Chapter 6
Seed-Based Production of Recombinant Proteins

Cristiano Lacorte, Amanda Lopes Ferreira, Aline Melro Murad, Nicolau Brito da Cunha, and Patricia Valle Pinheiro

Abstract The use of plants to produce recombinant proteins has become a promising alternative to current expression systems based on microorganisms and cell cultures. Producing recombinant proteins in plants presents advantages such as reduced upstream costs, lower probability of infection by mammals' pathogens, and easy scale-up production. Furthermore, plants can make posttranslational modifications and express large and complex proteins. Seeds stand out for their intrinsic characteristics among the plant tissues that can be used for recombinant protein production. In seeds, recombinant proteins are stored in protein storage vacuoles in the endosperm cells. They can be stably maintained for 4–6 years at room temperature without significant loss of protein and biological activity. Progress in methods for genome editing, alteration of the glycosylation pattern of the recombinant proteins produced, and use of protein fusions to aid downstream processes are some trends that shall contribute further to make molecular farming an attractive option for recombinant protein production. This chapter discusses the potential of seeds as a platform for producing recombinant proteins, across the scale-up of the production systems, basic approaches for the purification of recombinant proteins from plant cells, and biosafety issues.

C. Lacorte (✉)
Embrapa Recursos Genéticos e Biotecnologia, Brasília, Brazil
e-mail: cristiano.lacorte@embrapa.br

A. L. Ferreira · P. V. Pinheiro
Embrapa Arroz e Feijão, Santo Antônio de Goiás, Brazil
e-mail: patricia.pinheiro@embrapa.br

A. M. Murad
Universidade Federal de Santa Catarina, Florianópolis, Brazil

N. B. da Cunha
Universidade de Brasília, Brasília, Brazil

C. Kole et al. (eds.), *Tools & Techniques of Plant Molecular Farming*, Concepts and Strategies in Plant Sciences, https://doi.org/10.1007/978-981-99-4859-8_6

Keywords Plant molecular farming · Heterologous protein production · Foreign gene expression · Glycosylation · Fusion proteins · Recombinant protein purification · Biosafety

Abbreviations

AA	Amino acids
APHIS	Animal and Plant Health Inspection Service
ELP	Elastin-like polypeptides
ER	Endoplasmic reticulum
FDA	Food and Drug Administration
GFP	Green fluorescent protein
HIV	Human immunodeficiency virus
ITC	Inverse transition cycling
PB	Protein bodies
PSV	Protein storage vacuoles

6.1 Introduction

The production of high-value biopharmaceuticals and other recombinant proteins for diagnosis or industrial application is mainly based on microbial and cell cultures in large bioreactors (De Jaeger et al. 2002; Tschofen et al. 2016; Kesik-Brodacka 2018). However, research has demonstrated the expression of hundreds of different recombinant proteins in plants over the last decades. Some have reached commercial production, confirming the viability and potential of this approach (Schillberg et al. 2019; Liu and Timko 2022). The advantages of using plants as an alternative to the other, more established systems based on microbial and cell culture are the reduced costs, the lower probability of infection by pathogens capable of infecting humans and other mammals, and the easy scale-up production (Fischer and Buyel 2020; Schillberg and Finnern 2021). Biomass production from plants is more sustainable than bioreactor-based production once it requires soil, fertilizers, and water, instead of expensive complex culture media (Obembe et al. 2011). Besides, the technology for cultivation is largely available and demands less skilled staff (Fischer and Buyel 2020; Liu and Timko 2022).

Different plant-based systems have been described for the production of recombinant proteins (see also chapter "Molecular Farming of Pharmaceutical Proteins in Different Crop Systems: A Way Forward"). These include the transient expression in leaves and the stable integration of the foreign gene for the expression in seeds, leaves, fruits, tubers, roots, aquatic plants, moss, hairy root culture, and cell suspension culture (Xu et al. 2018). Each of these systems presents advantages and specific applications. However, to be competitive, a plant-based system must ensure a high-level expression of the recombinant protein. To this end, several factors that can

increase both gene expression and protein accumulation in the cell should be considered. These include the type of tissue or cell; selection of strong promoters and enhancer elements; and features of the target protein, subcellular targeting, and posttranslational modifications (Streatfield 2007; Ghag et al. 2021). Next, the downstream processes must be cost effective, and the final product must be safe, sufficiently pure, and biologically active (Wilken and Nikolov 2012).

This chapter discusses factors involved in producing recombinant proteins using seeds as a platform. The advantages, limitations, and challenges of seed-based production systems are commented, as well as biosafety issues and general aspects associated with the purification of heterologous proteins from seeds.

6.2 Seeds as Bioreactors for Producing and Storing Recombinant Protein

Expressing a recombinant protein in seeds has some unique advantages. During development, endosperm cells are committed to producing and storing proteins and other nutrients (oils, starch, carbohydrates, etc.) to nurture the developing plantlet after germination (Li and Berger 2012). Thus, the target protein expressed in the seed finds a cellular environment that favors protein accumulation (Robinson et al. 2005; Khan et al. 2012).

Product yield and protein quality are critical factors for any recombinant protein production system. Yield involves the efficiency of biosynthesis (i.e., transcription and translation) and the stability of the recombinant protein in the cell (Chen et al. 2020; Liu and Timko 2022). Quality may include the correct assembly and posttranslational modifications, which are involved in protein turnover processes, impacting yield, and are likely necessary for the protein to retain its biological function (Vitale and Boston 2008; Thomas and Walmsley 2015; Strasser et al. 2021). This balance between translation and turnover will affect protein accumulation and involves features of the recombinant protein, the subcellular location where it is directed, and the metabolic burden associated with transcription and translation of this protein (Thomas and Walmsley 2015).

In leaves, very high transcriptional and translational levels of a heterologous gene can be achieved, particularly in transient expression assays (Fischer et al. 1999; Gleba et al. 2004; Pogue et al. 2010). This high expression of a foreign gene, nonessential, may impose a metabolic imbalance on the cell and activate endogenous protection mechanisms that can limit the expression and accumulation of the heterologous protein or direct the cell to apoptosis (Thomas and Walmsley 2015). In seeds, the transcription level may not be as high as in leaves. Still, it can be steadily maintained throughout the developing endosperm and embryo, reducing the metabolic burden on these cells (Boothe et al. 2010). That does not imply that a seed-based production system will produce any recombinant protein. According to the type of recombinant proteins, some may be poorly expressed while others may not even be detected. There will also be differences in the target protein level between

different transgenic lines expressing the same gene (Streatfield 2007; Hood et al. 2012). In any case, transgenic cereals and legume grains have been shown to accumulate high amounts of different types of recombinant proteins at consistent levels throughout different generations and batches (Hudson et al. 2014; Mirzaee et al. 2022). Seed production may take a couple of months. Still, a recombinant protein therein is steadily accumulated and stably maintained throughout seed development and can be stored for 4–6 years after the harvest, allowing scale-up production (Oakes et al. 2009; Boothe et al. 2010).

Once the recombinant protein is stably maintained in the cell, the final yield and purity will also depend on the extraction and purification processes (Menkhaus et al. 2004; Janson 2011; Wilken and Nikolov 2012). Optimizing each of these aspects is an effort to make molecular farming an increasingly attractive alternative for recombinant protein production.

6.3 Setting a Seed-Based Platform for Recombinant Protein Production

The production of recombinant proteins in seeds implies that the foreign candidate gene is stably integrated into the genome of a transgenic plant and that this gene is expressed in the seed and inherited by the progeny. Hence, a key factor in setting a seed-based platform for recombinant protein production is the ability to transform the candidate crop. For many species, the transformation protocol may be lengthy and cumbersome, even intractable, in some cases. However, the years-long plant transformation experience has resulted in relatively efficient protocols for species with high-protein content seeds, including cereals and grain legumes, such as rice, maize, barley, pea, and soybean.

To assure that the candidate gene is expressed in the seed, the usual approach includes a seed-specific promoter in the expression cassette used for transformation (Fig. 6.1). Several seed-specific promoters have been identified and tested, both for monocots and dicots (Furtado et al. 2009; Joshi et al. 2015; Xu et al. 2016; Mirzaee et al. 2022). Promoter sequences of storage proteins, for example, follow a tissue-specific pattern and are strictly regulated in time during embryogenesis (Chen et al. 1989). Endosperm-specific promoters that have been used for transgene expression in cereals include the rice globulin, prolamin, glutelin GluB-4, *Gt13a* maize zein, and barley D hordein (Kawakatsu and Takaiwa 2010). In legume seeds, some endosperm-specific promoters tested are the soybean β-conglycinin α′ subunit, the pea legumin, and the arcelin and phaseolin, from common bean (Chen et al. 1989; De Jaeger et al. 2002; Mirzaee et al. 2022).

Other features that have also been shown to influence the expression levels in transgenic plants include the presence of introns, enhancer sequences, codon optimization, terminator sequence, and other 3′ flanking regions, such as scaffold matrix

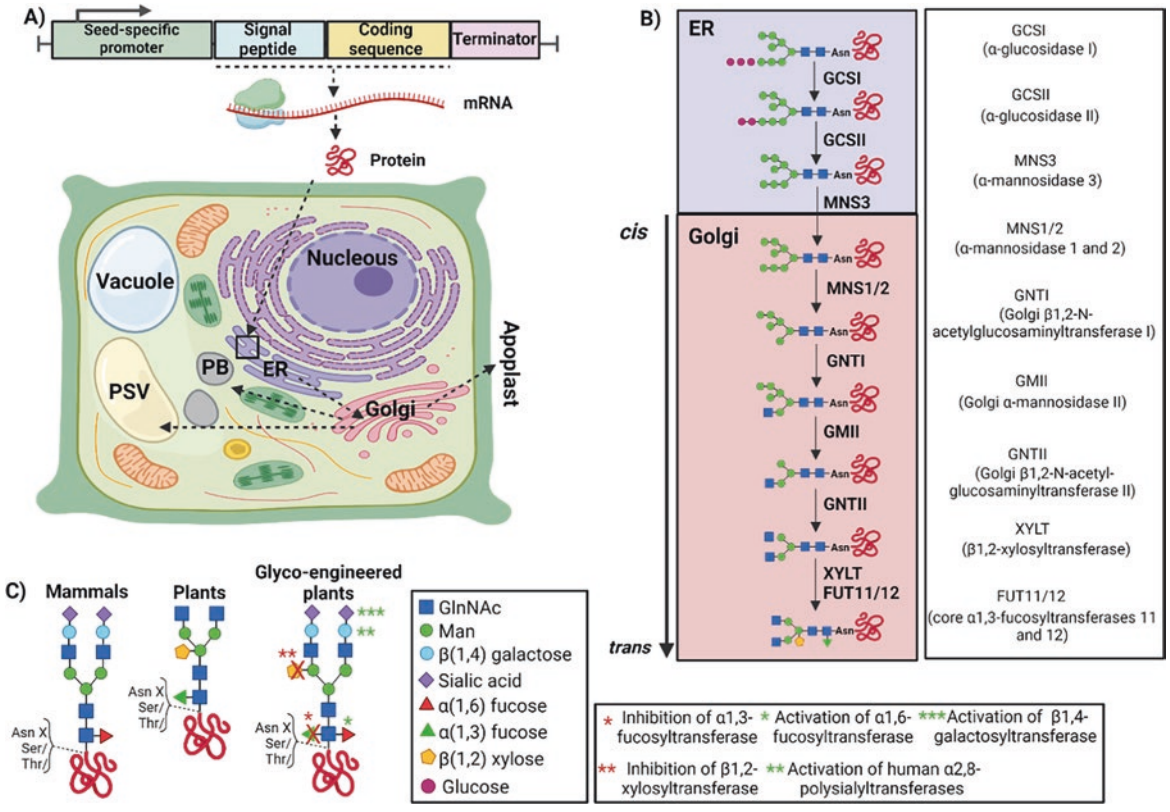

Fig. 6.1 Schematic overview of N-glycosylation in plants. (**a**) Following transcription in the nucleus, the protein is translated and enters the ER, where the glycosylation process begins. Next, the protein moves to the Golgi where it is further processed. The glycosylated protein may be secreted to the PSV, PB, or apoplast. (**b**) Glycosylation process in the ER and Golgi corresponding enzymes. (**c**) Differences in glycosylation pattern of plant and mammal proteins, and the approach for glycol-engineering protein in a transgenic plant

attachment regions (MARs) (Habibi et al. 2017; Webster et al. 2017; Diamos and Mason 2018).

Besides the promoter region, a cassette for protein expression in seeds includes a sequence for a signal peptide that will direct the protein to the endoplasmic reticulum (ER) and the secretory pathway (Fig. 6.1) (Arcalis et al. 2014). A signal peptide is a fragment of 20–30 amino acids present at the N- or C-terminal end of the target protein, which are recognized by specific complexes of RNA and proteins, called "signal recognition particles" (SRPs), that mediates the internalization of the target protein into the membranous organelles (Jolliffe et al. 2005; Robinson et al. 2005; Ashnest and Gendall 2018). Next, the signal peptide is cleaved, leaving the target protein at its intracellular destination (Bohnsack and Schleiff 2010). In the absence of signal peptides, the protein synthesized in free ribosomes accumulates in the cytoplasm, generally an unstable environment with high proteolytic activity (Obembe et al. 2011).

In seeds, storage proteins are directed to protein storage vacuoles (PSVs), which are derived from prevacuoles detached from the Golgi complex (Figs. 6.1 and 6.2). In *Poaceae*, which includes the cereals, the ER also forms protein bodies (PBs) that store mainly prolamin aggregates (Khan et al. 2012; Arcalis et al. 2014; Pedrazzini et al. 2016). PSVs are highly specialized vacuoles derived from the rough endoplasmic reticulum (Khan et al. 2012; Arcalis et al. 2014) and possibly from embryonic vacuoles (EVs), formed during seed development (Feeney et al. 2018). Their lumen has a pH close to neutral and practically no aminopeptidases, features that characterize them as a subcellular environment where protein degradation is minimal and an excellent target for addressing heterologous polypeptides (Takaiwa et al. 2007). In addition to providing a low oxidative environment, PSVs also harbor a high

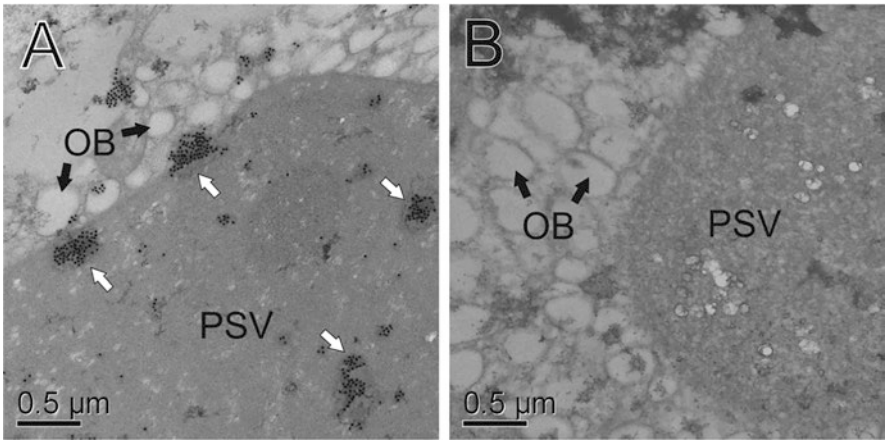

Fig. 6.2 Subcellular localization of recombinant FIX by immunocytochemistry in ultrathin sections of soybean cotyledons. (**a**) Immunogold-labeled FIX (white arrows) localizes to protein storage vacuoles (PSV) in transgenic soybean seeds. (**b**) Non-transgenic cotyledon. *OB* oil bodies. (Images: N.B. Cunha)

concentration of protease inhibitors, which increases their potential as a target for protein targeting aiming at increasing protein stability (Jolliffe et al. 2005; Oakes et al. 2009).

In cereal grains and legume seeds, the families of storage proteins represent the major part of the total seed protein content. For example, the family of glutelins in rice comprises 80% of the seed protein content; the glutelins, in wheat, 40%; and the zein, in maize, 60% (Kawakatsu and Takaiwa 2010). In soybeans, the globulins, glycinin, and β-conglycinin account for up to 80% of the total protein in the seed (Hudson et al. 2014). The expression of these storage proteins along the development of the seed is highly regulated and might constrain the accumulation of a heterologous protein.

A rebalancing of the seed storage protein of soybean was tested by partially suppressing the α/α′ subunit of β-conglycinin, resulting in the increased accumulation of glycinin, along with heterologous green fluorescent protein (GFP) regulated by glycinin promoter and terminator (Schmidt and Herman 2008). In a similar approach, Kim et al. (2014) found that the increase in the recombinant methionine-rich 11 kDa δ-zein in soybean was dependent on a sulfur-rich medium supplementation. In maize, Hood et al. (2012) crossed transgenic lines expressing a recombinant cellulase with high protein elite genotypes and selected for lines with higher content of cellulase. These approaches demonstrate that seed-based platforms have many possibilities to further optimize the production, both in yield and quality of recombinant proteins.

6.4 Posttranslational Modification in Plants and Its Relevance in Molecular Farming

As eukaryotic organisms, plants have the metabolic pathways for posttranslational modifications of proteins—glycosylation, acetylation, and phosphorylation, among others. These modifications occur in the ER and Golgi and are relevant for molecular farming (Fig. 6.1). It is estimated that 50–70% of human proteins are glycosylated (Walsh and Jefferis 2006). Likewise, about 50% of the biopharmaceuticals currently produced are represented by glycosylated proteins (Mizukami et al. 2018; Montero-Morales and Steinkellner 2018). Also, glycosylation and other posttranslational modifications are involved in protein stability and turnover, potentially impacting the accumulation and, hence, the recombinant protein's final yield (Thomas and Walmsley 2015; Varki 2017; Gupta and Shukla 2018).

The glycosylation pattern of proteins in plants, both N- and O-glycosylation, differs from that observed in insects, yeast, and animal cells. Glycosylated proteins from plants contain xylose, and arabinose residues (in *O*-glycans), which are not found in mammalian proteins. Moreover, plant *N*-glycans present α(1,3)-fucose, which is also present in mammalian cells but in β(1,6) linkage form. Plant glycans

lack galactose and terminal sialic acids, which are present in mammalian glycoproteins (Fig. 6.1) (Strasser et al. 2021; Bohlender et al. 2022).

In many cases, these differences in the glycosylation pattern may not interfere with the biological activity or the functionality of the recombinant protein, particularly in non-pharma proteins. However, glycostructures can influence the pharmacokinetics, stability, and immunogenicity for biopharmaceuticals (Gupta and Shukla 2018; Bohlender et al. 2022). For example, the presence of nonhuman glycans, particularly the fucose and xylose residues, may cause allergies and immunogenic responses in humans (Montero-Morales and Steinkellner 2018), and non-sialylated glycoproteins are rapidly cleared from serum (Walsh and Jefferis 2006; Bohlender et al. 2022).

To circumvent these problems, *Nicotiana benthamiana* plants were glycol-engineered to present a more "humanlike" glycosylation pattern (Fig. 6.1). Transgenic *N. benthamiana* plants expressing β(1-4)-galactosyltransferase were successfully tested. The α(1,3)-fucosyltransferase and β(1,2)-xylosyltransferase genes were knocked out in the moss *Physcomitrella*, in transgenic *Arabidopsis thaliana*, and *N. benthamiana*, which produced *N*-glycans lacking xylose and fucose residues (Strasser et al. 2004). Next, *N. benthamiana* plants were modified to express the α-1,6-fucosyltransferase and the pathways involved in the biosynthesis, activation, transport, and transfer of Neu5Ac to terminal galactose of heterologous proteins (Fig. 6.1) (Castilho et al. 2010, 2011; Kallolimath et al. 2016).

Expression platforms that are able to produce proteins with extensively modified glycosylation patterns, as for *N. benthamiana*, for example, have not yet been developed for cereals or legume seeds. However, in one attempt to engineer the glycosylation pattern, Wang et al. (2017) expressed the human α-1,6-fucosyltransferase (FUT8) in rice, controlled by an endosperm-specific promoter. After crossing with a plant expressing recombinant human 1-antitrypsin, they confirmed the presence of α(1,6)-fucose and a reduction of β(1,3)-fucose both in the recombinant protein and in globulins. Vamvaka et al. (2016b) demonstrated that recombinant heavy chain of the HIV-neutralizing monoclonal antibody 2G12 expressed in rice seeds was predominantly non-glycosylated, potentially less immunogenic, and more potent in HIV-neutralization assays than the 2G12 antibodies produced in *Nicotiana tabacum*. Similar results were reported by Zhang et al. (2012), which found that approximately 70% of their rice-derived recombinant human a1-anti-α-trypsin was aglycosylated. Indeed, plants appear to tolerate alterations in the glycosylation pathways, not showing phenotypical alterations and being well suited for the production of glyco-engineered recombinant proteins. That flexibility was further demonstrated by the transient co-expression in *N. benthamiana* of specific glycosyltransferases allowing the production of the glycoproteins omega-1 and kappa-5 of *Schistosoma mansoni* containing the helminth-like glycosylation pattern (Wilbers et al. 2017; see also chapter "Tobacco Plants as a Versatile Host for the Expression of Glycoproteins").

New available technologies, such as targeted genome editing, could be used to efficiently knock down β1,2-xylose and core α1,3-fucose residues, for example. In any case, developing a glycol-engineered seed-based platform is a promising yet challenging process (Buyel et al. 2021; see also Chaps. 3 and 4).

6.5 Comparison of Current Seed-Based Platforms

6.5.1 Cereals

Cereals are among the most cultivated and consumed crops worldwide, and most of them are considered a staple food for many countries. Currently, the main products obtained through the production of recombinant proteins in commercially available seed-based platforms are amylase, peroxidase, and cellobiohydrolase I (maize); growth factors and cytokines (barley); and a variety of enzymes (lactoferrin, albumin, transferrin, and lysozyme) and growth factors in rice (Fischer and Buyel 2020; Mirzaee et al. 2022).

Maize was the first seed-based platform used to produce an industrial reagent, avidin, by ProdiGene Inc. (USA). The product was indistinguishable from its counterpart from hen egg white and presented a high yield (2.3% of extractable protein from seed, on average) (Hood et al. 1997; Fischer and Buyel 2020; Moon et al. 2020).

Because maize is a cross-pollinating crop, working with transgenic maize is challenging and requires strict biosafety protocols to be followed. Despite that, maize has several advantages as a seed-based platform for the production of recombinant proteins, such as larger grain size, high yield, and lower production cost compared to other cereals. Moreover, maize has a higher endosperm proportion; a set of specific promoters for the seeds, which can be used to drive the expression of the transgene alone or in combinations; and easy genetic transformation, with established protocols (Watson and Ramstad 1987; Hood et al. 1997; Witcher et al. 1998). As a result, maize seeds have been used as a platform for the production of industrial reagents, such as enzymes and cosmetics, and also pharmaceuticals, such as antibodies to treat human and animal diseases (Rademacher et al. 2008; Egelkrout et al. 2020) and vaccines (Nahampun et al. 2015).

On the contrary, rice and barley are self-pollinating crops, making the risk of undesirable gene flow very low. The rice transformation system is effective in most varieties cultivated worldwide. However, some genotypes may be more suitable for producing a high level of recombinant proteins. The amount of proteins in rice seeds is 7–15% of the total seed weight (Takaiwa et al. 2007). The productivity of recombinant protein in relation to the total weight of the rice seed has reached high levels, from 1% of the total seed weight to 20% of dry seed weight (Vamvaka et al. 2016a). However, in some cases, the high content of recombinant proteins in rice seeds results in grains with an impaired phenotype, indicating that further research is required (Kusaba et al. 2003; Tada et al. 2003; Wakasa and Takaiwa 2013). Besides that, the ease of processing and scale-up production by well-established cultivation systems favor the choice of rice as a platform for producing recombinant proteins.

The rice seed-based platform has been explored for the production of biopharmaceuticals, such as vaccines, growth factors, and antiviral proteins (Takagi et al. 2005; Xie et al. 2008; Vamvaka et al. 2016a). As one of the main staple foods globally, it would be useful to exploit the rice seed-based platform to deliver pharmaceuticals as food. Some studies have shown that the recombinant protein produced in

rice seeds remains active in the seeds, even after processing them as a fine powder or crude extracts, for oral administration to mice or macaques (Takagi et al. 2005; Xie et al. 2008; Nochi et al. 2009). In one study, the oral administration of rice seeds expressing immunogens, processed as a fine powder, effectively inhibited allergy-associated immune responses in mice (Takagi et al. 2005). Another study developed a rice-based vaccine that expresses the B subunit of cholera toxin (CT), initially tested in mice with positive results. Subsequently, the vaccine was orally administered to macaques and induced CT-neutralizing IgG antibodies, confirming its effectiveness against cholera in nonhuman primates. This vaccine (MucoRice-CTB) has recently passed phase 1 human clinical trials (Nochi et al. 2009; Yuki et al. 2021).

Another interesting example was the use of crude extracts of rice seeds expressing a microbicide against human immunodeficiency virus (HIV) in cytotoxicity and antiviral assays with human cells (Vamvaka et al. 2016a). Results showed that the crude extracts had stronger binding activity to HIV than the wild-type rice seeds, similar to the purified protein, and were not toxic to human cell lines. Also, the crude extracts expressing the microbicide presented the same oligosaccharide-dependent binding properties as the same recombinant protein expressed in *Escherichia coli*. Altogether, these results show that it is possible to administer rice seeds expressing biopharmaceuticals without processing or with only minimal processing (see also chapter "The Use of Rice Seed as Bioreactor").

Barley is another cereal species that has been used as a seed-based platform for the production of recombinant proteins. The European regulatory agency (EFSA) has declared self-pollinating cereals such as wheat and barley as GRAS (generally recognized as safe) (Mirzaee et al. 2022; see also Chap. 14). In addition to being self-pollinating, barley has other characteristics of interest, such as its ability to regenerate, especially the cultivar Golden Promise, in which *transformation amenability* (TFA) alleles have been identified as responsible for its *Agrobacterium* transformation efficiency (Hisano and Sato 2016; Orman-Ligeza et al. 2020).

Other cereal seeds have been studied as a platform for producing heterologous proteins, however, on a smaller scale. For example, wheat seeds have been used to express TM-1 protein as an antigen to be used as an edible vaccine for chronic respiratory disease, a common disease in chickens, resulting in a significant level of protection (Shi et al. 2023).

6.5.2 Soybean

Among the plants that are candidates for seed-based stable accumulation of recombinant proteins, soybean plants, along with pea plants, present seeds with a high protein content (corresponding to approximately 40% of their weight) and, compared to other sources, represent a lower protein cost, due to their high seed yield (Mikschofsky and Broer 2012; Hudson et al. 2014; Vollmann 2016). Besides, recombinant proteins stored in soybean seeds were shown to remain stable and functional for long periods at room temperature (Oakes et al. 2009; Lobato Gómez

et al. 2021). A study showed that the production of a functional subunit vaccine for *Staphylococcal enterotoxin* B in soybean seeds was stable over several soybean generations, and biochemically and immunologically similar to commercial recombinant forms (Hudson et al. 2014). The expression of cyanovirin-N, a lectin with antiviral activity, was demonstrated in soybean seeds at levels up to 10% of total soluble protein. Attempts to express this protein in transient assays in leaves of *N. benthamiana* were unsuccessful, demonstrating the potential of soybean as an alternative to express and accumulate this recombinant protein (O'Keefe et al. 2015).

From a regulatory point of view, it has a reduced risk of pollen contamination since soybean is largely self-pollinated (Paul and Ma 2011; Paul et al. 2011). Furthermore, soybean seeds accumulate proteins in PSV, resulting in optimal conditions for long-term storage of immunogenic and fully active recombinant proteins (Fig. 6.2) (Cunha et al. 2011).

Although soybean has not been commercially used for the production of recombinant proteins yet, numerous studies show the potential of this plant species for the production of pharmaceutical proteins, such as human growth hormone, proinsulin antiviral lectins, coagulation factor IX, antigens, antibodies, as well as non-pharma proteins (Moravec et al. 2007; Yamada et al. 2008; Cunha et al. 2011, 2014; Hudson et al. 2014; O'Keefe et al. 2015). Some of these studies have tested the potential of soybean expressing recombinant proteins when administered orally to mice, as seed extracts, with promising results (Moravec et al. 2007; Hudson et al. 2014). This indicates that soybean seeds could be formulated into edible products for oral delivery of pharmaceutical proteins (Adelakun et al. 2013).

Another important characteristic of this legume is its sensitivity to the photoperiod, expressed in the temporal modulation of the vegetative phase of its phenological cycle as a function of daily time and light intensity (O'Keefe et al. 2015). The production of seeds per plant can be greatly increased under controlled conditions in a greenhouse reaching up to 1000 seeds per plant. This increase in the production scale can be exploited by the molecular farming industry and is particularly relevant for biocontainment and more controlled cultivation conditions (Kantolic and Slafer 2007; see also chapter "Legume Seed: A Useful Platform for the Production of Medical Proteins/Peptide").

6.6 Scale-Up of Seed-Based Production Systems

The commercial success of large-scale seed molecular agriculture depends on technology, economics, and public acceptance (see also Chap. 15). Factors important for the biopharmaceutical industry include the expected reduction in costs and the indirect effects on the biopharmaceutical market (Twyman et al. 2003). Seed-based platforms can be used to produce recombinant proteins at a significantly lower cost compared to other systems such as microbial fermentation and cell cultures (Giddings 2001; Hood et al. 2002; Twyman et al. 2003). It is estimated that the production costs of recombinant proteins in maize, for example, will be threefold

higher than that for the production of maize for food use (Mison and Curling 2000). Still, the savings in operational expenses can generate a considerable reduction in terms of capital investment (about 75–80%) and manufacture (50–60%) compared to microbial fermentation and cell culture production (Buyel 2019). In addition, the production of proteins that require lower purity, such as industrial enzymes and oral vaccines, is significantly reduced compared to other products that require expensive purification processes (Nikolov and Hammes 2002; see also chapters "Molecular Farming of Industrial Enzymes: Products and Applications" and "Plant Molecular Farming for Vaccine Development").

Field cultivation is the most reasonable option for the large-scale production of transgenic seeds. Although production in greenhouses may be feasible for transient transformations in short-cycle crops, generation time and space requirements would reduce the cost advantages of seed production (Boothe et al. 2010). The cost of production ($/g product) of the same recombinant protein produced in greenhouses is estimated as fivefold higher than that produced in an open-field system (Pogue et al. 2010). Besides, in terms of equipment for field production (from planting to harvesting), there is no difference between those used in conventional commodity crops. The existing technologies already meet this demand, enabling good scalability compared to other systems limited by the size of the culture reactor, for example. Furthermore, the long-term stability of the recombinant protein in the seeds allows the harvest to be decoupled from the purification process, generating greater flexibility and better stock management (Boothe et al. 2010).

Another important aspect of scale production is a high level of quality control to ensure high protein purity. Therefore, crop management must be carefully conducted, which is also necessary for biosafety reasons. Crops must provide high-quality seeds, which must present stable expression levels, and homogeneous products (e.g., glycosylation pattern and degradation levels) over generations and cultivation places to meet the established quality standards.

Finally, approximately 60–70% of the production cost of seed-based recombinant proteins is associated with downstream processing, so it is essential to develop techniques to improve it and make it increasingly efficient, reducing its cost (Dyr and Suttnar 1997). Therefore, understanding the conditions that affect the product's quantity and quality is essential to meet the market standards and competitiveness. An overview of the methods currently used for protein purification is presented in the following section.

6.7 Basic Approaches for the Purification of Recombinant Proteins from Plant Cells

The choice of an expression system should consider the particularities of the target protein to be purified. Expression based on the bacteria *Escherichia coli*, for example, could be advantageous in terms of production time and costs (Rosano and

Ceccarelli 2014; Lozano Terol et al. 2021). However, the lack of posttranslational modifications and contamination with bacterial endotoxins represent limitations for using this popular expression system (Sahdev et al. 2008). On the other hand, expression based on mammalian, yeast, and insect cells allows posttranslational modifications. However, these expression systems based on large bioreactors present a high upstream cost (Schillberg and Finnern 2021).

The production of biomass from plants requires soil and fertilizers, whereas bioreactor cultures demand complex media. That makes plant-based biomass production more sustainable than microbe and cell culture expression systems (Buyel et al. 2021), requiring fewer investments and demanding less specialized staff.

Downstream processes, independently of the expression system utilized, are critical in terms of cost and quality of the final product. It is estimated that downstream processing may represent 50–80% of the total costs, depending on the yield, recovery efficiency, and purity grade (Schillberg et al. 2019). Besides the costs, if purification of the recombinant protein is needed, one may be facing a technical challenge. Indeed, purification of a specific protein, either recombinant or endogenous, is generally not trivial, as each protein will have its specific physical characteristics. Plus, in the case of a recombinant protein, there will likely be no available protocol for purification from plant tissues. However, information on the purification method of the target protein from other expression systems may be helpful. Some factors do favor the downstream processing and purification of proteins from seeds since they are presented as a comparatively homogeneous starting material, have reduced water content, lack chlorophyll pigments and alkaloids, and confer high stability of the target protein stored in the seed storage vacuoles.

The source of the material to be processed needs to be handled carefully—plants need to be well cultivated so that the collected tissues are healthy. Next, some basic protein analyses are needed to confirm the presence of the heterologous protein and evaluate the expression level. These can be done from total extracts (i.e., tissues ground in extraction buffer), using well-established protocols for total protein quantification and detection, such as Western blot or ELISA. Finally, the amount of the recombinant protein can be expressed as a percentage of the total protein in the extract, or of the seed dry weight, as weight per mass of fresh or dry tissue (e.g., mg/g fresh leaves, μg/g dry seed weight).

The seed composition varies among species and requires specific extraction conditions regarding the contents of proteins, starch, oil, etc. For soybean seeds, for example, due to large amounts of oil (about 18–22% of seed weight), the material often requires homogenizing the seed meal with solvents, such as hexane. Furthermore, to ensure that the protein of interest maintains its stability and that its extraction process is efficient, it is essential that the extraction solutions present the appropriate pH and ionic strength (Robić et al. 2010). Hence, the extraction conditions consider pH, saline buffers, chaotropic agents/detergents, protease inhibitors, etc. (Janson 2011). Once extracted from the seeds, the target protein may be further purified by chromatography.

6.7.1 *Non-affinity Absorption and Affinity Techniques to Purify Proteins*

Proteins are made of amino acids (AAs) as basic building blocks assembled in a chain via amide bonds (peptide linkages). The 20 L-AAs found in proteins have four different ligand groups (an amino group, carboxyl group, hydrogen, and R-group). These groups and their interactions within the protein give unique biochemical characteristics and functions and influence their physiological and biological activities. The R-group of the 20 AA commonly found in proteins varies widely, especially their polarity at a biological pH (around pH 7.0), from polar and hydrophilic (water soluble) to nonpolar and hydrophobic (water insoluble) (Wu 2009). These physical properties can be exploited to aid the protein isolation and purification process.

In aqueous solutions, functional AA groups from folded proteins contribute significantly to the protein surface charge in a pH-dependent way. According to external pH, the overall charge may vary from positive (at low pH) to negative (at high pH). Therefore, separating a complex sample of proteins based on their surface charge helps purify a protein-rich sample with similar physicochemical characteristics in a reduced volume compared to the initial volume before separation (Bonner 2018). The purification may be optimized by using various techniques to exploit differences in the target protein's charge and biospecificity. Several successful cases in the literature explore these methodologies for efficient protein separation from seeds. For example, Zhang et al. (2012) purified the human alpha-antitrypsin protein expressed in transgenic rice seed using different anion exchange columns. In another work, cellobiohydrolase I was expressed in transgenic corn seeds and purified with ammonium sulfate precipitation (a fractionation technique also used to isolate proteins from a complex sample) together with both cationic and anionic exchange chromatography, yielding 63% of pure protein (Hood et al. 2014).

It is a common strategy to engineer the recombinant protein with affinity tags to facilitate affinity-based purification procedures. The most common examples found in the literature are proteins containing polyhistidine tails (6×His or 10×His) (Valdez-Ortiz et al. 2005). The sequence of six (or ten) consecutive histidine residues is currently one of the most used strategies worldwide to purify recombinant proteins for biochemical and structural studies.

Unlike ion exchange chromatography, affinity chromatography does not explore the physicochemical characteristics of proteins. The affinity chromatography technique is based precisely on the unique biospecificity of the protein engineered for isolation from a complex sample of proteins. Biospecificity involves the interaction between two immiscible phases, that is, the reversible interaction between a ligand (which can be a small molecule, enzymes, among others) immobilized on a resin (known as the stationary phase) and the recombinant target protein inserted into a solute (mobile phase) (Janson 2011). Menkhaus et al. (2004) compared various techniques, such as precipitation with polyethyleneimine cationic polyelectrolyte (PEI), anion/cation exchange, diafiltration (molecular exclusion), and immobilized

metal ion affinity chromatography (IMAC) for the purification of histidine-tagged β-glucuronidase from transgenic pea seeds. They observed an increased recovery of pure protein and higher enzyme activity when utilizing affinity chromatography (Menkhaus et al. 2004).

In some cases, proteins from the host can be present as contaminants in samples purified by affinity purification. To remove these contaminants, denaturation and refolding steps of proteins are sometimes necessary. Fujiwara et al. (2010) observed the need for two steps of affinity purification combined with a denaturation step with 6 M guanidine to remove protein contaminants from rice seed in the purification of IL-4 and -6 cytokines (Fujiwara et al. 2016). Similar results were previously obtained by Fujiwara et al. (2010) during the purification of human interleukin-10 (IL-10), also expressed in rice seeds. These results demonstrate the efficiency of combining different chromatographic and purification techniques to remove contaminants and consequently increase the yield and purity of the recombinant proteins obtained (Fujiwara et al. 2010).

6.7.2 Chromatography-Free Protein Purification

Another approach for recombinant protein purification is based on fusion proteins (FPs). The idea is to exploit the unique properties of the fusion partner, allowing an increase in stability and facilitating the purification process (Viana et al. 2013; Ki and Pack 2020). Examples of these fusion partners include the synthetic peptide elastin-like polypeptides (ELPs) (Ciofani et al. 2014), γ-zein (Torrent et al. 2009), and hydrophobin (Lahtinen et al. 2008).

Derived from its soluble precursor, tropoelastin, elastin has a hydrophobic motif composed of a repeated sequence of hydrophobic amino acids alanine (Ala) and valine (Val), in addition to the presence of other residues in significant amounts, such as glycine (Gly) and proline (Pro) (Partridge et al. 1955). At temperatures below 25 °C, the protein remains soluble; however, when the temperature is raised to 37 °C, a precipitated protein known as coacervate is observed. This process is fully reversible upon returning the protein to room temperature (Urry et al. 1969). Based on these properties, synthetic peptides were developed, known as the elastin-like polypeptides (ELPs), composed of the canonical sequence of the pentapeptide (Val-Pro-Gly-Xaa-Gly)*n*.

When investigating the strategy of using FP in conjunction with scalable purification processes, Phan et al. (2014) observed an enhancement in expression levels (about tenfold higher) of the avian influenza virus (H5N1) hemagglutinin subtype 5 (H5) protein fused with ELP at the C-terminus in transgenic tobacco seeds. The enhanced accumulation of HA, which is the major antigen of the influenza virus, by ELP-FP resulted in high concentrations of the ELPylated target protein in the aqueous crude extract. Further purification was facilitated by using optimized processes involving inverse transition cycling (ITC). In comparison with another strategy, also using FP (fungal hydrophobin I—HFBI), the same authors observed that only

ELPylation was able to increase HA expression in seeds, resulting in high-purity protein (Phan et al. 2014), demonstrating the efficiency of using ELPylated proteins in the processes of expression and purification simply and inexpensively (Khan et al. 2012).

In contrast, in another study, Yang et al. (2021) showed that the γ-zein system, a member of the major prolamin storage family in maize, was more efficient for the accumulation of GFP in immature soybean seeds than the ELP system. In addition, the use of the γ-zein system provided a 3.9-fold increase in the accumulation of fused GFP in comparison with unfused GFP protein, demonstrating that the γ-zein system is a promising FP for future enhancement in the expression and purification of recombinant proteins in plants.

Although fusion proteins are generally non-immunogenic and biologically compatible, they may interfere with the activity of native proteins (Shamji et al. 2007; Viana et al. 2013). That implies the removal of the FP, which is done after purification by specific proteolytic enzymes that recognize cleavage sites placed at the junction of the target protein and the fusion partner. This additional step for recombinant protein purification may result in unspecific degradations of the target protein, reducing the final yield and impacting downstream costs (Tian and Sun 2011).

A promising alternative to proteolytic enzymes for the removal of FP is the use of inteins (Viana et al. 2013; Ki and Pack 2020). These proteins can catalyze their self-cleavage and, through amino acid substitution, can be regulated to cleave at either the N- or the C-terminus in response to reducing agents or changes in solution pH (Xu and Perler 1996; Perler 1998; Gillies et al. 2009). Therefore, the self-cleavage property of inteins can be applied to replace the traditional proteolytic cleavage. By fusing the intein (Eitag) with the ELP-FP, for example, the recombinant proteins can be purified by applying both ITC followed by autocatalysis by changing the pH of the solution. Tian and Sun (2011) explored the use of the ELP-intein system to increase the accumulation of the recombinant lectin fused with ELP in transgenic rice and tested the capacity of autocatalysis of intein after ELP extraction from seeds. Furthermore, the presence of Eitag + ELP did not alter the N-glycosylation patterns of the recombinant protein, demonstrating the potential application of the ELP-intein fusion system for the expression and purification of recombinant proteins in plants, especially in seeds.

6.8 A Brief Overview of Biosafety and Risk Assessment of Seed-Based Expression Systems

The technology for the production of recombinant proteins using these platforms is developing fast and focuses on two product lines: pharmaceuticals and non-pharmaceuticals. Despite intense research on developing biopharmaceutical production, most plant-based products currently available on the market belong to the non-pharma field, mainly because regulation processes are faster and less expensive

for non-pharma products. These include products for the diagnosis, industrial reagents, and cosmetics, among others.

The biological safety assessment to produce recombinant proteins in seeds is an important issue. Biosafety involves several relevant issues, such as choice of plant platform, transgenic plants, field production, handling, harvesting, and transport. Therefore, one must consider plant biology from the perspective of productivity and how it impacts the environment, food security, and human health. Therefore, the best material based on technical aspects (e.g., seeds with better processing capacity, high protein content, stability, etc.) may not be the best choice if considering the regulatory issues of biosafety (Sparrow and Twyman 2009).

The use of seeds as a "bioreactor" has its risks of propagation in nature, contaminating non-transgenic plants, and being potentially hazardous to people and animals if used unintentionally as food and for insects and soil microorganisms (Lee et al. 2003). Most of the steps required to avoid mixing these seed-derived biopharmaceuticals in the food chain are relatively simple, such as meticulous planning and execution. The plants must be cultivated in an isolated area to avoid genetic and mechanical mixing of seeds containing biopharmaceuticals with those intended for food. Likewise, small-scale and large-scale field trials must be isolated from conventional practices with crops to avoid cross-pollination. Although these risks apply to all transgenic crops grown in the field, plants cultivated for molecular farming deserve special attention due to the nature of the recombinant proteins (i.e., biopharmaceuticals), consequently, with unpredictable, potentially hazardous outcomes (Basaran and Rodríguez-Cerezo 2008). Achieving an effective isolation level to avoid wind and insect pollination is challenging. If the plants are cultivated in confinement, the risks and threats to the environment would be reduced and would imply less strict regulatory issues.

Appropriate mitigation measures for recombinant protein-producing seeds will depend on several factors, including properties of the molecule, biology of the crop, and characteristics of the environment where it is being produced. Approaches of containment methods include identity preservation—using varieties that are visually distinct from traditional varieties (such as purple maize or black soybeans)—and use of marker genes, such as a fluorescent protein, barrier crops, and temporal barriers aiming to minimize undesirable crosses (Sparrow and Twyman 2009).

In general, regulatory guidelines for the production of recombinant proteins are similar across countries, but some specificities may apply. For example, in the United States, the production of biopharmaceuticals on transgenic plants is regulated by two agencies. The US Department of Agriculture (USDA), Animal and Plant Health Inspection Service (APHIS), focuses on the containment of these seed producers of protein. In contrast, the Food and Drug Administration (FDA) focuses on the manufacture of the drug or vaccine. APHIS reviews production license applications, assessing probable environmental impacts of these releases (Basaran and Rodríguez-Cerezo 2008). In the European Union, authorizations involve all member states and the European Commission (Breyer et al. 2009). In Brazil, GMO studies are only allowed in research institutions after authorization by the National Biosafety Technical Commission (CTNBio) (Mendonça-Hagler et al. 2008).

6.9 Conclusions and Perspectives

The demand for biopharmaceuticals continues to grow as new products are approved. To couple with this demand, a general trend for the production of recombinant proteins has been to increase yield and optimize upstream and downstream processes to reduce costs. In addition, bioequivalence is also very relevant for the biopharmaceutical industry, which pursues products that are as similar as possible to the original product.

As discussed in this chapter, several aspects of plant molecular farming are aligned with these demands. Producing recombinant protein in plants demands substantially lower costs for upstream processes as compared to the expression systems based on microorganisms and cell cultures. Concerning posttranslational modifications, efforts for glyco-engineering of plants have also achieved amazing progress. Although still restricted to model plants and transient assays using *N. benthamiana*, efforts toward a seed-based glyco-engineered platform are in progress. It is reasonable to consider its viability and availability before long.

As experience accumulates, methods for protein purification will become more efficient. These may compensate for limitations on the expression level, which can be very low, depending on the target protein. For seed-based expression, the final yield could be increased by applying more efficient extraction and purification, involving, for example, fusion proteins and fusion tags. Seed-based platforms, for their advantages in terms of protein content and long-term stability of the recombinant protein stored therein, offer great potential for new ideas to be implemented (see also Chap. 5). The progress witnessed in the last decades confirms the potential of molecular farming as an alternative system for expressing recombinant proteins and represents a field of opportunities.

References

Adelakun OE, Duodu KG, Buys E, Olanipekun BF (2013) Potential use of soybean flour (*Glycine max*) in food fortification. In: El-Shemy HA (ed) Soybean. Bio-active compounds. IntechOpen, Rijeka, Croatia, pp 513–520

Arcalis E, Ibl V, Peters J, Melnik S, Stoger E (2014) The dynamic behavior of storage organelles in developing cereal seeds and its impact on the production of recombinant proteins. Front Plant Sci 5:439

Ashnest JR, Gendall AR (2018) Trafficking to the seed protein storage vacuole. Funct Plant Biol 45:895–910

Basaran P, Rodríguez-Cerezo E (2008) Plant molecular farming: opportunities and challenges. Crit Rev Biotechnol 28:153–172

Bohlender LL, Parsons J, Hoernstein SNW, Bangert N, Rodriguez-Jahnke F et al (2022) Unexpected *Arabinosylation* after humanization of plant protein N-glycosylation. Front Bioeng Biotechnol 10:1–12

Bohnsack MT, Schleiff E (2010) The evolution of protein targeting and translocation systems. Biochim Biophys Acta 1803:1115–1130

Bonner P (2018) Protein purification, 2nd edn. Taylor & Francis, London, UK

Boothe J, Nykiforuk C, Shen Y, Zaplachinski S, Szarka S et al (2010) Seed-based expression systems for plant molecular farming. Plant Biotechnol J 8:588–606
Breyer D, Goossens M, Herman P, Sneyers M (2009) Biosafety considerations associated with molecular farming in genetically modified plants. J Med Plant Res 3:825–838
Buyel JF (2019) Plant molecular farming - integration and exploitation of side streams to achieve sustainable biomanufacturing. Front Plant Sci 9:1–17
Buyel JF, Stoger E, Bortesi L (2021) Targeted genome editing of plants and plant cells for biomanufacturing. Transgenic Res 30:401–426
Castilho A, Strasser R, Stadlmann J, Grass J, Jez J et al (2010) *In planta* protein sialylation through overexpression of the respective mammalian pathway. J Biol Chem 285:15923–15930
Castilho A, Gattinger P, Grass J, Jez J, Pabst M et al (2011) N-glycosylation engineering of plants for the biosynthesis of glycoproteins with bisected and branched complex N-glycans. Glycobiology 21:813–823
Chen ZL, Naito S, Nakamura I, Beachy RN (1989) Regulated expression of genes encoding soybean beta-conglycinins in transgenic plants. Dev Genet 10:112–122
Chen Q, Yu F, Xie Q (2020) Insights into endoplasmic reticulum-associated degradation in plants. New Phytol 226:345–350
Ciofani G, Genchi GG, Mattoli V, Mazzolai B, Bandiera A (2014) The potential of recombinant human elastin-like polypeptides for drug delivery. Expert Opin Drug Deliv 11:1507–1512
Cunha NB, Murad AM, Cipriano TM, Araújo ACG, Aragão FJL et al (2011) Expression of functional recombinant human growth hormone in transgenic soybean seeds. Transgenic Res 20:811–826
Cunha NB, Murad A, Vianna G, Rech E (2014) Recombinant biosynthesis of functional human growth hormone and coagulation factor IX in transgenic soybean seeds. BMC Proc 8:P112
De Jaeger G, Scheffer S, Jacobs A, Zambre M, Zobell O et al (2002) Boosting heterologous protein production in transgenic dicotyledonous seeds using *Phaseolus vulgaris* regulatory sequences. Nat Biotechnol 20:1265–1268
Diamos AG, Mason HS (2018) Chimeric 3′ flanking regions strongly enhance gene expression in plants. Plant Biotechnol J 16:1971–1982
Dyr JE, Suttnar J (1997) Separation used for purification of recombinant proteins. J Chromatogr B Biomed Sci Appl 699:383–401
Egelkrout E, Hayden C, Fake G, Keener T, Arruda P et al (2020) Oral delivery of maize-produced porcine epidemic diarrhea virus spike protein elicits neutralizing antibodies in pigs. Plant Cell Tissue Organ Cult 142:79–86
Feeney M, Kittelmann M, Menassa R, Hawes C, Frigerio L (2018) Protein storage vacuoles originate from remodeled preexisting vacuoles in *Arabidopsis thaliana*. Plant Physiol 177:241–254
Fischer R, Buyel JF (2020) Molecular farming - the slope of enlightenment. Biotechnol Adv 40:107519
Fischer R, Vaquero-Martin C, Sack M, Drossard J, Emans N et al (1999) Towards molecular farming in the future: transient protein expression in plants. Biotechnol Appl Biochem 30:113–116
Fujiwara Y, Aiki Y, Yang L, Takaiwa F, Kosaka A et al (2010) Extraction and purification of human interleukin-10 from transgenic rice seeds. Protein Expr Purif 72:125–130
Fujiwara Y, Yang L, Takaiwa F, Sekikawa K (2016) Expression and purification of recombinant mouse interleukin-4 and -6 from transgenic rice seeds. Mol Biotechnol 58:223–231
Furtado A, Henry RJ, Pellegrineschi A (2009) Analysis of promoters in transgenic barley and wheat. Plant Biotechnol J 7:240–253
Ghag SB, Adki VS, Ganapathi TR, Bapat VA (2021) Plant platforms for efficient heterologous protein production. Biotechnol Bioprocess Eng 26:546–567
Giddings G (2001) Transgenic plants as protein factories. Curr Opin Biotechnol 12:450–454
Gillies AR, Mahmoud RB, Wood DW (2009) PHB-intein-mediated protein purification strategy. Methods Mol Biol 498:173–183
Gleba Y, Marillonnet S, Klimyuk V (2004) Engineering viral expression vectors for plants: the 'full virus' and the 'deconstructed virus' strategies. Curr Opin Plant Biol 7:182–188

Gupta SK, Shukla P (2018) Glycosylation control technologies for recombinant therapeutic proteins. Appl Microbiol Biotechnol 102:10457–10468
Habibi P, Prado GS, Pelegrini PB, Hefferon KL, Soccol CR et al (2017) Optimization of inside and outside factors to improve recombinant protein yield in plant. Plant Cell Tissue Organ Cult 130:449–467
Hisano H, Sato K (2016) Genomic regions responsible for amenability to *Agrobacterium*-mediated transformation in barley. Sci Rep 6:37505
Hood EE, Witcher DR, Maddock S, Meyer T, Baszczynski C et al (1997) Commercial production of avidin from transgenic maize characterization of transformant, production, processing, extraction and purification. Mol Breed 3:291–306
Hood EE, Woodard SL, Horn ME (2002) Monoclonal antibody manufacturing in transgenic plants - myths and realities. Curr Opin Biotechnol 13:630–635
Hood EE, Devaiah SP, Fake G, Egelkrout E, Teoh KT et al (2012) Manipulating corn germplasm to increase recombinant protein accumulation. Plant Biotechnol J 10:20–30
Hood NC, Hood KR, Woodard SL, Devaiah SP, Jeoh T et al (2014) Purification and characterization of recombinant Cel7A from maize seed. Appl Biochem Biotechnol 174:2864–2874
Hudson LC, Garg R, Bost KL, Piller KJ (2014) Soybean seeds: a practical host for the production of functional subunit vaccines. Biomed Res Int 2014:340804–340804
Janson J-C (2011) Protein purification: principles, high resolution methods, and applications, 3rd edn. Wiley, Hoboken, NJ
Jolliffe NA, Craddock CP, Frigerio L (2005) Pathways for protein transport to seed storage vacuoles. Biochem Soc Trans 33:1016–1018
Joshi JB, Geetha S, Singh B, Kumar KK, Kokiladevi E et al (2015) A maize alpha-zein promoter drives an endosperm-specific expression of transgene in rice. Physiol Mol Biol Plants 21:35–42
Kallolimath S, Castilho A, Strasser R, Grünwald-Gruber C, Altmann F et al (2016) Engineering of complex protein sialylation in plants. Proc Natl Acad Sci U S A 113:9498–9503
Kantolic AG, Slafer GA (2007) Development and seed number in indeterminate soybean as affected by timing and duration of exposure to long photoperiods after flowering. Ann Bot 99:925–933
Kawakatsu T, Takaiwa F (2010) Cereal seed storage protein synthesis: fundamental processes for recombinant protein production in cereal grains. Plant Biotechnol J 8:939–953
Kesik-Brodacka M (2018) Progress in biopharmaceutical development. Biotechnol Appl Biochem 65:306–322
Khan I, Twyman RM, Arcalis E, Stoger E (2012) Using storage organelles for the accumulation and encapsulation of recombinant proteins. Biotechnol J 7:1099–1108
Ki MR, Pack SP (2020) Fusion tags to enhance heterologous protein expression. Appl Microbiol Biotechnol 104:2411–2425
Kim W-S, Jez JM, Krishnan HB (2014) Effects of proteome rebalancing and sulfur nutrition on the accumulation of methionine rich δ-zein in transgenic soybeans. Front Plant Sci 5:633
Kusaba M, Miyahara K, Iida S, Fukuoka H, Takano T et al (2003) Low glutelin content1: a dominant mutation that suppresses the glutelin multigene family via RNA silencing in rice. Plant Cell 15:1455–1467
Lahtinen T, Linder MB, Nakari-Setälä T, Oker-Blom C (2008) Hydrophobin (HFBI): a potential fusion partner for one-step purification of recombinant proteins from insect cells. Protein Expr Purif 59:18–24
Lee RWH, Pool AN, Ziauddin A, Lo RYC, Shewen PE et al (2003) Edible vaccine development: stability of *Mannheimia haemolytica* A1 leukotoxin 50 during post-harvest processing and storage of field-grown transgenic white clover. Mol Breed 11:259–266
Li J, Berger F (2012) Endosperm: food for humankind and fodder for scientific discoveries. New Phytol 195:290–305
Liu H, Timko MP (2022) Improving protein quantity and quality-the next level of plant molecular farming. Int J Mol Sci 23:1326

Lobato Gómez M, Huang X, Alvarez D, He W, Baysal C et al (2021) Contributions of the international plant science community to the fight against human infectious diseases – Part 1: Epidemic and pandemic diseases. Plant Biotechnol J 19:1901–1920
Lozano Terol G, Gallego-Jara J, Sola Martínez RA, Martínez Vivancos A, Cánovas Díaz M et al (2021) Impact of the expression system on recombinant protein production in *Escherichia coli* BL21. Front Microbiol 12:682001
Mendonça-Hagler L, Souza L, Aleixo L, Oda L (2008) Trends in biotechnology and biosafety in Brazil. Environ Biosaf Res 7:115–121
Menkhaus TJ, Pate C, Krech A, Glatz CE (2004) Recombinant protein purification from pea. Biotechnol Bioeng 86:108–114
Mikschofsky H, Broer I (2012) Feasibility of *Pisum sativum* as an expression system for pharmaceuticals. Transgenic Res 21:715–724
Mirzaee M, Osmani Z, Frébortová J, Frébort I (2022) Recent advances in molecular farming using monocot plants. Biotechnol Adv 58:107913
Mison D, Curling J (2000) The industrial production costs of recombinant therapeutic proteins expressed in transgenic corn. BioPharm 13:48–54
Mizukami A, Caron AL, Picanço-Castro V, Swiech K (2018) Platforms for recombinant therapeutic glycoprotein production. Methods Mol Biol 1674:1–14
Montero-Morales L, Steinkellner H (2018) Advanced plant-based glycan engineering. Front Bioeng Biotechnol 6:81
Moon K-B, Park J-S, Park Y-I, Song I-J, Lee H-J et al (2020) Development of systems for the production of plant-derived biopharmaceuticals. Plants 9:30
Moravec T, Schmidt MA, Herman EM, Woodford-Thomas T (2007) Production of *Escherichia coli* heat labile toxin (LT) B subunit in soybean seed and analysis of its immunogenicity as an oral vaccine. Vaccine 25:1647–1657
Nahampun HN, Bosworth B, Cunnick J, Mogler M, Wang K (2015) Expression of H3N2 nucleoprotein in maize seeds and immunogenicity in mice. Plant Cell Rep 34:969–980
Nikolov Z, Hammes D (2002) Production of recombinant proteins from transgenic crops. In: Hood EE, Howard JA (eds) Plants as factories for protein production. Springer, Dordrecht, Netherlands, pp 159–174
Nochi T, Yuki Y, Katakai Y, Shibata H, Tokuhara D et al (2009) A rice-based oral cholera vaccine induces macaque-specific systemic neutralizing antibodies but does not influence pre-existing intestinal immunity. J Immunol 183:6538–6544
O'Keefe BR, Murad AM, Vianna GR, Ramessar K, Saucedo CJ et al (2015) Engineering soya bean seeds as a scalable platform to produce cyanovirin-N, a non-ARV microbicide against HIV. Plant Biotechnol J 13:884–892
Oakes JL, Bost KL, Piller KJ (2009) Stability of a soybean seed-derived vaccine antigen following long-term storage, processing and transport in the absence of a cold chain. J Sci Food Agric 89:2191–2199
Obembe OO, Popoola JO, Leelavathi S, Reddy SV (2011) Advances in plant molecular farming. Biotechnol Adv 29:210–222
Orman-Ligeza B, Harwood W, Hedley PE, Hinchcliffe A, Macaulay M et al (2020) TRA1: a locus responsible for controlling *Agrobacterium*-mediated transformability in barley. Front Plant Sci 11:355
Partridge SM, Davis HF, Adair GS (1955) The chemistry of connective tissues. 2. Soluble proteins derived from partial hydrolysis of elastin. Biochem J 61:11–21
Paul M, Ma JKC (2011) Plant-made pharmaceuticals: leading products and production platforms. Biotechnol Appl Biochem 58:58–67
Paul M, van Dolleweerd C, Drake PMW, Reljic R, Thangaraj H et al (2011) Molecular Pharming: future targets and aspirations. Hum Vaccin 7:375–382
Pedrazzini E, Mainieri D, Marrano CA, Vitale A (2016) Where do protein bodies of cereal seeds come from? Front Plant Sci 7:1139

Perler FB (1998) Protein splicing of inteins and hedgehog autoproteolysis: structure, function, and evolution. Cell 92:1–4

Phan HT, Hause B, Hause G, Arcalis E, Stoger E et al (2014) Influence of elastin-like polypeptide and hydrophobin on recombinant hemagglutinin accumulations in transgenic tobacco plants. PLoS One 9:e99347

Pogue GP, Vojdani F, Palmer KE, Hiatt E, Hume S et al (2010) Production of pharmaceutical-grade recombinant aprotinin and a monoclonal antibody product using plant-based transient expression systems. Plant Biotechnol J 8:638–654

Rademacher T, Sack M, Arcalis E, Stadlmann J, Balzer S et al (2008) Recombinant antibody 2G12 produced in maize endosperm efficiently neutralizes HIV-1 and contains predominantly single-GlcNAc N-glycans. Plant Biotechnol J 6:189–201

Robić G, Farinas CS, Rech EL, Miranda EA (2010) Transgenic soybean seed as protein expression system: aqueous extraction of recombinant β-glucuronidase. Appl Biochem Biotechnol 160:1157–1167

Robinson DG, Oliviusson P, Hinz G (2005) Protein sorting to the storage vacuoles of plants: a critical appraisal. Traffic 6:615–625

Rosano GL, Ceccarelli EA (2014) Recombinant protein expression in *Escherichia coli*: advances and challenges. Front Microbiol 5:1–17

Sahdev S, Khattar SK, Saini KS (2008) Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. Mol Cell Biochem 307:249–264

Schillberg S, Finnern R (2021) Plant molecular farming for the production of valuable proteins - critical evaluation of achievements and future challenges. J Plant Physiol 258–259:153359

Schillberg S, Raven N, Spiegel H, Rasche S, Buntru M (2019) Critical analysis of the commercial potential of plants for the production of recombinant proteins. Front Plant Sci 10:720

Schmidt MA, Herman EM (2008) Proteome rebalancing in soybean seeds can be exploited to enhance foreign protein accumulation. Plant Biotechnol J 6:832–842

Shamji MF, Betre H, Kraus VB, Chen J, Chilkoti A et al (2007) Development and characterization of a fusion protein between thermally responsive elastin-like polypeptide and interleukin-1 receptor antagonist: sustained release of a local anti-inflammatory therapeutic. Arthritis Rheum 56:3650–3661

Shi Y, Habibi P, Haq ANU, Saeed M, Gulghutay Amjad N et al (2023) Seed-based system for cost-effective production of vaccine against chronic respiratory disease in chickens. Mol Biotechnol 65:570–580

Sparrow PA, Twyman RM (2009) Biosafety, risk assessment and regulation of plant-made pharmaceuticals. Methods Mol Biol 483:341–353

Strasser R, Altmann F, Mach L, Glössl J, Steinkellner H (2004) Generation of *Arabidopsis thaliana* plants with complex N-glycans lacking β1,2-linked xylose and core α1,3-linked fucose. FEBS Lett 561:132–136

Strasser R, Seifert G, Doblin MS, Johnson KL, Ruprecht C et al (2021) Cracking the "sugar code": a snapshot of n- and o-glycosylation pathways and functions in plants cells. Front Plant Sci 12:640919

Streatfield SJ (2007) Approaches to achieve high-level heterologous protein production in plants. Plant Biotechnol J 5:2–15

Tada Y, Utsumi S, Takaiwa F (2003) Foreign gene products can be enhanced by introduction into low storage protein mutants. Plant Biotechnol J 1:411–422

Takagi H, Hiroi T, Yang L, Tada Y, Yuki Y et al (2005) A rice-based edible vaccine expressing multiple T cell epitopes induces oral tolerance for inhibition of Th2-mediated IgE responses. Proc Natl Acad Sci U S A 102:17525–17530

Takaiwa F, Takagi H, Hirose S, Wakasa Y (2007) Endosperm tissue is good production platform for artificial recombinant proteins in transgenic rice. Plant Biotechnol J 5:84–92

Thomas DR, Walmsley AM (2015) The effect of the unfolded protein response on the production of recombinant proteins in plants. Plant Cell Rep 34:179–187

Tian L, Sun SM (2011) A cost-effective ELP-intein coupling system for recombinant protein purification from plant production platform. PLoS One 6:e24183
Torrent M, Llompart B, Lasserre-Ramassamy S, Llop-Tous I, Bastida M et al (2009) Eukaryotic protein production in designed storage organelles. BMC Biol 7:5
Tschofen M, Knopp D, Hood E, Stoger E (2016) Plant molecular farming: much more than medicines. Annu Rev Anal Chem (Palo Alto, Calif) 9:271–294
Twyman RM, Stoger E, Schillberg S, Christou P, Fischer R (2003) Molecular farming in plants: host systems and expression technology. Trends Biotechnol 21:570–578
Urry DW, Starcher B, Partridge SM (1969) Coacervation of solubilized elastin effects a notable conformational change. Nature 222:795–796
Valdez-Ortiz A, Rascón-Cruz Q, Medina-Godoy S, Sinagawa-García SR, Valverde-González ME et al (2005) One-step purification and structural characterization of a recombinant His-tag 11S globulin expressed in transgenic tobacco. J Biotechnol 115:413–423
Vamvaka E, Arcalis E, Ramessar K, Evans A, O'Keefe BR et al (2016a) Rice endosperm is cost-effective for the production of recombinant griffithsin with potent activity against HIV. Plant Biotechnol J 14:1427–1437
Vamvaka E, Twyman RM, Murad AM, Melnik S, Teh AY et al (2016b) Rice endosperm produces an underglycosylated and potent form of the HIV-neutralizing monoclonal antibody 2G12. Plant Biotechnol J 14:97–108
Varki A (2017) Biological roles of glycans. Glycobiology 27:3–49
Viana FCJ, Dias CS, Franco LO, Lacorte C (2013) Heterologous production of peptides in plants: fusion proteins and beyond. Curr Protein Pept Sci 14:568–579
Vitale A, Boston RS (2008) Endoplasmic reticulum quality control and the unfolded protein response: insights from plants. Traffic 9:1581–1588
Vollmann J (2016) Soybean versus other food grain legumes: a critical appraisal of the united nations international year of pulses 2016. Die Bodenkultur 67:17–24
Wakasa Y, Takaiwa F (2013) The use of rice seeds to produce human pharmaceuticals for oral therapy. Biotechnol J 8:1133–1143
Walsh G, Jefferis R (2006) Post-translational modifications in the context of therapeutic proteins. Nat Biotechnol 24:1241–1252
Wang X, Jiang D, Shi J, Yang D (2017) Expression of alpha-1,6-fucosyltransferase (FUT8) in rice grain and immunogenicity evaluation of plant-specific glycans. J Biotechnol 242:111–121
Watson SA, Ramstad PE (1987) Corn: chemistry and technology. American Association of Cereal Chemists, St. Paul, USA
Webster GR, Teh AY, Ma JK (2017) Synthetic gene design-the rationale for codon optimization and implications for molecular pharming in plants. Biotechnol Bioeng 114:492–502
Wilbers RH, Westerhof LB, van Noort K, Obieglo K, Driessen NN et al (2017) Production and glyco-engineering of immunomodulatory helminth glycoproteins in plants. Sci Rep 10:45910
Wilken LR, Nikolov ZL (2012) Recovery and purification of plant-made recombinant proteins. Biotechnol Adv 30:419–433
Witcher DR, Hood EE, Peterson D, Bailey M, Bond D et al (1998) Commercial production of β-glucuronidase (GUS): a model system for the production of proteins in plants. Mol Breed 4:301–312
Wu G (2009) Amino acids: metabolism, functions, and nutrition. Amino Acids 37:1–17
Xie T, Qiu Q, Zhang W, Ning T, Yang W et al (2008) A biologically active rhIGF-1 fusion accumulated in transgenic rice seeds can reduce blood glucose in diabetic mice via oral delivery. Peptides 29:1862–1870
Xu MQ, Perler FB (1996) The mechanism of protein splicing and its modulation by mutation. EMBO J 15:5146–5153
Xu R, Li D, Li H, Li J, Yang Y et al (2016) Isolation of four rice seed-specific promoters and evaluation of endosperm activity. Plant Cell Tissue Organ Cult 128:125–132
Xu J, Towler M, Weathers PJ (2018) Platforms for plant-based protein production. In: Pavlov A, Bley T (eds) Bioprocessing of plant in vitro systems. Springer, Cham, Switzerland, pp 1–40

Yamada Y, Nishizawa K, Yokoo M, Zhao H, Onishi K et al (2008) Anti-hypertensive activity of genetically modified soybean seeds accumulating. Peptides 29:331–337

Yang J, Xun H, Niu L, He H, Cheng Y et al (2021) Elastin-like polypeptide and gamma-zein fusions significantly increase recombinant protein accumulation in soybean seeds. Transgenic Res 30:675–686

Yuki Y, Nojima M, Hosono O, Tanaka H, Kimura Y et al (2021) Oral MucoRice-CTB vaccine for safety and microbiota-dependent immunogenicity in humans: a phase 1 randomised trial. Lancet Microbe 2:e429–e440

Zhang L, Shi J, Jiang D, Stupak J, Ou J et al (2012) Expression and characterization of recombinant human alpha-antitrypsin in transgenic rice seed. J Biotechnol 164:300–308