

# Opportunities for recombinant antigen and antibody expression in transgenic plants—technology assessment

Stefan Schillberg<sup>a,\*</sup>, Richard M. Twyman<sup>b</sup>, Rainer Fischer<sup>a,c</sup>

<sup>a</sup> *Fraunhofer-Institute for Molecular Biology and Applied Ecology, Worringerweg 1, 52074 Aachen, Germany*

<sup>b</sup> *Department of Biology, University of York, Heslington, York YO10 5DD, UK*

<sup>c</sup> *Institute for Molecular Biotechnology, RWTH Aachen, Worringerweg 1, 52074 Aachen, Germany*

Available online 18 November 2004

## Abstract

Plants are now gaining widespread acceptance as a general platform for the large-scale production of recombinant proteins. The principle has been demonstrated by the success of a diverse repertoire of proteins, with therapeutic molecules showing the most potential for added value. Over the past 10 years, several efficient plant-based expression systems have emerged. However, a number of issues remain to be addressed before plant bioreactors can be accepted and adopted widely in preference to the established microbial and mammalian platforms. Overcoming bottlenecks imposed by low yields, poor and inconsistent product quality and difficulties with downstream processing are the most important goals for researchers working in this field. The achievement of these goals is conditional on the development of extraction and processing steps that comply with GMP standards, including extensive quality assurance and control procedures. Such rigorous and validated standards should be combined with measures applied earlier in production to ensure product sustainability and quality, such as the use of master seed banking procedures. Moreover, there are several further challenges concerning topics of environmental impact, biosafety and risk assessment, which reflect the release of transgenic plants, as well the safety of the plant-derived products themselves.

We are facing a growing demand for protein diagnostics and therapeutics, but lack the capacity to meet those demands using established facilities. A shift to plant bioreactors may, therefore, become necessary within the next few years, making it more imperative that the technical and regulatory limitations are addressed and solved. The production of pharmaceutical proteins in plants will only realize its huge potential if the products are provided at consistent high quality levels, allowing the delivery of clinical grade proteins that will gain regulatory approval and which can be used routinely in clinical trials.

© 2004 Elsevier Ltd. All rights reserved.

**Keywords:** Plant-made pharmaceutical; Transgenic plant; Recombinant protein production

## 1. Introduction

The advent of recombinant protein technology has allowed the exploration of novel expression systems for pharmaceutical protein production. Plants have emerged as a convenient, safe and economical alternative to the mainstream expression systems, which are based on the large-scale culture of microbes or animal cells, or on transgenic animals [1–4]. Extensive research over the past two decades has shown that a wide range of valuable proteins can be expressed effi-

ciently in plants. Examples include human serum proteins and growth regulators, antibodies, vaccines, industrial enzymes, biopolymers and molecular biology reagents [5,6]. The success of these experiments suggests that plant systems could be used to produce recombinant proteins on a commercial scale. This is known as Molecular Farming, and despite initial scepticism, the first generation of recombinant proteins produced in transgenic plants is now reaching commercial status.

In this paper, we discuss the advantages and disadvantages of different production crops that have been used to produce antigens and antibodies, including leafy crops, cereals, legumes, fruits and vegetables. We also discuss produc-

\* Corresponding author.

E-mail address: [schillberg@ime.fraunhofer.de](mailto:schillberg@ime.fraunhofer.de) (S. Schillberg).

tion bottlenecks caused by constraints on the product yields, the control of raw material quality and challenges with extraction and downstream processing stages of production.

## 2. Plant transformation technology

Techniques have been developed for the expression of proteins in stably transformed plants (where the gene is incorporated into the plant genome), or in transiently transformed plant tissues. Transient gene expression is rapid compared to stable plant transformation and gives results in days. However, transient gene expression is limited in scale and is generally used to test the constructs used for protein expression before stable transformation is performed.

### 2.1. Expression in stably transformed plants

Stable plant transformation is defined as the genomic integration of a transgene. Both the nuclear [7] and plastid [8,9] genomes can be transformed through a variety of transgenesis methods. The principle transformation technologies currently used in plant biotechnology are *Agrobacterium*-mediated gene transfer to dicots, such as tobacco and pea [7], or biolistic delivery of genes to monocots, such as wheat, rice and corn [10]. Ease of transformation has been a critical bottleneck in plant biotechnology and still influences the choice of crops used in Molecular Farming. The basis of stable transformation is well understood but is uniformly time consuming. Three to nine months are needed in our hands, depending on the plant variety, to generate stably transformed plants that can be used for testing the function and characteristics of the expressed protein.

### 2.2. Transient expression systems

Transient expression can be used to test the function of an expression construct before progressing to large-scale production in stable transgenic lines, but it is also possible to rely completely on transient expression for protein production. A useful method is agroinfiltration, in which recombinant *Agrobacterium tumefaciens* are infiltrated into plant tissue [11]. The T-DNA is transferred to the nucleus in a large number of plant cells resulting in the production of milligram amounts of recombinant protein within a few days [12].

Viral vectors have also attracted interest because viral infections are rapid and systemic, and infected cells yield large amounts of virus and viral gene products [13,14]. Since plant viruses do not integrate into the genome, there is no stable transformation and the transgene is not passed through the germ line. However, plant viruses often have a wide host range, are easily transmissible by mechanical inoculation and can spread from plant to plant, making it possible to infect large numbers of plants rapidly. For example, plant viral vectors have been used to express scFvs [14–16] and full-size antibodies [14]. In the latter case, two tobacco mosaic virus

vectors were constructed carrying the heavy and light chains of the antibody. Tobacco plants were co-infected with the two vectors and both transgenes were expressed. Assembly of the antibody *in planta* was confirmed.

## 3. Choice of host species and production system

A large number of species are now amenable to Molecular Farming including model plants (tobacco, *Arabidopsis*), cereal crops (rice, wheat, maize), legumes (pea, soybean, alfalfa), fruit crops (tomato, banana) and solanaceous species (potato). Thus far, it has been difficult to evaluate the relative performance of different crops for industrial or pharmaceutical Molecular Farming because this requires the production of the same protein in a range of hosts, using a ‘standardized’ expression construct. The performance of an expression construct across species is itself difficult to judge since the same promoter (or other regulatory sequence) may have a different intrinsic level of activity in different genetic backgrounds. Therefore, different research groups and companies have concentrated instead on optimising their ‘favourite’ system, i.e. the host species for which there is the greatest in-house expertise and experience.

We have performed a valid comparison of different production hosts using a scFv antibody that recognizes the carcinoembryonic antigen (CEA), one of the best characterized tumour-associated antigens [17,18]. By focussing on promoters with the highest and most widespread activities in each species (generally the CaMV 35S promoter for dicots and the *ubi-1* promoter for monocots), the need to factor in promoter performance across species was eliminated. We have shown that across all the species tested, accumulation of the scFv in the endoplasmic reticulum by means of the KDEL retrieval signal significantly improved protein levels, compared to antibodies secreted into the apoplast. Moreover, the study demonstrated that although production levels (measured in terms of the amount of recombinant protein generated per kg biomass) varied among the different plants and systems, they were broadly comparable, so production level alone was not a critical factor in expression host choice. Much more importance could be attached to other biological and geographical factors, which should be evaluated on a case-by-case basis taking into account the market value of the recombinant protein. These factors include:

- *Set up costs*: the set-up cost for any transgenic crop is generally high, reflecting the requirement for controlled environment rooms and greenhouses during plant regeneration. The actual time required depends on the species, and can be several months in the case of cereals and legumes. These costs are minimal for transient expression systems, such as agroinfiltration or virus infection. For cell suspension cultures, the cost of fermenters and media has to be included.

- *Scale-up and maintenance costs*: the costs associated with scaling up and maintaining crops of transgenic plants are minimal because standard farming practices are sufficient. The choice of crop may, therefore, depend on geographical factors such as the available land, the cost of labour and the distance to the nearest processing plant or distribution centre. Conversely, scaling up and maintaining agroinfiltration or fermenter-based systems can be very expensive. The advantages of transgenic plants, therefore, increase in proportion to the scale of production.
- *Length of production cycle*: the disadvantages of low yields and high running costs can in some cases be balanced by shorter and more frequent production cycles. This is particularly the case for fermenter-based systems, where the costs are high but the production cycle can be reduced to a matter of days. When comparing different transgenic crops, it may be wise to consider the length of the growth cycle, the number of times per year a species can be grown and harvested. Tobacco and alfalfa, for example, can each be harvested several times in a year.
- *Biomass yield*: although the level of recombinant protein production varies only moderately among different species, the actual biomass produced per hectare of plants varies considerably. For example, a comparison of tomato and pea shows that, due to the high biomass yield of tomato, for every hectare of peas only one third of a hectare of tomatoes would be required to produce the same amount of protein, even though the yield of recombinant protein per unit biomass is lower in tomato than pea. Tobacco has the highest biomass yield of all crop plants used for Molecular Farming.
- *Costs of processing*: the downstream processing of recombinant proteins (extraction, purification and characterization) can make up 80% or more of total production costs [19]. Minimizing the costs of processing is, therefore, of prime importance and this can be achieved by choosing a host species for which processing infrastructure is already locally available, and choosing a host system from which protein extraction is easy. It is easier to extract proteins from watery tissues, such as tomato fruits than it is from dry material, such as cereal grains, which have to be milled, ground, homogenized and extracted by sedimentation, filtration and concentration. Ease of extraction is also a major advantage of cell suspension cultures, especially when the recombinant protein is secreted into the medium [20]. However, recombinant proteins can be modified by proteases and phenolic compounds released from watery plant tissue, while these tend to be present in lower amounts in cereal grains. Tobacco, while advantageous in terms of biomass yield, can contain toxic metabolites that have to be removed. This is not a problem with edible crops.
- *Edibility*: for certain recombinant proteins, such as vaccines and antibodies, it may be advantageous to use edible plant organs as the production vehicle (e.g. tomato fruits, potato tubers) as this allows direct oral administration with no processing costs at all. The same applies to industrial

proteins, such as phytase and  $\alpha$ -amylase, which are used to increase the nutritional value of animal feed.

- *Costs of storage and distribution*: this takes into consideration the nature of the processing and distribution network. If there is a short distance between the field and the processing plant, and the distribution network is small, storage and distribution costs are low. The converse applies if the distances are larger. If this is a critical economic consideration then the production of recombinant proteins in cereal seeds, which can be stored for months or even years at ambient temperatures without the protein losing activity, is better than production in tomatoes, which must be stored and transported in chilled containers, or tobacco and alfalfa leaves, which must be desiccated or frozen to preserve protein stability [21].
- *Costs of containment*: where containment is an issue, plant cell cultures grown in industrial scale fermenters are optimal since these are closed vessels. However, tomatoes are also attractive since these are grown in greenhouses. In any crop species, chloroplast-expression of the transgene will result in natural containment since functional chloroplast DNA is not transmitted through the pollen.

The compromise between production costs and profit is likely to be a key issue in selecting the most suitable species for Molecular Farming because most pharmaceuticals will be produced by industry. We predict that these costs will dictate which crop or crops become generally accepted for recombinant antibody production. Some properties of major transgenic crop species are compared in Table 1.

#### 4. Constraints on the product yields

The yield of recombinant proteins produced in a plant system depends on three factors: the intrinsic limitations of the production host and expression system, limitations imposed by the level of transgene expression, and limitations imposed by the stability of the recombinant protein. The intrinsic yield potential of the production crop depends partly on its biomass yield per hectare (see above), partly on the protein content of the plant tissue (which is highest in legume seeds) and partly on the extent to which the endogenous protein synthesis machinery can be diverted to produce the recombinant protein of interest [22]. However, when considering a production host, it is also necessary to take into account factors such as the availability of agricultural infrastructure and processing technology, the throughput and efficiency of first-generation transformation and regeneration, and the identification of high producer lines that perform stably in the field. Such lines should comply with all relevant biosafety and regulatory standards [23].

Factors affecting the level of transgene expression in plants include the promoter activity and specificity, the optimization of mRNA stability and translational efficiency, and the removal of spurious AU-rich sequences that may act as cryp-

Table 1  
Comparison of major transgenic crops used for Molecular Farming (data adapted from Stoger et al. [17])

	Tobacco	Rice	Tomato	Wheat	Maize	Pea
Favoured transformation method	<i>Agrobacterium</i>	Particle bombardment	<i>Agrobacterium</i>	Particle bombardment	Particle bombardment	<i>Agrobacterium</i>
Favoured expression strategy	Leaves	Seed	Fruit	Seed	Seed	Seed
Biomass yield (kg ha <sup>-1</sup> )	>100,000	6570	67,900	2870	8880	2400
Producer price (US\$ per tonne)	4180	160	180	140		300
Storage and distribution	Freeze dried or frozen	Ambient	Chilled (4 °C)	Ambient	Ambient	Ambient
Stability under normal storage conditions	Several months	>1 year	Several weeks	>1 year	>1 year	>1 year
Containment	Field, flowering may be avoided	Field, self-pollinating	Greenhouse	Field, self-pollinating	Field	Field, self-pollinating
Other comments	Contains toxic metabolites	Edible	Edible raw	Edible	Edible	Edible

tic splice sites, all of which can be controlled by expression cassette design [18]. Other factors, such as transgene copy number and organization are more difficult to control in this manner, and still the best method remains to generate a population of transgenic lines and select those with the best performance. Protein stability is probably the single most important factor limiting yields in Molecular Farming, and this can be addressed at least in part by appropriate sub-cellular targeting. For example, recombinant antibodies targeted to the secretory system generally accumulate to much higher levels (>1000-fold) than those synthesized in the cytosol, and further yield increases occur when they are retained in the endoplasmic reticulum rather than secreted to the apoplast. It is also important to remember that the ability of the host plant to produce a recombinant protein depends on its spare metabolic capacity. For example, a recombinant protein that is particularly demanding for a rare amino acid is unlikely to accumulate to high levels.

## 5. Control of raw material and product quality

The raw material in question is the fresh, harvested plant tissue, or tissue that has been pulped or shredded. Several quality aspects can be considered, including the biosafety of the raw material, and the quality and homogeneity of the product itself. Biosafety factors are in some cases dependent on species, for example, production in tobacco leaves comes with the associated risks of contamination with nicotine and other alkaloids. In other cases, they depend on agricultural practices—for example, the backcrossing in elite lines that show maximum performance in the field, influence of fertilizers and environmental conditions on recombinant protein yields as well as potential for contamination with agrochemicals or fertilizers. Finally, the presence of allergens is a risk where the protein is intended for therapeutic use. Strategies also have to be applied to avoid horizontal gene transfer and out-crossing events, and to monitor product flow to ensure traceability, using for example identity preserved varieties [23].

The quality of the product itself depends on its authenticity, i.e. similarity to the native counterpart. Unless a fusion tag of some description is used, the amino acid sequences of a plant-derived recombinant protein and the equivalent native protein are usually the same. However, certain forms of post-translational modification are not carried out in plants (e.g. hydroxylation of proline residues) and there are subtle differences in glycan chain structures, such as the absence of sialic acid and the presence of core  $\beta(1,3)$ xylose and  $\alpha(1,2)$ fucose residues [24]. Glycan structures affect the structure, folding and interactions of glycoproteins, and may, therefore, influence their distribution, longevity and activity. Foreign glycan structures may also be immunogenic. As well as the glycan structure per se, the homogeneity of plant-derived glycoproteins is also important for batch-to-batch consistency. This can be improved using bioreactor systems (such as *Chlamydomonas reinhardtii* or *Lemna minor*) where growth conditions can be controlled, but it also depends to a certain extent on species. For example, while tobacco produces very heterogeneous glycans, glycoproteins produced in alfalfa have homogenous glycan chains [25]. The presence of fusion tags also affects the activity and structure of recombinant proteins, and these are not considered acceptable for therapeutic proteins. Therefore, if a fusion is used to control targeting or facilitate purification, it must be removed during the processing steps before the product is used.

## 6. Extraction and processing

Not all recombinant proteins produced in plants need processing to the same extent. At one extreme are proteins intended for intravenous use in humans, which must be purified to the highest standards. At the other extreme are proteins that can be utilized in raw, unprocessed plant material (e.g. vaccine candidates and industrial enzymes used for food processing). Between the extremes are proteins that need greater or lesser degrees of purification for their intended uses. The purification strategies employed, and the bottlenecks encountered, depend on the expression host and tissue. Leaves, seeds, fruits and vegetables can all be processed in much the same

manner, but the processing of leaves must take place immediately after harvest (or they must be dried or frozen) to avoid protein degradation, while fruits can be chilled, and vegetables and seeds can be stored for long periods at room temperature without significant loss of activity.

To make protein production in plants competitive and sustainable, it is important to reduce the amount of bio-waste and to dispose properly of all the buffers and reagents used during extraction and processing. For example, leafy crops can be dried to restrict the buffer volumes necessary for extraction, and the reagents used in downstream processing could be regenerated and recycled, e.g., through the use of immobilized enzymes.

Purification strategies may be based on standard chromatographic or electrophoretic procedures, but methods can be devised to affinity purify specific products (e.g. using immobilized antigens to purify recombinant antibodies or the development of stable and cheap synthetic ligands) or to purify proteins with particular fusion tags. One of the most promising of these is the oleosin fusion platform developed at SemBioSys Genetics Inc., in which recombinant proteins are expressed fused to the oil-body-specific endogenous protein oleosin in rapeseed and safflower. This is followed by a simple and inexpensive purification scheme involving the separation of oil bodies. As discussed above, the presence of fusion proteins is incompatible with therapeutic use, and the tag must be removed. In the oleosin technology, this is achieved by *in vitro* endoproteolytic processing.

## 7. Conclusions

Proof of principle for the synthesis of pharmaceutical proteins in plants was demonstrated in transgenic tobacco, and this crop is at the forefront of emerging systems for plant-based commercial protein production. The high biomass yields, robust transformation technology and strong biosafety profile from this non-food/non-feed crop make it an attractive production vehicle. Other emerging systems include alternative leafy crops, such as alfalfa, which as a leguminous species requires limited fertilizer input, the use of cereal and legume seed crops for stable antibody accumulation in dry seeds, and the use of fruit and vegetable crops to combine storage with ease of processing and product administration.

Overcoming bottlenecks imposed by low yields, poor and inconsistent product quality and difficulties with downstream processing are the most important goals for researchers working in the field of Molecular Farming. However, achieving these goals is conditional on the development of extraction and processing steps that comply with GMP standards, including extensive quality assurance and control. Such rigorous and validated standards should be combined with measures applied earlier in production to ensure product sustainability and quality, such as the use of master seed banking procedures. In the future, it is hoped that re-

combinant proteins can be expressed in plants to consistent and high quality levels, allowing the production of pharmaceutical proteins that can be used routinely in clinical trials.

## References

- [1] Fischer R, Drossard J, Emans N, Commandeur U, Hellwig S. Towards molecular farming in the future: *Pichia pastoris*-based production of single-chain antibody fragments. *Biotechnol Appl Biochem* 1999;30:117–20.
- [2] Giddings G. Transgenic plants as protein factories. *Curr Opin Biotechnol* 2001;12:450–4.
- [3] Houdebine LM. Antibody manufacture in transgenic animals and comparisons with other systems. *Curr Opin Biotechnol* 2002;13:625–9.
- [4] Larrick JW, Thomas DW. Producing proteins in transgenic plants and animals. *Curr Opin Biotechnol* 2001;12:411–8.
- [5] Hood EE, Woodard SL, Horn ME. Monoclonal antibody manufacturing in transgenic plants—myths and realities. *Curr Opin Biotechnol* 2002;13:630–5.
- [6] Twyman RM, Stoger E, Schillberg S, Christou P, Fischer R. Molecular farming in plants: host systems and expression technology. *Trends Biotechnol* 2003;21:570–8.
- [7] Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT. A simple and general method for transferring genes into plants. *Science* 1985;227:1229–31.
- [8] Maliga P. Progress towards commercialization of plastid transformation technology. *Trends Biotechnol* 2003;21:20–8.
- [9] Staub JM, Garcia B, Graves J, Hajdukiewicz PT, Hunter P, Nehra N, et al. High-yield production of a human therapeutic protein in tobacco chloroplasts. *Nat Biotechnol* 2000;18:333–8.
- [10] Christou P. Particle bombardment. *Methods Cell Biol* 1995;50:375–82.
- [11] Kapila J, DeRycke R, VanMontagu M, Angenon G. An *Agrobacterium*-mediated transient gene expression system for intact leaves. *Plant Sci* 1997;122:101–8.
- [12] Vaquero C, Sack M, Chandler J, Drossard J, Schuster F, Monecke M, et al. Transient expression of a tumor-specific single-chain fragment and a chimeric antibody in tobacco leaves. *Proc Natl Acad Sci USA* 1999;96:11128–33.
- [13] Gleba Y, Marillonnet S, Klimyuk V. Engineering viral expression vectors for plants: the ‘full virus’ and the ‘deconstructed virus’ strategies. *Curr Opin Plant Biol* 2004;7:182–8.
- [14] Verch T, Yusibov V, Koprowski H. Expression and assembly of a full-length monoclonal antibody in plants using a plant virus vector. *J Immunol Methods* 1998;220:69–75.
- [15] Hendy S, Chen ZC, Barker H, Santa Cruz S, Chapman S, Torrance L, et al. Rapid production of single-chain Fv fragments in plants using a potato virus X episomal vector. *J Immunol Methods* 1999;231:137–46.
- [16] McCormick AA, Kumagai MH, Hanley K, Turpen TH, Hakim I, Grill LK, et al. Rapid production of specific vaccines for lymphoma by expression of the tumor-derived single-chain Fv epitopes in tobacco plants. *Proc Natl Acad Sci USA* 1999;96:703–8.
- [17] Stoger E, Sack M, Perrin Y, Vaquero C, Torres E, Twyman RM, et al. Practical considerations for pharmaceutical antibody production in different crop systems. *Mol Breed* 2002;9:149–58.
- [18] Schillberg S, Fischer R, Emans N. Molecular farming of recombinant antibodies in plants. *Cell Mol Life Sci* 2003;60:433–45.
- [19] Kusnadi AR, Hood EE, Witcher DR, Howard JA, Nikolov ZL. Production and purification of two recombinant proteins from transgenic corn. *Biotechnol Prog* 1998;14:149–55.
- [20] Doran PM. Foreign protein production in plant tissue cultures. *Curr Opin Biotechnol* 2000;11:199–204.

- [21] Khoudi H, Laberge S, Ferullo JM, Bazin R, Darveau A, Castonguay Y, et al. Production of a diagnostic monoclonal antibody in perennial alfalfa plants. *Biotechnol Bioeng* 1999;64:135–43.
- [22] Twyman RM. Host plants, systems and expression strategies for molecular farming. In: Fischer R., Schillberg S. (Eds.), *Molecular farming: plant-made pharmaceuticals and technical proteins*. Wiley VCH, 2004, p. 191–216.
- [23] Commandeur U, Twyman RM, Fischer R. The biosafety of molecular farming in plants. *AgBiotechNet* 2003;5:ABN 110.
- [24] Gomord V, Faye L. Posttranslational modification of therapeutic proteins in plants. *Curr Opin Plant Biol* 2004;7:171–81.
- [25] Fischer R, Stoger E, Schillberg S, Christou P, Twyman RM. Plant-based production of biopharmaceuticals. *Curr Opin Plant Biol* 2004;7:152–8.