Plant-based expression systems for protein and antimicrobial peptide production

Azadeh Niknejad
Received: 06.08.2018 / Revised: 11.09.2018 / Accepted: 22.09.2018 / Published: 21.12.2018

Department of Cell and Molecular Sciences, Faculty of Biological Sciences, Kharazmi University, Tehran, Iran
Correspondent author: niknejad@khu.ac.ir

Abstract. Molecular farming technology offers a unique advantage that almost any protein can be produced economically and safely under very controlled conditions. Besides traditional production systems, such as bacteria, yeasts, insects and mammalian cell lines, plants can now be used to produce eukaryotic recombinant proteins, especially therapeutic ones. Their advantages as hosts for protein production include correct post-translational modifications, low-cost maintenance and no risk of contamination by human pathogens. The system is widely applied in agriculture and industry, especially in life science and pharmaceutical industry. The application of transgenic plants in the production of vaccines, antibodies and pharmaceutical proteins has been playing a key role in plant genetic engineering in recent years. The production of recombinant proteins plays a critical role in the production of high amounts of high-quality proteins. In this review, common problems in the production of recombinant proteins and antimicrobial peptides in plant-based expression systems are discussed and strategies for their solution are suggested. Viral vector-mediated transient gene expression in plants enables rapid production of pharmaceutical proteins such as vaccine antigens and antibodies. To conclude, plant-based systems have the potential to bring unique efficacy-enhancing features to increase the utility and effectiveness of vaccines and therapeutics.

Keywords. antibodies, pharmaceutical proteins, therapeutic proteins, transgenic plant

INTRODUCTION

Among various expression systems such as bacteria, yeast, mammalian cells and insects, plants are becoming an attractive system for the production of recombinant proteins for pharmaceuticals, industrial proteins, monoclonal
antibodies and vaccine antigens, due to a large number of advantages including safe usage, rapid scale-up, long-term storage and low-production cost (Daniell et al., 2001, 2009b; Fischer and Emans, 2000; Fischer et al., 2003, 2004, 2013; Horn et al., 2004; Howard and Hood, 2005; Klimyuk et al., 2008; Ma et al., 2003, 2005; Rybicki, 2009; Schmidt, 2004 Thomas et al., 2011). In addition, plant cells have the ability to perform post-translational modifications (PTM) of proteins, which are typical of eukaryotic organisms (Gomord and Faye, 2004; Mett et al., 2008; Vitale and Pedrazzini, 2005). PTM affects the proteins activities and functions such as folding, stability, solubility as well as dynamic interactions with other molecules. The major type of PTM for plant-made recombinant proteins is glycosylation (Stulemeijer and Joosten, 2008; Webster and Thomas, 2012).

PTM is an action following translation for additional covalent modifications with changing tertiary and quaternary structures of proteins. A wide variety of promoters have been used for protein expression in plants, including constitutive, inducible and tissue-specific, native and synthetic promoters (Egelkrout et al., 2012). The most widely-used promoter is the cauliflower mosaic virus 35S promoter (CaMV35S), which promotes high-level of gene expression through the transcriptional terminator of the Agrobacterium nopaline synthase gene (nos) (Barampuram and Zhang, 2011; Lee et al., 2008). Several molecular events, such as gene transcription, mRNA translation and protein accumulation, affect overall success of protein expression (Table 1).

The recombinant proteins accumulation in plant system depends on several factors including a choice of suitable host tissue and subcellular location, understanding the nature of the foreign proteins and determine their possible effect on the host plant as well as the post-translational modifications and the degree of protein purification (Egelkrout et al., 2012; Fischer et al., 2004; Schiermeyer et al., 2004).

Recently the phenomenon of recombinant protein expression and its corresponding impact have sparked a heated debate. Although contested by many, the matter of plant system is highly beneficial. Therefore, there are several options for protein expression in plant systems (Table 2) that will be discussed in this review.

Table 1. Summary of factors affecting gene transcription, mRNA translation and protein accumulation Source: Modified from Egelkrout et al. (2012)

<table>
<thead>
<tr>
<th>Factors affecting transcription rate</th>
<th>Factors affecting translation rate</th>
<th>Factors affecting protein accumulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promoter selection</td>
<td>Post-transcriptional gene silencing</td>
<td>Protein stabilisation</td>
</tr>
<tr>
<td>Transcription factor binding (TFB) sites</td>
<td>Codon usage</td>
<td>Protein localization</td>
</tr>
<tr>
<td>Upstream regulatory sequences (URS)</td>
<td>mRNA stability</td>
<td>Protein modification</td>
</tr>
<tr>
<td>Intron-mediated sequences</td>
<td>Recruitment and retention of translation-related protein</td>
<td>Cellular milieu</td>
</tr>
<tr>
<td>Transcriptional gene silencing (TGS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-mRNA stability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transgene trait stacking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The impact of the site of integration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The number of gene copies</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. List of required options for plant expression systems (Egelkrout et al., 2012)

<table>
<thead>
<tr>
<th>options for plant expression system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choice of host plant</td>
</tr>
<tr>
<td>dicot or monocot</td>
</tr>
<tr>
<td>food or non-food</td>
</tr>
<tr>
<td>Type of transformation method</td>
</tr>
<tr>
<td>biological (viral, bacterial)</td>
</tr>
<tr>
<td>physical (biolistic, electroporation)</td>
</tr>
<tr>
<td>Expression parameters</td>
</tr>
<tr>
<td>stable or transient</td>
</tr>
<tr>
<td>constitutive or tissue-specific</td>
</tr>
<tr>
<td>The intracellular location</td>
</tr>
<tr>
<td>cytoplasm or apoplast</td>
</tr>
<tr>
<td>Organelle</td>
</tr>
<tr>
<td>Plastid</td>
</tr>
</tbody>
</table>

Host plants

Protein expression platforms have been developed for a variety of host systems including seed crops (e.g., corn, canola, soybeans and rice) leafy crops (e.g., tobacco, alfalfa and lettuce), plant cell cultures (e.g., tobacco, carrot and rice), hairy root cultures and aquatic plants (e.g., *Lemma minor*), in addition to moss and green algae (Wilken and Nikolov, 2012). One of the major plant systems used for the production of recombinant proteins is tobacco. Because of high leaf biomass yields, it can be cropped several times a year (Basaran and Rodriguez-Cerezo, 2008; Twyman et al., 2003; Twyman, 2004). Ma et al. (1995, 1998) reported the first successful expression of the secretory monoclonal antibody in transgenic *N. tabacum*. The two tobacco cell lines, BY-2 (bright yellow 2) and NT-1 (*N. tabacum* 1), are readily available for Agrobacterium-mediated genetic transformation and have been optimised for rapid growth (Doran, 2013).

**Nicotiana tabacum** L. (Tobacco)

The genus *Nicotiana* belongs to the family Solanaceae, which includes 76 species, originating in North America, South America, Australia and Africa (Chase et al., 2003; Knapp et al., 2004). *Nicotiana tabacum* is an amphidiploid/allotetraploid species with 48 chromosomes, generated from hybridisation between *N. sylvestris* and either *N. tomentosiformis Goodspeed* or *N. otophora* Grisebach, all of which originate in South America (Kitamura et al., 2001; Lewis, 2011). *Nicotiana tabacum*, as a model plant in research, provides many practical advantages for large-scale production of recombinant proteins, including high biomass yield, low maintenance and high speed production (Twyman, 2004). It is the non-feed crop which carries a reduced risk of transgenic material contaminating food chains and contains a high amount of nicotine and other alkaloids that are necessary to be removed (Stoger et al., 2005; Twyman et al., 2003; Twyman, 2004).

**Nicotiana benthamiana** Domin

The non-cultivated tobacco species, *Nicotiana benthamiana*, is a unique Australian native plant that may have resulted from hybridisation between *N. suaveolens* and *N. debneyi* (Goodspeed, 1954; Goodin et al., 2008). It is an amphidiploid species with 38 chromosomes (Goodspeed, 1954) that used as a host in the production of monoclonal antibodies and vaccines. It is a suitable host species to produce recombinant proteins using viral vectors such as tobacco mosaic virus (TMV) (Twyman, 2004; Yang et al., 2004). It can be transformed with high efficiency and maintained easily, due to its short stature, short regeneration time and high seed production. It is also useful for the transient expression of genes (Goodin et al., 2008).
**Transformation methods and expression systems**

Genetic transformation can be achieved by physical or biological methods. Physical methods are usually used to transfer non-viral vectors by using a gene gun (biolistics) and chemical polyethylene glycol (PEG) transformation or electroporation of protoplasts (Barampuram and Zhang, 2011; Rao et al., 2009). The biological method of plant cell genetic transformation by *Agrobacterium* involves the transfer and integration of a large tumor-inducing (Ti) or rhizogenic (Ri) plasmid resident in *Agrobacterium* into the plant nuclear genome. Although several species of *Agrobacterium* are known, two species are predominantly used for transformation: *Agrobacterium tumefaciens* (T-DNA transfer for generation of transgenic plants) and *Agrobacterium rhizogenes* (Ri-plasmid transfers necessary DNA to generate transgenic hairy roots) (Gelvin, 2003; Gleba et al., 2014; Rao et al., 2009).

The production of recombinant proteins in plants, also known as molecular farming, is classified into two expression systems, i.e., stable and transient expressions. Stable expression involves integration of foreign DNA into the plant genome and transfer to the next generations (Sharma et al., 2005). The storage organs such as seeds have been used for the stable expression of proteins of interest (Scheller and Conrad, 2004). The chloroplast and nuclear stable transformations require more than two month, on the one hand, and several months, on the other. (Patiño-Rodríguez et al., 2013).

Transient expression systems include *Agrobacterium*-mediated and viral-based transformation, wherein the foreign gene is not integrated in the genome and being expressed only for a few days after being introduced into the plant cell/tissue.

The strategy relies on *A. tumefaciens*, which contains transgenes; the bacteria are infiltrated directly into the organs of interest by injection or by using a vacuum chamber (agroinfiltration) (Klimyuk et al., 2014; Leckie and Stewart, 2011; Leuzinger et al., 2013; Vaghchhipawala et al., 2011). Transient expression is generally used to verify the activity of the expression constructs and generates small amounts of recombinant protein for functional analyses (Twyman et al., 2003). Many results have been obtained so far with viral backbones such as those of the RNA viruses, Tobacco mosaic virus (TMV), Potato virus X (PVX) and Cowpea mosaic virus (CPMV) (Gleba et al., 2014; Marillonnet et al., 2005). Transient expression, using viral vectors, is one of the best and most rapid methods to obtain a high yield of protein expression (Egelkrout et al., 2012).

**Expression systems**

The production of recombinant proteins in plants, also known as molecular farming, is classified into two expression systems including stable and transient expressions.

**Stable expression**

Stable expression involves the production of transgenic plants or transgenic cell cultures. The foreign DNA is integrated into the plant genome and is transferred to the following generations (Sharma et al., 2005). The storage organs of crop plants such as seeds have been used for the stable expression of the proteins of interest (Fischer et al., 2004).

**Transient expression using viral vectors**

Transient viral-based systems are amongst the best available systems and have the advantage of serving as a rapid method for large-scale production of proteins such as antibodies and vaccine antigens (Gleba et al., 2007; Pogue et al., 2002). The viral vector system is based on two strategies: ‘full virus’ vectors used to express long polypeptides (at least 140 amino acids) and ‘deconstructed’ virus vectors, which rely on *A. tumefaciens* to deliver DNA copies of one or more viral RNA replicons to plant cells (Gleba et al., 2007; Gleba and Giritch, 2012).

One of the main systems of deconstructed virus vectors is the ‘magICON™’ from ICON genetics, based on the tobacco mosaic virus (TMV), which, instead of supplying whole RNA or linear DNA, it is divided in different modules captured between the left and right border of T-plasmid of *A. tumefaciens* (Gleba et al., 2004; Gleba et al., 2014; Marillonnet et al., 2004).

The gene of interest is cloned in a 3’ vector module. Various 5’ modules are then used for targeting the expressed protein to different parts of the plant cell (cytosol, chloroplast or apoplast). An integrase module is used for the recombination of 3’ and 5’ modules in planta. These modules recombine within the plant cell, the resulting DNA is transcribed and the recombination sites (integrate) are spliced out in order to create an RNA replicon to produce the desired protein (Marillonnet et al., 2004). Three *A. tumefaciens* cultures deliver different vector modules containing different viral components, targeting gene sequences. These pro-vector 5’ and 3’ modules assemble within the plant cell in the form of a functional viral vector (Marillonnet et al., 2004).

After being delivered by agroinfiltration, they are assembled inside the cell with the help of a site-specific recombinase to create an RNA replicon to produce the desired protein (Fig. 1).
Fig. 1. Schematic showing the assembly of the three viral vector modules and production of protein. The 5’ and 3’ modules with integrase in the plant cell allow recombining via the AttP and AttB sites. The intron is then spliced out. Viral replication occurs and translation and interest protein processing occurs. P: ACT2 promoter, T: nos terminator, G: Gene of interest. Source: Modified from Marillonnet et al. (2004).

The advantage of this system is the high-yield production of the desired protein because of viral amplification (Marillonnet et al., 2004). Targeting the proteins into different subcellular compartments can alter protein stability (Gils et al., 2005) e.g., the improvement of stability by chloroplasts (Doran, 2006). The green florescence protein (GFP) was expressed in different subcellular compartments of N. benthamiana using this system and produced the recombinant protein at up to 80% of total soluble protein (Marillonnet et al., 2004, 2005). A considerable amount of research has been published on the use of the magnICON® system (Gleba et al., 2004, 2005, 2007; Gils et al., 2005; Giritch et al., 2006; Huang et al., 2010; Marillonnet et al., 2004, 2005; Niknejad et al., 2016; Santì et al., 2006, 2008; Werner et al., 2011).

The targeting of foreign proteins to the most appropriate subcellular compartment is important for protein accumulation, stability, folding and post-translational modification (Twyman, 2004). The choice depends on the structural characteristics of the protein and the use of appropriate signals (Benchabane et al., 2008; Faye et al., 2005). Several subcellular compartments including the cytosol, endoplasmic reticulum (ER), apoplast, vacuole and chloroplast have been considered (Benchabane et al., 2008; Ma et al., 2003).

Cytosol

Recombinant proteins are retained in the cytosol when no targeting signals are used. Several studies have revealed that recombinant proteins remain stable within the cytosol (De Jaeger et al., 1999; Michaud et al., 1998; Marusic et al., 2007). However, the cytosol may not be the best location

Subcellular location in a plant system for protein accumulation
for several reasons: high levels of protease and ubiquitin activity, the negative redox potential that may affect the correct folding of proteins via formation of disulphide bonds, and a lack of important co- and post-translational modifications, such as glycosylation, which may affect folding, assembly and the tertiary structural stability (Benchabane et al., 2008; Faye et al., 2005).

**Endoplasmic Reticulum (ER)**

Targeting proteins to the ER can lead to high yields of recombinant proteins in planta (Conrad and Fiedler, 1998; Wandelt et al., 1992). Entry of recombinant proteins into the endoplasmic reticulum (ER) can be retained by adding (SEK/H/K)DEL retention sequences to the C-terminal (Benchabane et al., 2008; Michaud et al., 1998; Mainieri et al., 2004), in most cases a (K/H) DEL sequence (Benchabane et al., 2008; Ma et al., 2003). The presence of chaperones in the ER helps to di-sulphide bonds formation and correct folding of the recombinant protein (Benchabane et al., 2008; Faye et al., 2005).

**Aoplast**

Proteins can be targeted into the apoplast via the secretory pathway (Benchabane et al., 2008). The plant apoplast includes the cell wall matrix and intercellular spaces in a wide range of physiological functions (Witzel et al., 2011). The lack of KDEL in endosperm-expressed protein results in targeting to the apoplast, where the recombinant protein mostly accumulates in the space under the cell wall (Christou et al., 2004).

**Vacuole**

The vacuole is one of the alternative destinations for recombinant proteins produced in plants (Benchabane et al., 2008). Additionally, it serves several key roles in planta e.g., control of cell turgidity, levels of macromolecules, accumulation of toxic secondary metabolites and storage of high-energy compounds in vegetative tissues (Benchabane et al., 2008; Twyman, 2004).

**Chloroplast**

The chloroplast is considered to be one of the cellular destinations in a plant expression system through the addition of an appropriate targeting signal peptide (Hyunjong et al., 2006). The recombinant proteins may be moved to the chloroplast. Chloroplast transformation offers uniform expression rates, a high copy number of the transformed genes (due to hundreds of plastids in a cell), co-expression of multiple genes from the same construct, minimal transgene escape in the environment due to the maternal inheritance of chloroplast DNA and minimal gene silencing (Daniell et al., 2004).

Post-translational modifications, such as multimerisation and disulphide bridge formation, can occur in the stroma (Daniell, 2006; Parachin et al., 2012). Thus, the chloroplast can be considered to be a suitable compartment with no need for the cell secretory pathway or modifications such as glycosylation, although, some endogenous proteases can affect the stability and accumulation of recombinant proteins (Adam and Clarke, 2002). However, this may not be relevant issue for proteins expressed at very high yield (Daniell, 2006).

**Plant systems for heterologous production of antimicrobial peptide**

Anti-microbial peptides or proteins (AMPs) identified from prokaryotes and eukaryotes are usually less than 100 amino acid residues long and often have a high frequency of positively charged amino acids (Silva et al., 2011). They are strongly cationic (pI: 8.9-10.7) having a net charge of 2+ to 7+, mainly due to the presence of arginine, lysine and histidine residues, heat-stable (100°C, 15 minutes), and additionally, they may have 50% hydrophobic amino acids and do not have any lysing/ killing effects on eukaryotic cells (Hancock and Chapple 1999; Li et al., 2012). AMPs have shown to serve as an important mechanism of natural resistance and also as a metabolically rapid inexpensive first line defence against pathogens (Broekaert et al., 1997; Egorov et al., 2005).

The antimicrobial activity of such peptides was observed to involve ionic interactions between their basic residues and the negatively charged bacterial membrane constituents, such as phospholipid head groups (Jenssen et al., 2006; Melo et al., 2009; Shai, 2002). On the other hand, the outermost layer of the eukaryotic membranes is composed of lipids with no net charge and zwitterionic head groups positioned within the inner layers; membrane-binding is controlled by hydrophobic rather than ionic interactions (Lee et al., 2011a).

During the past 15 years, a wide range of AMPs have been shown to play essential roles in plant defense systems. Their sizes range from 2 to 9 kDa and many of them have a cysteine-rich structure with 6 or 8 cysteine residues involved in disulphide bridges; e.g., thionins, defensins, non-specific lipid transfer proteins (nsLTPs), hevein - and knottin-like peptides, MBP1, IbAMP and snakins (Broekaert et
The hevein-like peptide with a 10 cysteine motif isolated from Triticum kiharae has shown antifungal and antibacterial activity (Dubovskii et al., 2011; Odintsova et al., 2009).

Table 3. AMPs heterologous produced in Nicotiana tabacum

<table>
<thead>
<tr>
<th>AMP transformed</th>
<th>Source</th>
<th>Objective</th>
<th>Resistance against</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>LFchimera</td>
<td>chimerical peptide</td>
<td>Production peptide and antimicrobial activity</td>
<td>Escherichia coli, Staphylococcus aureus, Erwinia amylovora, Ralstonia solanacearum</td>
<td>Chahardoli et al. (2018)</td>
</tr>
<tr>
<td>protegrin-1 (PG-1)</td>
<td>Synthesized peptide</td>
<td>Production peptide and antimicrobial activity</td>
<td>Klebsiella pneumoniae, Staphylococcus aureus, Escherichia coli, Mycobacterium bovis, Candida albicans</td>
<td>Patiño-Rodríguez et al. (2013)</td>
</tr>
<tr>
<td>Dermaseptin Lysozyme C4V3</td>
<td>Phyllomedusa sauvagii, Gallus gallus, Synthetic</td>
<td>Production and isolation of the peptide and vaccine</td>
<td>Anti-HIV activity, induction of mammalian immune response</td>
<td>Rubio-Infante et al. (2012)</td>
</tr>
<tr>
<td>EtMIC2</td>
<td>Eimeria tenella</td>
<td>Peptide production</td>
<td>Eimeria tenella (chicken coccidiosis)</td>
<td>Sathish et al. (2011)</td>
</tr>
<tr>
<td>Retrocyclin-101 and Protegrin-1</td>
<td>Artificial AMP based on rhesus monkey (Macaca mulatta) circular minidefensins</td>
<td>Chloroplast genome transformation aiming peptide production</td>
<td>Erwinia carotovora, tobacco mosaic virus</td>
<td>Lee et al. (2011b)</td>
</tr>
<tr>
<td>NmDEF02</td>
<td>Nicotiana megalosiphon</td>
<td>Proof of concept for crop improvement</td>
<td>Phytophthora parasitica var. nicotianae, Peronospora hyoscyami f.sp. tabacina, Alternaria solani, Phytophthora infestans</td>
<td>Portieles et al. (2010)</td>
</tr>
<tr>
<td>Trichokonins</td>
<td>Trichoderma pseudokoningii</td>
<td>Proof of concept for crop improvement</td>
<td>Tobacco mosaic virus</td>
<td>Luo et al. (2010)</td>
</tr>
<tr>
<td>Pal and Cpl-1</td>
<td>Phages infecting S.pneumoniae</td>
<td>Chloroplast genome transformation aiming peptide production</td>
<td>Streptococcus pneumoniae</td>
<td>Oey et al. (2009)</td>
</tr>
</tbody>
</table>
Antimicrobial peptides/proteins (AMPs) are unique biologically active molecules that comprise defense systems against numerous pathogens such as bacteria, fungi, parasites and/or viruses. Heterologous expression systems, including bacteria and fungi as host cells, have been used for the production of AMPs with different sizes, folds and complexities (Parachin et al., 2012). Plant systems have also been used as suitable sources for the production of AMPs and can be directly used for crop improvement without purifying the peptide or proteins (Desai et al., 2010; Giddings et al., 2000). The heterologous production of radish Rs-AFP2 defensin in tobacco protected the transgenic tobacco from Alternaria longipes (Terras et al., 1995), whereas the alfalfa anti-fungal peptide (aLFAPF) defensin reduced the infection in transgenic potato plants against the fungal pathogen Verticillium dahliae (Gao et al., 2000).

Due to their antimicrobial activity and complex structures, several AMPs cannot be produced in microbial systems. For example, retrocyclin-101 (RC101) and Protegrin-1 (PG1) can be used as therapeutic agents against bacterial and/or viral infections; none has yet been produced in microbial systems, but both have been expressed in the tobacco chloroplast with a yield of around 20% of total soluble protein and retention of antiviral activity (Lee et al., 2011b). The large-scale production of AMPs can be applied in numerous industries such as biotechnology, pharmaceuticals, cosmetics and food industry. However, the limiting factor is the high production costs of some AMPs due to their molecular mass and protein folding characteristics.

The AMP can be degraded by endopeptidases, which also limits their bioactivity. Thus, selecting the right system is essential for large-scale production and functional recombinant AMPs. Plant-based expression systems are promising platforms for further engineering and production of AMPs (Parachin et al., 2012). All AMPs heterologously produced in N. tabacum are listed in Table 3.

CONCLUSION

In conclusion, plants can act as ‘biofactories’ for the production of recombinant proteins, to reduce the production costs due to the low maintenance cost and high quality of the therapeutic proteins (Davies, 2010; Gomord and Faye, 2004; Gleba, 2007; Ling et al., 2010; Mett et al., 2008; Melnik and Stoger, 2013; Niknejad et al., 2016). Moreover, several remarkable advances make plant expression system an attractive high-tech platform for plant-made protein pharmaceuticals (PMPs).

ACKNOWLEDGMENTS

The author acknowledges Prof. Mrinal Bhave and Dr. Daine Webster, Swinburne University of Technology, Melbourne, Australia.

REFERENCES


Daniell, H., Cohill, P.R., Kumar, S. and Dufourmantel, N. 2004. Chloroplast genetic


endomembrane system as bioreactor. – Mol. Interv. 5: 216-225.


How to cite this article: