Expression of a new chimeric protein with a highly repeated sequence in tobacco cells

Amélie Saumonneau, Karine Rottier, Udo Conrad, Yves Popineau, Jacques J. Guéguen, Mathilde Francin-Allami

EXPRESSION OF A NEW CHIMERIC PROTEIN WITH A HIGHLY REPEATED SEQUENCE IN TOBACCO CELLS

Amélie Saumonneau[1], Karine Rottier[1], Udo Conrad[2], Yves Popineau[1], Jacques Guéguen[1*] and Mathilde Francin-Allami[1*]

[1] INRA, UR1268, Biopolymères Interactions Assemblages, 44300 Nantes, France

[2] Leibniz Institute of Plant Genetics & Crop Plant Research, IPK, 06466 Gatersleben, Germany

*Corresponding authors

E-mail: allami@nantes.inra.fr; gueguen@nantes.inra.fr

Abstract

In wheat, the high molecular weight (HMW) glutenin subunits are known to contribute to gluten viscoelasticity and show some similarities to elastomeric animal proteins as elastin. Combining the sequence of a glutenin with that of elastin is a way to create new chimeric functional proteins, which could be expressed in plants. The sequence of a glutenin subunit was modified by the insertion of several hydrophobic and elastic motifs derived from elastin (elastin-like peptide, ELP) into the hydrophilic repetitive domain of the glutenin subunit to create a triblock protein, the objective being to improve the mechanical (elastomeric) properties of this wheat storage protein.

In this study, we investigated an expression model system to analyze the expression and trafficking of the wild type HMW glutenin subunit (GS_W) and an HMW glutenin subunit mutated by the insertion of elastin motifs (GS_M -ELP). For this purpose, a series of constructs was made to express wild-type subunits and subunits mutated by insertion of elastin motifs in fusion with green fluorescent protein (GFP) in tobacco BY-2 cells. Our results showed for the first time the expression of HMW glutenin fused with GFP in tobacco protoplasts. We also expressed and localized the chimeric protein composed of plant glutenin and animal elastin-like peptides (ELP) in BY-2 protoplasts, and demonstrated its presence in protein body-like structures in the endoplasmic reticulum. This work therefore provides a basis for heterologous production of the glutenin-ELP triblock protein to characterize its mechanical properties.

Keywords: glutenin . elastin . biopolymer . triblock protein . endoplasmic reticulum . BY-2 cells.

Abbreviations

BiP	Binding Protein
BY-2	Bright Yellow-2
ELP	Elastin-Like Peptide
ER	Endoplasmic Reticulum
GFP	Green Fluorescent Protein
GS_W	Wild type Glutenin Subunit
GS_M	Modified Glutenin Subunit
GS_M-ELP	Modified Glutenin Subunit-Elastin Like Peptide
HMW	High Molecular Weight

PB	Protein Body
PDI	Protein Disulphide Isomerase

Introduction

Agricultural products, and particularly plant proteins, are candidates for the preparation of renewable and biodegradable materials. Among these, wheat gluten presents the advantage of being a commercially available raw material. The properties of wheat gluten films have been widely studied and the potential of these proteins is now well defined (Popineau et al. 2001; Mangavel et al. 2002). Wheat gluten proteins are prolamin storage proteins, comprised of monomeric gliadins and polymeric glutenins, synthesized and stored in the endosperm cells of the seeds. They exhibit unusual viscoelastic properties when hydrated and are largely responsible for the rheological behaviours of wheat flour dough (Shewry et al. 2002; Shewry and Halford 2002).

Wheat gluten proteins are typical secretory proteins, and processing within the endoplasmic reticulum (ER) would include folding and intra-chain disulphide bond formation, as well as formation of the inter-chain bonds that stabilize the glutenin polymers (Shewry et al. 2003). Non-covalent interactions (notably hydrogen bonds) between gliadins and glutenin polymers may also be established, resulting in protein precipitation and the formation of hydrated protein particles within the secretory pathway. It is probable that the folding and assembly of the gluten proteins is assisted by ER lumenal proteins such as the enzyme protein disulphide isomerase (PDI) and the molecular chaperone binding protein (BiP) although this is still not conclusively established (Grimwade et al. 1996; DuPont et al. 1998). It has been reported that some prolamins accumulate in the ER before being deposited into protein bodies (PBs) (Galili et al. 1993). These PBs are derived from the ER and it has been suggested that they are transported directly to the storage vacuole bypassing the Golgi (Levanony et al. 1992; Galili et al. 1993; Vitale and Galili 2001). However, other reports show the trafficking of some prolamins *via* the Golgi particularly during the early stages of seed development (Parker and Hawes 1982; Shy et al. 2001; Loussert et al. 2008; Tosi et al. 2009). In mature seeds, the prolamins form a matrix resulting from the disappearance of the PB structure (Rubin et al. 1992).

The glutenin proteins are elastomeric, similar to silk protein, abductin or elastin protein. In spite of their different amino acid sequences, all of these proteins present common characteristics such as short repeated sequence conferring elasticity and the ability to form stable intermolecular links (Tatham and Shewry 2000). Our experimental approach exploits a sequence of the animal elastic protein, elastin. This protein exhibits mechanical properties that

are already used in the polymer industry with applications in biomedicine and nanobiotechnology (Rodriguez-Cabello et al. 2009, Arias et al. 2006). As an additional beneficial effect, elastin-like peptide (ELP) fusions have also been shown to significantly enhance the accumulation of a range of different recombinant proteins in transgenic tobacco leaves (Patel et al. 2007; Floss et al. 2008) and seeds (Scheller et al. 2006). Although it is thought that ELP tags confer increased stability or solubility to their fusion partner, the means by which ELP increases the production yield of recombinant proteins *in planta* has not yet been established. The ability of ELP fusions to induce PB formation in plant leaves was recently shown (Conley et al. 2009).

The two proteins, elastin and glutenin, have similar mechanical properties, but differ in their molecular characteristics. Both proteins are proline-rich, with the ability to self-assemble and form supramolecular secondary structures consisting of helices and spirals as a result of their highly repetitive sequences (Shewry et al. 2002; Miao et al. 2003). The introduction of several elastic hydrophobic motifs of elastin into glutenin could modify gluten viscoelasticity by the formation of a triblock protein consisting of one hydrophobic elastin block into a hydrophilic domain. The structure of this triblock protein could result in a microstructure of separated phases where the hydrophobic parts could bind together separated by hydrophilic domains. Such a triblock protein could improve the water sensitivity and mechanical properties by the formation of elastic hydrophobic microdomains (Wellner et al. 2006; Wright and Conticello 2002) . This macrostructure type has been previously considered for elastin-mimetic triblock polymers (Wright and Conticello 2002). Triblock copolymers are already used as an interface agent to allow polymer mixing and to obtain stable composite materials (Kumar et al. 2006).

The aim of this work is to study the expression of a new chimeric triblock protein based on a HMW glutenin subunit and ELP, in fusion with green fluorescent protein (GFP) in BY-2 suspension-cultured tobacco cells. Previous studies demonstrated the possibility to express glutenins of high or low molecular weight (Shani et al. 1994; Robert et al. 1989; Shimoni et al. 1997; Lombardi et al. 2009) and elastin-like motifs (Patel et al. 2007; Floss et al. 2008; Conley et al. 2009) in tobacco plants. Our results demonstrated the ability to express and identify the subcellular localization of the new chimeric protein in fusion with GFP in a tobacco cell system.

Materials and methods

Plant Material

Tobacco *Nicotiana tabacum L. cv Bright Yellow* (BY-2) cells was maintained at 24°C on an orbital shaker (120 rpm) in Linsmaier and Skoog modified medium (Nagata et al. 1992) in the dark. This culture was maintained by weekly dilution of the cells by the medium (0.5:30 [v/v]).

Cloning of 1Dy10 cDNA

The cloning of 1Dy10 cDNA from a mRNA library obtained from endosperm of wheat grain cultivar Recital aged of 313°D was achieved by PCR with 1Dy10 specific primers (5'-AAGTCGACCACCCACAACACCGAGCACC-3'; 5'-AACCCGGGAGTTCTATCACTGG CTAGCC-3') carrying the restriction sites SalI and SmaI, respectively.

GS_W-GFP, GS_M-GFP, GS_M-ELP-GFP fused genetic constructs

To study the subcellular compartmentalization of 1Dy10 glutenin subunit (GS) fused proteins, three different constructs based on a GFP fusion were used. The pRTd35S-GFP vector was designed especially as a reporter gene with enhanced green fluorescence. Expression of the *GFP* gene is driven by a double 35S Cauliflower mosaic virus (CaMV) promoter, and GFP protein synthesis is induced by a viral translation enhancer. In order to introduce the GS cDNA in GFP upstream, the SalI and SmaI restriction sites are introduced in pRTd35S-GFP between GFP open reading frame and the viral translation enhancer sequence by PCR using primers 5'-AAGGAATCCAGAGTCGACACTCCCGGGTCCATGGGTAAAGGAGAAGA ACT-3' and 5'-ACTCTAGAGGATCCTTAGAGATCTAGTTCA-3'.

The first construct corresponded to the GS_W-GFP fusion, for which the complete cDNA of 1Dy10 encoded for a wild-type glutenin subunit (GS_W) was used. To produce the GS_W-GFP fusion protein, the GS cDNA without a STOP codon was amplified by PCR with the same forward primer previously described and the reverse primer 5'-AATCCCGGGCTGGCTAGCCG ACAATGCGTC-3' carrying the SmaI restriction site. The PCR product corresponding to GS without a STOP codon was cloned as a SalI-SmaI fragment

in the new SalI-SmaI sites of pRTd35S-GFP vector under the control of the CaMV 35S promoter and terminator.

For the second construct, the GS cDNA without a STOP codon was mutated by directed mutagenesis with the QuickChange® Site-Directed Mutagenesis Kit (Clontech) in order to introduce a BglII restriction site. The mutagenesis PCR was performed with the forward primer 5'-CCAAGACAATGGCAAAGATCTGGACAAGGGCAACAAG-3' and the reverse primer 5'-CTTGTTGCCCTTGTCCAGATCTTTGCCATTGTCTTGG-3', allowing the introduction of the BglII restriction site (underlined sequence) into the repeated domain of GS in order to obtain the GS mutated fragment corresponding to the GS_M-GFP fused protein.

For the third construct, the GS mutated fragment was used to introduce a fragment of 630pb encoding the 40xELP sequence. The ELP nucleotide sequence used in this work encodes the VPGVG pentapeptide (Scheller et al. 2004). It was provided into the pRTRA-40xELP plasmid containing two BglII sites, which are used for 40xELP sequence insertion into the modified glutenin subunit sequence, the GS_M-*GFP* construct. The correct sequence of constructs GS_W-GFP, GS_M-GFP, GS_M-ELP-GFP and the open reading frame conservation in the gene junction were checked by sequencing of both strands in each construct.

BY-2 protoplasts preparation and transient transformation

Protoplasts were isolated from tobacco BY-2 cells cultured as previously described, four days after subculture in fresh medium. Protoplast preparation and transformation using a polyethyleneglycol–based technique were performed essentially according to the method of Neuhaus and Boevink (2001). Overnight digestion of cell walls was achieved with 0.1% pectolyase and 1% cellulase (Duchefa, Onozuka R-10) at 28°C in the dark with gentle shaking. Approximately $7.5x10^5$ protoplasts were transformed with 20 µg of plasmid DNA and 20 µg of carrier DNA to a medium containing 40% PEG-6000, 0.1 M $Ca(NO_3)_2$, 0.4 M mannitol, and 0.1% Mes, pH 8.0. Observation by confocal microscopy was made 48 h after transformation, and protoplasts were maintained at 26°C in the dark without shaking.

Immunolocalization on BY-2 protoplasts

The transiently transfected protoplasts were recovered by three volumes of W5 medium (Neuhaus and Boevink, 2001) and spinning for 10 min at 60 *g*. The fixative buffer (3.7%

paraformaldehyde, 0.5 M mannitol, 50 mM HEPES pH 5.8) added to the protoplasts pellet. Protoplasts were kept in fixative buffer one hour at room temperature, gently mixed occasionally, and transferred to 4°C overnight. The fixed protoplasts were rinsed three times for few minutes and then for one hour in fresh mannitol buffer (0.5 M mannitol, 50 mM HEPES pH 5.8) before using them for the immunolabelling step.

The protoplasts were incubated for 10 min exactly in 0.5 % Triton X-100 at room temperature to perforate the membranes. The cells were washed quickly with IL (ImmunoLabelling) buffer (PBS 1x, 0.5 % BSA, 0.05 % Tween-20, filter sterilized through a 0.45 µm filter) and washed twice more for 5 min. The protoplasts were blocked with IL buffer plus 1 % BSA for 15 min at room temperature. The protoplasts suspensions were aliquoted into individual tubes, one for each primary antibody, rabbit polyclonal anti-R2-HMW (Denery-Papini et al. 1996), and rabbit anti-BiP (Pedrazzini et al. 1997). The supernatant was removed, and cells were incubated for 1 h at room temperature with primary antibody diluted at 1/500 (rabbit anti-glutenin or rabbit anti-BiP antibody). The protoplasts were rinsed quickly with 100 µl of IL buffer, and washes were repeated three times for 10 min each. After removing supernatant, the secondary antibody was added at a 1/100 dilution (Alexa Fluor 546 goat anti-rabbit) for 1 h at room temperature. The protoplasts were quickly rinsed with 100 µl of IL buffer, and the washes were repeated three times for 10 min each with 100 µl of fresh IL buffer. Cells were analyzed with a Zeiss LSM 410 microscope at 546 nm excitation for Alexa Fluor 546. Emissions were detected using a 550–614 nm band-pass filter.

Confocal Imaging

Confocal microscopy experiments were performed with the facilities of platform Biopolymers, Interactions and Structural Biology, INRA Nantes. An inverted Zeiss LSM 410 Axiovert confocal laser scanning microscope was used to examine the subcellular localization of GFP fluorescence. The samples were examined with a water-immersed x40/NA objective. GFP was excited at 488 nm with a blue argon ion laser, and emission was collected via a photomultiplier through a 522-nm band-pass filter. Alexa Fluor 546 was excited at 546 nm. Emissions were detected using a 550–614 nm band-pass filter.

Results

Construction of glutenin fluorescent proteins

HMW glutenin subunit (GS) 1Dy10 (Fig. 1a), has a signal peptide consisting of the first 21 amino acids, which is cleaved off during protein synthesis in the endosperm of wheat. It is highly homologous with the signal peptides of other HMW GS (Kreis et al. 1985; Halford et al. 1987). The primary structure of the mature 1Dy10 subunit consists of three domains, a non-repetitive N-terminal domain of 104 residues, a repetitive central domain of 502 residues and a non-repetitive C-terminal domain of 42 residues. Seven cysteine residues are present in 1Dy10 with five in the N-terminal region, one in the C-terminal region and one at the end of the repetitive region. The central repetitive domain of 1Dy10 consists of nonapeptides (consensus GYYPTSLQQ), hexapeptides (consensus PGQGQQ), and tripeptide (GQQ).

The hydrophobic domains responsible for elasticity of elastin contain a repeated motif VPGVG (Urry and Parker 2002). To modify the elasticity of the HMW glutenin, 40 tandem repeats of the elastin-like peptide motif VPGVG (ELPx40) (Fig. 1b) were inserted in the HMW glutenin sequence. Because the ELPx40 sequence was strictly repeated, PCR technology could not be used for the cloning and insertion into GS sequence, so cloning with restriction sites was used.

In order to introduce the ELPx40 sequence into the GS repetitive domain, the GS cDNA was modified by directed mutagenesis to insert one BglII restriction site. Because of the high degree of repetition in this domain, its modification by PCR resulted in the loss of 315 bp of repeated domain and in a new and unique restriction site BglII. This unique BglII site was used to introduce the ELPx40 nucleotide sequence of 630 pb in the modified GS (GS_M). The GS_M protein therefore consisted of 105 fewer amino acids than the wild-type GS (GS_W) protein (Fig. 1a between double arrows). This loss was replaced by 210 amino acid of ELPx40 (40xVPGVG, Fig. 1b) in the GS_M-ELP protein. The GS_M sequence was used for protoplasts transformation as control in comparison with the GS_W sequence.

The GS_W (Fig. 2a), GS_M (Fig. 2b) and GS_M-ELP (Fig. 2c) nucleotide sequences were obtained without STOP codons and fused upstream of the Green Fluorescent Protein (GFP) sequence in order to allow the accessibility and functionality of the glutenin N-terminal signal peptide. These three sequences were introduced into the pRTd35S-GFP plasmid under the control of the double cauliflower mosaic virus 35S promoter and terminator. These GFP

fusions were expressed in BY-2 protoplasts in order to determine their subcellular localization.

Expression and localization in BY-2 cells

We first investigated the possibility to express the GS_W, GS_M and GS_M-ELP proteins in fusion with GFP and then determined their subcellular localization in BY-2 cells. After transient transformation of BY-2 protoplasts and 48 h of incubation in the dark, expression of the fluorescent proteins was observed by confocal microscopy. The efficiency of the transient BY-2 protoplast transformation was checked with the expression of GFP alone (Fig. 3a). Confocal microscopy revealed greater fluorescence of the GFP alone in the nucleus than in the cytoplasm. This agrees with the rapid import of GFP into the nuclear compartment as described previously (Chiu et al. 1996).

For each fusion protein, the cellular sections illustrated were chosen after three-dimensional reconstruction of whole BY-2 cells in order to select the confocal plane that best revealed the nuclei and the ER network (Fig. 3b-g).

With GS_W-GFP expression, localization of the fluorescent signal was observed to be more concentrated around the nucleus, where the ER network is very dense (Fig. 3b). No fluorescence was observed in the vacuole and the nucleus. In addition, in the confocal plane passing through the protoplast surface, the fluorescence signal formed a network characteristic of the ER (Fig. 3c). In these two planes, a number of discrete spots were observed by the emission of high fluorescence intensity and identified as protein body-like structures. We estimated the size of these as about 0.5 µm. The probable localization of GS_W-GFP in the ER and the observation of body-like structure in BY-2 protoplasts are consistent with the endogenous trafficking of wheat prolamins in endosperm cells of wheat seeds.

The fluorescent signal observed with the expression of the GS_M-GFP in tobacco protoplasts was also concentrated around the nucleus and vacuoles (Fig. 3d) but absent from these compartments. On peripheral cell focal plane, the fluorescent signal strongly suggested localization of GS_M-GFP in the ER network (Fig. 3e). On the other hand, the fusion protein was mainly not aggregated into protein body-like structures (PB-like structures), contrasting with those observed with the full length GS_W-GFP fusion protein (Table 1, Fig. 4).

Expression of the GS_M-ELP-GFP fusion protein also showed localization of the fluorescent signal around the nucleus and vacuole, and organized into the ER network (Fig. 3f, g). PB-like structures were present in cells expressing the GS_M-ELP-GFP protein.

Although we observed a certain disparity between the different cells, PB-like structures were generally more numerous than those bodies observed with GS_M-GFP expression. However, they were less numerous, smaller and less individualized in comparison with those observed with GS_W-GFP expression (Table 1, Fig. 4). Interestingly the fluorescence was low with the GS_M-ELP-GFP construct compared to those obtained with the two others constructs, which could be due to a lower expression level (Fig. 5).

Checking *in vivo* integrity of fusion between GS proteins and GFP

In order to ensure that the observed green fluorescent signal corresponded to GS fusion proteins and to ensure of the integrity of the fusion proteins, we carried out an immunolocalization using specific HMW glutenin subunit (HMW GS) antibodies. Transient transformed protoplasts were first fixed before being subjected to immunolocalization (see Material and Methods).

GFP alone was expressed in BY-2 protoplasts and used as negative control. This showed no fluorescence after hybridization with the HMW GS antibody, which confirmed the absence of endogenous HMW-GS-like sequences in tobacco BY-2 cells and the specificity of the HMW GS antibody (Fig. 6a-c).

As previously demonstrated with non-fixed cells, protoplasts expressing all three glutenin fusion proteins (GS_W-GFP, GS_M-GFP and GS_M-ELP-GFP) showed that the fluorescent signal was more concentrated around the nucleus and at the cell periphery, strongly suggesting an ER localization (Fig. 6e, h, k). No fluorescence was present in the vacuole and the nucleus compartments. Some protein body-like structures were also observed with the GS_W-GFP but not in a significant manner with GS_M-ELP-GFP and GS_M-GFP expression. In this experiment, cells were submitted to a fixation treatment before being immunolabelled, which could explain that PB-like structures were less visible than in living cells, especially with GS_M-GFP and GS_M-ELP-GFP constructs with which PB-like structures were seen to be less numerous and smaller than with GS_W-GFP. In addition, weaker fluorescence intensity was also observed in the case of the GS_M-ELP-GFP as shown previously.

After immunolocalization with the HMW GS antibody, protoplasts transformed with the three GS constructs showed the same pattern of fluorescence as that obtained with the GFP fluorescence (Fig. 6d, g, j), and merging of these patterns showed perfect co-localization

between the green fluorescence fusion protein and the HMW GS antibody hybridization (Fig. 6f, i, l). This confirmed that the integrity of fusion proteins was preserved.

Co-localization of the GS fusion protein with BiP

In order to confirm that the fluorescence observed with all the GS constructs was specifically localized in the ER and therefore whether the 1Dy10 HMW glutenin and its variants were retained in the ER or transported out of this compartment when expressed in tobacco protoplasts, we performed immunolocalization with the transformed tobacco protoplasts using a specific ER marker (Fig. 7). We chose to use an anti-BiP antibody as an ER marker because the BiP is a typical ER-resident protein (Napier et al. 1992). BiP is a member of the heat-shock protein-70 family and is involved in the folding and assembly of proteins in the ER. BiP has been implicated in PB biogenesis (Zhang and Boston 1992; Li et al. 1993) and has been shown to interact with storage proteins containing PBs in a manner that is distinct from its normal chaperone activity (Frigerio et al. 2001). It is probable that the folding and assembly of the gluten proteins is assisted by ER lumenal proteins including BiP (Grimwade et al. 1996; DuPont et al. 1998).

Forty-eight hours after transformation of protoplasts with GS_W-GFP, GS_M-GFP or GS_M-ELP-GFP constructs, the cells were fixed and immunohybridized with anti-BiP antibody. They were then observed by confocal microscopy at two different wavelengths to detect the GFP signal in the green channel (Fig. 7e, h, k) and Alexa Fluor secondary antibody signal for BiP localization in the red channel (Fig. 7d, g, j).

As previously, we used the GFP alone as negative control. As foreseen, protoplasts transformed with GFP alone showed a difference in compartmentalization between GFP and BiP protein expression (Fig. 7a, b, c). We observed a clear fluorescent signal in the nucleus and a lower signal in the cytoplasm for GFP while BiP showed no fluorescent signal in the nucleus and a strong signal in ER compartment.

The three fluorescent fusion proteins were found to be confined to the ER region, which was also stained by the anti-BiP antibody (Fig. 7). The subcellular co-localization of GS_W-GFP (Fig. 7d- f), GS_M-GFP (Fig. 7g-i) and GS_M-ELP-GFP (Fig. 7j-l) proteins with the BiP therefore confirmed their ER unique localization. As seen previously, PB-like structures were especially formed in protoplasts transformed with GS_W-GFP. These PB-like structures co-localized with the BiP marker, indicating that BiP was also present in these bodies.

Whereas the GS_W-GFP protein was essentially present as PB-like structures, the BiP staining is more diffuse in the ER compared to the GS_W-GFP.

Discussion

There are currently few reports of prolamin fusions with fluorescent protein (Foresti et al. 2008; Saito et al. 2009), although GFP has been widely employed as a marker protein to demonstrate passage of protein through the plant secretory pathway. To our knowledge, this is the first reported example of fluorescent protein fusion with a HMW glutenin. We can conclude that fusion with GFP does not modify the glutenin subcellular localization in our cell model since glutenin fused to fluorescent protein is retained in the ER and forms PB-like structures that are similar to the reports of endogenously expressed prolamins in wheat endosperm cells (Levanony et al. 1992; Shewry et al. 1995). Whereas the size of the PB-like structures formed by GS subunit expression in tobacco cells was smaller than that observed in wheat endosperm cells, they could be compared to those which were present at the early stages of development in wheat seeds (Loussert et al. 2008). These similarities between this heterologous expression system and expression in wheat endosperm show that the expression in tobacco protoplast is a suitable model to study the HMW glutenin subunit trafficking. Moreover, the fusion protein GS-GFP retained its integrity during its trafficking in tobacco cells, since we demonstrated a perfect co-localization of the immunolabelling with glutenin and GFP antibodies in the transformed tobacco protoplasts.

As described in this work, the GS_W-GFP, GS_M-GFP and GS_M-ELP-GFP fusion proteins are located in the ER, despite the differences in the size and sequence of the repeated domain. However, differences were revealed between their forms of accumulation. After deletion of 105 amino acids from the repeated sequence of the GS_W, the truncated protein seemed to possess a weaker ability to aggregate, since a small amount of fluorescent protein body-like structures was observed. This deletion from the repeated sequence then reduces the ability of the glutenin to form protein body-like structures in the ER. Whereas mechanisms allowing the assembly of prolamins in the ER and the formation of protein bodies are still unknown, our results suggest the importance of an intact repetitive domain for glutenin accumulation. It has been suggested that accumulation in the ER lumen could be a consequence of the ability of glutenins to form insoluble aggregates (Shewry 1999). Non-covalent interactions, notably hydrogen bonds formed by the repeated glutamine-rich domain, may be established, resulting in protein precipitation and the formation of hydrated protein particles (Rubin et al. 1992). Other sequences present in cereal storage proteins may favour aggregation and protein body formation. For example, the PPPVHL repeat domain of Zera^{TM}, derived from maize γ-zein (Torrent et al. 2009), adopts a polyproline II conformation forming

an extended amphipathic helix, which is able to self-assemble and form cylindrical micelles (Kogan et al. 2002). This intramolecular interaction among seed storage proteins appears to be essential for their aggregation and the biogenesis of PBs (Kogan et al. 2001). Our study therefore suggested that the decrease in GS_M-GFP protein bodies could be due to the loss of repetitive motifs in the glutenin sequence, and thus the ability to form hydrogen bonds, as suggested by Wan et al. (2005) and Feeney et al. (2003).

The GS_M-ELP-GFP protein contains 40 repeats of the VPGVG pentapeptide replacing the 105 amino acids lost in the GS_M-GFP protein. While we observed a cell-to-cell variation in the subcellular distribution of the fusion proteins, our experimental and quantitative data supported the fact that in a general manner, more protein body-like structures were observed in the case of GS_M-ELP-GFP expression compared to the GS_M-GFP expression. This result suggested that the addition of 40xELP to the GS_M sequence seemed to improve its ability to aggregate. A recent report also showed that an elastin-like polypeptide (28xVPGVG) tagged with KDEL (an ER retention signal), significantly enhanced accumulation of recombinant protein in the ER (Conley et al. 2009, for review see Floss et al. 2010). They suggested that this induction of PB formation resulted from the accumulation of heterologous protein reaching a critical local concentration exceeding the normal solubility limit, which subsequently triggered the aggregation and assembly into spherical PBs. Floss et al. (2009) investigated the localization of recombinant antibodies fused to 100xELP after seed-specific expression in tobacco. Complete full-length ELPylated recombinant antibodies were localized in small, putative ER-derived protein bodies found throughout the cytosol but also in the protein storage vacuole. An ELPylated scFv expressed in seeds was shown to form protein bodies in the cytoplasm. The authors hypothesize that ELP retard the trafficking of the fusion proteins along the endomembrane system. Furthermore, the ELPs are hydrophobic sequences already reported to possess autoassembly, properties responsible in part for its entropic elastic force (Urry and Parker, 2002). The relatively high content of hydrophobic amino acids suggests that ELPs may aggregate with themselves as a means of avoiding the hydrophilic conditions in the aqueous lumen of the ER to protect the hydrophobic ELP from the aqueous phase, as seen previously in the case of maize zein (Kogan et al. 2002). Analogous to seed storage proteins, mammalian-derived ELP also possesses the ability to self-aggregate and undergo co-acervation *via* an ordered process, leading to the formation of a stable supramolecular structure (Miao et al. 2003). As the temperature rises, the soluble ELPs polypeptide collapses from an extended chain into an insoluble twisted filament structure consisting of β-spirals comprising type II β-turns (Reiersen et al. 1998; Lee et al. 2001). This

process could be also explained by the development of closer associations between single ß-spiral strands of ELP and the subsequent formation of interstrand ß-sheet structures. These structures might mediate close association or aggregation of different ELP chains (Serrano et al. 2007). The temperature used for plant cells culture by Conley et al. (2009) should not have induced the aggregation of a 28xELP tag, however, the concentration of ELP is also known to affect the precipitation behaviour (Meyer and Chilkoti 2004; Ge and Filipe 2006). Although not proven, high local concentrations and long stretches of ELP-fusion proteins in the ER may play a role in their aggregation and subsequent formation into novel PBs.

The GS_M-ELP-GFP protein seemed to form protein body-like structures but these were smaller and less numerous than those formed with GS_W-GFP protein expression. In our study, the propensity of ELP to induce PB-like structures seemed to be partially restored but could not completely compensate for the ability of the GS_W-GFP protein to form PB-like structures. The ELP motifs were inserted into the repeated sequence of the glutenin. This localization may explain the less numerous PB-like structures with the chimeric protein because in this case the ELP motifs were in a specific environment that could influence its structure and physical-chemical properties. The expected physical-chemical properties of the GS_M-ELP-GFP triblock protein could induce an aggregate structure that differs from that adopted by the wild-type glutenin subunit. Moreover we used a double 35S promoter that leads to an expression level weaker than in the work of Conley (2009), where they used the p19 suppressor of gene silencing, which has been found to significantly increase the production levels of recombinant proteins in plants (Voinnet et al. 2003; Sudarshana et al. 2006). The expression level of the 28xELP-KDEL was enhanced in the presence of p19 by approximately 30-fold after agroinfiltration in tobacco leaves compared to the expression without p19 (Conley et al. 2009). The use of p19 suppressor could lead to an increase in the expression level of our chimerical protein in tobacco cells and enhance the PB-like structures formation, in order to produce the GS_M-ELP protein and study its mechanical properties.

We also noted a lower fluorescence level for the protoplasts transformed with the GS_M-ELP-GFP construct compared to the protoplasts expressing the wild GS_W-GFP and GS_M-GFP, which could be related to the protein stability. This is the first example where the fusion of a protein with an ELP sequence reduces the accumulation level of the fusion protein. However, in contrast to other ELP fusion proteins described in the literature, the ELP sequence in our case was located into the glutenin sequence and not at its extremity, the objective being to obtain a triblock protein. This could explain the relatively low level of glutenin-ELP-GFP expression. Indeed, the protein stability of GS_M-ELP-GFP protein could

be disturbed by the presence of exogenous sequences inside the glutenin sequence, which can induce ER-associated protein degradation (ERAD). This quality control in the ER prevents the arrival of incorrectly or incompletely folded proteins at their final destination and targets permanently misfolded proteins for degradation. Such proteins have a high affinity for the ER chaperone BiP and are finally degraded via retrograde translocation from ER lumen back to the cytosol. ERAD is currently thought to constitute the main disposal route for misfolded proteins (Anelli and Sitia 2008). The protein degradability hypothesis could explain the low fluorescence and low amount of PB-like structures observed with the GS_M-ELP-GFP protein expression. Further investigations could focus on the location of ELP sequences at an extremity of the glutenin, and determine the expression level and PB formation resulting from such a fusion protein.

The ER-resident chaperone BiP has been shown to be associated with prolamin PBs in rice (Muench et al. 1997) and wheat (Levanony et al. 1992). In addition, BiP has been suggested to be involved in PB biogenesis because of its role in retention of prolamins in the ER lumen by facilitating their folding and assembly into insoluble PBs in developing seeds (Zhang and Boston 1992; Li et al. 1993) and leaves of transgenic plants (Bagga et al. 1997; Mainieri et al. 2004). Previous studies identified specific BiP binding motifs within storage proteins, such as zein (Randall et al. 2005) and phaseolin (Foresti et al. 2003), using BiP scoring software (Blond-Elguindi et al. *1993*). Furthermore, BiP has been shown to interact with seed storage proteins in a specific manner, which is unique from its normal chaperone activity (Mainieri et al. 2004; Frigerio et al. 2001; Randall et al. 2005). Although Conley et al. (2009) identified a strong BiP-binding motif repeated throughout the ELP sequence, the co-immunoprecipitation analysis revealed no specific interactions between ELP and the ER-chaperone BiP. In the present study, BiP was co-localized with the different fusion proteins, suggesting that BiP may play a role in the trafficking of the GS variants in tobacco protoplasts, as discussed previously (Zhang and Boston 1992; Li et al. 1993). Further investigations could focus on the involvement of BiP protein and other proteins such as the protein disulphide isomerase, which were suggested to play a major role in the prolamin assembly during their trafficking.

To conclude, in addition to show the ER localization and the formation of PB-like structures when an HMW-GS of wheat seed was expressed in fusion with a fluorescent protein in tobacco cells model, we demonstrated the expression of a new chimeric triblock protein based on glutenin and elastin-like motifs and showed its subcellular localization in

tobacco cells. These experiments open the way to investigate the possibility to perform high level production of this triblock protein in order to conduct structural and functional analysis, and to characterize its specific mechanical and *ex vivo* associative properties.

Acknowledgements The authors thank Denis Lourdin for the discussion of the triblock protein design and properties, and Peter Shewry and Paola Tosi for the critical review of the manuscript. We also thank Axelle Bouder for her technical assistance in molecular and cellular biology experiments, Brigitte Bouchet and Marie-Françoise Devaux for their assistance with confocal microscopy and fluorescence quantification, respectively. We gratefully acknowledge MESR for the post-doctoral grant of A. Saumonneau.

References

Anelli T and Sitia R (2008) Protein quality control in the early secretory pathway. Embo J 27:315-27

Arias FJ, Reboto V, Martin S, Lopez I and Rodriguez-Cabello JC (2006) Tailored recombinant elastin-like polymers for advanced biomedical and nano(bio)technological applications. Biotechnol Lett 28:687-95

Bagga S, Adams HP, Rodriguez FD, Kemp JD and Sengupta-Gopalan C (1997) Coexpression of the maize delta-zein and beta-zein genes results in stable accumulation of delta-zein in endoplasmic reticulum-derived protein bodies formed by beta-zein. Plant Cell 9:1683-96

Blond-Elguindi S, Cwirla SE, Dower WJ, Lipshutz RJ, Sprang SR, Sambrook JF and Gething MJ (1993) Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP. Cell 75:717-28

Chiu W, Niwa Y, Zeng W, Hirano T, Kobayashi H and Sheen J (1996) Engineered GFP as a vital reporter in plants. Curr Biol 6:325-30

Conley AJ, Joensuu JJ, Menassa R and Brandle JE (2009) Induction of protein body formation in plant leaves by elastin-like polypeptide fusions. BMC Biol 7:48

Denery-Papini S, Popineau Y, Quillien L, Van Regenmortel MHV (1996) Specificity of antisera raised against synthetic peptide fragments of high Mr glutenin subunits. Journal of Cereal Science 23:133‑144

Dupont FM, Hurkman WJ, Tanaka CK and Chan R (1998) BiP, HSP70, NDK and PDI in wheat endosperm. I. Accumulation of mRNA and protein during grain development. Physiologia Plantarum 102:70–79

Feeney KA, Wellner N, Gilbert SM, Halford NG, Tatham AS, Shewry PR and Belton PS (2003) Molecular structures and interactions of repetitive peptides based on wheat glutenin subunits depend on chain length. Biopolymers 72:123-31

Floss DM, Sack M, Stadlmann J, Rademacher T, Scheller J, Stoger E, Fischer R and Conrad U (2008) Biochemical and functional characterization of anti-HIV antibody-ELP fusion proteins from transgenic plants. Plant Biotechnol J 6; 379-91

Floss DM, Sack M, Arcalis E, Stadlmann J, Quendler H, Rademacher T, Stoger E, Scheller J, Fischer R, Conrad U (2009) Influence of elastin-like peptide fusions on the quantity and quality of a tobacco-derived human immunodeficiency virus-neutralizing antibody. Plant Biotechnol J 7:899-913

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

Foresti O, De Marchis F, De Virgilio M, Klein EM, Arcioni S, Bellucci M and Vitale A (2008) Protein domains involved in assembly in the endoplasmic reticulum promote vacuolar delivery when fused to secretory GFP, indicating a protein quality control pathway for degradation in the plant vacuole. Mol Plant 1:1067-76

Foresti O, Frigerio L, Holkeri H, De Virgilio M, Vavassori S and Vitale A (2003) A phaseolin domain involved directly in trimer assembly is a determinant for binding by the chaperone BiP. Plant Cell 15:2464-75

Frigerio L, Pastres A, Prada A and Vitale A (2001) Influence of KDEL on the fate of trimeric or assembly-defective phaseolin: selective use of an alternative route to vacuoles. Plant Cell 13:1109-26

Galili G, Altschuler Y and Levanony H (1993) Assembly and transport of seed storage proteins. Trends Cell Biol 3:437-42

Ge X and Filipe CD (2006) Simultaneous phase transition of ELP tagged molecules and free ELP: an efficient and reversible capture system. Biomacromolecules 7:2475-8

Grimwade B, Tatham AS, Freedman RB, Shewry PR and Napier JA (1996) Comparison of the expression patterns of genes coding for wheat gluten proteins and proteins involved in the secretory pathway in developing caryopses of wheat. Plant Mol Biol 30:1067-73

Halford NG, Forde J, Anderson OD, Green FC and Shewry PR (1987) The nucleotide and deduced amino acid sequences of an HMW glutenin subunit gene from chromosome 1B of bread wheat (Triticum aestivum L.) and comparison with those of genes from chromosomes 1A and 1D. Theor Appl Genet 75:117-126

Kogan MJ, Dalcol I, Gorostiza P, Lopez-Iglesias C, Pons M, Sanz F, Ludevid D and Giralt E (2001) Self-assembly of the amphipathic helix (VHLPPP)8. A mechanism for zein protein body formation. J Mol Biol 312:907-13

Kogan MJ, Dalcol I, Gorostiza P, Lopez-Iglesias C, Pons R, Pons M, Sanz F and Giralt E (2002) Supramolecular properties of the proline-rich gamma-Zein N-terminal domain. Biophys J 83:1194-204

Kreis M, Forde BG, Rahman S, Miflin BJ and Shewry PR (1985) Molecular evolution of the seed storage proteins of barley, rye and wheat. J Mol Biol 183:499-502

Kumar M, Sanford KJ, Cuevas WA, Du M, Collier KD and Chow N (2006) Designer protein-based performance materials. Biomacromolecules 7:2543-51

Lee J, Macosko CW and Urry DW (2001) Elastomeric polypentapeptides cross-linked into matrixes and fibers. Biomacromolecules 2:170-9

Levanony H, Rubin R, Altschuler Y and Galili G (1992) Evidence for a novel route of wheat storage proteins to vacuoles. J Cell Biol 119:1117-28

Li X, Wu Y, Zhang DZ, Gillikin JW, Boston RS, Franceschi VR and Okita TW (1993) Rice prolamine protein body biogenesis: a BiP-mediated process. Science 262:1054-6

Lombardi A, Barbante A, Cristina PD, Rosiello D, Castellazzi CL, Sbano L, Masci S and Ceriotti A (2009) A relaxed specificity in interchain disulfide bond formation characterizes the assembly of a low-molecular-weight glutenin subunit in the endoplasmic reticulum. Plant Physiol 149:412-23

Loussert C, Popineau Y and Mangavel C (2008) Protein bodies ontogeny and localization of prolamin components in the developing endosperm of wheat caryopses. J. Cereal Sc 47:445-456

Mainieri D, Rossi M, Archinti M, Bellucci M, De Marchis F, Vavassori S, Pompa A, Arcioni S and Vitale A (2004) Zeolin. A new recombinant storage protein constructed using maize gamma-zein and bean phaseolin. Plant Physiol 136:3447-56

Mangavel C, Barbot J, Bervas E, Linossier L, Feys M, Gueguen J and Popineau Y (2002) Influence of prolamin composition on mechanical properties of cast wheat gluten films. Journal of Cereal Science 36:157-166

Meyer D.E. and Chilkoti A. (2004) Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides. Biomacromolecules 5:846-51

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

Muench D.G, Wu Y, Zhang Y, Li X, Boston RS and Okita TW (1997) Molecular cloning, expression and subcellular localization of a BiP homolog from rice endosperm tissue. Plant Cell Physiol 38:404-12

Nagata T, Nemoto Y and Hasezawa S (1992) Tobacco BY-2 cell line as the "HeLa" celle in the cell biology of higher plants. Int Rev Cytol 132:1-30

Napier RM, Fowke LC, Hawes C, Lewis M and Pelham HR (1992) Immunological evidence that plants use both HDEL and KDEL for targeting proteins to the endoplasmic reticulum. J Cell Sci 102:261-71

Neuhaus JM and Boevink P (2001) The green fluorescent protein (GFP) as reporter gene in plant cells. Plant Cell Biology 127-142

Parker ML and Hawes CR (1982) The Golgi-apparatus in developing endosperm of wheat (triticum aestivum L). Planta 154:277-283

Patel J, Zhu H, Menassa R, Gyenis L, Richman A and Brandle J (2007) Elastin-like polypeptide fusions enhance the accumulation of recombinant proteins in tobacco leaves. Transgenic Res 16:239-49

Pedrazzini E, Giovinazzo G, Bielli A, De Virgilio M, Frigerio L, Pesca M, Faoro F, Bollini R, Ceriotti A and Vitale A (1997) Protein quality control along the route to the plant vacuole. Plant Cell 9:1869-80

Popineau Y, Deshayes G, Lefebvre J, Fido R, Tatham AS and Shewry PR (2001) Prolamin aggregation, gluten viscoelasticity, and mixing properties of transgenic wheat lines expressing 1Ax and 1Dx high molecular weight glutenin subunit transgenes. J Agric Food Chem 49:395-401

Randall JJ, Sutton DW, Hanson SF and Kemp JD (2005) BiP and zein binding domains within the delta zein protein. Planta 221:656-66

Reiersen H, Clarke AR and Rees AR (1998) Short elastin-like peptides exhibit the same temperature-induced structural transitions as elastin polymers: implications for protein engineering. J Mol Biol 283:255-64

Robert LS, Thompson RD and Flavell RB (1989) Tissue-specific expression of a wheat high molecular weight glutenin gene in transgenic tobacco. Plant Cell 1:569-78

Rodriguez-Cabello JC, Martin L, Alonso M, Arias FJ and Testera AM (2009) "Recombinamers" as advanced materials for the post-oil age. Polymer 50:5159-5169

Rubin R, Levanony H and Galili G (1992) Evidence for the Presence of Two Different Types of Protein Bodies in Wheat Endosperm. Plant Physiol 99:718-724

Saito Y, Kishida K, Takata K, Takahashi H, Shimada T, Tanaka K, Morita S, Satoh S and Masumura T (2009) A green fluorescent protein fused to rice prolamin forms protein body-like structures in transgenic rice. J Exp Bot 60:615-27

Scheller J, Henggeler D, Viviani A and Conrad U (2004) Purification of spider silk-elastin from transgenic plants and application for human chondrocyte proliferation. Transgenic Res 13:51-7

Scheller J, Leps M and Conrad U (2006) Forcing single-chain variable fragment production in tobacco seeds by fusion to elastin-like polypeptides. Plant Biotechnol J 4:243-9

Serrano V, Liu W, Franzen S (2007) An infrared spectroscopic study of the conformational transition of elastin-like polypeptides. Biophys J 93:2429-35

Shani N, Rosenberg N, Kasarda DD and Galili G (1994) Mechanisms of assembly of wheat high molecular weight glutenins inferred from expression of wild-type and mutant subunits in transgenic tobacco. J Biol Chem 269:8924-30

Shewry PR (1999) The synthesis, processing, and deposition of gluten proteins in the developing wheat grain. Cereal Foods World 44:587–589

Shewry PR and Halford NG (2002) Cereal seed storage proteins: structures, properties and role in grain utilization. J Exp Bot 53:947-958
Shewry PR, Halford NG, Arthus ST, Popineau Y, Lafiandra D and Belton PS (2003) The high molecular weight subunit of wheat glutenin and their role in determining wheat processing properties. Advances in food and nutrition research 45:219-302
Shewry PR, Halford NG, Belton PS and Tatham AS (2002) The structure and properties of gluten: an elastic protein from wheat grain. Philos Trans R Soc Lond B Biol Sci 357:133-42
Shewry PR, Napier JA and Tatham AS (1995) Seed storage proteins: structures and biosynthesis. Plant Cell 7:945-56
Shimoni Y, Blechl AE, Anderson OD and Galili G (1997) A recombinant protein of two high molecular weight glutenins alters gluten polymer formation in transgenic wheat. J Biol Chem 272:15488-95
Shy G., Ehler L, Herman E and Galili G (2001) Expression patterns of genes encoding endomembrane proteins support a reduced function of the Golgi in wheat endosperm during the onset of storage protein deposition. J Exp Bot 52:2387-8
Sudarshana MR, Plesha MA, Uratsu SL, Falk BW, Dandekar AM, Huang TK and Mcdonald KA (2006) A chemically inducible cucumber mosaic virus amplicon system for expression of heterologous proteins in plant tissues. Plant Biotechnol J 4:551-9
Tatham A.S. and Shewry P.R. (2000) Elastomeric proteins: biological roles, structures and mechanisms. Trends Biochem Sci 25:567-71
Torrent M, Llompart B, Lasserre-Ramassamy S, Llop-Tous I, Bastida M, Marzabal P, Westerholm-Parvinen A, Saloheimo M, Heifetz PB and Ludevid MD (2009) Eukaryotic protein production in designed storage organelles. BMC Biol 7:5
Tosi P, Parker M, Gritsch CS, Carzaniga R, Martin B and Shewry PR (2009) Trafficking of storage proteins in developing grain of wheat. J Exp Bot 60:979-91
Urry DW and Parker TM (2002) Mechanics of elastin: molecular mechanism of biological elasticity and its relationship to contraction. J Muscle Res Cell Motil 23:543-59
Vitale A and Galili G (2001) The endomembrane system and the problem of protein sorting. Plant Physiol 125:115-8
Voinnet O, Rivas S, Mestre P and Baulcombe D (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J 33:949-56
Wan Y, Yan Z, Liu K, Zheng Y, D'ovidio R, Shewry PR, Halford NG and Wang D (2005) Comparative analysis of the D genome-encoded high-molecular weight subunits of glutenin. Theor Appl Genet 111:1183-90
Wellner N, Marsh JT, Savage AW, Halford NG, Shewry PR, Clare Mills EN and Belton PS (2006) Comparison of repetitive sequences derived from high molecular weight subunits of wheat glutenin, an elastomeric plant protein. Biomacromolecules 7:1096-103
Wright ER and Conticello VP (2002) Self-assembly of block copolymers derived from elastin-mimetic polypeptide sequences. Adv Drug Deliv Rev 54:1057-73
Zhang F and Boston RS (1992) Increases in binding protein (BiP) accompany changes in protein body morphology in three high-lysine mutants of maize. Protoplasma 171:142-152

Figure legends

Fig. 1 1Dy10 subunit. **a** Amino acid sequence of 1Dy10 (GS_W) protein, the signal peptide and non-repeated sequence are framed respectively by a discontinue line and by a continue line. The repeated domain is underlined. The deleted sequence of the 1Dy10 mutated protein (GS_M) is localized between double arrows. **b** Amino acid sequence of ELP introduced into GS_M protein in replacement of sequence deleted between double arrows of amino acid sequence of GS_W protein **(a)**.

Fig. 2 Schematic representation of the expression cassettes and designation of chimerical proteins expressed in transient transformed tobacco cells. **a** GS_W-GFP: 1Dy10 protein composed of a long repeated domain and flanked by two non repeated domains in N and C terminal extremities of 1Dy10 with respectively 5 and 2 cysteines (short lines). **b** GS_M-GFP resulting of a loss of sequence represented between two double arrows on A. into the repeated domain by directed mutagenesis to introduce a restriction site. **c** GS_M-ELP-GFP results of ELP (Elastin Like-Peptide) coding for 40 VPGVG repetitions introduced into the repeated domain of mutated Dy10 using the new restriction site BglII. These three proteins were fused to GFP protein at their C-terminal extremity and placed under the control of cauliflower mosaic virus 35S promoter and in the upstream cauliflower mosaic virus 35S terminator.

Fig. 3 Expression and subcellular localization of glutenin subunits in BY2 cells. Tobacco protoplasts were transiently transformed with a plasmid encoding the GFP **(a)**, the GS_W-GFP **(b, c)**, the GS_M-GFP **(d, e)** or the GS_M-ELP-GFP **(f, g)** proteins and observed after 48 h transfection. The subcellular distribution of these proteins was examined by confocal laser scanning microscopy using the 488-nm blue line excitation, and emission of the fluorescence was collected via a photomultiplier through a 522-nm band-pass filter. The subcellular distribution was observed with two different confocal plans crossing the cells surface **(c, e, g)** or into the cells through the nucleus **(b, d, f)**. **n**, nucleus; **v**, central vacuole. Scale bar = 20 µm. *Experiment repetitions among* $n \geq 3$.

Fig. 4 Confocal image illustrations corresponding to the different cell categories defined in Table 1 ((a): numerous and well defined PB-like structures; (b): few and/or less individualized PB-like structures; (c): no PB-like structures) for each construct.

Fig. 5 Fluorescence intensity quantification from cells expressing the chimerical proteins. Fluorescence intensities were measured from confocal images of BY-2 protoplasts expressing GS_W-GFP, GS_M-GFP or GS_M-ELP-GFP proteins. They are expressed as the gray level mean per pixel.

Fig. 6 *In vivo* checking of fluorescence signal specificity due to fusion between the Dy10 chimerical protein and the GFP protein. Tobacco protoplasts were transiently transformed with a plasmid encoding the GFP **(a, b, c)**, the GS_W-GFP **(d-f)**, the GS_M-GFP **(g-i)** or the GS_M-ELP-GFP **(j-l)** proteins and fixed with 3% paraformaldehyde after 48 h transfection. The subcellular distribution of GFP protein and GS wild-type or mutated proteins was examined by confocal laser scanning microscopy using a combination of a rabbit anti-HMW antiserum. **a**, **d**, **g** and **j**: Reticular fluorescence of GS visualized by the anti-HMW and Alexa Fluor 546 goat anti-rabbit secondary antibody. **b**, **e**, **h** and **k**: Distribution pattern of the GFP fluorescent protein observed by confocal laser scanning microscopy using the 488-nm blue line excitation, and emission of the fluorescence was collected via a photomultiplier through a 522-nm band-pass filter. **c**, **f**, **i** and **l**: Overlaid images respectively of **a** (red) + **b** (green), **d** (red) + **e** (green), **g** (red) + **h** (green), **j** (red) + **k** (green), where the yellow color revealed the co-localization of green and red colors. **n**, nucleus; **v**, central vacuole. Scale bar = 20 µm. *Experiment repetitions among n≥3.*

Fig. 7 Co-localization of GS-GFP fusion proteins with an endoplasmic reticulum marker, the protein BiP. Tobacco protoplasts were transiently transformed with a plasmid encoding the GFP **(a-c)**, theGS_W -GFP **(d-f)**, the GS_M-GFP **(g-i)** or the GS_M-ELP-GFP **(j-l)** proteins and fixed with 3% paraformaldehyde after 48 h transfection. The subcellular distribution of fusion GFP-proteins and ER marker BiP was examined by confocal laser scanning microscopy using a combination of a rabbit anti-BiP antiserum. **a**, **d**, **g** and **j**: Reticular fluorescence of endogenous BiP visualized by the anti-BiP and Alexa Fluor 546 goat anti-rabbit secondary antibody. **b**, **e**, **h** and **k**: Distribution pattern of the heterologous fluorescent proteins observed by confocal laser scanning microscopy using the 488-nm blue line excitation, and emission of the fluorescence was collected via a photomultiplier through a 522-nm band-pass filter. **c**, **f**, **i** and **l**: Overlaid images respectively of **a** (red) + **b** (green), **d** (red) + **e** (green), **g** (red) + **h** (green), **j** (red) + **k** (green), where the yellow color revealed the co-localization of green and red colors. **n**, Nucleus; **v**, central vacuole. Scale bar = 20 µm. *Experiment repetitions among n≥3.*

Table

Table 1 Quantification of the transformed cells containing protein body-like structures

	cells containing:		
	numerous well defined PBs (a)	few and/or less individualized PBs (b)	no PBs (c)
GSw-GFP	100%	0%	0%
GSm-GFP	16.7%	33.3%	50%
GSm-ELP-GFP	33.3%	44.5%	22.2%

Analyzed cells were classified into three categories according to the protein body-like structures (PBs) they contained: the first category corresponds to cells containing numerous and well-defined PB-like structures; the second category corresponds to cells containing few and/or less individualized PB-like structures; and the third category corresponds to cells without PB-like structures.

Cells were provided from at least three individual protoplast transformation experiments, transformed with each of the three constructs. A total of 51, 48 and 54 cells were examined for the GS_W-GFP, GS_M-ELP and GS_M-ELP-GFP constructs respectively.

Figures

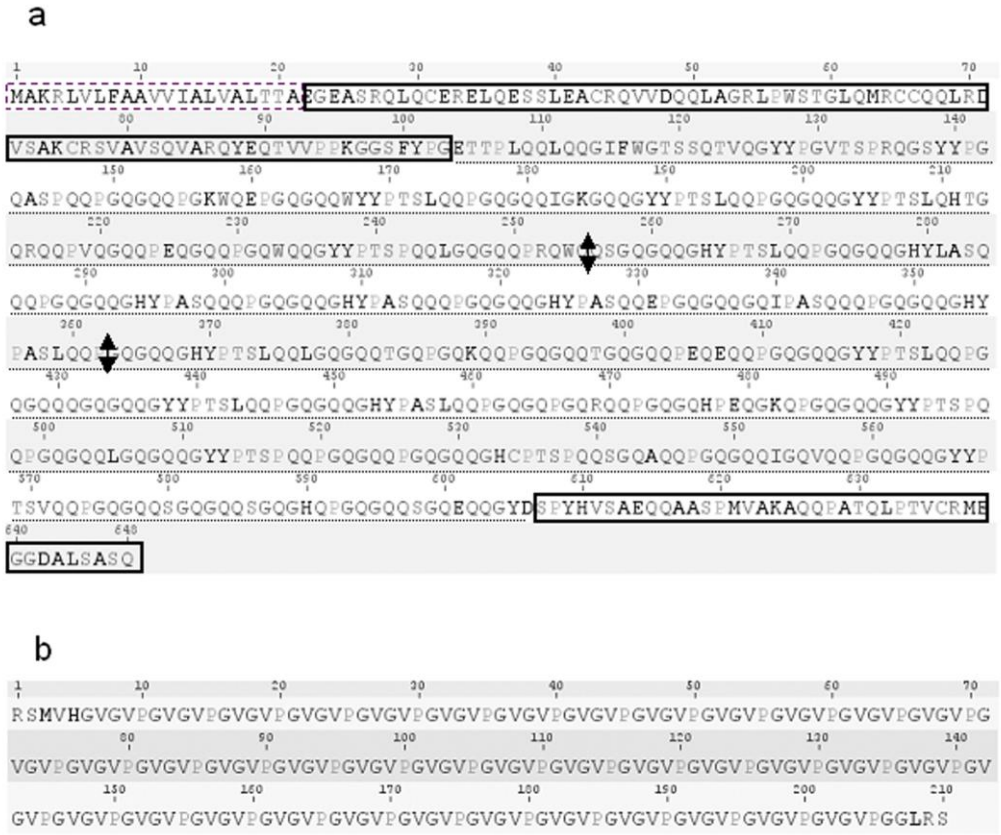

Figure 1

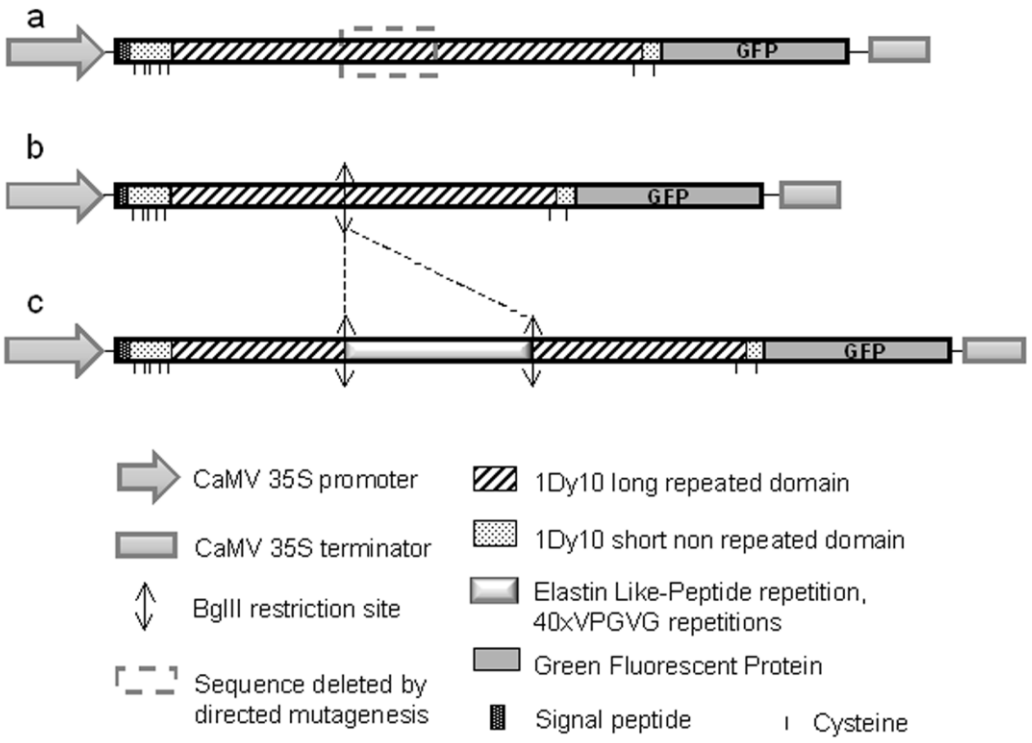
a
GFP
b
GFP
c
GFP
CaMV 35S promoter
CaMV 35S terminator
BglII restriction site
Sequence deleted by directed mutagenesis
1Dy10 long repeated domain
1Dy10 short non repeated domain
Elastin Like-Peptide repetition, 40xVPGVG repetitions
Green Fluorescent Protein
Signal peptide
Cysteine

Figure 2

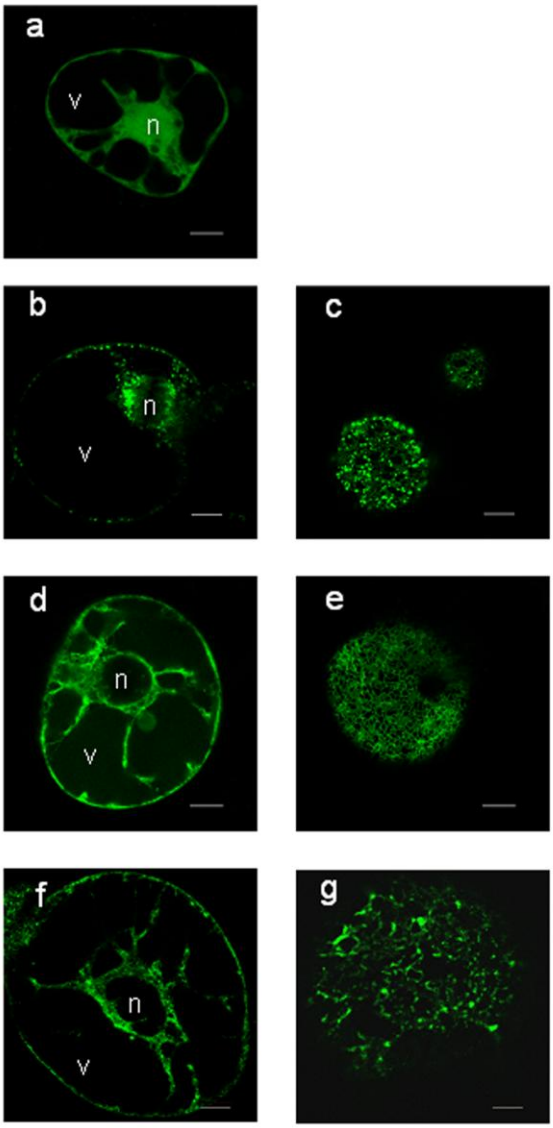

Figure 3

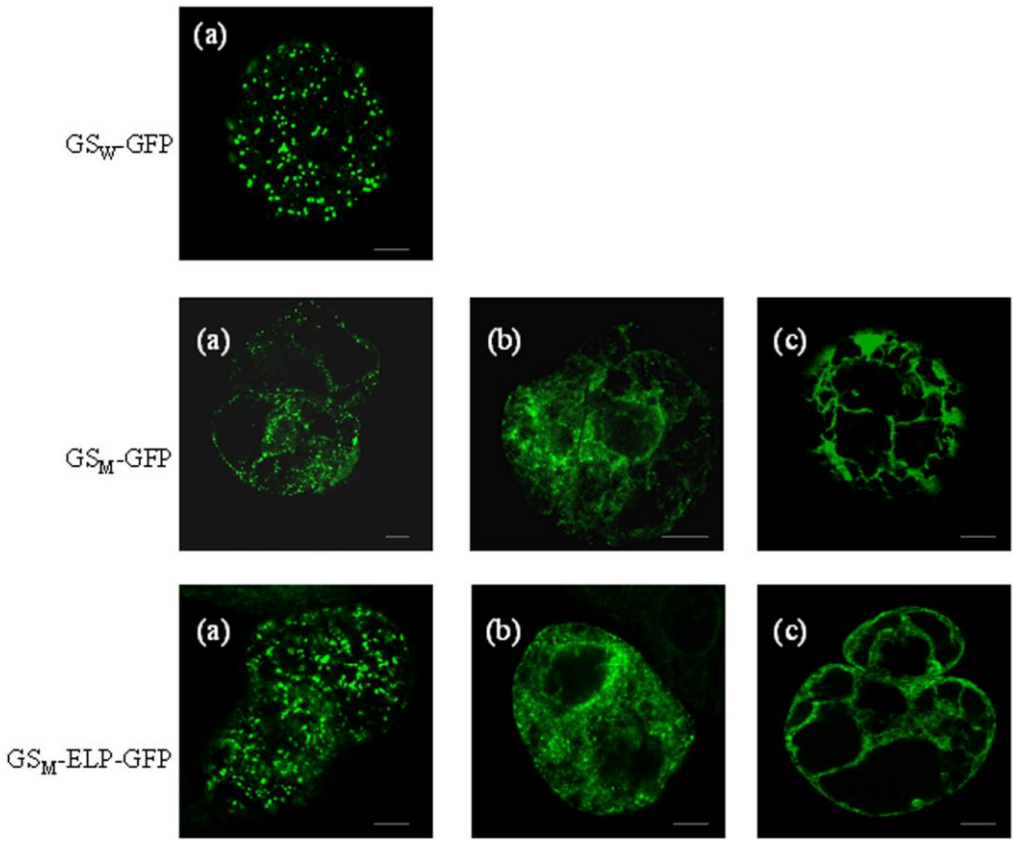

Figure 4

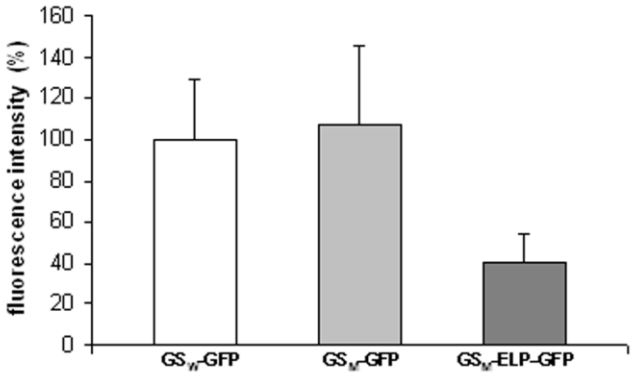
fluorescence intensity (%)
160
140
120
100
80
60
40
20
0
GS$_W$-GFP
GS$_M$-GFP
GS$_M$-ELP-GFP

Figure 5

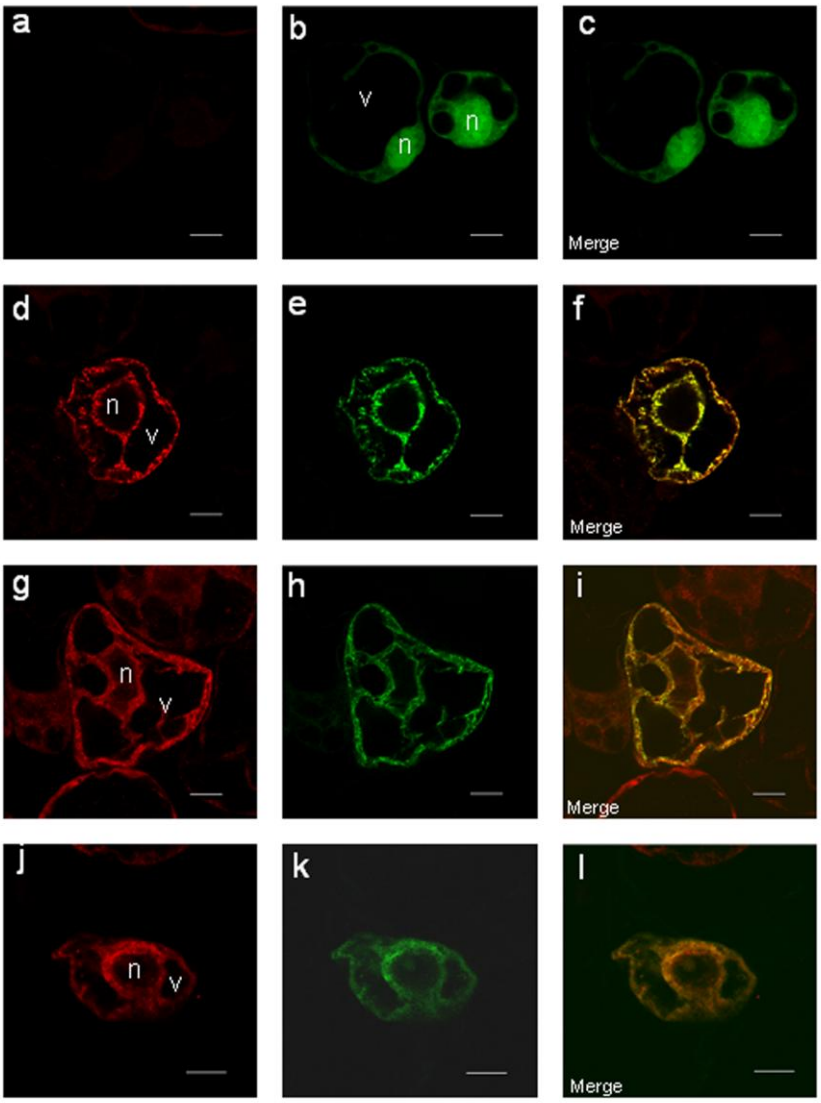

Figure 6

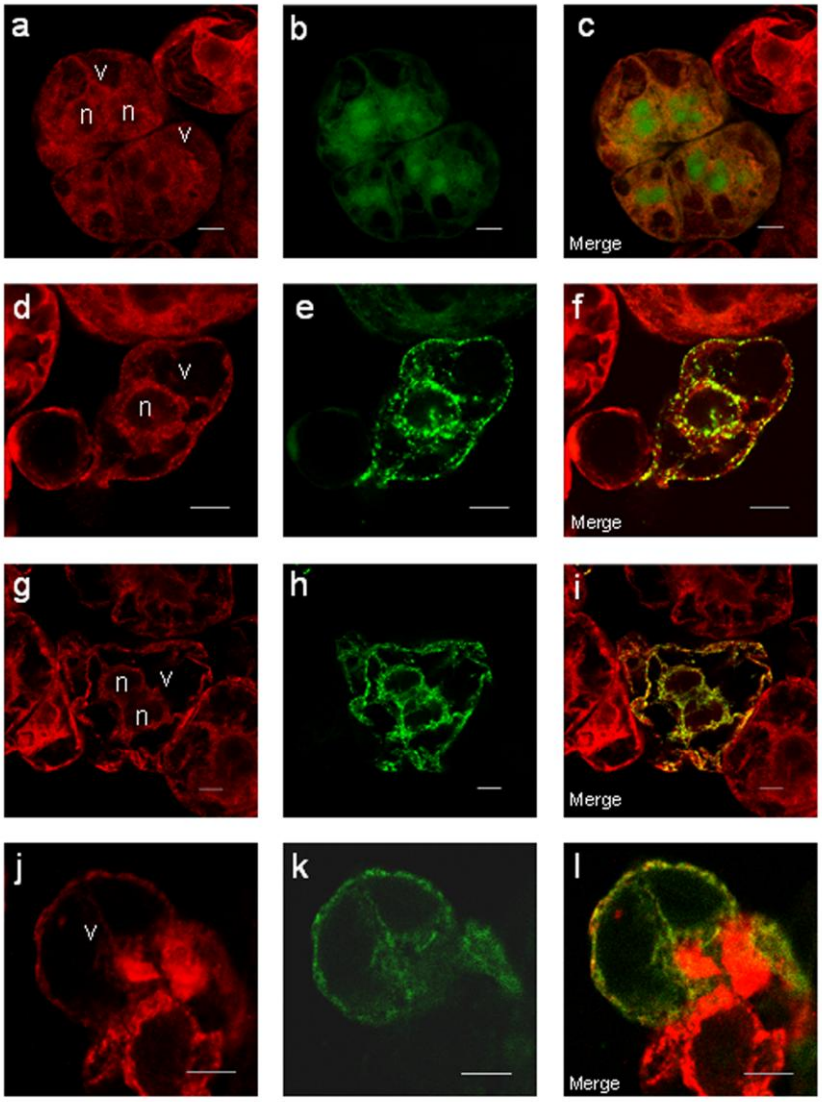

Figure 7