Proteolysis of recombinant proteins in bioengineered plant cells

Priyen Pillay1, Urite Schlüeter1, Stefan van Wyk1, Karl Josef Kunert2,* and Barend Juan Vorster2

1Department of Plant Science; Forestry and Agricultural Biotechnology Institute; University of Pretoria; Pretoria, South Africa; 2Department of Plant Production and Soil Science; Forestry and Agricultural Biotechnology Institute; University of Pretoria; Pretoria, South Africa

Keywords: recombinant protein production in plants, protein stability, proteases, protease inhibitors

Plants are increasingly used as alternative expression hosts for the production of recombinant proteins offering many advantages including higher biomass and the ability to perform post-translational modifications on complex proteins. Key challenges for optimized accumulation of recombinant proteins in a plant system still remain, including endogenous plant proteolytic activity, which may severely compromise recombinant protein stability. Several strategies have recently been applied to improve protein stability by limiting protease action such as recombinant protein production in various sub-cellular compartments or application of protease inhibitors to limit protease action. A short update on the current strategies applied is provided here, with particular focus on sub-cellular sites previously selected for recombinant protein production and the co-expression of protease inhibitors to limit protease activity.

Introduction

The world-wide demand for recombinant therapeutic and diagnostic proteins requires exploring plant-based protein expression platforms supplementing existing prokaryotic production systems. A number of valuable human recombinant proteins have already been successfully produced in plant-based systems, ranging from soil-grown plants to plant cells grown in a bioreactor. Cells may be used to transiently express the protein over a relatively short time period or be genetically engineered to stably express any recombinant protein. Plants offer the general advantage of a high plant biomass, the ability to perform post-translational modifications and protein yield, depending on the resident proteases. Proteolysis may occur in planta or during protein extraction and harvesting, often requiring protease inhibitors to be added to the extraction buffer to improve protein stability and yield. However, this strategy is expensive and is seldom economically viable with regards to large scale extractions.

The purpose of this short review is to give an overview on the current knowledge of protease action on recombinant proteins produced in plants and to provide an update of some current strategies applied to improve recombinant protein stability in plant-based production systems.

Proteases Act on Recombinant Proteins

Protease abundance in plant tissues represents a severe burden to effective recombinant protein production. The degree of proteolysis, either partial or complete, depends on the amino acid sequence of the recombinant protein, susceptibility of sites to proteolytic action and also the number of protease-susceptible sites. Studies on plant proteases have advanced substantially and a more detailed understanding of the role of proteases, particularly in growth, development and pest resistance, is emerging. Hundreds of plant genes encode for proteins involved in proteolysis. In the model plant Arabidopsis, about 1900 genes involved in peptide bond hydrolysis have already been identified, but only a small number of proteases has so far been characterized, with the biological function of only around 40 proteases elucidated. Plants with larger genomes are likely to also have a higher number of proteases with highly polymorphic activity profiles in different plant species. Protease functions include assembling and disassembling proteins as well as removing damaged, mis-folded or potentially harmful proteins. Based on their active site residues for catalysis, most proteases can be distinguished as serine, cysteine, aspartic, and metallo-types with serine proteases consisting of about 200 members, and the cysteine, aspartic, and metallo-type proteases about 100 members in each class (http://merops.sanger.ac.uk).
In *Nicotiana* species, often used for recombinant protein production, the majority of proteases are of aspartic or cysteine type (papain-like cysteine proteases) and to a lesser extent serine and metallo-type.16-18 When recombinantly expressing proteins in *Nicotiana*, the leaves of *N. benthamiana* is considered to contain lower protease activity compared with leaves of *N. tabacum*, consisting mostly of cathepsin L- and legumain-like cysteine proteases.19 Table 1 outlines the type of proteases and their localizations so far identified in the different plant species previously used for recombinant protein production.

### Table 1. Cellular locations for recombinant protein production in various plant host and types of proteases identified at these locations.

| Host | Protein | Compartment | Protease | References |
|------|---------|-------------|----------|------------|
| *Solanum tuberosum cv Desireé* | Sea anemone equistatin | Secretory pathway, lytic vacuole, ER | Arginine/lysine-specific, legumain-type Asn-specific cysteine | 1 |
| *Nicotiana tabacum cv Samsun NN* | Monoclonal mouse IgG1 | Apoplast | Cysteine, aspartic | 2 |
| *Nicotiana tabacum L. cv Samsun* | Glutathione reductase | Cytosol | Cysteine | 3 |
| *Solanum tuberosum* plants, cv Kennes | human W-1-antichymotrypsin | Cytosol | Aspartic, serine | 4 |
| *Oryza sativa L. cv Dongjin* | Human granulocyte-macrophage colony stimulating factor | Secretory pathway | Cysteine | 5 |
| *Oryza sativa L. cv Dongjin* | Human granulocyte-macrophage colony stimulating factor | Extracellular | Serine | 6 |
| *Solanum tuberosum L. cv Kennes* | Bovine aprotinin | Cytosol, ER, apoplast | Serine | 8 |
| *Nicotiana tabacum cv BY-2* | Human α-1-antichymotrypsin | ER, Golgi, apoplast, extracellular | Serine | 9 |
| *Nicotiana tabacum L.* | Oryzacystatin-1 | ER, chloroplast | Cysteine | 10 |
| *Nicotiana tabacum (var. santon)* | Human IgG, antibody | Apoplast | Cysteine, aspartic, serine | 11 and 12 |
| *Solanum tuberosum L. cv Kennes* | Tomato cathepsin D inhibitor (CDI), Bovine aprotinin | Cytosol, ER | Serine | 13 |
| *Nicotiana tabacum cv BYPV* | Spider dragline silk | ER | Serine | 14 |
| *Solanum tuberosum L. cv Kennes* | Tomato cathepsin D inhibitor (S252) | Cytosol | Aspartic, serine | 15 |
| *Nicotiana tabacum* | Monoclonal antibodies | Secretory pathway | Serine | 16 |
| *Solanum lycopersicum var PED* | Human α-1-protease inhibitor | ER, apoplast, vacuole, cytosol | Serine | 17 |

### Selecting the Cellular Compartment for Recombinant Production

A cellular localization with limited proteolytic activity may be interesting for recombinant protein stability and ultimately yield. Protease activity is pH dependent and proteases therefore reside in different cellular compartments favorable for their respective activities. These enzymes are found in various cellular compartments including the cytosol, the vacuole, the chloroplast, the mitochondria, and the lysosome.10-12 Figure 1 provides an overview of the different classes of proteases active in the different compartments of a plant cell. A number of proteases are extracellular, residing in the apoplast to which recombinant proteins can be secreted. For secretion, proteins travel from the endoplasmic reticulum (ER) through the Golgi apparatus to the cell surface.

The **cytosol and the vacuoles**

Undesired protein modifications changing protein structure may occur in the cytosol.23 The cytosolic ubiquitin–proteasome proteolytic pathway further degrades any improperly folded protein.24 Recombinant proteins are generally poorly accumulated when expressed in the cytosol,25 and this compartment is often regarded as unsuitable for effective recombinant protein production.

Lytic vacuoles are also unsuitable for recombinant protein deposition due to their high protease content. Protein storage vacuoles, abundant in seeds, are more suitable for protein accumulation. Targeting proteins to the storage vacuoles is achieved by a specific amino acid sequence, or sorting signal, within the primary sequence of the protein.26 An example of vacuolar accumulation of recombinant proteins achieved through vacuolar targeting can be found in the case of dog gastric lipase produced in transgenic tobacco plants.27

The **ER and the Golgi**

The ER has been the production site for several recombinant proteins of industrial and pharmaceutical value. Directing a
proteases act in the secretory pathway with pepsin-like (A1), papain-like (C1), trypsin chymotrypsin-like (S1), subtilisin-like (S8), and serine carboxypeptidase-like (S10), the most represented protease families. Unintended processing of recombinant proteins along this pathway by resident proteases has been reported by several research groups. Some examples include the systematic processing of mammalian antibodies and the partial trimming of the anti-inflammatory bovine aprotinin protein at the C- and N-termini when retained in the ER. Cleavage of the C-terminal region of the human α1-anti-chymotrypsin by intracellular and apoplastic proteases when targeted to the secretory pathway of BY-2 tobacco cells and subsequently detected in the culture medium is another example of unintended processing that may occur.

The apoplast and the chloroplast

A number of recombinant proteins have been successfully expressed in the apoplast with expression of the human interleukin 6 in N. benthamiana recently reported as an excellent example. However, abundance and post specificity of proteolytic enzymes in the apoplast is still a major obstacle. Intact bovine aprotinin was for instance detected in the apoplast of transgenic potato leaves, but final yields in planta were much lower when compared with retaining the protein in the ER. The value of ER retention has been previously demonstrated where a ER retention signal increased human anti-HIV 2G12 levels in N. benthamiana plants.

Higher expression of ER-retained proteins has also been demonstrated for a structural poly-protein, P1–2A, as well as for a 3C protease from FMDV serotype O when stably expressed in foliar tomato extract. In contrast, when a signal peptide such as the CTB signal peptide is absent, expression of the viral protein is not detectable possibly due to degradation within the cytoplasm during, or immediately after, synthesis. Proteins that are retained in the ER may also reside in protein bodies enhancing post-translational stability. Greater yield in the ER is very likely due to the action of chaperone proteins supporting proper protein folding. Folding and/or post-translational modification of recombinant proteins can however differ if post-translational processing occurs in the Golgi apparatus downstream of the ER. Recombinant proteins may be directed to this organelle via KDEL or HDEL signal peptides, however this may also result in undesired, structurally distinct, proteins due to non-native amino acid additions or non-authentic protein glycosylation patterns.

The disadvantage of ER retention is the existence of ER proteolytic pathways acting on misfolded proteins. Misfolded proteins are in some cases re-translocated into the cytosol by the ER machinery for proteasomal degradation. Several classes of proteases act in the secretory pathway with pepsin-like (A1), papain-like (C1), trypsin chymotrypsin-like (S1), subtilisin-like (S8), and serine carboxypeptidase-like (S10), the most represented protease families. Unintended processing of recombinant proteins along this pathway by resident proteases has been reported by several research groups. Some examples include the systematic processing of mammalian antibodies and the partial trimming of the anti-inflammatory bovine aprotinin protein at the C- and N-termini when retained in the ER. Cleavage of the C-terminal region of the human α1-anti-chymotrypsin by intracellular and apoplastic proteases when targeted to the secretory pathway of BY-2 tobacco cells and subsequently detected in the culture medium is another example of unintended processing that may occur.

The apoplast and the chloroplast

A number of recombinant proteins have been successfully expressed in the apoplast with expression of the human interleukin 6 in N. benthamiana recently reported as an excellent example. However, abundance and post specificity of proteolytic enzymes in the apoplast is still a major obstacle. Intact bovine aprotinin was for instance detected in the apoplast of transgenic potato leaves, but final yields in planta were much lower when compared with retaining the protein in the ER. In the apoplast proteins are generally exposed to a large number of proteases. The apoplast of N. tabacum leaves primarily contains

---

Figure 1: Protease locations within plant cell subcellular compartments.
aspartic-, cysteine- and serine-type proteases while the apoplast in *N. benthamiana* leaves show preferential activity of aspartic- and serine-type proteases.

A fairly new strategy to improve recombinant protein stability and achieve of higher yields is production in the chloroplast compartment of genetically engineered plants. Chloroplast engineering has several advantages including uniform protein expression rates, multiple copies of an integrated transgene and low gene silencing. High plant size number per cell and maternal inheritance of chloroplast DNA leading to minimal transgene escape are also among the advantages. Examples of chloroplast-based production of a recombinant protein include production of a cholera toxin B-pro-insulin fusion in transgenic lettuce and tobacco and production of the VP1 structural protein from the foot-and-mouth disease virus in tobacco chloroplasts. It was also recently reported that a chloroplast-derived vaccine candidate was stable at room temperature for 20 months. Despite this success, the inability to perform more complex post-translational modifications, such as glycosylation, or to perform protein sub-unit assembly and proper protein folding, are disadvantages in a chloroplast-based production system. In addition, endogenous proteases are present in the chloroplast that may compromise recombinant protein accumulation. For instance, high protein accumulation of the rotavirus VP6 protein was found in young tobacco leaves whereas in older leaves the amount of VP6 protein decreased possibly due to proteolytic degradation.

**Preventing Protease Action**

Different strategies have been proposed to minimize unintended proteolysis in plants. Some strategies involve the targeting of recombinant proteins to specific cellular locations using peptide sorting signals or the addition of a stabilizing fusion partner to the protein of interest. Elastin-like peptide (ELP) fusions for instance can be used to avoid proteolytic degradation by facilitating the sorting of recombinant proteins to specific cellular locations using peptide sorting signals. The advantage of ELP-fusion proteins is that they can be directly accessible for purification in a viscous and solvated state and they can increase the amounts produced of enzymatically active recombinant proteins. In addition, the in vivo expression of a tomato cathepsin D inhibitor (SCDI) resulted in an increase in leaf protein content with transient expression of human AACT (α1-anti-chymotrypsin) significantly higher in transgenic lines expressing the SCDI inhibitor. Co-expression of an aspartic/serine “companion” inhibitor also greatly increased leaf apoplast protein content with more murine diagnostic antibody (CS-1) co-secreted in the apoplast.

There is still very limited knowledge on cysteine protease inhibitors as “companions” possibly due to observations that less activity has been found in plant systems for cysteine proteases when compared with serine proteases. In a first attempt to demonstrate potential of such strategy, co-expression of the rice cystatin OC-1 decreased cysteine protease activity resulting in a stabilizing effect on isolated Rubisco. SLCY39, an inhibitor of papain- and legumain-like cysteine proteases, had no impact on apoplast-based production, but stabilized the CS-1 antibody in planta, presumably upstream in the secretory pathway. In our group, we investigated the use of an “in-built” protein stabilizing agent in genetically engineered tobacco plants expressing OC-1 in the cytosol. Constitutively expressing the rice cystatin in tobacco leaves lowered overall cysteine protease activity and increased the amounts produced of enzymatically active recombinant glutathione reductase, which was used as a model enzyme, when the enzyme was transiently produced in transgenic tobacco leaves after agroinfiltration.

**Challenges Ahead**

Proteolysis caused by plant endogenous proteases is still a key challenge severely compromising recombinant protein yield. However, there is a constant search for new production systems not only to ease purification but also to limit proteolysis. Targeting recombinant proteins either to oil bodies or roots for rhizo-secretion are recent attractive new strategies for easier purification as well as to limit proteolysis. The advantage of rhizo-secretion to leaf based-production is that after secretion the hydroponic culture medium has lower and less complex levels of proteolytic enzymes when compared with leaf extracts. Engineered carnivorous plants have also recently been suggested as production platform systems. Carnivorous plants express and transport digestive enzymes into the traps, where any enzymes would be directly accessible for purification in a viscous and sticky liquid without plant destruction allowing continuous
harvest. The limitation of this system is the presence of proteases in the juice which might affect recombinant protein stability. However, all these recent new technologies, although interesting, have so far resulted in insufficient protein yields to be considered commercially viable and the search for new innovative production systems for stable recombinant protein production should therefore be an ongoing activity.

Searching for plant species with low proteolytic activity providing better recombinant protein stability should also be relevant in future activities. In comparison to prokaryotic systems, pro tease-deficient mutant plants do not exist for plant species currently used for recombinant production. More intense screening of plant species useful for recombinant protein production for low protease activity is therefore urgently required. This is in addition to detailed studies on identifying plant endogenous proteases and investigating their expression profiles in various cellular locations.

Better knowledge of protease involvement in senescence processes might be particularly helpful to improve protein stability in any transient expression system, as almost all proteases have been associated with some aspects of plant senescence.55 There is evidence that leaf infiltration with Agrobacterium cells causes leaf senescence resulting in the expression of senescence-related proteases including cysteine proteases.56 Senescence, the final developmental stage of every plant organ, leads to cell death and senescence-associated proteolysis naturally enables the remobilization of nutrients but might also degrade recombinant proteins. Strategies for protein stabilization might also include application for induced inhibition or induced downregulation of proteases. In this regard, antisense or RNA silencing approaches could be of interest to contain proteolysis in the plant host. Indeed, first evidence that such a strategy might be successful has been recently demonstrated with rice cells where application of the RNA interference technology using a gene to express dsRNA of α-amylase and cysteine protease resulted in a 2.4 fold increase in the production of human granulocyte-macrophage colony-stimulating factor (SGM-CSF) after downregulation of cysteine protease expression.57 However, identification of particular proteases involved in recombinant protein degradation is required to avoid any protease involved in vital cellular processes required for growth and development from being targeted.

Future research might also focus on identifying the specific inhibitors of proteases involved in recombinant protein degradation. A better understanding of the exact nature of these inhibitors would allow the design of more active inhibitors. Design by amino acid mutagenesis would optimize their inhibitory activity for application as a “companion” protease inhibitor in either transient or stable expression of a recombinant protein in addition to co-expressing a “companion” protease inhibitor is, however, rather complex due to pleiotropic effects caused by inhibitor expression affecting plant growth and development.60 There is so far only limited knowledge about these protease inhibitor actions. Such a strategy might also be problematic when inhibition involves targeting proteases in the senescence pathway.61 Although such an inhibitor approach might be a major obstacle in a transgenic plant and/or seed approach, this might possibly be less problematic in short-term transient expression of a recombinant protein, which lasts only a few days.

Recent genomic and proteomic approaches have allowed the large-scale identification of proteases and the elucidation of their particular roles in cellular metabolism. This expanding knowledge will certainly help the isolation of inhibitors or new techniques for high-throughput analysis of protease activity and identification of target proteases. This will also advance our knowledge on recombinant protein stability and application of this knowledge in the future will be critical to significantly improving plant-based recombinant protein production.

Disclosure of Potential Conflicts of Interest

No potential conflict of interest was disclosed.

References
1. Yin J, Li G, Ren X, Horler G. Select what you need a comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes. J Biotechnol 2007; 127:375-471; PMID:16959596; http://dx.doi.org/10.1016/j. jbiotec.2006.07.012
2. Ma J, Cizmeci I, Beck R, Chirita P, Dale JD, Dixon PR, et al.; European Union Framework 6 Planta-Planta consortium. Molecular farming for new drugs and vaccines. Current perspectives on the production of pharmaceuticals in transgenic plants. EMBO Rep 2005; 6:319-323; PMID:15995064; http://dx.doi.org/10.1038/sj.embor.7408478
3. Davard H, Singh ND, Mason H, Streetfield SJ. Plant-made vaccines: antigens and biopharmaceuticals. Trends Biotechnol 2008; 26:669-79; PMID:18936295; http://dx.doi.org/10.1016/j.tibtech.2008.09.009
4. De Mercuri B, Narayan C, Beatty M. Production of antibodies in plants: status after twenty years. Plant Biotechnol J 2000; 8:92-103; PMID:12152531; http://dx.doi.org/10.1046/j.1467-7652.2000.00937.x
5. Refsby EP. Pharmacological vaccines for humans and animals. Plant Biotechnol J 2010; 8:020-37; http://dx.doi.org/10.1111/j.1467-7652.2010.00937.x
6. Kang SR, Kallos PT. The production of biopharmaceut-icals in transgenic plants. Biotechnol Adv 2009; 27:679-94; PMID:19447803; http://dx.doi.org/10.1016/j. biot.2009.08.002
7. Simonsen CR, VanderJeavon JS, Vayapalree SR. A model of Agrobacterium tumefaciens vacuolar protein infiltra- tion into harvested leaf tissue and subsequent in planta transient transgene expression. Biotechnol Bioeng 2009; 102:365-70; PMID:18895157; http:// dx.doi.org/10.1002/bit.22118
8. Etherton MA, Lewis PD, Bellflower J, Beckford N, Mardel M, Vitina LP. Transient exposure of antibodies in plants using spring fragmentation. Recombinant Protein From Plants, 2010:61-68
9. Rosado D, Augspurger R, Brownell R, Li-Vq, Vitina LP, Tsurpakan S, et al. In-silico protease inhibi- tor system for the protection of recombinant proteins recovered from transgenic plants. Plant Biotechnol J 2006; 4:359-68; PMID:17147641; http://dx.doi. org/10.1111/j.1467-7652.2006.00187.x
10. Rosado D, Augspurger R, Brownell R, Li-Vq, Vitina LP, Tsurpakan S, et al. An in-silico protease inhibi- tor system for the protection of recombinant proteins recovered from transgenic plants. Plant Biotechnol J 2006; 4:359-68; PMID:17147641; http://dx.doi.org/10.1111/j.1467-7652.2006.00187.x
11. Schaffr A. A cut above the rest: the regulatory function of plant proteins. Planta 2004; 219:1105-7; PMID:15191785; http://dx.doi.org/10.1007/s00425-004-0160-2
12. van der Heijden BAJ. Plant proteases from pheno- types to molecular mechanisms. Annu Rev Plant Biol 2008; 59:201-223; PMID:18277708; http://dx.doi.org/10.1146/annurev.arplant.59.032607.072859
13. Vranjes KD. Proteolysis in plants: mechanisms and functions. Plant Mol Biol 1996; 32:275-302; PMID:8806650; http://dx.doi.org/10.1007/BF00039386
14. Poens GC, Angiao JL, Eliau OD, Joanna KE. Inhibitors of the activity of serine, cysteine, and threo- mine proteases. Chem Biol 2002; 9:609-75; PMID:12472055; http://dx.doi.org/10.1016/S1074-5521(02)00106-3
15. van Slag MS, van den Berg E, Voskuil C. Plant caspases. Curr Opin Plant Biol 2006; 9:282-7; PMID:16422957; http://dx.doi.org/10.1016/j.pbi.2005.11.004
16. Davies GD, Caffrey CR. Recombinant protein production in tomatoes and tobacco (Nicotiana tabacum) with genetically improved resistance to tobacco mosaic virus and tobacco necrosis virus (Solanum species). J Insect Bioeng 2006; 3:41-55; PMID:16959596; http:// dx.doi.org/10.1039/b502622c
Arabidopsis thaliana, Haussuhl K, Manuell A, et al. Chloroplast and mitochondria-translation of plant nomenclature. Plant Physiol 2001; 125:1912-1917.

O. Canepari, G. Forgione, L. Capaccioli, et al. Preparation of mitochondria and recombinant enzyme activity of mitochondrial proteins in Arabidopsis thaliana: A proposed nomenclature. Plant Physiol 2010; 153:1912-1917.

C. Calle, Jr. Regulation of Protein Degradation. Plant Cell 1995; 7:940-957. PMID:12252390.

A. Adami, A. Dadiana, S. Nobakhtabadi, et al. Osmotin and mitochondrial proteins in Arabidopsis. A proposed nomenclature. Plant Physiol 2011; 155:1912-1917.

J. Cao, R. J. Dooley, S. M. McShane, et al. Osmotin and rapid transient protein production in plants. Curr Opin Biotechnol 2005; 16:167-173; PMID:15831382.

R. L. Zhang, H. J. Zhang, D. V. Shadrach, et al. Osmotin occurs in the vacuole. Planta 1995; 196:586-591; PMID:19137543.

C. Call, U. Fiedler U. Compartment-specific accumulation of recombinant proteins in plant cells: an essential tool for antibody production, protein purification and functional protein and pathway identification. Plant Methodol Biotechnol 2004; 7:445-450; PMID:15102724.

J. V. Drage, M. J. B. Mitchison, J. R. T. H. Ryan, et al. Osmotin in the vacuole is a signal for retention in the endoplasmic reticulum. Nature 2004; 429:117-121; PMID:15458686.

S. J. Steady, J. R. T. H. Ryan. Osmotin is essential for targetting to the endoplasmic reticulum and for targeting to the endoplasmic reticulum. Planta 2000; 196:586-591; PMID:19137543.

A. O. Schuett, I. Schäffel, H. Apel, S. Rieger, et al. Osmotin as a substrate for the processing of human interleukin-6 in BRF tobacco cultured cells. Plant Biotechnol J 2007; 5:848-856; PMID:17549997.

M. J. B. Mitchison, J. R. T. H. Ryan. Osmotin occurs in the vacuole. Planta 1995; 196:586-591; PMID:19137543.

M. J. B. Mitchison, J. R. T. H. Ryan. Osmotin in the vacuole is a signal for retention in the endoplasmic reticulum. Nature 2004; 429:117-121; PMID:15458686.