Molecular Farming: Strategies, Expression Systems and Bio-Safety Considerations

Kafeel AHMAD

Department of Biology, University of Leicester, Leicester, United Kingdom

Abstract


Molecular farming is an experimental application of biotechnology that involves the genetic modification of crops for the production of proteins and chemicals for medicinal and commercial purposes. The vast majority in the developing world cannot afford the high cost of therapeutics produced by existing methods. We need to produce not only new therapeutics but also cheaper versions of the existing ones. Molecular farming could offer a viable option for this growing need for biopharmaceuticals. Plant made therapeutics are cheaper, safer, can be abundantly produced and easily stored. Here, strategies and approaches utilized in plant molecular farming are discussed. Furthermore, the bio-safety considerations related to this emerging field are also discussed.

Keywords: biopharmaceuticals; recombinant proteins; rhizosecretion; therapeutics; transformation

Molecular farming can be described as an experimental application of biotechnology to genetically engineered crops in order to produce proteins and chemicals for pharmaceutical and other commercial interests (Franken et al. 1997). Evidence of using plants for medicinal purposes can be traced back to as early as the Neanderthal period (about 130 000 years ago), when plants were used to aid healing of wounds (Kleiner 1995). Romans used willow bark for the treatment of fever which was later identified to contain aspirin (Mewett et al. 2007). During the 16th century medicinal plants were grown for teaching medicine and as a source of treatment for various diseases (Akeele 1993). The use of plants and plant extracts for medicinal purposes flourished until the 17th century when more scientific pharmacological treatments were recommended (Trevelyan 1993; Winslow & Kroll 1998). One fourth of the currently used medicines still have a plant origin (Winslow & Kroll 1998).

Genetic engineering has recently opened up new opportunities for using plants as production factories for biopharmaceuticals. Human growth hormone was the first pharmaceutically important protein that was expressed in transgenic tobacco (Barta et al. 1986). Since then transgenic plants expressing vaccines, therapeutics, industrial enzymes, antibodies, nutraceuticals, and other pharmaceutical proteins have been produced (Krebbers et al. 1992; Whitelam et al. 1993; Ma et al. 1994; Herbers & Sonnewald 1999; Ma et al. 2005).

Historically both prokaryotic (Georgiou & Bowden 1991) and eukaryotic (Harashima 1994; Lubon et al. 1996) systems have been utilized to express recombinant proteins. Prokaryotic production systems are comparatively inexpensive and convenient compared to mammalian systems in terms of the technology/equipment required.

However, many mammalian proteins require post-translational modifications such as protein glycosylation for their biological activity which cannot be performed by prokaryotic production systems. The use of prokaryotic expression systems is therefore perhaps limited. On the other hand, the cost of producing proteins in mammalian cells is very high in terms of maintaining cell cultures and scale up. In contrast, pathways of protein synthesis and modification in plants are very similar to those in animals facilitating protein modification analogous to that seen in animal cells (Cabanes-Macheteau et al. 1999). Mass production of heterologous proteins and biopharmaceuticals in plants can be achieved at much lower costs and contamination or co-purification of human or animal pathogens.

Although protein glycosylation pathways are predominantly similar between humans and plants, minor differences do exist in glycosylation. These include, for example, the absence of terminal sialic acid residues in the plant derived glycan chains and the existence of plant based residues (CABANES-MACHETEAU et al. 1999). These differences could make a recombinant protein inactive, harmful or immunogenic. Hence, these glycosylation pathways need to be humanized for expression of therapeutic proteins. For example, human glycosyltransferase could be co-engineered into plants along with the transgene of interest (BAKKER et al. 2001). The recombinant protein could also be targeted to the endoplasmic reticulum, which could result not only in the lack of plant specific glycosylation but also in increased accumulation of recombinant protein (CABANES-MACHETEAU et al. 1999; LOOS & STEINKELLNER 2012; NAGELS et al. 2012).

**Strategies for optimum protein production**

**Selection of plant type for transformation.** Theoretically, any type of plant can be transformed; however, it is useful to use plants which are well studied and characterized allowing effective risk assessment and tracking of the transgene of interest. The appropriate type of plant selection is important for efficient protein production. Proper consideration should be given to use plants species with the least chance(s) of dissemination of transgene/gene product to other plants/environment. Use of self-pollinating plants can reduce chances of gene transfer to other plants. The choice of host plant also depends on the recombinant protein to be expressed, the host plant life cycle, biomass yield and production costs (SHARMA & SHARMA 2009). If the protein is to be expressed in green tissues, then leafy plants like tobacco may serve as an ideal plant with the huge biomass production capacity (FISCHER & EMANS 2000). Tobacco has been used as the system of choice for a number of plant-derived recombinant proteins (SEXTON et al. 2009). Tobacco has the advantage of producing huge quantities of green leaf material per acre. The agrobacterium-mediated transformation is highly efficient in tobacco. The plant has a prolific seed production which could facilitate biomass scale-up. Tobacco plants mainly self-pollinate, so there is little risk of transfer of genetic material to other plants. Tobacco is also a non-food crop, so there is little risk of food chain contamination (MA et al. 2003). An important issue in the use of tobacco as production system is the presence of undesirable high amounts of nicotine in tobacco leaves that could make the expressed proteins less usable. However, there are cultivars available in tobacco that produce reduced levels of unwanted secondary metabolites like nicotine (MENASSA et al. 2001). Alternatively, efficient purification systems for the removal of unwanted nicotine could be utilized to get nicotine-free recombinant proteins (FU et al. 2010). Other leafy crops that have been used for recombinant protein production include spinach (YUSIBOV et al. 2002), lettuce (RUHLMAN et al. 2007) and lucerne (KHOUDI et al. 1999). A disadvantage associated with leafy crops is that the expressed proteins may be unstable in the leaf environment which could interfere with yield and quality of the protein in question. Also the phenolic compounds released during an extraction process could be detrimental to downstream processing (MA et al. 2003). Cereals and legumes can also be used as alternative production systems to overcome some of these problems (RADENMACHER et al. 2008; TSUBOI et al. 2008). Seeds have specialised storage compartments which help in reducing protein degradation, the exposure of the recombinant protein to phenolic compounds is also avoided thus improving downstream processing (MA et al. 2003). A number of cereal crops like maize (RADENMACHER et al. 2008), rice (QIAN et al. 2008) and wheat (TSUBOI et al. 2008) have been utilized for recombinant protein production. Legumes like pea (PERRIN et al. 2000) and soybean (MORAVEC et al. 2007) have also been used to express foreign proteins. Oil crops offer another inexpensive platform for the expression of recombinant proteins. With oleosin fusion technology, developed by SemBioSys Genetics Inc. (http://www.sembiosys.com/), the recombinant protein gene sequence is fused to the sequence of an oil body specific endogenous protein oleosin in rapeseed and safflower, after purification the protein is separated by an endoprotease digestion (SCHILLBERG et al. 2005). Potato has also been exploited in the production of vaccines (MA et al. 2003). Other plants that have been used for biopharming include tomato (SANDHU et al. 2000), carrot (DANIELL et al. 2005), banana (TRIVEDI & NATH 2004) and papayas (CARTER & LANGRIDGE 2002).

**Sub-cellular targeting.** Consideration should be given to where the protein of interest should be
produced in plants. Targeting to sub-cellular compartments can help increase the production and recovery of the target proteins (FISCHER & EMANS 2000). The endoplasmic reticulum is an important site for the processing, disulphide bond formation, assembly and glycosylation of proteins (HELENIUS & AEBI 2001). Targeting proteins to the apoplast can result in a high level of expression and better downstream processing; however, the endoplasmic reticulum retention has resulted in 10- to 100-fold higher yields (CONRAD & FIEDLER 1998). Another important sub-cellular compartment with capability of several hundredfold accumulation of recombinant proteins is the chloroplast which is capable of correct folding and disulphide bond formation, but is not capable of glycosylation (RUF et al. 2001). Recombinant proteins have also been targeted to protein storage vacuoles (STOGER et al. 2000). Rhizosecretion or the secretion of recombinant protein in the hydroponic medium in roots is another strategy that can help in simplifying the downstream processing and increasing the protein yield (DRAKE et al. 2009).

**Plant transformation types**

**Nuclear transformation.** Stable transformation involves the integration of foreign gene/s in the genome of the plant. This can be achieved either by agrobacterium-mediated transformation of dicotyledonous plants or through biolistic delivery (gene-gun) methods in monocots. These transformations result in heritable expression of recombinant protein, which is stable from generation to generation, so it is good for long-term production of recombinant proteins. However, it is time consuming as to develop stably transformed plants takes months or years depending on the plant type used for recombinant protein expression. The strategy is also costly compared to transient expression. With nuclear transformation it is possible to target the protein of interest to various sub-cellular locations such as the nucleus, cytoplasm, endoplasmic reticulum, plastids, vacuole and apoplast. Correct sub-cellular targeting makes it possible for the correct posttranslational modifications to be carried out on the expressed protein. Another advantage of nuclear transformation is the high level of scalability that is possible as the stably transformed plants could be theoretically grown on huge acreages anywhere in the world. However, there is an inherent chance of transfer of genetic material to other crops, so special regulatory measures must be taken to prevent transgene escape to non-target crops.

**Transient expression.** During transient expression the foreign genetic material does not integrate into the genome of the plant. This can be achieved through agro-infiltration (using agro-bacteria), viral vectors or through biolistics. Transient expression offers some advantages, as data on whether a particular gene is being expressed or not can be obtained in days (KAPILA et al. 1996). Transient expression can be used in pilot experiments before proceeding to the time consuming and costly stable transformation. Problems in protein expression can be identified and corrected so that the chances of producing the desired protein through stable transformation are made more likely. Through agro-infiltration a number of genes can be expressed at the same time which can help in studying the combined effect of multiple transgenes which are expressed (JOHANSEN & CARRINGTON 2001). However, the yield of expressed protein is normally lower and the plant material requires processing immediately as due to the perishable nature of leaves degradation during the storage of plant tissue will result in further loss of the protein. Recently SAINSbury et al. (2009) have developed a number of binary vectors called pEAQ vectors for transient expression which could give high levels of transgene expression in a very short time. Furthermore, these vectors also allow the simultaneous expression of more than one protein.

**Chloroplast transformation.** The gene of interest can also be incorporated into the chloroplast genome. The commonly used method to transform chloroplasts involves using a gene gun to incorporate the transgene into the chloroplast genome (DANIELL 2006). A mature leaf cell contains up to 100 chloroplasts and each chloroplast can contain up to 100 copies of chloroplast genome. It is, therefore, likely that a higher yield of foreign protein could be produced from chloroplast transformation than from nuclear transformation (DANIELL 2006). The amount of recombinant protein could be as high as 46% of total soluble protein (DE COSA et al. 2001). Since the chloroplast is prokaryotic in nature, gene silencing is not observed unlike nuclear transformation. As the expressed protein is confined to the chloroplast, therefore it has no toxic effects on the host plant (DANIELL 2006). Chloroplasts are maternally inherited; therefore, transfer to other plants through pollen is avoided. The technique is most commonly used in tobacco. However, the chloroplast does not provide any post-translational modifications such as glycosylation and hence the technology is not adopted widely (BOEHM 2007). As with nuclear
transformation, chloroplast transformation is also time consuming as it requires the generation of stably transformed plants. Furthermore, different cloning vectors are required for transformation of different plant species (Verma & Daniell 2007).

Stability of recombinant DNA inside plant

One possible outcome of insertion of recombinant DNA into the plant genome is that the recombinant DNA can sometimes undergo inactivation, preventing its expression inside the plant cell. Recombinant DNA inactivation has been attributed to multiple copy integration, different base composition between recombinant DNA and the integration site, overexpression effects and detrimental effects of sequences adjacent to the recombinant DNA integration site (Finnegan & McElroy 1994). The presence of repeated homologous sequences, recombinant DNA methylation and co-suppression can also lead to recombinant DNA inactivation (Meyer & Saedler 1996). Recombinant DNA inactivation can be avoided by selecting lines with a single insertion of the transgene, not using repetitive homologous sequences, selecting stable recombinant lines and creating site-specific recombination systems (Finnegan & McElroy 1994).

Optimizing foreign gene sequences for expressing in plants

Plants have a different codon usage bias than animals, however foreign DNA can be optimized for expression in plants to increase translation and therefore obtain higher protein yields (Kusnadi et al. 1998). Expression can be increased further by the use of tissue specific promoters, improving transcript stability and the use of viral sequences for translational enhancement (Gallie 1998). The 35S promoter is suitable for dicotyledonous plants while the maize ubiquitin-1 promoter is normally used for monocot plants (Ma et al. 1995; Christensen & Quail 1996). The use of tissue specific promoters can also help in avoiding adverse effects on plant growth, development and environment (Cramer et al. 1999; Stoger et al. 2000).

Protein expression can be increased by the use of introns in the recombinant DNA molecule (Maas et al. 1991). However, the exact mechanism how introns enhance protein expression is not known. It has been found that transgene translation can be enhanced by the addition of untranslated leader sequence of alfalfa mosaic virus mRNA 4 (Datla et al. 1993). The stability of mRNA is also influenced by the polyadenylation sites in plant cells (Ingelbrecht et al. 1989; Hunt 1994). These sites protect enzymatic degradation of mRNA. There also exist specific recognition sites that result in RNA degradation (Sullivan & Green 1993). Some of these recognition sites have been discovered and are supposed to be involved in mRNA degradation as a result of their interaction with specific binding factors (Taylor & Green 1995). It may be useful to screen for these sites and remove them by modifying the gene of interest to increase gene expression. Specific protein initiation sequences such as the Kozak sequence have been found to help in efficient translation in animals; however, this may be different in plants (Lutcke et al. 1987; Cavener & Ray 1991). The level of amino acid can also be a limiting factor in the expression of particular proteins, it may, therefore, be necessary to alter amino acid synthesis pathways for the expression of some particular proteins (Matthews & Hughes 1993, Singh & Matthews 1994).

Expression of multiple genes

Traditionally, plants have been transformed with single genes to improve plant characteristics, to study plant gene expression and to express foreign proteins for industrial or pharmaceutical purposes. However, there are situations when the transformation of plants with multiple genes is desirable. These include improving plants for multigenic traits, improving or altering metabolic pathways, expressing multimeric foreign proteins and expressing multiple enzymes involved in the synthesis of various compounds. Multiple gene engineering can be achieved in different ways.

1) IRESs or internal ribosome entry sites have been used primarily for the expression of two genes in the form of a bicistronic message (Hellen & Sarnow 2001; Allera-Moreau et al. 2006; Sasaki et al. 2008). IRESs are nucleotide sequences that recruit eukaryotic ribosomes to mRNA to initiate protein translation in the middle of the mRNA molecule without the requirement for a 5’ cap that is normally needed for translation initiation (Pelletier & Sonenberg 1988). The main drawback of using IRES for multiple protein expression is that the expression of the IRES regulated gene is lower than the cap-dependent gene upstream of IRES sequence (Kaufman et al. 1991; Zhou et al. 1998; Houdebine & Attal 1999).

2) Independent transgenic lines expressing one gene can be developed and then all the transgenic lines can be crossed together to combine all the genes respon-
sible for the trait under study in a single plant (Ma et al. 1995). However, this strategy is time consuming and laborious and cannot be used for many genes as the occurrence of transgenes on different loci makes the process of obtaining and maintaining homozygous plants complicated. Furthermore, the strategy cannot be used for vegetatively propagated plants.

(3) A plant can be sequentially transformed with the transgenes of interest one by one or in units consisting of more than one gene. However, this is time consuming again and a different selection marker is required for each transformation event, and the number of available markers is limited.

(4) The multiple genes can be expressed in the form of multiple expression cassettes linked together, each expression cassette with its own promoter and terminator (Slater et al. 1999; Goderis et al. 2002). However, multiple copies of the same promoter used for the transgenes of interest can lead to gene silencing (Van Den Elzen et al. 1993; Matzke & Matzke 1998) and there is only a limited choice of different promoters available to overcome this problem.

(5) Co-transformation is another method for multiple gene expression that involves the simultaneous transformation of plant with the transgenes of interest through biolistics or through agrobacterium-mediated transformation (Zhang & Faquett 1998; Li et al. 2003). Again, the problem with co-transformation is that in many cases the plant cannot be efficiently transformed with all the genes of interest and variable integrations of the genes involved can occur, which makes the subsequent characterization difficult.

(6) Multiple genes can also be expressed in the chloroplast genome (Daniell & Dingra 2002); however, this method is not suitable for proteins that need to be targeted to cell compartments other than chloroplast.

Purification and downstream processing of recombinant proteins

Downstream processing refers to the recovery and purification of the recombinant protein from plants. Recovery usually involves processing/fractionating of the plant tissue, protein extraction, solid-liquid separation, and concentration, whereas purification consists of immunoprecipitation, liquid-liquid extraction, membrane filtration, chromatography, etc. Processing of leaves requires special attention. Leaves must be processed immediately after harvest or must be frozen to prevent degradation of proteins by proteases, while seeds can be stored for longer periods as there are fewer chances of degradation of recombinant proteins expressed in the seed. The use of cell secretion systems could also be beneficial as there is no need to disrupt plant cells during recovery, so the release of phenolic compounds could be avoided, however, in the culture medium the recombinant protein may not be stable (Fischer et al. 2004). Another way of facilitating protein recovery is the use of affinity tags. Protein tags should be removed after purification to restore the structure of the purified protein to its native state (Fischer et al. 2004). Oleosin fusion technology, developed by SemBioSys Genetics Inc. (http://www.SEMBIOSYS.com/), is another system in which the recombinant protein gene sequence is fused to the sequence of an oil body specific endogenous protein oleosin in rapeseed and safflower, after purification the protein is separated by an endoprotease digestion (Schillerberg et al. 2005). Problems encountered during protein extraction primarily include proteolytic degradation and structural modification due to the reaction with phenolic compounds. When devising a strategy for heterologous protein production in plants, proper consideration should be given to downstream processing feasibility of the recombinant protein to get optimum protein yield.

Economics of plant made pharmaceuticals

The economic prospects for plant made proteins are quite high in comparison with conventional systems. There is a huge demand for many pharmaceutical proteins. Thanks to the scalability of transgenic plants this demand can be fulfilled at fractional costs of the traditional systems. It is possible to obtain yields of kilogram quantities of recombinant protein from just one hectare of transgenic tobacco (Fischer & Emans 2000). Even if the expression level of protein in maize is 1% of dry weight and its recovery is only 50%, the cost of its production may still be only 2–10% of microbial systems and may be even lower than that for mammalian systems (Twyman et al. 2003; Chen et al. 2005). It has been estimated that at 20% total soluble seed protein expression level, one bushel (25 kg) of maize can produce the same amount of avidin as one tonne of chicken eggs and the cost is only 0.5% of that of chicken eggs (Hood 2004). Several companies at the moment are involved in the commercial production of plant made recombinant proteins. For example, Maltagen Forschung GmbH has been able to produce 3 g human albumin, 2 g lactoferrin and 1.5 g lysozyme per kg of barley...
seeds, respectively (http://www.maltagen.de/PDF/Products.pdf). Several plant molecular farming products are currently at the advanced stages of development and some have been produced commercially. Plants also have the advantage that several types of proteins like the recombinant subunit vaccines can be administered in the form of raw or partially processed fruits and vegetables (MASON et al. 2002). However, it has been quite hard for the industry to grow because of very strict regulations. The consumers’ concerns about the safety of plant made pharmaceuticals further add to the difficulties involved in commercializing molecular farming products. Once the regulatory hurdles, issues over public confidence can be overcome, and purification and downstream processing is simplified, then plant produced molecular farming products have the potential to bring about the cost effective and large-scale production of some important products, many with high medical value.

Bio-safety and monitoring

Public concern about the introduction of genetically modified crops represents one of the most challenging issues. Lack of communication among the authorities dealing with research, bio-safety and trade is an important issue that has hindered developments in molecular farming (RAMESSAR et al. 2008). The long-term impact of molecular farming products on the environment is difficult to assess. An important concern is the contamination of the food chain with plant made pharmaceuticals. This could happen as a result of transfer of genetic material from transgenic plants to food crops, using the same equipment for harvesting and processing of transgenic and food crops without proper decontamination, and growing food crops in the same field where a transgenic crop was grown previously and no decontamination was carried out (RIGANO & WALMSLEY 2005). To avoid the food chain contamination, strict regulation needs to be put in place such as geographically isolating the transgenic crop, growing in greenhouses instead of open fields, and harvesting and processing transgenic plants using separate equipment or properly decontaminating the equipment if the same equipment is also applied to food crops (RIGANO & WALMSLEY 2005). Containment can also be achieved by using male-sterile traits and using chloroplast transformation of plants. As the chloroplast genome is maternally inherited, the chances of transgene spread through pollen are controlled. It is also important to label genetically modified products so that the consumer has the choice to select according to his/her own preferences. Regulatory agencies are facing a number of challenges regarding the regulation of transgenic crops. Each molecular farming product and each host system is unique so each case needs to be handled separately. The 0.5% presence of transgenic material in non-transgenic food or feed has been allowed by the European Parliament and the Council of the European Union in cases where the presence of transgenic material is unavoidable and its negative effects are dominated by its benefits (European Parliament 2003). Efforts to confine transgenics and reduce environmental exposure have been made recently (MA et al. 2003; CHEN et al. 2005; SPARROW & TWYMAN 2009), however, these regulations are still in infancy and much more effort is needed to overcome problems regarding regulations of molecular farming products.

CONCLUSION

Till date a number of pharmaceutical proteins have been expressed in plants. Plants offer a cheaper and safer source of biopharmaceuticals. However, there are a number of technical limitations molecular farming is facing. These mostly include the low yields and recovery of the expressed proteins. Currently, much effort is being devoted to overcome these limitations. Furthermore, there are bio-safety and environmental issues concerning molecular farming. Efforts are being devoted to overcome this, however, a lot more has to be done to make molecular farming products a success.

References

of nopaline synthase human growth hormone chimaeric gene in transformed tobacco and sunflower callus tissue. Plant Molecular Biology, 6: 347–357.


Received for publication September 16, 2013
Accepted after corrections December 4, 2013

Corresponding author:
Dr. Kafeel Ahmad, Centre of Biotechnology and Microbiology, University of Peshawar, Peshawar Khyber Pakhtunkhwa, Pakistan; e-mail: kafeelpbg@gmail.com