11-28-2014 12:00 AM

Protein Body Biogenesis and Utility in Recombinant Protein Production in Nicotiana benthamiana

Reza Saberianfar
The University of Western Ontario

Supervisor
Dr. Rima Menassa
The University of Western Ontario Joint Supervisor
Dr. Susanne Kohalmi
The University of Western Ontario

Graduate Program in Biology
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy

PROTEIN BODY BIOGENESIS AND UTILITY IN RECOMBINANT PROTEIN PRODUCTION IN *NICOTIANA BENTHAMIANA*

(Thesis format: Integrated Article)

by

Reza Saberianfar

Graduate Program in Biology

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

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

Abstract

Protein bodies (PBs) are endoplasmic reticulum (ER) derived organelles found in seeds whose function is to accumulate seed storage proteins. It was shown that PB formation is not limited to seeds, and green fluorescent protein (GFP) fused to either elastin-like polypeptide (ELP), hydrophobin-I (HFBI) or Zera® fusion tags induces PBs in leaves of *Nicotiana benthamiana*. The mechanism by which fusion tags induce PBs is not well understood. To address how PBs form and develop in plant leaves, I studied the factors involved in their formation including recombinant protein concentration, effect of the fusion tags, PB sequestration patterns and destination *in vivo*.

In this study I showed that PB formation is a concentration-dependent mechanism and that proteins accumulating at levels higher than 0.2% of total soluble protein are capable of inducing PBs *in vivo*. The presence of fusion tags is not necessary for the formation of PBs but affects their distribution pattern and size. ELP-induced PBs are larger than HFBI-induced PBs and the size of both PBs increases over time along with accumulation levels of the recombinant protein. I found that in the process of PB formation, secretory and ER resident molecules are passively sequestered into the lumen of PBs. This property of PBs was proposed as a tool to increase accumulation levels of erythropoietin and human interleukin-10 by co-expression with PB-inducing proteins. To understand whether PBs are terminally-stored cytosolic organelles or if they are connected to the ER and to each other, I developed an EGFP-based photoconversion technique, which was successfully used to visualize the trafficking of proteins targeted to the cytosol, ER, apoplast, and chloroplast *in vivo*. Study of PBs with this technique suggested that PBs remain associated with the ER and communicate with one another via the ER.

Keywords

Protein body, PB, fusion tag proteins, elastin-like polypeptide, ELP, hydrophobin, HFBI, Zera, photoconversion, *Nicotiana benthamiana*.

Co-Authorship

The following thesis contains material from manuscripts in preparation which are co-authored by Reza Saberianfar (RS), Andrew J. Conley (AC), Jussi J. Joensuu (JJ), Amirali Sattarzadeh (AS), Warren R. Zipfel (WZ), Maureen Hanson (MH), Susanne E. Kohalmi (SK), and Rima Menassa (RM).

My supervisors Rima Menassa and Susanne Kohalmi provided insight and strategic direction for the projects and also edited the final manuscripts (Chapters 3-5).

Chapter 2. Author's contributions

RS designed the research, performed the experiments and drafted the manuscript. JJ and AC conceived the co-expression concept. AC performed the IL-10 co-expression experiment (Chapter 2, Figure 2.7b). RM conceived the study and participated in its design. RS, JJ, AC, RM edited the manuscript.

Chapter 3. Author's contributions

RS and AS designed the research, performed the experiments and drafted the manuscript. WZ characterized the biophysical properties of EGFP (Chapter 3, Figure 2.7 b-d). MH provided material and lab space. RM and MH conceived the study. RS, AS, MH and RM edited the manuscript.

Chapter 4. Author's contributions

RS designed the research, performed the experiments and drafted the manuscript. RM conceived the study and participated in its design. RS, SK, and RM edited the manuscript.

Dedication

To Maman, Baba, Noor and Mandana

for all of their love, support and encouragement.

Acknowledgments

I would like to express my sincere gratitude to my supervisor and mentor, Dr. Rima Menassa who helped me find my way through my PhD. I am deeply grateful for her guidance, encouragement, support and trust to work on this project. The time she spent on reading and polishing up my draft was enormous, for which I cannot be thankful enough.

Dr. Susanne Kohalmi has been my co-supervisor officially, but her contribution to my work went beyond that. She has been a great source of encouragement and inspiration during my studies. It has also been a privilege to have Dr. Greg Kelly and Dr. Richard Gardiner serve on my committee. I am very grateful to them for sharing their knowledge and expertise with me. I would also like to express my appreciation to Dr. Jussi Joensuu for his ideas and constructive suggestions, and to Dr. Maureen Hanson for kindly hosting me to work in her lab and helping me develop part of my research project.

A special thanks to Angelo Kaldis and Hong Zhu for their technical and intellectual assistance and for their never ending patience. I am thankful to Alex Molnar for his assistance with the preparation of figures and posters. I would also like to extend my thanks to all the past and present Menassa lab members: Mistianne Feeney, Sonia Gutiérrez, Igor Kolotilin, Eridan Pereira, Caroline Sepiol, Ruoyu Yan, Sean Miletic, Kira Liu, Jacqueline McDonald, Tanja Patry, Jenny Ge, Zein Khamis, and colleagues and friends from Agriculture and Agri-Food Canada: Mehran Dastmalchi, Alex Tromas, Behnaz Saatian, Preetam Janakirama, Jaya Joshi, Arun Kumaran, and friends from UWO: Rainer Bode, Hadi Ghofrani and Sooshiant Kiarasi. Thanks for your companionship through hard times and special thanks for all the great memories we will share for the rest of our lives.

Finally, I would like to acknowledge the Department of Biology at UWO, the Ontario Graduate Scholarship (OGS) program and the private award donors who generously provided financial support throughout my graduate studies. Their support was encouraging and kept me focused on my project during my degree.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations:

ATPS	Aqueous two-phase separation system
BiP	Immunoglobulin heavy chain binding protein
BSA	Bovine serum albumin
BY-2	Tobacco bright yellow-2 cell line
CaMV	Cauliflower mosaic virus
CFP	Cyan fluorescent protein
dpi	Days post infiltration
ECFP	Enhanced cyan fluorescent protein
EGFP	Enhanced green fluorescent protein
ELISA	Enzyme-linked immunosorbent assay
ELP	Elastin-like polypeptide
EPO	Erythropoietin
ER	Endoplasmic reticulum
GFP	Green fluorescent protein
HC-Pro	Helper component protein
HDEL	ER retrieval signal
HFBI	Hydrophobin-I
HFBs	Hydrophobins
hIL-10	Human interleukin-10
IB	Inclusion body
IL-10	Interleukin-10
ITC	Inverse transition cycling
KDEL	ER retrieval signal
MES	Morpholinoethanesulfonic acid

nos	Nopaline synthase
NT-1	*Nicotiana tabacum*-1 cell line
OB	Oil body
OD	Optical density
PB	Protein body
Pr1b	Pathogenesis-related protein 1b of tobacco
PSV	Protein storage vacuole
PTGS	Post transgenic gene silencing
PVDF	Polyvinylidene difluride
PVX	Potato virus X
RISC	RNA induced silencing complex
RdRp	RNA-dependent RNA polymerase
RFP	Red fluorescent protein
SQS1	Squalene synthase 1
TBSV	Tobacco bushy stunt virus
tCUP	Tobacco cryptic upstream promoter
TEV	Tobacco etch virus
Ti	Tumor inducing
TMV	Tobacco mosaic virus
TSP	Total soluble protein
T_t	Transition temperature
UTR	Untranslated region
YFP	Yellow fluorescent protein
Xyn	Xylanase

1 General Introduction

1.1 Plants as bioreactors

During the past two decades, recombinant protein production in plants has gained popularity over other conventional methods. Plants are now considered as a safe, efficient and inexpensive platform for production of a wide range of recombinant proteins including enzymes, vaccines, antibodies and other biopharmaceuticals (Gutiérrez and Menassa, 2014). In 2012, the first plant-produced therapeutic protein, taliglucerase alfa, for treatment of Gaucher disease was approved by the US Food and Drug Administration (Maxmen, 2012).

The second success story was recently reported upon the outbreak of Ebola. ZMapp™, a cocktail of three monoclonal antibodies produced in transgenic *Nicotiana benthamiana* is so far the only treatment successfully used against the Ebola virus. This cocktail combination was tested on infected Rhesus monkeys and 100% of the animals recovered (Qiu *et al.*, 2014). Even though ZMapp™ was produced by collaboration of three biotechnology companies (Mapp Biopharmaceutical Inc. (San Diego), LeafBio (San Diego), Defyrus Inc. (Toronto)), and the public health agency of Canada and the defense advanced research project agency (DARPA) of the U.S.A., the cocktail was only enough for treatment of six individuals, and another 50 treatment courses are expected to be produced by the end of 2014 (Strauss, 2014).

Although successful cases of plant-produced recombinant proteins are gradually emerging, plant biotechnology still faces two serious challenges: low production yield and lack of efficient purification methods that need to be addressed. To achieve the highest levels of recombinant protein production in plants, several factors should be taken into consideration, such as the expression system, the target tissue and the subcellular location for production of the protein of interest (Twyman *et al.*, 2013).

Currently, several types of plant expression systems are available for production of foreign proteins, which include transgenic plants with either nuclear or chloroplast transformed genomes, cell suspension cultures and transient expression. Similarly, recombinant proteins have been targeted to a wide range of subcellular locations such as

the ER, the apoplast, the cytosol and the chloroplast to optimize their expression and accumulation levels. Each one of these categories is explained in the following sections.

1.1.1 Nuclear transformed plants

Biological and physical methods can be used to transfer foreign DNA into a plant cell. The most studied biological method is *Agrobacterium*-mediated transformation. *Agrobacterium tumefaciens* is a soil microorganism that causes crown gall disease in many species of dicotyledonous plants (Shrawat and Lörz, 2006) and is well known for its ability to infect plants with a tumor-inducing (Ti) plasmid. A remarkable feature of the Ti plasmid is that, after infection, part of the plasmid molecule, called T-DNA, integrates into the plant chromosomal DNA (Zhu *et al.*, 2000). This strategy can be used to integrate any foreign gene into the T-DNA and infect the cells with the *Agrobacterium* containing the recombinant plasmid (Gelvin, 2003). Since *Agrobacterium*-mediated transformation is more successful in dicots than in monocots, other transformation methods are required to address the need for transforming monocot plants. Such methods are usually based on physical delivery of the genes into the plant cells.

Direct gene transfer via microprojectile bombardment (biolistics), polyethylene glycol (PEG) mediated transformation, electroporation, and inoculation following abrasion of the surface are examples of physical transfer of DNA into plant cells (Rao *et al.*, 2009). Biolistic delivery is the most common physical approach for introduction of a transgene into the nuclear genome of a host plant (Sanford *et al.*, 1987). A “gene gun” is usually used to shoot gold or tungsten particles, coated with “naked DNA”, into plant cells. This method has been successfully applied on a wide range of crops especially monocots such as wheat and maize (Shrawat and Lörz, 2006).

The prominent advantage of nuclear transformation of plants is their flexibility and efficiency in scaling up the production of recombinant proteins. Additionally, since foreign genes are integrated into the nuclear genome of plants, they can be inherited over generations which accounts for stable and predictable transgene expression in several generations (Egelkrout *et al.*, 2012).

On the other hand, nuclear transformation requires relatively long periods of time (approximately 6 to 9 months) to generate transgenic plants and select for lines producing high levels of recombinant proteins. Also, biosafety concerns are raised with transgenic plants due to the risk of gene leakage through cross pollination and seed mixing with wild type plants (Andow and Hilbeck, 2004; Breyer *et al.*, 2012). Since the integration of the T-DNA happens in a random fashion into unpredictable locations within the plant genome, expression levels may vary between different transformants (also known as "positional effect"). As a result, nuclear transformed plants generally provide low accumulation levels of recombinant proteins, which usually fall within the range of 0.001% to 1% of the total soluble protein (TSP) (Lossl and Waheed, 2011).

1.1.2 Chloroplast transformed plants

Integrating a transgene into the chloroplast genome carries two major advantages when compared with nuclear transformed plants: first, accumulation levels of the recombinant proteins are generally higher in transplastomic plants (plants whose chloroplast genome is transformed); second, since in most angiosperms plastids are inherited in a maternal fashion, it is highly unlikely that the foreign DNA will be transmitted through pollen which reduces the risk of gene flow from genetically modified plants to others (Maliga and Bock, 2011; Ruf *et al.*, 2007).

Chloroplast transformation is an attractive candidate for expression of a transgene since every chloroplast contains approximately 100 identical genomes and a typical plant leaf cell contains about 100 chloroplasts. Therefore, by transforming chloroplasts, a single gene can be represented up to 10,000 times within one cell. Considering this major advantage, it is not unusual to see accumulation levels of 5% to 20% of TSP in transplastomic plants (Chebolu and Daniell, 2009; Svab *et al.*, 1990). Unlike the random integration of the T-DNA in nuclear transformation which results in variable expression of the transgene, incorporation of the transgene occurs accurately through homologous recombination in chloroplasts. This precise integration of the transgene reduces labor-intensive and time-consuming screening steps required for selecting high expressing plants produced by nuclear transformation (caused by positional effects). Lack of

transgene silencing is another advantage of chloroplast expression systems in addition to the accumulation of transcripts at high levels compared with nuclear transgenic plants. Moreover, chloroplast transformation includes the ability of expressing several genes simultaneously (transgene stacking), and the ability to perform the complex post-translational modifications including disulfide bond bridging, protein lipidation, folding and assembly of recombinant proteins (New *et al.*, 2012; Verma and Daniell, 2007).

However, chloroplast transformation has its own limitations. Because of their prokaryotic endosymbiotic origins, chloroplasts have prokaryotic transcription and translation machineries and as such, they are unable to perform glycosylation which is necessary for many pharmaceutical glycoproteins such as monoclonal antibodies (Hefferon, 2013). So far, plastid transformation has been more successful in dicots and has remained limited to a few plant species such as tobacco, tomato, and lettuce (Clarke and Daniell, 2011).

1.1.3 Plant cell suspension cultures

Plant cell suspension cultures have the same advantages of whole plant cultivation, and also offer advantages found in microbial and animal cell cultures. This includes rapid doubling times, easy scale-up, post-translational modifications and product safety. Moreover, recombinant proteins produced in plant cell suspension cultures display better consistency than over other plant expression systems. This is mainly because of the highly controlled conditions in bioreactors leading to uniformity of cells grown in suspension cultures (De Muynck *et al.*, 2009). A number of good examples are tobacco BY-2 and NT-1 cell lines which have high growth rates, and also produce highly consistent protein products (Hellwig *et al.*, 2004; Xu *et al.*, 2011).

The main advantage of cell suspension cultures is their ability to secrete the recombinant protein to the cell culture medium, which eliminates the need to break the cells, and reduces the complexity of purification (Huang *et al.*, 2009). However, the major challenge for production of secreted recombinant proteins is relatively low accumulation (Xu *et al.*, 2011). The main reason for low accumulation levels in the cell culture medium is degradation by secreted proteases (Schiermeyer *et al.*, 2005). However, retaining

recombinant proteins intracellularly has been shown to produce high accumulation levels (Kaldis *et al.*, 2013).

1.1.4 Transient expression

Transient expression of a foreign gene stands out in several aspects when compared with other plant expression systems. Simplicity and ease of work, scalability, short generation and production timelines, high recombinant protein levels and biosafety are some of the major advantages of this expression system. To date, several breakthroughs have been reported in transient expression that allow for achieving impressive expression levels. These methods usually rely on two classes of plant pathogens; *Agrobacterium*, plant viruses or a combination of both (Magnifection).

1.1.4.1 *Agrobacterium*-mediated expression

The mechanism by which *Agrobacterium* transformation works is explained in section 1.1.1. Once the T-DNA is trafficked to the nucleus, only a small percentage of the T-DNA is integrated into the host genome which causes the formation of stable transformed cells that can subsequently be generated into transgenic plants (Zambryski, 1988). It is not clear what happens to the T-DNAs which do not integrate into the genome, but it has been shown that these T-DNAs are transcriptionally competent, and therefore could be employed for expression of the transgene for short periods of time (Voinnet *et al.*, 2003).

By using *Agrobacterium*-mediated transient expression, harvesting timelines have been reduced drastically. High yields of recombinant proteins can be achieved within 2-5 days, making this expression system an attractive option for production of recombinant proteins. This system is also very flexible as multiple genes can be expressed simultaneously. Since transient expression is independent of the problems involved with positional effects, it can be a reliable and reproducible source for transgene expression (Kapila *et al.*, 1997).

1.1.4.2 Virus-based expression

Viral-based expression systems rely on the idea of using viruses as vectors to deliver transgenes into plants. Tobacco mosaic virus (TMV) and potato virus X (PVX) are among the most commonly used viruses for transient expression. Virus-based systems have several advantages such as rapid delivery of the transgene and systemic infection of the whole plant after inoculation. As a result, they produce high levels of recombinant proteins. Plant virus-based expression systems have been used for production of several pharmaceutical proteins such as full-sized and fragments of antibodies (Gleba *et al.*, 2007; Yusibov *et al.*, 1997).

Viral expression systems have two major disadvantages: size limitations of the insert introduced into the viral genome, and safety concerns regarding the spread of the viral vectors into the environment (Yusibov *et al.*, 1997). To resolve these issues, a new system was designed by Icon Genetics, a biotechnology company, called "Magnifection" (also known as MagnICON®). Magnifection combines the high transfection efficiency of *Agrobacterium* with high expression yield of viral vectors (Gleba *et al.*, 2005). This method is based on a "deconstructed" TMV vector, optimized by deletion of viral sequences responsible for replication and spreading of the virus, and split into two modules. The first module carries the viral polymerase and the movement protein genes while the second module includes the gene of interest. Both modules are inserted into binary vectors and transformed into *Agrobacterium*. In this process, bacteria are responsible for the primary role of infection and delivery of modules into the plant cell nucleus. Also, a third *Agrobacterium* cell line is required to deliver a specific recombinase so the modules can assemble *in vivo*. Since TMV-based vectors cannot move systemically due to lack of a coat protein, production in the whole plant requires inoculation of all leaves. This was addressed by immersing all aerial parts of the plant into the bacterial solution (Marillonnet *et al.*, 2005). Major advantages of the MagnICON® system are: flexibility with the size of the transgene, inability to leak through the plant and into the environment, and accumulation of high quantities of recombinant protein. For instance, the company Kentucky BioProcessing (KBP) uses this

system to produce 25-75 grams of antibody per greenhouse on a biweekly basis (Pogue *et al.*, 2010).

1.1.4.3 Commercial scale *Agro*-infiltration

Transient expression by *Agro*-infiltration was initially used as a rapid technique to assess the ability and efficiency of constructs for production of recombinant proteins. For this purpose, transformation of only a portion of the plant tissue (e.g. leaf) with a syringe would suffice. Syringe infiltration has been used on a wide range of hosts including *Arabidopsis thaliana*, *Nicotiana tabacum*, tomato and lettuce (Wroblewski *et al.*, 2005). Due to the efficiency of this method, it has been scaled up for vacuum infiltration to fulfill the demands of the market. Vacuum infiltration enables the infiltration of the whole plant or kilograms of the tissue at once. It has been shown that syringe infiltration is approximately twice as efficient compared with vacuum infiltration (Vézina *et al.*, 2009). However, vacuum infiltration has been accepted and developed by several biotechnology companies such as Medicago Inc., Kentucky BioProcessing, and the Fraunhofer institute to infiltrate kilograms of plants per hour (D'Aoust *et al.*, 2010; Pogue *et al.*, 2010).

1.2 Strategies to increase expression and accumulation of recombinant proteins

1.2.1 Transcription

To ensure high levels of transcription, the most important factor is the promoter. To date, several promoters have been developed to obtain high levels of transcripts in plants. These promoters are usually obtained from plants or plant pathogens. Egelkrout *et al.* have reviewed 84 different promoters used for protein accumulation in various plant species (Egelkrout *et al.*, 2012). These promoters can generally be divided into constitutive, organ-specific and inducible promoters. Each one of these categories can be further divided into monocot- or dicot-specific promoters. Some of the most popular promoters are discussed in the following paragraphs.

The cauliflower mosaic virus (CaMV) 35S promoter has been used extensively since its discovery almost three decades ago (Gutiérrez *et al.*, 2013; Odell *et al.*, 1985). CaMV 35S is a strong constitutive promoter and the most popular choice for dicot plants. However, it has a much weaker activity in monocots. To obtain high expression levels in monocots, maize ubiquitin-1 (Christensen *et al.*, 1992) or *Agrobacterium tumefaciens nopaline synthase* (nos) (Shaw *et al.*, 1984) promoters are better choices.

To express the transgene in a particular tissue or at a specific developmental stage, tissue-specific promoters and inducible promoters were developed. For instance, maize *globulin-1*(Belanger and Kriz, 1989), *globulin-2* (Hood *et al.*, 2003) and *novel* (Streatfield *et al.*, 2010) are monocot embryo-specific promoters while *maize* α- and γ-zeins (Zhang *et al.*, 2009), rice *glutelin*, *globulin* and *prolamins* (Qu *et al.*, 2008) are examples of endosperm-specific promoters. These promoters restrict the expression of the foreign gene to certain tissues or areas of the seed. As a result, not only stability of the recombinant protein increases but also the whole plant is protected against the unfavorable effects (e.g. toxicity) of highly accumulated protein (Stoger *et al.*, 2002).

To further increase transcript levels of a transgene, other strategies have been developed such as incorporation of additional duplicated enhancer elements to CaMV 35S promoter. It has been shown that the activity of double enhanced CaMV 35S promoter is approximately 10 times stronger than the natural CaMV 35S (Kay *et al.*, 1987). Another strategy is to synthesize “hybrid promoters” by combining the most critical sequences of multiple well-characterized natural promoters. Combination of critical sequences of CaMV 35S and *Agrobacterium* Ti plasmid mannopine synthase promoters resulted in a much stronger activity compared with either of the two parental promoters (Comai *et al.*, 1990).

1.2.2 Suppression of post-transcriptional gene silencing (PTGS)

PTGS is considered one of the plants earliest self defense mechanisms against molecular parasites such as transposons, viruses and transgenes (Voinnet, 2001). PTGS is capable of generating sequence-specific signals that can move from cell to cell and activate the systemic silencing in faraway tissues which makes it an efficient antiviral system

(Palauqui *et al.*, 1997; Voinnet and Baulcombe, 1997). Highly expressed transgenes were shown to be susceptible to PTGS. This mechanism requires an RNA-dependent RNA polymerase (RdRp)-like protein to generate the complementary RNA of the target species (Dalmay *et al.*, 2000). Following the formation of double-stranded RNA (dsRNA), it will be recognized and cleaved by a nuclease into 21-23 nucleotide small dsRNAs (Voinnet *et al.*, 1999; Zamore *et al.*, 2000). These dsRNAs will incorporate into a conserved nuclease protein complex called RNA-induced silencing complex (RISC) and lead the complex to the complementary target mRNA for sequence-specific cleavage of target RNA transcript (Voinnet *et al.*, 1999).

Since suppression of gene silencing is critical for invading viruses, many plant viruses have evolved silencing suppressor proteins. These proteins were shown to target different steps of the silencing process (Qu and Morris, 2005; Voinnet *et al.*, 1999). To avoid PTGS, these virus-based suppressors of gene silencing can be expressed simultaneously with the gene of interest. For instance, co-expression of a murine anti-human IgG C5-1 with HC-Pro (from potato virus Y) resulted in a 5.3-fold increase in C5-1 antibody accumulation levels (Vézina *et al.*, 2009). Another well-studied example of PTGS suppressor is the tomato bushy stunt virus (TBSV) p19 protein. Co-expression of p19 along with the transgene was shown to increase accumulation levels of the target protein up to 50-fold (Voinnet *et al.*, 2003). Similarly, co-expression of p19 with trastuzumab, a therapeutic antibody used in the treatment of $HER2^{+}$ breast cancer, increased accumulation levels of trastuzumab in *N. benthamiana* up to 15-fold (Garabagi *et al.*, 2012). This strategy is compatible with different expression systems and can be easily used for both lab and industrial scale purposes.

When the p19 gene was incorporated in the same T-DNA cassette that contains the transgene in the pEAQ vector series (Peyret and Lomonossoff, 2013; Sainsbury *et al.*, 2009), the result was higher levels of recombinant protein than co-infiltrating a separate *Agrobacterium* culture containing p19 (Sainsbury *et al.*, 2009).

1.2.3 Translation

Maintaining optimal levels of transcript does not necessarily result in high accumulation levels of the recombinant protein. Approximately 20% to 40% of foreign protein accumulation is directly regulated by its corresponding mRNA transcript. Factors such as translation and degradation rates strongly influence the accumulation levels of recombinant proteins (Nie *et al.*, 2006). Regulation at the translational level has such importance that in some cases it has been counted as the most critical regulator of recombinant protein abundance. To boost the translation of foreign proteins, several strategies have been investigated including codon optimization, and the use of transcript untranslated regions (UTRs) (Desai *et al.*, 2010).

Different organisms have particular codon usages which can cause problems during translation such as pausing at disfavoured codons, and frame-shifting. Therefore, to increase the translation efficiency of a foreign gene, codon optimization is considered a priority, most importantly for production of prokaryotic proteins in eukaryotic expression systems. For instance, codon optimization of a synthetic *cholera toxin B* (CTB) gene based on that of tobacco resulted in approximately 15-fold increase in accumulation level of the recombinant protein compared with the bacterial CTB (Kang *et al.*, 2004). With the advances in the field of bioinformatics, currently several codon usage software packages are available online such as Gene Designer (Villalobos *et al.*, 2006), and Optimizer (Puigbò *et al.*, 2007) to help scientists optimize their transgene sequences at no cost. It is important to mention that any codon change may cause unwanted and problematic modifications at the RNA level leading to instability at the post-transcriptional processing level. Therefore, the codon-optimized sequence should be carefully examined for unexpected regulatory outcomes as well as possible issues with folding and stability of the transcript (Laguia-Becher *et al.*, 2010).

Another strategy to increase the rate of translation is to include the plant or plant viral 5'-untranslated sequences into the transgene construct to increase the rate of translation initiation. Several 5'-untranslated regions (5'-UTRs) have been studied and proved to be effective in increasing the accumulation levels of their respective proteins such as the 5'-

UTR of tobacco mosaic virus (Sleat *et al.*, 1987), potato virus X (Pooggin and Skryabin, 1992), tobacco tCUP (Wu *et al.*, 2001), and *Oryza sativa* alcohol dehydrogenase that was used to increase recombinant protein accumulation in both monocots and dicots (Sugio *et al.*, 2008). Liu *et al.* (2010) have also described the role of six different 5'-UTRs of seed-storage genes in enhancing the expression of foreign genes in stable transgenic rice.

1.2.4 Optimization of protein accumulation and stability

To achieve high accumulation levels of heterologous proteins in plants not only boosting transcription and translation are required, but also it is important to protect the produced protein against degradation. Proteases are the main factor responsible for degradation of heterologous proteins. Several strategies have been proposed to avoid or reduce the effects of proteases in plants including the use of protease-deficient plants, co-expression with protease inhibitors, targeting of the recombinant protein to a specific organelle, or fusion to other proteins (Benchabane *et al.*, 2008).

1.2.4.1 Proteolysis prevention via co-expression of protease inhibitors

The major function of proteases is to degrade the incorrectly folded or abnormal proteins to ensure the final quality of the synthesized proteins; meanwhile, proteases directly affect the yield of foreign proteins expressed in plants (Desai *et al.*, 2010). Protease-deficient plants were proposed to avoid protease-mediated degradation. This strategy was successfully developed in bacteria (Jiang *et al.*, 2002) and yeast (Macauley-Patrick *et al.*, 2005) several years ago, but as there are approximately over 1900 genes directly or indirectly involved in the hydrolysis of peptide bonds in plants, production of these plants is highly unlikely (Benchabane *et al.*, 2008). Another strategy is the co-expression of protease inhibitors along with the heterologous protein. For instance, expression of a cathepsin D inhibitor, S/CD1, was shown to increase protein levels in a proteasome-independent manner for intrinsic and transgenic proteins expressed in the cytosol (Goulet *et al.*, 2010). Similarly, expression of the secretory S/CD1 negatively affected the activities of pepsin-like (A1 family) and trypsin/chemotrypsin-like (S1 family) proteases

in the apoplast. This strategy was used to increase accumulation levels of C5-1, a murine antibody, secreted to the apoplast (Goulet *et al.*, 2012).

1.2.4.2 Subcellular targeting

A wide range of proteases are found in plant cells in different organelles including cytoplasm, vacuole, endoplasmic reticulum (ER), and apoplast, although the number, quantity and activity of proteases in these organelles are significantly different from one another (Benchabane *et al.*, 2008). In general, the cytoplasm, lytic vacuole and apoplast contain a high proteolytic content which results in low accumulation levels of foreign proteins in these organelles (Goulet *et al.*, 2006). Alternatively, the ER is considered a mild environment for the production of recombinant proteins due to low abundance of proteases (Faye *et al.*, 2005). Therefore, to minimize the unintended proteolysis of the recombinant proteins, the addition of the tetra-peptide KDEL (Lys-Asp-Glu-Leu) or HDEL (His-Asp-Glu-Leu) is used to retrieve the protein to the ER. Aside from the availability of fewer proteases in the ER, which provides a protective environment for recombinant proteins, the abundance of chaperones in the ER also helps with folding and assembly of heterologous proteins (Desai *et al.*, 2010; Nuttall *et al.*, 2002). Also, foreign proteins can acquire post-translational modifications in the ER such as disulfide bond formation and glycosylation (Ma *et al.*, 2003). ER retrieval has been used to increase accumulation levels of several antibodies, vaccines and cytokines in transgenic plants (Conley *et al.*, 2009c; Joensuu *et al.*, 2006; Kaldis *et al.*, 2013). Other subcellular compartments have also been used for targeting proteins including the chloroplast (Van Molle *et al.*, 2007), protein bodies, and protein storage vacuoles (Stoger *et al.*, 2005) which will be discussed in details in the next section. It is important to note that although several successful examples of production of recombinant proteins are mentioned above, it is often not possible to predict which subcellular compartment will be best suited for accumulating the recombinant proteins of interest and this is usually empirically determined by targeting the protein of interest to several organelles (Pereira *et al.*, 2014; Van Molle *et al.*, 2007).

1.2.4.3 Targeting to storage organelles

Seeds offer highly specialized organelles for protein storage. These include oil bodies, protein storage vacuoles and protein bodies. These organelles have been studied for their potential role in storage of high amounts of recombinant proteins in seeds (Khan *et al.*, 2012), but also they have been induced in leaves (Conley *et al.*, 2011a; Feeney *et al.*, 2013).

In seeds, both oil bodies and protein bodies have an ER origin. ER is the biggest organelle in the plant cell, and is composed of several domains such as cisternal ER, tubular ER, the nuclear envelope, and plasmodesmata. The tubular domain of the ER gives rise to oil and protein bodies, as well as protein precursor vesicles, and precursor accumulation vesicles (Herman, 2008). Also, ER is the most widespread organelle in the cell that comes into contact, and exchanges content, with several other organelles such as chloroplasts, Golgi, and plasma membrane (Stefano *et al.*, 2014). Other features of the ER such as presence of chaperone molecules, the favorable environment for disulfide bond formation and the absence of significant protease activity makes it an ideal organelle for protein expression and accumulation (Khan *et al.*, 2012).

1.2.4.3.1 Oil bodies

The major form of seed lipid storage is in oil bodies (OBs). Seeds store triacylglycerols (TAGs) as a carbon source to use during germination and post-germination growth in OBs. OBs are spherical organelles ranging between 0.5 to 1 µm and are composed of a core of TAGs surrounded by a layer of phospholipids and structural proteins known as oleosins (Hsieh and Huang, 2004).

To target recombinant proteins to OBs, oleosin fusions have been used which accumulate their fused protein on the membrane of OBs. Recombinant proteins can subsequently be recovered using a liquid-liquid phase separation by centrifugation and then released by endoprotease digestion (Boothe *et al.*, 2010). This approach was used for production of Apolipoprotein A1 (Nykiforuk *et al.*, 2011) and human insulin (Fischer *et al.*, 2012) in safflower (*Carthamus tinctorius*) seeds. OB targeting of recombinant proteins is an appealing strategy because of the relatively low cost of purification; however, this

technology has several drawbacks: it is limited to seeds and cannot be used in leafy crops, protein accumulation levels in seed are still not high enough to meet the economical production requirements, and the efficacy of oleosin-based separation and the post-recovery proteolytic cleavage have been challenged by several studies (Boothe *et al.*, 2010; Conley *et al.*, 2011a; Kuhnel *et al.*, 2003).

1.2.4.3.2 Protein storage vacuoles

During seed development, storage proteins accumulate in protein storage vacuoles (PSVs) which will be degraded and used by the embryo during germination. PSVs are composed of three distinct regions: the matrix, crystalloid and globoid (Jiang *et al.*, 2001). Proteins can be trafficked to PSVs through different pathways depending on the tissue (e.g. embryonic or endosperm) or plant species. These pathways include trafficking through the trans-Golgi network (TGN), precursor-accumulating vesicles (PAC), multi-vesicular bodies (MVB) or pre-vacuolar compartment (PVC) (Hara-Nishimura *et al.*, 1998; Hohl *et al.*, 1996). Interestingly, a recent study by Feeney *et al.* (Feeney *et al.*, 2013) has shown the possibility of induction of PSVs in *Arabidopsis* leaves. PSVs appeared in the vegetative tissue upon overexpression of a key transcriptional regulator of seed development gene, *LEAFY COTYLEDON2* (LEC2).

Seed PSVs have been used for accumulation of several recombinant proteins in a wide range of plants; *Aspergillus niger* phytase in pea (*Pisum sativum*) (Drakakaki *et al.*, 2006), human coagulation factor IX (hFIX) in soybean (*Glycine max*) (Cunha *et al.*, 2011), human acid β-glucosidase (GCase) in tobacco (*Nicotiana tabacum*) (Reggi *et al.*, 2005).

1.2.4.3.3 Protein bodies

Protein bodies can be found in different forms in nature. In bacteria they are known as inclusion bodies (IBs) which are essentially aggregates that form upon high expression levels of recombinant proteins in their insoluble state and are found in *Escherichia coli* (Mayer and Buchner, 2004; Williams *et al.*, 1982). Insoluble aggregates usually form by over-expression of eukaryotic proteins in prokaryotes. This is because prokaryotic proteins usually fold post-translationally, whereas eukaryotic proteins require co-

translational folding. Misfolded proteins initiate a core which grows concentrically by addition of similar misfolded proteins which will eventually form an equilibrium with the soluble form of the protein (Baneyx and Mujacic, 2004). Although the content of IBs can contain up to 95% of a single protein in some cases, they usually include a mixture of unfolded and folded molecules (García-Fruitós *et al.*, 2005).

More than 100 years ago, William Russell, a Scottish physician, noticed and reported unusual protein deposits in mammalian epithelial cells which he originally named fuchsine bodies (Russell, 1890). These structures are now known as Russell bodies (RBs) and usually form due to accumulation of a certain protein within the ER, for instance a mutated immunoglobulin that cannot be degraded or secreted due to excessive synthesis. It is suggested that RB formation is a defense mechanism by the cell against high amounts of non-transportable molecules that exceed the degradation capacity of the cell in which the excess molecules are encapsulated by an ER membrane and stored (Mattioli *et al.*, 2006).

Protein bodies can be found in a similar form in plant seeds as well. Seeds as the major plant storage tissue contain several classes of storage proteins; albumins (water soluble), globulins (dilute saline soluble), prolamins (alcohol soluble) and glutelins (dilute acid or base soluble) (Osborne, 1919). Prolamins are the major class of proteins found in cereals such as maize, rice and wheat (Shewry *et al.*, 1995), and are generally deposited as dense accretions in the ER termed protein bodies (PBs) (Larkins and Hurkman, 1978; Pompa and Vitale, 2006). PBs generally form within the ER lumen but they may bud off upon reaching a sufficient size and remain in the cytosol or may end up in PSVs by autophagy (Herman and Larkins, 1999; Levanony *et al.*, 1992).

Because of their ER origin, and their strong role in protein storage, PBs are desirable organelles to target and store recombinant proteins. Therefore, several attempts have been made to target recombinant proteins to seed PBs by exploiting motifs that promote protein aggregation or the ER retrieval signal (KDEL) (Arcalis *et al.*, 2004; Takagi *et al.*, 2010; Takaiwa *et al.*, 2009; Takaiwa *et al.*, 2007).

1.2.4.3.4 Induced protein bodies

Several studies have shown the positive impact of protein fusion tags on the accumulation of recombinant proteins. A few of these fusion tags share another feature, which is the induction of protein bodies *in vivo* and include Zera®, elastin-like polypeptide (ELP) and hydrophobin-I (HFBI) (Conley *et al.*, 2011a; Schmidt, 2013).

1.2.4.3.4.1 Zera fusion tag

Zera® (developed by ERA Biotech, Barcelona, Spain) is the N-terminal proline-rich domain of γ-zein, a maize prolamin. The Zera sequence is composed of 112 amino acids which includes the γ-zein signal peptide, the non-proline region containing a Cys-Gly-Cys motif, the proline-rich repeat region containing $(PPVHL)_8$, and the Pro-X sequence containing four cysteines (Torrent *et al.*, 2009b). Zera-fused proteins accumulate in the ER regardless of the presence of an ER retrieval signal peptide (HDEL or KDEL). It was shown that the two N-terminal cysteine residues in the Cys-Gly-Cys motif are required for the oligomerization of Zera molecules through intermolecular disulfide bonds (the first step in PB formation), and that the hexapeptide repeats are critical in PB formation through facilitating the lateral protein-protein interaction and alignment of Zera molecules (Llop-Tous *et al.*, 2010). Zera PBs were shown to be surrounded with membrane-bound ribosomes indicating their rough-ER origin (Llop-Tous *et al.*, 2010). Zera was shown to induce PBs in several expression systems including mammalian cells (Chinese hamster ovary (CHO) cells), fungal cells (*Trichoderma reesei*), insect cells (*Spodoptera frugiperda*) and plant leaf cells (*Nicotiana tabacum*, and *Nicotiana benthamiana*) (Torrent *et al.*, 2009a).

Zera-induced PBs can be purified with a density based centrifugation method. Zera-fusions have been used for production and purification of several recombinant proteins including epidermal growth factor, human growth hormone, and *Streptomyces* derived xylanases (Llop-Tous *et al.*, 2010; Llop-Tous *et al.*, 2011).

1.2.4.3.4.2 Elastin-like polypeptide fusion tag

Elastin-like polypeptide (ELP) is composed of VPGXG pentapeptides originally identified in the mammalian protein elastin. The guest amino acid (X) can be any amino acid except proline (Urry, 1988). ELPs are thermally responsive molecules which undergo a reversible phase transition from soluble protein to insoluble aggregates when heated above their transition temperature (T_t). This property of ELPs can be transferred to their fused protein partner, and used for the rapid, non-chromatographic purification method known as inverse transition cycling (ITC) (Meyer and Chilkoti, 1999). The number of ELP pentapeptides repeats can vary based on the experimental design and purification preferences. Fewer ELP repeats are more beneficial for protein accumulation while larger tags are more efficient for purification purposes. An ELP size of 30-40 pentapeptide repeats was found to be a good compromise for both accumulation and purification (Conley *et al.*, 2009a).

ELP fusions have been used to increase accumulation levels of several recombinant proteins including spider silk proteins (Patel *et al.*, 2007; Scheller *et al.*, 2004), murine interleukin-4 (Patel *et al.*, 2007), human interleukin-10 (Kaldis *et al.*, 2013), and anti-HIV antibody 2F5 (Floss *et al.*, 2008). Study of the GFP-ELP fusion showed a significant increase in the amount of GFP (up to 11% of TSP), and also showed the formation of PBs in *N. benthamiana* leaves. GFP-ELP-induced PBs are highly mobile organelles which move along the actin cytoskeleton of the cell. GFP-ELP PBs contain other ER resident proteins such as BiP and are surrounded with a membrane studded with ribosomes, both of which may indicate an ER origin for these PBs (Conley *et al.*, 2009b).

1.2.4.3.4.3 Hydrophobin fusion tag

Hydrophobins are a family of small surface-active proteins produced by filamentous fungi. In fungi, different hydrophobins are expressed at different stages of fungal life cycle with a wide range of biological roles such as assembly into fungal cell walls, covering fungal spores or coating fungal surfaces (Linder, 2009). The overall shape of hydrophobin molecules is globular with an exposed hydrophobic patch which in theory destabilizes its structure. To challenge destabilization effects of the hydrophobic patch,

all hydrophobins contain eight cysteine residues in their sequence which form four intramolecular disulfide bridges responsible for the structural stability of the molecule. The hydrophobic patch enables hydrophobins to self-assemble at hydrophilic-hydrophobic interfaces (Wösten and de Vocht, 2000). Hydrophobins can transfer this property to their fusion partners which can be used for purification using a surfactant based aqueous two-phase separation system (ATPS) (Linder, 2009).

Hydrophobin-I (HFBI) fusion tag was used to increase the accumulation levels of glucose oxidase (GOx), an enzyme that could not be expressed in other conventional expression systems (Bankar *et al.*, 2009). HFBI fusion also increased the accumulation levels of GFP up to 51% of TSP when transiently expressed in *N. benthamiana* leaves (Joensuu *et al.*, 2010), and induced the formation of PBs which cluster together (Conley *et al.*, 2011b; Joensuu *et al.*, 2010). Similar to Zera- and ELP-induced PBs, HFBI PBs are surrounded with a distinct ribosome-studded membrane. GFP-HFBI-induced PBs are mobile and move around the cell similar to GFP-ELP PBs.

1.3 Research goals and objectives

Scalability, rapid regeneration, high production rates, ability to perform co- and post-translational modifications, and biosafety are some of the numerous advantages of plant expression systems. Regardless of the fact that plant biotechnology is still a young and growing technology, it has already been used to produce a wide range of vaccines, biopharmaceuticals and industrial enzymes. Nevertheless, much room still exists for significant advances to emerge in the near future. Two of the major challenges hindering the large scale production of recombinant proteins in plants are low accumulation levels and lack of efficient purification methods.

PB formation by Zera, ELP and HFBI fusion tags might be one solution to increase low levels of recombinant protein accumulation in plants and to facilitate protein purification at the same time. However, not much is known about PBs. The subcellular origin of PBs as well as the mechanism by which the fusion tags induce the formation of PBs and consequently increase the accumulation levels of recombinant proteins are not well understood and form the basis of this thesis. The specific objectives of my project were:

1- To study the biogenesis of PBs. The effects of recombinant protein concentration (accumulation levels), subcellular targeting of the recombinant protein, and the role of fusion tags were examined in the process of PB formation.

2- To study the role of PB size in increasing accumulation levels of recombinant proteins. The size of PBs induced by ELP and HFBI fusion tags were measured and correlated with their respective accumulation levels.

3- To study the means of sequestration of recombinant proteins into PBs. The role of fusion-tag-induced PBs in increasing the accumulation levels of unfused recombinant proteins, erythropoietin and interleukin-10, were assessed.

4- To address the question whether PBs are terminally stored cytoplasmic organelles or if they remain as part of the ER. A GFP photoconversion method was developed which enabled the tracking of photoconverted molecule movement between PBs and through the ER.

5- To confirm the ER origin of PBs induced by Zera, ELP and HFBI fusion tags. The ER origin of PBs was examined by co-expression with an ER-transmembrane protein.

1.4 References

Andow, D.A. and Hilbeck, A. (2004) Science-based risk assessment for nontarget effects of transgenic crops. *BioScience* **54**, 637-649.

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

Baneyx, F. and Mujacic, M. (2004) Recombinant protein folding and misfolding in Escherichia coli. *Nat. Biotechnol.* **22**, 1399-1408.

Bankar, S.B., Bule, M.V., Singhal, R.S. and Ananthanarayan, L. (2009) Glucose oxidase — An overview. *Biotechnol. Adv.* **27**, 489-501.

Belanger, F.C. and Kriz, A.L. (1989) Molecular characterization of the major maize embryo globulin encoded by the glb1 gene. *Plant Physiol.* **91**, 636-643.

Benchabane, M., Goulet, C., Rivard, D., Faye, L., Gomord, V. and Michaud, D. (2008) Preventing unintended proteolysis in plant protein biofactories. *Plant Biotechnol. J.* **6**, 633-648.

Boothe, J., Nykiforuk, C., Shen, Y., Zaplachinski, S., Szarka, S., Kuhlman, P., Murray, E., Morck, D. and Moloney, M.M. (2010) Seed-based expression systems for plant molecular farming. *Plant Biotechnol. J.* **8**, 588-606.

Breyer, D., De Schrijver, A., Goossens, M., Pauwels, K. and Herman, P. (2012) Biosafety of Molecular Farming in Genetically Modified Plants. In: *Molecular Farming in Plants: Recent Advances and Future Prospects* (Wang, A. and Ma, S. eds), pp. 259-274. Springer Netherlands.

Chebolu, S. and Daniell, H. (2009) Chloroplast-derived vaccine antigens and biopharmaceuticals: expression, folding, assembly and functionality. *Curr. Top. Microbiol. Immunol.* **332**, 33-54.

Christensen, A., Sharrock, R. and Quail, P. (1992) Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol. Biol.* **18**, 675-689.

Clarke, J.L. and Daniell, H. (2011) Plastid biotechnology for crop production: present status and future perspectives. *Plant Mol. Biol.* **76**, 211-220.

Comai, L., Moran, P. and Maslyar, D. (1990) Novel and useful properties of a chimeric plant promoter combining CaMV 35S and MAS elements. *Plant Mol. Biol.* **15**, 373-381.

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

Conley, A.J., Joensuu, J.J., Menassa, R. and Brandle, J.E. (2009b) Induction of protein body formation in plant leaves by elastin-like polypeptide fusions. *BMC Biol.* **7**, 48.

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

Conley, A.J., Mohib, K., Jevnikar, A.M. and Brandle, J.E. (2009c) Plant recombinant erythropoietin attenuates inflammatory kidney cell injury. *Plant Biotechnol. J.* **7**, 183-199.

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

Cunha, N.B., Murad, A.M., Ramos, G.L., Maranhão, A.Q., Brígido, M.M., Araújo, A.C.G., Lacorte, C., Aragão, F.J., Covas, D.T. and Fontes, A.M. (2011) Accumulation of functional recombinant human coagulation factor IX in transgenic soybean seeds. *Transgenic Res.* **20**, 841-855.

D'Aoust, M., Couture, M.M., Charland, N., Trépanier, S., Landry, N., Ors, F. and Vézina, L. (2010) The production of hemagglutinin-based virus-like particles in plants: A rapid, efficient and safe response to pandemic influenza. *Plant Biotechnol. J.* **8**, 607-619.

Dalmay, T., Hamilton, A., Mueller, E. and Baulcombe, D.C. (2000) Potato virus X Amplicons in Arabidopsis mediate genetic and epigenetic gene silencing. *Plant Cell* **12**, 369-379.

De Muynck, B., Navarre, C., Nizet, Y., Stadlmann, J. and Boutry, M. (2009) Different subcellular localization and glycosylation for a functional antibody expressed in *Nicotiana tabacum* plants and suspension cells. *Transgenic Res.* **18**, 467-482.

Desai, P.N., Shrivastava, N. and Padh, H. (2010) Production of heterologous proteins in plants: strategies for optimal expression. *Biotechnol. Adv.* **28**, 427-435.

Drakakaki, G., Marcel, S., Arcalis, E., Altmann, F., Gonzalez-Melendi, P., Fischer, R., Christou, P. and Stoger, E. (2006) The intracellular fate of a recombinant protein is tissue dependent. *Plant Physiol* **141**, 578-586.

Egelkrout, E., Rajan, V. and Howard, J.A. (2012) Overproduction of recombinant proteins in plants. *Plant Sci.* **184**, 83-101.

Faye, L., Boulaflous, A., Benchabane, M., Gomord, V. and Michaud, D. (2005) Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. *Vaccine* **23**, 1770-1778.

Feeney, M., Frigerio, L., Cui, Y. and Menassa, R. (2013) Following Vegetative to Embryonic Cellular Changes in Leaves of Arabidopsis Overexpressing LEAFY COTYLEDON2. *Plant Physiol.* **162**, 1881-1896.

Fischer, R., Schillberg, S., Hellwig, S., Twyman, R.M. and Drossard, J. (2012) GMP issues for recombinant plant-derived pharmaceutical proteins. *Biotechnol. Adv.* **30**, 434-439.

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

Garabagi, F., Gilbert, E., Loos, A., McLean, M.D. and Hall, J.C. (2012) Utility of the P19 suppressor of gene-silencing protein for production of therapeutic antibodies in *Nicotiana expression* hosts. *Plant Biotechnol. J.* **10**, 1118-1128.

García-Fruitós, E., González-Montalbán, N., Morell, M., Vera, A., Ferraz, R.M., Arís, A., Ventura, S. and Villaverde, A. (2005) Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. *Microb. Cell Fact.* **4**, 27.

Gelvin, S.B. (2003) *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiol Mol Biol Rev* **67**, 16-37, table of contents.

Gleba, Y., Klimyuk, V. and Marillonnet, S. (2005) Magnifection—a new platform for expressing recombinant vaccines in plants. *Vaccine* **23**, 2042-2048.

Gleba, Y., Klimyuk, V. and Marillonnet, S. (2007) Viral vectors for the expression of proteins in plants. *Curr. Opin. Biotechnol.* **18**, 134-141.

Goulet, C., Benchabane, M., Anguenot, R., Brunelle, F., Khalf, M. and Michaud, D. (2010) A companion protease inhibitor for the protection of cytosol-targeted recombinant proteins in plants. *Plant Biotechnol. J.* **8**, 142-154.

Goulet, C., Khalf, M., Sainsbury, F., D'Aoust, M.A. and Michaud, D. (2012) A protease activity-depleted environment for heterologous proteins migrating towards the leaf cell apoplast. *Plant Biotechnol. J.* **10**, 83-94.

Goulet, C., Michaud, D. and Teixeira da Silva, J. (2006) Degradation and stabilization of recombinant proteins in plants. *Floriculture, ornamental and plant biotechnology*, 35-40.

Gutiérrez, S. and Menassa, R. (2014) Protein body inducing fusions for recombinant protein production in plants. In: *Plant-derived pharmaceuticals: principles and applications for developing countries* (Hefferon, K. ed) pp. 9-19. Wallingford, UK: CABI.

Gutiérrez, S.P., Saberianfar, R., Kohalmi, S.E. and Menassa, R. (2013) Protein body formation in stable transgenic tobacco expressing elastin-like polypeptide and hydrophobin fusion proteins. *BMC Biotechnol.* **13**, 40.

Hara-Nishimura, I., Shimada, T., Hatano, K., Takeuchi, Y. and Nishimura, M. (1998) Transport of Storage Proteins to Protein Storage Vacuoles Is Mediated by Large Precursor-Accumulating Vesicles. *Plant Cell.* **10**, 825-836.

Hefferon, K. (2013) Plant-derived pharmaceuticals for the developing world. *Biotechnol. J.* **8**, 1193-1202.

Hellwig, S., Drossard, J., Twyman, R.M. and Fischer, R. (2004) Plant cell cultures for the production of recombinant proteins. *Nat. Biotechnol.* **22**, 1415-1422.

Herman, E. M. (2008) Endoplasmic reticulum bodies: solving the insoluble. *Curr. Opin. Plant Biol.***11**:672-679.

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

Hohl, I., Robinson, D.G., Chrispeels, M.J. and Hinz, G. (1996) Transport of storage proteins to the vacuole is mediated by vesicles without a clathrin coat. *Journal of Cell Science* **109**, 2539-2550.

Hood, E.E., Bailey, M.R., Beifuss, K., Magallanes-Lundback, M., Horn, M.E., Callaway, E., Drees, C., Delaney, D.E., Clough, R. and Howard, J.A. (2003) Criteria for high-level expression of a fungal laccase gene in transgenic maize. *Plant Biotechnol. J.* **1**, 129-140.

Hsieh, K. and Huang, A.H. (2004) Endoplasmic reticulum, oleosins, and oils in seeds and tapetum cells. *Plant Physiol.* **136**, 3427-3434.

Huang, T.K., Plesha, M.A., Falk, B.W., Dandekar, A.M. and McDonald, K.A. (2009) Bioreactor strategies for improving production yield and functionality of a recombinant human protein in transgenic tobacco cell cultures. *Biotechnol. Bioeng.* **102**, 508-520.

Jiang, L., Phillips, T.E., Hamm, C.A., Drozdowicz, Y.M., Rea, P.A., Maeshima, M., Rogers, S.W. and Rogers, J.C. (2001) The protein storage vacuole a unique compound organelle. *J. Cell Biol.* **155**, 991-1002.

Jiang, X., Oohira, K., Iwasaki, Y., Nakano, H., Ichihara, S. and Yamane, T. (2002) Reduction of protein degradation by use of protease-deficient mutants in cell-free protein synthesis system of *Escherichia coli*. *J. Biosci. Bioeng.* **93**, 151-156.

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

Joensuu, J.J., Kotiaho, M., Teeri, T.H., Valmu, L., Nuutila, A.M., Oksman-Caldentey, K.M. and Niklander-Teeri, V. (2006) Glycosylated F4 (K88) fimbrial adhesin FaeG expressed in barley endosperm induces ETEC-neutralizing antibodies in mice. *Transgenic Res.* **15**, 359-373.

Kaldis, A., Ahmad, A., Reid, A., McGarvey, B., Brandle, J., Ma, S., Jevnikar, A., Kohalmi, S.E. and Menassa, R. (2013) High-level production of human interleukin-10 fusions in tobacco cell suspension cultures. *Plant Biotechnol. J.* **11**, 535-545.

Kang, T.J., Loc, N.H., Jang, M.O. and Yang, M.S. (2004) Modification of the cholera toxin B subunit coding sequence to enhance expression in plants. *Mol Breed* **13**, 143-153.

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

Kay, R., Chan, A.M.Y., Daly, M. and McPherson, J. (1987) Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Science (New York, NY)* **236**, 1299.

Khan, I., Twyman, R.M., Arcalis, E. and Stoger, E. (2012) Using storage organelles for the accumulation and encapsulation of recombinant proteins. *Biotechnology J.* **7**, 1099-1108.

Kuhnel, B., Alcantara, J., Boothe, J., van Rooijen, G. and Moloney, M. (2003) Precise and efficient cleavage of recombinant fusion proteins using mammalian aspartic proteases. *Protein Eng.* **16**, 777-783.

Laguia-Becher, M., Martin, V., Kraemer, M., Corigliano, M., Yacono, M.L., Goldman, A. and Clemente, M. (2010) Effect of codon optimization and subcellular targeting on *Toxoplasma gondii* antigen SAG1 expression in tobacco leaves to use in subcutaneous and oral immunization in mice. *BMC Biotechnol.* **10**, 52.

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

Levanony, H., Rubin, R., Altschuler, Y. and Galili, G. (1992) Evidence for a novel route of wheat storage proteins to vacuoles. *J. Cell Biol.* **119**, 1117-1128.

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

Liu, W.X., Liu, H.L., Chai, Z.J., Xu, X.P., Song, Y.R. and Qu, L.Q. (2010) Evaluation of seed storage-protein gene 5′ untranslated regions in enhancing gene expression in transgenic rice seed. *Theor. Appl. Genet.* **121**, 1267-1274.

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

Llop-Tous, I., Ortiz, M., Torrent, M. and Ludevid, M.D. (2011) The expression of a xylanase targeted to ER-protein bodies provides a simple strategy to produce active insoluble enzyme polymers in tobacco plants. *PLoS ONE* **6**, e19474.

Lossl, A.G. and Waheed, M.T. (2011) Chloroplast-derived vaccines against human diseases: achievements, challenges and scopes. *Plant Biotechnol. J.* **9**, 527-539.

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

Macauley-Patrick, S., Fazenda, M.L., McNeil, B. and Harvey, L.M. (2005) Heterologous protein production using the Pichia pastoris expression system. *Yeast* **22**, 249-270.

Maliga, P. and Bock, R. (2011) Plastid biotechnology: food, fuel, and medicine for the 21st century. *Plant Physiol.* **155**, 1501-1510.

Marillonnet, S., Thoeringer, C., Kandzia, R., Klimyuk, V. and Gleba, Y. (2005) Systemic Agrobacterium tumefaciens-mediated transfection of viral replicons for efficient transient expression in plants. *Nat. Biotechnol.* **23**, 718-723.

Mattioli, L., Anelli, T., Fagioli, C., Tacchetti, C., Sitia, R. and Valetti, C. (2006) ER storage diseases: a role for ERGIC-53 in controlling the formation and shape of Russell bodies. *J. Cell Sci.* **119**, 2532-2541.

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

Mayer, M. and Buchner, J. (2004) Refolding of Inclusion Body Proteins. In: *Molecular Diagnosis of Infectious Diseases* (Decler, J. and Reischl, U. eds), pp. 239-254. Humana Press.

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

New, J., Westerveld, D. and Daniell, H. (2012) Chloroplast-derived therapeutic and prophylactic vaccines. In: *Molecular Farming in Plants: Recent Advances and Future Prospects* (Wang, A. and Ma, S. eds), pp. 69-87. Springer Netherlands.

Nie, L., Wu, G. and Zhang, W. (2006) Correlation between mRNA and protein abundance in Desulfovibrio vulgaris: A multiple regression to identify sources of variations. *Biochem. Biophys. Res. Commun.* **339**, 603-610.

Nuttall, J., Vine, N., Hadlington, J.L., Drake, P., Frigerio, L. and Ma, J.K. (2002) ER-resident chaperone interactions with recombinant antibodies in transgenic plants. *Eur J Biochem* **269**, 6042-6051.

Nykiforuk, C.L., Shen, Y., Murray, E.W., Boothe, J.G., Busseuil, D., Rhéaume, E., Tardif, J.C., Reid, A. and Moloney, M.M. (2011) Expression and recovery of biologically active recombinant Apolipoprotein AIMilano from transgenic safflower (*Carthamus tinctorius*) seeds. *Plant Biotechnol. J.* **9**, 250-263.

Odell, J.T., Nagy, F. and Chua, N.H. (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature* **313**, 810-812.

Osborne, T.B. (1919) *The vegetable proteins*:Longmans, Green, and Company.

Palauqui, J.-C., Elmayan, T., Pollien, J.-M. and Vaucheret, H. (1997) Systemic acquired silencing: transgene-specific post-transcriptional silencing is transmitted by grafting from silenced stocks to non-silenced scions. *EMBO J.* **16**, 4738-4745.

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

Pereira, E.O., Kolotilin, I., Conley, A.J. and Menassa, R. (2014) Production and characterization of in planta transiently produced polygalacturanase from *Aspergillus niger* and its fusions with hydrophobin or ELP tags. *BMC Biotechnol.* **14**, 59.

Peyret, H. and Lomonossoff, G.P. (2013) The pEAQ vector series: the easy and quick way to produce recombinant proteins in plants. *Plant Mol. Biol.* **83**, 51-58.

Pogue, G.P., Vojdani, F., Palmer, K.E., Hiatt, E., Hume, S., Phelps, J., Long, L., Bohorova, N., Kim, D., Pauly, M., Velasco, J., Whaley, K., Zeitlin, L., Garger, S.J., White, E., Bai, Y., Haydon, H. and Bratcher, B. (2010) Production of pharmaceutical-grade recombinant aprotinin and a monoclonal antibody product using plant-based transient expression systems. *Plant Biotechnol. J.* **8**, 638-654.

Pompa, A. and Vitale, A. (2006) Retention of a bean phaseolin/maize γ-zein fusion in the endoplasmic reticulum depends on disulfide bond formation. *Plant Cell* **18**, 2608-2621.

Pooggin, M.M. and Skryabin, K.G. (1992) The 5′-untranslated leader sequence of potato virus X RNA enhances the expression of a heterologous gene *in vivo*. *Molecular and General Genetics MGG* **234**, 329-331.

Puigbò, P., Guzmán, E., Romeu, A. and Garcia-Vallvé, S. (2007) OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. *Nucleic Acids Res.* **35**, W126-W131.

Qiu, X., Wong, G., Audet, J., Bello, A., Fernando, L., Alimonti, J.B., Fausther-Bovendo, H., Wei, H., Aviles, J. and Hiatt, E. (2014) Reversion of advanced Ebola virus disease in nonhuman primates with ZMapp. *Nature*. doi:10.1038/nature13777

Qu, F. and Morris, T.J. (2005) Suppressors of RNA silencing encoded by plant viruses and their role in viral infections. *FEBS Lett.* **579**, 5958-5964.

Qu, L.Q., Xing, Y.P., Liu, W.X., Xu, X.P. and Song, Y.R. (2008) Expression pattern and activity of six glutelin gene promoters in transgenic rice. *J. Exp. Bot.* **59**, 2417-2424.

Rao, A.Q., Bakhsh, A., Kiani, S., Shahzad, K., Shahid, A.A., Husnain, T. and Riazuddin, S. (2009) The myth of plant transformation. *Biotechnol. Adv.* **27**, 753-763.

Reggi, S., Marchetti, S., Patti, T., De Amicis, F., Cariati, R., Bembi, B. and Fogher, C. (2005) Recombinant human acid *beta*-glucosidase stored in tobacco seed is stable, active and taken up by human fibroblasts. *Plant Mol. Biol.* **57**, 101-113.

Ruf, S., Karcher, D. and Bock, R. (2007) Determining the transgene containment level provided by chloroplast transformation. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 6998-7002.

Russell, W. (1890) An address on a characteristic organism of cancer. *Br. Med. J.* **2**, 1356.

Sainsbury, F., Thuenemann, E.C. and Lomonossoff, G.P. (2009) pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnol. J.* **7**, 682-693.

Sanford, J.C., Klein, T.M., Wolf, E.D. and Allen, N. (1987) Delivery of substances into cells and tissue using particle bombardment process. *Particul. Sci. Technol.* **5**, 27-37.

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

Schiermeyer, A., Schinkel, H., Apel, S., Fischer, R. and Schillberg, S. (2005) Production of *Desmodus rotundus* salivary plasminogen activator alpha1 (DSPAalpha1) in tobacco is hampered by proteolysis. *Biotechnol. Bioeng.* **89**, 848-858.

Schmidt, S. (2013) Protein bodies in nature and biotechnology. *Mol. Biotechnol.* **54**, 257-268.

Shaw, C.H., Carter, G.H. and Watson, M.D. (1984) A functional map of the nopaline synthase promoter. *Nucleic Acids Res.* **12**, 7831-7846.

Shewry, P.R., Napier, J.A. and Tatham, A.S. (1995) Seed storage proteins: Structures and biosynthesis. *Plant Cell* **7**, 945-956.

Shrawat, A.K. and Lörz, H. (2006) *Agrobacterium*-mediated transformation of cereals: a promising approach crossing barriers. *Plant Biotechnol. J.* **4**, 575-603.

Sleat, D.E., Gallic, D.R., Jefferson, R.A., Bevan, M.W., Turner, P.C. and Wilson, T.M.A. (1987) Characterisation of the 5′-leader sequence of tobacco mosaic virus RNA as a general enhancer of translation in vitro. *Gene* **60**, 217-225.

Stefano, G., Hawes, C. and Brandizzi, F. (2014) ER – the key to the highway. *Curr. Opin. Plant Biol.* **22**, 30-38.

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

Stoger, E., Sack, M., Perrin, Y., Vaquero, C., Torres, E., Twyman, R.M., Christou, P. and Fischer, R. (2002) Practical considerations for pharmaceutical antibody production in different crop systems. *Mol. Breed.* **9**, 149-158.

Strauss, S. (2014) Biotech drugs too little, too late for Ebola outbreak. *Nat. Biotechnol.* **32**, 849-850.

Streatfield, S.J., Bray, J., Love, R.T., Horn, M.E., Lane, J.R., Drees, C.F., Egelkrout, E.M. and Howard, J.A. (2010) Identification of maize embryo-preferred promoters suitable for high-level heterologous protein production. *GM crops* **1**, 162-172.

Sugio, T., Satoh, J., Matsuura, H., Shinmyo, A. and Kato, K. (2008) The 5′-untranslated region of the *Oryza sativa* alcohol dehydrogenase gene functions as a translational enhancer in monocotyledonous plant cells. *J. Biosci. Bioeng.* **105**, 300-302.

Svab, Z., Hajdukiewicz, P. and Maliga, P. (1990) Stable transformation of plastids in higher plants. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 8526-8530.

Takagi, H., Hiroi, T., Hirose, S., Yang, L. and Takaiwa, F. (2010) Rice seed ER-derived protein body as an efficient delivery vehicle for oral tolerogenic peptides. *Peptides* **31**, 1421-1425.

Takaiwa, F., Hirose, S., Takagi, H., Yang, L. and Wakasa, Y. (2009) Deposition of a recombinant peptide in ER-derived protein bodies by retention with cysteine-rich prolamins in transgenic rice seed. *Planta* **229**, 1147-1158.

Takaiwa, F., Takagi, H., Hirose, S. and Wakasa, Y. (2007) Endosperm tissue is good production platform for artificial recombinant proteins in transgenic rice. *Plant Biotechnol. J.* **5**, 84-92.

Torrent, M., Llompart, B., Lasserre-Ramassamy, S., Llop-Tous, I., Bastida, M., Marzabal, P., Westerholm-Parvinen, A., Saloheimo, M., Heifetz, P.B. and Ludevid, M.D. (2009a) Eukaryotic protein production in designed storage organelles. *BMC Biol.* **7**, 5.

Torrent, M., Llop-Tous, I. and Ludevid, M.D. (2009b) Protein body induction: a new tool to produce and recover recombinant proteins in plants. *Methods Mol. Biol.* **483**, 193-208.

Twyman, R.M., Schillberg, S. and Fischer, R. (2013) Optimizing the yield of recombinant pharmaceutical proteins in plants. *Curr. Pharm. Des.* **19**, 5486-5494.

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.

Van Molle, I., Joensuu, J.J., Buts, L., Panjikar, S., Kotiaho, M., Bouckaert, J., Wyns, L., Niklander-Teeri, V. and De Greve, H. (2007) Chloroplasts assemble the major subunit FaeG of *Escherichia coli* F4 (K88) fimbriae to strand-swapped dimers. *J. Mol. Biol.* **368**, 791-799.

Verma, D. and Daniell, H. (2007) Chloroplast vector systems for biotechnology applications. *Plant Physiol.* **145**, 1129-1143.

Vézina, L.P., Faye, L., Lerouge, P., D'Aoust, M.A., Marquet-Blouin, E., Burel, C., Lavoie, P.O., Bardor, M. and Gomord, V. (2009) Transient co-expression for fast and high-yield production of antibodies with human-like N-glycans in plants. *Plant Biotechnol. J.* **7**, 442-455.

Villalobos, A., Ness, J., Gustafsson, C., Minshull, J. and Govindarajan, S. (2006) Gene Designer: a synthetic biology tool for constructing artificial DNA segments. *BMC Bioinformatics* **7**, 285.

Voinnet, O. (2001) RNA silencing as a plant immune system against viruses. *Trends Genet.* **17**, 449-459.

Voinnet, O. and Baulcombe, D.C. (1997) Systemic signalling in gene silencing. *Nature* **389**, 553.

Voinnet, O., Pinto, Y.M. and Baulcombe, D.C. (1999) Suppression of gene silencing: A general strategy used by diverse DNA and RNA viruses of plants. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 14147-14152.

Voinnet, O., Rivas, S., Mestre, P. and Baulcombe, D. (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant Journal* **33**, 949-956.

Williams, D.C., Van Frank, R.M., Muth, W.L. and Burnett, J.P. (1982) Cytoplasmic inclusion bodies in *Escherichia coli* producing biosynthetic human insulin proteins. *Science* **215**, 687-689.

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

Wroblewski, T., Tomczak, A. and 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. and 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.

Xu, J., Ge, X. and Dolan, M.C. (2011) Towards high-yield production of pharmaceutical proteins with plant cell suspension cultures. *Biotechnol. Adv.* **29**, 278-299.

Yusibov, V., Modelska, A., Steplewski, K., Agadjanyan, M., Weiner, D., Hooper, D.C. and Koprowski, H. (1997) Antigens produced in plants by infection with chimeric plant viruses immunize against rabies virus and HIV-1. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5784-5788.

Zambryski, P. (1988) Basic processes underlying *Agrobacterium*-mediated DNA transfer to plant cells. *Ann. Rev. Genet.* **22**, 1-30.

Zamore, P.D., Tuschl, T., Sharp, P.A. and Bartel, D.P. (2000) RNAi: Double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25-33.

Zhang, J., Martin, J.M., Beecher, B., Morris, C.F., Curtis Hannah, L. and Giroux, M.J. (2009) Seed-specific expression of the wheat puroindoline genes improves maize wet milling yields. *Plant Biotechnol. J.* **7**, 733-743.

Zhu, J., Oger, P.M., Schrammeijer, B., Hooykaas, P.J., Farrand, S.K. and Winans, S.C. (2000) The bases of crown gall tumorigenesis. *J Bacteriol* **182**, 3885-3895.

2 Protein body formation in leaves of *Nicotiana benthamiana*: a concentration dependent mechanism influenced by the presence of fusion tags

Reza Saberianfar[1, 2], Jussi J Joensuu[3], Andrew J Conley[3], Rima Menassa*[1, 2]

[1]Department of Biology, University of Western Ontario, London, ON, Canada

[2]Agriculture and Agri-Food Canada, London, ON, Canada

[3]VTT Technical Research Centre of Finland, Espoo, Finland

A version of this chapter is in revision at the Plant Biotechnology Journal.

2.1 Introduction

During the past two decades, plants have been extensively studied as efficient green bioreactors for production of recombinant proteins (Egelkrout *et al.*, 2012). In 2012, the first plant-produced pharmaceutical protein was approved by the US Food and Drug Administration (Maxmen, 2012). Several host plants and production systems including seed- and leaf-based systems have been used for producing recombinant proteins, but *Nicotiana benthamiana* has emerged as the best platform for scalable rapid recombinant protein production in transient expression systems (Tremblay *et al.*, 2010).

Seeds are naturally suited to produce and store high levels of proteins. High protein content, low protease activity and low water content make seeds very attractive for producing and accumulating target proteins (Benchabane *et al.*, 2008). Seeds have developed specialized protein storage organelles such as protein bodies (PBs), protein storage vacuoles (PSVs) and oil bodies (Khan *et al.*, 2012). Over the years, several plant-made pharmaceuticals and industrial recombinant proteins have been produced in plant seeds (Lau and Sun, 2009). However, seed-based expression systems face issues related to biocontainment due to the possibility of gene leakage into the environment through distribution of seed or pollen. As well, relatively long periods of time are required to generate transgenic plant lines and screen them for high expressing lines (Ma and Wang, 2012).

Unlike seed-based systems, leaves can be harvested before flowering therefore minimizing the risk of gene leakage to the environment. In addition, leaves produce high biomass yield along with high soluble protein content on a hectare (or surface) basis. Among leaf-based expression systems, *Agrobacterium*-mediated transient expression in leaves of *Nicotiana benthamiana*, a relative of tobacco, has become widely used as a rapid and high-yielding platform for production of high levels recombinant proteins. This is due to the flexibility of this system as it allows for the expression of multiple genes simultaneously, is amenable to transferring large genes (greater than 2 kb) and allows for rapid (within a week) production of recombinant proteins (Pogue *et al.*, 2010).

Although leaf-based expression systems have advantages for biocontainment and speed, the production yield of recombinant proteins is often low due to proteolytic degradation of foreign proteins (Benchabane *et al.*, 2008; Fischer *et al.*, 2004). Several fusion tags such as Zera®, elastin-like polypeptides (ELPs) and hydrophobins (HFBs) have been used to improve accumulation levels of recombinant fusion partners in leaves (Conley *et al.*, 2011). These three fusion tags were hypothesized to improve accumulation of recombinant proteins by sequestering them in novel organelles reminiscent of seed PBs that protect target proteins from proteolytic degradation. The use of these fusion tags were shown not only to induce PBs but also to facilitate the process of recombinant protein purification (Conley *et al.*, 2011; Khan *et al.*, 2012; Torrent *et al.*, 2009b).

Zera® is a proline-rich region at the N-terminal region of a maize storage protein known as γ-zein (Llop-Tous *et al.*, 2010). Proteins fused to Zera® were found in the form of PBs in a wide range of host cells including mammalian, insect, fungi and plants (Torrent *et al.*, 2009b). Zera-fused proteins accumulate in the endoplasmic reticulum (ER) independent of the conventional ER retrieval signal (HDEL/KDEL) and it is thought that the interaction of Zera® with the ER membrane induces the budding of PBs from the ER. The Zera-induced PBs can be purified by gradient density-centrifugation (Torrent *et al.*, 2009a).

ELPs are synthetic biopolymers of repeating pentapeptides, Val-Pro-Gly-Xaa-Gly, similar to repetitive pentapeptides of the mammalian protein elastin, in which the guest amino acid (Xaa) can be any amino acid except proline. ELP tags can be used for the rapid, non-chromatographic purification method called inverse transition cycling (ITC) (Urry and Parker, 2002). This method has been used to increase accumulation levels and to assist with the purification of several recombinant proteins (Floss *et al.*, 2010). An ER-retrieved GFP-ELP fusion was shown to induce the formation of PBs as well as boost accumulation levels of the fused recombinant protein both in transient and transgenic expression (Conley *et al.*, 2009b; Gutierrez *et al.*, 2013).

Hydrophobins are surface-active proteins originally found in fungi (Wessels, 1997). Hydrophobins share eight conserved cysteine residues responsible for the formation of

four intramolecular disulfide bridges which maintain the structural stability of the molecule. The rigid core enables hydrophobins to display a group of hydrophobic amino acids on their surface forming a “hydrophobic patch” responsible for their amphiphilic nature (Hakanpää *et al.*, 2004). Therefore, hydrophobins can self-assemble at hydrophilic-hydrophobic interfaces, a property that can be exploited for purification using a surfactant based aqueous two-phase separation system (ATPS) (Linder *et al.*, 2001). Hydrophobin-I (HFBI) fusion has been used to increase the accumulation levels of several recombinant proteins, including GFP, and to induce PB formation *in vivo* similar to ELP by transient and transgenic expression (Gutierrez *et al.*, 2013; Joensuu *et al.*, 2010; Lahtinen *et al.*, 2008).

In this study I compared ELP and HFBI fusion tags in the process of PB formation. I showed that PB formation is not exclusively promoted by the fusion tags and that protein accumulation level is a critical element involved in this process. My results indicate that PB formation is not a selective mechanism and proteins targeted to the secretory pathway can be passively sequestered into PBs. I have used this property of PBs as a novel approach to increase the accumulation levels of erythropoietin and human interleukin-10.

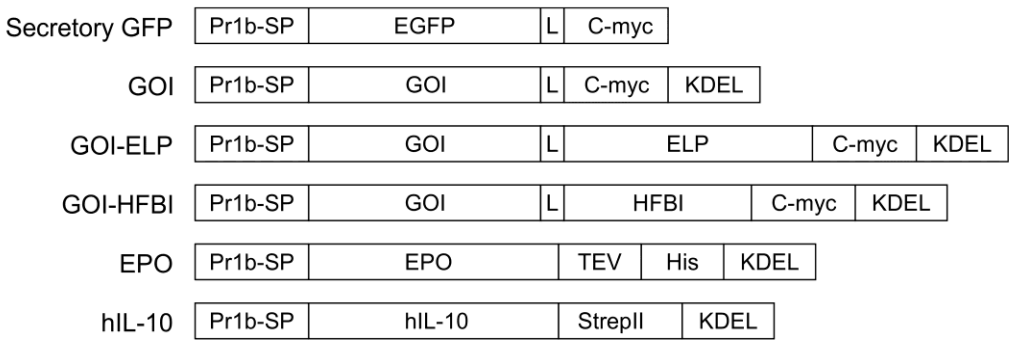
2.2 Results

2.2.1 Elastin-like polypeptides induce larger protein bodies than hydrophobin-I

To compare the effects of ELP and HFBI fusion tags on PB formation and accumulation levels of their fused proteins, constructs (Conley *et al.*, 2009b; Joensuu *et al.*, 2010) targeting GFP to the endomembrane system (Figure 2.1) were co-agroinfiltrated into *N. benthamiana* with a construct containing the suppressor of gene silencing, p19, from *Cymbidium ringspot* virus (CymRSV) (Silhavy *et al.*, 2002). Previous independent reports had shown that GFP fusion with ELP or HFBI increased the accumulation levels of the GFP and gave rise to clusters of PBs with variable distribution patterns (Conley *et al.*, 2009b; Joensuu *et al.*, 2010). To compare PBs induced by ELP or HFBI and consequent GFP accumulation levels, the sizes of PBs induced by GFP-ELP and GFP- HFBI

Figure 2. 1 Schematic representation of constructs used for *Agrobacterium*-mediated transient expression in *N. benthamiana* leaves.

Pr1b, secretory signal peptide from the tobacco pathogenesis-related protein 1; EGFP, enhanced green fluorescent protein; GOI, gene of interest; L, linker $(GGGS)_3$; TEV, tobacco etch virus protease recognition site; c-myc, detection tag; KDEL, endoplasmic reticulum retrieval tetra peptide; ELP, elastin like polypeptide; HFBI, hydrophobin-I; EPO, human erythropoietin; HIS, $(His)_6$ purification tag; hIL-10, human interleukin-10; Strep ll, purification tag. In GOI, GOIs include EGFP, xyn10A, or xyn11A. In GOI-ELP, GOIs include EGFP or RFP. In GOI-HFBI, GOIs include EGFP, RFP, or xyn11A. Schematic not drawn to scale.


Secretory GFP
Pr1b-SP
EGFP
L
C-myc
GOI
Pr1b-SP
GOI
L
C-myc
KDEL
GOI-ELP
Pr1b-SP
GOI
L
ELP
C-myc
KDEL
GOI-HFBI
Pr1b-SP
GOI
L
HFBI
C-myc
KDEL
EPO
Pr1b-SP
EPO
TEV
His
KDEL
hIL-10
Pr1b-SP
hIL-10
StrepII
KDEL

were measured and the amount of GFP produced in every treatment was quantified 3, 4 and 5 days post infiltration (dpi) (Figure 2.2). To obtain accurate measurements of all PBs within a cell, protoplasts were isolated from the infiltrated leaf tissue and used for imaging by confocal microscopy. Z-stack assemblies of 1.5 to 2 µm slices spanning entire protoplasts and covering approximately 70 to 80 µm were used for three-dimensional visualization. With this technique, PBs within the cell were visualized and their detailed information was analyzed using the Imaris software (Figure 2.2a-b) (Supplementary Movie 2.1).

Day to day comparison of ELP- and HFBI-induced PBs suggests that ELP-induced PBs were significantly larger in size than HFBI-induced PBs (Figure 2.2c). At 3 and 4 dpi, 75% of the ELP-induced PBs were larger than 1.7 µm in diameter with the median of approximately 2.4 µm, while 75% of the HFBI-induced PBs were smaller than 1.7 µm in diameter with the median of approximately 1.2 µm. At 5 dpi, the PBs were bigger in both the ELP and HFBI fusions, with 75% of the ELP-induced PBs larger than 2 µm in diameter with the median of 2.7 µm while 50% of the HFBI-induced PBs were also within this range with the median located at 2 µm. The maximum size of GFP-ELP-induced PBs increased from 5.5 µm at 3 dpi to 9 µm at 5 dpi while the maximum size of HFBI-induced PBs increased from 4 µm at 3 dpi to 7.7 µm at 5 dpi.

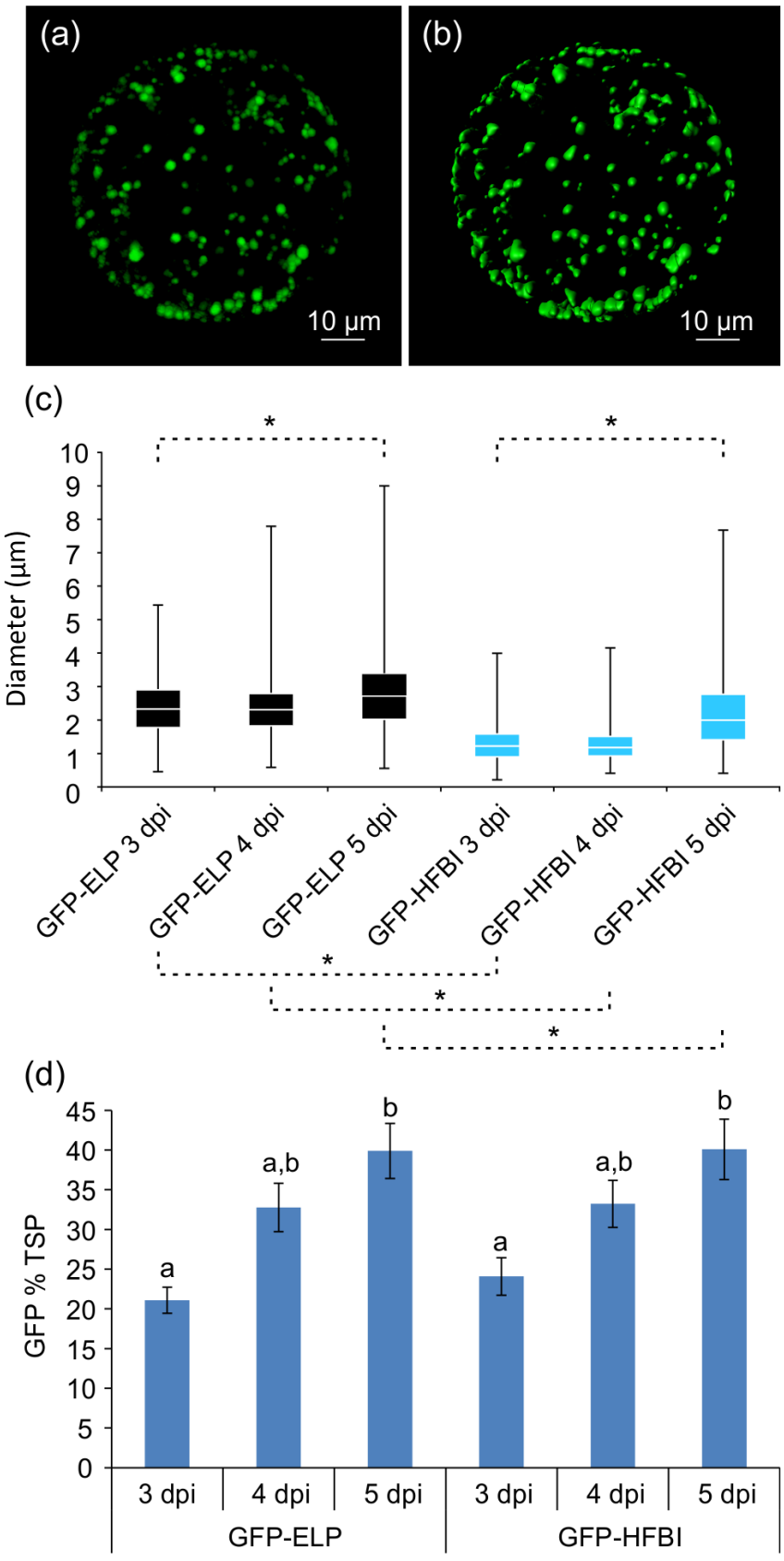
I found that PB size and protein accumulation levels increase simultaneously, as the size increase of ELP- or HFBI-induced PBs was accompanied with an increase in GFP accumulation from 21 and 24% of TSP at 3 dpi for GFP-ELP and GFP-HFBI respectively to 40% of TSP at 5 dpi for both proteins (Figure 2.2d).

2.2.2 Protein body formation is a concentration-dependent mechanism

Previous studies have shown that by addition of fusion tags, the accumulation levels of recombinant proteins can increase (Conley *et al.*, 2009a; Joensuu *et al.*, 2010; Kaldis *et al.*, 2013; Torrent *et al.*, 2009a). Conley *et al.* (2009b) have also shown that by co-agroinfiltrating GFP-ELP and p19, a suppressor of post-transcriptional gene

Figure 2. 2 Protein body size and GFP accumulation in protoplasts expressing GFP-ELP and GFP-HFBI increase simultaneously.

(a) 2D representation of the Z-stacks acquired from a protoplast expressing GFP-ELP at 4 dpi. (b) 3D-rendered Z-stack image of the protoplast shown in (a). (c) Size distribution pattern of GFP-ELP- and GFP-HFBI-induced PBs based on their respective diameters (μm) from 3 to 5 days post infiltration (dpi). The box represents the middle 50% of the protein body size distribution, and upper and lower "whiskers" represent the entire spread of the data acquired from 10 protoplasts per day per treatment; dotted lines indicate pairwise significant differences at * $p < 0.0001$ using Kruskal-Wallis one-way analysis of variance test. (d) GFP accumulation levels as a percentage of total soluble protein (TSP). Four biological replicates (each replicate containing four leaf discs of the infiltrated tissue from a different plant) were used for fluorometry. Columns not denoted with the same letter are significantly different from each other ($p < 0.05$) using one-way ANOVA. Error bars represent the standard error of the mean from biological replicates.


(a)
10 µm
(b)
10 µm
(c)
Diameter (µm)
0
1
2
3
4
5
6
7
8
9
10
*
*
*
*
*
GFP-ELP 3 dpi
GFP-ELP 4 dpi
GFP-ELP 5 dpi
GFP-HFBI 3 dpi
GFP-HFBI 4 dpi
GFP-HFBI 5 dpi
(d)
GFP % TSP
0
5
10
15
20
25
30
35
40
45
a
a,b
b
a
a,b
b
3 dpi
4 dpi
5 dpi
3 dpi
4 dpi
5 dpi
GFP-ELP
GFP-HFBI

silencing (Silhavy *et al.*, 2002), recombinant protein levels increase and PB formation is induced in a higher percentage of cells. Surprisingly, using the same experimental setup with unfused GFP, PBs were found in over 40% of the examined cells, suggesting a potential role for increased protein accumulation in the process of PB formation (Conley *et al.*, 2009b).

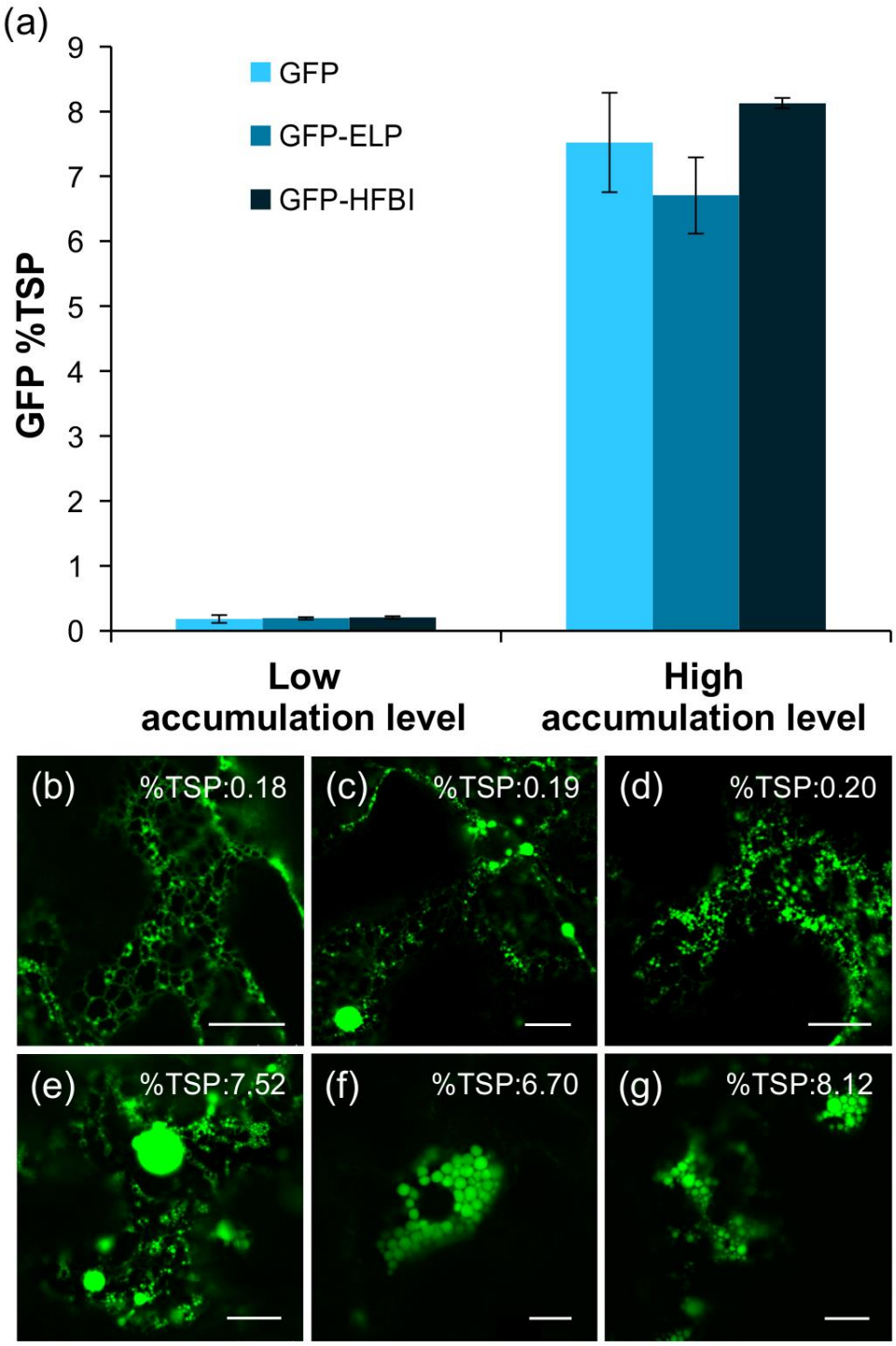
To investigate the effect of recombinant protein accumulation on PB formation, I conducted an experiment with a gradient of GFP levels. The GFP gradient was achieved by co-infiltrating a dilution series matrix of *Agrobacterium* containing the respective GFP constructs at an OD at A_{600} of 0.015 to 0.05 and the p19 suppressor of post-transcriptional gene silencing at OD A_{600} of 0.01 to 0.05. As expected, GFP accumulation levels increased with increasing *Agrobacterium* optical density (data not shown).

To compare the effects of fusion tags on PB formation at similar concentrations, I divided the treatments based on their respective accumulation levels into low (~ 0.2% of TSP), and high (> 6.5% of TSP) categories (Figure 2.3a) and looked for the effect of protein concentration on PB formation in each category (Figure 2.3b-g). At low accumulation levels, ER-targeted GFP highlighted the typical ER network pattern with minuscule PBs appearing along the ER network (Figure 2.3b). At a similar accumulation level of around 0.2% of TSP, both GFP-ELP and GFP-HFBI produced PBs. GFP-ELP-induced PBs were larger in size (in some cases reaching 5 µm) compared to GFP-HFBI-induced PBs, although the number of GFP-HFBI-induced PBs was larger (Figure 2.3c-d).

On the other hand, PBs were present in all treatments when accumulation levels were high (> 6.5% of TSP), irrespective of the presence or absence of the fusion tags (Figure 2.3 e-g). As shown in figure 2.3e, ER-targeted GFP induced PBs with sizes ranging from very small (~ 0.5 µm) to very large (~ 6 µm). At similar accumulation levels, GFP-ELP and GFP-HFBI induced higher numbers of homogenously sized PBs that were clustered together (Figure 2.3f-g). GFP-ELP generally induced PBs that were larger than GFP-HFBI. The appearance of very small PBs at 0.2% of TSP in the unfused GFP treatment, and their increase in number and size at higher accumulation levels indicates that levels of GFP, and possibly any other protein accumulating at high levels, may play a major

Figure 2. 3 High recombinant protein concentration leads to protein body formation.

(a) Leaf tissue samples were divided into low and high expressing groups and monitored for PB formation. Error bars represent the standard deviation of the mean from technical replicates quantified by dot blots. (b-d) PB formation in leaf tissues with low accumulation levels (~ 0.2% of TSP). Leaf tissue was infiltrated with GFP, GFP-ELP and GFP-HFBI in (b), (c) and (d), respectively. (e-g) PB formation in leaves with high accumulation levels (> 6.5% of TSP). Leaf tissue was infiltrated with GFP, GFP-ELP and GFP-HFBI in (e), (f) and (g), respectively. Bar, 10 µm.


(a)
GFP
GFP-ELP
GFP-HFBI
GFP %TSP
0
1
2
3
4
5
6
7
8
9
Low accumulation level
High accumulation level
(b) %TSP:0.18
(c) %TSP:0.19
(d) %TSP:0.20
(e) %TSP:7.52
(f) %TSP:6.70
(g) %TSP:8.12

role in PB biogenesis. The different patterns of PB appearance, i.e. their clustering and size uniformity may be due to the physico-chemical properties of the ELP and HFBI tags.

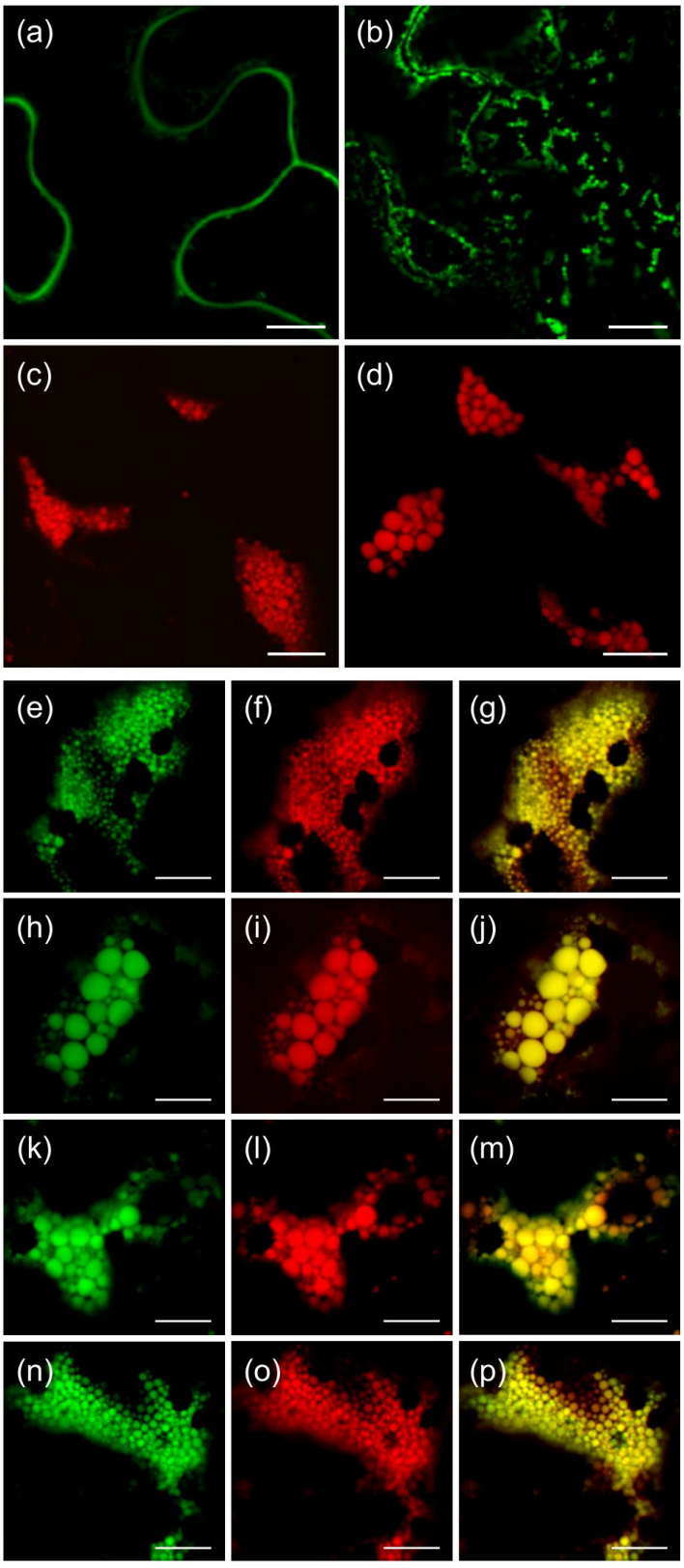
2.2.3 Secretory or ER-targeted proteins are sequestered passively into protein bodies

A proteomics study of the composition of Zera-induced PBs has shown the presence of a broad range of proteins other than Zera, including secretory pathway proteins, in PBs (Joseph *et al.*, 2012). Conley *et al.* (2009b) also showed that the ER-resident chaperone BiP is found in GFP-ELP-induced protein bodies. Therefore, I hypothesize that during the formation of PBs, secretory or ER resident proteins can be sequestered passively into the core of the PBs. To understand if this is the case, I co-expressed secretory GFP or ER-targeted GFP along with RFP-HFBI and RFP-ELP. All of the treatments were performed in the presence of p19 to ensure high accumulation levels.

Secretory GFP highlighted the apoplastic space between the cells without formation of PBs (Figure 2.4a), while ER-targeted GFP induced the formation of very small PBs along the ER network (Figure 2.4b). As expected, both RFP-HFBI and RFP-ELP gave rise to clusters of homogeneous PBs with RFP-ELP PBs being larger in size (Figure 2.4c-d). Co-expression of secretory GFP and RFP-HFBI (Figure 2.4e-g) or RFP-ELP (Figure 2.4h-j) resulted in the sequestration of secretory GFP into HFBI- or ELP-induced PBs, shown by co-localization of both green and red fluorescent proteins into the same PBs. Co-expression of ER-targeted GFP with RFP-HFBI and RFP-ELP also resulted in the sequestration of GFP into PBs induced by HFBI and ELP (Figure 2.4k-m, and 2.4n-p). In all of the co-expression treatments, the typical HFBI- or ELP-induced clustered patterns of PBs were observed (Figure 2.4e-p). My results indicate that proteins targeted to the secretory pathway, or retrieved to the ER (Figure 2.1) can be sequestered into HFBI- or ELP-induced PBs.

Figure 2. 4 Secretory and ER-targeted GFP are sequestered into RFP-HFBI and RFP-ELP induced protein bodies.

(a) Secretory GFP localizes to the apoplast. (b) ER-targeted GFP induces the formation of small PBs along the ER. (c) ER-targeted RFP-HFBI induces the formation of clusters of small PBs. (d) ER-targeted RFP-ELP induces the formation of large PB clusters. (e-g) Co-expression of Sec-GFP and RFP-HFBI results in sequestration of secretory GFP into RFP-HFBI-induced PBs. (h-j) Co-expression of Sec-GFP and RFP-ELP results in sequestration of secretory GFP into RFP-ELP-induced PBs. (k-m) Co-expression of GFP and RFP-HFBI results in co-localization of both proteins into the same PBs. (n-p) Co-expression of GFP and RFP-ELP results in co-localization of both proteins into the same PBs. All images were acquired in sequential mode. Bar, 10 µm.


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
(k)
(l)
(m)
(n)
(o)
(p)

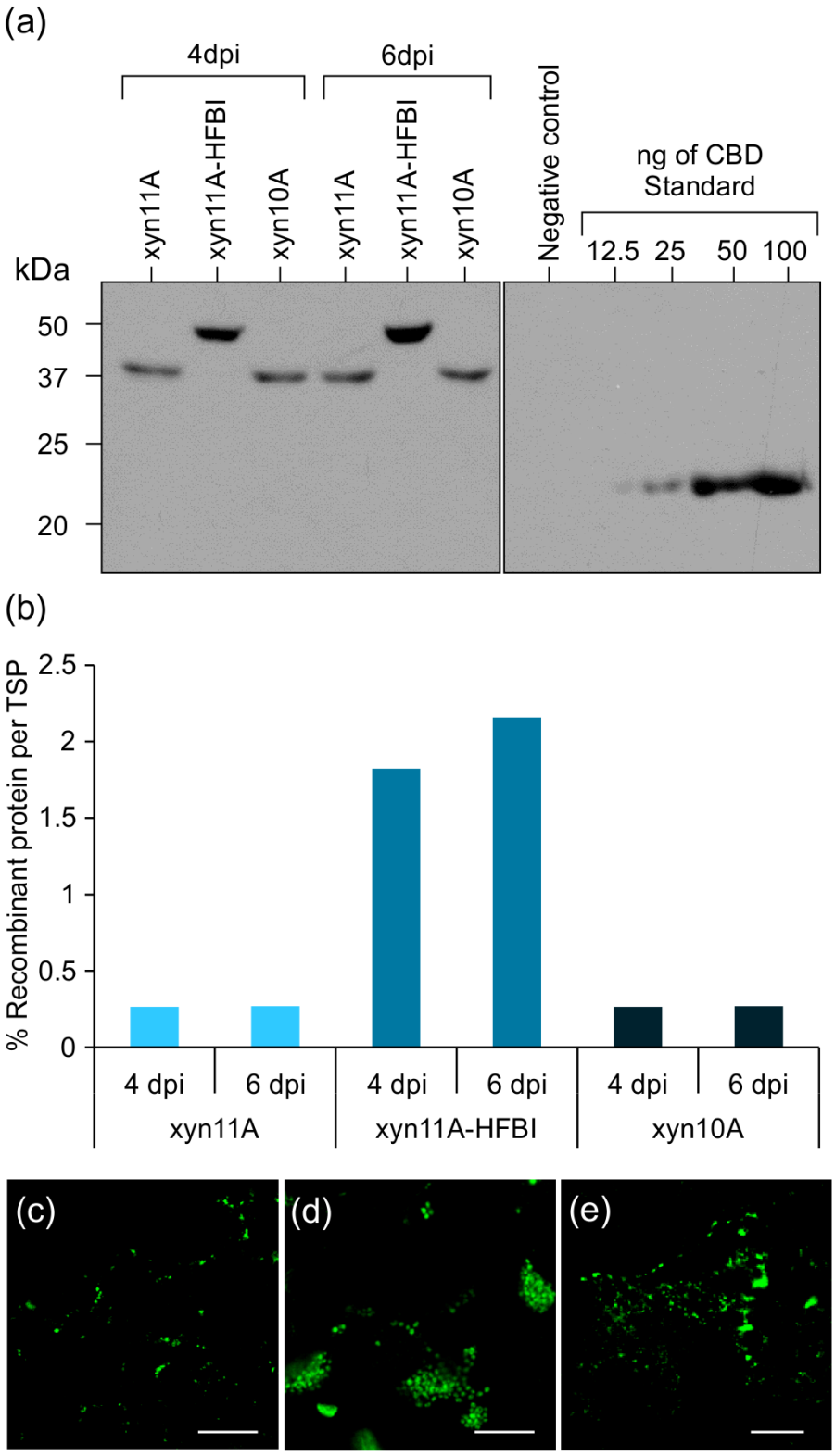
2.2.4 Fungal xylanases induce the formation of protein bodies

To determine if PB formation is limited to GFP or occurs upon overexpression of other proteins, I examined the possibility of PB formation with two other recombinant proteins accumulating at or higher than 0.2% of TSP. Two fungal xylanases from *Aspergillus niger* known as xyn11A (Nordberg *et al.*, 2014) with and without the HFBI tag, and xyn10A (Nordberg *et al.*, 2014) were selected and targeted to the ER with the ER-retrieval tetrapeptide KDEL (Figure 2.1). Accumulation levels of the three recombinant proteins were quantified by quantitative Western blot analysis at 4 and 6 dpi (Figure 2.5a-b). Xyn11A-HFBI accumulated to 2.1% of TSP at 6 dpi, almost 10-fold higher than the unfused enzyme. Based on my results with GFP, PB formation was expected in all three groups. The subcellular localization of xyn11A, xyn11A-HFBI and xyn10A, was determined by the whole-mount immunolocalization of the infiltrated leaf tissue using anti c-myc monoclonal mouse antibody as the primary antibody and donkey anti-mouse IgG antibody fused to Alexa Fluor® 488 dye as the secondary antibody, followed by confocal microscopy analysis. As shown in figure 2.5c and 2.5e, xyn11A and xyn10A induced the formation of dispersed PBs irregular in size, whereas higher accumulating xyn11A-HFBI induced numerous clustered PBs homogeneous in size (Figure 2.5d). This property of xyn11A-HFBI resembles the typical pattern of HFBI-induced PBs previously observed with GFP-HFBI. According to these results, I conclude that PB formation is not limited to GFP or its fusion partners and can be seen with other recombinant proteins as well.

Since xyn11A and xyn11A-HFBI induce the formation of PBs, and since I observed that secretory GFP is sequestered in RFP-HFBI-induced PBs, I asked if secretory GFP would be also sequestered in xyn11A-induced PBs. If so, simple co-localization of secretory GFP would provide a tool for visualizing PBs induced by non-fluorescent proteins without having to conduct immunolocalization, a lengthy and arduous process. I found that co-expression of xyn11A and xyn11A-HFBI with sec-GFP indeed allowed me to visualize the xyn11A-induced PBs. As shown in figure 2.6a, secretory GFP was localized to the apoplastic space of *N. benthamiana* leaf epidermal cells. Co-expression of secretory GFP and xyn11A resulted in sequestration of secretory GFP into xyn11A-induced

Figure 2. 5 Fungal xylanases accumulate in PBs.

xyn11A, xyn11A-HFBI and xyn10A xylanase proteins were co-expressed with p19 in *N. benthamiana* leaves. Leaf discs from 4 individual plants at 4 and 6 dpi (a-b) Equal amounts of TSP (3 µg) was loaded per lane. TSP from p19-infiltrated *N. benthamiana* tissue was used as negative control. Known amounts of a c-myc tagged protein were used as reference (12.5-100 ng). (a) Western blot analysis. (b) Quantification of (a). by densitometry. (c-e) PBs were visualized by confocal microscopy after whole-mount immunolocalization. Bar, 10 µm.


(a)
4dpi
6dpi
xyn11A
xyn11A-HFBI
xyn10A
xyn11A
xyn11A-HFBI
xyn10A
Negative control
ng of CBD Standard
12.5 25 50 100
kDa
50
37
25
20
(b)
% Recombinant protein per TSP
2.5
2
1.5
1
0.5
0
4 dpi
6 dpi
4 dpi
6 dpi
4 dpi
6 dpi
xyn11A
xyn11A-HFBI
xyn10A
(c)
(d)
(e)

PBs (Figure 2.6b). These large PBs were found in all of the cells producing PBs in the infiltrated tissue (Table 2.1). Similarly, co-expression of the secretory GFP and xyn11A-HFBI resulted in sequestration of secretory GFP into xyn11A-HFBI-induced PBs. As expected, these PBs were all small and homogenous in size and formed clusters (Figure 2.6c). This pattern was consistently observed in all of the cells forming PBs (Table 2.1).

Also, since I observed the sequestration of ER-targeted GFP into RFP-HFBI-induced PBs, I examined the possibility of sequestration of ER-targeted GFP into xyn11A-induced PBs. Even though ER-targeted GFP induced the formation of minuscule PBs along the ER network (Figure 2.6d), its co-expression with xyn11A resulted in the appearance of large PBs (Figure 2.6E). Co-expression of GFP-ER and xyn11A-HFBI, gave rise to clusters of homogenously-sized small PBs along with a number of large PBs (Figure 2.6f), (Table 2.1). My results suggest that in the process of PB formation, proteins present in the ER, such as proteins on their way to secretion (secretory GFP) or ER-retrieved proteins (GFP-KDEL) are sequestered into PBs.

2.2.5 PB induction can be used as a tool to increase accumulation levels of valuable proteins

My results indicate that proteins accumulating at levels higher than 0.2% of TSP are capable of PB formation *in vivo*, and that proteins targeted to the secretory pathway or retrieved to the ER using a KDEL retrieval signal can be sequestered into PBs. Based on these observations, and because I believe that PBs may protect recombinant proteins from turnover, I hypothesized that I can use PB formation as a strategy to help increase the accumulation levels of other recombinant proteins that do not accumulate at high levels. I first tested this hypothesis by co-expressing erythropoietin (EPO) with other proteins already shown to accumulate at high levels and induce PBs such as GFP, GFP-ELP, and GFP-HFBI (Lee *et al.*, 2012). EPO is a glycoprotein best known for its regulatory roles in production of red blood cells and its applications in the treatment of anemia caused by renal failure, chemotherapy and AIDS (Lee *et al.*, 2012).

Figure 2. 6 Secretory and ER-retrieved GFP localize to protein bodies induced by fungal xylanases.

(a) Secretory GFP highlights the apoplast of *N. benthamiana* epidermal cells. (b-c) Co-expression of secretory GFP with xyn11A (b) and xyn11A-HFBI (c) results in localization of secretory GFP to PBs. (d) ER targeted GFP highlights the ER network and induces the formation of very small PBs along the ER. (e-f) Co-expression of GFP with xyn11A (e) and xyn11A-HFBI (f) results in localization of GFP to PB clusters. Bar, 10 μm.

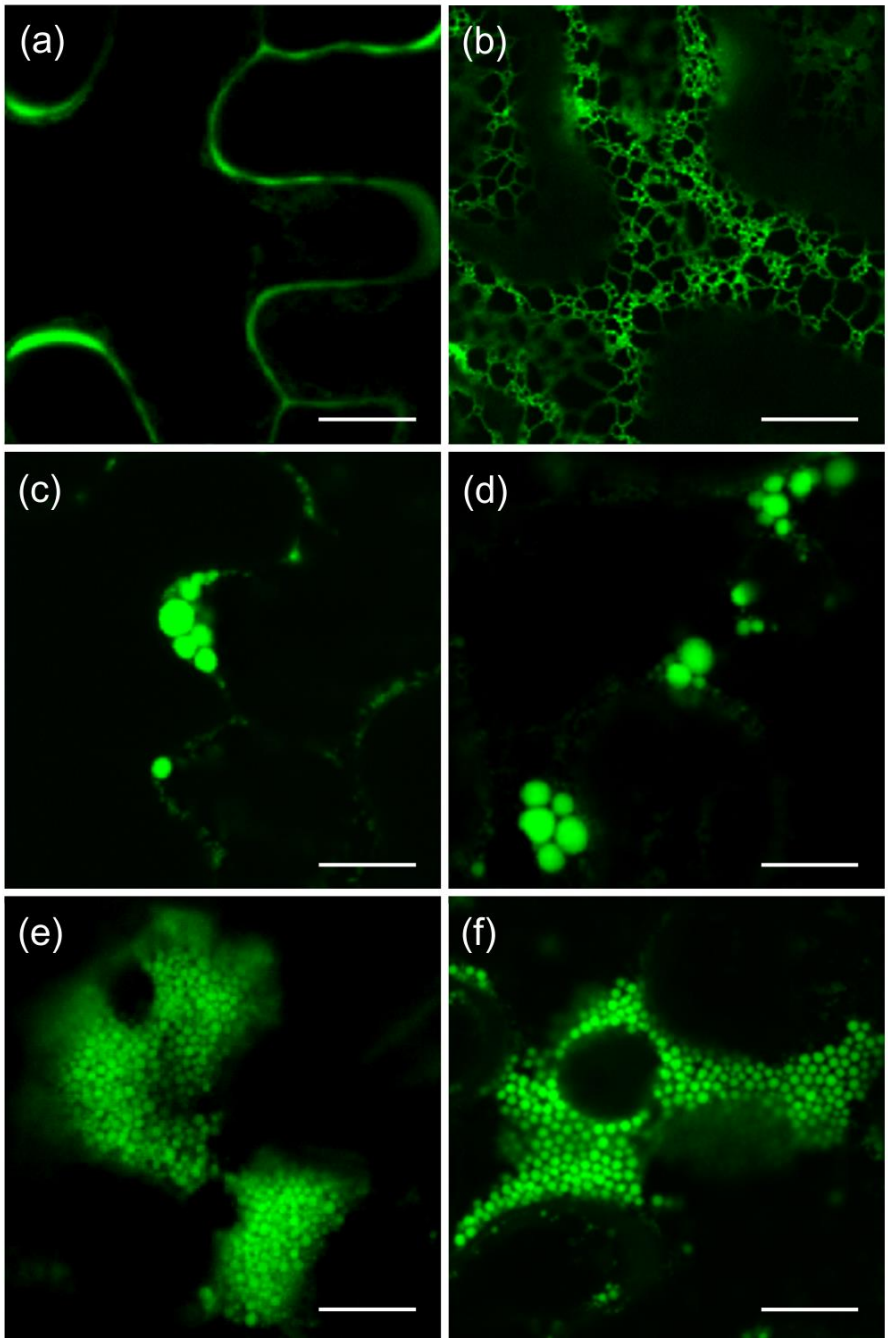

(a)
(b)
(c)
(d)
(e)
(f)

Table 2. 1 Occurrence and size distribution of PBs in cells transiently transformed with different constructs.

Treatments [a]	Protein body		Protein body size	
	Presence (%)	Absence [b] (%)	Small [c] (%)	Large [d] (%)
SecGFP	0	100	0	0
SecGFP / xyn11A	61	39	0	100
SecGFP / xyn11A-HFBI	48	52	100	0
GFP	71	29	90	10
GFP / xyn11A	90	10	43	57
GFP / xyn11A-HFBI	87	13	51	49

[a]To avoid any bias, presence or absence of PBs in 100 cells in four biological replicates (25 cells per replicate) were evaluated per treatment with confocal microscopy at 4 dpi.

[b]No PBs was observed in these cells.

[c]Maximum PB diameter was <1.0 µm.

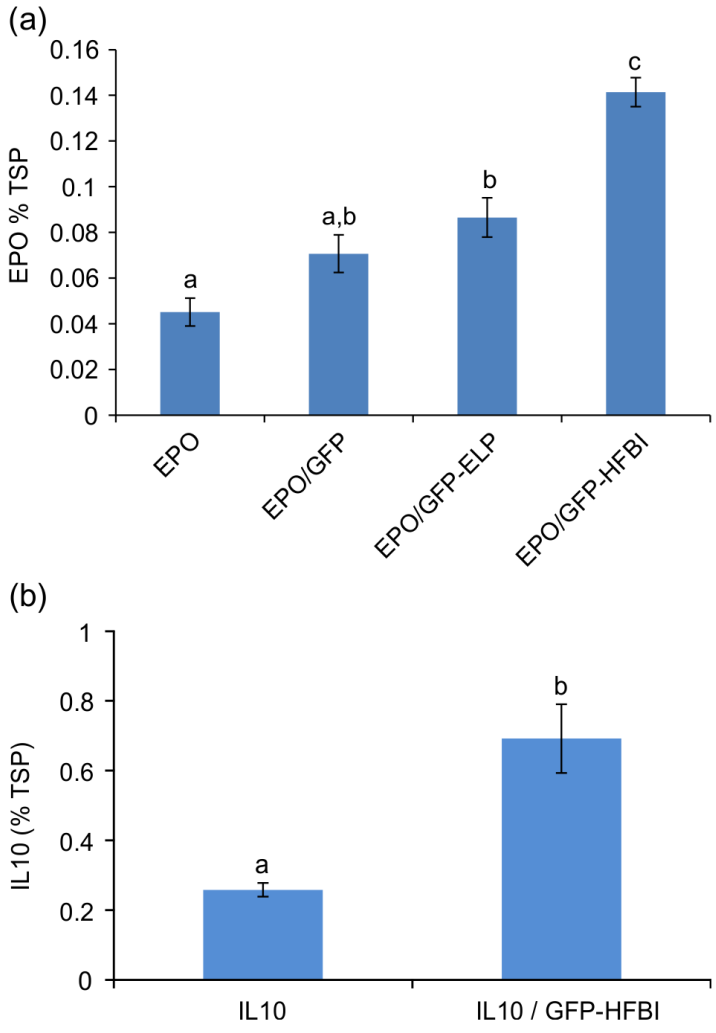
[d]Maximum PB diameter was >1.0 µm.

EPO transiently co-expressed with p19 in *N. benthamiana* leaves accumulated at 0.045% (± 0.012%) of TSP, almost a five-fold increase compared to the very low transient expression of the same EPO construct without p19 using binary vectors (Figure 7a) (Conley *et al.*, 2009a; 2009c). Therefore, p19 was used in all further experiments to obtain high accumulation levels of the recombinant proteins. Using the same construct as Conley *et al.* (Conley *et al.*, 2009c), co-expression of EPO in this work with GFP-ELP and GFP-HFBI, EPO accumulation levels were enhanced significantly by two-fold up to 0.086% (± 0.017%) of TSP with GFP-ELP, and more than three-fold up to 0.141% (± 0.012%) of TSP with GFP-HFBI (Figure 7a). Co-expression of EPO with GFP showed an increase in the mean EPO accumulation but not a significant difference from EPO expressed alone (Figure 2.7a).

To further validate the idea that accumulation of valuable unfused proteins can be increased by co-expression with a PB-inducing fusion, the human cytokine interleukin-10 (hIL-10) was co-expressed with GFP-HFBI. Human IL-10 is an anti-inflammatory cytokine with multiple roles in the regulation of the immune responses (Denes *et al.*, 2010). Several previous attempts to express unfused hIL-10 in plants resulted in low accumulation levels including expression in transgenic tobacco (accumulating at 0.0055% of TSP) (Menassa *et al.*, 2001), BY-2 cell suspension cultures (accumulating at 0.046% of TSP) (Kaldis *et al.*, 2013), and transient expression in *N. benthamiana* (accumulating at approximately 0.05% of TSP) (Conley *et al.*, 2009a). Transient co-expression of hIL-10 with p19 increased hIL-10 accumulation levels to 0.258% (± 0.056) of TSP. Upon co-expression of IL-10 with GFP-HFBI and p19, accumulation levels of IL-10 were significantly increased almost three-fold, up to 0.692% (± 0.279) of TSP (Figure 2.7b). This is approximately 125-fold higher than transgenic tobacco plants (Menassa *et al.*, 2001), and 15-fold higher than BY-2 cell culture (Kaldis *et al.*, 2013) or *N. benthamiana* transient expression in the absence of p19 (Conley *et al.*, 2009a).

Figure 2. 7 Co-expression of EPO and IL-10 with PB-inducing proteins increases EPO and IL-10 accumulation.

(a) Co-expression of erythropoietin (EPO) with GFP, GFP-ELP and GFP-HFBI. Enzyme-linked immunosorbent assay (ELISA) of recombinant EPO was used to quantify EPO accumulation levels in agroinfiltrated *N. benthamiana* leaves. (b) Co-expression of interleukin-10 (IL-10) with GFP-HFBI. IL-10 accumulation levels were determined by ELISA. Each column represents the mean value of eight biological replicates. Columns denoted with different letters are significantly different ($p < 0.002$) using the student *t*-test. The error bars represent the standard error of the mean.


(a)
EPO % TSP
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
a
a,b
b
c
EPO
EPO/GFP
EPO/GFP-ELP
EPO/GFP-HFBI
(b)
IL10 (% TSP)
1
0.8
0.6
0.4
0.2
0
a
b
IL10
IL10 / GFP-HFBI

2.3 Discussion

2.3.1 Protein body size and recombinant protein accumulation increase simultaneously

Characterization of PBs induced by GFP-ELP and GFP-HFBI revealed a clear size difference in PBs induced by the fusion tags in *N. benthamiana* leaf protoplasts. I have shown that ELP-induced PBs are larger than HFBI-induced PBs. Comparison of the data with the sizes of Zera-induced PBs (Llop-Tous *et al.*, 2010) suggests that both ELP and HFBI induce PBs larger than Zera. In this study, I used a novel technique by combining confocal imaging and 3D analysis to accurately measure the sizes of all of the PBs inside the protoplasts. The results presented here confirm that the size of ELP- and HFBI-induced PBs increase over time as well as their associated amount of fused recombinant protein. Previous studies by Llop-Tous *et al.* (2010) indicated that the size of Zera-ECFP-induced PBs increase over time from 2 to 10 dpi. This property of Zera-induced PBs was attributed to the continuous biosynthesis of Zera-ECFP polypeptides by the ER active ribosomes (Llop-Tous *et al.*, 2010). In 2012, Joseph *et al.*, showed that Zera-DsRed PBs were surrounded with a typical ER membrane decorated with ribosomes. Similarly, GFP-ELP- and GFP-HFBI-induced PBs were shown to be surrounded with membranes studded with ribosomes (Conley *et al.*, 2009b; Joensuu *et al.*, 2010). My results, in accordance with previous observations, suggest a possible role for ribosomes attached to membranes surrounding PBs in the increase in sizes of GFP-ELP- and GFP-HFBI-induced PBs from 3 to 5 dpi and the simultaneous increase in the amount of GFP during this time period. Meanwhile, in a series of fluorescent recovery after photobleaching (FRAP) experiments, Conley *et al.* (2009b) have shown recovery of GFP-ELP into photobleached PBs (within 5 minutes), which suggests the possibility of trafficking of GFP-ELP from surrounding PBs into the photobleached PB. Although protein trafficking between PBs or between the ER and the PBs needs further confirmation, it can still be considered a potential factor responsible for increasing the sizes of GFP-ELP- and GFP-HFBI-induced PBs.

2.3.2 Role of protein accumulation level in protein body formation

The occurrence of PBs in unfused ER-retrieved GFP and the increase in PB size at higher GFP accumulation levels suggests a direct role of protein accumulation in PB formation. This observation is consistent with a previous study in which a threshold level of protein accumulation was observed at 0.2% of TSP in transgenic tobacco plants expressing GFP, GFP-ELP and GFP-HFBI (Gutierrez *et al.*, 2013). Below 0.2% of TSP, no PBs were observed, while above 0.2% of TSP PBs were observed in almost all plants, regardless of the fusion tag. I confirmed the formation of PBs with fungal xylanases xyn11A and xyn10A, both of which reached accumulations slightly above 0.2% of TSP. Addition of the HFBI fusion tag to xyn11A not only increased accumulation levels by 10-fold, but also produced more PBs with the typical small and clustered HFBI-associated PB pattern. This pattern was previously observed with GFP-HFBI-induced PBs in independent studies (Gutierrez *et al.*, 2013; Joensuu *et al.*, 2010). Although I have shown that PBs form regardless of the availability of fusion tags above the threshold level, the distribution pattern and sizes of PBs were affected by the fusion tags, with obvious clustering and with larger PBs produced by ELP as compared to HFBI.

2.3.3 Protein body formation, a new tool in increasing recombinant protein accumulation

A proteomics study of Zera-DsRed-induced PBs has previously shown that aside from Zera-DsRed, which composed the majority of PB content, other ER resident and secretory proteins were found in PBs (Joseph *et al.*, 2012). In an independent study, the co-expression of ER-resident molecular chaperone binding protein (BiP) fused with CFP (BiP-CFP-KDEL) and ER-targeted YFP-ELP (YFP-ELP-KDEL) showed the sequestration of the ER chaperone BiP into the YFP-ELP-induced PBs (Conley *et al.*, 2009b). Furthermore, co-immunoprecipitation analysis of these two proteins suggested that BiP does not specifically interact with the ELP (Conley *et al.*, 2009b). This might have been due to the passive sequestration of BiP into the YFP-ELP-induced PBs. My results also confirmed the passive sequestration of ER-retrieved and secretory proteins

into PBs. This was shown separately by sequestration of secretory GFP and ER-retrieved GFP into RFP-HFBI- , RFP-ELP-, xyn11A-HFBI- and xyn11A-induced PBs.

Based on the observations of passive sequestration of ER resident proteins into PBs, I designed a series of experiments to study the potential of PB induction as a new approach to increase recombinant protein accumulation in plants. EPO and hIL-10 were chosen since they are both highly valuable pharmaceutical proteins that accumulate at low levels when expressed in plants. Although recent reports have shown accumulation levels of EPO fusions to high levels; EPOFc to 0.18% of TSP and EPO-ELD to approximately 2% of TSP by using viral based magnICON® vectors (Castilho *et al.*, 2013; Jez *et al.*, 2013), the highest accumulation level achieved in plants using binary vectors approached 0.01% of TSP in transient expression and up to 0.05% of TSP in transgenic tobacco (Conley *et al.*, 2009c). Here I have shown that by co-expressing EPO with a PB-inducing construct and p19, I could increase EPO accumulation levels up to 0.14 % of TSP, 14-fold higher than the previously reported transient expression levels using binary vectors (Conley *et al.*, 2009c).

Human IL-10 was previously expressed in several plant systems including transgenic tobacco, transgenic BY-2 cell lines and transiently in *N. benthamiana*, and in all cases accumulated below 0.05% of TSP (Conley *et al.*, 2009a; Kaldis *et al.*, 2013; Menassa *et al.*, 2001). By co-expressing hIL-10 and GFP-HFBI, hIL-10 accumulated up to 0.69% TSP on average, which is approximately a 14-fold increase from the highest previously reported hIL-10 accumulation level. This increase in accumulation levels of ER-targeted EPO or hIL-10 might be due to non-selective sequestration of ER-targeted proteins into PBs. I hypothesize that by co-expressing low-accumulating valuable proteins such as EPO and hIL-10 with high accumulating PB-inducing proteins, both recombinant proteins are sequestered in PBs. It has already been proposed that PBs protect recombinant proteins from physiological turnover within the cell (Conley *et al.*, 2011; Khan *et al.*, 2012; Müntz and Shutov, 2002), which might consequently explain the increased accumulation levels of EPO and hIL-10 in this work. It is worth noting that hIL-10-ELP fusion expressed in BY-2 cell lines were shown to accumulate to 3% of TSP; however the activity of the fused hIL-10 was significantly reduced when fused to ELP

(Kaldis *et al.*, 2013). With my proposed co-expression technique, the need for addition of fusion tags that might affect the proper folding and activity of their fused partner is eliminated.

2.4 Conclusions

In conclusion, the main objectives of this study were to understand the role of fusion tags and recombinant protein accumulation levels in the process of PB formation, and to investigate the potential utility of PBs as a tool to facilitate recombinant protein production in leaves of *N. benthamiana*. I have shown that the ELP fusion tag induces PBs larger in size compared to HFBI and that the PB size correlates with protein accumulation levels. I have also shown that presence of ELP or HFBI fusion tags is not necessary for PB formation, which occurs with several ER-targeted genes, but that ELP and HFBI affect the distribution patterns and size of PBs. I have also shown that PB formation is not limited to the conventional fluorescent proteins (e.g. GFP or RFP) as fungal xylanases also induce PBs. In addition, my results indicate that in the process of PB formation, secretory and ER resident proteins can passively integrate into the lumen of PBs. This property of PBs was used as a tool to increase accumulation levels of EPO and hIL-10 by co-expression with PB-inducing proteins.

2.5 Experimental procedures

2.5.1 Construct design and cloning

Secretory GFP, GFP, GFP-ELP, GFP-HFBI, and EPO plant expression vectors were previously published (Conley *et al.*, 2009b; 2009c; Joensuu *et al.*, 2010). The coding sequence of RFP was PCR amplified using primers RS1 (5'GGGGACAAGTTTGTACAAAAAAGCAGGCTTGATGGCCTCCTCCGAGGACG T3') and RS2 (5'GGGGACCACTTTGTACAAGAAAGCTGGGTCGGCGCCGGTG-GAGTGGC3'). The PCR construct was then inserted into the Gateway® donor vector pDONR/Zeo™ (Invitrogen, Carlsbad, USA). The integrity of the PCR construct was confirmed with sequencing analysis. Using Gateway technology, RFP was recombined

into the pCamGate-ER, pCamGate-HFBI-ER and pCamGate-ELP-ER (Figure 2.1) expression vectors (Conley *et al.*, in preparation) under the control of a double-enhanced cauliflower mosaic virus (CaMV) 35S promoter (Kay *et al.*, 1987; Wu *et al.*, 2001), a tCUP translational enhancer (Malik *et al.*, 2002) and a nopaline synthase (nos) terminator (Bevan *et al.*, 1983). The final expression constructs were then transformed into *Agrobacterium tumefaciens* strain EHA105. The xyn11A and xyn11A-HFBI plant expression vectors were obtained from Ruoyu Yan (Yan, 2011).

2.5.2 *Nicotiana benthamiana* growth and maintenance

Nicotiana benthamiana plants were grown for 7 weeks and used for transient expression. Plants were grown in a growth chamber at 22˚C with a 16 hour photoperiod at a light intensity of 110 µmol m^{-2} s^{-1}. Plants were always watered with the water soluble fertilizer (N : P : K = 20 : 8 : 20) at 0.25 g/L (Plant Products, Brampton, ON, Canada).

2.5.3 Transient expression in *Nicotiana benthamiana* plants

Agrobacterium tumefaciens cultures were grown to an optical density at 600 nm (OD_{600}) of 0.5-0.8, and collected by centrifugation at 1000 g for 30 minutes. The pellets were then resuspended in Agro-infiltration solution (3.2 g/L Gamborg's B5 medium and vitamins, 20 g/L sucrose, 10 mM MES pH 5.6, 200 µM 4′-Hydroxy-3′,5′-dimethoxyacetophenone) to a final OD_{600} of 0.30, followed by incubation at room temperature (21˚C) with gentle agitation for 1 hour. The suspension was then used for infiltration of the abaxial leaf epidermis through the stomata of *N. benthamiana* plants with a 1 ml syringe (Kapila *et al.*, 1997).

2.5.4 Tissue sampling and protein extraction

N. benthamiana leaf samples were collected at 4-6 dpi depending on the experiment. Four leaf discs were collected from at least three biological replicates per sample. Total soluble protein was extracted and quantified as previously described (Conley *et al.*, 2009a).

2.5.5 Protoplast preparation

Protoplasts were prepared as described by Sheen (2002). Briefly, 1 g of infiltrated leaf tissue was cut into strips of 0.5-1 mm, covered with enzyme solution (1.5% cellulase R10 (216016, Yakult pharmaceutical, Tokyo, Japan), 0.4% macerozyme® R10 (202051, Yakult pharmaceutical, Tokyo, Japan), 0.4 M mannitol, 20 mM KCl, 20 mM morpholinoethanesulfonic acid (MES) (pH 5.7), 10 mM $CaCl_2$, 5 mM β–mercaptoethanol and 0.1% BSA (Sigma A-6793)), and incubated overnight in the dark at 21˚C. Protoplasts were filtered with a 100 µm mesh, and centrifuged for 10 minutes at 80 x g at 4˚C. The supernatant was removed, replaced with an equal volume of cold washing and incubation solution (WI) (0.5 M mannitol, 4 mM MES (pH 5.7) and 20 mM KCl) and centrifuged again for 10 minutes at 80 x g at 4˚C. The solution was then transferred gently onto flotation medium (0.6 M sucrose, 10 mM MES (pH 5.7) and 20 mM KCl) to remove debris and centrifuged for 10 minutes at 80 x g at 4˚C. Protoplasts were removed from the top layer of the flotation medium and transferred to cold WI solution.

2.5.6 Confocal microscopy and image analysis

To visualize protoplasts, Z-stack confocal images of protoplasts were acquired with a Leica TCS SP2 confocal laser scanning inverted microscope (Leica Microsystems, Wetzlar, Germany) equipped with a 63X water immersion objective. The surpass module of Imaris® x64 software (version 7.6.1, Bitplane Scientific Software) was used to generate 3D reconstructed images and videos. To measure the size and volume of PBs, the surface module of Imaris software was used to detect separate PBs.

To visualize leaf samples, the abaxial leaf epidermis was imaged with a Leica TCS SP2 CLSM. For the imaging of GFP and chlorophyll autofluorescence, excitation with a 488 nm argon laser was used and fluorescence was detected at 500-525 nm and 630-690 nm, respectively. The excitation for RFP was 543 nm and its emission was recorded at 553-630 nm. Collected images were analyzed with the Leica Application Suite for Advanced Fluorescence (LAS AF, version 2.3.5) (Leica Microsystem, Germany). To avoid crosstalk between the fluorescent channels, images were acquired in the sequential

mode at all times. To visualize whole-mounted tissue, 488 nm argon laser was used to detect Alexa Fluor® 488 dye, and green fluorescence was detected at 495 to 525 nm.

2.5.7 Recombinant protein quantification

GFP quantification was performed with either fluorometry or immunodot blot as described by Gutierrez *et al.* (2013). For Western blot analysis, 3 µg of total soluble protein were resolved by 10% sodium dodecylsulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to polyvinylidene difluride (PVDF) membrane. The recombinant proteins were detected with the primary mouse anti-c-myc monoclonal antibody (GenScript, Cat. No. A00864) diluted 1:5000 in blocking solution and HRP-conjugated goat anti-mouse IgG secondary antibody diluted 1:3000 (Biorad, Cat. No. 170-6516). Membranes were visualized with the enhanced chemiluminescence (ECL) detection kit (GE Healthcare, Mississauga, ON, Canada) as described by the manufacturer. Quantification was performed using TotalLab TL 100 software (Nonlinear Dynamics, Durham, USA) with intensity analysis of specific bands, with known amounts of a c-myc-tagged standard protein as a standard.

Quantification of prEPO from *N. benthamiana* leaf extracts were performed by sandwich ELISA as described previously (Conley *et al.*, 2009c). At least 4 independent biological samples were quantified per treatment. Leaf tissue infiltrated with p19 was used as control.

The concentration of plant recombinant IL-10 was determined by comparison with a human IL-10 standard curve according to the manufacturer's instructions (BD Biosciences, Mississauga, Canada).

2.5.8 Immunocytochemistry

Whole-mounting of *N. benthamiana* leaves was performed as described by Sauer *et al.*, (Sauer *et al.*, 2006). Briefly, pieces of infiltrated and non-infiltrated leaves were fixed in 4% paraformaldehyde (Electron Microscopy Sciences, PA, USA) in PBS supplemented with 0.1% Triton™ X-100 (Sigma, Cat. No. T-9284), and permeabilized in Driselase

(Sigma, Cat. No. D-9515) and IGEPAL® (Sigma, Cat. No. CA-630) plus 10% dimethyl sulfoxide (DMSO) (Sigma, Cat. No. D8418). For immunocytochemistry, samples were pre-blocked in 3% BSA (Fisher Scientific, Cat. No. BP1600-100), incubated with c-myc monoclonal mouse antibody (1:1000 dilution) as the primary antibody and donkey anti-mouse IgG antibody conjugated to Alexa Fluor® 488 dye (1:130 dilution) as the secondary antibody (Life Technologies, Cat. No. A-21202). Confocal microscopy was performed as described above.

2.5.9 Statistical analysis

Statistical analysis of the protein body size comparison experiment was performed with Prism 6 (GraphPad Software, CA, USA). The Kolmogorov-Smirnov test was used to confirm the normal distribution of data. Since the data were not normally distributed, the Kruskal-Wallis non-parametric test of variance was used to assess if there were statistical differences between the median values of different treatments. Other statistical analyses were performed with SPSS Statistics V 22.0 for Windows (IBM, NY, USA). Depending on the number of comparisons involved, a student *t*-test or one-way ANOVA test was performed followed by Tukey-Kramer's test to find means significantly different from one another. A Post-Hoc test was then performed to analyze for significance of mean difference at 95% confidence intervals (the level of statistical difference was defined as $p < 0.05$).

2.6 Appendices

Supplementary Movie 2. 1 Four dimensional (4D) visualization of protein bodies.

Z-stack images of a GFP-ELP infiltrated protoplast were acquired by confocal microscopy and assembled into a 4D illustration with the Imaris software. Chlorophyll autofluorescence highlights chloroplasts inside the cell. PBs were detected in the green channel and the respective information of every PB was exported using the same software.

2.7 References

Benchabane, M., Goulet, C., Rivard, D., Faye, L., Gomord, V. and Michaud, D. (2008) Preventing unintended proteolysis in plant protein biofactories. *Plant Biotechnol. J.* **6**, 633-648.

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

Castilho, A., Neumann, L., Gattinger, P., Strasser, R., Vorauer-Uhl, K., Sterovsky, T., Altmann, F. and Steinkellner, H. (2013) Generation of biologically active multi-sialylated recombinant human EPOFc in plants. *PLoS One* **8**, e54836.

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

Conley, A.J., Joensuu, J.J., Menassa, R. and Brandle, J.E. (2009b) Induction of protein body formation in plant leaves by elastin-like polypeptide fusions. *BMC Biol.* **7**, 48.

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

Conley, A.J., Mohib, K., Jevnikar, A.M. and Brandle, J.E. (2009c) Plant recombinant erythropoietin attenuates inflammatory kidney cell injury. *Plant Biotechnol. J.* **7**, 183-199.

Denes, B., Fodor, I. and Langridge, W.H. (2010) Autoantigens plus interleukin-10 suppress diabetes autoimmunity. *Diabetes Technol. Ther.* **12**, 649-661.

Egelkrout, E., Rajan, V. and Howard, J.A. (2012) Overproduction of recombinant proteins in plants. *Plant Sci.* **184**, 83-101.

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

Floss, D.M., Schallau, K., Rose-John, S., Conrad, U. and Scheller, J. (2010) Elastin-like polypeptides revolutionize recombinant protein expression and their biomedical application. *Trends Biotechnol.* **28**, 37-45.

Gutierrez, S.P., Saberianfar, R., Kohalmi, S.E. and Menassa, R. (2013) Protein body formation in stable transgenic tobacco expressing elastin-like polypeptide and hydrophobin fusion proteins. *BMC Biotechnol* **13**, 40.

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

Jez, J., Castilho, A., Grass, J., Vorauer-Uhl, K., Sterovsky, T., Altmann, F. and Steinkellner, H. (2013) Expression of functionally active sialylated human erythropoietin in plants. *Biotechnol. J.* **8**, 371-382.

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

Joseph, M., Ludevid, M.D., Torrent, M., Rofidal, V., Tauzin, M., Rossignol, M. and Peltier, J.B. (2012) Proteomic characterisation of endoplasmic reticulum-derived protein bodies in tobacco leaves. *BMC Plant Biol.* **12**, 36.

Kaldis, A., Ahmad, A., Reid, A., McGarvey, B., Brandle, J., Ma, S., Jevnikar, A., Kohalmi, S.E. and Menassa, R. (2013) High-level production of human interleukin-10 fusions in tobacco cell suspension cultures. *Plant Biotechnol. J.* **11**, 535-545.

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

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

Khan, I., Twyman, R.M., Arcalis, E. and Stoger, E. (2012) Using storage organelles for the accumulation and encapsulation of recombinant proteins. *Biotechnol. J.* **7**, 1099-1108.

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

Lau, O.S. and Sun, S.S. (2009) Plant seeds as bioreactors for recombinant protein production. *Biotechnol. Adv.* **27**, 1015-1022.

Lee, J.S., Ha, T.K., Lee, S.J. and Lee, G.M. (2012) Current state and perspectives on erythropoietin production. *Appl. Microbiol. Biotechnol.* **95**, 1405-1416.

Linder, M., Selber, K., Nakari-Setala, T., Qiao, M., Kula, M.R. and Penttila, 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.

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

Ma, S. and Wang, A. (2012) Molecular farming in plants: an overview. In: *Molecular Farming in Plants: Recent Advances and Future Prospects* pp. 1-20. Springer. New York. USA.

Malik, K., Wu, K., Li, X.-Q., Martin-Heller, T., Hu, M., Foster, E., Tian, L., Wang, C., Ward, K. and Jordan, M. (2002) A constitutive gene expression system derived from the tCUP cryptic promoter elements. *Theor. Appl. Genet.* **105**, 505-514.

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

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

Müntz, K. and Shutov, A.D. (2002) Legumains and their functions in plants. *Trends Plant Sci.* **7**, 340-344.

Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., Smirnova, T., Grigoriev, I.V. and Dubchak, I. (2014) The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res.* **42**, D26-D31.

Pogue, G.P., Vojdani, F., Palmer, K.E., Hiatt, E., Hume, S., Phelps, J., Long, L., Bohorova, N., Kim, D. and Pauly, M. (2010) Production of pharmaceutical-grade recombinant aprotinin and a monoclonal antibody product using plant-based transient expression systems. *Plant Biotechnol. J.* **8**, 638-654.

Sauer, M., Paciorek, T., Benková, E. and Friml, J. (2006) Immunocytochemical techniques for whole-mount in situ protein localization in plants. *Nature Prot.* **1**, 98-103.

Sheen, J. (2002) A transient expression assay using *Arabidopsis* mesophyll protoplasts. http://molbio.mgh.harvard.edu/sheenweb/protocols_reg.html

Silhavy, D., Molnar, A., Lucioli, A., Szittya, G., Hornyik, C., Tavazza, M. and Burgyan, 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., Lasserre-Ramassamy, S., Llop-Tous, I., Bastida, M., Marzabal, P., Westerholm-Parvinen, A., Saloheimo, M., Heifetz, P.B. and Ludevid, M.D. (2009a) Eukaryotic protein production in designed storage organelles. *BMC Biol.* **7**, 5.

Torrent, M., Llop-Tous, I. and Ludevid, M.D. (2009b) Protein body induction: a new tool to produce and recover recombinant proteins in plants. *Methods Mol. Biol.* **483**, 193-208.

Tremblay, R., Wang, D., Jevnikar, A.M. and Ma, S. (2010) Tobacco, a highly efficient green bioreactor for production of therapeutic proteins. *Biotechnol. Adv.* **28**, 214-221.

Urry, D.W. and Parker, T.M. (2002) Mechanics of elastin: molecular mechanism of biological elasticity and its relationship to contraction. *J. Muscle Res. Cell Motil.* **23**, 543-559.

Wessels, J.G.H. (1997) Hydrophobins: proteins that change the nature of the fungal surface. In: *Advances in Microbial Physiology* pp. 1-45. Academic Press. London. UK.

Wu, K., Malik, K., Tian, L., Hu, M., Martin, T., Foster, E., Brown, D. and 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.

Yan, R. (2011) Expression of active fungal xylanases in *N. benthamiana* for hemicellulose degradation. M.Sc. thesis pp. 72. *Department of Biology.* University of Western Ontario.

3 Green to red photoconversion of GFP for protein tracking

Amirali Sattarzadeh*[1], Reza Saberianfar*[2,4], Warren R. Zipfel[3], Rima Menassa[2,4], and Maureen R. Hanson[1]

[1] Cornell University, Department of Molecular Biology and Genetics, Ithaca, NY, USA

[2] University of Western Ontario, Department of Biology, London, ON, Canada,

[3] Cornell University, Department of Biomedical Engineering, Ithaca, NY, USA

[4] Agriculture and Agri-Food Canada, London, ON, Canada

*These authors contributed equally to this work.

A version of this chapter is being prepared for publication in Nature Communications.

3.1 Introduction

Fusions of Green Fluorescent Protein (GFP) and its derivatives are extremely valuable tools for examining gene expression, protein and RNA localization, protein-protein interactions, protein synthesis and degradation, and organelle movement. The development of photoconvertible fluorescent proteins such as EosFP (Wiedenmann *et al.*, 2004), Kaede (Ando *et al.*, 2002), KikGR (Tsutsui *et al.*, 2005), Dendra (Gurskaya et al., 2006), and mCherry (Subach et al., 2009), allowed new questions to be asked concerning dynamic processes in living cells. However, the use of some photoconvertible proteins has drawbacks (Adam *et al.*, 2014). For example, non-GFP photoconvertible fluorescent proteins sometimes exhibit unexpected subcellular localization behavior (Hanson and Sattarzadeh, 2013). Another issue is the necessity to introduce new transgenes expressing the photoconvertible proteins into the genotypes of interest. Since many proteins have already been tagged with EGFP or GFP modified by an S65T mutation, being able to photoconvert these common forms of GFP would allow the use of existing transgenic lines to study dynamic changes.

EGFP has been known to photoconvert from green-state to red-emitting state upon exposure to blue light after excitation with 488 nm illumination under low oxygen conditions (Elowitz *et al.*, 1997), low oxygen with added flavin adenine dinucleotide (FAD) (Matsuda *et al.*, 2010) or by added oxidizing agents such as potassium ferricyanide, benzoquinone (BQ) or 3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyl tetrazolium bromide (MTT). The process of green to red photoconversion is also referred to as “redding” (Bogdanov *et al.*, 2009). The fluorescence spectra of the photoconverted species varies depending on the conditions used to catalyze the conversion. The “red” emission maximum was at 560 nm in the FAD-mediated conversion (Matsuda *et al.*, 2010), while three peaks were reported (560, 590 and 600 nm) under anaerobic conditions (no added FAD) (Elowitz *et al.*, 1997). The addition of chemical oxidants (e.g. BQ or MTT) resulted in 600 nm emissions with the exception of one report of a 675 nm peak (Saha *et al.*, 2013).

In this chapter, the previously uncharacterized irreversible photoconversion of EGFP from green-state to red-state is described. Unlike previous studies, the photoconversion of

EGFP *in vitro* and *in vivo* occurs spontaneously with no need for special treatments such as low oxygen environments or addition of electron acceptors.

3.2 Results

3.2.1 EGFP photoconverts from green to red in the absence of electron acceptors

While performing photobleaching studies on plant cells labeled with GFP, I made the unexpected observation that EGFP, mGFP4-T, and GFP modified by an S65T mutation (Figure 3.1a) can be photoconverted by 405 nm light *in vivo* without adding electron acceptors or using special conditions. To verify whether EGFP could be photoconverted *in vitro* from its normal green-state to a red-state, irradiation of purified EGFP was investigated. By using the Zeiss 710 bleaching mode with a 405 nm laser for photoconversion, rather than 488 nm as used in the previous reports (Bogdanov *et al.*, 2009; Elowitz *et al.*, 1997), purified EGFP was imaged before and after application of the 405 nm light. The green-state of purified EGFP was imaged with excitation at 488 nm, and the emission was acquired between 500-525 nm. The photoconverted EGFP (red-state) was monitored using 561 nm excitation and detected between 580-670 nm. Photoconversion occurred, as demonstrated by the decrease in green fluorescence and the simultaneous increase in red fluorescence (Figure 3.2a). Study of the emission spectra of photoconverted EGFP showed a maximum emission peak at around 612 nm when excited at 561 nm (Figure 3.2b).

Comparison of the absorption spectra of EGFP showed a higher absorption (~ 2.6 times) at pH 5.0 compared with pH 8.0 at 405 nm (Figure 3.2c). EGFP absorption peaks at 488 nm at pH 8.0, but the red species of EGFP was not excited by the 488 nm light, therefore the 488 nm illumination could not be used for photoconversion.

Time course comparison of the relative fluorescence of red-state EGFP at different pHs also indicated an increased yield (~ 2.3 times) of red photo-product at pH 5.0 compared to pH 8.0 (Figure 3.2d). This finding indicates that cellular compartments with higher pH may be less suitable for imaging of photoconverted GFP. The photoconversion appears to

Figure 3. 1 Comparison of different strains of GFP variants and schematic representation of constructs.

(a) Alignment of the coding sequence of the GFP variants used for photoconversion. The alignment was created using GeneDoc (http://www.nrbsc.org/gfx/genedoc/). (b) Schematic representation of the GFP transgenes used in photoconversion experiments. Pr1b-SP, secretory signal peptide from the tobacco pathogenesis-related protein 1; EGFP, enhanced green fluorescent protein; L, linker $(GGGS)_3$; KDEL, endoplasmic reticulum retrieval signal peptide; Htt^{Q103}, exon 1 of human Huntington polyQ gene; Cytosolic mGFP4-T in tobacco cell culture and cytosolic S65T-GFP in *Drosophila* gut cells. In cytosolic mGFP4-T, V (Valine) was replaced by A (Alanine), and Q (Glutamine) was replaced by R (Arginine) compared with EGFP. In cytosolic S65T-GFP, L (Leucine) was replaced by F (Phenylalanine) compared with EGFP.

a

```
                  *         20        *         40        *         60        *         80        *        100        *
GFP       : M-SKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG : 116
S65T-GFP  : MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG : 117
mGFP4-T   : MASKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMRRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG : 117
EGFP      : MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEG : 117

                 120        *        140        *        160        *        180        *        200        *        220        *
GFP       : DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGM : 233
S65T-GFP  : DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGM : 234
mGFP4-T   : DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGM : 234
EGFP      : DTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLGM : 234

GFP       : DELYK : 238
S65T-GFP  : DELYK : 239
mGFP4-T   : DELYK : 239
EGFP      : DELYK : 239
```

b

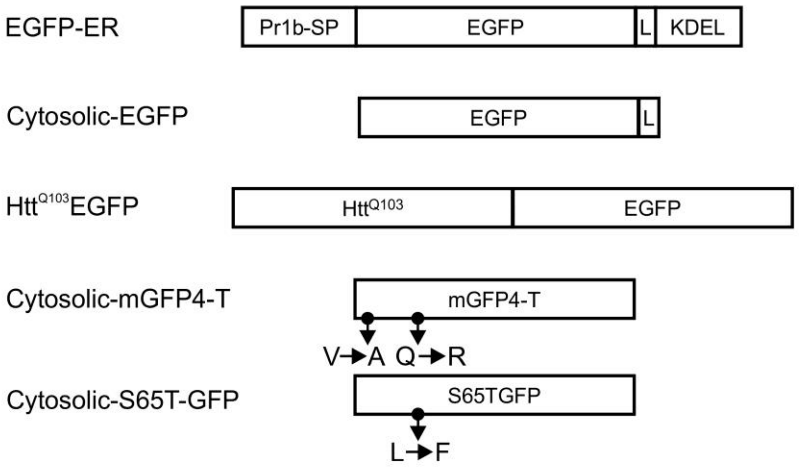
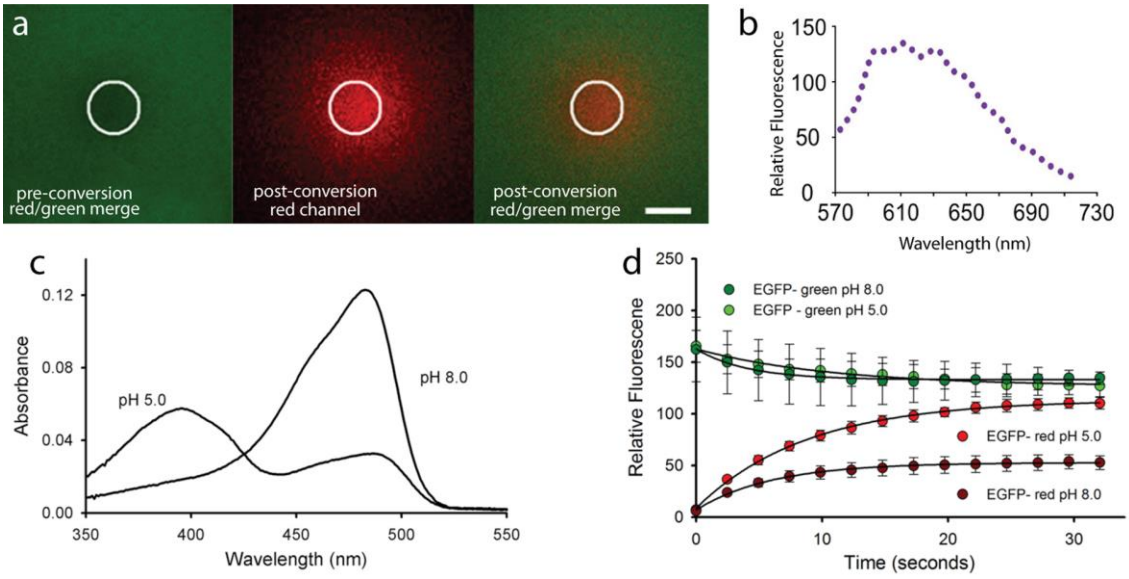


Figure 3. 2 Photoconversion of purified EGFP *in vitro*.

(a) Typical pre- and post-photoconversion images of a thin (~ 10 μm) layer of 30 μM EGFP using 405 nm excitation. Bar 50 μm. (b) Emission spectra of photoconverted red-state EGFP acquired in 5 nm bandwidth intervals. (c) Absorption spectra of EGFP (~ 2 μM) at pH 5.0 and pH 8.0. EGFP absorption at 405 nm is approximately 2.6 times higher in pH 5.0 compared to pH 8.0. (d) Time course comparison of the increase in red fluorescence with successive 405 nm irradiations within the region of irradiation (ROI) shown in white. EGFP (30 mM) was subjected to a sequence of 10 “bleach” iterations over the ROI. Plotted data points are the average pixel values across the entire field of view which estimates the total photo-product produced. Error bars represent the standard error of the mean of three trials at pH 5.0 and two at pH 8.0.


a
pre-conversion
red/green merge
post-conversion
red channel
post-conversion
red/green merge
b
Relative Fluorescence
150
100
50
0
570
610
650
690
730
Wavelength (nm)
c
Absorbance
0.12
0.08
0.04
0.00
pH 5.0
pH 8.0
350
400
450
500
550
Wavelength (nm)
d
Relative Fluorescene
250
200
150
100
50
0
EGFP- green pH 8.0
EGFP - green pH 5.0
EGFP- red pH 5.0
EGFP- red pH 8.0
0
10
20
30
Time (seconds)

be irreversible since the red-state EGFP was visible several hours after photoconversion in the *in vitro* experiments (data not shown).

3.2.2 EGFP photoconversion, a valuable tool for protein tracking *in vivo*

To determine whether GFP could be converted from green to red *in vivo*, a photoconversion experiment was performed first in transgenic tobacco cells in which GFP was localized in the cytoplasm (Köhler *et al.*, 1997). Converting GFP from a green- to a red-state resulted in rapid diffusion of the red form within the plant cytoplasm. In less than 3 seconds after irradiation, red-GFP was detected in the red channel. The photoconverted red-GFP then diffused through the whole cell and was readily visible throughout the cell within 2 minutes (Figure 3.3a and Supplementary Movie 3.2). Fluorescence intensity was measured in the region of irradiation (ROI) indicated by a white circle. There is a noticeable decrease in intensity of the green signal and a concomitant increase in intensity of the red signal, which is an indication of photoconversion from green to red (Figure 3.3b).

The photoconversion of EGFP was also observed in other eukaryotic model organisms, including *Drosophila* and rat cells. A *Drosophila* line in which S65T-GFP was expressed in gut cells from 3rd instar larvae (Reichhart and Ferrandon, 1998), and a rat PC12 cell line expressing EGFP fused to exon 1 of the human Huntington polyQ (Htt^{Q103}) gene (Aiken *et al.*, 2004) were obtained (Figure 3.1b). Upon irradiation of *Drosophila* and rat cells using conditions similar to plant cells, photoconversion from green- to red-state was observed (Figure 3.3c-d).

In addition to photoconversion of stably transformed cells, photoconversion was also observed in *Nicotiana benthamiana* epidermal leaf cells transiently expressing EGFP. Irradiation of a small area of the apoplast-targeted EGFP resulted in instant photoconversion of the protein localized in the extracellular space between the cells. To track the photoconverted protein and avoid photobleaching, irradiation of ROI was performed with time intervals (every 42 seconds in this case) to allow replacement of photoconverted protein with newly secreted protein in the ROI (Figure 3.4a). As shown

Figure 3. 3 Photoconversion of GFP from green to red state *in vivo*.

Single snapshots from pre-photoconversion and post-photoconversion are shown. Photoconversion was performed at 50% laser power and 30 iterations, except rat cells were irradiated at 70% power with 30 iterations. (a) Tobacco cell suspension culture in which GFP is targeted to the cytosol. (b) Changes in fluorescence intensity in the region of irradiation (ROI) in "(a)" over time indicate increasing red and decreasing green fluorescence. (c) *Drosophila* gut cells from 3rd instar larvae expressing S65T-GFP. (d) Rat PC12 cells expressing EGFP fusion to exon 1 of human Htt^{Q103} gene. Bar 10 μm.

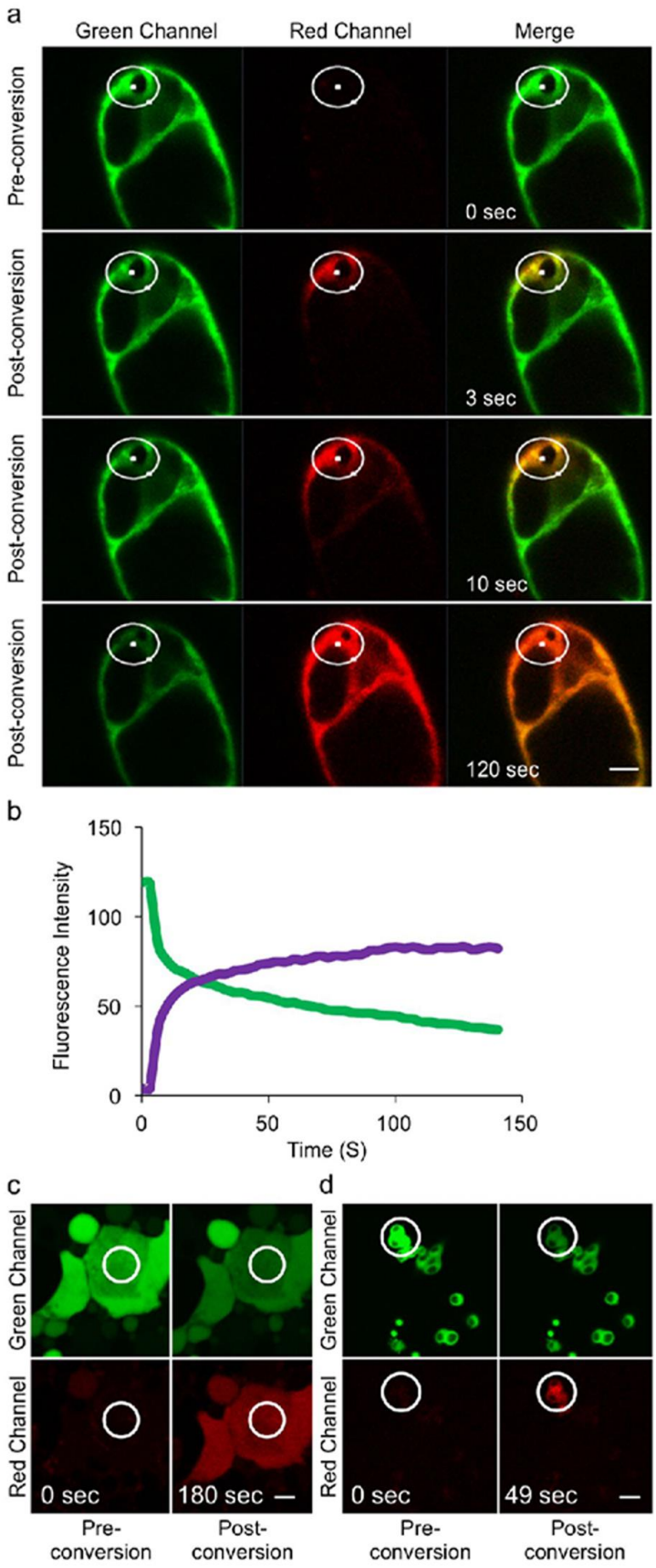
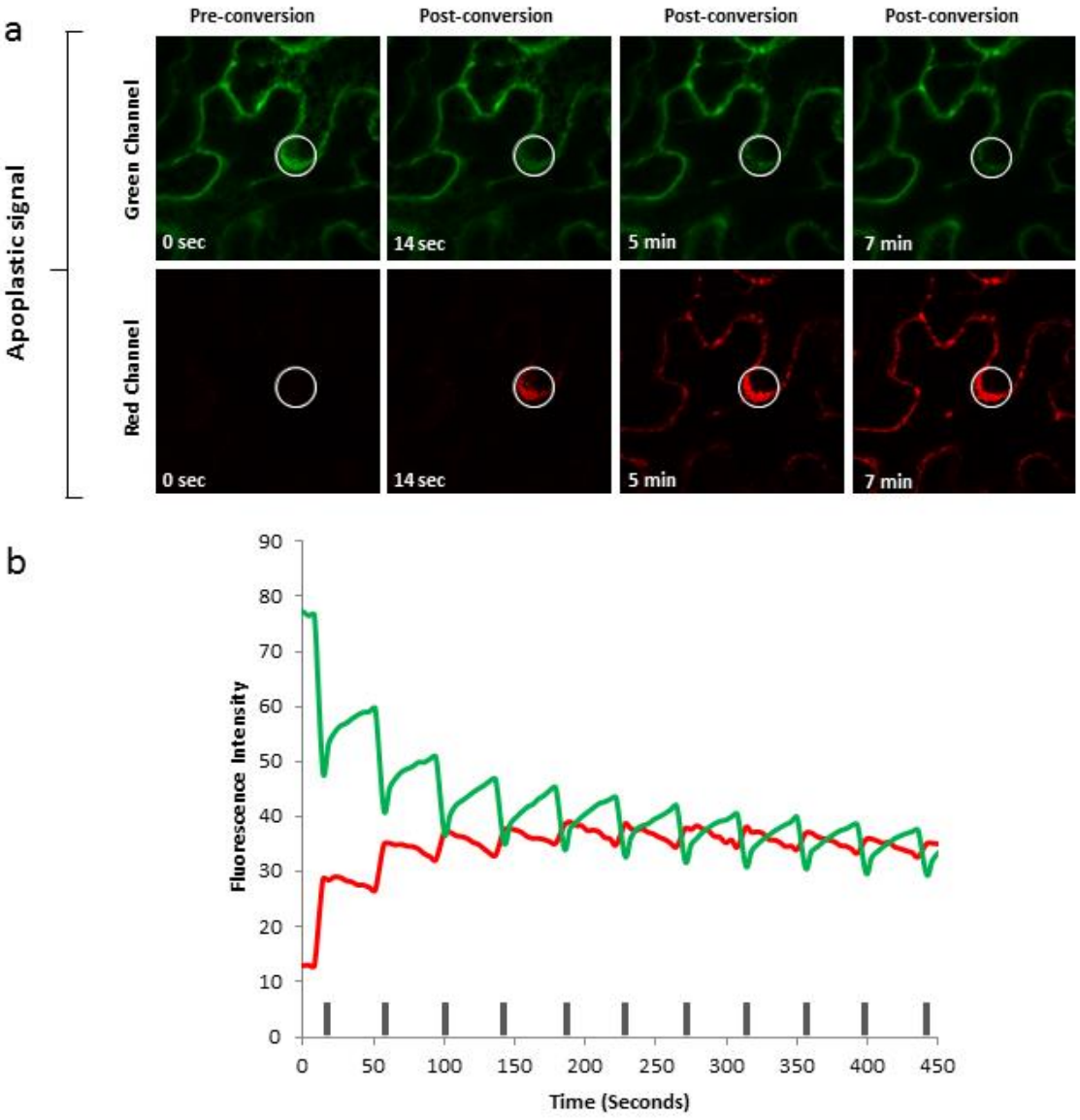

a
Green Channel
Red Channel
Merge
Pre-conversion
Post-conversion
Post-conversion
Post-conversion
0 sec
3 sec
10 sec
120 sec
b
Fluorescence Intensity
150
100
50
0
0
50
100
150
Time (S)
c
d
Green Channel
Red Channel
0 sec
180 sec
0 sec
49 sec
Pre-conversion
Post-conversion
Pre-conversion
Post-conversion

Figure 3. 4 EGFP photoconversion occurs in *N. benthamiana* cells transiently expressing EGFP targeted to apoplastic space.

(a) Secretory EGFP targeted to the apoplastic space between the cells photoconverts from green to red upon excitation with 405-nm laser at 35% of laser power and 20 iterations. The photoconverted red protein travels within the apoplastic space between the cells. (b) Change in fluorescence intensity in the ROI shown with a white circle in (a) over time indicate increasing red and decreasing green fluorescence. Grey bars represent the irradiation times. Bar 20 µm.


a
Pre-conversion
Post-conversion
Post-conversion
Post-conversion
Apoplastic signal
Green Channel
Red Channel
0 sec
14 sec
5 min
7 min
0 sec
14 sec
5 min
7 min
b
Fluorescence Intensity
Time (Seconds)
0
10
20
30
40
50
60
70
80
90
0
50
100
150
200
250
300
350
400
450

in Figure 3.4b, a significant drop in fluorescence intensity of green-state EGFP happened simultaneously with a significant increase in the fluorescence intensity of red-state EGFP. Trafficking of new EGFP to the ROI is confirmed by the increase in green fluorescence intensity between irradiations while trafficking of red-state EGFP out of ROI is shown by the decrease in the amount of fluorescence intensity of red-state EGFP (Figure 3.4b).

Upon photoconversion of a small area of the ER-targeted EGFP, the green signal instantly switched from green to red and spread throughout the ER within 5 minutes to areas more than 80 μm away from the irradiated area (Figure 3.5a, Supplementary Movie 3.3). A similar observation was made with irradiation of a small area of a cell expressing cytosolic EGFP in which the signal spread across the cytosol within 7 minutes after irradiation (Figure 3.5b, Supplementary Movie 3.4). On the other hand, chloroplasts in transplastomic tobacco cells expressing EGFP was completely photoconverted upon irradiation, but the photoconverted red-state EGFP did not fill the whole chloroplast or move out of the chloroplast, neither was the green-state EGFP replaced in the ROI (Figure 3.5c).

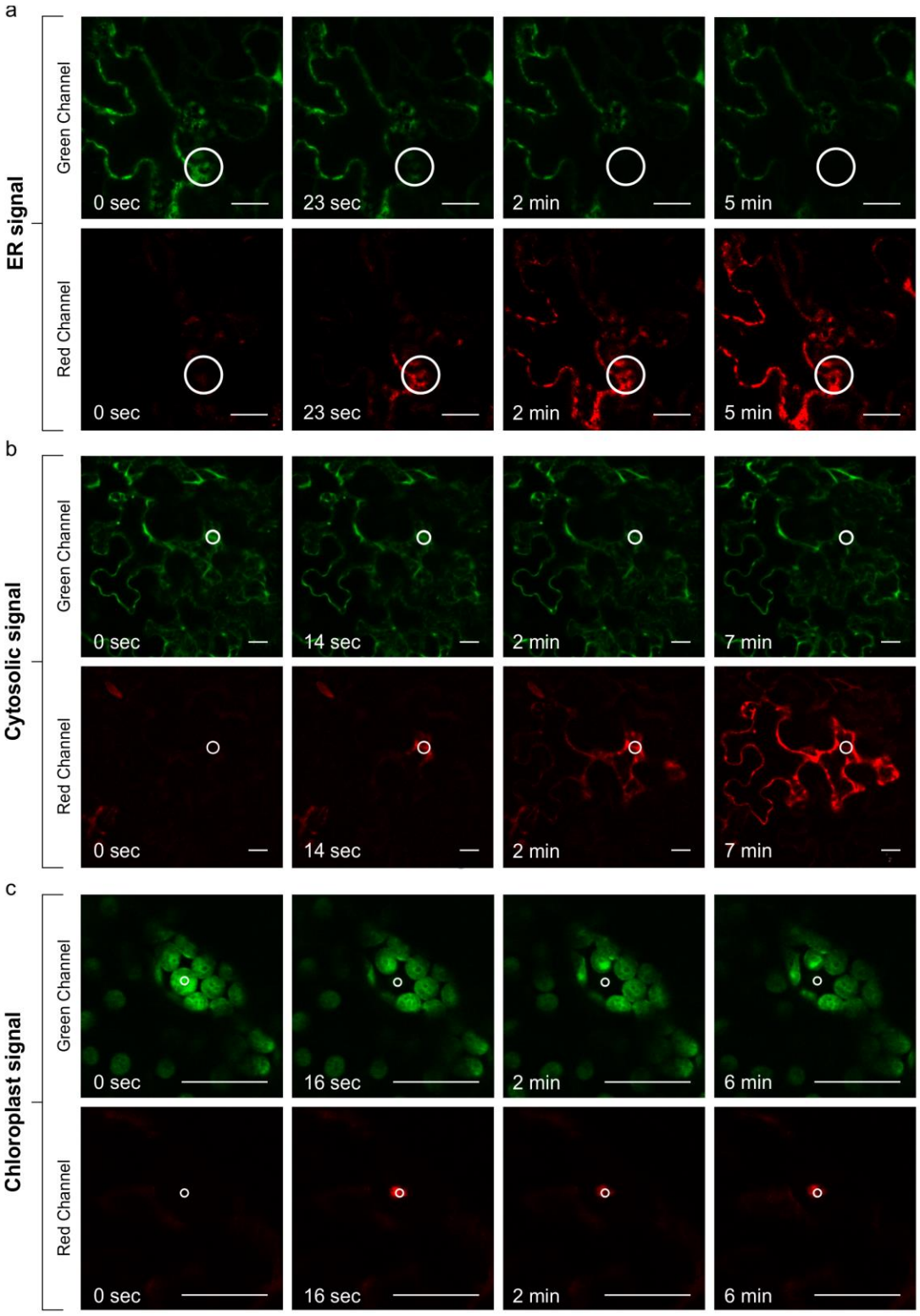
3.3 Discussion and conclusion

The results presented in this chapter describe a previously unnoticed ability of GFP which enables regional, non-invasive photoconversion of GFP without the need for providing a low oxygen environment. Because many proteins have already been tagged with GFP in transgenic cells, being able to photoconvert GFP, rather than other photoconvertible fluorescent proteins, would allow the use of existing transgenic lines for studies of dynamic processes in living cells.

GFP photoconversion described here was successfully tested in a range of model organisms, such as *Drosophila*, rat, tobacco cell culture and *N. benthamiana* leaf epidermal cells, in different subcellular locations such as the cytosol, ER, apoplast and chloroplast. By carefully selecting the conditions to avoid photobleaching, the flow of photoconverted molecules in several subcellular locations of leaf cells were observed.

Figure 3. 5 EGFP photoconversion occurs in *N. benthamiana* cells transiently expressing EGFP.

(a) ER-targeted EGFP photoconverts from green to red upon excitation with 405-nm laser at 60% of laser power and 100 iterations. The photoconverted red protein travels within the ER network of the cell. (b) Cytosolic EGFP photoconverts and spreads within the cytosolic space of the cell. Photoconversion was performed at 70% of laser power and 30 iterations. (c) Photoconversion of a chloroplast in transplastomic tobacco leaf cells expressing EGFP. The photoconverted signal does not spread to neighboring chloroplasts. Photoconversion was performed at 100% of laser power and 90 iterations. White circles represent the irradiated area. Bar 20 µm.


a
ER signal
Green Channel
Red Channel
0 sec
23 sec
2 min
5 min
b
Cytosolic signal
Green Channel
Red Channel
0 sec
14 sec
2 min
7 min
c
Chloroplast signal
Green Channel
Red Channel
0 sec
16 sec
2 min
6 min

The absorption spectrum of EGFP at pH 5.0 was found to be higher than at pH 8.0, suggesting that photoconversion may be more easily detectable in cellular compartments with lower pH values such as the trans Golgi network (TGN) (pH ~ 6.1) and vacuole (pH ~ 6.0) in plant cells and TGN (pH ~ 5.9), late endosome (pH ~ 6.0) and lysosome (pH ~ 5.5) in animal cells (Martiniere *et al.*, 2013a; Martiniere *et al.*, 2013b). On the other hand, photoconversion in organelles with higher pH values might not be easily performed. For instance, to fully photoconvert the EGFP in the chloroplast of transplastomic tobacco, high laser power (100%) was required, and the photoconverted protein was not visible in the whole chloroplast neither showed any movement within or out of the chloroplast. This might have been due to the high pH of chloroplast stroma (pH ~ 8.0) (Shen *et al.*, 2013) or due to the photobleaching of the protein since photoconversion of chloroplasts was only observed at full laser power (100%) and at high iterations (n=90).

Photoconversion of EGFP offers several advantages compared to other non-GFP-based photoconverting proteins. For instance, transgenic *Arabidopsis thaliana* plants expressing cytosolic mEosFP grown in bright fluorescent white light, up to 25% of hypocotyl epidermal cells contained red nuclei (due to auto-photoconversion) without any intentional photoconversion (Mathur *et al.*, 2010). Another concern involved with the use of mEosFP is the partial photoconversion under short exposure times. Photoconversion of mEosFP happens in a concentric manner, therefore producing variability in shades which is problematic since both partial photoconversion and co-localization of the green and red, will produce yellow hues (Mathur *et al.*, 2010).

Furthermore, the maximum emission peak of the photoconverted GFP (red-state) is at 612 nm, which is different from that of the most commonly used photoconvertible proteins such as mEOS2 (red-state) at 584 nm and Dendra2 (red-state) at 573 nm. This enables the imaging of red-state EGFP with other fluorescent proteins such as EBFP, ECFP, EYFP with maximum emission peaks at 445, 476, 527 respectively, and many other fluorescent proteins. In conclusion, my results demonstrate that GFP photoconversion can be added to the cell biologist’s toolkit for monitoring protein dynamics.

3.4 Experimental procedures

3.4.1 Construct design and transgenes

The GFP variants used in this study are shown in figure 3.1b. ER-EGFP and cytosolic-EGFP plant expression vectors (Conley *et al.*, 2009), PC12 cell line expressing Htt^{Q103}-EGFP generously provided by Dr. R. Cumming, University of Western Ontario (Aiken *et al.*, 2004), and the *Drosophila* strain expressing S65T-GFP (Reichhart and Ferrandon, 1998) were previously published.

3.4.2 Strains used and culture conditions

For transient expression experiments, 7-week old *N. benthamiana* plants grown at 22°C with a 16 hour photoperiod at 110 μmol m^{-2} s^{-1} were used. Plants were fertilized with the water soluble fertilizer (N:P:K=20:8:20) (Plant Products, Brampton, ON, Canada) at 0.25 g/L. *Agrobacterium tumefaciens* strain EHA105 containing the cytosolic-EGFP and ER-EGFP were used for infiltration as described by Conley *et al.* (2009). Rat PC12 cells were grown in Dulbecco's modified Eagle's medium (DMEM) and supplemented with 10% FBS and 1% penicillin/streptomycin at 37°C in 5% CO_2. To induce the expression of Htt^{Q103}-EGFP, cells were treated with 2.5 μM tebufenocide (Sigma, St. Louis, MI) for 24 to 48 hours. Tobacco cell suspensions were grown in NT1 medium with regular shaking at 22°C (Reichhart and Ferrandon, 1998). *Drosophila* larvae were grown at room temperature in standard corn meal agar.

3.4.3 Microscopy and imaging

A Zeiss LSM 710 confocal microscope equipped with a 10X dry or 40X water immersion objective was used for the imaging and photoconversion of GFP. The Zeiss automated bleaching mode was used for photoconversion. A 405 nm laser (2.5 mW at 100% power setting) was used to photoconvert. The photoconversion was monitored using 488/561 nm excitation with the emission collected between 500-525 nm (green) and 580-670 nm (red). To establish the baseline fluorescence intensity and distribution, five pre-photoconversion scans (~4 sec) were recorded before starting the photoconverting irradiation. Images were recorded and processed using Zeiss Zen software. In cells with

high expression levels of GFP, lower laser powers (20-30% corresponding to ~ 500 μW) and shorter exposure times were used for photoconversion. In cells with lower GFP expression, sequential and simultaneous imaging of cells during photoconversion was used in order to prevent photo-damage in the cells under irradiation.

To track the *in vivo* movement of proteins, multiple irradiations with time intervals (from seconds to minutes) were used to follow trafficking of the photoconverted protein and to photoconvert newly synthesized/trafficked proteins in the irradiated area (Figure 3.2, Supplementary Movie 3.3, 3.4).

3.5 Appendices

Supplementary Movie 3. 1 Time-lapse series demonstrating photoconversion of purified EGFP.

Supplementary Movie 3. 2 Time-lapse series demonstrating photoconversion of a tobacco suspension cell.

Tobacco suspension cell culture labeled with S65T-GFP in the cytoplasm undergoing photoconversion from the green state to the red state.

Supplementary Movie 3. 3 Time-lapse series demonstrating photoconversion and trafficking of ER-targeted EGFP in a *N. benthamiana* epidermal cell.

Confocal imaging of cells expressing ER targeted EGFP in *N. benthamiana* was performed to demonstrate the photoconversion and trafficking of EGFP in the ER network. The ROI (shown with white circles) was irradiated with 405 at 60% of laser power and 100 iterations every 30 seconds.

Supplementary Movie 3. 4 Time-lapse series demonstrating photoconversion and trafficking of cytosolic-targeted EGFP in a *N. benthamiana* epidermal cell.

Confocal imaging of cells expressing cytosolic EGFP in *N. benthamiana* was performed to demonstrate photoconversion and trafficking of EGFP in the cytosol. The ROI (shown

with white circles) was irradiated with 405 at 70% of laser power and 30 iterations every 22 seconds.

3.6 References

Adam, V., Berardozzi, R., Byrdin, M. and Bourgeois, D. (2014) Phototransformable fluorescent proteins: Future challenges. *Curr. Opin. Chem. Biol.* **20**, 92-102.

Aiken, C.T., Tobin, A.J. and Schweitzer, E.S. (2004) A cell-based screen for drugs to treat Huntington's disease. *Neurobiol. dis.* **16**, 546-555.

Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H. and Miyawaki, A. (2002) An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc. Natl. Acad. Sci.* **99**, 12651-12656.

Bogdanov, A.M., Mishin, A.S., Yampolsky, I.V., Belousov, V.V., Chudakov, D.M., Subach, F.V., Verkhusha, V.V., Lukyanov, S. and Lukyanov, K.A. (2009) Green fluorescent proteins are light-induced electron donors. *Nat. Chem. Biol.* **5**, 459.

Conley, A.J., Joensuu, J.J., Menassa, R. and Brandle, J.E. (2009) Induction of protein body formation in plant leaves by elastin-like polypeptide fusions. *BMC biol.* **7**, 48.

Elowitz, M.B., Surette, M.G., Wolf, P.-E., Stock, J. and Leibler, S. (1997) Photoactivation turns green fluorescent protein red. *Curr. Biol.* **7**, 809-812.

Gurskaya N.G., Verkhusha V.V., Shcheglov A.S., Staroverov D.B., Chepurnykh T.V., Fradkov A.F., Lukyanov S., and Lukyanov K. (2006) Engineering of monomeric green-to-red photoactivable fluorescent protein induced by blue light. *Nat. Biotechnol*. **24**, 461-465.

Hanson, M.R. and Sattarzadeh, A. (2013) Trafficking of proteins through plastid stromules. *Plant cell* **25**, 2774-2782.

Köhler, R.H., Zipfel, W.R., Webb, W.W. and Hanson, M.R. (1997) The green fluorescent protein as a marker to visualize plant mitochondria *in vivo*. *Plant J.* **11**, 613-621.

Martiniere, A., Bassil, E., Jublanc, E., Alcon, C., Reguera, M., Sentenac, H., Blumwald, E. and Paris, N. (2013a) *In vivo* intracellular pH measurements in tobacco and *Arabidopsis* reveal an unexpected pH gradient in the endomembrane system. *Plant cell* **25**, 4028-4043.

Martiniere, A., Desbrosses, G., Sentenac, H. and Paris, N. (2013b) Development and properties of genetically encoded pH sensors in plants. *Front. Plant Sci.* **4**, 523.

Mathur, J., Radhamony, R., Sinclair, A.M., Donoso, A., Dunn, N., Roach, E., Radford, D., Mohaghegh, P.M., Logan, D.C. and Kokolic, K. (2010) mEosFP-based green-to-red photoconvertible subcellular probes for plants. *Plant Physiol.* **154**, 1573-1587.

Matsuda, A., Shao, L., Boulanger, J., Kervrann, C., Carlton, P.M., Kner, P., Agard, D. and Sedat, J.W. (2010) Condensed mitotic chromosome structure at nanometer resolution using PALM and EGFP-histones. *PloS one* **5**, e12768.

Reichhart, J. and Ferrandon, D. (1998) Green balancers. *Dros. Inf. Serv* **81**, 201-202.

Saha, R., Verma, P.K., Rakshit, S., Saha, S., Mayor, S. and Pal, S.K. (2013) Light driven ultrafast electron transfer in oxidative redding of green fluorescent proteins. *Sci. Rep.* **3**, 1580.

Shen, J., Zeng, Y., Zhuang, X., Sun, L., Yao, X., Pimpl, P. and Jiang, L. (2013) Organelle pH in the *Arabidopsis* Endomembrane System. *Mol. Plant* **6**, 1419-1437.

Subach, F.V., Patterson G.H., Manley S., Gillette J.M., Lippincott-Schwartz J., and Verkhusha V.V. (2009) Photoactivable mCherry for high-resolution two-color fluorescence microscopy. *Nat. Methods*. **6**, 153-159.

Tsutsui, H., Karasawa, S., Shimizu, H., Nukina, N. and Miyawaki, A. (2005) Semi-rational engineering of a coral fluorescent protein into an efficient highlighter. *EMBO Rep.* **6**, 233-238.

Wiedenmann, J., Ivanchenko, S., Oswald, F., Schmitt, F., Röcker, C., Salih, A., Spindler, K.-D. and Nienhaus, G.U. (2004) EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 15905-15910.

4 Comparative study of Zera-, elastin like polypeptide-, and hydrophobin-induced protein bodies in *Nicotiana benthamiana* leaves

Reza Saberianfar[1,2], Susanne E. Kohalmi[1], Rima Menassa[1,2]

[1]Department of Biology, University of Western Ontario, London, ON, Canada

[2]Agriculture and Agri-Food Canada, London, ON, Canada

A version of this chapter is being prepared for publication in BMC Biology journal.

4.1 Introduction

Plants have been accepted as a growing platform for production of a wide range of recombinant proteins including vaccines, antibodies, pharmaceuticals and industrial enzymes. The popularity of plants as "green bioreactors" is due to advantages such as scalability, cost and safety compared to other systems. Moreover, plants offer other benefits, including the ability to introduce post-translational modifications such as disulfide bonds and glycosylation and the low risk of mammalian pathogen contamination (Egelkrout *et al.*, 2012).

Transient expression in leaves of *Nicotiana benthamiana* has become one of the most popular platforms for recombinant protein production and offers several advantages including flexibility and rapid scalability for protein production (D'Aoust *et al.*, 2010). Despite the advantages of this system, two major problems still need to be resolved: low production yield of some short-lived proteins and lack of efficient purification methods (Conley *et al.*, 2011a). Introduction of fusion tags has had a significant impact in tackling both these problems. Fusion-based approaches were shown to improve protein yield, prevent proteolysis, increase solubility and stability of the recombinant proteins and ease the isolation and purification process of the target proteins (Tremblay *et al.*, 2011). Zera®, elastin-like polypeptides (ELP) and hydrophobins (HFBI) fusions have been recently used for recombinant protein production in plants (Conley *et al.*, 2011a; Khan *et al.*, 2012). These three tags share a number of common features, including increased accumulation levels of their fusion partners, compatibility with large-scale purification methods, and induction of the formation of round structures called protein bodies (PB) usually found in seeds. PBs are endoplasmic reticulum (ER) derived organelles involved in folding, assembly and stabilization of seed proteins during seed development (Torrent *et al.*, 2009b).

Zera® is derived from the N-terminal region of γ-zein, a prolamin usually found in maize. Zera consists of a non proline rich region with a Cys-Gly-Cys motif, a proline rich region containing eight Pro-Pro-Pro-Val-His-Leu repeats, and a proline-X sequence with four cysteine residues. Cysteine residues enable the inter-molecular disulfide bond formation between Zera molecules and the amphipathicity of the $(PPPVHL)_8$ region helps the self-

assembly of Zera molecules, both of which are necessary for Zera PB formation (Llop-Tous *et al.*, 2010; Torrent *et al.*, 2009b). Zera-induced PBs have been reported in several expression systems including mammalian cultured cells (Chinese hamster ovary), plant leaves (*Nicotiana tabacum*), fungi (*Trichoderma reesei*), and insect cells (*Spodoptera frugiperda* (Sf9)) (Torrent *et al.*, 2009a). Zera-fused proteins can be purified using a simple density centrifugation method. Several recombinant proteins have been produced and purified using Zera fusions including epidermal growth factor (up to 0.5 g/kg LFW), human growth hormone (up to 3.2 g/kg LFW), and a *Streptomyces* derived xylanase (up to 1.6 g/kg LFW) (Llop-Tous *et al.*, 2010; 2011).

ELPs are composed of repeating pentapeptides Val-Pro-Gly-Xaa-Gly similar to repetitive pentapeptides found in mammalian elastin proteins. The guest amino acid (Xaa) can be any amino acid except proline (Urry and Parker, 2002). The size of ELP repeats can vary based on the experimental design (Conley *et al.*, 2009a). ELP-fused proteins can be purified using inverse transition cycling (ITC), a non-chromatographic protein purification method (Meyer and Chilkoti, 1999). ELP fusions have been successfully used to increase accumulation levels of interleukin-10 (up to 4.5% TSP) and green fluorescent protein (GFP) (up to 11% TSP) in *N. benthamiana* leaves (Conley *et al.*, 2009b). Similar to Zera, the ELP tag induces PB formation and allows higher accumulation of recombinant proteins (Conley *et al.*, 2009b).

Hydrophobins are a family of amphipathic proteins found in filamentous fungi. Eight cysteine residues positioned in a special pattern throughout the hydrophobin sequence are responsible for the formation of four intramolecular disulfide bridges which maintain the structural stability of the molecule. This enables hydrophobins to self-assemble at hydrophilic-hydrophobic interfaces (Joensuu *et al.*, 2010a). Similar to ELP, hydrophobin's properties can be exploited for purification using a surfactant-based aqueous two-phase separation system (ATPS) (Linder *et al.*, 2004). HFBI was used as a fusion tag to increase the accumulation levels of glucose oxidase (GOx), an enzyme that could not be produced in other conventional expression systems (Bankar *et al.*, 2009; Joensuu *et al.*, 2010b). Accumulation levels of GFP-HFBI reached up to 51% TSP when transiently expressed in *N. benthamiana* leaves, and the recombinant protein fusion was

found in PBs (Joensuu *et al.*, 2010a). Interestingly, the formation of PBs was correlated with a reduction in tissue necrosis, and it was hypothesized that PBs protect the leaf tissue from toxicity of highly expressed proteins (Conley *et al.*, 2011b; Joensuu *et al.*, 2010a).

The process by which PBs are induced by the fusion tags have been studied to some extent but due to the complexity of this process many questions remain unanswered. It has been shown that recombinant protein concentration is a critical element for PB formation in transgenic *N. tabacum* leaves, and by transient expression in *N. benthamiana* leaves. A threshold accumulation level of the recombinant protein (at 0.2% of total soluble protein) was found to be required for PB initiation (Gutiérrez *et al.*, 2013; Llop-Tous *et al.*, 2010) (Chapter 2). PB formation is shown to be a non-selective mechanism and proteins targeted to the secretory pathway are passively sequestered to ELP- and HFBI-induced PBs. It has also been shown that fusion tags are not necessarily required for PB formation but their presence affects the size and distribution patterns of PBs (Chapter 2). Interestingly, PBs induced by the Zera, ELP and HFBI fusion tags are not exactly similar. The Zera-induced PBs are more electron dense compared to ELP or HFBI PBs (Conley *et al.*, 2009b; Joensuu *et al.*, 2010a; Torrent *et al.*, 2009b), and ELP-induced PBs are larger in size compared to HFBI or Zera PBs (Chapter 2). The similarities and differences between the various fusion-tag-induced PBs are not fully understood and require further characterization.

In this study I compared PBs induced by Zera, ELP and HFBI fusion tags. I localized secretory and ER-targeted proteins to Zera-induced PBs and found that localization patterns are different from ELP- and HFBI-induced PBs. I confirmed the ER origin of all the fusion-tag-induced PBs. My results show that ELP- and HFBI-induced PBs are not terminally stored cytosolic organelles but actively communicating with the ER.

4.2 Results

4.2.1 Secretory and ER-targeted proteins are sequestered into Zera-induced PBs, but their localization pattern is different from ELP- or HFBI-induced PBs

A proteomic study of Zera-induced PBs showed the presence of secretory pathway proteins in PBs (Joseph *et al.*, 2012). I showed in chapter 2 that PBs induced by ELP and HFBI also can capture secretory and ER-targeted proteins. This property of PBs was used as a tool to increase accumulation levels of low-accumulating proteins such as erythropoietin (EPO) and interleukin-10 (IL-10).

To test if recombinant proteins can also be targeted into Zera-induced PBs, I co-expressed secretory GFP and ER-targeted GFP with Zera-DsRed. When expressed alone, secretory GFP localized to the apoplastic space between the cells (Figure 4.1a), ER-targeted GFP highlighted the ER and induced the formation of very small PBs along the ER network (Figure 4.1b), and Zera-DsRed gave rise to distinct PBs (Figure 4.1c). Upon co-expression of secretory GFP and Zera-DsRed, secretory GFP was found in Zera-induced PBs, but it did not mix with Zera-DsRed which localized to the center of PBs; GFP was seen at the periphery of PBs (Figure 4.1d-f). Co-expression of ER-targeted GFP and Zera-DsRed showed a similar pattern (Figure 4.1g-i) (Supplementary Movie 4.1). In these co-expression experiments, PBs were aligned along the ER network. This pattern is noticeable in the 4D compilation of Z-stack images (Supplementary Movie 4.1).

I showed in chapter 2 that co-expressing EPO with GFP-ELP or GFP-HFBI increased accumulation levels of EPO. To test if co-expression with Zera-DsRed would result in a similar increase, I co-expressed EPO with Zera-DsRed and found that accumulation of EPO was reduced (Figure 4.2).

Figure 4. 1 Secretory and endoplasmic reticulum-targeted GFP are sequestered in Zera induced PBs.

(a) Secretory GFP highlights the apoplastic space between the cells. (b) ER-targeted GFP forms small PBs along the ER network. (c) Zera-DsRed induces the formation of PBs. (d-f) Co-expression of secretory GFP and Zera-DsRed results in sequestration of secretory GFP into Zera-DsRed PBs. (g-i) Co-expression of ER-targeted GFP and Zera-DsRed results in sequestration of GFP into Zera-DsRed PBs. All images were acquired in sequential mode. Bar, 10 µm.

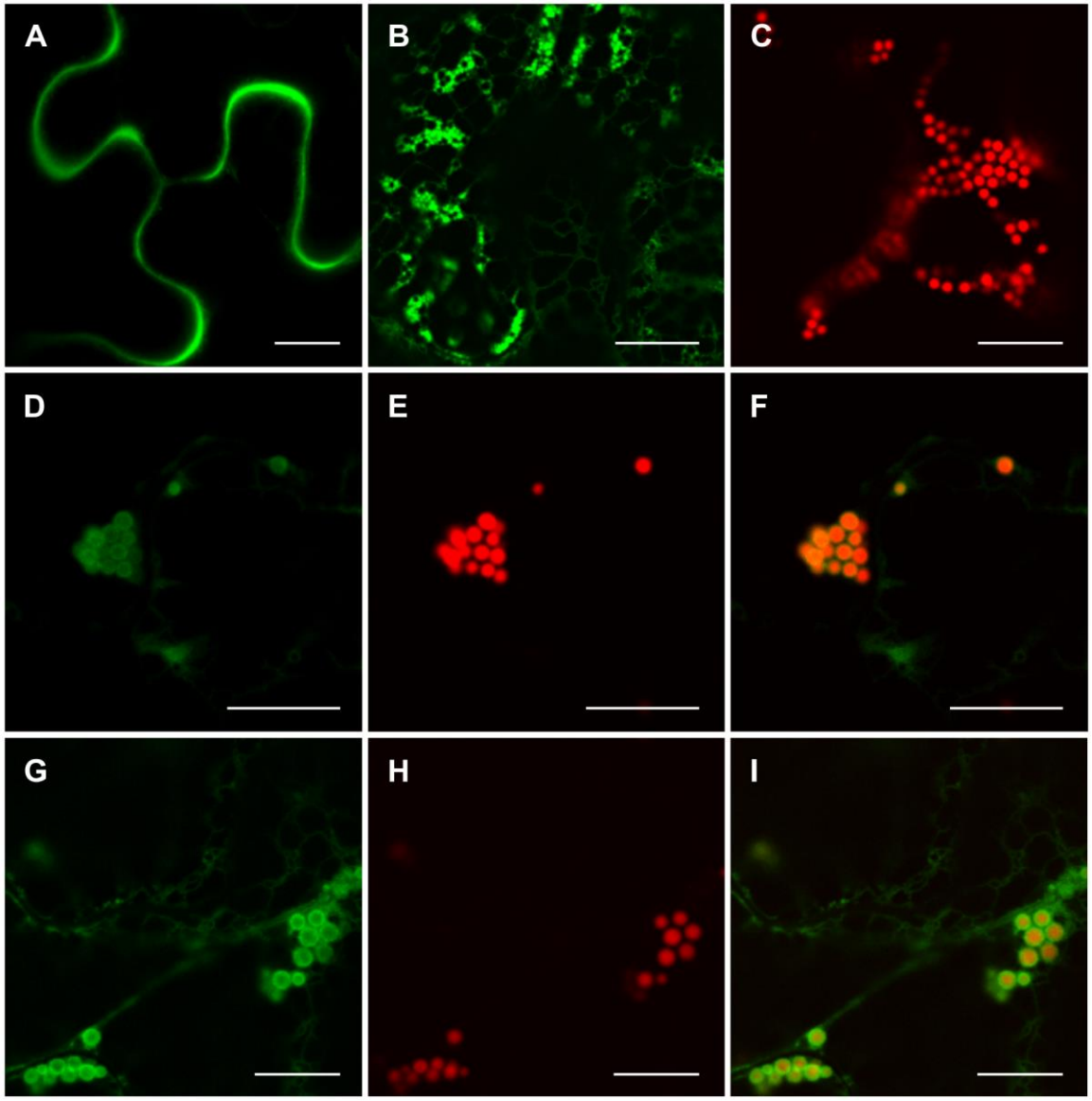
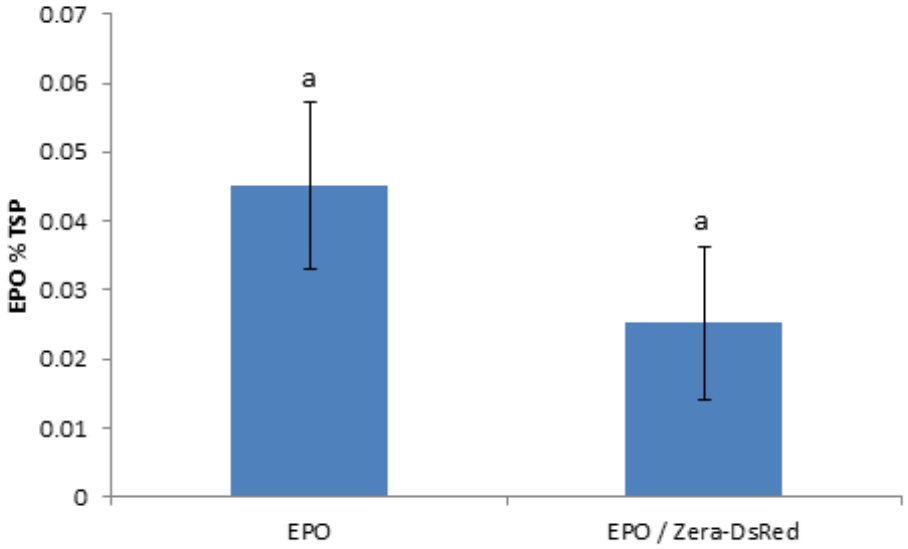

A
B
C
D
E
F
G
H
I

Figure 4. 2 Co-expression of erythropoietin (EPO) and Zera-DsRed.

Enzyme-linked immunosorbent assay (ELISA) of recombinant EPO was used for quantification of EPO in agroinfiltrated *N. benthamiana* leaves. Each column represents the mean value of 4 biological replicates. Columns denoted with the same letter were not significantly different ($p < 0.05$) using one-way ANOVA. The error bars represent the standard deviation of the mean.


0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
EPO % TSP
a
a
EPO
EPO / Zera-DsRed

4.2.2 ELP- and HFBI-fused proteins can be targeted to the same PBs unlike Zera-fused proteins

The previous comparison of PBs induced by ELP and HFBI showed that ELP-induced PBs were larger in size compared to HFBI-induced PBs (Chapter 2). Zera-induced PBs were shown to range mostly between 0.5 and 1 µm (and in some cases as large as 2 µm) (Joseph *et al.*, 2012), which is smaller than the average size ELP-induced PBs.

To characterize the relationship or distinctness of Zera-, ELP- and HFBI-induced PBs co-expression analyses were performed. When GFP-ELP, RFP-HFBI, GFP-HFBI or Zera-DsRed were expressed alone, each induced the formation of PBs (Figure 4.3a, 4.3b, 4.3c, and 4.3d, respectively). ELP-induced PBs were larger in size compared to HFBI- or Zera-induced PBs as expected. Upon co-expression of GFP-ELP and RFP-HFBI, both proteins co-localized into the same PBs (Figure 4.3e-g). This result suggests that the two fusion-induced PBs might share a similar origin of sequestering proteins.

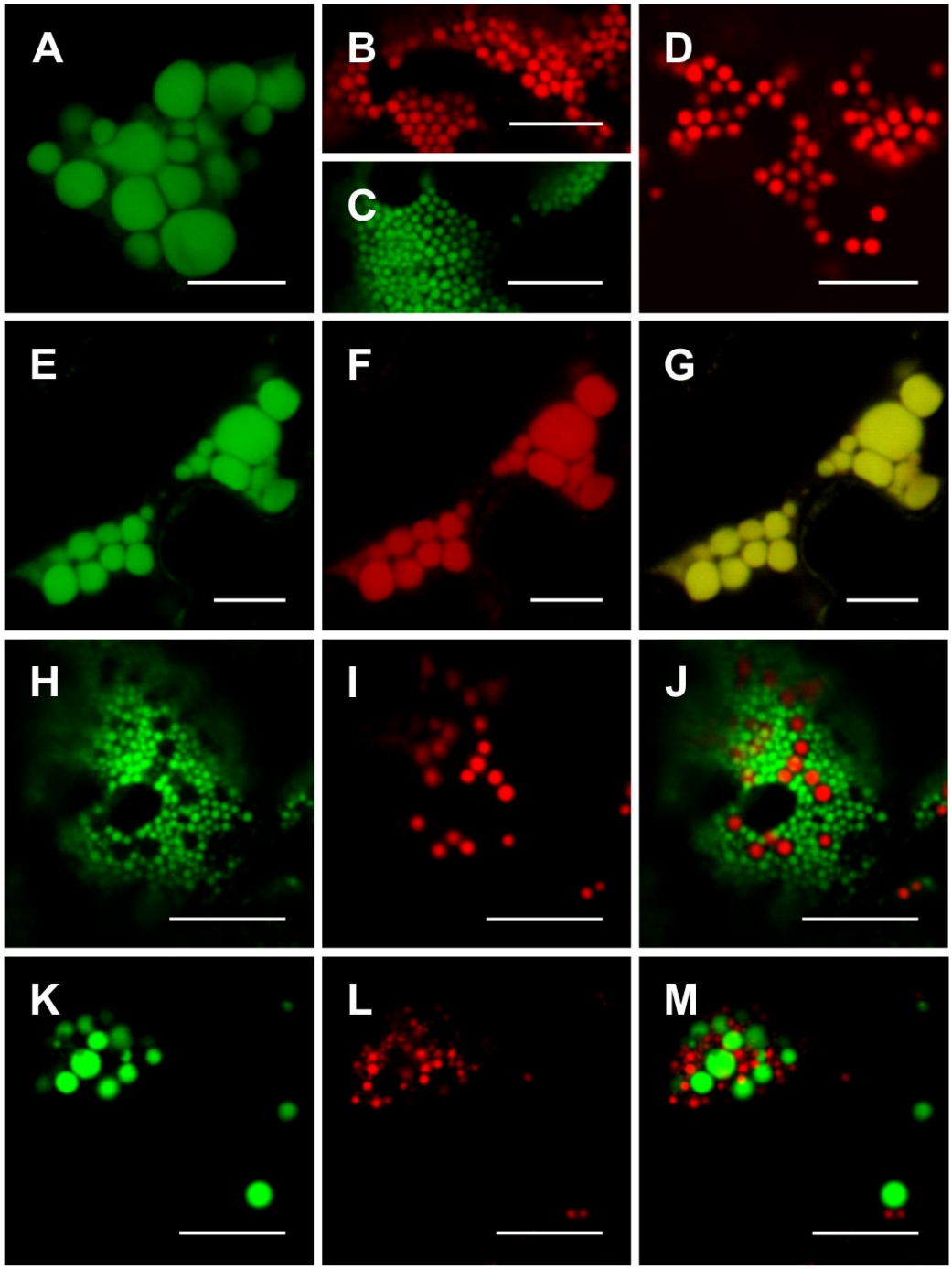
Unlike ELP and HFBI, co-expression of either fusion tag with Zera did not result in co-localization of the proteins into the same PBs, and gave rise to the formation of separate PBs. This can be easily seen as the gaps seen between GFP-HFBI PBs were filled with Zera-DsRed PBs (Figure 4.3h-j) (Supplementary Movie 4.2). Co-expression of GFP-ELP and Zera-DsRed proteins resulted in the formation of large GFP-ELP-induced PBs and smaller distinct Zera-DsRed-induced PBs (Figure 4.3k-m) (Supplementary Movie 4.3).

4.2.3 ELP-, HFBI- and Zera-induced protein bodies are surrounded by ER membrane

Previous reports have suggested an ER origin for the fusion-tag-induced PBs. ELP, HFBI and Zera were shown by transmission electron microscopy to be surrounded with a membrane studded with ribosomes therefore assumed to be an ER-derived membrane (Conley *et al.*, 2009b; Joensuu *et al.*, 2010a; Joseph *et al.*, 2012). Joseph *et al.*, also claimed the existence of an ER membrane around Zera-induced PBs. To ascertain if PBs induced by ELP, HFBI and Zera share an ER origin, I co-expressed RFP-ELP, RFP-HFBI and Zera-DsRed with the same construct used by Joseph *et al.* (2012) in which an ER transmembrane domain (C-terminus) of squalene synthase 1 (SQS1) of *Arabidopsis*

Figure 4. 3 Co-expression of ELP-, HFBI-, and Zera-fused fluorescent proteins.

(a) GFP-ELP promotes the formation of large PBs, (b) RFP-HFBI forms clusters of PBs, (c) GFP-HFBI forms clusters of PBs, and (d) Zera-DsRed induces the formation of PBs when expressed alone. (e-g) Co-expression of GFP-ELP and RFP-HFBI results in co-localization of the recombinant proteins into the same PBs. (h-j) Co-expression of GFP-HFBI and Zera-DsRed forms separate PBs. (k-m) Co-expression of GFP-ELP and Zera-DsRed forms distinct PBs. All images were acquired in sequential mode. Bar, 10 μm.


A
B
C
D
E
F
G
H
I
J
K
L
M

was fused to GFP (GFP-SQS1) at the N-terminus facing the cytosolic side of the ER membrane. The C-terminus region of SQS is comprised of a highly hydrophobic amino acid sequence, which helps to anchor the protein into the ER membrane (Kribii *et al.*, 1997). GFP-SQS highlights the ER membrane when expressed alone (Figure 4.4a), and RFP-ELP (Figure 4.4b), RFP-HFBI (Figure 4.3b) and Zera-DsRed (Figure 4.3d) induced the formation of PBs. Upon transient co-expression of RFP-ELP, RFP-HFBI and Zera-DsRed with GFP-SQS1, PBs were surrounded with a distinctive ER membrane highlighted by GFP-SQS1 (Figure 4.4c-e, f-h, i-k). In cases where several PBs were aligned side by side, the ER membrane showed a continuous (uninterrupted) pattern around each PB. It is also important to note that the membrane surrounding PBs appears to be continuous with the ER surrounding the cluster of PBs (Figure 4.4e, 4.4h, and 4.4k).

4.2.4 Protein bodies communicate with one another *in vivo*

The above results indicating the presence of an ER membrane around PBs that may be connected to the ER network prompted me to investigate if PBs completely bud off and leave the ER as terminally stored cytosolic organelles or if they remain connected with the ER, allowing them to exchange their content with other PBs and the surrounding ER.

Co-expression of YFP-KDEL with Zera-induced PBs previously had shown the localization of YFP to the periphery of Zera-induced PBs. Photobleaching of YFP in these PBs by using the fluorescence recovery after photobleaching (FRAP) technique resulted in rapid recovery of YFP, within 80 seconds, indicating the connection of Zera PBs with the ER (Joseph *et al.*, 2012). Conley *et al.* (2009b) also had shown the recovery of GFP-ELP-induced PBs after photobleaching within 5 minutes suggesting the trafficking of GFP from other parts of the cell to the bleached PB. Since the FRAP technique used in these experiments did not allow visualization of the photobleached protein, I used instead photoconversion, as described in chapter 3, which enabled me to track the photoconverted protein and also measure the fluorescence recovery with the newly trafficked protein into the irradiated area. Therefore, I used GFP fusions of HFBI and ELP to track the protein's movement once it was photoconverted in a specific PB or group of PBs (Figure 4.5 and 4.6).

Figure 4. 4 Protein bodies are surrounded with an endoplasmic reticulum-derived membrane.

(a) GFP-SQS1 highlights the ER membrane. The large circle highlights the nucleus. (b) RFP-ELP induces PBs clustered together. (c-e) Co-expression of GFP-SQS1 and RFP-ELP results in the formation of RFP-ELP PBs surrounded with the ER membrane highlighted by GFP-SQS1. (f-h) Co-expression of RFP-HFBI and GFP-SQS1 results in PBs surrounded by GFP-SQS1. (i-k) Co-expression of Zera-DsRed and GFP-SQS1 results in the formation of Zera-DsRed PBs surrounded by GFP-SQS1. All images were acquired in sequential mode. Bar, 5 μm.

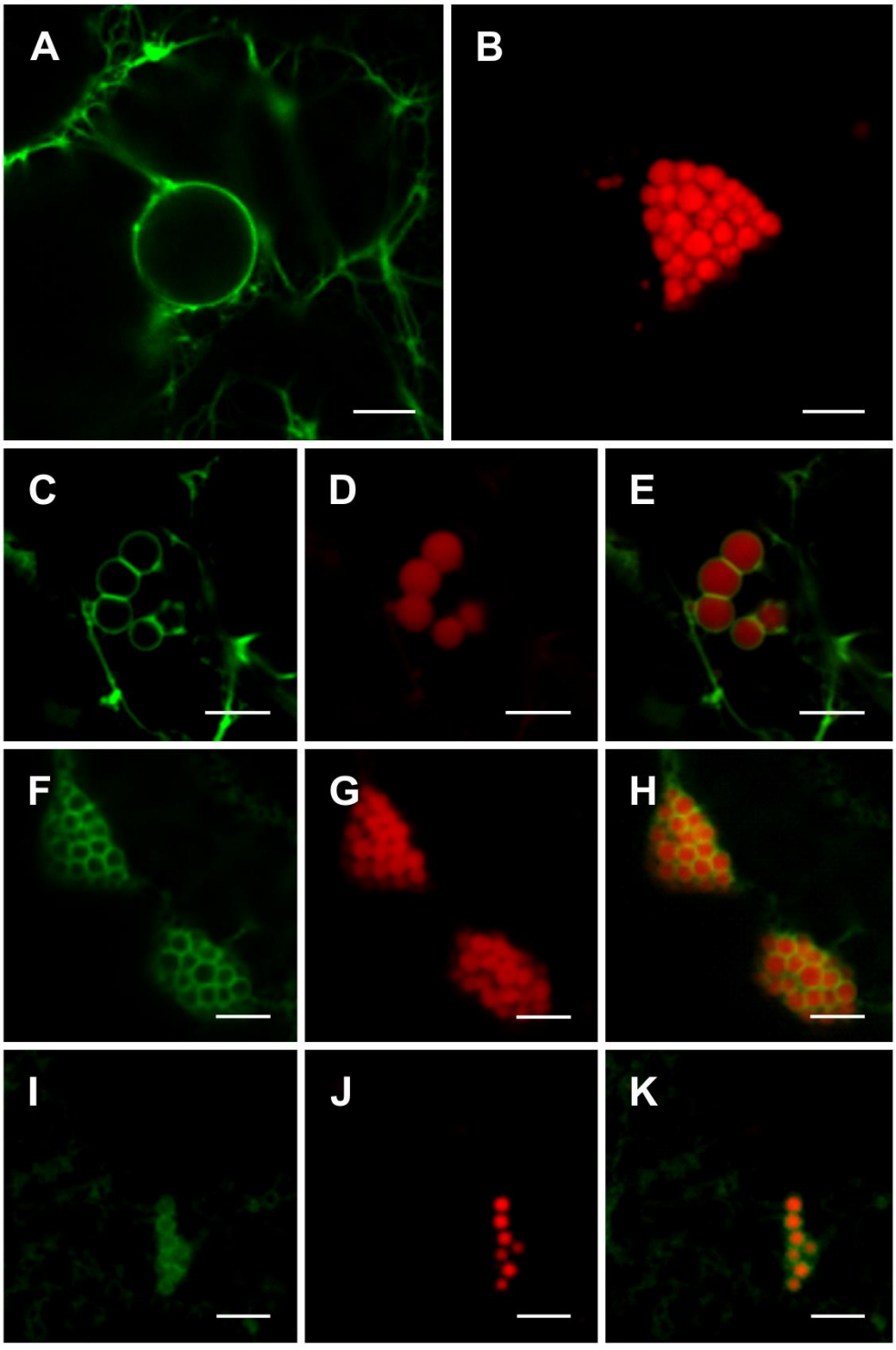
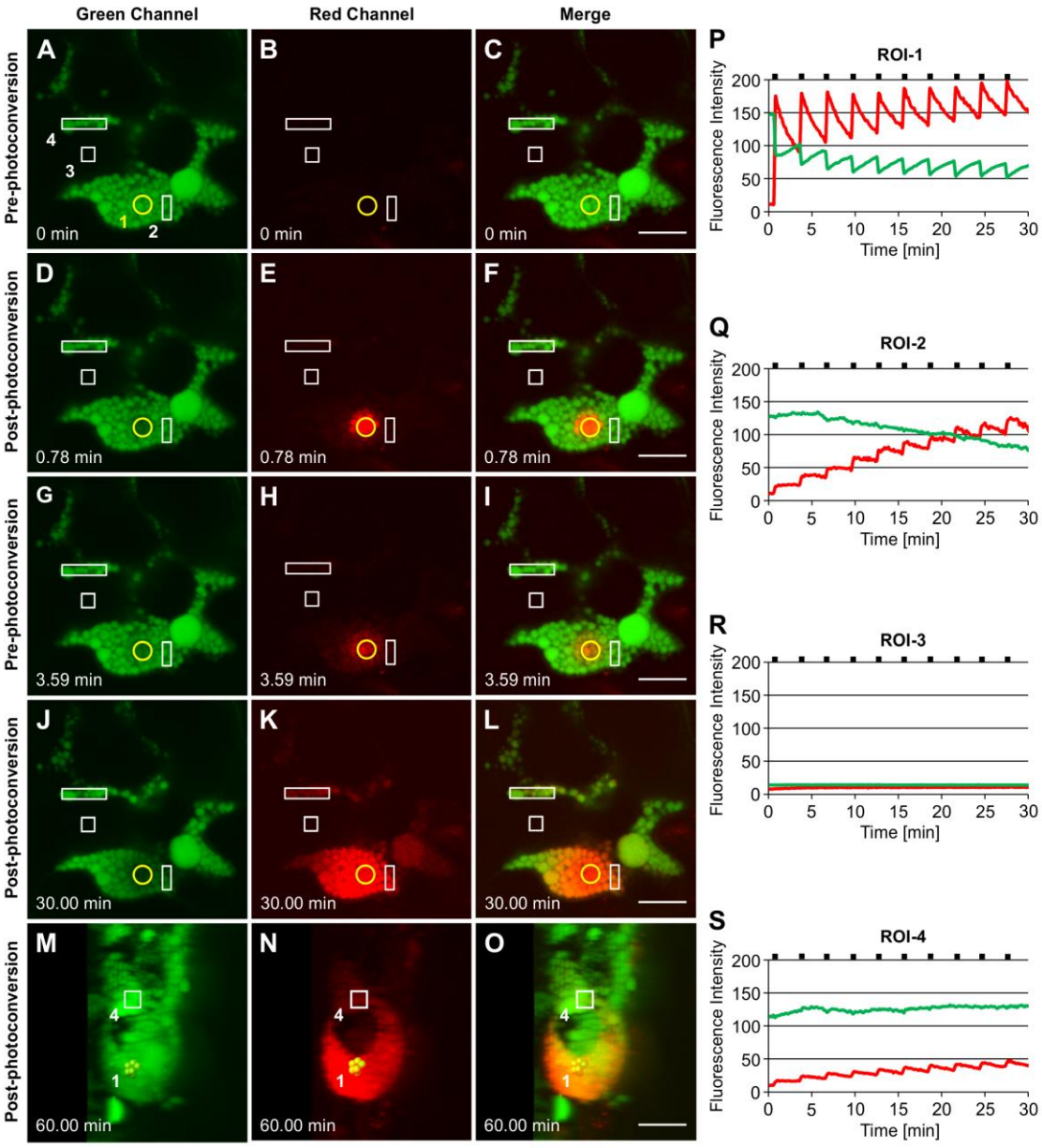

A
B
C
D
E
F
G
H
I
J
K

Figure 4. 5 Trafficking of proteins between GFP-HFBI-induced protein bodies.

(a-c) GFP-HFBI-induced PBs form clusters and can only be visualized in the green channel before photoconversion. The yellow circle represents the region of irradiation (ROI-1). Regions of interest (ROI 2-4) are shown in white. (d-f) Upon irradiation of ROI-1, the GFP in PBs within the circle photoconverts to red and starts spreading into neighboring PBs. (g-i) Within approximately 3 minutes from the first irradiation, GFP fluorescent recovers at ROI-1 and the red signal is spread to neighboring and distant PBs more than 10 µm away from ROI-1. (j-l) After multiple irradiations (with 3 minute time intervals) of ROI-1, sufficient red fluorescence is produced that shows the trafficking of the photoconverted protein from ROI-1 to ROI-4 and PBs further away (approximately 40 µm from ROI-1). (m-o) 3D representation of the PBs from a 90 degree perspective. PBs irradiated in ROI-1 are highlighted in yellow. ROI-4 is shown with a square. PBs between ROI-1 and ROI-4 are linked with a chain of PBs at a different focal plane within the cell. (p-s) Changes in fluorescence intensity of the green and red signal in ROI-1 (p), ROI-2(q), ROI-3 (r) and ROI-4 (s). Black bars indicate the irradiation period. All images were acquired in sequential mode. Bar, 10 µm.


Green Channel
Red Channel
Merge
Pre-photoconversion
Post-photoconversion
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
0 min
0.78 min
3.59 min
30.00 min
60.00 min
1
2
3
4
P
Q
R
S
ROI-1
ROI-2
ROI-3
ROI-4
Fluorescence Intensity
Time [min]
0
50
100
150
200
0
5
10
15
20
25
30

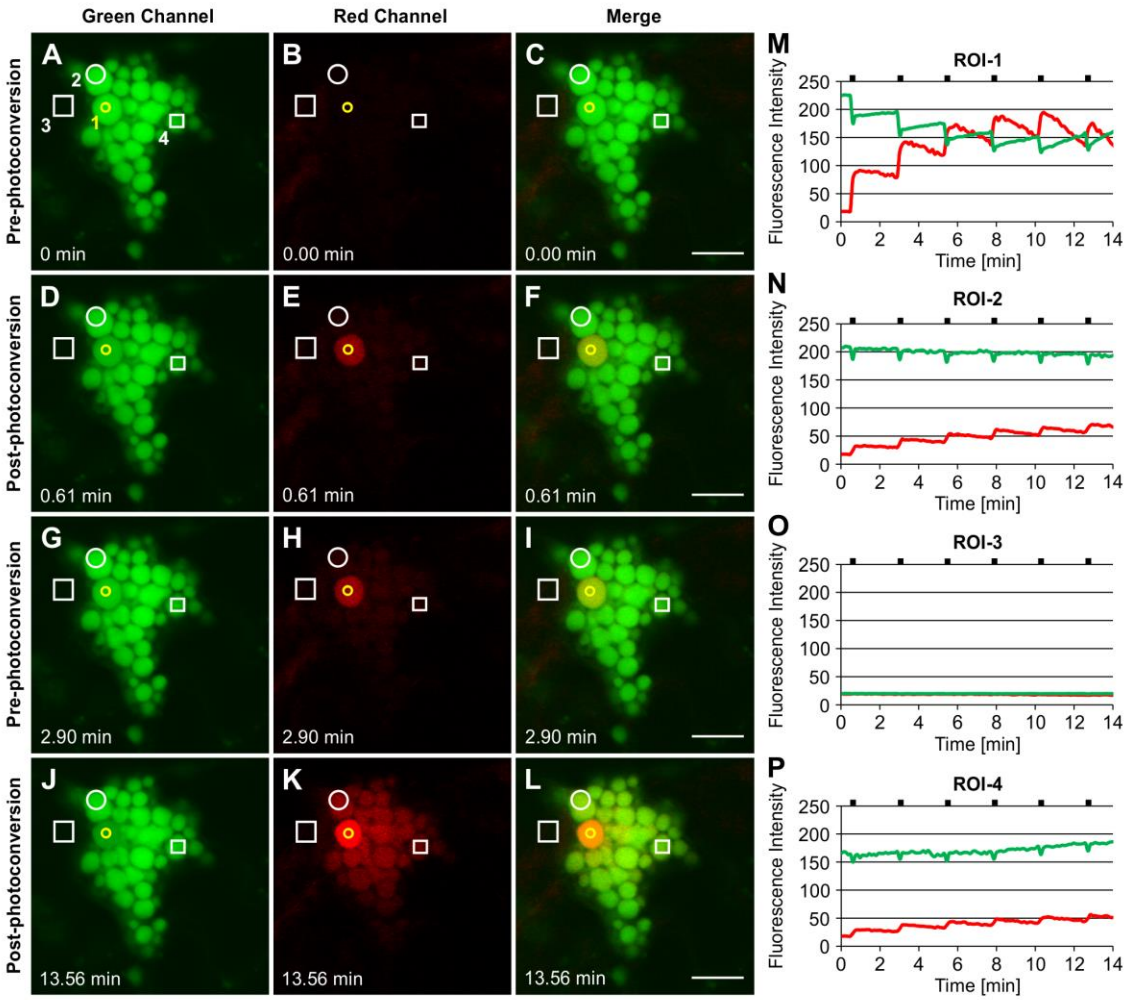
Irradiation of GFP-HFBI PBs resulted in immediate photoconversion of the irradiated PBs and their neighboring PBs in and around the region of irradiation-1 (ROI-1) (Figure 4.5a-f). This is also demonstrated by the sudden decrease of the green fluorescence and the simultaneous increase of red fluorescence at ROI-1 (Figure 4.4p). Photoconverted protein (red-state) continued to spread into the surrounding PBs. Approximately 3 minutes after the first irradiation, fluorescence intensity of the red signal at the ROI-1 was decreased by half (Figure 4.5p) as the photoconverted red signal was trafficked to the adjacent PBs whereas GFP increased, presumably due to trafficking to the ROI-1 (Figure 4.5g-i and 4.5p). I noticed that by repetitive irradiation of GFP trafficked to ROI-1 with 3 minutes intervals more photoconverted protein (red signal) could be produced which enabled me to track the signal as it travelled to the neighboring (ROI-2) and distant PBs (ROI-4) located more than 20 µm away from ROI-1 (Figure 4.5j-l, and 4.5s) (Supplementary Movie 4.4). It is important to note that the GFP at ROI-1 recovered gradually after every irradiation (Figure 4.5p). This recovery can be due to trafficking of GFP from neighboring PBs as shown in ROI-2 (Figure 4.5q) in which GFP gradually decreased over time and was replaced with the red signal coming from the ROI-1.

Interestingly, photoconverted protein only travelled to other PBs and through the ER. A perfect example is shown by appearance of the red signal in PBs at ROI-4, which did not seem to be connected to ROI-1. I noticed, however that PBs at ROI-4 were linked to ROI-1 PBs through a chain of PBs located in a different focal plane (Figure 4.5m-o) (Supplementary Movie 4.5). Considering no changes in neither green nor red fluorescence intensities at ROI-3, located in the area between ROI-1 and ROI-4 (Figure 4.5 a-l and 4.5r), I conclude that trafficking of proteins between PBs only happens through the ER connecting PBs.

Next I tested the possibility of trafficking of proteins among ELP-induced PBs (Figure 4.6). The ROI-1 in this case was much smaller than the large PB I irradiated (Figure 4.6a-f). Interestingly, upon irradiation of ROI-1, the entire PB photoconverted to red (Figure 4.6e) (Supplementary Movie 4.6) suggesting the high mobility of GFP-ELP within PBs similar to what was previously reported (Conley *et al.*, 2009b). The photoconverted signal gradually spread to neighboring PBs (ROI-2) while the green at ROI-1 recovered

Figure 4. 6 Trafficking of proteins between ELP-induced protein bodies.

(a-c) GFP-ELP-induced PBs form clusters and can only be visualized in the green channel before photoconversion. The yellow circle represents the region of irradiation (ROI-1). Regions of interest (ROI 2-4) are shown in white. (d-f) Photoconversion of the whole PB upon irradiation of a specific region within PB. (g-i) Spread of the photoconverted red signal from ROI-1 into neighboring (ROI-2) and distant (ROI-4) PBs within 2.3 minutes. (j-l) After multiple irradiations (with 2.3 minute time intervals) of ROI-1, sufficient red fluorescence is produced that shows the trafficking of the photoconverted protein from ROI-1 to other PBs. (m-p) Changes in fluorescence intensity of the green and red signal in ROI-1 (m), ROI-2(n), ROI-3 (o) and ROI-4 (p). Black bars indicate the irradiation period. All images were acquired in sequential mode. Bar, 10 μm.


Green Channel
Red Channel
Merge
Pre-photoconversion
Post-photoconversion
Pre-photoconversion
Post-photoconversion
A
B
C
D
E
F
G
H
I
J
K
L
1
2
3
4
0 min
0.00 min
0.00 min
0.61 min
0.61 min
0.61 min
2.90 min
2.90 min
2.90 min
13.56 min
13.56 min
13.56 min
M
N
O
P
ROI-1
ROI-2
ROI-3
ROI-4
Fluorescence Intensity
Time [min]
0
50
100
150
200
250
0
2
4
6
8
10
12
14

within approximately 2 minutes (Figure 4.6g-i and m-n). I repeated the photoconversion of ROI-1 with 2.3 minute intervals and monitored the trafficking of the red signal into neighboring (ROI-2) and distant (ROI-4) PBs (Figure 4.6j-l, n and p). I also noticed a sudden minor drop in the green fluorescence intensities at ROI-2 and ROI-4 upon irradiation of ROI-1 (Figure 4.6n and 4.6p, respectively). This might be due to immediate trafficking of GFP to the photobleached area at ROI-1 from the other PBs. Monitoring the fluorescence intensity at ROI-3 showed no changes of the green or red signal and therefore confirms the trafficking of the photoconverted protein amongst the PBs and not through the cytoplasm (Figure 4.6a-l and 4.6o).

To understand if protein trafficking occurs via PBs or if it occurs through the ER network I analyzed 3D images generated by Z-stacks of GFP-ELP acquired 45 minutes post-photoconversion (Figure 4.7). As expected, the red signal was observed in all the PBs clustered together at, above and below the focal plane at which the photoconverted PB was located (Figure 4.7b). I also noticed a number of PBs located away from the cluster of PBs which also turned red (denoted as 1 and 2 in Fig 4.7a-d) (Supplementary Movie 4.7). A high magnification 3D image of these PBs showed that they are surrounded by the ER network, which was continuous with the ER network around the cluster of PBs. I could only visualize the ER network in the green channel since the photoconverted red was not bright enough to highlight the ER (Figure 4.7c-d). It is likely that the ER works as a bridge between the cluster of PBs and the PBs located away from the cluster.

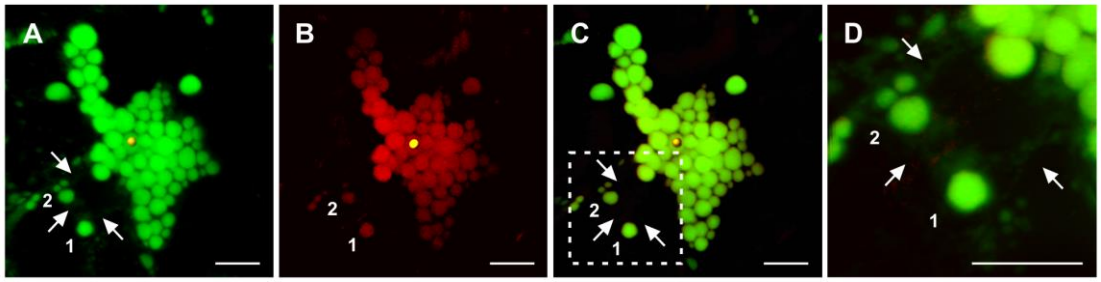
4.3 Discussion

4.3.1 Proteins targeted to the secretory pathway are sequestered passively into Zera-induced PBs

Previous proteomic analysis of Zera-induced PBs had shown the presence of ER resident proteins, such as Bip, calnexin and calreticulin, and secretory proteins such as cell wall proteins in these PBs (Joseph *et al.*, 2012). In agreement with these results, we found that upon co-expression of secretory GFP or GFP-KDEL with Zera, GFP was sequestered into Zera-induced PBs but mainly localized at the periphery of PBs. Interestingly, this pattern was different from what was previously seen by co-expression

Figure 4. 7 Protein bodies communicate with one another through the endoplasmic reticulum.

(a). 3D visualization of GFP-ELP PBs in the green channel. The region of irradiation (ROI) is shown by a yellow sphere. PBs marked as 1 and 2 are located away from the ROI and the cluster of PBs and connected to the rest of PBs only through the ER network shown with arrows. (b). 3D visualization of photoconverted GFP-ELP in the red channel. PBs marked as 1 and 2 photoconverted to red even though located far from the ROI. (c). Merged image of the green and red channels. (d). Close-up of the dotted square area in (c). Arrows point to the ER network. All images were acquired in sequential mode. Bar, 10 μm.


A
B
C
D
2
1
2
1
2
1
2
1

of secretory or ER-targeted GFP with ELP- or HFBI-induced PBs in which the GFP was evenly distributed throughout the PBs (Chapter 2). Considering that both secretory GFP and GFP-KDEL ended up in Zera-induced PBs, even though localized to the periphery of PBs, I conclude that proteins targeted to the secretory pathway with a Pr1b signal peptide, are sequestered passively into Zera-induced PBs. It is important to notice that the physico-chemical properties of Zera protein seem to prevent other recombinant proteins from penetrating into the core of PBs. This observation is in agreement with the study of molecular dynamics of Zera by Llop-Tous *et al.* (2010) in which they proposed that Zera molecules exhibit a sticklike conformation, and that the amphipathicity of the $(PPPVHL)_8$ repeat regions of Zera imposes lateral protein-protein interaction among Zera molecules and therefore hydrophobic packing of Zera-PBs (Llop-Tous *et al.*, 2010). This might be the reason why the GFP molecules were pushed away from the core of Zera PBs and seen as circles surrounding Zera-DsRed.

4.3.2 Co-expression of low accumulating proteins with Zera-induced PBs is not as efficient as with ELP- or HFBI-induced PBs

Co-expression of proteins targeted to the secretory pathway with PB-inducing fusion tags such as ELP and HFBI was found to increase the accumulation levels of low accumulating proteins (Chapter 2) (Joensuu *et al.*, 2014). It is assumed that PBs prevent recombinant protein degradation and also protect the cell from potential toxic effects of high level recombinant protein accumulation (Joensuu *et al.*, 2014). This idea can be applied to co-expression of unfused, low accumulating proteins with PB-inducing fusions. This can also be done by addition of an ER retrieval signal (KDEL/HDEL) to maximize the amount of ER-retrieved proteins since PB formation initiates in the ER. Indeed, co-expression of EPO with ELP- and HFBI- PBs-inducing fusions showed an increase in EPO accumulation levels in both cases, with HFBI-induced PBs being more effective and increasing EPO levels up to three-fold. Similarly, co-expression of IL-10 with HFBI a PB-inducing fusion increased IL-10 accumulation levels almost up to three-fold (Chapter 2).

In this study I co-expressed EPO with a PB-inducing Zera fusion and unlike previous results with ELP- and HFBI-induced PBs, I did not observe an increase in EPO accumulation levels. I hypothesize that the physico-chemical properties of Zera prevent efficient integration of EPO molecules into the lumen of Zera-induced PBs, similar to what was observed with the co-expression of GFP and Zera. Co-expression experiments by using PB induction seem to be more efficient by using HFBI or ELP compared to the Zera fusion tag.

4.3.3 Proteins can be targeted to different PBs

I found that ELP and HFBI fusion proteins co-localize to the same PBs while these protein fusions do not co-localise with zera fusions. The mechanism by which ELP, HFBI and Zera fusion tags induce the formation of separate PBs is not understood. Considering that Zera-fused proteins accumulate in the ER and form PBs without addition of the conventional ER retrieval signal suggests the possibility of Zera PBs originating from separate ER subdomains (Lynes and Simmen, 2011; Staehelin, 1997). Several previous reports of seed storage proteins also have shown the possibility of the formation of prolamin and globulin-like glutelins on separate subdomains of the ER (Choi *et al.*, 2000; Hamada *et al.*, 2003a; Hamada *et al.*, 2003b).

It is also possible that the reason behind the formation of Zera PBs separate from ELP and HFBI PBs might lie within the physico-chemical properties of the Zera molecules and their very strong affinity towards one another. Zera is a derivative of gamma zein which belongs to a family of maize seed storage proteins called prolamins. Prolamins are known to induce PBs by forming large aggregates in the ER due to their hydrophobicity and disulfide bond formation. It was suggested that these prolamin-induced aggregates are excluded from transport to the Golgi complex by COP-II vesicles, due to their large size and therefore induce PB formation (Kawagoe *et al.*, 2005; Pompa and Vitale, 2006; Vitale and Ceriotti, 2004). The hydrophobic region of Zera molecules is shown to enable lateral protein-protein interactions, which result in the sticklike alignment of Zera molecules. This structure is additionally stabilized by intermolecular disulfide bonds formed between cysteine residues (Llop-Tous *et al.*, 2010). All of these features contribute to PB formation and simultaneously prevent the integration of other proteins

into the core of Zera PBs. For instance, Llop-Tous *et al.* have shown the possibility of targeting GFP into the core of Zera-induced PBs only by its fusion to a Zera fusion tag (Llop-Tous *et al.*, 2010). Therefore, the affinity of Zera molecules might prevent the integration of ELP- or HFBI-fused proteins to Zera PBs as well. This feature of these three fusion tags can potentially be used for simultaneous expression of different proteins *in vivo* and their targeting into the same or separate PBs based on our preference, especially since Zera, ELP and HFBI can be isolated and purified by different purification strategies (Linder *et al.*, 2001; Torrent *et al.*, 2009b; Urry and Parker, 2002).

4.3.4 Protein bodies remain part of the ER and communicate with one another

The mechanism behind PB formation and maturation is complex and unclear. Several assumptions have been made regarding the fate of PBs depending on their seed- or leaf-based origins (Chapter 2). PBs were originally found as storage organelles in plant seeds, but have also been induced in plant leaves and other eukaryotic species (Shewry and Halford, 2002; Torrent *et al.*, 2009a). PB formation in seeds initiates in the ER lumen and depending on the plant species they might remain within the lumen of the rough ER as dense aggregates or bud off the ER, bypass the Golgi apparatus and be absorbed by protein storage vacuoles (Galili, 2004; Khan *et al.*, 2012; Lending and Larkins, 1989; Müntz and Shutov, 2002).

Two major assumptions were made regarding fusion tag-induced PBs in leaves. Zera-induced PBs are believed to originate from the ER, go through maturation and remain as part of the ER (Llop-Tous *et al.*, 2010; Torrent *et al.*, 2009a). This is based on the morphological studies of Zera PBs by electron microscopy in which an ER-membrane was observed surrounding PBs at 7dpi (Llop-Tous *et al.*, 2010). On the other hand, ELP- and HFBI-induced PBs were suggested to bud off the ER while surrounded with an ER membrane and become terminally stored within the cytoplasmic space. This assumption was made based on the observation of membranes studded with ribosomes surrounding ELP and HFBI PBs, that PBs did not appear to be connected with the ER in electron microscopy micrographs, and because they were seen to be mobile and move along the actin cytoskeleton (Conley *et al.*, 2009b; Joensuu *et al.*, 2010a). In the same study by

Conley *et al.* (2009b) a modified GFP-ELP with N-glycosylation sites was shown to produce PBs and to be fully susceptible to digestion by peptide N-glycosidase F and endoglycosidase H, confirming the glycosylation pattern of typical ER-retained glycoproteins. Moreover, co-expression of this construct with a Golgi marker did not show any co-localization, which indicates that GFP did not proceed to the Golgi apparatus, and that the PBs did not originate in the Golgi. Conley *et al.* (2009b) also showed the recovery of GFP into ELP-induced PBs upon bleaching in FRAP experiments, although the reason for the recovery was not clear and was attributed to either trafficking from other PBs or synthesis by ribosomes found on the PB membrane.

In this chapter I have shown that all three types of fusion-tag-induced PBs were surrounded by a membrane originating in the ER, therefore providing proof for the ER origin of PBs. I have shown, for the first time, by using a FRAP technique based on photoconversion of GFP, the trafficking of proteins between PBs in ELP- and HFBI-induced PBs separately. I have also shown that trafficking of PBs happens via the ER. Based on these observations I propose that ELP and HFBI PBs form and remain within the ER. I also believe that PBs are not inert storage organelles as they actively exchange content with other PBs and the ER.

4.4 Conclusion

The main objectives of this study were to understand and compare the similarities and differences of PBs induced by Zera, ELP and HFBI fusion tags in leaves. My results confirmed the ER origin of all three fusion-tag-induced PBs. I have shown that Zera-fused proteins do not co-localize with ELP- or HFBI-fused proteins into the same PBs, while ELP and HFBI-fused proteins appeared in the same PBs. I used the GFP photoconversion approach to track the movement of protein between PBs *in vivo* for the first time, and my results confirmed that ELP- or HFBI-induced PBs are not terminally stored cytosolic organelles, and that ELP and HFBI PBs form and remain part of the ER and communicate with one another via the ER.

4.5 Experimental procedure

4.5.1 Construct design and cloning

Secretory GFP, GFP, GFP-ELP (Conley *et al.*, 2009b), GFP-HFBI (Joensuu *et al.*, 2010a), RFP-HFBI, and RFP-ELP (Chapter 2) plant expression vectors were previously published. Zera-DsRed, and GFP-SQS were generously provided by Dr. Dolors Ludevid (Joseph *et al.*, 2012).

4.5.2 Transient expression in *N. benthamiana* leaves

N. benthamiana plants were grown at 22˚C with a 16 hour photoperiod at a light density of 110 µmol m^{-2} s^{-1} for 7 weeks before infiltration. Plants were watered with the water soluble fertilizer (N:P:K = 20:8:20) at 0.25 g/L (Plant Products, Brampton, ON, Canada). *Agrobacterium tumefaciens* cultures were prepared as previously described (Chapter 2).

4.5.3 Tissue sampling and protein extraction

N. benthamiana leaf samples were collected at 4 dpi. Four leaf discs were collected from three biological replicates per treatment. Protein extraction and total soluble protein quantification was performed as previously described (Conley *et al.*, 2009b).

4.5.4 Recombinant protein quantification

Quantification of EPO was performed by sandwich ELISA as previously described (Conley *et al.*, 2009c). Four biological replicates were quantified per treatment. *N. benthamiana* leaf tissue infiltrated with p19 was used as control.

4.5.5 Confocal microscopy and image analysis

To visualize the leaf samples, the abaxial epidermal cells were imaged with a Leica TCS SP2 CLSM. GFP was imaged by excitation with a 488 nm argon laser and detection between 500-525 nm. RFP and DsRed were imaged by excitation with 543 nm He/Ne laser and detection between 553-630 nm and 550-600, respectively.

All photoconversion experiments were performed with a Zeiss LSM 510 confocal microscope. A 405 nm laser (50 mW at 100% power setting) was used for

photoconversion. PB photoconversion was performed as described in Chapter 3 by using 20-40% of the laser power with 30-40 iterations. Green-state GFP was imaged by excitation at 488 nm and detection between 500-525 nm, whereas red-state GFP was excited by 543 nm laser and detected at 580-640 nm. To visualize the movement of proteins *in vivo*, multiple iterations with time intervals were used to allow trafficking of new GFP into the region of irradiation. Images were processed with the Zeiss Zen software 2009. Z-stack confocal images were used to generate 3D images and videos by using Imaris® software (version 7.6.1, Bitplane AG, Switzerland).

In co-expression experiments, all images were acquired in the sequential mode to avoid crosstalk between the fluorescent channels. All images were acquired at 4 dpi.

4.5.6 Statistical analysis

Minitab Express™ software (Minitab Inc., PA, USA) was used to perform the statistical analysis. Kolmogorov-Smirnov test was used to confirm the normal distribution of the data. A one-way analysis of variance (one-way ANOVA) was performed followed by Tukey-Kramer's test to find the means significantly different from one another (statistical difference was defined as $p < 0.05$).

4.6 Appendices

Supplementary Movie 4. 1 Endoplasmic reticulum-targeted GFP is sequestered into Zera induced PBs.

Z-stack images of cells co-expressing ER-targeted GFP and Zera-DsRed were acquired by confocal microscopy and assembled into 4D illustration using the Imaris software. ER-targeted GFP was sequestered to Zera-induced PBs and also highlighted the ER. Zera PBs were visible in the red channel. Simultaneous visualization of both green and red channels showed that ER-targeted GFP surrounded Zera-DsRed PBs. The signals in green and red channels were detected with the Imaris software and highlighted to show the accurate localization of each signal. Z-stack images were acquired in sequential mode. Bar, 5 µm.

Supplementary Movie 4. 2 Co-expression of HFBI- and Zera-fused fluorescent proteins.

Co-expression of GFP-HFBI and Zera-DsRed induced the formation of separate PBs. GFP-HFBI PBs were visualized in the green channel and Zera-DsRed were detected in the red channel. Simultaneous visualization of both channels showed HFBI and Zera PBs as separate PBs. Z-stack images were acquired by confocal microscopy and used for 4D illustrations. Green and red signals were detected and highlighted with the Imaris software to show the accurate localization of PBs. Z-stack images were acquired in sequential mode. Bar, 5 µm.

Supplementary Movie 4. 3 Co-expression of ELP- and Zera-fused fluorescent proteins.

Co-expression of GFP-ELP and Zera-DsRed induced the formation of separate PBs. GFP-ELP PBs were visualized in the green channel and Zera-DsRed were detected in the red channel. Simultaneous visualization of both channels showed HFBI and Zera PBs as separate PBs. Z-stack images were acquired by confocal microscopy and used for 4D illustrations. Green and red signals were detected and highlighted with the Imaris software to show the accurate localization of PBs. Z-stack images were acquired in sequential mode. Bar changes between 3-10 µm depending on different magnifications.

Supplementary Movie 4. 4 Trafficking of proteins between GFP-HFBI-induced protein bodies.

Time-lapse confocal imaging represents the photoconversion and trafficking of GFP-HFBI between PBs. The yellow circle represents the region of irradiation (ROI-1) which was irradiated with 405 nm laser at 20% of laser power and 40 iterations every 3 minutes. The white rectangular and square boxes represent the control regions described in Figure 4. Photoconverted GFP-HFBI (red state) can be seen in the red channel. Green channel represents the GFP-HFBI. Simultaneous photoconversion of the GFP-HFBI from green to red can be easily seen in the merge channel. Bar, 10 µm.

Supplementary Movie 4. 5 GFP-HFBI traffics to distant protein bodies via the endoplasmic reticulum.

GFP-HFBI-induced PBs are shown approximately 1 hour post-photoconversion. The irradiated PBs were detected with the Imaris software and highlighted with the yellow spheres. Green channel represents the GFP-HFBI. Red channel shows photoconverted GFP-HFBI (red-state). Bar, 10 µm.

Supplementary Movie 4. 6 Trafficking of proteins between GFP-ELP-induced protein bodies.

Time-lapse confocal imaging represents the photoconversion and trafficking of GFP-ELP between PBs. The yellow circle represents the region of irradiation (ROI-1) which was irradiated with 405 nm laser at 40% of laser power and 30 iterations every 2.30 minutes. The white rectangular and square boxes represent the control regions described in Figure 4.6. Photoconverted GFP-ELP (red state) can be seen in the red channel. Green channel represents the GFP-HFBI. Simultaneous photoconversion of the GFP-ELP from green to red can be easily seen in the merge channel. Bar, 10 µm.

Supplementary Movie 4. 7 GFP-ELP traffics to distant protein bodies via the endoplasmic reticulum.

GFP-ELP-induced PBs are shown approximately 45 minutes post-photoconversion. The irradiated region within the PB was detected with the Imaris software and highlighted with a yellow sphere. Green channel represents the GFP-ELP. Red channel shows photoconverted GFP-ELP (red-state). Bar, 10 µm.

4.7 References

Bankar, S.B., Bule, M.V., Singhal, R.S. and Ananthanarayan, L. (2009) Glucose oxidase — An overview. *Biotechnol. Adv.* **27**, 489-501.

Choi, S.B., Wang, C., Muench, D.G., Ozawa, K., Franceschi, V.R., Wu, Y. and Okita, T.W. (2000) Messenger RNA targeting of rice seed storage proteins to specific ER subdomains. *Nature* **407**, 765-767.

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

Conley, A.J., Joensuu, J.J., Menassa, R. and Brandle, J.E. (2009b) Induction of protein body formation in plant leaves by elastin-like polypeptide fusions. *BMC Biol.* **7**, 48.

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

Conley, A.J., Mohib, K., Jevnikar, A.M. and Brandle, J.E. (2009c) Plant recombinant erythropoietin attenuates inflammatory kidney cell injury. *Plant Biotechnol. J.l* **7**, 183-199.

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

D'Aoust, M., Couture, M.M., Charland, N., Trépanier, S., Landry, N., Ors, F. and Vézina, L. (2010) The production of hemagglutinin-based virus-like particles in plants: A rapid, efficient and safe response to pandemic influenza. *Plant Biotechnol. J.* **8**, 607-619.

Egelkrout, E., Rajan, V. and Howard, J.A. (2012) Overproduction of recombinant proteins in plants. *Plant Sci.* **184**, 83-101.

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

Gutiérrez, S.P., Saberianfar, R., Kohalmi, S.E. and Menassa, R. (2013) Protein body formation in stable transgenic tobacco expressing elastin-like polypeptide and hydrophobin fusion proteins. *BMC Biotechnol.* **13**, 40.

Hamada, S., Ishiyama, K., Choi, S.B., Wang, C., Singh, S., Kawai, N., Franceschi, V.R. and Okita, T.W. (2003a) The transport of prolamine RNAs to prolamine protein bodies in living rice endosperm cells. *Plant Cell* **15**, 2253-2264.

Hamada, S., Ishiyama, K., Sakulsingharoj, C., Choi, S.B., Wu, Y., Wang, C., Singh, S., Kawai, N., Messing, J. and Okita, T.W. (2003b) Dual regulated RNA transport pathways to the cortical region in developing rice endosperm. *Plant Cell* **15**, 2265-2272.

Joensuu, J., Mustalahti, E. and Conley, A. (2014) Improving the production of foreign proteins. *US2014/0212926A1. US Patent*.

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

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

Joseph, M., Ludevid, M.D., Torrent, M., Rofidal, V., Tauzin, M., Rossignol, M. and Peltier, J.B. (2012) Proteomic characterisation of endoplasmic reticulum-derived protein bodies in tobacco leaves. *BMC Plant Biol.* **12**, 36.

Kawagoe, Y., Suzuki, K., Tasaki, M., Yasuda, H., Akagi, K., Katoh, E., Nishizawa, N.K., Ogawa, M. and Takaiwa, F. (2005) The critical role of disulfide bond formation in protein sorting in the endosperm of rice. *Plant Cell* **17**, 1141-1153.

Khan, I., Twyman, R.M., Arcalis, E. and Stoger, E. (2012) Using storage organelles for the accumulation and encapsulation of recombinant proteins. *Biotechnology J.* **7**, 1099-1108.

Kribii, R., Arró, M., Arco, A., González, V., Balcells, L., Delourme, D., Ferrer, A., Karst, F. and Boronat, A. (1997) Cloning and characterization of the *Arabidopsis thaliana* SQS1 gene encoding squalene synthase. *Eur. J. Biochem.* **249**, 61-69.

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

Linder, M., Selber, K., Nakari-Setala, T., Qiao, M., Kula, M.R. and Penttila, 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., Hyytia, T., Nakari-Setala, T. and Penttila, M.E. (2004) Efficient purification of recombinant proteins using hydrophobins as tags in surfactant-based two-phase systems. *Biochemistry* **43**, 11873-11882.

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

Llop-Tous, I., Ortiz, M., Torrent, M. and Ludevid, M.D. (2011) The Expression of a Xylanase Targeted to ER-Protein Bodies Provides a Simple Strategy to Produce Active Insoluble Enzyme Polymers in Tobacco Plants. *PLoS One* **6**, e19474.

Lynes, E.M. and Simmen, T. (2011) Urban planning of the endoplasmic reticulum (ER): how diverse mechanisms segregate the many functions of the ER. *Biochim. Biophys. Acta* **1813**, 1893-1905.

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

Müntz, K. and Shutov, A.D. (2002) Legumains and their functions in plants. *Trends Plant Sci.* **7**, 340-344.

Pompa, A. and Vitale, A. (2006) Retention of a bean phaseolin/maize γ-Zein fusion in the endoplasmic reticulum depends on disulfide bond formation. *Plant Cell* **18**, 2608-2621.

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

Staehelin, L.A. (1997) The plant ER: a dynamic organelle composed of a large number of discrete functional domains. *Plant J.* **11**, 1151-1165.

Torrent, M., Llompart, B., Lasserre-Ramassamy, S., Llop-Tous, I., Bastida, M., Marzabal, P., Westerholm-Parvinen, A., Saloheimo, M., Heifetz, P.B. and Ludevid, M.D. (2009a) Eukaryotic protein production in designed storage organelles. *BMC Biol.* **7**, 5.

Torrent, M., Llop-Tous, I. and Ludevid, M.D. (2009b) Protein body induction: a new tool to produce and recover recombinant proteins in plants. *Methods Mol. Biol.* **483**, 193-208.

Tremblay, R., Diao, H., Huner, N., Jevnikar, A.M. and Ma, S. (2011) The development, characterization, and demonstration of a novel strategy for purification of recombinant proteins expressed in plants. *Transgenic Res.* **20**, 1357-1366.

Urry, D.W. and Parker, T.M. (2002) Mechanics of elastin: molecular mechanism of biological elasticity and its relationship to contraction. *J. Muscle Res. Cell Motil.* **23**, 543-559.

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

5 General Discussion

Plants are on their way to becoming a major platform for production of recombinant proteins since they offer exclusive advantages compared to traditional expression systems such as microbial, yeast, insect or mammalian cell cultures. These advantages include post-translational modifications that allow proper folding and assembly of complex proteins, rapid scalability, and the absence of human pathogens (Conley *et al.*, 2011; Twyman *et al.*, 2013). Meanwhile, the low accumulation levels of foreign recombinant proteins in plants have remained the most challenging problem to be addressed (Ahmad *et al.*, 2010).

A critical factor determining the final yield of recombinant proteins in plants is their subcellular destination since targeting to appropriate compartments within the cell affects the structure and stability of heterologous proteins (Streatfield, 2007). Targeting to organelles such as chloroplasts, PSVs, or PBs was proposed as a strategy to avoid the exposure of foreign protein to the host's proteases, and also to protect the host against the potential harmful effects of high accumulation levels of the foreign protein (Benchabane *et al.*, 2008; Conley *et al.*, 2009; Joensuu *et al.*, 2010; Torrent *et al.*, 2009). PBs are ideal organelles for heterologous protein targeting since they efficiently serve both of these purposes and also offer other advantages. Being ER-derived organelles, PBs contain low amounts of proteases, which provides a protective environment for foreign proteins, have high chaperone content, which helps with folding and assembly of recombinant proteins, and are contained within a membrane that eliminates the potential toxic effects of highly accumulated foreign proteins in cells (Desai *et al.*, 2010; Joensuu *et al.*, 2010).

Although PBs are naturally found in plant seeds, it is also possible to induce PB formation in plant leaves, for instance by addition of Zera, ELP or HFBI fusion tags to the protein of interest (Conley *et al.*, 2011). The mechanism of PB formation and the role of fusion tags in PB induction are not well understood. The purpose of this study was to investigate the driving forces involved in this process.

5.1 A working model for PB formation

5.1.1 PB formation initiates upon reaching a threshold level

In this thesis, recombinant protein accumulation levels and simultaneous monitoring of the cells for PB formation suggested that a recombinant protein threshold level of 0.2% of TSP is required for PB formation as shown in a conceptual model in Figure 5.1. This observation was based on transient expression of fluorescent proteins (GFP, GFP-ELP, GFP-HFBI) followed by confocal microscopy, as well as non-fluorescent proteins (xyn11A, xyn11A-HFBI, and xyn10A) and subsequent immunocytochemistry in *N. benthamiana* leaves (Chapter 2, Figure 2.3, and 2.5). These results are in agreement with the threshold level previously observed for PB formation (at 0.2% of TSP) by expression of GFP, GFP-ELP and GFP-HFBI in transgenic tobacco leaves (Gutiérrez *et al.*, 2013).

It is important to note that although the presence of fusion tags is not necessary for PB formation, their presence affects the distribution pattern and size of PBs (Chapter 2, Figure 2.3). This can be due to the affinity of fusion tags towards one another, which assists with aggregate formation. Indeed, all three types of fusion tags share a common important property: the availability of hydrophobic regions, which enables fusion tags and their fusion partners to self-assemble into aggregates and therefore facilitate the process of PB formation (Kogan *et al.*, 2002; Linder, 2009; Llop-Tous *et al.*, 2010; Miao *et al.*, 2003).

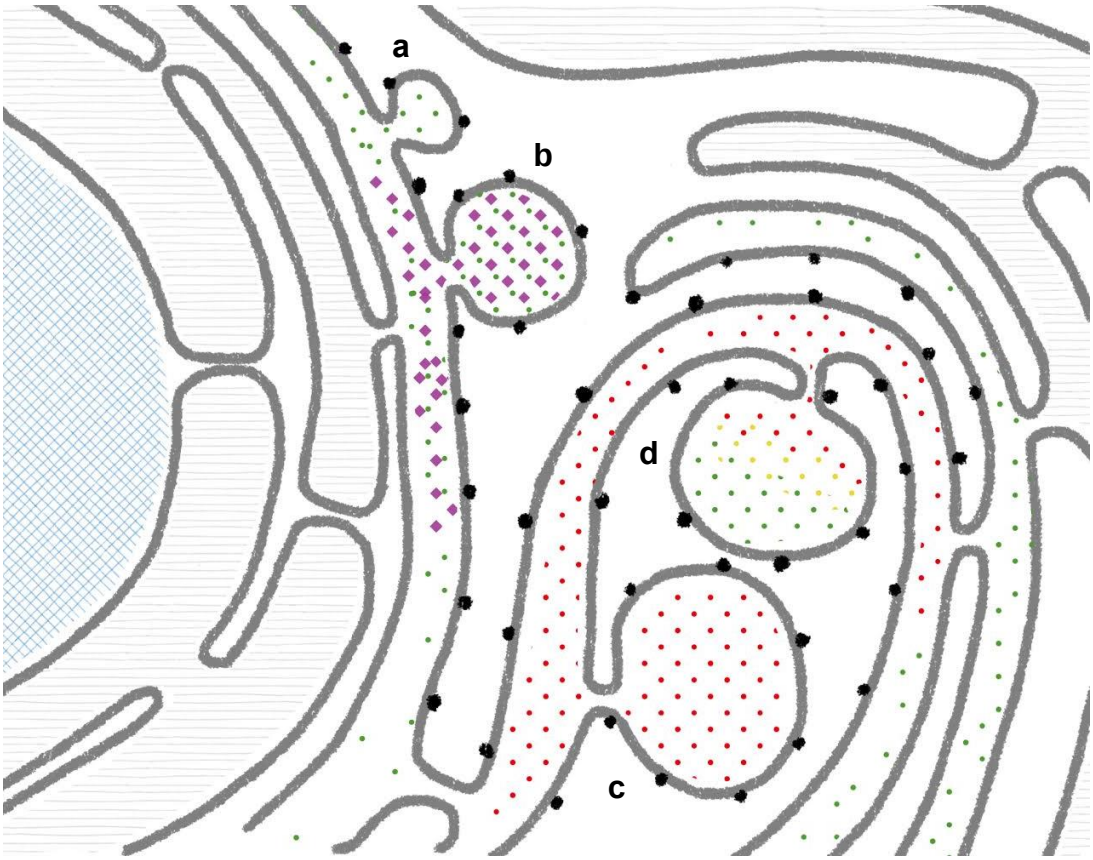
Once PBs start forming, they grow over time (Figure 5.1b). This is based on the observation of GFP-ELP- and GFP-HFBI-induced PBs from 3 to 5 dpi (Chapter 2, Figure 2.2). A similar observation was made for the size of Zera-ECFP-induced PBs from 2 to 10 dpi (Llop-Tous *et al.*, 2010). In all cases, PB size growth was accompanied by an increase in accumulation levels of the respective recombinant proteins.

5.1.2 Proteins are sequestered passively into PBs

A previous study of ELP-induced PBs showed the co-localization of the ER chaperone BiP in ELP-induced PBs (Conley *et al.*, 2009), and a proteomic study of Zera-induced PBs

Figure 5. 1 A working model of protein body formation and development.

(a) Proteins are synthesized on the rough ER by ribosomes and transferred into the ER lumen co-translationally. PBs form upon high concentrations of the recombinant proteins in the ER lumen. High concentration of proteins is represented by several GFP-HFBI molecules in an area. Once PBs form, they grow in size over time. (b) Co-expression of high value recombinant proteins (in this case EPO) with GFP-HFBI results in passive sequestration of EPO molecules into GFP-HFBI-induced PBs. (c) After reaching their final size, PBs remain connected with the ER and exchange their content with other PBs via the ER. In this example, GFP-HFBI photoconverts to red-state GFP-HFBI upon irradiation. Photoconverted proteins move from one PB to another through the ER. (d) Co-localization of green- and red-state GFP in the same PB is shown as yellow.


a
b
c
d
GFP-HFBI
EPO
Red-state GFP-HFBI
Co-localization of green and red GFP

also showed the presence of the ER chaperones BiP, calnexin and calreticulin in Zera PBs (Joseph *et al.*, 2012).

The results presented in this work confirm the sequestration into PBs of several proteins targeted to the secretory pathway with a signal peptide. For instance, secretory GFP and ER-targeted GFP both were sequestered into RFP-ELP-, RFP-HFBI-, xyn11A-, and xyn11A-HFBI-induced PBs upon co-expression (Chapter 2, Figure 2.4 and 2.6). Therefore, co-expression of valuable and low-accumulating proteins such as EPO and IL-10 with PB-inducing proteins such as GFP-HFBI was proposed as a strategy to increase the accumulation levels of low accumulating proteins (Chapter 2, Figure 2.7). As shown in figure 5.1b, EPO molecules are sequestered into the core of GFP-HFBI-induced PBs where they are prevented from progressing through the secretory pathway, and over time, more recombinant protein accumulates in PBs. This strategy is important for two reasons; it helps to increase the accumulation levels of difficult-to-express recombinant proteins, and eliminates the need for the addition of fusion tags which might affect the proper folding and activity of the protein of interest.

Although co-expression of EPO and IL-10 with GFP-ELP and GFP-HFBI increased their yield, co-expression with Zera-DsRed did not help with increasing accumulation levels (Chapter 4, Figure 4.2). This is attributed to the strong affinity of Zera molecules to one another, which results in a condensed and stick-like alignment of Zera molecules while forming protein aggregates, as described by Llop-Tous *et. al.* (2010), and therefore excluding other molecules from the core of PBs. This was also confirmed with co-expression of secretory GFP or ER-targeted GFP with Zera PBs in which the GFP signal localized to the periphery of PBs and not in their core (Chapter 4, Figure 4.1).

A patent application based on the co-expression technology using GFP-HFBI as a PB-inducing factor, was recently filed for the production of foreign recombinant proteins vulnerable to the intracellular and extracellular proteolytic activity of proteases (Joensuu *et al.*, 2014).

5.1.3 ER is the initiation point and the final destination of PBs

Previous studies on PBs have interpreted the subcellular localization of PBs differently. Zera-induced PBs were described as highly dense aggregates, which remain connected in the lumen of the ER rather than forming independent organelles (Llop-Tous *et al.*, 2010; Torrent *et al.*, 2009). Alternatively, ELP- and HFBI-induced PBs were described as terminally-stored storage organelles, which bud off the ER and are highly mobile in the cytosol (Conley *et al.*, 2009; Joensuu *et al.*, 2010). Conley *et al.* (2009) conducted FRAP experiments and found that fluorescence of the irradiated PB recovered within minutes after irradiation, and interpreted fluorescence recovery as due either to protein synthesis by ribosomes available on the PB membrane, or to trafficking of GFP from neighboring PBs into the irradiated area.

In this study I investigated whether ELP or HFBI PBs bud off the ER and become terminally stored cytosolic organelles or remain connected with the ER network. To address this question, a novel technique based on photoconversion of EGFP without requirement of additional treatments (e.g. oxygen scavengers) was described for the first time in this work. First, the ability of EGFP photoconversion from green state to red state was confirmed *in vitro* (Chapter 3, Figure 3.2). The photoconversion of EGFP was also shown to be irreversible. The technique was then tested on a range of different model organisms including tobacco suspension cell culture, *N. benthamiana*, *Drosophila* and rat cells, and proved functional for studying protein trafficking *in vivo* (Chapter 3, Figure 3.3, and 3.4). Considering the popularity of EGFP as a fluorescent protein marker in cell and molecular biology labs, and its widespread availability in different transgenic lines among several organisms, the ability to photoconvert this protein from green to red provides a valuable new tool to study protein trafficking *in vivo* without requirement of special treatments.

EGFP photoconversion was then tested on ELP- and HFBI-induced PBs. This technique allowed me to observe protein trafficking out of, and into, the irradiated PBs *in vivo* (Chapter 4, Figures 4.5 and 4.6). Results from photoconversion experiments of ELP and HFBI PBs suggested that these PBs are not terminally stored cytosolic organelles. In both cases irradiated proteins within PBs were first transferred to surrounding PBs, and then to

distant PBs within the cell (Figure 5.1c-d). All the photoconverted PBs were shown to be linked to each other through the ER (Chapter 4, Figure 4.7). Fluorescence recovery of irradiated PBs appears to be due to trafficking of GFP from surrounding PBs, although the role of ribosomes located on the PB membrane in synthesizing and transferring new protein to PBs cannot yet be ruled out.

5.2 Future prospects

It is not fully clear whether high accumulation of proteins causes the formation of PBs or if the formation of PBs is responsible for accumulation of high amounts of recombinant proteins. Understanding the mechanism of PB formation will help us to use this technology for increasing the yield of foreign proteins from plant expression systems which is a necessary requirement for developing plants as one of the major recombinant protein production platforms (Gutiérrez and Menassa, 2014).

The purpose of this thesis was to investigate the mechanism of PB formation and the factors involved in this process. I showed that PB formation is dependent on recombinant protein accumulation levels, and influenced by the addition of fusion tags. Co-expressed proteins targeted to the secretory pathway were passively sequestered into PBs, although localization of proteins targeted to the lumen of Zera PBs was different from their localization in ELP or HFBI PBs.

The differences between the fusion tags and how they direct their fused protein to PBs need further investigation. Zera fusion tag does not require an ER retrieval signal to accumulate in the ER and form PBs unlike ELP and HFBI (Conley *et al.*, 2012). I showed that Zera-induced PBs did not co-localize with either ELP- or HFBI-induced PBs. I also showed that all three types of PBs had ER origins. This behavior of Zera PBs is attributed to the strong affinity of Zera molecules to assemble together and avoid integration of other molecules. The possibility of Zera-induced PBs originating from a different subdomain of the ER compared to ELP- or HFBI-induced PBs has yet to be assessed, for instance by co-expression with ER subdomain-specific proteins, several of which are described by Lynes and Simmen (2011) and Stefano *et al.* (2014). Although PBs induced by fusion tags show differences, this feature might be useful for

simultaneous expression of different proteins *in vivo* and their targeting into the same or separate PBs, especially since Zera, ELP and HFBI can be isolated and purified by different purification strategies.

It is yet to be assessed if Zera PBs exchange their content with other PBs and if they use the ER for protein trafficking between PBs in a similar pattern as ELP and HFBI PBs. This can be answered by photoconversion experiments using an EGFP-Zera fusion. Also, the role of *de novo* protein synthesis by ribosomes attached to the PB membrane of Zera-, ELP- and HFBI-induced PBs needs to be studied. This question can be addressed by a photoconversion-based FRAP experiment in the presence and absence of an inhibitor of protein synthesis (e.g. cycloheximide), in which fluorescence recovery is measured.

In conclusion, induction of PB formation enables the storage of high amounts of recombinant proteins in the limited intracellular space without imposing excessive stress to the ER, and may be a coping mechanism that eukaryotic cells have evolved to prevent ER stress and cell death. PBs work as extensions of the ER and provide sufficient space for accumulation of more protein. In this thesis, I tried to understand the controlling factors in charge of PB formation. I believe understanding this process will enable us to efficiently challenge the production bottleneck of low recombinant protein accumulation levels in plants especially leaf-based expression systems.

5.3 References

Ahmad, A., Pereira, E.O., Conley, A.J., Richman, A.S. and Menassa, R. (2010) Green biofactories: recombinant protein production in plants. *Recent Pat. Biotechnol.* **4**, 242-259.

Benchabane, M., Goulet, C., Rivard, D., Faye, L., Gomord, V. and Michaud, D. (2008) Preventing unintended proteolysis in plant protein biofactories. *Plant Biotechnol J* **6**, 633-648.

Conley, A.J., Joensuu, J.J., Menassa, R. and Brandle, J.E. (2009) Induction of protein body formation in plant leaves by elastin-like polypeptide fusions. *BMC Biol.* **7**, 48.

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

Desai, P.N., Shrivastava, N. and Padh, H. (2010) Production of heterologous proteins in plants: strategies for optimal expression. *Biotechnol. Adv.* **28**, 427-435.

Gutiérrez, S. and Menassa, R. (2014) Protein body inducing fusions for recombinant protein production in plants. In: *Plant-derived pharmaceuticals: principles and applications for developing countries* (Hefferon, K. ed) pp. 9-19. Wallingford, UK: CABI.

Gutiérrez, S.P., Saberianfar, R., Kohalmi, S.E. and Menassa, R. (2013) Protein body formation in stable transgenic tobacco expressing elastin-like polypeptide and hydrophobin fusion proteins. *BMC Biotechnol.* **13**, 40.

Joensuu, J., Mustalahti, E. and Conley, A. (2014) Improving the production of foreign proteins. *US2014/0212926A1. US Patent.*

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

Joseph, M., Ludevid, M.D., Torrent, M., Rofidal, V., Tauzin, M., Rossignol, M. and Peltier, J.B. (2012) Proteomic characterisation of endoplasmic reticulum-derived protein bodies in tobacco leaves. *BMC Plant Biol.* **12**, 36.

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

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

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

Lynes, E.M. and Simmen, T. (2011) Urban planning of the endoplasmic reticulum (ER): How diverse mechanisms segregate the many functions of the ER. *Biochim. Biophys. Acta* **1813**, 1893-1905.

Miao, M., Bellingham, C.M., Stahl, R.J., Sitarz, E.E., Lane, C.J. and Keeley, F.W. (2003) Sequence and structure determinants for the self-aggregation of recombinant polypeptides modeled after human elastin. *J Biol Chem* **278**, 48553-48562.

Stefano, G., Hawes, C. and Brandizzi, F. (2014) ER – the key to the highway. *Curr. Opin. Plant Biol.* **22**, 30-38.

Streatfield, S.J. (2007) Approaches to achieve high-level heterologous protein production in plants. *Plant Biotechnol. J.* **5**, 2-15.

Torrent, M., Llompart, B., Lasserre-Ramassamy, S., Llop-Tous, I., Bastida, M., Marzabal, P., Westerholm-Parvinen, A., Saloheimo, M., Heifetz, P.B. and Ludevid, M.D. (2009) Eukaryotic protein production in designed storage organelles. *BMC Biol.* **7**, 5.

Twyman, R.M., Schillberg, S. and Fischer, R. (2013) Optimizing the yield of recombinant pharmaceutical proteins in plants. *Curr. Pharm. Des.* **19**, 5486-5494.

Curriculum Vitae

Reza Saberianfar

EDUCATION

2010-2014	Ph.D. Biology (Cell and Molecular Biology) The University of Western Ontario London, Ontario, Canada
2007-2009	M.Sc. Biology (Cell and Molecular Biology) The University of Western Ontario London, Ontario, Canada
2002-2006	B.Sc. Biology Tarbiat Moallem University of Tehran Tehran, Iran

MAJOR AWARDS AND SCHOLARSHIPS

2014	Dr. Irene Uchida's fellowship in life sciences
2014	Oral presentation award (second place), The 9th Canadian plant biotechnology conference,
2014	Robert and Ruth Lumsden graduate fellowship
2014	Department of Biology travel award
2013-2014	Ontario graduate scholarship (OGS)
2013	Ruth Horner Arnold fellowship
2013	Department of Biology travel award
2012	Best poster presentation (second place), Biology graduate research forum (BGRF)
2012	Best oral presentation award, The 8th Canadian plant biotechnology conference
2012	Dr. René R. Roth memorial award for academic and research excellence
2011	Best poster presentation award, Biology graduate research forum (BGRF)

2011	Best poster presentation award, The 8th Canadian plant genomics workshop
2007-2014	Western graduate research scholarship (WGRS)

WORK EXPERIENCE

2007-2014	Teaching Assistant, Department of Biology, The University of Western Ontario
2010-2014	Research Assistant, Agriculture and Agri-Food Canada
2007-2009	Research Assistant, Department of Biology, The University of Western Ontario

REFEREED PUBLICATIONS

R. Saberianfar, S. E. Kohalmi, and R. Menassa. Comparative study of Zera-, elastin-like polypeptide-, and hydrophobin-induced protein bodies in Nicotiana benthamiana leaves. (in preparation for submission to BMC Biology)

A. Sattarzadeh*, **R. Saberianfar***, W. R. Zipfel, R. Menassa and M. Hanson. Green to red photoconversion of GFP for protein tracking. (in preparation for submission to Nature Communications), (Authors shown with * contributed equally to this work)

R. Saberianfar, J. J. Joensuu, A. J. Conley and R. Menassa. Protein body formation in leaves of *Nicotiana benthamiana*: a concentration dependent mechanism influenced by the presence of fusion tags. Plant Biotechnology Journal. (in revision)

S. P. Gutiérrez, **R. Saberianfar**, S. Kohalmi and R. Menassa. Protein body formation in transgenic tobacco expressin elastin-like polypeptide and hydrophobin fusion proteins. BMC Biotechnology, 13:40, 2013

S. Rentas, **R. Saberianfar**, C. Grewal, R. Kanippayoor, M. Mishra, D. McCollum and J. Karagiannis. The SET domain protein, Set3p, promotes the reliable execution of cytokinesis in *Schizosaccharomyces pombe*. PLoS ONE, 7(2):e31224, 2012

R. Saberianfar, S. Cunningham-Dunlop and J. Karagiannis. Global Gene Expression Analysis of fission yeast mutants impaired in Ser-2 phosphorylation of the RNA Pol II carboxy terminal domain. PLoS ONE, 6(9):e24694, 2011

X. Li, P. Gao, D. Cui, L. Wu, I. Parkin, **R. Saberianfar**, R. Menassa, H. Pan, N. Westcott and M. Y. Gruber. The *Arabidopsis* tt19-4 mutant differentially accumulates proanthocyanidin and anthocyanin through a 3' amino acid substitution in glutathione s-transferase. Plant, Cell and Environment, 34: 374-388, 2011

SELECTED NON-REFEREED CONTRIBUTIONS

R. Saberianfar, J. J. Joensuu and R. Menassa. Role of elastin-like polypeptide and hydrophobin for protein body formation in *Nicotiana benthamiana* leaves. The 9th Canadian Plant Biotechnology Conference, McGill University, Montreal, QC, Canada, May 2014

R. Saberianfar, J. J. Joensuu and R. Menassa. Investigation of the role of elastin-like polypeptides, hydrophobin fusion tags and recombinant protein accumulation levels as critical elements of protein body biogenesis in *Nicotiana benthamiana* leaves. Plant-Based Vaccines, Antibodies and Biologics, University of Verona, Italy, June 2013

R. Saberianfar, J. J. Joensuu, S. Kohalmi and R. Menassa. Comparative study of the role of elastin-like polypeptides and hydrophobin fusion tags and recombinant protein accumulation levels on protein body biogenesis in *Nicotiana benthamiana* leaves.The 3rd Annual Biology Graduate Research Forum, University of Western Ontario, London, ON, Canada, October 2012

R. Saberianfar, S. Gutierrez, J. J. Joensuu, S. Kohalmi and R. Menassa. Characterizing the critical elements of protein body biogenesis by transient expression of elastin-like polypeptide and hydrophobin fusion tags in *Nicotiana benthamiana* leaves. The 8th Canadian Plant Biotechnology Conference, University of Guelph, Guelph, ON, Canada, May 2012

S. Gutierrez, **R. Saberianfar**, A. J. Conley, J. J. Joensuu and R. Menassa. Protein body formation in stable transgenic plants of *Nicotiana tabacum* expressing elastin-like polypeptide and hydrophobin fusions. The 8th Canadian Plant Biotechnology Conference, University of Guelph, Guelph, ON, Canada, May 2012

R. Saberianfar, J. J. Joensuu, A. J. Conley and R. Menassa. Elastin-like polypeptide and hydrophobin induced protein bodies share a similar sequestration mechanism from the endoplasmic reticulum, useful for high-level recombinant protein production. Biology Graduate Research Forum, University of Western Ontario, London, ON, Canada, October 2011

R. Saberianfar, S. Gutierrez, S. Kohalmi, J. J. Joensuu, A. J. Conley and R. Menassa. Protein body induction by elastin-like polypeptide and hydrophobin, new approach to high-level recombinant protein production. The 8th Canadian Plant Genomics Workshop, Niagara Falls, ON, Canada August 2011

S. Gutierrez, **R. Saberianfar**, A. J. Conley, J. J. Joensuu and R. Menassa. Protein body formation in stable transgenic plants of *Nicotiana tabacum* expressing elastin-like polypeptide and hydrophobin fusions. The 8th Canadian Plant Genomics Workshop, Niagara Falls, ON, Canada, August 2011

J. J. Joensuu, A. Ritala, M. Lienemann, A. J. Conley, **R. Saberianfar**, E. Pereira, R. Yan, S. Gutierrez, M. Linder and R. Menassa. Optimizing hydrophobin fusion technology for

protein expression and purification in plants and plant cell culture. Plant Based Vaccines and Antibodies, Porto, Portugal, June 2011

R. Menassa, A.J. Conley, J. J. Joensuu and **R. Saberianfar**. Protein fusions for high level expression and purification of pharmaceutical and industrial proteins in plants. The 14th International Biotechnology Symposium and Exhibition, Rimini, Italy, September 2010

R. Saberianfar, Stephen Truax and J. Karagiannis. Global gene expression analysis of fission yeast mutants impaired in Ser-2 phosphorylation of the RNA Polymerase II Carboxy Terminal Domain. Yeast Genetics and Molecular Biology Meeting, University of British Columbia, Vancouver, BC, Canada, July 2010