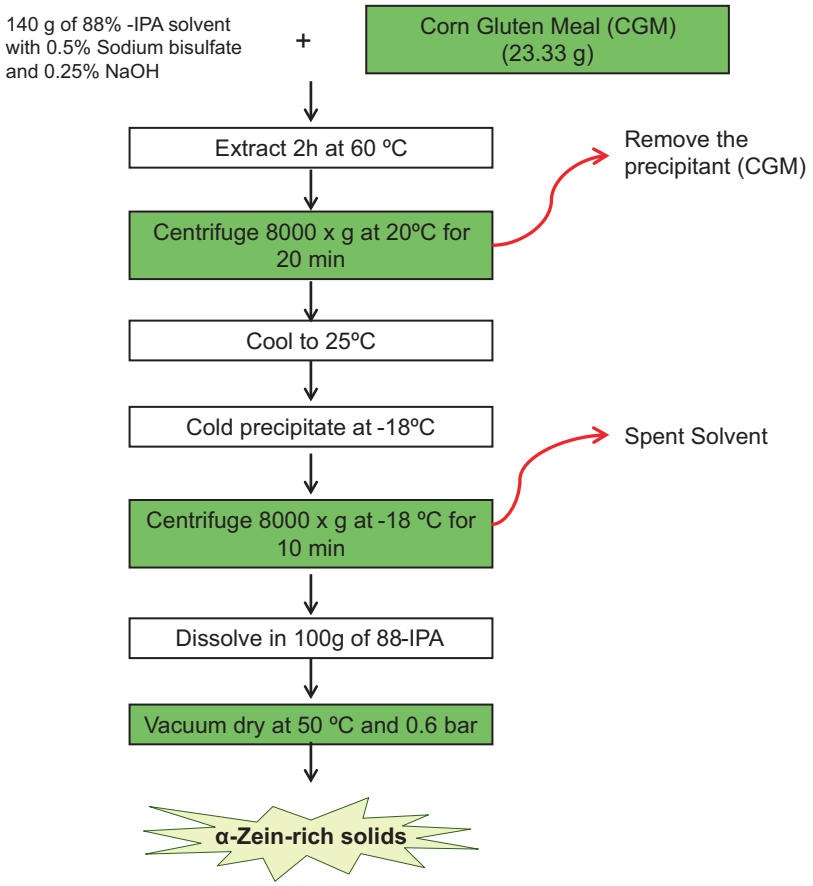


Fig. 8.7 Flow diagram for Carter and Reck (1977) zein extraction procedure. (Adopted from Anderson and Lamsal (2011))

pure corn zein. Precipitation and centrifugation procedures are collectively used to isolate corn zein for industrial applications.

8.5.2 Wheat Gluten Isolation

Commercial wheat gluten is the by-product of the wheat milling industry. After centrifugal removal of most of the starch from wet-milled endosperm, the supernatant is a fine, protein-rich, slurry called **wheat gluten meal**, from which gluten is extracted. The water-insoluble wheat glutens aggregate and form low-density but larger-sized particles than starch. These properties of wheat gluten are utilized to isolate and purify it from the starch. Several factors influence the commercial production of wheat gluten, including the protein content in the raw material,

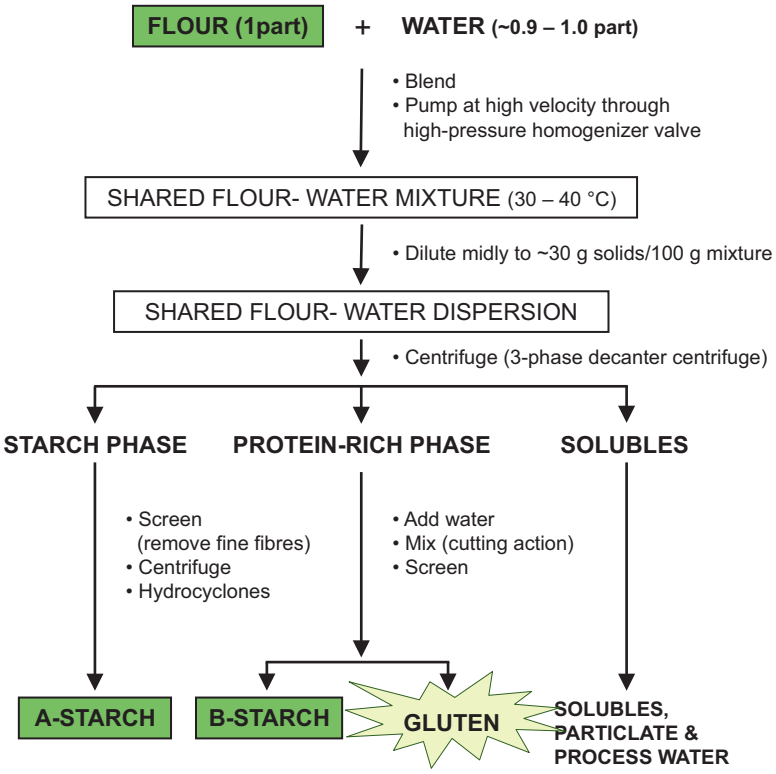


Fig. 8.8 Flow diagram for high-pressure disintegration process on a sheared flour-water dispersion. (Adopted from Sayaslan (2004))

starch-protein production ratio, handling of the processing effluent water, yield and purity of gluten and the cost of production. A number of processes have also been developed which use dry-milled flour as the raw material for these processes. The separation of starch and gluten particles is the initial step in gluten isolation. The extracted gluten granules are further water washed, aggregated, purified and flash dried to yield about 80% proteins.

Extraction processes are typically named after the company or the person who developed or patented the process. Industrially, four processes, including the Martin, Raisio/Alfa-Laval, hydrocyclone and high-pressure disintegration process, are the most popular to produce commercial wheat gluten (Sayaslan 2004). High-pressure disintegration (Fig. 8.8) is the most recent process for starch and gluten isolation from wheat flour. It was initially developed for potato starch extraction and later modified for corn and wheat. Comparatively, this process is the most efficient in terms of water consumption. It consumes 3–4 parts water per part flour compared to 4–5 parts in the HC and 5–7 parts in the modified Martin and Batter processes. This process separates the starch and gluten based on density. Low-gluten flour, such as soft wheat flour, can be used as raw material.

8.5.3 Soy Protein Isolation

Soy proteins are extracted from defatted soy flakes or meal derived from the soybean processing industry. Soy protein concentrates are extracted with 60–80% aqueous alcohol and soy protein isolates with aqueous mild alkali-extraction followed by isoelectric precipitation (Cho et al. 2007). In brief, the defatted soy flakes or meal are dispersed and moderately agitated in heated water at a ratio of between 1:10 and 1:20 and a pH of 7.5–9.0 for 45 min to 1 h. The mixture is centrifuged, and the pulp and extract are collected. The pulp, which contains mostly starch, is rewashed in water for maximum protein extraction. Extracts are filtered to remove any solids, acidified, centrifuged or filter separated. The precipitated curd is collected and washed with water, suspended in a minimum volume of water at neutralized pH and spray-dried in powder form for commercial use (Fig. 8.9).

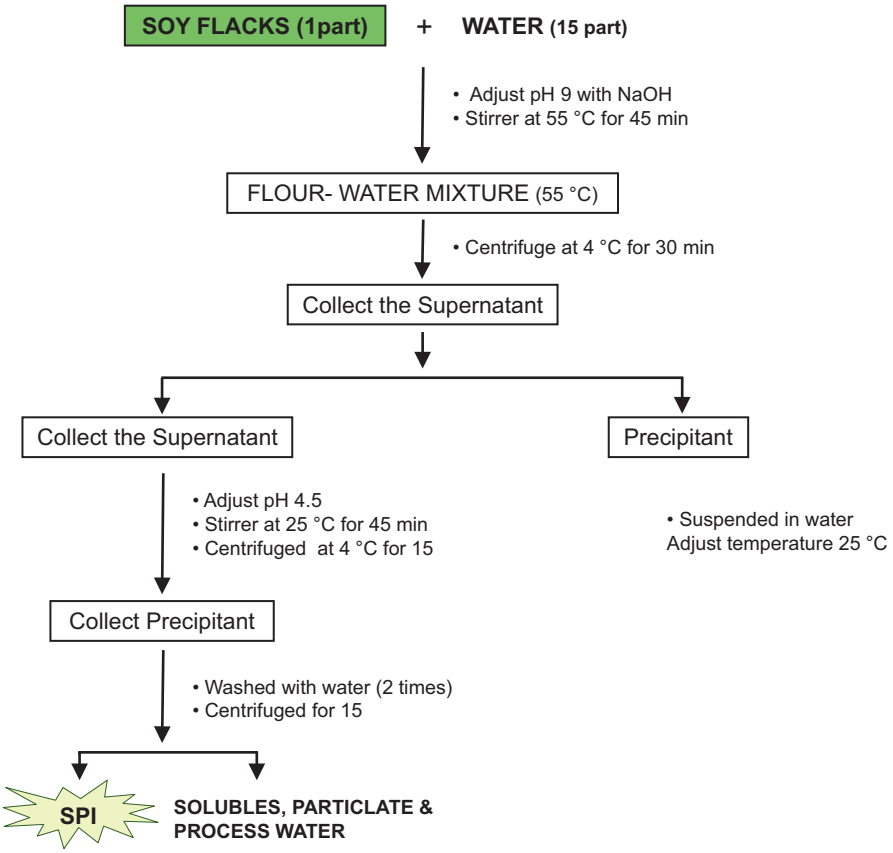


Fig. 8.9 Flow diagram of soy protein isolates extraction conventional procedure from defatted soy flakes meal

8.6 Closing Comments

Plant proteins are very diverse in nature in terms of their physiochemical and structural properties. The combination of different amino acids can generate a huge number of diverse secondary and tertiary structures. In addition, post-translational modifications increase the diversity of proteins. This variation lends them distinct functional characteristics that can be leveraged for the production of a wide range of bioproducts. Major crops could be a source of plant proteins. Zein, gluten and globulin are major proteins in corn, wheat and soybean, respectively.

References

Anderson TJ, Lamsal BP (2011) Development of new method for extraction of α-zein from corn gluten meal using different solvents. Cereal Chem 88:356–362

Burgess RR (2008) Protein purification. In: Nothwang HG, Pfeiffer SE (eds) Proteomics of the nervous system. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Carter R, Reck DR (1977) Low temperature solvent extraction process for producing high purity zein. US Patent 4716218

Cho SY, Park JW, Batt HP, Thomas RL (2007) Edible films made from membrane processed soy protein concentrates. LWT- Food Sci Technol 40:418–423

Coffman CW, Garcia VV (1977) Functional properties and amino acid content of a protein isolate from mungbean flour. J Food Technol 12:473–484

Damodaran S (1996) Amino acids, peptides, and proteins. In: Fennema OR (ed) Food chemistry, 3rd edn. Marcel Dekker, New York

Damodaran S (2008) Amino acids, peptides, and proteins. In: Damodaran S, Parkin KL, Fennema OR (eds) Food chemistry, 4th edn. CRC Press, Boca Raton

Dangaran K, Tomasula PM, Qi P (2011) Structure and function of protein-based films and coatings. In: Plackett D (ed) Biopolymers - new materials for sustainable films and coatings, 1st edn. Wiley, West Sussex

Danielsson CE (1949) Seed globulins of the Gramineae and Leguminosae. Biochem J 44:387–400

Delcour JA, Joye IJ, Pareyt B, Wilderjans E, Brijs K, Lagrain B (2012) Wheat gluten functionality as a quality determinant in cereal-based food products. Annu Rev Food Sci Technol 3:469–492

Earle FR (1977) Protein and oil in corn: variation by crop years from 1907 to 1972. Cereal Chem 54:70–79

Golenkov VF (1985) Comparative amino acid composition of proteins of wheat, rye and triticale grain. In: Lásztity R, Hidvégi M (eds) Amino acid composition and biological value of cereal proteins. Springer, Dordrecht

Gupta RB, Batey IL, MacRitchie F (1992) Relationships between protein composition and functional properties of wheat flours. Cereal Chem 69:125–131

Heldt HW, Piechulla B (2010) Plant biochemistry, 4th edn. Academic Press, London

Herman EM, Larkins BA (1999) Protein storage bodies and vacuoles. Plant Cell 11:601–614

Herman EM (2008) Endoplasmic reticulum bodies: solving the insoluble. Curr Opin Plant Biol 11:672–679

Kovalenko IV, Rippke GR, Hurburgh CR (2006) Determination of amino acids composition of soybean (*Glycine max*) by near-infrared spectroscopy. J Agric Food Chem 54:3485–3491

Krochta JM (2002) Proteins as raw materials for films and coatings: definitions, current status, and opportunities. In: Gennadios A (ed) Protein-based films and coatings. CRC Press, Boca Raton

Kumar R, Liu D, Zhang L (2008) Advances in proteinous biomaterials. J Biobaased Mater Bioenergy 2:1–24

Kuntz ID Jr, Kauzmann W (1974) Hydration of proteins and polypeptides. Adv Protein Chem 28:239–345
Kwanyuen P, Wilson RF, Burton JW (1998) Soybean protein quality. In: Koseoglu SS, Rhee KC, Wilson RF (eds) Emerging technologies, current practices, quality control, technology transfer and environmental issues. AOCS Press, Champaign
Lawrence MC, Suzuki E, Varghese JN, Davis PC, van Donkelaar A, Tulloch PA, Colman PM (1990) The three-dimensional structure of the seed storage protein phaseolin at 3 Å resolution. EMBO J 9:9–15
McMurry J (1994) Quimica organica. grupo editorial iberoamerica, Mexico
Mierovich H, Scheraga HA (1980) Empirical studies of hydrophobicity. 2. Distribution of the hydrophobic, hydrophilic, neutral, and ambivalent amino acids in the interior and exterior layers of native proteins. Macomolecules 13:1406–1414
Mori T, Nakamura T, Utsumi S (1981) Gelation mechanism of soybean 11S globulin: formation of soluble aggregates as transient intermediates. J Food Sci 47:26–30
Mosse J (1961) Monographie sur une protéine du mais: la zéine. Ann Physiol 3:105–139. (French)
Moure A, Sineiro J, Dominguez H, Parajo JC (2006) Functionality of oilseed protein products: a review. Food Res Int 39:945–963
Moran LA, Horton HR, Scrimgeour KG, Perry MD (2012) Principles of biochemistry, 5th edn. Pearson, Toronto
Mundi S, Aluko RE (2012) Physiochemical and functional properties of kidney bean albumin and globulin protein fractions. Food Res Int 48:299–306
Nelson DL, Cox MM (2017) Lehninger principles of biochemistry, 7th edn. W. H. Freeman and Company, New York
Ogawa T, Tayama E, Kitamora K, Kaizama N (1989) Genetic improvement of seed storage protein using three variant alleles of 7S globulin sub-unit in soybean (*Glycine max*). Japan J Breed 39:137–147
Otegui MS, Herder R, Schulze J, Jung R, Staehelin LA (2006) The proteolytic processing of seed storage proteins in *Arabidopsis* embryo cells starts in the multivesicular bodies. Plant Cell 18:2567–2581
Padua GW, Wang Q (2002) Formation and properties of corn zein films and coatings. In: Gennadios A (ed) Protein-based films and coatings. CRC Press, Boca Raton
Phillips LG, Whitehead DM, Kinsella JE (1994) Structure-function properties of food proteins. Academic Press, San Diego
Pommet M, Redl A, Morel MH, Domenek S, Guilbert S (2003) Thermoplastic processing of protein-based bioplastics: chemical engineering aspects of mixing, extrusion and hot molding. Macromol Symp 197:207–217
Rhim JW, Gennadios A, Handa A, Weller CL, Hanna MA (2000) Solubility, tensile, and color properties of modified soy protein isolate films. J Agric Food Chem 48:4937–4931
Sasmal A, Sahoo D, Nanda R, Nayak P, Nayak PL, Mishra JK, Chang Y, Yoon J (2008) Biodegradable nanocomposites from maleated polycaprolactone/soy protein isolate blend with organoclay: preparation, characterization, and properties. Polym Compos 30:708–714
Sathe SK (2002) Dry bean protein functionality. Crit Rev Biotechnol 22:175–223
Sayaslan A (2004) Wet-milling of wheat flour: industrial processes and small-scale test methods. LWT- Food Sci Technol 37:499–515
Shewry PR, Tatham AS, Forde J, Kreis M, Miflin BJ (1986) The classification and nomenclature of wheat gluten proteins: a reassessment. J Cereal Sci 4:97–106
Shimada K, Matsushita S (1980) Relationship between thermocoagulation of proteins and amino acid compositions. J Agric Food Chem 28:13–17
Shukla R, Cheryan M (2001) Zein: the industrial protein from corn. Ind Crop Prod 13:171–192
Stevens MP (1999) Polymer chemistry. An introduction. Oxford Univ Press, New York
Thanh VH, Shibasaki K (1976) Heterogeneity of betaconglycinin. Biochim Biophys Acta 439:326–338
Thompson GA, Larkins BA (1989) Structural elements regulating zein gene expression. BioEssays 10:109–113

Totosaus A, Montejano JG, Salazar JA, Guerrero I (2002) A review of physical and chemical protein-gel induction. Int J Food Sci Technol 37:589–601
Veraverbeke WS, Delcour JA (2002) Wheat protein composition and properties of wheat glutenin in relation to bread making functionality. Crit Rev Food Sci Nutr 42:179–208
Voet D, Voet JG, Pratt CW (2016) Fundamentals of biochemistry: life at the molecular level, 5th edn. Wiley, Hoboken
Webster DE, Thomas MC (2012) Post-translational modification of plant-made foreign proteins; glycosylation and beyond. Biotechnol Adv 30:410–418
Wilson CM (1988) Electrophoretic analyses of various commercial and laboratory-prepared zeins. Cereal Chem 65:72–73
Wolf WJ, Briggs DR (1956) Ultracentrifugal investigation of the effect of neutral salts on the extraction of soybean proteins. Arch Biochem Biophys 63:40–49
Wolf WJ, Tamura T (1968) Heat denaturation of soybean 11S protein. Cereal Chem 46:331–344
Yin SW, Tang CH, Wen QB, Yang XQ, Li L (2008) Functional properties and in vitro trypsin digestibility of red kidney bean (*Phaseolus vulgaris* L.) protein isolate: effect of high-pressure treatment. Food Chem 110:938–945