



**Aalto University
School of Chemical
Technology**

**School of Chemical Technology
Degree Programme of Chemical Technology**

Kaisa Göransson

**PRODUCTION AND CHARACTERIZATION OF SPHERICAL
BIOLOGICAL DESIGNER COLLOIDS**

**Master's thesis for the degree of Master of Science in Technology
submitted for inspection, Espoo, 18 April, 2016.**

Supervisor Professor Alexander Frey

Instructor Docent Jussi Joensuu

Author Kaisa Göransson

Title of thesis Production and characterization of spherical biological designer colloids

Department Biotechnology and chemical technology

Professorship Bioprocess engineering and food engineering **Code of professorship** KE-70

Thesis supervisor Professor Alexander Frey

Thesis advisor(s) / Thesis examiner(s) Docent Jussi Joensuu

Date 18.4.2016**Number of pages** 79**Language** English

Abstract

Design of protein interactions provides a modern approach for novel material development. Modification of the protein sequence results in changes in the protein's physical properties and its intra- and intermolecular interactions. Genetic engineering offers tools for altering the protein sequence for this purpose, enabling, for example, design of specific process of protein self-assembly. Some cereal seed storage proteins self-assemble into spherical structures called ER-derived protein bodies (PB), being a potential material for this kind of material development. Similar PBs can be induced by maize γ -zein derived ZERA-peptide and fungal hydrophobin HFBI. These proteins have been studied for various composite material applications and represent potential fusion protein partners for large-scale protein production. This work was conducted in order to increase the knowledge of the interactions involved in the PB assembly and to characterize PBs. Multiple gene construct of ZERA, HFBI and ZERA-HFBI were cloned as fusion with green and red fluorescent proteins. The interactions between ZERA and HFBI were studied by co-expressing these constructs in transiently transformed *Nicotiana benthamiana* tobacco plants and observing leaf samples by confocal microscopy. PBs were isolated by subcellular fractionation for characterization. Also, the expression levels of constructs were determined. The results support the hypothesis that ZERA and HFBI have different mechanism for PB formation. Co-expression experiments showed that interaction between ZERA and HFBI is weaker than interaction between ZERA and ZERA or HFBI and HFBI. The determination of expression levels supported hypothesis that PB formation depends on the protein accumulation level in the cell. The isolation procedure was shown to be suitable also for HFBI PBs. The electron microscopy of ZERA and HFBI PBs revealed a porous structure and confirmed the spherical form for PBs. Novel PB structures were observed including radially assembled PBs and Januslike PBs in which the proteins were present in two separate domains within the PB.

Keywords Seed storage protein, maize zein, ZERA, hydrophobin, HFBI, protein body, protein storage organelle, protein self-assembly, composite materials, protein nanoparticles



Tekijä Kaisa Göransson

Työn nimi Geneettisesti tuotettujen proteiinikolloidien vuorovaikutusten karakterisointi

Laitos Biotekniikan ja kemian tekniikan laitos

Professori Biotekniikka ja elintarviketekniikka

Professuurikoodi KE-70

Työn valvoja Professori Alexander Frey

Työn ohjaaja(t)/Työn tarkastaja(t) Dosentti Jussi Joensuu

Päivämäärä 18.4.2016

Sivumäärä 79

Kieli englanti

Tiivistelmä

Proteiinien vuorovaikutusten suunnitelmallinen muokkaaminen geenitekniikan avulla on nykyaikainen lähestymistapa uusien materiaalien kehityksessä. Proteiinisekvenssin muokkaaminen vaikuttaa proteiinin fysikaalisiin ominaisuuksiin sekä molekyylin sisäisiin ja ulkoisiin vuorovaikutuksiin mahdollistaen ennalta suunnitellun itsejärjestäytyvän rakenteen luomisen. Eräiden viljojen jyvissä esiintyvät varastoproteiinit varastoituvat jyvien pyöreiksi proteiinijyväiksi järjestäytyneinä. Esimerkiksi maissin varastoproteiinin zeinin johdannaisen ZERA:n sekä *Trichoderma reesei* –homeen tuottaman hydrofobiin I (HFBI) –proteiinin on todettu itsejärjestyvän vastaaviksi proteiinijyväiksi, mikä tarjoaa mielenkiintoisen lähtökohdan nykyaikaiselle materiaalikehitystyölle. Näitä proteiineja on tutkittu useita materiaalisovelluksia varten ja niitä voidaan hyödyntää fuusioproteiinien osana teollisessa proteiinituotannossa. Työssä selvittiin vuorovaikutuksia, jotka johtavat näiden proteiinien järjestäytymiseen pyöreiksi proteiinijyväiksi sekä tuotettiin ja puhdistettiin proteiinijyväsiä. Työssä kloonattiin useita geenikonstrukteja, jotka sisälsivät ZERA:n, HFBI:n tai ZERA-HFBI:n fuusiona fluoresoivan proteiinin kanssa. ZERA:n ja HFBI:n välisiä vuorovaikutuksia tutkittiin tuottamalla yhtä tai useampia geenikonstrukteja samanaikaisesti *Nicotiana benthamiana* tupakkakasveissa sekä mikroskopoimalla kasvisoluja. Proteiinijyväsien ominaisuuksien tutkimiseksi niitä tuotettiin tupakkakasveissa ja eristettiin lehtisolukosta fraktioimalla. Lisäksi geenikonstruktioiden proteiinituottotasot määritettiin. Tulosten perusteella ZERA ja HFBI järjestäytyvät proteiinijyväiksi eri mekanismein. Tuottokokeet osoittivat, että ZERA:n ja HFBI:n väliset vuorovaikutukset ovat heikompia kuin ZERA-ZERA vuorovaikutukset ja HFBI-HFBI vuorovaikutukset. Tuottokokeet tukevat havaintoja siitä, että ZERA- ja HFBI-proteiinijyväsien muodostuminen riippuu kyseisen proteiinin konsentraatiosta solussa. Lisäksi todettiin, että ZERA-proteiinijyväsien eristykseen käytetty menetelmä soveltuu myös HFBI-proteiinijyväsien eristämiseen. Elektronimikroskopiolla todettiin, että ZERA ja HFBI proteiinijyväset ovat rakenteeltaan pallomaisia ja huokoisia. Lisäksi tuottokokeissa havaittiin uudenlaisia proteiinijyväsiä: säteittäin järjestäytyneitä proteiinijyväsiä sekä Janus-partikkelien kaltaisia polaarisia proteiinijyväsiä.

Avainsanat Jyvän varastoproteiinit, maissi, zein, ZERA, hydrofobiini, HFBI, proteiinijyväsien, aleuronijyväsien, itsejärjestäytyvät proteiinit, komposiittimateriaalit, proteiininanopartikkelit

Preface

This master thesis was conducted at VTT Technical Research Centre of Finland Ltd within Academy of Finland's Centre of Excellence in Molecular Engineering of Biosynthetic Hybrid Materials research programme (HYBER). I am grateful to VTT Ltd for offering this opportunity and want to express my sincere gratitude to my instructor Docent Jussi Joensuu, especially for his patient guidance in the laboratory and with academic writing. I want to thank my supervisor Professor Alexander Frey at Aalto University School of Chemical Technology for keeping record of the progress and offering valuable feedback.

I am grateful to be given the opportunity to meet and work with motivated professionals throughout the organization at VTT Ltd and in HYBER group. I thank Dr. Nonappa in Molecular materials group in Department of Applied Physics in Aalto University School of Science for his high quality electron microscopy pictures presented in this work. The secrets of confocal microscopy would have remained cryptic without practical guidance by Ritva Heinonen and Lauri Reuter at VTT Ltd. I want to also express my gratitude for Professor Merja Penttilä and Protein Production team leader Markku Saloheimo and the team members for the warm and supporting working atmosphere. Thank you also for international members for making this experience unique with tasty and joyful events!

This work would not have been possible without the never-ending support of my loving family and friends. I owe this work to my dad, a chemist himself too. With his help and guidance during my early studies in mathematics, physics and chemistry, and his genuine interest in science, this career path became possible. I express my deep gratitude to my mom for her support throughout my student years and for her practical tips that only a truly experienced quality engineer can give.

Espoo 6.3.2016

Kaisa Göransson

Table of contents

1 Introduction	1
2 Literature review	4
2.1 Storage proteins in common cereal seeds.....	4
2.1.1 Cereal seed and seed storage proteins.....	4
2.1.2 Characteristics of most abundant cereal storage proteins.....	6
2.1.3 Seed protein storage organelles and transport to storage organelles	8
2.1.4 ER-derived protein body formation in maize.....	11
2.2 Inducible protein bodies	13
2.2.1 γ -zein and ZERA peptide	13
2.2.2 Hydrophobin I	16
2.2.3 Other protein body inducing proteins	19
2.2.4 Production of inducible PBs in plants	21
2.3 Material applications of PB inducing peptides	23
2.3.1 Zein composite materials.....	24
2.3.2 ZERA and seed PBs	25
2.3.3 HFBI applications.....	26
3 Materials and methods	29
3.1 Molecular cloning and bacterial transformations	29
3.1.1 Material for cloning.....	29
3.1.2 Golden Gate cloning.....	32

3.1.3 Transformation of <i>E. coli</i>	33
3.1.4 Sequencing and storing to strain collection.....	34
3.1.5 Transformation of <i>Agrobacterium tumefaciens</i>	34
3.2 Production system: plant transformation and growth conditions	35
3.2.1 Plant production	35
3.2.2 Plant agroinfiltration	35
3.3 Protein body isolation by subcellular fractionation.....	36
3.4 SDS-PAGE.....	38
3.5 Western blotting	39
3.6 Expression levels	40
3.6.1 Experimental design.....	40
3.6.2 Sample preparation and protein concentration measurement	41
3.7 Imaging.....	42
3.7.1 Confocal microscopy	42
3.7.2 Electron microscopy.....	43
4 Results	44
4.1 Plasmid constructs	44
4.2 Verification of expression	45
4.2.1 ZERA-mRFP, ZERA-eGFP and eGFP-HFBI	46
4.2.2 mRFP-HFBI, ZERA-eGFP-HFBI and eGFP	46
4.2.3 Expression levels	48
4.3 Co-expression experiments.....	50
4.3.1 ZERA and HFBI	50

4.3.2 ZERA-mRFP, ZERA-HFBI and eGFP-HFBI.....	50
4.3.3 ZERA-mRFP and ZERA-eGFP-HFBI	51
4.3.4 mRFP-HFBI and ZERA-eGFP-HFBI	51
4.4 Protein body isolation by subcellular fractionation.....	54
4.4.1 Homogenization of leaf tissue	54
4.4.2 Subcellular fractionation.....	56
4.4.3 Electron microscopy.....	60
5 Discussion.....	62
5.1 Mechanism of PB assembly.....	62
5.1.1 Expression levels and PB formation of constructs.....	62
5.1.2 Interactions between ZERA and HFBI	63
5.2 Isolation and characteristics of PBs	65
5.3 Further research and improvements	66
5.3.1 Further research suggestions.....	66
5.3.2 Improvements for production and isolation of PBs.....	67
6 Conclusions	69
Literature cited.....	70

Abbreviations

ATPS	aqueous two-phase separation
CBD	cellulose binding domain
COPI	coat protein complex I
COPII	coat protein complex II
DDW	double distilled water
dpi	days post -infiltration
DV	dense vesicle
eGFP	enhanced green fluorescent protein
ELP	elastin-like polypeptide
ER	endoplasmic reticulum
FW	fresh weight
HAp	hydroxylapatite
HFBI	hydrophobin I
HMW	high molecular weight
HPV	human papilloma virus
kDa	kilo-Dalton
LB	Luria Bertani cell culture media
LB repeat	left border repeat
mRFP	monomeric red fluorescent protein
MW	molecular weight
MVB	multivesicular body
o/n	overnight culture
PAC	precursor-accumulating compartment
PB	protein body
PLGA	poly(lactide-co-glycolide)
PSV	protein storage vacuole
RB repeat	right border repeat
rER	rough endoplasmic reticulum
rpm	rounds per minute
S-poor	sulphur-poor
S-rich	sulphur-rich
TSP	total soluble proteins
ZERA	N-terminal proline-rich domain of maize γ -zein

1 Introduction

The modern material research and development focuses greatly on bottom-up strategies in which nanoparticles and macromolecules and their self-assembly mechanism is utilized for building larger functional assemblies. For example, the self-assembly of proteins has been widely studied for material development purposes, in order to create either novel materials or to improve existing ones, such as ceramics and polymers. (Niemeyer 2001; Hu *et al.* 2012). Design of protein interactions provides a modern approach for novel material development enabling, for example, a specific protein self-assembly or enhanced affinity to other molecules. Genetic engineering offers tools for altering the protein sequence for this purpose. Modification of the protein sequence results in changes in the protein's physical properties and its intra- and intermolecular interactions. Hence, genetic engineering can be utilized as a designing tool to introduce specific interactions in proteins. A classic example of this kind of biopolymer development is spider silk protein. Due to its extraordinary physical properties, such as high tensile strength and toughness, it has been intensely studied to produce artificial silk with these physical properties (Wang *et al.* 2014; Hu *et al.* 2012).

In a similar way, seed storage proteins are potential starting materials for this kind of modern material development due to their self-assembly mechanism. Some of the cereal seed storage proteins accumulate in large amounts within endoplasmic reticulum (ER) and self-assemble into dense spherical protein assemblies called ER-derived protein bodies (PB) (for example, Arcalis *et al.* 2014). The self-assembly into spherical PBs is not limited only to seeds: similar PBs can be induced also by other protein sequences and in other host organisms (Schmidt 2013). These proteins or recombinant proteins can be modified and potentially produced in a large scale for novel material application purposes by the means of molecular biotechnology.

These insights have been the original inspiration for this research. In this research, the focus is on two protein sequences: maize seed storage protein γ -zein derived peptide called ZERA and a small fungal protein hydrophobin I (HFBI). Both of the proteins accumulate intracellularly and induce self-assembly of spherical ER-derived PBs (Llop *et al.* 2006; Joensuu *et al.* 2010). Zeins and HFBI have already been studied for various applications, including composite materials, and both proteins have some properties that are useful in large-scale protein production. For example, γ -zein and ZERA PBs are rather dense and they can be expressed in a variety of eukaryotic host cells (Schmidt 2013) whereas HFBI have remarkable amphipathic properties (Linder *et al.* 2005). These properties make ZERA and HFBI attractive molecules for novel material development and recombinant protein production. The aim of this work is to increase the knowledge of the interactions involved in the PB assembly and to produce and isolate PBs for protein composite material research purposes.

The literature review focuses on three main topics. Chapter 2.1 *Storage proteins in common cereal seeds* presents characteristics of the most common cereal seed storage proteins and their accumulation in seed storage organelles. Chapter 2.2 *Inducible protein bodies* emphasizes on ZERA and HFBI protein sequences and their production in plant based systems, and other PB inducing peptides sequences are shortly presented. In chapter 2.3 *Material applications of PB inducing peptides* few recent zein, ZERA and HFBI material applications are reviewed the focus being on protein composite materials and PB applications.

Recently, Dr. Reza Saberianfar (Agriculture and Agrifood Canada, unpublished results) has studied the self-assembly mechanism of ZERA and HFBI protein bodies by co-expressing these proteins in the same plant cell as fusion with fluorescent protein sequences and by visualizing the PBs by confocal microscopy. The researcher observed that ZERA and HFBI formed separate PBs and the proteins did not co-localize within the same PB whereas HFBI co-expressed with another PB inducing protein sequence, elastin like polypeptide

(ELP), co-localized within the same protein body. The result indicates that ZERA and HFBI protein sequences have different self-assembly and ER-retaining mechanism. These observations were the starting point for experiments in this research. This research tries to repeat and confirm the observation of Dr. Saberianfar and to answer following questions regarding the self-assembly of ZERA and HFBI protein sequences:

- Do ZERA and HFBI proteins have different mechanisms to self-assemble into PBs? In other words, do ZERA and HFBI self-assemble into same PBs when co-expressed within the same cell?
- Which one has stronger interactions in the self-assembly, ZERA or HFBI? For example:
 - o Does ZERA-HFBI fusion protein self-assemble within the same PB with ZERA protein when co-expressed in the same cell?
 - o Does ZERA-HFBI fusion self-assemble into same PB with HFBI protein when co-expressed in the same cell?
- What kind of PBs ZERA-HFBI fusion protein forms?

In addition, ZERA and HFBI PBs are produced in plants for isolation experiment, the goal being on testing the isolation procedure for both proteins, and to produce material for PB characterization by electron microscopy.

2 Literature review

2.1 Storage proteins in common cereal seeds

2.1.1 Cereal seed and seed storage proteins

Seeds are highly specialized in accumulating various compounds, such as starch, proteins and lipids, within the cells for the future use of germinating plant. Cereal seed endosperm tissues are highly differentiated and they are categorized to aleurone cells, sub-aleurone cells, central starchy endosperm (or starchy endosperm) tissue, embryo, embryo-surrounding region and embryo transfer cells (Figure 1). Aleurone cells form the outer surface of the seed endosperm and central starchy endosperm tissue comprises the core of the seed, accounting of 80-90 % of the seed weight (Zheng and Wang 2014; Reyes *et al.* 2011). The main storage components in central starchy endosperm are starch and proteins, while the major storage components in aleurone cells are lipids and minerals (Xiong *et al.* 2013). Sub-aleurone cells form few cells thick layer between aleurone cells and central starchy endosperm tissue, and it is richer in proteins than the central starchy endosperm. (Zheng and Wang 2014; Olsen 2004).

Seed proteins can be divided to three groups (Shewry and Halford 2002): storage proteins that are stored in order to serve as building blocks for the germinating seed; structural and metabolic proteins, including for example enzymes that catalyze the degradation of storage compounds during the germination; and seed protective proteins. Storage proteins locate mainly in central starchy endosperm tissue, however, storage proteins are found also in aleurone and sub-aleurone tissues (Zheng and Wang 2014; Arcalis *et al.* 2014).

Seed storage proteins are traditionally divided in four classes based on their solubility and extractability: water soluble albumins, dilute saline soluble globulins, alcohol soluble prolamins and dilute acid or base soluble glutelins (Osborne 1916). Although there are

some prolamins that form alcohol insoluble polymers, all prolamins are alcohol soluble in their reduced form (Shewry and Halford, 2002). Prolamin polymers are stabilized by internal disulfide bonds which is the reason for the alcohol insolubility (Kawakatsu and Takaiwa, 2010). Glutelins are actually 11-12S globulins, and they are nowadays categorized as a subgroup of globulins (Shewry and Halford, 2002).

The storage protein distribution in seed endosperm varies between the cell tissues and different cereal species. The central starchy endosperm of oat and rice store mainly 11-12S globulins, glutelins. Other cereals, such as maize, wheat and barley, store mainly prolamins instead of globulins in the central starchy endosperm tissue. Maize prolamins are called zeins, with subgroups of α -, β -, γ -, and δ -zeins. (Zheng and Wang 2014). α -zein accounts for 80-90 % of total prolamins of maize (Shewry 1995). Wheat prolamins are called gliadins. Barley storage prolamins are called hordeins, which is the major endosperm storage protein in barley (Zheng and Wang, 2014). Albumins are the main storage proteins in dicot seed. In cereals, which belong to monocot family, albumins are not common seed storage proteins. (Shewry *et al.* 1995).

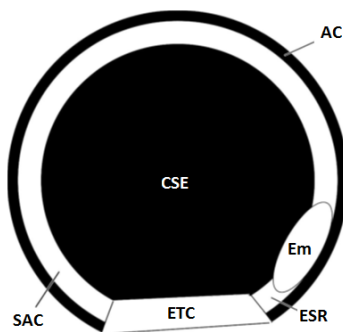


Figure 1. Location of different cereal seed endosperm tissues: aleurone cells (AC), sub-aleurone cells (SAC), central starchy endosperm (CSE), embryo (Em), embryo surrounding region (ESR), embryo transfer cells (ETC). (Zheng and Wang, 2014).

2.1.2 Characteristics of most abundant cereal storage proteins

Globulins are categorized into two subgroups: 11S globulin (belongs to glutelins) and 7S globulins. All globulins are deficient in cysteine and methionine, and proteins in both groups have remarkable variation in their sequences due to post-translational modification (Shewry *et al.* 1995). 11S globulins are present mainly in dicots, however, in cereals they are the main storage proteins of oat and rice. They are hexamers composed of six dimers organized in a ring structure as a disc (Shewry *et al.* 1995; Patel *et al.* 1994). Each dimer is a pair of acidic (40 kilo-Dalton, kDa) and basic subunit (20 kDa) that are linked together by disulfide bonds. The pairs within the ring structure are attached non-covalently. 7S globulins differ greatly from 11S globulins, although it is suggested that they are developed from the same origin, since they share conserved sequences. 7S globulins are trimers (150 – 190 kDa) organized into a disc shape. The most characteristic feature of 7S globulins is the lack of cysteine residues, hence lacking intra- or intermolecular disulfide bonds. (Shewry *et al.* 1995).

Prolamins are rich in proline and glutamine (Kawakatsu and Takaiwa, 2010), however, variation occurs depending on cereal species. Differences between prolamins of most common cereals are summarized in Table 1. Prolamins of wheat, barley and rye, called *Triticeae* family prolamins, are divided in three subgroups: sulfur-rich (S-rich), sulfur-poor (S-poor) and high molecular weight (HMW) prolamins. The S-rich group of prolamins consist 80-90% of total prolamins amount of these species. The protein sequence has a characteristic repetitive N-terminal sequence of proline and glutamine rich motifs and non-repetitive C-terminal sequence, which is instead proline and glutamine poor. The N-terminal sequence folds to polyproline-II helix and C-terminal sequence has a globular folding. In addition, the C-terminus includes abundant cysteine residues, which form intermolecular and intramolecular disulfide bonds resulting polymerization of single prolamins. Similarly to the S-rich group prolamins, S-poor group prolamins have repeats of proline and glycine rich motifs, however, the repetitive sequence comprises almost the whole sequence excluding few residues in N- and C-terminal. In addition, the whole

Table 1. Characteristics of prolamins in most common cereal species according to Shewry *et al.* (1995).

Cereal species	Category	Sequence characteristics			Other	M _r
		N-terminal	Middle	C-terminal		
Triticeae (rye, wheat, barley)	S-rich	repetitive, Pro and Gln rich motif	-	non-repetitive, Cys rich, Pro and Gln poor	Cys residues, polymerization	30 000 - 90 000
	S-poor	repetitive, Pro and Gln rich motif, lack of Cys residues			Cys deficient, no polymerization	
	HMW	non-repetitive	highly repetitive, diverse motifs	non-repetitive	N- and C-terminal Cys residues, polymerization	
Rice		conserved sequences with Triticeae			-	
Oat	avenins	conserved sequences with Triticeae	repetitive, Pro and Gln rich motif	conserved sequences with Triticeae	Middle repetitive sequence	
Maize	zeins (β, γ, δ)	conserved sequences with Triticeae			Cys and/or Met rich, polymerization	
	α-zein	unique	repetitive, no clear motif	unique	No homology with Triticeae prolamins	19 000 - 20 000

sequence of S-poor group prolamins lacks cysteine residues, thus the polymer formation does not occur. The protein sequence of HMW group prolamins differ from other groups: an extensively repetitive protein sequence is surrounded by non-repetitive N- and C-terminal sequences. The repetitive motif varies, however, the usual pattern includes amino acids with neutral-nonpolar side group and neutral-polar side groups, and the sequence forms a spiral structure. Cysteine residues are located in non-repetitive domains enabling intermolecular disulfide bonding and formation of an elastic net. (Shewry *et al.* 1995).

Prolamins of oat, rice and maize share conserved sequences with *Triticeae* family prolamins. In addition, avenins have proline and glutamine rich repetitive sequences, which lack from rice prolamins and maize zeins (β-, γ-, and δ-zeins). However, these zeins are rich in cysteine and/or methionine being able to form intra- and intermolecular

disulfide bonds. Maize α -zein differs greatly from typical prolamins: it has unique N- and C-terminal sequences and a repetitive sequence with no clear motif sequence. (Shewry *et al.* 1995).

2.1.3 Seed protein storage organelles and transport to storage organelles

During the development and the germination, the seeds go through a fast developmental reorganization resulting in dramatic changes in the structure and the composition of the seed. For example, as the endosperm tissue of cereal seed serves as nutrient storage for the developing plant, the central starchy endosperm cells go through apoptosis during the seed maturation (Arcalis *et al.* 2014; Brown and Lemmon, 2007). The trafficking of storage proteins in protein storage compartments alters depending on the developmental stage of the seed, and it is a dynamic process, rather than stable routes for accumulation of storage proteins (Arcalis *et al.* 2014). This chapter focuses on storage protein trafficking to protein storing compartments in maturing seed before the germination.

The storage proteins are deposited in separate intracellular membrane bound storage organelles in cytoplasm of endosperm cells which prevent the degradation of the proteins. These protein storage compartments are divided into two groups that are different in their morphology: protein storage vacuoles (PSV) and ER-derived protein bodies (PB). (Arcalis *et al.* 2014; Zheng and Wang, 2014). The trafficking from ER to storage organelles is an active transport system mediated by vesicles. The vesicle formation, budding, transport to target organelle, membrane recognition, fusion and release of content are highly regulated set of events. (Xiang *et al.* 2013).

The protein sorting to storage organelles initiates in the ER. Transmembrane proteins and soluble proteins contain short hydrophobic N-terminal amino acid sequence, called ER signal sequence, which translocates the protein into the lumen of ER. Upon translocation to ER, the signal sequence is cleaved by ER signal peptidases. Within the ER the protein is subjected to quality control and the misfolded proteins are retained in ER lumen for re-folding and correction. (Xiang *et al.* 2013; Ibl and Stoger, 2012).

There are few known protein transport routes from ER to protein storage organelles: routes from ER to PSV either via Golgi apparatus or bypassing it, and a direct route from ER to PBs (Figure 2). In Golgi-including route, there are two vesicle systems that mediate trafficking between ER and Golgi apparatus: the coat protein complex I (COPI) and coat protein complex II (COPII) systems (Xiang *et al.* 2013). The trafficking from ER to Golgi is mediated by COPII vesicles and reverse trafficking from Golgi to ER is mediated by COPI vesicles. Between Golgi and PSV, the trafficking includes another vesicular sorting: proteins aggregate within Golgi and are packed in small dense vesicles (DV) that can directly fuse to PSV, or alternatively to multivesicular bodies (MVB) that are fused to PSV. In the Golgi-bypassing route, protein precursors are trafficked from ER to PSV within vesicles called precursor-accumulating compartments (PAC). They are fused to PSV either by membrane fusion or by autophagy resembling mechanism. Another main route for storage protein accumulation is aggregation of insoluble proteins within ER and formation of ER-derived PBs. (Xiang *et al.* 2013; Ibl and Stoger, 2012). Interestingly, also these

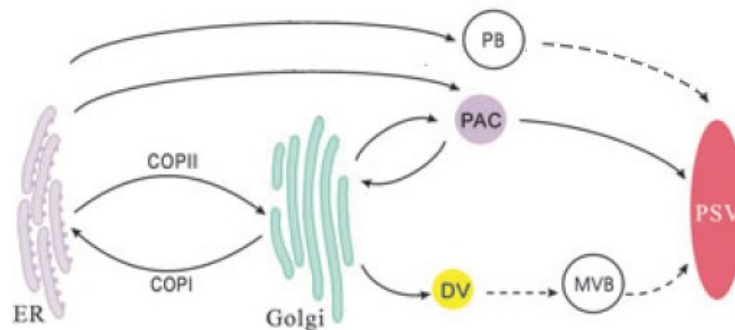


Figure 2. Summary of plant storage protein trafficking to storage organelles. Markings: endoplasmic reticulum (ER), coat protein complex I (COPI), coat protein complex II (COPII), protein body (PB), protein storage vacuole (PSV), precursor accumulating compartment (PAC), dense vesicle (DV), multivesicular body (MVB). Modified from Xiang *et al.* (2013).

storage organelles might finally fuse with PSVs by an autophagy type mechanism (Ibl *et al.* 2014).

The transportation mechanism for different storage proteins varies between plant species and seed tissues. Table 2 summarizes the presence of seed storage proteins in seed tissues of maize, rice, wheat and barley. Soluble storage proteins, such as albumin and globulins, and insoluble prolamins have typically different trafficking routes to protein storage compartments: albumins and globulins are transported from ER via Golgi to PSVs, while insoluble prolamins form aggregates within ER and form ER-derived PBs (Zheng and Wang, 2014; Arcalis *et al.* 2014; Xiang *et al.* 2013; Ibl and Stoger, 2012).

The trafficking route ER-Golgi-PSV applies to aleurone cells of mature cereal seed. The major storage protein in cereal aleurone cells is 7S globulin and it accumulates inside PSVs (aleurone granules) (Zheng and Wang, 2014; Arcalis *et al.* 2014). In aleurone cells of different cereal species, there is no variation in storage protein deposition: maize, wheat and barley globulins are all deposited into aleurone PSVs. Prolamins are absent or occur only as a minor storage protein in aleurone cells (Table 2). Interestingly, maize zein forms an exception to trafficking route of prolamins: a small minority of maize zein is present in aleurone PSV instead of trafficked into PBs. (Arcalis *et al.* 2014; Zheng and Wang, 2014). According to Reyes *et al.* (2011), zeins in maize aleurone cells are deposited in prevacuolar compartments, which are fused by autophagy to PSV.

A common hypothesis has been that the ER-PB storage protein trafficking route applies predominantly to prolamins storage proteins in central starchy endosperm (Zheng and Wang, 2014; Xiong *et al.*, 2013). However, the storage protein trafficking routes of prolamins in the starchy endosperm tissue vary greatly between the cereal species (Table 2). In the central starchy endosperm tissue of wheat and barley, the main storage organelle of prolamins (gliadins and hordeins, respectively) is actually large central PSV.

Also gliadins are trafficked via Golgi into PSV and only a low amount of wheat gliadins form PBs. Interestingly, these PBs are deposited within the central vacuole. In barley,

Table 2. The seed storage protein distribution and location in common cereals based on microscopy findings of Arcalis *et al.* (2014) and review of Zheng and Wang (2014). Globulins are present mainly in aleurone PSVs, while variation in storage protein location occurs in starchy endosperm tissue. Maize zein forms the greatest amount of ER-derived PBs. Markings in the table: protein storage vacuole (PSV), ER-derived protein body (PB), main storage protein in the seed tissue/location (***) , low amount of storage protein in the seed tissue/location (*), seed storage protein not present or the presence not known (-), PBs are deposited within the central PSV (+).

Cereal species	Storage protein	Aleurone tissue		Starchy endosperm tissue	
		PSV	PB	PSV	PB
Maize	zein	*	-	*	***
	globulin	***	-	*	-
Rice	prolamin	-	-	-	*
	globulin (glutelin)	***	-	***	-
Wheat	gliadin	-	-	***	* +
	globulin	***	-	*	-
Barley	hordein	-	-	***	-
	globulin	***	-	-	* +

prolamin PBs are totally absent. There is also another interesting exception in barley seeds: low amount of barley globulins are present in starchy endosperm tissue and are stored into multiphased PBs within the central vacuole. (Arcalis *et al.* 2014). According to Ibl *et al.* (2014) the trafficking of hordeins to PSVs partially bypasses Golgi and hordeins are deposited to PSVs by autophagy-like process. Based on the microscopy results of Arcalis *et al.* (2014), prolamins are dominant in ER-derived PBs only in maize and rice. Prolamin content of rice is low (Arcalis *et al.* 2014), thus maize zeins have the greatest potential to form ER-derived PBs among prolamins of common cereal species.

2.1.4 ER-derived protein body formation in maize

From storage proteins of common cereals, maize zein has an interesting feature to form intracellular ER-derived PBs that are present in the cell in high amounts and evenly

distributed (Arcalis *et al.* 2014). Protein re-trafficking from Golgi to ER within COPI vesicles is determined by short C-terminal H/KDEL peptide motif, called the ER retention signal (Xiang *et al.* 2013). However, prolamins lack the H/KDEL ER retention signal and therefore the aggregation of zeins in ER is probably the result of intrinsic protein properties. According to Shewry *et al.* (1995) the protein stabilizing disulfide bond formation occurs within the lumen of ER. Maize β -, γ - and δ -zeins are rich in cysteine and have the ability to form disulfide bonds between these residues, either intra- or intermolecularly (see Table 1). Therefore, it is probable that the ability of β -, γ - and δ -zeins to assemble into disulfide-stabilized insoluble polymers is the reason for ER retention and PB assembly.

Lending and Larkins (1989) have originally proposed a model for the distribution of maize zeins in the PB during the PB formation which is based on light and electron microscopy findings (Figure 3). More recently, Guo *et al.* (2013) have studied the function of zein family proteins in protein body formation in maize endosperm tissue by silencing the expression of zein genes by using RNA interference. The initiation of PB formation requires only 27 kDa γ -zeins which form the core of early PB aggregation within the lumen of rough ER (Guo *et al.* 2013). In addition to 27 kDa γ -zeins, the early PB core includes β -zeins and other γ -zeins (Guo *et al.* 2013), which are probably incorporated to the 27 kDa γ -zein core by strong interactions between the 27 kDa, 50 kDa and 16 kDa γ -zeins and the 15 kDa β -zein proteins (the latter belonging actually to γ -zein family) (Kim *et al.* 2002). The initiation is followed by formation of α - and δ -zein inclusions within the γ - and β -zein core (Guo *et al.* 2013). Based on the research of Kim *et al.* (2002), also this is due specific interactions: 16 kDa γ -zein and 15 kDa β -zein that interact with α - and δ -zeins, instead of 27 kDa and 50 kDa γ -zeins that do not have the interaction. The mature PB has an even spherical shape and a diameter of 1-2 μm , and it is surrounded by rough ER (rER). The increase in size is due the accumulation of 19 kDa α -zein within the core of the PB. The core of a mature PB consist of 19 kDa α -zein and δ -zein, and it is surrounded by a layer of 22 kDa α -zein and an outer layer of γ - and β -zeins. Lower amount of certain zeins affects the amount of PBs in the cell, the size of the PBs or the spherical morphology of the PBs.

Thus, zeins are needed in correct stoichiometric ratio to generate correctly formed spherical zein PBs in maize endosperm. (Guo *et al.* 2013). However, notable is that the 27 kDa γ -zein has the most critical role in PB initiation and formation in maize by stabilizing other zeins.

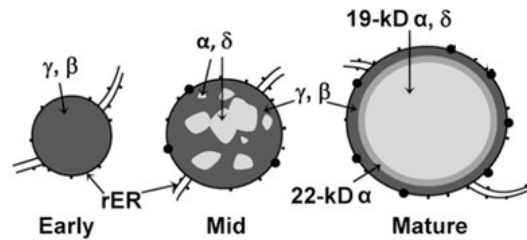


Figure 3. Maize zein distribution in the protein body during the protein body formation. The protein body is surrounded by rough endoplasmic reticulum (rER). Figure modified from Guo *et al.* (2013).

2.2 Inducible protein bodies

ER-derived PBs can be induced also in non-seed organisms and tissues in which PBs are not naturally present (Schmidt, 2013; Torrent *et al.* 2009a). In this chapter, proteins that have been experimentally demonstrated to induce PBs are presented with their reported production systems. Examples of their applications are presented in the chapter 2.3 *Material applications of PB inducing peptides.*

2.2.1 γ -zein and ZERA peptide

Geli *et al.* (1994) discovered by mutant screening that the ability of maize γ -zein to be retained in ER is associated to N-terminal proline-rich repeat domain, and the self-assembly into PBs is due to the N-terminal proline-rich repeat domain and the C-terminal cysteine-rich domain in the protein. To utilize the ER retaining and PB forming abilities of maize γ -zein, researchers of ERA Biotech Sa., Spain (nowadays ZIP Solutions Sl., Spain)

have developed and patented a technology to accumulate recombinant proteins in ER-derived PBs *in vivo* (Anon. 2016). The technology is based on ZERA sequence which is derived from the N-terminal proline-rich region of 27 kDa γ -zein ZERA (Llop *et al.* 2006) and which forms spherical ER-derived PBs that are approx. 1 μ m in diameter (Schmidt, 2013). In the recombinant protein production, the ZERA sequence is used as PB formation and purification tag by fusing the protein of interest to the C-terminus of ZERA sequence (Llop *et al.* 2006).

The ZERA sequence (Figure 4A) has 112 amino acids that include γ -zein signal peptide and the first 93 amino acids of maize γ -zein. The N-terminal sequence contains two cysteine residues, the repeat region has eight repeat sequences of three proline residues and one residue of valine, histidine and lysine (PPPVHL₈), and the Pro-X region contains four cysteine residues. ZERA peptide is thought to have a stick-like tertiary structure, in which the PPPVHL₈ repeat region is folded to an amphipathic polyproline II conformation. (Llop-Tous *et al.* 2010). According to Llop-Tous *et al.* (2010), the minimum features of ZERA sequence that are necessary for inducing the PB formation are two N-terminal cysteine residues (Cys⁷ and Cys⁹) and eight PPPVHL repeats. Even though, 4-6 proline repeats are sufficient to form smaller multimers (Llop-Tous *et al.* 2010).

Llop-Tous *et al.* (2010) suggest that the PB self-assembly is determined by the equilibrium of the protein's aggregation capacity, the protein concentration and the activity of ER-resident chaperons. On the contrary, Hofbauer *et al.* (2014) have shown that the formation and budding of PB do not require ER-specific factors. The N-terminal part of 27 kDa γ -zein (similar to the ZERA sequence) that was targeted to the cytosol, ER, or plastid, resulted in the induction of PBs within each of these locations. In addition, budding of PBs from the outer membrane of plastid envelope was observed to occur in a similar manner as the budding from the ER membrane. (Hofbauer *et al.* 2014). Therefore, it seems that ER-resident chaperons are not critical in the formation of 27 kDa γ -zein based PBs, and the mechanism of ZERA self-assembly to PBs depends mostly on the ZERA - ZERA interactions.

According to Llop-Tous *et al.* (2010), the mechanism for self-assembly is driven by hydrophobic interactions and disulfide bond formation (Figure 4B): the hydrophobic protein-protein interaction between PPPVHL₈ regions of ZERA sequences initiates the assembly by aligning the molecules into small oligomers, and intermolecular disulfide binding between Cys residues stabilizes the structure by enabling increased packing into a larger polymer. (Llop-Tous *et al.* 2010). The discovery by Hofbauer *et al.* (2014) supports the proposed mechanisms: in their experiment a construct that contained the PPPVHL₈ repeat region but lacked N-terminal cysteines was unable to form PBs. Thus it seems that the PB formation is mainly dependent on the stabilizing effect of covalent disulfide bonds

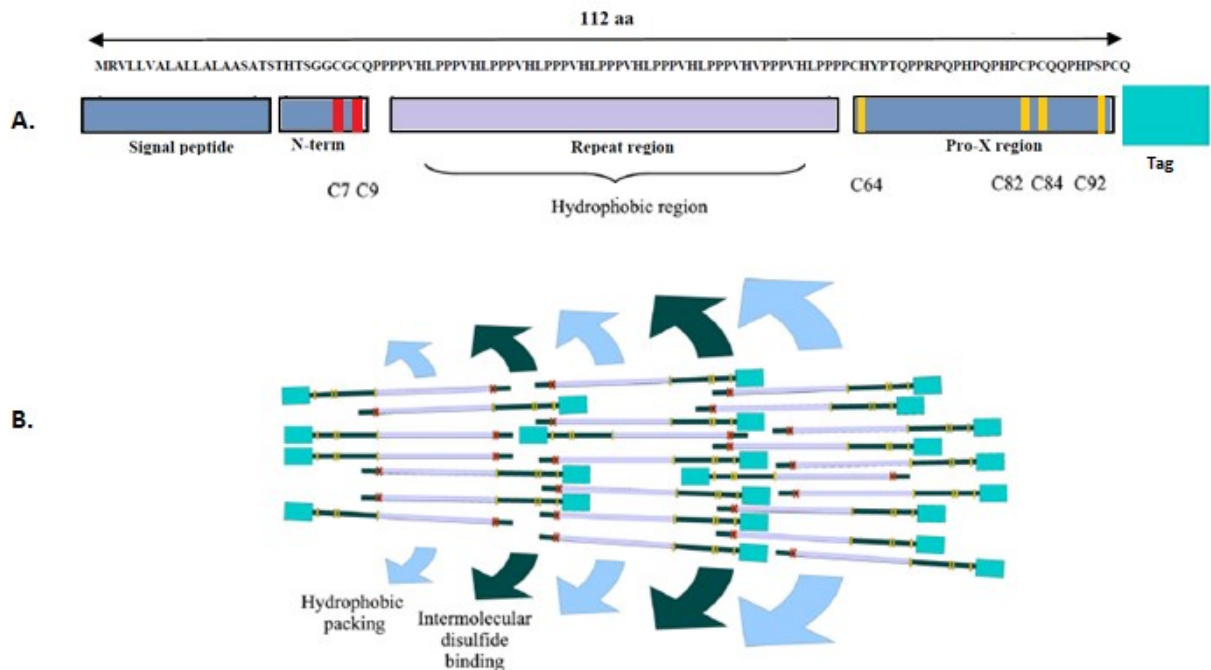


Figure 4. A. ZERA amino acid sequence with 112 aminoacids (aa) and marked cysteine residues (C7, C9, etc.). B. The mechanism of ZERA PB packing initiates by hydrophobic packing of hydrophobic repeat region followed by intermolecular disulfide bonding of N-terminal cysteine residues. Plastid targeting sequence locates in Pro-X region, aa 51 – 93 (Hofbauer *et al.* 2014). Figure modified from Llop-Tous *et al.* (2010).

between cysteine residues (Hofbauer *et al.* 2014). Clearly, the probability of hydrophobic interactions and disulfide bond formation between ZERA proteins increase as the protein concentration increases.

ZERA is widely studied and it has potential to be a versatile fusion tag for recombinant protein production. ZERA PBs have been produced successfully in various host organisms including mammalian cells, insect cells, fungi, plant leaves and plant protoplasts (Schmidt 2013). In addition, the interactions involved in ZERA PB self-assembly are strong enough to induce the PB formation even in the presence of a large fusion protein at the C-terminus. For example, ZERA (12,3 kDa) fused with human growth hormone (22 kDa) or with F1-V antigen (53,2 kDa), an antigen for vaccine against plague caused by *Yersinia pestis*, have shown to induce the PB formation (Alvarez *et al.* 2010; Schmidt 2013). As mentioned previously, PBs can also be induced in various subcellular compartments in addition to ER. Interestingly, Hofbauer *et al.* (2014) have discovered a plastid targeting element within the N-terminal part of 27 kDa γ -zein which translocates the protein into plastid without the aid of any additional targeting signal. The researchers suggest that the plastid targeting peptide motif is in the cysteine rich Pro-X region (amino acid residues 51-93), since the Pro-X region includes a large proportion of hydrophobic and basic amino acids that are usually prevalent in plastid targeting signal sequences. However, the morphology of PBs varies depending on the subcellular location, due the variety in the chemical composition and chaperon content between the subcellular compartments. (Hofbauer *et al.* 2014).

2.2.2 Hydrophobin I

Hydrophobins are small amphipathic fungal proteins that are produced by filamentous fungi, and they have various roles in fungal development. For example, they form coatings on spores and are secreted from fungal hyphae to facilitate the penetration of the air-water interface. (Szilvay *et al.* 2006). The amphipathic structure of hydrophobins enables the self-assembly at hydrophobic/hydrophilic interphase, such as air-water interface, and

allows changing a hydrophobic surface to hydrophilic and *vice versa* (Wösten and Scholtmeijer, 2015; Szilvay *et al.* 2006). Actually, hydrophobins are a group of the most surface active proteins known (Linder *et al.* 2005). Due to these properties, they have been intensely studied to reveal the structure-function relationship and to utilize them in various applications.

Hydrophobins are divided into two classes, class I and class II, primarily based on the pattern of hydrophobic and hydrophilic clusters in the amino acid sequence and the solubility of the protein aggregates (Wessels, 1994). HFBI is naturally produced by a fungus *Trichoderma reesei* and it has a proven ability to induce ER-derived PBs. Joensuu *et al.* (2010) were the first to discover that HFBI formed similar kind of ER-derived PBs as ZERA. In their research, PBs were formed when HFBI was fused with green fluorescent protein (GFP) and expressed in leaves of *Nicotiana benthamiana*. HFBI belongs to class II hydrophobins and it forms soluble aggregates in water (Wösten and Scholtmeijer, 2015). The protein sequence of HFBI (Figure 5A, 5B) includes 75 amino acids with a molecular weight of 7,5 kDa (Khalesi *et al.* 2015). The protein sequence folds into a typical conformation for class II hydrophobins: a compact globular tertiary structure with a flat hydrophobic patch on the otherwise hydrophilic surface. (Szilvay *et al.* 2006; Linder *et al.* 2005). The hydrophobic patch is formed from aliphatic amino acid residues, which are highly conserved in class II hydrophobins (Linder *et al.* 2005; Szilvay *et al.* 2006). The structure of HFBI is cross-linked by four disulfide bonds between Cys residues, which is characteristic for all hydrophobins (Hakanpää *et al.* 2006).

The surface activity of HFBI is mainly due to these structural features. A large hydrophobic patch prevents efficiently water molecules from maintaining hydrogen bonding around the hydrophobic patch (Szilvay *et al.* 2006; Chandler 2005) and cross-linking stabilizes the protein into a rigid structure, thus preventing the hydrophobic patch to be hidden within the molecule in a hydrophilic environment (Szilvay *et al.* 2006). In addition, a patch of charged residues in the hydrophilic part of the protein affect significantly to HFBI surface

activity increasing the hydrophilic properties of the protein. However, the role of these residues to the self-assembly remains somewhat unclear. (Lienemann *et al.* 2013).

A common view is that the mechanism of hydrophobin self-assembly is driven by hydrophobic-hydrophilic interactions between the protein and the solution in the events of film formation at interfaces, oligomerization and PB formation. In the two latter cases, it is suggested that hydrophobic interactions turn the hydrophilic patches of molecules within the oligomer to be protected from the hydrophilic solution. (Joensuu *et al.* 2010; Hakanpää *et al.* 2004). In addition, Lienemann *et al.* (2013) suggest that the dissociation of

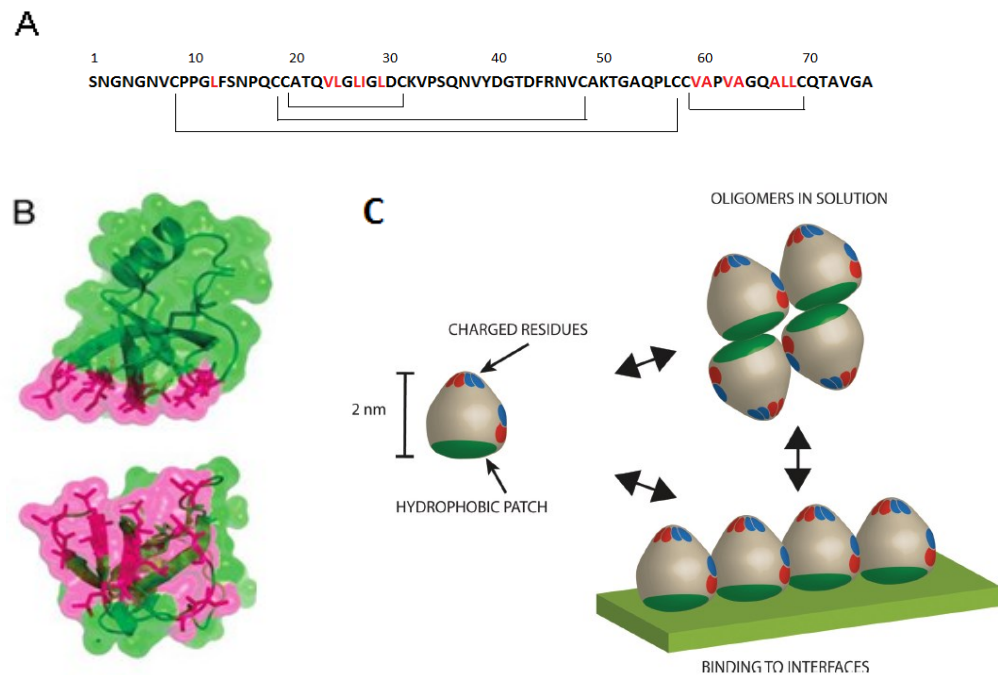


Figure 5. Hydrophobin I. A-B. Conserved aliphatic residues of HFBI amino acid sequence form a large hydrophobic patch (in red color) and disulfide bonds between cysteine residues stabilize the tertiary structure. Modified from Szilvay *et al.* (2006), Hakanpää *et al.* (2006) and Linder *et al.* (2005). C. The oligomerization of monomers to dimers and tetramers and binding to interfaces is affected by the stability of oligomers (Lienemann *et al.* 2013).

oligomers to monomers and the film formation at the interface is affected by the stability of oligomers: the greater the stability of oligomers, the lower the binding to surface and *vice versa* (Figure 5C).

Recently, Saberianfar *et al.* (2015) showed that the PB induction of GFP-HFBI fusion protein is merely due to HFBI rather than GFP. Although, according to Conley *et al.* (2009), unfused GFP is also able to form protein aggregations when expressed at high levels in *N. benthamiana*. Also according to research of Saberianfar *et al.* (2015) and Gutierrez *et al.* (2013), the PB formation depends directly on the protein accumulation level the threshold value for GFP-HFBI being 0,2 % of total soluble proteins (TSP). Kisko *et al.* 2008 have shown that the polymerization of HFBI in solution depends on the protein concentration. In dilute solutions (few $\mu\text{g/ml}$) hydrophobins are present mainly in monomers and as the concentration increases the dimerization occurs. At higher concentration (0,5 – 10 mg/ml) hydrophobins form tetramers and larger assemblies. (Kisko *et al.* 2008). Compared to ZERA PB induction, notable is that the PB induction of HFBI was achieved by targeting the HFBI to ER by H/KDEL signal sequence as HFBI would otherwise be secreted from the cell (Joensuu *et al.* 2010; Schmidt 2013). Instead, ZERA does not require H/KDEL signal sequence to be retained in the ER.

In the case that self-assembly is driven by above mentioned hydrophobic interactions, the effect of protein concentration to the self-assembly is reasonable. Since H/KDEL signal sequence is needed to retain HFBI in the ER and induce the PB formation, and the oligomerization of HFBI is concentration dependent, it seems that a high local accumulation of HFBI in the ER is needed for PB induction.

2.2.3 Other protein body inducing proteins

Protein bodies can be induced by using truncated seed storage proteins as fusion partners with recombinant protein. In addition to maize γ -zein and ZERA, also rice prolamins and glutenins can be used for this purpose (Saito *et al.* 2009; Saumonneau *et al.* 2011). One

intensely studied PB inducing fusion protein is zeolin. Zeolin consists of 89 amino acids of γ -zein N-terminus and the whole amino acid sequence of phaseolin, a major vacuolar storage protein of beans that is deficient in cysteine. Mainieri *et al.* (2004) have showed that zeolin forms ER-derived PBs in *N. benthamiana* leaves in a similar way than γ -zein. Since phaseolin lacks cysteine residues, the disulfide bonding of γ -zein part of the zeolin seems to be essential for PB induction (Pompa and Vitale, 2006). Zeolin was originally developed to balance the nutritional value of cysteine rich γ -zein. (Schmidt 2013). Later, zeolin has been studied for using it as a tag to accumulate recombinant proteins in plants, similarly as ZERA and γ -zein fusions. De Virgilio *et al.* (2008) have showed the potential of zeolin for recombinant protein production: human immunodeficiency virus antigen Nef, produced with zeolin, accumulated in PBs. More recently, de Marchis *et al.* (2011) have studied zeolin production in plant plastids in order to use plastids as bioreactors for recombinant protein production.

PB induction can be achieved also by using H/KDEL ER-retention signal sequence as in the case of HFBI. Elastin-like polypeptide (ELP) is another protein which has been proven to induce spherical PB formation in plants when fused with KDEL ER-retention signal sequence (Conley *et al.* 2009). ELP is a highly repeated polypeptide consisting of VPGXG repeat pattern and it has an interesting reversible feature to form insoluble hydrophobic aggregates when heated to a critical temperature called transition temperature. The phenomenon is called inverse phase transition and it can be utilized for protein purification, thus making ELP very potential fusion protein tag. (Schmidt 2013; Urry *et al.* 1991; Meyer and Chilkoti 1999). ELP has been shown to increase the protein yield even by 100 folds (Conley *et al.* 2009). However, as the H/KDEL retention depends on receptors, the retention mechanism is saturating and can limit the protein accumulation in ER (Semenza *et al.* 1999).

2.2.4 Production of inducible PBs in plants

Transformed whole plants, plant cell suspension culture and tissue cultures can be used for recombinant protein production in plants. The host organism can be either stably (either in nucleus or plastid) or transiently transformed. (Krenek *et al.* 2015; Xu *et al.* 2012). Also plastid transformation has become attractive alternative for protein production, due to the maternal inheritance of plastid genome that can solve various regulatory concerns in field cultivation (Waheed *et al.* 2015). Inducible PBs have been produced in different plant expression platforms, including transient transformation and stable transformation (whole-plant systems), and plant cell culture (for example Hofbauer *et al.* 2014; Alvarez *et al.* 2010; Torrent *et al.* 2009a, respectively). Examples of plant species that have been used for the PB production are tobacco plants *N. benthamiana* and *N. tabacum*, and alfalfa *Medicago sativa* (for example, Alvarez *et al.* 2010).

The most common method for PB production is *Agrobacterium*-mediated transient transformation of *N. benthamiana* and *N. tabacum* plants. *Agrobacterium*-mediated gene transfer is based on the tumor induction ability of *Agrobacterium* which is related to crown gall disease caused by *Agrobacteria* (Yadav *et al.* 1980; Willmitzer *et al.* 1980). The virulence of *Agrobacteria* is due to the Ti (tumor inducing) plasmid of the bacterium: part of the plasmid DNA, called transferred DNA (T-DNA), is copied to a single stranded DNA and delivered to the plant cell and further to the plant chromosome. (Yadav *et al.* 1980; Willmitzer *et al.* 1980; Gelvin 2003). In the Ti plasmid, the T-DNA is surrounded by two short repeats, left border (LB) and right border (RB), which determine the sequence that is copied and transferred (Gelvin 2003). In the transformation, the natural mechanism can be utilized via binary vector system (Krenek *et al.* 2015): the Ti plasmid is divided into a nonvirulent plasmid including T-DNA and border sequences, and to another small plasmid including virulence *vir* genes of *Agrobacterium*. The recombinant DNA construct is cloned to T-DNA region and the plasmid is transformed to *Agrobacterium* cells (Gelvin 2003). The advantage of transient transformation production system is short production time from cloning of a gene construct to expression in plants (Musiychuk *et al.* 2007).

In laboratory, the transfection of plant cells by *Agrobacteria* is easily achieved by agroinfiltration method, in which the bacterial suspension is infiltrated by pressure into the intercellular space of plant leaves either by syringe or by vacuum. (Bechtold and Pelletier 1998; Schöb *et al.* 1997). The method is suitable to transiently transform partial leaves or whole plants. For large-scale production, researchers at Kentucky Bioprocessing LLC (Owensboro, KY, USA) have developed a vacuum infiltration pilot process which enables to infiltrate even 1.2 tons of plant biomass per day (Gleba *et al.* 2013). In addition, Giritich *et al.* (2012) have patented an interesting method called agrospraying for fast and economical method for large-scale transfection, in which tobacco plant leaves are sprayed with *Agrobacterium* suspension in the presence of a surfactant.

The efficiency of *Agrobacterium* transfection can be improved by using viral vectors with *Agrobacterium* binary vector system. The method is called agroinfection and is developed by Grimsley *et al.* (1986). The virus genome is cloned to T-DNA together with the expressed DNA construct to enable initiation of viral infection in the plant, thus spreading the construct systemically to the whole plant (Grimsley *et al.* 1986). More recently, Gleba *et al.* (2005) have described an efficient variation of this method called magniffection, in which *Agrobacterium* vector delivers DNA copies of viral DNA or RNA replicons in plant cells. In addition, in *N. benthamiana*, the delivery of T-DNA depends on the strain: the highest deliveries are achieved by CryX strain developed by Nomad Bioscience GmbH (Germany), which enables 100 – 1000 folds higher T-DNA delivery to *N. benthamiana* (Gleba *et al.* 2013). Reavy *et al.* (2007) have also noticed that one of the T-DNA transfer related protein, VirD2, is degraded by a caspase-like protease in tobacco, thus decreasing the amount of infected cells. By mutagenesis of VirD2 protein, the transfer of T-DNA was improved since caspase-like protease activity was decreased (Reavy *et al.* 2007).

In the production of PBs by whole-plant production system, the greatest bottleneck is down-stream processing of the leaf material. Torrent *et al.* (2009b) have described a protocol to isolate ZERA PBs from leaf extract by ultracentrifugation procedure that

separates PBs by their density (1,20 g/cm³) from other leaf material. The method is applicable also for PB isolation from other host organisms. After the isolation of ZERA PBs, the protein of interest can be further purified from PBs by simple chromatography methods (Llompart *et al.* 2010). However, ultracentrifugation based separation and chromatography purification steps limit the scalability of the down-stream procedure mainly to laboratory scale. It can be considered as suitable method only if relatively small quantities of pure PBs are needed, for example in the case of high-value materials. In order to produce PB in large amounts, a large-scale isolation process of PB from plant leaves needs to be developed, since no method for a large-scale ZERA PB isolation is reported in the literature. Also in the case of HFBI PBs, any successful isolation of HFBI PBs even in a laboratory scale has not been reported in the literature. If the whole leaf is a suitable end product form in the PB production, an interesting alternative would be storing plant leaves as silage, as described by Hahn *et al.* (2012). In their research, cellulases were produced in plant leaves which were stored as silage allowing direct usage as enzymatic additive and resulting no significant loss of cellulase enzyme activity (Hahn *et al.* 2012).

2.3 Material applications of PB inducing peptides

Proteins are biopolymers that naturally act as building blocks in tissues and have special biological activities. Fibrous proteins, such as elastins, collagens, silks, keratins and maize zeins, are commonly used to create protein-based composite materials by mixing them with diverse already existing materials, for example, natural polymers, synthetic polymers, inorganic materials and drugs. They are preferred substances in material development due to their special properties, for example flexibility and strength in collagens and elastins. Using these proteins in a mixture with other materials allows introducing novel properties into the composite materials and development of new applications. In addition, a major advantage of using proteins for composite material development is the possibility

to alter the protein sequence by genetic engineering in order to design specific interactions between composite material molecules. Examples of protein composite formats are hydrogels, sponges or scaffolds, fibers or tubes, microspheres and thin films. (Wang *et al.* 2014). In this chapter, some recent zein and ZERA, as well as hydrophobin HFBI, based material applications are presented.

2.3.1 Zein composite materials

Zein is a very versatile protein for material use, and it has been widely used in material applications as pure zein film, plasticized film, chemically modified film, hybrid composite film with other natural polymer (such as whey protein or other protein), microsphere particles and coatings in nanoparticles (Zhang *et al.*, 2015). There are many potential applications for zein as biomaterial within food industry, health care and pharma.

For food packaging applications, zein is an attractive material since it is a non-toxic protein. Recent researches of zein based food packaging applications have focused on improvement of physical properties of zein films. Shakeri and Radmanesh (2014) have produced zein/nanocellulose composite polymer to create a potential biodegradable material for food packaging purposes. Adding of cellulose nanofibrils to zein/cellulose composite improved thermal resistance compared to zein films (with 3 % and 5 % of cellulose nanofibril addition), probably due to the increase of hydrogen bonding in the material. (Shakeri and Radmanesh, 2014). In the research of Zhang *et al.* (2013), zein was added to wheat gluten to form a composite film. The addition of zein increased, for example, tensile strength and elongation strength and decreased water vapour permeability. Most interestingly, the antibacterial properties improved compared to wheat gluten films. (Zhang *et al.* 2013).

For biomaterial applications, the biocompatibility and blood compatibility of multiple zein/starch composite films with variable amount of zein (0-100 % (w/w)) were evaluated by Liu *et al.* (2015). The composite films presented anticoagulant ability that increased as the amount of zein increased in the films. The films that included over 50 % of zein also

showed compatibility with the cells tested. Cells were able to grow on these films, most probably due to smooth surface for cell attachment and zein's anticoagulant properties. (Liu *et al.* 2015). Recently, Lin *et al.* (2015) studied zein based composite material, zein/hydroxylapatite/poly(lactide-co-glycolide) (zein/HAp/PLGA) scaffold, *in vitro* and *in vivo* for its biocompatibility in tissue engineering. The material had excellent suitability with cartilage cell tissue (a specific flexible tissue that is present, for example, in nose, ears and in the joints close to bones), since there was no significant difference between proliferation of cartilage cells growing on zein/HAp/PLGA scaffold and growing without the scaffold. (Lin *et al.* 2015).

Zein microspheres and zein based nanoparticles have a potential for drug delivery, controlled drug release and various enzymatic reactions. Zein can be used for coating of nanoparticles (Zhang *et al.* 2015) or as part of composite carrier particles, which are microscale particles that include two separate components, the carrier core and another substance attached to the core (Baars *et al.* 2015). Baars *et al.* (2015) presented a novel method to create composite carrier particles from zein. Various particles were formed from positively charged zein and negatively charged nanoparticles, such as magnetite, hematite, gold and silver. Three different particle morphologies were presented: spherical zein carrier of approx. 200-300 nm in diameter with nanoparticles organized either within the zein carrier, on the surface of the carrier, or inside the carrier as an internal shell. Additionally, composite particles were coated with silica to enable surface modification of the particles. The coating enables an additional modification of the particle surface, allowing, for example, attaching of enzymes or specific targeting of the nanoparticles. (Baars *et al.* 2015)

2.3.2 ZERA and seed PBs

ZERA and other seed storage proteins have been applied to vaccine development, especially interest on edible vaccines. Wakasa *et al.* (2015) report a new formulation for oral vaccination by using seed protein bodies. Protein bodies were used for allergen

protein storage compartments in rice seed to produce oral tolerogen against japanese cedar pollen. The allergens were expressed in rice seed endosperm under rice endosperm-specific promoters to produce protein bodies, and concentrated PB powder was formulated from these seeds. The powder was shown to be stable up to 10 months at room temperature and it was able to induce immune tolerance in mice when orally administered. (Wakasa *et al.* 2015). Whitehead *et al.* (2014) have produced human papillomavirus (HPV) type 16E7SH protein as fusion with ZERA peptide in *N. benthamiana* plant leaves for vaccination experiments in mice. High expression level of ZERA-16E7SH fusion protein was reached and the fusion protein vaccine enhanced the immune response in mice causing tumour regression. Interestingly, significant tumour regression was observed also after co-inoculation of individual ZERA PBs and 16E7SH protein. In addition, researchers noticed that ZERA protein had also an adjuvanting effect. (Whitehead *et al.* 2014).

2.3.3 HFBI applications

Due to remarkable surfactant properties, hydrophobins are suitable for various applications, for example, to act as coating agent in surface modification, to disperse hydrophobic solids into hydrophilic solutions, to stabilize foams and emulsions, to immobilize cells and molecules or prevent their binding, and aid purification of recombinant proteins (Wösten and Scholtmeijer 2015; Khalesi *et al.* 2015).

For pharmaceutical use, Valo *et al.* (2010) have used HFBI to render lipophilic drug molecules to hydrophilic drug nanoparticles in order to enhance their solubility in aqueous solution. In their protocol, HFBI was mixed with the drug in an aqueous solution, resulting in coating of the drug with HFBI molecules. According to the study, the surface of the coated drug nanoparticle could be further modified with fusion proteins. (Valo *et al.* 2010). Indeed, in their later research (Valo *et al.* 2011), lipophilic pharmaceutical molecules, such as itraconazole, were immobilized on cellulose nanofibrils by HFBI fusion protein that had two cellulose binding domains (CBD). The immobilization to cellulose

nanofibrils enhanced the stability of the drug for up to 10 months and interestingly, also enhanced its *in vivo* performance (Valo *et al.* 2011). Relatively recently, Valo *et al.* (2013) have studied the bioavailability of the drug from HFBI and double CBD fusion protein coated nanoparticles. Nanoparticles were tested in four different nanofibrillar cellulose aerogels resulting with different release patterns indicating that HFBI coated drug nanoparticles are potential carriers for controlled drug release applications. (Valo *et al.* 2013). However, regulatory issues limit the usage of hydrophobins in food industry and pharma. In food industry, hydrophobins are not yet authorized as Novel Food in the EU, although hydrophobins can be utilized as processing aid in case they are removed from the final product (Khalesi *et al.* 2015).

Takatsuji *et al.* (2013) have used HFBI for enzyme immobilization. In their research, glucose oxidase was immobilized on solid surface by applying surface active properties of HFBI for attaching the enzyme on the surface instead of using conventional methods of chemical cross-linking or non-specific adsorption. An enhanced enzyme activity was achieved by a recombinant fusion protein of HFBI and glucose oxidase that had a flexible linker between the proteins. Researchers suggest that the higher enzyme activity was due to the enhanced mobility of the immobilized enzyme, since conventional immobilizing methods restrict the orientation and mobility of the enzyme. In addition, by changing the mixing ratio of the fusion protein and free HFBI, the amount of adsorbed enzyme on the solid surface could be controlled, which allows maximal enzymatic activity by minimal usage of the enzyme on the surface. (Takatsuji *et al.* 2013).

Joensuu *et al.* 2010 have optimized a simple purification method of HFBI fusion proteins from tobacco leaf extract that is based on HFBI surface activity. By aqueous two-phase separation (ATPS), fusion protein GFP-HFBI was purified with only one step from leaf extract, with a selective recovery of up to 91 % and a concentration of up to 10 mg/ml. The method is easily scalable for large-scale production. (Joensuu *et al.* 2010). Reuter *et al.* 2014 have developed another ATPS method for large-scale purification of GFP-HFBI

from tobacco bright yellow 2 (BY-2) cell culture. The cultivation was performed in 600 L pilot scale bioreactor and the GFP-HFBI was purified from 20 L of cell lysate. The purification resulted in threefold concentration and 60 % recovery providing an efficient method for large-scale purification of HFBI fusion proteins from plant cell suspension. (Reuter *et al.* 2014).

3 Materials and methods

The outline of experimental procedures of this work is presented in Figure 6.

3.1 Molecular cloning and bacterial transformations

3.1.1 Material for cloning

For cloning of pJJJ751-754, gene fragments were obtained from *E. coli* strains of VTT Ltd strain collection (Table 3; enhanced green fluorescent protein (eGFP) and monomeric red fluorescent protein (mRFP)). For ZERA-eGFP-HFBI (pJJJ755), gene fragment 2-linker1-eGFP-linker1-3 was ordered as gBlocks® Gene Fragment from Integrated DNA Technologies Inc. (Belgium). Plasmid pJJJ178 was used as destination vector in cloning (Figure 7).

Table 3. Entry vectors for cloning.

Entry vector	Gene fragment	Code in VTT Ltd strain collection (<i>E. coli</i>)
pJJJ743	1-ZERA-2	B6384
pJJJ745	1-mRFP-2	B6386
pJJJ335	2-linker1-3	B4466
pJJJ746	3-mRFP-4	B6387
pJJJ337	3-eGFP-4	B4468
pJJJ334	3-HFBI-4	B4465
pJJJ178	destination vector	B3918

Before cloning, plasmid maps for new constructs were created by Geneious 7.1.8 software (Biomatters Ltd, New Zealand). *E. coli* strains with the desired entry vectors were inoculated to 5 ml of Luria Bertani (LB) + ampicillin (Amp) 50 µg/ml culture media in 10 ml tube. The *E. coli* strain with the destination vector (pJJJ178) was inoculated in 5 ml of LB + kanamycin (Kan) 50 µg/ml culture media in a 10 ml tube. All samples were cultivated overnight (o/n) at 37 °C and 220 rounds per minute (rpm) shaking. Bacterial cells were

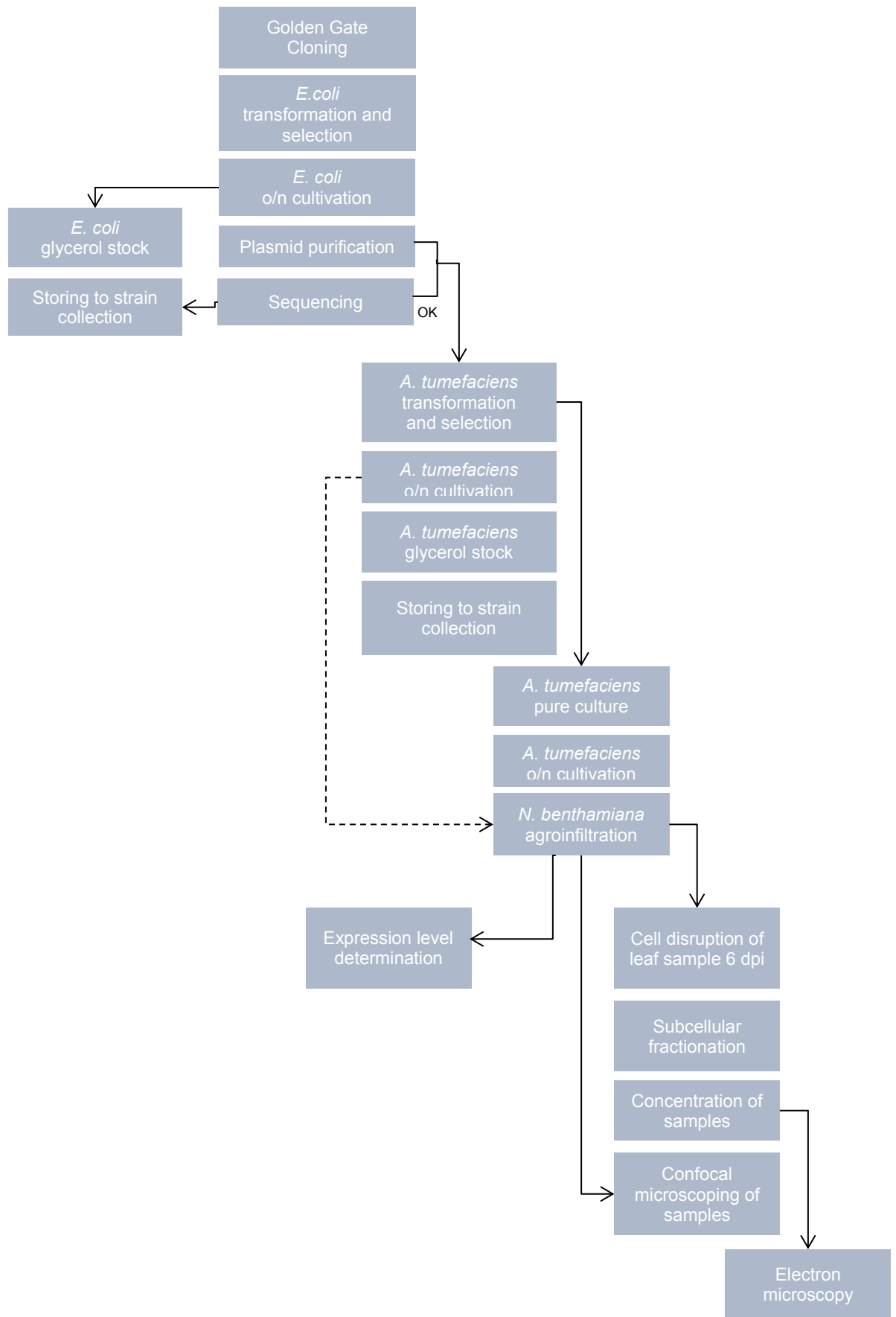


Figure 6. Outline of experiment procedures

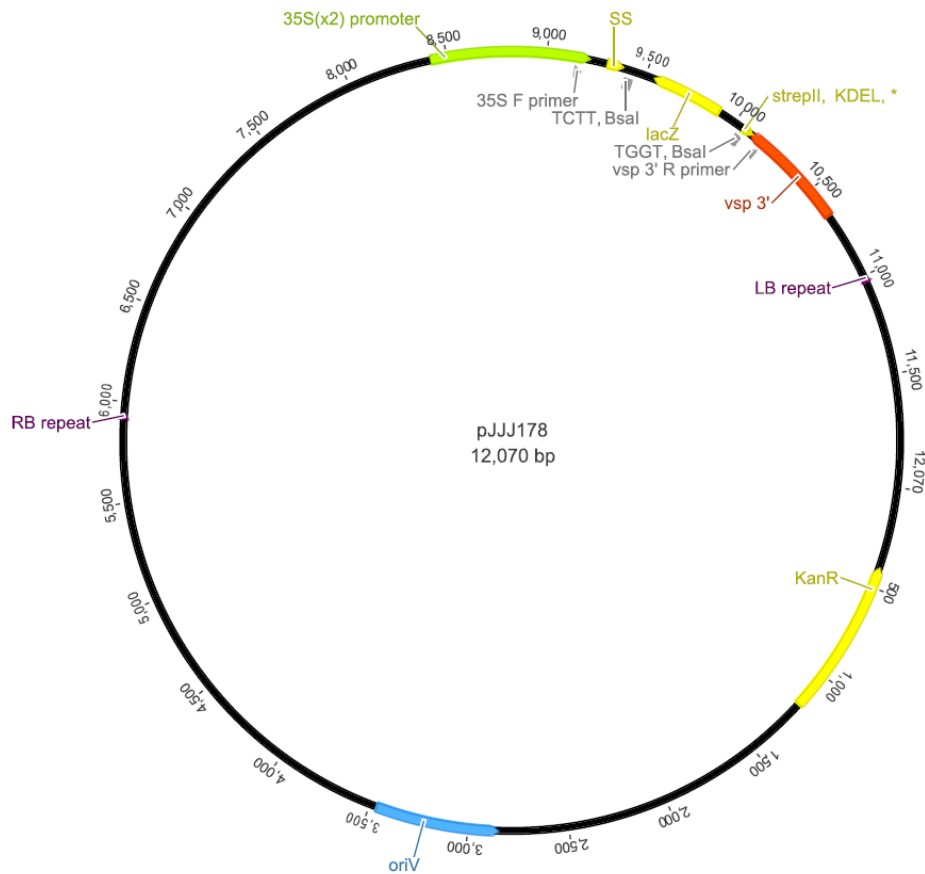


Figure 7. A. Plasmid map of pJJJ178 destination vector. In successful cloning, lacZ (β -galactosidase) gene is replaced by insert genes. Kanamycin resistance (KanR) gene is essential for plasmid selection. Left border repeat (LB repeat) and right border repeat (RB repeat) are important for *Agrobacterium* –mediated transient expression experiments in *N. benthamiana*: the gene sequence between LB and RB repeat is transferred to *N. benthamiana*. oriV sequence shows the origin of plasmid replication. Essential sequences for PB induction are signal sequence (SS), which traffics the expressed protein to ER, and KDEL ER-retention signal, which traps the protein in ER preventing the protein from being secreted out of the cell. Streptactin II (StrepII) tag is used in immunoblotting of proteins. Promoter and terminator of inserted gene fragment are 35S(x2) and vsp 3', respectively.

harvested by centrifugation at 3200 g and 4 °C for 10 min (Eppendorf Centrifuge 5810R, Eppendorf, Germany). Plasmids were purified from harvested *E. coli* cells by using QIAprep Spin Miniprep Kit (Anon., 2006) and concentration of plasmid was determined by using Nanodrop 2000 (Thermo Scientific, USA).

3.1.2 Golden Gate cloning

The cloning was done by using Golden gate cloning method, by which it is possible to digest and ligate directionally multiple DNA fragments in the same reaction. The method is based on type IIS restriction enzymes (RE), which have separate recognition and cutting sites on the DNA. In these experiments, the *BsaI* restriction enzyme was used for cloning. *BsaI* cuts any DNA sequence located at 1 nucleotide towards 3' end of the recognition site leaving a specific sticky end to the cutting site (Figure 8) (Engler *et al.* 2008; Engler *et al.* 2009; Weber *et al.* 2011). In plasmid pJJJ178, the *BsaI* recognition sites are located upstream and downstream of β -galactosidase (*lacZ*) gene. In successful cloning, the *lacZ* gene is replaced by the insert gene fragments.

The composition of Golden Gate reaction mixtures is shown in Table 4. All entry vectors were pipetted to same test tube with enzymes. Due to error, reactions of pJJJ751-753 were pipetted on room temperature (RT) instead of on ice. Plasmids were digested with *BsaI* (New England Biolabs Inc., USA) and ligated with T4 DNA ligase (Promega Corporation, USA). Reactions of pJJJ754-755 included 100 μ g/ml of bovine serum albumin (BSA) and reactions were pipetted on ice. In the cloning of pJJJ755 the amount of entry vector with 2-linker1-eGFP-linker1-3 gene fragment was 30 ng instead of 100 ng.

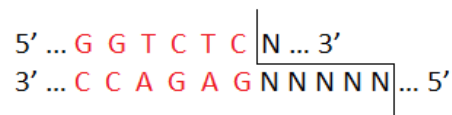


Figure 8. *BsaI* recognition site (in red) and cutting site (modified from Engler *et al.* 2009).

The Golden Gate assembly protocol (Table 5) was run with BIO-RAD MINI Personal Thermal Cycler (Bio-Rad Laboratories, Inc., USA). A control sample was used to monitor the effectiveness of cloning and transformation. The control sample was handled as other cloning samples except one entry vector was replaced with same volume of distilled water (DDW).

Table 4. Golden Gate reaction mixtures.

A. Reaction mixture		B. Reaction mixture with BSA addition	
Sample	amount	Sample	amount
Destination vector	250 ng	Destination vector	250 ng
Entry vectors	à 100 ng	Entry vectors	à 100 ng
Bsal (NEB)	10 U	Bsal (NEB)	10 U
10x ligase buffer (Promega)	1x	BSA	2 µg
T4 DNA ligase (Promega)	3 U	10x ligase buffer (Promega)	1x
add DDW		T4 DNA ligase (Promega)	3 U
Total volume	20 µl	add DDW	
		Total volume	20 µl

Table 5. Golden Gate assembly protocol.

Temperature (°C)	Time (min)
37	90
50	5
80	10
10	∞

3.1.3 Transformation of *E. coli*

The cloning product and control sample was transformed into chemically competent *E. coli* DH5α cells. Cells were thawed on ice for 20 min. 5 µl of cloning reaction mixture was added to 100 µl of chemically competent *E. coli* DH5α cells and incubated on ice for 30 min and heat shocked 30 s at 42 °C. 1 ml of LB was added and cells were cultivated for 1-1,5 h at 37 °C and in 220 rpm shaking. Cells were harvested by centrifugation with Eppendorf Centrifuge 5424 (Eppendorf, Germany) at 2500 g for 3 min. The supernatant

was discarded, cells were resuspended gently in 100 µl LB, and two amounts (10 µl and 90 µl) of cells were plated on LB + Kan 50 µg/ml + 100 µl X-galactose (X-gal) (20 mg/ml) plates for selection. Plates were incubated o/n at 37 °C in the dark.

3.1.4 Sequencing and storing to strain collection

A single white *E. coli* colony from selection plate was cultivated o/n in 10 ml LB + Kan 50 µg/ml in 50 ml tube at 37 °C, 220 rpm shaking. Bacterial cells were harvested by centrifugation at 3200 g, at 4 °C for 10 min (Eppendorf Centrifuge 5810R, Eppendorf, Germany). Plasmids were purified from harvested *E. coli* cells using QIAGEN QIAprep Spin Miniprep Kit (Anon., 2006) and concentration of plasmid was determined using Nanodrop 2000 (Thermo Scientific, USA).

The insert was sequenced with primers 35S F (TTCATTTGGAGAGGACACGC) and vsp3' R (ATAGTGCATATCAGCATACC) which anneal outside of the insert region in the pJJ178 plasmid. Sequencing was performed by GATC Biotech LightRun Sequencing service (Germany, <http://www.gatc-biotech.com/lightrun>). The sequence data was analyzed by comparing the sequencing data to corresponding plasmid map by using Geneious 7.1.8 software (Map to reference tool).

500 µl of o/n cell culture was added to 500 µl 40 % glycerol for storing the strain (two samples of each strain). Samples were placed to VTT Ltd strain collection after the DNA sequences were confirmed.

3.1.5 Transformation of *Agrobacterium tumefaciens*

Purified pJJ751-pJJ755 constructs were transformed into *A. tumefaciens* strain EHA 105 (Hood *et al.* 1993) by electroporation. Electrocompetent *A. tumefaciens* EHA 105 cells were thawed on ice. Approx. 100-150 ng of purified plasmid were added to 40 µl of cells. After 1 min incubation on ice, cells were electroporated in pre-cooled electroporation cuvette (200 Ω, 25 µFD and 2,5 kV) with BIORAD (USA) electroporation equipment (Capacitance Extender Plus, Pulse Controller Plus and Gene Pulser II). Cells were

transferred to a 15 ml tube containing 1 ml SOC media (10 mM MgCl₂, 10 mM MgSO₄, 0,36 % (w/v) glucose in SOB media; SOB is 2 % (w/v) tryptone, 0,5 % (w/v) yeast extract, 10 mM NaCl, 2 mM KCl). Cells were incubated for 1h at 28 °C and in 220 rpm. Cells were harvested by centrifugation 3 min in 2500 g with Eppendorf centrifuge 5424 (Eppendorf, Germany). 90 % of the supernatant was discarded and the pellet was gently resuspended to approx. 100 µl of supernatant that was left in the tube. Two amounts of cells, 10 µl and approx. 90 µl, were plated on LB + Kan 50 µg/ml + Rifampicin (Rif) 10 µg/ml selection plates and incubated 3 days at 28 °C.

For storing to VTT Ltd strain collection, a single colony from selection plate was cultivated o/n in 10 ml LB + Kan 50 µg/ml + Rifampicin (Rif) 10 µg/ml in 50 ml tube, at 28 °C, in shaking. 500 µl of o/n cell culture was added to 500 µl 40 % glycerol for storing the strain (two samples of each strain).

3.2 Production system: plant transformation and growth conditions

3.2.1 Plant production

N. benthamiana tobacco seeds were planted into pots of KekkiläGarden W R8015 peat mixture (Kekkilä Ltd, Finland) which was pretreated with Nemasys® *Steinernema feltiae* nematodes (Becker Underwood Inc., UK) for biocontrol of fungus gnat (*Bradysia* spp.) flies. Plants were cultivated hydroponically for 6 weeks with photoperiod of 16 h, in the light of 120 µEm⁻²s⁻¹ by 250 W Plantastar Inter lamps (Osram GmbH, Germany), at 20 °C temperature and 50-60 % relative humidity. The hydroponic substrate included 0,15 % (v/w) Hydroflex L (Everris, Australia) and 0,15 % (v/w) YaraLiva Calcinit (Yara International ASA, Norway) fertilizers. Plants were grown first 1-2 weeks with substrate, which included half of the aforementioned amount of fertilizers, and then with full amount of fertilizers.

3.2.2 Plant agroinfiltration

To infect *N. benthamiana* plants, an *A. tumefaciens* liquid culture was made: a single colony of the appropriate *A. tumefaciens* strain picked from a freshly grown *A.*

tumefaciens plate culture was cultivated in 10 ml LB + Kan 50 µg/ml + Rif 10 µg/ml in 50 ml tube, o/n, at 28 °C and in shaking.

A. tumefaciens cultures were diluted in 10 mM MgSO₄ - 10 mM MES (4-morpholineethanesulfonic acid), pH 5.0 according to how many different *A. tumefaciens* strains were co-expressed in the plant, and appropriate cultures were mixed to acquire final cell density of approx. OD_{600nm} = 0,3 - 0,4 for single strain in the mixture (Table 6). Absorbance at 600 nm was measured with UV/Vis spectrophotometer (Ultrospec 2100 Pro, Amersham Biosciences, UK) and DDW was used as blank. To infect *N. benthamiana* plant with *A. tumefaciens* culture, 3 upper leaves of 5-6 weeks old *N. benthamiana* plants were agroinfiltrated by injection of culture mixture into the abaxial side of the leaf. Plants were kept in the dark for 1-2 h to recover from infiltration before returning them to greenhouse. Depending on the experiment, *A. tumefaciens* strains were infiltrated either in the presence or absence of *A. tumefaciens* strain of plasmid p19. Plasmid p19 includes silencing suppressor gene which enhances the production of the recombinant protein in the plant (Voinnet *et al.* 2003).

Table 6. Dilutions and mixing ratio for *A. tumefaciens* cultures before agroinfiltration.

Amount of <i>A. tumefaciens</i> strains	Cell density after dilution (OD _{600nm})	Mixing ratio	Cell density of single strain in the mixture (OD _{600nm})
1	0,3-0,4	-	0,3-0,4
2	0,7-0,9	1:1	0,3-0,4
3	0,9-1,2	1:1:1	0,3-0,4
4	approx. 1,2	1:1:1:1	approx. 0,3

3.3 Protein body isolation by subcellular fractionation

One agroinfiltrated (5-8 dpi) or wild-type *N. benthamiana* leaf (negative control sample, from 7 weeks old plant) was homogenized in mortar on ice with 10 % (w/w) sucrose in HB. Approx. 3,5 ml buffer was used for 1 g of leaf tissue. Homogenized leaf tissue extract was filtrated through Merck Miracloth 22-24 µm (Millipore, USA) filter, which was folded in four layers, or alternatively filtrated four times through one layer of Miracloth filter.

To determine whether PBs remained spherical during the homogenization procedure, two samples (NTOT and N1) were taken from leaf extracts (Table 7). Sample N1 was centrifuged 20 000 g, 10 min, in 4 °C (Eppendorf Centrifuge 5430R, Eppendorf, Germany), and samples were taken from supernatant (N1L) and pellet (N1P, resuspended to 120 µl HB). SDS-PAGE and western blot were prepared from samples NTOT, N1L and N1P.

Table 7. Samples from homogenization step.

Sample	Volume (µl)	From
NTOT	150	homogenized leaf tissue
N1	200	homogenized leaf tissue
N1L	120	after centrifugation of sample N1; supernatant
N1P	120	after centrifugation of sample N1; pellet resuspension with HB

Subcellular fractionation of *N. benthamiana* leaf extracts was done by sucrose gradient ultracentrifugation based on the procedure by Torrent *et al.* (2009b). Sucrose gradients (Figure 9) were made from 20, 30, 42 and 56 % (w/w) sucrose in homogenization buffer (HB) (100 mM Tris-HCl pH 8, 50 mM KCl, 6 mM MgCl₂, 1 mM EDTA, 0,4 M NaCl) to Ultra-Clear™ 14x89 mm (Beckman, USA) ultracentrifuge tubes.

The filtrate was centrifuged for 5 min at 50 g and 4 °C with Eppendorf Centrifuge 5810R (Eppendorf, Germany). 1 ml supernatant was loaded on top of fresh sucrose gradient and 10 % sucrose in HB was added on top of the sample to fill and balance ultracentrifuge tubes. Sucrose gradients were centrifuged for 2 h at 80 000 g and 4 °C with Beckman Optima LE-80K ultracentrifuge (Beckman, USA) and Beckman SW 41Ti rotor (Beckman, USA). Rotor was cooled in +4 °C before use.

After ultracentrifugation, samples were taken from sucrose gradient interphases and from the pellet (Figure 9). Approx. 1 ml samples were collected from supernatant (S) and interphases (I1-I3) by puncturing the ultracentrifuge tube with an injection needle. The pellet was resuspended to 200 µl 56 % sucrose in HB. 10 µl sample of each gradient

fraction was taken for SDS-PAGE and western blot, and rest was stored in -20 °C for microscopy.

Fractionation samples were observed with confocal microscope without further preparation, and after concentration. Samples for microscopy were concentrated by mixing 100 µl of sample with 1 ml of HB and centrifuged 20 000 g, 10 min, 4 °C (Eppendorf Centrifuge 5430R, Eppendorf, Germany). If visible pellet was formed, supernatant was discarded and pellet was resuspended to 20 µl of HB.

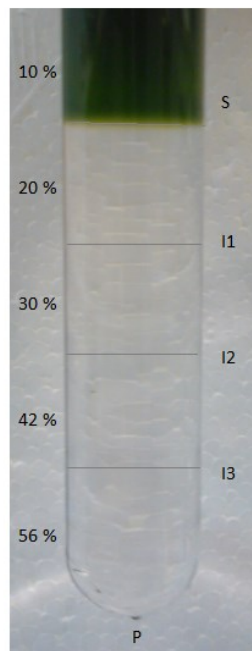


Figure 9. Sucrose gradient showing sucrose concentrations and sample fractions S, I1, I2, I3 and P.

3.4 SDS-PAGE

All SDS-PAGE samples were diluted 3:1 in 4x Laemmli sample buffer with 10 % (v/v) β -merkaptoethanol (LSB-BME) and samples were heated for 5 min to 95 °C. In case of the samples were frozen for storage, they were reheated similarly before loading to SDS-PAGE gel. 5 µl Precision Plus Protein™ Dual Xtra Standard (BIORAD, USA) protein standard was

used in all gels. Criterion TXG Precast Gel Any kD gels (BIORAD, USA) were used. Gels were run at 200V for 30-60 min. Before staining, gels were incubated 2 times for 30 min in reverse osmosis (R/O) water. Gels were stained with GelCode Blue Stain Reagent (ThermoScientific, USA) and scanned with GS 710 Calibrated Imaging Densitometer (BIORAD, USA).

3.5 Western blotting

Western blotting was made from non-stained SDS-PAGE gel immediately after running the gels. Proteins were transferred from gel to Transblot Turbo™ Transfer Pack Midi Format 0,2 µm nitrocellulose membranes (BIORAD, USA) by Trans-Blot Turbo Transfer System (BIORAD, USA) using Preprogrammed Protocol for Mixed MV (7 min, 1 Midi Gel). Due to error in procedure, gels were incubated 2 times for 30 min in (R/O) water before blotting instead of immediate blotting of gels. After transfer, the membrane was blocked with 50 ml 5 % (w/v) fat free milk powder in TBST (0,1 M Tris-HCl pH 8.0, 0,15 mM NaCl, 0,05 % (v/v) Tween 20), 90 min at RT, or o/n at 4 °C. Blocking solution was rinsed shortly 2 times with TBST. For detection, the membrane was incubated in a solution of 1/2000 Strep-Tactin-alkaline phosphatase (AP) conjugate (IBA GmbH, Germany) in TBST and 3 % (w/v) fat free milk, 1 h at RT. Membrane was rinsed shortly 2 times with TBST and washed 2 times 5 min with TBST. AP was detected by incubating 15 min at RT in fresh AP detection solution (66 µl NBT (Promega, USA) and 33 µl BCIP (Promega, USA) in 10 ml 100mM Tris-HCl pH 9.5, 100mM NaCl and 5 mM MgCl₂). Reaction was stopped by rinsing with R/O water and membrane was dried between Whatman paper o/n. Membrane was scanned with GS 710 Calibrated Imaging Densitometer (BIORAD, USA).

3.6 Expression levels

3.6.1 Experimental design

Three top leaves of *N. benthamiana* were divided to 6 agroinfiltration spots, labelled 1 to 6. Each spot (Figure 10) was located in different plant, either in lowest, middle or highest (oldest to youngest) leaf. A single or multiple *A. tumefaciens* strains were agroinfiltrated to each leaf spot. Each single construct was co-expressed with p19 protein. For one analysis, a leaf disc was collected at 6 dpi from each spot and combined to one sample. Four parallel samples were collected from same agroinfiltration spots to have four technical replicates.

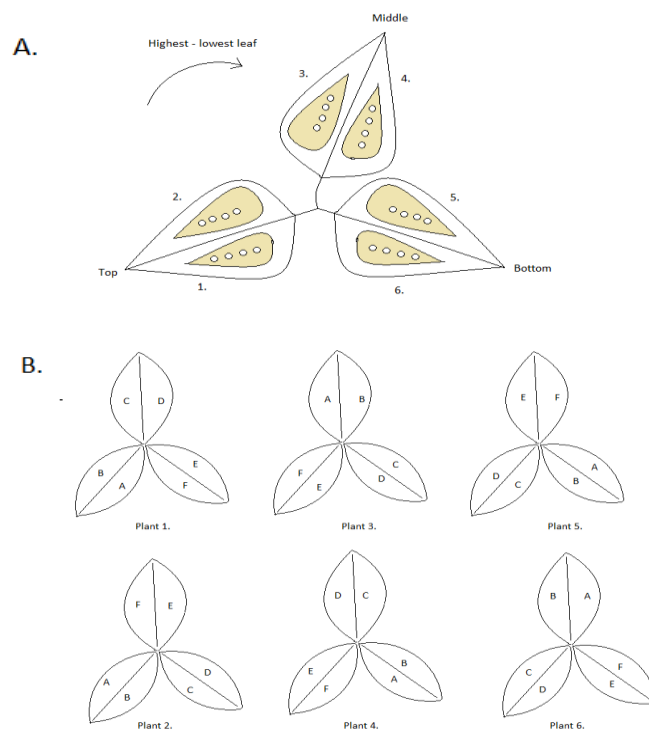


Figure 10. Sample rotation in expression level experiment (A) and agroinfiltration plan for single constructs (B). Letters A-F represent expressed constructs pJJ751-755.

The expression level was calculated as mean value with standard deviation from four parallel determinations. Expression of eGFP-HFBI (pJJJ247) was used as internal control in all experiments. To determine the expression level of pJJJ751-755, three top leaves of six plants were agroinfiltrated with pJJJ751-755 and pJJJ247 *A. tumefaciens* strains according to Figure 10B.

3.6.2 Sample preparation and protein concentration measurement

For one measurement, 6 leaf discs with 5 mm diameter were collected to one 2 ml microtube, which included 3 ceramic pearls for sample homogenization. Sample was immediately frozen in liquid nitrogen and stored to -80 °C until homogenization. Homogenization blocks were cooled to -80 °C. Samples were dipped in liquid nitrogen and placed to homogenization blocks. Frozen samples were homogenized in microtubes with Retsch MM01 mill (Retsch GmbH, Germany) two times for 1 min and 30 Hz, reverting the block in the holder before the second round. Homogenized samples were centrifuged for 1 min at 3200 g and 4 °C with Eppendorf Centrifuge 5810R (Eppendorf, Germany). Samples were kept on ice and 300 µl cold buffer (0,1 % (v/v) Tween-20, 1 mM EDTA, 2 % (w/v) ascorbic acid in phosphate buffered saline (PBS) buffer (12 mM Na₂HPO₄ * 4 H₂O – 3 mM NaH₂PO₄ * H₂O – 0,15 M NaCl)) was added on samples and mixed in vortex. 100 µl 4x LSB+BME was added to samples and heated 5 min at 95 °C.

SDS-PAGE and western blot were made to determine protein concentrations of genetic constructs that were expressed. Purified eGFP-HFBI protein (990 ng/µl) was used to generate a standard curve (1000 ng, 500 ng, 250 ng, 125 ng, 62,5 ng). Four microliters of each sample was loaded to SDS-PAGE gel, so that the volume of homogenized sample was 3 µl (+ 1 µl of loading dye) in the gel, except for samples pJJJ751II and pJJJ754II. For these samples the amount of homogenized sample in the gel was 1,8 µl due to the error in sample preparation. This was taken into account when calculating the final protein concentration of samples.

The protein amount of target proteins was quantified from western blot by ImageQuant TL software (GE Healthcare Europe GmbH, Germany) by using Precision Plus Protein™ Dual Xtra Standard (BIORAD, USA). The target protein amounts were quantified by densitometry using the above mentioned software and eGFP-HFBI standards. The target protein concentration (mg/g leaf fresh weight) was calculated by equation 1,

$$C = \frac{m_{sample,blot}}{0,75 \times V_{sample,blot}} \times \frac{V_{sample,tot}}{m_{leaf}} \quad (1)$$

in which $m_{sample,tot}$ is quantified protein amount (mg) from western blot, $V_{sample,blot}$ is sample volume in SDS-PAGE gel (μ l), $V_{sample,tot}$ is total volume of sample (400 μ l) and m_{leaf} is average mass of six 5 mm leaf discs (g). m_{leaf} was calculated as mean value from four samples including six 5 mm leaf discs.

3.7 Imaging

3.7.1 Confocal microscopy

N. benthamiana leaf tissue samples were prepared by cutting a 7 mm leaf discs from agroinfiltrated leaves and placing them on microscope slide abaxial side upwards. Vaseline was used to surround the leaf disc and water droplet was placed on top of the leaf. Cover glass was pressed on the leaf and vaseline so that water drop remained between the leaf sample and the cover glass.

Subcellular fractionation samples were observed as collected or after concentration. A drop of sample was pipetted on objective slide and covered with cover glass. Cover glass was sealed by nail polish.

Samples were observed with Carl ZEISS LSM 710 laser scanning microscope (Carl ZEISS Microscopy GmbH, Germany) and with either EC Plan-Neofluar 40x/0.75 M27 or LCI Plan Neofluar 63x/1.3 Imm Korr DIC M27 (Carl ZEISS Microscopy GmbH, Germany) objectives. The former objective was used only for agroinfiltration I experiment. Water immersion was used with 63x objective for all samples. Samples were observed max. 3 hours after

preparation. Microscopy pictures were acquired and processed by using ZEN 2011 Imaging software (Carl ZEISS Microscopy GmbH, Germany).

In all experiments, excitation of eGFP and mRFP was done with 488 nm and 543 nm lasers, respectively. The emission window to gather the signal varied between experiments. eGFP and mRFP signals were gathered from approx. 490 – 540 nm and 600 – 650 nm emission windows, respectively. In co-expression experiments, eGFP and mRFP emission signals were gathered separately by using two settings: one for excitation and signal gathering of eGFP and another for mRFP by using aforementioned wavelength settings. The absence of background of eGFP and mRFP signals was confirmed by observing wild-type leaf sample with these settings. By observing eGFP leaf sample with mRFP settings, it was confirmed that eGFP signal is not leaking to mRFP channel. Similarly, it was also confirmed that mRFP signal is not leaking to eGFP channel by observing mRFP leaf sample with eGFP settings.

3.7.2 Electron microscopy

Electron microscopy was performed by Dr. Nonappa at the Molecular materials group in Department of Applied Physics in Aalto University School of Science.

Samples for electron microscopy were prepared from I3 fraction of subcellular fractionation samples. 200 μ l of collected fraction was mixed with 1 ml of HB and centrifuged 20 000 g, 10 min, 4 °C (Eppendorf Centrifuge 5430R, Eppendorf, Germany). Most supernatant was discarded and 50 μ l of sample was left to the bottom of the tube and 50 μ l of HB was added.

4 Results

4.1 Plasmid constructs

Plasmids that were used in the expression experiments of this study are shown in Table 8. New plasmids ZERA-mRFP (pJJJ751), ZERA-eGFP (pJJJ752), ZERA-HFBI (pJJJ753), mRFP-HFBI (pJJJ754) and ZERA-eGFP-HFBI (pJJJ755) were designed and cloned for the experiments of this master thesis. Plasmids pJJJ751-755 were transformed to *E. coli* and *A. tumefaciens* strains and stored to VTT Ltd strain collection (Table 8). Figure 11 shows the plasmid maps of pJJJ751-755, pJJJ247 and pJJJ147.

Table 8. Plasmids that were used in the experiments. Plasmids in the gray background were already available in *E. coli* and *A. tumefaciens* cells at VTT Ltd strain collection.

Plasmid	Gene fragment of interest	Code in VTT Ltd strain collection	
		<i>E. coli</i>	<i>A. tumefaciens</i>
pJJJ751	ZERA-mRFP in pJJJ178	B6454	B6457
pJJJ752	ZERA-eGFP in pJJJ178	B6455	B6458
pJJJ753	ZERA-HFBI in pJJJ178	B6456	B6459
pJJJ754	mRFP-HFBI in pJJJ178	B6747	B6748
pJJJ755	ZERA-eGFP-HFBI in pJJJ178	B6749	B6750
pJJJ247	ZERA-eGFP in pJJJ178	-	B5764
pJJJ147	eGFP in pJJJ178	B4047	B4048
p19	silencing suppressor gene	-	B4090

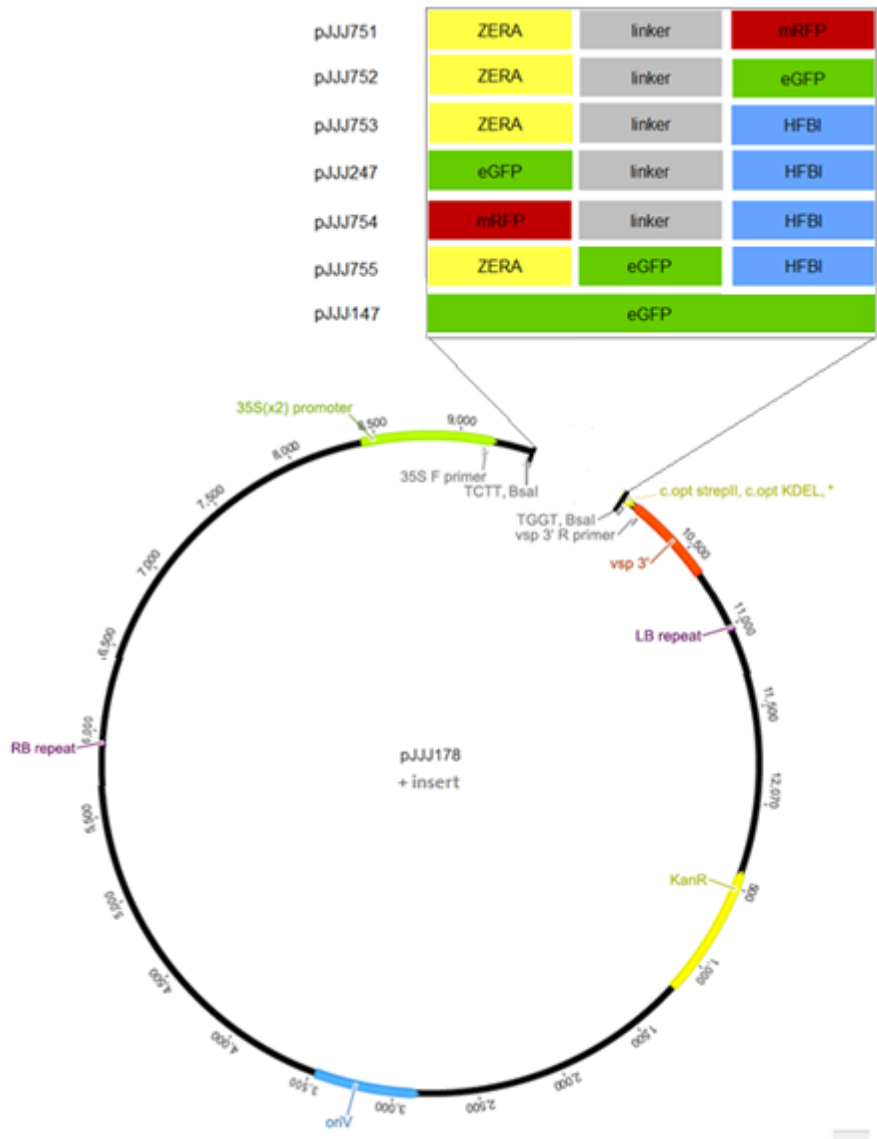


Figure 11. Plasmid maps of plant expression vectors that were used in this study. StrepII tag lacks from plasmid pJJJ147.

4.2 Verification of expression

The coding region of the cloned plant expression vectors was verified by sequencing. The constructs were agroinfiltrated to *N. benthamiana* for transient expression experiments and the target protein accumulation was confirmed by immunoblotting in the expression

level experiment. The PB formation was followed by confocal microscopy (except for pJJJ753 ZERA-HFBI).

4.2.1 ZERA-mRFP, ZERA-eGFP and eGFP-HFBI

ZERA-mRFP (pJJJ751) and ZERA-eGFP (pJJJ752) were first expressed without plasmid p19. Agroinfiltrated leaf samples were observed at 3 dpi. The expression level of samples was low and only a small amount of PBs formed. Upon the microscopy, it was noticed that ZERA-mRFP agroinfiltration culture was cross contaminated by ZERA-eGFP *A. tumefaciens* cells (Figure 12, A-C). As a result of the contamination, it seemed that ZERA-mRFP and ZERA-eGFP proteins accumulate in the same PBs when co-expressed. Expression of ZERA-mRFP fusion protein was repeated to confirm that *A. tumefaciens* strain pJJJ751 produces only ZERA-mRFP. Leaf sample was observed at 6 dpi with confocal microscope. Two different channels were used for excitation and gathering of eGFP and mRFP signal. PBs were observed using mRFP channel and no PBs were visible when observed with eGFP channel. Therefore, the earlier contamination with ZERA-eGFP happened most probably during the agroinfiltration procedure.

To have higher protein accumulation, the expression of ZERA-mRFP and ZERA-eGFP was repeated by expressing them with gene silencing suppressor p19. For comparison, also eGFP-HFBI (pJJJ247) was co-expressed with p19. Samples were observed at 4 dpi. High amount of spherical PBs were observed in ZERA-eGFP (Figure 13A), ZERA-mRFP (Figure 13B) and eGFP-HFBI (Figure 13D) samples. The diameter of these PBs was approx. 1 μm . Co-expression with plasmid p19 clearly increased the accumulation of target protein.

4.2.2 mRFP-HFBI, ZERA-eGFP-HFBI and eGFP

mRFP-HFBI (pJJJ754) and ZERA-eGFP-HFBI (pJJJ755) were expressed in the presence of p19. Agroinfiltrated leaves were observed at 5 dpi. eGFP and mRFP signals were gathered by using separate settings. In the ZERA-eGFP-HFBI sample, there were high amount of symmetrical spherical PBs (Figure 13F). The diameter of PBs was approx. 2 μm , and these PBs appeared to be more clearly defined spheres than PBs observed in case of expressing

ZERA or HFBI alone. In mRFP-HFBI sample, only small amount of PBs were visible (Figure 13E). Sample seemed to have low expression level. For comparison, also eGFP (pJJJ147) was expressed with p19 and the leaf sample was observed 3 dpi. Instead of clear PB formation, eGFP was mainly visible in the web of ER membrane and some small asymmetrical aggregates were found (Figure 13C).

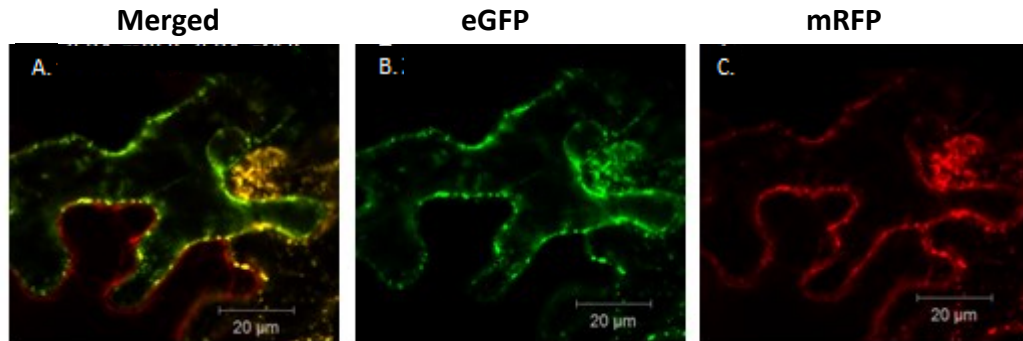


Figure 12. ZERA-mRFP with ZERA-eGFP cross-contamination: A. combined eGFP and mRFP channels, B. eGFP channel, C. mRFP channel. Yellow color represents presence of eGFP and mRFP.

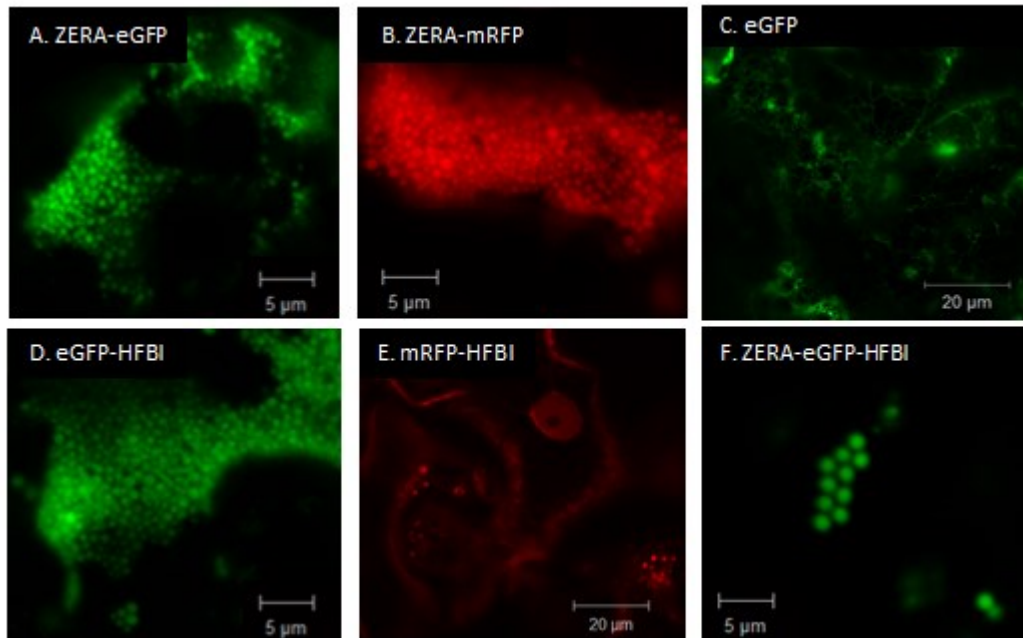


Figure 13. PB formation in *N. benthamiana* leaves: A. ZERA-eGFP at 4 dpi, B. ZERA-mRFP at 4 dpi, C. eGFP at 3 dpi, D. eGFP-HFBI at 4 dpi, E. mRFP-HFBI at 5 dpi, F. ZERA-eGFP-HFBI at 5 dpi.

4.2.3 Expression levels

The expression levels of constructs are shown in Figure 14. ZERA-eGFP had highest expression level ($1,66 \pm 0,31$ mg/g leaf fresh weight (FW)), which is approx. four times higher than ZERA-mRFP ($0,39 \pm 0,21$ mg/g FW). The expression levels of ZERA-eGFP-HFBI and eGFP-HFBI were low ($0,14 \pm 0,03$ and $0,21 \pm 0,06$ mg/g FW, respectively). The amount of ZERA-HFBI (pJJJ753) was too low to be quantified from the western blot membrane (Figure 15), although there were faint bands visible at 25 kDa corresponding to the expected molecular weight (MW) of ZERA-HFBI (21 kDa). This indicates that ZERA-HFBI is expressed in *N. benthamiana* in low amounts. Also the amount of mRFP-HFBI was too low to be quantified from immunoblot membrane (Figure 15), which confirmed the low amount and only few minuscule mRFP-HFBI PBs observed upon microscopy experiments. Notable is that due to limitations of processing the signal from western blot membrane, this experiment gives a qualitative estimate of expression levels, illustrating the differences between the strains.

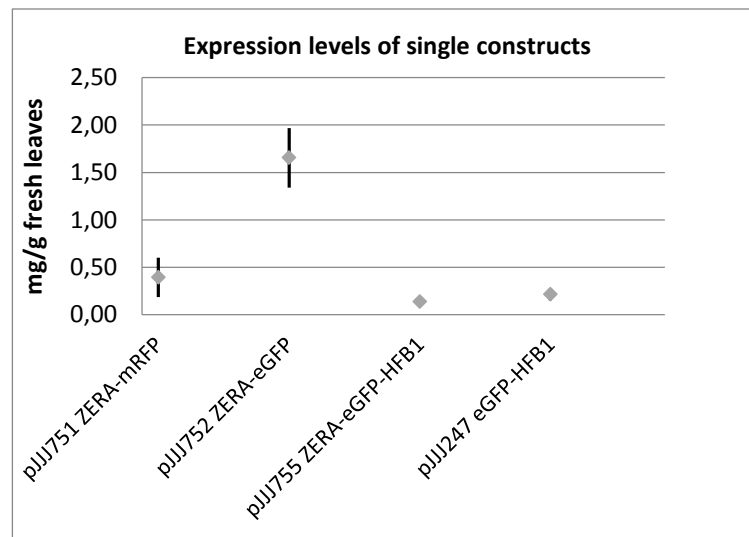


Figure 14. Expression levels of single constructs (mg/g FW) with standard deviation. Protein amounts of mRFP-HFBI and ZERA-HFBI were too low to be quantified.

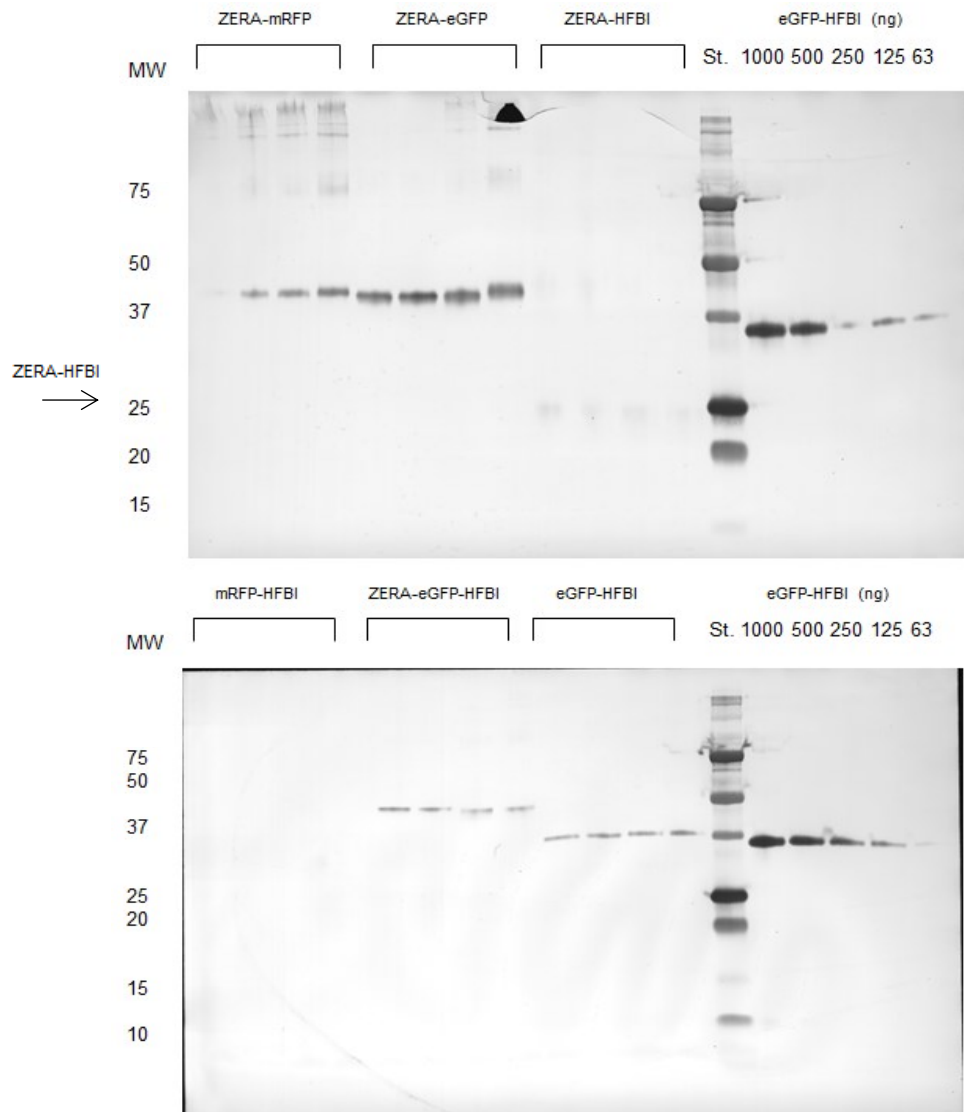


Figure 15. Western blots of expression level experiment. Expected MWs of fusion proteins are: ZERA-mRFP (39 kDa), ZERA-eGFP (40 kDa) and ZERA-HFBI (21 kDa), mRFP-HFBI (36 kDa), ZERA-eGFP-HFBI (49 kDa) and eGFP-HFBI (38 kDa). ZERA-HFBI is faintly visible in the membrane at 25 kDa marker.

4.3 Co-expression experiments

4.3.1 ZERA and HFBI

ZERA-mRFP and eGFP-HFBI were co-expressed to repeat the result of previous research (Dr. Reza Saberianfar, Agriculture and Agrifood Canada; unpublished results) in which ZERA and HFBI formed separate PBs in *N. benthamiana*. The co-expression was done in the presence of the gene silencing suppressor p19 and the sample was observed at 4 dpi. As it was expected, ZERA-mRFP and eGFP-HFBI accumulated in separate PBs (Figure 16A). Interestingly, there were also Janus particle resembling PBs having both eGFP-HFBI and ZERA-mRFP fusion proteins unmixed (Figure 16B). The co-expression of ZERA-mRFP and eGFP-HFBI was repeated to confirm the result of separate PB formation. Sample was observed at 6 dpi. Again, separate ZERA-mRFP and eGFP-HFBI PBs was formed and some Januslike particles were found.

Based on these results the question was, whether ZERA and HFBI form separate PBs when fluorescent tags are exchanged. For this purpose, mRFP-HFBI construct was cloned. The co-expression experiment was repeated with ZERA-eGFP and mRFP-HFBI (in the presence of p19) and the sample was observed 5 dpi. Clear ZERA-eGFP PBs were visible. Unexpectedly, clear mRFP-HFBI PBs were not observed. Instead, mRFP-HFBI signal seemed to surround ZERA-eGFP derived PBs (Figure 16). Based on this observation, the expression levels of single constructs were determined to understand whether the difference in expression levels could explain the phenomenon.

4.3.2 ZERA-mRFP, ZERA-HFBI and eGFP-HFBI

Since ZERA-mRFP and ZERA-eGFP co-localize in the same PB (based on the observation in the cross contamination), and ZERA-mRFP and eGFP-HFBI do not co-localize, it was tested whether ZERA-HFBI fusion protein can change the behaviour and make ZERA-mRFP and eGFP-HFBI to accumulate in the same PBs. ZERA-mRFP, ZERA-HFBI, eGFP-HFBI were co-expressed in *N. benthamiana* and was observed at 4 dpi. The triple co-expression of ZERA-

mRFP, eGFP-HFBI and ZERA-HFBI answered positive to the question: both fluorescent proteins seemed indeed to co-localize in the same PBs when ZERA-HFBI was present (Figure 17A). Nevertheless, the resolution in the confocal microscopy was not sufficient to determine how ZERA-mRFP, eGFP-HFBI and ZERA-HFBI were arranged within the PB.

4.3.3 ZERA-mRFP and ZERA-eGFP-HFBI

Since ZERA-HFBI alone is not visible on confocal microscope, a new construct with fluorescent tag, ZERA-eGFP-HFBI (pJJJ755), was cloned. ZERA-ZERA interaction was studied more closely by co-expressing ZERA-mRFP with ZERA-eGFP-HFBI and observing at 5 dpi by confocal microscopy. As a result, ZERA-mRFP and ZERA-eGFP-HFBI seemed to locate in same PBs (Figure 17B). This confirmed the hypothesis that ZERA-ZERA interaction is strong enough to assemble two ZERA fusion proteins to the same PB regardless of the presence of HFBI.

4.3.4 mRFP-HFBI and ZERA-eGFP-HFBI

HFBI-HFBI interaction was studied in the similar way as ZERA-ZERA interaction. mRFP-HFBI and ZERA-eGFP-HFBI were co-expressed in *N. benthamiana* and leaf samples were observed at 5 dpi. Two kinds of PBs were found: similar PBs as produced in the expression of ZERA-eGFP-HFBI (Figure 17C) and, unexpectedly, “hollow” PBs that seemed to have a brighter outer surface (Figure 17D). Additionally, mRFP-HFBI signal was observed in the ER membrane (visible in the Figure 17C). Both fusion proteins seemed to be present within the same PBs which confirmed that HFBI fusion proteins can self-assemble into the same PB regardless of the presence of ZERA.

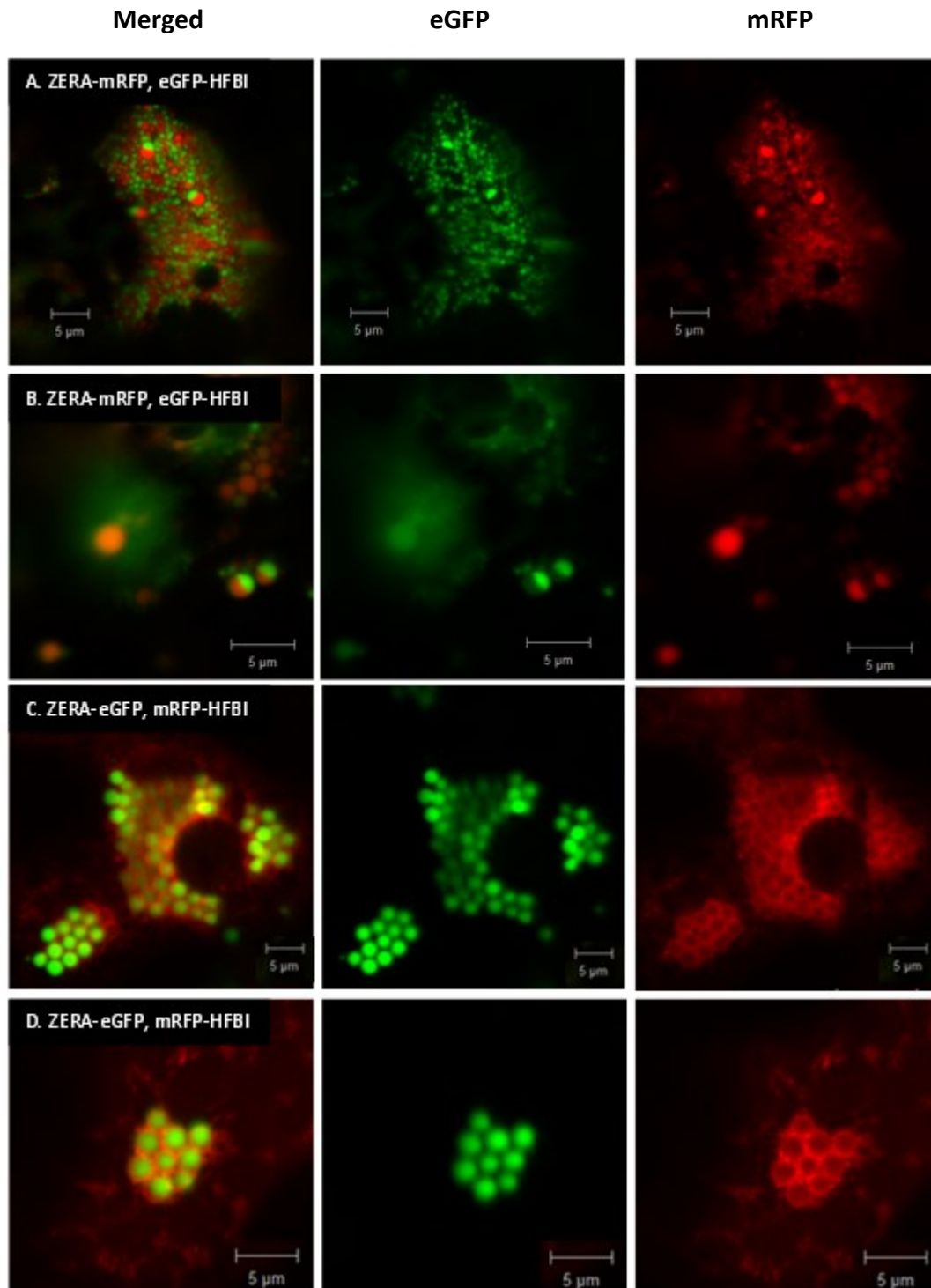


Figure 16. Co-expression of ZERA-mRFP and eGFP-HFBI at 4 dpi (A, B). Co-expression of ZERA-eGFP and mRFP-HFBI 5 at dpi (C, D). From left to right: combined eGFP and mRFP channels, only eGFP channel and only mRFP channel, respectively. Yellow color represents the presence of eGFP and mRFP.

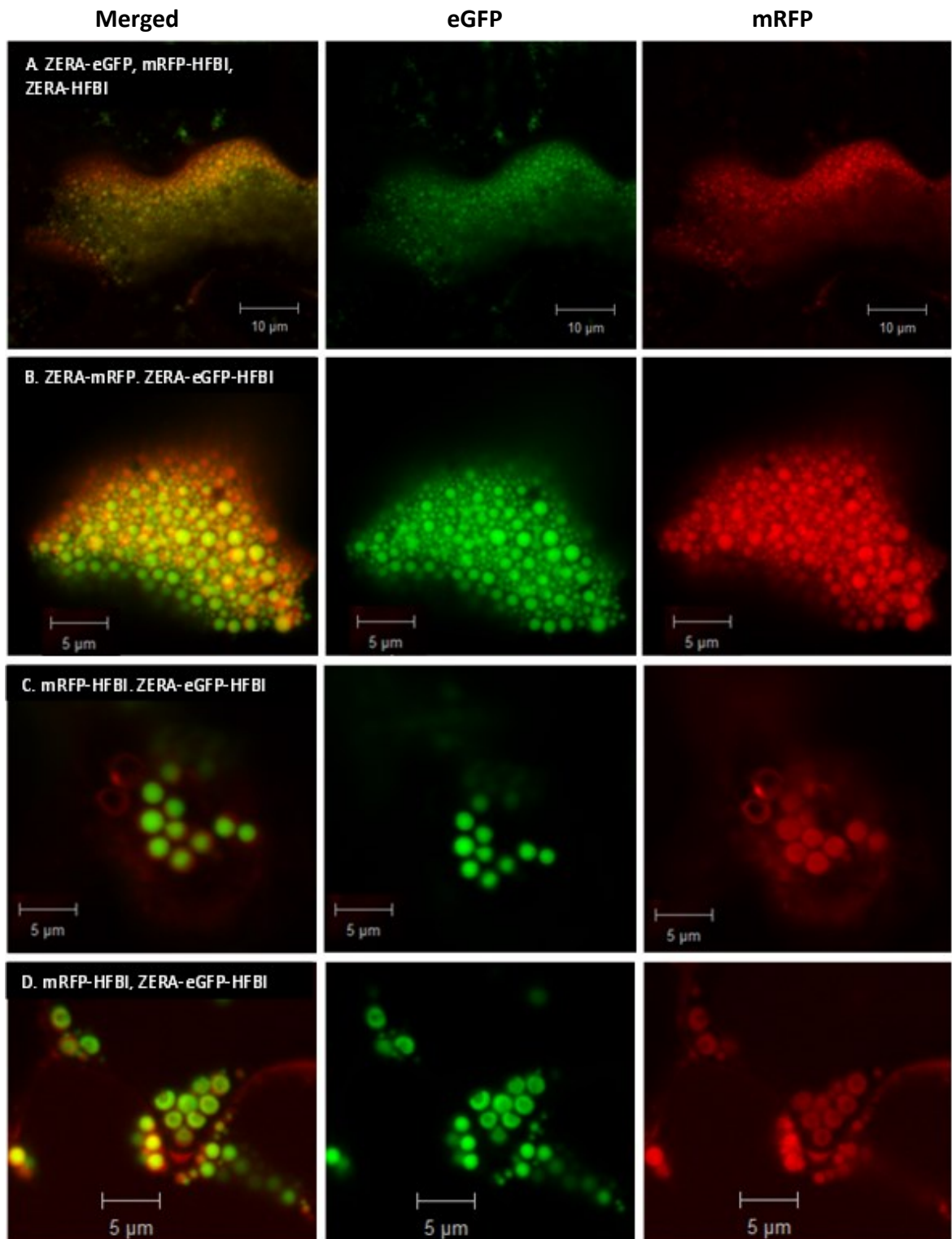


Figure 17. Confocal microscopy of co-expressions. A. ZERA-mRFP, ZERA-HFBI and eGFP-HFBI at 4 dpi. B. ZERA-mRFP and ZERA-eGFP-HFBI at 5 dpi. C-D. mRFP-HFBI and ZERA-eGFP-HFBI at 5 dpi: round PBs (C) and “hollow” PBs (D). Yellow color represents the presence of eGFP and mRFP.

4.4 Protein body isolation by subcellular fractionation

4.4.1 Homogenization of leaf tissue

Preliminary control experiment of the homogenization step was conducted to determine whether the PBs remained spherical during the homogenization. In case the membrane surrounding the PB is disrupted during the homogenization of the leaf sample, the soluble proteins within PB dissolve and remain in the supernatant, whereas intact PBs sediment to the insoluble pellet during the centrifugation.

In this experiment, the amounts of fusion protein in SDS-PAGE and western blot (Figure 18) were compared between supernatant (soluble proteins; N1L) and pellet (insoluble proteins; N1P) from centrifuged homogenized leaf tissue sample. The protein amount of these samples was compared to the total protein amount in unprocessed homogenized leaf tissue sample (total proteins; Ntot). Results are summarized in Figure 19.

In single expression samples of ZERA-mRFP and eGFP-HFBI, there were no difference between supernatant and pellet samples in the western blot probably due to signal saturation, and no conclusion can be made based on this result. The distribution of these fusion proteins was more clearly visible in SDS-PAGE, whereas eGFP-HFBI (expected MW 38 kDa) was present in equal amounts in supernatant and pellet samples. This indicates that part of the eGFP-HFBI PBs were disrupted, although some PBs remained intact. ZERA-eGFP (expected MW 40 kDa) was mainly present in the pellet sample, suggesting high stability of ZERA-eGFP PBs during the homogenization.

In the co-expression experiment of ZERA-mRFP and eGFP-HFBI, eGFP-HFBI seemed to be present in the pellet and supernatant based on the western blot. However, the amount of eGFP-HFBI in homogenized leaf sample (Ntot) was low, and therefore the presence of eGFP-HFBI in pellet remains somewhat unclear. Based on this, perhaps some eGFP-HFBI PBs remained intact during the homogenization. ZERA-mRFP (expected MW 39 kDa) seems to be equally present in pellet and supernatant in the western blot, indicating that

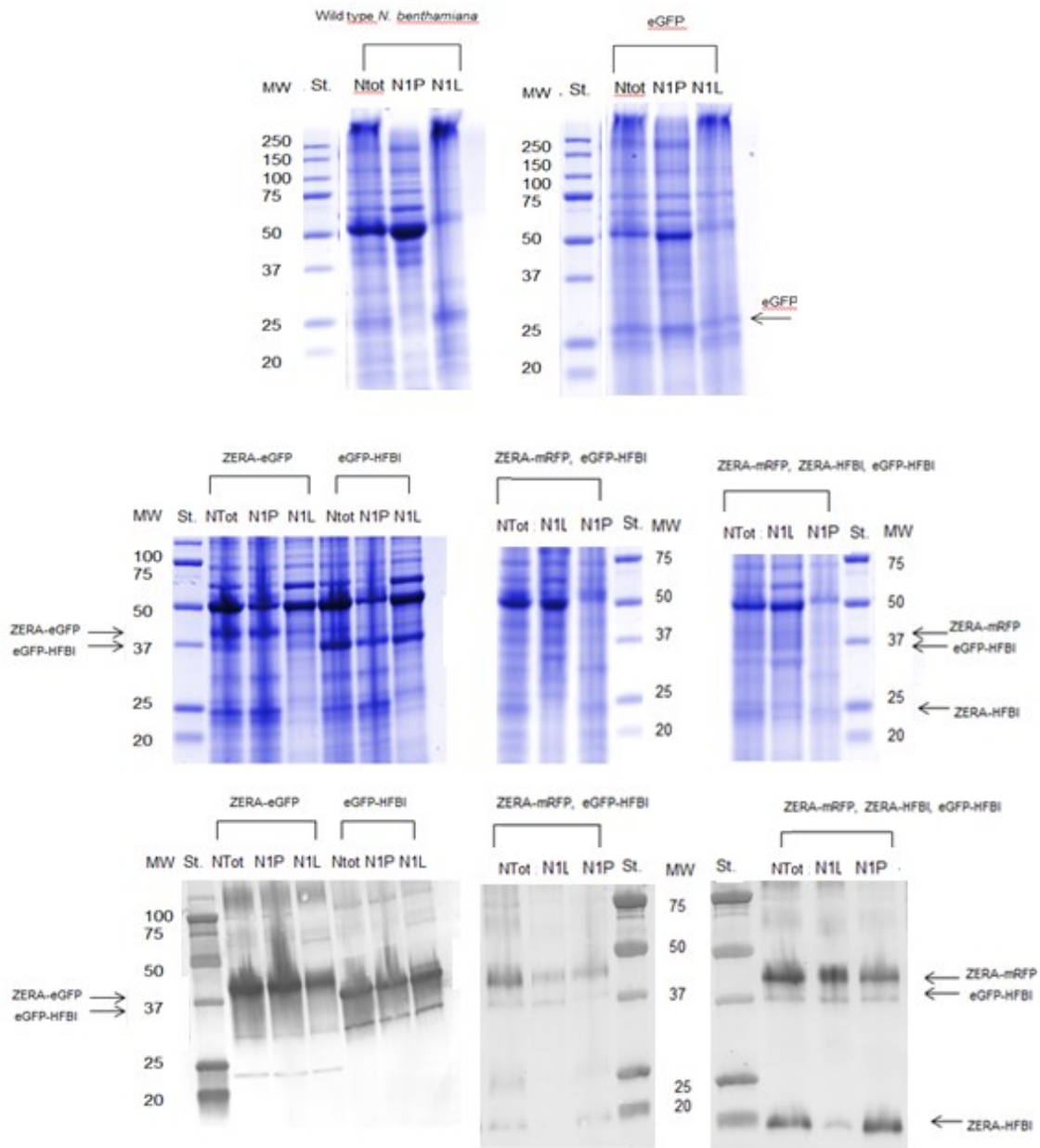


Figure 18. Total protein stained SDS-PAGE gel and western blot of solubility samples: homogenized leaf tissue (Ntot), pellet after centrifugation of homogenized leaf tissue (N1P), supernatant after centrifugation of homogenized leaf tissue (N1L), protein standard (St) and corresponding molecular weights (MW) in kDa.

perhaps half of ZERA including PBs remained undisrupted. In SDS-PAGE, the protein bands of eGFP-HFBI and ZERA-mRFP were not visible when compared to wild type *N. benthamiana* samples.

The protein bands of ZERA-mRFP, eGFP-HFBI and ZERA-HFBI were not visible on total protein stain of SDS-PAGE when compared to wild type *N. benthamiana* samples. In western blot, ZERA-mRFP and eGFP-HFBI were equally present in supernatant and pellet, and ZERA-HFBI (expected MW 21 kDa) was present mainly in the pellet. This result indicates that some PBs remained intact during the homogenization procedure. These PBs include ZERA-HFBI, ZERA-mRFP or eGFP-HFBI, or combinations of these proteins.

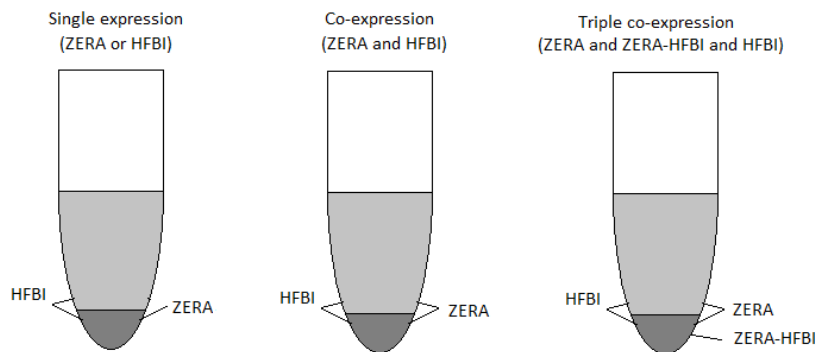


Figure 19. Summary of the presence of HFBI and ZERA in the solubility experiment. Dark and light grey colors represent pellet and supernatant fractions in the test tube, respectively.

4.4.2 Subcellular fractionation

Subcellular fractionation experiments were conducted to see whether PBs can be isolated from *N. benthamiana* leaves and whether PBs remain spherical after fractionation. The fractionation was done from single expression of eGFP-HFBI (5 dpi) and ZERA-eGFP (5 dpi), co-expression of ZERA-mRFP and eGFP-HFBI (7 dpi), and triple co-expression of ZERA-mRFP, eGFP-HFBI and ZERA-HFBI (7 dpi). eGFP (8 dpi) and wild type *N. benthamiana* (8 dpi) samples were fractionated for comparison. In the case of co-expressions, the interest was to see whether the PBs include one or more fluorescent proteins within the same PB

after fractionation. Fractionation samples were analyzed by SDS-PAGE and western blot to preliminary screen which fractions included expressed proteins (Figure 20), and then observed by confocal microscope (Figure 21).

By confocal microscopy, it was confirmed that there were no fluorescent particles in any fraction of wild type *N. benthamiana*. In the case of eGFP, low amounts of green fluorescent aggregates were observed by confocal microscopy in all fractions, except the supernatant fraction. The highest amount was observed in fractions I2 and I3, as it was also detected in SDS-PAGE samples in which eGFP (expected MW 27 kDa) seems to be present at 25 kDa.

In total protein staining of SDS-PAGE gel and western blot of eGFP-HFBI samples, eGFP-HFBI (expected MW 38 kDa) was present in fractions I2-I3 and pellet fraction with highest amounts of eGFP-HFBI in I3 and pellet fractions. In confocal microscopy, spherical eGFP-HFBI PBs were indeed observed. They were mainly in fraction I3. There were also some small fluorescent particles in the pellet fraction. After concentration of fractions, small fluorescent particles were observed also from fractions I1 and I2. Fractions S and I3 could not be concentrated, since there was no visible pellet after centrifugation.

In total protein staining of SDS-PAGE gel and western blot of ZERA-eGFP samples, ZERA-eGFP (expected MW 45) was present in fractions I2-I3 and pellet fraction. Similarly to eGFP-HFBI, clear spherical ZERA-eGFP PBs were observed in fraction I3. In pellet, PBs were observed with other fluorescent particles. After concentration, the highest amount of spherical PBs was observed in fraction I3 and some spherical PBs were observed also in concentrated fraction I2.

Co-expressed ZERA-mRFP and eGFP-HFBI sample was fractioned at 7 dpi. In western blot, ZERA-mRFP and eGFP-HFBI seemed to be present in almost all fractions: S, I2, I3 and pellet fraction, however, ZERA-mRFP and eGFP-HFBI could not be distinguished from total protein staining of SDS-PAGE. In confocal microscopy, separate red and green fluorescent PBs were observed in fractions I2 and I3. This result confirmed the previous co-expression

experiment in which ZERA-mRFP and eGFP-HFBI were accumulated in separate PBs. Also in the pellet fraction, there were a lot of red and green fluorescent particles that were probably cell debris with disrupted PBs. After sample concentration, again red and green fluorescent PBs were observed in fraction I2, and also in small amount in fraction I3.

Co-expressed ZERA-mRFP, ZERA-HFBI and eGFP-HFBI sample was fractioned at 7 dpi. In western blot, ZERA-mRFP, eGFP-HFBI and ZERA-HFBI were present in I3 and pellet fractions. The amount of ZERA-HFBI was low in these fractions. From the total protein stained SDS-PAGE gel, the fusion proteins could not be determined. In confocal microscopy, there were some red and green fluorescent spherical PBs in fraction I3. Some of spherical PBs seemed to include both fluorescent proteins, indicating the co-localization of ZERA-mRFP and eGFP-HFBI. There were not PBs in fractions S, I1 and I2. In the pellet, there were a lot of red and green particles. In the concentrated samples, there were PBs in fractions I2 and I3. Also in these samples, part of the PBs seemed to include red and green fluorescent proteins, similarly as in non-concentrated sample. This result confirmed the observation of co-localization of ZERA-mRFP and eGFP-HFBI in triple co-expression experiments. Again, no PBs were observed in fraction I1, and in the pellet fraction, there was lot of red and green fluorescent aggregates. Supernatant fraction could not be concentrated.

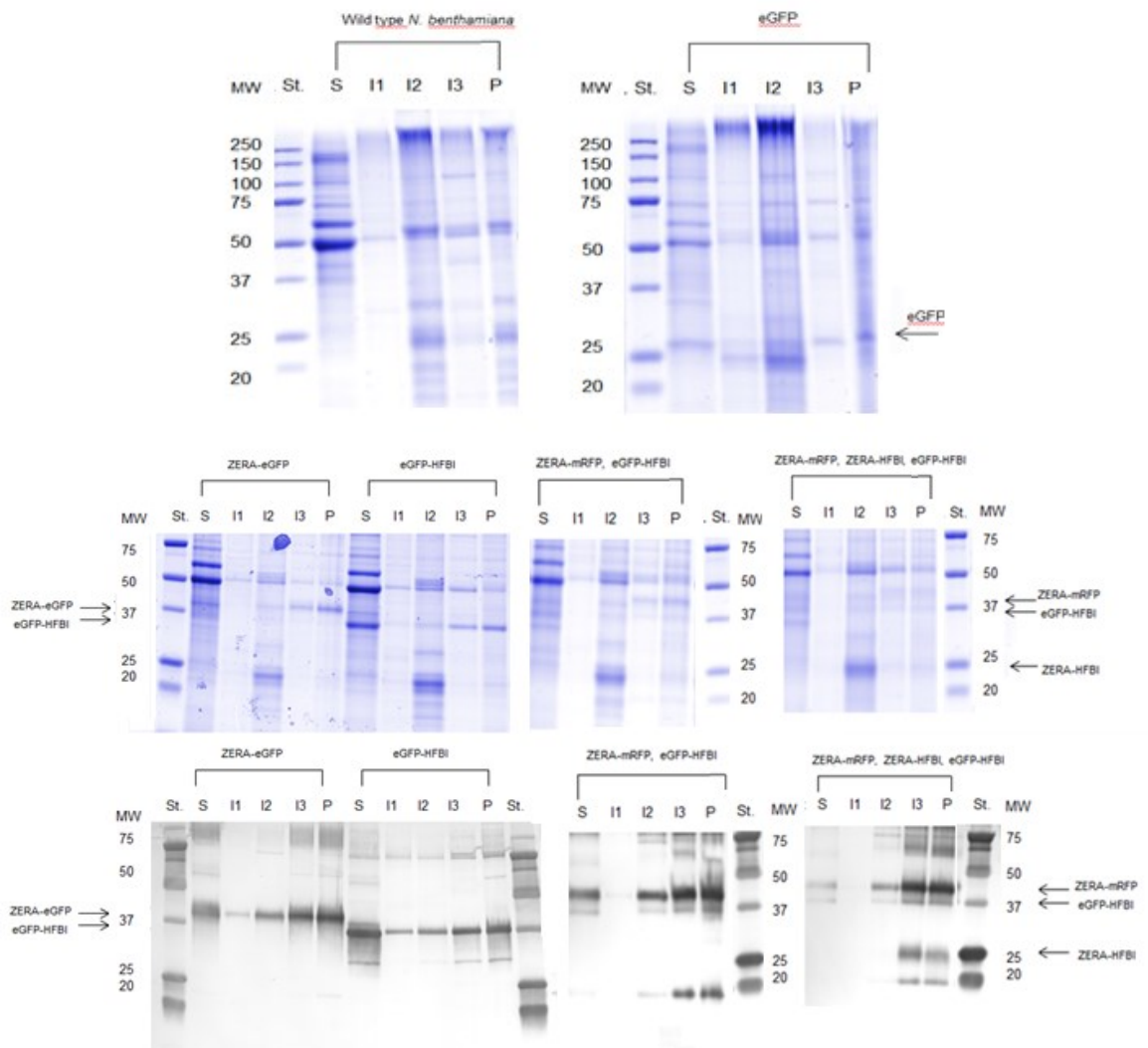


Figure 20. Total protein stained SDS-PAGE gel and western blot of sucrose gradient samples: supernatant (S) fraction, fractions I1, I2, I3 and pellet (P) fraction. Protein standard (St) and corresponding molecular weights (MW) in kDa.

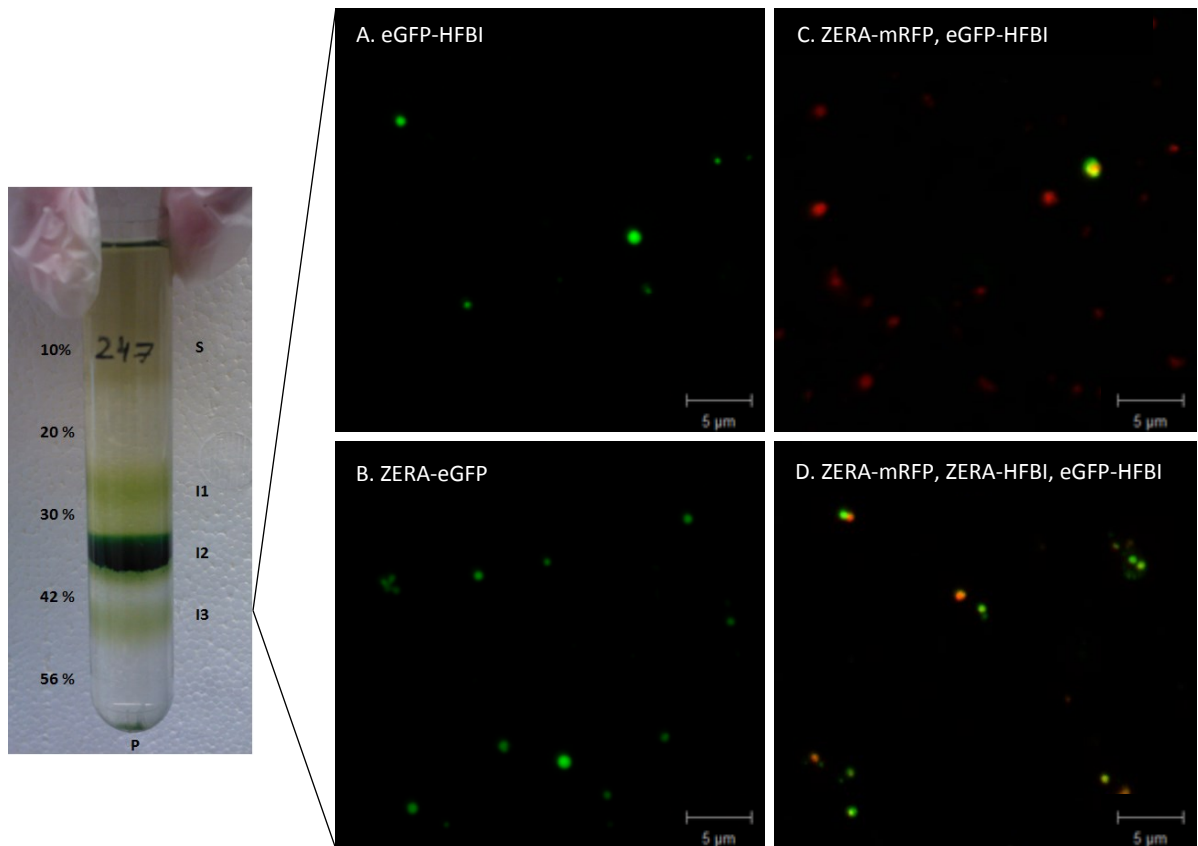


Figure 21. Sucrose gradient and confocal microscopy of fraction I3 after fractionations. A. eGFP-HFBI fraction I3 before concentration, B. ZERA-eGFP fraction I3 before concentration, C. co-expressed ZERA-mRFP and eGFP-HFBI fraction I3 before concentration, D. co-expressed ZERA-mRFP, ZERA-HFBI and eGFP-HFBI fraction I3 after concentration.

4.4.3 Electron microscopy

Isolated protein bodies were studied by electron microscopy to have a better understanding of PB structure. Figure 22 shows the structures of ZERA-eGFP and eGFP-HFBI PBs visualized by electron microscopy. The structure of both PB was indeed spherical, and interestingly, the tomographic reconstruction from transmission electron microscopy data revealed that the PBs were porous. ZERA-eGFP PB had round or oval cavities in the center of PB (Figure 22B). Also eGFP-HFBI PB had cavities within the PB (Figure 22D), however, those were not as clearly defined compartments as in ZERA-eGFP PB. In eGFP-HFBI PB, the pores were smaller than in ZERA-eGFP PB.

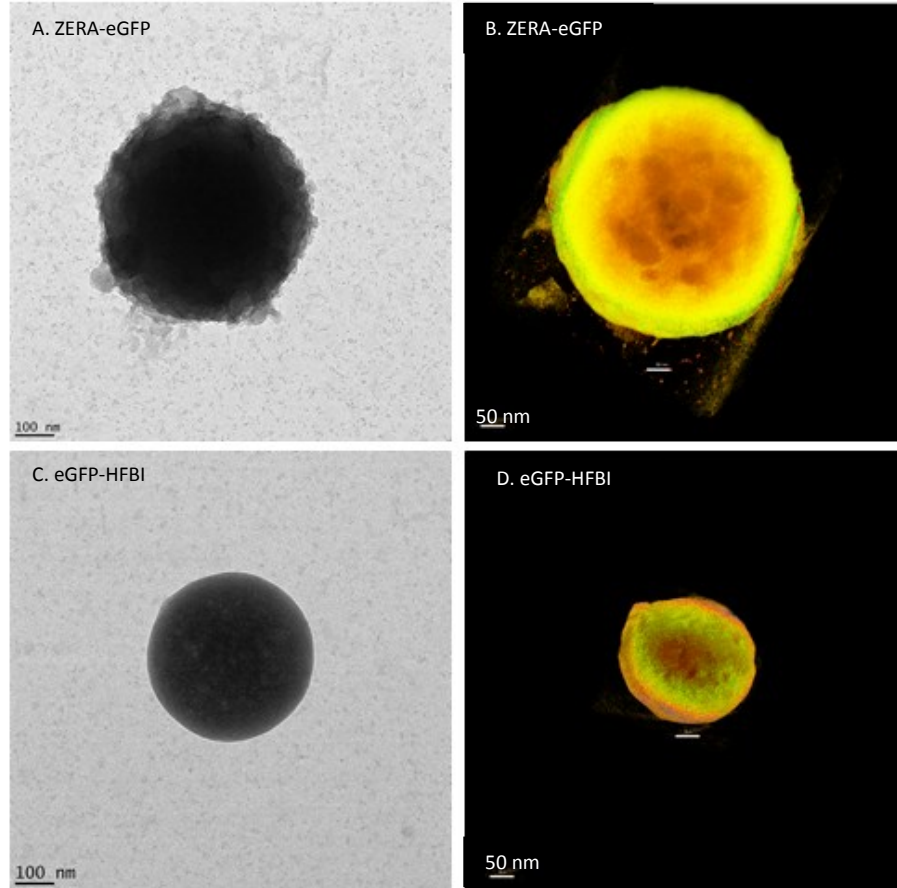


Figure 22. The structure of protein bodies visualized by electron microscopy. A. ZERA-eGFP PB, B. tomographic reconstruction of ZERA-eGFP PB, C. eGFP-HFBI PB, D. tomographic reconstruction of eGFP-HFBI PB.

5 Discussion

5.1 Mechanism of PB assembly

5.1.1 Expression levels and PB formation of constructs

In *N. benthamiana*, the TSP is reported to be approx. 6-8 mg/g of leaf FW (Gleba *et al.* 2013). By using an approximation of 7 mg/g FW of TSP in *N. benthamiana*, the expression levels within this study (mg/g FW) resulted in approx. 2, 3, 6 and 23 % of TSP for ZERA-eGFP-HFBI, eGFP-HFBI, ZERA-mRFP and ZERA-eGFP, respectively. Each of these constructs showed PB induction in confocal microscopy studies. Saberianfar *et al.* (2015) and Gutierrez *et al.* (2013) suggest that PB assembly depends on the recombinant protein concentration in the cell, the threshold value being 0,2 % of TSP for eGFP-HFBI. Here, the observed expression levels were at least one magnitude higher than the reported threshold value for PB induction, thus showing supporting evidence for the concentration-dependent PB induction mechanism.

Gutierrez *et al.* (2013) have presented confocal microscopy pictures of eGFP-HFBI that have fusion protein concentration of 1,17 % and 1,68 % of TSP in the cell. PBs seemed to be more abundant in the confocal microscopy pictures of eGFP-HFBI strain in this study (Figure 13D) than in the previously mentioned eGFP-HFBI pictures in the research by Gutierrez *et al.* (2013) (all pictures were presented in the same scale). In this study, the approximated protein concentration for eGFP-HFBI was 3 % of TSP.

Here, the expression level could not be quantified for mRFP-HFBI construct and only few minuscule mRFP-HFBI PBs were observed while otherwise the protein was mainly present in the ER membrane (Figure 13E). There are clearly more eGFP-HFBI PBs in confocal microscopy picture in which the protein concentration was 0,39 % TSP in the research of Gutierrez *et al.* (2013) than in Figure 13E of this study (both pictures had same scale). Therefore it is probable that the expression level of mRFP-HFBI construct was less than 0,4

% TSP, which is relatively near the threshold value 0,2 % TSP of eGFP-HFBI PB induction. This would explain the observation that mRFP-HFBI PBs were almost non-detectable in expression experiments.

5.1.2 Interactions between ZERA and HFBI

Based on the literature review, the following hypothesis of ZERA and HFBI interactions were formed:

- I. ZERA and HFBI interaction: ZERA and HFBI do not co-localize in the same PBs, thus ZERA and HFBI PB assembly is driven by different intermolecular interactions
- II. ZERA—ZERA interaction: ZERA-eGFP (or ZERA-mRFP) and ZERA-HFBI do not form separate PBs, thus the ZERA—ZERA interaction is present in ZERA-HFBI PB assembly
- III. HFBI—HFBI interaction: eGFP-HFBI (or mRFP-HFBI) and ZERA-HFBI do not form separate PBs, thus the HFBI—HFBI interaction is present in ZERA-HFBI PB assembly

The results indicate that ZERA and HFBI do not interact with each other in order to form PBs, providing evidence for hypothesis I. ZERA-mRFP and eGFP-HFBI were observed to accumulate in separate PBs within the same plant cell (Figure 16A). Also Januslike bipolar particles were observed (Figure 16B). In these PBs, ZERA and HFBI seemed to form two domains within a spherical PB. However, it is not certain whether these domains were separated by ER membrane. In the fractionation of ZERA-mRFP and eGFP-HFBI co-expression sample, separate ZERA-mRFP and eGFP-HFBI PBs were observed after PB isolation of the co-expression leaf sample (Figure 21C). In the same fractionation sample, also some PBs seemed to include both fusion proteins in separate sections. However, it was not possible to confirm whether these PBs were Januslike particles. When the fluorescent tags were exchanged in the co-expression experiment, mRFP-HFBI was visible in the ER membrane surrounding ZERA-eGFP PBs and proteins were not observed within

the same PB (Figure 16 C,D). The reason for this was probably the low expression level of mRFP-HFBI construct compared to ZERA-eGFP (Figure 14). However, it is not certain whether ZERA-eGFP and mRFP-HFBI would co-localize in the same PB if the expression level of mRFP-HFBI would be higher. Regardless of this, other before mentioned observations indicate that ZERA and HFBI do not directly interact with each other in PB assembly.

The results also indicate that a combining factor is needed in order to acquire PB assembly that includes ZERA and HFBI, providing evidence for hypothesis II and III. When ZERA-mRFP, eGFP-HFBI and ZERA-HFBI were co-expressed within the same cell, ZERA-mRFP and eGFP-HFBI were observed to co-localize within the same PB assembly (Figure 17A). In the fractionation experiment, sample from this co-expression seemed to include ZERA-mRFP and eGFP-HFBI proteins within the same PB (Figure 21D), which confirmed the confocal microscopy observation. Additional proof for these hypotheses was obtained from the co-expression experiment of ZERA-mRFP and ZERA-eGFP-HFBI. These experiments showed that ZERA-mRFP and ZERA-eGFP-HFBI were co-localized within the same PB (Figure 17B). Similarly, mRFP-HFBI and ZERA-eGFP-HFBI co-localized within the same PB (Figure 17 C,D). These observations suggest that the PB assembly of ZERA-mRFP and eGFP-HFBI was truly mediated by ZERA-HFBI in the co-expression experiment of these three proteins. Based on the results of co-expression and fractionation experiments (summarized in Figure 23), it seems that the interaction between ZERA and HFBI is weaker than interaction between ZERA and ZERA or between HFBI and HFBI.

The presented results support the observation (Dr. Reza Saberianfar, Agriculture and Agrifood Canada; unpublished results) that HFBI and ZERA do not co-localize within same PB assembly when these proteins are expressed within the same cell. Both of these proteins have hydrophobic parts in protein sequence: amphipathic polyproline II sequence in ZERA and the hydrophobic patch in HFBI.

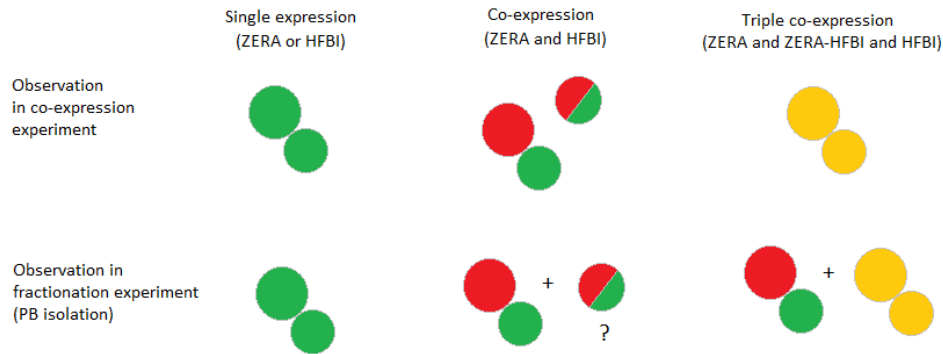


Figure 23. Summary of results in co-expression experiments and leaf tissue extract fractionation (PB isolation).

However, based on these results, the hydrophobic interaction between ZERA and HFBI is not as strong as ZERA-ZERA or HFBI-HFBI hydrophobic interaction, and cannot induce the PB assembly. Therefore, it seems that the PB assembly of these proteins has different mechanisms driven by interactions that are specific for these protein sequences.

5.2 Isolation and characteristics of PBs

The PB isolation procedure was proved to be suitable for ZERA, HFBI and ZERA-HFBI PB isolation from leaf extract (Figure 21). Previously, PBS buffer was used unsuccessfully in isolation of eGFP-HFBI PBs at VTT Ltd (Dr. Jussi Joensuu, unpublished results). In this isolation experiment, 10 % (w/w) sucrose was used in homogenization buffer, and probably osmotic pressure prevented the disruption of eGFP-HFBI PBs during the homogenization of the leaf sample.

The experiments showed that ZERA and HFBI constructs formed spherical PB assemblies with a diameter of approx. 1 μm (Figure 13 A,B and Figure 13D, respectively), similar as reported earlier in the literature. The structure of ZERA and HFBI PBs were shown to be spherical and porous by electron microscopy, ZERA PB having larger cavities within the PB than HFBI PB (Figure 23). Interestingly, a novel ZERA-eGFP-HFBI construct differed from

these by assembling into clearly defined large spherical PB with diameter of approx. 2 μm (Figure 13F). However, the resolution in the confocal microscopy was not sufficient to determine how proteins with fluorescent tag were organized within the ZERA-HFBI or ZERA-eGFP-HFBI PBs in the co-expression experiments.

In addition, novel types of PBs were observed in the experiments. In the co-expression of ZERA and HFBI, Januslike particles were observed (Figure 16B). These PBs included separate halves of both proteins and had a diameter of approx. 2,5 μm . Unexpectedly, another novel type of PBs was observed in the co-expression experiment of mRFP-HFBI and ZERA-eGFP-HFBI: “hollow” PBs that seemed to have more dense and bright surface than the inner part of the PB, thus having a radial PB assembly (Figure 17D). The PB included both mRFP-HFBI and ZERA-eGFP-HFBI and had a diameter of approx. 1 μm . Perhaps, these radially assembled PBs could be “early” PB assemblies of ZERA-eGFP-HFBI and mRFP-HFBI as the result of low protein concentration in the cell. These types of PBs were not earlier reported in the literature.

5.3 Further research and improvements

5.3.1 Further research suggestions

To increase the knowledge of PB assembly mechanism and confirm the shown results, few further experiments could be conducted. To confirm that the PBs are surrounded by ER-membrane, a membrane dissolving experiment could be conducted by dissolving the membrane from isolated PBs with detergent and centrifuging the solution similarly as in fusion protein interaction experiment. The comparison of fusion protein distributions with and without detergent addition could confirm whether the membrane prevents the breakage of PBs and leaking of fusion proteins. In addition, by cloning of ZERA-mRFP-HFBI construct, the ZERA-eGFP-HFBI and mRFP-HFBI co-expression experiment could be repeated by having the fluorescent tags vice versa. The advantage is also that the expression level of eGFP-HFBI is relatively high compared to mRFP-HFBI. By this approach,

the relationship between HFBI—HFBI interaction and ZERA—HFBI interaction could be confirmed. This experiment could reveal more information about the effect of protein concentration in PB assembly.

For material application research purposes, isolation and electron microscopy imaging of ZERA-eGFP-HFBI PBs would be especially interesting in order to compare the structure with ZERA and HFBI PBs. Since HFBI PBs isolation was successful in this research, further research to utilize them as pure PBs for material purposes would be of great interest. For example, since HFBI coated drug nanoparticles have already been studied to utilize them as carriers for controlled drug release (Valo *et al.* 2013), a logical step would be studying the suitability of ZERA-HFBI PBs for this purpose. Inspired by the observation of radially assembled spherical PBs, it would be interesting to study the effect of protein concentration to the PB assembly: in case that the unequal expression levels of two or more PB inducing proteins (when co-expressed within the same cell) would result in different PB morphology, specific morphologies could be achieved by altering the concentration of single PB production. This could be useful tool for nanoparticle design and production.

5.3.2 Improvements for production and isolation of PBs

In the confocal microscopy, one major limitation was observed related to the transient transformation protocol. The plant cell infection by *A. tumefaciens* was unequal, which made microscopy difficult, since all the infected plant cells did not express the construct to the same level, or in the case of co-expression, both constructs. Since the similar transient transformation method was used also for expression level experiment, the unequal infection might have affected the result, regardless the experiment design that was used to minimize the error caused by the unequal cell infection, the effect of the leaf age and the agroinfiltrated tissue to the expression, and the variability between the plants. To overcome these limitations, a plant viral replicon based expression system

could be used in transient transformation protocol to acquire systematic infection of plant cells (Marillonnet *et al.* 2005).

The PB isolation procedure used in this study showed to be feasible only for laboratory-scale due to the ultracentrifugation step. More investigations should be made to find a suitable large-scale isolation procedure. The choice of upstream and downstream process methods depends greatly on production quantities, the end usage and the end product formula. Thus these should be taken into account in order to produce PBs more routinely. For example, by using viral replicons in T-DNA and agrospraying even 90 % of the leaves could be transfected producing the protein of interest even 50 % of TSP (Gleba *et al.* 2013). Combining the process with silage processing described by Hahn *et al.* (2012) could be advantageous for low-cost bulk production, such as simple biomaterials or plant-produced oral vaccinations. For example, the protein instability during storing is considered to be limitation when oral administered protein pharmaceuticals are produced in lettuce or alfalfa (Fischer *et al.* 2004). In this case, silage procedure and expression of pharmaceutical protein as fusion with PB inducing sequence could increase the protein stability during storing and also offer a solution for formulation. Fusion with PB inducing sequence would be also suitable solution, since PBs serve as protein storing compartments in seeds that nature has designed for storing proteins in an active form for long periods of time. In addition, regulatory issues related to PBs could be addressed, especially in the case of animal vaccinations, by using edible maize zein or ZERA as PB inducing sequence. Indeed, Alvarez *et al.* (2010) have produced pharmaceutical protein as fusion with ZERA sequence in alfalfa, which is considered suitable production host for animal oral vaccines (Fischer *et al.* 2004).

6 Conclusions

Multiple gene constructs of ZERA and HFBI, and ZERA-HFBI were cloned as fusion with green and red fluorescent proteins and the expression levels of constructs were determined. The interactions between ZERA and HFBI were studied by co-expressing these constructs in transiently transformed *N. benthamiana* tobacco plants and by observing leaf samples by confocal microscopy. Co-expression experiments supported the hypothesis that ZERA and HFBI have different mechanism for PB formation: ZERA and HFBI were not observed to self-assemble within the same PB when co-expressed in the same plant cell. The co-expression experiments also indicated that interactions between ZERA and HFBI are weaker than ZERA—ZERA interactions or HFBI—HFBI interactions: ZERA-HFBI fusion protein was observed to self-assemble within the same PB with ZERA or with HFBI when ZERA-HFBI was co-expressed with ZERA or with HFBI within the same plant cell, respectively. Additionally, comparison between expression levels and PB formation supported hypothesis that PB formation depends on the protein accumulation level in the cell.

Subcellular fractionation of leaf extract was used to isolate ZERA and HFBI PBs, showing that the used isolation procedure was suitable for HFBI PB isolation from *N. benthamiana* leaf extract. The electron microscopy of isolated PBs revealed that ZERA PBs and HFBI PBs have porous ultrastructure. By confocal microscopy, ZERA-eGFP-HFBI was observed to assemble into larger spherical PBs than ZERA and HFBI. However, the ultrastructure of ZERA-eGFP-HFBI could not be determined by confocal microscopy and it remained to be revealed later by electron microscopy. Interestingly, novel PB structures were observed in co-expression experiment, including spherical radially assembled PBs and bipolar Januslike PBs, in which the proteins were present in two separate domains within the PB. These findings reveal novel insights for PB assembly.

Literature cited

Alvarez, M., Topal, E., Martin, F., Cardineau, G., Higher accumulation of F1-V fusion recombinant protein in plants after induction of protein body formation, *Plant Mol Biol*, **72** (1-2) (2010) 75-89.

Anonymous, Protocol: Plasmid DNA purification using the QIAprep Spin Miniprep Kit and a microcentrifuge, in *QIAprep Miniprep Handbook for purification of molecular biology grade DNA*, ed. Anonymous, QIAGEN, 2006, pp. 22-23.

Anonymous, <http://www.zipsolutions.cat/index.php> 2.2.2016.

Arcalis, E., Ibl, V., Peters, J., Melnik, S., Stoger, E., The dynamic behavior of storage organelles in developing cereal seeds and its impact on the production of recombinant proteins, *Front Plant Sci*, **5** (2014) 118-128.

Baars, R., van Leeuwen, Y., Hendrix, Y., Velikov, K., Kegel, W., Philipse, A., Morphology-controlled functional colloids by heterocoagulation of zein and nanoparticles, *Colloids Surf A Physicochem Eng Asp*, **483** (2015) 209-215.

Bechtold, N., Pelletier, G., In planta *Agrobacterium*-mediated transformation of adult *Arabidopsis thaliana* plants by vacuum infiltration, *Methods Mol Biol*, **82** (1998) 259-266.

Brown, R., Lemmon, B., The developmental biology of cereal endosperm, in *Endosperm*, ed. Olsen, O-A., Springer Berlin Heidelberg 2007, pp. 1-20, ISBN 978-3-540-71234-3.

Chandler, D., Interfaces and the driving force of hydrophobic assembly, *Nature*, **437** (7059) (2005) 640-647.

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

De Marchis, F., Pompa, A., Mannucci, R., Morosinotto, T., Bellucci, M. A plant secretory signal peptide targets plastome-encoded recombinant proteins to the thylakoid membrane, *Plant Mol Biol*, **76** (3-5) (2011) 427-441.

De Virgilio, M., De Marchis, F., Bellucci, M., Mainieri, D., Rossi, M., Benvenuto, E., Arcioni, S., Vitale, A., The human immunodeficiency virus antigen Nef forms protein bodies in leaves of transgenic tobacco when fused to zeolin, *J Exp Bot*, **59** (10) (2008) 2815-2829.

Engler, C., Gruetzner, R., Kandzia, R., Marillonnet, S., Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes, *PLoS one*, **4** (5) (2009) e5553.

Engler, C., Kandzia, R. & Marillonnet, S., A one pot, one step, precision cloning method with high throughput capability, *PLoS one*, **3** (11) (2008) e3647.

Fischer, R., Stoger, E., Schillberg, S., Christou, P., Twyman, R., Plant-based production of biopharmaceuticals, *Curr Opin Plant Biol*, **7** (2) (2004) 152-158.

Geli, M., Torrent, M., Ludevid, D., Two Structural Domains Mediate Two Sequential Events in γ -Zein Targeting: Protein Endoplasmic Reticulum Retention and Protein Body Formation, *Plant Cell*, **6** (12) (1994) 1911-1922.

Gelvin, S., Agrobacterium-mediated plant transformation: the biology behind the "gene-jockeying" tool, *Microbiol Mol Biol Rev*, **67** (1) (2003) 16-37.

Gleba, Y., Tusé, D., Giritch, A., Plant viral vectors for delivery by *Agrobacterium*, in *Plant Viral Vectors*, ed. Palmer, K., Gleba, Y., Springer Berlin Heidelberg 2013, pp. 155-192, ISBN 978-3-642-40828-1.

Gleba, Y., Klimyuk, V., Marillonnet, S., Magniffection—a new platform for expressing recombinant vaccines in plants, *Vaccine*, **23** (17) (2005) 2042-2048.

Grimsley, N., Hohn, B., Hohn, T., Walden, R., "Agroinfection", an alternative route for viral infection of plants by using the Ti plasmid, *Proc Nat Acad Sci U S A*, **83** (10) (1986) 3282-3286.

Giritch, A., Symonenko, Y., Hahn, S., Tiede, D., Shvarts, A., Roemer, P., Gleba, Y., Process of Transfecting Plants, WO/2012/019660, 2012.

Guo, X., Yuan, L., Chen, H., Sato, S., Clemente, T., Holding, D, Nonredundant function of zeins and their correct stoichiometric ratio drive protein body formation in maize endosperm, *Plant Physiol*, **162** (3) (2013) 1359-1369.

Gutierrez, S., Saberianfar, R., Kohalmi, S., Menassa, R., Protein body formation in stable transgenic tobacco expressing elastin-like polypeptide and hydrophobin fusion proteins, *BMC biotechnol*, **3** (40) (2013).

Hahn, S., Giritch, A., Gleba, Y., Production, storage and use of cell wall-degrading enzymes, WO 2013056829A1, 2012.

Hakanpää, J., Szilvay, G.R., Kaljunen, H., Maksimainen, M., Linder, M., Rouvinen, J., Two crystal structures of *Trichoderma reesei* hydrophobin HFBI—the structure of a protein amphiphile with and without detergent interaction, *Protein Sci*, **15** (9) (2006) 2129-2140.

Hakanpää, J., Paananen, A., Askolin, S., Nakari-Setälä, T., Parkkinen, T., Penttilä, M., Linder, M., Rouvinen, J., Atomic resolution structure of the HFBI hydrophobin, a self-assembling amphiphile, *J Biol Chem*, **279** (1) (2004) 534-539.

Hofbauer, A., Peters, J., Arcalis, E., Rademacher, T., Lampel, J., Eudes, F., Vitale, A., Stoger, E., The induction of recombinant protein bodies in different subcellular compartments reveals a cryptic plastid-targeting signal in the 27-kDa γ -zein sequence, *Front Bioeng Biotech*, **2** (67) (2014).

Hood, E., Gelvin, S., Melchers, L., Hoekema, A., New *Agrobacterium* helper plasmids for gene transfer to plants, *Transgenic Res*, **2** (4) (1993) 208-218.

Hu, X., Cebe, P., Weiss, A., Omenetto, F., Kaplan, D., Protein-based composite materials, *Mater Today*, **15** (5) (2012) 208-215.

Ibl, V., Stoger, E. The formation, function and fate of protein storage compartments in seeds, *Protoplasma*, **249** (2) (2012) 379-392.

Ibl, V., Kapusi, E., Arcalis, E., Kawagoe, Y., Stoger, E., Fusion, rupture, and degeneration: the fate of in vivo-labelled PSVs in developing barley endosperm, *J Exp Bot*, **65** (12) (2014) 3249-3261.

Joensuu, J., Conley, A., Lienemann, M., Brandle, J., Linder, M., Menassa, R., Hydrophobin fusions for high-level transient protein expression and purification in *Nicotiana benthamiana*, *Plant Physiol*, **152** (2) (2010) 622-633.

Kawakatsu, T., Takaiwa, F., Cereal seed storage protein synthesis: fundamental processes for recombinant protein production in cereal grains, *Plant Biotechnol J*, **8** (9) (2010) 939-953.

Khalesi, M., Gebruers, K., Derdelinck, G., Recent advances in fungal hydrophobin towards using in industry, *Protein J*, **34** (4) (2015) 243-255.

Kim, C., Woo Ym, Y., Clore, A., Burnett, R., Carneiro, N., Larkins, B., Zein protein interactions, rather than the asymmetric distribution of zein mRNAs on endoplasmic reticulum membranes, influence protein body formation in maize endosperm, *Plant Cell*, **14** (3) (2002) 655-672.

Kisko, K., Szilvay, G., Vainio, U., Linder, M, Serimaa, R., Interactions of hydrophobin proteins in solution studied by small-angle X-ray scattering, *Biophys J*, **94** (1) (2008) 198-206.

Krenek, P., Samajova, O., Luptovciak, I., Daskocilova, A., Komis, G., Samaj, J., Transient plant transformation mediated by *Agrobacterium tumefaciens*: Principles, methods and applications, *Biotechnol Adv*, **33** (6) (2015) 1024-1042.

Lending, C., Larkins, B., Changes in the zein composition of protein bodies during maize endosperm development, *Plant Cell*, **1** (10) (1989) 1011-1023.

Lienemann, M., Gandier, J., Joensuu, J., Iwanaga, A., Takatsuji, Y., Haruyama, T., Master, E., Tenkanen, M., Linder, M., Structure-function relationships in hydrophobins: probing the role of charged side chains, *Appl Environ Microbiol*, **79** (18) (2013) 5533-5538.

Lin, Y., Ding, Z., Zhou, X., Li, S., Xie, D., Li, Z., Sun, G., In vitro and In vivo evaluation of the developed PLGA/HAp/Zein scaffolds for bone-cartilage interface regeneration, *Biomed Environ Sci*, **28** (1) (2015) 1-12.

Linder, M., Szilvay, G., Nakari-Setälä, T., Penttilä, M., Hydrophobins: the protein-amphiphiles of filamentous fungi, *FEMS Microbiol Rev*, **29** (5) (2005) 877-896.

Liu, X., Xie, Y., Li, W., Sheng, W., Li, Y., Tong, Z., Ni, H., Huselstein, C., Wang, X., Chen, Y., Structure, physical properties, hemocompatibility and cytocompatibility of starch/zein composites, *Biomed Mater Eng*, **25** (1) (2015) 47-55.

Llompart, B., Llop-Tous, I., Marzabal, P., Torrent, M., Pallissé, R., Bastida, M., Ludevid, M., Walas, F., Protein production from recombinant protein bodies, *Process Biochem*, **45** (11) (2010) 1816-1820.

Llop, I., Torrent, M., Marzabal, P., Bastida, M., Llompart, B., Ludevid, D., Zera[®], a novel technology for stable accumulation and easy recovery of recombinant proteins in eukaryotic protein-production hosts, *Microb Cell Fact*, **5** (S42) (2006).

Llop-Tous, I., Madurga, S., Giralt, E., Marzabal, P., Torrent, M., Ludevid, M., Relevant elements of a maize gamma-zein domain involved in protein body biogenesis, *J Biol Chem*, **285** (46) (2010) 35633-35644.

Mainieri, D., Rossi, M., Archinti, M., Bellucci, M., De Marchis, F., Vavassori, S., Pompa, A., Arcioni, S., Vitale, A., Zeolin. A new recombinant storage protein constructed using maize gamma-zein and bean phaseolin, *Plant Physiol*, **136** (3) (2004) 3447-3456.

Marillonnet, S., Thoeringer, C., Kandzia, R., Klimyuk, V., Gleba, Y., Systemic *Agrobacterium tumefaciens*–mediated transfection of viral replicons for efficient transient expression in plants, *Nat Biotechnol*, **23** (6) (2005) 718-723.

Meyer, D., Chilkoti, A., Purification of recombinant proteins by fusion with thermally-responsive polypeptides, *Nat Biotechnol*, **17** (11) (1999) 1112-1115.

Musiychuk, K., Stephenson, N., Bi, H., Farrance, C., Orozovic, G., Brodelius, M., Brodelius, P., Horsey, A., Ugulava, N., Shamloul, A., A launch vector for the production of vaccine antigens in plants, *Influenza Other Respir Viruses*, **1** (1) (2007) 19-25.

Niemeyer, C., Nanoparticles, proteins, and nucleic acids: biotechnology meets materials science, *Angew Chem Int Ed Engl*, **40** (22) (2001) 4128-4158.

Olsen, O., Nuclear endosperm development in cereals and *Arabidopsis thaliana*, *Plant Cell*, **16** (2004) S214-27.

Osborne, T. *The vegetable proteins*, Longmans, Green and Company, London 1916.

Patel, S., Cudney, R., McPherson, A., Crystallographic characterization and molecular symmetry of edestin, a legumin from hemp, *J Mol Biol*, **235** (1) (1994) 361-363.

Pompa, A., Vitale, A., Retention of a bean phaseolin/maize gamma-zein fusion in the endoplasmic reticulum depends on disulfide bond formation, *Plant Cell*, **18** (10) (2006) 2608-2621.

Reavy, B., Bagirova, S., Chichkova, N., Fedoseeva, S., Kim, S., Vartapetian, A., Taliansky, M., Caspase-resistant VirD2 protein provides enhanced gene delivery and expression in plants, *Plant Cell Rep*, **26** (8) (2007) 1215-1219.

Reuter, L., Bailey, M., Joensuu, J., Ritala, A., Scale-up of hydrophobin-assisted recombinant protein production in tobacco BY-2 suspension cells, *Plant Biotechnol J*, **12** (4) (2014) 402-410.

- Reyes, F., Chung, T., Holding, D., Jung, R., Vierstra, R., Otegui, M., Delivery of prolamins to the protein storage vacuole in maize aleurone cells, *Plant Cell*, **23** (2) (2011) 769-784.
- Saberianfar, R., Joensuu, J., Conley, A., Menassa, R., Protein body formation in leaves of *Nicotiana benthamiana*: a concentration-dependent mechanism influenced by the presence of fusion tags, *Plant Biotechnol J*, **13** (7) (2015) 927-937.
- Saito, Y., Kishida, K., Takata, K., Takahashi, H., Shimada, T., Tanaka, K., Morita, S., Satoh, S., Masumura, T., A green fluorescent protein fused to rice prolamin forms protein body-like structures in transgenic rice, *J Exp Bot*, **60** (2) (2009) 615-627.
- Saumonneau, A., Rottier, K., Conrad, U., Popineau, Y., Guéguen, J., Francin-Allami, M., Expression of a new chimeric protein with a highly repeated sequence in tobacco cells, *Plant Cell Rep*, **30** (7) (2011) 1289-1302.
- Schmidt, S., Protein bodies in nature and biotechnology, *Mol Biotechnol*, **54** (2) (2013) 257-268.
- Schöb, H., Kunz, C., Meins Jr, F., Silencing of transgenes introduced into leaves by agroinfiltration: a simple, rapid method for investigating sequence requirements for gene silencing, *Mol Gen Genet*, **256** (5) (1997) 581-585.
- Semenza, J., Hardwick, K., Dean, N., Pelham, H., ERD2, a yeast gene required for the receptor-mediated retrieval of luminal ER proteins from the secretory pathway, *Cell*, **61** (7) (1999) 1349-1357.
- Shakeri, A., Radmanesh, S., Preparation of cellulose nanofibrils by high-pressure homogenizer and zein composite films, *Adv Mat Res*, **15** (829) (2014) 534-538.
- Shewry, P., Halford, N., Cereal seed storage proteins: structures, properties and role in grain utilization, *J Exp Bot*, **53** (370) (2002) 947-958.
- Shewry, P., Napier, J., Tatham, A., Seed storage proteins: structures and biosynthesis, *Plant Cell*, **7** (7) (1995) 945-956.

Szilvay, G., Behavior of *Trichoderma reesei* hydrophobins in solution: Interactions, dynamics, and multimer formation, *Biochemistry*, **45** (28) (2006) 8590-8598.

Takatsuji, Y., Yamasaki, R., Iwanaga, A., Lienemann, M., Linder, M., Haruyama, T., Solid-support immobilization of a “swing” fusion protein for enhanced glucose oxidase catalytic activity, *Colloids Surf B Biointerfaces*, **112** (2013) 186-191.

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

Torrent, M., Llop-Tous, I., Ludevid, M., Protein body induction: a new tool to produce and recover recombinant proteins in plants, in *Recombinant Proteins from Plants*, ed. Faye, L., Gomord, V., Humana Press 2009b, pp. 193-208. ISBN 978-1-58829-978-9.

Urry, D., Luan, C., Parker, T., Gowda, D., Prasad, K., Reid, M., Safavy, A., Temperature of polypeptide inverse temperature transition depends on mean residue hydrophobicity, *J Am Chem Soc*, **113** (11) (1991) 4346-4348.

Valo, H., Laaksonen, P., Peltonen, L., Linder, M., Hirvonen, J., Laaksonen, T., Multifunctional hydrophobin: toward functional coatings for drug nanoparticles, *ACS Nano*, **4** (3) (2010) 1750-1758.

Valo, H., Kovalainen, M., Laaksonen, P., Häkkinen, M., Auriola, S., Peltonen, L., Linder, M., Järvinen, K., Hirvonen, J., Laaksonen, T., Immobilization of protein-coated drug nanoparticles in nanofibrillar cellulose matrices—enhanced stability and release, *J Control Release*, **156** (3) (2011) 390-397.

Valo, H., Arola, S., Laaksonen, P., Torkkeli, M., Peltonen, L., Linder, M., Serimaa, R., Kuga, S., Hirvonen, J., Laaksonen, T., Drug release from nanoparticles embedded in four different nanofibrillar cellulose aerogels, *Eur J Pharm Sci*, **50** (1) (2013) 69-77.

- Voinnet, O., Rivas, S., Mestre, P., Baulcombe, D., An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus, *Plant J*, **33** (5) (2003) 949-956.
- Waheed, M., Ismail, H., Gottschamel, J., Mirza, B., Lössl, A., Plastids: the green frontiers for vaccine production, *Front Plant Sci*, **6** (1005) (2015).
- Wakasa, Y., Takagi, H., Watanabe, N., Kitamura, N., Fujiwara, Y., Ogo, Y., Hayashi, S., Yang, L., Ohta, M., Thet Tin, W., Sekikawa, K., Takano, M., Ozawa, K., Hiroi, T., Takaiwa, F., Concentrated protein body product derived from rice endosperm as an oral tolerogen for allergen-specific immunotherapy--a new mucosal vaccine formulation against Japanese cedar pollen allergy, *PLoS one*, **10** (3) (2015) e0120209.
- Wang, F., Yang, C., Hu, X., Advanced Protein Composite Materials, In *Lightweight Materials from Biopolymers and Biofibers*, ed. Yang, Y., Xu, H., Yu, X., American Chemical Society, Washington DC. 2014. pp. 177-208, ISBN 9780841229907
- Weber, E., Engler, C., Gruetzner, R., Werner, S., Marillonnet, S., A modular cloning system for standardized assembly of multigene constructs, *PLoS one*, **6** (2) (2011) e16765.
- Wessels, J., Developmental regulation of fungal cell wall formation, *Annu Rev Phytopathol*, **32** (1) (1994) 413-437.
- Whitehead, M., Ohlschlager, P., Almajhdi, F., Alloza, L., Marzabal, P., Meyers, A., Hitzeroth, I., Rybicki, E., Human papillomavirus (HPV) type 16 E7 protein bodies cause tumour regression in mice, *BMC Cancer*, **14** (367) (2014).
- Willmitzer, L., De Beuckeleer, M., Lemmers, M., Van Montagu, M., Schell, J., DNA from Ti plasmid present in nucleus and absent from plastids of crown gall plant cells, *Nature*, **287** (1980) 359-361.
- Wösten, H., Scholtmeijer, K., Applications of hydrophobins: current state and perspectives, *Appl Microbiol Biotechnol*, **99** (4) (2015) 1587-1597.

Xiang, L., Etxeberria, E., den Ende, W., Vacuolar protein sorting mechanisms in plants, *FEBS J*, **280** (4) (2013) 979-993.

Xiong, F., Yu, X., Zhou, L., Wang, Z., Wang, F., Xiong, A., Structural development of aleurone and its function in common wheat, *Mol Biol Rep*, **40** (12) (2013) 6785-6792.

Xu, J., Dolan, M., Medrano, G., Cramer, C., Weathers, P., Green factory: plants as bioproduction platforms for recombinant proteins, *Biotechnol Adv*, **30** (5) (2012) 1171-1184.

Yadav, N., Postle, K., Saiki, R., Thomashow, M., Chilton, M., T-DNA of a crown gall teratoma is covalently joined to host plant DNA, *Nature*, **287** (1980) 458-461.

Zhang, C., Chang, N., Li, C., Li, X., Effects of zein on film-forming ability and properties of wheat gluten films, *Adv Mat Res*, **750** (2013) 1582-1585.

Zhang, Y., Cui, L., Che, X., Zhang, H., Shi, N., Li, C., Chen, Y., Kong, W., Zein-based films and their usage for controlled delivery: Origin, classes and current landscape, *J Control Release*, **206** (2015) 206-219.

Zheng, Y., Wang, Z., Protein accumulation in aleurone cells, sub-aleurone cells and the center starch endosperm of cereals, *Plant Cell Rep*, **33** (10) (2014) 1607-1615.