Beak and feather disease virus candidate vaccine development

Lucian Duvenage

This dissertation is submitted in fulfilment of the requirements for the degree of

Master of Science in the Department of Molecular and Cell Biology,

University of Cape Town

September 2012

Abstract

Psittacine beak and feather disease, caused by a circovirus known as beak and feather disease virus (BFDV), is a threat to both wild and captive psittacine species. There is currently no vaccine against BFDV and safe and affordable vaccine candidates are needed to alleviate the disease burden caused by this virus. Production of the BFDV's major antigenic determinant, the capsid protein (CP), in the inexpensive and highly scalable plant expression system, could satisfy these requirements as a potential subunit vaccine. In this work, truncated CP (ΔN40 CP) was first expressed in *E. coli* to successfully generate anti-CP polyclonal antibodies. ΔN40 CP and full-length CP transient expression in tobacco (*Nicotiana benthamiana*) was optimised as fusions to elastin-like polypeptide (ELP). Fusion of CP or ΔN40 CP to ELPs of different lengths was shown to increase yield relative to unfused CP/ΔN40 CP. Free ELP and a GFP-ELP fusion could be purified by inverse transition cycling (ITC), using centrifugation and membrane filtration methods. A ΔN40 CP-ELP fusion expressed in plants could be partially purified and represents low-cost vaccine candidate against BFDV. A candidate DNA vaccine expressing ΔN40 CP was also evaluated for expression of the antigen *in vitro* and may prove useful in a prime-boost regimen together with one of the plant-produced vaccine candidates.

Declaration

I, Lucian Duvenage, declare that this dissertation is my own work.

I know the meaning of plagiarism and declare that all of the work in the document, save for that which is properly acknowledged, is my own.

I have used the Harvard (author, date) convention to cite other sources of information.

_____

Lucian Duvenage

Acknowledgements

I would like to thank the following for their contributions:

- My supervisors Prof. Edward Rybicki, Dr Inga Hitzeroth and Dr Ann Meyers for their advice and support. Dr Meyers provided the GFP-ELP28 clone.
- The members of the vaccine group at the Department of Molecular and Cell Biology (MCB), UCT, for their help with various aspects of this project. The culture collection of this group was the source of all genes and expression vectors. Guy Regnard provided the pPROEX CP, pPROEX-ΔN40 CP and pTH-ΔN40 CP clones.
- Faezah Davids for her help with tissue culture work done at MCB.
- Bruce Allan, Craig Adams and Dr Niki Douglass for their assistance with primary cell work done at the Institute for Infectious Diseases and Molecular Medicine (UCT).
- UCT postgraduate funding office, the National Research Foundation and the Poliomyelitis Research Foundation for providing funding.
- My family, for supporting me for the duration of this MSc

Contents

Abbreviations

ΔN40	truncation of first 40 N-terminal amino acids
ATPS	aqueous two-phase extraction
BFDV	beak and feather disease virus
BSA	bovine serum albumin
CEF	chicken embryonic fibroblast
CAV	chicken anaemia virus
CP	capsid protein
CP-ELP	BFDV capsid protein fused to a C-terminal ELP
DMEM	Dulbecco's modified Eagle medium
DNA	deoxyribonucleic acid
dpi	days post-infiltration
ELP	elastin-like polypeptide
ER	endoplasmic reticulum
FCS	foetal calf (bovine) serum
IB	inclusion body
ITC	inverse transition cycling
His-tag	6 x histidine sequence
cITC	ITC by centrifugation
mITC	ITC by membrane filtration
ORF	open reading frame
MES	4-morpholineethanesulfonic acid

PBFD	psittacine beak and feather disease
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PCV	porcine circovirus
PTGS	post-transcriptional gene silencing
Rep	replication-associated protein
RNA	ribonucleic acid
Rubisco	1,5-ribulose bisphosphate carboxylase oxygenase
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
TSP	total soluble protein
VLP	virus-like particle

Chapter 1: Literature Review

Psittacine beak and feather disease

Birds belonging to the order Psittaciformes include parrots (*Psittacidae*) and cockatoos (*Cacatuidea*) and are collectively referred to as psittacines. Infectious diseases represent the greatest threat to both wild and captive psitaccine species. Of these, psittacine beak and feather disease (PBFD), caused by beak and feather disease virus (BFDV), is one of the most common (Shearer et al., 2008). The symptoms of PBFD consist of feather loss (Figure 1) and abnormalities in feather and beak shape resulting from damage to the epidermis of the feathers and beak (Pass & Perry, 1984).

Figure 1: BFDV infection in the Cape Parrot

A healthy Cape parrot (*Poicephalus robustus)* in the wild is shown in the left panel. Shown in the right panel is another Cape parrot infected with BFDV, showing signs of feather loss, a common symptom of BFDV infection (courtesy of Dr Steve Boyes, personal communication).

Another symptom is acquired immunodeficiency, which often leads to fatal secondary infections (Latimer et al., 1991). BFDV is highly infectious. Viral shedding occurs via the feathers and faeces, as is evident from the presence of inclusion bodies of the virus in the feathers and the cells of the alimentary tract (Jacobson et al., 1986; Latimer et al., 1990). The pathology of PBFD is highly variable depending

on the host. The disease may manifest itself in a peracute form in young African Grey parrots (*Psittacus erithacus erithacus*), for example, which is characterised by sudden illness, leucopenia (depletion of white blood cells) and death (Schoemaker et al., 2000). Other birds may serve as carriers without any detectable symptoms (Khalesi et al., 2005).

Circoviridae and beak and feather disease virus

The family *Circoviridae* is comprised the genera *Gyrovirus*, of which chicken anaemia virus (CAV) is the only species, and *Circovirus,* of which beak and feather disease virus (BFDV) is a member. The viruses of the genus *Circovirus* share similarities with members of the plant virus family *Nanoviridae*, and, to a lesser extent, with those of the plant virus family *Geminiviridae* (ICTV, 2005). Several new circovirus species, infecting a number of avian species, have recently been classified. Among the newest members of genus *Circovirus* are duck circovirus, gull circovirus and finch circovirus (ICTV, 2011). Porcine circovirus (PCV) type 1 and PCV-2 are the only viruses in the genus *Circovirus* which do not infect birds.

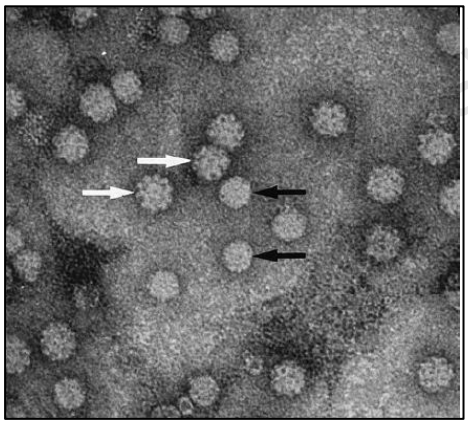


Figure 2: Negative contrast electron microscopy of CAV and BFDV.

"Purified virus preparations were mixed and stained with uranyl acetate. The diameters of CAV and BFDV were estimated to be 26.5 ± 1.2 and 21.7 ± 1.4 nm, respectively. CAV particles (white arrows) exhibit distinctive surface structure in comparison with BFDV particles (black arrows)" (Todd, 2000, with permission).

The circoviruses are the only known viruses infecting animals with a single-stranded DNA genome. These are also the smallest genomes of any virus known to infect animals (Todd, 2000). The single-stranded DNA genome of BFDV is approximately 2000 bases in length, covalently-closed circular and contains two major open reading frames in opposite orientation (De Kloet & De Kloet, 2004). These ORFs encode the replication-associated protein (Rep) and capsid protein (CP). A third ORF has been identified, but its function remains unknown. The *Rep* gene is located on the virion DNA strand and the *CP* gene is located on the complementary strand, once double stranded DNA is formed in the host. The capsid is icosahedral, non-enveloped and approximately 20 nm in diameter (Todd et al., 1991). The surface of the BFDV virion, as viewed by electron microscopy, has a smooth appearance, in contrast to the surface structure of CAV (Figure 2) (Todd, 2000).

Circovirus cell entry and replication

The mechanism of BFDV cell entry has not been studied in detail. The related porcine circovirus type 2 (PCV-2) is known to bind to cell surface glycoproteins heparin sulphate and chondroitin sulphate B. The entry of the virus into PK-15 cells is clathrin-dependent (Nauwynck et al., 2012). Virion disassembly and release of PCV-2 components is thought to occur by acidification within endosomes, but serine proteases also play an important role (Misinzo et al., 2008).

Once inside the cell, the capsid protein transports BFDV DNA and the replication associated protein (Rep) to the nucleus (Heath et al., 2006). The viral single-stranded DNA is converted to double stranded form by host enzymes, after which Rep and CP are transcribed and translated. In circoviruses, Rep protein initiates replication of the viral DNA by rolling circle amplification (RCA), the origin of which lies within a conserved stem-loop structure (Finsterbusch & Mankertz, 2009). The mechanism of circovirus assembly and the export of virions from the cell has not yet been studied.

BFDV detection

Detection of BFDV-specific antibodies in sera from infected birds is done by the haemagglutination-inhibition assay (HI) (Raidal & Cross, 1994), but a successful blocking ELISA has also been developed by Shearer et al. (2009a). Excreted virus in feather and faecal samples can be detected by the haemagglutination assay (HA) and BFDV genes can be detected by PCR in blood or feather samples using primers specific for the *Rep* or *CP* genes (Shearer et al., 2009b). Ideally, combining the results of HI, HA and PCR tests can provide information on the progress of the infection and the immune state of the bird (Khalesi et al., 2005). Rolling circle amplification (RCA) of the entire BFDV genome, followed by restriction enzyme digestion and sequence analysis of cloned fragments can also identify BFDV (Varsani et al., 2010). This is useful for isolating new strains of the virus for which no sequence data is available.

Recombinant expression of BFDV CP and potential BFDV vaccines

The majority of work on circovirus vaccines has focused on PCV-2. A capsid protein subunit vaccine has been evaluated for their immunogenicity in mice (Fan et al., 2008). PCV-2 virus-like particles have been produced in *E. coli* (Yin et al., 2010), as well as a parvovirus chimaeric virus-like particle displaying the neutralising epitope from PCV-2 capsid protein, produced in mammalian cells . The immunogenicity of PCV-2 DNA vaccines have also been tested in mouse models (Shen et al., 2008; Aravindaram et al., 2009)

All attempts to propagate BFDV in cell culture have so far been unsuccessful. Vaccine development therefore relies on recombinant DNA technology to produce the viral antigens in a variety of expression systems. The focus of this approach is production of the major antigenic determinant of the virus, the capsid protein. The capsid protein may be used as a subunit vaccine, or could self-assemble to produce virus-like particles (VLPs). VLP vaccines are typically more immunogenic than their subunit vaccine counterparts, as VLPs mimic the surface structure of the native virus and preserve conformational epitopes (Liu et al., 2012a).

BFDV CP has been expressed using the baculovirus-insect cell system, in which the recombinant capsid protein was observed to self-assemble into VLPs (Stewart et al.,

2007). This BFDV VLP preparation was used to vaccinate long-billed corellas (*Cacatua tenuirostris*) and was shown to prevent PBFD symptoms upon challenge with the virus isolated from infected birds (Bonne et al., 2009). The N-terminal region of BFDV CP contains a nuclear localisation signal that is rich in arginine residues. Deletion of portions of the N-terminus of BFDV CP resulted in increased expression of truncated CP over the full-length CP (Heath et al., 2006). In particular, deletion of the first 40 N-terminal amino acids, resulting in the truncated protein ΔN40 CP, was beneficial to yield.

To date, the expression systems used for production of BFDV CP has been limited to *E. coli* and insect cells. There is no published literature describing the expression of BFDV proteins in plants or other cell culture expression systems. No subunit vaccines consisting of recombinant BFDV CP have yet been tested for the ability to protect against lethal challenge. However, the immunogenicity of PCV-2 CP subunit vaccines suggests that this is a valid approach to vaccination. The immunogenicity of truncated BFDV capsid proteins such as ΔN40 CP relative to the full-length CP has not been investigated, even though antibodies have been raised in chickens against bacterial-expressed ΔN40 CP (Johne et al., 2004). This ΔN40 CP antigen was also used successfully in serological tests to screen for the presence of BFDV-specific antibodies in psittacine sera.

Plants as a production system for biopharmaceuticals including vaccines

The use of proteins is widespread in research, industry and medicine. The production of biopharmaceuticals, which are commercially valuable proteins with important medical applications, is the fastest-growing sector of the pharmaceutical industry (Karg & Kallio, 2009). The use of plants as a system for the production these recombinant proteins is now being recognised as an alternative to traditional animal and bacterial cell production systems (Rybicki et al., 2012). The production costs for plant systems would be substantially lower, owing to the simple growth requirements of whole transgenic plants grown in the field. Up-scaling this process simply requires more land. Plants perform all the necessary post-translational modifications (including N- and O-glycosylation, phosphorylation and disulfide bridge

formation) to produce functional proteins (Karg & Kallio, 2009). The risk of contamination with human or animal pathogens is avoided with plant-produced proteins. Recombinant proteins produced in plants include hormones, enzymes, monoclonal antibodies and vaccine antigens (Sharma & Sharma, 2009). The major capsid protein (L1) of human papillomavirus type 16, which was transiently expressed in tobacco, is one example of a successfully plant-produced vaccine antigen (Varsani et al., 2006; Maclean et al., 2007)

Important considerations when choosing the host plant for expression include the total protein content, biomass yield per acre and the ease of transformation. It is desirable if the recombinant protein can be expressed in a plant whose native proteins differ in their physiochemical properties. For example, the majority of native proteins in tobacco leaves are acidic (Balasubramaniam et al., 2003). The extraction of a basic recombinant protein under acidic conditions will therefore be contaminated with fewer proteins, simplifying the purification. Leafy crops such as tobacco, lettuce and alfalfa have been used for recombinant protein production. One drawback of using these is that the protein is unstable in the aqueous leaf environment due to protease activity (Benchabane et al., 2008). This necessitates immediate processing of crops after harvesting. In contrast, seeds are desiccated and the proteins within may remain stable for long periods at room temperature (Fischer & Emans, 2000). In addition, seeds do not contain the same level of phenolic compounds found in leaves (Stoger et al., 2005). Tissue-specific accumulation of the recombinant protein is achieved by placing the transgene under the control of tissue-specific promoters. Several promoters can be used to target expression of the protein to different parts of seeds. For example, the rice glutelin promoter *Gt-1* restricts expression to the endosperm of monocot seeds (Stoger et al., 2005). Cereal crops such as maize, rice, barley and wheat have been investigated for the production of recombinant proteins (Fischer & Emans 2000).

Tobacco (*Nicotiana benthamiana or N. tabacum* species) has been the most widely used as a model system for the production of recombinant proteins at the laboratory research scale. The genetic transformation of tobacco is relatively efficient and the gene-transfer technology is well-established. Toxic alkaloids and phenolic

compounds are present in the leaves that can interfere with the purification, although there are low-alkaloid cultivars that may be used (Fischer & Emans, 2000). Tobacco has a high biomass yield (up to 1000 tonnes of leaf biomass per hectare for close-cropped tobacco) and a high soluble protein content compared to other leafy crops (Tremblay et al., 2010). Because tobacco is a non-food crop, the risk of food or feed contamination with any recombinant proteins is avoided. These traits make tobacco a suitable candidate for commercial production of recombinant proteins. Tobacco leaves must be processed immediately after harvesting, or kept in cold storage, to minimise protease activity.

The production of recombinant proteins by rhizosecretion is another alternative that shares several advantages with plant cell culture. Plant roots are infected by *Agrobacterium rhizogenes* and tissues at the site of infection become differentiated into so-called hairy roots. The plants are cultivated in a hydroponic system and can secrete the protein from the roots (Borisjuk et al. 1999). The purification of the protein is simplified and does not require harvesting of the plant.

Plant cell cultures have been used for recombinant protein production. The production times are shorter than for whole plants. Plant cell culture requires simple media, but the operating costs are still significantly greater than for agriculture. The potential of scale-up is also reduced. However, if the heterologous protein is secreted into the culture medium, the purification of proteins from the culture media is far simpler and cheaper than from homogenised biomass (Shih & Doran, 2009). Glucocerebrosidase, used in therapy for Gaucher disease, is being produced in transgenic carrot cell suspension by Protalix (Israel) and has passed phase 3 clinical trials (Yusibov et al., 2011).

Transient and transgenic expression in plants

Expression of recombinant proteins in whole plants may be achieved by two means: generation of transgenic plants with the gene of interest stably integrated into the plant genome, or transient expression of the desired protein by means of *Agrobacterium* infiltration. The creation of homozygous transgenic plant lines requires several months. Transformation in the case of dicotyledonous plants such as tobacco is most often achieved through *Agrobacterium* gene transfer. Monocotyledons such as maize and rice, which are resistant to *Agrobacterium* gene transfer, are typically transformed by biolistic methods (Lowe et al., 2009). However, once transgenic lines are established, the production of recombinant protein is highly scalable. The production capacity would be limited only by the available acreage. In practice, transgenic plants are most often contained within greenhouses to comply with regulations. A problem that is often observed in transgenic plants is silencing the transgene in later generations, especially if multiple copies of the transgene have integrated into the genome. The epigenetic phenomena behind these types of gene silencing is an active area of research (Reddy et al., 2003; Lowe et al., 2009).

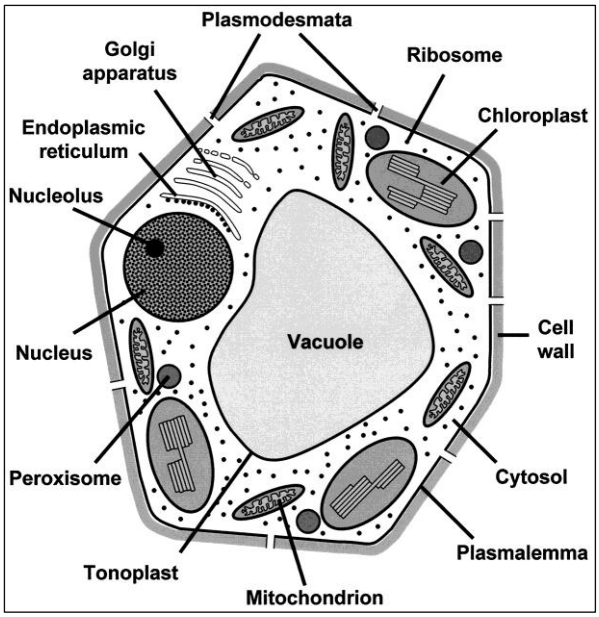
Transient expression has a much faster development timescale, because it does not require stable transformation of plants. Rather, recombinant *Agrobacterium* deliver DNA into plant cells for expression of the protein of interest over a period of several days. In the case of tobacco leaves, the *Agrobacterium* cell suspension is infiltrated into the air spaces surrounding mesophyll cells, using a needleless syringe for small-scale confirmation of protein expression. This process is termed Agroinfiltration. For larger-scale operations, plant leaves may be infiltrated by placing the whole plant in an *Agrobacterium* suspension under a reversible vacuum, or by spraying *Agrobacterium* directly onto leaves (Marillonnet et al., 2005). Although transient expression may potentially be less scalable than using field-grown transgenic plants, facilities have been established that carry out Agroinfiltration on a large scale, capable of processing several tonnes of biomass by vacuum infiltration per hour (Yusibov et al., 2011).

A major factor influencing the yield of plant-produced recombinant proteins is post-transcriptional gene silencing (PTGS). This is a form of ‘adaptive immunity’

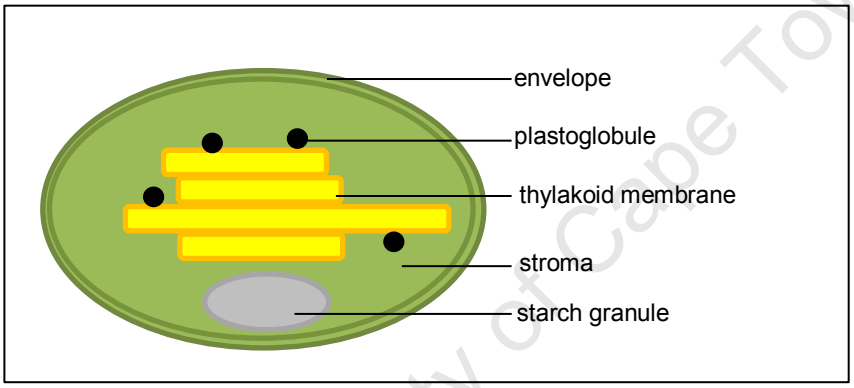
conserved amongst eukaryotes, including plants. The vast majority of viruses infecting plants have a single-stranded RNA genome, and they replicate through a double-stranded RNA intermediate. The recognition of double stranded RNA by the plant's PTGS response is followed by the subsequent destruction of RNA sharing sequence homology to the foreign RNA (Scholthof, 2007). By the same means, mRNA transcripts of the recombinant gene of interest are similarly destroyed upon induction of PTGS, leading to a lower protein yield. Plant viruses have overcome the effects of PTGS by producing silencing suppressor proteins that interfere with various stages of the PTGS pathway. The genes encoding silencing suppressors can be co-infiltrated with the gene of interest into plants for increased transient protein expression (Voinnet et al., 2003). Plants stably transformed with silencing suppressor genes typically exhibit negative phenotypic effects (Siddiqui et al., 2008). However, Saxena et al. (2011) stably transformed *N. benthamiana* with a mutant of the silencing suppressor p19 (from tomato bushy stunt virus) that had only minor developmental effects on the plant.

Sub-cellular protein targeting

The recombinant protein may be targeted to a specific part of the plant cell to increase its stability and accumulation (Figure 3). The choice of subcellular compartment will influence the protein yield as well as the purification strategy employed. Secretion is the default pathway for proteins entering the endomembrane system. In the absence of any further targeting information, proteins containing an N-terminal signal sequence are secreted by the cell, causing the protein to accumulate in the apoplastic space (Ma et al., 2003). Non-secretory proteins lacking this sequence are retained in the cytosol. Retention of recombinant proteins in the cytosol is generally associated with lower yields for most proteins, compared to other subcellular compartments such as the endoplasmic reticulum (ER) (Fischer and Emans, 2000). Proteins containing the N-terminal signal sequence, in addition to the C-terminal tetrapeptides KDEL or HDEL accumulate in the ER. The ER is a protease-free environment and contains chaperone proteins to assist in the correct folding of proteins. Proteins may accumulate to high levels in the ER of plant cells without adversely affecting cell function or development (Sharma & Sharma 2009).



A



B

Figure 3: Representation of a typical plant cell with its subcellular compartments

A: Representation of a typical photosynthetic plant cell with its subcellular compartments, including the nucleus, ER and chloroplasts (Lunn, 2007). **B**: Representation of a chloroplast, showing the association of plastoglobules with the thylakoid membrane.

There are as many as 100 chloroplasts in a plant cell and each one contains about 100 copies of the chloroplast genome (Sharma and Sharma, 2009). Transformation of the chloroplast genome therefore will allow thousands of copies of the transgene to be present in each cell. This leads to higher levels of expression and production of the recombinant protein. Plant chloroplast genomes may be stably transformed, generating transplastomic plant lines. Transplastomic plants have an additional advantage of biocontainment; because plastids like chloroplasts are maternally inherited, there is no spread of the transgene through pollen (Lössl & Waheed, 2011). The transformation of the chloroplast genome also avoids the problem of position effect often encountered with nuclear transformation because the site of

gene insertion can be controlled (Dufourmantel et al., 2004). There is one limitation to transplastomic expression, in that the translational machinery of the chloroplast is similar to that of prokaryotes. Therefore proteins requiring post-translational modifications such as glycosylation are not suited for expression in the choloroplast (Ma et al., 2003). Recombinant proteins expressed transiently by Agroinfiltration may be targeted for accumulation in the chloroplast by a N-terminal chloroplast transit peptide (Rybicki, 2010). Such a protein is produced by translation in the cytoplasm and is then imported into the chloroplast, where it may accumulate to high levels. Using this strategy, Maclean et al. (2007) expressed human papillomavirus protein L1 in tobacco plants, with a yield of 11% of the total soluble protein.

Targeting of proteins to the chloroplast has so far mostly been limited to the stroma, (Figure 3B). Recently, plastoglobules have been evaluated as sites for protein accumulation (Vidi et al., 2007). Plastoglobules are low-density lipid-containing stuctures present in all plastids. In chloroplasts they are associated mainly with the thylakoid membranes (Figure 3B). A recent study of the proteome of *A. thaliana* plastoglobules identified several proteins of the fibrillin family, specific to plastoglobules (Ytterberg et al., 2006). Vidi et al. (2007) targeted the accumulation of yellow fluorescent protein to plastoglobules by fusion to one of these fibrillin proteins. The plastoglobules were isolated by flotation centrifugation. This system simplifies purification from plant leaves and results in efficient enrichment of the target protein in a single step.

Protein recovery and purification from plants

The first phase of downstream processing is protein recovery. This involves processing of the plant biomass, solid-liquid separation, protein extraction and concentration (Kusnadi et al. 1997). The initial recovery steps are largely dependent on the host plant. After harvesting, the tissues of the plant expressing the recombinant protein (e.g. seeds or leaves) are separated from the rest of the plant. Grinding is required to reduce particle size. This breaks cell walls and releases intracellular components, in preparation for protein extraction. For leaves, grinding and protein extraction may be performed simultaneously by wet grinding (Menkhaus

et al., 2004). This process is usually not suitable for seed crops, which must be dry ground and subsequently mixed with aqueous extraction buffer.

The ionic strength and especially the pH of extraction buffer are important in determining the amount of extracted protein. In one protein extraction study, the amount of recombinant lysozyme extracted from tobacco leaves varied from 1.0% to 1.6% (w/w) for extraction buffers of pH 3 to 9 (Balasubramaniam et al., 2003). High concentrations of secondary metabolites can be detrimental to the purification. Of these, phenolic compounds (occurring in high concentrations particularly in tobacco) are the most problematic. They interact with proteins and can cause aggregation of proteins or changes in protein conformation (Farinas et al., 2005). The level of phenolic contamination can be minimized by changing the pH or ionic strength of the extraction buffer. The degradation of proteins by host proteases is of major concern, especially for protein extracted from leaf biosmass. Protease inhibitors such as phenylmethanesulphonyl fluoride (PMSF) are usually added to prevent degradation of the protein (Kusnadi et al. 1997). Clarification of the protein extract is then performed. The residual solids are removed by centrifugation or other methods such as rotary vacuum filtration (Evangelista et al., 1998). The clarified protein extract may be purified by various methods, including several types of chromatography. However, further clarification of the protein extract by dead-end filtration or centrifugation steps is often necessary before it may be applied to a chromatography column.

As a pre-purification step, the use of aqueous two-phase extraction (ATPE) has been applied to the recovery of proteins from tobacco (Platis & Labrou, 2006). When polymers such as polyethylene glycol (PEG) are mixed with salt solutions of a particular concentration, two immiscible layers are formed. The partitioning of proteins between the two phases is based on their charge and hydrophobicity. In one study using ATPE, lysozyme in tobacco leaf extracts was concentrated by a factor of 14 and was partially purified (enrichment factor of 4) (Balasubramaniam et al., 2003). In addition, it was found that ATPE can be used to effectively remove phenolic compounds and alkaloids, which interfere with the purification (Platis & Labrou,

2006). ATPE can be applied cost-effectively on a large scale and reduces the load on subsequent purification steps.

The final purification steps will depend on the recombinant protein and the host plant. Ion-exchange chromatography is often used because of its low cost, high binding capacity and reusability of the resin (Platis et al., 2008). Affinity-based purification is a powerful purification technique. Fusion proteins can be purified to near homogeneity in a single step and at a high yield – often over 90% (Arnau et al., 2006). It is widely used and invaluable in laboratory-scale research. Affinity chromatography is typically not used for protein purification on an industrial scale because of the high cost of affinity resins and loss of the bound ligand from the column with repeated use (Platis and Labrou, 2006). Therapeutic proteins most often require removal of the affinity tag which requires expensive proteases. For the purification of lower-value products, the costs of affinity chromatography will outweigh the reduction in cost by using plant expression systems. Therefore, more cost-effective and technically simple protein purification methods are needed that can be easily scaled up.

Non-chromatographic protein purification

Highly efficient extraction and purification processes are essential for recombinant protein production, especially if those proteins have pharmaceutical applications. Such proteins would need to be 95-98% pure (Menkhaus et al., 2004). Most of the total cost is incurred by downstream processing (Evangelista et al., 1998). In some cases, the cost of downstream processing alone may determine if a particular plant production system is economically feasible. Factors such as the choice of crop and the targeting of recombinant protein accumulation to specific tissues or sub-cellular compartments will influence the design of the purification process. Currently, chromatography is part of most purification schemes. The use of affinity fusions and affinity chromatography, although highly effective as research tools, may be unsuitable for large-scale purifications because of the high cost of affinity resins (Nilsson et al., 1997). The success of large-scale production will therefore depend on robust and inexpensive purification methods. The extraction of protein from plant leaves on the industrial scale produces very large volumes of crude extract.

Chromatography is limited in its scalability and the inevitable fouling of chromatography resin by volumes of this magnitude would increase the cost of the purification process. Non-chromatographic purification methods are therefore very attractive for the purification of recombinant proteins from plants. Some of the methods described here have been developed in recent years, but not all of them may yet be suitable for use at the industrial scale.

Oleosin fusions

Oil bodies are ER-derived organelles that store oils in plant seeds. They are composed of a matrix of triacylglycerols bounded by a phosopholipid monolayer. Embedded on the surface are proteins unique to oil bodies, such as oleosins. Oleosins have a central hydrophobic domain that associates with the interior of the oil body and amphipathic N- and C-termini that are exposed to the cytosol (Bhatla et al., 2010). The target protein may be fused to the N- or C-terminus of the oleosin, separated by a linker peptide containing an endoprotease recognition sequence. This fusion does not affect the targeting of oleosin to oil bodies. The heterologous protein is tightly associated to the oil body and is exposed to the cytosol. The oil bodies are low-density and are easily separated from other cell components by centrifugation methods. The protein may then be cleaved from its oleosin partner by digestion with the specific endoprotease. This method has been used to produce and purify hirudin, a blood anticoagulant protein, in *Brassica carinata* seeds (Chaudhary et al., 1998) and biologically functional insulin in *Arabidopsis thaliana* seeds (Nykiforuk et al., 2006).

Zein fusions

Zeins are the major storage proteins of maize endosperm cells. The formation of accretions known as protein bodies is dependent on the presence of γ-zein (Coleman et al., 1996). A sequence derived from the proline-rich, N-terminal domain of γ-zein, can be fused to the target protein. This results in aggregation of the fusion protein to form protein bodies, which are retained in membrane-bound storage organelles derived from the ER (Torrent et al., 2009). This sequence, known as the Zera® domain (Torrent et al., 2009), has been shown to induce protein body formation in several eukaryotic expression systems, including plants and mammalian

cultured cells. Another study using a similar γ-zein sequence fused to bean (*Phaseolus vulgaris*) phaseolin concluded that the presence of this sequence alone in the fusion is sufficient for protein body formation (Mainieri et al., 2004). The Zera® domain contains a repetitive sequence $(PPPVHL)_8$ which was found to self-assemble *in vitro* (Kogan et al., 2002). Protein bodies are insoluble and can be recovered from crude cell or tissue extracts using density gradient centrifugation (Mainieri et al., 2004). Sequestration of the heterologous protein in protein bodies will isolate it from physiological processes of the cell. The benefits of this include the protection of the protein from degradation and the prevention of its possible toxic effects on the host. The recovery of Zera®-tagged proteins is simplified as the protein storage organelles are recovered in a specific fraction following density gradient centrifugation. The organelle membrane is dissolved to release the fusion protein. The protein bodies are then de-aggregated under mild reducing conditions (Torrent et al., 2009) and the Zera® tag is separated from the target protein by protease cleavage. The Zera® fusion strategy allows for simultaneous purification and concentration of the recombinant protein.

Hydrophobin fusions

The formation of protein bodies allows for higher accumulation of the recombinant protein plant cells. Recently it has been shown that fusion of the protein of interest to hydrophobin and expression in plants induces the formation of protein bodies in the ER (Joensuu et al., 2010), as with the Zera® tag. Hydrophobins are small (12-17 kDa) proteins produced by filamentous fungi. These proteins have a conserved hydrophobic “patch” that is responsible for their amphipathic nature, and they therefore have a propensity to form protein membranes at hydrophobic-hydrophilic interfaces (Wang et al., 2005). Hydrophobin fusions can be purified by an aqueous two-phase separation process with the addition of a specific surfactant. The hydrophobin fusions form micelles that are concentrated in the surfactant phase, but absent from the aqueous phase containing the majority of contaminating proteins (Mustalahti et al., 2011). This process is simple and highly scalable. The utility of hydrophobin fusions has been evaluated for the system of transient expression in tobacco plants for green fluorescent protein and Glc oxidase enzyme (Joensuu et al., 2010).

Elastin-like polypeptide (ELP) fusions

Elastin is a component of the extracellular matrix of vertebrate connective tissue. It imparts elasticity and strength to tissues that are required to deform reversibly, such as lung and skin (Floss et al., 2010b). Elastin is deposited in elastic fibres by the complex process known as elastogenesis. Tropoelastin is the soluble precursor of elastin. It is secreted into the extracellular space, where the action of lysyl oxidase forms covalent cross-links between tropoelastin molecules. This forms the mature elastin, which is an insoluble polymer (Debelle & Tamburro, 1999). Tropoelastin contains large hydrophobic domains that are composed mainly of proline, alanine, valine, leucine, isoleucine and glycine residues (Debelle & Tamburro, 1999). These residues occur as tetra-, penta- and hexa-peptide repeats. The physical properties of these repeat sequences have been studied using synthetic polypeptides such as poly-(VPGXG) (X is referred to as the "guest residue", being any amino acid except proline) (Li et al., 2001).

Polypeptides based on the structure of elastin and its precursors, such as poly-(VPGXG), are called elastin-like polypeptides (ELPs). These peptides are soluble in aqueous solution below a specific temperature, known as the inverse transition temperature (T_t). When the temperature is raised above T_t the peptides are desolvated and aggregate, making them insoluble (Meyer & Chilkoti, 1999). This is termed the inverse phase transition and occurs over a narrow temperature range (2-3°C) and is reversible. The process can be monitored by measuring the turbidity of the solution at 350 nm in real time. At high ELP concentrations, the increase in turbidity upon heating is visible to the naked eye, as is the decrease in turbidity when the ELP solution is again cooled to below T_t. The T_t is influenced by the length (i.e. number of repeat units) of the ELP; longer polypeptides have a lower T_t (Urry, 1997). The composition of guest residues in the ELP also influences T_t, as do factors such as the ionic strength and pH of the solution (Urry, 1997). It should be noted that ELPs can be designed in regard to length and composition of guest residues to give a desired T_t under specific conditions. ELP-tagged proteins are purified in a process known as inverse transition cycling (ITC) (Figure 4) (Meyer & Chilkoti, 1999). The tagged protein is first separated from other contaminants in the aqueous cell or

tissue extract by heating to above the T_t. Typically the T_t is first depressed by the addition of salt. This allows the ELP aggregation to occur at temperatures that do not denature the target protein; usually 30-40 °C. Once the protein solution is heated to above the transition temperature, selective aggregation of the ELP-tagged protein occurs. The precipitated ELP fusion is removed by centrifugation or microfiltration (Ge et al., 2006). The fusion protein aggregates are then resolubilised by cooling to below the T_t.

The concentration of ELP is also an important factor in determining the T_t and the efficacy of ITC: with more concentrated ELP protein solutions, the recovery of the ELP or ELP fusion is more efficient. In order to purify poorly expressed ELP fusions, free ELP protein – with a T_t matching that of the fusion protein – may be added in order to increase the overall ELP concentration. The free ELP co-aggregates with the fusion protein during ITC and improves its recovery. Using this method, Christensen et al. (2007) purified picomolar quantities of a thioredoxin ELP fusion from *E. coli* cell lysate.

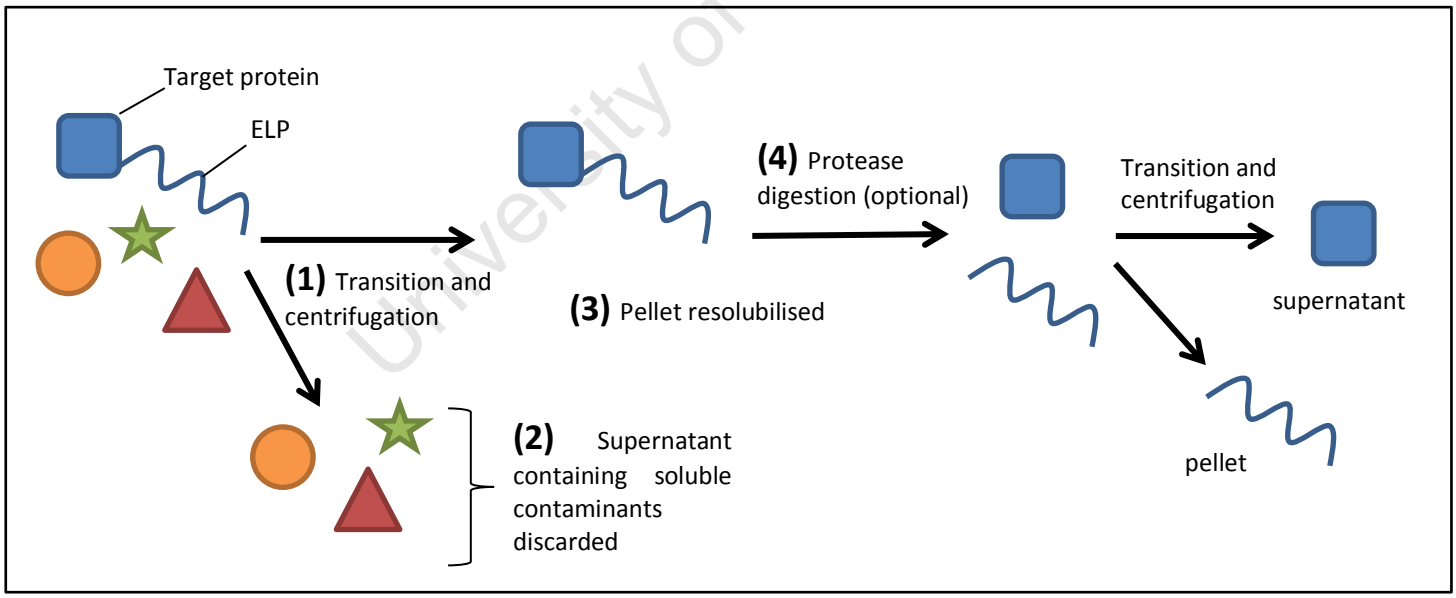


Figure 4: Schematic of inverse transition cycling

1) The ELP fusion aggregates specifically above its transition temperature, allowing it to be pelleted by centrifugation at elevated temperature.
2) The supernatant containing soluble contaminating proteins is discarded.
3) The purified ELP pellet is resolubilised in cold buffer. Steps 1 to 3 can be repeated several times to increase purity.
4) The ELP can be separated from the target protein by protease digestion. The target protein is then recovered by removal of the ELP tag by ITC.

At this point the ELP tag may be removed from the target protein by endoprotease digestion, which may be necessary for proteins with therapeutic applications. This requires the insertion of a specific protease recognition site between the ELP and the target protein. The use of ELP-tagged proteases has been recently described by Lan et al. (2011) for the removal of proteases by ITC once digestion is complete. However, proteases are expensive and may also exhibit non-specific cleavage. As an alternative to proteases, an intein sequence may be introduced between the target protein and the ELP tag. The intein is a self-cleaving protein element that will allow the separation of the ELP tag from the target protein by a mild pH shift, or the addition of reducing agents, depending on the specific intein used (Wu et al., 2006). The cleaved ELP tag is precipitated by another round of ITC. After centrifugation, the pellet containing the precipitated ELPs is discarded and the target protein is recovered in the supernatant. Purification using the ELP tag system, as an alternative to chromatography, is more economical and can be implemented on a large scale.

Apart from cheaper purification, the use of ELP fusions can increase the yield of the recombinant protein. The yields of several heterologous proteins, including human interleukin-10, expressed in tobacco leaves as ELP fusions were higher than that of the un-fused proteins (Patel et al., 2007). The ELP used was a 27-mer of VPGVG. In another study, the accumulation of single-chain variable fragments in tobacco seeds was increased 40-fold when expressed as an ELP fusion (Scheller et al., 2006). Here the ELP was a 100-mer. The yield of a thioredoxin-ELP fusion expressed in *Escherichia coli* was greater when a shorter ELP tag (20-mer) is used as opposed to a longer one (90-mer). Shortening the ELP tag did not compromise its ability to be purified by ITC, but the phase transition behavior was more complex. The optimal length of the ELP may vary for different target proteins (Meyer et al., 2001). The presence of the ELP tag can lead to the formation of novel protein bodies, as investigated by Conley et al. (2009b). Tobacco leaves (*Nicotiana benthamiana)* were agroinfiltrated, resulting in transient expression of green fluorescent protein (GFP). When the expression of a GFP-ELP fusion was targeted to the endoplasmic reticulum by the C-terminal KDEL tetrapeptide sequence, spherical particles

resembling protein bodies were observed by confocal microscopy. The formation of these protein bodies allows for higher accumulation of the target protein as well as protecting it from proteolytic degradation.

For some proteins expressed in *E. coli,* fusion to ELP may decrease the yield of the target protein, especially for high molecular weight ELPs (Trabbic-Carlson et al., 2004a). To harness the purification potential of ELP, affinity capture methods have been designed that allow protein purification without direct fusion of ELP to the protein of interest. The ELP is expressed as a fusion to an affinity capture partner that binds to the protein of interest. The target protein binds to its partner and then ITC may be used to purify the entire complex. The ELP fusion may then be re-used to purify more target protein. The success of this method has been demonstrated for the purification of antibodies (Kim et al., 2005). The ELP affinity capture system is extremely versatile and can be implemented for any protein with fusion to a small affinity tag, as has been done for various dockerin-tagged proteins expressed in *E. coli* (Liu et al., 2012b).

The applications of ELPs are not limited to recombinant protein purification. By modifying the pentapeptide repeat sequence of ELP and by constructing ELP block co-polymers, a variety of structures can be formed including membranes, micelles and hydrogels (MacEwan & Chilkoti, 2010). Polymers derived from elastin have been used in the construction of nanoparticles for drug delivery. Macromolecular drug carriers have a longer half-life in plasma, as they are more slowly cleared by the kidneys than free drugs, and allow for controlled release of the drug over an extended period (Altunbas & Pochan, 2012). Due to the environmental responsiveness of ELP, concentrations of the drug-ELP conjugate can be increased by ELP aggregation at an area of local hyperthermia, for example at a tumour site (McDaniel et al., 2012). Micelles displaying protein on their surface can be formed through temperature-dependent aggregation of ELP block copolymers (Hassouneh et al., 2012). This could have great potential for vaccine development, as multivalent antigen display could potentially increase immunogenicity, which is true in case of virus-like particle (VLP) vaccines.

ELPs have found applications in tissue engineering by acting as scaffolds for cell attachment and proliferation. In one study, an ELP membrane was used to allow for reversible attachment of intact cell sheets (Mie et al., 2008). ELP has also been used for the coating of synthetic cardiovascular devices to improve biocompatibility (Woodhouse et al., 2004).

In summary, ELP fusions may be especially useful for the expression and purification plant-produced recombinant proteins. Plant systems, which typically exhibit lower recombinant protein yields compared to other expression systems, can greatly benefit from the increase in protein yield by fusion to ELP. The most expensive step in recombinant protein production – purification – may be reduced in cost by avoiding the use of chromatography resins. The highly scalable ELP purification system is able to cope with the massive volumes of plant biomass extract that would be generated for industrial recombinant protein production in plants such as tobacco.

Project aims and outline

There is currently no vaccine against beak and feather disease virus and any vaccine candidate should be safe and affordable. The aim of this work was to produce beak and feather disease virus vaccine candidates using different approaches. Firstly, ΔN40 CP was produced in bacteria, using the pPROEX expression system. Bacterial over-expression, with its typically high recombinant protein yields and short timescale, was used to produce ΔN40 CP as a potential subunit vaccine. ELP fusions of ΔN40 CP were also expressed in order to characterise the purification conditions of these ELP fusions by ITC. A secondary aim for bacterial expression was to generate enough CP antigens for the production of CP-specific antibodies. These antibodies are not commercially available and would facilitate the detection of BFDV CP produced in a variety of expression systems by our laboratory.

As described in Chapter 3, the principal aim of this project was to produce candidate BFDV CP subunit vaccines expressed in tobacco leaves (*N. benthamiana*) by

transient expression. Plant expression was chosen because of its low cost and safety. Fusion of BFDV CP to ELPs of different lengths was done to increase its accumulation in plants, and to provide an inexpensive and scalable means of purification. GFP-ELP fusions and a free ELP were also produced using this system. The ITC purification system was then used to purify ELP fusions. Due to their high expression levels, the GFP-ELP51 fusion and the free ELP51 were first purified in order to test the robustness of the ITC system. Two ITC methods, one based on centrifugation and the other on membrane filtration, were compared. This was followed by purification experiments of a truncated CP-ELP fusion (ΔN40 CP-ELP51) which are described in Chapter 4.

As an alternative to a subunit vaccine candidate, a eukaryotic expression vector containing the ΔN40 CP gene was evaluated as a potential DNA vaccine. The capacity of this candidate DNA vaccine to express ΔN40 CP in transfected cells was tested for two cell types, human embryonic kidney cells and chicken embryonic fibroblast primary cells. Proof of truncated CP expression was required in order to proceed with future animal trials. This DNA vaccine could potentially be used in a low-cost 'prime-boost' vaccine regimen together with one of the plant-expressed vaccine candidates.

Chapter 2: Expression of BFDV capsid protein in *E. coli*

2.1 Introduction

Escherichia coli is the most used organism for the expression of recombinant proteins, because of its extensive genetic characterisation and availability of mutant strains. Various promoter systems have been developed in order to maximise the expression of heterologous proteins (for a review, see Terpe (2006)). Among them, the synthetic *trc* promoter, derived from the *E. coli trp* and *lac*UV5 promoters, is commonly used for its ability to drive high-level transcription of genes (Tegel et al., 2011). The expression vector used in this chapter, pPROEX-HT (Life Technologies, USA), contains the *trc* promoter and lacI gene, which is induced by isopropyl-beta-d-thiogalactopyranoside (IPTG). In order to achieve maximum protein yield and solubility, expression must be optimised for individual proteins, using different promoters, bacterial strains and culture conditions.

All attempts to propagate BFDV in cell culture have so far been unsuccessful. BFDV proteins must therefore be produced by expression of the recombinant DNA in hosts such as bacteria, or eukaryotic systems. As the capsid protein (CP) is the major antigenic determinant of the virus, efforts for BFDV vaccine development are focused on the production of CP. In addition, CP can be useful for diagnosis of BFDV infection when used as an antigen in ELISA-based detection methods. The N-terminus of circovirus capsid proteins such as BFDV CP and PCV-2 CP is rich in basic amino acids that forms part of the nuclear localisation signal (Heath et al., 2006). The arginine codons AGG and AGA are rare in prokaryotes, but are the most commonly used ones in eukaryotes (Alexandrova et al., 1995). Inefficient translation of these codons occurs in prokaryotes due to a lack of the corresponding tRNA species. Problems arising from codon bias include frameshift events and premature termination of translation (Sørensen & Mortensen, 2005). Groups of these codons may also compete with the native Shine-Dalgarno sequence, thereby interfering with

translation (Alexandrova et al., 1995). Codon optimisation of the PCV-2 CP sequence by the replacement of rare arginine codons results in a marked increase in PCV-2 capsid protein expression in *E. coli* (Marcekova et al., 2009). Deletion of sequence in the BFDV CP gene encoding the first 40 N-terminal amino acids resulted in high-level expression of truncated BFDV CP (ΔN40 CP), while full-length CP could not be detected when expressed in *E. coli* (Johne et al., 2004).

The expression of BFDV ΔN40 CP proteins in *E. coli* as described in this chapter was done primarily to acquire enough protein for anti-BFDV ΔN40 CP antibody production. A truncated BFDV CP (ΔN40 CP) gene was previously generated in our laboratory and showed higher yields relative to full-length CP when expressed in insect cells (Heath et al., 2006). Therefore this ΔN40 CP was used in the bacterial expression described here, with the aim of maximising protein yield. Each protein was injected into a rabbit and serum was harvested as described in 2.2.7. No commercial anti-BFDV CP is available. The serum would facilitate the detection of BFDV CP proteins produced by different expression systems, including plants. ELP fusions of ΔN40 CP were created with the aim of characterising the purification conditions by ITC (see Chapter 4). Expression of recombinant proteins in *E. coli* is rapid and inexpensive. The typically high yield of recombinant protein over-expressed in microbial systems, compared to eukaryotic systems, makes this option attractive for the rapid characterisation of proteins, including its physical properties.

2.2 Materials and Methods

2.2.1 Genes and Plasmids

Full-length and truncated (40 amino acids from N-terminus) beak and feather disease virus coat protein genes (BFDV CP, Genbank accession number AY450451.1) were both maintained in pGEM®-T Easy vectors (Promega, USA). The full-length capsid protein (UniprotKB accession number Q9YUC8) is designated BFDV CP and the truncated version is designated ΔN40 CP. These were obtained from the culture collection of the vaccine group at the Department of Molecular and Cell Biology, University of Cape Town, South Africa. The ELP28 sequence,

previously synthesized by GENEART (Germany), was maintained in the vector pMK-TEV-ELP (Appendix C). The binary plant expression vector pTRAkc-ERH was provided by Rainer Fischer of the Fraunhofer Institute for Molecular Biology and Applied Ecology, IME, Germany. The pPROEX-HTb bacterial expression vector was obtained from Qiagen (Qiagen, USA) (Appendix B). The antibiotic selection for each of the plasmids is summarised in Table 1.

Table 1: Antibiotic selection of plasmids

Plasmid	**Antibiotic and concentration**
pGEM®-T Easy	Ampicillin, 100 µg/ml
pMK-TEV-ELP	Kanamycin, 50 µg/ml
pTRAkc-ERH, pTRA-ERH-ELP	Ampicillin, 100 µg/ml
pPROEX-HTb	Ampicillin, 100 µg/ml

2.2.2 Construction of pTRA-ERH-ELP28

The PCR product consisting of the TEV protease recognition site, ELP28 sequence comprising 28 VPGVG repeats, His tag and SEKDEL sequences was amplified using the primers TEV-ELP For and ELP-HK Rev listed in Table 2, from the template pMK-TEV-ELP (Appendix C). This PCR product was digested with *NcoI* and *XhoI* and cloned into pTRAkc-ERH (Appendix B) to generate the vector termed pTRA-pTRA-ERH-ELP28, using procedures described in Appendix A. This vector creates C-terminal ELP28 fusions of genes cloned into it using the *Nco*I and *Not*I sites.

Table 2: Oligonucleotide primers used in PCR.

'For' designates a forward primer. 'Rev' designates a reverse primer. Restriction enzyme sites are underlined.

Primer name	Primer sequence	Features
*Afl*III-BFDVN40 For	5'-GGAACGCGTTAGGTACATGT TCTCAACCAATAGAATTTACAC-3'	*Afl*III site
*Not*I-BFDVN40 Rev	5'-GAGCGGCCGCAGTACTTGGA TTGTTGGGGGC-3'	*Not*I site interrupting stop codon
TEV-ELP For	5'-AACCATGGAAGCTCTTTGCGG CCGCAGAGAACCTTTACTT-3'	5' *Nco*I, *Not*I sites, TEV protease site coding sequence
ELP-HK Rev	5-'TTCTCGAGCTAGAGCTCAT CTTTCTCAGAGTGATGGTGA TGGTGATGAGCGGCGGCACC-3'	*Xho*I site, 6X His, SEKDEL coding sequence
*Afl*III-BFDV CP For	5'-GGAACGCGTTAGGTACATGT GGGGCACCTCTAACTGC-3'	*Afl*III site
*Xho*I-ERH Rev	5'-GACTCGAGCTAGAGCTCATCT TTCTCAG-3'	*Xho*I site
NcoI-TEV For	5-'GACCATGGTCTACTTCCAA GGTG-3'	*Nco*I site, interruption of TEV recognition sequence

2.2.3 Construction of pTRA-ERH-ELP-51:

The completed pTRA-ERH-ELP28 vector was transformed into chemically competent dam-/dcm- *E. coli* (Bioline, USA). The plasmid was isolated from an overnight culture, digested with *Pfl*MI and the products were separated by gel electrophoresis. The digested pTRA-ERH-ELP28 vector and 0.5-1 kb area of the gel were excised and digestion products were extracted from the gel. The 0.5-1 kb products were allowed to self-ligate for one hour at 37 °C by the addition of

Fermentas Rapid DNA Ligase, with the intention of generating ELP fragments of random lengths. The ELP ligation product was then ligated into the digested pTRA-ERH-ELP28 vector, and transformed into chemically competent *E. coli* DH5α cells (E. cloni®, Lucigen, USA)) (see Appendix A for general cloning procedures). Transformants were screened by colony PCR, to determine the size of the ELP sequence insert, using pTRA vector-specific sequencing primers forward: 5'-CATTTCATTTGGAGAGGACACG-3', reverse: 5'-GAACACTCACACATTATTCT GG-3'. Clones with inserts larger than those in pTRA-ERH-ELP28 were analysed by restriction enzyme digests and were verified by sequencing. The pTRA-ERH-ELP51 vector contains an ELP sequence encoding 51 repeats of VPGVG.

2.2.4 Generation of pTRA and pPROEX clones

pTRAkc-ERH and pTRA-ERH-ELP clones

Both BFDV CP and ΔN40 CP genes each required a *Not*I restriction site that interrupts the stop codon, in order to allow read-through of the C-terminal His-tag (in the case of pTRAkc-ERH vector) or ELP (in the case of pTRA-ERH-ELP vectors). This was achieved by PCR, using BFDV CP-specific primers *Afl*III-BFDVN40 For or *Afl*III-BFDV CP For, with the reverse primer, *Not*I-BFDVN40 Rev, adding the *Not*I site (Table 2).

CP or ΔN40 CP PCR products were digested with *Afl*III and *Xho*I and ligated into pTRAkc-ERH vector (Appendix B), digested with *Nco*I and *Xho*I, using the procedures described in Appendix A. Similarly, to generate CP-ELP or ΔN40 CP-ELP fusions, CP or ΔN40 CP PCR products were cloned into pTRA-ERH-ELP vectors. Transformants were screened for the presence of inserts by colony PCR using pTRA vector-specific sequencing primers.

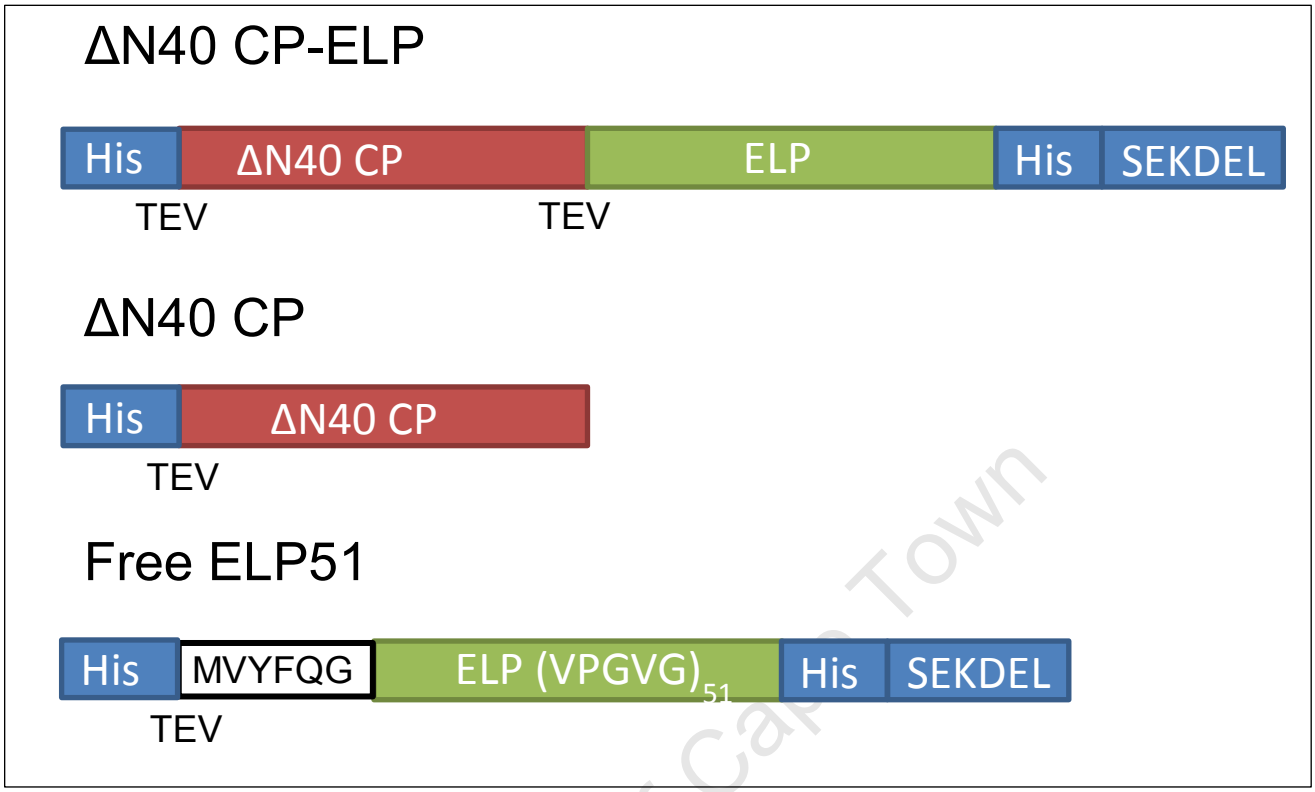


Figure 5: Schematic of proteins expressed by pPROEX-HTb

The proteins expressed by pPROEX-HTb are summarised. His: His-tag, ΔN40 CP: truncated BFDV CP, ELP: 28 x VPGVG or 51 x VPGVG sequence, TEV: TEV protease recognition sequence at the junction of segments. The free ELP51 contains a linker peptide MVYFQG separating the N-terminal His tag and the ELP. Note that the N-terminal His tag and TEV recognition sequence are original features of the pPROEX-HTb vector.

pPROEX-HTb expression constructs

The pPROEX CP and pPROEX-ΔN40 CP clones were provided by Guy Regnard (Vaccine Group, Department of Molecular and Cell Biology, UCT). These were created by subcloning the respective CP genes from the pGEM®-T Easy vectors into pPROEX-HTb, using the 5' *Afl*III and 3' *Xho*I sites flanking the CP genes and the compatible 5' *Nco*I and 3' *Xho*I sites in the multiple cloning site of pPROEX-HTb (Appendix D).

pPROEX constructs containing ΔN40 CP-ELP fusions were prepared. PCR was again used to generate ΔN40 CP-ELP inserts for the pPROEX vector, using the respective pTRA-ΔN40 CP-ELP clones as the templates. BFDV ΔN40 CP-specific forward primer *Afl*III-BFDVN40 For and the vector-specific reverse primer, *Xho*I ERH Rev, were used. These PCR products were digested with *Afl*III and *Xho*I and ligated into the pPROEX vector which had been digested with *Nco*I and *Xho*I. The free ELP51 PCR product was generated from the pTRA-ERH-ELP51 vector template using the primers *Nco*I-TEV For and *Xho*I ERH Rev. The free ELP51 PCR product was digested with *Nco*I and *Xho*I and cloned directly into similarly digested pPROEX. The proteins expressed by pPROEX are summarised in Figure 5.

2.2.5 Protein expression with pPROEX-HTb

Protein expression was performed using *E. coli* DH5α cells. Ten millilitre LB medium starter cultures were inoculated with a single colony, supplemented with 100 µg/ml ampicillin and incubated overnight with shaking at 37 °C. Erlenmeyer flasks containing 1/5th total volume of LB medium with 100 µg/ml ampicillin was inoculated with 1/00th volume of starter culture. These cultures were grown at 37 °C in an orbital shaker (120 rpm) until the optical density of the culture at 600 nm reached 0.5-0.8. Protein expression was induced by the addition of 0.6 mM IPTG, and the cultures were grown at 37 °C in an orbital shaker for a further 3 hours.

2.2.6 Protein extraction from *E. coli* cells

Frozen cell pellets were resuspended in 1/10th culture volume of 50 mM Tris-HCl pH 8.0. The cells were sonicated using a Sonicator 3000 (MISONIX, USA) with microtip for 8 x 15 second pulses at an output of 6 W. Triton X-100 was added to the cell lysate to 0.1% (v/v) and incubated with gentle shaking at 25 °C for 30 minutes. Cell debris was pelleted (insoluble fraction) by centrifugation at 15,000 x g at 4 °C and the soluble protein fraction was recovered in the supernatant. The extraction of inclusion body protein using Bugbuster® solution (Novagen, USA) and the purification of inclusion bodies from *E. coli* cells was performed using the manufacturer's protocol. To prepare soluble protein to be used as a positive control for immunoblots, washed IB proteins were solubilised using 50 mM Tris-HCl buffer pH 8.0 containing 0.2% (w/v) N-lauroyl sarcosine (NLS) (sodium salt, Sigma-Aldich

USA). The soluble protein was collected by recovering the supernatant following centrifugation at 15 000 x g for 20 minutes. The NLS was removed by a twofold dilution of the sample in PBS, followed by dialysis against PBS overnight at 4 °C, with stirring. In cases of TEV protease digestion, ΔN40 CP was digested with AcTEV protease (Invitrogen, USA) using the manufacturer's recommended reaction setup at 27 °C for 16 hours.

2.2.7 Dialysis and endotoxin testing of antigens

Dialysis of sample was done using dialysis tubing with a molecular weight cut-off of 10 kDa (Thermo Fisher Scientific, USA) against 5L sterile Dulbecco's PBS (Sigma-Aldrich, USA) for 20h at 4 °C, with constant stirring. Levels of endotoxin were measured using a Limulus Amebocyte assay QCL-100 kit (Cambrex, USA) as per the manufacturer's instructions. The absorbance readings were performed using *Immuno* 96-microwell clear solid plates (Thermo Fisher Scientific, USA) and a BioTek Powerwave XS plate reader (BioTek, USA).

2.2.8 Polyclonal antibody production in rabbits: inoculation schedule

The amount of protein was divided equally into 4 tubes for each antigen (for the initial and three booster injections; see below) containing a 0.5 mL suspension of the inclusion bodies in PBS. The total amount of antigen injected in each dose was approximately 40 µg for ΔN40-CP and 17 µg for ΔN40 CP-ELP51.

Production of rabbit immune sera was done by Deltamune (South Africa) as follows: Ten millilitres pre-immune blood was collected from the marginal ear vein, prior to administration of the antigen.

Initial injection: Animals were injected subcutaneously on the back at one site with 0.5 mL of the antigen in the presence of Freund's incomplete adjuvant (IFA) made according to the protocol in the Difco catalogue (#263910) (Difco, USA). This was followed by 3 subsequent booster injections which were also administered subcutaneously:

Booster 1: the first booster followed two weeks after the primary (initial) injection (as above) with IFA.

Booster 2 the second booster followed one week after the first booster (as above) with IFA.

Booster 3: the third and last booster injection followed one week after the second booster (as above) with IFA.

The pre-immune serum and the final serum samples received from Deltamune were stored in 1 mL aliquots at -20 °C. All protocols were approved by the Deltamune Animal Ethics Committee

2.2.9 SDS PAGE and Immunoblotting:

5 x SDS sample buffer (25% (v/v) glycerol, 0.5 M DTT, 5% (w/v) SDS, 0.01% (w/v) bromothymol blue) was added to protein samples to 1 x concentration. Protein samples were then incubated at 85 °C for 5 minutes and loaded onto 12% polyacrylamide gels, followed by electrophoresis at 20 mA. Alternatively, gels cast and used with the Bio-Rad Tetra Cell system (Bio-Rad Laboratories, USA) were electrophoresed at 120 V, as per the manufacturer's protocol. Gels were stained with Coomassie Brilliant Blue stain for 2 hours at 37°C and destained overnight. PageRuler™ pre-stained protein ladder #SM6071 (Fermentas, Canada) was used for all gels as a molecular weight marker.

For immunoblots, protein transfer from polyacrylamide gels to nitrocellulose membrane was done with a Bio-Rad Trans-Blot® Semi-dry transfer cell at 15 V for 1.5 hours. Blots were blocked with blocking buffer (PBS pH 7.4, 0.1 % Tween-20, 5 % skim milk) for 30 minutes and then incubated with a one in 2000 dilution of anti-His-tag mouse IgG antibody (Serotech, USA), or rabbit serum at the specified dilution, in blocking buffer overnight at room temperature, with shaking (50 rpm). Blots were washed for 4 x 15 minutes and then incubated with 1/10,000 diluted anti-mouse (or anti-rabbit) secondary antibody conjugated to alkaline phosphatase (Sigma-Aldrich, USA) in blocking buffer for 2 hours at room temperature. After a final four 15 minute washes in blocking buffer (without skim milk), blots were developed with NBT/BCIP substrate (Roche, USA).

The amounts of ΔN40 CP and ΔN40 CP-ELP51 antigens on Coomassie gels were estimated using a BSA (bovine serum albumin) (Fermentas, USA) standard curve.

The quantitation of bands was performed using a Syngene Gene Genius imaging system and GeneTools software (Synoptics Inc., UK).

2.2.10 Pre-absorption of rabbit sera

Where specified, rabbit sera were pre-absorbed against cell lysate from uniduced *E. coli*, or uninfiltrated plant extract, in order to reduce background. Nitrocellulose membrane was covered with *E. coli* cell lysate or plant extract and incubated at room temperature on an orbital shaker for 2 hours. The cell lysate or plant extract was discarded and the membrane was washed 4 x15 minutes with blocking buffer. The membrane was then incubated with rabbit serum, diluted in blocking buffer, for 2 hours. The membrane was discarded and the pre-absorbed rabbit serum was then used for immunoblotting or stored at -20 °C.

2.3 Results

The pTRAkc-ERH vector was successfully modified to create pTRA-ERH-ELP28, as confirmed by restriction enzyme digests (Figure 6) and sequencing.

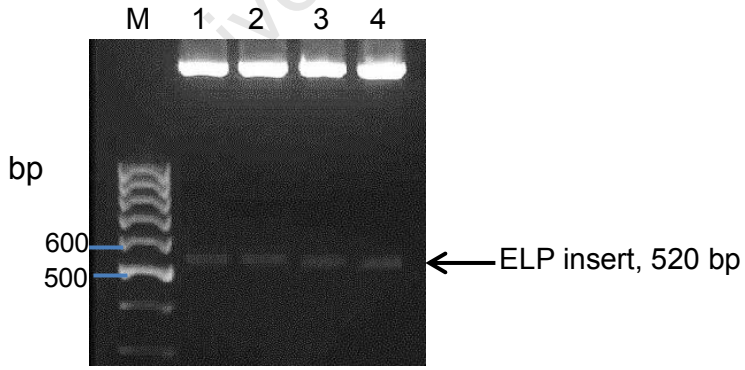


Figure 6: Restriction enzyme digest verification of pTRA-ERH-ELP28 clones

M: DNA bp marker, 1-4: four putative pTRA-ERH-ELP28 clones digested with *Nco*I and *Xho*I to excise the ELP insert (approximately 520 bp, indicated by the arrow).

CP or ΔN40 CP genes were cloned into pTRA-ERH-ELP28 as described in 2.2.4 to create CP-ELP28 or ΔN40 CP-ELP28 fusions, which were confirmed by restriction enzyme digests (data not shown) and sequencing.

A longer ELP (consisting of more than 28 VPGVG repeats) was needed to investigate the effect of different ELP lengths on the yield and purification efficiency of the ELP fusions. This required a pTRA-ERH-ELP vector with a longer ELP insert. A library of pTRA-ERH-ELP constructs with random ELP lengths had been generated, as described in 2.2.3. Clones were screened by colony PCR for the size of the ELP insert using pTRA-specific primers. As shown in Figure 7, lane 1, a clone was identified which contained an insert approximately 300 bp larger than that of the original pTRA-ERH-ELP28 vector (approximately 800 bp). This plasmid was later screened by restriction enzyme digests and was verified by sequencing. Sequence analysis confirmed that the new vector contained the coding sequence for 51 VPGVG repeats. This confirmed that the pTRA-ERH-ELP51 vector had been constructed successfully. CP or ΔN40 CP genes were cloned into pTRA-ERH-ELP51as described in 2.2.4 to create CP-ELP51or ΔN40 CP-ELP51 fusions, which were confirmed by restriction enzyme digests (data not shown) and sequencing.

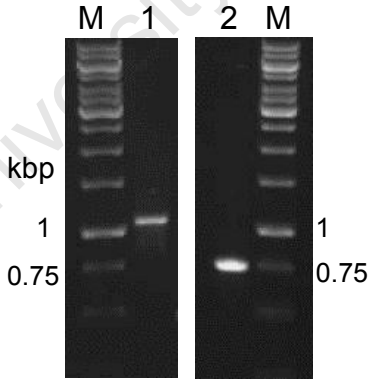


Figure 7: Colony PCR identifying the pTRA-ERH-ELP51 vector clone

M: DNA bp marker, 1: Putative pTRA-ERH-ELP construct with 300 bp larger insert (subsequently identified as the coding sequence for 51 VPGVG repeats), 2: original pTRA-ERH-ELP28 positive control DNA

The insertion of ΔN40 CP, ΔN40 CP-ELP28 and ΔN40 CP-ELP51 into pPROEX-HTb was confirmed by *Mlu*I and *Xho*I restriction enzyme digests (Figure 8) and sequencing with the 23-base M13 reverse primer.

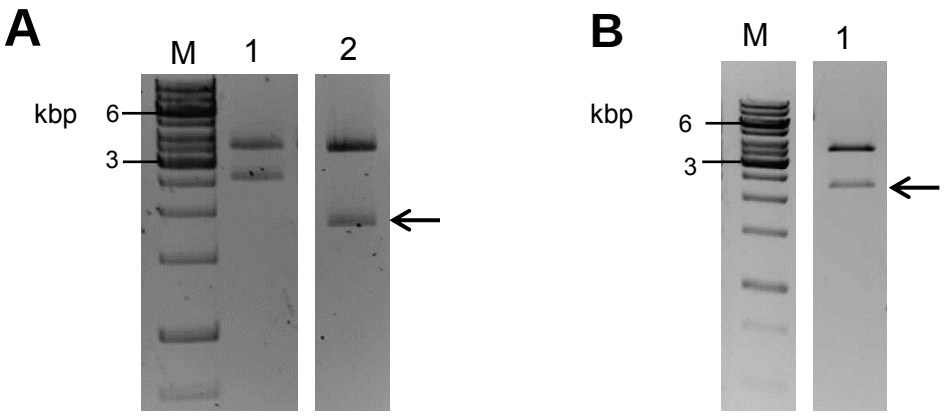


Figure 8: *Mlu*I-*Xho*I restriction enzyme digests of pPROEX clones to confirm inserts

A: M: DNA bp standards, 1: pPROEX-ΔN40 CP-ELP51 2: pPROEX-ΔN40 CP, **B**: M: DNA bp standards, 1: pPROEX-ΔN40 CP-ELP28. The vector backbone is 3670 bp and the inserts are indicated by the arrows. *Mlu*I is a unique restriction site in the vector and the *Xho*I site is preserved in cloning of the digested PCR product into pPROEX-HT.

Small-scale protein expression trials were conducted with harvesting of cell samples at 1, 2 and 3 hours post-induction with IPTG. The cells were sampled at different time points in order to determine the effect of incubation time on protein yield. SDS-PAGE and immunoblotting of induced whole cells showed that the full-length BFDV CP (Figure 9, lanes 2-4) was not expressed in *E. coli*, in contrast to ΔN40 CP (Figure 9, lanes 5-7). Lack of IPTG induction produced no ΔN40 CP (Figure 9, lane 4). Incubation time of induced cells had no noticeable effect on the yield of ΔN40 CP: the minimum incubation time of 1 hour yielded as much ΔN40 CP (Figure 9, lane 5) as for 3 hours (Figure 9, lane 7).

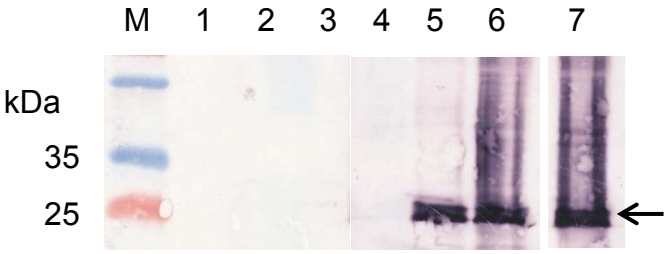


Figure 9: Anti-His immunoblot of *E. coli* whole cell lysates

M: molecular weight marker, 1: CP uninduced cells, 2: CP cells 1hour post-induction, 3: CP cells 2 hours post-induction, 4: ΔN40 CP uninduced cells, 5: ΔN40 CP cells 1 hour post-induction, 6: ΔN40 CP cells 2 hours post-induction, 7: ΔN40 CP cells 3 hours post-induction (ΔN40 CP band, 25 kDa, indicated by the arrow)

The ΔN40 CP and both the ΔN40 CP-ELP28 and ΔN40 CP-ELP28 proteins were expressed in *E coli*, but all of these were almost completely insoluble. The recombinant proteins in the insoluble fraction (obtained in the pellet of cell debris after cell lysis) were clearly visible by Coomassie Blue staining of polyacrylamide gels (Figure 10A, lanes 2, 4 and 6) and their identity was confirmed by immunoblot (Figure 10B). ΔN40 CP, ΔN40 CP-ELP28 and ΔN40 CP-ELP51 were the expected sizes of 27, 40 and 50 kDa, respectively. The amount of recombinant protein in the soluble fraction was too low to be seen by Coomassie Blue staining. However, the presence of ΔN40 CP protein could be detected in the soluble fraction by immunoblotting (Figure 10 B, lanes 1, 3 and 5).

Free ELP51 was produced in bacteria to facilitate purification of ΔN40 CP-ELP51 fusion proteins by inverse transition cycling, by acting as co-aggregate. This free ELP was expressed in bacteria as for the ΔN40 CP and ΔN40 CP-ELP proteins. In contrast to the ΔN40 CP proteins, ELP51 was more soluble, as the ELP band in the soluble fraction could be identified on a Coomassie stained gel (Figure 10A, lane 7). Immunoblotting confirmed the identity of the protein, which was the expected size of approximately 26 kDa (Figure 10B, lane 7). A proportion of this ELP protein was insoluble (Figure 10B, lane 8).

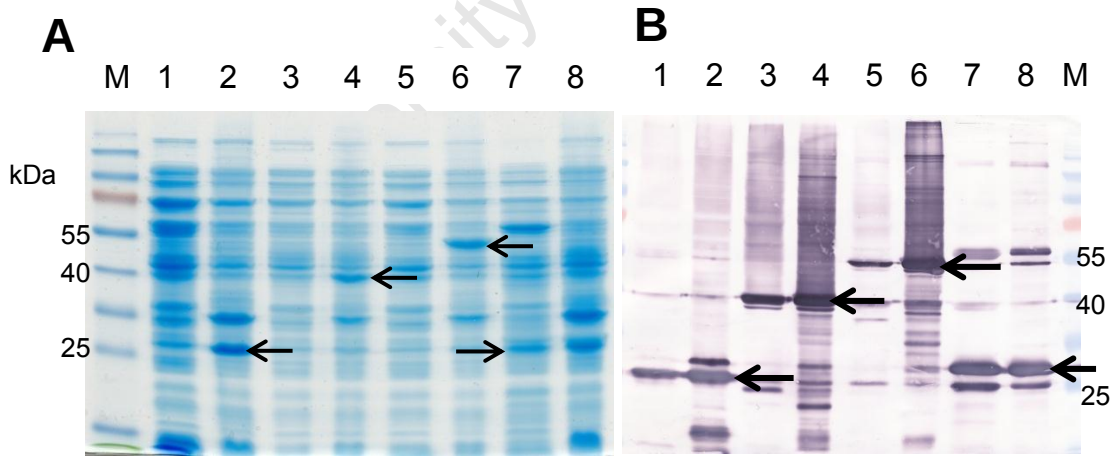


Figure 10: Expression of BFDV, ELP and CP-ELP proteins using pPROEX-HT

The respective protein bands are indicated by arrows. **A**: Coomassie-Blue stained gel for which equal volumes of all samples were loaded, 1: ΔN40 CP soluble fraction, 2: ΔN40 CP insoluble fraction, 3: ΔN40 CP-ELP28 soluble fraction, 4: ΔN40 CP-ELP28 insoluble fraction, 5: ΔN40 CP-ELP51 soluble fraction, 6: ΔN40 CP-ELP51 insoluble fraction, 7: ELP51 soluble fraction, 8: ELP51 insoluble fraction. **B** is the cognate anti-His immunoblot. The volume loaded of insoluble fractions is half that of the soluble fractions. Note the presence of dimers for free ELP51 (approximately 55 kDa, lanes 7 and 8).

Given the insoluble nature of ΔN40 CP proteins expressed in bacteria, it was most likely that most of the the proteins had been sequestered into insoluble inclusion bodies (IBs). IBs typically contain the recombinant protein with a high degree of purity (Ventura & Villaverde, 2006). It is therefore possible to acquire partially purified recombinant protein by successive washes of the insoluble IBs. ΔN40 CP and ΔN40 CP-ELP51 were produced in 250 mL cultures and the inclusion bodies were processed with Bugbuster™ (Novagen, USA), using the manufacturer's protocol. Only one of the ELPylated proteins (ΔN40 CP-ELP51 and not ΔN40 CP-ELP28) was used for antibody production, because these two antigens were similar and so they would likely evoke the same antibody responses. Bugbuster™ is a mixture of detergents that achieves efficient lysis of cells. After ten wash steps, the purity of the IB preparation was analysed by SDS-PAGE and Coomassie Blue staining. (Figure 11, lanes 1-4). Because the expressed ΔN40 CP proteins were insoluble, they could not be quantified using a total soluble protein assay, such as the Bradford assay. The amount recombinant protein was estimated using serial dilutions of bovine serum albumin (BSA) loaded for SDS-PAGE and stained with Coomassie Blue (Figure 11, lanes 5-8).

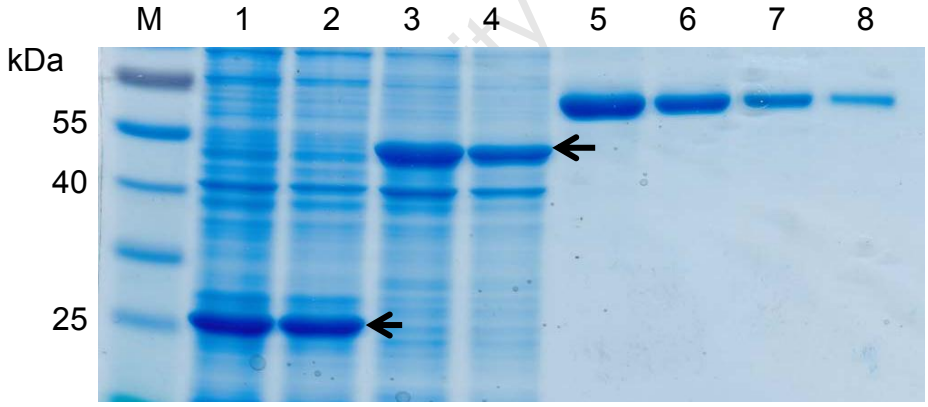


Figure 11: Partial purification of *E. coli* inclusion body samples

1: ΔN40 CP undiluted, ΔN40 CP diluted 1:2 (indicated by the arrow), 3: ΔN40 CP-ELP51 undiluted, 4: ΔN40 CP-ELP51 diluted 1:2 (indicated by the arrow), 5-8: BSA standards of 6.4, 3.2, 1.6 and 0.8 µg protein, respectively.

The inclusion body suspensions were dialysed against sterile Dulbecco's PBS (Sigma-Aldrich, USA) overnight at 4 °C in order to remove bacterial endotoxins from the protein sample. SDS-PAGE and Coomassie Blue staining was done for the samples after dialysis (Figure 12, lanes 1 and 2). It was noticed that some of the ΔN40 CP-ELP51 (Figure 12, lane 2) protein had been lost during the dialysis process, possibly during the transfer from the dialysis tubing. However, the amount of background proteins were reduced as well, and most were undetectable on the gel. This could have caused fewer non-specific antibodies to be raised in the inoculated rabbits. The final antigen doses were 40 µg per 0.5 mL injection for ΔN40 CP and 17 µg per 0.5 mL injection for ΔN40 CP-ELP51. The protein samples were then tested for their levels of bacterial endotoxin using the Limulus Amebocyte assay QCL-1000 kit (Cambrex, USA). The endotoxin levels of both antigen samples were reduced by the dialysis step (data not shown). The inoculation of rabbits was done by Deltamune (South Africa), according to the inoculation schedule described in 2.2.7.

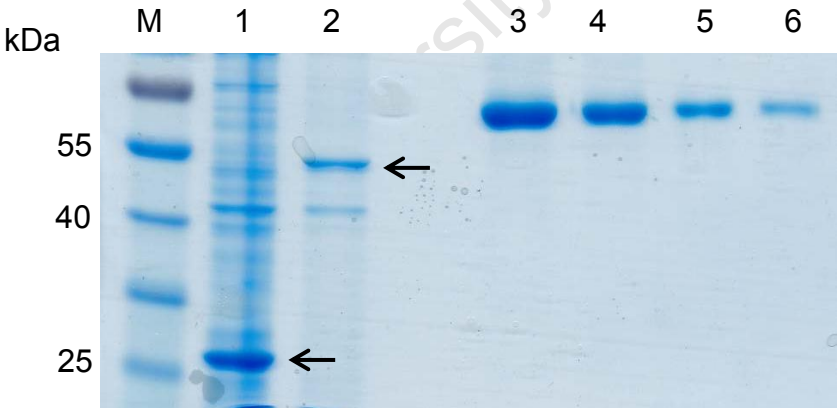


Figure 12: Inclusion body proteins after dialysis

Coomassie-stained gel of final inclusion body samples submitted for polyclonal antibody production in rabbits. 1: ΔN40 CP, 2: ΔN40 CP-ELP51 (indicated by arrows), 3-6: BSA standards of 6.4, 3.2, 1.6 and 0.8 µg protein, respectively.

The use of N-lauroyl sarcosine (NLS) detergent to solubilise insoluble protein aggregates produced in *E. coli* has been previously reported (Frankel et al., 1991). Inclusion body proteins were partially purified by washing as before, but these were then solubilised in buffer containing 0.2% (w/v) NLS. Centrifugation pelleted the remaining insoluble debris. The supernatant was dialysed against PBS to remove the detergent and filtered through a 0.2 µm cellulose acetate syringe filter. The result was partially purified, soluble ΔN40 CP protein that could be used as standards for immunoblots (as used in Figure 13A and B)

The serum received from Deltamune was used in immunoblots as described in 2.2.9 to probe for ΔN40 CP or ΔN40 CP-ELP proteins expressed in bacteria. Each serum was used at a 1/5000 dilution, after pre-absorbtion against proteins from uninduced *E. coli* as described in 2.2.9. There is no commercially available anti-BFDV ΔN40 CP antibody. Therefore all BFDV CP proteins in in our laboratory have been expressed as His tag fusions for detection by anti-His immunoblots. The His-tag had to be removed from ΔN40 CP in order to determine if the sera recognised ΔN40 CP without a His-tag. This was done by TEV protease digestion of ΔN40 CP obtained from solubilised inclusion bodies (Figure 13A and B). The TEV digestion was not complete, but the lower band corresponding to ΔN40 CP without a His-tag showed that the anti-ΔN40 CP serum recognised BFDV ΔN40 CP alone (Figure 13A, lane 2), as did the anti-ΔN40 CP-ELP51 serum (Figure 13B, lane 2). Bands of approximately 50 and 75 kDa were also detected (Figure 13A), corresponding to dimers and trimers, respectively, of ΔN40 CP. As shown in Figure 13B, an additional 40 kDa protein was detected (lanes 1 and 2) by the anti-ΔN40 CP-ELP51 serum. It was suspected that this was a result of an *E. coli* host protein being recognised by the sera. This may correspond to the contaminating *E. coli* protein of approximately 40 kDa which can be seen in the ΔN40 CP-ELP51 antigen sample (Figure 12, lane 2). The sera were expected to contain antibodies against *E. coli* host proteins, as the antigen used to inoculate rabbits was made in *E. coli*, although the pre-absorption technique and appropriate serum dilution had greatly reduced this background.

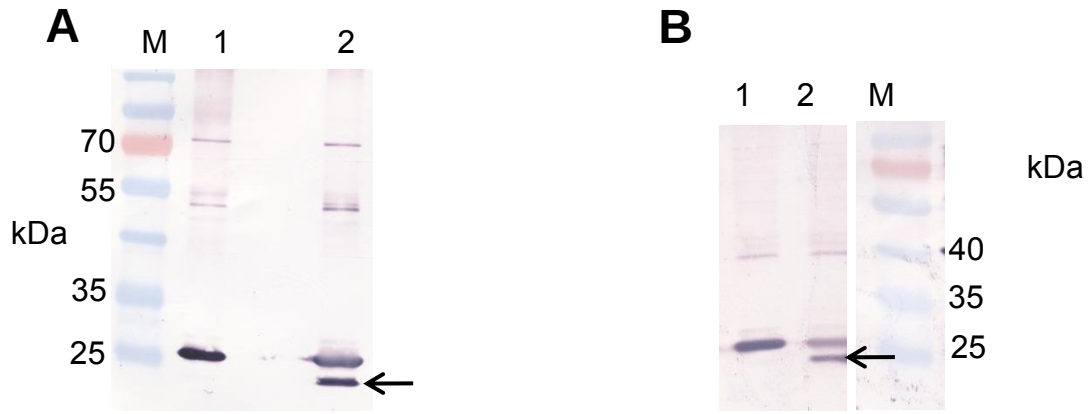


Figure 13: Partial TEV digestion of ΔN40 CP produced in bacteria and detection with anti-ΔN40 CP serum

Solubilised ΔN40 CP protein was digested with TEV protease overnight at 27 °C. **A**: anti-ΔN40 CP serum (1/5000 dilution). M: molecular weight marker, 1: undigested control, 2: TEV digested sample. **B**: anti-ΔN40 CP-ELP51 serum (1/5000 dilution). 1: undigested control, 2: TEV digested sample, M: molecular weight marker. The ΔN40 CP with the His tag is 26 kDa and ΔN40 CP with the His tag removed is indicated by the arrow (approx. 24 kDa). Note that bands of approximately 50 and 75 kDa in **A** (lanes 1 and 2) correspond dimers of ΔN40 CP and the 40 kDa band in **B** is suspected to be recognition of *E. coli* host protein by non-specific antibodies.

A final test of the rabbit sera was carried out to determine if they could detect proteins expressed by a different host, namely, tobacco plants. The anti-ΔN40 CP-ELP51 serum was used at a dilution of 1/10,000 to detect plant-expressed ΔN40-ELP proteins targeted to the ER or cytoplasm (plant expression described in Chapter 3). ΔN40-ELP51 targeted to the cytoplasm was detected (Figure 14A, lane 2), as was the ΔN40-ELP28 targeted to the ER (Figure 14A, lane 3) and ΔN40-ELP51 targeted to the ER (Figure 14A, lane 4). ΔN40-ELP28 targeted to the cytoplasm was loaded in lane 1 (Figure 14A), but the amount of protein was presumably too low to be detected. Degradation products of lower molecular weight than the intact proteins could be detected for all proteins (Figure 14, ~40kDa band in lanes 2 and 3, and ~45 kDa band in lane 4,). In another immunoblot, ΔN40-ELP51 targeted to the ER (Figure 14B, lane 5) was subjected to one round of cITC purification (see Chapter 4), and degradation products ranging from 25-35 kDa in size were detected in the purified sample (Figure 14B, lane 6). The anti-ΔN40 CP-ELP51 serum could in fact detect any ELPylated protein, including GFP-ELP28 (Figure 14B, lane1), GFP-ELP51 (Figure 14 B, lane 3) and free ELP51 (Figure 14 B, lane 2), all produced in plants.

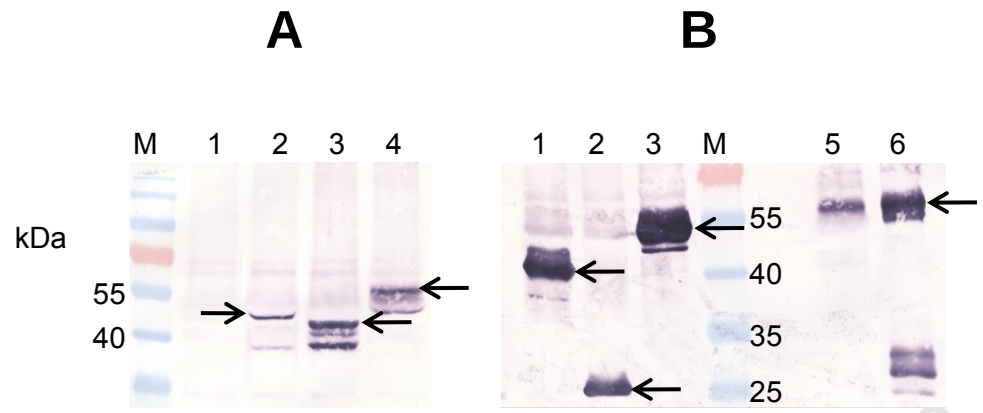


Figure 14: Detection of ELPylated proteins produced in plants using rabbit serum

A: Detection of ΔN40 CP-ELP proteins using anti-ΔN40 CP-ELP51 rabbit serum at 1/10,000 dilution. 1: ΔN40 CP-ELP28 targeted to the cytoplasm, 2: ΔN40 CP-ELP51 targeted to the cytoplasm, 3: ΔN40 CP-ELP28 targeted to the ER, 4: ΔN40 CP-ELP51 targeted to the ER. **B**: Detection of other ELPylated proteins using anti-ΔN40 CP-ELP51 rabbit serum at 1/8,000 dilution. 1: GFP-ELP28, 2: free ELP51, GFP-ELP51, 5: ΔN40 CP-ELP51 crude extract from plants., 6: ΔN40 CP-ELP51 cold spin supernatant from ITC purification (see Chapter 4). The relevant protein bands are indicated by arrows.

The anti-ΔN40 CP serum was used at a dilution of 1/2000 to detect ΔN40 CP proteins produced in plants. It could detect ΔN40 CP produced in plants, which had been targeted to the cytoplasm (Figure 15, lane 1). ΔN40 CP produced in bacteria was used as a positive control for which a 50 kDa ΔN40 CP dimer was also observed (Figure 15, lane 2). The ΔN40 CP-ELP51 produced in plants (Figure 15, lane 3) was not detected as strongly as ΔN40 CP, even though the yield of the ELPylated protein was significantly higher (see Chapter 3, Figure 23). This suggests that the ELP interferes with antibody binding to ΔN40 CP epitopes. The ΔN40 CP-ELP51 was purified by one round of mITC (see Chapter 4), which concentrated the protein, giving rise to a cleaner band (Figure 15, lane 4). Degradation products of ΔN40 CP-ELP51, presumably containing ΔN40 CP epitopes, were detected as 25 kDa bands (Figure 15, lanes 3 and 4).

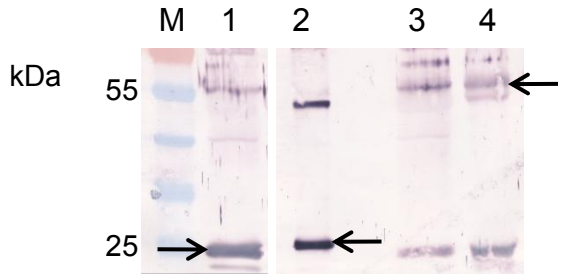


Figure 15: Recognition of BFDV ΔN40 CP expressed in plants by anti-ΔN40 CP serum

Anti-CP serum used at 1/2,000 dilution to detect plant-expressed ΔN40 CP proteins. M: molecular weight marker, 1: ΔN40 CP expressed in plants (25 kDa, indicated by the arrow), 2: ΔN40 CP expressed in bacteria, 3: ΔN40 CP ELP-51 crude extract, 4: ΔN40 CP ELP-51(55 kDa, indicated by the arrow) after one round of purification by mITC (see Chapter 4).

2.4 Discussion

Bacterial expression of BFDV CP proteins was done in order to obtain sufficient antigen for antibody production. Truncated BFDV CP, namely ΔN40 CP, was successfully expressed in *E. coli.* Elastin-like polypeptide fusions of ΔN40 CP were also expressed (Figure 10). However, the ΔN40 CP proteins expressed in *E. coli* were almost completely insoluble. The insolubility of bacterial-produced ΔN40 CP has been described previously (Johne et al., 2004). Strategies that could be tested to increase the amount of soluble protein include: lower incubation temperature, lower concentration of the inducer (IPTG) or the increased co-expression of molecular chaperones (De Marco et al., 2005). In an attempt to produce more soluble ELP fusion protein, ΔN40 CP-ELP51 induced cultures were incubated at 22 °C (instead of 37 °C) for 20 hours. There was no change in the expression level or the solubility of the protein (data not shown).

Expression of full-length CP with pPROEX yielded no detectable protein in induced cells, as determined by anti-His immunoblot. The N-terminal nuclear translocation signal of the full-length CP is rich in arginine residues, some of them coded for by the codons AGA and AGG (Heath et al., 2006). These codons are rarely used in prokaryotes and can lead to premature termination of translation (Sørensen & Mortensen, 2005). Not even truncated proteins were observed for full-length BFDV CP expression, suggesting that translation was inhibited at the early stages. This problem may be rectified by codon optimisation, i.e. replacing rare codons in the coding sequence with codons commonly used in *E. coli*. This strategy worked for the translation of full-length PCV-2 CP in *E. coli* (Marcekova et al., 2009). Instead of codon optimisation or truncation, commercially available bacterial strains such as Rosetta (DE3) (Novagen, Germany), that co-express eukaryotic-type tRNAs, could be used. This strain has proved useful for the expression of human genes containing rare arginine codons (Tegel et al., 2010), and may therefore be useful for the expression of BFDV CP genes that have not been codon optimised for bacteria.

The free ELP51 was expressed in *E. coli* for use as a co-aggregate during ITC purification of plant-produced CP-ELP proteins (see Chapter 4). This free ELP51 was more soluble than ΔN40 CP or ΔN40 CP-ELP fusion proteins. The band corresponding to the ELP protein in the soluble fraction was clearly visible on a Coomassie-stained gel (Figure 10). This showed that the ELP was not responsible for the insoluble nature of expressed CP-ELP. Fusion to ELP has previously been used to increase the solubility of target proteins expressed in *E. coli* (Solomon et al., 2004; Trabbic-Carlson et al., 2004). The ΔN40 CP-ELP proteins, like ΔN40 CP, were mostly insoluble, i.e. fusion to ELP did not increase the solubility of ΔN40 CP. Although a large proportion of free ELP51 protein was trapped in the insoluble fraction, the amount of soluble ELP that could be recovered, concentrated and purified by ITC was typically up to 1 mg/mL (see Chapter 4). This was sufficient for further purification experiments.

ΔN40 CP and ΔN40 CP-ELP51 inclusion bodies were partially purified from *E. coli* using successive washes in PBS. This protein was injected into rabbits for antibody production. Inclusion bodies have long been considered a ‘dead-end’ by-product of bacterial protein expression, but the quality and purity of recombinant proteins within IBs has recently been noted (Ventura & Villaverde, 2006; García-Fruitós et al.,

2012). However, the solubilisation and refolding of proteins from inclusion bodies is time-consuming and a large proportion of protein may be lost. Inclusion bodies are insoluble and can be separated from the bulk of contaminating host proteins by successive washes in PBS. The protein losses using this method are minimal. Therefore it was decided to successively wash the IBs and use these in order to preserve the antigen dose.

Bacterial-expressed ΔN40 CP from solubilised inclusion bodies could be recognised by anti-ΔN40 CP (Figure 13 A) and anti-ΔN40 CP-ELP (Figure 13B) rabbit sera. It has been shown that serum from animals immunised with proteins in the form of inclusion bodies is reactive against soluble forms of the protein (Chen et al., 2005). Indeed, both sera recognised soluble ΔN40 CP-ELP51 expressed in plants (Figures 14, 15 and 35B).

The anti-ΔN40 CP serum had non-detectable anti-His-tag activity, as determined by loading a large amount of His-tagged free ELP51 on the same gel for immunoblotting (Figure 36B, lanes 2 and 3). Plant-expressed BFDV ΔN40 CP could be detected using this serum (Figure 15, lane 1). Interestingly, the recognition of BFDV ΔN40 CP as a fusion to ELP reduced the signal when probed with the anti-CP serum, relative to unfused CP (compare lane 1 to lanes 3 and 4 in Figure 15). The ELP may interfere with antibody binding to the relevant CP epitopes. The anti-ΔN40 CP serum could detect 25 kDa degradation products of ΔN40-ELP51, presumably containing ΔN40 CP epitopes, as sharp bands on immunoblots (Figure 15).

The anti-ΔN40 CP-ELP serum recognised ΔN40 CP alone (Figure 13B) as well as any ELPylated proteins (Figure 14). This serum could be used to probe for other ELPylated proteins expressed in plants, such as GFP-ELP. However, this serum exhibited a higher anti-His-tag activity, probably because the ΔN40 CP-ELP51 antigen used to inoculate rabbits had both N- and C-terminal His tags (data not shown). As both antigens used for antibody production were only partially purified, antibodies were raised against *E. coli* host proteins, as seen by the 50 and 70 kDa bands on immunoblots (Figure 13A and 13B). The sera were therefore pre-absorbed against proteins from untransformed *E. coli* cell lysate before use, which minimised the non-specific bands on immunoblots. Both antisera could be useful for detection

of BFDV CP proteins produced in our laboratory, as no anti-BFDV CP antibodies are commercially available.

To summarise the results, truncated BFDV ΔN40 CP proteins produced in *E. coli* were mostly in the form of insoluble inclusion bodies. In the case of full-length BFDV CP, no expression was detected. The free ELP51 exhibited a higher solubility than the ΔN40 CP proteins. Due to the insolubility of ΔN40 CP proteins, inclusion bodies were partially purified by successive washes in PBS. The inclusion body proteins were used to raise antibodies in rabbits. The anti-ΔN40 CP serum could detect ΔN40 CP protein produced in bacteria as well as ΔN40 CP produced in plants. The anti-ΔN40 CP-ELP51 serum could be used to detect ΔN40 CP as well as any ELPylated proteins. The ΔN40 CP-ELP51 serum did exhibit some anti-His activity as well.

ΔN40 CP produced in bacteria with a high yield could serve as a vaccine candidate after purification and endotoxin removal. Sequence analysis of PCV-2 capsid protein (Marcekova et al., 2009) and similarly, BFDV CP, reveals putative sites for N-glycosylation and phosphorylation. These post-translational modifications would be carried out only in eukaryotic expression systems. Whether BFDV CP in native virions carries these modifications, and their role in eliciting protective immune responses in the case of a CP subunit vaccine, is unclear. ΔN40 CP produced in prokaryotic systems could therefore be successful in imitating the epitopes of the native virus and provide an effective immune response. This is true assuming that no important epitopes are removed in the N-terminal truncation. None of the immunogenic epitopes of PCV-2 identified by Mahé et al. (2000) occurred within the first 40 amino acids, but this information is lacking for BFDV CP. The immunogenic epitopes of BFDV CP may be difficult to determine as there is no cell culture system for the propagation of mutant viruses. The use of truncated BFDV CP to raise antibodies that were subsequently used to detect BFDV in psittacine sera (Johne et al., 2004) suggests that a truncated CP subunit vaccine could be capable of eliciting at least a humoral response against the virus.

Chapter 3: Optimisation of BFDV capsid protein expression in *N. benthamiana*

3.1 Introduction

Proteins naturally produced in eukaryotes requiring post-translational modifications such as glycosylation must be produced by eukaryotic expression systems. Yeast fermentation and insect- and mammalian cell culture systems have traditionally been used to produce the majority of these therapeutic proteins. However, whole plants and plant cell cultures are gaining popularity as eukaryotic expression systems (Rybicki, 2010). Major advantages of the plant expression system include low cost and no risk of contamination with animal pathogens. Transient expression of recombinant proteins in plants such as tobacco by infiltration of *Agrobacterium* suspension into the leaves of the plant (agroinfiltration) has the advantages of a shorter production timescale and typically higher yields than for transgenic plants (Paul & Ma, 2011). Several vaccine antigens have been expressed in plants and evaluated in clinical trials (for a review, see Thanavala et al. (2006)).

Research efforts have been focused on increasing the yield of recombinant proteins expressed in plants. Plant viruses are adapted to induce expression of large amounts of viral protein in plant cells, as part of their replication. The genetic elements of plant viruses, such as promoters, have therefore been incorporated into many plant expression vectors. Among these, the cauliflower mosaic virus (CaMV) 35S promoter is well-studied and often used for its ability to confer high-level transcription (Mushegian & Shepherd, 1995).

Replicating expression vectors give rise to a replicon derived from a particular plant virus in order to increase the number of copies of the transgene, which may lead to higher protein expression. These replicons may also contain a gene encoding viral movement protein, which leads to systemic spread of the transgene via plasmodesmata connecting plant cells (Regnard et al., 2010). Non-replicating vectors often contain virus genetic elements, but do not rely on increasing the

transgene copy number in order to drive high-level expression of recombinant protein.

Cowpea mosaic virus, like the majority of plant viruses, has a positive sense, single-stranded RNA genome. The virus replicates via a double-stranded RNA intermediate and its RNA genome is directly translated by plant ribosomes The genome is segmented into two RNA molecules, RNA1 and RNA2 (Verver et al., 1998). The genes encoded by RNA2 are flanked by 5' and 3; untranslated regions (UTR's), which are important for high-level translation of viral polyproteins. Sainsbury and Lomonossoff (2008) discovered that the use of these UTR's flanking a gene of interest, in conjunction with the CaMV 35S promoter, could drive high-level expression the recombinant proteins in plants. This lead to the construction of pEAQ binary expression vectors (Sainsbury et al., 2009). The pEAQ vector has been used for the transient expression of enzymes (Vardakou et al., 2012) and monoclonal antibodies (Sainsbury et al., 2010) in tobacco plants.

The aim of the work presented here was to optimise the expression of the BFDV CP antigen in plants towards the development of low-cost vaccine candidates. Truncated CP (ΔN40 CP) was also expressed with the aim of maximising protein yield. In addition, CP and ΔN40 CP were each fused to elastin-like polypeptides (ELPs) of 28 or 51 pentapeptide repeats each. The fusion to ELP was done in an attempt to increase the yield of the CP-ELP of ΔN40 CP-ELP fusion proteins relative to their unfused counterparts. This would produce four new vaccine candidates, called CP-ELP28, CP-ELP51, ΔN40 CP-ELP28 and ΔN40 CP-ELP51. Fusion to ELP has been shown to increase the yield of several proteins expressed in plants (Patel et al., 2007; Floss et al., 2008a; Conley et al., 2009a). Free ELP51 (the ELP51 not fused to another protein) was similarly expressed in plants to characterise the purification conditions of ELP51. Green fluorescent protein was also expressed in plants as ELP28 or ELP51 fusions in order to serve as a model ELP fusion protein. The purification of all ELP fusions could then be attempted using the non-chromatographic method known as inverse transition cycling (ITC) (see Chapter 4).

This chapter describes the transient expression of BFDV CP and ΔN40 CP their respective ELP fusion proteins in tobacco (*N. benthamiana)* plants by agroinfiltration. Two expression vectors were tested. Initially, the pTRAkc-ERH vector was used for

expression of ΔN40CP-ELP28 fusion protein, which targeted protein accumulation the ER. The pEAQ-HT vector was then used to express protein targeted to either the endoplasmic reticulum (ER) or the cytoplasm. The pEAQ-HT vector was used for the majority of plant expression because of higher protein yields and because it did not require co-infiltration with a silencing suppressor plasmid. Expression time trials were conducted to optimise the length of transient protein expression. This optimisation work is the essential first step towards the goal of a plant-produced BFDV CP subunit vaccine, based on purified BFDV CP or its ELP fusion protein.

3.2 Materials and Methods

3.2.1 Cloning from pTRAkc-ERH and pTRA-ERH-ELP to pEAQ-HT

CP, free ELP51 and CP-ELP inserts from pTRAkc-ERH or pTRA-ERH-ELP clones described in 2.2.4 were prepared for cloning into the pEAQ-HT vector using PCR. The primers listed in Table 3 were used to amplify the respective genes. For cytoplasmic protein targeting, BFDV CP-specific forward primers BFDV CP-*Age*I For or BFDVN40-*Age*I Rev were used with the *Xho*I-h Rev reverse primer. For targeting to the endoplasmic reticulum, the LPH-*Age*I For and *Xho*I-ERH Rev primer pair was used. This added a 5' *Age*I site and preserved the 3' *Xho*I site for all genes.

The pTRA-ERH-ELP28 clone GFP-ELP28 (enhanced green fluorescent protein, accession: AY952326.1, with C-terminal ELP28) was provided by Ann Meyers (Department of Molecular and Cell Biology, UCT). The GFP-ELP28 insert was isolated by PCR using the primers LPH-*Age*I For and *Xho*I-ERH Rev for cloning into pEAQ-HT. To generate a GFP-ELP51 clone, the GFP segment was isolated by PCR using the LPH-*Age*I For primer and the GFP-specific reverse primer GFP-*Not*I Rev. This PCR product was digested with *Age*I and *Not*I. An existing pEAQ-HT clone containing CP-ELP-51 was digested with *Age*I and *Not*I and the vector, still containing the ELP51 segment, was isolated. This digested vector was ligated with the digested GFP PCR product to generate the pTRA-ERH-ELP51 clone, GFP-ELP51. The GFP-ELP51 and free ELP51 PCR inserts were isolated using the LPH-*Age*I For and *Xho*I-ERH Rev primer pair, from their respective pTRA-ERH-ELP templates, for cloning into pEAQ-HT.

Cloning into pEAQ-HT was performed as described in Appendix A. All PCR products were digested with *Age*I and *Xho*I, and ligated into similarly digested pEAQ-HT vector. Colonies were screened by PCR using pEAQ For2 and pEAQ Rev vector-specific primers (Table 3) and putative positive clones were checked by restriction enzyme digest with *Age*I and *Xho*I. pEAQ-HT constructs were then verified by sequencing (Macrogen, Netherlands).Targeting protein accumulation to the ER required an N-terminal signal peptide, which was already present for all pTRAkc-ERH constructs (murine IgG heavy chain variable region signal peptide, (Maclean et al., 2007)). When subcloning from pTRAkc-ERH to the pEAQ-HT vector, the PCR

product included the signal peptide upstream of the gene of interest, as well as the SEKDEL sequence required for ER retention. Cytoplasmic targeting excluded the signal peptide and mutated the S codon of the SEKDEL sequence (TCT) to a stop codon (TGA). A schematic of the proteins expressed in pEAQ-HT is shown in Figure 16.

Table 3: Oligonucleotide primers used for pEAQ-HT subcloning

'For' designates a forward primer and 'Rev' designates a reverse primer. Restriction enzyme sites are underlined.

Primer name	Sequence	Features
BFDV CP-*Age*I For	5'-GAACCGGTATGTGGGGC ACCTCTAAC-3'	*Age*I site
BFDVN40-*Age*I Rev	5'-GAACCGGTATGGAGTGGAGC TGGATC-3'	*Age*I site
LPH-*Age*I For	5'-GAACCGGTATGGAGTGG AGCTGGATC-3'	*Age*I site, anneals to start of 5' LPH signal sequence
*Xho*I-h Rev	5'-GACTCGAGCTAGAGCTCA TCTTTCTCCTAGTGATGGTG-3'	*Xho*I site, mutation of S codon of SEKDEL sequence into stop codon
*Xho*I-ERH Rev	5'-GACTCGAGCTAGAG CTCATCTTTCTCAG-3'	*Xho*I site
pEAQ For2	5'-GACGAACTTGGAGAAAGA TTGTTAAGC-3'	Forward sequencing primer
pEAQ Rev	5'-GACCGCTCACCAAACA TAGAAATG-3'	Reverse sequencing primer
GFP-*Not*I Rev	5'-ATGCGGCCGCCTTGTACAGC TCGTCCATGCC-3'	*Not*I site interrupting stop codon

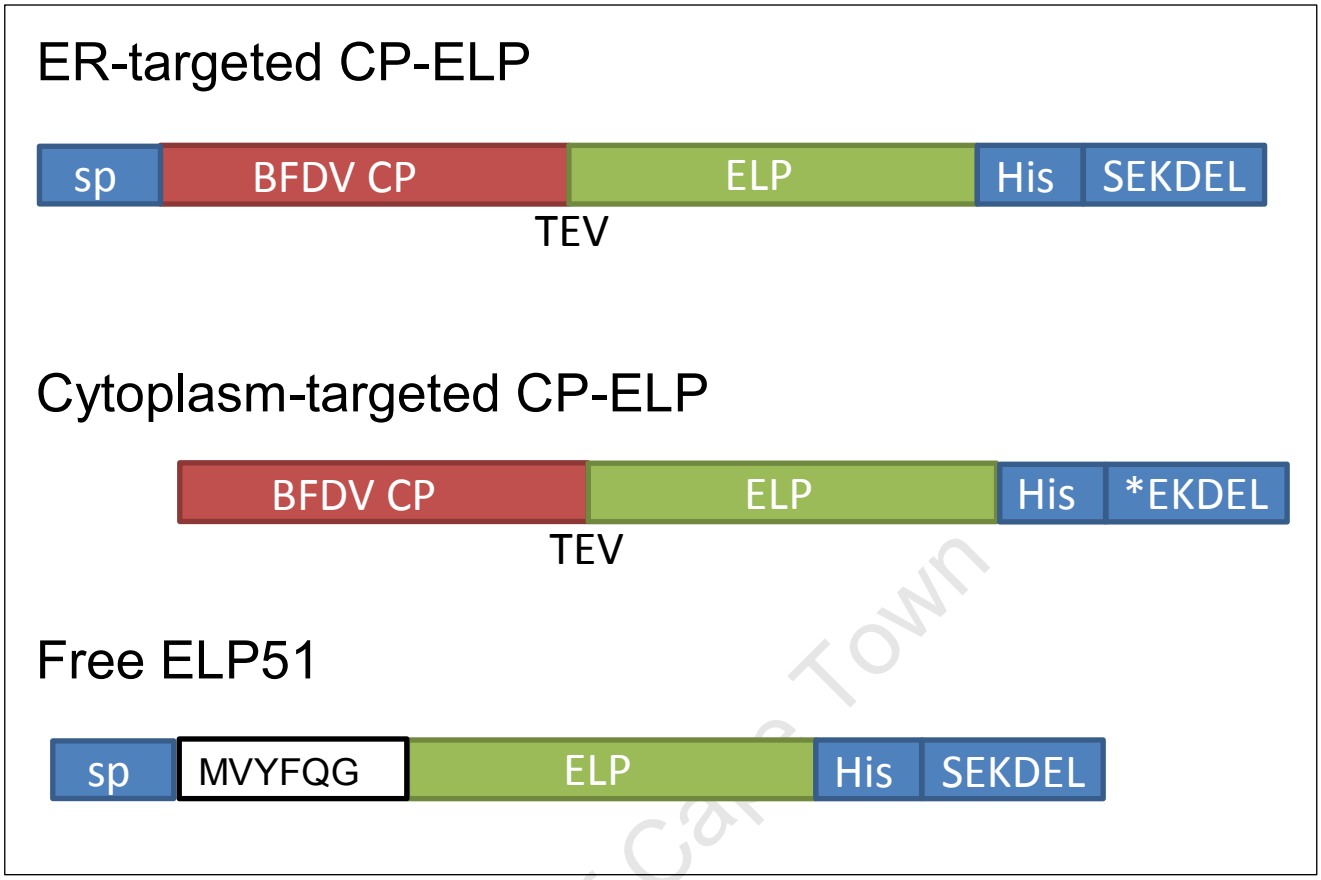


Figure 16: Schematic of proteins expressed in pEAQ-HT

Representation of proteins expressed in pEAQ-HT using different subcellular protein targeting. sp: signal peptide, BFDV CP: full-length or truncated capsid protein (replaced by GFP at this position for GFP-ELP proteins), ELP: 28 x VPGVG or 51 x VPGVG (absent in the case of non-ELPylated BFDV CP), TEV: protease recognition site at the junction of segments, His: His tag, SEKDEL: ER retention signal, (*) indicates a stop codon. The free ELP51 includes a linker peptide MVYFQG between the signal peptide and the ELP. Note that the free ELP51 has no TEV protease recognition site.

3.2.2 Transformation of *A. tumefaciens*

Electrocompetent *Agrobacterium tumefaciens* GV3101::pMP90RK or LBA4044 cells were prepared (Shen & Forde, 1989). A 100µl aliquot of competent cells was mixed with 400ng of recombinant plasmid DNA in a 1 mm gap electroporation cuvette and electroporated using a Bio-Rad GenePulser (1.8kV, 25F, 200Ω) (Bio-Rad Laboratories, USA). The cells were incubated in LB broth for 2 hours at 27°C and were then plated on LB agar plates containing the relevant antibiotics (see Table 4). The plates were incubated at 27°C for 2-3 days. The LBA4044 strain containing the

silencing-suppressor gene *NSs* (of tomato spotted wilt virus) in pBIN (see Table 4) was obtained from culture collection.

Table 4: Antibiotic selection of *Agrobacterium* strains containing different plasmids

A. tumefaciens strain	Plasmids	Antibiotics
GV3101::pMP90RK	pTRAkc-ERH	30 µg/ml kanamycin, 50 µg/ml carbenicillin and 50 µg/ml rifampicin
LBA4044	pBIN-NSs	30 µg/ml kanamycin and 50 µg/ml rifampicin
LBA4044	pEAQ-HT	30 µg/ml kanamycin and 50 µg/ml rifampicin

Colonies were picked from transformed A*grobacterium* plates with a sterile toothpick and transferred to LB broth containing the relevant antbiotics. For the LBA4044 strain, 2 mM $MgSO_4$ was added to prevent cell clumping. Ten millilitre cultures were grown overnight at 27°C. Plasmid DNA was isolated and 'back-transformed' into chemically competent *E. coli* DH5α cells (E. cloni®, Lucigen) and was plated with antibiotic selection to confirm transformation of *Agrobacterium* with the relevant expression construct. *Agrobacterium* cultures were also tested directly by PCR for the presence of the respective constructs using pEAQ-HT-specific primers listed in Table 3.

3.3.3 Transient protein expression in *N. benthamiana* leaves by agroinfiltration

One millilitre of *Agrobacterium* culture was used to inoculate 50 ml of induction media (LB broth with 10 mM MES, 2 mM $MgCl_2$, 1 mM acetosyringone antibiotics, pH 5.6) and was allowed to grow for 16-24 hours in an orbital shaker at 27°C, 120 rpm. *Agrobacterium cells* were pelleted by centrifugation at 4000 x g for 10 min and resuspended in infiltration media (3% (w/v) sucrose, 10 mM acetosyringone, 10 mM MES, 2 mM $MgCl_2$, pH 5.6). For expression with pTRAkc-ERH vector,, the optical density of the cell suspension at 600 nm (OD_{600}) was adjusted to 0.5 with infiltration

media and equal volumes of the two *Agrobacterium* strains harbouring the pTRA and pBIN-NSs constructs (see Table 4) were mixed together. For expression with pEAQ-HT, the OD_{600} of the culture was adjusted to 1.0 with infiltration medium. The cultures were left at room temperature for 1 hour to allow induction of *vir* genes. Leaves from 6-8-week-old *Nicotiana benthamiana* plants were infiltrated with *Agrobacterium* suspensions by injecting the bacterial suspension into the abaxial air spaces from the underside of the leaf, using a 1 ml syringe. The plants were grown at 22 °C under conditions of 16 h light and 8 h dark.

3.2.4 Protein extraction

Three 7mm leaf disks (fresh weight approximately 20mg in total) was homogenised in liquid nitrogen in a 1.5 ml microcentrifuge tube and resuspended in 200 µl extraction buffer (see Table 5 for a list of buffers used. All extraction buffers were supplemented with Roche EDTA-free Complete protease inhibitor cocktail (Roche, USA) before use). Tubes were then centrifuged at 10,000 rpm for 5 minutes in a benchtop Eppendorf 5424 microcentrifuge (Eppendorf, Germany) and the supernatant was recovered. Samples were prepared immediately for SDS-PAGE or stored at – 20 °C.

Expression time trials were carried out to monitor the levels of protein expression over time. Two plants per construct were infiltrated on the same day for each expression time trial. Protein was extracted from plants at 2, 4 and 7 days post-infiltration (dpi), using 200 µL Tris-HCl pH 8.0, 0.1 % (v/v) TritonX-100 (Tris-Triton), per 3 leaf disks. These samples were stored at -20 °C until used for SDS-PAGE.

3.2.5 Total soluble protein assay of plant crude extracts

The TSP assay was performed on crude plant extracts using the Bio-Rad DC protein assay kit according to the manufacturer's instructions, using a BioTek Powerwave XS microplate reader. The standard curve was created by serial dilutions of a 1.43 mg/mL bovine IgH standard solution (Bio-Rad) in 50 mM Tris-HCl pH 8.0, 0.1% (v/v) TritonX-100, which is the same buffer used for protein extraction from plants. The densitometry analysis of immunoblots was done by manual band selection and quantitation using Sygene GeneTools software (Synoptics, UK). The raw volume of any particular band is given in arbitrary units.

Table 5: Extraction buffers used

Buffer abbreviation	Buffer composition
PBS	10 mM phosphate buffered saline pH 7.4, 0.14 M NaCl
PBST	10 mM phosphate-buffered saline pH 7.4, 0.14 M NaCl, 0.1% Tween-20
Tris-Triton	50 mM Tris-HCl pH 8.0, 0.1% Triton X-100
Urea	8 M urea

3.2.6 SDS-PAGE and immunoblotting

SDS-PAGE and immunoblotting were done are described in section 2.2.9. Note that all of the immunoblots shown in this chapter are anti-His-tag immunoblots, done using commercial anti-His-tag primary antibody (Serotech, USA) (at a dilution of 1in 2000) and anti-mouse secondary antibody (at a dilution of 1 in 10,000), conjugated to alkaline phosphatase, as described in 2.2.9.

3.3 Results

The expression of ΔN40 CP-ELP28 in plants was initially performed using pTRAkc-ERH expression vector. This vector targets accumulation of the expressed protein to the endoplasmic reticulum (ER) through the N-terminal signal sequence and C-terminal SEKDEL ER retention signal. ΔN40 CP-ELP28 was expressed in initial experiments with the pTRA vector because the pTRA clone expressing ΔN40 CP-ELP28 was among the first plant expression constructs created. Plants were co-infiltrated with *Agrobacterium* cultures containing the pTRA and pBIN-NSs constucts and protein was extracted at 3 and 5 days post-infiltration (dpi). Expressed ΔN40 CP-ELP28 protein was detected by anti-His immunoblotting, but the soluble protein levels were low. The maximum protein was extracted using 8 M urea (Figure 17, lanes 3 and 6), but extraction was not as effective using other extraction buffers

(Figure 17, lanes 1, 2, 4 and 5). Protein extracted with 8 M urea cannot be used in inverse transition cycling experiments (see Chapter 4), because heating of urea solutions causes carbamylation of proteins (Schokker et al., 1998). The pEAQ-HT vector had been reported to give high yields of recombinant protein (Sainsbury & Lomonossoff, 2008). The pEAQ-HT construct targeting ΔN40 CP-ELP28 to the ER (Figure 19 panel 2 dpi, lane 4 and panel 4 dpi, lane 4) gave higher yields of ΔN40 CP-ELP28, in a shorter time, than the pTRA construct (Figure 17, lane 5), using mild extraction buffers such as Tris-Triton. Using pEAQ-HT is also more convenient because it contains the p19 silencing suppressor gene on the same T-DNA, which means that it does not need to be co-infiltrated with a silencing suppressor plasmid as for pTRA vectors. Therefore, the majority of plant expression work was carried out using the pEAQ-HT vector.

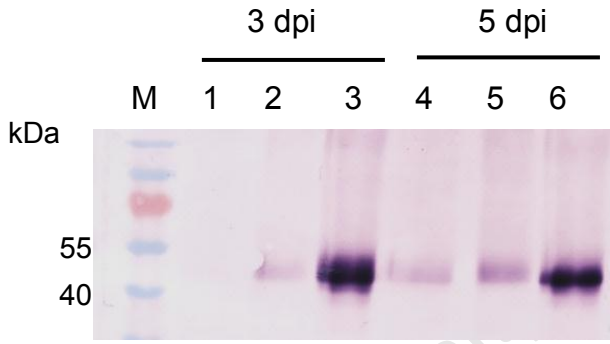


Figure 17: pTRA-ΔN40CP-ELP28 expression time trial

Anti-His-tag immunoblot of pTRA-ΔN40CP-ELP28 expression time trial, which targets ΔN40CP-ELP28 protein to the ER. M: molecular weight marker, lanes 1 and 4: extraction with PBS, lanes 2 and 5: extraction with Tris-Triton buffer, lanes 3 and 6: extraction with 8 M urea. Proteins were extracted at 3 dpi and 5 dpi, as shown. See Table 5 for buffer compositions.

All pEAQ-HT constructs were verified by restriction enzyme digests and sequencing. *Agrobacterium tumefaciens* LBA4044 was transformed with pEAQ-HT constructs by electroporation. The back-transformation as described in 3.2.2 confirmed the presence of pEAQ-HT constructs. *N. benthamiana* leaves were infiltrated with recombinant *A. tumefaciens* containing pEAQ-HT constructs. Initial extractions were done for ΔN40 CP and ΔN40-CP-ELP51 at 2 dpi to optimise the extraction buffer (Figure 18). This showed that extaction with 50 mM TrisHCl buffer with 0.1% (v/v)

TritonX-100 (Tris-Triton) maximised the amount of recombinant protein in crude extracts.

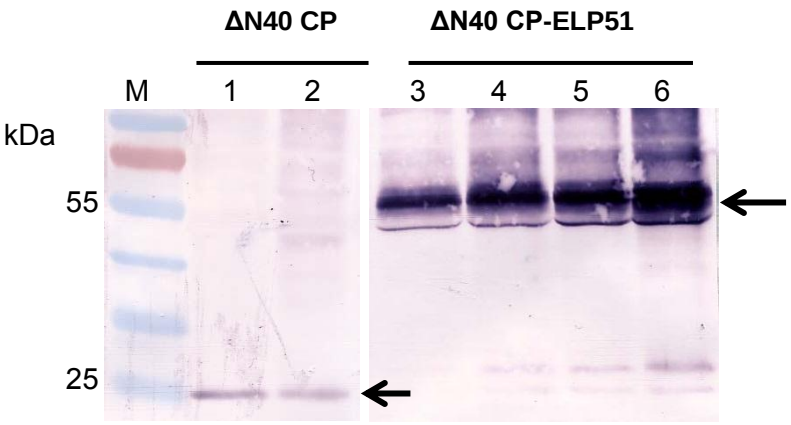


Figure 18: Optimisation of BFDV ΔN40CP protein extraction

ΔN40 CP and ΔN40 CP-ELP51 (indicated by arrows) were extracted at 2dpi as described using different extraction buffers shown in Table 5. Equal volumes of crude extract were analysed by SDS-PAGE and anti-His immunoblot. M: molecular weight marker, Lanes 1 and 2: ΔN40 CP extracted with PBST or Tris-Triton, respectively, Lanes 3-6: ΔN40 CP-ELP51 extracted with PBS, PBST, Tris and Tris-Triton buffers, respectively.

Thereafter, expression time trials were performed for all pEAQ-HT constructs. An expression time trial involves extracting the expressed protein at different time points, over a one-week period, in order to monitor the levels of recombinant protein in the plant leaf over time. Protein was extracted for SDS-PAGE, followed by immunoblots using anti-His-tag antibody. The ΔN40 CP protein time trial is shown in Figure 19 and the full-length CP protein time trial is shown in Figure 20. In both time trials, rapid expression resulted in detection of proteins by two days post-infiltration (dpi). The sizes of unfused BFDV CP proteins were as expected, but the ELPylated proteins exhibited a higher molecular weight (by approximately 5 kDa) than the calculated theoretical weight. Full-length and truncated BFDV CP, each fused to ELP, exhibited the same sizes, being approximately 45 kDa for ELP28 fusions and 55 kDa for ELP51 fusions. This suggests that the ELP has a major influence on the electrophoretic mobility of the fusion protein. The ER-targeted fusions appeared to be slightly larger than their cytoplasmic-targeted counterparts. This may be due to

the signal peptide not being removed correctly, or glycosylation of the protein in the ER. As the time trial progressed, degradation products, of lower molecular weight than the intact proteins indicated by arrows, were evident for all of the expressed proteins, although in general more degradation products were observed for cytoplasm-targeted proteins.

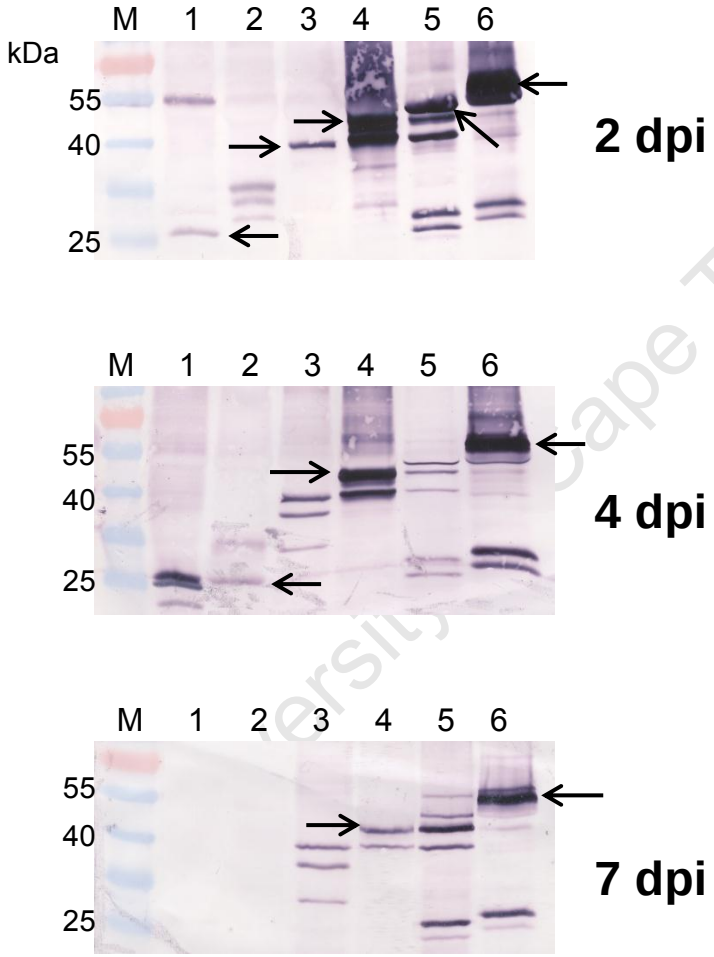


Figure 19: Expression time trial of ΔN40 CP proteins produced in plants

Expression time trial of ΔN40 CP proteins expressed in plants. Protein was harvested at 2 , 4 and 7 dpi as described in 3.2.4 and equal volumes of crude extract were subjected to SDS-PAGE and anti-His-tag immunoblot. For the 4 dpi blot, an equal amount of total soluble protein (70 µg) was loaded for all samples. 1: ΔN40 CP targeted to the cytoplasm, 2: ΔN40 CP targeted to the ER, 3: ΔN40 CP-ELP28 targeted to the cytoplasm, 4: ΔN40 CP-ELP28 targeted to the ER, 5: ΔN40 CP-ELP51 targeted to cytoplasm, 6: ΔN40 CP-ELP51 targeted to the ER. The relevant protein bands are indicated by arrows, where applicable

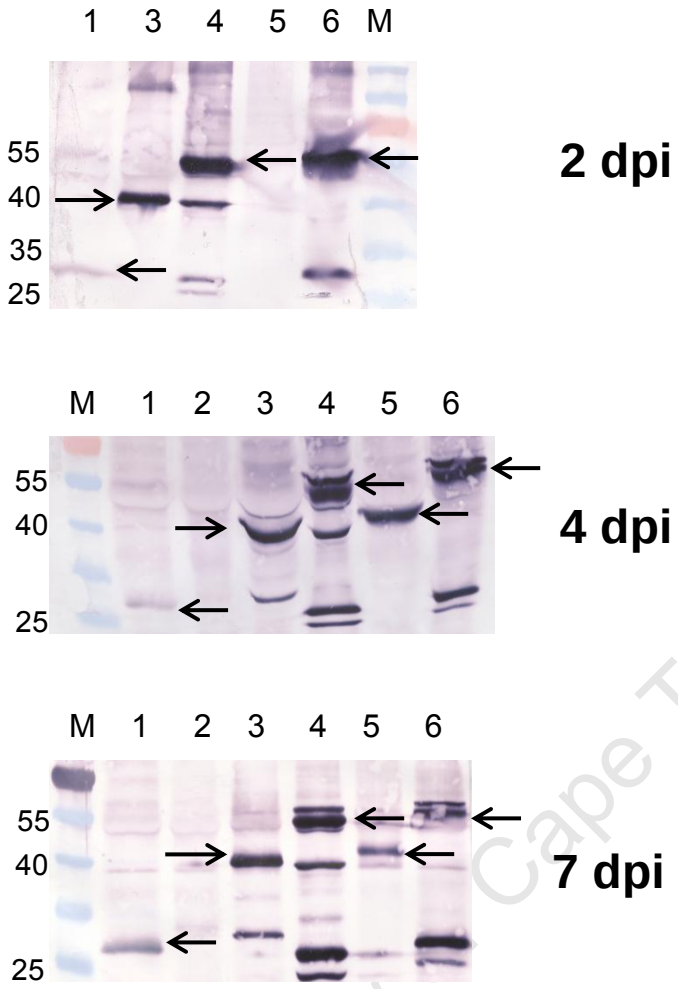


Figure 20: Expression time trial of full-length BFDV CP proteins produced in plants

Expression time trial of full-length BFDV CP proteins expressed in plants. Protein was harvested at 2, 4 and 7 dpi as described in 3.2.4 and equal volumes of crude extract were subjected to SDS-PAGE and anti-His-tag immunoblot. 1: CP targeted to the cytoplasm, 2: CP targeted to the ER, 3: CP-ELP28 targeted to the cytoplasm, 4: CP-ELP51 targeted to the cytoplasm, 5: CP-ELP28 targeted to the ER, 6: CP-ELP51 targeted to the ER. The relevant protein bands are indicated by arrows, where applicable.

The highest yields of ELPylated BFDV CP proteins were compared to that of their unfused counterparts by analysing equal amounts of total soluble protein (70 µg) by anti-His immunoblot, followed by analysis with densitometry software. Expression time trials provided evidence that targeting to the ER decreased the expression of ΔN40 CP and CP but increased that of ΔN40 CP-ELP fusions only. Therefore ΔN40 CP-ELP targeted to the ER and CP-ELP fusions targeted to the cytoplasm were then compared to ΔN40 CP and CP without an ELP tag targeted to the cytoplasm. Three

biological repeats were used for each comparison by harvesting leaf discs from three different infiltrated (upper) leaves of the plant. The results showed that fusion to ELP significantly increased the yield of both full-length and truncated BFDV CP (Figures 21 and 22). Fusion to ELP28 increased the yield of ΔN40 CP approximately three-fold, while fusion to ELP51 increased the yield of ΔN40 CP 4.5-fold (Figure 23). Fusion of full-length CP to either ELP28 or ELP51 increased the yield of CP by approximately three-fold. There was a significant difference in yield between ELP28 and ELP51 fusions for truncated BFDV CP (Figure 21), but this difference was not significant for full-length BFDV CP (Figure 22).

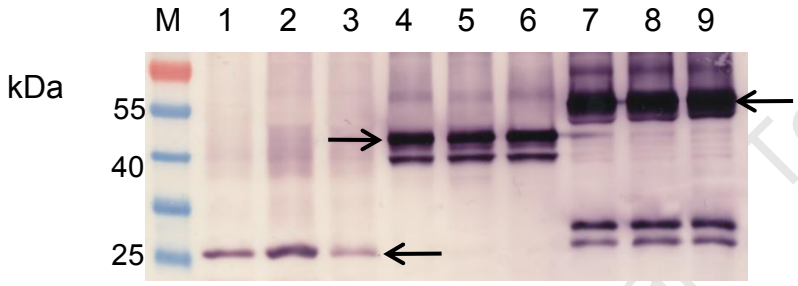


Figure 21: Comparison of ΔN40 CP proteins expressed in plants

Anti-His immunoblot of selected ΔN40 CP proteins extracted 4 dpi. M: molecular weight marker, 1-3: ΔN40 CP targeted to the cytoplasm, 4-6: ΔN40 CP-ELP28 targeted to the ER, 7-9: ΔN40 CP-ELP51 targeted to the ER. Seventy micrograms of total soluble protein was loaded for each lane. Each protein is represented by three different biological repeats. The arrows indicate the bands corresponding to each (undegraded) protein.

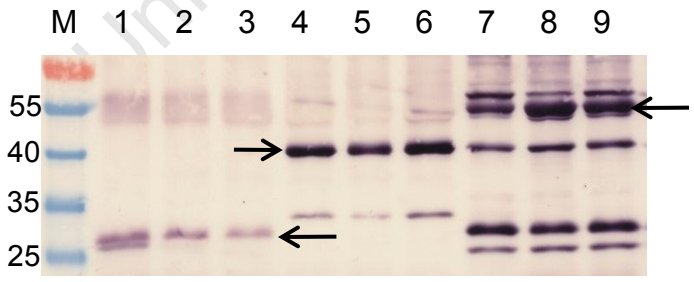


Figure 22: Comparison of BFDV CP proteins expressed in plants

Anti-His immunoblot of selected full-length CP proteins extracted 4 dpi. M: molecular weight marker, 1-3: BFDV CP targeted to the cytoplasm, 4-6: BFDV CP-ELP28 targeted to the cytoplasm, BFDV CP-ELP51 targeted to the cytoplasm. Seventy micrograms of total soluble protein was loaded for each lane. Each protein is represented by three different biological repeats. The arrows indicate the bands corresponding to each (undegraded) protein.

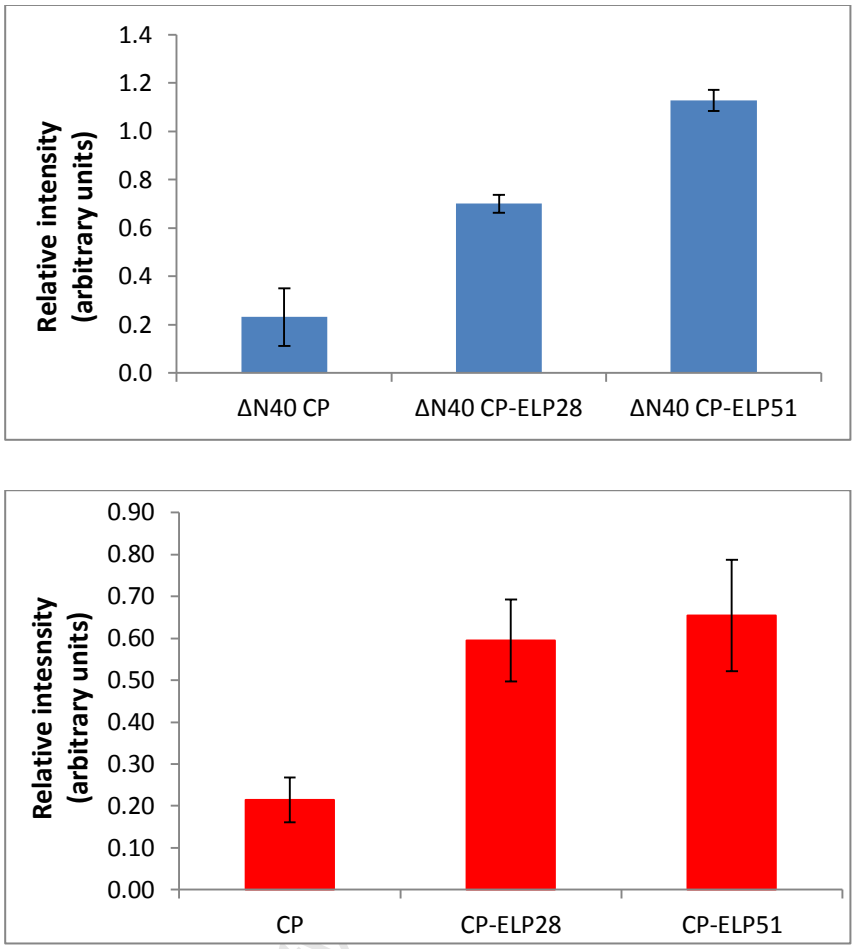


Figure 23: Densitometry analysis of BFDV CP and ΔN40 CP proteins expressed in plants

The bands indicated by the arrows on the immunoblots in Figure 21 and 22 were quantitated using densitometry software. The top graph shows the average relative intensities of the bands for the ΔN40 CP proteins and the bottom graph shows the same for the full-length BFDV CP proteins. The error bars indicate standard deviation.

GFP is a favourite reporter protein because of its high yield when expressed in most systems. pEAQ-HT constructs expressing GFP as a fusion to ELP28 or ELP51 were created (described in 3.2.1), which targeted protein accumulation to the ER. These GFP proteins were expressed because of their anticipated high yield relative to CP-ELP. With greater yield, the total ELP concentration increases, and therefore GFP-ELP51 could possibly be used as a model for purification using the ELP51 fusions by ITC. GFP-ELP28 and GFP-ELP51 were expressed successfully and could be

detected by immunoblot using anti-ΔN40 CP-ELP51 serum (Figure 14B, lanes 1 and 3). GFP-ELP51 was later purified by ITC (see Chapter 4, Figure 27) and used as a standard His-tagged protein in order to quantify the yield of ΔN40 CP-ELP51 targeted to the ER (Figure 24). Using this method, the yield of ΔN40 CP-ELP51 targeted to the ER and extracted 4dpi was calculated at 0.74 ± 0.19 % of the total soluble protein.

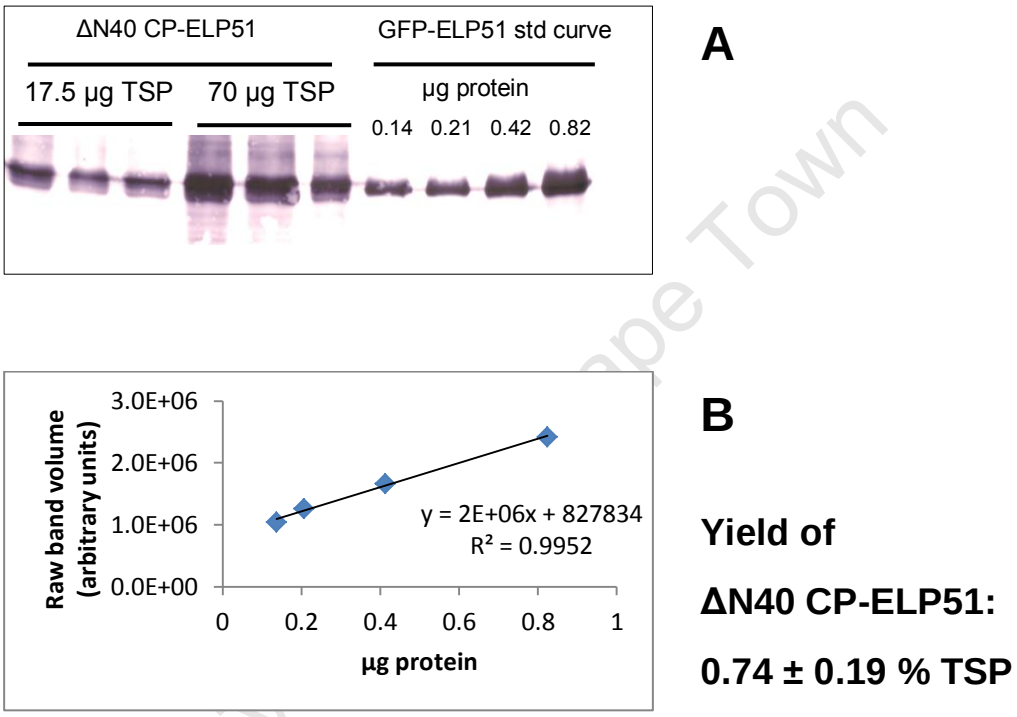


Figure 24: Quantitation of ΔN40 CP-ELP51 by immunoblot and densitometry analysis

A: ΔN40 CP-ELP51 targeted to the ER was extracted 4 dpi and the crude extracts were analysed by SDS-PAGE and anti-His immunoblot. Serial dilutions of purified GFP-ELP51 were analysed concurrently. **B**: Bands were quantified by densitometry software to to create a GFP-ELP51 standard curve and calculate the yield of ΔN40 CP-ELP51.

Based on the successful results of free ELP51 expression in bacteria (Chapter 2), free ELP51 was produced in plants and was intended to assist in the purification of CP-ELP (Chapter 4). In this way, the production and purification of CP-ELP could be exclusively plant-based. The pEAQ-HT construct expressing ELP51 targeted protein accumulation to the ER in order to increase its yield and stability. The protein was

detected by Coomassie staining of a SDS-PAGE gel (approximate size of 26 kDa, Figure 25), and its identity was confirmed by anti-His immunoblot (data not shown).

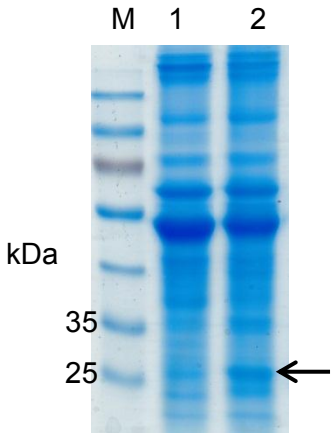


Figure 25: Free ELP51 expressed in plants

Free ELP51 (26 kDa, indicated by the arrow) was extracted from plants at different time points. 1: crude extract from 2 dpi, 2: crude extract from 4 dpi. The crude extracts were analysed by SDS-PAGE and Coomassie staining. The identity of this band was confirmed by immunoblots (data not shown).

3.4 Discussion

Transient expression by agroinfiltration is the fastest way for evaluating the expression of new proteins in plants. Recombinant protein yields using this method are typically higher than those achieved with transgenic plants (Rybicki et al., 2012; Paul & Ma, 2011). The use of this method allowed for the rapid expression of CP-ELP proteins and may also have contributed to increased yield, as transient expression typically gives higher protein yields than transgenic plants (Rybicki, 2010).

The expression of BFDV CP proteins in plants was optimised first by testing two different vector systems, the pTRA system and the pEAQ-HT system. Expression of ΔN40 CP-ELP28 using pTRA targeted protein accumulation to the ER. The amount of protein that could be extracted with Tris-HCl buffer was very low (Figure 17, lanes 2 and 5). The pEAQ-HT vector was then tested to express the same ΔN40 CP-ELP28 protein targeted to the ER. As shown in Figure 19 (panels 2dpi and 4dpi, lane

2), a greater amount of ΔN40 CP-ELP28 was obtained using the same extraction protocol. Therefore the pEAQ-HT vector was used for all further plant expression work.

The plant expression vector pEAQ developed by Sainsbury et al. (2009) has been used for rapid and high-level expression of several proteins in tobacco, including the human papillomavirus coat protein L1 (Matić et al., 2012). The use of the pEAQ-HT vector gave higher yields of L1 protein than produced by a replicating vector based on tobacco mosaic virus (TMV). The situation was reversed when GFP was expressed, suggesting that the particular protein being produced, in conjunction with its expression vector, has a major influence on yield. In this work, ΔN40 CP-ELP and CP-ELP proteins were expressed rapidly using pEAQ-HT, reaching a high level at only 2 days post-infiltration (Figures 19 and 20). The pEAQ-HT vector includes the silencing suppressor gene p19 from tomato bushy stunt virus on the same T-DNA. This system is more efficient than the pTRAkc-ERH system, for which silencing suppressor T-DNA must be provided by co-infiltration with another *Agrobacterium* strain. The p19 protein acts as a potent inhibitor of the plant's post-transcriptional gene silencing (PTGS) defence by sequestering small RNA molecules in a sequence non-specific manner (Lakatos et al., 2004). The suppression of PTGS in this manner has been shown to greatly increase the yield of recombinant proteins (Gelvin, 2003).

The choice of *Agrobacterium* strain also influences the efficacy of DNA transfer and ultimately, all other factors being equal, the protein yield. The pEAQ constructs for expression of CP-ELP proteins were initially transformed into *Agrobacterium* strain GV3101, resulting in successful expression. In previous studies that have made use of the pEAQ vector, the *Agrobacterium* strain LBA4044 was chosen to mediate the transient protein expression in tobacco leaves (Sainsbury et al., 2009; Vardakou et al., 2012). The use of the LBA4044 strain increased the yield of all CP-ELP proteins tested (data not shown) and was adopted for all further work with pEAQ constructs.

The choice of extraction buffer is another important factor influencing the yield of protein. The calculated isoelectric point of CP-ELP proteins is approximately 9.5, therefore buffers with pH range of 7.5 – 8.5 were used in extraction of these proteins to promote solubility, given their slightly negative charge in this pH range. When added to the extraction buffer, the non-ionic detergents Tween-20 and TritonX-100 at

a concentration of 0.1% (v/v) increased the extraction efficiency. It was determined in initial extraction experiments that 50 mM Tris HCl pH 8.0 with 0.1% (v/v) TritonX-100 gave the highest amount of CP-ELP protein (Figure 18).

Recombinant expression of a CP-His-tag fusion and targeting to the ER did not yield any detectable protein when expressed in plants (Figure 19). The expression levels of unfused CP proteins, when targeted to the cytoplasm, were low. For truncated BFDV CP targeted to the ER, some His-tagged proteins were detected, of which the ~26 kDa band may have been ΔN40 CP, but nevertheless the yield was very low and these proteins were not detectable at 7 dpi (Figure 20).

Fusion of ΔN40 CP to ELP, and targeting to the ER, led to a significant increase in the yield of the respective fusion proteins (Figures 21). This could be a result of the formation of ELP protein bodies observed previously in ER-targeted GFP-ELP fusions (Conley et al., 2009b). For both ΔN40 CP-ELP and CP-ELP fusions, cytoplasmic accumulation resulted in more degradation products observed than for ER-targeted proteins (Figures 19 and 20). This could be because proteins in the cytoplasm are more exposed to proteases than in the ER. This was especially true for ΔN40 CP fusions, and may be explained by the observation that truncated CP alone was unstable in the cytoplasm, as it could not be detected at 7 dpi (Figure 19, panel 7 dpi, lane 1). The expression of ΔN40 CP in the cytoplasm reached its peak at 4 dpi, although evidence of degradation was already apparent. Thereafter it was degraded and was undetectable at 7 dpi. The degradation of ΔN40 CP-ELP fusions targeted to the cytoplasm resulted in several degradation products being observed which were smaller than the expected protein sizes. Furthermore, recombinant proteins in the cytoplasm are more likely to have a toxic effect on the host, as opposed to ER accumulation, which segregates the protein from the rest of the cell and renders it physiologically inert (Conley et al., 2011). In contrast to ΔN40 CP proteins, different subcellular targeting seemed to have no effect on the yield of full-length CP fusions. Full-length CP in the cytoplasm continued to accumulate gradually up to 7 dpi, which suggests that it is less susceptible to degradation in the cytoplasm. This would explain the relatively high yield of full-length CP-ELP fusions in the cytoplasm as well.

Based on the results of CP-ELP expression and subcellular localisation, the free ELP51 expressed in plants was targeted to the ER to maximise yield. This protein was to be used in assisting purification of CP-ELP fusions (see Chapter 4). The yield of ELP51 was high at 4 dpi and could be detected on a Coomassie-stained gel (Figure 25, lane 2). There is some evidence that the expression profile of a free ELP is different to respective ELP fusion proteins, and is influenced by ELP length (Conley et al., 2009a). However, expression free ELP51 seemed to follow a similar trend as the CP-ELP51 fusions.

GFP-ELP fusion proteins were expressed in order to demonstrate that the ELP fusion strategy could be used to purify proteins other than BFDV CP-ELP (see Chapter 4 for purification data). GFP was chosen as it is a widely used model protein, known to express at a high level in most systems, and it is highly stable and an easily detectable fluorophore (Grebenok et al., 1997; Harper et al., 1999). GFP-ELP28 and GFP-ELP51 fusion proteins were targeted to the ER to increase yield, as GFP fused to an ELP of 28 VPGVG repeats, and targeting to the ER, has been shown to form novel protein bodies in the ER (Conley et al., 2009b). These proteins were rapidly expressed, reaching a maximum at 4 dpi (Figure 14B, lanes 1 and 3), but remaining stable in the plant leaf up to 7 dpi (data not shown). GFP-ELP51 could be purified by ITC (see Chapter 4) and used as a standard His-tagged protein for immunoblots (Figure 24).

There are no published reports on BFDV or PCV-2 capsid protein expression in plants. Although the ELP fusion strategy has been used for many plant-expressed proteins, ranging from full-size antibodies to spider silk proteins (for a review, see Floss et al., 2010b), there are not many reports on the production of antigen-ELP fusions (for the purpose of creating a vaccine) in any expression system. In one study, *Mycobacterium tuberculosis* antigens were each fused to ELP and expressed in transgenic tobacco (Floss et al., 2010a). Interestingly, the ELP was not removed from the antigen, and it was found that the ELP had no effect on the immune response against the antigens in mice. However, more work needs to be done in order to assess the immunogenicity, or potential adjuvant effect, of ELP when fused to an antigen. The CP-ELP fusions produced in plants could serve as vaccine candidates against BFDV, following purification. The removal of ELP may not be necessary, which would reduce the cost of the vaccine.

Chapter 4: Purification of elastin-like polypeptide fusion proteins

4.1 Introduction

By exploiting the temperature-dependent phase transition of elastin-based polymers, McPherson et al (1996) successfully purified an ELP produced in *E. coli* consisting of 251 repeats of VPGVG. The use of elastin-like polypeptide fusions for purification of recombinant proteins was first described for proteins produced in *E. coli* (Meyer & Chilkoti, 1999), when it was noted that the fusion protein retained the phase transition behaviour of the ELP. Later, this strategy was applied to the plant expression system, to increase the yield and stability of recombinant proteins (Patel et al., 2007; Floss et al., 2008b) while providing a means for simple and scalable purification (Conley et al., 2009a; Phan & Conrad, 2011).

The factors influencing the inverse transition temperature (T_t) of ELPs include: ionic strength of the solution, the length of the ELP and hydrophobicity of the guest residues (Meyer & Chilkoti, 1999). The total ELP concentration is also a critical factor influencing T_t and the overall successful recovery of the fusion protein. This of considerable importance to plant expressed ELP fusions, for which yields may be lower than those produced by microbial systems. Fusion to longer ELPs increases the purification efficiency by ITC, but decreases the yield of the fusion protein (Meyer et al., 2001). Therefore, choosing the length of the ELP fusion partner requires a compromise between yield of the target protein and the purification efficiency.

This chapter describes the purification of elastin-like polypeptide fusions expressed in *E. coli* and plants, using ITC. These include ELP fusions of GFP or BFDV CP proteins and free ELP51 expressed in both *E. coli* and plants. Free ELP51 was expressed in order optimise the conditions required for the phase transition of this ELP51 and its fusions. Free ELP51 was also used as a co-aggregate during purification of CP-ELP fusions, increasing the total ELP concentration and improving the purification efficiency.

4.2 Materials and Methods

4.2.1 Proteins

The proteins that were used in ITC experiments are: ΔN40 CP-ELP51 and free ELP51 produced in bacteria (Figure 5) using the relevant pPROEX-HT constructs (bacterial protein production and extraction described in 2.2.5 and 2.2.6) and GFP-ELP51, free ELP51 and ΔN40 CP-ELP51 (ER-targeted) expressed in plants (Figure 16), using the relevant pEAQ-HT constructs (agroinfiltration described in 3.3.3). Only ELP51 and its fusions were purified by ITC, because ELP51 (which is longer than ELP28) has a lower transition temperature than ELP28, which means that a lower temperature is required for ELP51 purification. This is an important factor in preventing thermal denaturation of ELP fusion proteins during purification.

4.2.2 Bulk extraction of ELP fusion proteins from plants

For large-scale extractions, leaves from infiltrated plants were pooled, weighed and homogenised in liquid nitrogen using a mortar and pestle. Two volumes of extraction buffer (w/v) were added. Samples were then centrifuged at 20,000 x g in a Beckman-Coulter Avanti® J25-I centrifuge (Beckman-Coulter, USA) for 20 minutes. The supernatant was filtered through two layers of Miracloth™ (EMD Chemicals, USA). The sample was centrifuged again at 20,000 x g for 20 minutes to remove any remaining insoluble matter. Protein samples were stored at – 20 °C. Total soluble protein concentrations were determined using a Bio-Rad Protein Assay Kit (Bio-Rad Laboratories, USA) as per the manufacturer's instructions, using a IgG standard curve and a BioTek microplate reader (BioTek, USA).

4.2.3 Inverse transition cycling (ITC) by centrifugation (cITC)

Stock salt solution (3 M ammonium sulphate or 5 M sodium chloride) was added to the protein sample in 1 mL increments, with constant mixing, until the desired salt concentration was achieved (the optimised salt concentration of 0.4 M ammonium sulphate was most often used, unless specified otherwise). For 'spiked' ITC, in which free ELP51 was added to the sample as a co-aggregate, a final ELP51 concentration

in excess of 5 µM was used. In these cases, the free ELP51 was added to the sample prior to the addition of salt solution. The sample was then pre-incubated in a water bath at the specified temperature (the optimised temperature of 37 °C was most often used, unless specified otherwise) for 10 minutes to trigger the phase transition of the ELP. Samples were then centrifuged in a Beckman Coulter Optima L-100 XP ultracentrifuge pre-heated to the specified temperature for 40 minutes at 20,000 x g. This step is termed 'hot spin'. The supernatant was removed and the pellet resuspended in 1/10th sample volume of ice-cold 50 mM Tris-HCl-pH 8.0, 0.1% TritonX-100. In some cases, mechanical grinding of the pellet with a pestle in a 1.5 mL plastic tube was required for complete resuspension. To further break up protein aggregates, the resuspended pellet was sonicated on ice using a Misonix 3000 sonicator (Misonix, USA) for four 15 second pulses at 12 W. The sample was then incubated at 4 °C on an orbital shaker for approximately 30 minutes to allow for solubilisation of ELP proteins. The resuspended pellet sample was centrifuged at 15,000 x g at 4 °C for 10 minutes to remove insoluble debris. This step is termed 'cold spin' (CS). The supernatant resulting from the cold spin step was recovered. The pellet was resuspended in 1/10th sample volume of 50 mM Tris-HCl-pH 8.0, 0.1% TritonX-100 and kept for further analysis. Where specified, ELP fusions were digested using AcTEV protease (Invitrogen, USA), with the manufacturer's recommended reaction setup at 27 °C for 16 hours.

4.2.4 Membrane filtration-based ITC (mITC)

Five millilitres of crude extract clarified by centrifugation was pre-filtered through 0.2 µm Acrodisc® syringe filters (with Supor® membrane (hydrophilic polyethersulfone)) (Pall Corporation, USA). Free ELP51 was added in the case of spiked ITC as described previously. Ammonium sulphate was added to a final concentration of 0.4 M. Samples were incubated in a 37 °C water bath for 15 minutes to induce ELP aggregation. Samples were then filtered through 0.2 µm syringe filters, using a sterile syringe, and the filtrate was kept for further analysis. An equal volume of pre-warmed 0.5 M ammonium sulphate solution was passed through the membrane in a washing step. Ice-cold 50 mM Tris-HCl pH 8, 0.1 % (v/v) TritonX-100 was added to the syringe, without a plunger, fixed to the filter. This was allowed to pass through the filter by gravity flow for 4-16 hours at 4 °C. The elution volume was 1/5 of the crude

extract volume; if this volume had not passed through the filter by gravity flow after the specified time, a plunger was used to force the solution through the filter until the correct volume had been eluted.

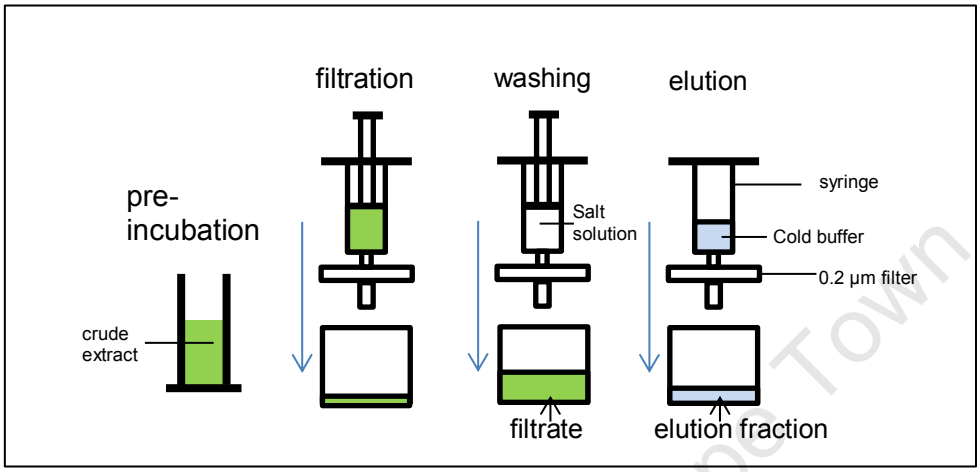


Figure 26: Flow diagram of mITC protocol

The crude extract is first pre-incubated at 37 °C in the presence of ammonium sulphate (0.4 M) to cause ELP aggregation. Filtration through the membrane captures ELP aggregates on the membrane while contaminating proteins pass through in the filtrate. The membrane is washed by passing a salt solution through the membrane. The ELP-tagged proteins are then eluted from the membrane by passing cold buffer through it. The purified ELP fusion protein is now located in the elution fraction.

4.2.5 SDS-PAGE and immunoblotting

SDS-PAGE and immunoblotting were done as described in section 2.2.9.

4.3 Results

4.3.1 Purification of GFP-ELP51

GFP-ELP51 could be purified using cITC, which was recovered to a high degree in the cold spin supernatant. However, it is likely that some Rubisco large subunit protein, which is similar in size to GFP-ELP51, remained (Figure 27 A, lane 2). The

purified protein was digested with TEV protease, to separate GFP from its ELP51 fusion partner. The digestion was partial, with the the 55 kDa band being reduced in volume, and the unfused GFP could be seen on a Coomassie stained gel (Figure 27B, lane 2). The size of ELP51 (26 kDa) is similar to that of GFP (30 kDa), but ELP51 probably is not contributing to the intensity of the band indicated by the arrow in Figure 27B, as it was later found that free ELP51 it is poorly visualised on Coomassie-stained gels (Figure 30, lane 2). Preliminary experiments done at 37 °C, 0.4 M ammonium sulphate, showed that a substantial amount of GFP-ELP51 was left in the hot spin supernatant, indicating that not all of the protein was precipitated efficiently, leading to a low recovery of GFP-ELP51 (data not shown). The centrifugation temperature and the salt concentration were increased in order to cause more of the GFP-ELP51 to aggregate and pellet during centrifugation. The use of a high salt concentration (1 M ammonium sulphate) and 40 °C resulted in higher recovery of GFP-ELP51 after purification (Figure 27 A, lane 2). However, these conditions may have increased the amount of contaminating proteins, especially Rubisco large subunit protein, that co-precipitated with GFP-ELP51 during purification.

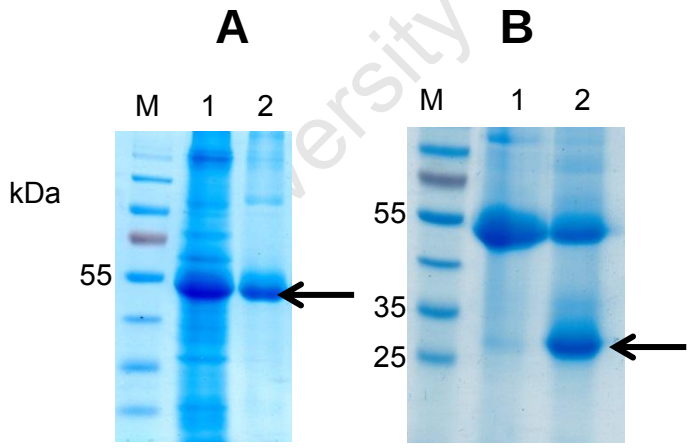


Figure 27: Plant-produced GFP-ELP51 purification by cITC

GFP-ELP51 was purified by centrifugation ITC, using 1 M ammonium sulphate and 40 °C centrifugation temperature. **A:** M: molecular weight marker, 1: GFP-ELP51 crude extract, 2: supernatant from the cold spin step. GFP-51 band is indicated by the arrow. **B:** The supernatant from the cold spin step was partially digested by overnight incubation with TEV protease. 1: undigested control, 2: TEV digest. Note the GFP band of approximately 30 kDa (indicated by the arrow).

ELP fusions may also be purified by membrane filtration, instead of centrifugation, to capture ELP aggregates. ITC by membrane filtration is referred to here as 'mITC'. This method was used to purify GFP-ELP51 to homogeneity in only one purification cycle (Figure 28). This was done as described in 4.2.4, except that the pre-incubation temperature used was 42 °C. The pre-incubation temperature was increased from 37 °C to 42 °C based on the results of cITC experiments of GFP-ELP51. Contaminating proteins passed though the membrane in the filtrate (Figure 28A, lane 3), while GFP-ELP51 was successfully captured on the membrane and subsequently eluted with cold buffer (Figure 28A, lane 2). The identity of GFP-ELP51 was confirmed by anti-His immunoblot (Figure 28B). The immunoblot showed that much of the GFP-ELP51 protein was lost in the filtrate. The purification of GFP-ELP51 was not optimised further (to reduce this loss), because it could not be used as a model for ΔN40 CP-ELP51 purification as originally intended. Nevertheless, purification of GFP-ELP51 demonstrated the proof of concept for the purification of ELP51 fusions by ITC. The results suggest that GFP-ELP51 aggregates more effectively at higher temperatures (42 °C) and higher salt concentrations. These conditions are different to those required for effective ΔN40 CP-ELP51 aggregation, as determined by later empirical experiments.

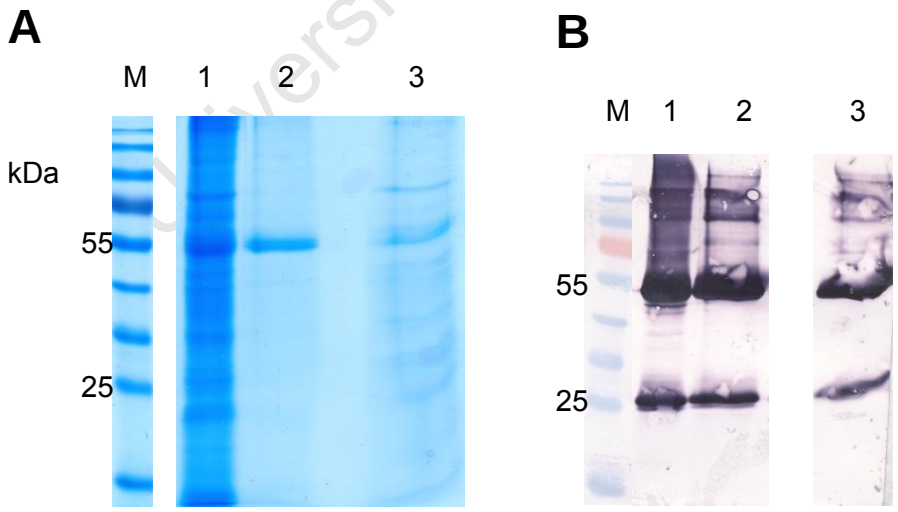


Figure 28: Purification of GFP-ELP51 by mITC

GFP-ELP51 expressed in plants was purified by one cycle of the membrane ITC protocol as described, with the addition of 0.5 M ammonium sulphate and pre-incubation at 42 °C. **A**: Coomassie-stained gel showing: lane 1: the crude extract, lane 2: the elution fraction, lane 3: the filtrate. **B** is the cognate anti-His immunoblot.

4.3.2 Purification of free ELP51

The free ELP51 produced in bacteria (Figure 29) or plants (Figure 30) could both be purified by ITC. The purification of free ELP51 was done in order to determine the purification conditions later used for ΔN40 CP-ELP51. These ITC conditions were determined as 0.4 M ammonium sulphate and 37 °C centrifugation temperature. ELP51 extracted from *E. coli* was subjected to one round of ITC, in triplicate experiments (Figure 29). In the case of bacterial-expressed ELP51, there was a loss of protein – due to irreversible aggregation – during ITC (Figure 29, lanes 4-6). This could be ELP51 that was improperly folded or processed. All soluble ELP recovered after the first ITC cycle (Figure 29, lanes 1-3) exhibited high recovery in subsequent ITC cycles. The ELP51 purified from plants could be visualised on a Coomassie-stained gel (Figure 30), although it was poorly stained in comparison to other proteins. This phenomenon has been described previously (McPherson et al., 1996). The ELP51 produced in bacteria had more extraneous sequences on the N-terminus (an additional His tag), which may be responsible for its superior staining on a Coomassie gel (Figure 29, lanes 1-3). The purification of plant-produced free ELP51 confirmed that ammonium sulphate was the more effective salt to use during purification, compared to NaCl. A reduced amount of ELP51 was recovered when NaCl was used (Figure 30, lane 7), compared to when ammonium sulphate was used (Figure 30, lanes 2, 4 or 6).

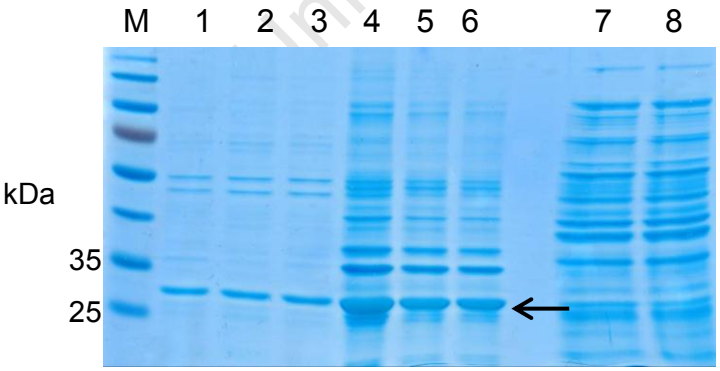


Figure 29: Free ELP51 purification from *E. coli* cell lysate

Free ELP51 purified from *E. coli* cell lysate by one round of cITC: M: molecular weight marker, lanes 1-3: soluble proteins after purification, lanes 4-6: insoluble proteins after purification, lanes 7 and 8: crude lysates. The purification was performed in triplicate. Equal volumes of all samples were loaded.

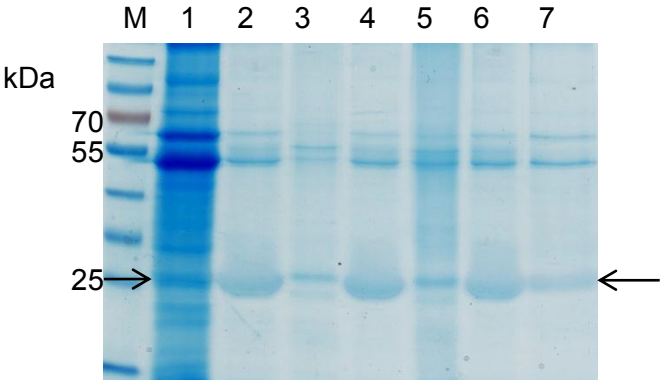


Figure 30: Plant-produced ELP51 expression and purification

Optimisation of free ELP51 cITC purification produced in plants at 37 °C, using three different salt conditions. M: molecular weight marker, 1: crude extract (ELP band indicated by the arrow), 2: (0.3 M ammonium sulphate) soluble proteins, 3: (0.3 M ammonium sulphate) insoluble proteins, 4 and 6: (0.4 M ammonium sulphate) soluble proteins, 5: (0.4 M ammonium sulphate) insoluble proteins, 7: (2.5 M NaCl) soluble proteins. The ELP51 band is indicated by arrows. Equal volumes of all samples were loaded.

4.3.3 Purification of ΔN40 CP-ELP51

In order to establish purification conditions for ΔN40 CP-ELP51, this protein was expressed in bacteria. The ΔN40 CP-ELP51 protein expressed in bacteria was mostly insoluble (Figure 10A, lane 6). The protein band for ΔN40 CP-ELP51 in the crude lysate is not clearly visible on a Coomassie-stained gel, in contrast to the insoluble fraction (the pellet of cell debris obtained after centrifugation of the cell lysate) (Figure 31, lane 2). ITC purification requires soluble protein. However, an attempt at ITC was made using the soluble fraction of expressed ΔN40 CP-ELP51, using the ITC conditions previously determined for free ELP51. The ΔN40 CP-ELP51 could be concentrated and purified by ITC, and the protein was recovered in the pellet of the cold spin step (Figure 31, lanes 3 and 4). The pellet from the cold spin step was resuspended in 1/40th of the soluble fraction volume. However, the aggregation was irreversible, as none of the ΔN40 CP-ELP51 remained soluble after ITC. The identity of the bands indicated by the arrow in Figure 31 was confirmed to be ΔN40 CP-ELP51 by anti-His immunoblot (data not shown)

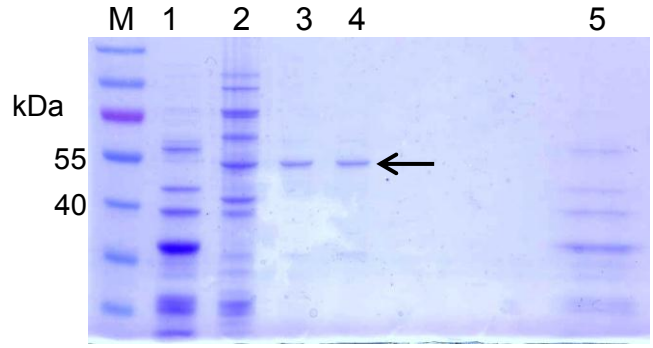


Figure 31: Purification of ΔN40 CP-ELP51 from *E. coli* cell lysate

Purification of ΔN40 CP-ELP51 from *E. coli* cell lysate by cITC: M: molecular weight marker, 1: soluble cell lysate, 2: insoluble proteins from cell lysate, 3 and 4: resuspended pellets from the cold spin step of purification, 5: supernatant from the hot spin step. Equal volumes of all samples were loaded.

The majority of purification experiments of CP-ELP from plants were performed using the ER-targeted ΔN40 CP-ELP51. The yield of this protein in crude extracts was consistently one of the highest of the CP-ELP proteins. Also, the longer ELP of this fusion protein (51 x VPGVG) reduces the temperature necessary for the ELP aggregation (Meyer & Chilkoti, 1999), which could be important in preventing temperature-induced denaturation of BFDV ΔN40 CP. Under ITC conditions of 0.8 M ammonium sulphate and 40 °C, the fusion protein was shown to be pelleted effectively, as a large amount of the protein as detected in the pellet (Figure 32, lane 2). However, the resolubilisation of the pellet was poor, as there was a reduced amount of the protein in the cold spin supernatant (Figure 32, lane 3). The protein level in the cold spin supernatant was very low and while it could be detected by immunoblot, it was not detectable by Coomassie staining. Some of the ΔN40 CP-ELP51 had not been pelleted during the hot spin step and remained in the supernatant (Figure 32, lane 4).

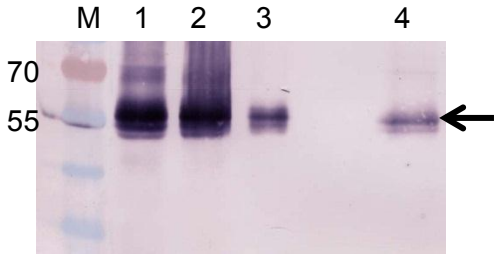


Figure 32: Plant-produced ΔN40 CP-ELP51 ITC

ΔN40 CP-ELP51 produced in plants was subjected to cITC followed by SDS-PAGE and anti-His immunoblot. M: molecular weight marker, 1: crude extract, 2: resuspended pellet after heated centrifugation, 3: supernatant after cold spin step, 4: supernatant after hot spin step. The ΔN40 CP-ELP51 band (55 kDa) is indicated by the arrow. Equal volumes of all samples were loaded.

The combination of elevated temperatures and high salt concentrations may cause the denaturation of CP. This would lead to irreversible aggregation, should the CP target proteins come into contact with each other. Therefore, milder ITC conditions were tested. The salt concentration was lowered to 0.4 M and the centrifugation temperature was set to 37 °C. The resuspended pellet was also sonicated at 4 °C to disrupt any persistent ELP aggregates. The first round of purification was performed with four technical repeats. The cold spin supernatants resulting from the first round of purification were pooled for the second round of purification. This revised protocol led to a greater recovery of purified soluble ΔN40 CP-ELP51, as is evident from the greater amount of ΔN40 CP-ELP51 in the cold spin supernatants (Figure 33B, lanes 2 and 3). The greater proportion of ΔN40 CP-ELP51 that was recovered using the new purification conditions (Figure 33B, from lane 1 to lane 2) can be compared with the low recovery of previous conditions (Figure 32, from lane 1 to lane 3). The majority of contaminants were removed during the first ITC cycle, and remained in the hot spin supernatant (Figure 33A, lane 5). Some ΔN40 CP-ELP51 was not pelleted and remained in the hot spin supernatant (Figure 33B, lane 5). However, irreversible aggregation remained the major cause of ΔN40 CP-ELP51 protein loss. This is evident from the amount of ΔN40 CP-ELP51 remaining in the insoluble debris of the cold spin pellet, Figure 33B, lane 4).

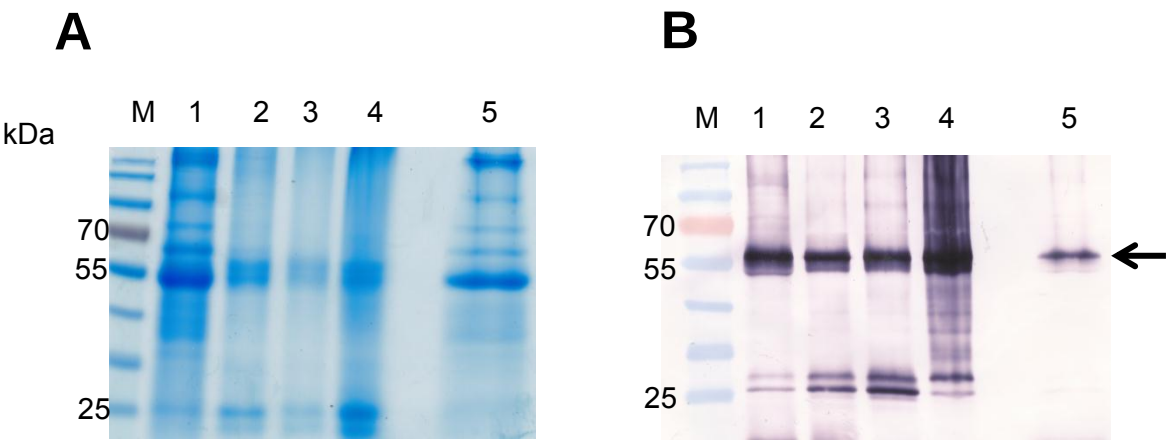


Figure 33: Purification of ΔN40-CP-ELP51 using optimised ITC protocol

ΔN40-CP-ELP51 extracted from plants was purified by two rounds of cITC using the optimised protocol, which involved centrifugation at 37 °C and sonication of the ELP pellet. **A:** Coomassie-stained gel; M: molecular weight marker, 1: crude extract, 2: supernatant from first cold spin, 3: supernatant from the second cold spin, 4: resuspended pellet of insoluble matter from the second cold spin, 5: supernatant from the first hot spin. **B** is the cognate anti-His immunoblot, which was loaded with half the volume of all samples in **A**. The position of the ΔN40-CP-ELP51 band is indicated by the arrow. Equal volumes of all samples were loaded.

'Spiked' ITC of ΔN40 CP-ELP51

The purification conditions of free ELP51 and ΔN40 CP-ELP51 were matched (37 °C, 0.4 M ammonium sulphate) so that co-aggregation could occur if the two proteins were mixed. Free ELP51expressed in bacteria was purified by two ITC cycles and was added to crude plant extracts containing ΔN40 CP-ELP51 protein. However, the free ELP51 (Figure 34A, lane 2 (26 kDa)) expressed in bacteria exhibited a band of approximately 55 kDa when analysed by anti-His immunoblot (Figure 34B, lane 2), which is similar in size to the ΔN40 CP-ELP51 fusion (Figure 34B, lane 1). It is suspected that this is an ELP dimer that failed to dissociate under the conditions of denaturing SDS-PAGE. Higher bands indicate the presence of higher molecular weight ELP multimers as well (Figure 34B, lane 2 and 3). Therefore it was not possible to distinguish between the ELP dimer band and the band corresponding to ΔN40 CP-ELP51 protein on an anti-His immunoblot, when the two proteins were mixed. The ELP51 expressed in plants exhibited similar multimers on SDS-PAGE gels. It should be noted that the addition of free ELP51 increased the pelleting efficiency of ΔN40 CP-ELP51, as the amount of ΔN40 CP-ELP51 present in the hot

spin supernatant (Figure 33B, lane 5) was reduced relative to that of unspiked purification experiments.

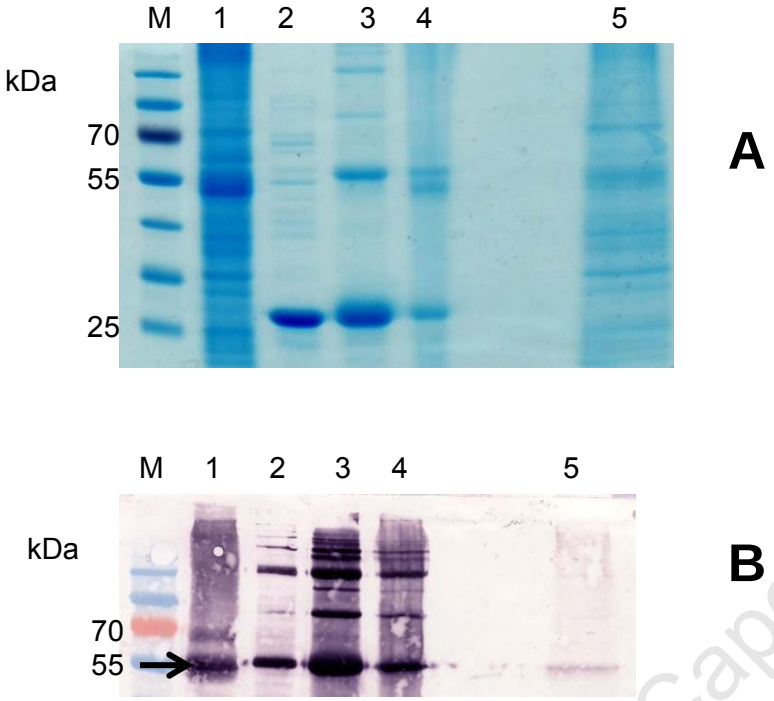


Figure 34: ΔN40 CP-ELP51 ITC spiked with bacterial ELP51

cITC of ΔN40 CP-ELP51 produced in plants with added free ELP51 produced in *E. coli*. **A**: Coomassie-stained gel; M: molecular weight marker, 1: ΔN40 CP-ELP51 crude extract (protein indicated by the arrow), 2: free ELP51 added, purified by two rounds of ITC from *E. coli* cell lysate, 3: supernatant from the cold spin step, 4: pellet from the cold spin step, 5: supernatant from the hot spin step. **B**: cognate anti-His immunoblot. Note that the 55 kDa band for the ELP51 is a dimer.

The spiked ITC was repeated to determine the effect of free ELP51 in the purification of ΔN40 CP-ELP51. Purification of ΔN40 CP-ELP51 was performed without the addition of free ELP51 ('unspiked'), or free ELP purified by two ITC cycles (Figure 35A, lane 9) was added ('spiked'). Samples of the pellets resulting from the hot spin step were analysed, as well as the cold spin supernatants. The results showed that the ΔN40 CP-ELP51 was pelleted in unspiked purification (Figure 35B, lanes 2 and 4), but was lost in the cold spin step, as there is reduced amounts of the protein in the cold spin supernatants (Figure 35B, lanes 3 and 5). The ΔN40 CP-ELP51 band could be seen on a Coomassie-stained gel (Figure 35A, lane 2, indicated by the red arrow). For spiked purification, there is a similar band in the cold spin supernatant,

(Figure 35A, lane 7). This was thought to be ΔN40 CP-ELP51, but this could not be confirmed as there is an ELP51 dimer of similar size (Figure 35B, lane 8 and 9). In conclusion, the effect of free ELP51 on the purification of ΔN40 CP-ELP51 could only be determined by immunoblot, using an antibody that could detect ΔN40 CP but not the free ELP51 dimer. The anti-ΔN40 CP serum was used for this purpose (Figure 36 B).

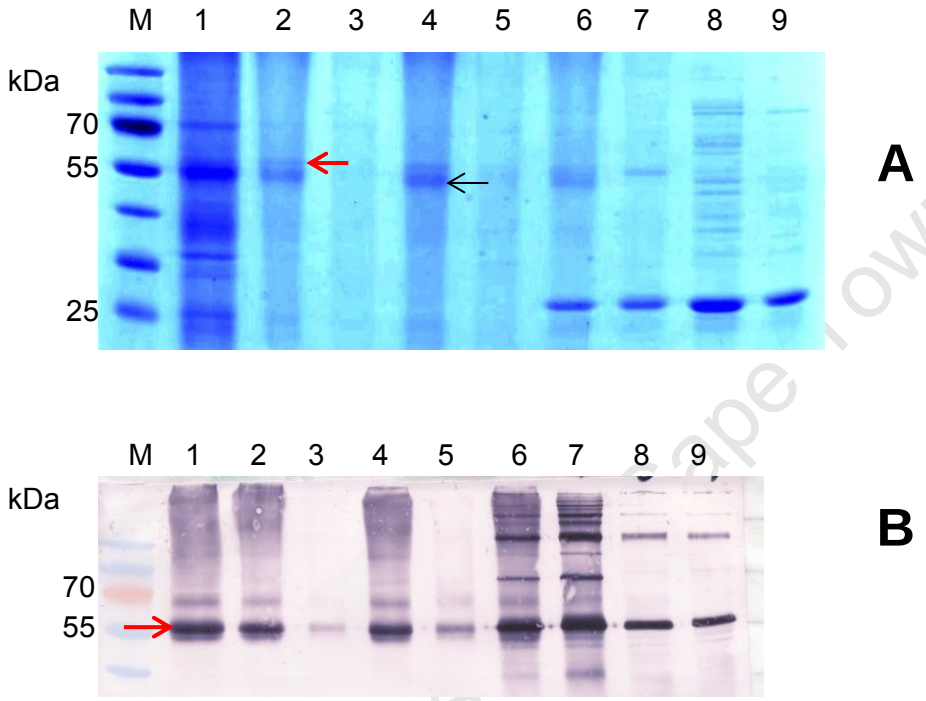


Figure 35: Plant produced ΔN40-ELP51: spiked or unspiked ITC and free ELP51 purification

cITC of ΔN40 CP-ELP51 produced in plants without added free ELP51 (”unspiked”) or with added free ELP51 produced in *E. coli* (“spiked”). The unspiked purification was performed in duplicate. **A**: Coomassie-stained gel; M: molecular weight marker, 1: ΔN40 CP-ELP51 crude extract, 2 and 4: (unspiked) resuspended pellet after hot spin step, 3 and 5: (unspiked) supernatant from cold spin step, 6: (spiked) resuspended pellet from cold spin step, 7: (spiked) supernatant from cold spin step. The red arrow indicates the ΔN40 CP-ELP51 band and the black arrow indicates the Rubisco large subunit band. 8 and 9 represent one and two cycles, respectively, of cITC purification of free ELP51 (25 kDa) from *E. coli* cell lysate. **B** is the cognate anti-His immunoblot. Equal volumes were loaded for each sample in lanes 1-7.

Spiked purification of ΔN40 CP-ELP51 was performed by adding free ELP51 (expressed in plants) to ΔN40 CP-ELP51 crude extract. The cold spin supernatant and the insoluble proteins from the cold spin step were analysed by Coomassie staining and immunoblot using the anti-ΔN40 CP serum. The anti-ΔN40 CP serum could be used to detect exclusively ΔN40 CP-ELP51 in a mixture of free ELP51 and ΔN40 CP-ELP51 (Figure 36B). There was no detection of ELP51 (circled in Figure 36A) by the serum (Figure 36B, lane 2). The results showed that the majority of the ΔN40 CP-ELP51 was insoluble after purification, as most of it was located in the cold spin pellet (Figure 36B, lane 4). Only a small amount of ΔN40 CP-ELP51 was detected in the cold spin supernatant after purification, (Figure 36B, lane 3). Therefore the addition of free ELP51, which itself exhibited high recovery, could not reduce the irreversible aggregation of ΔN40 CP-ELP51.

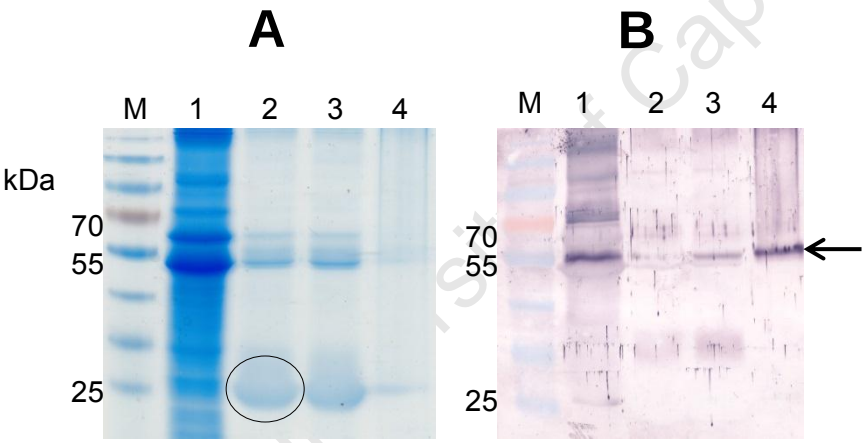


Figure 36: cITC of ΔN40 CP-ELP51 spiked with plant-produced free ELP51

cITC of ΔN40 CP-ELP51 produced in plants with added plant-produced free ELP51. **A**: Coomassie-stained gel; M: molecular weight marker, 1: crude ΔN40 CP-ELP51, 2: purified ELP51 (the free ELP51 band is circled), 3: supernatant from the cold spin step, 4: resuspended pellet from the cold spin step. **B** is the cognate anti-ΔN40 CP immunoblot. The ΔN40 CP-ELP51 band is indicated by the arrow. Note the absence of free ELP51 detection by anti-ΔN40 CP serum. Equal volumes of all samples were loaded.

The mITC method was employed for the purification of ΔN40-CP-ELP51 produced in plants (as described in 4.2.3). The bulk of contaminating proteins were removed in

the filtrates (Figure 37A, lanes 4 and 5) and ΔN40-CP-ELP51 was concentrated by eluting in 1/5th of the sample volume (Figure 37B, lanes 2 and 3). Using this method, the amount of soluble ΔN40-CP-ELP51 recovered was increased, because the concentration of the protein was evident (compare Figure 37B lane 1 to lanes 2 and 3), in contrast to the cITC method. However, a substantial amount of ΔN40-CP-ELP51 was lost in in the filtrates (Figure 37B, lanes 4 and 5). In two separate experiments, free ELP51 was added to the ΔN40-CP-ELP51 prior to mITC (i.e. spiked mITC) In these cases, no His-tagged protein was detected in the filtrate (Figure 37B, lanes 6 and 7) suggesting that more stable particles were formed, and prevented detectable loss of protein in the filtrate. The presence of ΔN40 CP-ELP51 in the elution fraction of the spiked mITC experiment was confirmed by anti-ΔN40 CP immunoblot (Figure 15).

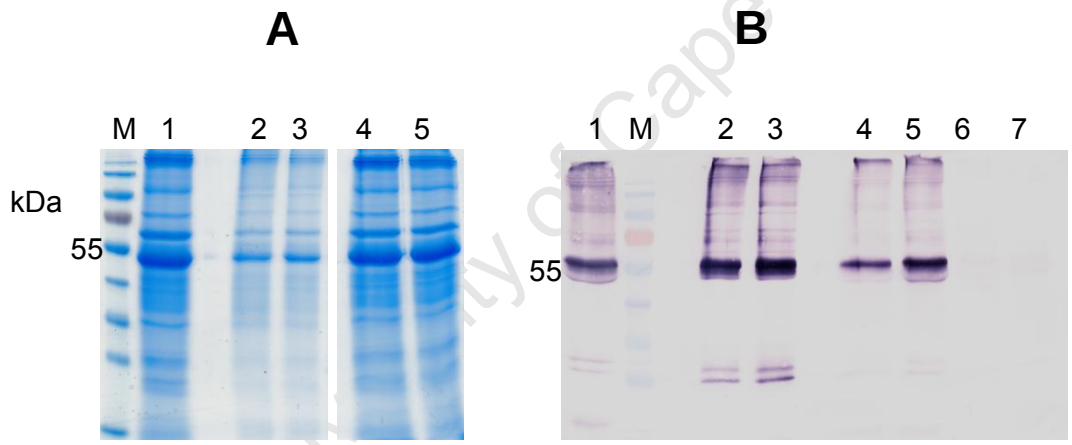


Figure 37: Purification of ΔN40 CP-ELP51 by mITC

Plant-produced ΔN40 CP-ELP51 was subjected to mITC. The experiment was performed in duplicate. **A**: Coomassie-stained gel: M: molecular weight marker, 1: crude extract, 2 and 3: elution fractions from the two experiments, 4 and 5: filtrates from the two expermiments. **B**: anti-His immunoblot analysis. Lanes 6 and 7 are filtrates from two separate mITC experiments in which free ELP51 was added as a co-aggregate. The elution fraction resulting from one of these spiked mITC experiments was analysed by anti-ΔN40 CP immunoblot as shown in Figure 15, lane 4. Equal volumes of all samples were loaded.

4.4 Discussion

GFP-ELP51 was expressed in order to serve as a model ELP51 fusion protein. This protein was expressed at a high level and was highly stable, and therefore a large amount of this protein could be generated in order to test the purification conditions of a typical ELP51 fusion protein. GFP-ELP51 could be efficiently purified by ITC. Similar plant-expressed GFP-ELP51 fusions described in the literature also exhibit a relatively high recovery after ITC (Conley et al., 2009a). This is probably because of the high stability and solubility of GFP and the high yield of the fusion protein. One cycle of cITC produced a single band on a Coomassie-stained gel, but it is likely that some Rubisco large subunit protein is also present, which is approximately the same size of GFP-ELP51 (Figure 26). The conditions used for purification of GFP-ELP51 were a high salt concentration of 1 M ammonium sulphate and 40 °C centrifugation temperature. The salt concentration and temperature are greater than those used to later purify free ELP51 or ΔN40 CP-ELP51. It is known that GFP increases the transition temperature $(T)_t$ of an ELP fusion partner due to the prevalence of hydrophilic residues on the surface of GFP (Trabbic-Carlson et al., 2004b). Therefore, GFP-ELP51 could not be used as a model for the purification of ΔN40 CP-ELP51 because GFP-ELP51 required a higher temperature for its purification. Nevertheless, the successful purification of GFP-ELP51 verified the utility of the ELP51 fusion in protein purification. ITC by membrane filtration was more successful than cITC, as GFP-ELP51 was purified to homogeneity in only one cycle (Figure 28) and this protein was used as a standard ELP- and His-tagged protein for immunoblots, and also for the optimisation of TEV cleavage reactions. Although the purification of GFP-ELP51 by membrane ITC was not optimised, the degree of purity after one cycle of mITC is much greater than that of other plant-expressed GFP-ELP proteins described by Conley et al. (2009a). The authors used cITC to purify GFP fused to ELPs of varying lengths, but the majority of contaminants still remained after one round of purification.

The pioneering work on the environmentally-responsive nature of protein polymers based on elastin was carried out by Dan Urry (Urry et al., 1991; Urry, 1997). Free

ELP expressed in bacteria could be purified to a high degree after only one ITC cycle (McPherson et al., 1996). When this was later applied to the purification of ELP fusions, it was noted that purification efficiency of the fusion was typically less than that of the free ELP (Meyer & Chilkoti, 2002b). This suggests that the target protein fused to the ELP can have an effect of the purification efficiency, and the degree of this effect may vary for different target proteins. The bacterial-expressed free ELP51 described in this chapter exhibited irreversible aggregation during the first cycle of ITC (Figure 29), but in later cycles exhibited high recovery (Figure 35A, lanes 8 and 9). This suggests that a proportion of the ELP expressed in bacteria was not properly folded, as irreversible aggregation is not among the expected physical characteristics of an ELP. Free ELP expressed in plants was observed by Conley et al. (2009a) to have a higher purification efficiency than ELP fusion proteins. The purification of free ELP51 produced in plants (Figure 30) was very efficient and exhibited a higher recovery during the first ITC cycle than the bacterial-produced ELP (Figure 29,). Also shown in Figure 30 is that the use of ammonium sulphate to depress T_t leads to much greater recovery than when of sodium chloride is used (Figure 30, lane 7). Fong et al. (2009) described the use of low concentrations of ammonium sulphate during ITC, instead of high NaCl concentrations, to increase ELP recovery and minimise co-precipitation of impurities.

The purification of ΔN40 CP-ELP51 produced in plants presented the same problem as in the attempt to purify bacterial-expressed ΔN40 CP-ELP51. The fusion protein aggregated during ITC and was pelleted effectively, but the protein in the pellet remained largely insoluble. The protein was not recovered in the cold spin supernatant and therefore additional rounds of purification were not possible. The efficient purification of free ELP suggests that BFDV CP is the cause of this problem. BFDV capsid proteins naturally interact to form the capsid of the virion. The strong interaction between ΔN40 CP proteins to form a virus-like particle (VLP) could be a factor that leads to irreversible aggregation of ΔN40 CP-ELP51, when at high concentrations in the pellet. Modification of the purification conditions and the addition of a sonication step led to increased recovery of ΔN40 CP-ELP51 (Figure 33). When over-expressed in bacteria, the ΔN40 CP protein alone was mostly insoluble, in contrast to free ELP51 protein. This suggests that ΔN40 CP in higher concentrations contributes to the insolubility of ΔN40 CP-ELP51. BFDV CP

expressed in the baculovirus insect cell system was also found to be largely insoluble (Stewart et al., 2007).

The addition of a free ELP to successfully assist in the purification of poorly-expressed ELP fusions was described for protein expressed in bacteria (Christensen et al., 2007). An ELP with matched T_t is added to the crude extract containing the target ELP fusion before commencing with ITC. The rationale behind this strategy is to increase the total ELP concentration and thus promote the aggregation of all ELP proteins, including the target ELP fusion. The target ELP fusion can then be concentrated and purified effectively. Once protease digestion (or intein cleavage) has separated the ELP from the target protein, all ELP can be separated from it by one additional ITC cycle. The spiked experiments of plant-produced ΔN40 CP-ELP51 described in this chapter showed that addition of free ELP51 reduced the amount of ΔN40 CP-ELP51 left behind in the hot spin supernatant (in the case of cITC) or filtrate (in the case of mITC). This is evidence that an increased concentration of total ELP leads to more efficient purification. In this case, the ΔN40 CP-ELP51 co-aggregated with the free ELP51 and formed hybrid aggregates more efficiently than could be achieved with ΔN40 CP-ELP51 alone.

This strategy was used to recover as little as 10 pM of thioredoxin-ELP fusion from crude extracts, which, without the addition of free ELP, would have been impossible by ITC (Ge & Filipe, 2006). In addition, The free ELP protein may increase the efficiency of purification by preventing contact between target proteins, which may result in irreversible aggregation (Meyer & Chilkoti, 2002b). It is possible that contact between BFDV CP proteins is responsible for the irreversible aggregation of CP-ELP fusions. Free ELP51 was added in an attempt to mitigate this problem – the more free ELP51 in the aggregates, the less chance of contact between BFDV CP proteins. However, the addition of free ELP did not reduce the amount of irreversible aggregation of ΔN40 CP-ELP51, as most of the protein was located in the cold spin pellet after purification (Figure 36B, lane 4).

Spiked ITC was done with the assumption that the free ELP51 and ΔN40 CP-ELP51 had the same transition temperature (T_t). Trabbic-Carlson et al. (2004b) showed that the mean surface hydrophobicity of the target protein fused to an ELP can have an effect on the T_t of the ELP. Proteins with a high proportion of surface hydrophobic

residues tend to depress T_t, meaning that a lower temperature is required for aggregation of the ELP fusion relative to the free ELP. Conversely, more charged proteins elevate T_t. It is not clear whether BFDV CP could modulate the T_t of the fused ELP51, as the 3D structure of BFDV CP has not been solved. Studies have shown that mixtures of ELPs with different transition temperatures aggregate independently of each other as a function of temperature (Meyer & Chilkoti, 2002a). However, the use of the same conditions to pellet both free ELP51 and CP-ELP51 suggests that the use of ELP51 to act as a coaggregate is valid.

The sizes of the aggregates formed by ELP molecules above T_t are greater than one micron in diameter (Meyer et al., 2001). This presents another method in which to purify ELP-tagged proteins. Filtration by 0.2 µm membrane may be used to capture ELP aggregates, while contaminating proteins are passed through in the filtrate (Ge et al., 2006). When cold buffer is passed over the membrane, the aggregation of ELP is reversed and the protein may be eluted from the membrane. This method has been used for plant-expressed ELP fusions and was found to reduce the irreversible aggregation of ELP fusions that was observed for ITC by centrifugation (Phan & Conrad, 2011; Tian & Sun, 2011). Phan & Conrad purified influenza antigens haemagglutinin (H5) and neuraminidase (N1), each fused to an ELP of 100 repeats and expressed transiently in tobacco leaves. Tian & Sun purified a lectin protein, fused to an ELP of 60 repeats, produced in transgenic rice. This method, termed 'mITC' (for membrane ITC), was used in the purification of ΔN40 CP-ELP proteins, in an attempt to mitigate the irreversible aggregation that was observed for ITC by centrifugation.

GFP-ELP51 was first used to test the mITC purification method and could be purified efficiently (Figure 28). ΔN40 CP-ELP51 purification was then attempted under similar conditions (Figure 37). The recovery of soluble ΔN40 CP-ELP51 was greater than for centrifugation ITC. However, there were significant protein losses in the filtrate, which suggests that a proportion of the ΔN40 CP-ELP did not aggregate completely to form micron-sized particles, or that the aggregation was partially reversed upon filtration. Filtration was carried out at room temperature immediately after removing the sample from the 37 °C water bath, but some cooling may have occurred, which could have caused some of the ΔN40 CP-ELP aggregation to be reversed. When free ELP51 was added to act as a co-aggregate, no His-tagged protein was

observed in the filtrate (Figure 37, lanes 6 and 7), suggesting that the co-aggregation increased the proportion of stable micron-sized particles. Membrane ITC with added free ELP has been used to concentrate and purify ELP fusions expressed at very low levels in *E. coli* (Christensen et al., 2007). This was shown by diluting ^{14}C-labelled bacterial-expressed proteins to 15-20 pmol per liter and tracking their purification by scintillation counting. The labelled proteins were recovered after purification by mITC only when free ELP51 was added.

Based on these results, future production of the ΔN40 CP-ELP51 vaccine candidate should be performed by transient expression in tobacco and the purification should be done by several cycles of mITC 'spiked' with free ELP51. The amount of crude extract processed should be scaled up by using membranes with larger surface areas and by using vacuum pumps to increase the filtration pressure. Fortunately, membrane filtration is already well-established in industry and this process is easily scalable (Ge et al., 2006). The purification of ELP fusions of full-length BFDV CP could also be optimised in future work, based on the results presented here for ΔN40 CP-ELP51. The removal of ELP from the ΔN40 CP-ELP51 vaccine candidate may not be necessary, as ELPs are biocompatible (Rincón et al., 2006), which would further reduce the cost of the vaccine.

Chapter 5: Towards the development of a candidate BFDV DNA vaccine

5.1 Introduction

With the advent of recombinant DNA technology, the evolution of vaccines has progressed from inactivated and attenuated virus to subunit vaccines and virus-like particles. The latter have several advantages including safety and rapid response (El-Attar et al., 2009; Toussaint et al., 2011). Delivery of recombinant DNA encoding an antigen directly to the host is another strategy that can be used for vaccination. DNA vaccines consist of a bacterial plasmid DNA backbone containing an antigen-encoding gene under the control of a strong eukaryotic promoter (Doria-Rose, 2003; Hitzeroth et al., 2009).

There are several modes of delivery of DNA vaccines including injection or electroporation. Intramuscular electroporation has become the preferred route of administration for animals, as it has proven to be more immunogenic than other methods (Van Drunen Littel-van den Hurk & Hannaman, 2010). The plasmid is taken up by muscle cells and monocytes and the antigen is expressed. Antigens are processed and displayed on the cell surface by MHC class I molecules, which initiates the adaptive immune response. DNA vaccines offer several advantages, including flexibility in design and rapid preparation of plasmid DNA, which allows for a rapid response against emerging viral strains (Saade & Petrovsky, 2012).

DNA immunisation of mice against PCV-2 has proven to be effective in providing protection (Fu et al., 2011). This suggests that a DNA vaccine against BFDV could be a good vaccine candidate. The expression of HIV and *Plasmodium falciparum* antigens using pTH eukaryotic expression vector has shown to induce CTL responses in mice (Hanke et al., 1998). A pTH clone containing the BFDV *ΔN40 CP* gene (pTH-ΔN40 CP) had been previously constructed in our laboratory by Guy Regnard, but had yet not been evaluated for its ability to express ΔN40 CP in transfected cells.

This chapter describes the expression of truncated BFDV CP (ΔN40 CP) in the human embryonic kidney cell line HEK293T and in chicken embryonic fibroblast primary cells, by transfection with pTH-ΔN40 CP. ΔN40 CP was expressed instead of full-length CP, as it was expected that ΔN40 CP would be expressed at a higher level. The expression of BFDV ΔN40 CP by the pTH vector *in vitro* is a requirement for further challenge experiments of the DNA vaccine in animals.

5.2 Materials and Methods

5.2.2 Plasmid

The truncated BFDV CP, including the N-terminal His-tag, was isolated from pPROEX-ΔN40 CP, with the addition of 5' *Hind*III and 3' *EcoR*I sites, by PCR. The PCR product was digested with *Hind*III and *Eco*RI and ligated into the similarly digested pTH expression vector (Appendix B) (Hanke et al., 1998). The ligation product was transformed into *E. coli* DH5α cells (E. cloni®, Lucigen). This pTH-ΔN40 CP clone was provided by Guy Regnard. The insertion of ΔN40 CP gene into the pTH mammalian expression vector was confirmed by restriction digests, after which the sequence was verified (Macrogen, Netherlands) using vector specific primers. Endotoxin-free plasmid DNA was prepared using the NucleoBond® Xtra Midi EF kit (Macherey-Nagel, USA) according to the manufacturer's protocol.

5.2.3 Cell culture

HEK293T human embryonic kidney cells were obtained from the ATCC (ATCC® number CRL-11268™). These cells were grown in Dulbecco's modified Eagle medium (DMEM) with GlutaMAX™, 5% glucose (Invitrogen, USA), supplemented with 10% foetal calf serum (FCS) (Gibco, USA). Penicillin-streptomycin 100x (Sigma-Aldrich, USA) was added to media to a final concentration of 10 mL/L to control bacterial contamination. HEK293T cells were grown in a 37°C incubator with CO_2 control. All work was carried out aseptically in a BSL-2 Ebsco laminar flow hood.

Chicken embryonic fibroblast (CEF) primary cells were isolated from 11-day old embryonated chicken eggs. Foetuses were removed from the eggs and the head limbs were amputated using sterilised scissors and forceps. The foetus body was washed in sterile PBS twice and was then incubated in 0.125% Trypsin-EDTA (Sigma-Aldrich, USA) at 37 °C for 20 minutes. The foetus was removed and the cell suspension was filtered through sterile cheesecloth. The cells were pelleted by centrifugation at 1000 x g. The supernatant was discarded and the cell pellet was resuspended in DMEM containing 5% FCS. The cells were counted and $2x10^6$ were seeded in 75 cm^2 tissue culture flask (Corning, USA), with 10 ml of DMEM, 5% FCS. The cells were grown for 3 days in a CO_2-controlled incubator at 37°C, before passaging. The preparation of primary cells of from embryonated eggs complied with animal ethics regulations (animal ethics number 010/012).

5.2.4 Transfection and cell harvesting

Cells were seeded in 6-well plates in a total volume of 2 ml DMEM, 10% FCS in order to obtain 70-90% confluency in 24 hours. Fugene® 6 transfection reagent (Roche, USA) was used in initial transfection experiments, using the manufacturer's protocol. Thereafter, cells were transfected with ExtremeGene HP transfection reagent (Roche, USA), using the manufacturer's quick protocol, with a ratio of 3 µl reagent to 1µg pTH plasmid DNA. For negative controls, the DNA was replaced with sterile water. All transfections were performed in duplicate. As a positive control, the reporter construct pLuc-SV40 (luciferase gene under the control of SV40 promoter) was transfected separately, using 3 µl of transfection reagent and 50 ng plasmid DNA per well. Cells were incubated for 72 hours before harvesting. Cells were dislodged from the wells by washing with PBS, and pelleted by centrifugation. The cells were resuspended in 1 x Glo Lysis buffer (Promega, USA) supplemented with EDTA-free protease inhibitor (Roche). These samples were then prepared directly for SDS-PAGE by the addition of SDS loading buffer.

5.2.5 SDS-PAGE and immunoblotting

SDS-PAGE and immunoblots were done as described in 2.2.9. Note that all the immunoblots presented in this chapter are anti-His-tag immunoblots.

5.3 Results

The insertion of ΔN40 CP into expression vector pTH was confirmed by restriction digests with *Nco*I and *Xho*I (Figure 38). The sequence was then verified.

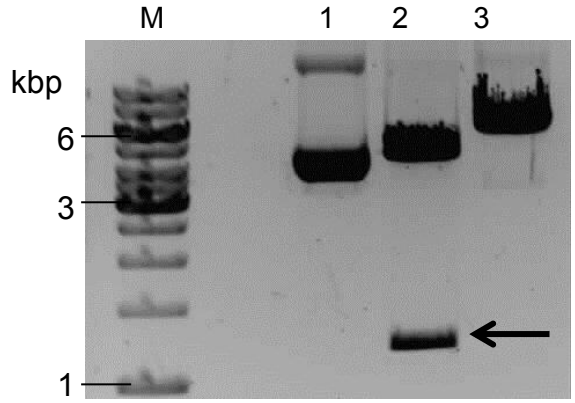


Figure 38: Restriction enzyme digests of pTH-ΔN40 CP to confirm insert and plasmid integrity

1: undigested pTH DNA, 2: *Nco*I digest, excising insert of approximately 1.3 kbp (indicated by the arrow), 3: *Xho*I digest, generating linearised vector of approximately 6 kbp.

HEK293T cells were transfected with pTH-ΔN40 CP and whole cell lysates were subjected to anti-6xHis immunoblot analysis (Figure 39). Optimisation of transfection was performed by varying the transfection reagent (Fugene™ 6, Roche, USA) to DNA ratio. Ratios of Fugene® (µl) to DNA (µg) of 3:1 and 3:2 did not yield detectable expression of CP. There was no detectable expression of CP after 24 or 48 hours (data not shown). Only with a 6:1 ratio and 72 hours incubation was CP expression detected (Figure 39, lane 3). For negative controls, the cells were transfected with a mixture that had sterile water added instead of DNA. The transfection was repeated with X-tremeGene HP tranfection reagent (Roche) using a 3:1 reagent to DNA ratio, as recommended by the manufacturer. In this case, expression was detected after 5 days incubation (Figure 40, lanes 3 and 4).

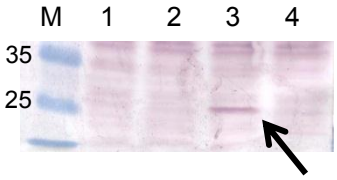


Figure 39: Optimisation of transfection in HEK293T cells

Anti-His immunoblot analysis of 72 hours after transfection of HEK293T cells with Fugene® 6 transfection reagent using different ratios of transfection reagent (µl) to DNA (µg). 1: 3:1 ratio, 2: 3.2 ratio, 3: 6:1 ratio, 4: negative control. The ΔN40 CP band is indicated by the arrow (25 kDa).

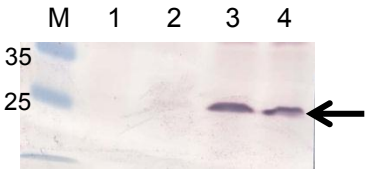


Figure 40: HEK293T cells 5 days post-transfection

Anti-His Immunoblot analysis of HEK293T cells transfected with with X-tremeGENE HP transfection reagent. The transfections were performed in duplicate. Whole cell lysates were analysed five days post-transfection, 1 and 2: negative controls, 3 and 4: transfected cells. The ΔN40 CP band is indicated by the arrow (25 kDa).

The prepared CEF cells were seeded for transfection after no more than two passages. Based on the previous work in HEK293T cells, the CEF cells were transfected with a 3:1 ratio of X-tremeGENE HP reagent to pTH-ΔN40 CP plasmid DNA. Additional cells were transfected separately with the reporter construct pLuc-SV40 using 3 µl of transfection reagent and 50 ng plasmid DNA per well. This was done in order to measure the efficiency of the transfection procedure and the capacity of the prepared CEF cells for protein expression. High luciferase readings were obtained at 3 days post-transfection (data not shown).

The expression of ΔN40 CP in transfected CEF cells was detected by anti-His immunoblot of whole cell lysates after 4 days incubation (Figure 41, lanes 3 and 4). Soluble proteins in cell lysates taken three days post transfection did not contain any detectable ΔN40 CP (Figure 41, lanes 1 and 2), suggesting that the expressed ΔN40 CP was insoluble. The transfection was repeated and only whole cell lysates were analysed by SDS-PAGE and anti-His immunoblot, based on the previous result.

Expression of ΔN40 CP was detected in CEF whole cell lysate at two days post transfection (Figure 42, lanes 1 and 2). After five days post transfection, the cell layer had begun to detach from the culture plate. Cells harvested five days after transfection did not contain detectable ΔN40 CP (Figure 42, lanes 3 and 4).

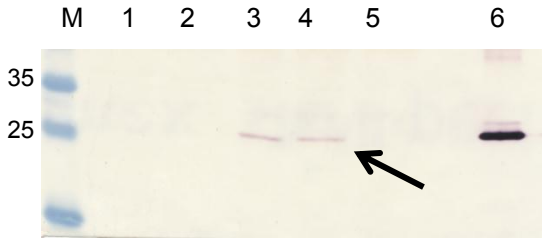


Figure 41: Transfection of CEF cells

Anti-His immunoblot analysis of transfection of CEF cells. M: molecular weight marker, 1 and 2: soluble proteins in cell lysate 3 days post-transection, 3 and 4: whole cells 4 days post-transfection, 5: negative control, 6: positive control (ΔN40 CP expressed in *E. coli*). The ΔN40 CP band is indicated by the arrow (25 kDa).

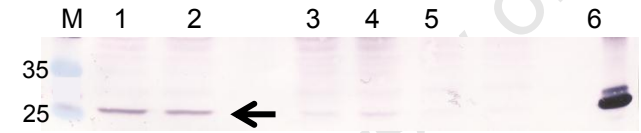


Figure 42: Repeat transfection of CEF cells

Anti-His immunoblot analysis of transfection of CEF cells. M: molecular weight marker, 1 and 2: whole cells 2 days post-transfection, 3 and 4: whole cells 5 days post transfection, 5: negative control, 6: positive control (ΔN40 CP expressed in *E. coli*). The ΔN40 CP band is indicated by the arrow (25 kDa).

5.4 Discussion

DNA vaccines expressing the coat protein of PCV-2 have been shown to be immunogenic in mice (Shen et al., 2008; Fu et al., 2011). The capsid protein is the major antigenic determinant of PCV-2 and BFDV. Therefore DNA vaccines expressing CP should be promising vaccine candidates for BFDV as well. In one study of PCV-2 vaccination, DNA vaccination yielded comparable humoral responses to recombinant subunit CP vaccination, but the DNA vaccine generated a superior CD8 T+ cell response (Shen et al., 2008). There are currently no published reports of DNA vaccination for BFDV. The success of PCV-2 DNA vaccines suggests that this may be a viable vaccination strategy for BFDV.

Transfection of cells *in vitro* with candidate DNA vaccines is done to confirm the expression of the antigen. Truncated BFDV CP is expressed as an N-terminal His-tag fusion. The human embryonic kidney cell line HEK293T was transfected with the candidate BFDV DNA vaccine pTH-ΔN40 CP, and ΔN40 CP was detected by anti-His immunoblot after 72 hours. As there is no cell culture system for psitaccine species, chicken embryonic fibroblasts (CEF) were also used for confirmation of CP expression, as these are more closely related to cells naturally infected by BFDV. The CEF cells were prepared from chicken embryos by trypsin treatment of the embryo surface. Therefore a mixture of dermal cells and fibroblasts may have been isolated, as there was no selection applied for either cell type. The fibroblasts have a typical elongated and branching morphology. After successive passages, cells with this type of morphology predominated, suggesting that the culture conditions favoured the proliferation of the fibroblasts.

Resuspended whole cells were loaded directly (with the addition of SDS sample buffer) for SDS-PAGE followed by anti-His immunoblots, because no protein was detected in the soluble part of the lysate at three days post-transfection when the lysate was cleared by centrifugation (Figure 41). This suggests that the majority of the expressed ΔN40 CP was insoluble. The expression of full-length BFDV CP in insect cells also yielded protein that was mostly insoluble, with most of the CP detected in the cell pellet after lysis (Stewart et al., 2007). ΔN40 CP was detected after 2 and 4 days post-transfection. In a repeated transfection of CEF cells, no

protein was detected after 5 days, but this is probably due to most of the cells having died and detached from the wells. ΔN40 CP was detected in HEK cells after 5 days when the viable cell monolayer was still present. Anti-ΔN40 CP serum was used to probe for the expressed protein, but due to high background and the low level of ΔN40 CP as determined by anti-His immunoblots, the ΔN40 CP band could not be identified (data not shown).

DNA vaccines can be manufactured on a large scale, and the simplicity of plasmid purification reduces the cost of downstream processing. Immune responses elicited by DNA vaccines may be 'boosted' by subsequent inoculation with a subunit or virus-like particle vaccines (Doria-Rose, 2003). This has been shown to elicit broader neutralising antibody responses in the case of H5N1 influenza (Lin et al., 2012). The BFDV ΔN40 CP DNA vaccine described here could be used in such a 'prime-boost' regimen in conjunction with a BFDV CP subunit vaccine or VLP vaccine. Plant-produced BFDV CP, or one of the CP-ELP vaccine candidates described in Chapter 3, could be used as a subunit vaccine. The cost of a DNA vaccine with a plant-produced subunit vaccine boost would be highly affordable.

Chapter 6: Conclusions

The aim of the project was to produce potential vaccine candidates against BFDV. Three strategies were employed: Expression truncated CP (ΔN40 CP) in *E. coli*, expression of CP and ΔN40 CP in plants as ELP fusions, and a DNA vaccine expressing ΔN40 CP. The truncated CP (ΔN40 CP), formed by removal of the first 40 N-terminal amino acids, was used in all of these strategies because of its expected high yield relative to the full-length CP, based on previous work done in our laboratory.

ΔN40 CP was produced in bacteria with or without an ELP tag, but these proteins were mostly insoluble. ΔN40 CP and ΔN40 CP-ELP51 inclusion body proteins were partially purified by successive washes and injected into rabbits for polyclonal antibody production. The ΔN40 CP-ELP51 antigen sample in particular contained relatively few impurities, as assessed by SDS-PAGE and Coomassie staining (Figure 12). Additional ΔN40 CP-ELP51 prepared in this way could potentially be used in animal trials to evaluate whether it could produce an immune response against BFDV. The rabbit sera could recognise BFDV CP proteins produced in bacteria and plants, indicating a targeted immune response against ΔN40 CP.

The expression of BFDV CP was optimised in tobacco by using different expression vectors, pTRAkc-ERH and pEAQ-HT. The pEAQ-HT expression vector was mainly used to transiently express BFDV CP or ΔN40 CP, and their respective ELP fusions, in tobacco plants because of the convenience of single-culture infiltration and the large increase in yield using this vector system. Next, fusion to ELP of 28 or 51 repeats was investigated to increase the yield of BFDV CP and ΔN40 CP. Fusion to either ELP increased the yield of BFDV CP and ΔN40 CP proteins by a minimum of threefold, and non-fused BFDV CP/ ΔN40 CP had a characteristically lower yield. Targeting accumulation of CP-ELP fusions to the ER also increased the stability of the fusion proteins, especially the truncated BFDV CP fusions. Expression was rapid and high levels of protein were observed 2 days post-infiltration. The yield of ΔN40 CP-ELP51 targeted the ER, harvested at 4 dpi, was calculated at 0.74 ± 0.19 % TSP.

Purification of GFP-ELP51 expressed in plants showed that the ELP fusion strategy could be used to successfully purify ELP51 fusion proteins. ΔN40 CP-ELP51 produced in bacteria was first used to characterise the conditions necessary for ΔN40 CP-ELP51 purification from plants. However, ITC of ΔN40 CP-ELP51 produced in bacteria resulted in irreversible aggregation. Nevertheless, the resulting insoluble ΔN40 CP-ELP51 was purified to a high degree with one cycle and was concentrated despite low levels in the soluble cell lysate. The resulting insoluble and purified ΔN40 CP-ELP51 antigen could be used as a vaccine candidate, following endotoxin removal. Future work involving the vaccination of animals with this antigen in its insoluble form could evaluate its immunogenicity.

Inverse transition cycling was used in attempts to purify the ΔN40 CP-ELP51 produced in plants. This protein had the highest yield of all the BFDV CP fusions and therefore was the best candidate for ITC, as purification efficiency increases with increasing ELP concentration. As with the bacterial-expressed protein, irreversible aggregation was observed. An alternate method of ITC using membrane filtration was used to mitigate the irreversible aggregation. This method allowed concentration and partial purification of ΔN40 CP-ELP51, but with significant protein losses in the filtrate, probably due to incomplete aggregation. Free ELP51 was produced in bacteria and plants and both of these proteins could be purified by ITC. These were added to ΔN40 CP-ELP51 crude extract to act as a co-aggregate during purification. The results showed that the free ELP51 could be recovered, but the ΔN40 CP-ELP51 still aggregated irreversibly. However, the free ELP51 was able to increase the aggregation efficiency of ΔN40 CP-ELP51, which was especially useful for mITC purification.

Future work could involve the optimisation of ELP length and guest residue composition (any amino acid ‘X’ in the VPGXG repeating sequence of the ELP) in order to prevent the irreversible of aggregation of ΔN40 CP-ELP during purification. Designing an ELP fusion with a lower transition temperature reduces the temperature required for purification. This could be useful if ΔN40 CP-ELP is thermally unstable. It has been reported that longer ELP tags increase the recovery of ELP fusion after purification, although this may compromise the yield of the fusion protein (Meyer & Chilkoti, 1999; Conley et al., 2009a).

A candidate DNA vaccine, pTH-ΔN40 CP, was evaluated for its capacity to induce BFDV ΔN40 CP expression in transfected cells. A human CMV promoter is used to drive expression of the protein. HEK293T cells were transfected with pTH-ΔN40 CP and protein expression was detected by anti-His immunoblot after 72 hours. Chicken embryonic fibroblast cells were prepared to serve as model avian cells, which are more closely related to psittacine cells than HEK293T. These cells were transfected and protein expression was similarly detected after 2-4 days. The expression of BFDV CP *in vitro* by this DNA vaccine approves it for future vaccination studies. This DNA vaccine may be used in a 'prime-boost' regimen together with a BFDV CP subunit or VLP vaccine. If the subunit vaccine could be produced in plants, such as ΔN40 CP-ELP51 expressed in tobacco, the combination of this DNA vaccine and subunit vaccine boost would be highly affordable.

Production of BFDV CP in bacteria can give higher yields of the protein, but rare codons in the coding DNA sequence requires truncation of the gene or the use of special strains that co-express eukaryotic-type tRNAs. The BFDV Δ40 CP expressed in bacteria were mostly insoluble, the purification of which would require time-consuming denaturation and refolding steps. Recombinant expression of BFDV CP has yielded lower levels of protein in eukaryotic systems such as insect cells (Heath et al., 2006; Stewart et al., 2007). However, these are most likely to produce proteins with the correct post-translational modifications. BFDV CP was transiently expressed in plants, although the yields were low. The yield was significantly increased by fusion to elastin-like polypeptide. The production of BFDV CP fused to ELP in plants is a viable method of antigen production, due to the large increase in yield. The use of ITC is an economical and scalable first step in the purification of CP-ELP. Based on the results, mITC with added free ELP51 as a co-aggregate is the purification method giving the highest recovery of ΔN40 CP-ELP51. Future work on purification of CP-ELP fusions should focus on up-scaling this membrane filtration protocol. For animal studies, the protein may need to be purified further by immobilized metal affinity chromatography (IMAC), by making use of the His tag on the ELP. Removal of the ELP from BFDV CP may not be necessary, as there is evidence of ELP having no immune-modulating effects (Floss et al., 2010a). The production of BFDV CP in plants as an ELP fusion therefore has great promise for the development of an affordable vaccine against BFDV.

References

Alexandrova, R., Eweida, M., Georges, F., Dragulev, B., Abouhaidar, M.G. & Ivanov, I. (1995). Domains in human interferon alpha-1 gene containing tandems of arginine codons AGG play the role of translational initiators in E. coli. *The International Journal of Biochemistry & Cell Biology*. 27 (5), 469–473.

Altunbas, A. & Pochan, D.J. (2012). Peptide-based and polypeptide-based hydrogels for drug delivery and tissue engineering. In: T. Deming (ed.). *Topics in Current Chemistry*. Springer, 135–167.

Aravindaram, K., Kuo, T.-Y., Lan, C., Yu, H., Wang, P., Chen, Y.-S., Chen, G.H. & Yang, N.-S. (2009). Protective immunity against porcine circovirus 2 in mice induced by a gene-based combination vaccination. *The Journal of Gene Medicine*. 11 (4), 288–301.

Arnau, J., Lauritzen, C., Petersen, G.E. & Pedersen, J. (2006). Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expression and Purification*. 48 (1), 1–13.

Balasubramaniam, D., Wilkinson, C., Van Cott, K. & Zhang, C. (2003). Tobacco protein separation by aqueous two-phase extraction. *Journal of Chromatography A*. 989 (1), 119–129.

Benchabane, M., Goulet, C., Rivard, D., Faye, L., Gomord, V. & Michaud, D. (2008). Preventing unintended proteolysis in plant protein biofactories. *Plant Biotechnology Journal*. 6 (7), 633–648.

Bhatla, S.C., Kaushik, V. & Yadav, M.K. (2010). Use of oil bodies and oleosins in recombinant protein production and other biotechnological applications. *Biotechnology Advances*. 28 (3), 293–300.

Bonne, N., Shearer, P., Sharp, M., Clark, P. & Raidal, S. (2009). Assessment of recombinant beak and feather disease virus capsid protein as a vaccine for psittacine beak and feather disease. *Journal of General Virology*. 90 (3), 640–647.

Borisjuk, N. V, Borisjuk, L.G., Logendra, S., Petersen, F., Gleba, Y. & Raskin, I. (1999). Production of recombinant proteins in plant root exudates. *Nature Biotechnology*. 17 (5), 466–469.

Chaudhary, S., Parmenter, D.L. & Moloney, M.M. (1998). Transgenic Brassica carinata as a vehicle for the production of recombinant proteins in seeds. *Plant Cell Reports*. 17 (3), 195–200.

Chen, D., Duggan, C., Texada, D.E., Reden, T.B., Kooragayala, L.M. & Langford, M.P. (2005). Immunogenicity of enterovirus 70 capsid protein VP1 and its non-overlapping N- and C-terminal fragments. *Antiviral Research*. 66 (2-3), 111–117.

Christensen, T., Trabbic-Carlson, K., Liu, W. & Chilkoti, A. (2007). Purification of recombinant proteins from Escherichia coli at low expression levels by inverse transition cycling. *Analytical Biochemistry*. 360 (1), 166–168.

Coleman, C.E., Herman, E.M., Takasaki, K. & Larkins, B.A. (1996). The maize gamma-zein sequesters alpha-zein and stabilizes its accumulation in protein bodies of transgenic tobacco endosperm. *The Plant Cell*. 8 (12), 2335–2345.

Conley, A.J., Joensuu, J.J., Jevnikar, A.M., Menassa, R. & Brandle, J.E. (2009a). Optimization of elastin-like polypeptide fusions for expression and purification of recombinant proteins in plants. *Biotechnology and Bioengineering*. 103 (3), 562–573.

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

Conley, A.J., Joensuu, J.J., Richman, A. & Menassa, R. (2011). Protein body-inducing fusions for high-level production and purification of recombinant proteins in plants. *Plant Biotechnology Journal*. 9 (4), 1–11.

Debelle, L. & Tamburro, A.M. (1999). Elastin: molecular description and function. *The International Journal of Biochemistry & Cell Biology*. 31 (2), 261–272.

Doria-Rose, N. (2003). DNA vaccine strategies: candidates for immune modulation and immunization regimens. *Methods*. 31 (3), 207–216.

Van Drunen Littel-van den Hurk, S. & Hannaman, D. (2010). Electroporation for DNA immunization: clinical application. *Expert Review of Vaccines*. 9 (5), 503–517.

Dufourmantel, N., Pelissier, B., Garçon, F., Peltier, G., Ferullo, J.-M. & Tissot, G. (2004). Generation of fertile transplastomic soybean. *Plant Molecular Biology*. 55 (4), 479–489.

El-Attar, L., Oliver, S.L., Mackie, A., Charpilienne, A., Poncet, D., Cohen, J. & Bridger, J.C. (2009). Comparison of the efficacy of rotavirus VLP vaccines to a live homologous rotavirus vaccine in a pig model of rotavirus disease. *Vaccine*. 27 (24), 3201–3208.

Evangelista, R., Kusnadi, A., Howard, J. & Nikolov, Z. (1998). Process and economic evaluation of the extraction and purification of recombinant beta-glucuronidase from transgenic corn. *Biotechnology Progress*. 14 (4), 607–614.

Fan, H., Xiao, S., Tong, T., Wang, S., Xie, L., Jiang, Y., Chen, H. & Fang, L. (2008). Immunogenicity of porcine circovirus type 2 capsid protein targeting to different subcellular compartments. *Molecular Immunology*. 45 (3), 653–660.

Farinas, C.S., Leite, A. & Miranda, E.A. (2005). Aqueous extraction of maize endosperm: Insights for recombinant protein hosts based on downstream processing. *Process Biochemistry*. 40 (10), 3327–3336.

Finsterbusch, T. & Mankertz, A. (2009). Porcine circoviruses--small but powerful. *Virus Research*. 143 (2), 177–183.

Fischer, R. & Emans, N. (2000). Molecular farming of pharmaceutical proteins. *Transgenic Research*. 9 (4-5), 279–299.

Floss, D.M., Mockey, M., Zanello, G., Brosson, D., Diogon, M., Frutos, R., Bruel, T., Rodrigues, V., Garzon, E., Chevaleyre, C., Berri, M., Salmon, H., Conrad, U. & Dedieu, L. (2010a). Expression and immunogenicity of the mycobacterial Ag85B/ESAT-6 antigens produced in transgenic plants by elastin-like peptide fusion strategy. *Journal of Biomedicine & Biotechnology*, doi: 10.1155/2010/274346.

Floss, D.M., Sack, M., Stadlmann, J., Rademacher, T., Scheller, J., Stöger, E., Fischer, R. & Conrad, U. (2008a). Biochemical and functional characterization of anti-HIV antibody-ELP fusion proteins from transgenic plants. *Plant biotechnology journal*. 6 (4), 379–391.

Floss, D.M., Sack, M., Stadlmann, J., Rademacher, T., Scheller, J., Stöger, E., Fischer, R. & Conrad, U. (2008b). Biochemical and functional characterization of anti-HIV antibody-ELP fusion proteins from transgenic plants. *Plant Biotechnology Journal*. 6 (4), 379–391.

Floss, D.M., Schallau, K., Rose-John, S., Conrad, U. & Scheller, J. (2010b). Elastin-like polypeptides revolutionize recombinant protein expression and their biomedical application. *Trends in Biotechnology*. 28 (1), 37–45.

Fong, B.A., Wu, W. & Wood, D.W. (2009). Optimization of ELP-intein mediated protein purification by salt substitution. *Protein Expression and Purification*. 66 (2), 198–202.

Frankel, S., Sohn, R. & Leinwand, L. (1991). The use of sarkosyl in generating soluble protein after bacterial expression. *Proceedings of the National Academy of Sciences of the United States of America*. 88 (4), 1192–1196.

Fu, F., Li, X., Lang, Y., Yang, Y., Tong, G., Li, G., Zhou, Y. & Li, X. (2011). Co-expression of ubiquitin gene and capsid protein gene enhances the potency of DNA immunization of PCV2 in mice. *Virology Journal*. 8 (264).

García-Fruitós, E., Vázquez, E., Díez-Gil, C., Corchero, J.L., Seras-Franzoso, J., Ratera, I., Veciana, J. & Villaverde, A. (2012). Bacterial inclusion bodies: making gold from waste. *Trends in Biotechnology*. 30 (2), 1–6.

Ge, X. & Filipe, C.D.M. (2006). Simultaneous phase transition of ELP tagged molecules and free ELP: an efficient and reversible capture system. *Biomacromolecules*. 7 (9), 2475–2478.

Ge, X., Trabbic-Carlson, K., Chilkoti, A. & Filipe, C.D.M. (2006). Purification of an elastin-like fusion protein by microfiltration. *Biotechnology and Bioengineering*. 95 (3), 424–432.

Gelvin, S.B. (2003). Agrobacterium-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiology and Molecular Biology Reviews*. 67 (1), 16–37.

Grebenok, R.J., Pierson, E., Lambert, G.M., Gong, F.C., Afonso, C.L., Haldeman-Cahill, R., Carrington, J.C. & Galbraith, D.W. (1997). Green-fluorescent protein fusions for efficient characterization of nuclear targeting. *The Plant Journal*. 11 (3), 573–586.

Hanke, T., Schneider, J., Gilbert, S.C., Hill, A. V & McMichael, A. (1998). DNA multi-CTL epitope vaccines for HIV and Plasmodium falciparum: immunogenicity in mice. *Vaccine*. 16 (4), 426–435.

Harper, B.K., Mabon, S.A., Leffel, S.M., Halfhill, M.D., Richards, H.A., Moyer, K.A. & Stewart, C.N. (1999). Green fluorescent protein as a marker for expression of a second gene in transgenic plants. *Nature Biotechnology*. 17 (11), 1125–1129.

Hassouneh, W., Fischer, K., MacEwan, S.R., Branscheid, R., Fu, C.L., Liu, R., Schmidt, M. & Chilkoti, A. (2012). Unexpected multivalent display of proteins by temperature triggered self-assembly of elastin-like polypeptide block copolymers. *Biomacromolecules*. 13 (5), 1598–1605.

Heath, L., Williamson, A.-L. & Rybicki, E.P. (2006). The capsid protein of beak and feather disease virus binds to the viral DNA and is responsible for transporting the replication-associated protein into the nucleus. *Journal of Virology*. 80 (14), 7219–7225.

Hitzeroth, I.I., Passmore, J.S., Shephard, E., Stewart, D., Müller, M., Williamson, A., Rybicki, E.P. & Kast, W.M. (2009). Immunogenicity of an HPV-16 L2 DNA vaccine. *Vaccine*. 27 (46), 6432–6434.

ICTV (2005). Circoviridae. In: C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger, & L. A. Ball (eds.). *Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses*. Elsevier Inc., 327–334.

ICTV (2011). *Virus Taxonomy: 2011 Release (current)*. [Online]. 2011. International Committee on Taxonomy of Viruses. Available from: http://www.ictvonline.org/virusTaxonomy.asp. [Accessed: 30 August 2012].

Jacobson, E.R., Clubb, S., Simpson, C., Walsh, M., Lothrop, C.D., Gaskin, J., Bauer, J., Hines, S., Kollias, G. V & Poulos, P. (1986). Feather and beak dystrophy and necrosis in cockatoos: clinicopathologic evaluations. *Journal of the American Veterinary Medical Association*. 189 (9), 999–1005.

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

Johne, R., Raue, R., Grund, C., Kaleta, E.F. & Müller, H. (2004). Recombinant expression of a truncated capsid protein of beak and feather disease virus and its application in serological tests. *Avian Pathology*. 33 (3), 328–336.

Karg, S.R. & Kallio, P.T. (2009). The production of biopharmaceuticals in plant systems. *Biotechnology Advances*. 27 (6), 894–897.

Khalesi, B., Bonne, N., Stewart, M., Sharp, M. & Raidal, S. (2005). A comparison of haemagglutination, haemagglutination inhibition and PCR for the detection of psittacine beak and feather disease virus infection and a comparison of isolates obtained from loriids. *The Journal of General Virology*. 86 (11), 3039–3046.

Kim, J.-Y., Mulchandani, A. & Chen, W. (2005). Temperature-triggered purification of antibodies. *Biotechnology and Bioengineering*. 90 (3), 373–379.

De Kloet, E. & De Kloet, S.R. (2004). Analysis of the beak and feather disease viral genome indicates the existence of several genotypes which have a complex psittacine host specificity. *Archives of Virology*. 149 (12), 2393–2412.

Kogan, M.J., Dalcol, I., Gorostiza, P., Lopez-Iglesias, C., Pons, R., Pons, M., Sanz, F. & Giralt, E. (2002). Supramolecular properties of the proline-rich gamma-Zein N-terminal domain. *Biophysical Journal*. 83 (2), 1194–1204.

Kusnadi, A.R., Nikolov, Z.L. & Howard, J.A. (1997). Production of recombinant proteins in transgenic plants: Practical considerations. *Biotechnology and Bioengineering*. 56 (5), 473–484.

Lakatos, L., Szittya, G., Silhavy, D. & Burgyán, J. (2004). Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombusviruses. *The EMBO Journal*. 23 (4), 876–884.

Lan, D., Huang, G., Shao, H., Zhang, L., Ma, L., Chen, S. & Xu, A. (2011). An improved nonchromatographic method for the purification of recombinant proteins using elastin-like polypeptide-tagged proteases. *Analytical Biochemistry*. 415 (2), 200–202.

Latimer, K.S., Rakich, P.M., Kircher, I.M., Ritchie, B.W., Niagro, F.D., Steffens, W.L. & Lukert, P.D. (1990). Extracutaneous Viral Inclusions in Psittacine Beak and Feather Disease. *Journal of Veterinary Diagnostic Investigation*. 2 (3), 204–207.

Latimer, K.S., Rakich, P.M., Steffens, W.L., Kircher, I.M., Ritchie, B.W., Niagro, F.D. & Lukert, P.D. (1991). A Novel DNA Virus Associated with Feather Inclusions in Psittacine Beak and Feather Disease. *Veterinary Pathology*. 28 (4), 300–304.

Li, B., Alonso, D.O. & Daggett, V. (2001). The molecular basis for the inverse temperature transition of elastin. *Journal of Molecular Biology*. 305 (3), 581–592.

Lin, S.-C., Lin, Y.-F., Chong, P. & Wu, S.-C. (2012). Broader Neutralizing Antibodies against H5N1 Viruses Using Prime-Boost Immunization of Hyperglycosylated Hemagglutinin DNA and Virus-Like Particles. *PloS One*. 7 (6), e39075.

Liu, F., Ge, S., Li, L., Wu, X., Liu, Z. & Wang, Z. (2012a). Virus-like particles: potential veterinary vaccine immunogens. *Research in Veterinary Science*. 93 (2), 553–559.

Liu, F., Tsai, S.-L., Madan, B. & Chen, W. (2012b). Engineering a high-affinity scaffold for non-chromatographic protein purification via intein-mediated cleavage. *Biotechnology and bioengineering*, doi: 10.1002/bit.24545.

Lowe, B.A., Shiva Prakash, N., Way, M., Mann, M.T., Spencer, T.M. & Boddupalli, R.S. (2009). Enhanced single copy integration events in corn via particle bombardment using low quantities of DNA. *Transgenic Research*. 18 (6), 831–840.

Lunn, J.E. (2007). Compartmentation in plant metabolism. *Journal of Experimental Botany*. 58 (1), 35–47.

Lössl, A.G. & Waheed, M.T. (2011). Chloroplast-derived vaccines against human diseases: achievements, challenges and scopes. *Plant Biotechnology Journal*. 9 (5), 527–539.

Ma, J.K.-C., Drake, P.M.W. & Christou, P. (2003). The production of recombinant pharmaceutical proteins in plants. *Nature Reviews Genetics*. 4 (10), 794–805.

MacEwan, S.R. & Chilkoti, A. (2010). Elastin-like polypeptides: biomedical applications of tunable biopolymers. *Biopolymers*. 94 (1), 60–77.

Maclean, J., Koekemoer, M., Olivier, a J., Stewart, D., Hitzeroth, I.I., Rademacher, T., Fischer, R., Williamson, A.-L. & Rybicki, E.P. (2007). Optimization of human papillomavirus type 16 (HPV-16) L1 expression in plants: comparison of the suitability of different HPV-16 L1 gene variants and different cell-compartment localization. *The Journal of General Virology*. 88 (5), 1460–1469.

Mahé, D., Blanchard, P., Truong, C., Arnauld, C., Le Cann, P., Cariolet, R., Madec, F., Albina, E. & Jestin, A. (2000). Differential recognition of ORF2 protein from type 1 and type 2 porcine circoviruses and identification of immunorelevant epitopes. *The Journal of General Virology*. 81 (Pt 7), 1815–1824.

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

Marcekova, Z., Psikal, I., Kosinova, E., Benada, O., Sebo, P. & Bumba, L. (2009). Heterologous expression of full-length capsid protein of porcine circovirus 2 in Escherichia coli and its potential use for detection of antibodies. *Journal of Virological Methods*. 162 (1-2), 133–141.

De Marco, A., Vigh, L., Diamant, S. & Goloubinoff, P. (2005). Native folding of aggregation-prone recombinant proteins in Escherichia coli by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones. *Cell stress & Chaperones*. 10 (4), 329–339.

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

Matić, S., Masenga, V., Poli, A., Rinaldi, R., Milne, R.G., Vecchiati, M. & Noris, E. (2012). Comparative analysis of recombinant Human Papillomavirus 8 L1 production in plants by a variety of expression systems and purification methods. *Plant Biotechnology Journal*. 10 (4), 410–412.

McDaniel, J.R., Macewan, S.R., Dewhirst, M. & Chilkoti, A. (2012). Doxorubicin-conjugated chimeric polypeptide nanoparticles that respond to mild hyperthermia. *Journal of Controlled Release*. 159 (3), 362–367.

McPherson, D.T., Xu, J. & Urry, D.W. (1996). Product purification by reversible phase transition following Escherichia coli expression of genes encoding up to 251 repeats of the elastomeric pentapeptide GVGVP. *Protein Expression and Purification*. 7 (1), 51–57.

Menkhaus, T.J., Bai, Y., Zhang, C., Nikolov, Z.L. & Glatz, C.E. (2004). Considerations for the recovery of recombinant proteins from plants. *Biotechnology Progress*. 20 (4), 1001–1014.

Meyer, D.E. & Chilkoti, A. (2002a). Genetically encoded synthesis of protein-based polymers with precisely specified molecular weight and sequence by recursive directional ligation: examples from the elastin-like polypeptide system. *Biomacromolecules*. 3 (2), 357–367.

Meyer, D.E. & Chilkoti, A. (2002b). Protein Purification by Inverse Transition Cycling. In: E. A. Golemis (ed.). *Protein-Protein Interactions: A Molecular Cloning Manual*. New York: Cold Spring Harbor Laboratory Press, 329–344.

Meyer, D.E. & Chilkoti, A. (1999). Purification of recombinant proteins by fusion with thermally-responsive polypeptides. *Nature Biotechnology*. 17 (11), 1112–1115.

Meyer, D.E., Trabbic-Carlson, K. & Chilkoti, A. (2001). Protein purification by fusion with an environmentally responsive elastin-like polypeptide: effect of polypeptide length on the purification of thioredoxin. *Biotechnology Progress*. 17 (4), 720–728.

Mie, M., Mizushima, Y. & Kobatake, E. (2008). Novel extracellular matrix for cell sheet recovery using genetically engineered elastin-like protein. *Journal of Biomedical Materials Research Part B: Applied Biomaterials*. 86 (1), 283–290.

Misinzo, G., Delputte, P.L. & Nauwynck, H.J. (2008). Inhibition of endosome-lysosome system acidification enhances porcine circovirus 2 infection of porcine epithelial cells. *Journal of Virology*. 82 (3), 1128–1135.

Mushegian, A.R. & Shepherd, R.J. (1995). Genetic elements of plant viruses as tools for genetic engineering. *Microbiological Reviews*. 59 (4), 548–578.

Mustalahti, E., Saloheimo, M. & Joensuu, J.J. (2011). Intracellular protein production in Trichoderma reesei (Hypocrea jecorina) with hydrophobin fusion technology. *New Biotechnology*, doi: 10.1016/j.nbt.2011.09.006.

Nauwynck, H.J., Sanchez, R., Meerts, P., Lefebvre, D.J., Saha, D., Huang, L. & Misinzo, G. (2012). Cell tropism and entry of porcine circovirus 2. *Virus Research*. 164 (1-2), 43–45.

Nilsson, J., Ståhl, S., Lundeberg, J., Uhlén, M. & Nygren, P.A. (1997). Affinity fusion strategies for detection, purification, and immobilization of recombinant proteins. *Protein Expression and Purification*. 11 (1), 1–16.

Nykiforuk, C.L., Boothe, J.G., Murray, E.W., Keon, R.G., Goren, H.J., Markley, N.A. & Moloney, M.M. (2006). Transgenic expression and recovery of biologically active recombinant human insulin from Arabidopsis thaliana seeds. *Plant Biotechnology Journal*. 4 (1), 77–85.

Pass, D.A. & Perry, R.A. (1984). The pathology of psittacine beak and feather disease. *Australian veterinary journal*. 61 (3), 69–74.

Patel, J., Zhu, H., Menassa, R., Gyenis, L., Richman, A. & Brandle, J. (2007). Elastin-like polypeptide fusions enhance the accumulation of recombinant proteins in tobacco leaves. *Transgenic Research*. 16 (2), 239–249.

Paul, M. & Ma, J.K.-C. (2011). Plant-made pharmaceuticals: leading products and production platforms. *Biotechnology and Applied Biochemistry*. 58 (1), 58–67.

Phan, H.T. & Conrad, U. (2011). Membrane-based inverse transition cycling: an improved means for purifying plant-derived recombinant protein-elastin-like polypeptide fusions. *International Journal of Molecular Sciences*. 12 (5), 2808–2821.

Platis, D., Drossard, J., Fischer, R., Ma, J.K.-C. & Labrou, N.E. (2008). New downstream processing strategy for the purification of monoclonal antibodies from transgenic tobacco plants. *Journal of Chromatography A*. 1211 (1-2), 80–89.

Platis, D. & Labrou, N.E. (2006). Development of an aqueous two-phase partitioning system for fractionating therapeutic proteins from tobacco extract. *Journal of Chromatography A*. 1128 (1-2), 114–124.

Raidal, S.R. & Cross, G.M. (1994). The haemagglutination spectrum of psittacine beak and feather disease virus. *Avian Pathology*. 23 (4), 621–630.

Reddy, M.S.S., Dinkins, R.D. & Collins, G.B. (2003). Gene silencing in transgenic soybean plants transformed via particle bombardment. *Plant Cell Reports*. 21 (7), 676–683.

Regnard, G.L., Halley-Stott, R.P., Tanzer, F.L., Hitzeroth, I.I. & Rybicki, E.P. (2010). High level protein expression in plants through the use of a novel autonomously replicating geminivirus shuttle vector. *Plant Biotechnology Journal*. 8 (1), 38–46.

Rincón, A.C., Molina-Martinez, I.T., De Las Heras, B., Alonso, M., Baílez, C., Rodríguez-Cabello, J.C. & Herrero-Vanrell, R. (2006). Biocompatibility of elastin-like polymer poly(VPAVG) microparticles: in vitro and in vivo studies. *Journal of Biomedical Materials Research Part A*. 78 (2), 343–351.

Rybicki, E.P. (2010). Plant-made vaccines for humans and animals. *Plant Biotechnology Journal*. 8 (5), 620–637.

Rybicki, E.P., Chikwamba, R., Koch, M., Rhodes, J.I. & Groenewald, J.-H. (2012). Plant-made therapeutics: an emerging platform in South Africa. *Biotechnology Advances*. 30 (2), 449–459.

Saade, F. & Petrovsky, N. (2012). Technologies for enhanced efficacy of DNA vaccines. *Expert Review of Vaccines*. 11 (2), 189–209.

Sainsbury, F. & Lomonossoff, G.P. (2008). Extremely high-level and rapid transient protein production in plants without the use of viral replication. *Plant Physiology*. 148 (3), 1212–1218.

Sainsbury, F., Sack, M., Stadlmann, J., Quendler, H., Fischer, R. & Lomonossoff, G.P. (2010). Rapid transient production in plants by replicating and non-replicating vectors yields high quality functional anti-HIV antibody. *PloS ONE*. 5 (11), e13976.

Sainsbury, F., Thuenemann, E.C. & Lomonossoff, G.P. (2009). pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. *Plant Biotechnology Journal*. 7 (7), 682–93.

Saxena, P., Hsieh, Y.-C., Alvarado, V.Y., Sainsbury, F., Saunders, K., Lomonossoff, G.P. & Scholthof, H.B. (2011). Improved foreign gene expression in plants using a virus-encoded suppressor of RNA silencing modified to be developmentally harmless. *Plant Biotechnology Journal*. 9 (6), 703–712.

Scheller, J., Leps, M. & Conrad, U. (2006). Forcing single-chain variable fragment production in tobacco seeds by fusion to elastin-like polypeptides. *Plant Biotechnology Journal*. 4 (2), 243–249.

Schoemaker, N.J., Dorrestein, G.M., Latimer, K.S., Lumeij, J.T., Kik, M.J., Van der Hage, M.H. & Campagnoli, R.P. (2000). Severe leukopenia and liver necrosis in young African grey parrots (Psittacus erithacus erithacus) infected with psittacine circovirus. *Avian Diseases*. 44 (2), 470–478.

Schokker, E.., Wagenberg, A.C.. van & Boekel, M.A.J.. van (1998). A note on the use of urea in studying the mechanism of thermal inactivation of extracellular proteinase from pseudomonas fluorescens 22F. *Enzyme and Microbial Technology*. 42 (8), 628–698.

Scholthof, H.B. (2007). Heterologous expression of viral RNA interference suppressors: RISC management. *Plant Physiology*. 145 (4), 1110–1117.

Sharma, A.K. & Sharma, M.K. (2009). Plants as bioreactors: Recent developments and emerging opportunities. *Biotechnology Advances*. 27 (6), 811–832.

Shearer, P.L., Bonne, N., Clark, P., Sharp, M. & Raidal, S.R. (2008). Beak and feather disease virus infection in cockatiels (Nymphicus hollandicus). *Avian Pathology*. 37 (1), 75–81.

Shearer, P.L., Sharp, M., Bonne, N., Clark, P. & Raidal, S.R. (2009a). A blocking ELISA for the detection of antibodies to psittacine beak and feather disease virus (BFDV). *Journal of Virological Methods*. 158 (1-2), 136–140.

Shearer, P.L., Sharp, M., Bonne, N., Clark, P. & Raidal, S.R. (2009b). A quantitative, real-time polymerase chain reaction assay for beak and feather disease virus. *Journal of Virological Methods*. 159 (1), 98–104.

Shen, H.-G., Zhou, J.-Y., Huang, Z.-Y., Guo, J.-Q., Xing, G., He, J.-L., Yan, Y. & Gong, L.-Y. (2008). Protective immunity against porcine circovirus 2 by vaccination with ORF2-based DNA and subunit vaccines in mice. *The Journal of General Virology*. 89 (8), 1857–1865.

Shen, W.J. & Forde, B.G. (1989). Efficient transformation of Agrobacterium spp. by high voltage electroporation. *Nucleic Acids Research*. 17 (20), 8385.

Shih, S.M.-H. & Doran, P.M. (2009). Foreign protein production using plant cell and organ cultures: Advantages and limitations. *Biotechnology Advances*. 27 (6), 1036–1042.

Siddiqui, S.A., Sarmiento, C., Truve, E., Lehto, H. & Lehto, K. (2008). Phenotypes and functional effects caused by various viral RNA silencing suppressors in transgenic Nicotiana benthamiana and N. tabacum. *Molecular Plant-Microbe Interactions*. 21 (2), 178–187.

Solomon, J.S., Nixon, C.P., McGarvey, S.T., Acosta, L.P., Manalo, D. & Kurtis, J.D. (2004). Expression, purification, and human antibody response to a 67 kDa vaccine candidate for schistosomiasis japonica. *Protein Expression and Purification*. 36 (2), 226–231.

Stewart, M.E., Bonne, N., Shearer, P., Khalesi, B., Sharp, M. & Raidal, S. (2007). Baculovirus expression of beak and feather disease virus (BFDV) capsid protein capable of self-assembly and haemagglutination. *Journal of Virological Methods*. 141 (2), 181–187.

Stoger, E., Ma, J.K.-C., Fischer, R. & Christou, P. (2005). Sowing the seeds of success: pharmaceutical proteins from plants. *Current Opinion in Biotechnology*. 16 (2), 167–173.

Sørensen, H.P. & Mortensen, K.K. (2005). Advanced genetic strategies for recombinant protein expression in Escherichia coli. *Journal of Biotechnology*. 115 (2), 113–128.

Tegel, H., Ottosson, J. & Hober, S. (2011). Enhancing the protein production levels in Escherichia coli with a strong promoter. *The FEBS Journal*. 278 (5), 729–739.

Tegel, H., Tourle, S., Ottosson, J. & Persson, A. (2010). Increased levels of recombinant human proteins with the Escherichia coli strain Rosetta(DE3). *Protein Expression and Purification*. 69 (2), 159–167.

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology*. 72 (2), 211–222.

Thanavala, Y., Huang, Z. & Mason, H.S. (2006). Plant-derived vaccines: a look back at the highlights and a view to the challenges on the road ahead. *Expert Review of Vaccines*. 5 (2), 249–260.

Tian, L. & Sun, S.S.M. (2011). A cost-effective ELP-intein coupling system for recombinant protein purification from plant production platform. *PloS ONE*. 6 (8), e24183.

Todd, D. (2000). Circoviruses: immunosuppressive threats to avian species: a review. *Avian Pathology*. 29 (5), 373–394.

Todd, D., Niagro, F.D., Ritchie, B.W., Curran, W., Allan, G.M., Lukert, P.D., Latimer, K.S., Steffens, W.L. & McNulty, M.S. (1991). Comparison of three animal viruses with circular single-stranded DNA genomes. *Archives of Virology*. 117 (1-2), 129–135.

Torrent, M., Llompart, B., Lasserre-Ramassamy, S., Llop-Tous, I., Bastida, M., Marzabal, P., Westerholm-Parvinen, A., Saloheimo, M., Heifetz, P.B. & Ludevid, M.D. (2009). Eukaryotic protein production in designed storage organelles. *BMC Biology*. 7 (5), doi: 10.1186/1741–7007–7–5.

Toussaint, N.C., Maman, Y., Kohlbacher, O. & Louzoun, Y. (2011). Universal peptide vaccines - Optimal peptide vaccine design based on viral sequence conservation. *Vaccine*. 29 (47), 8745–8753.

Trabbic-Carlson, K., Liu, L.L., Kim, B. & Chilkoti, A. (2004a). Expression and purification of recombinant proteins from Escherichia coli: Comparison of an elastin-like polypeptide fusion with an oligohistidine fusion. *Protein Science*. 13 (12), 3274–3284.

Trabbic-Carlson, K., Meyer, D.E., Liu, L., Piervincenzi, R., Nath, N., LaBean, T. & Chilkoti, A. (2004b). Effect of protein fusion on the transition temperature of an environmentally responsive elastin-like polypeptide: a role for surface hydrophobicity? *Protein Engineering, Design & Selection*. 17 (1), 57–66.

Tremblay, R., Wang, D., Jevnikar, A.M. & Ma, S. (2010). Tobacco, a highly efficient green bioreactor for production of therapeutic proteins. *Biotechnology Advances*. 28 (2), 214–221.

Urry, D.W. (1997). Physical Chemistry of Biological Free Energy Transduction As Demonstrated by Elastic Protein-Based Polymers. *The Journal of Physical Chemistry B*. 101 (51), 11007–11028.

Urry, D.W., Luan, C.H., Parker, T.M., Gowda, D.C., Prasad, K.U., Reid, M.C. & Safavy, A. (1991). Temperature of polypeptide inverse temperature transition depends on mean residue hydrophobicity. *Journal of the American Chemical Society*. 113 (11), 4346–4348.

Vardakou, M., Sainsbury, F., Rigby, N., Mulholland, F. & Lomonossoff, G.P. (2012). Expression of active recombinant human gastric lipase in Nicotiana benthamiana using the CPMV-HT transient expression system. *Protein Expression and Purification*. 81 (1), 69–74.

Varsani, A., De Villiers, G.K., Regnard, G.L., Bragg, R.R., Kondiah, K., Hitzeroth, I.I. & Rybicki, E.P. (2010). A unique isolate of beak and feather disease virus isolated from budgerigars (Melopsittacus undulatus) in South Africa. *Archives of Virology*. 155 (3), 435–439.

Varsani, A., Williamson, A.-L., Stewart, D. & Rybicki, E.P. (2006). Transient expression of Human papillomavirus type 16 L1 protein in Nicotiana benthamiana using an infectious tobamovirus vector. *Virus Research*. 120 (1-2), 91–96.

Ventura, S. & Villaverde, A. (2006). Protein quality in bacterial inclusion bodies. *Trends in Biotechnology*. 24 (4), 179–185.

Verver, J., Wellink, J., Van Lent, J., Gopinath, K. & Van Kammen, A. (1998). Studies on the movement of cowpea mosaic virus using the jellyfish green fluorescent protein. *Virology*. 242 (1), 22–27.

Vidi, P.-A., Kessler, F. & Bréhélin, C. (2007). Plastoglobules: a new address for targeting recombinant proteins in the chloroplast. *BMC Biotechnology*. 7 (4).

Voinnet, O., Rivas, S., Mestre, P. & 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. *The Plant Journal*. 33 (5), 949–956.

Wang, X., Shi, F., Wösten, H. a B., Hektor, H., Poolman, B. & Robillard, G.T. (2005). The SC3 hydrophobin self-assembles into a membrane with distinct mass transfer properties. *Biophysical Journal*. 88 (5), 3434–3443.

Woodhouse, K. a, Klement, P., Chen, V., Gorbet, M.B., Keeley, F.W., Stahl, R., Fromstein, J.D. & Bellingham, C.M. (2004). Investigation of recombinant human elastin polypeptides as non-thrombogenic coatings. *Biomaterials*. 25 (19), 4543–4553.

Wu, W.-Y., Mee, C., Califano, F., Banki, R. & Wood, D.W. (2006). Recombinant protein purification by self-cleaving aggregation tag. *Nature Protocols*. 1 (5), 2257–2262.

Yin, S., Sun, S., Yang, S., Shang, Y., Cai, X. & Liu, X. (2010). Self-assembly of virus-like particles of porcine circovirus type 2 capsid protein expressed from Escherichia coli. *Virology Journal*. 7 (166), doi: 10.1186/1743–422X–7–166.

Ytterberg, A.J., Peltier, J.-B. & Van Wijk, K.J. (2006). Protein profiling of plastoglobules in chloroplasts and chromoplasts. A surprising site for differential accumulation of metabolic enzymes. *Plant Physiology*. 140 (3), 984–997.

Yusibov, V., Streatfield, S.J. & Kushnir, N. (2011). Clinical development of plant-produced recombinant pharmaceuticals: Vaccines, antibodies, and beyond. *Human Vaccines*. 7 (3), 313–321.

Appendix A: General procedures for PCR, cloning and plasmid preparation

PCR

GoTaq® DNA Polymerase (Promega, USA) or Accuzyme™ DNA Polymerase (Bioline, USA) were used as per the manufacturer's recommendations. All PCR was performed with a Bio-Rad MyCycler™ (Bio-Rad Laboratories, USA) thermal cycler. The concentration of Mg^{2+} was 2.0 mM for all reactions. All PCR products were resolved by electrophoresis on 0.8% agarose gels. Colony PCR was often used to screen transformants. A single colony was inoculated into the reaction mixture in each individual 0.2 ml PCR tube, using a sterile toothpick.

Cloning

The relevant bands in agarose gels were excised under long wavelength UV illumination and DNA was extracted using a Qiagen Gel Extraction Kit (Qiagen, USA). Restriction enzymes used were: *Afl*III, *Mlu*I, *Nco*I, *Not*I, *Xho*I (Roche, USA), *Age*I, *Eco*RI, *Eco*RV and *Xho*I (Fermentas, Canada). All restriction enzyme digests were carried out at 37°C for 2 hours. Clean-up of restriction enzyme digests was performed with a Qiagen PCR Purification Kit. Ligation of the insert into the vector (similarly digested at compatible enzyme sites) was performed using Rapid DNA Ligation Kit (Fermentas, Canada) as per the manufacturer's recommendations. All digested vector DNA was dephosphorylated with Rapid Alkaline Phosphatase Kit (Roche, USA) prior to ligation. The ligation products were then transformed into *E. coli* DH5α chemically competent cells (E. cloni® (Lucigen, USA)) as per the manufacturer's instructions and were incubated overnight at 37°C on LB agar supplemented with the relevant antibiotics.

Plasmid preparation

A single colony on a LB agar plate was used to inoculate 10 mL of liquid LB media, supplemented with the relevant antibiotics. The culture was grown overnight at 37 °C

in an orbital shaker. The cell pellet from 4 mL of culture was collected by centrifugation. Plasmid DNA was prepared using a Qiagen Spin Miniprep Kit as per the manufacturer's protocol. Plasmid DNA was sequenced by Macrogen Inc. (Netherlands).

Appendix B: Vectors used in cloning and expression

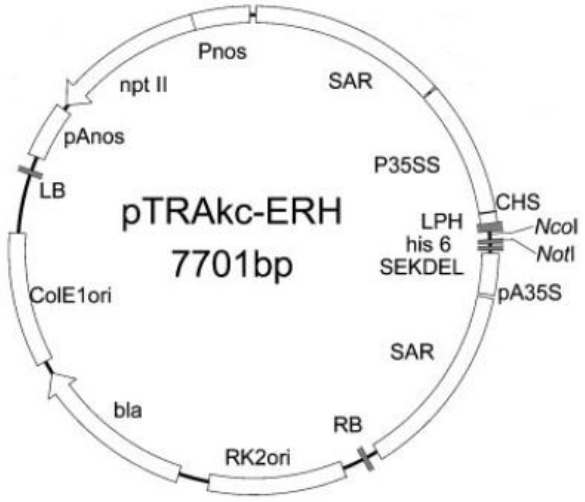


Figure 43: pTRAkc-ERH binary expression vector features:

"CaMV 35S promoter with duplicated transcriptional enhancer; CHS, chalcone synthase 5' untranslated region; pA35S, CaMV 35S polyadenylation signal; SAR, scaffold attachment region of the tobacco Rb7 gene; LB and RB, the left and right borders for T-DNA integration; ColE1ori, origin of replication for E. coli; RK2ori, origin of replication for Agrobacterium; bla, ampicillin/carbenicillin-resistance bla gene; LPH, signal-peptide sequence from the murine mAb24 heavy chain; his6, 6X His tag sequence; SEKDEL, ER-retention signal sequence; npt II, kanamycin-resistance npt II gene; Pnos and pAnos, promoter and polyadenylation signal of the nopaline synthase gene" (adapted from Maclean et al. (2007)).

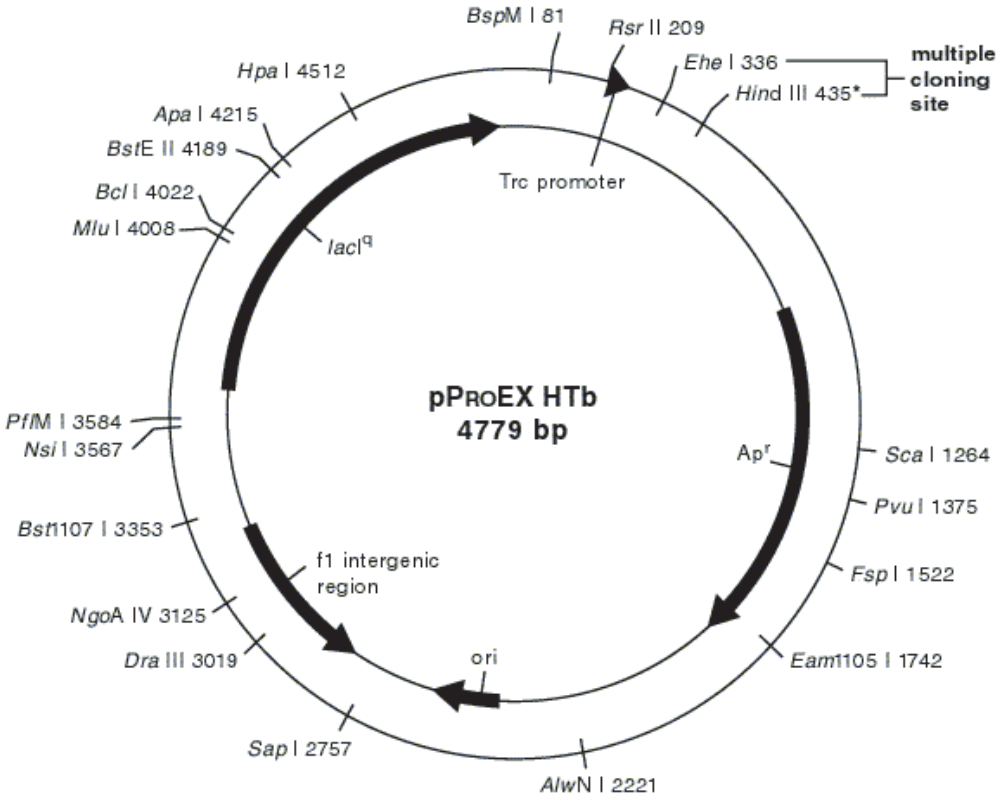


Figure 44: pPROEX-HTb features (Life Technologies, USA)

Schematic of pPROEX-HTb vector. Included here are the trc promoter, lacI gene induced by IPTG, ampicillin resistance gene (Apr), origin of replication (ori) and the multiple cloning site (Invitrogen-Life Technologies). The sequence of the multiple cloning site and primer binding regions is attached in Appendix D.

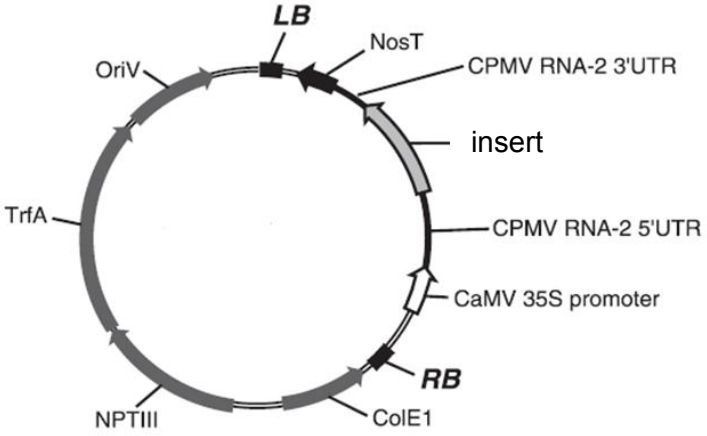


Figure 45: Features of pEAQ-HT

"Light grey arrows, regions non-essential for plasmid growth or transient transformation; dark grey arrows, regions essential for plasmid growth and replication; black boxes, T-DNA left (LB) and right (RB) borders; white arrows, promoter sequences (CaMV, Cauliflower mosaic virus); solid black lines, CPMV RNA-2 UTRs; bordered grey arrows, GFP coding sequences; and black arrows, terminator sequences (nos, nopaline synthase). OriV, pRK2 replication origin; ColEI, pBR322 replication origin; NPT, neomycin phosphotransferase; TrfA, replication-essential locus", insert: position of cloned gene(s). Adapted from Sainsbury et al. (2009)

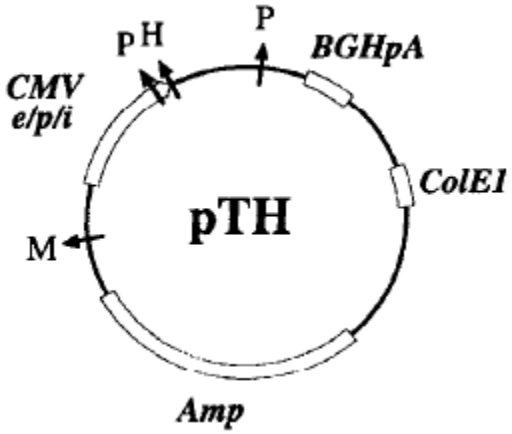
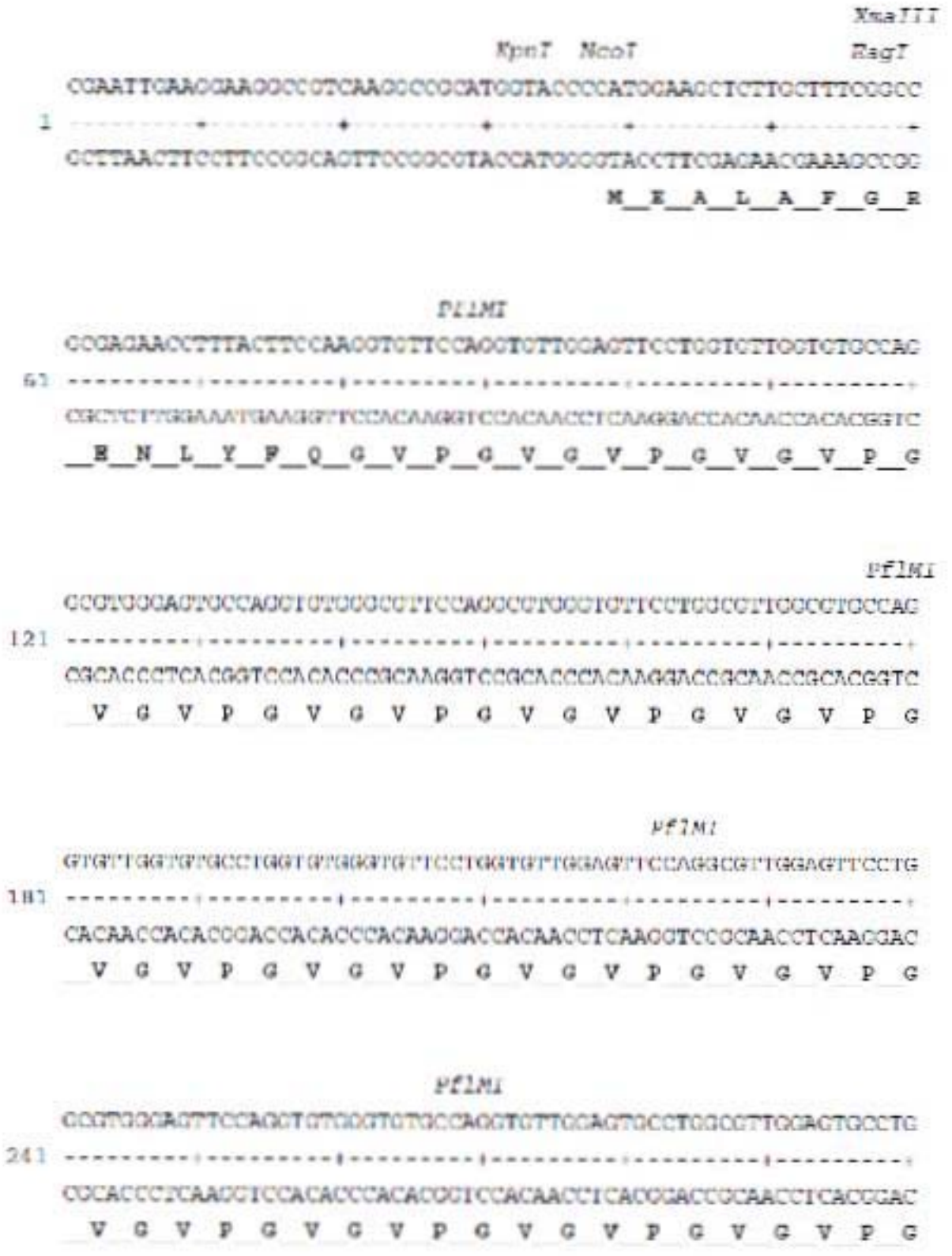
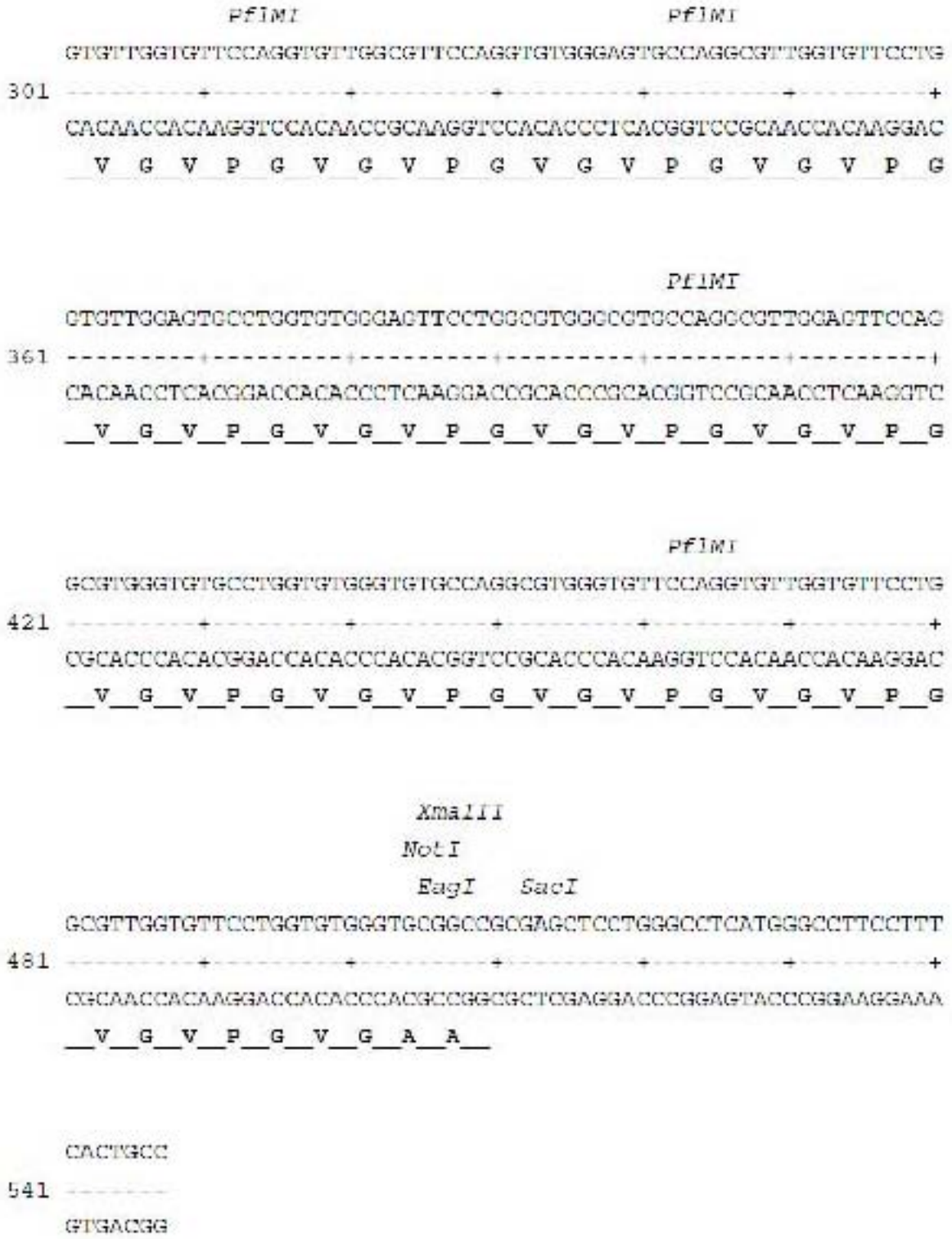


Figure 46: Features of the pTH expression vector

"CMVe/p/i: human cytomegalovirus enhancer/promoter/intron region; BGHpA: bovine growth hormone polyadenylation signal; Amp: β-lactamase gene; ColE1: origin of plasmid replication; Unique restriction sites: M: *Mlu*I, P: *Pst*I, H: *Hind*III". Adapted from Hanke et al. (1998)

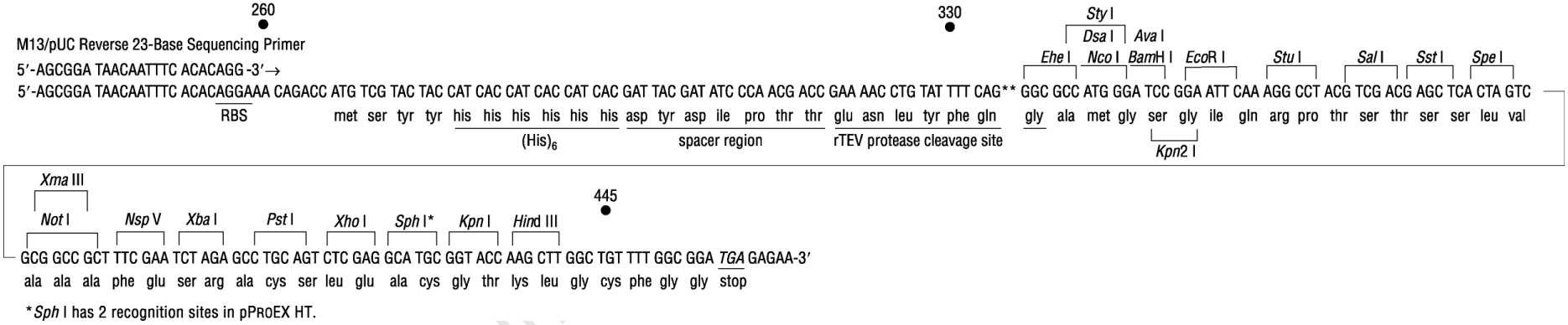
Appendix C: ELP28 gene sequence in pMK-TEV-ELP synthesized by GeneArt (Germany)




PflMI
PflMI
GTGTTGGTGTTCCAGGTGTTGGCGTTCCAGGTGTGGGAGTGCCAGGCGTTGGTGTTCCTG
301
CACAACCACAAGGTCCACAACCGCAAGGTCCACACCCTCACGGTCCGCAACCACAAGGAC
V G V P G V G V P G V G V P G V G V P G
PflMI
GTGTTGGAGTGCCTGGTGTGGGAGTTCCTGGCGTGGGCGTGCCAGGCGTTGGAGTTCCAG
361
CACAACCTCACGGACCACACCCTCAAGGACCGCACCCGCACGGTCCGCAACCTCAAGGTC
V G V P G V G V P G V G V P G V G V P G
PflMI
GCGTGGGTGTGCCTGGTGTGGGTGTGCCAGGCGTGGGTGTTCCAGGTGTTGGTGTTCCTG
421
CGCACCCACACGGACCACACCCACACGGTCCGCACCCACAAGGTCCACAACCACAAGGAC
V G V P G V G V P G V G V P G V G V P G
XmaIII
NotI
EagI SacI
GCGTTGGTGTTCCTGGTGTGGGTGCGGCCGCGAGCTCCTGGGCCTCATGGGCCTTCCTTT
481
CGCAACCACAAGGACCACACCCACGCCGGCGCTCGAGGACCCGGAGTACCCGGAAGGAAA
V G V P G V G A A
CACTGCC
541
GTGACGG

Appendix D: Sequence of multiple cloning site and promer binding reqions for pPROEX-HTb (Invitrogen™, Life Technologies)

pPROEX HTb multiple cloning site and primer binding region: 235-463



*Sph I has 2 recognition sites in pPROEX HT.