

2.2 Dynamic patterns of protein accumulation during seed development

Protein accumulation during soybean seed development is temporally regulated rather than continuous. Early stages contain relatively higher proportions of low-molecular-weight or nonstorage proteins, whereas the major storage proteins accumulate later as seeds enter active filling and maturation; correspondingly, storage-protein transcripts appear before visible protein deposition, indicating transcriptional control ahead of bulk reserve accumulation. The dominant storage fractions are 7S globulins and 11S glycinin, and their accumulation generally intensifies from mid seed filling toward late filling or maturation, although individual subunits differ in timing. More recent proteomic and compositional studies further show that 7S β -subunit and several 11S glycinin subunits increase steadily across filling stages, globulins remain higher than albumins throughout development, and globulin accumulation often peaks around R6 rather than rising linearly to maturity (Islam et al., 2021; Montanha et al., 2023).

2.3 Dynamic patterns of oil accumulation during seed development

Oil accumulation also shows a marked developmental pattern, beginning later than residual carbohydrate deposition and increasing rapidly during seed filling. A water-relations analysis indicated that residual accumulation starts first, followed by protein and then oil, while classical compositional studies showed that oil percentage rises sharply from about 24 to 40 days after flowering before stabilizing later in development. This rapid lipid deposition is supported by coordinated metabolic reprogramming during developing seed filling. Genes involved in carbon fixation, photosynthesis, glycolysis, and fatty acid biosynthesis are up-regulated during the critical period of oil accumulation, and later developmental transitions are associated with strong activation of carbohydrate degradation, triacylglycerol biosynthesis, and phospholipid signaling pathways, consistent with intensified flux toward storage lipid production in the late filling stages (Niu et al., 2025). Soybean seed development therefore involves a tightly staged transition from morphogenesis to reserve filling and finally desiccation, with protein and oil accumulation following distinct but overlapping temporal programs shaped by developmental state, carbon-nitrogen partitioning, and late-stage metabolic regulation.

3 Molecular Mechanisms Regulating Protein and Oil Biosynthesis in Developing Soybean Seeds

3.1 Genetic regulation of storage protein biosynthesis

Storage protein biosynthesis in developing soybean seeds is controlled by a hierarchical genetic program that integrates seed maturation regulators with structural genes encoding the major storage proteins. The principal storage proteins are β -conglycinin (7S) and glycinin (11S), and their accumulation is regulated not only by the expression of their own gene families but also by upstream AFL-type regulators, including LEC1, LEC2, FUS3, and ABI3, that coordinate the broader maturation state required for reserve deposition (Qi et al., 2026). Genetic mapping further shows that variation in storage protein composition is polygenic, with multiple loci affecting glycinin, β -conglycinin, total storage protein subunits, and the glycinin: β -conglycinin ratio, while polymorphism in the Gy1 promoter is specifically associated with 11S glycinin content (Zhang et al., 2021).

Recent transcriptomic and functional studies indicate that protein accumulation is also regulated through networks linking hormone signaling, nitrogen allocation, and storage protein processing. In contrasting soybean genotypes, candidate pathways associated with seed protein metabolism include photosynthesis, the TCA cycle, and starch and sucrose metabolism, and 40 days after flowering appears to be a critical stage for protein accumulation (Hu et al., 2025). At the gene level, *GmGASA12* acts as a molecular hub: its knockout increases water-soluble protein content, upregulates amino acid transporters and storage protein genes, and modulates the cooperative biosynthesis of β -conglycinin and glycinin through interaction with GmCG-6 (Yang et al., 2025).

3.2 Molecular regulation of oil biosynthesis and fatty acid metabolism

Oil biosynthesis in developing soybean seeds is governed by transcriptional circuits that channel carbon into fatty acid synthesis in plastids and triacylglycerol assembly in the endoplasmic reticulum. WRINKLED1 is a central regulator of this process, directly controlling numerous genes involved in lipid biosynthesis and participating in a positive feedback relationship with LEC1 that promotes the onset and stability of the fatty acid and TAG