식물조직배양학회지 제27권 제4호
Korean J. Plant Tissue Culture
Vol. 27, No. 4, 283~297 (2000)

Seeds as Repositories of Recombinant Proteins in Molecular Farming

Maurice M. Moloney*

Dept. of Biological Sciences, University of Calgary, 2500 University Drive N.W. Calgary, AB T2N 1N4 Canada

ABSTRACT Seeds are an ideal repository for recombinant proteins in molecular farming applications. However, in order to use plant seeds efficiently for the production of such proteins, it is necessary to understand a number of fundamental biological properties of seeds. This includes a full understanding of promoters which function in a seed-specific manner, the subcellular targeting of the desired polypeptide and the final form in which a protein is stored. Once a biologically active protein has been deposited in a seed, it is also critical that the protein can be extracted and purified efficiently. In this review, these issues are examined critically to provide a number of approaches which may be adopted for production of recombinant proteins in plants. Particular attention is paid to the relationship between subcellular localization and protein extraction and purification. The robustness and flexibility of seed-based production is illustrated by examples close to or already in commercial production.

Key words: Extraction, gene expression, molecular farming, oil-body, oleosin, pharmaceutical, protein purification, recombinant protein, secretion pathway, seeds

Seeds as natural repositories for proteins

The seed, as a means of plant propagation, utilizes sophisticated and versatile storage mechanisms. Seeds store lipids or carbohydrates as a carbon source, protein for carbon and nitrogen, and some inorganics such as phosphate which is sequestered in phytic acid. The seed is capable of storing these products for extended periods of time often under extreme conditions of cold, heat and water stress. The deposition and storage of proteins occurs in a variety of different tissues (e.g. cotyledons, perisperm, endosperm) within seeds. In addition, the subcellular localization of the stored proteins varies from species to species.

Recognizing that the seed is the site of storage protein deposition in plants leads to questions about which other proteins could be stored in seeds of transgenic plants expression recombinant proteins. Using gene-transfer techniques, it should be possible to modify the type of proteins in seeds so that they accumulate other proteins which can be used for therapeutic or industrial purposes.

The concept of using plants as a hold for the production of valuable proteins has been dubbed "molecular farming" (Goodman et al. 1987). A wide range of proteins and polypeptides have already been expressed in diverse plant tissues and organs, but the features of seeds which make them stable repositories of stored proteins provide unique advantages in molecular farming applications.

Modified seed-storage proteins

Early experiments to test the possibility of using seeds as hosts for molecular farming focused on the production of modified seed-storage proteins. These experiments were predicated on the idea that modified storage proteins might be most suitable for accumulation in large amounts in seeds, due to their deposition in protein bodies. Hoffman et al. (1988) constructed a modified β-phaseolin gene with an additional short sequence encoding a methionine-rich peptide. This sequence was inserted in the third exon of a β-phaseolin

*Corresponding author Tel (403) 220-6823 Fax (403) 220-0704
E-mail mmmolone@ucalgary.ca

gene, into a region predicted to encode an α-helical structure. This positioning was chosen to avoid disrupting the overall secondary structure of the protein when this modified storage protein was expressed in tobacco seeds. Unlike unmodified phaseolin, only a small amount of the modified phaseolin accumulated in the transgenic seeds. It was concluded that this was because modified storage protein failed to complete normal trafficking from the ER, through Golgi, into vacuolar protein bodies. Incomplete targeting may result in degradation of the modified storage protein. This early study illustrated that the structural requirements of a native protein may be so rigorous that any modification to it may impair subcellular targeting, leading to destruction of the desired polypeptide.

A second study by Vandekerckhove et al. (1989) on production of a modified seed-storge protein in transgenic seeds considered in more detail the processing of these proteins *in vivo*. They attempted to modify 2S albumins from *Arabidopsis* to allow them to act as carriers for heterologous peptides. Tagging the protein at the C-terminal end was not possible, because all 2S albumins are cleaved close to their C-termini as part of their processing into α- and β-subunits (Krebbers et al. 1988). Similarly, because of signal sequences and consequent splicing, tagging the N-terminus was equally impractical. Through a comparison of several 2S albumins, it was concluded that a loop between the 6th and 7th cysteines in the *Arabidopsis* 2S albumin is hypervariable and thus might accommodate modifications more readily than any other part of the protein. Vandekerckhove et al. (1989) produced an internal gene fusion of a Leuenkephalin coding sequence and an *Arabidopsis* 2S albumin gene. The enkephalin comprised a pentapeptide fragment, thus minimizing the potential for imposed strain in the storage protein loop. This fusion protein was expressed in transgenic *Arabidopsis* and *Brassica napus* seeds, and was found to undergo subcellular targeting to the protein bodies. To recover the desired pentapeptide the storage protein was subjected to tryptic cleavage which occurred after the flanking lysines included in the construct.

This work demonstrated that it is possible to exploit the properties of storage proteins to accumulate novel peptides and proteins in developing seeds. However, it is also clear that such modifications may be severely limited by the architecture of the storage protein carrier and that this may impose size restrictions upon the desired recombinant protein.

Oleosin proteins as carriers of recombinant proteins

If a seed is intended to act as a platform for molecular farming, the best vehicles for peptide/protein carriers are probably proteins which express highly in seeds. In addition to storage proteins, a second group of highly-expressed seed proteins have been exploited as vehicles for making recombinant proteins in seeds. These proteins are ***oleosins***. These are nearubiquitous seed proteins, present in all common oilseeds such as canola, sunflower, soybean, safflower and peanuts. They are also constituents of many other seeds not cultivated primarily for their oil, including corn and cotton. Oleosins are strictly not storage proteins although in canola, for example, they comprise between 8% (Huang 1992) to 20% (Murphy et al. 1987) of the total seed protein. The fundamental role of oleosins appears to be in encapsulating oil-bodies and probably controlling oil-body surface-to-volume ration, a property which determines ease of lipolysis of storage lipid during germination.

Unlike seed storage proteins, oleosins undergo subcellular targeting to oil-bodies without any structural modifications such as cleavage of N- or C-termini. Furthermore, they accumulate only on oil-bodies and can therefore be easily separated form other cellular contents by floatation centrifugation of aqueous seed extracts. These properties have led to their use as carriers for recombinant proteins in oilseeds (van Rooijen and Moloney 1995a).

Regulation of seed-specific expression

One of the major constraints of using seeds as a host for molecular farming is the need for optimization of gene expression levels. In fact, seeds lend themselves readily to this application, because seeds frequently store individual polypeptides at very high levels. A corollary of this is that genes encoding certain storage proteins often include highly active promoters which can be used to drive expression of heterologous genes in seeds.

The promoters of several seed-specific genes have

been studied in detail and these help to exemplify the potential of seed-specific regulatory sequences.

Phaseolin

One of the strongest seed promoters is that derived from ß-phaseolin (Sengupta-Gopalan et al. 1985). Using a 0.8 kb fragment of the 5' flanking region of the ß-phaseolin gene, Bustos et al. (1989) established the temporal and spatial regulatory pattern of this promoter in transgenic tobacco. This 800 bp fragment supported strong expression of a transgene (ß-glucuronidase) throughout the embryonic axis of tobacco. Dissection of the key regulatory elements of the phaseolin promoter showed that at least two domains of upstream sequence in the phaseolin promoter at -295 to -109 (UAS1) and -468 to -391 (UAS2) determine the seed specificity of this promoter. These two regions apparently determine cotyledonary and shoot meristematic transcriptional activity (UAS1) and hypocotyl localized expression (UAS2) (Bustos et al. 1991). The phaseolin promoter also contains negative cis-acting elements which may help to determine the temporal onset of phaseolin expression. Studies using the ß-phaseolin promoter deriving ß-glucuronidase (GUS) in transgenic plants (Bustos et al. 1989) indicate that phaseolin promoter is essentially seed-specific (Frisch et al. 1995). In molecular farming applications where biologically-active molecules may be expressed at high levels, such specificity may become important. For example, Hood et al. (1997) reported the creation of male-sterile corn plants as a collateral effect when an avidin gene was expressed under the control of a constitutive (i.e. not seed-specific) promoter. In some cases, such secondary effects are deleterious to the host and result in an uneconomical production system.

Oleosin promoters

The ß-phaseolin promoter exemplifies the seed-storage protein promoters, which as a group are the strongest promoters available for seed-specific expression. In oilseeds, however, seed metabolism is heavily weighted towards the synthesis and deposition of oil into oilbodies. As noted above, oleosins are significant protein components in many common oilseeds. The transcriptional activity necessary to produce oleosin to at least 8% of total cellular protein in *Brassica* spp. (Huang 1992) indicates that oleosin promoters may be useful in seed-specific molecular farming applications. Oil biosynthesis frequently precedes the major phase of storage protein production in oilseeds, thus oleosin promoters become activated earlier in seed development. This prolongs the time during which a desired heterologous protein can be deposited.

The functional architecture of an oleosin promoter was investigated by Plant et al. (1994). Dissection of a 2.5 kb region upstream of the coding sequence of the *Arabidopsis* 18 kDa oleosin suggested that there are modulating sequences at least 1050 bp distal to the transcriptional start site. Sequences responsible for the high amplitude of expression were found between -1100 and -600 and at -400 to -200 and there was also a region between -600 to -400 which acted as a negative regulatory element. The *Arabidopsis* 18 kDa seed oleosin promoter was highly seed-specific with only background expression detectable in leaves, stems, flower buds, roots and other vegetative organs. Expression was easily detected at early stages of embryogenesis (late heart to early torpedo stage) and became intense as the cotyledons began to expand. It is interesting to note that the *Arabidopsis* oleosin promoter is functional in a wide range of dicotyledonous seeds including flax, soybean, safflower, sunflower, cotton and castor bean (Abenes et al. 1997). Surprisingly, it also expressed at easily detectable levels in corn embryos indicating that this promoter from a dicotyledonous plant retains its activity in monocots. This promoter has been used in several molecular farming applications (Parmenter et al. 1995; Liu et al. 1997) including the production of hirudin, xylanase, and cyprinid growth hormone (Seyed and Moloney, unpublished; Seyed 1999).

Strategies for protein targeting

The successful production of recombinant proteins in seeds may be determined by a number of factors including the extent of gene expression, protein stability and correct folding or processing of the desired polypeptide. While gene expression levels can be manipulated to some extent using promoters of high transcriptional activity, the accumulation of the final product may depend on targeting or deposition the protein of interest

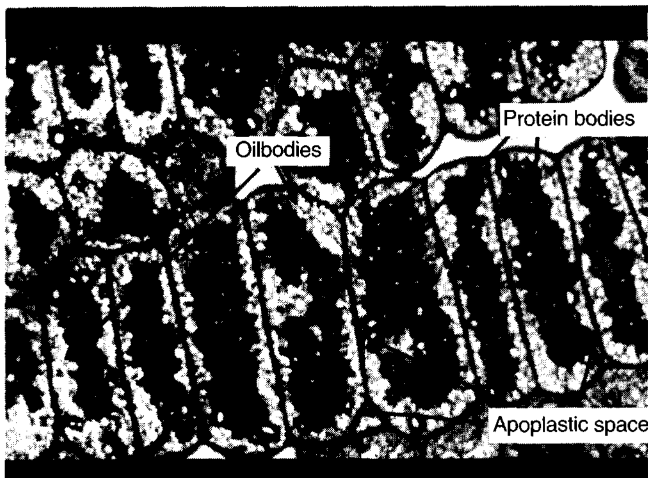

Figure 1. Transverse section of hypocotyl cells of developing *Brassica napus* seeds, indicating the protein bodies (dark staining), oil bodies lighter-stained spheres and apoplastic space. These subcellular locations have each been shown to act as suitable sites for recombinant protein deposition in seeds.

in a suitable subcellular compartment. This has not been investigated exhaustively, but several examples illustrate that protein accumulation may be strongly influenced by the cellular compartment to which it is targeted. Figure 1 shows a micrograph of developing canola cells, indicating some potential targets for recombinant protein localization.

Importance of the secretory pathway

In the area of molecular farming, one of the first indicators of the important influence of protein targeting was the accumulation of assembled bivalent antibodies in plants (Hiatt et al. 1989). The authors attempted to express a murine catalytic antibody in tobacco leaves by co-expressing two constructs specifying the heavy and light chains of the desired antibody. When the light and heavy chains were expressed without the inclusion of a secretion leader sequence or the transgene, no assembly or accumulation of the antibody occurred. Inclusion of leader sequences for either of the chains resulted in its very much higher accumulation, which was further enhanced when both heavy and light chains (with secretion leaders) were co-expressed. In the absence of ER retention signals in these antibody proteins, the two polypeptides entered the secretion pathway via the ER, underwent Golgi trafficking and were secreted into the apoplast (Hein et al. 1991).

In contrast to this, Conrad et al. (1998) found that high and stable accumulation of a single-chain, antigen-binding fragment (scFv) was favoured by targeting the polypeptide to the endoplasmic reticulum (ER) and including and ER-retention signal. It is not clear whether this is a general rule for scFvs, but these results suggest that in order to obtain appropriate accumulation of a desired protein it may be necessary to evaluate more than one target compartment. Some proteins may pose greater problems than others in this respect.

Targeting to oil bodies

Most seeds harbour cellular organelles called oleosomes or oil bodies for the storage of neutral lipid, which is most frequently triacylglycerol. In the common oilseeds such as sunflower (*Helianthus annuus*), safflower (*Carthamus tinctorium*), canola (*Brassica napus*) and flax (*Linum usitatissimum*) triacylglycerol may comprise as much as 40% of dry weight of the seed, and therefore, up to 50% of seed volume may constitute oil bodies. Oil bodies in seeds comprise three components: a neutral lipid lumen, a phospholipid half-unit membrane and a protein "shell" comprising one or more isoforms of the oleosin protein family (Huang 1992). Oil bodies in seeds occur in relatively large numbers and are typically 0.5-2 microns in diameter. Thus, a cell in a developing *Brassica* embryo may contain hundreds of oil bodies of average size 0.5 microns. Other proteins may adhere to oil bodies *in vivo* and, certainly, on extraction from seeds other proteins not usually resident on oil bodies can adhere (Kalinski et al. 1992). Despite this it is clear that oleosins associate with oil bodies in a manner very distinct from other cellular proteins. By using relatively harsh treatments such as washing with 8M urea or 0.5M sodium carbonate it is possible to remove all surface-adhering proteins except oleosins from oil bodies. Furthermore, it is now clear that the topology of oleosins is such that their highly lipophilic core is embedded in the lumen of the oil body while the N- and C-termini are cytoplasmically oriented (Huang 1992; Abell et al. 1997). That oleosins form an apparent "shell" around oil bodies is evident from experiments with phospholipase C, which has no effect on the disruption of oil bodies unless they are pre-treated with

a protease such as trypsin. Given the great number of small oil bodies displaying a large surface area in many oilseeds, the oleosin content of many seeds approaches those normally associated with storage proteins (8-20% in canola). When oilseeds undergo aqueous extraction, the oil bodies readily float, especially when subjected to moderate g forces (e.g. 5000 x g). When the resulting oil bodies are washed the result is a preparation essentially uncontaminated by other cellular proteins including storage proteins. Thus oleosins might provide a vehicle for the accumulation and purification of recombinant proteins in oilseeds. First, however, it must be determined whether creating recombinant oleosins with additional domains disrupts targeting and whether there is a severe limitation on the length of the polypeptide that can be produced. Both of these problems were addressed in a study by van Rooijen and Moloney (1995a) in which an oleosin-ß-glucuronidase (GUS) fusion gene was expressed in a seed-specific manner in *Brassica napus.* The oleosin gene used specified an 18kDa *Arabidopsis* oleosin. The GUS coding region used specified a protein of 68KDa. The resulting 86KDa fusion protein was shown to target efficiently to the oil body and to permit accumulation of the GUS marker. The ß-glucuronidase retained its enzymatic activity as an oleosin fusion. This permitted a simple estimation of GUS activity throughout the subcellular fractions of developing seeds and, typically, about 85% of the GUS activity could be accounted for on oil bodies (van Rooijen and Moloney 1995b). Of the remaining 15%, about half was associated with the microsomal fraction. This suggests that recombinant oleosins can target proteins of moderately high molecular weight to oil bodies with high efficiency. To test the idea that a recombinant protein can be easily purified using this procedure a four amino-acid proteolytic cleavage site was incorporated between the oleosin and GUS protein domains. This labile linker which was specific for the endo-protease, thrombin, was contiguous with the native C-terminus of *Arabidopsis* oleosin and was thus expected to be accessible to protease treatment. This configuration is shown in figure 2. When washed oleosin-GUS oil bodies were resuspended in fresh buffer and thrombin-treated, GUS was released from the oil bodies and into the aqueous phase. After cleavage of the protein, the oil bodies remained physically intact and could be floated off by one further round of centrifugation. This left an aqueous phase under the oil bodies ("the undernatant") which was enriched for GUS and had catalytic amounts of thrombin as its major contaminant. A diagram of these steps in schematic form is shown in figure 3.

These studies using oleosin-GUS fusions provided a proof of concept for the use of such a targeting system not only for protein production, but to assist in purification of the target protein.

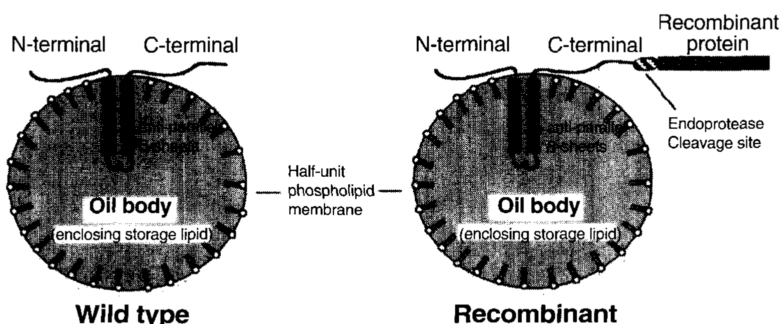

Figure 2. Schematic diagram showing a comparison of the configuration of wild type oleosins on seed oil bodies with recombinant proteins created as fusions with oleosins (recombinant).

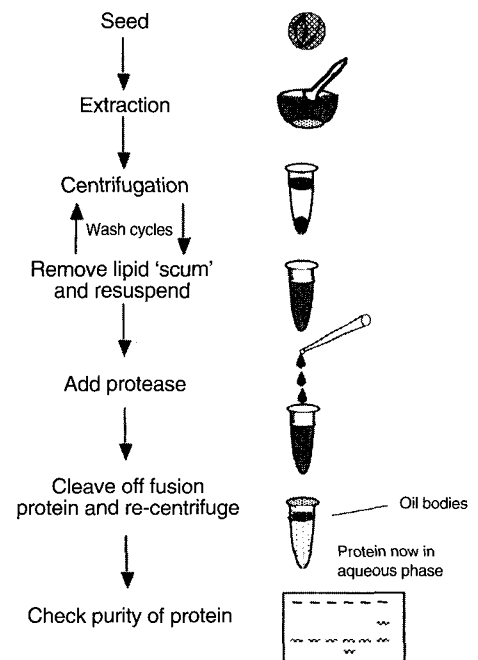

Figure 3. Flow diagram showing the key steps in extraction and recovery of oil body associated recombinant proteins expressed in oilseeds.

Plastid targeting

In the mesophyll and palisade cells of leaves, plastids contain a large fraction of the total cellular protein. It is not unreasonable, therefore, to consider plastids as a potential repository of recombinant proteins. The plastid genome ("plastome") itself encodes roughly 100 proteins (Sugiura 1989) but the majority of plastid proteins are encoded in the nucleus and targeted to one or other plastid subcompartment. Thus, to produce and accumulate substantial amounts of a desired recombinant protein in plastids, it is necessary to design a gene in which an appropriate chloroplast transit peptide is encoded upstream of the coding sequence for the mature protein. For simple targeting to the chloroplast stroma, numerous transit peptides are known which are capable of targeting heterologous proteins (Keegstra et al. 1989). In theory, it is possible to express a gene comprising a plastid-transit sequence and a recombinant-protein-coding sequence in developing seeds, giving rise to a seed-plastid localized protein. This approach has not yet been used for recombinant protein production, although non-protein products such as polyhydroxybutyrate (PHB) have been induced to accumulate in plastids, including seed plastids, in this manner by targeting three biosynthetic enzymes to these organelles to allow conversion of acetyl CoA into PHB (Nawrath et al. 1994).

The alternative approach of transforming the plastome with a gene encoding the desired polypeptide has become feasible in recent years (Maliga et al. 1993). However, the approach of these authors does not immediately lend itself to seed-specific, plastid-based expression. The difficult was that no naturally occurring factors determining seed-specific, plastid-based gene expression have been identified. MacBride et al. (1994) devised an elegant procedure to address this problem by creating a two-gene expression system that links nuclear and plastid transcription. To do this, a plastid-integrated T7 gene 10 promoter upstream of a GUS coding sequence was activated by a nuclear-encoded T7 RNA polymerase expressed cytoplasmically, but delivered to the plastid by means of an ribulose bisphosphate carboxylase small subunit (RuBPssu) transit peptide (McBride et al. 1994). Using engineered plastid genomes resulted in plastid-gene expression under the control of a nuclear-encoded gene. This approach can easily be extended to provide seed-specificity while obtaining the high levels of gene expression possible using engineered plastomes. Plastome transformation has the potential to deliver very high levels of expression because of the large numbers of copies of the heterologous gene present per cell which is a natural consequence of the gene's localization to a high copy number genome such as the plastome.

Implications of subcellular targeting for recombinant protein production

The delivery of recombinant proteins in seeds to different subcellular compartments has a number of critical implications including effects on protein stability, folding and post-translational processing of which on eof the most important is glycosylation.

In plants, like other eukaryotes, N-linked glycosylation of secreted proteins can occur on certain asparagine residues while O-linked glycosylation may occur on serine, threonine or hydroxyproline. These glycosylation events take place primarily in the ER, Golgi and vacuoles of all eukaryotes. Despite the broad similarities between the eukaryotes in the glycosylation of polypeptides, plants show a number of unique features compared to mammals and yeast. Plants and mammals share very similar "simple" glycosylation patterns in the ER in the production of "high mannose" type glycosyl residues. These comprise mannose residues in a dendric arrangement of α 1-6, α 1-3 and α 1-2 linkages attached to an N-acetyl glycosamine (GlcNAc) stem. However, trafficking of the simple glycosylated polypeptides to the Golgi generally results in further modifications to produce "complex" glycans. In plants, those differ significantly from their mammalian counterparts. Differences include the incorporation of fucosyl residues onto GlcNAc in positions other than stem GlcNAcs, and the incorporation xylosyl residues onto chain-branching mannose residues by a β1-2 linkage (Chrispeels and Faye 1996).

The impact of such glycosylation on the use of seeds for molecular farming applications may prove to be severe as many of these plant-specific glycosylation patterns are immunogenic to mammals. In consequence, if a seed were being used to generate, for example, a blood protein for therapeutic use, there might be drastic

immunological consequences which would obviate the use of such a product (Garcia-Casado et al. 1996).

Despite these fundamental problems, which render plant glycosylation as a major hurdle for production of certain proteins in seeds, there are several potential strategies available for the "humanization" of plant glycosylation. Von Schaewen et al. (1993) isolated an *Arabidopsis* mutant that lacks GnT-1, a glycosyl transferase and the first key enzyme in the modification pathway. It is interesting to note that this mutation does not result in any deleterious phenotype, whereas similar defects in the synthesis of N-glycans in mammals may result in extremely debilitating phenotypes (Fukuda et al. 1990). However, it is clear that the use of a plant host with a similar defect in complex glycosylation can result in the production of therapeutic polypeptides which are immunocompatible with mammals. This notwithstanding, certain mammalian proteins may still prove difficult to produce as functional entities in plants. This is exemplified in a study by Matsumoto et al. (1995), who attempted to produce erythropoietin in transgenic tobacco. In this case, the lack of sialic acid residues capping the glycosyl chains of the protein resulted in its rapid breakdown when administered to rats. Plants do not make sialic acid nor create sialyl-capped N-linked glycans. In consequence, it will be necessary to engineer plants to synthesize and incorporate sialyl residues if they are to be used for the synthesis of many important therapeutic proteins. This is no trivial task, and might involve incorporation of three or four functional mammalian genes into the plant host.

Localization to the cytoplasm of recombinant proteins avoids the necessity for glycosylation, and it is still possible to produce correctly-folded proteins therein, with appropriate disulfide linkages. This was shown by Parmenter et al. (1996) by the production of hirudin on seed oil bodies. This protein, which requires three disulfude bridges for functionality, accumulated on oil bodies and could be recovered as described above. Specific activity measurements of the hirudin indicated that it was essentially properly folded with appropriate disulfide linkages. It is not clear whether this disulfide bridging took place *in vivo* or whether it occurred spontaneously on extraction in air as an oxidizing environment. Nevertheless, for the production of proteins which require disulfide linkages it is not obligatory to use the secretion pathway to obtain a functional protein.

At this time, the stability of proteins in different subcellular compartments cannot be easily predicted. For many proteins, secretion into the apoplast has proven to be the most efficient means for their accumulation. However, this is not a general rule. Some secreted proteins, such as trout growth hormone, do not accumulate well in the seed apoplast (Bosch et al. 1994). Nevertheless, a carp growth hormone, expressed as an oleosin fusion (Moloney and Seyed, unpublished results; Seyed, 1999) accumulates efficiently in the apoplast even though the polypeptide was never processed through the conventional seed secretion pathway. As more examples are reported of specific proteins being targeted to different organelles or subcellular locations, more general predictive rules will probably emerge, which will help in the design of desired recombinant protein production systems.

Production of recombinant proteins in seeds

Enzymes

Seeds have proven to be a useful vehicle for the production of a number of enzymes for use in food, feed or as industrial catalysts. These products can accumulate stably in seeds and withstand extended periods of storage. The requirements for purification may be extensive or minimal, but in either case seeds have been shown to be an excellent repository for such products.

Phytase

An excellent example of this is the production of a fungal phytase in tobacco seeds (Pen et al. 1993). Phytase catalyses the hydrolysis of phytate to myo-inositol and inorganic phosphate (Pi). It is an important enzyme in seeds to release Pi from phytate following germination for use in the metabolism of growing seedlings. Phytase is an extremely useful additive to feed prepared for monogastric animals since they are unable to metabolize phytate. For this reason, in many feed formulations inorganic phosphate must be added. Phytate may also bind essential cations making them less available to the animal. However, the addition of Pi poses significant environmental issues since excreted phytate and unabsorbed Pi can accumulate in surface waters leading to

eutrophication. The addition of a microbial phytase to feed also improves feeding efficiency in monogastrics (Simons et al. 1990).

It was reasoned by Pen et al. (1993) that seed expression of an *Aspergillus* phytase would produce a simple feed additive, namely flaked seeds, which could be incorporated into feed rations and permit the animal to digest the phytate component of a traditional feed meal. Expression in tobacco seed of an *Aspergillus* phytase-coding sequence, preceded by a leader sequence specifying secretion obtained from a tobacco pathogenesis-related protein, resulted in the accumulation of the enzyme in the apoplast. Accumulation of the phytase occurred at approximately 1% of soluble protein. It is noteworthy that the phytase underwent glycosylation and that the glycosylation pattern was distinct from the fungal N-glycans. This resulted in an enzyme of about 67 kDa in tobacco seed instead of the 80 kDa protein normally produced in *Aspergillus*. When chemically deglycosylated, the enzymes from both sources had the same molecular mass. Feeding trials using milled phytase-containing seed as an additive demonstrated a major effect on feeding efficiency. When fed to chickens, phytase resulted in a 75% increase in weight gain over a 4-week period. This was equivalent to the weight gain obtained by the usual optimal Pi addition to such feed rations. This was an elegant, early use of molecular farming technology in that the functional product did not require any separation or purification. It also illustrates the idea of convenient enzyme production in food or feed sources to permit a monogastric animal to benefit from enzymes normally associated with ruminants and thus increase substantially feeding efficiency.

Xylanase

Xylanase catalyses the breakdown of xylan cross-links between cellulosic strands in plant cell walls. It is used in non-chemical pulp bleaching ("bio-bleaching"), de-inking and recycling of paper and possibly, in combination with other celluloytic enzymes, as a feed additive for monogastric animals.

A feed-additive application, similar to that for phytase, was envisaged by Liu et al. (1997) in the production in *Brassica* seeds of a xylanase from the rumen fungus, *Neocallimastix patriciarum*. The xylanase retained its enzymatic activity when expressed as an oleosin fusion and its Km was essentially unchanged from the native soluble xylanase. This resulted in an enzyme that was effectively immobilized on a liquid substrate and that could be easily recovered using flotation centrifugation. Attempts to use and recycle the enzyme in this manner showed that it was possible to recover the enzyme quantitatively over several cycles without loss of activity. While the xylanase produced on seed oil bodies was catalytically equivalent to the native xylanase, it showed a broader pH tolerance than the same enzyme expressed in a bacterial host such as *E. coli*. This property could have significant benefit when the enzyme is used as a feed additive. Xylans are a component of secondary cell walls and monocot primary cell walls, but exist only as a minor component of dicot primary cell walls. Herbers et al. (1995) expressed a thermostable xylanase from *Clostridium thermocellum* in tobacco seeds directing its accumulation through the secretory pathway to the apoplast. Not only did the xylanase accumulate in the apoplast but it was fully active. This implies that, at least in dicots, xylanases can reside in the apoplastic space without causing detriment to overall plant growth or reproduction through degradation of specific wall-components. In the gut of a monogastric animal, xylanase would assist in the breakdown of xylan-containing fibre derived from a monocot component such as corn meal. In combination with other cellulytic enzymes, such enzymic additions would increase feeding efficiency significantly.

Therapeutic factors

The oleosin system has proven to be one of the most versatile means for seed-based production of potential pharmaceutical proteins. Studies on this system have been directed towards the production of therapeutically-useful proteins such as hirudin, a blood anticoagulant, which occurs naturally as a secreted protein in the salivary glands of the medicinal leech (*Hirudo medicinalis*). Hirudin owes its anticoagulant properties to its ability to act as a potent inhibitor of thrombin in the blood coagulation cascade. Thrombin cleaves fibrinogen to yield an aggregate-forming protein fibrin. Mature hirudins are typically 65 amino acids in length. Parmenter et al. (1995, 1996) used oleosin fusions as a means of accumulating hirudin in seeds with the objective of producing and purifying to homogeneity the plant-derived

anti-coagulant and estimating its eventual specific activity. They used an *Arabidopsis* oleosin-coding sequence fused to a domain specifying the factor Xa cleavage site and the coding sequence of hirudin variant HV2. The coding sequence for hirudin was synthesized, correcting unusual codons for preferred plant codon usage, and the oleosin-hirudin gene fusion was expressed under the seed-specific control of the same *Arabidopsis* oleosin gene. The oleosin-hirudin fusion accumulated in substantial amounts (0.2-1% of total seed protein) and was predominantly oil-body associated.

To recover the hirudin polypeptide, it was necessary to separate and wash the oil-bodies and cleave the protein. Washing yielded a fraction which was essentially free of contaminating proteins. The washed oil-bodies were treated with factor Xa to release the hirudin. The cleavage site of factor Xa leaves no residual amino acids at the N-terminus of the fused protein. The resulting cleavage product was subjected to reversed-phase chromatography which yielded two peaks at retention times close to those of hirudin. The proteins contained in these peaks had potent antithrombin activity. Mass spectrometry and sequence analysis revealed that the peaks represented a full length (65 amino-acid) hirudin and a second form truncated by two amino acids at the C-terminus. Both forms had near-equivalent specific activities of 7500 antithrombin units per mg protein, very similar to the specific activity of yeast-derived commercial recombinant hirudin.

This example demonstrates the possibility of producing a protein biopharmaceutical in oilseeds and using partitioning as the primary method of purification. The oleosin system simplifies significantly the purification protocol and allows for recovery of highly purified product.

Diagnostic reagents

The first commercialized protein synthesized in transgenic seeds was the avian glycoprotein, avidin, which appeared in the Sigma chemical catalogue in 1997. Avidin is a protein with extremely high affinity for biotin. The tight binding between avidin and biotin is the basis for a wide range of diagnostics and affinity purification schemes in which biotinylated probes are detected with an enzyme-linked (covalent or non-covalent) form of avidin. Avidin is normally produced from egg whites, but this starting material can be costly and economic viability depends on co-extraction and purification of other egg-white components (e.g. lysozyme). In consequence, Hood et al. (1997) developed a line of transgenic corn expressing avidin in its seeds and other plant parts using a promoter derived from a corn ubiquitin gene. Accumulation of the secreted avidin occurred preferentially in the embryo even though this structure corresponds to less than 15% of the total seed weight. However, overall accumulation of avidin was 2% of total soluble seed protein. An interesting feature of the constitutively expressed avidin was the occurrence of male-sterility in plants displaying a high expression. This was probably a consequence of biotin starvation in developing anthers or pollen. Sequencing and characterization of the corn-seed-derived avidin indicated that the plant protein had a similar molecular weight and an identical sequence to the native avian protein. Although glycosylation residues were almost certainly different between the avidin produced in the corn and chicken, this did not result in substantially different molecular weights. The biotin binding activity of the seed-derived product was also identical. The plant-derived protein was easily purified using its affinity for immobilized biotin and thus a simple production and purification scheme was defined.

Antibodies (plantibodies) in seeds

The remarkable achievement of Hiatt et al. (1989) to produce assembled, functional antibodies in the leaves of plants has been developed further, and transgenic plants are now able to assemble complex secretory antibodies (Ma et al. 1995). While there is a variety of reasons to express a functional antibody (Ab) in plants (e.g. for phytoprotection or to inactivate a metabolic step), production of Abs in plants for diagnostic or therapeutic use requires a robust system. Seeds lend themselves readily to such 'plantibody' production as was demonstrated first by Fiedler and Conrad (1995). Rather than express an assembled bivalent antibody, this group constructed a single-chain antigen binding protein (scFv) that recognized the hapten, oxazolone. The coding sequence specified a signal sequence, an scFv consisting of variable heavy-linkervariable light domains and a *myc* immunological tag. The promoter was a strong legumin promoter derived from *Vicia faba*. Expression of this

scFv construct in tobacco seeds resulted in extensive mRNA accumulation in developing seeds of many of the transformants. The highest protein expression achieved was about 0.7% of total soluble seed protein. The resulting single chain antibody was functional in an ELISA and indistinguishable from the same scFv expressed in bacteria. One control that was performed was the expression of the scFv construct lacking a secretion leader sequence. Despite adequate amounts of mRNA, no accumulated scFv was detected in transformed seeds expressing the leaderless construct. This again emphasizes the importance of the subcellular localization of the desired protein in transgenic systems. Subsequent studies by the same group (Fiedler et al. 1997) found that incorporation of a KDEL ER- retention signal into the polypeptide sequence resulted in a 5-6 fold increase in the accumulation of the scFv against oxazolone. The reasons for this increase is not clear, although it is likely that the scFv is protected from proteolysis in the ER whereas in the apoplastic space it is degraded more rapidly.

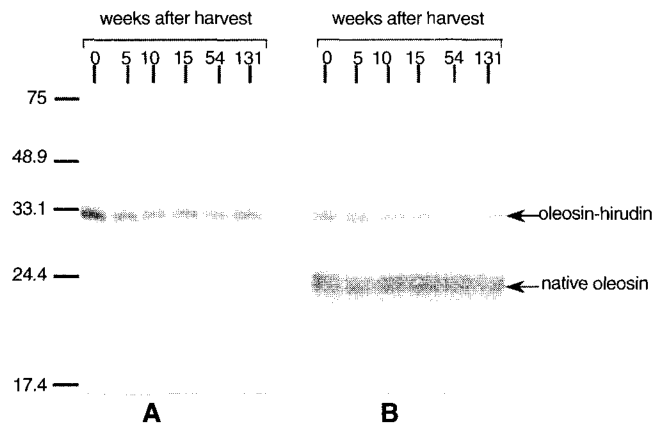

Figure 4. Long-term stability of recombinant proteins in seeds. Hirudin was stored as an oleosin fusion in *Brassica napus* seeds and compared with the stability of natural oleosin proteins. Panel A is an equivalent immunoblot probed with an anti-hirudin antibody. Panel B is an immunoblot visualized with an oleosin polyclonal antibody. The relative intensity of the oleosin and hirudin bands shows that the hirudin displays stability profiles similar to the native oleosin over the timescale of this test.

Why use seeds for molecular farming?

Long-term storage and stability

The concept of "molecular farming" has existed almost since the first experiments in making transgenic plants were successfully completed (Goodman et al. 1987). Nevertheless, it has taken about ten years for the idea to become a practical proposition. For large scale production of many proteins of pharmaceutical or industrial use, a reliable system of storage, extraction and purification needs to be designed. Seeds fulfill a number of criteria which render them particularly valuable repositories for recombinant proteins.

It was interesting to know whether recombinant proteins like storage proteins might remain stable over long periods in mature seeds. That this is indeed the case has been demonstrated using several species and seed types. For example, Pen et al. (1993) showed that phytase in transgenic tobacco seeds is stable at room temperature for over one year. This stability was the same in stored whole seeds or milled dry seed. Parmenter et al. (1996) found that hirudin stored as an oleosin fusion in canola seed survived intact and unmodified for at least three years. The stability of hirudin in seeds is shown in figure 4. In corn, avidin remained stable for at least three months even when the kernels were cracked and flaked (Hood et al. 1997). ß-Glucuronidase in corn seeds was stable for more than one year at room temperature (Kusnadi et al., 1998b). Thus it is clear that seeds offer a very stable means of storing a recombinant protein for long periods. This has huge economic implications for processing and extraction as it renders processing independent of the growing season and offers an extremely inexpensive means of storing an "inventory" of transgenically-produced compounds.

Expression levels

Unlike many plant organs, seeds exhibit very high expression of a relatively select group of proteins (typically storage proteins and oleosins). This is rendered possible by the existence of transcriptional regulators (promoters) which are among the most active known in eukaryotes. Promoters from genes of seed storage protein such as phaseolin (Bustos et al. 1989) and *Vicia* legumins (Fiedler and Conrad 1995) and oleosins (Plant et al. 1994) and, surprisingly, certain lipid-modifying enzymes such as the *Lesquerella* hydroxylase (Broun et al. 1997) which are also expressed in a seed-specific manner, have been used successfully to regulate high

levels of recombinant gene expression.

While the final accumulation of a desired protein depends on more than just transcriptional activity, high transcription rates are needed to ensure levels of polypeptide synthesis which will render the overall production system economically viable.

Reduced microbial contamination

In the production of polypeptides destined for use in pharmaceutical or food applications, a great deal of attention needs to be paid to the biotic load which accompanies the product through the extraction and purification system. Angiosperm seeds have a beneficial property in this respect since they develop in what is essentially a sterile compartment (the ovary), unlike leaves, stems or tubers which are exposed to a wide variety of biotic contaminants during the growth of the plant. While it is unrealistic to perform industrial-scale harvesting under aseptic conditions, most dry seeds may be subjected to surface sterilization without any detriment to seed contents. Most of these sterilization agents (hypochlorite, metabisulfite or ethanol) would impart some damage on leaves and stems in larger-scale preparations. Insect pests frequently deposit microorganisms below the cuticle of leaves. Under these circumstances surface sterilization will not be effective and the microorganisms will be released during the extraction procedures. In general, seeds are relatively free of these problems and thus carry a low microbial load.

Recovery of recombinant proteins from seeds

Besides these benefits which favour seeds as vehicles for recombinant-protein production, seeds also lend themselves readily to the recovery and purification of the transgene product. To date there have been only a few studies performed on recovery of recombinant proteins from seeds (Parmenter et al. 1996; Hood et al. 1997; Kusnadi et al. 1998a,b). A significant advantage of seeds as hosts for recombinant proteins is that seed processing is a very sophisticated industry. In consequence, there is a plethora of existing processing equipment and technology available.

Aqueous extraction systems

Where no purification is required a technique as simple as dry milling may be adequate for the production of a recombinant protein formulation. This was the approach used by Pen et al. (1993) for phytase. Normally, however, it will be essential to extract and at least partially purify the desired protein. Where seeds contain substantial quantities of oil, it is customary to employ hexane extraction to recover maximal amounts of the oil (see Figure 5). In most instances involving recombinant proteins, this step is not possible due to the denaturant property of hexane and many organic solvents. Therefore, extraction and purification of recombinant seeds will most likely require aqueous extraction (see Figure 6). This has been performed successfully at scale by Kusnadi et al. (1998 a,b) for the extraction of avidin and ß-glucuronidase (GUS) from corn seed. They noted that in spite of using a constitutive (i.e. non tissue-specific) promoter up to 98% of the GUS activity was deposited in the germ (embryo). This observation led to an advantageous extraction scheme in which the kernel separated into the endosperm and the embryo. The full-fat germ was then extracted using aqueous buffers to obtain a protein fraction, or alternatively the germ was first defatted with hexane at 60°C. Surprisingly, the GUS enzyme was not inactivated by this solvent treatment, although this may be a property peculiar to GUS which is generally considered to be a

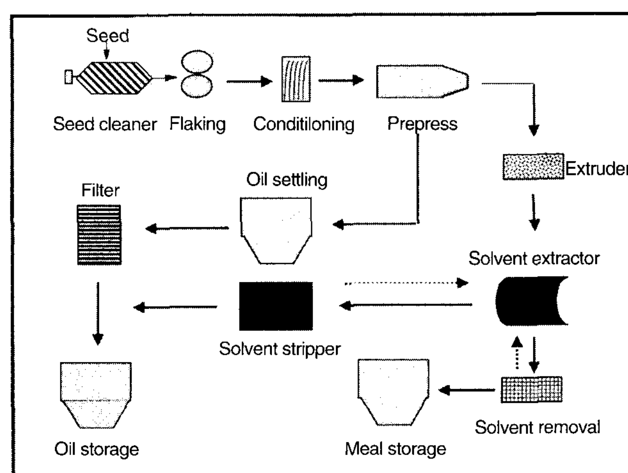

Figure 5. Process diagram depicting traditional extraction of oleaginous seeds which typically involves a heating step (conditioning) and hexane extraction. Without significant modification, it would be difficult to extract biologically active proteins using this traditional procedure.

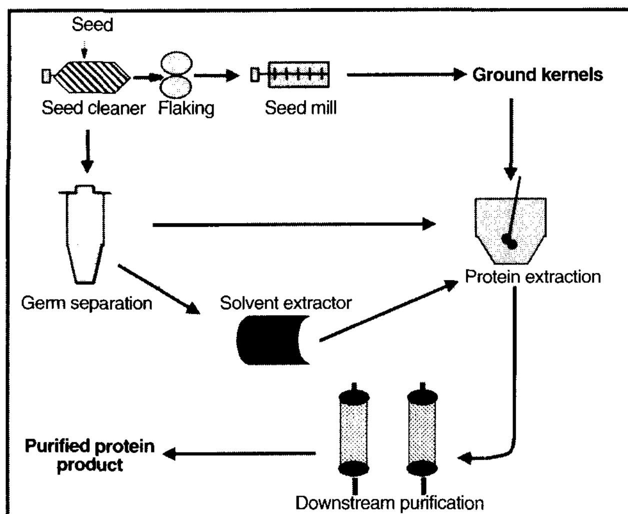

Figure 6. Process diagram for a modified procedure to permit extraction of recombinant proteins from corn while recovering some traditional by-products (after Kusnadi et al. 1998 a, b).

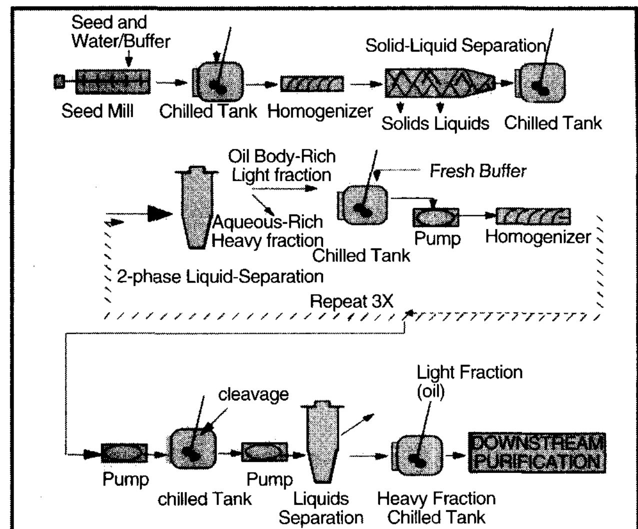

Figure 7. Process diagram showing a two phase partitioning system used with aqueous extraction to recover recombinant proteins from oilseeds particularly where the desired protein is associated with oil bodies or oleosins.

very stable enzyme. The aqueous protein-rich fraction was subjected to four rounds of chromatography including two ion exchanges, one hydrophobic interaction and one size-exclusion step. The overall yield of purified protein was 10%. This purification scheme would be quite costly on a large scale and with such a low recovery would only be economical for high-value proteins. Nevertheless, the scheme is simple and lends itself readily to optimization of each individual step of the process.

Two-phase extraction systems

An alternative extraction scheme for separation and purification to homogeneity of a protein is based on oil-body partitioning (Parmenter et al. 1996). In this system the host-seed that has been used is canola, and thus apart from dehulling, little pre-fractionation is possible. However, aqueous extraction of the whole seed followed by centrifugation to separate oilbodies has proven to be a major enrichment step. This alternative process is depicted in figure 7. Once the oilbodies are washed only minor amounts of other seed proteins remain. The cleavage of recombinant hirudin-oleosin protein from oilbodies provides a further enrichment step. This extract has been subjected to anion and reverse-phase chromatography and yielded an extraction efficiency of purified hirudin of about 30%. Minimization of chromatography steps and early-stage enrichment of the desired protein assist greatly in overall rates of recovery.

Recovery and purification of expressed recombinant proteins from transgenic plants is probably the most critical factor in establishing plants as a practical alternative system for protein production. While a variety of schemes might be envisaged, it is essential that the number of processing steps be minimized and that each step be carried out at much higher efficiencies than have thus far been reported. This is, however, an area that has not yet received much attention and it is likely that greater efficiencies will be forthcoming.

Conclusions

It is now evident that a wide range of proteins can be expressed in seeds and that many will accumulate stably. The level of accumulation of these foreign proteins is sensitive to a large number of factors including cell and tissue specificity of transcription, translatability of RNA and codon usage, subcellular targeting, assembly of subunits and availability of molecular chaperones, protein breakdown due to proteases, long-term stability and extraction and recovery protocols. Given that there are numerous alternatives to plant-based recombinant protein production, all of these factors will require optimization if plants are to become routine, economic vehicles for this technology. If these efficiency criteria are met, plants may prove to be a preferred vehicle for pro-

tein production in many instances. Plants are free of human or mammalian pathogens in their cells. This means that a plant-synthesized blood products, for example, could circumvent the problems associated with mammalian cell or blood-based sources of these vital therapeutics. If plant-based antibody and scFv production can be optimized, the cost of antibodies will be sufficiently low to permit a range of novel uses which currently cannot be envisaged. This might include using antibodies for large-scale purification of proteins or stereoisomers of smaller molecules. It would also render the use of antibodies in therapeutic applications much more feasible due to economics and ease of supply.

The implementation of such plant-based systems will be limited not only by technical progress. The use of seeds in this application will invoke a high degree of scrutiny in the regulatory arena. First, there is the question of this specialist use of transgenic plants and the potential for ecological damage, involuntary dispersal through pollen and possible accidental mixing of food-quality and pharma-quality seed materials. These issues are new to seed-based agriculture and will require careful management to avoid problems or the perception of unacceptable risk. The second regulatory issue is that of product quality of plant-derived materials. This will be particularly important in pharmaceutical applications where inadvertent carry-over of small amounts of a plant allergen might produce serious problems for some patients. Anticipation of these potential difficulties and the reduction of risk through exhaustive quality control and data collection will be required to establish seeds as routine sources of key pharmaceutical proteins.

These are significant challenges facing plant biologists in establishing 'molecular farming' but there is every reason for optimism as we explore the versatility of seed-based protein production of recombinant proteins.

References

Abell BM, Holbrook LA, Abenes M, Murphy DJ, Hills MJ, and Moloney MM (1997) Role of the proline knot motif in oleosin endoplasmic reticulum topology and oil body targeting. Plant Cell **9**:1481-93

Abenes M, Holbrook L, and Moloney MM. (1997) Transient expression and oil body targeting of an *Arabidopsis* oleosin-GUS reporter fusion protein in a range of oilseed embryos. Plant Cell Reports **17**: 1-7

Bosch D, Smal J, and Krebbers E. (1994) A trout growth hormone is expressed, correctly folded and partially glycosylated in the leaves but not the seeds of transgenic plants. Transgenic Research **3**: 304-310

Broun P, Boddupalli S, and Somerville C. (1998) A bifunctional oleate 12-hydroxylase: desaturase from *Lesquerella fendleri*. Plant Journal **13**: 201-210.

Bustos MM, Guiltinan MJ, Jordano J, Begum D, Kalkan FA, and Hall TC. (1989) Regulation of ß-glucuronidase expression in transgenic tobacco plants by an A/T-rich, cis-acting sequence found upstream of a French bean beta-phaseolin gene. Plant Cell **1**: 839-53

Bustos MM, Begum D, Kalkan FA, Baltraw MJ, and Hall TC. (1991) Positive and negative cis-acting DNA domains are required for spatial and temporal regulation of gene expression by a seed storage protein promoter. EMBO Journal **10**: 1469-1479.

Chrispeels MJ and Faye L. (1996) Production of recombinant glycoproteins with defined non-immunogenic glycans, in MRC Owen and J Pen (eds.), Transgenic plants: a production system for industrial and pharmaceutical proteins, John Wiley and Sons, pp 99-114.

Conrad U, Fiedler U, Artsaenko O, and Phillips J. (1998) High-level and stable accumulation of single-chain Fv antibodies in plant storage organs. Journal of Plant Physiology **152**: 708-711

Fiedler U and Conrad U. (1995). High level production and long-term storage of engineered antibodies in tobacco seeds. Biotechnology **13**: 1090-1093

Fiedler U, Phillips J, Artsaenko O, and Conrad U. (1997) Optimization of scFv antibody production in transgenic plants. Immunotechnology **3**: 205-216

Frisch DA, vanderGeest AHM, Dias K, and Hall TC. (1995) Chromosomal integration is required for spatial regulation of expression from the ß-phaseolin promoter. The Plant Journal **7**: 503-512

Fukuda MN, Masri KA, Dell A, Luzzatto L, and Moremen KW. (1990) Incomplete synthesis of N-glycans in congenital dyserythropoietic anemia type II caused by a defect in the gene encoding alpha-mannosidase II. Proceeds of the National Academy of Science USA **87**: 7443-7447

Garcia-Casado G, Sanchez-Mong R, Chrispeels MJ, Armentia A, Salcedo G, and Gomez L. (1996) Role of complex asparagine-linked glycans in the allergenicity of plant-glycoproteins. Glycobiology **6**: 471-477

Goodman RM, Knauf VC, Houck CM, and Comai L. (1987) Molecular Farming. European Patent Application. WO 87/00865, PCT/US86/01599

Hein MB, Tang Y, McLeod DA, Janda KD, and Hiatt A. (1991) Evaluation of immunoglobulins from plant cells. Biotechnology

Progress **7**: 455-461

Herbers K, Wilke I, and Sonnewald U. (1995) A thermostable xylanase from *Clostridium thermocellum* expressed at high levels in the apoplast of transgenic tobacco has no detrimental effects and is easily purified. Bio/Technology **13**: 63-66

Hiatt A, Cafferkey R, and Bowdish K. (1989) Production of antibodies in transgenic plants. Nature **342**: 76-78

Hoffman LM, Donaldson DD, and Herman EM. (1988) A modified storage protein is synthesized, processed and degraded in the seeds of transgenic plants. Plant Molecular Biology **11**: 717-729

Hood EE, Witcher DR, Maddock S, Meyer T, Baszczynski C, Bailey M, Flynn P, Register J, Marshall L, Bond D, Kulisek E, Kusnadi A, Evangelista R, Nikolov Z, Wooge C, Mehigh RJ, Hernan R, Kappel WK, Ritland D, Li CP and Howard JA (1997). Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. Molecular Breeding **3**: 291-306

Huang AHC. (1992) Oilbodies and oleosins in seeds. Annual Reviews of Plant Physiology and Plant Molecular Biology **43**: 177-200

Kalinski A, Melroy DL, Dwivedi RS, and Herman EM. (1992). A soybean vacuolar protein (p34) related to thiol proteases is synthesized as a glycoprotein precursor during seed maturation. Journal of Biological Chemistry **267**: 12068-12076.

Keddie JS, Tsiantis M, Piffanelli P, Cella R, Hatzopoulos P, and Murphy DJ. (1994) A seed-specific *Brassica napus* oleosin promoter interacts with a G-box-specific protein and may be bi-directional. Plant Molecular Biology **24**: 327-340

Keegstra K, Olsen LJ, and Theg SM. (1989) Chloroplastic precursors and these transport across the envelope membranes. Annual Reviews of Plant Physiology and Molecular Biology. **40**: 471-501

Krebbers E, Herdies L, Clercq A, Seurinck J, Leemans J, Damme J, Segura M. Gheysen G, and Van Montagu M. (1988) Determination of the processing sites of an *Arabidopsis* 2S albumin and characterization of the complete gene family. Plant Physiology **87**: 859-866

Kusnadi AR, Hood EE, Witcher DR, Howard JA, and Nikolov ZL. (1998a) Production and purification of two recombinant proteins from transgenic corn. Biotechnology Progress **14**: 149-155

Kusnadi AR, Evangelista RL, Hood EE, Howard JA, and Nikolov ZL. (1998b) Processing of transgenic corn seed and its effect on the recovery of recombinant beta-glucuronidase. Biotechnology and Bioengineering **60**: 44-52

Liu J-H, Selinger LB, Cheng K-J, Beauchemin KA, and Moloney MM. (1997) Plant seed oil-bodies as an immobilization matrix for a recombinant xylanase from the rumen fungus *Neocallimastix patriciarum*. Molecular Breeding **3**: 463-470

Ma JK-C, Hiatt A, Hein MB, Vine ND, Wang F, Stabila P, van Dolleweerd C, Mostov K and Lehner T. (1995) Generation and assembly of secretory antibodies in plants. Science **268**: 716-719

Mahmoud SS. (1999) Production of recombinant carp growth hormone in the seed of *Brassica napus*. Ph.D Dissertation, University of Calgary

Maliga P, Carrer H, Kanevski I, Staub J, Svab Z (1993) Plastid engineering in land plants: a conservative genome is open to change. Philosophical Transactions of the Royal Society London Biological Science **342**: 203-208

Mason HS, Lam DM-K, and Arntzen DJ. (1999) Expression of hepatitis B surface antigen in transgenic plants. Proceeds of the National Academy of Science USA **89**: 11745-11749

Matsumoto S, Ikura K, Ueda M, and Sasaki R. (1995) Characterization of a human glycoprotein (erythropoietin) produced in cultured tobacco cells. Plant Molecular Biology **27**: 1163-1172

McBride KE, Schaaf DJ, Daley M, and Stalker DM (1994) Controlled expression of plastid transgenes in plants based on a nuclear DNA-encoded and plastid-targeted T7 RNA polymerase. Proceeds of the National Academy of Science USA **91**: 7301-7305

McGarvey PB, Hammond J, Dienelt MM, Hooper DC, Fu ZF, Dietzschold B, Koprowski H, and Michaels FH. (1995) Expression of the rabies virus glycoprotein in transgenic tomatoes. Bio/Technology **13**: 1484-1487

Moloney MMM and van Rooijen GJH (1996) Recombinant proteins *via* oleosin partitioning. INFORM **7**: 107-113

Murphy DJ, Cummins I, and Kang SA. (1989) Synthesis of the major oil body membrane protein in developing rapeseed (*Brassica napus*) embryos. Biochemistry Journal **258**: 285-293

Nawrath C, Poirier Y, and Somerville C. (1994) Targeting of the polyhydroxybutyrate biosynthetic pathway to the plastids of *A. thaliana* results in high-levels of polymer accumulation. Proceeds of the National Academy of Science USA **91**: 12760-12764

Parmenter DL, Boothe JG, van Rooijen GJH, Yeung EC, and Moloney MM (1995) Production of biologically active hirudin in plant seeds using oleosin partitioning. Plant Molecular Biology **29**: 1167-1180

Parmenter DL, Boothe JG, and Moloney MM. (1996) Production and purification of recombinant hirudin from plant seeds, in MRL Owen and J Pen (eds.), Transgenic plants: a production system for industrial and pharmaceutical proteins, John Wiley and Sons Ltd., pp 261-280

Pen J, Verwoerd TC, van Paridon PA, Beudeker RF, van den Elzen PJM, Geerse K, van der Klis JD, Versteegh HAJ, van Ooyen AJJ, and Hoekema A. (1993) Phytase-containing transgenic seeds as a novel feed additive for improved phosphorus utilization. Bio/Technology **11**: 811-814

Plant AL, van Rooijen GJH, Anderson CP, and Moloney MM. (1994) Regulation of an *Arabidopsis* oleosin gene promoter in transgenic *Brassica napus*. Plant Molecular Biology **25**: 193-205

Sengupta-Gopalan C, Reichert NC, Burker RF, Hall TC, Kemp JD.

(1985) Developmentally regulated expression of the bean ß-phaseolin gene in tobacco seeds. Proceedings of the National Academy of Science USA **82**: 3320-3324

Sijmons PC, Dekker BMM, Schrammeijer B, Verwoerd TC, VandenElzen PJM, and Hoekema A (1990) Production of correctly processed human serum albumin in transgenic plants. Bio/Technology **8**: 217-221

Simons PCM, Versteegh HAJ, Jongbloed AW, Kemme PA, Slump P, Bos KD, Wolters MGE, Beudeker RF, and Verschoor GJ. (1990) Improved phosphorus availability by microbial phytase in broilers and pigs. British Journal of Nutrition **64**: 525-540

Sugiura M. (1989) The chloroplast chromosomes of land plants. Annual Reviews of Cell Biology **5**: 51-70

Vandekerckhove J, Damme J, Lijsebettens M, Botterman J, Block M, Vandewiele M, Clercq A, Leemans J, Van Montagu M, and Krebbers E. (1989) Enkephalins produced in transgenic plants using modified 2S seed storage proteins. Bio/Technology **7**: 929-932

van Rooijen GJH and Moloney MM. (1995a) Plant seed oil-bodies as carriers for foreign proteins. Bio/Technology **13**: 72-77

van Rooijen GJH and Moloney MM. (1995b) Structural requirements of oleosin domains for subcellular targeting to the oil body. Plant Physiology **109**: 1353-1361

von Schaewen A, Sturm A, O'Neill J, and Chrispeels MJ. (1993) Isolation of a mutant *Arabidopsis* plant that lacks N-acetyl glucosaminyl transferase I and is unable to synthesize Golgi-modified complex N-linked glycans. Plant Physiology **102**: 1109-18

Witcher DR, Hood EE, Peterson D, Bailey M, Bond D, Kusnadi A, Evangelista R, Nikolov Z, Wooge C, Mehigh R, Kappel W, Register J and Howard JA. (1998) Commercial production of ß-glucuronidase (GUS): A method system for the production of proteins in plants. Molecular Breeding **4**: 301-312