8-15-2012 12:00 AM

Protein Body Formation in Stable Transgenic Plants of Nicotiana Tabacum Expressing Elastin-Like Polypeptide and Hydrophobin Fusion Proteins

Sonia P. Gutierrez, *The University of Western Ontario*

Supervisor: Dr. Rima Menassa, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Biology

PROTEIN BODY FORMATION IN STABLE TRANSGENIC PLANTS OF NICOTIANA TABACUM EXPRESSING ELASTIN-LIKE POLYPEPTIDE AND HYDROPHOBIN FUSION PROTEINS

Spine title: FORMATION OF PROTEIN BODIES IN TOBACCO STABLE TRANSGENIC PLANTS

Thesis format: Monograph

by

Sonia Patricia Gutiérrez Delgadillo

Graduate Program in Biology

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

THE UNIVERSITY OF WESTERN ONTARIO
School of Graduate and Postdoctoral Studies

CERTIFICATE OF EXAMINATION

Supervisor

Dr. Rima Menassa

Co- Supervisor

Dr. Susanne Kohalmi

Supervisory Committee

Dr. Norm Hüner

Dr. Aiming Wang

Examiners

Dr. Abdelali Hannoufa

Dr. Greg Kelly

Dr. Denis Maxwell

The thesis by

Sonia Patricia Gutiérrez Delgadillo

entitled:

PROTEIN BODY FORMATION IN STABLE TRANSGENIC PLANTS OF NICOTIANA TABACUM EXPRESSING ELASTIN-LIKE POLYPEPTIDE AND HYDROPHOBIN FUSION PROTEINS

is accepted in partial fulfillment of the
requirements for the degree of
Master of Science

Date

Chair of the Thesis Examination Board

ABSTRACT

Plants are recognized as an efficient and inexpensive system to produce valuable recombinant proteins. However, the use of plants still faces two main limitations: the low accumulation levels of some recombinant proteins and the lack of efficient protein purification methods. Two fusion partners, elastin-like polypeptides (ELP) and hydrophobin I (HFBI) were previously found to increase the accumulation of recombinant proteins and induce the formation of protein bodies (PBs) when targeted to the ER in transient expression assays. In this study I examined the effect of these tags in stable transgenic plants of two *Nicotiana tabacum* cultivars when fused to green fluorescent protein (GFP) and targeted to the ER. The ELP and HFBI fusions increased the recombinant protein accumulation and induced the formation of PBs. Furthermore, my findings suggest that these tags induce the formation of PBs in a concentration-dependent manner, where a specific level of accumulation of recombinant protein has to be reached for PBs to appear.

Keywords

Molecular farming, elastin-like polypeptides, hydrophobins, green fluorescent protein, recombinant proteins, stable transgenic plants, *Nicotiana tabacum*, protein bodies.

Dedicated to my lovely and amazing family, Tito Gutiérrez, Patricia Delgadillo, Ana María Gutiérrez, María Camila Gutiérrez and Gustavo Chaparro.

ACKNOWLEDGMENTS

First, I would like to thank my supervisor, Dr Rima Menassa, for accepting me as a graduate student in her scientific group. Thanks to you I came to Canada to one of the most prestigious universities and one of the best labs ever. You have been a guide to me in every possible way throughout the years, always supporting me. I have only gratitude for you. Thanks for sharing all your knowledge and incredible ideas with me.

I would like to thank to my co-supervisor, Dr. Susanne Kohalmi. Thanks for all the great ideas that you brought to this project and all your support during the harder days. To me you are an example for all the women in science. An inspiring scientist that is smart, energetic and very happy. Thanks for all!

I would also like to thank my advisors Dr. Norm Hüner and Dr. Aiming Wang for all the advice and new concepts that you kindly offered me; they really helped me to improve my project and grow as a scientist.

I would also like to thank my lab mates, which I consider my friends. Reza Saberianfar, thanks for being my canadian brother! You have been an incredible source of knowledge and laughs, one of the best graduate students I have ever met. I really admire you. I would also like to thank Igor Kolotilin and Angelo Kaldis for all the help they offered me during my research and all the interesting talks, stories and jokes that made my day. I would also like to thank Mistianne Fenney for all the ideas and support. Also for all the advice and little details that she shared with me and made me feel special. I would also like to thank Hong Zhu for teaching me in the best possible way and always helping me with my lab work. I would also like to thanks Eridan Pereira for all his help and support. It has been a pleasure to work with all of you and I am really going to miss you all.

Lastly, I would like to thank my friends Ana Maria Aguirre, Nadia Morales, Ruoyu Yan, Mandana Miri, Behnaz Saatian, Behnaz Nayafi and Gabriela Navarro for being with me during the happiest and harder days.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

* Standard SI units not listed

APS	ammonium persulfate
ATPS	surfactant-based aqueous two-phase system
BA	N^6-benzyladenine
BiP	chaperon binding protein
CaMV	cauliflower mosaic virus
CFP	cyan fluorescent protein
CymRSV	cymbidium ringspot tombusvirus
Cys	cysteine
DNA	deoxyribonucleic acid
dpi	days post infiltration
EDTA	ethylenediaminetetraacetic acid
ELP	elastin-like polypeptide
EM	electron microscopy
ER	endoplasmic reticulum
FDA	food and drug administration
GFP	green fluorescent protein
HFBI	hydrophobin I
HRP	horseradish peroxidase
ITC	inverse transition cycling
KDEL	lysine-aspartate-glutamic-acid-leucine ER retrieval signal
MES	2-(*N*-morpholino)ethanesulfonic acid
MS	Murashige and skoog media
NAA	napthalene acetic acid

OD	optical density
PEB	plant protein extraction buffer
PBS	phosphate buffered saline
PB	protein body
PMSF	phenylmethanesulfonylfluoride
PTGS	post-transcriptional gene silencing
PVPP	polyvinylpolypyrrolidone
PSV	protein storage vacuole
rER	rough endoplasmic reticulum
RNA	ribonucleic acid
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SSP	seed storage proteins
TEMED	tetramethylethylenediamine
TBS	tris buffered saline
TSP	total soluble protein
T_t	transition temperature
YEB	yeast extract and beef
YFP	yellow fluorescent protein

1 INTRODUCTION

1.1 Plant molecular farming

Plants are considered as a safe, efficient and inexpensive alternative system to produce a wide variety of recombinant proteins such as industrial enzymes, vaccines, antibodies and other biopharmaceuticals (Knäblein, 2005; Ma *et al*., 2005). This area of plant biotechnology is defined as molecular farming (Lico *et al.*, 2005).

Plants possess important advantages over conventional expression systems such as bacteria, yeast, mammalian and insect cells. Their major benefit lies in the low cost of large-scale production (Ma *et al.*, 2003; Twyman *et al.*, 2003), in part because the agricultural systems needed for cultivation are already available, reducing the capital and operating costs (Lico *et al.*, 2005). As the protein synthesis is conserved between animals and plants, recombinant proteins can undergo proper post-translational modifications, folding and assembly (Twyman *et al.*, 2003). Furthermore, direct oral administration of unprocessed or partially processed plant material is possible due to the absence of human pathogens and endotoxins (Fischer *et al*., 2004; Commandeur *et al*., 2003).

The use of plants as bioreactors for the production of recombinant proteins has recently achieved its first success (Maxmen, 2012). The plant-produced taliglucerase alfa for treatment of Gaucher disease was approved in May 2012 by the FDA. This is the first example of an approval for a plant-made therapeutic protein and the first time a major pharmaceutical company embraced this technology.

1.2 *Nicotiana tabacum* as a plant-based production platform

Although a wide range of host plants have been developed and tested for molecular

farming, tobacco still remains one of the favorite hosts for the commercial production of recombinant proteins (Conley *et al*., 2011b; Commandeur *et al*., 2003). Methods for genetic manipulation and transformation of tobacco are well established, and large-scale production of recombinant proteins at low cost is facilitated by its high biomass yield (more than 100,000 kg of tissue per hectare) (Twyman *et al.*, 2003; Sheen, 1983). Moreover, different *Nicotiana* varieties have been characterized according to their agronomic properties and ability to accumulate recombinant proteins at high levels, therefore facilitating the search for the most effective tobacco host for recombinant protein production (Conley *et al.*, 2011a). For example, cultivar I64 has been classified as a high biomass cultivar (Conley *et al.*, 2011a), whereas cultivar 81V9 has been classified as a low alkaloid cultivar that has been used for direct oral administration of the protein products (Conley *et al.*, 2009a; Menassa *et al*., 2007; Menassa *et al*., 2001).

Because the tobacco expression platform is based on leaves, harvest occurs before flowering, therefore the possibility of gene leakage into the environment through pollen or seed dispersal is generally reduced (Rymerson *et al*., 2002). Furthermore, tobacco is a non-food, non-feed crop, which eliminates the risk of plant made recombinant proteins entering the human and animal food supply (Twyman *et al.*, 2003; Menassa, *et al.*, 2001).

Although tobacco is inherently biosafe, the commercial viability of molecular farming with this plant species has been limited by two main factors: the low accumulation levels of some recombinant proteins (usually due to proteolytic degradation or poor stability), and the lack of efficient and scalable protein purification methods (Conley *et al*., 2011a,b; Twyman *et al.*, 2003). Because of low protein expression levels in plants, yield is the biggest hurdle for successful production of recombinant proteins

(Doran, 2006). Therefore, it is crucial to address this issue for the commercial success of a tobacco-based expression system.

1.3 Plant seeds: organs naturally designed for protein storage

Plant seeds are known for their natural ability to synthesize and store large amounts of complex proteins over long periods of time in a very stable environment (Lau and Sun, 2009; Torrent *et al.*, 2009a). These characteristics are highly desirable for a molecular farming system because high accumulation levels (yield gain) and long term storage of recombinant proteins in a small volume can be obtained (Lau and Sun, 2009; Stöger *et al*., 2005). Therefore, studying the accumulation of seed storage proteins (SSPs) is expected to provide information that can be used to increase the amount of recombinant proteins in seeds (Kawakatsu and Takaiwa, 2010).

Cereal seeds from rice, barley, wheat and mainly maize, are currently well-developed platforms used for the production of recombinant proteins (Lau and Sun, 2009). The main reasons for their success are their high grain yield, their relatively high seed protein content and the existence of well-developed molecular tools for their genetic manipulation (Lau and Sun, 2009; Ma *et al.*, 2003).

Based on their solubility, plant SSPs have been classified into albumins (water soluble), globulins (salt soluble), prolamins (aqueous alcohol soluble) and glutelins (dilute acid or alkali soluble) (Kawakatsu and Takaiwa, 2010; Shewry and Halford, 2002; Stöger *et al.*, 2001). It has been established that albumins and globulins are the major SSPs of dicots, whereas prolamins and glutelins are the major SSPs in monocots (Shewry

and Halford, 2002; Stöger *et al.*, 2001; Mandal and Mandal, 2000; Herman and Larkins, 1999).

SSPs are synthesized on rough endoplasmic reticulum (rER) membranes and after co-translational cleavage of their N-terminal signal peptide, they are translocated to the ER lumen (Shewry and Halford, 2002). Most SSPs are deposited into protein storage vacuoles (PSVs). Alternatively, prolamins accumulate within the ER lumen and form ER-derived accretions called protein bodies (PB), which are surrounded by a membrane of ER origin (Kawakatsu and Takaiwa, 2010; Shewry and Halford, 2002; Herman and Larkins, 1999; Coleman *et al*., 1996; Larkins and Hurkman, 1978). Prolamins lack the classical C-terminal KDEL or HDEL motif for retention within ER. Therefore, it appears they have intrinsic structural features responsible for their retention in the ER (Shewry and Halford, 2002; Herman and Larkins, 1999; Coleman *et al*., 1996; Munro and Pelham, 1987). The ability to sequester proteins in these large and dense PBs seems to be the key to enhance accumulation of proteins in seeds (Torrent *et al*., 2009a Galili, 2004). Unfortunately, the mechanisms behind transport and deposition of prolamins in PBs are still not clear (Llop-Tous *et al.,* 2010).

Although the use of seeds as platforms for molecular farming was initially popular, concerns by the public and regulatory agencies have led to an effective worldwide decrease in their use (Commandeur *et al*., 2003). Specifically, regulators are concerned about the spread of transgenes to wild plants and transgenic seeds entering human or animal food chains. To circumvent this, the use of non-food leaf-based production systems for the generation of recombinant proteins has been suggested (Rybicki, 2010; Knäblein, 2005; Commandeur *et al*., 2003). However, the use of the natural mechanism

of protein accumulation in seeds, combined with leaf-based production systems, offers a valuable option for an enhanced production of recombinant proteins.

1.4 Adapting seed mechanisms to improve leaf-based production systems

The mechanisms behind prolamins aggregation, which are responsible for PB assembly, as well as the mechanisms that determine their ER localization are only partially understood (Llop-Tous *et al*., 2010; Torrent *et al*., 2009a). Prolamins are characterized by their high content in proline and their solubility in alcohol, which reflects their general hydrophobic nature (Shewry and Halford, 2002). According to these properties, it was assumed that they aggregate in a nonspecific manner (Argos *et al*., 1982). However, studies in maize showed that PBs were directly formed in the lumen of the ER and contained four structurally different maize prolamins: α-, β-, δ-, and γ-zein (Shewry and Halford, 2002; Herman and Larkins, 1999; Lending and Larkins, 1989; Larkins and Hurkman, 1978). Zeins are distributed heterogeneously in PBs. Initially the small PBs consists of β- and γ-zein (sulphur-rich prolamins) aggregates (Lending and Larkins, 1989). This is followed by penetration and accumulation of α- and δ-zeins into the center of the PB, expanding it until it reaches a diameter of 1 to 2 μm. As the center of the PB is filled, β- and γ-zeins are concentrated toward the periphery of PBs (Kim *et al*., 2002; Herman and Larkins, 1999; Lending and Larkins, 1989). This spatial organization of zeins in PBs suggests that each of these proteins has specific properties, responsible for ER retention and localization within the PBs (Coleman *et al*., 1996).

Further characterization of γ-zein showed that it facilitates the assembly of the other PB-associated zeins and stabilizes them (Coleman *et al*., 1996). Moreover,

heterologous expression of γ-zein in leaves promotes the formation of PB-like structures inside the ER, suggesting that structural motifs within γ-zein may be responsible for PB induction. It was later shown that the proline-rich repetitive sequences $(PPPVHL)_8$ present at its N-terminus and a central Pro-X domain are responsible for γ-zein's ability to self-assemble and be retained within the ER (Torrent *et al*., 2009b; Geli *et al*., 1994). Therefore, a synthetic peptide of the N-terminal proline-rich repetitive sequence $(PPPVHL)_8$ of γ-zein was synthesized and tested *in vitro*. Surprisingly, this peptide was able to self-assemble and form cylindrical micelles, which are globular aggregates where the hydrophilic regions of the peptide are exposed to the surface and the hydrophobic regions of the peptide are clumped together in the center of the aggregate (Kim and Hu, 2008; Kogan *et al.*, 2002).

Consequently, based on this natural ability of the seeds, a fusion protein-based system designed to accumulate recombinant proteins in ER-derived PBs in plants was developed (Torrent *et al*., 2009a). The approach relies on fusion proteins that contain the γ-zein signal peptide, the γ-zein proline-rich domain (Zera®; developed by ERA Biotech, Barcelona, Spain) and the protein of interest fused to the C-terminus of the proline-rich region (Torrent *et al*., 2009b). The formation of γ-zein induced PBs was confirmed in plant vegetative tissue and in several non-plant eukaryotic systems (i.e. fungal, insect and mammalian cells), where enhanced recombinant protein accumulation by using Zera was also evidenced (Llop-Tous *et al.*, 2010; Torrent *et al*., 2009b; StorPro organelles http://www.erabiotech.com). Therefore, this new system allows the expression of heterologous recombinant proteins in non-seed tissues with three main advantages: proteins remain stable inside membrane-bound PBs, the induced PBs are dense organelles

which can be easily purified by centrifugation on density gradients and, the presence of these PBs does not affect normal growth and development of the hosts (Conley *et al*., 2011b; Torrent *et al*., 2009a,b).

1.5 Fusion partners as enhancers of recombinant protein accumulation and PB inducers

Meanwhile, a wide variety of strategies have been tested in plant-based systems to increase stability and yield of recombinant proteins, including protein targeting to different sub-cellular compartments, the use of different promoters and expression vectors, the use of different expression hosts and expression in different tissues (Schillberg *et al*., 2003). Among these approaches, subcellular targeting is one of the most important factors influencing yield of recombinant proteins because the environment in the compartment is essential for proper folding, assembly and addition of post-translational modifications (Fischer *et al*., 2004). As the point of entry to the secretory pathway, the ER is an attractive compartment for foreign protein expression, as it provides an oxidizing environment with few proteases and a high number of chaperones and stabilizing agents that help in folding and assembly of proteins. Additionally, quality control of the incoming or newly assembled proteins occurs because ER-resident proteins target misfolded polypeptides for disposal (Fischer *et al*., 2004; Hadlington and Denecke, 2000).

In recent years, fusion protein technology has been also used to enhance recombinant protein accumulation in heterologous systems. Among these, ER-targeted fusion proteins with elastin-like polypeptides (ELP) and hydrophobin (HFBI) have been

of particular interest as they allow to increase the accumulation and stability of recombinant proteins in plants, and assist in the subsequent purification process of the recombinant proteins (Conley *et al*., 2011b; Joensuu *et al*., 2010).

1.5.1 Elastin-like polypeptides as fusion partners

ELPs are synthetic biopolymers composed of a repeating pentapeptide sequence 'VPGXG' where the "guest residue" X can be any amino acid except proline. These repeats occur naturally in all mammalian elastin proteins (Raju and Anwar, 1987). In aqueous solution, ELPs undergo a reversible inverse phase transition from soluble protein to insoluble hydrophobic aggregates that form β-spiral structures when heated above their transition temperature (T_t) (Urry, 1988). This thermally responsive property of ELPs is also transferred to fusion partners, providing a simple, rapid and scalable non-chromatographic method for protein purification called 'Inverse Transition Cycling' (ITC) (Meyer and Chilkoti, 1999).

It was previously shown by ELP expression in bacteria that the T_t of an ELP tag varies with its sequence, length and concentration. A significant increase in T_t was observed with short ELP tags (Meyer and Chilkoti, 2004). Therefore, long ELPs (i.e. 90-180 pentapeptide repeats) have been used for ITC purification studies.

Protein fusions with synthetic ELP tags retrieved to the ER using a C-terminal peptide (H/KDEL) have been successfully produced in plants and purified with ITC (Floss *et al*., 2009; Hadlington and Denecke, 2000). Some examples include fusions with cytokines (Conley *et al.*, 2009a*;* Patel *et al.*, 2007; Lin *et al.*, 2006), antibodies (Conley *et al.*, 2009a; Joensuu *et al*., 2009; Floss *et al*., 2008) and spider silk (Patel *et al.*, 2007;

Scheller *et al*., 2004). All of these ELP fusions were tested in stable transgenic plants of *Nicotiana tabacum* and it was observed that ELP tags significantly enhanced accumulation in tobacco leaves (Conley *et al.*, 2009a; Floss *et al*., 2008; Patel *et al.*, 2007). However, expression of ELP fusion proteins both in plants and in bacteria showed that the concentration of recombinant protein was inversely proportional to the size of ELP tag (Conley *et al.*, 2009a; Meyer and Chilkoti, 1999). Therefore, studies to determine the optimal ELP size for the accumulation of recombinant proteins and their subsequent purification from leaf tissue were carried out. Different ELP tags were fused to several proteins and were transiently expressed in tobacco leaves (Conley *et al.*, 2009a). It was established that 30 pentapeptide repeats are the best compromise between the positive effects of smaller ELP tags on the accumulation of recombinant proteins and the positive effect of larger ELP tags on purification efficiency during ITC (Conley *et al.*, 2009a). Furthermore, it was found that the position of the ELP tag at the C-terminal of the fusion promotes the enhancing effect on accumulation of the recombinant proteins (Christensen *et al.*, 2009; Conley *et al.*, 2009a).

Although optimized conditions for ELP use have been found, ITC purification and recovery rates of some recombinant proteins with very low accumulation levels in plants are not significant. Therefore, the use of other purification techniques, such us affinity chromatography, are still required in these cases (Conley *et al.*, 2011b).

To better understand the role of ELP tags in the enhanced accumulation of fusion proteins in tobacco plants, recent studies of transient expression assays were performed (Conley et al., 2009b). These assays incorporated the p19 suppressor of post-transcriptional gene silencing (PTGS) (from *Cymbidium* ringspot tombusvirus)

(CymRSV), which had been found to significantly increase the production levels of recombinant proteins in plants (Sylhavy *et al.*, 2002).

Transient expression of an ER-targeted green fluorescent protein (GFP)-ELP fusion showed that the ELP not only increased accumulation levels of the recombinant protein, but also induced formation of PB-like structures in leaves (Figure 1A and 1B; Conley *et al.*, 2009b). To confirm that the observed structures were ER-derived, several tests were performed. This included simultaneous expression in *Nicotiana benthamiana* of an ER-resident molecular chaperone binding protein (BiP) fused with cyan fluorescent protein (CPF) and an ER-targeted yellow fluorescent protein (YFP)-ELP fusion. This experiment showed co-localization of both fluorescent proteins in PBs, supporting the idea that these accretions are ER-derived. Moreover, electron microscopy (EM) and immunogold labelling were used to analyze leaf tissue expressing ER-targeted GFP-ELP fusion proteins. These experiments revealed that PBs are surrounded by a membrane studded with ribosomes, further suggesting that they are ER-derived (Figure 1C). In addition, PBs were labeled with immunogold particles targeting ELP and GFP proteins to confirm the presence of GFP-ELP inside the induced PBs (Figure 1D, Conley *et al.*, 2009b). Together, these findings are consistent with the idea that ELP-induced PBs are ER-derived. Therefore, this new property provided new insights into the mechanisms behind the positive effect of ELP on accumulation of recombinant proteins.

Based on experimental evidence to date, it is believed that ELP tags increase the stability and solubility of target proteins, leading to an increase in the accumulation levels of the recombinant proteins. Consequently, an increase in the accumulation levels induces the formation of PBs (Conley *et al.*, 2009a,b). This effect seems to be explained

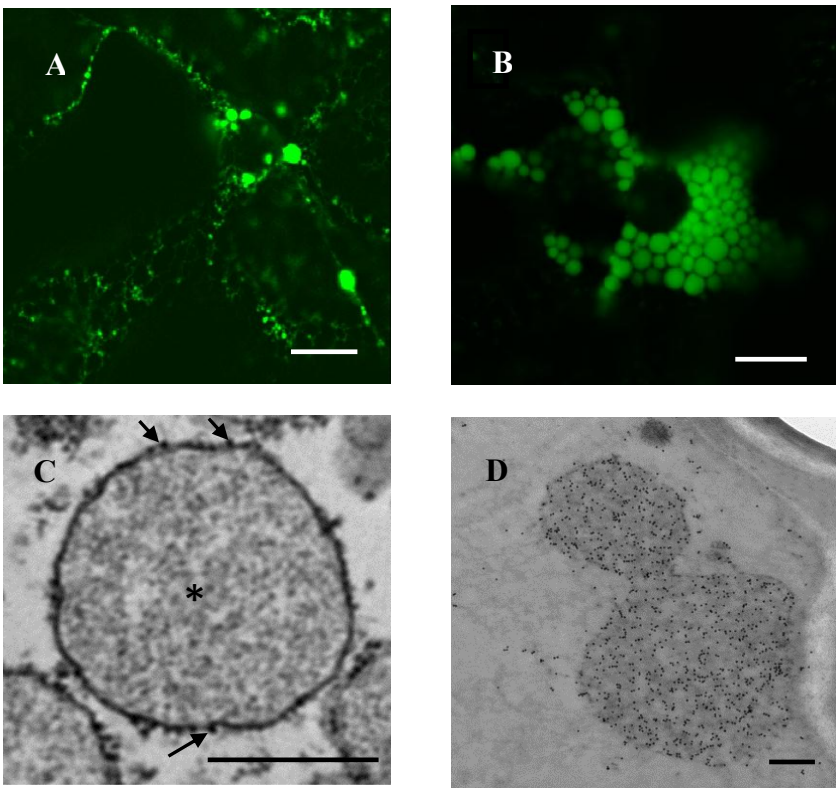


Figure 1. Protein bodies induced in tobacco leaves by the expression of an ER-targeted GFP-ELP fusion.

A. The novel PBs were closely associated with the ER tubules as small punctuate structures early on in the PB-formation process. B. Different sizes of the generated PBs. The PBs tend to cluster together within the cell, although the distribution pattern was variable. C. PB (indicated by an asterisk) membrane decorated with ribosomes (indicated with arrows). The presence of the ribosomes in the membrane suggests that they were originally derived from the rough ER. D. Immunogold localization confirmed the presence of green fluorescent protein (GFP)-ELP fusion protein inside the PBs using anti-ELP antibodies. Confocal microscopy images A and B of leaf cells courtesy of Reza Saberianfar. Electron microscopy images C and D courtesy of Dr. Jussi Joensuu. Bars= 10µm (A), 10µm (B), 500nm (C), 500nm (D).

by specific biochemical properties of ELP, reminiscent of seed prolamins like γ-zein. ELPs and γ-zein proteins are hydrophobic and proline-rich, characteristics responsible of the self-assembly properties of prolamins in a wide range of hosts (Llop-Tous *et al.*, 2010). Therefore, the ELP-induced PBs enhance the accumulation levels of the fusion proteins probably by inhibiting their physiological turnover (Conley *et al.*, 2009a,b).

1.5.2 Hydrophobin I as a fusion partner

Hydrophobins are a group of small, surface-active proteins originally identified in filamentous fungi (Nakari-Setälä *et al.*, 1996). These proteins play an important role in fungal growth and development, involving adaptation of fungi to their environment by controlling surface interactions (Linder, 2009; Hakanpää *et al*., 2004). One remarkable feature of these proteins is that one part of their surface is made of hydrophobic aliphatic side chains that form an exposed 'hydrophobic patch' on one side of the protein. This is an uncommon characteristic as hydrophobic residues are usually buried in the core of hydrophobic proteins, thus stabilizing their conformation (Linder, 2009). Alternatively to a core stabilized by hydrophobic interactions, hydrophobins possess a characteristic pattern of eight conserved cysteine (Cys) residues, which form four intramolecular disulfide bridges that convey a high degree of protein stability. Based on these characteristics it is clear that hydrophobins are amphipathic molecules, with a hydrophilic and a hydrophobic region, similar to surfactants structure (Joensuu *et al*., 2010; Conley *et al*., 2011b; Linder, 2009).

Due to their structural properties, hydrophobins can self-assemble into an amphipathic protein membrane at hydrophilic-hydrophobic interfaces (Wang *et al.*, 2005;

Wösten and Vocht, 2000). Furthermore, it has been shown that they have a very high affinity for surfactants. Therefore, purification of HFBI can be facilitated by using a two-step surfactant-based aqueous two-phase system (ATPS) (Linder *et al.*, 2001). During ATPS, a surfactant is added to the crude protein extract, which concentrates the hydrophobin fusions inside micellar structures and partitions them towards the surfactant phase (Lahtinen *et al*., 2008). Later, the fusion protein can be recovered from the surfactant phase with a non-denaturing organic solvent, such as isobutanol (Linder *et al.*, 2001). The selective and high yield extraction of fusion proteins from fungal culture supernatants, that contain other naturally secreted proteins, validated the efficiency of this technique (Linder *et al.*, 2004; 2001).

Importantly, when fused to other proteins, hydrophobins can alter the hydrophobicity of the fusion partner allowing for simple, rapid, efficient, scalable, and inexpensive purification using ATPS (Linder *et al.*, 2004). Recovery of an ER targeted GPF-hydrophobin fusion protein from plant leaf extracts showed that it selectively recovered up to 91% of the fusion protein (Joensuu *et al*., 2010).

Hydrophobins have been divided into two classes (class I and class II) based on the presence of hydrophilic and hydrophobic amino acid residues in their protein sequence. Members of class I are characterized by their highly insoluble aggregates in aqueous solution, whereas aggregates of the members of class II are easier to dissolve (Linder, 2009). Hydrophobin I (HFBI) from *Trichoderma reesei* is a class II hydrophobin (Nakari-Setälä *et al.*, 1996) and has been used as a fusion tag to over express and purify recombinant proteins from *Trichoderma* sp., insect cells and plant tissues (Joensuu *et al*., 2010; Lahtinen *et al*., 2008; Linder *et al.*, 2004).

A recent study where an ER-targeted GPF-HFBI fusion protein was transiently expressed in the presence of p19 (Sylhavy *et al.*, 2002) in *N. benthamiana* plants by agro-infiltration, showed that the HFBI tag not only significantly increased the expression of the fusion protein, but also induced formation of PBs comparable to the ones observed with the ELP tag (Figure 2). Transmission electron microscopy and immunogold labelling revealed that similar to PBs seen with the ELP tag, PBs were surrounded by a membrane studded with ribosomes. Also, immunogold labeling for GFP revealed specific gold decoration of the PBs (Joensuu *et al*., 2010).

Interestingly, a protective effect of the HFBI tag in the infiltrated tissue was observed (Joensuu *et al*., 2010). Although overexpression in transient systems usually leads to necrotic lesions four days post infiltration (dpi), with the HFBI tag the tissue remained healthy up to 10 dpi. Therefore, recombinant protein accumulation could continue. A possible explanation of this protective effect may be the sequestration of the fusion proteins in PBs. While being stored, possible negative effects of overexpression of a foreign protein in the plant cell can be prevented (Joensuu *et al*., 2010).

1.6 Fusion partners and stable transgenic plants

Subsequent analysis of ELP and HFBI fusions for protein recovery and biological activity showed that the proteins of interest do maintain their function (Joensuu *et al*., 2010; Conley *et al*., 2009a; Floss *et al*., 2009). These promising results have triggered a renewed interest in high level production and purification of recombinant proteins in plants (Floss *et al*., 2009). However, most studies to date have been carried out using transient expression by agro-infiltration while the main goal of using plants as bioreactors

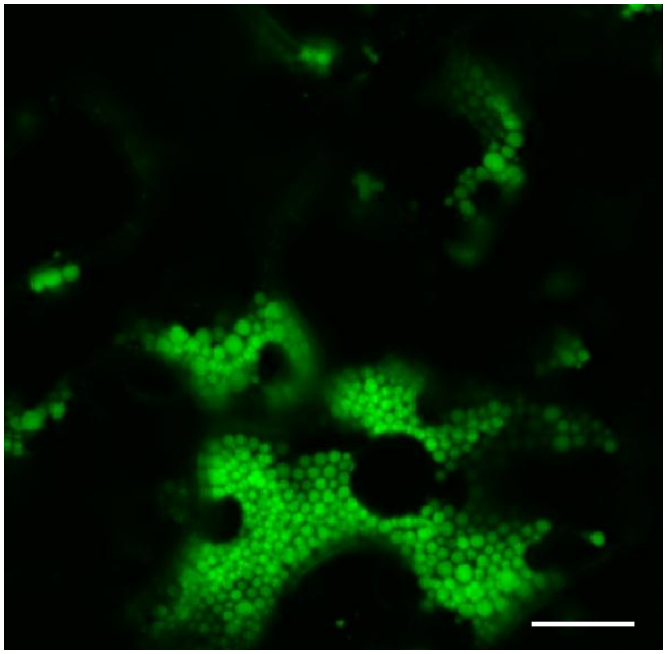


Figure 2. Expression of an ER-targeted GFP-HFBI fusion induces the formation of protein bodies in tobacco leaves.

Confocal microscopy image of GFP-HFBI transient expression assay in *N. benthamiana*. Leaf epidermal cell accumulating GFP-HFBI fusion protein in PBs. Bar = 10 µm. Confocal microscopy images courtesy of Reza Saberianfar.

as a cheap and economic source of recombinant proteins is to generate stable transgenic plants for large-scale production. Hence, it was necessary to evaluate the effect of the ELP and HFBI tags, targeted to the ER, for recombinant protein accumulation and PBs formation in stable transgenic plants.

1.7 Hypothesis and thesis objectives

I hypothesize that PBs will be formed in stable transgenic tobacco plants expressing ER-targeted GFP-ELP and GFP-HFBI recombinant proteins. Furthermore, I hypothesize that PB formation in tobacco plants will depend on the accumulation levels of the recombinant protein.

The objectives of this thesis were:

1- To generate tobacco stable transgenic plants of *N. tabacum* to express ER-targeted GFP, GFP-HFBI and GFP-ELP recombinant proteins to determine if PBs will form.

2- To evaluate and compare accumulation levels of recombinant proteins and the presence of PBs in leaves of stable transgenic plants expressing ER-targeted GFP, GFP-HFBI and GFP-ELP recombinant proteins to determine if there is a threshold value of recombinant protein accumulation needed for the formation of PBs.

3- To compare two tobacco cultivars, 81V9 and I64, for accumulation levels and PB formation to determine if the behavior of these tags changes between these two cultivars.

2 MATERIALS AND METHODS

2.1 Media, buffers and polyacrylamide gels

2.1.1 Media

Antibiotics

Stock solutions of timentin (250 mg/ml) and kanamycin (100 mg/ml) were prepared by dissolving each antibiotic in milliQ water. The stock solution of rifampicin (10 mg/ml) was prepared by dissolving the antibiotic in 100% methanol. Once dissolved, each antibiotic solution was filter-sterilized and stored at -20ºC.

Hormones

Stock solutions of 2-Naphthalene-acetic acid (NAA) (2.8 mg/ml) and N^6-Benzyladenine (BA) (1 mg/ml) were prepared by dissolving each compound in 1 N NaOH. Once dissolved, each stock was filter-sterilized and stored at 4ºC.

Murashige and Skoog medium

To prepare 1 liter of media, 4.4 g of Murashige and Skoog (MS) (Sigma, Cat. No. M5519) salt and 40 g of sucrose were dissolved in 750 ml of milliQ water. The pH was adjusted to 5.8 with HCl before 10 g of agar were added. The final volume was adjusted to 1 liter with milliQ water. This medium was used for growing wild type tobacco plants for stable transformation (MST1 medium). For co-cultivation for *Agrobacterium* transformation, 1 ml of BA and 35 µl of NAA were added to the MST1 preparation before adjusting the pH (MST2 medium). For regeneration and shoot induction of

tobacco explants, 2 ml of timentin and 1 ml of kanamycin stock solutions were added to the MST2 preparation after autoclaving (MST3 medium). For selection and root induction of tobacco plants, 2 ml of timentin and 1 ml of kanamycin were added to the MST1 preparation after autoclaving (MST4). This media was poured into phytotray boxes (SIGMA Aldrich). For root induction without selection, 18 µl of NAA were added to the MST1 original preparation before adjusting the pH. After autoclaving, 2 ml of timentin were added (MST5 medium). MST1 media was poured into magenta boxes. MST2 and MST3 media were poured into deep petri dishes. MST4 and MST5 media were poured into phytotray boxes (SIGMA Aldrich).

Yeast extract and beef medium

To prepare 1 liter of liquid media, 1 g of yeast extract, 5 g of beef extract, 5 g of peptone, 5 g of sucrose and 2 ml of 1 M $MgSO_4$ were dissolved in 900 ml of milliQ water, and the final volume was adjusted to 1 liter with milliQ water. To prepare 1 liter of solid media, 15 g of agar were added. After autoclaving, 1 ml of rifampicin and 500 µl of kanamycin stock solutions were added.

Agrobacterium induction medium

To prepare 50 ml of media, 500 µl of (2-(N-morpholino)ethanesulfonic acid) (MES) (1 M MES stock in milliQ water, pH adjusted to 5.6 with KOH), 25 µl of acetosyringone (200 mM stock), 25 µl of kanamycin and 50 µl of rifampicin stock solutions were dissolved in 49.4 ml of YEB liquid media.

Gamborg solution (*Agrobacterium* infiltration solution)

To prepare 100 ml of solution, 0.32 g of Gamborg`s B5 media with vitamins (Research Products International Corp., Order No. G20200), 2.0 g of sucrose, 1 ml of MES (1M MES stock, pH 5.6) and 100 µl of acetosyringone (200 µM stock) were dissolved in 98.9 ml of milliQ water.

2.1.2 Buffers and solutions

10X Phosphate buffered saline (PBS)

To prepare 1 liter of buffer, 80 g of sodium chloride, 2 g of potassium chloride and 11.6 g of disodium hydrogen phosphate were dissolved in 800 ml of milliQ water. The pH was adjusted to pH 7.4 with HCl and the final volume was adjusted to 1 liter with milliQ water.

1X PBS-T

To prepare 1 liter of buffer, 100 ml of 10X PBS and 1 ml of Tween-20 were diluted in 899 ml of milliQ water.

Plant protein extraction buffer (PEB)

To prepare 100 ml of buffer, 100 µl of Tween-20 (0.1%) and 2 g of polyvinylpolypyrrolidone (2%) (PVPP) were dissolved in 96 ml of 1X PBS. The solution was kept overnight at 4ºC to allow the PVPP to get wet. Just prior to using, 200 µl of ethylenediaminetetraacetic acid (EDTA) (0.5 M stock), 1 ml of phenylmethanesulfonylfluoride (PMSF) (100 mM stock), 100 µl of leupeptin (1 mg/ml

stock) and 2 g of sodium L-ascorbate were added to the solution. The resulting solution was kept on ice and used immediately.

10X Tris-buffered saline (TBS)

To prepare 1 liter of buffer, 24.2 g of Tris-base and 175.3 g of NaCl were dissolved in 800 ml of milliQ water. The pH was adjusted to pH 7.5 with concentrated HCl and the final volume was adjusted to 1 liter with milliQ water.

1X TBS-T

To prepare 1 liter of buffer, 100 ml of 10X TBS and 1 ml of Tween-20 were dissolved in 899 ml of milliQ water.

Separating gel buffer

To prepare 100 ml of 1.5 M Tris-HCl buffer, 18.15 g of Tris-base were dissolved in 95 ml of milliQ water. The pH was adjusted to 8.8 with HCl and the final volume was adjusted to 100 ml with milliQ water.

Stacking gel buffer

To prepare 100 ml of 0.5 M Tris-HCl buffer, 6 g of Tris-base were dissolved in 95 ml of milliQ water. The pH was adjusted to 6.8 with HCl and the final volume was adjusted to 100 ml with milliQ water.

5X Electrode running buffer

To prepare 1 liter of buffer , 15 g Tris-base, 72 g glycine and 5 g sodium dodecyl

sulphate (SDS) were dissolved in 1 liter of milliQ water.

Blocking solution

To prepare 200 ml of solution, 10 g of Carnation skim milk powder were dissolved in 200 ml of 1X TBS-T buffer.

1X Semi-dry trans blot - Protein transfer buffer

To prepare 1 liter of buffer, 3.03 g of Tris-base and 14.415 g of glycine were dissolved in 800 milliQ water.

5X Reducing sample buffer

Fisher, Cat. No. PI39000.

2.1.3 Polyacrylamide gels

Separating gel

To prepare a 10 % gel (10 ml final volume), 3.94 ml of milliQ water, 2.50 ml of separating gel buffer, 100 µl of 10% SDS, 3.40 ml of 30% acrylamide (Bio-Rad), 50 µl of 10% ammonium persulfate (APS) and 15 µl tetramethylethylenediamine (TEMED) were mixed. APS and TEMED were added immediately prior to pouring the gel.

Stacking gel

To prepare a 4 % gel (5 ml final volume), 3 ml of milliQ water, 1.25 ml of stacking gel buffer, 50 µl of 10% SDS, 666 µl of 30% acrylamide, 25 µl of 10% APS and 10 µl TEMED were mixed. APS and TEMED were added immediately prior to pouring the gel.

2.2 Strains, plants and plasmids

2.2.1 Bacterial strains and growth conditions

Agrobacterium tumefaciens EHA 105 (Hood *et al*., 1993) strain was used for stable transformation of tobacco and for transient expression assays. *A. tumefaciens* was grown at 28ºC in liquid or solid YEB media supplemented with appropriate antibiotics. When grown in solid media, cultures were incubated for two days. When grown in liquid media, cultures were incubated at 250 rpm shaking overnight.

2.2.2 Plasmids

Binary expression vectors, previously designed and built in our laboratory by Dr. Andrew Conley and Dr. Jussi Joensuu (Joensuu *et al*., 2010; Conley *et al*., 2009b), were used for tobacco stable transformation and tobacco transient expression assays (Figure 3). The constructs were designed to express ER-targeted GFP, GFP-HFBI and GFP-ELP fusion proteins in tobacco plants under the control of the dual-enhancer cauliflower mosaic virus (CaMV) 35S promoter (Kay *et al*., 1987), a tCUP translational enhancer (Wu *et al*., 2001), and the nopaline synthase (nos) terminator (Bevan *et al*., 1983) in the plant binary expression vector pCaMterX (Harris and Gleddie, 2001).

2.2.3 Plant growth

Stable transformations were carried out in *N. tabacum* plants of cv. 81V9 and I64. Transient expression assays of p19 protein were carried out in *N. tabacum* stable transgenic plants of cv. 81V9 and I64 expressing GFP-HFBI recombinant protein. Transient expression assays for GFP-HFBI recombinant protein purification were carried

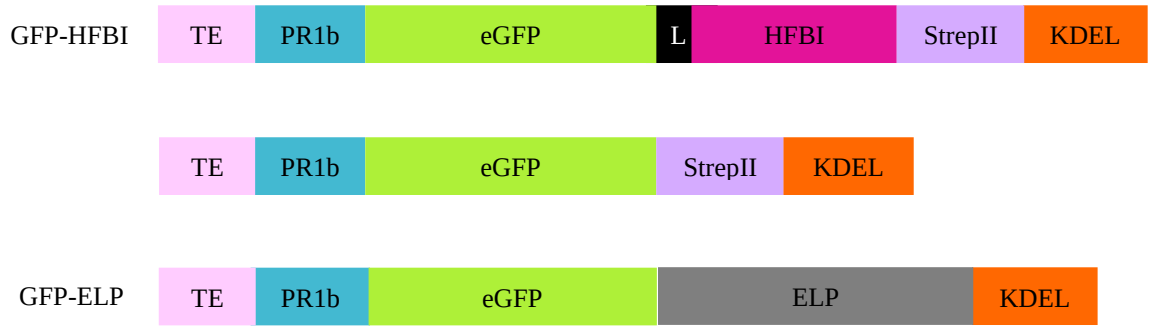


Figure 3. Schematic representation of DNA constructs used to generate stable transgenic plants of *N. tabacum*.

From right to left: TE, Translational enhancer from the tobacco tCup promoter. PR1b, Secretory signal peptide from the pathogenesis related protein 1. eGFP, Enhanced green fluorescent protein. L, Linker $(GGGS)_3$. HFBI, Hydrophobin I tag. ELP, Elastin-like polypeptide tag. StrepII, StrepII tag for detection and purification of recombinant protein. KDEL, ER retention peptide. Size of each element of the DNA constructs is not proportional to the actual sequence length (Joensuu *et al*., 2010; Conley *et al*., 2009b).

out in *N. benthamiana* plants.

2.2.3.1 Seed sterilization and germination

Seeds of *N. tabacum* cultivars 81V9 and I64 were individually placed in the middle of a kimwipe tissue. The tissue was folded to form a small bag and later submerged in 70% ethanol for 1 minute. The bag was then submerged in milliQ water for 2 min. This step was repeated one more time. The bag was then opened and the seeds were allowed to dry. The seeds were then placed on MST1 plates and incubated in a growth chamber at 24ºC with a 16 hours photoperiod for 1 week.

2.2.3.2 Tobacco plants for stable and transient transformation

N. tabacum plants of cv. 81V9 and I64 were grown in sterile magenta boxes containing 70 ml of MST1 medium. The plants were maintained in sterile conditions in a growth cabinet at 24ºC with a 16 hours photoperiod. *N. benthamiana* plants were individually grown in pots with professional growing medium (PROMIX BX Mycorrhizae). The plants were maintained in a growth room at 22ºC with a 18 hours photoperiod.

2.2.3.3 Tobacco stable transgenic plants

Transgenic plants of *N. tabacum* cv. 81V9 and I64 were individually grown in 7” pots with professional growing medium (PROMIX BX Mycorrhizae). The plants were maintained in the greenhouse.

2.3 Plant transformation

2.3.1 Tobacco stable transformation

Plant transformations were carried out using *A. tumefaciens* (Horsch *et al*., 1985). *A. tumefaciens* transformants carrying the different expression constructs (Table 1) were streaked on YEB solid media supplemented with appropriate antibiotics and incubated for two days at 28ºC. Single colonies were picked from each plate and used to inoculate liquid 20 ml of YEB media supplemented with appropriate antibiotics. Cultures were shaken at 250 rpm overnight at 28ºC then centrifuged at 1000*g* for 30 minutes. Each pellet was resuspended in 40 ml of MST1 liquid medium.

Individual leaves from sterile *N. tabacum* plant of cv. I64 and 81V9 were cut and immersed in the appropriate *A. tumefaciens* suspension in a sterile petri dish for 30 seconds. The leaf discs were then blotted dry on sterile filter paper and placed on MST2 medium with abaxial side down. The explants were co-cultivated with *A. tumefaciens* on MST2 medium for 2 to 3 days, after which they were transferred to MST3 medium for callus induction and shoot formation. This step was repeated every 10 days until a callus with shoots was evident. Individual shoots were cut off from the callus and transferred to MST5 medium for rooting. For every combination of construct and cultivar a total of forty independent transformed shoots were transferred to MST5 medium.

When the shoots rooted and their leaves were dark green in colour they were transferred to MST4 medium for further selection. Once the plants were about 4 cm in height, they were transplanted into soil. The plants were placed and maintained in the greenhouse. To protect the plants from drying out in the greenhouse they were covered with a plastic bag for 4 days. Each day the bag was lifted until the plants were completely

Table 1. Tobacco stable transgenic lines produced with the three expression constructs.

	Construct name	Description of recombinant protein	Number of transgenic lines
Cultivar I64	1. GFP	GFP, targeted to ER	24
	2. GFP-ELP	GFP with ELP tag, targeted to ER	24
	3. GFP-HFBI	GFP with HFBI tag, targeted to ER	24
Cultivar 81V9	4. GFP	GFP, targeted to ER	24
	5. GFP-HFBI	GFP with HFBI tag, targeted to ER	24

exposed to the greenhouse environment. A total of twenty four independent stable transgenic plants from each cultivar, transformed with different constructs, were maintained for further analyses.

2.3.2 Tobacco transient expression

For transient expression analysis in *N. tabacum* stable transgenic plants, leaves of plants expressing GFP-HFBI recombinant protein were pressure infiltrated (Kapila *et al*., 1997) with an *A. tumefaciens* culture expressing p19 from CymRV (Havelda, 2003; Sylhavy *et al.*, 2002). An *A. tumefaciens* 3 ml culture was grown in YEB medium supplemented with appropriate antibiotics overnight at 28ºC in an orbital shaker set at 250 rpm. A 20 µl aliquot of this starter culture was used to inoculate 20 ml of YEB liquid medium supplemented with appropriate antibiotics. The culture was grown to an optical density at 600nm (OD_{600}) of 0.5-0.8. Cultures were centrifuged at 1000*g* for 30 minutes and the pellet was resuspended in agroinfiltration solution to a final OD_{600} of 0.3 and then incubated at room temperature with gentle agitation for 1 hour. A young leaf, a medium sized leaf and an old leaf from different stable transgenic plants were pressure infiltrated with the bacterial suspension using a needle-less syringe. Leaf tissue samples from transient transformants were collected 4 days after infiltration for confocal microscopy and protein extraction.

For transient expression analysis in *N. benthamiana* plants, leaves of 6 week-old plants were co-infiltrated with an *A. tumefaciens* culture expressing GFP-HFBI recombinant protein and an *A. tumefaciens* culture expressing p19 from CymRV (Havelda, 2003; Sylhavy *et al.*, 2002). Strains were growth following the previous

description and for co-infiltration, equal volumes of *A. tumefaciens* cultures were combined to maintain a final OD_{600} of 0.3. Leaves were pressure infiltrated with the bacterial suspension using a needle-less syringe. Leaf tissue samples from transient transformants were collected 4 days after infiltration for protein extraction.

2.4 Seed collection from stable transgenic plants

Stable transgenic tobacco plants placed in the green house were self-pollinated. For this, the first inflorescence of each plant was covered with a mesh bag. Once the pods inside the bag were mature and dry, they were cut from the plant and placed in a paper bag. The seeds were then collected in paper envelopes. Each envelope was labeled and stored at room temperature in a cardboard box for later use.

2.4.1 Confocal microscopy of seeds from stable transgenic and wild type plants

Dry seeds were imbibed in milliQ water for 30 minutes before removing the seed coat using a needle and forceps. The seeds were placed over a drop of milliQ water previously put on a coverslip and immediately covered with another coverslip.

2.5 Tissue sample collection and storage

2.5.1 Tissue sampling of stable transgenic plants for protein extraction

Leaf tissue samples were collected for each stable transgenic plant using a cork borer. Samples were collected when the plant had 8 fully expanded leaves. In total, 8 leaf discs of 5 mm in diameter were collected per plant, each from a different leaf. The discs

were placed in a 2 ml Eppendorf tube containing three ceramic beads of 2.3 mm and stored at -80ºC for later use.

Tissue samples from stable transgenic plants infiltrated with an *A. tumefaciens* culture expressing p19 were collected for each transient transformant following the previous description. In total 4 leaf discs, 5 mm in diameter were collected from infiltrated and non-infiltrated tissue. The discs were placed in a 2 ml Eppendorf tube containing three ceramic beads of 2.3 mm and stored at -80ºC for later use.

2.5.2 Tissue sampling of stable transgenic plants for confocal microscopy

Tissue samples were collected from fully expanded and developed leaves using a cork borer. In total 3 leaf discs, 4 mm in diameter, were collected per plant, each from a different size leaf (one from a young leaf, one from a medium size leaf and one from an old leaf). The discs were placed over a drop of milliQ water previously put on a coverslip and covered with another drop of milliQ water before laying another coverslip over the sample. The samples were imaged by confocal microscopy immediately.

2.6 Protein extraction

Leaf tissue samples to be processed were homogenized frozen for two times 30 second pulses using a Tissue Lyser (Qiagen). Then 600 µl of protein extraction buffer were added to each tube and the tubes were vortexed for 15 seconds to extract the total soluble proteins (TSP). The samples were centrifuged at 14,000*g* for 5 minutes at 4ºC and the supernatant was transferred to a new tube. The previous step was repeated and the supernatant was transferred again to a new tube. The extracted proteins were kept on ice

and used for quantitation of GFP by dot blot analysis. An aliquot of the extract from each sample was kept on ice and used the same day for determination of TSP concentration using a Bradford assay (Bio-Rad, Bradford, 1976). A bovine serum albumin (BSA) protein standard curve was used to quantify the amount of TSP present in each extract. The standard curve was prepared with seven different concentrations: 1.0 µg, 2.0 µg, 4.0 µg, 6.0 µg, 8.0 µg, 10 µg and 12 µg BSA.

For *N. benthamiana* plants, infiltrated with GFP-HFBI, 6.77 g of fresh leaf tissue were collected 4 days post infiltration (dpi). The tissue was frozen using liquid nitrogen and manually homogenized by grinding until a fine powder was obtained. Then protein extraction buffer (3 ml per gram of fresh tissue) was added and the sample was mixed manually. The sample was centrifuged at 20,000*g* for 30 minutes at 4ºC and the supernatant was transferred into a new tube. The previous step was repeated and the extracted protein was kept on ice and immediately used for purifying GFP-HFBI by the aqueous two-phase system (ATPS) method (Linder *et al.*, 2004).

2.7 Aqueous two phase system

The GFP-HFBI recombinant protein was purified by aqueous two phase system (ATPS) (Linder *et al.*, 2004) as follows: The total soluble protein extracted from leaf tissue of transiently transformed *N. benthamiana* plants was pre-warmed at 24ºC for 10 minutes and mixed with the surfactant Agrimul to a final concentration of 8% (w/v). The sample was mixed by vortexing and rotated gently on an ORBITRON rotator (Boekel Scientific, Model 260250) at room temperature for 10 minutes. The tube was then incubated at 24ºC for 1 hour and subsequently centrifuged at 3000*g* for 5 minutes at

24ºC. The upper clear phase, containing the GFP-HFBI recombinant protein was then transferred to a new tube and isobutanol (400 µl isobutanol per 1 ml upper phase) was added. The sample was gently mixed by inversion and rotated on an ORBITRON rotator at room temperature for 10 minutes. After this, the tube was centrifuged at 3000*g* for 5 minutes at 24ºC. Three layers were observed; the upper and middle phases were discarded and the bottom phase containing the GFP-HFBI recombinant protein was transferred to a new tube. Two aliquots of 80 µl each were mixed with 5X reducing sample buffer and stored at -80ºC prior to sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) analysis. The remaining sample was divided in aliquots of 10 µl to be used as GFP standard in dot blot analyses.

2.8 Protein quantification

Quantification of the recombinant proteins was performed using the immunodot blot technique. Briefly, each extracted sample was serially diluted three times using 1X PBS. Later, 2 µl of each undiluted sample and from the three serial dilutions were spotted in triplicate onto a nitrocellulose membrane in an 8 x 12 format (Bio-Rad). Five different concentrations of purified GFP-HFBI were also spotted in triplicate: 4 ng, 8 ng, 12 ng, 16 ng and 20 ng. After loading the standards and the samples, the membranes were blocked overnight at 4ºC with 1X TBS-T containing 3% (w/v) powdered skim milk. The membranes were then moved to hybridization tubes and incubated for 1 hour at room temperature with 1 µl of mouse monoclonal anti GFP primary antibody (Living Colors® A.v. monoclonal antibody (JL-8), Clontech) diluted 1/5000 in 1X TBS-T containing 3% (w/v) powdered skim milk. The membranes were then washed three times for 15 minutes with 1X TBS-T and then incubated for an hour at room temperature with a

horseradish peroxidase (HRP) - conjugated goat-anti mouse IgG secondary antibody (Bio-Rad) diluted in 1/3000 in 1X TBS-T containing 3% (w/v) powdered skim milk. Finally, the membranes were washed three times for 10 minutes with 1X TBS-T containing 3% (w/v) powdered skim milk and visualized using the enhanced chemiluminescence (ECL) detection system (GE Healthcare, Mississauga, Canada) according to the manufacturer's protocol, and scanned using a ChemidocXRS (Bio-Rad). Scanned membranes were analyzed using TotalLab TL 100 software (Nonlinear Dynamics, Durham, USA). The GFP standard was used to generate a standard curve which was used to extrapolate the amount of recombinant protein in each sample.

Quantification of GFP-HFBI, extracted by ATPS, was performed by SDS-PAGE and Coomassie staining. Briefly, an aliquot of GFP-HFBI was serially diluted using 1X PBS. Undiluted GFP-HFBI and six dilutions were separated on a 10% polyacrylamide gel and visualized by GelCode^{TM} blue stain reagent (Thermo Scientific) staining. The gel was scanned and analyzed by Totallab TL100 Software Array V 2008. A BSA protein standard curve (0.5 µg, 0.15 µg, 1.0 µg, 2.0 µg, 3.0 µg and 4.0 µg BSA) was used as reference to quantify the amount of GFP-HFBI present.

2.9 Confocal microscopy

A Leica TCS SP2 confocal laser scanning inverted microscope (Leica Microsystems, Wetzlar, Germany) equipped with a 63X water immersion objective was used to examine the subcellular localization of GFP in the different tissue samples and seeds of stable transgenic plants of *N. tabacum*. To visualize GFP fluorescence, excitation with a 488 nm argon laser was used and its emission was detected at

500-530 nm. Chlorophyll autofluorescence was also detected with emission at 630-680 nm. Collected images were analysed using the Leica Application Suite for Advanced Fluorescence (LAS AF) (Leica Microsystem, V2.3.5 build 5379).

3 RESULTS

3.1 Generation of stable transgenic *N. tabacum* plants

Previous studies using transient expression assays have demonstrated that ELP and HFBI fusion tags possess the ability to enhance the accumulation levels of recombinant proteins when targeted to the ER (Conley *et al*., 2011b; Joensuu *et al*., 2010; Conley *et al*., 2009a). Furthermore, it was shown that they induce the formation of PBs, which seems to be the key mechanism to increase accumulation levels of recombinant proteins in plants (Conley *et al*., 2011b; Joensuu *et al*., 2010).

Although transient expression systems are fast and effective, generating stable transgenic plants still remains one of the long-term goals for producing recombinant proteins in plants. Therefore, the first objective of this project was to generate stable transgenic plants of *N. tabacum* expressing ER-targeted GFP, GFP-HFBI and GFP-ELP recombinant proteins to evaluate the effect of ELP and HFBI tags on the accumulation levels of GFP.

To achieve this, plant expression vectors previously constructed to produce and target the three recombinant proteins to the ER (Figure 3) were used for *Agrobacterium*-mediated transformation of leaf discs. Cultivars I64 and 81V9 of *N. tabacum* were transformed and twenty four independent transgenic plants for each combination of construct and cultivar were regenerated (Table 1). All transgenic plants showed a normal phenotype compared with nontransgenic plants.

3.2 Quantification of recombinant proteins in stable transgenic lines of *N. tabacum*

Once the different transgenic lines were obtained, protein extraction and quantification of recombinant proteins were carried out for each of the twenty four independent transformants to estimate accumulation levels of the different recombinant proteins in each transgenic plant. This was done to compare GFP accumulation levels in cultivars I64 and 81V9, both transformed with GFP-HFBI, and between the ELP and HFBI fusion constructs, both transformed into cultivar I64.

Before starting to analyze protein levels of the transgenic plants, a GFP-HFBI standard protein was generated. For this, *Agrobacterium*-mediated transient expression was used to produce GFP-HFBI recombinant protein in *N. benthamiana* plants. The recombinant protein was purified from infiltrated leaf tissue using ATPS and quantified using SDS-PAGE (Figure 4). Standards of 4, 8, 12, 16, 20 and 24 ng of GFP-HFBI per µl were prepared and used to create a standard curve for extrapolating concentrations of GFP in transgenic plants.

Protein extractions of transgenic plants were quantified and recombinant protein levels were determined using immunodot blotting and a GFP antibody (Figure 5). Aliquots for each original extract were saved and used for total soluble protein (TSP) quantification.

Transgenic plants of cultivars I64 expressing GFP were first analyzed. These plants had accumulation levels ranging between 0.01% and 1.2% of recombinant protein per TSP (Table 2). In general, most plants had low levels of accumulation (under 0.1% of

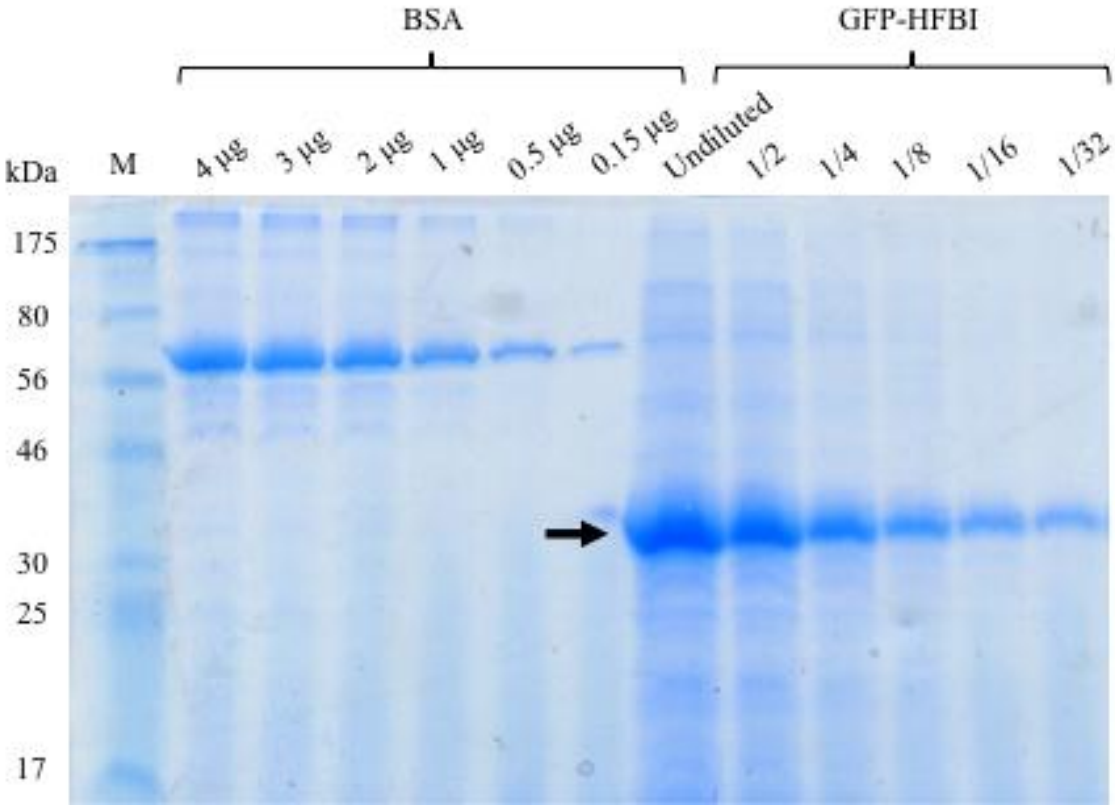


Figure 4. Quantification of purified GFP-HFBI by SDS-PAGE analysis.

GFP-HFBI extracted by ATPS was serially diluted (1/2, 1/4, 1/8, 1/16 and 1/32), separated on a 10% SDS-PAGE, and subsequently stained with GelCode^{TM} blue stain reagent. Using the BSA 1 µg standard and 1/16 dilution of ATPS extract, the concentration of purified GFP-HFBI was calculated to be 1.6 µg/µl. Arrow, GFP-HFBI (32.7 kDa). M, protein marker.

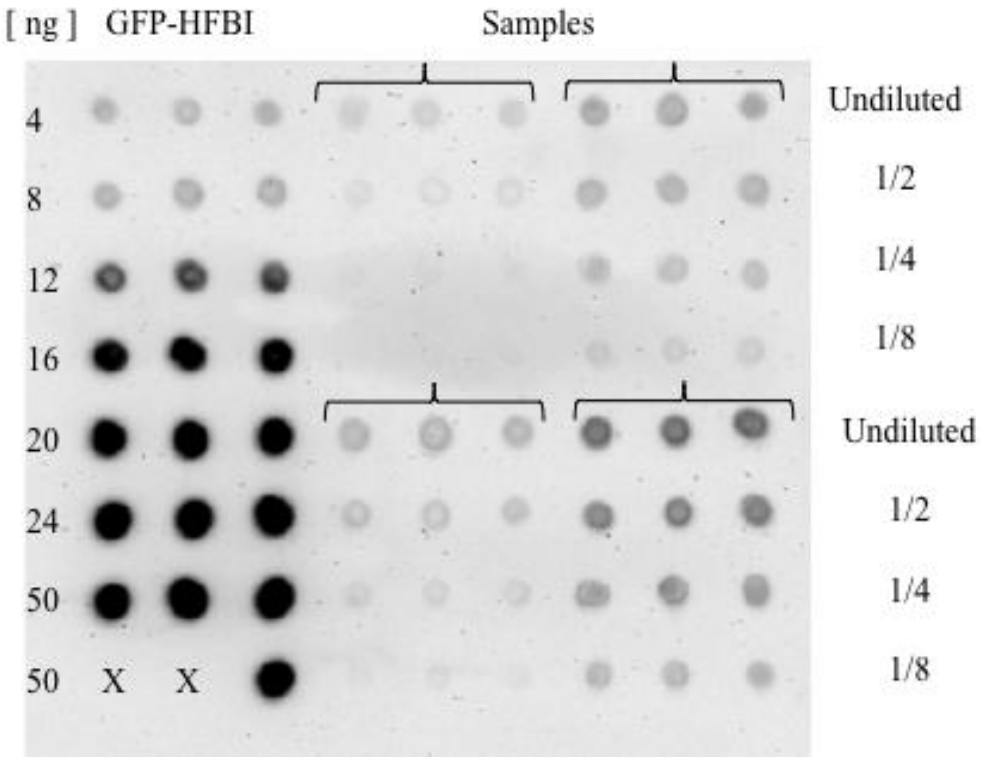


Figure 5. Quantification of recombinant proteins by immunodot blotting.

Immunodot blot was performed with GFP antibody. Left, standard of purified GFP-HFBI. Each concentration was loaded in triplicate. Right, leaf tissue samples. After protein extraction, each sample was serially diluted (1/2, 1/4 and 1/8) and 2 µl of each dilution were loaded in triplicate. Results for four different samples are shown.

Table 2. Accumulation levels of ER-targeted GFP protein in *N. tabacum* cv. I64 stable transgenic plants.

Twenty four independent transformants were evaluated. The quantity of recombinant protein present in each sample was calculated as the mean of three technical replicates for a particular dilution.

Plant	**Recombinant protein [% per TSP]**	**Standard deviation [x 10^{-4}]**[a]	**PBs**[b]
1	0.014	±0.35	A
2	0.018	±1.13	A
3	**0.022**	**±8.23**	**A**
4	0.031	±0.35	ND
5	0.031	±0.09	A
6	0.032	±1.04	A
7	0.041	±1.45	ND
8	0.049	±0.56	A
9	0.054	±0.56	A
10	0.060	±1.31	A
11	0.074	±0.99	ND
12	0.090	±1.49	ND
13	0.106	±11.2	ND
14	0.112	±11.2	A
15	0.227	±4.16	A

Table 2. continued. Accumulation levels of ER-targeted GFP protein in *N. tabacum* cv. I64 stable transgenic plants.

16	0.237	±13.9	P
17	0.321	±2.85	P
18	**0.354**	**±7.07**	**P**
19	0.457	±17.5	ND
20	0.460	±9.68	A
21	0.631	±8.73	ND
22	0.810	±15.1	ND
23	1.142	±14.3	ND
24	1.201	±3.18	P

a, Standard deviation of the mean value of three technical replicates.

b, Presence (P) or absence (A) of PBs in stable transgenic lines determined by confocal microscopy analysis. ND, not determined.

Dotted line separates transgenic lines without PBs from those with PBs.

Bolded rows correspond to the plants shown in pictures from Results section.

recombinant protein per TSP).

Analysis of the protein quantification results obtained from stable transgenic plants of cultivar I64 expressing GFP-ELP revealed that these plants had accumulation levels ranging between 0.02% and 2.1% of recombinant protein per TSP (Table 3). Compared to the plants expressing GFP, most of the plants had higher levels of recombinant protein accumulation (Figure 6).

Most transgenic plants of cultivar I64 and 81V9 expressing GFP-HFBI had accumulation levels ranging between 0.01% and 2.5% of recombinant protein per TSP (Table 4 and 5). Only a single transgenic plant of the cultivar 81V9 had accumulation levels of 5.2% of recombinant protein per TSP. Overall however, no obvious differences were found between these two cultivars (Figure 6).

When comparing GFP-ELP transgenic lines to GFP-HFBI lines, all in cultivar I64, the distribution of accumulation levels was overall higher in lines transformed with GFP-HFBI.

A comparison between transgenic plants expressing GFP and transgenic lines expressing GFP-ELP and GFP-HFBI, showed that transgenic plants expressing GFP presented the lowest levels of accumulation of the recombinant protein (Figure 6). Therefore, these finding suggests that the fusion tags increase the accumulation levels of the recombinant protein in transgenic plants.

3.3 Protein body formation in stable transgenic plants of *N. tabacum*

To begin the evaluation of the sub-cellular localization of ER-targeted GFP, GFP-HFBI

Table 3. Accumulation levels of ER-targeted GFP-ELP protein in *N. tabacum* cv. I64 stable transgenic plants.

Twenty four independent transformants were evaluated. The quantity of recombinant protein present in each sample was calculated as the mean of three technical replicates in dot blot for a particular dilution.

Plant	**Recombinant protein [% per TSP]**	**Standard deviation [x 10^{-4}]**[a]	**PBs**[b]
1	0.020	±19.7	A
2	**0.031**	**±1.95**	**A**
3	0.043	±16	A
4	0.074	±11.8	A
5	0.085	±2.48	A
6	0.094	±19.7	ND
7	0.106	±0.1	ND
8	0.108	±59.3	A
9	0.126	±6.06	A
10	0.193	±0.42	ND
11	0.210	±0.12	P
12	**0.224**	**±2.19**	**P**
13	0.236	±8.43	P
14	**0.303**	**±0.52**	**P**
15	0.325	±0.17	P

Table 3. continued. Accumulation levels of ER-targeted GFP-ELP protein in *N. tabacum* cv. I64 stable transgenic plants.

16	0.404	±1.12	P
17	0.599	±53.5	P
18	0.715	±0.35	ND
19	0.813	±0.08	P
20	0.877	±1.31	ND
21	1.085	±11.2	A
22	1.449	±9.38	ND
23	1.722	±52.6	ND
24	2.100	±25.9	P

a, Standard deviation of the mean value of three technical replicates.

b, Presence (P) or absence (A) of PBs in stable transgenic lines determined by confocal microscopy analysis. ND, not determined.

Dotted line separates transgenic lines without PBs from those with PBs.

Bolded rows correspond to the plants shown in pictures from Results section.

Figure 6. Distribution *N. tabacum* transgenic plants expressing GFP, GFP-ELP and GFP-HFBI recombinant proteins.

Light green bars, plants without PBs with accumulation levels between 0.01 - 0.2 % of recombinant protein per TSP. Dark green bars, plants with PBs with accumulation levels between 0.2 - 5.2% of recombinant protein per TSP. Gray bars, plants where presence or absence of PBs were not determined. Each column represents the mean value of three technical replicates, and error bars indicate the standard deviation of the mean. Dotted line separates plants without and with PBs and shows that plants had protein accumulation levels of at least 0.2% of recombinant protein per TSP before PBs could be visualized. Arrows identifying transgenic plants which were analyzed by confocal microscopy.

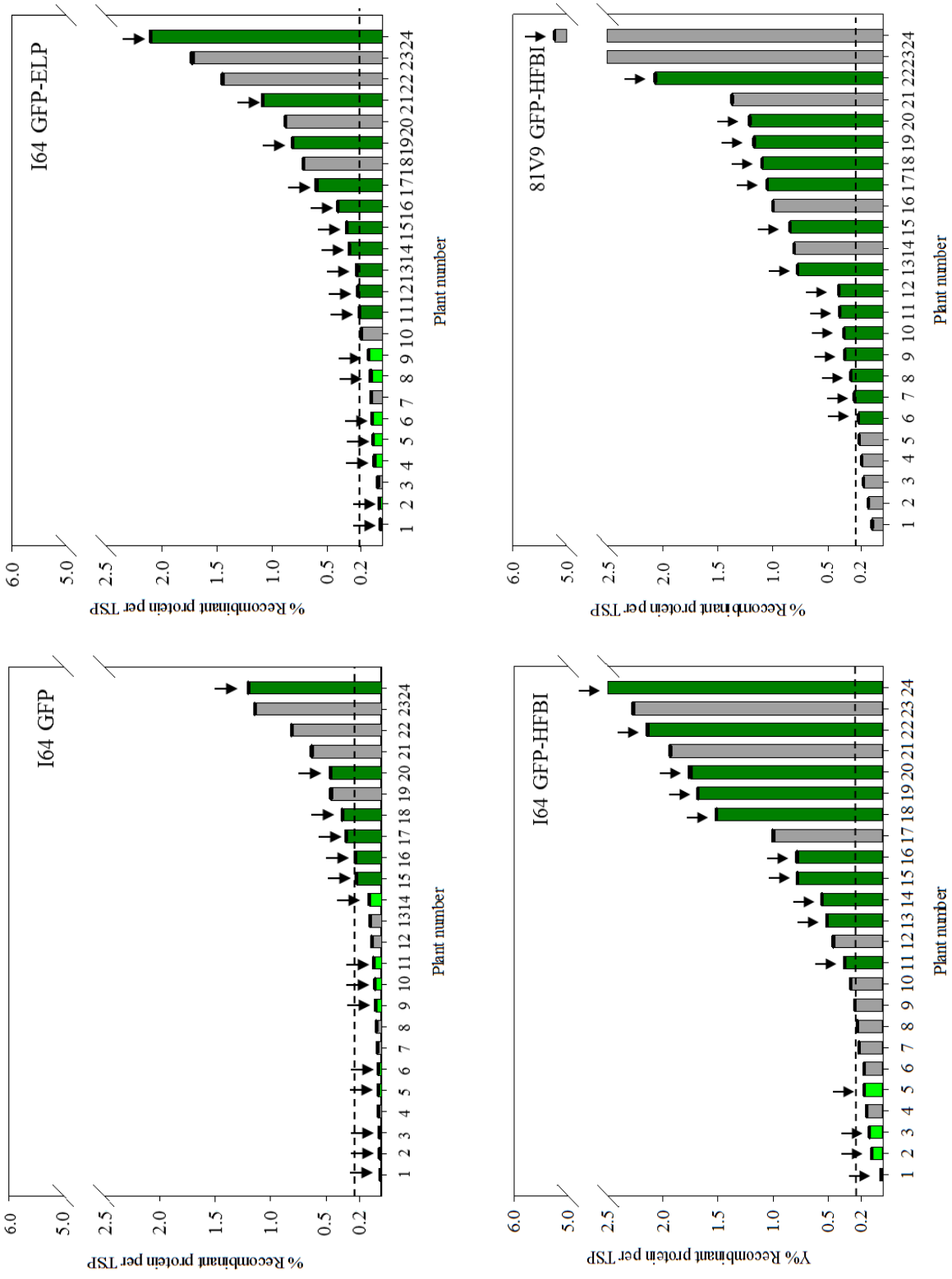

I64 GFP-ELP
81V9 GFP-HFBI
I64 GFP
I64 GFP-HFBI
% Recombinant protein per TSP
Plant number

Table 4. Accumulation levels of ER-targeted GFP-HFBI protein in *N. tabacum* cv. I64 stable transgenic plants.

Twenty four independent transformants were evaluated. The quantity of recombinant protein present in each sample was calculated as the mean of three technical replicates in dot blot for a particular dilution.

Plant	**Recombinant protein [% per TSP]**	**Standard deviation [x 10^{-4}]**[a]	**PBs**[b]
1	0.018	±1.75	A
2	**0.100**	**±5.41**	**A**
3	0.122	±2.81	A
4	0.149	±3.43	ND
5	0.169	±7.57	A
6	0.172	±2.11	ND
7	0.219	±5.59	ND
8	0.232	±4.81	ND
9	0.255	±2.72	ND
10	0.294	±1.70	ND
11	0.349	±0.17	P
12	0.451	±9.65	ND
13	0.510	±11.7	P
14	**0.553**	**±6.15**	**P**

Table 4. continued. Accumulation levels of ER-targeted GFP-HFBI protein in *N. tabacum* cv. I64 stable transgenic plants.

15	0.779	±0.13	P
16	0.782	±9.36	P
17	0.999	±0.45	ND
18	1.514	±0.16	P
19	**1.684**	**±1.13**	**P**
20	1.754	±0.95	P
21	1.935	±7.99	ND
22	2.141	±0.53	P
23	2.271	±0.23	ND
24	2.521	±0.15	P

[a], Standard deviation of the mean value of three technical replicates.

[b], Presence (P) or absence (A) of PBs in stable transgenic lines determined by confocal microscopy analysis. ND, not determined.

Dotted line separates transgenic lines without PBs from those with PBs.

Bolded rows correspond to the plants shown in pictures and Table 6 from Results section.

Table 5. Accumulation levels of ER-targeted GFP-HFBI protein in *N. tabacum* cv. 81V9 stable transgenic plants.

Twenty four independent transformants were evaluated. The quantity of recombinant protein present in each sample was calculated as the mean of three technical replicates in dot blot for a particular dilution.

Plant	**Recombinant protein** [% **per TSP**]	**Standard deviation** [**x** 10^{-4}]**[a]**	**PBs[b]**
1	0.102	±3.42	ND
2	0.136	±0.03	ND
3	0.179	±0.14	ND
4	0.196	±0.03	ND
5	0.218	±4.38	ND
6	0.225	±1.46	P
7	0.263	±3.08	P
8	0.294	±4.88	P
9	**0.351**	**±2.90**	**P**
10	0.357	±6.13	P
11	**0.397**	**±5.52**	**P**
12	0.404	±6.70	P
13	**0.776**	**±17.1**	**P**
14	0.808	±12.9	ND

Table 5. continued. Accumulation levels of ER-targeted GFP-HFBI protein in *N. tabacum* cv. 81V9 stable transgenic plants.

15	0.846	±1.13	P
16	1.0	±6.82	ND
17	1.051	±11.4	P
18	**1.098**	**±11.5**	**ND**
19	**1.171**	**±15.6**	**P**
20	1.212	±29.8	A
21	1.371	±11.3	ND
22	2.066	±13.3	P
23	2.507	±23.3	ND
24	5.226	±12.7	P

[a], Standard deviation of the mean value of three technical replicates.

[b], Presence (P) or absence (A) of PBs in stable transgenic lines determined by confocal microscopy analysis. ND, not determined.

Dotted line separates transgenic lines without PBs from those with PBs.

Bolded rows correspond to the plants shown in pictures and Table 6 from Results section.

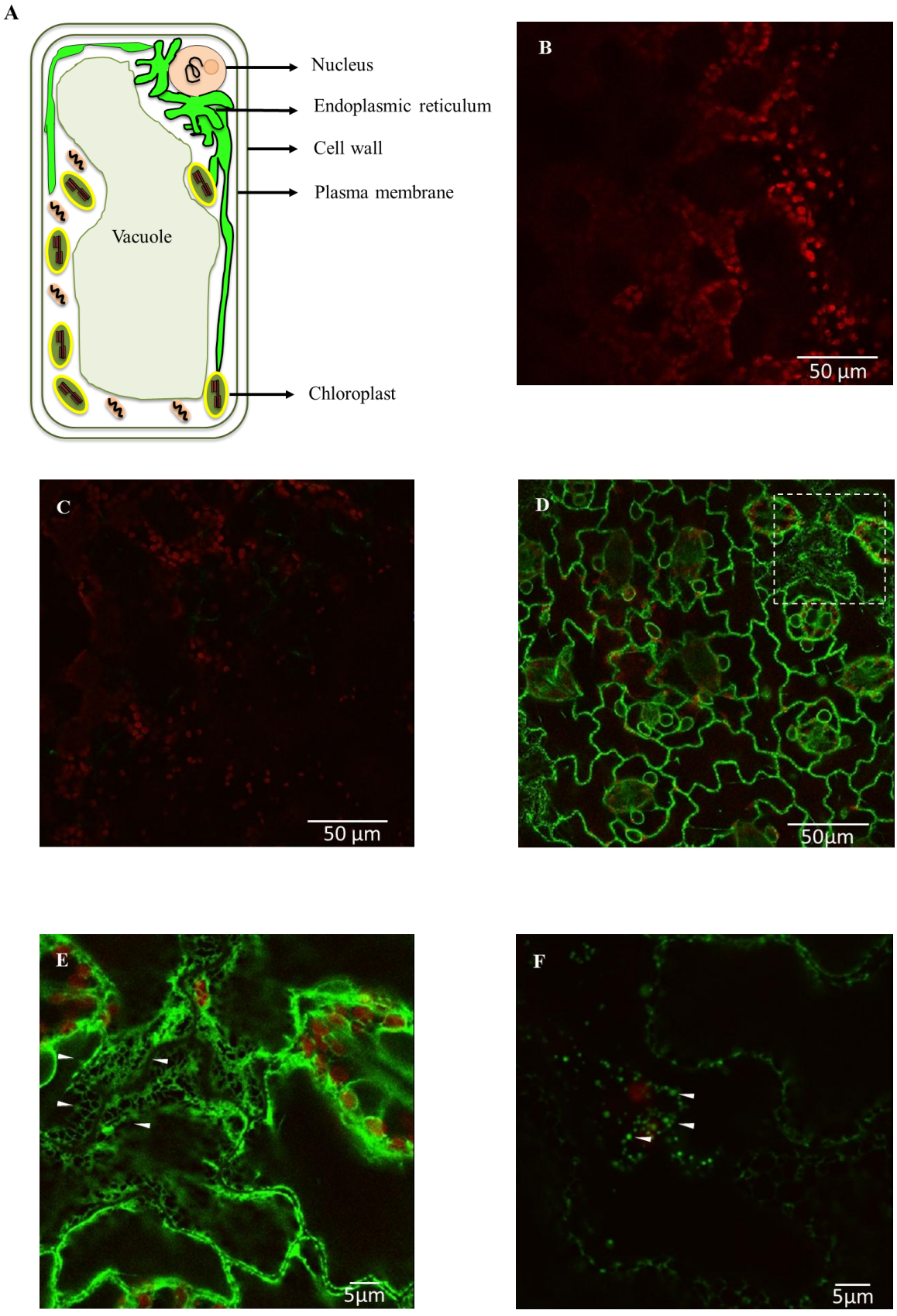
and GFP-ELP recombinant proteins, confocal laser scanning microscopic analyses were carried out on some of the stable transgenic plants. In total, three leaf samples per plant were analyzed including one young leaf, one medium sized leaf and one old leaf. The plants that were analyzed are identified in Figure 6 by an arrow, but only select representative images are shown in this thesis.

When analyzing transgenic plants of cultivar I64 expressing GFP (fifteen plants), fluorescence was detected (Figure 7A, 7C and 7D), indicating that GFP was being expressed. However, fluorescence levels varied among plants, ranging from very low fluorescence (Figure 7C) to high levels of fluorescence (Figure 7D). The accumulation levels visualized by confocal microscopy corresponded to those determined by protein quantification (Table 2). Therefore, in plants with low GFP accumulation levels (below 0.1%) a low fluorescent signal was detected and leaf cells were barely visible. As untransformed plants showed no evidence for fluorescence, the observed fluorescence in GFP transgenic lines can be attributed to the expression of the GFP construct (Figure 7B).

Analysis of plants with higher accumulation levels of GFP showed a strong fluorescence signal that allowed the visualization of cells and a fluorescent reticulated pattern which is characteristic of the ER network (Figure 7E), indicating expected localization of the recombinant protein. Although GFP (without tag) was expected to be localized to the ER, surprisingly, some plants showed formation of very small spherical particles with a diameter of 0.2 – 0.5 µm, with very few larger particles that were up to 2 µm in size, which could be PBs (Figure 7F). No differences were observed between young, medium and old leaves in a particular plant.

Figure 7. Leaf tissues from stable transgenic *N. tabacum* I64 plants expressing ER-targeted GFP.

A. Overview of a simplified plant cell cross-section expressing ER-targeted GFP protein. Plant cells possess a large central vacuole that pushes the cytosol and organelles to the periphery of the cell, creating a visible GFP outline around it. ER showed in light green. B. Leaf cells from a wild type cultivar I64 *N. tabacum* plant. No fluorescence was observed. C. Low levels of fluorescence in leaf cells from transgenic plant number 3 (Table 2, 0.022% recombinant protein per TSP) of cultivar I64 expressing GFP targeted to the ER. D. Fluorescent leaf cells from transgenic plant number 18 (Table 2, 0.354% recombinant protein per TSP). E. Close-up of the area in the white panel in picture D showing in detail the reticulated pattern characteristic of the ER. Arrowhead, ER reticulated pattern. F. View of small spherical particles found in some leaf cells of plant number 18. Arrowhead, PBs. All pictures were acquired with the same gain settings. Red autofluorescence of chloroplasts is observed in all pictures.


A
Nucleus
Endoplasmic reticulum
Cell wall
Plasma membrane
Vacuole
Chloroplast
B
50 µm
C
50 µm
D
50µm
E
5µm
F
5µm

Analysis of seventeen transgenic plants of cultivar I64 expressing GFP-ELP showed a very different fluorescence pattern compared to plants expressing GFP. Leaf cells contained PBs which varied in size and numbers between the analyzed plants (Figure 8A and 8B). These PBs ranged from 0.5 – 1.0 µm in diameter to large bright spheres of the same size or larger than chloroplasts (2.0 – 5.0 µm diameter) (Figure 8C and 8D). In these plants, the simultaneous visualization of PBs and ER was difficult due to the brightness of the large PBs. No obvious differences were observed between leaves of different ages. As for the GFP transgenic lines, plants with different levels of fluorescence were identified and again fluorescence levels and the presence and size of PBs levels correlated well with accumulation levels of the recombinant protein (Table 3, Figure 8).

Analysis of twenty eight transgenic plants expressing GFP-HFBI from cultivars 81V9 and I64 (fourteen plants for each transformed cultivar), showed similar results to the ones obtained for plants expressing GFP-ELP. Fluorescent cells with abundant PBs were observed in different transgenic plants (Figure 9A and 9B). As in the other transgenic lines, no obvious differences were observed between the different leaf ages. Nevertheless, the size and frequency of these PBs were also variable between transgenic plants. Most plants had PBs with a diameter of 0.5 – 2.0 µm (Figure 9B). Plants with higher levels of accumulation also had larger PBs ranging from 2.0 – 3.0 µm (Figure 9). No obvious differences were found between the two transformed cultivars.

A comparison between the frequency and sizes of the observed PBs in plants with different accumulations levels of GFP-HFBI further supported the idea that the characteristics of the HFBI tag were very similar in the two evaluated cultivars (Table 6).

Figure 8. Leaf tissues from stable transgenic *N. tabacum* I64 plants expressing GFP-ELP show numerous protein bodies.

A. Leaf cells from cultivar I64 transgenic plant 14 (Table 3, 0.303% recombinant protein per TSP) expressing GFP-ELP recombinant protein. Formation of protein bodies is evident. Arrowhead, PBs. B. Close-up of the area in the white rectangle in picture A showing protein bodies in detail. C, D. Observed PBs with a diameter between 0.5 µm – 5 µm. Due to the size and brightness of the PBs, the ER is only faintly visible. Arrowhead, ER reticulated pattern. Arrow, small PBs. E. Leaf cells from cultivar I64 transgenic plant 12 (Table 3, 0.224% recombinant protein per TSP) expressing GFP-ELP recombinant protein show formation of PBs. Arrowhead, small PBs. F. Close-up of picture E showing protein bodies in detail. PBs heterogeneous in size are observed. Arrowhead, ER reticulated pattern. Arrow, PBs. All pictures were acquired with the same gain settings

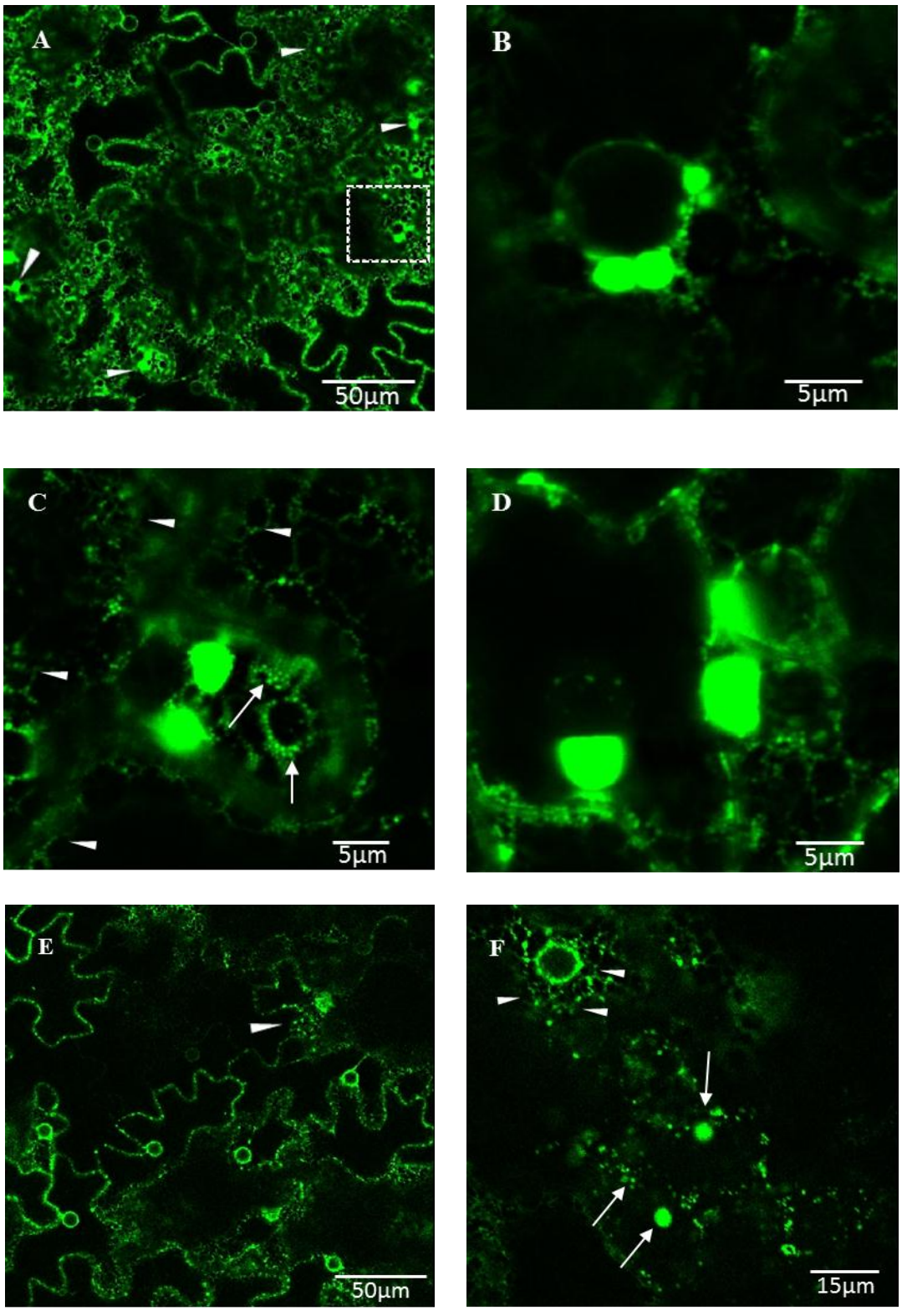

A
50µm
B
5µm
C
5µm
D
5µm
E
50µm
F
15µm

Figure 9. Leaf tissues from a stable transgenic *N. tabacum* 81V9 plant expressing GFP-HFBI induces the formation of protein bodies.

A. Leaf cells from cultivar 81V9 transgenic plant 11 (Table 5, 0.397% recombinant protein per TSP) expressing GFP-HFBI recombinant protein. Formation of protein bodies is evident. B. Close-up of the area in the white rectangle in picture A allowing to see protein bodies. PBs had an average size between 0.5 - 2.0 µm. C. Leaf cells from cultivar 81V9 transgenic plant 19 (Table 5, 1.171% recombinant protein per TSP) expressing GFP-HFBI recombinant protein. Formation of protein bodies is evident. D. Close-up of picture C allowing to visualize protein bodies. Clusters of PBs homogeneous in size are observed. Due to the size and brightness of the PBs, the ER is only faintly visible. E. Close-up of the area in the white rectangle in picture D allowing to visualize PBs in more detail. Arrowhead, PBs. All pictures were acquired with the same gain settings.

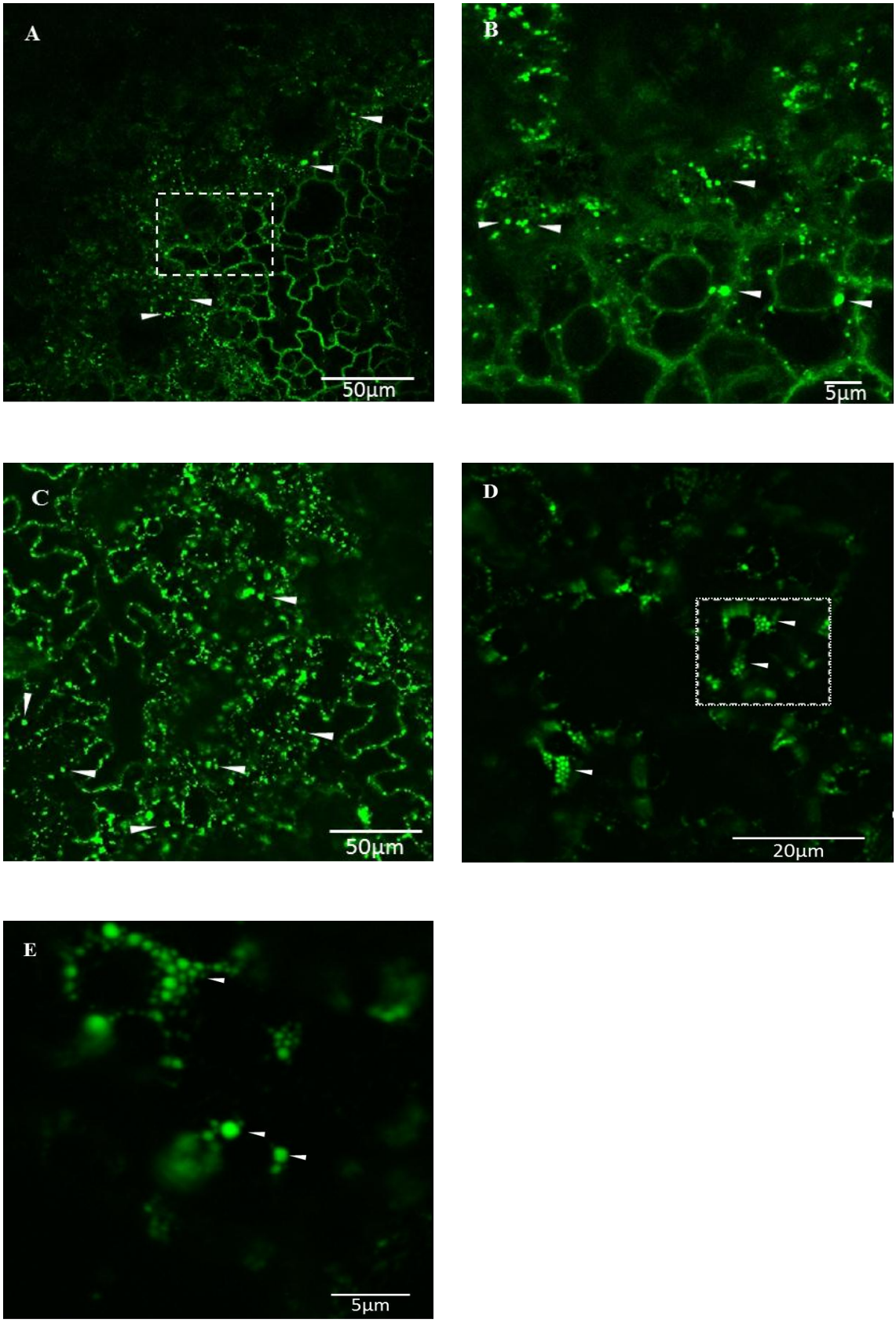

A
50µm
B
5µm
C
50µm
D
20µm
E
5µm

Table 6. Frequency and size of PBs present in leaf cells of *N. tabacum* cv. 81V9 and cv. I64 stable transgenic plants with different levels of accumulation of GFP-HFBI.

Analyzed plants[a]	Protein bodies		Protein body size	
	Absence[b]	Presence	Small[c]	Large[d]
I64 GFP-HFBI - 0.553% recombinant protein per TSP (plant 14).	0	100	97%	3%
I64 GFP-HFBI - 1.684% recombinant protein per TSP (plant 19).	0	100	37%	63%
81V9 GFP-HFBI – 0.776% recombinant protein per TSP (plant 13)	0	100	87%	13%
81V9 GFP-HFBI - 1.171% recombinant protein per TSP (plant 19).	0	100	48%	52%

[a], In total, five leaves per plant were sampled. Two leaf discs per leaf were collected and analysed by laser scanning confocal microscopy. Ten cells were analysed per leaf disc, for a total of one hundred cells.

[b], Cells with a bright fluorescent ER reticulated pattern.

[c], PB diameter <1.0 µm.

[d], PB diameter >1.0 µm.

PBs were present in 100% of the cells in each plant. However, large PBs (>1.0 µm) were more frequent in cells of plants with higher accumulation levels of the recombinant protein.

When comparing confocal results obtained for transgenic plants expressing GFP-HFBI and GFP-ELP, PBs in plants expressing GFP-HFBI were smaller and more homogenous in size (0.5 – 3.0 µm) (Figure 9) than the ones observed in plants expressing GFP-ELP (0.5 – 5.0 µm) (Figure 8). Furthermore, most of the analyzed stable transgenic plants that formed PBs regardless of the cultivar and the construct had high numbers of PBs in guard cells of stomata. Sometimes, having more PBs than their surrounding epidermal cells (Figure 10).

A comparison of the confocal laser scanning microscopy results and the obtained accumulation levels for the different analyzed stable transgenic lines revealed a relationship between PB formation and recombinant protein accumulation. Essentially, higher accumulation levels of recombinant protein were associated with the presence of PBs (Figure 6, Tables 2-5).

Based on these findings, all transgenic lines were divided into two different classes according to the accumulation levels and presence of PBs. The first class represents plants with low accumulation levels, ranging between 0.01% and 0.2% of recombinant protein per TSP, and absence of PBs. These plants showed low fluorescence levels where cells were barely visible or only the outline and ER of the cells was visible. The second class represents plants with higher protein accumulation levels, starting at 0.2% of recombinant protein per TSP, and presence of PBs. Plants in this class showed bright and

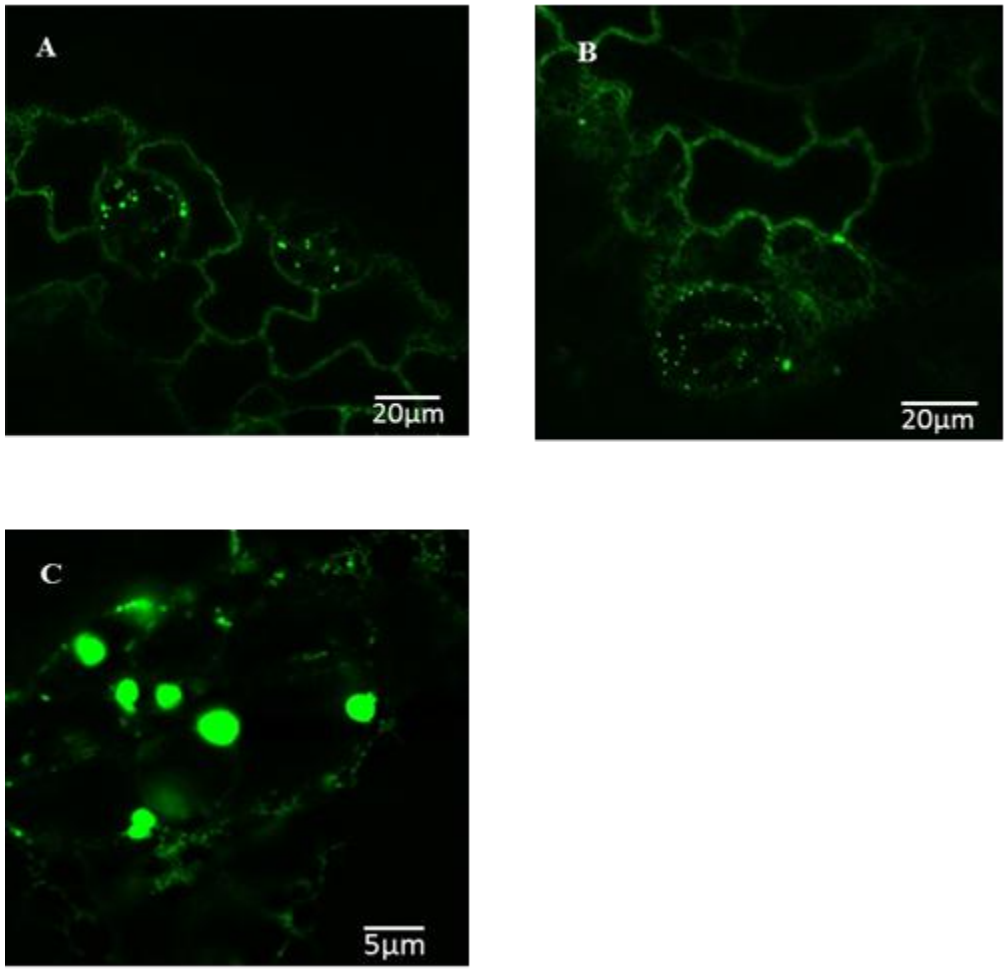


Figure 10. Guard cells from leaf tissue of a stable transgenic plant of *N. tabacum* expressing GFP-HFBI contain protein bodies.

A, B and C. Protein bodies in guard cells from cultivar 81V9 transgenic plant 18 (Table 5, 0.354% recombinant protein per TSP) expressing GFP-HFBI. All pictures were acquired with the same gain settings.

and clearly visible cells. Furthermore, the characteristic reticulated pattern of the ER was visible and PBs were formed in almost every cell. Furthermore, PB frequency and size increased with accumulation levels (Table 6). According to the analyses, 0.2% of recombinant protein per TSP would be the threshold level needed for the formation of PBs.

Based on the two different classes, 58.33% of the transgenic plants of cultivar I64 expressing GFP had low accumulation levels of the recombinant protein (Table 2, Figure 6). 41.66% of plants of cultivar I64 expressing GFP-ELP had low accumulation levels (Table 3, Figure 6), whereas from transgenic lines from cultivars I64 and 81V9 expressing GFP-HFBI, 25% and 16.6% respectively, had low accumulation levels (Table 3 and 4, Figure 6).

A comparison between the distributions of the transgenic lines in the two classes demonstrated that the plants expressing GFP had the lowest accumulation levels and most transgenic plants were part of the low accumulation class, followed by transgenic plants expressing GFP-ELP. Transgenic plants of cultivars I64 and 81V9 expressing GFP-HBFI had the highest numbers of plants outside the lower level accumulation category.

In general, these findings demonstrate that the two tags increase the accumulation levels of the recombinant protein and are associated with the formation of PBs in the different transgenic plants. However, PBs can be also formed in plants without the use of these tags if a certain level of accumulation of the recombinant protein is reached.

3.4 Effect of a viral suppressor of gene silencing on accumulation levels of recombinant protein and PB formation in stable transgenic plants

In all the stable transgenic lines generated in this study various transgenic plants had low accumulation levels and in some plants no fluorescence was detected. The presence of plants with such low accumulation levels is frequently explained by positional effects of transgene insertion in the genome or silencing mechanisms of the plant (Schubert *et al*., 2004). The latter is usually attributed to post-transcriptional gene silencing (PTGS), a ribonucleic acid (RNA) based silencing mechanism that can be activated by plant pathogens, transposons and transgenes (Fagard and Vaucheret, 2000).

Plant pathogenic viruses have evolved strong PTGS suppressors that can act at different levels of the silencing pathway. One of the best characterized suppressors, expressed by members of the *Tombusvirus* family, is the p19 protein. p19 has been successfully used in transient expression assays as a strategy to increasing the yield of several recombinant proteins (Havelda *et al*., 2003; Fischer *et al.*, 2004).

To evaluate if GFP-HFBI transgenic plants had low accumulation levels due to an active PTGS mechanism, transgenic plants with low accumulation levels (plants 2 and 3 in Table 4, plant 2 in Table 5) were transiently transformed with an *Agrobacterium* strain expressing the p19 gene from CymRSV (Sylhavy *et al.*, 2002).

Analysis of recombinant protein accumulation of infiltrated and non-infiltrated leaf tissue four days after infiltration showed that in two transgenic plants (plants 2 and 3 in Table 4) there was an increase in the accumulation levels of the recombinant protein in the infiltrated areas of the leaf (Figure 11A, 11B and 11C). Furthermore, confocal

microscopy analysis of infiltrated leaf tissue samples revealed an increase in the level of fluorescence and showed the formation of PBs in one plant (Figure 11D and 11E). Therefore, this data is consistent with the idea that in certain cases, low accumulation levels of the recombinant protein in stable transgenic plants may be attributed to PTGS. Assays in plant 3 (Table 4) are not shown but accumulation levels without p19 had an average of 0.12% of recombinant protein per TSP (standard deviation of the mean 1.00×10^{-1}) and with p19 had an average of 0.24% of recombinant protein per TSP (standard deviation of the mean 2.75×10^{-4}).

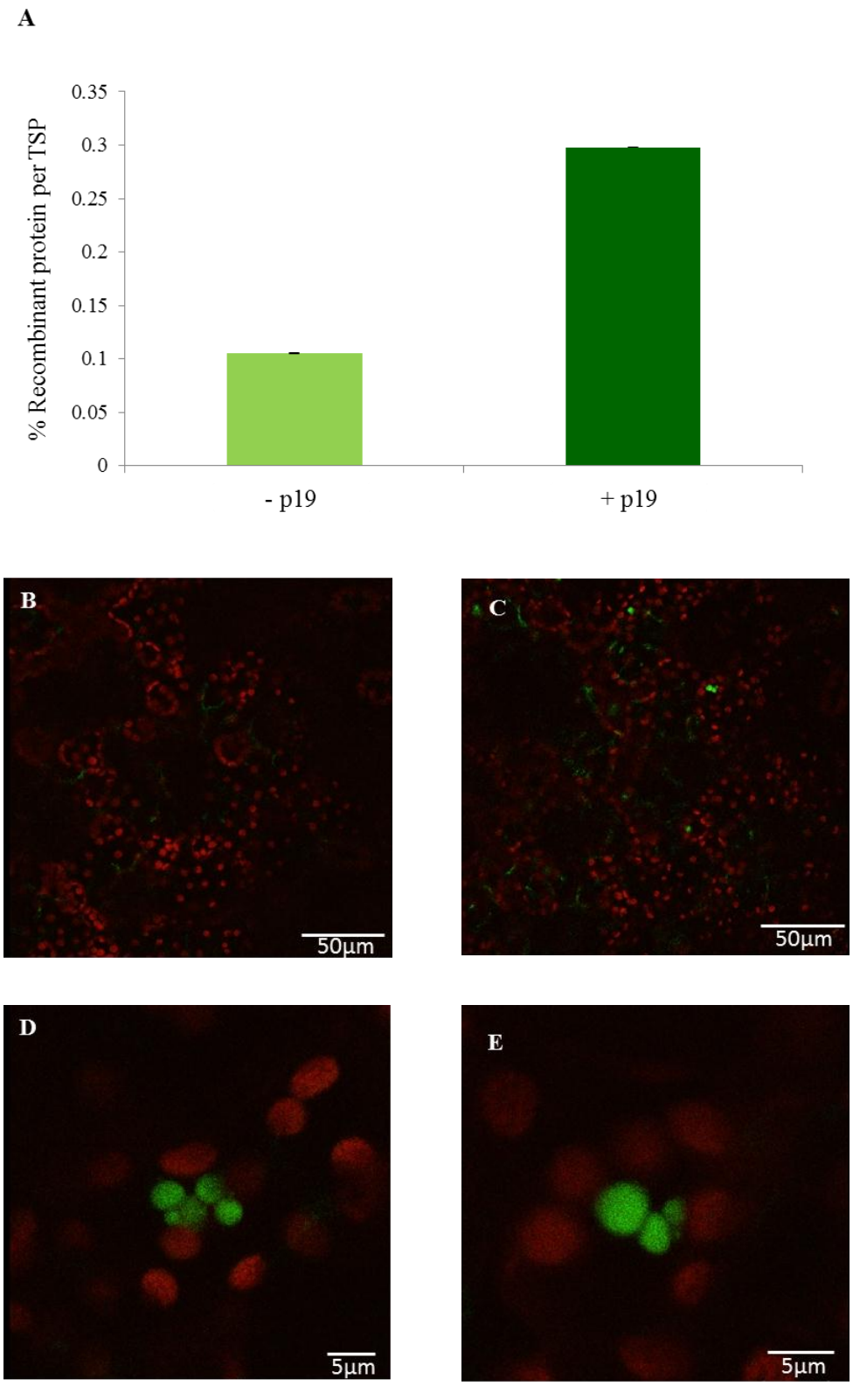
Stable transgenic plants with higher accumulation levels of GFP-HFBI were also transiently transformed with p19 (plant 11 in Table 4 and plants 9 and 11 in Table 5). Comparison of the recombinant protein accumulation of infiltrated and non-infiltrated leaf tissue did not reveal an increase in the accumulation levels of the recombinant protein in infiltrated areas of the leaf (data not shown). However, confocal microscopy analysis of infiltrated leaf tissue showed an increase in frequency and size of PBs (Figure 12). Interestingly, this increase was mostly notable in the guard cells of the stomata, where bright and large PBs were found. This finding is interesting because it implies that more p19 find its way to the guard cells than to other epidermal cells. It is important to note that *Agrobacterium* penetrates the epidermal cells trough stomata. This should be documented in the literature.

3.5 Analysis of GFP expression in seeds

Once the expression of the different recombinant proteins was confirmed in transgenic plants and formation of PBs was observed, I wondered if PBs were also

Figure 11. Effect of p19 silencing suppressor over a stable transgenic plant of cv. I64 with low levels of accumulation of GFP-HFBI.

A. Accumulation levels of recombinant protein with and without the p19. The experiment was performed with three different plants but only results of one assay are shown (Table 4, Plant 2, 0.1% of recombinant protein per TSP). Each column represents the average of two technical replicates. Error bars represent the standard deviation of the mean. B. View of leaf cells without p19. C. View of leaf cells with p19. An increase in GFP fluorescence levels was observed. D, E. formation PB in leaf cells with p19. All pictures were acquired with the same gain settings. Red autofluorescence of chloroplasts is observed in all pictures.


A
% Recombinant protein per TSP
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
- p19
+ p19
B
50µm
C
50µm
D
5µm
E
5µm

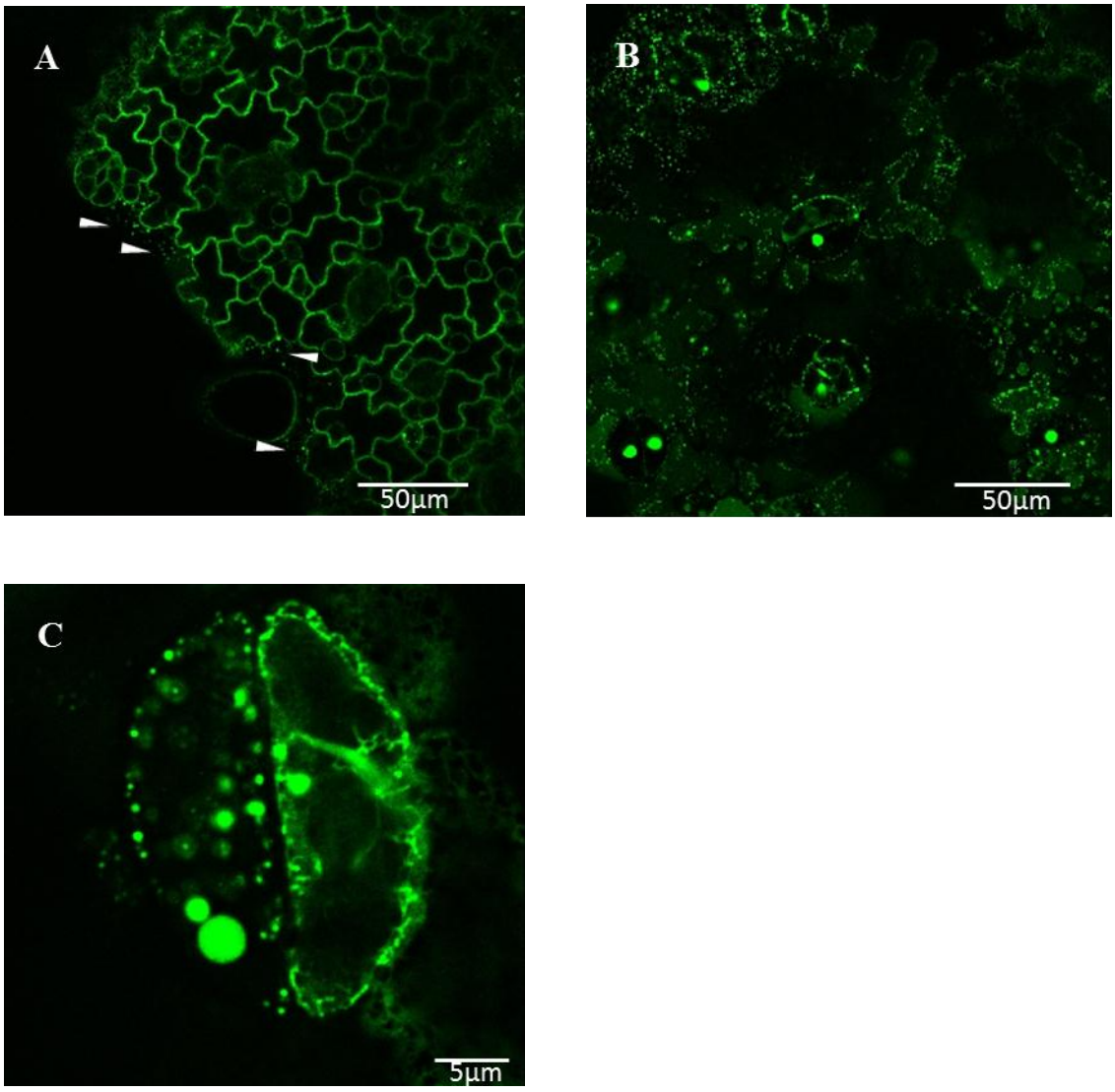


Figure 12. Effect of the p19 silencing suppressor in a stable transgenic plant of cv. 81V9 with higher accumulation levels of GFP-HFBI.

A.View of leaf cells from cultivar 81V9 transgenic plant 9 (Table 5, 0.351% of recombinant protein per TSP) without p19. PBs are present in the cells. Arrowhead, small PBs B. View of leaf cells with p19. The observed PBs are larger, compared with PBs in picture A. C. Close up to protein bodies present in guard cells. All pictures were acquired with the same gain settings.

induced in tobacco seeds because in dicot plants proteins are accumulated in PSVs (Peters and Stöger, 2011). Therefore, seeds from *N. tabacum* wild type plants of cultivars 81V9 and I64, and various of the transgenic plants expressing GFP, GFP-HFBI (seeds from plant 19, Table 5 and plant 11, Table 4) and GFP-ELP (seeds from plant 11, Table 3) were analyzed by laser scanning confocal microscopy.

Confocal microscopy analysis of seeds from wild type plants of cultivars I64 and 81V9 did not show any fluorescence, demonstrating that none of the organelles present in the seeds autofluoresce under the established confocal parameters (data not shown).

Confocal microscopy analysis of the seeds from transgenic plants expressing GFP recombinant protein showed bright fluorescent endosperm, where individual cells could be identified (Figure 13B). A close-up view of some of these cells showed the presence of numerous round organelles of different sizes. According to their shape and size, they might be PSVs or oil bodies, which are common in tobacco seeds (Peters and Stöger, 2011; Kawakatsu *et al*., 2010; Tzen *et al*., 1993) (Figure 13B and 13C).

Confocal microscopy analysis of seeds from transgenic plants expressing GFP-HFBI (Figure 14A) and GFP-ELP (data not shown) also showed a bright fluorescent endosperm, where individual cells were clearly visible. A close-up view of some of these cells not only showed the presence of numerous round black organelles, surrounded by a fluorescent signal but also the presence of some small bright fluorescent particles that resemble PBs (Figure 14B).

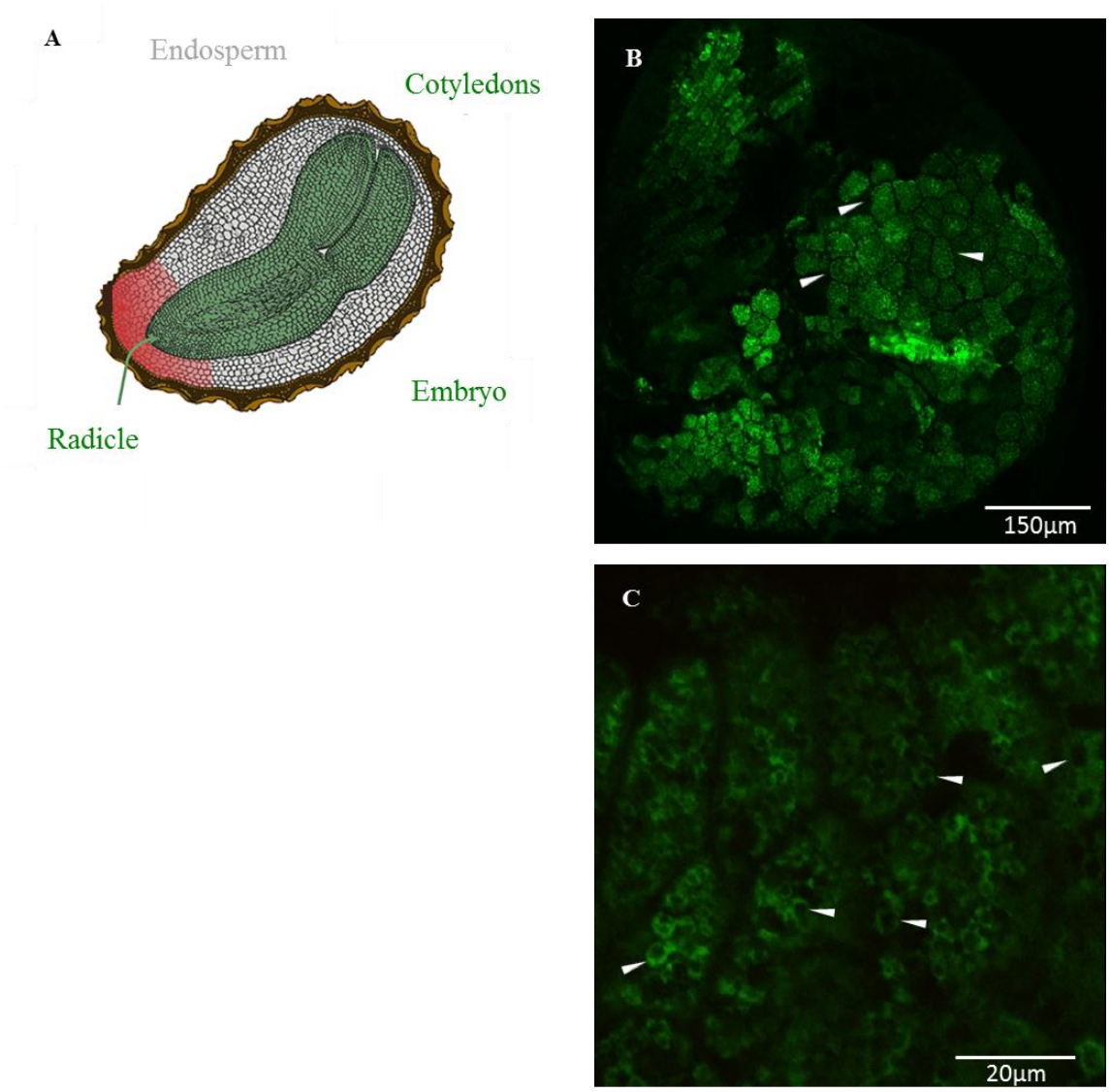


Figure 13. Analysis of the seeds of *N. tabacum* cv. I64 stable transgenic plant expressing GFP.

A. Overview of a simplified mature tobacco seed cross section showing the main parts of the seed. Modified from 'The Seed Biology Place' (http://www.seedbiology.de). B. View of seed. Arrowhead, endosperm cells. C. Close-up of picture B. Presence of round organelles. Arrowhead, round organelles. All pictures were acquired with the same gain settings.

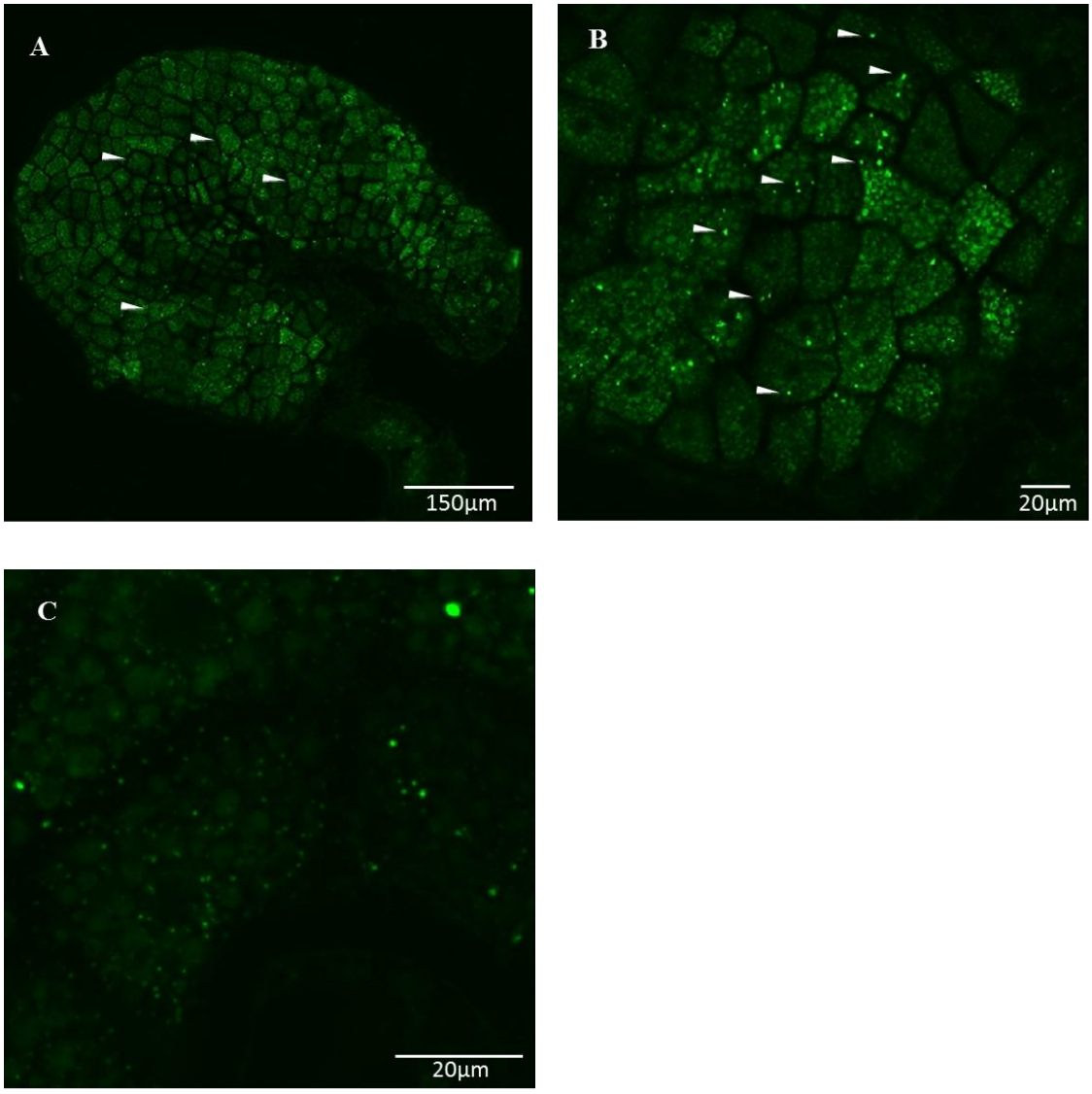


Figure 14. Analysis of the seeds of *N. tabacum* cv. 81V9 stable transgenic plant expressing GFP-HBFI.

A. View of seed. Arrowhead, endosperm cells. (seeds from plant 19, Table 5) B. Close-up of picture A. Presence of spherical particles that resemble PBs. Arrowhead, fluorescent particles. C. Close up of picture B showing in more detail the fluorescent spherical particles. All pictures were acquired with the same gain settings.

4 DISCUSSION

4.1 Elastin-like polypeptide and hydrophobin fusions improve accumulation levels of GFP in stable transgenic plants of *N. tabacum*

The use of fusion partners for the expression of recombinant proteins in plants has been used as a strategy to solve two major problems: low accumulation levels and lack of efficient purification methods for plant made proteins (Conley *et al.*, 2011b). The ELP and HFBI fusion tags have been tested in transient expression experiments, and shown to increase recombinant protein accumulation levels (Joensuu *et al*., 2010; Conley *et al.*, 2009b). ELP was shown to effectively concentrate the recombinant proteins, although a second affinity step is needed for purifying the protein (Conley *et al*., 2009a). On the other hand, HFBI was shown to recover 90% of the protein in one ATPS step (Joensuu *et al*., 2010). ELP was shown to increase the accumulation of recombinant proteins in stable transgenic plants but no PBs were found (Patel, *et al*., 2007), while HFBI had not been tested in a stable plant system and neither were shown to produce PBs in stable transgenic plants.

Analysis of the different stable transgenic plants of *N. tabacum* demonstrated that ELP and HFBI tags can be efficiently used in a constitutive environment. The use of any of the fusion tags at least doubled the amount of the recombinant protein, compared with the expression of the protein alone (Figure 6, Tables 2 - 5). These findings are important for different reasons.

Importantly, this is the first time that the HFBI tag has been used in a stable plant transgenic transformation. Consequently, this is the first report showing that this tag is

functional in stable transgenic plants, where it positively impacts the accumulation of the protein of interest. Furthermore, ATPS can be effectively used for the purification of the recombinant protein from a crude plant extract, which is a simple and scalable purification system (Linder *et al*., 2004). This technique has been used in this study for the purification of GFP-HFBI to produce and purify the standard protein used for quantification of the GFP accumulation in the transgenic plants by immunodot blot technique (Figure 4).

Additionally, no differences were found between cultivars I64 and 81V9 transformed with the HFBI tag (Figure 6). This finding opens the possibility of using different cultivars with different properties that can be beneficial for the expression of a particular protein. I64 produces high biomass, while 81V9 cultivar is a low alkaloid cultivar (Conley *et al*., 2011a), a property that can be exploited for the expression of proteins that can be orally administered to humans and animals, such us vaccines. This is a very important aspect because if expensive purification processes can be avoided, lower costs of the product can be achieved (Joensuu *et al*., 2008; Menassa *et al*., 2007). Furthermore, these findings also agree with a previous study where the accumulation levels of four different recombinant proteins showed no differences between 16 different tested cultivars of *Nicotiana* plants (Conley *et al.*, 2011a).

In regard to the stable transgenic plants of cultivar I64 expressing GFP-ELP, this tag increased the accumulation levels of GFP (Figure 6). However, these plants had in general lower accumulation levels, compared with the GFP-HFBI plants. Therefore, it seems that the ELP tag may be less efficient in this cultivar (Table 3 and 4, Figure 6). In order to evaluate if this result is due to an interaction between the tag and cultivar I64

plants or due to an arbitrary accumulation of the proteins, it will be important to evaluate a larger group of stable transgenic plants using this cultivar. The use of a different, non-fluorescent fusion partner would be also recommended because properties of the fused protein may affect the behavior of the tag.

A comparison between the results of this study and previous studies where GFP-ELP was transiently expressed in *N. benthamiana* and *N. tabacum* plants without p19 co-expression, revealed that the accumulation levels in some of the stable transgenic plants were higher (0.35% GFP per TSP reported value; Conley *et al.*, 2009a,b). This suggests that transient expression assays cannot predict accumulation levels in stable transgenic plants although they can predict the functionality of a particular construct in plants and allow the comparison of different constructs (Conley *et al.*, 2011a; Conley *et al.*, 2009a; Fischer *et al.*, 2004). However it is important to remember that these analyses were performed in primary transformants. Therefore, the observed distribution of recombinant protein accumulation in the different stable transgenic lines (Figure 6) reflect the typical trend of a transgenic population, where the variability in accumulation levels across the transgenic lines is expected and can be explained by positional effects of the transgene insertion within the genome, number of copies of the transgene and/or silencing mechanisms that can be activated in the plant to suppress the expression of the protein of interest (Stam *et al*., 1997; Finnegan and Mc Elroy, 1994).

Despite this, the results of this study demonstrate that it is possible to generate stable transgenic plants of *N. tabacum* expressing recombinant proteins with these two tags, which is one of the long term goals in plant biotechnology for different reasons. First, the need for labor-intensive agro-infiltration for commercial large-scale production

can be avoided. Second, *N. tabacum* is usually easier to grow and it produces higher biomass compared to *N. benthamiana,* which is often used for transient transformations. Therefore, these characteristics and the fact that in this study plants with high accumulation levels of the recombinant proteins were found, indicates that it is possible to obtain stable transgenic plants with high accumulation levels of the protein of interest. For example, compared to previous reports of accumulation levels found in tobacco stable transgenic plants expressing different types of cellulases (Endo-1,4-β-D-glucanase, 0.1% per TSP and cellobiohydrolase, 0.002% per TSP; Ziegelhoffer et al. 1999), a recombinant human erythropoietin (EPO), human interleukin-10 (IL-10) and a synthetic antibody against *Pseudomonas aeruginosa* (APA) (0.02% EPO per TSP, 0.16% IL-10 per TSP and 2.5% APA per TSP; Conley *et al.*, 2011a) demonstrate that the levels found in this study are promising.

However, screening a larger group of primary transformants is recommended because this will increase the possibility of finding plants with higher levels of accumulation. Furthermore, it is also important to analyze the progeny of the selected primary transformants to determine copy number of the transgene and establish if the expression and accumulation levels of the protein of interest are stable and maintained in these lines.

4.2 Elastin-like polypeptide and hydrophobin fusions induce the formation of protein bodies in stable transgenic plants of *N. tabacum*

The ER is the first gateway of the protein secretory pathway. In plant seeds, the endomembrane system is able to generate multiple ER-derived compartments or protein

bodies, primarily used to store reserve compounds, such as storage proteins, lipids, carbohydrates, minerals, etc. Consequently, they have an appropriate environment for folding, assembling, and long-term storage of massive amounts of proteins (Galili, 2004; Vitale and Ceriotti, 2004).

Confocal laser scanning microscopy analyses of GFP, GFP-ELP and GFP-HFBI stable transgenic plants showed that ELP and HFBI tags, when targeted to the ER, promote the formation of multiple spherical particles in the leaves of the mentioned plants. A comparison between the confocal microscopy images and images of PBs induced by the transient expression of different fusion proteins with ELP and HFBI tags (Joensuu *et al.*, 2010; Conley *et al.*, 2009b; Torrent *et al.*, 2009a), indicates that the spherical structures found in this study may be PBs. Their size and morphology are clearly similar. However, PBs generated in this study should be characterized in more detail by EM and immunogold localization of ER luminal and membrane proteins to determine their ER origin.

Furthermore, HFBI and ELP tags share specific physicochemical properties with maize seed storage proteins (Conley *et al.*, 2011b), hydrophobicity being one of the most important as it is the property that is predicted to allow for their self-assembly and formation of PBs, which bud from the ER (Joensuu *et al.*, 2010).

Analysis of the transgenic plants expressing GFP, GFP-ELP and GFP-HFBI recombinant proteins revealed a new important finding. A comparison between the accumulation levels of the recombinant protein and formation of PBs indicates that there is a close relationship between these two variables. Specifically, it appears that there is a

minimum level of accumulation of recombinant protein needed for the formation of these PBs in *N. tabacum* plants. According to my results, 0.2% of recombinant protein per TSP is the minimum required amount. To my knowledge, this is the first report that demonstrates that there is a minimum threshold value needed for the formation of PBs in leaves.

PBs were found in plants expressing GFP-ELP. However, compared to the PBs found in plants expressing GFP-HFBI, these PBs were usually larger and more heterogeneous in size (0.5 – 5 μm), whereas in plants expressing GFP-HFBI they were smaller and more homogeneous (0.5 – 3 μm). The reason for this difference is not known, however it may be influenced by the fusion partner and properties of each tag. Therefore, further characterization of how the PBs are induced and formed by each tag in these plants is necessary. Immunogold localization of ER proteins, EM and proteomic analysis should provide information about their protein composition and the proteins involved in PB formation.

With regards to the transgenic plants of cultivar I64 expressing GFP, these plants had the lowest levels of accumulation of the recombinant protein and the largest number of plants in the low level class. However, despite these results, PBs were still found in some transgenic plants. This is an important finding as it suggests that the ELP and HFBI tags are not essential for the formation of PBs, although they increase accumulation levels and therefore enhance the size and numbers of PBs. This result agrees with and confirms my previous proposal that if a recombinant protein reaches an accumulation level of 0.2% of recombinant protein per TSP or higher, PBs will be formed.

Previous studies of transient expression assays in *N. benthamiana* plants where GFP-ELP and GFP-HFBI were co-expressed with the p19 suppressor of gene silencing, showed the formation of PBs in agro-infiltrated leaves (Joensuu *et al*., 2010 ; Conley *et al*., 2009b). Although the achieved levels of these recombinant proteins were much higher than levels reported here, their results are similar to mine to some extent. Their findings, similar to this study, suggest that both of these tags have a positive effect on GFP accumulation levels. Furthermore, they showed formation of PBs with both of these tags and with the GFP protein alone in agro-infiltrated leaves, suggesting that these tags enhance PB formation, but that they are not an essential factor for PB formation (Joensuu *et al*., 2010; Conley *et al*., 2009a).

No visible differences were found in size and number of PBs between the three different stages of leaf development, suggesting that the accumulation levels of the recombinant protein were not affected by the stage of leaf development. This result is potentially important due to the fact that harvesting of the recombinant product can be performed at once, regardless of the age or size of the leaves. However, to further demonstrate this, it will be important to process samples from different stages of development and compare their accumulation levels of the recombinant protein.

Overall, my results demonstrate that the use of both of these tags as fusion partners is a good alternative for enhancing the accumulation levels of recombinant proteins in stable transgenic plants of *N. tabacum*. In this study, a plant with 5.2% of GFP per TSP was obtained, a significant value for stable transgenic plants. Moreover, these fusions induce formation of PBs in leaves, a phenomenon that happens naturally only in seeds and has been considered as key for long-term storage of proteins (Torrent *et al*., 2009a).

Likely, these tags increase the accumulation levels of the recombinant proteins thus enhancing the formation of PBs, where the proteins are protected from environmental factors that can be detrimental for them, such as degradation enzymes and pH differences, etc (Conley *et al.*, 2011b). Furthermore, each tag provides a simple and inexpensive purification method that has been successfully used previously to purify different types of recombinant proteins (Joensuu *et al*., 2010; 2009; Floss *et al*., 2008; Lin *et al.*, 2006; Scheller *et al*., 2004).

It is important to mention that being able to use *N. tabacum* for a leaf-based system with these tags is an advantage because the proteins will be recovered from the leaves, and these plants produce a high biomass. Moreover, the use of a leaf-based system reduces the risk of dispersion of the transgenic product to the environment (Rymerson *et al*., 2002; Sheen, 1983). However, it should be considered that the behavior of these tags may change according to their fused protein. Therefore, it is recommended to test the recombinant protein of interest in transient expression assays and then screen a large group of stable transgenic plants.

4.3 Expression of the p19 PTGS suppressor increases accumulation levels of GFP-HFBI in low-expressing stable transgenic plants of *N. tabacum*

The observed variation in accumulation levels of the recombinant proteins in different stable transgenic plants was expected. This is because the generation of stable transgenic plants using *Agrobacterium*-mediated transformation can result in the integration of the transgene in unfavorable or favorable locations in the chromosome and sometimes more than one copy of the transgene are inserted. These factors will influence

expression levels of the protein of interest (Stam *et al*., 1997). However, it is also known that this variation can be induced by other factors, such us the activation of plant PTGS mechanism. PTGS is a homology-dependent RNA degradation system that was evolved by plants to protect themselves from invasive or mobile foreign genetic elements, including transposons, transgenes and viruses (Havelda *et al*., 2003). Therefore, if expression levels of the transgene are high and an overproduction of RNA occurs, this mechanism will silence expression of the transgene (Wroblewski *et al*., 2005; Rybicki, 2009).

Results from the agro-infiltration of p19, a known PTGS suppressor, in three GFP-HFBI plants with low accumulation levels of the recombinant protein demonstrated that some of these plants might have been silenced. This was shown by the observed increase in the accumulation of the recombinant protein and the fluorescence levels before and after the infiltration. Furthermore, in one of the agro-infiltrated plants formation of PBs was induced. This finding is expected, because all of the constructs used in this study have one of the strongest constitutive promoters for plants, the double enhanced 35S promoter from Cauliflower Mosaic virus (CaMV) (Kay *et al*., 1987), increasing the possibility of the activation of PTGS in transgenic plants.

In the case of the agro-infiltrated plants with higher accumulation levels of the recombinant protein no difference in accumulation levels was observed, but an increase in number and size of the PBs was evident. An explanation for the outcome of this assay is not clear. More inoculations in plants with different accumulation levels may generate more information.

4.4 Analysis of seeds from stable transgenic plants reveals the presence of small fluorescent particles

Confocal microscopy analysis of the seeds from stable transgenic plants expressing GFP showed a bright fluorescent endosperm, where individual cells containing fluorescent organelles, possibly PSVs or oil bodies, were clearly visible. This result is very interesting since CaMV 35S is a constitutive promoter that seems to be regulated developmentally and tissue-specifically, and usually is not active in seeds (Sunilkumar *et al.*, 2002). Previous studies where expression of ER-targeted GFP under the control of the CaMV 35S promoter was followed during germination of stable transgenic seeds of *N. tabacum* (Hraska *et al*., 2008) and cotton (Gossypium *hirsutum* L.) (Sunilkumar *et al.*, 2002) are examples of this.

In tobacco, very low levels of GFP expression were present in the dry mature seed, and strong green fluorescence appeared after 24h of germination (Hraska *et al*., 2008). In cotton, no fluorescence was detected during early seed embryo developmental stages, indicating a complete lack of promoter expression or very low levels of expression (Sunilkumar *et al.*, 2002). Furthermore, hygromycin phosphotransferase (*HPT*) gene expression under the control of CaMV 35S in transgenic rice was not detected in developing seeds (Wasaka *et al*., 2009).

In this study, confocal microscopy analysis of seeds from stable transgenic plants expressing recombinant proteins with or without the ELP and HFBI tags showed a bright fluorescent endosperm. Some individual cells showed small spherical fluorescent particles that resemble protein bodies with both fusion tags. This is interesting since PBs

are characteristics of cereal seeds. It was expected that in tobacco seeds, GFP-HFBI recombinant protein would be deposited in PSVs or secreted (Peters and Stöger, 2011; Abranches *et al*., 2008).

To define the identity of these particles, further analyses have to be conducted. Including the use of different stains to identify and discriminate between compounds that are usually stored in the seeds like lipids, proteins, starch, etc. Also, the use of immunolabelling to specific membrane proteins could be informative to identify the different sub-cellular compartments naturally present in seeds. Furthermore, it is of our particular interest to estimate the accumulation levels of the recombinant proteins in these seeds, to determine if the defined threshold for the PB formation in leaves applies also for their formation in seeds.

5 CONCLUSIONS AND FUTURE STUDIES

According to my work, the use of ELP and HFBI tags as fusions partners in stable transgenic plants of *N. tabacum* is feasible and promising. My results demonstrate that in a constitutive environment these tags increase the accumulation levels of the recombinant protein and induce the formation of PBs, a key mechanism for safe and efficient storage of recombinant proteins. Furthermore, the accumulation levels reported in this study demonstrate that stable transgenic plants have the potential for the commercial production of recombinant proteins. However, it is important to evaluate larger groups of transgenic lines because this increases the possibility of finding higher expressing lines. Moreover, my results were similar in two different cultivars, which further support my conclusions and suggest that both of these tags can be used for transformation of different tobacco cultivars.

On the other hand, I found and established a threshold value for the accumulation of the recombinant protein that has to be reached in the plant in order to induce the formation of PBs. Future studies should be conducted in order to evaluate if the threshold value found in these transgenic lines applies to other fusion proteins and different expression platforms.

Additionally, I showed that the use of a PTGS suppressor can increase the levels of accumulation of the recombinant protein and induce formation of PBs, however more experiments have to be done to have a better sense of this phenomenon. Finally, it was very interesting to find spherical fluorescent particles that resemble cereal seed PBs. Nevertheless, further experiments have to be done to establish their identity and formation process in tobacco seeds.

6 REFERENCES

Abranches, R., Arcalis, E., Marcel, S., Altmann, F., Ribeiro-Pedro, M., Rodriguez, J., Stöger, E (2008) Functional specialization of *Medicago truncatula* leaves and seeds does not affect the subcellular localization of a recombinant protein. *Planta* **227**: 649-658.

Arcalis, E., Marcel, S., Altmann, F., Kolarich, D., Drakakaki, G., Fischer, R., Christou, P., Stöger, E (2004) Unexpected deposition patterns of recombinant proteins in post-endoplasmic reticulum compartments of wheat endosperm. *Plant Phys* **136**: 3457-3466.

Argos, P., Pedersen, K., Marks, M.D., Larkins, B. A (1982) A structural model for maize zein proteins. *J Biol Chem* **257**: 9984-9990.

Bevan, M., Barnes, W.M., Chilton, MD (1983) Structure and transcription of the nopaline synthase gene region of T-DNA. *Nucleic Acids Res* **11**: 369–385.

Bradford, M.M (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of dye-binding. *Anal Biochem* **72**: 248-254.

Christensen, T., Amiram, M., Dagher, S., Trabbic-Carlson, K., Shamji, M.F., Setton, L.A., Chilkoti, A (2009) Fusion order controls expression level and activity of elastin-like polypeptide fusion proteins. *Protein Sci* **16**: 1377-1387.

Coleman, C.E., Herman, E.M., Takasaki, K., Larkins, B.A (1996) The maize gamma-zein sequesters alpha-zein and stabilizes its accumulation in protein bodies of transgenic tobacco endosperm. *Plant Cell* **8**: 2335-2345.

Commandeur, U., Twyman, R.M, Fischer, R (2003) The biosafety of molecular farming in plants. *AgBiotechNet*. **5**: 1-9.

Conley, A.J., Joensuu, J.J., Jevnikar, A.M., Menassa, R., Brandle, J.E (2009a) Optimization of elastin-like polypeptides and purification of recombinant proteins in plants. *Biotechnol Bioeng* **103**: 562-573.

Conley, A.J., Joensuu, J.J., Menassa, R., Brandle, J.E (2009b) Induction of protein body formation in plant leaves by elastin-like polypeptide fusions. *BMC Biol* **7**: 1-18.

Conley, A.J., Zhu, H., C.L., Jevnikar, A.M., Lee, B.H., Brandle, J.E., Menassa, R (2011a) Recombinant protein production in a variety of *Nicotiana* hosts: a comparative analysis. *Plant Biotechnol J* **9**: 434-444.

Conley, A.J., Joensuu, J.J, Richman, A., Menassa, R (2011b) Protein body-inducing fusions for high-level production and purification of recombinant proteins in plants. *Plant Biotechnol J* **9**: 419-433.

Doran, P.M (2006) Foreign protein degradation and instability in plants and plant tissue cultures. *Trends Biotechnol* **24**: 426-432.

Fagard, M., Vaucheret, H (2000) (Trans)Gene silencing in plants: How many mechanism?. *Annu Rev Plant Physiol Plant Mol* **51**: 167-194.

Finnegan, J., McElroy, D (1994) Transgene inactivation: Plants fight back!. *Nat Biotechnol* **12**: 883-888.

Fischer R., Stöger, E., Schillberg, S., Christou, P., Twyman, R.M (2004) Plant-based production of biopharmaceuticals. *Curr Opin Plant Biol* **7**: 152-158.

Floss, D.M., Sack, M., Stadlamann, J., Rademacher, T., Scheller, J., Stöger, E., Fischer, R., Conrad, U (2008) Biochemical and functional characterization of anti-HIV antibody-ELP fusion proteins from transgenic plants. *Plant Biotechnol J* **6**: 379-391.

Floss, D.M., Schallau, K., Rose-John, S., Conrad, U., Scheller, J (2009) Elastin-like polypeptides revolutionize recombinant protein expression and their biomedical application. *Trends Biotechnol* **28**: 37-45.

Galili, G (2004) ER-derived compartments are formed by highly regulated processes and have special functions in plants. *Plant Physiol* **136**: 3411-3413.

Geli, M.I., Torrent, M., Ludevid, D (1994) Two structural domains mediate two sequential events in γ-zein targeting: protein endoplasmic reticulum retention and protein body formation. *Plant Cell* **6**: 1911-1922.

Hadlington, J.L., Denecke, J (2000) Sorting of soluble proteins in the secretory pathway of plants. *Curr Opin Plant Biol* **3**: 461-468.

Hakanpää, J., Paananen, A., Askolin, S., Nakari-Setälä, T., Parkkinen, T., Penttilä, M., Linder, M.B., Rouvinen, J. (2004) Atomic resolution structure of the HFBII hydrophobin, a self-assemble amphiphile. *J Biol Chem* **279**: 534-539.

Harris, L.J, Gleddie, S.C (2001) A modified Rpl3 gene from rice confers tolerance of the Fusarium graminearum mycotoxin deoxynivalenol to transgenic tobacco. *Physiol Mol Plant Pathol* **58**: 173–181.

Havelda, Z., Hornyik, C., Crescenzi, A., Burgyán, J (2003) In situ characterization of *cymbidium ringspot tombusvirus* infection-induced posttranscriptional gene silencing in *Nicotiana benthamiana*. *J Virol* **77**: 6082-6086.

Herman, E.M., Larkins, B.A (1999) Protein storage bodies and vacuoles. *Plant Cell* **11**: 601-613.

Hood, E. E., Gelvin, S. B., Melchers, L. S., Hoekema, A (1993) New *Agrobacterium* helper plasmids for gene transfer to plants. *Transgenic Res* **2**: 208-218.

Horsch, R.B., Fry, J.E., Hoffmann, N.L., Eichholtz, D., Rogers, S.G., Fraley, R.T (1985) A simple and general method for transferring genes into plants. *Science* **227**:1229-1231.

Hraska, M., Rakousky, S., Curn, V (2008) Tracking of the CaMV-35S promoter performance in GFP transgenic tobacco, with a special emphasis on flowers and reproductive organs, confirmed its predominant activity in vascular tissues. *Plant Cell Tiss Organ Cult*. **94**: 239-251.

Joensuu, J.J., Niklander-Teeri, V., Brandle, J.E (2008) Transgenic plants for animal health: plant made vaccine antigens for animal infectious disease control. *Phytochem Rev* **7**: 553-577.

Joensuu, J.J., Brown, K.D., Conley, A.J., Clavijo, A., Menassa, R., Brandle, J. E (2009) Expression and purification of an anti-Foot-and-mouth disease virus single chain variable antibody fragment in tobacco plants. *Transgenic Res* **18**: 685-696.

Joensuu, J.J., Conley, A.J., Lienemann, M., Brandle, J.E., Linder, M.B., Menassa, R. (2010) Hydrophobin fusions for high-level transient protein expression and purification in *Nicotiana benthamiana*. *Plant Physiol* **152**: 622-633.

Kapila, J., De Rycke, R., Montagu, M.V., Angenon, G (1997). An *Agrobacterium*-mediated transient gene expression system for intact leaves. *Plant Sci* **122**: 101-108.

Kawakatsu, T., Takaiwa, F (2010) Cereal seed storage protein synthesis: fundamental processes for recombinant protein production in cereal grains. *Plant Biotechnol J* **8**: 939-953.

Kawakatsu, T., Hirose S., Yasuda, H., Takaiwa, F (2010) Reducing rice seed storage protein accumulation leads to changes in nutrient quality and storage organelle formation. *Plant Physiol* **154**: 1842–1854.

Kay, R. Chan, A., Daly, M., McPherson, J (1987) Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Science* **236**: 1299-1302.

Kim, C.S., Woo, Ym.Y.M., Clore, A.M., Burnett, R.J., Carneiro, N.P., Larkins, B.A. (2002) Zein protein interactions, rather than the asymmetric distribution of zein mRNAs on endoplasmic reticulum membranes, influence protein body formation in maize endosperm. *Plant Cell* **14**: 655-672.

Kim, S., Xu, J (2008) Aggregate formation of zein and its structural inversion in aqueous ethanol. *J Cereal Sci* **47**: 1-5.

Knäblein, J (2005) Plant-based expression of biopharmaceuticals. In *Encyclopedia of Molecular Cell Biology and Molecular Medicine* (R.A Meyers Ed.) Wiley-VCH, Weinheim **10**: 385-410.

Kogan, M.J., Dalcol, I., Gorostiza, P., Lopez-Iglesias, C., Pons, R., Pons, M., Sanz, F., Giralt, E (2002) Supramolecular properties of the proline-rich γ-zein N-terminal domain. *Biophys J* **83**: 1194-1204.

Lahtinen, T., Linder, M.B., Nakari-Setälä, T., Oker-Blom, C (2008) Hydrophobin (HFBI): A potential fusion partner for one-step purification of recombinant proteins form insect cells *Protein Expr purif* **59**:18-24.

Larkins, B.A., Hurkman, W.J (1978) Synthesis and deposition of zein in protein bodies of maize endosperm *Plant Physiol* **62**: 256-263.

Lau, O.S., Sun, S.S.M (2009) Plant seeds as bioreactors for recombinant protein. *Biotechnol adv* **27**:1015-1022.

Lending, C.R., Larkins, B.A (1989) Changes in the zein composition of protein bodies during maize endosperm development. *Plant Cell* **1**: 1011-1023.

Lico, C., Desiderio, A., Banchieri, S., Benvenuto, E (2005) Plants as biofactories: Production of pharmaceutical recombinant proteins. In *In the wake of the double helix: from the green revolution to the gene revolution* (Tuberosa, R., Phillips, R.L., Gale, M., eds). Bologna: Avenue Media: 577-593.

Lin, M., Rose-John, S., Grötzinger, J., Conrad, U., Scheller, J (2006) Functional expression of a biologically active fragment of soluble gp130 as an ELP-fusion protein in transgenic plants: purification via inverse transition cycling *Biochem J* **398**: 577–583.

Linder, M., Selber, K., Nakari-Setälä, T., Qiao, M., Kula, M.R., Penttilä, M (2001) The hydrophobins HFBI and HFBII from *Trichoderma reesei* showing efficient interactions with nonionic surfactants in aqueous two-phase systems. *Biomacromolecules* **2**: 511-517.

Linder, M.B., Qiao, M., Laumen, F., Selber, K., Hyytiä, T., Nakari-Setälä, T., Penttilä, M.E (2004) Efficient purification of recombinant proteins using hydrophobins as tags in surfactant-based two-phase systems. *Biochemistry* **43**: 11873-11882.

Linder, M.B (2009) Hydrophobins: proteins that self assemble at interfaces. *Curr Opin Colloid In***14**: 356-363.

Llop-Tous, I., Madurga, S., Giralt, E., Marzabal, P., Torrent, M., Ludevid, M.D (2010) Relevant elements of a maize γ-zein domain involved in protein body biogenesis. *J Biol Chem* **285**: 35633-35644.

Ma, J.K., Drake, P.M., Christou, P (2003) The production of recombinant pharmaceutical proteins in plants *Nat rev Genet* **4**:794-805.

Ma, J.K., Barros, E., Bock, R., Christou, P., Dale, P.J., Dix, P.J., Fischer, R., Irwin, J., Mahoney, R., Pezzotti, M., Schillberg, S., Sparrow, P., Stöger, E. and Twyman,

R.M (2005) Molecular farming for new drugs and vaccines. Current perspectives on the production of pharmaceuticals in transgenic plants. *EMBO rep* **6**: 593-599.

Mandal, S., Mandal, R.K (2000) Seed storage proteins and approaches for improvement of their nutritional quality by genetic engineering. *Curr Sci* **79**: 576-589.

Maxmen, A (2012) Drug-making plant blooms. *Nature* **485**: 160.

Menassa, R., Nguyen, V., Jevnikar, A., Brandle, J (2001) A self-contained system for the field production of plant recombinant interleukin-10. *Mol Breed* **8**: 177-185.

Menassa, R., Du, C., Yin, Z.Q., Ma, S., Poussier, P., Brandle, J. E., Jevnikar, A.M (2007) Therapeutic effectiveness of orally administered transgenic low-alkaloid tobacco expressing human interleukin-10 in a mouse model of colitis. *Plant Biotechnol J* **5**: 50-59.

Meyer, D.E. and Chilkoti, A. (1999) Purification of recombinant proteins by fusion with thermally-responsive polypeptides. *Nat Biotech* **17**:1112-1115.

Meyer, D.E., Chilkoti, A (2004) Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides. *Biomacromolecules* **5**: 846-851

Munro, S., Pelham H.R.B. (1987) A C-terminal signal prevents secretion of luminal ER proteins. *Cell* **48**: 899-907.

Nakari-Setälä, T., Aro, N., Kalkkinen, N., Alatalo, E., Penttilä, M (1996) Genetic and biochemical characterization of the *Trichoderma reesei* hydrophobin HFBI. *Eur J Biochem* **235**: 248-255.

Patel, J., Zhu, H., Menassa, R., Gyenis, L., Richman, A., Brandle, J (2007) Elastin-like poylpeptide fusions enhance the accumulation of recombinant proteins in tobacco leaves. *Transgenic Res* **16**: 239-249.

Peters, J., Stöger, E (2011) Transgenic crops for the production of recombinant vaccines and anti-microbial antibodies. *Hum Vaccin* **7**: 367-374.

Raju, K. and Anwar R.A (1987) Primary structures of bovine elastin a, b, and c deduced from the sequences of cDNA clones. *J of Biol Chem* **262**: 5755-5762.

Rybicki, E.P. (2009) Plant-produced vaccines: promise and reality. *Drug Discov Today* **14**:16-24.

Rybicki, E.P (2010) Plant-made vaccines for human and animals. *Plant Biotechnol J* **8**: 620-637.

Rymerson, R.T., Menassa, R., Brandle J.E (2002) Tobacco, a platform for the production of recombinant proteins. In *Molecular farming of plants and animals for human*

and veterinary medicine (Erickson L., Brandle J., Rymerson R.T., eds) Amsterdam: Kluwer: 1-32.

Scheller, J., Henggeler, D., Viviani, A., Conrad, U (2004) Purification of spider silk-elastin from transgenic plants and application for human chondrocyte proliferation. *Transgenic Res* **13**: 51-57.

Schillberg, S., Fischer, R., Emans, N (2003) Molecular farming of recombinant antibodies in plants. *Cell Mol Life Sci* **60**: 433-445.

Schubert, D., Lechtenberg, B., Forsbach, A., Gils, M., Bahadur, S., Schmidt, R (2004) Silencing in arabidopsis T-DNA transformants: The predominant role of a gene-specific RNA sensing mechanism versus position effects. *Plant Cell* **16**: 2561–2572.

Sheen, S.J (1983) Biomass and chemical composition of tobacco plants under high density growth. *Beitr Tabakforsch Int* **12**: 35–42.

Shewry P.R., Halford, N.G (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. *J Exp Bot* **53**: 947-958.

Stam, M., Mol, J.N.M., Kooter, J.M (1997) The silence of genes in transgenic plants. *Annals of Botany* **79**: 3-12.

Stöger, E., Parker, M., Christou, P., Casey, R (2001) Pea legumin overexpressed in wheat endosperm assembles into an ordered paracrystalline matrix. *Plant Physiol* **125**: 1732-1742.

Stöger, E., Ma, J.K., Fischer, R., Christou, P (2005) Sowing the seeds of success: pharmaceutical proteins from plants. *Curr Opin Biotech* **16**: 167-173.

Sunilkumar, G., Mohr, L.A., Lopata-Finch, E., Emani, C., Rathore, K. S (2002) Developmental and tissue-specific expression of CaMV35 promoter in cotton as revealed by GFP. *Plant Mol Biol* **50**: 463-474.

Sylhavi, D., Molnár, A., Lucioli, A., Szittya, G., Hornyik, C., Tavazza M., Burgyán, J (2002) A viral protein suppresses RNA silencing and binds silencing-generated, 21- to 25-nucleotide double stranded RNAs. *EMBO J* **21**: 3070-3080.

Torrent, M., Llompart, B., Laserre-Ramassamy, S., Llop-Tous, I., Bastida, M., Marzabal, P., Westerholm-Parvinen, A., Saloheimo, M., Heifetz, P.B., Ludevid, M.D (2009a) Eukaryotic protein production in designed storage organelles. *BMC Biol* **7**:1-14.

Torrent, M., Llop-Tous, I., Ludevid, M.D (2009b) Protein body induction: a new tool to produce and recover recombinant proteins in plants. In *Recombinant Proteins from Plants. Methods in Molecular Biology*. (Gomord V, Faye L. Heidelberg, Springer Verlag, Humana Press) **483**: 193-208.

Twyman, R.M., Stöger, E., Schillberg, S., Christou, P., Fischer, R (2003) Molecular farming in plants: host systems and expression technology. *Trends Biotechnol* **21**: 570- 578.

Tzen, J.T.C., Cao, Y., Laurent, P., Ratnayake, C., Huang, A.H.C (1993) Lipids, proteins, and structure of seed oil bodies from diverse species. *Plant Physiol* **101**: 267-276.

Urry, D.W (1988) Entropic elastic processes in protein mechanisms. I. Elastic structure due to an inverse temperature transition and elasticity due to internal chain dynamics. *J Protein Chem* **7**: 1-34.

Vitale, A., Ceriotti, A (2004) Protein quality control mechanisms and protein storage in the endoplasmic reticulum. A conflict of interests?. *Plant Physiol* **136**: 3420-3426.

Wösten, H.A.B., Vocht, M.L (2000) Hydrophobins, the fungal coat unravelled. *Biochim Biophys Acta* **1469**: 79-86.

Wang, X., Shi, F., Wösten, H.A.B., Hektor, H., Poolman, B., Robillard, G.T (2005) The SC3 hydrophobin self-assembles into a membrane with distinct mass transfer properties. *Biophys J* **88**: 3434-3443.

Wasaka, Y., Ozawa, K., Takaiwa, F (2009) Higher-level accumulation of foreign gene products in transgenic rice seeds by the callus-specific selection system. *J Biosci Bioeng* **107**: 78-83.

Wroblewski, T., Tomczak, A., Michelmore, R (2005) Optimization of *Agrobacterium*-mediated transient assays of gene expression in lettuce, tomato and *Arabidopsis*. *Plant Biotechnol J* **3**: 259-273.

Wu, K., Malik, K., Tian, L., Hu, M., Martin, T., Foster, E., Brown, D., Miki, B (2001) Enhancers and core promoter elements are essential for the activity of a cryptic gene activation sequence from tobacco, tCUP. *Mol Genet Genomics* **265**: 763-770.

Ziegelhoffer, T., Will, J., Austin-Phillips, S (1999) Expression of bacterial cellulase genes in transgenic alfalfa (*Medicago sativa* L.), potato (*Solanum tuberosum* L.) and tobacco (*Nicotiana tabacum* L.). *Mol Breed* **5**: 309–318.

7 APPENDICES

Appendix A Statistical analysis - Normality test. The statistical analysis was performed using Minitab 15. The Kolmogorov-Smirnov normality test (Lilliefor's test) demonstrated that the data estimated for all the transgenic lines does not have a normal distribution ($P < 0.010$).

Transgenic lines	KS
I64 GFP	0.262
I64 GFP-ELP	0.243
I64 GFP-HFBI	0.219
81V9 GFP-HFBI	0.232

Appendix B Statistical Analysis - Kruskal-Wallis test. The statistical analysis was performed using Minitab 15. The Kruskal-Wallis test demonstrated that at least one of the generated transgenic lines was significantly different from the other transgenic lines ($\alpha < 0.001$, $X^2 = 16.266$).

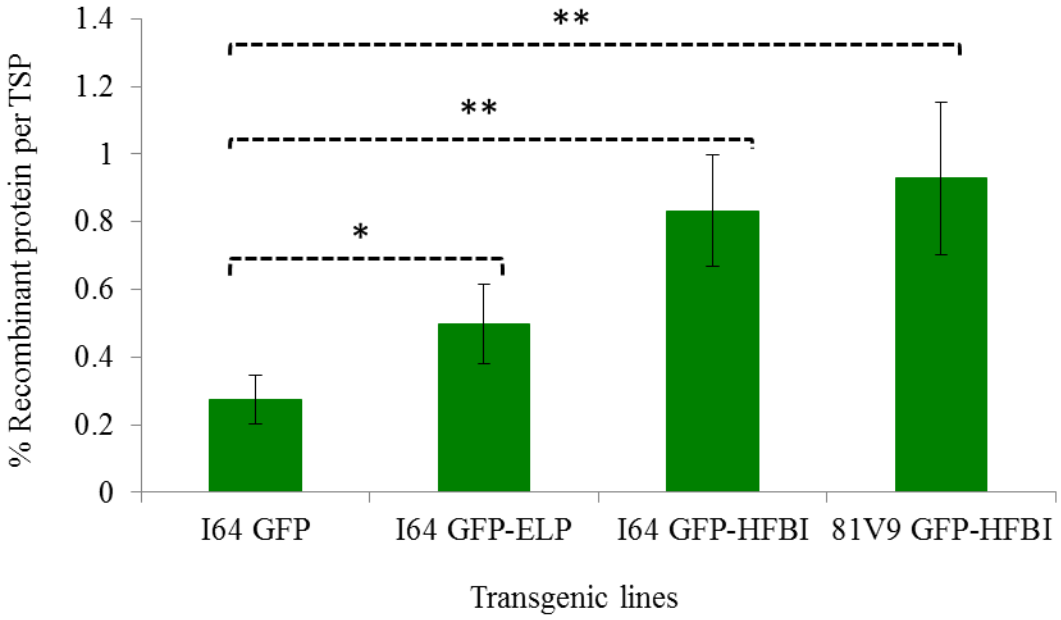
Kruskal-Wallis Test

Treatment	N	Median	Ave Rank	Z
I64 GFP	24	0.09834	31.6	-3.44
I64 GFP-ELP	24	0.23076	44.3	-0.86
I64 GFP-HFBI	24	0.48124	57.3	1.79
81V9 GFP-HFBI	24	0.59035	60.8	2.50
Overall	96		48.5	

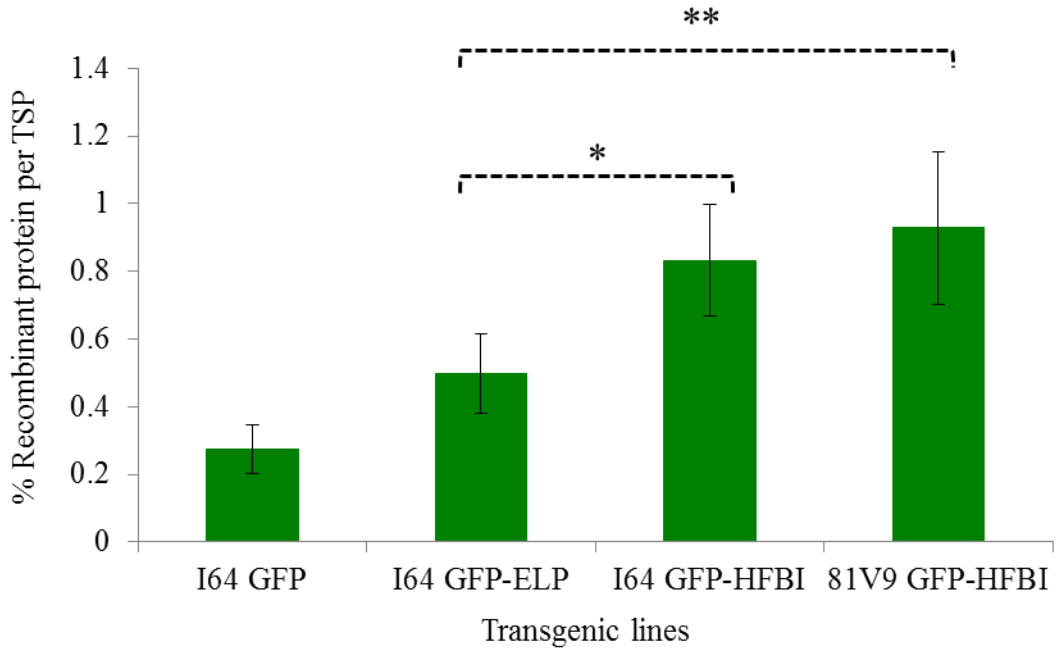
H = 16.54 DF = 3 **P = 0.001**

H = 16.54 DF = 3 P = 0.001 (adjusted for ties)

Appendix C Statistical Analysis - Wilcoxon-Mann-Whitney test. The statistical analysis was performed using Minitab 15. The Wilcoxon-Mann-Whitney test demonstrated that I64 GFP transgenic lines were significantly different from the I64 GFP-ELP, I64 GFP-HFBI and 81V9 GFP-HFBI transgenic lines (n=24; standard error is represent with error bars; significantly different at *P < 0.1; **P < 0.05).



Appendix D Statistical Analysis - Wilcoxon-Mann-Whitney test. The statistical analysis was performed using Minitab 15. The Wilcoxon-Mann-Whitney test demonstrated that I64 GFP-HFBI transgenic lines were not significantly different from the 81V9 GFP-HFBI transgenic lines. The Wilcoxon-Mann-Whitney test demonstrated that I64 GFP-HFBI and 81V9 GFP-HFBI transgenic lines were significantly different from the I64 GFP-ELP transgenic lines (n=24; standard error is represent with error bars; significantly different at *P < 0.1; **P < 0.05).



8 CURRICULUM VITAE

SONIA PATRICIA GUTIERREZ DELGADILLO

EDUCATION

2010-2012 **Master of Science Candidate**

Department of Biology

Cell and Molecular Biology
The University of Western Ontario, London, Ontario, Canada

2008-2009 **Master of Science in Biological Sciences. Phytopathology.**

Universidad de los Andes, Bogotá, Colombia 2009

2001-2007 **Bachelor of Science in Microbiology.**

Minor in **Industrial Microbiology.**

Universidad de los Andes, Bogotá, Colombia 2008

2001-2007 **Bachelor of Science in Biology**

Minor in **Molecular Biology.**

Universidad de los Andes, Bogotá, Colombia 2008

THESES

2012 **Protein body formation in stable transgenic plants of *Nicotiana tabacum* expressing elastin-like polypeptide and hydrophobin fusions.** M. Sc. Thesis. The University of Western Ontario**.** Project Supervisor: Rima Menasa Ph.D, Co-supervisor: Susane Kohalmi PhD.

2007-2009 **Identification of type three effector proteins of *Xanthomonas axonopodis* pv. manihotis using reporter translocation systems.** M. Sc. Thesis. Laboratory of Mycology and Phytopathology (LAMFU). Universidad de los Andes, Bogotá. Suopervisor: Adriana Bernal Ph.D, Co-supervisor Silvia Restrepo PhD.

2007 **Development of a screening system for the identification of effector proteins of *Xanthomonas axonopodis* pv. manihotis.** Bachelor of Science thesis**.** Laboratory of Mycology and Phytopathology (LAMFU). Universidad de los Andes, Bogotá. Supervisor: Adriana Bernal Ph.D, Co-supervisor: Silvia Restrepo PhD.

RESEARCH EXPERIENCE

2009-2010 **Graduate Assistant. Laboratory of Mycology and Phytopathology (LAMFU). Universidad de los Andes, Bogotá.** Project director: Adriana Bernal Ph.D and Silvia Restrepo PhD. I worked in a molecular biology project in *Pseudomonas syringae* pv. tomate, where isolates from South America are being used for a comparative evolutionary

genomics investigation, proposed and carried out by Boris A Vinatzer in Virginia Tech.

2008 **Scientific Exchange. ECOS NORD Program. Laboratoire de Génome et Développment des Plantes, Institute de Recherche pour le Dévelopment (IRD), Montpellier, France.** Project Director: Valerie Verdier, Boris Szureck and Ralph Roebnik. I worked in a molecular biology project in *Xanthomonas oryzae* pv. *oryzae*. I used Homologue Recombination to obtain mutants of genes that codify proteins involved in the formation of the type three secretion apparatus. Plant inoculations in susceptible varieties of rice to characterized strains and different mutants of *Xanthomonas oryzae* pv. *oryzae*.

TEACHING EXPERIENCE

2010-2012 **Graduate Teaching Assistant.** Fourth year Fungi Biology Laboratory, Undergraduate Course. The University of Western Ontario, Canada.

2010-2012 **Graduate Teaching Assistant.** First year Cell Biology Laboratory, Undergraduate Course. The University of Western Ontario, Canada**.**

2007-2009 **Graduate Teaching Assistant.** Plant Pathology Laboratory, Undergraduate Course. Universidad de los Andes, Bogotá.

2007 **Teaching Assistant.** Plant Pathology Laboratory. Universidad de los Andes, Bogotá.