

Review article

Approaches to achieve high-level heterologous protein production in plants

Stephen J. Streatfield*†

Applied Biotechnology Institute, Building 36, California Polytechnic State University, San Luis Obispo, CA 93407, USA

Received 6 March 2006;

revised 20 June 2006;

accepted 26 June 2006.

*Correspondence (fax (302) 369-8955;
e-mail sstreatfield@fraunhofer-cmb.org)

†Present address: Fraunhofer USA Center for
Molecular Biotechnology, 9 Innovation Way,
Suite 200, Newark, DE 19711, USA

Summary

Plants offer an alternative to microbial fermentation and animal cell cultures for the production of recombinant proteins. For protein pharmaceuticals, plant systems are inherently safer than native and even recombinant animal sources. In addition, post-translational modifications, such as glycosylation, which cannot be achieved with bacterial fermentation, can be accomplished using plants. The main advantage foreseen for plant systems is reduced production costs. Plants should have a particular advantage for proteins produced in bulk, such as industrial enzymes, for which product pricing is low. In addition, edible plant tissues are well suited to the expression of vaccine antigens and pharmaceuticals for oral delivery. Three approaches have been followed to express recombinant proteins in plants: expression from the plant nuclear genome; expression from the plastid genome; and expression from plant tissues carrying recombinant plant viral sequences. The most important factor in moving plant-produced heterologous proteins from developmental research to commercial products is to ensure competitive production costs, and the best way to achieve this is to boost expression. Thus, considerable research effort has been made to increase the amount of recombinant protein produced in plants. This research includes molecular technologies to increase replication, to boost transcription, to direct transcription in tissues suited for protein accumulation, to stabilize transcripts, to optimize translation, to target proteins to subcellular locations optimal for their accumulation, and to engineer proteins to stabilize them. Other methods include plant breeding to increase transgene copy number and to utilize germplasm suited to protein accumulation. Large-scale commercialization of plant-produced recombinant proteins will require a combination of these technologies.

Keywords: foreign gene expression, heterologous gene expression, industrial protein production, plant-based vaccines, plant-produced pharmaceuticals, recombinant protein production.

Introduction

Genetically engineered microorganisms are a very important source of industrial and medicinal proteins. Over the last 30 years, the economic and safety advantages of recombinant microbial systems have allowed them to replace many native sources of proteins (Cronin, 1997; Frank, 1998). Recombinant microbes have also enabled the production of alternative engineered protein forms with modified activities and properties (Crabb and Bolin, 1999). However, microbes have certain limitations in the classes of proteins that can be economically produced, and in the post-translational process-

ing that can be achieved. Thus, insect and mammalian cell cultures have also been utilized for eukaryotic protein production (Lubiniecki and Lupker, 1994; Kost and Condreay, 1999), and transgenic animals (Houdebine, 2002) are also being evaluated. However, with these systems, in particular cell cultures, production costs are prohibitively high for many proteins.

Over the past decade plants have been developed as recombinant protein production systems (Howard and Hood, 2005). Potentially, they have great advantages over microorganisms, and especially animal cell systems, for the costs of production (Daniell *et al.*, 2001a). Moreover, unlike bacteria, they are capable of eukaryotic post-translational modifications,

Table 1 Plant systems for the bulk production of recombinant proteins

Approach to generate plant system	Tissues most commonly used for expression	Examples of high-level protein accumulation
Nuclear transformation	Leaves of dicots, seeds of dicots and monocots	Avidin (Hood <i>et al.</i> , 1997); single-chain variable fragment (Fiedler <i>et al.</i> , 1997; De Jaeger <i>et al.</i> , 2002)
Plastid transformation	Leaves of <i>Nicotiana tabacum</i>	Somatotropin (Staub <i>et al.</i> , 2000); receptor binding subunit of cholera toxin (Daniell <i>et al.</i> , 2001b)
Inoculation with plant viral vector	Leaves of dicots (especially <i>Nicotiana</i> species)	Trichosanthin (Kumagai <i>et al.</i> , 1993); follicle-stimulating hormone (Dirnberger <i>et al.</i> , 2001)

most importantly glycosylation (Hood *et al.*, 1997; Karnoup *et al.*, 2005). This can be achieved without the hyperglycosylation that is often observed with certain yeast systems (Nakamura *et al.*, 1993; Montesino *et al.*, 1998). Plants are particularly well suited to the production of edible oral vaccines (Streatfield and Howard, 2003) and industrial enzymes for biomass treatment (Xue *et al.*, 2003; Teymouri *et al.*, 2004). Protein purification is not required and, in the case of industrial enzymes, volumes are anticipated to be very large.

Several plant systems have been evaluated for the economic production of recombinant proteins. These can be categorized by the approach used to generate the production vehicle and by the type of tissue in which recombinant protein is produced (Table 1). Three broad approaches have been followed to make heterologous proteins in plants. The most common approach, and one that is applicable to a wide range of species, has been to use biolistics or agrobacteria to generate stably transformed plants (Davey *et al.*, 1989; Gelvin, 2003). Recombinant gene sequences are integrated into the host plant's nuclear genome, and the approach can be routinely applied, even to multisubunit proteins, such as antibodies (Hiatt *et al.*, 1989). Where possible, agrobacterial infection is generally preferred, as the resulting transgenic plants are less subject to silencing (De Wilde *et al.*, 2000). A second approach has been to use biolistics to generate transplastomic plants in which the plastid genome is modified to incorporate recombinant sequences (Svab and Maliga, 1993). Homologous recombination targets the transgene to a specific locus, and transplastomic plants are not subject to silencing. Until recently, very few tissue types and species were amenable to this technology, but it has now been extended to more important crops (Daniell *et al.*, 2005). Mitochondrial transformation is even more limited, with no reliable methodology yet developed, and so the production of recombinant proteins in plant mitochondria has not been pursued. As a third approach to produce foreign proteins in plants, some species have been inoculated with recombinant plant viral genomes. The engineered viral sequences are

transiently expressed in infected plant tissues, usually leaves (Porta and Lomonossoff, 2002; Canizares *et al.*, 2005). Transient recombinant protein production can also be accomplished using biolistics (Godon *et al.*, 1993) or agrobacteria (Hellens *et al.*, 2005), but, in these cases, expression tends to be impractically low.

Depending on which of these approaches is followed to introduce recombinant sequences into plants, and depending on the species of plant utilized, foreign proteins can be produced in various plant tissue types and can even be secreted into surrounding media, greatly simplifying purification (Borisjuk *et al.*, 1999; Fischer *et al.*, 1999). Thus, the tissue type for expression can greatly influence the costs of protein production (Kusnadi *et al.*, 1997; Horn *et al.*, 2004). Other factors, such as the options for post-translational processing, potential for contamination with harmful secondary metabolites and the confidence for biological containment, differ between alternative plant expression systems and target tissues. However, production costs are the key issue for recombinant protein production, and several factors affect these costs. These include the upfront costs of generating recombinant plant material, the costs of growing and harvesting plant material, and the downstream processing and purification costs. The latter tend to increase with tissue complexity. High levels of recombinant gene expression can greatly reduce the amounts of plant biomass required and can simplify processing and purification of the product. Thus, an overriding concern for all plant systems is the achievement of high expression (Nandi *et al.*, 2005).

Some protein products, such as industrial enzymes for applications such as bleaching and edible vaccines for delivery to animals, may require only limited processing of plant material (Hood *et al.*, 2003; Lamphear *et al.*, 2004). However, these products are relatively inexpensive, and so must be produced at high levels in plants to ensure economic viability. At the other extreme, more expensive products, such as pharmaceuticals for administration to humans and enzymes for pharmaceutical processing, require protein extraction and purification, greatly

adding to the production costs (Samyn-Petit *et al.*, 2001; Woodard *et al.*, 2003). These costs can be minimized through increasing expression. Thus, high-level expression is essential for economic recombinant protein production. In addition, for protein products that are not extracted and purified prior to use, expression must be sufficiently high to ensure efficacy. For example, for an edible vaccine, a sufficient dose of antigen to confer protection must be delivered in a quantity of plant tissue that can be practically ingested at a single sitting (Lamphear *et al.*, 2004; Tacket *et al.*, 2004).

Strategies have been developed to increase the levels of recombinant proteins for plant production systems. These focus on the synthesis and stability of recombinant nucleic acids and encoded proteins. The type of system used for protein production specifies the range of approaches that can be applied. With stable nuclear transgenics, fairly high levels of expression have been achieved for some recombinant proteins (Hood *et al.*, 1997; Streatfield *et al.*, 2002), and

there are many approaches available to boost expression. With transplastomic expression, very high levels of a few recombinant proteins have been achieved (Daniell *et al.*, 2001b; Kumar and Daniell, 2004), presumably due in part to the presence of up to 100 genomes per plastid and up to 100 plastids per cell. However, options to further boost expression are currently more limited than with nuclear transgenics. With transient viral systems, fairly high levels of expression have been achieved (Kumagai *et al.*, 1993), and tools are available to further boost expression. In this article, the various approaches used to boost recombinant protein production levels in plants are reviewed (Table 2).

Molecular approaches to boost expression

Several molecular approaches have been applied to boost heterologous gene expression. These target replication, transcription, transcript stability and translation.

Table 2 Approaches to boost recombinant protein accumulation in plants

Approach to increase protein accumulation	Applications of the approach with examples
Expression of replicases to replicate sequences encoding recombinant proteins	Transiently expressing plant viral systems (Canizares <i>et al.</i> , 2005); viral-based systems with sequences incorporated into the host genome (Yusibov <i>et al.</i> , 2002)
Use of strong promoters to transcribe sequences encoding recombinant proteins (options for promoter stacking and chimeric promoters)	Plant viral (Fiedler <i>et al.</i> , 1997), agrobacterial (Stefanov <i>et al.</i> , 1991) and plant (Dai <i>et al.</i> , 2000) promoters for nuclear transformed plants; plastid promoters for transplastomic plants (Daniell <i>et al.</i> , 2001b); viral promoters (Chapman <i>et al.</i> , 1992) for plant viral systems
Inclusion of scaffold attachment regions on expression constructs	Nuclear transformed plants (Allen <i>et al.</i> , 1993; Li <i>et al.</i> , 2002)
Co-expression of transcription factors	Nuclear transformed plants (Yang <i>et al.</i> , 2001); transplastomic plants with nuclear-expressed factors (McBride <i>et al.</i> , 1994); transiently expressing plant viral systems expressing factors to activate nuclear transformed sequences (Hull <i>et al.</i> , 2005)
Use of heterologous 3' untranslated regions to stabilize message, and heterologous 5' untranslated regions to boost translation	General applicability (Hood <i>et al.</i> , 1997; Richter <i>et al.</i> , 2000; Staub <i>et al.</i> , 2000; Shivprasad <i>et al.</i> , 1999)
Optimization of sequence surrounding the translation start, and codon optimization of foreign gene sequences	General applicability (Streatfield <i>et al.</i> , 2001)
Increasing the gene copy number through the crossing of transgenic lines, and further increasing protein accumulation through introgression into germplasm suitable for high-level expression	Nuclear transformed plants (Zhong <i>et al.</i> , 1999; Hood <i>et al.</i> , 2003)
Targeting to subcellular sites suitable for high-level protein accumulation	Nuclear transformed plants (Streatfield <i>et al.</i> , 2003); transiently expressing plant viral systems (Kumagai <i>et al.</i> , 2000)
Expression in tissues and at times suitable for protein accumulation	Use of tissue-specific (Hood <i>et al.</i> , 2003), temporal and inducible (Padidam, 2003) promoters in nuclear transformed plants, and regulatory leader sequences in transplastomic plants (Fernandez-San Millan <i>et al.</i> , 2003); timing of inoculation with transiently expressing plant viral sequences
Protein fusions to improve recombinant protein stability	General applicability, in particular fusions to coat proteins for transiently expressing plant viral systems (Canizares <i>et al.</i> , 2005)

Boosting the replication of transiently transformed sequences

With transiently expressing plant viral systems, recombinant proteins are encoded on an engineered viral vector. Most attention has focused on the use of positive-sense RNA viruses, in which, in addition to replication, the inoculating genome can be directly translated. *In vitro*-synthesized transcripts derived from cDNA clones of viral vectors are inoculated on to susceptible plant tissues. Foreign genes are generally inserted as additional reading frames into viral genomes rather than replacing native genes (Canizares *et al.*, 2005). This overcomes the problems observed in maintaining and expressing viral sequences when native genes are deleted. However, with many viral systems, there appear to be size constraints on the foreign protein expressed. *Cis*-acting elements can have a major effect on expression. For example, viral heterologous sequences can influence the expression of reporter genes (Shivprasad *et al.*, 1999).

Sequences encoding proteins required for viral replication are generally present on the inoculating transcripts. However, these replication functions can alternatively be supplied *in trans* from sequences stably integrated into the nuclear genome of the host plant (Taschner *et al.*, 1991; Sanchez-Navarro *et al.*, 2001; Yusibov *et al.*, 2002). This approach allows plants to be developed that constitutively express sufficiently high levels of the viral proteins required for replication, and has the added benefit of conferring a degree of biological containment on the inoculating viral sequences. Post-transcriptional gene silencing can be avoided by expressing replicase functions from inducible promoters (Mori *et al.*, 2001). As an extension of this approach to incorporate viral sequences into the host genome, efforts are underway to use viral replication machinery to boost expression from viral cassettes integrated into the genomes of transgenic plants. The aim is for sequences encoding viral replicases and sequences encoding viral vector-expressed foreign proteins to be integrated into the host plant genome, allowing RNA transcribed from the integrated viral vector to be replicated to high levels.

Similarly, a geminivirus expression system has been engineered that avoids repeated infection cycles each time recombinant protein is to be produced. In this case, viral sequences encoding foreign proteins have been co-transformed with a vector directing the constitutive expression of viral replicases by biolistic bombardment of a tobacco cell line (Hefferon and Fan, 2004). High-level replication can be transiently achieved and foreign protein expression enhanced. Ongoing work with this system is focused on the development of stably transformed lines with the replicase

functions under inducible control and encoded by the plant genome, and the recombinant viral sequences for foreign protein expression also inserted into the genome but subject to excision and replication.

As an alternative to the engineering of the genome of the host plant to provide viral replication functions, viral replicons can be delivered through infection with *Agrobacterium*, a process termed 'magnification' (Gleba *et al.*, 2005). Using this system, foreign protein can be transiently expressed at up to 80% of total soluble protein.

Boosting transcription

Increasing the levels of transcription of stably transformed sequences has probably received the most attention of the various approaches to boost expression. Transcription is an early step in the process of generating recombinant protein, and thus improvements here are viewed as having great potential to increase yield. Several promoters have been tested, of both plant pathogen and plant origin. In addition, some synthetic promoters have been developed.

As tobacco and *Arabidopsis* are generally relatively easily transformed, and leaf tissues provide most of the biomass with these species, promoters that drive relatively high levels of expression in dicot leaves were the tools of choice for much early phase research. The promoters tested included those of plant viruses and highly expressed dicot leaf genes. The cauliflower mosaic virus 35S promoter (Odell *et al.*, 1985) was the mainstay of early phase work focused on expressing recombinant proteins in dicot leaves. Using this promoter, fairly high levels of accumulation of recombinant proteins have been achieved (Fiedler *et al.*, 1997; Gutierrez-Ortega *et al.*, 2005). Other plant viral promoters have also been tested, including a cassava vein mosaic virus promoter (Verdaguer *et al.*, 1996), the C1 promoter of cotton leaf curl Multan virus (Xie *et al.*, 2003) and the promoter of component 8 of Milk vetch dwarf virus (Shirasawa-Seo *et al.*, 2005). Most of the viral promoters tested also show activity in monocots (Shirasawa-Seo *et al.*, 2005). Some plant viral promoters show tissue specificity (Mazithulela *et al.*, 2000). However, the strongest expressing plant viral promoters are generally not tissue specific (Odell *et al.*, 1985), and the expression of recombinant proteins in tissues other than the target tissue can be problematical for plant health and can make biological containment difficult. *Agrobacterium tumefaciens* promoters have also been used to express recombinant proteins in plants. These include the nopaline synthase (Shaw *et al.*, 1984) and mannopine synthetase promoters. The nopaline synthase promoter is less

active than the cauliflower mosaic virus 35S promoter (Sanders *et al.*, 1987; Harpster *et al.*, 1988), but the mannopine synthetase promoter shows comparable or increased activity (Stefanov *et al.*, 1991). Leaf-specific expression has been achieved using plant leaf promoters, such as that of the small subunit of ribulose-bisphosphate carboxylase, giving expression levels in excess of 1% of total soluble protein (Dai *et al.*, 2000). Dicot seed promoters have also been assessed, and very high levels of foreign gene expression have been achieved using the arcelin 5-I promoter of common bean (De Jaeger *et al.*, 2002). Using this system, an antibody single-chain variable fragment accumulated to approximately 36% of total soluble protein.

As a result of the relative difficulty of transformation, less work has focused on monocot systems, even though the seed tissues of cereals are very well suited to protein accumulation and storage. However, several groups have pursued monocot systems, often using high-expressing plant promoters, such as the maize polyubiquitin-1 promoter (Christensen *et al.*, 1992). Engineered versions of this promoter (Streatfield *et al.*, 2004) have achieved fairly high levels of expression in seeds. However, polyubiquitin-1 is a constitutive promoter, and therefore several other seed-specific monocot promoters have also been assessed, including the embryo-specific maize globulin-1 promoter (Belanger and Kriz, 1991) and the endosperm-specific maize 27-kDa zein and barley D hordein promoters (Russell and Fromm, 1997; Horvath *et al.*, 2000). The expression levels achieved in the seed tissues of monocots using seed-specific promoters have exceeded those using constitutive promoters (Hood *et al.*, 2003).

Transplastomic expression requires the use of promoters that are active in the plastid, which is somewhat akin to a prokaryotic expression system. The promoters tested here are of highly expressed plastid genome sequences, including the constitutive promoter of the rRNA operon. Very high levels of expression have been achieved using this promoter (Daniell *et al.*, 2001b; Tregoning *et al.*, 2003). However, the high expression level here is primarily considered to be a consequence of having up to 10 000 plastid genomes per leaf cell.

Foreign genes inserted into plant viral vectors for transient expression are under the transcriptional control of native plant viral promoters. As an alternative, foreign genes may be placed under the control of a viral subgenomic promoter. A subgenomic promoter can be duplicated to increase expression (Chapman *et al.*, 1992). In addition, separate subgenomic promoters can be used for the foreign gene and the viral coat protein. In this way, deletion of sequences as a result of recombination between identical regulatory elements can be avoided (Donson *et al.*, 1991).

Several strategies have been followed to further boost transcription over that achieved with plant or plant viral promoters. Multiple copies of enhancer sequences from a highly active promoter, such as cauliflower mosaic virus 35S, can be stacked to boost expression (Guerineau *et al.*, 1992). Further synthetic promoters have been developed that combine the most active sequences of multiple well-characterized natural promoters. For example, elements of the cauliflower mosaic virus 35S promoter and the *Agrobacterium* Ti plasmid mannopine synthetase promoter have been combined. In this case, the synthetic promoter is several-fold more active than either component (Comai *et al.*, 1990). Similar results have been achieved with chimeric plant viral promoters (Rance *et al.*, 2002).

Another approach to boost transcript levels has been to stack transcription units. These units can be driven by repeats of the same promoter and terminator or, preferably, by different controlling elements to reduce the potential for recombination or silencing. The stacking of transcription units typically increases expression over the use of a single unit, and allows for multiple proteins to be expressed (During *et al.*, 1990). However, transcriptional interference usually results in promoter activities being less than fully additive (Thompson and Myatt, 1997). The orientation of the transcription units relative to one another affects the degree of interference observed, and interference can be eliminated using transcriptional blocker sequences between the units (Padidam and Cao, 2001). An alternative way to express multiple copies of the same gene is to use synthetic bidirectional promoters with enhancer elements flanked on either side by core promoter elements (Xie *et al.*, 2001; Li *et al.*, 2004).

A further way to increase transcriptional activity is to include global regulatory sequences next to promoters on expression cassettes. These scaffold attachment or matrix attachment regions have been characterized as interacting with plant nuclear scaffolds *in vitro*, and are considered to place surrounding loci in locations suitable for the recruitment of transcription factors to promoters (Spiker and Thompson, 1996). These global sequences can boost the expression of plant genes by an order of magnitude (Allen *et al.*, 1993; Li *et al.*, 2002).

In addition to approaches to boost transcription that focus on *cis*-acting elements, other strategies have been followed that utilize *trans*-acting factors. Transcription factors capable of activating expression from chosen *cis*-acting promoter sequences are co-expressed in transgenic plants (Yang *et al.*, 2001). The transcription factors may bind directly to the promoters of the recombinant genes, or may interact with other factors binding these promoters. However, for coexpressed factors to have a positive effect on transcription, native levels of

these factors must be limiting in the host plant. The expression unit encoding a transcription factor can be co-transformed together with the transcription unit for recombinant protein expression, an approach that has boosted expression by two- to fourfold (Yang *et al.*, 2001). Alternatively, a transgenic line can be developed expressing elevated levels of a transcription factor, and this line can be transformed with a recombinant protein expression unit. As a further alternative, separate lines expressing a transcription factor and a recombinant expression unit can be crossed. A transcription factor can also be provided transiently by the inoculation of a transgenic line with plant viral sequences encoding the factor (Hull *et al.*, 2005). Moreover, by expressing a transcription factor from a promoter that the factor can itself bind to, very high levels of the transacting factor can potentially be achieved through positive feedback (Schwechheimer *et al.*, 2000). In addition to stable nuclear transgenic systems, the co-expression of *trans*-acting factors has been applied to transplastomic plants. For example, T7 RNA polymerase has been expressed from the nuclear genome and then targeted to the plastid, where it can transcribe sequences integrated into the plastid genome (McBride *et al.*, 1994).

Stabilizing the message and ensuring correct message processing

As a recombinant message is transcribed, the stability of that message is important to ensure high-level expression. The non-translated region located downstream of the translation stop codon is critical for processing and should include signals targeting the message for polyadenylation. Message destabilizing sequences in the downstream region can greatly affect stability (Green, 1993; Newman *et al.*, 1993), and these sequences must be avoided when preparing gene constructs to express high levels of recombinant proteins. Several plant viral and plant 3' untranslated regions have been utilized to process messages for recombinant proteins without destabilizing these messages. They include the cauliflower mosaic virus 35S terminator and the potato proteinase inhibitor II terminator (Hood *et al.*, 1997).

For plant-based expression systems, eukaryotic genes are generally re-synthesized or engineered to remove introns. Sequences such as consensus intron splice sites, message destabilizing sequences and transcript termination sequences should be avoided in constructing these mini-genes. However, some plant intron sequences contribute positively to the expression level observed for their native genes, and these sequences can boost expression if inserted as synthetic introns into genes (Fiume *et al.*, 2004).

Boosting translation

There are several approaches to increase translational activity for recombinant sequences. Several plant and plant viral 5' non-translated regions have been used with the aim of increasing the rate of translation initiation. For example, the tobacco mosaic virus and potato virus X leader sequences boost recombinant protein levels (Gallie *et al.*, 1987; Pooggin and Skryabin, 1992). In order to optimize translation, any sequences located immediately around the translation start site should be modified to fit the consensus initiation sequence, which varies between plant species (Joshi *et al.*, 1997).

Many of the genes expressed in recombinant plant systems, particularly those expressed by industrial groups, are synthesized *de novo* from oligonucleotides. This allows for fully codon-optimized genes to be designed to suit the expression host. Experiments with tobacco-expressed green fluorescent protein have demonstrated the benefit of codon optimization in plants (Rouwendal *et al.*, 1997). Preferred codon usage differs for monocots vs. dicots, and is greatly different for genes expressed from plastid genomes. Codon optimization within approximately 40 amino acids of the N-terminus is particularly important for increasing recombinant protein production (Batard *et al.*, 2000). Engineering the sequence here allows for most of the benefits of codon optimization, whilst minimizing gene construction costs. As another alternative to synthesizing whole genes, specific rare codons can be removed by site-specific mutagenesis. However, when optimizing a gene, extensive runs or localized concentrations of a specific codon should be avoided, so that the corresponding tRNA does not become rate limiting. Extensive predicted mRNA secondary structures that might hinder translation should also be avoided, as should internal ribosome entry sites that might compete for translation of the full message.

Polycistronic messages offer an alternative to multiple transcription units for the expression of more than one recombinant protein in transgenic plants. Co-expression of two reporter genes together with the cauliflower mosaic virus translational activator allowed for comparable levels of each foreign protein to accumulate (Futterer and Hohn, 1991). The expression of multiple foreign genes as polycistronic messages has also been achieved in plastids (Staub and Maliga, 1995; Quesada-Vargas *et al.*, 2005).

Genetic approaches to boost expression

Several genetic approaches have been applied to boost expression, including increasing the transgene copy number

through selfing or crossing, and introducing foreign genes into germplasm suited to their over-expression.

Increasing transgene copy number

With stable transgenics, taking a single transgene to homozygosity through selfing doubles transgene representation and typically boosts expression (Zhong *et al.*, 1999). Moreover, crossing high-expressing transgenic lines arising from independent transformation events can boost transgene copy number and expression levels. This approach to achieve high gene copy numbers appears to have a reduced risk of gene silencing in comparison with making multiple gene insertions during transformation.

Selection of germplasm

Certain germplasm is well suited to the expression of proteins at high levels. For example, maize lines have been developed over many generations with elevated protein or oil levels in seeds (Bhattaramakki and Kriz, 1996; Ting *et al.*, 1996). High-protein lines have the potential for the production of more recombinant protein in the seed on a weight basis. In high-oil lines, the increased proportion of oil is primarily a result of increased embryo size. Therefore, the level of recombinant proteins expressed in the embryo is also increased on a seed weight basis. As maize embryos have a much higher proportion of soluble to insoluble protein compared with endosperm, the use of high-oil lines can also increase recombinant protein expression on a percentage of total soluble protein basis. The use of high-oil lines has been shown to boost recombinant protein expression approximately fourfold in maize (Hood *et al.*, 2003). Direct transformation of these lines may not be a viable strategy, as maize transformation is very inefficient for these lines. Rather, a transgene introduced into a maize line receptive to transformation can then be introgressed into the desired germplasm. Moreover, the introgression of a transgene into a high-yielding agronomic line can significantly boost foreign protein expression (Hood *et al.*, 2003), in addition to the anticipated increase in yield on a per acre basis.

Certain mutants and engineered transgenic lines may also be well suited to recombinant protein production. For example, a transgenic maize line expressing a cytokinin-synthesizing enzyme under the control of a senescence-inducible promoter develops two normal-sized embryos, one at the expense of endosperm tissue (Young *et al.*, 2004). Thus, as with high-oil lines, this transgenic line has the potential for increased production of embryo-expressed recombinant proteins.

Protein accumulation and stability

Several approaches have been followed that focus on the stability of foreign proteins expressed in plants. These include avoiding amino acid sequences that target proteins for degradation, targeting proteins to an organellar environment suitable for accumulation, directing expression in a tissue type or at a time suited to protein accumulation, and making fusions to improve stability.

Avoidance of sequences directing protein degradation

Specific amino acid sequences identifying proteins for rapid turnover should be avoided when pursuing high-level expression. For example, a version of the N-end rule elucidated in microbial and mammalian systems, by which cytoplasmic or nuclear proteins with certain amino acids at their N-terminus are rapidly degraded, is also relevant to plants (Vierstra, 1996). Amino acids that target degradation should be avoided at the N-terminus, or recombinant proteins should be targeted to alternative organelles.

Subcellular targeting of foreign proteins

Recombinant proteins encoded by transgenes integrated into the nuclear genome can be targeted to specific subcellular sites by including signal sequences on the expression constructs. In the absence of a signal sequence, the recombinant protein will accumulate in the cytoplasm, and this has been the most common means of expression pursued. However, signal sequences targeting recombinant protein to the apoplast, the plastid, the mitochondrion, vacuolar compartments, and the nucleus have also been used (Hood *et al.*, 1997; Logan and Leaver, 2000; Streatfield *et al.*, 2003). Combining an apoplast targeting sequence with an endoplasmic reticulum retention sequence results in the recombinant protein being sequestered in the endoplasmic reticulum (Streatfield *et al.*, 2003). For some organelles, including plastids and mitochondria, alternative signal sequences offer suborganellar options for targeting recombinant protein.

In cases in which recombinant proteins have been targeted to alternative subcellular sites, considerable variation in protein accumulation has been observed. For example, the receptor binding subunit of the heat-labile toxin of *Escherichia coli* has been expressed in six alternative subcellular locations in maize seed, with expression varying across four orders of magnitude between organelles (Streatfield *et al.*, 2003). Factors that may affect the degree of accumulation in different subcellular locations include the biochemical environment of

the compartment, for example pH, and the available space in the compartment for protein storage. Space may be a constraint for a highly expressed recombinant protein. For example, the expression of hepatitis B surface antigen in soybean cell culture, with the recombinant protein localized to the endoplasmic reticulum, resulted in dilation of this membrane network, possibly indicating an accumulation limit being reached (Smith *et al.*, 2002).

Signal sequences specifying subcellular targets are not always removed from recombinant proteins. This may be a consequence of vector design, as with the C-terminal endoplasmic retention signal, which is not cleaved from native proteins that carry it. However, even if a signal sequence is removed from its native protein, it may not be cleaved when fused to a recombinant protein with a different amino acid sequence immediately adjoining the fusion site. For example, adding the signal sequence of soybean vegetative storage protein to the N-terminus of hepatitis B surface antigen for expression in tobacco cell culture resulted in the accumulation of the antigen with attached signal peptide in the endoplasmic reticulum (Sojikul *et al.*, 2003). Non-removal or imprecise cleavage of signal sequences may affect recombinant protein activity and may raise safety concerns for products that must pass through a regulatory path for licensing. Moreover, the use of native signal sequences of recombinant proteins in plant expression systems can result in unexpected targeting, as seen with the receptor binding subunit of the heat-labile toxin of *E. coli* expressed from its native signal sequence in maize endosperm tissue (Chikwamba *et al.*, 2003). Apoplast targeting was anticipated, but protein accumulated in starch granules, suggestive of amyloplast targeting.

The selection of organelle targets can also affect the post-translational processing of recombinant proteins. For example, proteins targeted through the endoplasmic reticulum can be glycosylated. If a recombinant protein is most strongly expressed in an organelle in which it will be glycosylated, but glycosylation would be detrimental to activity, one option is to mutate the normally glycosylated sites to alternative amino acids. Differences between animal and plant glycosylation patterns are being addressed in some recombinant plant systems. Anti-sense potato and tobacco lines have been generated that do not add a plant-specific moiety (Wenderoth and von Schaewen, 2000), and transgenic tobacco lines have been developed with the capacity to add mammalian moieties normally absent from plant proteins (Bakker *et al.*, 2001).

For recombinant proteins synthesized in the plastid from sequences integrated into the plastid genome, suborganellar targeting to particular locations within this organelle can be considered. Given their endosymbiotic origin, plastids may be

particularly well suited to the production of bacterial proteins. Recombinant proteins are not glycosylated, but correct disulphide bonding can be accomplished (Staub *et al.*, 2000). In addition, other post-translational modifications, such as lipidation, can be achieved, as for the outer surface lipoprotein A of *Borrelia burgdorferi* expressed from the tobacco plastid genome (Glenz *et al.*, 2006).

With regard to protein processing, plant viral expression systems are more akin to nuclear than plastid systems. They allow for targeting and for glycosylation of recombinant proteins (Kumagai *et al.*, 2000; Dirnberger *et al.*, 2001).

Tissue-specific, temporal and inducible expression

The choice of tissue for expression can greatly influence the accumulation and stability of recombinant proteins. Factors such as protein content, water content and presence of proteases and protease inhibitors affect foreign protein accumulation. The choice of tissue may also be influenced by the anticipated product application. For example, an edible tissue is suitable for an oral vaccine. As discussed above, tissue-specific promoters allow for the expression of recombinant proteins only in preferred target tissues, such as tobacco leaves (Dai *et al.*, 2000) and cereal seeds (Hood *et al.*, 2003).

The timing of expression can also affect protein accumulation and stability. For example, during the development of cereal seeds, most of the storage tissues are laid down just prior to desiccation. Expression from seed promoters active at this time will favour a similar accumulation of foreign proteins prior to desiccation, and this should favour protein stability. The expression of recombinant proteins at this late stage is also less likely to negatively affect plant health or embryo development. An alternative strategy, applicable to fresh tissues such as developing seedlings and leaves, is to harvest tissue expressing recombinant protein when accumulation of the protein is optimal (Rodriguez, 1999; Daniell *et al.*, 2001a).

The use of inducible promoters allows for expression in a target tissue only following a specific treatment, such as chemical spraying (Aoyama and Chua, 1997; Zuo and Chua, 2000; Padidam, 2003). In this case, protein accumulation can be restricted to a fine timeframe to limit detrimental effects on plant health. In addition, some regulatory sequences boost activity under certain growth conditions, such that the bulk of foreign protein accumulates in a brief period prior to harvest. For example, the plastid untranslated leader sequence of the 32-kDa D1 polypeptide of photosystem II enhances translational activity on switching growth conditions from a light/dark cycle to continuous light, and this has been utilized to boost yields from plants grown under

controlled light conditions (Fernandez-San Millan *et al.*, 2003; Koya *et al.*, 2005).

Protein fusions

Protein fusions can address issues of subcellular targeting, stability, purification and activity of product. Expression of a foreign protein in plants as a fusion to a plant protein present at high levels in a subcellular storage compartment within seed tissues allows for the foreign protein to be targeted to the same location as the native protein, where it can stably accumulate. For example, foreign proteins have been fused to oleosins for expression in *Brassica napus* seeds, where they accumulate in oil bodies (van Rooijen and Moloney, 1995). Fusion of a foreign protein to the C-terminus of a single ubiquitin coding unit has been used to stabilize expressed recombinant proteins in several recombinant systems. In plants, this has increased recombinant protein levels 10-fold (Garbarino *et al.*, 1995). The ubiquitin unit is subsequently removed *in vivo*. As a further alternative, polyproteins have been co-expressed in plants with a plant viral protease. Cleavage by the protease at a recognition site engineered between sequences encoding polypeptide chains yields distinct recombinant proteins (Dasgupta *et al.*, 1998). These proteins can be targeted to different organelles.

Fusion of a foreign protein or peptide to a second recombinant protein that has been shown to be stably expressed in plants can act to stabilize the target protein or peptide. The fusion of a tuberculosis antigen to the receptor binding subunit of the heat-labile toxin of *E. coli* follows this strategy (Rigano *et al.*, 2004). An added advantage in this case is that the heat-labile toxin subunit can direct the fused antigen to ganglioside receptors on the surface of the gut, so facilitating delivery of the product to the target tissue. In addition, fusion of a foreign protein to a recombinant protein with defined binding characteristics or to an affinity tag can simplify purification. This has been applied in insect cells and other recombinant systems with fusions to molecules such as avidin (Airenne *et al.*, 1999). However, later cleavage from the protein carrier or peptide tag may be necessary to release the desired product, so adding to the downstream processing costs.

With plant viral expression systems, a peptide or protein is often expressed as a fusion to a viral coat protein (Canizares *et al.*, 2005). Such fusions serve to stabilize the expressed protein or peptide, and also simplify purification, as plant viral particles can easily be separated from plant tissues. With plant-based vaccines the plant viral particles may also enhance the immunogenicity of the incorporated antigens. However, there is generally a strict size limit on the length of

foreign sequence that can be fused without disrupting the assembly of the viral protein coat. Typically, only up to a few tens of amino acids can be fused to the coat protein (Porta *et al.*, 2003). The isoelectric point of the sequence fused to the viral coat protein also affects the ability of the fusion to systemically infect plant tissues. With cowpea mosaic virus, the addition of a peptide with a high isoelectric point has a negative impact on the yield (Porta *et al.*, 2003). A strategy to overcome the size limit on viral coat fusions, but still to obtain particle assembly, is to design a fusion with the 16-amino-acid 2A peptide of foot-and-mouth disease virus positioned between the foreign protein and the coat protein. A proportion of the fusion is co-translationally cleaved to yield coat protein, allowing for the assembly of particles containing both coat protein and uncleaved fusion molecules (Cruz *et al.*, 1996; Smolenska *et al.*, 1998). Use of the 2A peptide has also been extended to stable transgenic plants to yield multiple proteins from a single transgene that can be independently targeted to different organelles (El Amrani *et al.*, 2004).

Expression of zymogens

The expression of proteases in plants offers a particular challenge, as successful expression is likely to result in the degradation of plant proteins and poor plant health. Organellar targeting and tissue-specific expression may serve to limit negative effects, but an alternative strategy was followed to express commercial levels of bovine trypsin in maize (Woodard *et al.*, 2003). In this case, expression was achieved without excessive proteolysis of plant proteins and poor plant health by producing the zymogen form of the protease. This strategy potentially calls for an additional proteolytic cleavage step during product purification, but, in the case of trypsinogen, the zymogen form of the enzyme was converted to active trypsin either in seed tissues or during protein extraction from seed (Woodard *et al.*, 2003).

Future prospects

The expression of foreign proteins in plants with a view to commercial production has attracted considerable attention over the past decade. However, only a few small-scale products have so far reached the market. These are pure protein products that require cost-effective processing and purification schemes for commercial viability, and currently they are not produced economically on a large scale (Hood *et al.*, 1997; Woodard *et al.*, 2003). Advances are required to significantly boost expression further, but, to date, many of the

available tools discussed above have not been put together to optimize the expression of a single protein product. This stacking of approaches will probably be required to produce commercial products, although this can be limited by access to the necessary intellectual property. Moreover, when implementing new strategies and combining currently available approaches, care must be taken to minimize potential negative influences, in particular gene silencing (De Wilde *et al.*, 2000).

In the near term, enzymes for large-scale industrial processes and antigens for oral animal vaccines are the most likely plant-expressed products to be commercially viable. Purification is not necessary for these products, and regulatory paths are less stringent than for pharmaceuticals (Streatfield, 2005). With edible vaccines, expression levels of some antigens are already sufficient for economic products, provided that promising early-stage efficacy data are confirmed in later trials (Lamphear *et al.*, 2004). With industrial enzymes, further increases in expression will be necessary.

Much more attention will also need to be applied to technical hurdles to ensure correct post-translational processing and protein stability in plant tissues. Post-translational processing is likely to be particularly important for pharmaceuticals, for which safety and efficacy concerns over issues such as plant-specific glycosylation may arise. Moreover, recombinant protein stability will be affected by host proteases. For recombinant protein production in microbes, host lines have been developed lacking proteases (Meerman and Georgiou, 1994). A similar approach with plants, achieved through gene knock-out or silencing, should serve to increase recombinant protein stability.

Foreign protein expression in plants has largely focused on the expression of one or, in a few cases such as antibodies, a few molecules at one time (Hiatt *et al.*, 1989; During *et al.*, 1990). However, in some cases it may be necessary to express several proteins in the same plant material, for example for an edible vaccine comprising several antigens along with a protein adjuvant. As discussed above, strategies are available to express multiple proteins from single vectors or to combine expressing lines. However, the development of plant artificial chromosomes should allow for many more proteins to be co-expressed (Brown *et al.*, 2000).

References

- Airenne, K.J., Laitinen, O.H., Alenius, H., Mikkola, J., Kalkkinen, N., Arif, S.A., Yeang, H.Y., Palosuo, T. and Kulomaa, M.S. (1999) Avidin is a promising tag for fusion proteins produced in baculovirus-infected insect cells. *Protein Expr. Purif.* **17**, 139–145.
- Allen, G.C., Hall, G.E. Jr, Childs, L.C., Weissinger, A.K., Spiker, S. and Thompson, W.F. (1993) Scaffold attachment regions increase reporter gene expression in stably transformed plant cells. *Plant Cell*, **5**, 603–613.
- Aoyama, T. and Chua, N.-H. (1997) A glucocorticoid-mediated transcriptional induction system in transgenic plants. *Plant J.* **11**, 605–612.
- Bakker, H., Bardor, M., Molthoff, J.W., Gomord, V., Elbers, I., Stevens, L.H., Jordi, W., Lommen, A., Faye, L., Lerouge, P. and Bosch, D. (2001) Galactose-extended glycans of antibodies produced by transgenic plants. *Proc. Natl. Acad. Sci. USA*, **98**, 2899–2904.
- Batard, Y., Hehn, A., Nedelkina, S., Schalk, M., Pallett, K., Schaller, H. and Werck-Reichhart, D. (2000) Increasing expression of P450 and P450-reductase proteins from monocots in heterologous systems. *Arch. Biochem. Biophys.* **379**, 161–169.
- Belanger, F.C. and Kriz, A.L. (1991) Molecular basis for allelic polymorphism of the maize globulin-1 gene. *Genetics*, **129**, 863–872.
- Bhatramakki, D. and Kriz, A.L. (1996) Nucleotide sequence analysis of a novel globulin1 null allele from the Illinois high protein strain of maize. *Plant Mol. Biol.* **32**, 1215–1219.
- Borisjuk, N.V., Borisjuk, L.G., Logendra, S., Petersen, F., Gleba, Y. and Raskin, I. (1999) Production of recombinant proteins in plant root exudates. *Nat. Biotechnol.* **17**, 466–469.
- Brown, W.R., Mee, P.J. and Hong Shen, M. (2000) Artificial chromosomes: ideal vectors? *Trends Biotechnol.* **18**, 218–223.
- Canizares, M.C., Nicholson, L. and Lomonosoff, G.P. (2005) Use of viral vectors for vaccine production in plants. *Immunol. Cell Biol.* **83**, 263–270.
- Chapman, S., Kavanagh, T. and Baulcombe, D. (1992) Potato virus X as a vector for gene expression in plants. *Plant J.* **2**, 549–557.
- Chikwamba, R.K., Scott, M.P., Mejia, L.B., Mason, H.S. and Wang, K. (2003) Localization of a bacterial protein in starch granules of transgenic maize kernels. *Proc. Natl. Acad. Sci. USA*, **100**, 11 127–11 132.
- Christensen, A.H., Sharrock, R.A. and Quail, P.H. (1992) Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol. Biol.* **18**, 675–689.
- Comai, L., Moran, P. and Maslyar, D. (1990) Novel and useful properties of a chimeric plant promoter combining CaMV 35S and MAS elements. *Plant Mol. Biol.* **15**, 373–381.
- Crabb, W.D. and Bolin, J. (1999) Protein technologies and commercial enzymes. *Curr. Opin. Biotechnol.* **10**, 321–323.
- Cronin, M.J. (1997) Pioneering recombinant growth hormone manufacturing: pounds produced per mile of height. *J. Pediatr.* **131**, S5–S7.
- Cruz, S.S., Chapman, S., Roberts, A.G., Roberts, I.M., Prior, D.A. and Oparka, K.J. (1996) Assembly and movement of a plant virus carrying a green fluorescent protein overcoat. *Proc. Natl. Acad. Sci. USA*, **93**, 6286–6290.
- Dai, Z., Hooker, B.S., Anderson, D.B. and Thomas, S.R. (2000) Expression of *Acidothermus cellulolyticus* endoglucanase E1 in transgenic tobacco: biochemical characteristics and physiological effects. *Transgenic Res.* **9**, 43–54.
- Daniell, H., Streatfield, S.J. and Wycoff, K. (2001a) Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. *Trends Plant Sci.* **6**, 219–226.
- Daniell, H., Lee, S.-B., Panchai, T. and Wiebe, P.O. (2001b) Expression of the native cholera toxin B subunit gene and assembly of

- functional oligomers in transgenic tobacco chloroplasts. *J. Mol. Biol.* **311**, 1001–1009.
- Daniell, H., Kumar, S. and Dufourmantel, N. (2005) Breakthrough in chloroplast genetic engineering of agronomically important crops. *Trends Biotechnol.* **23**, 238–245.
- Dasgupta, S., Collins, G.B. and Hunt, A.G. (1998) Coordinated expression of multiple enzymes in different subcellular compartments in plants. *Plant J.* **16**, 107–116.
- Davey, M.R., Rech, E.L. and Mulligan, B.J. (1989) Direct DNA transfer to plant cells. *Plant Mol. Biol.* **13**, 273–285.
- De Jaeger, G., Scheffer, S., Jacobs, A., Zambre, M., Zobell, O., Goossens, A., Depicker, A. and Angenon, G. (2002) Boosting heterologous protein production in transgenic dicotyledonous seeds using *Phaseolus vulgaris* regulatory sequences. *Nat. Biotechnol.* **20**, 1265–1268.
- De Wilde, C., Van Houdt, H., De Buck, S., Angenon, G., De Jaeger, G. and Depicker, A. (2000) Plants as bioreactors for protein production: avoiding the problem of transgene silencing. *Plant Mol. Biol.* **43**, 347–359.
- Dirnberger, D., Steinkellner, H., Abdennebi, L., Remy, J.-J. and van de Wiel, D. (2001) Secretion of biologically active glycoforms of bovine follicle stimulating hormone in plants. *Eur. J. Biochem.* **268**, 4570–4579.
- Donson, J., Kearney, C.M., Hilf, M.E. and Dawson, W.O. (1991) Systemic expression of a bacterial gene by a tobacco mosaic virus-based vector. *Proc. Natl. Acad. Sci. USA*, **88**, 7204–7208.
- During, K., Hippe, S., Kreuzaler, F. and Schell, J. (1990) Synthesis and self-assembly of a functional monoclonal antibody in transgenic *Nicotiana tabacum*. *Plant Mol. Biol.* **15**, 281–293.
- El Amrani, A., Barakate, A., Askari, B.M., Li, X., Roberts, A.G., Ryan, M.D., Halpin, C. (2004) Coordinate expression and independent subcellular targeting of multiple proteins from a single transgene. *Plant Physiol.* **135**, 16–24.
- Fernandez-San Millan, A., Mingeo-Castel, A.M., Miller, M. and Daniell, H. (2003) A chloroplast transgenic approach to hyper-express and purify human serum albumin, a protein highly susceptible to proteolytic degradation. *Plant Biotechnol. J.* **1**, 71–79.
- Fiedler, U., Phillips, J., Artsaenko, O. and Conrad, U. (1997) Optimization of scFv antibody production in transgenic plants. *Immunotechnology*, **3**, 205–216.
- Fischer, R., Liao, Y.C. and Drossard, J. (1999) Affinity purification of a TMV-specific recombinant full-size antibody from a transgenic tobacco suspension culture. *J. Immunol. Methods*, **226**, 1–10.
- Fiume, E., Christou, P., Giani, S. and Breviario, D. (2004) Introns are key regulatory elements of rice tubulin expression. *Planta*, **218**, 693–703.
- Frank, B.H. (1998) The development of recombinant human insulin. *Dev. Biol. Stand.* **96**, 91–95.
- Futterer, J. and Hohn, T. (1991) Translation of a polycistronic mRNA in the presence of the cauliflower mosaic virus transactivator protein. *EMBO J.* **10**, 3887–3896.
- Gallie, D.R., Sleat, D.E., Watts, J.W., Turner, P.C. and Wilson, T.M.A. (1987) The 5'-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts *in vitro* and *in vivo*. *Nucleic Acids Res.* **15**, 3257–3273.
- Garbarino, J.E., Oosumi, T. and Belknap, W.R. (1995) Isolation of a polyubiquitin promoter and its expression in transgenic potato plants. *Plant Physiol.* **109**, 1371–1378.
- Gelvin, S.B. (2003) Agrobacterium-mediated plant transformation: the biology behind the 'gene-jockeying' tool. *Microbiol. Mol. Biol. Rev.* **67**, 16–37.
- Gleba, Y., Klimyuk, V. and Marillonnet, S. (2005) Magnification – a new platform for expressing recombinant vaccines in plants. *Vaccine*, **23**, 2042–2048.
- Glenz, K., Bouchon, B., Stehle, T., Wallich, R., Simon, M.M. and Warzecha, H. (2006) Production of a recombinant bacterial lipoprotein in higher plant chloroplasts. *Nat. Biotechnol.* **24**, 76–77.
- Godon, C., Caboche, M. and Daniel-Vedele, F. (1993) Transient plant gene expression: a simple and reproducible method based on flowing particle gun. *Biochimie*, **75**, 591–595.
- Green, P.J. (1993) Control of mRNA stability in higher plants. *Plant Physiol.* **102**, 1065–1070.
- Guerineau, F., Lucy, A. and Mullineaux, P. (1992) Effect of two consensus sequences preceding the translation initiator codon on gene expression in plant protoplasts. *Plant Mol. Biol.* **18**, 815–818.
- Gutierrez-Ortega, A., Sandoval-Montes, C., de Olivera-Flores, T.J., Santos-Argumedo, L. and Gomez-Lim, M.A. (2005) Expression of functional interleukin-12 from mouse in transgenic tomato plants. *Transgenic Res.* **14**, 877–885.
- Harpster, M.H., Townsend, J.A., Jones, J.D., Bedbrook, J. and Dunsmuir, P. (1988) Relative strengths of the 35S cauliflower mosaic virus, 1', 2', and nopaline synthase promoters in transformed tobacco, sugarbeet and oilseed rape callus tissue. *Mol. Gen. Genet.* **212**, 182–190.
- Hefferon, K.L. and Fan, Y. (2004) Expression of a vaccine protein in a plant cell line using a geminivirus-based replicon system. *Vaccine*, **23**, 404–410.
- Hellens, R.P., Allan, A.C., Friel, E.N., Bolitho, K., Grafton, K., Templeton, M.D., Karunairetnam, S., Gleave, A.P. and Laing, W.A. (2005) Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods*, **1**, 13.
- Hiatt, A., Cafferkey, R. and Bowdish, K. (1989) Production of antibodies in transgenic plants. *Nature*, **342**, 76–78.
- Hood, E.E., Witcher, D.R., Maddock, S., Meyer, T., Baszczyński, C., Bailey, M., Flynn, P., Register, J., Marshall, L., Bond, D., Kulisek, E., Kusnadi, A., Evangelista, R., Nikolov, Z., Wooge, C., Mehig, R.J., Hernan, R., Kappel, W.K., Ritland, D., Li, C.-P. and Howard, J.A. (1997) Commercial production of avidin from transgenic maize: characterization of transformant, production, processing, extraction and purification. *Mol. Breed.* **3**, 291–306.
- Hood, E.E., Bailey, M.R., Beifuss, K., Magallanes-Lundback, M., Horn, M.E., Callaway, E., Drees, C., Delaney, D.E., Clough, R. and Howard, J.A. (2003) Criteria for high-level expression of a fungal laccase gene in transgenic maize. *Plant Biotechnol. J.* **1**, 129–140.
- Horn, M.E., Woodard, S.L. and Howard, J.A. (2004) Plant molecular farming: systems and products. *Plant Cell Rep.* **22**, 711–720.
- Horvath, H., Huang, J., Wong, O., Kohl, E., Okita, T., Kannangara, C.G. and von Wettstein, D. (2000) The production of recombinant proteins in transgenic barley grains. *Proc. Natl. Acad. Sci. USA*, **97**, 1914–1919.
- Houdebine, L.M. (2002) Antibody manufacture in transgenic animals and comparisons with other systems. *Curr. Opin. Biotechnol.* **13**, 625–629.
- Howard, J.A. and Hood, E. (2005) Bioindustrial and biopharmaceutical products produced in plants. *Adv. Agron.* **85**, 91–124.
- Hull, A.K., Yusibov, V. and Mett, V. (2005) Inducible expression in plants by virus-mediated transgene activation. *Transgenic Res.* **14**, 407–416.

- Joshi, C.P., Zhou, H., Huang, X. and Chiang, V.L. (1997) Context sequences of translation initiation codon in plants. *Plant Mol. Biol.* **35**, 993–1001.
- Karnoup, A.S., Turkelson, V. and Anderson, W.H. (2005) O-linked glycosylation in maize-expressed human IgA1. *Glycobiology*, **15**, 965–981.
- Kost, T.A. and Condreay, J.P. (1999) Recombinant baculoviruses as expression vectors for insect and mammalian cells. *Curr. Opin. Biotechnol.* **10**, 428–433.
- Koya, V., Moayeri, M., Leppla, S.H. and Daniell, H. (2005) Plant-based vaccine: mice immunized with chloroplast-derived anthrax protective antigen survive anthrax lethal toxin challenge. *Infect. Immun.* **73**, 8266–8274.
- Kumagai, M.H., Turpen, T.H., Weinzettl, N., della-Cioppa, G., Turpen, A.M., Donson, J., Hilf, M.E., Grantham, G.L., Dawson, W.O., Chow, T.P., Piatak, M. Jr and Grill, L.K. (1993) Rapid, high-level expression of biologically active alpha-trichosanthin in transfected plants by an RNA viral vector. *Proc. Natl. Acad. Sci. USA*, **90**, 427–430.
- Kumagai, M.H., Donson, J., della-Cioppa, G. and Grill, L.K. (2000) Rapid, high-level expression of glycosylated rice α -amylase in transfected plants by an RNA viral vector. *Journal*, **245**, 169–174.
- Kumar, S. and Daniell, H. (2004) Engineering the chloroplast genome for hyperexpression of human therapeutic proteins and vaccine antigens. *Methods Mol. Biol.* **267**, 365–383.
- Kusnadi, A.R., Nikolov, Z.L. and Howard, J.A. (1997) Production of recombinant proteins in transgenic plants: practical considerations. *Biotechnol. Bioeng.* **56**, 473–484.
- Lamphear, B.J., Jilka, J.M., Kesi, L., Welter, M., Howard, J.A. and Streatfield, S.J. (2004) A corn-based delivery system for animal vaccines: an oral transmissible gastroenteritis virus vaccine boosts lactogenic immunity in swine. *Vaccine*, **22**, 2420–2424.
- Li, X.-G., Zeng, Q.-C., Chen, S.-B., Xu, J.-W., Chang, T.-J. and Zhu, Z. (2002) Influence of matrix attachment regions from maize on transgene expression level in tobacco. *Acta Bot. Sinica*, **44**, 804–808.
- Li, Z.T., Jayasankar, S. and Gray, D.J. (2004) Bi-directional duplex promoters with duplicated enhancers significantly increase transgene expression in grape and tobacco. *Transgenic Res.* **13**, 143–154.
- Logan, D.C. and Leaver, C.J. (2000) Mitochondria-targeted GFP highlights the heterogeneity of mitochondrial shape, size and movement within living plant cells. *J. Exp. Bot.* **51**, 865–871.
- Lubiniecki, A.S. and Lupker, J.H. (1994) Purified protein products of rDNA technology expressed in animal cell culture. *Biologicals*, **22**, 161–169.
- Mazithulela, G., Sudhakar, D., Heckel, T., Mehlo, L., Christou, P., Davies, J.W. and Boulton, M.I. (2000) The maize streak virus coat protein transcription unit exhibits tissue-specific expression in transgenic rice. *Plant Sci.* **155**, 21–29.
- McBride, K.E., Schaaf, D.J., Daley, M. and Stalker, D.M. (1994) Controlled expression of plastid transgenes in plants based on a nuclear DNA-encoded and plastid-targeted T7 RNA polymerase. *Proc. Natl. Acad. Sci. USA*, **91**, 7301–7305.
- Meerman, H.J. and Georgiou, G. (1994) Construction and characterization of a set of E. coli strains deficient in all known loci affecting the proteolytic stability of secreted recombinant proteins. *Biotechnology (NY)*, **12**, 1107–1110.
- Montesino, R., Garcia, R., Quintero, O. and Cremata, J.A. (1998) Variation in N-linked oligosaccharide structures on heterologous proteins secreted by the methylotrophic yeast *Pichia pastoris*. *Protein Expr. Purif.* **14**, 197–207.
- Mori, M., Fujihara, N., Mise, K. and Furusawa, I. (2001) Inducible high-level mRNA amplification system by viral replicase in transgenic plants. *Plant J.* **27**, 79–86.
- Nakamura, S., Takasaki, H., Kobayashi, K. and Kato, A. (1993) Hyperglycosylation of hen white lysozyme in yeast. *J. Biol. Chem.* **268**, 12 706–12 712.
- Nandi, S., Yalda, D., Lu, S., Nikolov, Z., Misaki, R., Fujiyama, K. and Huang, N. (2005) Process development and economic evaluation of recombinant human lactoferrin expressed in rice grain. *Transgenic Res.* **14**, 237–249.
- Newman, T.C., Ohme-Takagi, M., Taylor, C.B. and Green, P.J. (1993) DST sequences, highly conserved among plant SAUR genes, target reporter transcripts for rapid degradation in tobacco. *Plant Cell*, **5**, 701–714.
- Odell, J.T., Nagy, F. and Chua, N.H. (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature*, **313**, 810–812.
- Padidam, M. (2003) Chemically regulated gene expression in plants. *Curr. Opin. Plant Biol.* **6**, 169–177.
- Padidam, M. and Cao, Y. (2001) Elimination of transcriptional interference between tandem genes in plant cells. *Biotechniques*, **31**, 328–334.
- Pooggin, M.M. and Skryabin, K.G. (1992) The 5'-untranslated leader sequence of potato virus X RNA enhances the expression of a heterologous gene *in vivo*. *Mol. Gen. Genet.* **234**, 329–331.
- Porta, C. and Lomonosoff, G.P. (2002) Viruses as vectors for the expression of foreign sequences in plants. *Biotechnol. Genet. Eng. Rev.* **19**, 245–291.
- Porta, C., Spall, V.E., Findlay, K.C., Gergerich, R.C., Farrance, C.E. and Lomonosoff, G.P. (2003) Cowpea mosaic virus-based chimaeras. Effects of inserted peptides on the phenotype, host range, and transmissibility of the modified viruses. *Virology*, **310**, 50–63.
- Quesada-Vargas, T., Ruiz, O.N. and Daniell, H. (2005) Characterization of heterologous multigene operons in transgenic chloroplasts: transcription, processing, and translation. *Plant Physiol.* **138**, 1746–1762.
- Rance, I., Norre, F., Gruber, V. and Theisen, M. (2002) Combination of viral promoter sequences to generate highly active promoters for heterologous therapeutic protein over-expression in plants. *Plant Sci.* **162**, 833–842.
- Richter, L.J., Thanavala, Y., Arntzen, C.J. and Mason, H.S. (2000) Production of hepatitis B surface antigen in transgenic plants for oral immunization. *Nat. Biotechnol.* **18**, 1167–1171.
- Rigano, M.M., Alvarez, M.L., Pinkhasov, J., Jin, Y., Sala, F., Arntzen, C.J. and Walmsley, A.M. (2004) Production of a fusion protein consisting of the enterotoxigenic *Escherichia coli* heat-labile toxin B subunit and a tuberculosis antigen in *Arabidopsis thaliana*. *Plant Cell Rep.* **22**, 502–508.
- Rodriguez, R.L. (1999) Process for protein production in plants. US Patent 5889189.
- van Rooijen, G.J.H. and Moloney, M.M. (1995) Plant seed oil-bodies as carriers for foreign proteins. *Biotechnology (NY)*, **13**, 72–77.
- Rouwendal, G.J., Mendes, O., Wolbert, E.J. and Douwe de Boer, A. (1997) Enhanced expression in tobacco of the gene encoding green fluorescent protein by modification of its codon usage. *Plant Mol. Biol.* **33**, 989–999.

- Russell, D.A. and Fromm, M.E. (1997) Tissue-specific expression in transgenic maize of four endosperm promoters from maize and rice. *Transgenic Res.* **6**, 157–168.
- Samyn-Petit, B., Gruber, V., Flahaut, C., Wajda-Dubos, J.P., Farrer, S., Pons, A., Desmaizieres, G., Slomianny, M.C., Theisen, M. and Delannoy, P. (2001) N-glycosylation potential of maize: the human lactoferrin used as a model. *Glycoconj. J.* **18**, 519–527.
- Sanchez-Navarro, J., Miglino, R., Ragozzino, A. and Bol, J.F. (2001) Engineering of Alfalfa mosaic virus RNA 3 into an expression vector. *Arch. Virol.* **146**, 923–939.
- Sanders, P.R., Winter, J.A., Barnason, A.R., Rogers, S.G. and Fraley, R.T. (1987) Comparison of cauliflower mosaic virus 35S and nopaline synthase promoters in transgenic plants. *Nucleic Acids Res.* **15**, 1543–1558.
- Schwechheimer, C., Corke, F.M., Smith, C.H. and Bevan, M. (2000) Transactivation of a target gene through feedforward loop activation in plants. *Funct. Integr. Genomics*, **1**, 35–43.
- Shaw, C.H., Carter, G.H., Watson, M.D. and Shaw, C.H. (1984) A functional map of the nopaline synthase promoter. *Nucleic Acids Res.* **12**, 7831–7846.
- Shirasawa-Seo, N., Sano, Y., Nakamura, S., Murakami, T., Gotoh, Y., Naito, Y., Hsia, C.N., Seo, S., Mitsuhashi, I., Kosugi, S. and Ohashi, Y. (2005) The promoter of Milk vetch dwarf virus component 8 confers effective gene expression in both dicot and monocot plants. *Plant Cell Rep.* **24**, 155–163.
- Shivprasad, S., Pogue, G.P., Lewandowski, D.J., Hidalgo, J., Donson, J., Grill, L.K. and Dawson, W.O. (1999) Heterologous sequences greatly affect foreign gene expression in tobacco mosaic virus-based vectors. *Virology*, **255**, 312–323.
- Smith, M.L., Mason, H.S. and Shuler, M.L. (2002) Hepatitis B surface antigen (HBsAg) expression in plant cell culture: kinetics of antigen accumulation in batch culture and its intracellular form. *Biotechnol. Bioeng.* **80**, 812–822.
- Smolenska, L., Roberts, I.M., Learmonth, D., Porter, A.J., Harris, W.J., Wilson, M.A. and Santa Cruz, S. (1998) Production of a functional single chain antibody attached to the surface of a plant virus. *FEBS Lett.* **441**, 379–382.
- Sojikul, P., Buehner, N. and Mason, H.S. (2003) A plant signal peptide-hepatitis B surface antigen fusion protein with enhanced stability and immunogenicity expressed in plant cells. *Proc. Natl. Acad. Sci. USA*, **100**, 2209–2214.
- Spiker, S. and Thompson, W.F. (1996) Nuclear matrix attachment regions and transgene expression in plants. *Plant Physiol.* **110**, 15–21.
- Staub, J.M. and Maliga, P. (1995) Expression of a chimeric *uidA* gene indicates that polycistronic mRNAs are efficiently translated in tobacco plastids. *Plant J.* **7**, 845–848.
- Staub, J.M., Garcia, B., Graves, J., Hajdukiewicz, P.T.J., Hunter, P., Nehra, N., Paradkar, V., Schlittler, M., Carroll, J.A., Spatola, L., Ward, D., Ye, G. and Russell, D.A. (2000) High-yield production of a human therapeutic protein in tobacco chloroplasts. *Nat. Biotechnol.* **18**, 333–338.
- Stefanov, I., Ilubae, S., Feher, A., Margoczi, K. and Dudits, D. (1991) Promoter and genotype dependent transient expression of a reporter gene in plant protoplasts. *Acta Biol. Hung.* **42**, 323–330.
- Streatfield, S.J. (2005) Regulatory issues for plant-made pharmaceuticals and vaccines. *Expert Rev. Vaccines*, **4**, 591–601.
- Streatfield, S.J. and Howard, J.A. (2003) Plant production systems for vaccines. *Expert Rev. Vaccines*, **2**, 763–775.
- Streatfield, S.J., Jilka, J.M., Hood, E.E., Turner, D.D., Bailey, M.R., Mayor, J.M., Woodard, S.L., Beifuss, K.K., Horn, M.E., Delaney, D.E., Tizard, I.R. and Howard, J.A. (2001) Plant-based vaccines: unique advantages. *Vaccine*, **19**, 2742–2748.
- Streatfield, S.J., Mayor, J.M., Barker, D.K., Brooks, C., Lamphear, B.J., Woodard, S.L., Beifuss, K.K., Vicuna, D.V., Massey, L.A., Horn, M.E., Delaney, D.D., Nikolov, Z.L., Hood, E.E., Jilka, J.M. and Howard, J.A. (2002) Development of an edible subunit vaccine in corn against enterotoxigenic strains of *Escherichia coli*. *In Vitro Cell. Dev. Biol.-Plant*, **38**, 11–17.
- Streatfield, S.J., Lane, J.R., Brooks, C.A., Barker, D.K., Poage, M.L., Mayor, J.M., Lamphear, B.J., Drees, C.F., Jilka, J.M., Hood, E.E. and Howard, J.A. (2003) Corn as a production system for human and animal vaccines. *Vaccine*, **21**, 812–815.
- Streatfield, S.J., Magallanes-Lundback, M.E., Beifuss, K.K., Brooks, C.A., Harkey, R.L., Love, R.T., Bray, J., Howard, J.A., Jilka, J.M. and Hood, E.E. (2004) Analysis of the maize polyubiquitin-1 promoter heat shock elements and generation of promoter variants with modified expression characteristics. *Transgenic Res.* **13**, 299–312.
- Svab, Z. and Maliga, P. (1993) High-frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene. *Proc. Natl. Acad. Sci. USA*, **90**, 913–917.
- Tacket, C.O., Pasetti, M.F., Edelman, R., Howard, J.A. and Streatfield, S.J. (2004) Immunogenicity of recombinant LT-B delivered orally to humans in transgenic corn. *Vaccine*, **22**, 4385–4389.
- Taschner, P.E.M., van der Kuyl, A.C., Neeleman, L. and Bol, J.F. (1991) Replication of incomplete alfalfa mosaic virus genome in plants transformed with viral replicase genes. *Virology*, **181**, 445–450.
- Teymouri, F., Alizadeh, H., Laureano-Perez, L., Dale, B. and Sticklen, M. (2004) Effects of ammonia fiber explosion treatment on activity of endoglucanase from *Acidothermus cellulolyticus* in transgenic plant. *Appl. Biochem. Biotechnol.* **113–116**, 1183–1191.
- Thompson, A.J. and Myatt, S.C. (1997) Tetracycline-dependent activation of an upstream promoter reveals transcriptional interference between tandem genes within T-DNA in tomato. *Plant Mol. Biol.* **34**, 687–692.
- Ting, J.T., Lee, K., Ratnayake, C., Platt, K.A., Balsamo, R.A. and Huang, A.H. (1996) Oleosin genes in maize kernels having diverse oil contents are constitutively expressed independent of oil contents. Size and shape of intracellular oil bodies are determined by the oleosins/oils ratio. *Planta*, **199**, 158–165.
- Tregoning, J.S., Nixon, P., Kuroda, H., Svab, Z., Clare, S., Bowe, F., Fairweather, N., Ytterberg, J., van Wijk, K.J., Dougan, G. and Maliga, P. (2003) Expression of tetanus toxin fragment C in tobacco chloroplasts. *Nucleic Acids Res.* **31**, 1174–1179.
- Verdaguer, B., de Kochko, A., Beachy, R.N. and Fauquet, C. (1996) Isolation and expression in transgenic tobacco and rice plants of the cassava vein mosaic virus (CVMV) promoter. *Plant Mol. Biol.* **31**, 1129–1139.
- Vierstra, R.D. (1996) Proteolysis in plants: mechanisms and functions. *Plant Mol. Biol.* **32**, 275–302.
- Wenderoth, I. and von Schaewen, A. (2000) Isolation and characterization of plant N-acetyl glucosaminyltransferase I (GntI) cDNA sequences. Functional analyses in the Arabidopsis cgl mutant and in antisense plants. *Plant Physiol.* **123**, 1097–1108.
- Woodard, S.L., Mayor, J.M., Bailey, M.R., Barker, D.K., Love, R.T., Lane, J.R., Delaney, D.E., McComas-Wagner, J.M., Mallubhotla, H.D., Hood, E.E., Dangot, L.J., Tichy, S.E. and Howard, J.A. (2003) Maize-derived bovine trypsin: characterization of the first

- large-scale, commercial protein product from transgenic plants. *Biotechnol. Appl. Biochem.* **37**, 1–8.
- Xie, M., He, Y. and Gan, S. (2001) Bidirectionalization of polar promoters in plants. *Nat. Biotechnol.* **19**, 677–679.
- Xie, Y., Liu, Y., Meng, M., Chen, L. and Zhu, Z. (2003) Isolation and identification of a super strong plant promoter from cotton leaf curl Multan virus. *Plant Mol. Biol.* **53**, 1–14.
- Xue, G.P., Patel, M., Johnson, J.S., Smyth, D.J. and Vickers, C.E. (2003) Selectable marker-free transgenic barley producing a high level of cellulase (1,4-beta-glucanase) in developing grains. *Plant Cell Rep.* **21**, 1088–1094.
- Yang, D., Wu, L., Hwang, Y.-S., Chen, L. and Huang, N. (2001) Expression of the REB transcriptional activator in rice grains improves the yield of recombinant proteins whose genes are controlled by a *Reb*-responsive promoter. *Proc. Natl. Acad. Sci. USA*, **98**, 11 438–11 443.
- Young, T.E., Giesler-Lee, J. and Gallie, D.R. (2004) Senescence-induced expression of cytokinin reverses pistil abortion during maize flower development. *Plant J.* **38**, 910–922.
- Yusibov, V., Hooper, D.C., Spitsin, S.V., Fleysh, N., Kean, R.B., Mikheeva, T., Deka, D., Karasev, A., Cox, S., Randall, J. and Koprowski, H. (2002) Expression in plants and immunogenicity of plant virus-based experimental rabies vaccine. *Vaccine*, **20**, 3155–3164.
- Zhong, G.-Y., Peterson, D., Delaney, D.E., Bailey, M., Witcher, D.R., Register, J.C. III, Bond, D., Li, C.-P., Marshall, L., Kulisek, E., Ritland, D., Meyer, T., Hood, E.E. and Howard, J.A. (1999) Commercial production of aprotinin in transgenic maize seeds. *Mol. Breed.* **5**, 345–356.
- Zuo, J. and Chua, N.-H. (2000) Chemical-inducible systems for regulated expression of plant genes. *Curr. Opin. Biotechnol.* **11**, 146–151.